JOURNAL OF VIROLOGY, June 2007, p. 5850–5863 Vol. 81, No. 11
0022-538X/07/$08.00+0 doi:10.1128/JVI.02403-06

A Distinct Group of Hepacivirus/Pestivirus-Like Internal Ribosomal Entry Sites in Members of Diverse *Picornavirus* Genera: Evidence for Modular Exchange of Functional Noncoding RNA Elements by Recombination▽†

Christopher U. T. Hellen* and Sylvain de Breyne

Department of Microbiology and Immunology, State University of New York Downstate Medical Center, Brooklyn, New York 11203

Received 1 November 2006/Accepted 20 March 2007

The 5′ untranslated regions (UTRs) of the RNA genomes of *Flaviviridae* of the *Hepacivirus* and *Pestivirus* genera contain internal ribosomal entry sites (IRESs) that are unrelated to the two principal classes of IRESs of *Picornaviridae*. The mechanism of translation initiation on *h*epacivirus/*p*estivirus (HP) IRESs, which involves factor-independent binding to ribosomal 40S subunits, also differs fundamentally from initiation on these picornavirus IRESs. Ribosomal binding to HP IRESs requires conserved sequences that form a pseudoknot and the adjacent IIId and IIIe domains; analogous elements do not occur in the two principal groups of picornavirus IRESs. Here, comparative sequence analysis was used to identify a subset of picornaviruses from multiple genera that contain 5′ UTR sequences with significant similarities to HP IRESs. They are avian encephalomyelitis virus, duck hepatitis virus 1, duck picornavirus, porcine teschovirus, porcine enterovirus 8, Seneca Valley virus, and simian picornavirus. Their 5′ UTRs are predicted to form several structures, in some of which the peripheral elements differ from the corresponding HP IRES elements but in which the core pseudoknot, domain IIId, and domain IIIe elements are all closely related. These findings suggest that HP-like IRESs have been exchanged between unrelated virus families by recombination and support the hypothesis that RNA viruses consist of modular coding and noncoding elements that can exchange and evolve independently.

Positive-sense RNA virus genomes contain elements that play key roles during steps in infection, such as translation initiation, replication, and encapsidation. Their activities are independent of coding potential but depend on functional sequence motifs and structural integrity. These requirements impose constraints that limit sequence variation; thus, similar functional elements in related viruses contain conserved sequences, despite the high error rate of RNA genome replication (61). One such element is the 300-nucleotide (nt)-long internal ribosomal entry site (IRES) in the 5′ untranslated region (UTR) of hepatitis C virus (HCV), the type species of the hepacivirus genus of *Flaviviridae*. It mediates end-independent translation initiation by recruiting ribosomes to the initiation codon, has three domains (see Fig. 4A), and is the most conserved part of the genome (59, 70). Domain II is a 75-nt-long bent hairpin that enhances IRES function (32, 51). The essential 205-nt-long domain III structure consists of branching hairpins (IIIa-IIIf), many of which contain conserved, functionally important motifs; a series of imperfect stems (III_1 to III_4); and a pseudoknot that incorporates domain IIIf (47, 54). The initiation codon is part of the domain IV hairpin (14). The adjacent conserved coding region enhances IRES function (51). Initiation on this IRES differs from that on all eukaryotic cellular mRNAs in that the IRES binds to the small (40S) ribosomal subunit without the involvement of eukaryotic initiation factors (eIFs); ribosomal initiation complexes assemble at the initiation codon independently of the cap-binding complex eIF4F (45). The 40S subunit binds to noncontiguous parts of the IRES, including domain IIId, the pseudoknot, and nucleotides flanking the initiation codon (22, 26, 32), and eIF3 binds directly to the apical half of domain III (45, 62).

The 5′ UTRs of members of the pestivirus genus of *Flaviviridae*, such as classical swine fever virus (CSFV) and bovine viral diarrhea virus (BVDV), as well as of the HCV-like flavivirus GB virus B (GBV-B) contain sequences and structures similar to elements in the HCV IRES (1, 54) (see Fig. 4I). These IRESs contain regions of extensive similarity but also differ due to insertions (hairpin $IIId_2$ in pestivirus IRESs and hairpins IIb and IIc in the GBV-B IRES) and substitutions that, for example, result in the domain IV equivalent in pestiviruses being unstructured (62). These regions of the BVDV, CSFV, and GBV-B 5′ UTRs are IRESs (47, 48, 52, 53), and pestivirus IRESs have been found to mediate initiation by a mechanism like that used by the HCV IRES (42, 45). Moreover, conserved elements in HCV and pestivirus IRESs play identical roles in initiation (7, 26, 27, 45, 53, 71). Because of the structural and functional conservation, we shall refer to *h*epacivirus and *p*estivirus IRESs collectively as HP IRESs.

Two other mechanisms of IRES-mediated initiation that occur on structurally distinct IRESs have been characterized. One is exemplified by the ~200-nt-long intergenic region discistrovirus IRESs, which bind directly to ribosomes and initiate translation without the involvement of initiator tRNA or ini-

* Corresponding author. Mailing address: Dept. of Microbiology and Immunology, SUNY Downstate Medical Center, 450 Clarkson Avenue, Box 44, Brooklyn, NY 11203. Phone: (718) 270-1034. Fax: (718) 270-2656. E-mail: christopher.hellen@downstate.edu.

† Supplemental material for this article may be found at http://jvi.asm.org/.

▽ Published ahead of print on 28 March 2007.

TABLE 1. Viral sequences identified using BLAST as matches at the indicated levels of identity to the indicated HCV, BVDV, CSFV, and PTV-1 IRES sequences

Parental sequence	Result for indicated virus group					
	Contiguous identical nucleotides	Matching sequences (total)	Pestiviruses	GBV-B viruses	Picornaviruses	Other viruses
HCV1a nt 287–302 (identical to GBV-B nt 391–406)	≥12	968[a]	913	3	11	40[a]
	≥11	1,335[a]	1,163	3	12	155[a]
HCV G903 nt 188–203	≥14	57[a]	0	0	57	0[a]
	≥13	60[a]	0	0	1	2[a]
	≥14/15	18[a]	0	3	10	5[a]
CSFV nt 304–319	≥11	87[a,b]	38[a,b]	0	6	41[a,b]
	≥11	1,280[a,b]	861[a,b]	3	15	401[a,b]
BVDV nt 319–334	≥11	126[a,c]	89[a,c]	3	9	25[a,c]
	≥11	467[a,c]	240[a,c]	3	12	212[a,c]
PTV-1 nt 352–367	≥12	88[a]	0	0	60	28[a]
	≥13/14	19[a]	0	0	11	8[a]

[a] To simplify analysis, HCV sequences were excluded from consideration as indicated.
[b] To simplify analysis, CSFV sequences were excluded from consideration as indicated.
[c] To simplify analysis, BVDV sequences were excluded from consideration as indicated.

tiation factors (43, 46). The other is used by IRESs from the *Aphthovirus* and *Cardiovirus* genera of *Picornaviridae* and involves binding of the eIF4G and eIF4A subunits of eIF4F to an IRES domain upstream of the initiation codon, followed by ATP-dependent recruitment by them of the 43S preinitiation complex (44). The identification of three distinct mechanisms of IRES-mediated initiation raises the question of whether other, less well-characterized IRESs employ variants of one of these mechanisms or use novel mechanisms. Computational methods can theoretically be used to identify mRNAs that contain IRESs that are related to members of these three groups, but a search of this type for intergenic region IRES-like RNAs by use of a pattern-matching algorithm failed to identify any strong candidates (12). On the other hand, recent studies have noted homology between the HCV IRES and the 5′ UTRs of the picornaviruses porcine teschovirus (PTV) and porcine enterovirus 8 (PEV-8) (29, 47) and have suggested that PTV, PEV-8, and simian picornavirus (SPV) contain HP-like IRESs (2, 3, 46, 47).

Here, we report that database searches for mRNAs containing sequences related to conserved elements in HP IRESs identified closely related sequences in the 5′ UTRs of seven different picornaviruses. Modeling suggested that they have a structure that resembles that of HP IRESs, including the functionally important pseudoknot and domains IIId and IIIe. The HP and HP-like IRESs can be divided into three groups: (i) pestivirus IRESs; (ii) HCV, GBV-B, putative avian encephalomyelitis virus (AEV), and Seneca Valley virus (SVV) IRESs; and (iii) a third novel group, consisting of the other five (variously putative and proven) picornavirus IRESs. Nucleotide sequence covariation in the 5′ UTRs from isolates of each of these picornaviruses favors the individual proposed structures. These observations, taken together with mechanistic data concerning these mRNAs (2, 3, 46, 47), suggest that functional noncoding RNA elements have been exchanged by recombination between members of distinct enveloped and nonenveloped RNA viruses. The implications of these findings for viral evolution are discussed below.

MATERIALS AND METHODS

Sequences. Sequences analyzed were from the AEV Calnek (GenBank accession no. AJ225173), L2Z (AY275539), and van Reokel (AY517471) strains; duck picornavirus (DPV; strain TW90A; AY563023); duck hepatitis virus 1 (DHV-1) strains 03D (DQ249299), H (DQ249300), 5886 (DQ249301), R85952 (DQ226541), DRL-62 (DQ219396), JX (EF0935902), C80 (DQ864514), and A66 (DQ886445); PEV-8 strain V13 (AF406813); PEV-8 isolates 38-V-II (AY392545), 26-T-XII (AY392544), 16-S-X (AY392543), Po 5116 (AY393538), and Sek 1562/98 (AY392556) and 14 other PEV-8 isolates (28); PTV-1 Talfan (AB038528); PTV-2 isolates T80 (AF296087) and DS 756/93 (AY392534), PTV-3 isolate O2b (AF296088), PTV-4 isolate PS36 (AF296089), PTV-5 isolate F26 (AF296090), PTV-6 isolate PS37 (AF296091), PTV-7 isolate F43 (AF296092), PTV-8 isolate UKG 173/74 (AF296093), PTV-9 isolate Vir 2899/84 (AF296094), PTV-10 isolates Vir 460/88 (AF296095) and UKG/170/80 (AY392539), PTV-11 isolate Dresden (AF296096), and other PTV isolates in the NCBI database; SVV (DQ641257); and SPV-1 (formerly simian virus 2 [SV2; AY064708]), SPV-3 (formerly SV16; AY064715), SPV-4 (formerly SV18; AY064716), SPV-9 (formerly SV42; AY064717), SPV-11 (formerly SV44; AY064718), SPV-12 (formerly SV45; AY064719), and SPV-15 (formerly SV49; AY064720). Sequences analyzed were from HCV types 1a (NC_004102), 2a (AB047639), 3a (D28917), 4a (Y11604), 5a (AF064490), and 6a (AY859526), HCV isolate JFH-1 (AB047639), and HCV strain G903 (AY236366); GBV-B (NC_001655); border disease virus 1 (BDV-1) isolate X818 (NC_003679), BDV-2 (reindeer-1 V60-Krefeld) (NC_003677), and BDV-4 (Chamois-Spain02) (AY641529); CSFV Alfort (J04358); BVDV-1 NADL (M31182) and BVDV-2 strain C413 (AF002227); pronghorn antelope pestivirus (AY781152); ovine pestivirus strain SN2T (AF461996); and giraffe pestivirus H138 (AB040131).

Sequence alignment. Picornaviruses containing HP IRES-like sequences in their 5′ UTRs were identified with BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST/) of viral sequences in the GenBank database. For short (16-nt) sequences, the parameters used were as follows: *E* (expect) = 1,000, word size = 7, reward match/mismatch = 1/−3, gap extension = 5, gap penalty = 2. Only those matches with the greatest percent identities were considered, as described in the text. Searches done with complete picornavirus 5′ UTRs used the following parameters: *E* = 10, word size = 11, reward match/mismatch = 1/−3, gap extension = 5, gap penalty = 2. Nucleotide sequences were aligned with CLUSTAL-W (http://www.ebi.ac.uk/clustalw/index.html) (67), using the default parameters (DNA gap open penalty = 15.0, DNA gap extension penalty = 6.66,

DNA matrix = identity, DNA ENDGAP = −1, DNA GAPDIST = 4), and with T-COFFEE (http://www.ebi.ac.uk/t-coffee/) (37), using the default parameters.

Modeling secondary and tertiary IRES structures. Secondary/tertiary structure elements in picornavirus 5′ UTRs were modeled using Pfold (http://www.daimi.au.dk/~compbio/rnafold/) (25) and ILM (http://cic.cs.wustl.edu/RNA/) (55, 56). Pfold and ILM analyses were done using individual DPV and SVV sequences, aligned CSFV Alfort and BVDV NADL sequences, aligned sequences of HCV types 1a to 6a, aligned sequences of the PTV type 1 to 11 isolates listed above, and for AEV, DHV, PEV8, and SPV, aligned sequences of the isolates listed above. Sequences were submitted for analysis in FASTA format.

Mfold (75) was used to verify secondary structures of RNA sequences and to identify stable structures not identified by Pfold, using default parameters (37°C; 1 M NaCl, no divalent ions; 5% suboptimality; maximum interior/bulge loop size = 30 nt; maximum asymmetry of an interior/bulge loop = 30 nt; no limit on the maximum distance between paired bases). The ILM (iterated loop matching) algorithm was used with the following parameters: minimum loop length (L) = 3, minimum virtual loop length (V) = 3, minimum helix length (H) = 3, number of helices per iteration (N) = 1, and number of iterations before stop (I) = 0. The number of sequences analyzed for each virus was small (<10), so a 1:3 mutual-information-score-to-extended-helix-plot-score ratio was used to generate the scoring matrix.

Covariance analysis. AEV, DHV, PEV-8, PTV, and SPV sequences were aligned against the sequences shown in Fig. 4B, D, E, G, and H. Sequences in each alignment were then scored for substitutions that maintained or disrupted the predicted base pairing. Substitutions were categorized as "covariant" (in which paired changes, such as A-U⇔G-C substitutions, maintain base pairing), "neutral" (in which single changes, such as A-U⇔G-U and G-U⇔G-C, occur without disrupting base pairing), "disruptive" (in which single changes, such as A-U⇔A-C, disrupt base pairing), or "unpaired" (occurring in unpaired/single-stranded regions). To calculate the expected numbers of random occurrences for these categories of changes, the RNA sequence corresponding to the pseudoknot, domains IIId and IIIe, and helices III_1 and III_2 of PEV-8 strain V13 was subjected to in silico evolution to generate a family of related sequences that diverged ~10% or less from the parental sequence, using Rose (66) (http://bibiserv.techfak.uni-bielefeld.de/rose/welcome.html) with the Kimura two-parameter substitution model, a transition bias of 1.0, a transition-to-transversion ratio of 10, and no allowance of insertions or deletions. Substitutions in 1,000 sequences generated in this way from 10 independent experiments were scored manually for substitutions that maintained or disrupted base pairing in the predicted domain III structure.

RESULTS

Sequence similarities between HCV-like IRESs and the 5′ UTRs of diverse picornaviruses. Domain IIIe is part of the most conserved region in HP IRESs (54). BLAST searches of viral sequences that matched domain IIIe and the 4 upstream nucleotides of the functionally and structurally characterized HCV, CSFV, BVDV, and GBV-B IRESs at 11 or more contiguous nucleotides identified sequences in other HP IRESs, in the 5′ UTRs of the picornaviruses AEV, DHV-1, PEV-8, PTV, and SVV, and in an assortment of other viruses (Table 1 and Fig. 1A to C). BLAST searches done in this way to identify matches for equivalent 16-nt-long elements from these picornaviruses identified sequences from heterologous picornaviruses, HP IRESs, and a variety of unrelated viruses; BLAST searches that allowed a single mismatch in 13 or more nucleotides of this motif identified isolates of these five picornaviruses and SPV types 1, 3, 4, 9, 11, 12, and 15 (Table 1 and data not shown). HCV sequences were excluded from some analyses as indicated (Table 1) because of the overwhelming number of matches at these cutoff levels. Matching sequences were not considered if they occurred in bacteriophages, in the negative-sense strand, in the principal open reading frame, or in the 3′ UTR. These criteria eliminated all but the HP and picornavirus sequences.

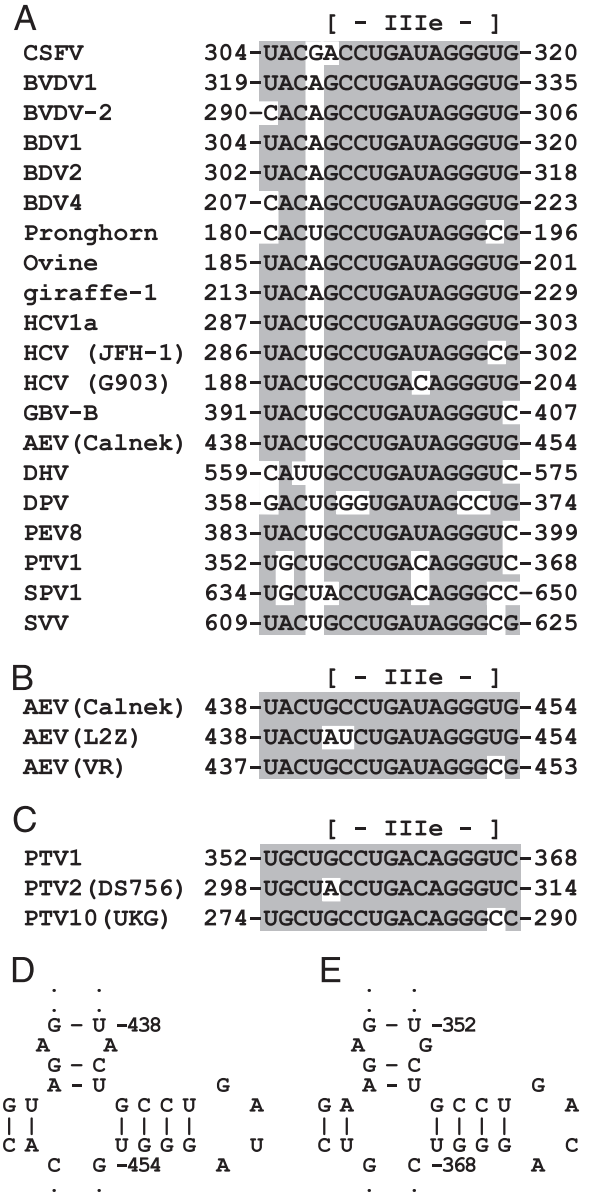


FIG. 1. Sequence variation in the domain IIIe structures of hepacivirus, pestivirus, and picornavirus HP-like IRESs. (A) Alignment of sequences corresponding to domain IIIe and flanking nucleotides (for accession numbers, see Materials and Methods). (B to E) Structurally neutral substitutions in isolates of AEV (B and D) and PTV (C and E), shown as aligned sequences (B and C) and mapped onto partial secondary structure models (D and E) of AEV (Calnek) (D) and PTV-1 IRES domain IIIe (E). PTV-2 (DS 756) refers to PTV-2 (DS 756/93). PTV-10 (UKG) refers to PTV-10 (UKG 170/80). (A to C) The 5′ and 3′ borders of each element are numbered; identical nucleotides are shaded.

To determine whether sequence homology extended beyond this motif, BLAST searches were undertaken to identify matches for the complete 5′ UTR of each picornavirus. These analyses confirmed and extended relationships between the 5′ UTRs of these picornaviruses and between the picornavirus 5′ UTRs and HP IRESs. For example, a search done with the PEV-8 (strain V13) 5′ UTR identified matches at a significance level (E) of <0.01 in 88 PEV-8, PTV, DHV, and SPV 5′

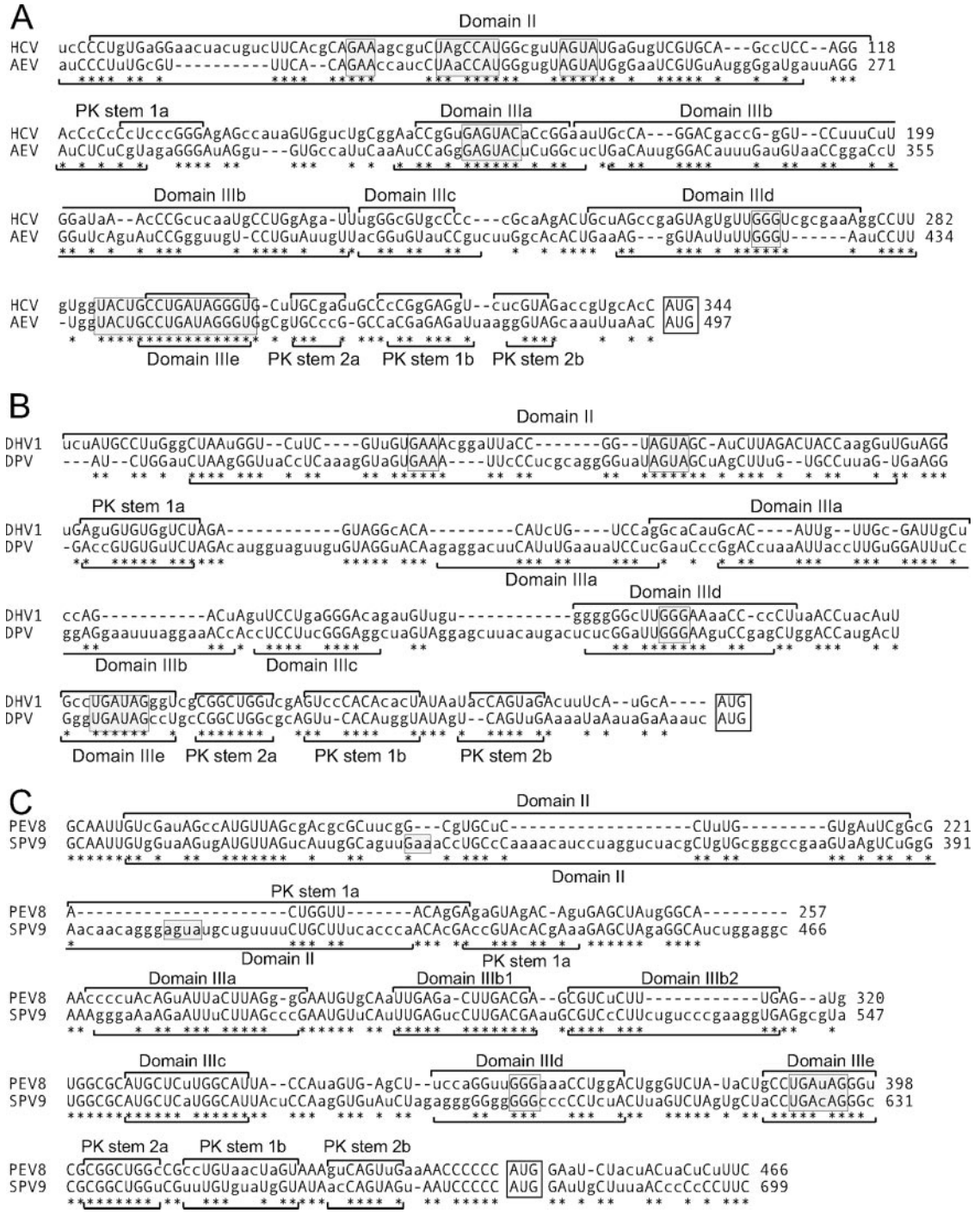


FIG. 2. T-COFFEE sequence alignments of the HCV IRES and the AEV 5′ UTR (A), the DHV-1 and DPV 5′ UTRs (B), and the PEV-8 and SPV-9 sequences, including the 5′ UTRs and adjacent coding regions (C); identical nucleotides are in uppercase and marked with asterisks. Sequences were annotated after alignment to show the boundaries of predicted structural elements.

UTR sequences but not in a single other viral sequence. The 100 next-best matches comprised 7 picornavirus sequences (the 5′ UTRs of DPV-1, DHV, PEV-8, and SPV isolates), 88 HP IRESs, and 5 unrelated virus sequences (that were discounted based on criteria described above). Similar results were obtained in searches done using other HP-like picornavirus 5′ UTRs, including that of DPV. As a further preliminary test of potential similarity with HP IRESs, the 3′-terminal 100 nt of the 5′ UTRs (including the initiation codon) were compared by pairwise CLUSTAL-W alignment. The region of homology between all picornavirus 5′ UTRs and HP IRESs began in the immediate vicinity of the initiation codon. The degree of identity is high within virus families (40 to 65% for picornaviruses) and is as high between picornaviruses and fla-

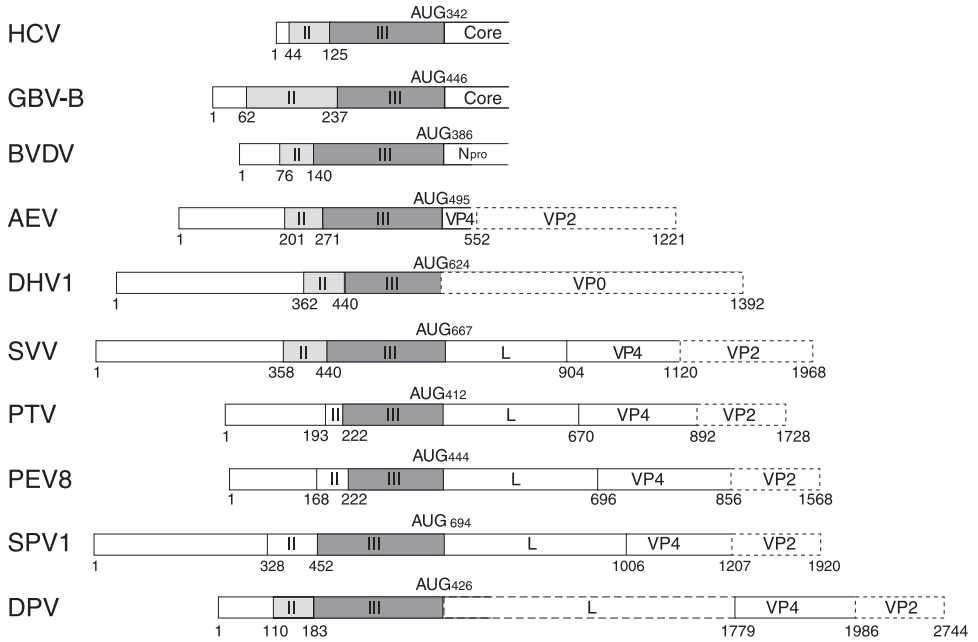


FIG. 3. Schematic representation of the 5′ UTRs and adjacent regions of HCV, GBV-B, and BVDV (flaviviruses) and of AEV, DHV-1, DPV, PTV, PEV-8, SPV-1, and SVV (picornaviruses). 5′ UTRs are divided into domains and are shaded to show those that are homologous to HCV domain II (light gray) and domain III (dark gray). The picornavirus L (leader), VP4, and VP2 (capsid) proteins, the HCV and GBV-B core proteins, and the BVDV N^{pro} protease (not shown to scale) are labeled. The initiation codon for each polyprotein is indicated; nucleotides at the boundaries between 5′ UTR domains and protein-coding regions within each polyprotein are numbered. Note that the sequences of the 5′ UTRs of PEV-8 isolates (28) and PTV isolates (20) upstream of the short poly(C) tract are known to be incomplete and that the presence of a 5′-terminal poly(C) tract in DPV-1 TW90A (GenBank accession number AY563023) suggests that it is likely also incomplete.

viviruses (39 to 58%) as between all but the most closely related flaviviruses (data not shown).

Sequence alignment of picornavirus 5′ UTRs and HP IRESs. Sequence similarities between individual picornavirus RNAs and representative HP IRESs extended for 200 to 300 nt upstream of the initiation codon and, in some genomes, from the 5′ terminus to the coding region (Fig. 2). Sequence similarities were greatest for PTV types 1 to 11 (>97% nucleotide identity with PTV-1 nt 160 to 411) (see the supplemental material), consistent with a prior report concerning sequences downstream of nt 300 (74). They will therefore be treated as variants of a single entity. SPV types 4, 11, and 15 and types 1, 3, 9, and 12 formed two closely related groups (see the supplemental material).

Other significant similarities were noted. AEV nt 201 to 497 aligned with sequences that form HCV IRES domains II and III and GBV-B IRES domain III (57% and 43% identity, respectively), and homologous sequences formed homologous structural domains (Fig. 3A and data not shown). Similar levels of sequence identity were apparent in pairwise comparisons of the HCV IRES, DPV nt 110 to 428, SVV nt 356 to 669, and DHV-1 nt 362 to 626 (Fig. 3B). Homology was particularly strong between sequences that are predicted to form domains IIId and IIIe and elements of the pseudoknot (see below); pairwise comparisons indicated that homologous sequences in some subsets of these viruses formed other homologous elements, such as PEV-8/SPV domains $IIIb_1$ and IIIc (Fig. 3C). Sequence identities between HCV domain II and the corresponding elements in these and SPV 5′ UTRs were centered on the opposed GAA and AGUA motifs (which form a loop E motif in the domain II structures of all HP IRESs [29, 32, 54]). By contrast, equivalent PEV-8 and PTV sequences lack these motifs and their sequence identities with HCV domain II are lower. Sequence identity between PEV-8 and SPV genomes extended into the coding region, but whereas the 3′-terminal ~180 nt of the PEV-8 5′ UTR aligned well with the SPV-9 5′ UTR (~70% nucleotide identity) and contained only a few small gaps, alignment of the upstream regions revealed larger gaps (Fig. 3C). This alignment of SPV and PEV-8 sequences, the predicted position of the pseudoknot of SPV IRESs relative to the principal open reading frame (see below), and in vitro translation data for SPV-1 (2) all indicate that the initiation codon in, e.g., the SPV-9 genome is A_{676} (rather than AUG_{724} as previously suggested) (39).

Bioinformatic determination of HP IRES structure. The sequence similarities with HP IRESs suggest that picornavirus 5′ UTRs might have similar secondary/tertiary structures, but the presence of large insertions/deletions in sequence alignments indicates that they cannot simply be modeled using HP IRESs as structural templates. To model the structures of the HP IRES-like picornavirus sequences, we used Pfold (25), which is based on an explicit evolutionary model and yields a probabilistic model of secondary structure (5), and ILM (iterated loop matching) (55, 56), which can utilize both thermodynamic and comparative information to predict RNA secondary structures and pseudoknots. To validate this approach, Pfold was first used to fold the domain III structures of BVDV, CSFV, and HCV IRESs, which have previously been determined by chemical/enzymatic probing, X-ray crystallography, nuclear magnetic resonance, phylogenetic comparison, and

TABLE 2. Structural elements identified using Pfold and ILM in HCV and CSFV/BVDV IRESs and in AEV, DHV, DPV, PEV8, PTV, SPV, and SVV 5′UTRs[a]

Structural element	Result for indicated virus and method																	
	HCV		CSFV/BVDV		AEV		DHV		DPV		PEV8		PTV		SPV		SVV	
	Pfold	ILM	Pfold	ILM	Pfold	ILM	Pfold	ILM	Pfold	ILM	Pfold	ILM	Pfold	ILM	Pfold	ILM	Pfold	ILM
Domain II											+[b]	+[b]		+	+	+[b]	+	
PK1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
PK2		+		+						+		+				+		
IIIa	+	+	+	+	+		+		+		+	+	+		+	+	+	
IIIb	+	+	+	+	+[c]						+[c]	+[c]	+	+[c]	+	+		
IIIc	+		+		+												+	
IIId			+	+	+	+	+		+		+	+	+	+	+	+	+	+
$IIId_2$	NA	NA	+	+	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
IIIe			+		+						+							
IIIf							+				+							
III_1		+	+	+			+		+		+	+	+		+	+		
III_2	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+		
III_3	+		+	+	+					+	+*	+*	+		+	+	+	
III_4	+		+	+	NA						+				+		+	

[a] Structural elements, including helices in domain III, are as indicated in Fig. 4A to I. NA, not applicable.
[b] Only the apical hairpin of domain II was identified.
[c] Pfold identified AEV domain $IIIb_2$, Pfold and ILM identified PEV-8 domain $IIIb_2$ and helix III_3 (Fig. 4H, inset panel), and Pfold and ILM identified PTV domain IIIb, as seen in Fig. 4D, inset panel.

mutational analysis (47). Pfold identified the majority of individual structural elements in the domain III structures of HCV types 1a to 6a and of BVDV/CSFV except, as expected, in pseudoknot stem 2 (PK2), for which no alternative was predicted (Table 2). ILM identified both PK1 and PK2 of many of these confirmed/putative IRESs but identified a smaller subset of other structural elements than Pfold (Table 2).

Structural models of picornavirus IRESs. Having established the validity of using Pfold and ILM to determine the structures of HP IRESs, we then used them and Mfold (75) (which generates complementary structural predictions on the basis of energy minimization) to characterize HP-like picornavirus sequences. Individual elements identified using Pfold and ILM are indicated in Table 2; individual elements in each of these IRES-like elements and their sizes are given in Table 3. Elements shown but not identified using Pfold and ILM were instead predicted using Mfold, with the exception of some PK2 helices, as described below. The predicted picornavirus structures resemble HP IRESs in key respects (Fig. 4B to H and Table 3), notably in containing domain III (a basal pseudoknot, a series of helices [III_1 to III_4], and a series of branching hairpins, including domains IIId and IIIe), an upstream bent hairpin-like structure (domain II), and several conserved internal and apical loops. The loops indicated by

TABLE 3. Structural elements and sequence motifs in group A, group B, and group C HP-like IRESs

IRES group and virus	Data for indicated structural element												
	Domain II			Pseudoknot				Domain III					
	Length (nt)	Presence of:		Length (nt) of:				Presence of:					III_1 mispair
		Loop E motif	UANCCAU loop	Stem 1	Stem 2	Linker	Spacer	IIIa AGUA loop	IIIc	IIId G loop	IIId2	IIIe	
A													
HCV	75	+	+	9/9	6/6	1	11	+	+	+	−	+	A-A
GBVB	174	+	(+)	9/9	6/6	1	11	−	+	+	−	+	A-A
AEV	64	+	+	9/9	5/5	3	11	+	+	+	−	+	A-A
SVV	79	+	−	11/11[a]	5/5	1	13	+	+	+	+	+	A-A
B													
PEV-8	54	−	−	12/12	8/8	3	10	−	−	+	−	+	A-A
PTV-1	29	−	−	12/12	8/8	3	9	−	−	+	−	+	A-G
SPV-9	119	+	−	12/12	8/8[a]	3	9	−	−	+	−	+	A-G
DHV-1	76	+	−	12/12[a]	8/8	5	11	−	−	+	−	+	A-A
DPV	66	+	−	12/12[a]	8/8	3	17	−	+	+	−	+	A-A
C													
CSFV	59	+	+	14/16[a]	7/7	7	12	+	+	+	+	+	A-A
BVDV	64	+	+	13/15[a]	7/7	6	12	+	+	+	+	+	A-A

[a] Numbers defining the lengths of helices in the pseudoknot include internal mispairs and internal bulges as appropriate. "Spacer" refers to the nucleotides between the pseudoknot and the initiation codon.

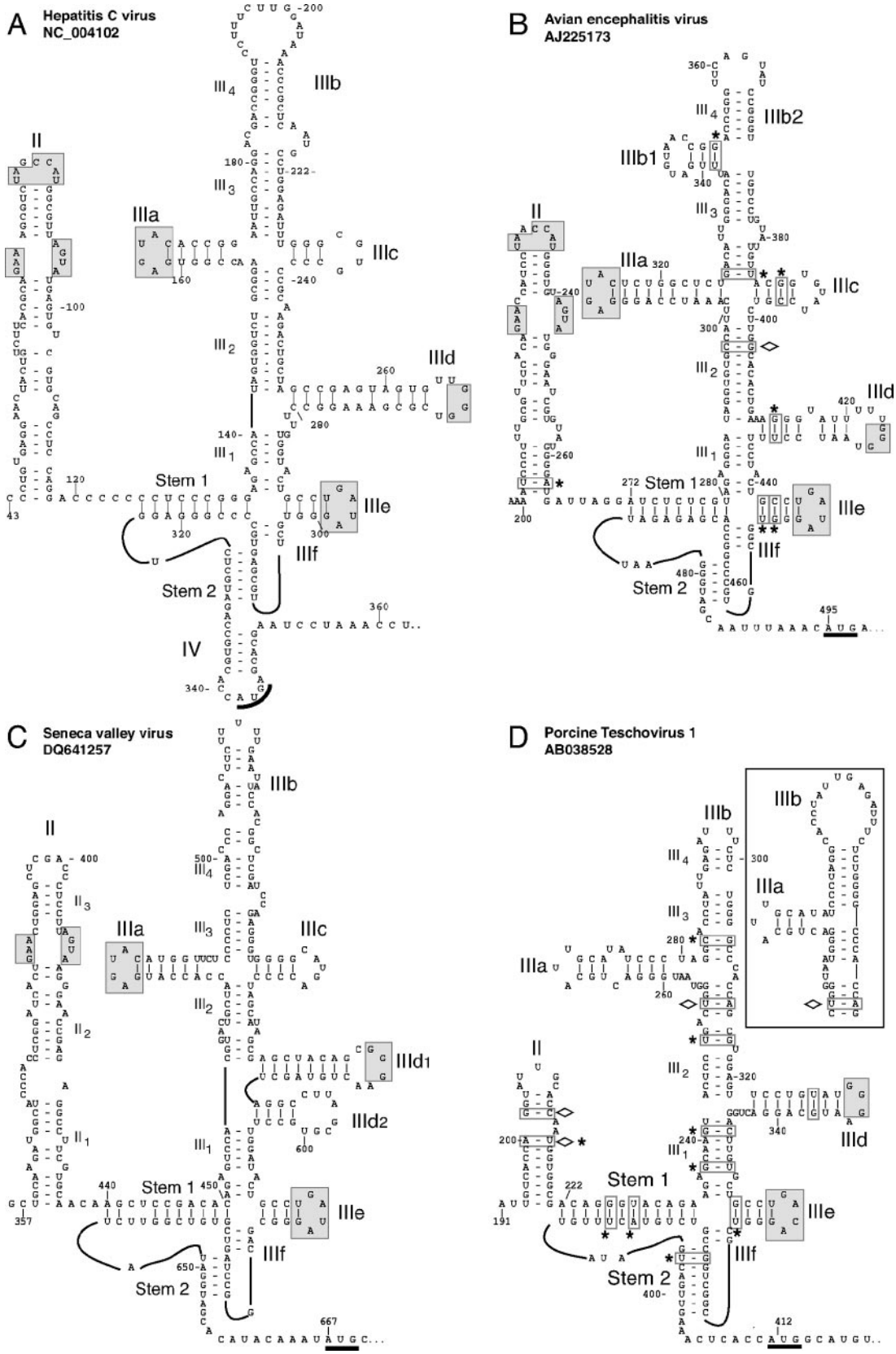
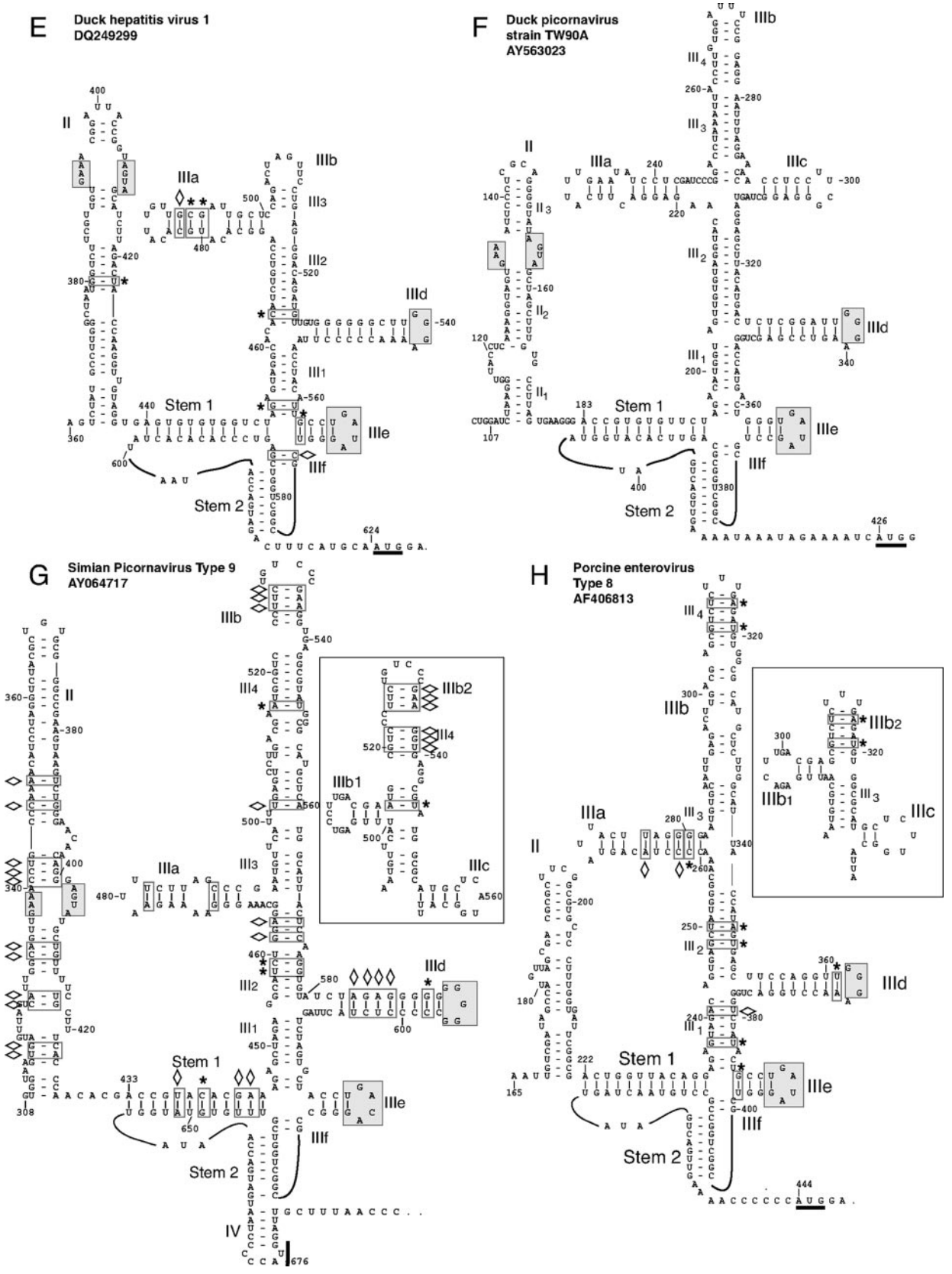


FIG. 4. Models of predicted secondary and tertiary structures of HCV and various picornavirus IRESs. Models of the IRESs of HCV (A), AEV (B), SVV (C), PTV-1 (D), DHV-1 (E), DPV (F), SPV-9 (G), and PEV-8 (H) were derived as described in the text. Domains are labeled II, II, and IV; individual helical segments are labeled II_1, II_2, III_1, and III_2, etc.; and individual hairpins are labeled IIIa and IIIb, etc., and are analogous to similarly labeled structures in the HCV IRES. To maintain the continuity of the current nomenclature, we have designated additional hairpin elements $IIIb_1$ in the AEV IRES and $IIId_2$ in the SVV IRES. Stem 1 and stem 2 are elements of the pseudoknot. AUG triplets marked with a solid black bar represent the translation initiation site for the viral polyprotein. Lightly shaded rectangles indicate base pairs that are maintained



despite sequence variation between types, strains, and isolates of designated viruses; single-site substitutions that do not disrupt base pairing are indicated by asterisks, and base pairs maintained by paired covariant substitutions are indicated by lozenges. Unpaired bases in the domain II structures of individual picornavirus IRESs that are conserved in the HCV IRES are indicated by gray shading, as are the conserved apical loops of domain IIIa (as appropriate) and domain IIIe and the apical G-rich element of the domain IIId loop. The insets for panels D, G, and H represent potential alternative structures of the apical IIIab regions of PTV, SPV-9, and PEV-8 IRESs identified by Pfold and ILM, as described in the text, whereas the structures formed by the equivalent residues shown in the main panel were determined using Mfold.

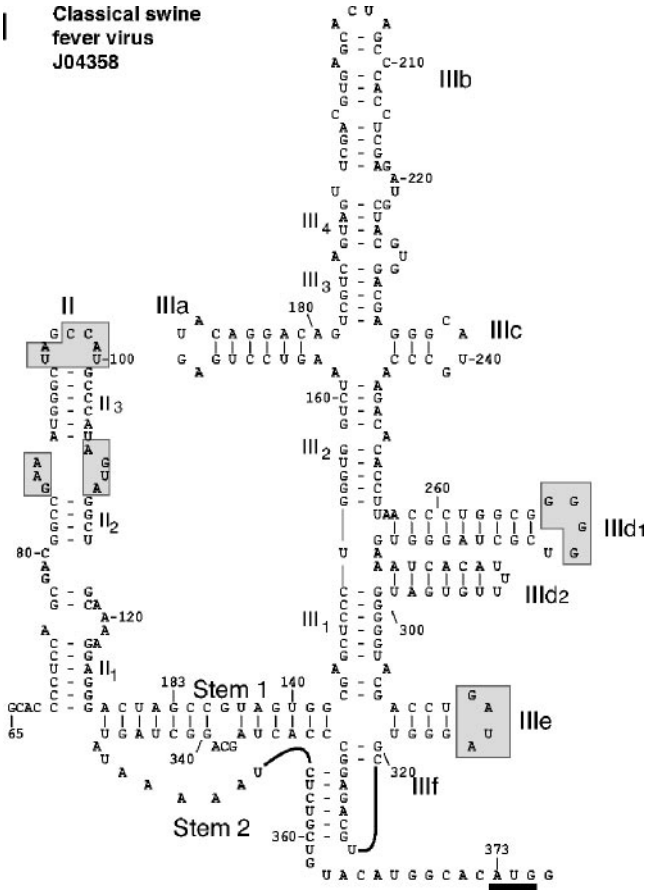


FIG. 4—*Continued.*

shading in Fig. 4B to I and in Table 3 are analogous to the HCV domain II E-loop $GAA_{71\text{-}73}$/$AGUA_{93\text{-}96}$ motifs and the apical HCV domain II $UANCCAU_{80\text{-}86}$ motif (which has the slightly divergent sequence GAGCCAAC in GBV-B) and to the $GAGUAA_{161\text{-}166}$, $GGG_{266\text{-}268}$, and $GA(U/C)A_{294\text{-}299}$ motifs in HCV domains IIIa, IIId, and IIIe, respectively. Table 3 also indicates the sequence of the conserved mispair in helix III_1 of domain III and whether the predicted structures contain domains IIIc, $IIId_2$, and IIIe. Individual picornavirus HP IRES-like structures differed from canonical HP IRESs most significantly in terms of the size and structure of domain II and the apical region of domain III and by the absence of a domain IV-like hairpin in all viruses but SPV.

Pfold identified the majority of HP IRES-like structural elements in these RNAs, except for the PK2 helices, which were identified in DPV, PEV-8, PTV, SPV, and SVV 5′ UTRs by using ILM. Alignment with other HP IRES-like sequences (Fig. 5) indicates that AEV and DHV 5′ UTRs can nevertheless potentially form PK2 helices and, consequently, that all seven picornavirus 5′ UTRs can form an HP-like pseudoknot 9 to 11 nt upstream of the initiation codon (except in SVV and DPV, which are separated by 13 nt and 17 nt, respectively). Elements of the IRES-like fold that were not identified by Pfold or ILM (such as domain IIIe and some helices in the IIIabc region of AEV, DHV, DPV, and SVV) were identified using Mfold. Mfold was also used to test the predicted 5′ UTR structural models by determining whether the helical and stem-loop substructures identified by Pfold (excluding elements of the pseudoknot) were the most thermodynamically stable structures within larger fragments of each 5′ UTR. Mfold analysis supported the results of Pfold analysis in all instances, except for the differences in predicted structures in the "IIIabc" regions of DPV, PEV-8, PTV, and SPV IRESs noted below. The resulting predicted structures are shown in Fig. 4B to H (together with the HCV [Fig. 4A] and CSFV [Fig. 4I] IRESs).

Domain II in AEV, DHV, DPV, and SVV 5′ UTRs resembles domain II in conventional HP IRESs, whereas the SPV domain II elements were significantly larger than those in HP IRESs, and the PEV-8 and PTV domain II elements were significantly smaller and lacked the GAA and AGUA motifs present in HP IRESs and all other picornavirus HP IRES-like elements (Fig. 4A to I). The structure of PTV domain II proposed here differs significantly from that in a recent proposal (3) but was independently identified by Pfold, ILM, Mfold, and Alifold (13) analyses of multiple PTV sequences and is supported by analysis of covariant nucleotide substitutions. Its 5′ border does not correspond to the experimentally determined 5′ border of the PTV IRES at nt 126 (46). Bioinformatic analyses suggested that PTV nucleotides between nt 126 and domain II have the potential to form additional hairpins that may thus be functionally important (data not shown).

The numbers and sizes of helices that formed the apical half of domain III differed significantly between these different picornaviruses, and in contrast to basal elements of domain III, their structures were not all predicted unambiguously by Mfold, ILM, and Pfold. The DPV IIIabc region predicted by Pfold is shown in Fig. 4F: Mfold analysis indicated that this corresponds to the most stable conformation for this region and for the 11-base-pair III_2 helix, except that base pairing of UCC245-247 and GGA314-316 was indicated. The relative sizes of PTV domains IIIa and IIIb predicted by Mfold (Fig. 4D) and by Pfold and ILM (Fig. 4D, inset) differed slightly, and whereas Pfold and ILM suggested that the apical regions of the PEV-8 and SPV domain III structures can form a four-way junction and that domain IIIb consists of two subdomains (Fig. 4G and H, insets), Mfold suggested that these regions fold most stably to form Y-shaped structures (Fig. 4G and H), indicating that, like PTV, SPV and PEV-8 would lack domain IIIc. This conclusion is consistent with suggestions by Chard et al. (3) that PTV domain III may lack a stable four-way junction near its apex.

Conservation of structure in picornavirus HP-like IRESs. The predicted 5′ UTR structural models were also tested by assessing their compatibility with sequences of closely related viruses and, when possible, of divergent viral isolates that had not been used to build the models. The few insertions/deletions in the aligned sequences are limited to the apexes of PEV-8 domain IIIb and SPV domain IIId and to unpaired elements of SPV domains $IIIb_1$ and $IIIb_2$. Most substitutions (Table 4) occurred in unpaired regions ("unpaired" substitutions) or did not disrupt base pairing (e.g., A-U⇔G-U and G-U⇔G-C transitions ["neutral" transitions]). Some substitutions in unpaired regions which were not separately classified were predicted to lead to the formation of additional base pairs: nine such substitutions in the domain II structures of SPV isolates were equivalent to $A_{349}G$, $C_{387}U$, $U_{406}A$, and $U_{406}G$ substitutions in

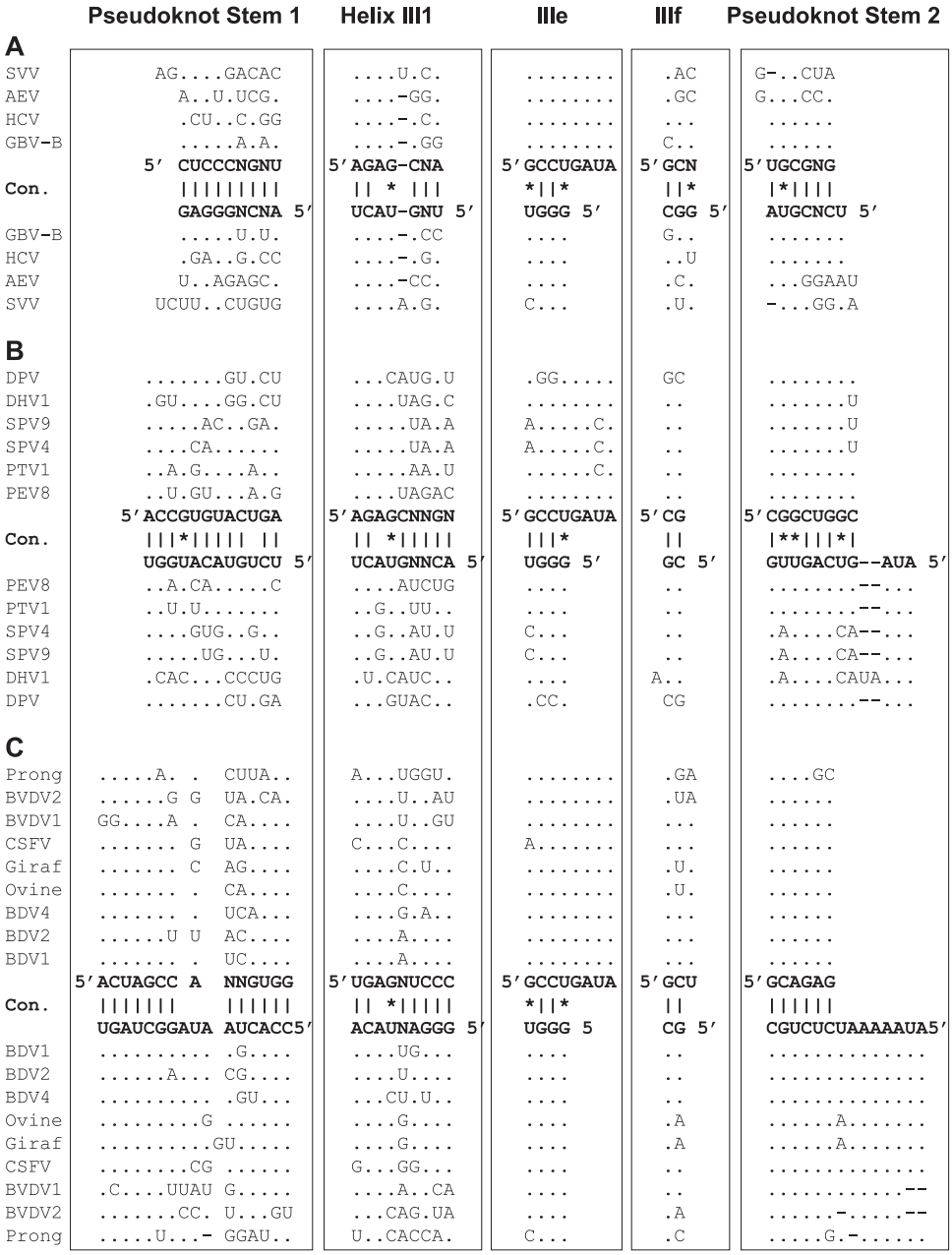


FIG. 5. Segregation of IRESs into three groups on the basis of length and sequence covariation of the pseudoknot and adjacent elements. The boundaries of elements in confirmed/putative picornavirus IRESs are as shown in Fig. 4B to H and references 1 and 54; the accession numbers for each genome are given in Materials and Methods. SPVs 1, 3, and 12 are identical to SPV-9, and SPVs 11 and 15 are identical to SPV-4 in the regions shown and were therefore excluded. All differences from consensus nucleotide sequences ("con.") in panels A to C are indicated by a single-letter code; N indicates that there is no consensus. Vertical lines indicate Watson-Crick base pairing, and asterisks indicate GU base pairing.

SPV-9 and would stabilize the long subapical helix. A small number of substitutions were transitions that disrupted base pairing (e.g., A-U→A-C; "disruptive" substitutions). The majority occur at only a few loci: for example, 7 of 10 such substitutions in SPV isolates were the equivalent of SPV-9 $U_{502}G/U_{502}A$ and $G_{503}A$ substitutions in domain $IIIb_1$ (Fig. 4G). A significant proportion of substitutions are covariant; therefore, base pairing in predicted helices is maintained by, e.g., A-U⇔G-C double substitutions. The locations of covariant and neutral substitutions in each predicted IRES structure are shown in Fig. 4B to H. The frequencies of disruptive mutations in all sets of picornavirus sequences were lower, and (except for PTV) the frequencies of covariant substitutions were higher than those in a control set of PEV-8 IRES domain III-derived sequences generated by random evolution in silico. PTV genomes contained an average of only two substitutions, the majority of which were "unpaired." The observed pattern of sequence variation provided strong support for the proposed individual structural elements and for the proposed overall structure of the picornavirus IRESs. More generally,

TABLE 4. Relative frequencies of nucleotide substitutions in the 5′ UTRs of AEV, DHV1, PEV8, PTV, and SPV isolates[a]

Virus	No. of isolates	No. (%) of substitutions in indicated category				
		Total	Unpaired	Covariant	Neutral	Disruptive
AEV	3	14	4 (29)	2 (14)	7 (50)	1 (7%)
DHV-1	8	29	5 (17)	4 (14)	14 (48)	6 (21%)
PEV-8	6	102	61 (60)	22 (22)	14 (14)	5 (5%)
PTV	60	124	73 (59)	6 (5)	41 (33)	4 (3%)
SPV	7	249	44 (18)	181 (73)	14 (6)	10 (4%)
PEV-8 in silico	1,000		13%	3%	13%	71%

[a] The categories of substitutions are defined in the text. Substitution frequencies were determined relative to those for designated type sequences (Fig. 4), from the 5′ border of the predicted domain II structure to the nucleotide preceding the initiation codon, and are compared with those for a control data set derived by in silico evolution of PEV-8 domain III.

this pattern is consistent with conclusions that secondary/tertiary structure within IRES groups is conserved despite sequence variation that can be extensive (16, 21).

Although only one DPV-1 sequence has been reported, we noted that although the core region of its proposed IRES and the corresponding region of the PTV-1 IRES (nt 222 to 242/nt 323 to 405) are only 60% identical, only 2 predicted base pairs in the corresponding predicted structures are absent, whereas 18 base pairs are maintained by sequence covariation. This observation strongly suggests that these two RNAs have nearly identical folds. Similarly, the PEV-8 strain V13 domain III sequence is ~70% identical to members of both principal SPV subgroups, and substitutions at 75% of positions (excluding those for the $IIIb_2$ domain) are covariant substitutions or structurally neutral transitions and 15% occur at unpaired positions. There is therefore significant structural similarity between the >200 nt upstream of the initiation codon in PEV-8 and SPV 5′ UTRs.

Sequence and structural properties of coding sequences adjacent to the initiation codon. The coding region immediately adjacent to the IRES encodes the core protein in HCV and GBV-B and the N-terminal protease (N^{pro}) in pestiviruses. Equivalent regions in the picornavirus genomes encode different proteins. The AEV 5′ UTR is followed immediately by the VP4 and VP2 capsid protein-coding regions (35), and the coding region downstream of the DHV-1 5′ UTR encodes VP0, an uncleaved VP4-VP2 precursor (23). The other picornaviruses all encode VP4 and VP2 proteins preceded by "leader" proteins of differing sizes (Fig. 3); the function of these proteins is unknown.

Among HP IRESs, domain IV occurs only in HCV and GBV-B, and the equivalent regions of pestivirus genomes are unstructured (14, 62, 63). Of the picornavirus sequences analyzed here, only those from SPVs could form stable domain IV structures that in SPV-9 contained five base pairs ($\Delta G = -3.5$ kJ mol^{-1}) (Fig. 4G) and in SPV-3 forms an interrupted 12-base-pair helix ($\Delta G = -11.3$ kJ mol^{-1}) (data not shown). HCV and CSFV IRES activity is enhanced by the proximal 40 nt of coding sequence, including a conserved A-rich sequence (6, 51). This element is preceded by a pyrimidine-rich element in HCV, GBV-B, and all pestiviruses; similar pyrimidine-rich elements are present in all of the picornaviruses considered here, but in some, particularly AEV and SVV, the A-rich element is weak or absent (data not shown).

IRESs from five distinct picornaviruses constitute a third novel group of HP-like IRESs. Analysis of length and sequence variation in the structural core of these IRESs (i.e., the pseudoknot, domain IIIe helix III_1, and the domain IIIf helix) (Fig. 5) indicated that HCV-like IRESs (group A [HCV, GBV-B, AEV and SVV, an outlier]) and pestivirus IRESs (group C) are distinct from the other picornavirus IRESs (group B). This analysis excluded the variable "spacer" between the pseudoknot and the initiation codon, domain II, the nonconserved pestivirus/SVV-specific $IIId_2$ element, and distal sequences in domain III. Domain IIIe is highly conserved in all HP-like IRESs; the adjacent element of helix III_1 has the identical consensus sequence 5′-AGAG/5′-U(A/G)CU in IRES groups A and B, and the proximal half of PK2 has the consensus sequence 5′-GAG/5′CUC in IRES groups A and C. Other core elements differ significantly in terms of length and have distinct consensus sequences; formation of each helix is maintained by covariant nucleotide substitutions, indicative of functional importance. Individual elements in each IRES group also have distinctive lengths. Thus, PK1 is 9 nt long in group A (except in SVV), is 12 nt long in all members of group B, and in group C consists of 7-nt and 6-nt helices linked by an asymmetric bulge. PK2 is 8 nt long in all members of group B but only 6 nt long in groups A (except in SVV [5 nt]) and C (which differ from each other in that group C IRESs have an A-rich linker between PK1 and PK2). These differences in the lengths of stems and connecting loops are maintained within each group, likely because each combination represents a set of compatible elements that together satisfy structural and thermodynamic requirements for stable pseudoknot formation (11).

DISCUSSION

The 5′ UTRs of picornaviruses from several genera contain HP IRES-like elements at different locations relative to the 5′ end of the genome and next to a variety of coding sequences (Fig. 3). These observations and the fact that several of these elements have been found to mediate initiation by an HP IRES-like mechanism rather than by a conventional picornavirus-like mechanism (see below) suggest that HP-like IRESs have on several occasions been exchanged by recombination between members of *Picornaviridae* and *Flaviviridiae*.

Structural models of picornavirus HP-like IRESs. Probabilistic models of seven distinct HP IRES-like structures were supported by thermodynamic considerations and by analysis of additional sequence variants. Sequence differences between the proposed IRESs of different types and isolates of each virus were predominantly structurally neutral transitions or substitutions in unpaired regions and thus had little or no effect on the overall fold or on base pairing in individual helices. Structurally disruptive substitutions were rarer than covariant substitutions, which uniformly served to maintain the proposed pattern of base pairing (Fig. 4B to H). Sequence covariation in noncoding RNAs is indicative of a necessity to maintain functionally important structures (58, 61).

The proposed structures are consistent with and significantly extend previous models for the structures of fragments of these 5′ UTRs, which included SPV domains IIIa, IIIb, and IIId (39), PTV-1 domain IIIe (46), and the pseudoknot region (47). The PTV-1 IRES structure proposed here differs in detail, particularly regarding domain II, but is generally similar to another recent model (3). However, the structures proposed here are incompatible with previous models for the PTV-1 IRES (72) and for elements of DHV, PEV-8, PTV-1, and SPV-9 5′ UTRs (28, 39, 46, 68, 69), possibly because they were derived by analysis of 5′ UTR fragments or by use of algorithms that cannot identify tertiary (pseudoknot) interactions. By contrast, Pfold and ILM, which were used to derive the models described here, were first validated by their accuracy in predicting the secondary and tertiary structures of HCV, CSFV, and BVDV IRES domains, which have previously been independently determined by phylogenetic, biochemical, and biophysical means.

An important corollary to the identification of sequence and structural similarities between these seven picornavirus 5′ UTRs and HP IRESs is proof that the former do in fact mediate internal initiation by an HP-like mechanism rather than by the fundamentally different conventional picornavirus mechanism (44, 47). Recent reports indicate that the PTV-1, PEV-8, and SPV-1 elements described here are IRESs that initiate translation independently of eIF4A (2, 19) and that the PTV-1 IRES mediates initiation by an HP-like mechanism (3, 46).

Implications for the mechanism of IRES-mediated initiation of translation. The IRESs identified here differ from one another significantly, primarily in terms of the size and number of subdomains in domain III and, for some, in the structure of domain II, but they all contain a structurally related core comprising the pseudoknot, domains IIId and IIIe, and helices III_1 and III_2. The high level of sequence and/or structural conservation of this core is indicative of its functional importance, consistent with the results for mutational analyses of HP IRESs (7, 18, 26, 27, 33, 49, 52, 53, 71) and, to a limited extent, the PTV-1 IRES (3). The most variable elements of these picornavirus IRESs are domain II and the apical IIIabc region, which, during initiation, play roles in interacting with the ribosomal E site and stabilizing binding of the coding region of mRNA in the ribosomal mRNA-binding cleft (domain II) (26, 65) and in binding to eIF3 (IIIabc) (45, 62). It is not yet known whether the high level of structural variation in these various IRESs is paralleled by subtle mechanistic differences in mediating internal initiation or whether the tertiary structures of these regions may, despite sequence differences, orient functionally important motifs in similar ways so that critical interactions with components of the translation apparatus can still occur. Alternatively, specific functions played by peripheral IRES domains may be fulfilled in different ways. Differences in domains II and IIIabc of HP and picornavirus HP-like IRESs might be due to the presence of virus-specific *cis*-acting elements that influence replication in addition to translation initiation (8–10, 50). The absence of elements directly equivalent to HCV domain II and/or IIIabc in some picornavirus HP-like IRESs may therefore be due to the locations of replication signals elsewhere in their genomes and a consequent loss of selective pressure to maintain these elements in their IRESs.

Exchange of IRESs between RNA viruses by recombination. PEV-8 and SPVs 1, 3, 4, 9, 11, 12, and 15 are related and may constitute a new picornavirus genus (28, 38, 39), to which DPV may also be assigned. PTV-1 is the type member of the *Teschovirus* genus (74), whereas AEV is a member of the *Hepatovirus* genus (35). The DHV-1 polyprotein is most closely related to polyproteins encoded by members of the *Parechovirus* genus, but DHV-1 may constitute a new genus (23, 69) and the SVV polyprotein is most closely related to *Cardiovirus* polyproteins. Phylogenetic analysis indicates that SPV and PEV-8 (and thus probably DPV) cluster with the *Enterovirus* and *Rhinovirus* genera, teschoviruses (e.g., PTV-1) cluster with the *Cardiovirus*, *Aphthovirus*, and *Erbovirus* genera, and AEV clusters with the *Hepatovirus* genus (15). HP-like IRESs are therefore present in members of distinct and deeply branched picornavirus genera as well as in distinct (*Hepacivirus* and *Pestivirus*) genera of *Flaviviridae*. These observations and the observations that these IRESs are located at various distances from the 5′ end of the genome and are juxtaposed next to a variety of coding sequences (Fig. 3) might be due to repeated capture of a cellular RNA element by different viruses but can more plausibly be accounted for by multiple recombination events between viral genomes. The feasibility of replacing the IRES of a picornavirus by the HCV IRES without loss of viability has been demonstrated using synthetic poliovirus chimeras (30).

Recombination is a common event that contributes significantly to the genetic diversity and evolution of RNA viruses (73) and has been reported for several virus groups, including pestiviruses and picornaviruses (17, 57). Recombination sites occur most frequently in the nonstructural protein-coding region of the picornavirus genome, and although less common, recombination has been reported at the 5′ UTR/VP4 capsid protein junction (31, 41, 57, 60). The preferential recombination of related sequences for yielding viable progeny accounts for observations that recombination between enteroviruses is mostly limited to members of the same species (40, 41). Nonhomologous recombination yields viable recombinants at a 100-fold-lower frequency than homologous recombination (24), but it has nevertheless been detected in picornaviruses and pestiviruses, as is evident by the "capture" of cellular RNA elements (4, 36) and by the apparent exchange of genetic elements between distinct viral species. For example, the coding regions and 3′ noncoding regions of human enteroviruses cluster into four clades (HEV species A to D), whereas their 5′ UTR sequences form only two clades (the poliovirus-like 5′ UTRs of species A and B and the coxsackievirus B-like

5′ UTRs of species C and D) (57). Similarly, most circulating HEV B strains are recombinants in which only the capsid protein region groups with the sequence of the prototype strain, whereas the 5′ UTR and the nonstructural proteins group with each other but not with the capsid proteins (31). Enterovirus capsid proteins were therefore suggested to be independent genetic entities that can combine freely with a variety of 5′ UTRs and nonstructural protein genes from related viruses so that elements encoding structural and nonstructural proteins can evolve relatively independently, accounting for the relatively high genetic variability of capsid proteins (31, 60). The observations reported here can be considered an extension of this model: at a lower frequency, functional noncoding elements may also recombine independently of coding regions. Whereas data derived from phylogenetic analysis of enteroviruses are suggestive of genetic exchange between species, the observations reported here are particularly significant because they constitute one of the few indications that genetic material can be exchanged between virus families. Moreover, unlike previous reports (such as the suggestion that the presence of the hemagglutinin-esterase gene in the positive-strand single-stranded RNA genomes of toroviruses and coronaviruses and the negative-strand single-stranded RNA genome of influenza virus C is due to nonhomologous recombination) (34, 64), data reported here suggest that nonhomologous recombination can also involve the modular exchange of functional noncoding RNA elements between viruses.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant AI-51340.

REFERENCES

1. **Brown, E. A., H. Zhang, L. H. Ping, and S. M. Lemon.** 1992. Secondary structure of the 5′ nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. Nucleic Acids Res. **20:**5041–5045.
2. **Chard, L. S., M. E. Bordeleau, J. Pelletier, J. Tanaka, and G. J. Belsham.** 2006. Hepatitis C virus-related internal ribosome entry sites are found in multiple genera of the family Picornaviridae. J. Gen. Virol. **87:**927–936.
3. **Chard, L. S., Y. Kaku, B. Jones, A. Naya, and G. J. Belsham.** 2006. Functional analyses of RNA structures shared between the internal ribosome entry sites of hepatitis C virus and the picornavirus porcine teschovirus 1 Talfan. J. Virol. **80:**1271–1279.
4. **Charini, W. A., S. Todd, G. A. Gutman, and B. L. Semler.** 1994. Transduction of a human RNA sequence by poliovirus. J. Virol. **68:**6547–6552.
5. **Dowell, R. D., and S. R. Eddy.** 2004. Evaluation of several lightweight stochastic context free grammars for RNA secondary structure prediction. BMC Bioinformatics **5:**71–84.
6. **Fletcher, S. P., I. K. Ali, A. Kaminski, P. Digard, and R. J. Jackson.** 2002. The influence of viral coding sequences on pestivirus IRES activity reveals further parallels with translation initiation in prokaryotes. RNA **8:**1558–1571.
7. **Fletcher, S. P., and R. J. Jackson.** 2002. Pestivirus internal ribosome entry site (IRES) structure and function: elements in the 5′ untranslated region important for IRES function. J. Virol. **76:**5024–5033.
8. **Friebe, P., V. Lohmann, N. Krieger, and R. Bartenschlager.** 2001. Sequences in the 5′ nontranslated region of hepatitis C virus required for RNA replication. J. Virol. **75:**12047–12057.
9. **Frolov, I., M. S. McBride, and C. M. Rice.** 1998. cis-Acting RNA elements required for replication of bovine viral diarrhea virus-hepatitis C virus 5′ nontranslated region chimeras. RNA **4:**1418–1435.
10. **Grassmann, C. W., H. Yu, O. Isken, and S. E. Behrens.** 2005. Hepatitis C virus and the related bovine viral diarrhea virus considerably differ in the functional organization of the 5′ non-translated region: implications for the viral life cycle. Virology **333:**349–366.
11. **Gultyaev, A. P., F. H. van Batenburg, and C. W. Pleij.** 1999. An approximation of loop free energy values of RNA H-pseudoknots. RNA **5:**609–617.
12. **Hatakeyama, Y., N. Shibuya, T. Nishiyama, and N. Nakashima.** 2004. Structural variant of the intergenic internal ribosome entry site elements in dicistroviruses and computational search for their counterparts. RNA **10:**779–786.
13. **Hofacker, I. L.** 2003. Vienna RNA secondary structure server. Nucleic Acids Res. **31:**3429–3431.
14. **Honda, M., E. A. Brown, and S. M. Lemon.** 1996. Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. RNA **2:**955–968.
15. **Hughes, A. L.** 2004. Phylogeny of the Picornaviridae and differential evolutionary divergence of picornavirus proteins. Infect. Genet. Evol. **4:**143–152.
16. **Jackson, R. J., and A. Kaminski.** 1995. Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. RNA **1:**985–1000.
17. **Jones, L. R., and E. L. Weber.** 2004. Homologous recombination in bovine pestiviruses. Phylogenetic and statistic evidence. Infect. Genet. Evol. **4:**335–343.
18. **Jubin, J., N. E. Vantuno, J. S. Kieft, M. G. Murray, J. A. Doudna, J. Y. Lau, and B. M. Baroudy.** 2000. Hepatitis C virus internal ribosome entry site (IRES) stem loop IIId contains a phylogenetically conserved GGG triplet essential for translation and IRES folding. J. Virol. **74:**10430–10437.
19. **Kaku, Y., L. S. Chard, T. Inoue, and G. J. Belsham.** 2002. Unique characteristics of a picornavirus internal ribosome entry site from the porcine teschovirus-1 Talfan. J. Virol. **76:**11721–11728.
20. **Kaku, Y., A. Sarai, and Y. Murakami.** 2001. Genetic reclassification of porcine enteroviruses. J. Gen. Virol. **82:**417–424.
21. **Kanamori, Y., and N. Nakashima.** 2001. A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. RNA **7:**266–274.
22. **Kieft, J. S., K. Zhou, R. Jubin, and J. A. Doudna.** 2001. Mechanism of ribosome recruitment by hepatitis C IRES RNA. RNA **7:**194–206.
23. **Kim, M.-C., Y.-K. Kwon, S.-J. Joh, A. M. Lindberg, J.-H. Kwon, J.-H. Kim, and S.-J. Kim.** 2006. Molecular analysis of duck hepatitis virus type 1 reveals a novel lineage close to the genus *Parechovirus* in the family *Picornaviridae*. J. Gen. Virol. **87:**3307–3316.
24. **Kirkegaard, K., and D. Baltimore.** 1986. The mechanism of RNA recombination in poliovirus. Cell **47:**433–443.
25. **Knudsen, B., and J. Hein.** 2003. Pfold: RNA secondary structure prediction using stochastic context-free grammars. Nucleic Acids Res. **31:**3423–3428.
26. **Kolupaeva, V. G., T. V. Pestova, and C. U. T. Hellen.** 2000. An enzymatic footprinting analysis of the interaction of 40S ribosomal subunits with the internal ribosomal entry site of hepatitis C virus. J. Virol. **74:**6242–6250.
27. **Kolupaeva, V. G., T. V. Pestova, and C. U. T. Hellen.** 2000. Ribosomal binding to the internal ribosomal entry site of classical swine fever virus. RNA **6:**1791–1807.
28. **Krumbholz, A., M. Dauber, A. Henke, E. Birch-Hirschfeld, N. J. Knowles, A. Stelzner, and R. Zell.** 2002. Sequencing of porcine enterovirus groups II and III reveals unique features of both virus groups. J. Virol. **76:**5813–5821.
29. **Locker, N., L. E. Easton, and P. J. Lukavsky.** 2007. HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. EMBO J. **26:**795–805.
30. **Lu, H. H., and E. Wimmer.** 1996. 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. USA **93:**1412–1417.
31. **Lukashev, A. N., V. A. Lashkevich, O. E. Ivanova, G. A. Koroleva, A. E. Hinkkanen, and J. Ilonen.** 2005. Recombination in circulating human enterovirus B: independent evolution of structural and non-structural genome regions. J. Gen. Virol. **86:**3281–3290.
32. **Lukavsky, P. J., I. Kim, G. A. Otto, and J. D. Puglisi.** 2003. Structure of HCV IRES domain II determined by NMR. Nat. Struct. Biol. **10:**1033–1038.
33. **Lukavsky, P. J., G. A. Otto, A. M. Lancaster, P. Sarnow, and J. D. Puglisi.** 2000. Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. Nat. Struct. Biol. **7:**1105–1110.
34. **Luytjes, W., P. J. Bredenbeek, A. F. Noten, M. C. Horzinek, and W. J. Spaan.** 1988. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. Virology **166:**415–422.
35. **Marvil, P., N. J. Knowles, A. P. Mockett, P. Britton, T. D. Brown, and D. Cavanagh.** 1999. Avian encephalomyelitis virus is a picornavirus and is most closely related to hepatitis A virus. J. Gen. Virol. **80:**653–662.
36. **Meyers, G., N. Tautz, E. J. Dubovi, and H. J. Thiel.** 1991. Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. Virology **180:**602–616.
37. **Notredame, C., D. Higgins, and J. Heringa.** 2000. T-Coffee: a novel method for multiple sequence alignments. J. Mol. Biol. **302:**205–217.
38. **Oberste, M. S., K. Maher, and M. A. Pallansch.** 2002. Molecular phylogeny and proposed classification of the simian picornaviruses. J. Virol. **76:**1244–1251.
39. **Oberste, M. S., K. Maher, and M. A. Pallansch.** 2003. Genomic evidence that simian virus 2 and six other simian picornaviruses represent a new genus in Picornaviridae. Virology **314:**283–293.
40. **Oberste, M. S., K. Maher, and M. A. Pallansch.** 2004. Evidence for frequent recombination within species *Human enterovirus B* based on complete genomic sequences of all thirty-seven serotypes. J. Virol. **78:**855–867.
41. **Oberste, M. S., S. Penaranda, and M. A. Pallansch.** 2004. RNA recombina-

tion plays a major role in genomic change during circulation of coxsackie B viruses. J. Virol. **78:**2948–2955.
42. **Pestova, T. V., and C. U. T. Hellen.** 1999. Internal initiation of translation of bovine viral diarrhea virus RNA. Virology **258:**249–256.
43. **Pestova, T. V., and C. U. T. Hellen.** 2003. Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. Genes Dev. **17:**181–186.
44. **Pestova, T. V., V. G. Kolupaeva, I. B. Lomakin, E. V. Pilipenko, I. N. Shatsky, V. I. Agol, and C. U. T. Hellen.** 2001. Molecular mechanisms of translation initiation in eukaryotes. Proc. Natl. Acad. Sci. USA **98:**7029–7036.
45. **Pestova, T. V., I. N. Shatsky, S. P. Fletcher, R. J. Jackson, and C. U. T. Hellen.** 1998. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. **12:**67–83.
46. **Pisarev, A. V., L. S. Chard, Y. Kaku, H. L. Johns, I. N. Shatsky, and G. J. Belsham.** 2004. Functional and structural similarities between the internal ribosome entry sites of hepatitis C virus and porcine teschovirus, a picornavirus. J. Virol. **78:**4487–4497.
47. **Pisarev, A. V., N. E. Shirokikh, and C. U. T. Hellen.** 2005. Translation initiation by factor independent binding of eukaryotic ribosomes to internal ribosomal entry sites. C. R. Biol. **328:**589–605.
48. **Poole, T. L., C. Wang, R. A. Popp, L. N. Potgieter, A. Siddiqui, and M. S. Collett.** 1995. Pestivirus translation initiation occurs by internal ribosome entry. Virology **206:**750–754.
49. **Psaridi, L., U. Georgopoulou, A. Varaklioti, and P. Mavromara.** 1999. Mutational analysis of a conserved tetraloop in the 5′ untranslated region of hepatitis C virus identifies a novel RNA element essential for the internal ribosome entry site function. FEBS Lett. **453:**49–53.
50. **Reusken, C. B., T. J. Dalebout, P. Eerligh, P. J. Bredenbeek, and W. J. Spaan.** 2003. Analysis of hepatitis C virus/classical swine fever virus chimeric 5′NTRs: sequences within the hepatitis C virus IRES are required for viral RNA replication. J. Gen. Virol. **84:**1761–1769.
51. **Reynolds, J. E., A. Kaminski, H. J. Kettinen, K. Grace, B. E. Clarke, A. R. Carroll, D. J. Rowlands, and R. J. Jackson.** 1995. Unique features of internal initiation of hepatitis C virus RNA translation. EMBO J. **14:**6010–6020.
52. **Rijnbrand, R., G. Abell, and S. M. Lemon.** 2000. Mutational analysis of the GB virus B internal ribosome entry site. J. Virol. **74:**773–783.
53. **Rijnbrand, R., T. van der Straaten, P. A. van Rijn, W. J. Spaan, and P. J. Bredenbeek.** 1997. Internal entry of ribosomes is directed by the 5′ noncoding region of classical swine fever virus and is dependent on the presence of an RNA pseudoknot upstream of the initiation codon. J. Virol. **71:**451–457.
54. **Rijnbrand, R. C., and S. M. Lemon.** 2000. Internal ribosome entry site-mediated translation in hepatitis C virus replication. Curr. Top. Microbiol. Immunol. **242:**85–116.
55. **Ruan, J., G. D. Stormo, and W. Zhang.** 2004. An iterated loop matching approach to the prediction of RNA secondary structures with pseudoknots. Bioinformatics **20:**58–66.
56. **Ruan, J., G. D. Stormo, and W. Zhang.** 2004. ILM: a web server for predicting RNA secondary structures with pseudoknots. Nucleic Acids Res. **32:**W146–W149.
57. **Santti, J., T. Hyypia, L. Kinnunen, and M. Salminen.** 1999. Evidence of recombination among enteroviruses. J. Virol. **73:**8741–8749.
58. **Shapiro, B., A. Rambaut, O. G. Pybus, and E. C. Holmes.** 2006. A phylogenetic method for detecting positive epistasis in gene sequences and its application to RNA virus evolution. Mol. Biol. Evol. **23:**1724–1730.
59. **Simmonds, P.** 2004. Genetic diversity and evolution of hepatitis C virus—15 years on. J. Gen. Virol. **85:**3173–3188.
60. **Simmonds, P.** 2006. Recombination and selection in the evolution of picornaviruses and other mammalian positive-stranded RNA viruses. J. Virol. **80:**11124–11140.
61. **Simmonds, P., and D. B. Smith.** 1999. Structural constraints on RNA virus evolution. J. Virol. **73:**5787–5794.
62. **Sizova, D. V., V. G. Kolupaeva, T. V. Pestova, I. N. Shatsky, and C. U. T. Hellen.** 1998. Specific interaction of eukaryotic translation initiation factor 3 with the 5′ nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. J. Virol. **72:**4775–4782.
63. **Smith, D. B., J. Mellor, L. M. Jarvis, F. Davidson, J. Kolberg, M. Urdea, P. L. Yap, P. Simmonds, et al.** 1995. Variation of the hepatitis C virus 5′ non-coding region: implications for secondary structure, virus detection and typing. J. Gen. Virol. **76:**1749–1761.
64. **Snijder, E. J., J. A. den Boon, M. C. Horzinek, and W. J. Spaan.** 1991. Comparison of the genome organization of toro- and coronaviruses: evidence for two nonhomologous RNA recombination events during Berne virus evolution. Virology **180:**448–452.
65. **Spahn, C. M. T., J. S. Kieft, R. A. Grassucci, P. A. Penczek, K. Zhou, J. A. Doudna, and J. Frank.** 2001. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. Science **291:**1959–1962.
66. **Stoye, J., D. Evers, and F. Meyer.** 1998. Rose: generating sequence families. Bioinformatics **14:**157–163.
67. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
68. **Thurner, C., C. Witwer, I. L. Hofacker, and P. F. Stadler.** 2004. Conserved RNA secondary structures in Flaviviridae genomes. J. Gen. Virol. **85:**1113–1124.
69. **Tseng, C. H., N. J. Knowles, and H. J. Tsai.** 2007. Molecular analysis of duck hepatitis virus type 1 indicates that it should be assigned to a new genus. Virus Res. **123:**190–203.
70. **Tsukiyama-Kohara, K., N. Iizuka, M. Kohara, and A. Nomoto.** 1992. Internal ribosome entry site within hepatitis C virus RNA. J. Virol. **66:**1476–1483.
71. **Wang, C., S. Y. Le, N. Ali, and A. Siddiqui.** 1995. An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5′ noncoding region. RNA **1:**526–537.
72. **Witwer, C., S. Rauscher, I. L. Hofacker, and P. F. Stadler.** 2001. Conserved RNA secondary structures in Picornaviridae genomes. Nucleic Acids Res. **29:**5079–5089.
73. **Worobey, M., and E. C. Holmes.** 1999. Evolutionary aspects of recombination in RNA viruses. J. Gen. Virol. **80:**2535–2543.
74. **Zell, R., M. Dauber, A. Krumbholz, A. Henke, E. Birch-Hirschfeld, A. Stelzner, D. Prager, and R. Wurm.** 2001. Porcine teschoviruses comprise at least eleven distinct serotypes: molecular and evolutionary aspects. J. Virol. **75:**1620–1631.
75. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.