Review

Viral IRES RNA structures and ribosome interactions

Jeffrey S. Kieft

Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Mail stop 8101, PO Box 6511, Aurora, CO 80045, USA

In eukaryotes, protein synthesis initiates primarily by a mechanism that requires a modified nucleotide 'cap' on the mRNA and also proteins that recruit and position the ribosome. Many pathogenic viruses use an alternative, cap-independent mechanism that substitutes RNA structure for the cap and many proteins. The RNAs driving this process are called internal ribosome-entry sites (IRESs) and some are able to bind the ribosome directly using a specific 3D RNA structure. Recent structures of IRES RNAs and IRES–ribosome complexes are revealing the structural basis of viral IRES' 'hijacking' of the protein-making machinery. It now seems that there are fundamental differences in the 3D structures used by different IRESs, although there are some common features in how they interact with ribosomes.

Cap-dependent and cap-independent translation in eukaryotes

Eukaryotic translation initiation is a tightly regulated process. The canonical mechanism begins with the recognition of the 5′ modified nucleotide cap on the mRNA, it requires many initiation factor (eIF) proteins, uses ribosome scanning (movement of the ribosomal subunit in a 5′ to 3′ direction along the mRNA) to find the start codon and culminates with an 80S ribosome located on the mRNA, ready to begin protein synthesis (reviewed recently in [1]). The 80S ribosome contains a large (50S in bacteria, 60S in eukaryotes) and a small (30S in bacteria, 40S in eukaryotes) ribosomal subunit (Figure 1). During translation on the 80S ribosome, the RNA moves through the decoding groove on the small subunit where it is 'read' by tRNAs that move through the space between the two subunits. There are three tRNA-binding sites on the ribosome, designated as the aminoacyl-site (A-site), the peptidyl- site (P-site) and the exit-site (E-site) (Figure 1). In canonical translation initiation, the initiator tRNA is in the P-site and addition of a tRNA to the A-site leads to peptide-bond formation and translocation. Overall, it is important to remember that the canonical pathway of translation initiation begins when the 5′ cap is recognized.

Although the canonical cap- and scanning-dependent mechanism accounts for protein synthesis on the majority of eukaryotic mRNAs, it cannot account for translation initiation on all messages, including certain viral RNAs. For example, poliovirus (PV) is a single-stranded positive-sense RNA virus with no 5′-cap but with a genome-linked peptide (VPg) on its 5′ end, hence, the canonical signal for ribosome recruitment is not present. The mystery of precisely how translation initiates on this important human pathogen's RNA was solved in 1988, when it was reported that both PV and encephalomyocarditis virus (EMCV) RNA use a cap-independent, end-independent, internal-translation initiation mechanism [2,3] (Box 1). Rather than being dependent on the 5′ cap and the proteins that interact with the cap, this mechanism is driven by specific RNA sequences that are found in the untranslated region (UTR) of the viral RNA, called internal ribosome-entry sites (IRESs) (reviewed in [4,5]). These IRESs can functionally replace both the cap and also many (in some cases, all) of the proteins needed to recruit the ribosome to the start codon in a process that is RNA dependent. Circularized IRES-containing mRNAs can be translated, demonstrating the robustness of IRES function [6]. Since their discovery, IRESs have been identified in at least 39 viral RNAs and 85 cellular mRNAs [7], making this RNA-driven mechanism an important part of the repertoire of translation initiation and regulation in eukaryotes. Furthermore, it is interesting to consider where IRES-driven translation fits into the evolutionary development of translation initiation; are IRESs molecular fossils representing an ancient process that preceded the cap-dependent pathway [8]?

The importance of IRESs in eukaryotic translation initiation and in other pathogenic viruses, such as hepatitis A virus (HAV) [9], hepatitis C virus (HCV) [10], foot-and-mouth-disease virus (FMDV) [11], HIV [12,13] and many others, has led to a significant effort to understand IRES function. Because IRES-driven translation is RNA-driven, the structure of the IRES RNA is a primary focus of these studies. The picture that has emerged is one of structural and mechanistic diversity; as a result, viral IRESs are placed currently into one of four groups (Box 1).

Of these groups, two include IRESs that can bind directly to the ribosome in the absence of any other co-factors, such as host-cell initiation factors and this presents an interesting question: what are the structures that confer this ability and how do they work? In fact, although a variety of probing, biochemical and biophysical methods have been applied to IRESs from all groups, the only IRESs for which structures have been solved are those that bind directly to the ribosome. Specifically, cryo-electron microscopy, X-ray crystallography and NMR spectroscopy have been used to analyze these IRES RNAs, both in the free

Corresponding author: Kieft, J.S. (Jeffrey.kieft@uchsc.edu).

0968-0004/$ – see front matter doi:10.1016/j.tibs.2008.04.007

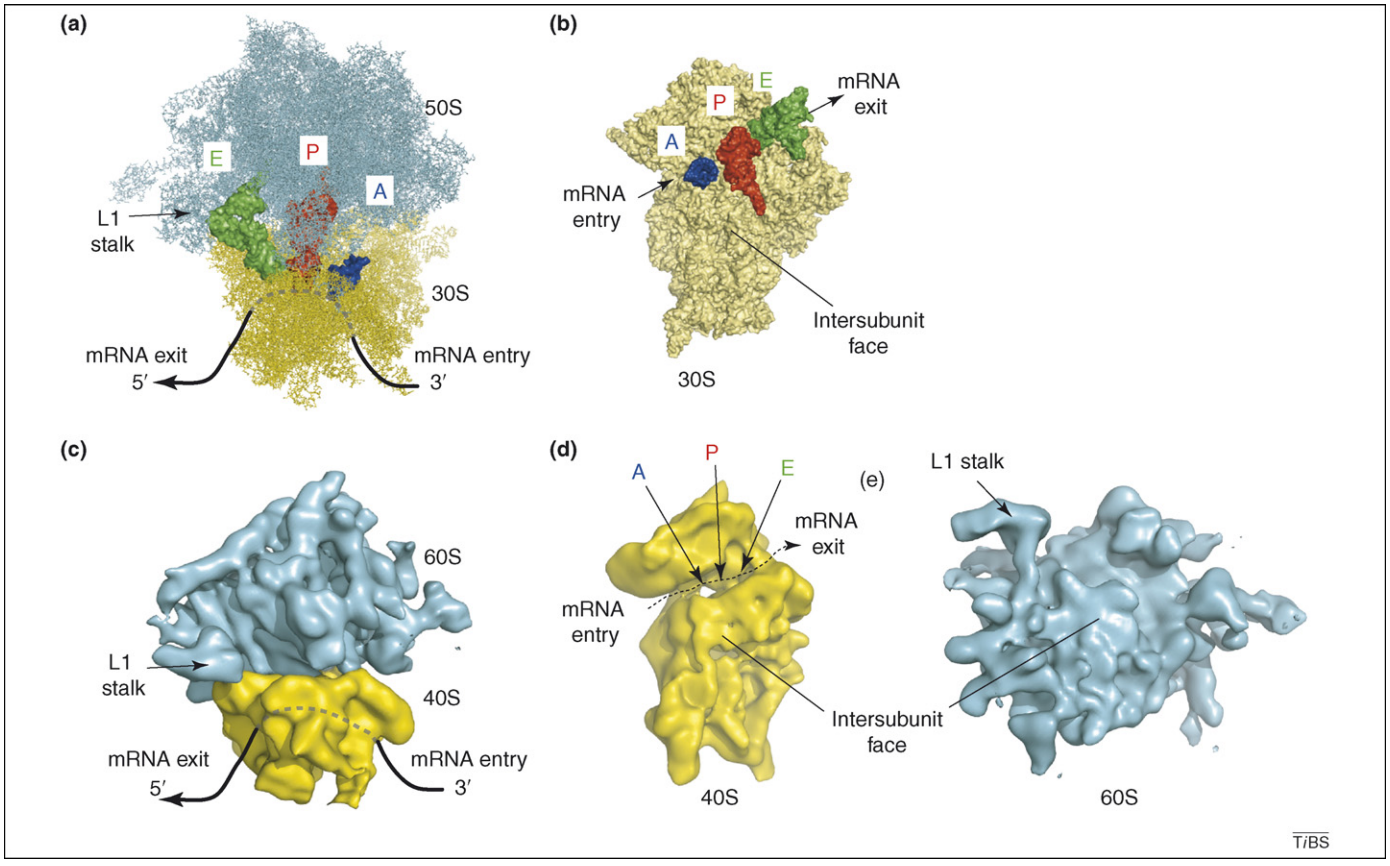

Figure 1. Structural features of ribosomes and ribosome-containing complexes. **(a)** The crystal structure of the bacterial 70S ribosome from *Thermus thermophilus* with three tRNAs bound is shown with the 50S (large) subunit in cyan and the 30S (small) subunit in yellow [58]. This view is sometimes called a 'top' view. The A-, P- and E-site tRNAs are labeled and shown in blue, red and green, respectively. Only a portion of the A-site tRNA is shown because the rest of this tRNA is not visible in the crystal structure. The approximate pathway of mRNA through the decoding groove on the 30S subunit and out of the exit tunnel is shown with a black line and the L1 stalk of the 50S subunit is indicated. **(b)** Structure of the 30S subunit from the same structure shown in (a) and using the same coloring scheme but turned to reveal the inter-subunit face of the 30S subunit and the positions of the three tRNAs along the decoding groove. The mRNA entry and exit positions are shown. **(c)** Cryo-EM reconstruction of a human 80S ribosome in the same orientation as the bacterial ribosome in (a). This reconstruction is from a ribosome bound to the CrPV IRES [29]; here, the IRES density has been removed computationally. The 60S subunit is cyan, the 40S subunit is yellow, the path of the mRNA is shown with a black line and the L1 stalk is indicated. **(d)** Cryo-EM reconstruction of the 40S subunit portion of the ribosome in (c), rotated to show the intersubunit face. The A-, P- and E-sites in the decoding groove are indicated and the approximate pathway of mRNA is shown with a broken line. **(e)** 60S subunit of the cryo-EM reconstruction shown in (c), rotated to show the intersubunit face and the location of the L1 stalk.

form and in ribosome-containing pre-initiation complexes. The complementary data obtained from these structural methods, combined with biochemical and functional results, led to new testable mechanistic hypotheses and models for IRES action. This review focuses on recent advances in understanding the 3D structures of ribosome-binding viral IRES RNAs and their interactions with the eukaryotic translation machinery.

Structures of Dicistroviridae IGR IRESs in the free and bound forms

The first IRES RNA structures solved were of individual stem-loops of the HCV IRES (member of group 2) and it is only in the last 2 years that structures of domains from group 1 IRESs were solved. However, because the group 1 IRESs use the simplest form of cap-independent initiation yet discovered (in terms of co-factor requirement) and because the structural data available for this group are most complete, I will discuss group 1 IRESs first.

Crystal structures of unbound group 1 IRES RNAs

The first IRES group is found in the single-stranded positive-sense Dicistroviridae viruses, specifically in the intergenic region (IGR) between two open-reading frames in the viral RNA (reviewed in [14]). These IRESs initiate translation in a highly atypical fashion. First, initiation is from a non-AUG start codon; second, it occurs from the ribosome's aminoacyl-tRNA-site A-site rather than the P-site, as is the case for all other initiation mechanisms; third, it occurs without the involvement of any initiation factors or initiator methionyl tRNA in a process that drives a ribosome-translocation event before a peptide bond has been made (which is termed 'pseudotranslocation') [15–21]. Thus, these IRESs functionally replace many protein factors in addition to initiator tRNA, leading to the hypothesis that these IRESs use specific structures to bind and manipulate the ribosome.

Indeed, RNA structure is important. These IRESs have an intricate secondary and tertiary structure with three essential pseudoknots (PKI, II and III) and two conserved stem-loops (SL IV and V) [22] that make direct contact with the 40S subunit [23,24] (Figure 2a,b). Not only do the Dicistroviridae IGR IRESs (hereafter referred to as the 'IGR IRESs') have an intricate base-pairing scheme but they have tightly folded compact structures with two independently folded domains [25] and, to date, these are the

Box 1. Cap-dependent translation initiation and four IRES groups

The top of Figure I shows a simplified diagram of the cap-dependent ribosome-recruitment process, in which several factors (including tRNA, eIF1 and eIF1A) are not shown for simplicity. Defining characteristics of this pathway are the need for the 5′ cap (m7G) and for a full complement of initiation factors (eIFs). IRESs bypass the need for the cap and some or all of the factors, although they differ from one another in many ways. Based on their need for factors, the proposed secondary structure of the IRESs, the location of the start codon relative to the IRES and the ability of the IRES to operate in rabbit reticulocyte extract with or without supplementation, the viral IRESs are divided into four groups (Figure I):

- Group 1: IRES RNAs that bind directly to the ribosome and do not require any protein factors to do so. These IRESs also do not require initiator methionyl-tRNA$_i$. Members of this group include Cricket paralysis virus (CrPV) [18], Plautia stali intestine virus (PSIV) [20] and Taura syndrome virus (TSV) [59,60].
- Group 2: IRES RNAs that bind to the 40S subunit and also use a subset of the canonical eIFs (eIF3, eIF2) as well as Met-tRNA$_i$. Members of this group include HCV [10], classical swine fever virus (CSFV, a pestivirus) [61] and porcine teschovirus 1 (PTV-1, a picornavirus) [62].
- Group 3: IRES RNAs that require some canonical eIFs, Met-tRNA$_i$ and additional proteins called IRES *trans*-activating factors (ITAFs). These IRESs function in rabbit reticulocyte lysate (RRL) and translation starts at the 3′ end of the IRES. Members include EMCV [2], FMDV [11] and Theiler's Murine Encephalomyelitis virus (TMEV) [63].
- Group 4: IRES RNAs that require some canonical eIFs, Met-tRNA$_i$ and ITAFs and function efficiently in RRL only when it is supplemented with extracts from other cell types and also initiate at an AUG codon somewhat downstream of the IRES. Members include PV [3] and rhinovirus [64].

There is no evidence of direct IRES RNA–ribosome interactions in the last two groups (except for the start codon, by necessity). Hence, they might operate by binding initiation factors and/or ITAFs that then interact with the ribosome and might best be regarded as IRES ribonuclear protein complexes (RNPs).

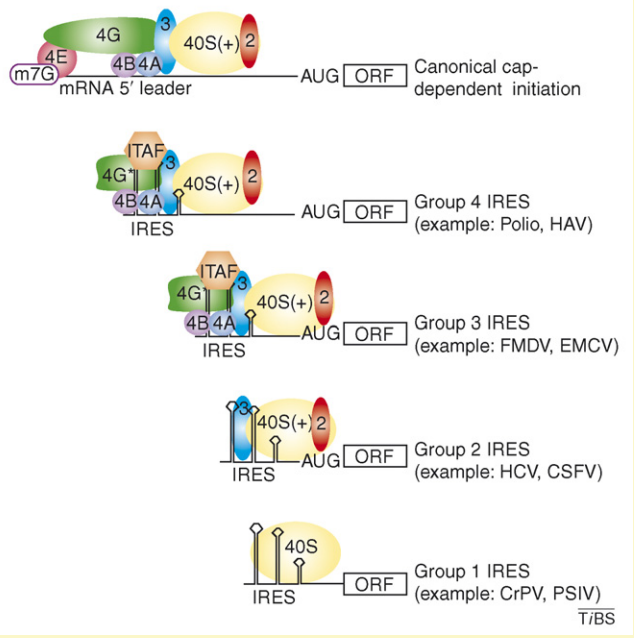

Figure I. Comparison of cap-dependent translation and four IRES groups.

only IRESs known to have this type of tightly folded, globular structure. Note that I use the term 'region' for a collection of logically connected secondary structure elements but reserve the term 'domain' for entities shown to fold into an independent structure experimentally. Hence, the larger domain contains two smaller regions (regions 1 and 2) that fold together and this folded domain exhibits affinity for the ribosome [23–25]. The smaller domain (domain 3) comprises a pseudoknot that docks into the P-site of the ribosome's mRNA-decoding groove (Figure 1), helping to drive initiation from the A-site [16,17]. Because compact folded RNAs are amenable to crystallization, X-ray crystallography was used successfully to solve the structures of each domain separately (Figure 2c,d).

The recent crystal structure of the ribosome-binding domain (regions 1 + 2) of the Plautia stali intestine virus (PSIV) IGR IRES indicates that the elements of the IRES that make specific interactions with the ribosome are largely pre-positioned by the fold before ribosome binding [26] (Figure 2c). Evidence for this is that, within the structure, the two stem-loops that make direct contact to the 40S subunit (SL IV and SL V) lie adjacent to each other and extend away from the folded IRES. This occurs within a stable structure that is held together by specific intramolecular interactions involving highly conserved bases. This fold centers on an underwound helix (termed P2.2) that interacts with multiple parts of the IRES RNA, seemingly serving as the core element that stabilizes the structure of the ribosome-binding domain (Figure 2c). Interestingly, some parts of the IRES are dynamic or disordered in the crystal and are therefore partially or totally invisible within the structure. Specifically, the disordered parts are a conserved loop (L1.1) and parts of a large loop (L1.2) that are adjacent to a functionally crucial pseudoknot (PK II). The fact that these structures are dynamic in the crystal does not prove that they are dynamic in solution, although it might indicate that these IRES elements change structure in response to interactions with the ribosome and that this could be important for the overall IRES mechanism [26,27].

The other domain of the IGR IRESs is the smaller domain 3 or 'P-site domain'. The recent crystal structure of domain 3 from the Cricket Paralysis Virus (CrPV) IGR IRES reveals that this single RNA domain mimics precisely the intermolecular interaction between the initiator tRNA anticodon loop and the AUG start codon within the P-site of the ribosome [28] (Figure 2d). This mimicry was proposed as a result of previous functional experiments and now the structure provides confirmation [15,16]. The mimicry indicates that domain 3 forms intermolecular interactions that precisely mirror those normally made among the tRNA anti-codon, mRNA codon and the ribosome. Because IGR IRES-driven translation starts from the A-site at the codon immediately downstream of this P-site-docked structure, it seems probable that this mimicry places that codon precisely into the A-site. This structure is the first demonstration of direct molecular mimicry in a viral IRES, raising the question of whether other IRESs also use structural mimicry. Indeed, it has been suggested that the ability of RNase P to specifically cleave RNA near

Box 2. Do IRESs mimic tRNA structure?

Structural mimicry of tRNAs is observed in several different contexts, including transfer-mRNA (tmRNA, a noncoding bacterial RNA that rescues stalled ribosomes) [65], tRNA-like sequences at the 3′ end of some viral genomes [66] and protein factors involved in translation [67]. Evidence indicating tRNA structural mimicry in IRESs came from RNase P cleavage assays in which the enzyme purportedly cleaves IRES RNAs 3′ of putative tRNA-like structures [68–70]. However, because RNase P cleaves authentic pre-tRNAs 5′ of the tRNA structure, it would be unusual for RNase P to cleave 3′ of a tRNA-like structure (reviewed in [71]). In fact, RNase P cleaves a large number of RNA substrates, including the precursor of 4.5S RNA, bacteriophage M3 RNA, mRNA, riboswitches and even small single- and double-stranded RNAs, some with no structural relationship to tRNA (reviewed in [67]). Because of the broad specificity of RNase P, one must be cautious in interpreting RNase P cleavage as indicative of tRNA-like structure. X-ray and NMR structures show no tRNA-like structures in the ribosome-binding domain of the PSIV IRES or any HCV IRES elements, despite the fact that they are cleaved by RNase P. To date, there is direct structural evidence for tRNA mimicry only in domain 3 of the CrPV IRES, which mimics the anticodon loop–codon interaction [28]. However, RNase P recognizes the T-loop and stem of tRNA and not the anticodon loop and hence this structure would not be recognized by the enzyme in the same mode as tRNA. Despite a paucity of direct evidence for IRES tRNA structural mimicry, there is evidence that the Dicistroviridae IGR IRESs might act in part by mimicking the overall position and ribosomal contacts made between a tRNA and ribosome. The overall placement of the IGR IRES on the ribosome is reminiscent of a P/E hybrid state tRNA and this has led to the proposal that hybrid-state mimicry can help to explain translocation without peptide-bond formation. Evidence for this idea is found in the behavior of the PSIV IRES and tRNA in the presence of elongation factors, the conformation of the ribosome and the location of the IRES on the ribosome [28,33].

many IRES sequences indicates tRNA-like structures in many IRESs. However, the broad specificity and characteristics of RNase-P cleavage cast doubt on this interpretation (Box 2). When the crystal structures of the two independently folded domains of the IGR IRESs are combined, they yield the first complete high-resolution structural picture of an IRES RNA (Figure 2e) that guide future functional experiments aimed at a deeper understanding of how the structures of the group 1 IRES drive function.

Cryo-EM of IGR IRES–ribosome complexes

For the IGR IRESs, cryo-EM structures of both the human 40S ribosomal subunit–CrPV IRES complex (at ~20 Å resolution) and 80S ribosome–CrPV IRES complexes (at ~17.3 Å for a human ribosome and ~7.3 Å for a yeast ribosome) have been determined, lending insight into ribosome–IRES interactions [27,29] (Figure 2f). The IRES binds to the ribosome as a two-domain structure lying over the decoding groove, makes contacts with both ribosomal subunits and interacts with ribosome features known to be associated with tRNA binding. This location agrees with directed probing experiments that were used to map the position of the IRES on the 40S subunit [30]. That the CrPV IRES binds as a compact two-domain structure has significance for understanding the mechanism by which the IRES works. Specifically, this observation confirms that the IRES pre-folds before interacting with the ribosome and indicates that adoption of an evolutionary conserved fold is crucial for function. Indeed, biochemical and biophysical examination of different members of the IGR IRES family show a similar 3D fold despite substantial sequence variation [31]. In addition, the overall shape of the CrPV IRES density and the PSIV IRES crystal structure are similar, indicating that there are only local changes in IRES structure when it binds the 40S subunit (Figure 2e).

The quality of the recent 7.3 Å cryo-EM structure of the yeast 80S ribosome–CrPV IRES complex was such that a model of the IRES could be built. This model, and the docked crystal structure of the PSIV IRES ribosome-binding domain, reveal specific contacts between the IRES RNA and the ribosome [29] (Figure 2e). Specifically, SL V and SL IV of the IRES are positioned to contact the small ribosomal subunit protein (rp) S5 and an unidentified protein rpSX, dynamic loop L1.1 of the IRES contacts the large ribosomal subunit's L1 stalk (which contains rpL1 and the 28S rRNA helices, H76-H78) and L1.2 and probably also PK II contact rpL11. In addition, crosslinking experiments show that rpS25 crosslinks to SL IV of the PSIV IRES, although the precise location of rpS25 within the small subunit is not known [32]. A full description of the roles of these ribosomal proteins and what their interactions with the IRES might mean for IRES function remains unclear and speculation is beyond the scope of this review. Nonetheless, one interesting preliminary observation is that the ribosome-binding domain of the CrPV IRES interacts with ribosome components that contact E-site tRNA (namely, rpS5 and the L1 stalk) and this hints at the mechanism by which these group 1 IRESs might function.

Whereas the larger ribosome-binding domain interacts with the E-site, the smaller domain 3 extends over the A- and P-sites on the 40S subunit [17] (Figure 2f). Within the 40S subunit's decoding groove, domain 3 is positioned to contact rRNA helix h18 and h34 of the 18S rRNA [29] (Figure 2e) (note that lower case 'h' signifies a rRNA helix on the small subunit, uppercase 'H' on the large subunit). However, the density of domain 3 is less well defined than most of the ribosome-binding domain density, indicating domain 3 is mobile or structurally dynamic [27,29]. Indeed, probing of the free RNA showed that the pseudoknot interaction (PK I) within domain 3 is dynamic [23–25] and probing of the RNA within the context of the 40S subunit and 80S ribosome showed local changes in the domain 3 structure as the ribosome assembles on the IRES [28]. Perhaps domain 3 is not fully docked into position until the A-site aminoacyl-tRNA is delivered [33], however, once docked it mimics the P-site tRNA–mRNA interaction precisely [28] (Figure 2d). That domain 3 contacts the P-site and features of the E-site simultaneously indicates that this IRES might mimic a P/E hybrid state tRNA [28,33]. Hybrid-state tRNAs occur during the normal elongation cycle; following peptide-bond formation, the tRNA located in the P-site enters a hybrid state, in which it essentially occupies the P-site on the small subunit and part of the E-site on the large subunit (this an authentic pre-translocation intermediate) [34]. Recent biochemical and functional data regarding the binding of tRNAs to the A-site of an 80S ribosome–IGR IRES complex and the action of elongation factors on translocation are consistent with the P/E hybrid-state hypothesis [35]. P/E hybrid-state mimicry could help explain how translocation occurs before

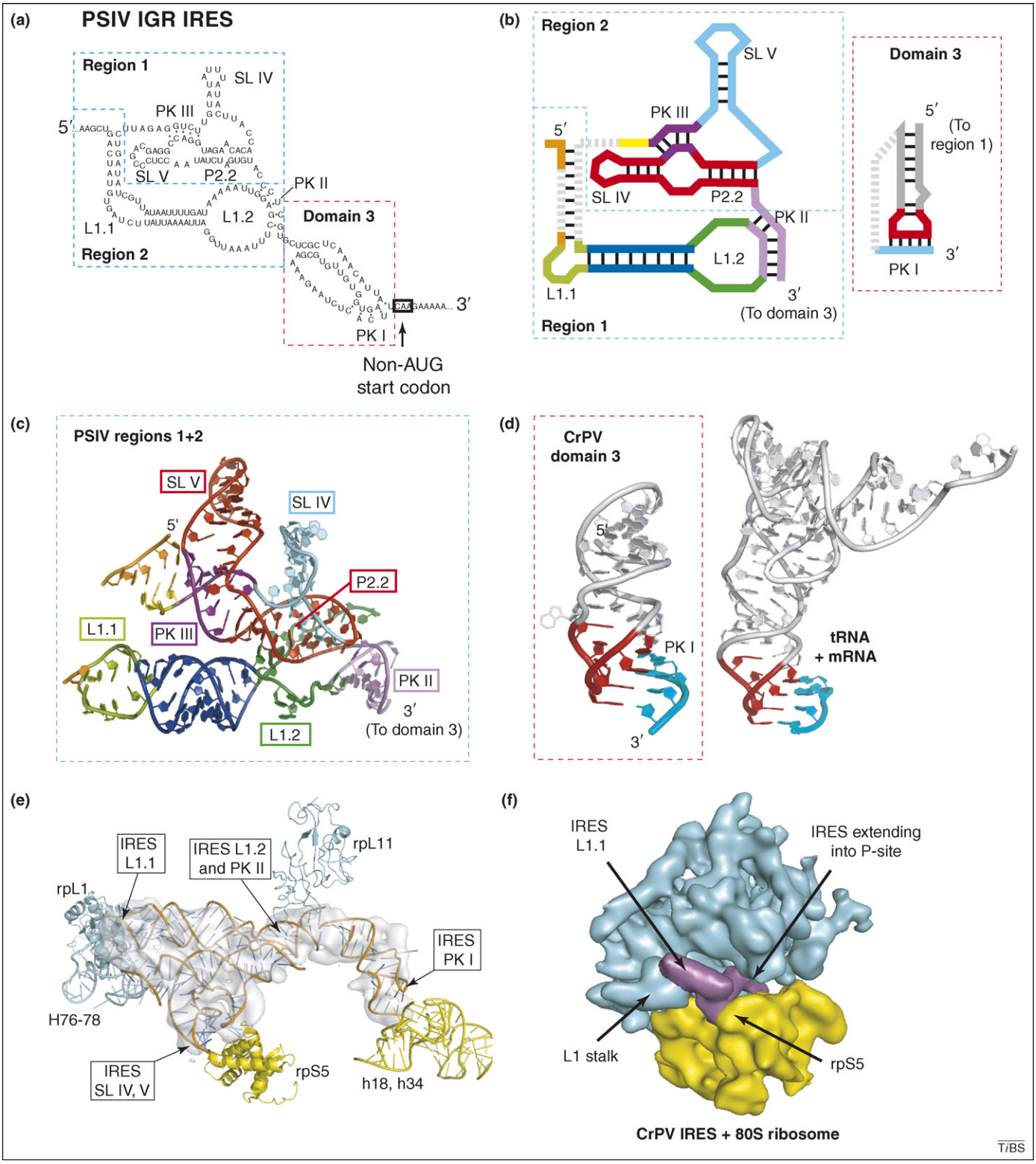

Figure 2. Structures of the *Dicistroviridae* intergenic region (IGR) IRES RNAs. (a) Proposed secondary structure of the Plautia stali intestine virus (PSIV) intergenic region IRES, which includes three pseudoknots and two conserved stem-loops. Important structural elements are labeled: SL, stem loop; PK, pseudoknot; L, loop; P, paired (helix). Regions 1 + 2 are boxed in blue; these regions fold together to comprise the ribosome-binding domain of the IRES. Domain 3 is the part of the IRES that docks into the P-site of the ribosome and is boxed in red. The start codon for protein synthesis is shown; in these IRESs, it is not an AUG codon as in canonical translation initiation. **(b)** Representation of the secondary structures of the two independently folded domains that make up the IGR IRES and for which crystal structures have been solved. On the left, regions 1 and 2 are boxed in blue as in (a), with secondary structural elements colored and labeled. Portions of the structures that were not visible in the crystal structure are shown with broken gray lines. On the right is domain 3, boxed in red as in (a). Parts of the structure that are involved in crystal packing in such a way that they are not relevant biologically are shown with dashed gray. **(c)** Ribbon representation of the crystal structure of the ribosome-binding domain (regions 1 + 2) of the PSIV IGR IRES [26]. The structure is color coded to match (b), with important secondary structure elements and pseudoknots labeled. **(d)** Ribbon representation of the crystal structure of domain 3 of the CrPV IGR IRES colored to match (b) [28]. To the right of the domain 3 structure is the crystal structure of a P-site tRNA–mRNA interaction [58] with the anticodon loop in red and the mRNA in cyan. Comparing these structures shows the mimicry between the tRNA–mRNA interaction and domain 3. **(e)** Cryo-EM difference density of the CrPV IRES bound to the 80S ribosome, at 7.3 Å resolution from which the ribosome density has been removed computationally [29]. The crystal structures of the ribosome-binding domain from PSIV [26] and domain 3 from CrPV [28] are shown docked into the structure. The good fit indicates that the unbound

peptide-bond formation because the ribosome would essentially be in a pre-translocation state immediately on association of the IRES and delivery of the first aminoacyl-tRNA to the A-site. This model will certainly be the subject of studies in the near future.

How dynamic are the structures of the IRES and ribosome during initiation? In the cryo-EM structures, it can be seen that initial binding of the CrPV IRES to the 40S subunit induces a conformational change in the subunit that is similar to that observed with the HCV IRES [36] (discussed later in this review). In addition, on subunit joining, the IRES undergoes an allosteric change involving the position of domain 3 and some changes in the position and shape of the ribosome-binding domain of the IRES [27]. Part of this change involves CrPV IRES loop L1.1 (which was disordered in the PSIV IRES crystal structure) [26]. The cryo-EM density of this loop is weak in the 40S subunit-bound form of the CrPV IRES but this becomes better defined in the 80S ribosome–CrPV IRES complex when interacting with the L1 stalk, perhaps driving the overall allosteric shift in IRES structure that is associated with subunit joining. Furthermore, the L1 stalk is known to move during the translocation cycle and, thus, dynamic parts of the IRES are interacting with dynamic parts of the ribosome [37]. These data indicate that the mechanism of IGR IRES action involves active manipulation of the translational machinery and the IRES responds structurally to the state of the pre-initiation complex.

Structure of the group 2 IRESs in the free and bound forms

Unlike the group 2 IRESs, a complete high-resolution structural model of the group 2 IRESs does not yet exist. Nonetheless, existing structures from NMR, X-ray crystallography and cryo-EM are lending insight into the mechanism by which this RNA recruits and manipulates the ribosome. Structures of HCV IRES elements were solved first and now they can be re-examined in light of lessons learned from the IGR IRESs.

NMR and crystal structures from the unbound group 2 IRES RNAs

The HCV IRES secondary and tertiary structure is very different from the IGR IRESs [38] (Figure 3a) and it does not adopt a globular fold [39]. Rather, the free HCV IRES RNA is overall in an extended conformation and this extended architecture mandates a 'divide and conquer' structural biology approach in which secondary structure elements are solved separately. This approach is necessary because extended RNAs are difficult to crystallize (owing to their inherent flexibility) and because the complete RNA is too large for NMR spectroscopy. Nine structures of HCV IRES RNA elements have been reported so far, with the majority having been solved by NMR (Figure 3b). Specifically, elements IIa (by both NMR and X-ray), IIb and domain II as a whole [40,41], an internal loop in IIIb [42], junction IIIabc (by X-ray) [43] and stem-loops IIId [44,45], IIIc [46] and IIIe [44] have been solved. In addition, internal loop IIa from the classical swine fever virus (CSFV) IRES was solved by NMR [47] (Figure 3b). Many of these (stem-loops IIId, IIIe, IIIc, IIb) are known to make direct contact with the ribosome [48,49] and efforts are being made to integrate these structures into a single cohesive model [50].

An in depth-analysis of each HCV IRES RNA structure is beyond the scope of this review, however, as an example of the type of structural information available, I focus on domain II. Domain II contacts the 40S subunit at the E-site for tRNA binding [36] and is important for progression of the small subunit-bound HCV IRES complex to the 80S ribosome–HCV IRES complex [47,51,52]. This domain adopts an overall bent conformation that is observed in both the isolated IIa structure and the complete structure of domain II (Figure 3c). This bent conformation matches the conformation observed in the cryo-EM structures of the HCV IRES bound to the ribosome, indicating that the bend pre-forms to position domain II precisely in the E-site. Loop IIa from the CSFV IRES also induces a bend, however, it is less strongly bent than domain II of the HCV IRES. This might indicate that the conformation of domain II of the CSFV differs from HCV or perhaps that the CSFV IRES only adopts its native conformation when bound to the ribosome. However, because cryo-EM reconstructions of the CSFV IRES (or any other group 2 IRES) bound to the ribosome do not yet exist, the question of whether these features are conserved remains open.

Cryo-EM of HCV IRES–ribosome complexes

The first cryo-EM structures of the HCV IRES complexes were of full-length (nt 42–372) IRES RNA and a domain II deletion mutant bound to the 40S ribosomal subunit (resolution of ~20 Å). These structures show the HCV IRES on the solvent-exposed side of the subunit in an extended conformation [36] (Figure 4a). Thus, as with the Dicistroviridae IGR IRESs, the overall architecture of the free HCV IRES is maintained when bound and this extended structure enables contact with multiple sites on the ribosomal subunit. Indeed, crosslinking experiments with HCV IRES RNA and human 40S subunits identified many ribosomal proteins as being in contact with the IRES [53].

Using cell-free extracts as a model for HCV IRES-driven translation initiation, it has been shown that, after the 40S subunit–IRES complex forms, the HCV IRES recruits eIF3 and the eIF2–Met-tRNA$_i$–GTP ternary complex to form the 48S*–IRES complex, which then progresses to yield the 80S ribosome [51,52]. A cryo-EM structure of the

structure is similar to the bound form, although differences between the cryo-EM density and the docked crystal structures reveal probable local changes in structure when the IRES binds to the ribosome. The docked structure enables predictions of what IGR IRES RNA structural features interact with specific ribosomal proteins or rRNA [26–29]. The putative interactions are shown, with the relevant rRNA helices, ribosomal proteins and key IRES features labeled. Although rpS25 crosslinks to the PSIV IRES SL IV, its precise location is not known, therefore it is not shown in this figure. Ribosomal proteins or rRNA from the 60S subunit are in cyan and from the 40S subunit are in yellow. Note that, in this view, the ribosome-binding domain is flipped 180 degrees around the horizontal axis relative to panel (c). **(f)** Cryo-EM reconstruction of the CrPV IGR IRES bound to a human 80S ribosome [27], with the IRES density in magenta, the 60S subunit in cyan and the 40S subunit in yellow. The location of the L1 stalk and rpS5 are indicated. The view is from the 'top' of the ribosome, as in Figure 1c. From this perspective, the view is down the length of the IRES structure in (e), looking at the L1.1 end.

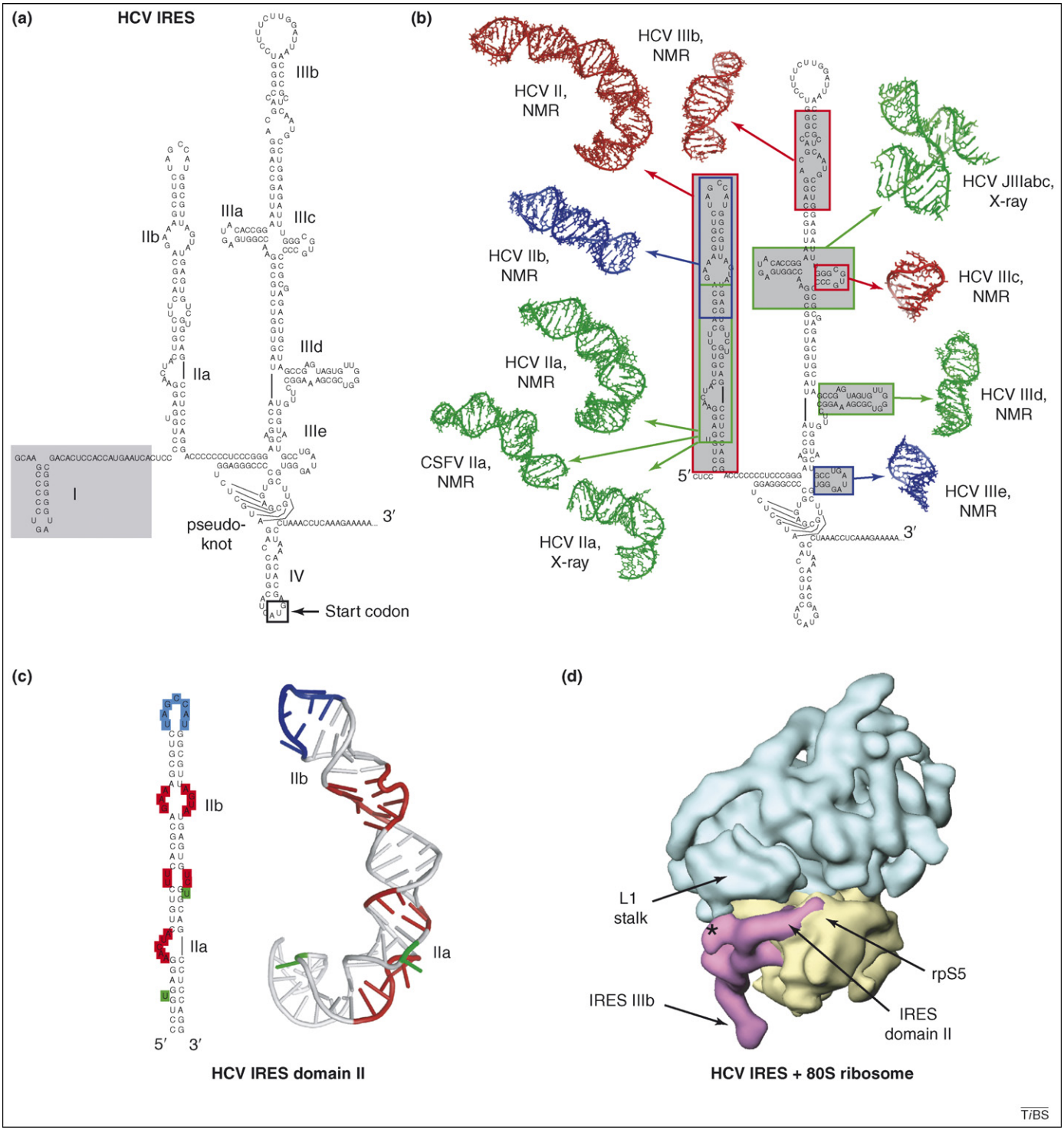

Figure 3. Structures of the hepatitis C virus (HCV) IRES RNA. **(a)** Proposed secondary structure of the HCV IRES, with its stem-loops labeled. Stem-loop I and the portion of the RNA that is shaded can be removed without affecting IRES function. The extended stem-loop structure that contains internal loops IIa and IIb is generally referred to as 'domain II'. In the HCV IRES, the AUG start codon is located in a stem-loop structure (IV), however, this stem-loop is not found in all related IRESs (such as that of CSFV). **(b)** At the left is the secondary structure of the HCV IRES surrounded by structures of secondary-structure elements solved by NMR and X-ray crystallography [40–47]. The location of each element in the structure is shown with arrows and boxes that correspond to the color of the structure. **(c)** Secondary structure of domain II of the HCV IRES and its 3D structure as solved by NMR [41]. Blue denotes the apical loop that is placed into the 40S subunit's E-site, red denotes internal loops IIa and IIb and green denotes bulged bases. These features cause the RNA to adopt an overall bent conformation. **(d)** HCV IRES RNA bound to a human 80S ribosome, with the IRES density in magenta, the 60S subunit in cyan and the 40S subunit in yellow. The location of the L1 stalk and rpS5 are indicated [50]. The asterisk denotes a putative interaction between domain I (which is not needed for IRES function) of the HCV IRES and the base of the L1 stalk.

48S*–IRES complex does not exist but structures of eIF3 and of the HCV IRES bound to eIF3 have been used to construct a model of the 40S–IRES–eIF3 complex [54] (Figure 4a). In this model, the IRES lies between the 40S subunit and eIF3, consistent with mechanistic models that propose that the IRES binds the subunit before recruiting eIF3 [52]. Finally, the cryo-EM structure of the 80S–IRES complex shows IRES density that is similar

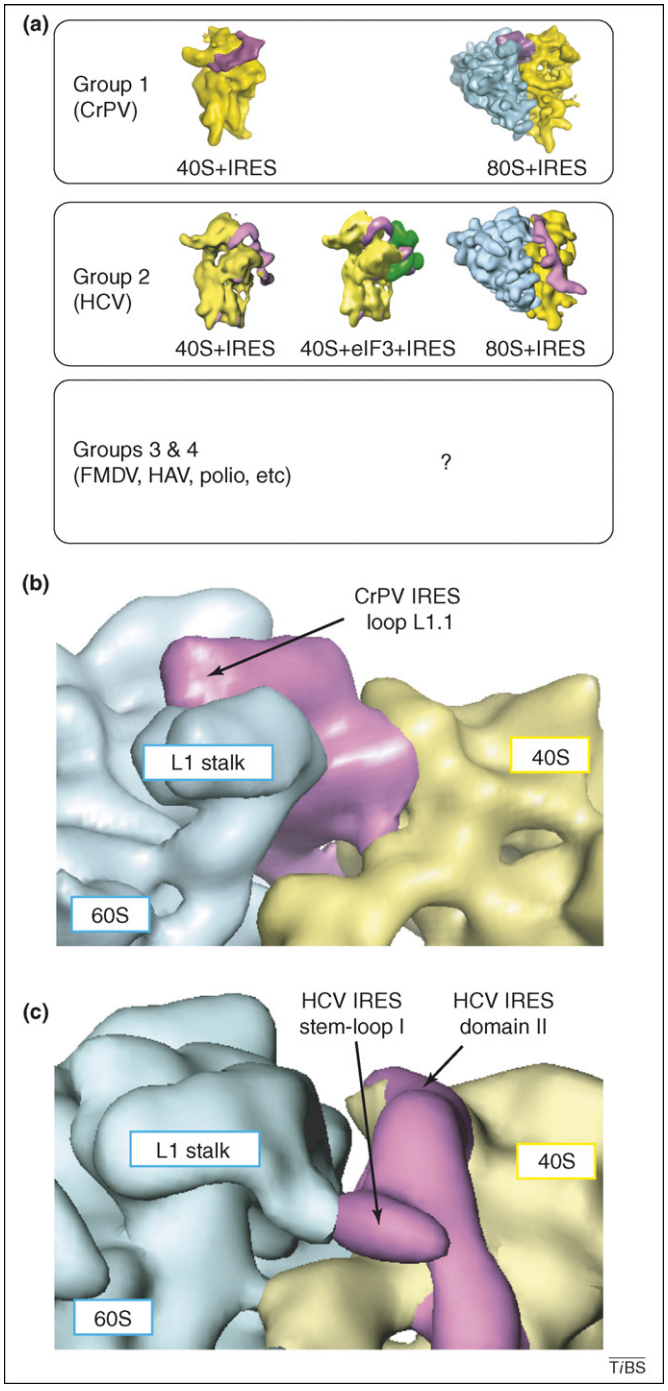

Figure 4. Comparison of cryo-EM reconstructions of the CrPV IGR IRES and the HCV IRES. **(a)** Cryo-EM reconstructions of IRES-containing complexes studied to date. In all reconstructions, the IRES RNA is shown in magenta, the 60S subunit in cyan and the 40S subunit in yellow. At top left is the CrPV IGR IRES RNA bound to a human 40S subunit and at top right is the CrPV IGR IRES bound to the 80S ribosome [27]. The structure of the CrPV IGR IRES bound to a yeast 80S ribosome at higher resolution is not shown [29]. In the middle are structures of the HCV IRES bound to 40S subunits from rabbit, a model of the 40S subunit–eIF3–HCV IRES complex constructed from other reconstructions (eIF3 in green) [54] and the HCV IRES bound to a human 80S ribosome [50]. Cryo-EM reconstructions of members of the group 3 or 4 IRESs do not yet exist but will probably be a goal of future efforts. **(b)** A close up view of the CrPV IRES bound to a human 80S ribosome, with the L1 stalk and loop L1.1 of the IRES labeled [27]. This view is rotated ~90 degrees around both the horizontal axis relative to Figure 2f, with the view looking into the E-site over the L1 stalk. **(c)** Close-up view of the HCV IRES bound to the 80S ribosome, with the IRES mostly outside the intersubunit space, but reaching into the E-site within domain II. The location of the L1 stalk is labeled. The view is rotated 90 degrees around both the horizontal and vertical axes, relative to Figure 3d.

in shape to that observed in the 40S subunit-bound form, as well as additional contacts to the 60S subunit's L1 stalk [50] (Figures 3d and 4a). Parts of the HCV IRES cryo-EM density have been assigned to specific structural elements of the IRES RNA [36] and these assignments guide placement of high-resolution structures to generate a more detailed model of the HCV IRES–ribosome complex. Specifically, Boeringer *et al.* placed domains II, IIId, part of IIIb and the IIIabc junction into the density of the 80S ribosome-bound HCV IRES, enabling prediction of specific intermolecular interactions [50]. The full functional implications of these predicted contacts remain to be explored.

As with the CrPV IGR IRES, the HCV IRES-binding changes the conformation of the 40S subunit [36]. This conformational change depends on the IRES domain II, a result corroborated by toeprinting and crosslinking experiments [55]. These results imply that the HCV IRES is an active manipulator of the translation machinery and this is owing, in part, to IRES contacts near the E-site. It is interesting to note that binding of two elongation factors involved in start codon selection (eIF1 and eIF1A) also induces a change in 40S subunit structure [56], at least raising the possibility that the IRES mimics this function. Cryo-EM reconstructions of the 80S–IRES complex show the 40S subunit changes conformation again on subunit joining, especially in areas that interact with mRNA [50]. Although the overall conformation of the HCV IRES is similar in the 40S subunit and 80S-ribosome bound forms, there is evidence of subtle restructuring of the pseudoknot of the IRES and four-way junction. The conformation of the 80S ribosome is somewhat different in the HCV IRES-bound-form than in the free form [50] (Figure 4b,c). The mechanistic implication of these changes remains unknown; they could be important for proper placement of the HCV-coding sequence during initiation. Additional discussion of conclusions from cryo-EM reconstructions of HCV IRES complexes can be found in a recent review [57].

Differences and similarities between HCV and IGR IRES ribosome binding

Comparing structures of the HCV IRES and the CrPV IRES in free and bound forms now enables us to start answering the question: are there common features to these IRES groups? The overall architectures (extended versus compact) of the two IRESs are different and no similarities in their high-resolution structures are apparent. The locations of the CrPV and HCV IRESs on the ribosome are different in that the former is tucked into the inter-subunit space and the latter is mostly on the solvent-accessible side of the 40S subunit (Figure 4a), although they are similar in that both contact the 40S subunit at the E-site in the vicinity of ribosomal protein S5 (rpS5) and both IRESs induce a similar conformational change in the ribosome. On the large subunit, both of these IRESs contact the L1 stalk, although the locations of the interactions are different (Figure 4b,c). In the case of the CrPV IGR IRES, the interaction is in a similar location to that contacted by an E-site tRNA. The L1 stalk is a highly mobile element that helps remove the E-site tRNA from the ribosome during translocation [37], which leads to the

hypothesis that contact with the L1 stalk helps move IRES RNAs in the E-site during IRES-driven initiation or that this contact is important for manipulation of the ribosome. In addition, there is another subtle similarity between the IGR and HCV IRESs: both use only a portion of their structure to provide affinity for the ribosome and these are the most structurally compact portions of the IRES RNA (regions 1 + 2 of the IGR IRESs and domains IIIa, IIIc-f and the pseudoknot of the HCV IRES). But, in both IRESs, ribosome binding is not sufficient for function; both need at least one other domain to interact in the decoding groove (CrPV domain 3 in the P-site and HCV IRES domain II in the E-site). Are these similarities profound or coincidental? The answer to this question remains to be answered but the fact that we can now ask this question illustrates the power of structure to guide studies aimed at uncovering any fundamental rules of IRES function.

Concluding remarks and future perspectives

Viral IRESs are crucial for successful infection by many viruses but the structural basis of their function is only beginning to be understood. In the last decade, X-ray, NMR and cryo-EM of IRES RNAs that bind to the ribosome directly are informing new mechanistic models and helping to guide the next round of experiments. It is probable that more detailed exploration of the dynamic nature of these IRES–ribosome complexes will include the application of single-molecule methods and possibly crystallography applied to IRES-containing ribosomal complexes. The latter will be particularly challenging given that high-resolution structures of eukaryotic ribosomes do not yet exist. In addition, although structures of the ribosome-binding IRESs are proving useful, RNA or RNA–protein complex structures from IRESs that require initiation factors or additional protein factors to bind the ribosome (that is IRESs from groups 3 and 4) have yet to be solved. Solving the structures of these targets is an important component of studies aimed at understanding the full breadth and depth of the structural basis of IRES function.

Acknowledgements

I would like to thank all members of my laboratory for helpful discussions, J. Pfingsten and R. Batey for critical reading of the manuscript, C. Spahn, H. Stark, J. Doudna and C. Fraser for supplying cryo-EM density files and N. Pace for a useful discussion about RNase P cleavage assays. IRES RNA studies in the Kieft Laboratory are supported by grants R01 GM072560 and R03 AI072187 from the National Institutes of Health and Research Scholar Grant 0805801GMC from the American Cancer Society.

References

1 Pestova, T. *et al.* (2007) The mechanism of translation initiation in eukaryotes. In *Translational Control in Biology and Medicine* (Mathews, M.B. *et al.*, eds), pp. 87–128, Cold Spring Harbor Laboratory Press
2 Jang, S.K. *et al.* (1988) A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.* 62, 2636–2643
3 Pelletier, J. and Sonenberg, N. (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334, 320–325
4 Doudna, J.A. and Sarnow, P. (2007) Translation initiation by viral internal ribosome entry sites. In *Translational Control in Biology and Medicine* (Mathews, M.B. *et al.*, eds), pp. 129–153, Cold Spring Harbor Laboratory Press
5 Elroy-Stein, O. and Merrick, W.C. (2007) Translation initiation via cellular internal ribosome entry sites. In *Translational Control in Biology and Medicine* (Mathews, M.B. *et al.*, eds), pp. 155–172, Cold Spring Harbor Laboratory Press
6 Chen, C.Y. and Sarnow, P. (1995) Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* 268, 415–417
7 Baird, S.D. *et al.* (2006) Searching for IRES. *RNA* 12, 1755–1785
8 Hernandez, G. (2008) Was the initiation of translation in early eukaryotes IRES-driven? *Trends Biochem. Sci.* 33, 58–64
9 Glass, M.J. and Summers, D.F. (1992) A cis-acting element within the hepatitis A virus 5′-non-coding region required for *in vitro* translation. *Virus Res.* 26, 15–31
10 Tsukiyama-Kohara, K. *et al.* (1992) Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* 66, 1476–1483
11 Kuhn, R. *et al.* (1990) Functional analysis of the internal translation initiation site of foot-and-mouth disease virus. *J. Virol.* 64, 4625–4631
12 Buck, C.B. *et al.* (2001) The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. *J. Virol.* 75, 181–191
13 Brasey, A. *et al.* (2003) The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J. Virol.* 77, 3939–3949
14 Jan, E. (2006) Divergent IRES elements in invertebrates. *Virus Res.* 119, 16–28
15 Jan, E. *et al.* (2003) Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *Proc. Natl Acad. Sci. U. S. A.* 100, 15410–15415
16 Jan, E. *et al.* (2001) Initiator Met-tRNA-independent translation mediated by an internal ribosome entry site element in cricket paralysis virus-like insect viruses. *Cold Spring Harb. Symp. Quant. Biol.* 66, 285–292
17 Wilson, J.E. *et al.* (2000) Initiation of protein synthesis from the A site of the ribosome. *Cell* 102, 511–520
18 Wilson, J.E. *et al.* (2000) Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol. Cell. Biol.* 20, 4990–4999
19 Sasaki, J. and Nakashima, N. (2000) Methionine-independent initiation of translation in the capsid protein of an insect RNA virus. *Proc. Natl Acad. Sci. U. S. A.* 97, 1512–1515
20 Sasaki, J. and Nakashima, N. (1999) Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus *in vitro*. *J. Virol.* 73, 1219–1226
21 Pestova, T.V. and Hellen, C.U. (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
22 Kanamori, Y. and Nakashima, N. (2001) A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA* 7, 266–274
23 Jan, E. and Sarnow, P. (2002) Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J. Mol. Biol.* 324, 889–902
24 Nishiyama, T. *et al.* (2003) Structural elements in the internal ribosome entry site of Plautia stali intestine virus responsible for binding with ribosomes. *Nucleic Acids Res.* 31, 2434–2442
25 Costantino, D. and Kieft, J.S. (2005) A preformed compact ribosome-binding domain in the cricket paralysis-like virus IRES RNAs. *RNA* 11, 332–343
26 Pfingsten, J.S. *et al.* (2006) Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. *Science* 314, 1450–1454
27 Spahn, C.M. *et al.* (2004) Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes; the IRES functions as an RNA-based translation factor. *Cell* 118, 465–475
28 Costantino, D.A. *et al.* (2008) tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat. Struct. Mol. Biol.* 15, 57–64
29 Schuler, M. *et al.* (2006) Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat. Struct. Mol. Biol.* 13, 1092–1096
30 Pestova, T.V. *et al.* (2004) Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. *EMBO Rep.* 5, 906–913
31 Pfingsten, J.S. *et al.* (2007) Conservation and diversity among the three-dimensional folds of the Dicistroviridae intergenic region IRESes. *J. Mol. Biol.* 370, 856–869

32 Nishiyama, T. *et al.* (2007) Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a dicistroviral intergenic internal ribosome entry site. *Nucleic Acids Res.* 35, 1514–1521
33 Yamamoto, H. *et al.* (2007) Binding mode of the first aminoacyl-tRNA in translation initiation mediated by Plautia stali intestine virus IRES. *J. Biol. Chem.* 282, 7770–7776
34 Dorner, S. *et al.* (2006) The hybrid state of tRNA binding is an authentic translation elongation intermediate. *Nat. Struct. Mol. Biol.* 13, 234–241
35 Yamamoto, H. *et al.* (2007) Binding mode of the first aminoacyl-tRNA in translation initiation mediated by Plautia stali intestine virus internal ribosome entry site. *J. Biol. Chem.* 282, 7770–7776
36 Spahn, C.M. *et al.* (2001) Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* 291, 1959–1962
37 Valle, M. *et al.* (2003) Locking and unlocking of ribosomal motions. *Cell* 114, 123–134
38 Brown, E.A. *et al.* (1992) Secondary structure of the 5′ nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res.* 20, 5041–5045
39 Kieft, J.S. *et al.* (1999) The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J. Mol. Biol.* 292, 513–529
40 Dibrov, S.M. *et al.* (2007) Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angew. Chem. Int. Ed. Engl.* 46, 226–229
41 Lukavsky, P.J. *et al.* (2003) Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol.* 10, 1033–1038
42 Collier, A.J. *et al.* (2002) A conserved RNA structure within the HCV IRES eIF3-binding site. *Nat. Struct. Biol.* 9, 375–380
43 Kieft, J.S. *et al.* (2002) Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nat. Struct. Biol.* 9, 370–374
44 Lukavsky, P.J. *et al.* (2000) Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat. Struct. Biol.* 7, 1105–1110
45 Klinck, R. *et al.* (2000) A potential RNA drug target in the hepatitis C virus internal ribosomal entry site. *RNA* 6, 1423–1431
46 Rijnbrand, R. *et al.* (2004) Mutational and structural analysis of stem-loop IIIC of the hepatitis C virus and GB virus B internal ribosome entry sites. *J. Mol. Biol.* 343, 805–817
47 Locker, N. *et al.* (2007) HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J.* 26, 795–805
48 Kieft, J.S. *et al.* (2001) Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* 7, 194–206
49 Kolupaeva, V.G. *et al.* (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
50 Boehringer, D. *et al.* (2005) Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. *Structure* 13, 1695–1706
51 Ji, H. *et al.* (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc. Natl Acad. Sci. U. S. A.* 101, 16990–16995
52 Otto, G.A. and Puglisi, J.D. (2004) The pathway of HCV IRES-mediated translation initiation. *Cell* 119, 369–380
53 Otto, G.A. *et al.* (2002) Ribosomal proteins mediate the hepatitis C virus IRES–HeLa 40S interaction. *RNA* 8, 913–923
54 Siridechadilok, B. *et al.* (2005) Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* 310, 1513–1515
55 Pestova, T.V. *et al.* (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
56 Passmore, L.A. *et al.* (2007) The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol. Cell* 26, 41–50
57 Fraser, C.S. and Doudna, J.A. (2007) Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat. Rev. Microbiol.* 5, 29–38
58 Selmer, M. *et al.* (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313, 1935–1942
59 Hatakeyama, Y. *et al.* (2004) Structural variant of the intergenic internal ribosome entry site elements in dicistroviruses and computational search for their counterparts. *RNA* 10, 779–786
60 Mari, J. *et al.* (2002) Shrimp Taura syndrome virus: genomic characterization and similarity with members of the genus Cricket paralysis-like viruses. *J. Gen. Virol.* 83, 915–926
61 Rijnbrand, R. *et al.* (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
62 Pisarev, A.V. *et al.* (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
63 Pilipenko, E.V. *et al.* (1994) Starting window, a distinct element in the cap-independent internal initiation of translation on picornaviral RNA. *J. Mol. Biol.* 241, 398–414
64 Borman, A. and Jackson, R.J. (1992) Initiation of translation of human rhinovirus RNA: mapping the internal ribosome entry site. *Virology* 188, 685–696
65 Moore, S.D. and Sauer, R.T. (2007) The tmRNA system for translational surveillance and ribosome rescue. *Annu. Rev. Biochem.* 76, 101–124
66 Fechter, P. *et al.* (2001) Novel features in the tRNA-like world of plant viral RNAs. *Cell. Mol. Life Sci.* 58, 1547–1561
67 Kristensen, O. *et al.* (2002) Is tRNA binding or tRNA mimicry mandatory for translation factors? *Curr. Protein Pept. Sci.* 3, 133–141
68 Lyons, A.J. and Robertson, H.D. (2003) Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. *J. Biol. Chem.* 278, 26844–26850
69 Piron, M. *et al.* (2005) Characterizing the function and structural organization of the 5′ tRNA-like motif within the hepatitis C virus quasispecies. *Nucleic Acids Res.* 33, 1487–1502
70 Serrano, P. *et al.* (2007) Characterization of a cyanobacterial RNase P ribozyme recognition motif in the IRES of foot-and-mouth disease virus reveals a unique structural element. *RNA* 13, 849–859
71 Kirsebom, L.A. (2007) RNase P RNA mediated cleavage: substrate recognition and catalysis. *Biochimie* 89, 1183–1194