Published in final edited form as:

Science. 2006 December 1; 314(5804): 1450–1454. doi:10.1126/science.1133281.

Structural Basis for Ribosome Recruitment and Manipulation by a Viral IRES RNA

Jennifer S. Pfingsten, David A. Costantino, and Jeffrey S. Kieft*

Department of Biochemistry and Molecular Genetics, University of Colorado at Denver and Health Sciences Center, Mail Stop 8101, Post Office Box 6511, Aurora, CO 80045, USA

Abstract

Canonical cap-dependent translation initiation requires a large number of protein factors that act in a stepwise assembly process. In contrast, internal ribosomal entry sites (IRESs) are *cis*-acting RNAs that in some cases completely supplant these factors by recruiting and activating the ribosome using a single structured RNA. Here we present the crystal structures of the ribosome-binding domain from a Dicistroviridae intergenic region IRES at 3.1 angstrom resolution, providing a view of the prefolded architecture of an all-RNA translation initiation apparatus. Docking of the structure into cryo-electron microscopy reconstructions of an IRES-ribosome complex suggests a model for ribosome manipulation by a dynamic IRES RNA.

In eukaryotes, there are two known mechanisms for the initiation of protein synthesis (Fig. 1A). The canonical mechanism requires a modified nucleotide cap on the 5' end of the mRNA, which is recognized by an initiation factor protein (eIF4E). This protein recruits other factors that assemble the ribosome on the mRNA in a stepwise process (1). In contrast, internal initiation of translation does not require a cap or recognition of the mRNA 5' end. Rather, structured RNA sequences called internal ribosomal entry sites (IRESs) recruit and activate the translation machinery, functionally replacing many protein factors (2). IRESs are essential for infection by many medically and economically important viruses such as hepatitis C (HCV), hepatitis A, polio, foot-and-mouth disease, rhinovirus, coxsackievirus-B3, and HIV-1 (3). IRESs also drive the translation of eukaryotic mRNAs, encoding factors involved in development, growth regulation, apoptosis, transcription, translation, and other important cellular processes (3). The molecular rules underlying this RNA structure-driven mechanism remain elusive.

Ideal model systems for understanding IRES RNA-driven translation are the mechanistically streamlined intergenic region (IGR) IRESs of the virus family Dicistroviridae (4). The IGR IRESs drive the association of the ribosomal subunits without any of the protein factors that comprise the canonical translation initiation apparatus (Fig. 1A) (5). Hence, this one structured RNA (molecular size ~66 kD) supplants over 1000 kD of structured initiation factor proteins, operating as an all-RNA translation initiation apparatus (6-9). The full-length IGR IRES folds in solution into two structurally independent domains (10-13). The larger domain (regions 1 and 2, Fig. 1A and fig. S1) is the ribosome-binding domain. It folds into a compact structure

*To whom correspondence should be addressed. E-mail: Jeffrey.Kieft@uchsc.edu

(10) that binds directly to the 40S subunit (10,12,13). Cryo-electron microscopy (cryo-EM) reconstructions of an IGR IRES bound to the ribosome reveal that the IGR IRES binds over the mRNA-binding groove, making contact to and changing the structure of both ribosomal subunits (40S and 60S) (14). However, these cryo-EM structures do not reveal the structure of the IRES, how the IRES structure creates a ribosome-binding site, or which IRES structural features specifically contact and manipulate the ribosome.

To address these questions and develop a model for the structural basis of IGR IRES-driven translation, we have solved the structure of the ribosome-binding domain of the *Plautia stali* intestine virus (PSIV) IGR IRES RNA using x-ray crystallography to a resolution of 3.1 Å with $R_{\text{work}} = 25.3\%$ and $R_{\text{free}} = 29.4\%$ (Fig. 1B). In the structure, regions 1 and 2 pack side-by-side with stem-loop IV (SL IV) and SL V emerging from the same side of the structure (Fig. 1B). Both stem-loops have been shown through mutagenesis and footprinting experiments to make functionally critical direct contact to the 40S subunit (10,12,13), and isolated region 2 has been shown to bind to the 40S subunit (12). Thus, the position of SL IV and SL V identifies the side of the structure that is the 40S subunit-binding interface (“top” of the structure in Fig. 1B) and indicates that the 40S subunit recognition surface prefolds before subunit binding. The foundation of this prefolded IRES architecture is helix P2.2 (red, Fig. 1B), which is contacted by bases from multiple IRES elements (J2.3, J2.2, L1.2A, and L1.2B, cyan and green, Fig. 1B). This folded core corresponds to areas of strong protection in hydroxyl-radical probing experiments (fig. S2), providing additional evidence that we have captured the authentic prefolded, unbound form of the IRES RNA (10). Many of the most-conserved bases in the IGR IRESs cluster in this highly structured core region. Our structure suggests that these bases are kept invariant to maintain the specific intramolecular contacts that knit the structure together.

The two stem-loops are positioned together to contact a relatively small area on the ribosome by a set of specific intramolecular contacts involving pseudoknot III (PK III) and helix P2.2, which form a continuous helical stack that extends into PK II (Fig. 1B). The placement of SL IV is facilitated by underwinding of helix P2.2, which opens the normally deep and narrow major groove (Fig. 2A). The degree of underwinding is such that a single turn of helix covers a total rise of ~ 42 Å, as compared to ~ 34 Å for canonical A-form RNA. This feature is induced by four key bases that stack into the helix, forming two noncanonical base pairs and anchoring the 3' end of SL IV (J2.3, Fig. 2A). The 5' end of SL IV is anchored by a G-U wobble pair (G6110-U6082, Fig. 2B) (12). From these anchors in the P2.2 major groove, the stem of SL IV is closed with a U-U pair, the bases of which are flipped out from the P2.2 major groove (U6083-U6096, Fig. 2B). Whereas SL IV nestles in the major groove, SL V is positioned by several conserved bases that force the stem of SL V to emerge from the minor groove, placing it at nearly a right angle to P2.2 and adjacent to SL IV (Fig. 2C). Classic pseudoknot folding is characterized by two stacked helices and two single-stranded loops; the 5'-most loop crosses the major groove, whereas the other crosses the minor groove (15). Hence, in the IGR IRES, this classic pseudoknot architecture is maintained, despite the fact that both loops contain a large amount of embedded structured RNA (fig. S3).

Region 1 does not participate directly in the interactions that create the 40S subunit-binding site; rather, it packs against region 2 in position to contact the 60S subunit (vide infra). Regions 1 and 2 pack in an interaction in which the large L1.2 loop (green, Fig. 2D) cradles one strand of helix P2.2 (red, Fig. 2D). One strand of this loop (L1.2A) lies along the minor groove of P2.2, forming stabilizing A-minor interactions [for discussions of A-minor interactions, see (16)]. The other strand of the loop (L1.2B) forms the other half of the cradle; apparently poised to enter the major groove, it loops back to the minor groove to form more stabilizing A-minor interactions. The tight, complex packing of RNA that is essential for function is especially evident in this region, where five strands of RNA trace in close proximity. The other parts of regions 1 and 2 (helix P1.1 and loop L1.1) are only weakly visible in the electron density. We

built RNA structure into this weak density, but the structure is poorly defined in these regions, indicating conformational flexibility. Hence, although we can clearly see where L1.1 lies in relation to the rest of the structure, we cannot report its high-resolution structure. This flexibility corresponds to the fact that the RNA crystallized as a domain-swapped dimer in which the native interactions between regions 1 and 2 are preserved (for a detailed discussion, see fig. S4). Hence, the IRES contains regions of stable, highly structured RNA and other regions of flexible, less stable structure. We examined the functional significance of this observation by combining our structure with existing cryo-EM reconstruction data (14).

To identify the specific IGR IRES RNA structures that contact the ribosome, we docked the crystal structure into published cryo-EM reconstructions (Fig. 3). To generate a model for docking, we returned the SL IV and SL V stems to their wild-type lengths [these were changed to induce crystallization (17)]. With the use of the location of the 3' end of the IRES RNA, assignments of region 3 and regions 1 and 2 into the cryo-EM maps (14), footprinting and directed hydroxyl-radical probing data (12,18), and the overall agreement of the crystal structure with the shape of the cryo-EM density, the docking orientation was unambiguous (fig. S5) (12,14,18). Based on the good, but not perfect, match of the structure to the cryo-EM density, we conclude that the IRES does not need to undergo a global structural rearrangement to match our structure. Rather, local structural shifts (such as a change in the relative angle of helices or a shift of region 1 relative to region 2) occur upon binding. The docking and fit are robust enough to identify the IRES RNA structural domains that contact the 40S and 60S subunits and to suggest a dynamic mechanism of IRES action.

Cryo-EM showed that small ribosomal protein S5 (rpS5) makes contact with the IGR IRES (14), and our docked structure reveals that SL IV and SL V are the features that interact with rpS5 (Fig. 3A). These are the only IRES contacts to rpS5, indicating that this is the keystone interaction that drives 40S subunit binding, induces conformational change, and begins the translation initiation process. This observation is supported by data showing that mutation of these loops abrogates 40S subunit binding and translation initiation activity (10,12,13). The apical loop sequences of SL IV and SL V are conserved almost universally among the IGR IRESs, underscoring that their specific interaction with rpS5 is critical for function (fig. S6). The HCV IRES also contacts rpS5, and the HCV IRES domain that contacts rpS5 also changes the conformation of the 40S subunit (19-21), suggesting that this interaction is central to this mode of IRES RNA-driven translation.

In the docked structure, IRES loop L1.1 (and perhaps adjacent helices) is positioned to make direct contact with the large subunit's L1 stalk, which contains rPL1 and ribosomal RNA helix H77 (Fig. 3B) (14). That L1.1 and the adjacent P1.1 helix are poorly ordered suggests this part of the IRES is not stably folded until it interacts with the 60S subunit. During translation, the L1 stalk interacts with the T loop of E site-bound tRNA (22,23), which suggests that IGR IRES loop L1.1, when bound in the 80S ribosome, adopts a structure that mimics this portion of tRNA structure. Like the 40S subunit-binding stem-loops, IRES loop L1.1 is highly conserved, demonstrating the importance and specificity of its interactions with the L1 stalk (fig. S6) (4, 11). Our observation that L1.1 makes an important contact to the 60S subunit predicts that mutating L1.1 will block translation initiation after 40S subunit binding. To test this prediction, we constructed two mutants and used them in preinitiation complex assays analyzed on sucrose gradients (Fig. 4, A to C). Both mutants produce IRES-40S subunit complexes but fail to progress to 80S ribosomes, demonstrating that L1.1 is critical for recruitment of the 60S subunit. Neither of these mutations globally misfold the RNA (Fig. 4B), which suggests that the effect is due to the loss of the direct L1.1 interaction with the 60S subunit.

The IRES interactions with rpS5 and the L1 stalk are the only intimate contacts to regions 1 and 2 in our docked structure. This does not preclude that other interactions may occur,

especially upon proposed subtle conformational changes to the IRES. However, the fact that the conservation of nucleotides in regions 1 and 2 of the IRES is fully accounted for by the folded core and contacts with rpS5 and the L1 stalk suggests that other contacts may be nonspecific in nature (fig. S6).

Our structure of the PSIV IGR IRES ribosome-binding domain, combined with a wealth of published biochemical, functional, and low-resolution structural data, suggests a mechanistic model for the structural basis of IGR IRES-driven initiation that involves programmed regions of stable and flexible structure (Fig. 4D). We propose that P1.1 and L1.1 remain flexible when the IRES binds to the 40S subunit through structured SL IV and SL V (Fig. 4D, step 1). Support for this idea comes from a close examination of the IRES structure docked into the cryo-EM density of the 40S subunit-bound IRES. The part of the density that corresponds to L1.1 and P1.1 is weak or missing (14) (red oval, Fig. 4D). In the 40S-bound form, IRES region 3 overlaps with the P and A sites (14). 60S subunit binding (Fig. 4D, step 2) results in structural changes in the IRES and a shift that withdraws region 3 from the A and P sites, as well as a change in the structure of the ribosome's L1 stalk that is similar to changes associated with elongation factor binding (14,24,25). These IRES structural changes are explained by a 60S subunit binding-induced organization of L1.1 and perhaps P1.1, evident in the appearance of additional cryo-EM density around L1.1 and P1.1 upon 60S subunit binding (Fig. 4D) (14). Furthermore, the fact that regions 1 and 2 of the IRES make relatively few inter-region contacts suggests that the two regions can shift relative to one another. Change in the L1.1 and P1.1 structure thus could be linked to PK II and domain 3 through the P2.1 helix. Other ribosome features positioned near the IRES may also be part of this overall mechanism (14). Pestova *et al.* have reported that the IGR IRES RNAs also can recruit 80S ribosomes directly (18) in what must be a coupled series of events within the context of the assembled 80S ribosome.

Although there is great diversity in IRES structure, combining stable and flexible regions may be a strategy used by other IRESs. This structural characteristic is observed in the HCV IRES despite a very different overall RNA architecture (26). Thus, the structure presented here provides the basis for experiments aimed at understanding the basic tenets of RNA-based translation initiation.

Acknowledgements

We acknowledge the staff at beamline ALS 8.2.1 for assistance, R. Zhao for managing the University of Colorado (UC) at Denver and Health Sciences Center x-ray facility, D. Farrell for computer administration, and M. Churchill, R. Batey, and A. Ferré-D'Amaré for useful discussions and advice. We especially thank C. Spahn for supplying various cryo-EM density files and R. Batey for the iridium (III) hexammine. R. Batey, M. Churchill, D. Bentley, L. Krushel, and R. Zhao provided a critical reading of this manuscript. This work was supported by a grant from NIH and funding from the UC Cancer Center in support of the x-ray facility. Structure factors and coordinates have been deposited in the Protein Data Bank under accession code 2IL9.

References and Notes

1. Hershey, JWB.; Merrick, WC. Translational Control of Gene Expression. Sonenberg, N.; Hershey, JWB.; Mathews, MB., editors. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2000. p. 33-88.
2. Hellen CU, Sarnow P. Genes Dev 2001;15:1593. [PubMed: 11445534]
3. Bonnal S, Boutonnet C, Prado-Lourenco L, Vagner S. Nucleic Acids Res 2003;31:427. [PubMed: 12520041]
4. Jan E. Virus Res 2005;119:16. [PubMed: 16307820]
5. Pestova TV, Hellen CU. Genes Dev 2003;17:181. [PubMed: 12533507]
6. Sasaki J, Nakashima N. J. Virol 1999;73:1219. [PubMed: 9882324]
7. Sasaki J, Nakashima N. Proc. Natl. Acad. Sci. U.S.A 2000;97:1512. [PubMed: 10660678]

8. Jan E, et al. Cold Spring Harbor Symp. Quant. Biol 2001;66:285. [PubMed: 12762030]
9. Wilson JE, Pestova TV, Hellen CU, Sarnow P. Cell 2000;102:511. [PubMed: 10966112]
10. Costantino D, Kieft JS. RNA 2005;11:332. [PubMed: 15701733]
11. Kanamori Y, Nakashima N. RNA 2001;7:266. [PubMed: 11233983]
12. Nishiyama T, et al. Nucleic Acids Res 2003;31:2434. [PubMed: 12711689]
13. Jan E, Sarnow P. J. Mol. Biol 2002;324:889. [PubMed: 12470947]
14. Spahn CM, et al. Cell 2004;118:465. [PubMed: 15315759]
15. Hilbers CW, Michiels PJ, Heus HA. Biopolymers 1998;48:137. [PubMed: 10333742]
16. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. Proc. Natl. Acad. Sci. U.S.A 2001;98:4899. [PubMed: 11296253]
17. Materials and methods are available as supporting material on *Science* Online.
18. Pestova TV, Lomakin IB, Hellen CU. EMBO Rep 2004;5:906. [PubMed: 15332113]
19. Fukushi S, et al. J. Biol. Chem 2001;276:20824. [PubMed: 11331271]
20. Spahn CM, et al. Science 2001;291:1959. [PubMed: 11239155]
21. Kieft JS, Zhou K, Jubin R, Doudna JA. RNA 2001;7:194. [PubMed: 11233977]
22. Agrawal RK, et al. J. Cell Biol 2000;150:447. [PubMed: 10931859]
23. Yusupov MM, et al. Science 2001;292:883. [PubMed: 11283358]
24. Gomez-Lorenzo MG, et al. EMBO J 2000;19:2710. [PubMed: 10835368]
25. Agrawal RK, Heagle AB, Penczek P, Grassucci RA, Frank J. Nat. Struct. Biol 1999;6:643. [PubMed: 10404220]
26. Kieft JS, et al. J. Mol. Biol 1999;292:513. [PubMed: 10497018]
27. Jucker FM, Pardi A. RNA 1995;1:219. [PubMed: 7585251]

Fig. 1. IRES-driven initiation and the structure of the PSIV IGR IRES. **(A)** Ribosome recruitment strategies used in cap-dependent and IGR-IRES-dependent translation initiation. ORF, open reading frame; GTP, guanosine triphosphate; GDP, guanosine diphosphate. At right is the protein-independent pathway used by this IRES RNA. The inset is a cartoon of the secondary structure of the ribosome-binding domain, colored to match the structure of **(B)**. Parts colored gray did not appear in the crystal structure and were not built into the final model. The secondary structure consists of two regions (regions 1 and 2), which contain two functionally critical pseudoknots (10,12,13). Because PK III is nested inside PK II, this forms an RNA tertiary structure called a double-nested pseudoknot (15). Figure S1 contains a detailed secondary structure with sequence information. **(B)** Structure of the ribosome-binding domain, colored to match the inset in **(A)**. J, junction; P, paired/helix; L, loop. The gray hexagon shows structure that was weakly visible and hence conformationally flexible. The RNA crystallized in a domain-swapped dimer (fig. S4) in which the functionally essential structural features are preserved.

Fig. 2.

Structural details of the IRES. (A) Underwound helix P2.2 (red) is shown from the minor groove side. A6105 (red) stacks into the helix, forming a noncanonical A·A N7-amino base pair with A6098 of J2.3 (cyan). A6135 (red) stacks on A6098 and forms a reverse (parallel) Watson-Crick A-U pair with conserved base U6097 of J2.3 (cyan). In this and subsequent panels, base pairing is indicated with double-ended arrows and stacking with a thick dashed line. (B) Within the P2.2 major groove, both U6097 (cyan) and U6082 (purple) pair with bases in P2.2 (red), whereas U6083 and U6096 pair with each other at the end of SL IV (cyan). Region 1 bases (green) buttress this structure through A-minor interactions. (C) The U6082-G6110 pairing extrudes U6130 from the helix and induces a sharp turn in the backbone. Bases A6129 and A6128 stack on the minor groove of the PK III helix, starting the base stacking that extends into SL V. (D) The two strands of L1.2 splay apart, with L1.2A lying in the minor groove of P2.2, stabilizing the inter-region packing through A-minor interactions (e.g., A6033 and A6034). Bases U6044 and U6045 of L1.2B continue to stack on the PK II stack, whereas A6046-A6048 form a u-turn-like structure (27), allowing A6049 and A6050 to reach the minor groove of P2.2.

Fig. 3. Interaction of the IGR IRES RNA with the ribosome. **(A)** The PSIV IGR IRES ribosome affinity domain structure docked into the cryo-EM representation of the IRES bound to the 80S ribosome, with the 60S ribosome density computationally removed. The 40S subunit is in yellow, the cryo-EM density of the IRES is in gray, and the IRES crystal structure is colored as in Fig. 1. The positions of rpS5 (40S subunit) and SL IV and SL V (IGR IRES) are shown. **(B)** Detailed view of the interaction of the IGR IRES to the 60S subunit within the 80S ribosome-IRES complex, with the 40S subunit density computationally removed. The L1 stalk of the 60S subunit contacts IRES loop L1.1 and perhaps P1.1. For both (A) and (B), the orientation is indicated in the insets.

Fig. 4. IRES structural changes and a proposed mechanism of ribosome recruitment. **(A)** Sequences of two L1.1/P1.1 mutants. COMP, Watson-Crick complement. **(B)** Native gel analysis of these mutants shows that they do not globally misfold, because they run very close to wild-type (WT) RNA. Previously published (10) native gel analysis of mutants, in which SL IV and V were changed to GAAA tetraloops (which does not change the fold) and in which PK II was altered (which causes global misfolding), is shown for comparison at right. **(C)** Assembly assays of L1.1/P1.1 mutants analyzed on a sucrose gradient. The locations of 40S- and 80S-bound IRESs are indicated. CPM, counts per minute. **(D)** Interactions with the 40S subunit are shown in yellow, and those with the 60S subunit are shown in blue. The IRES is shown as a space-filling representation inside the corresponding gray cryo-EM density, colored as in previous figures. The folded, unbound IRES binds to the 40S subunit through SL IV and SL V interacting with rpS5, inducing a conformational change in the subunit and docking region 3 into the P site (step 1). Subsequent 60S subunit binding induces a series of conformational changes in both the IRES and the ribosome, including a putative organization of IRES structures L1.1 and P1.1

(step 2). The appearance of additional IRES density in the cryo-EM map of the 80S ribosome-bound IRES (indicated with the red oval) supports the existence of this structural switch.