Article No. jmbi.1999.3095 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 292, 513–529

JMB

The Hepatitis C Virus Internal Ribosome Entry Site Adopts an Ion-dependent Tertiary Fold

Jeffrey S. Kieft[1], Kaihong Zhou[1], Ronald Jubin[2], Michael G. Murray[2], Johnson Y. N. Lau[2] and Jennifer A. Doudna[1]*

[1]*Department of Molecular Biophysics and Biochemistry and Howard Hughes Medical Institute, Yale University, New Haven, CT 06520-8114, USA*

[2]*Schering-Plough Research Institute, Kenilworth, NJ 07033-0539, USA*

**Corresponding author*

Hepatitis C virus (HCV) contains an internal ribosome entry site (IRES) located in the 5′ untranslated region of the genomic RNA that drives cap-independent initiation of translation of the viral message. The approximate secondary structure and minimum functional length of the HCV IRES are known, and extensive mutagenesis has established that nearly all secondary structural domains are critical for activity. However, the presence of an IRES RNA tertiary fold and its functional relevance have not been established. Using chemical and enzymatic probes of the HCV IRES RNA in solution, we show that the IRES adopts a unique three-dimensional structure at physiological salt concentrations in the absence of additional cofactors or the translation apparatus. Folding of the IRES involves cooperative uptake of magnesium and is driven primarily by charge neutralization. This tertiary structure contains at least two independently folded regions which closely correspond to putative binding sites for the 40 S ribosomal subunit and initiation factor 3 (eIF3). Point mutations that inhibit IRES folding also inhibit its function, suggesting that the IRES tertiary structure is essential for translation initiation activity. Chemical and enzymatic probing data and small-angle X-ray scattering (SAXS) experiments in solution show that upon folding, the IRES forms an extended structure in which functionally important loops are exposed. These results suggest that the 40 S ribosomal subunit and eIF3 bind an HCV IRES that is prefolded to spatially organize recognition domains.

Keywords: RNA structure; translation; internal ribosomal entry site (IRES); hepatitis C virus (HCV); chemical and enzymatic probing

Introduction

Hepatitis C virus (HCV), a single-stranded RNA virus of the family Flaviviridae, is the major infectious agent leading to non-A, non-B hepatitis worldwide (Alter *et al.*, 1989; Choo *et al.*, 1989, 1990; Kuo *et al.*, 1989). The 9.5 kb genome encodes a precursor polypeptide that is processed by viral and cellular proteases to yield mature viral proteins (reviewed by Clarke, 1997). The viral RNA contains a 341 nt-long 5′ untranslated region (UTR) that is highly conserved among isolates and strains (Han *et al.*, 1991; Bukh *et al.*, 1992). Within the 5′ UTR, an internal ribosome entry site (IRES) mediates translation initiation of the viral message *via* a 5′ cap-independent mechanism (Tsukiyama-Kohara *et al.*, 1992; Wang *et al.*, 1993). A viral IRES was first identified in the picornaviruses, and they remain the most studied of the viral IRES elements (Jang *et al.*, 1988, 1989; reviewed by Jackson & Kaminski, 1995). Internal translation initiation has also been observed in the pestiviruses (Poole *et al.*, 1995), which have many secondary structural features similar to the HCV IRES (Brown *et al.*, 1992; Le *et al.*, 1995). A number of eukaryotic mRNAs have also been shown to use an IRES-mediated translation initiation mechanism (Macejak & Sarnow, 1991; Oh *et al.*, 1992; Teerink *et al.*, 1995; Vagner *et al.*, 1995; Gan & Rhoads, 1996; Bernstein *et al.*, 1997).

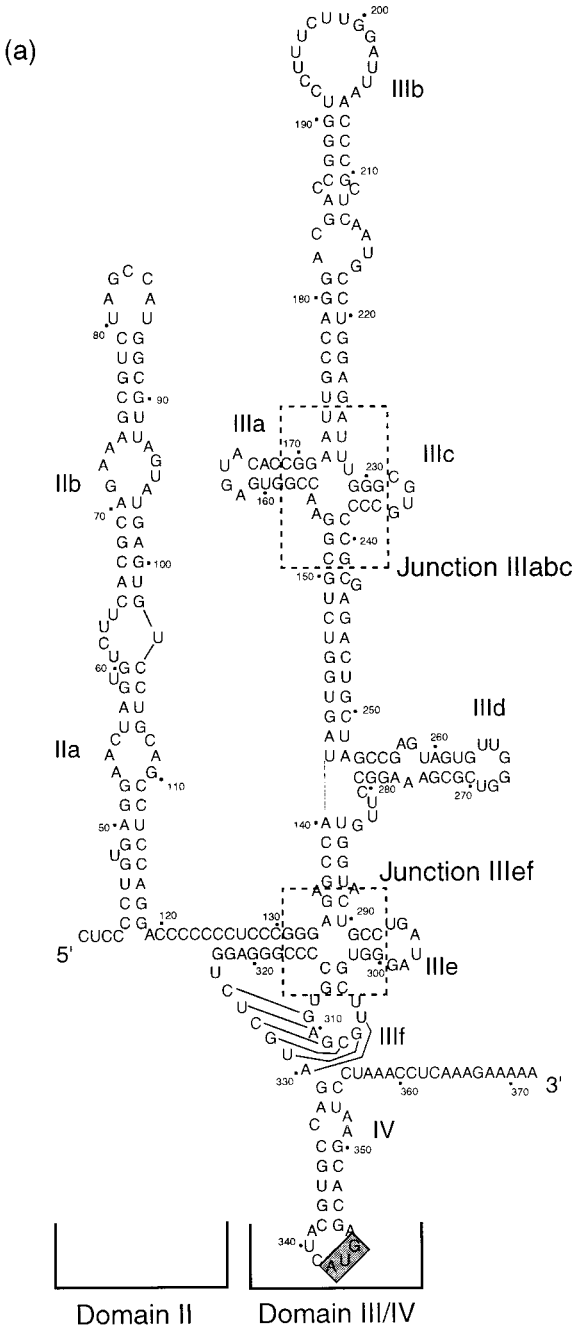
Abbreviations used: HCV, hepatitis C virus; UTR, untranslated region; IRES, internal ribosome entry site; eIF3, eukaryotic initiation factor 3; SAXS, small-angle X-ray scattering.

E-mail address of the corresponding author: doudna@csb.yale.edu

0022-2836/99/380513–17 $30.00/0

The proposed secondary structure of the HCV 5′ UTR (Brown *et al*., 1992) is highly conserved between strains and isolates (Figure 1(a)). Mutation and deletion analyses both *in vitro* and *in vivo* show that nearly the entire 5′ UTR (approximately residues 40-372) is necessary for full translation initiation activity (Fukushi *et al*., 1994; Rijnbrand *et al*., 1995; Honda *et al*., 1996b; Reynolds *et al*., 1996; Kamoshita *et al*., 1997; Tang *et al*., 1999). Within this minimal IRES, several secondary structural domains are critical for IRES function, including virtually all stem-loops, a helix between domains II and III, a proposed pseudoknot involving loop IIIf and nucleotides 325-330, the single-stranded regions that flank domain IIb, and a stem-loop containing the start AUG codon (Wang *et al*., 1994, 1995; Honda *et al*., 1996a, 1999; Varaklioti *et al*., 1998).

The detailed mechanism of IRES-directed translation initiation is unknown. Recently, Pestova and co-workers reported that assembly of preinitiation complexes on the HCV IRES involved only canonical initiation factors and the ribosomal subunits (Pestova *et al*., 1998; Sizova *et al*., 1998). Toeprinting experiments showed that parts of the HCV IRES are sites of direct



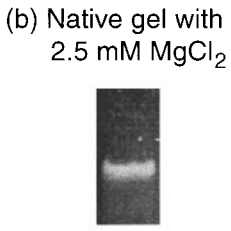


Figure 1. HCV IRES RNA sequence used for probing experiments. (a) Nucleotides 40-372 of the proposed secondary structure of the HCV IRES RNA, genotype 1b (Brown *et al*., 1992). The start codon is in the shaded box. Junction IIIabc and junction IIIef are delineated with broken boxes. (b) Portion of a native polyacrylamide gel of the sequence shown in (a) in the presence of 2.5 mM $MgCl_2$. The RNA runs as a single band, indicating an absence of aggregation or stable alternate structure. Gel conditions are described in Materials and Methods.

binding for both the 40 S ribosomal subunit and eukaryotic initiation factor 3 (eIF3) (Pestova *et al*., 1998). Other studies have identified the start AUG codon as the actual location for ribosome entry (Reynolds *et al*., 1996). Additional cellular proteins that bind the HCV IRES have been identified (Ali & Siddiqui, 1995, 1997; Yen *et al*., 1995; Fukushi *et al*., 1997; Hahm *et al*., 1998), but their role in initiating translation is unclear.

Given the high degree of sequence and secondary structure conservation of the HCV IRES RNA, it is likely that both are important functional determinants. However, it is not known how these elements recognize and position the translation machinery. One possibility is that the HCV IRES RNA is unstructured in the absence of the appropriate factor or ribosome subunit, forming a defined structure only upon binding to the translational apparatus. Alternatively, the IRES RNA might fold into a stable tertiary structure that provides an essential scaffold for the assembly of the translation machinery.

Using chemical and enzymatic probing, we show here that the unliganded HCV IRES folds into a specific three-dimensional conformation in the presence of physiological concentrations of magnesium ions, and in the absence of any additional factors. The structure contains at least two independently folded domains that coincide with the putative binding sites of the 40 S ribosomal subunit and eIF3 (Buratti *et al*., 1998; Pestova *et al*., 1998). Point mutations that prevent folding of one or more domains disrupt IRES translation initiation activity, providing strong evidence that HCV IRES function requires formation of this tertiary structure. These data together with small-angle X-ray scattering (SAXS) results suggest that the HCV IRES RNA forms an extended scaffold, in which recognition sites for ribosome and cofactor binding emanate from compactly folded helical junctions. Our results provide a structural framework for understanding interaction of the translation machinery with the prefolded IRES RNA.

Results

HCV IRES sequence forms a single species

The HCV IRES sequence used in the following experiments contains nucleotides 40-372 of genotype standard 1b (Figure 1(a)). This sequence contains the minimum IRES sequence required for translation, including 30 nt of the coding region (Fukushi *et al*., 1994; Rijnbrand *et al*., 1995; Reynolds *et al*., 1995, 1996; Honda *et al*., 1996b; Kamoshita *et al*., 1997). Native polyacrylamide gel electrophoresis in the presence of 2.5 mM magnesium shows that this sequence runs as a single, tight band, indicating a lack of aggregation or stable alternate structures (Figure 1(b)).

The HCV IRES adopts a cation-dependent tertiary structure

We tested for the presence of a stable three-dimensional fold in the HCV IRES using both chemical and enzymatic probes to examine the accessibility of the RNA backbone under different solution conditions. We anticipated that any tertiary structure of the free HCV IRES RNA was likely to depend on the presence of cations (Pan *et al*., 1993).

The Fe(II)-EDTA complex generates hydroxyl radicals that cleave the solvent-accessible backbone of nucleic acids, providing a useful probe of the change of solvent accessibility that accompanies folding of RNA molecules (Latham & Cech, 1989; Celander & Cech, 1991). If the HCV IRES RNA folds in an ion-dependent manner, we expect to see areas of the backbone become protected from hydroxyl radical cleavage upon addition of cations. Due to the small size of the hydroxyl radical, protection from cleavage requires tight backbone packing. Secondary structure alone is not sufficient to induce protection. Thus, Fe(II)-EDTA probing is a very reliable method of detecting RNA structures involving close packing of the phosphodiester backbone (Cate *et al*., 1996).

Fe(II)-EDTA-mediated cleavage of the HCV IRES in the absence of added magnesium produces a relatively uniform ladder of products (Figure 2(a), lane 7). Upon addition of 0.5-5 mM $MgCl_2$, backbone protections become visible (Figure 2(a), lanes 8-10). Ion-induced backbone protections ranging from 2 to 2.5-fold compared to the level of cleavage in the absence of magnesium ion cluster around two four-way helical junctions: the backbone within junction IIIabc, and areas within and proximal to the IIIef junction (Figure 2(b)). The protection is more modest than that reported for tRNA (four to sixfold) or the *Tetrahymena* group I intron (tenfold) (Latham & Cech, 1989), but similar to that observed for ribonuclease P RNA (Pan, 1995). We also observe a very subtle, yet reproducible, protection area in an internal loop of the IIb domain. We did not observe large areas of protection in helical regions, indicating that although the IRES is folding, this event is not accompanied by extensive regions of helical packing. Rather, tight backbone packing is occurring locally.

To complement hydroxyl radical probing we employed RNase T_1, which cuts after G residues in single-stranded regions and is considerably larger than the hydroxyl radical (11,000 Da *versus* 17 Da). While hydroxyl radical probing reveals solvent accessibility of the RNA backbone, RNase T_1 is more sensitive than hydroxyl radicals to global changes in structure that could introduce steric hindrance. The HCV IRES was subjected to partial RNase T_1 digestion under buffer conditions identical with those used for Fe(II)-EDTA cleavage reactions. Without added cation, much of the backbone is cleaved by the enzyme, indicating that much of the secondary structure is not formed or is unstable

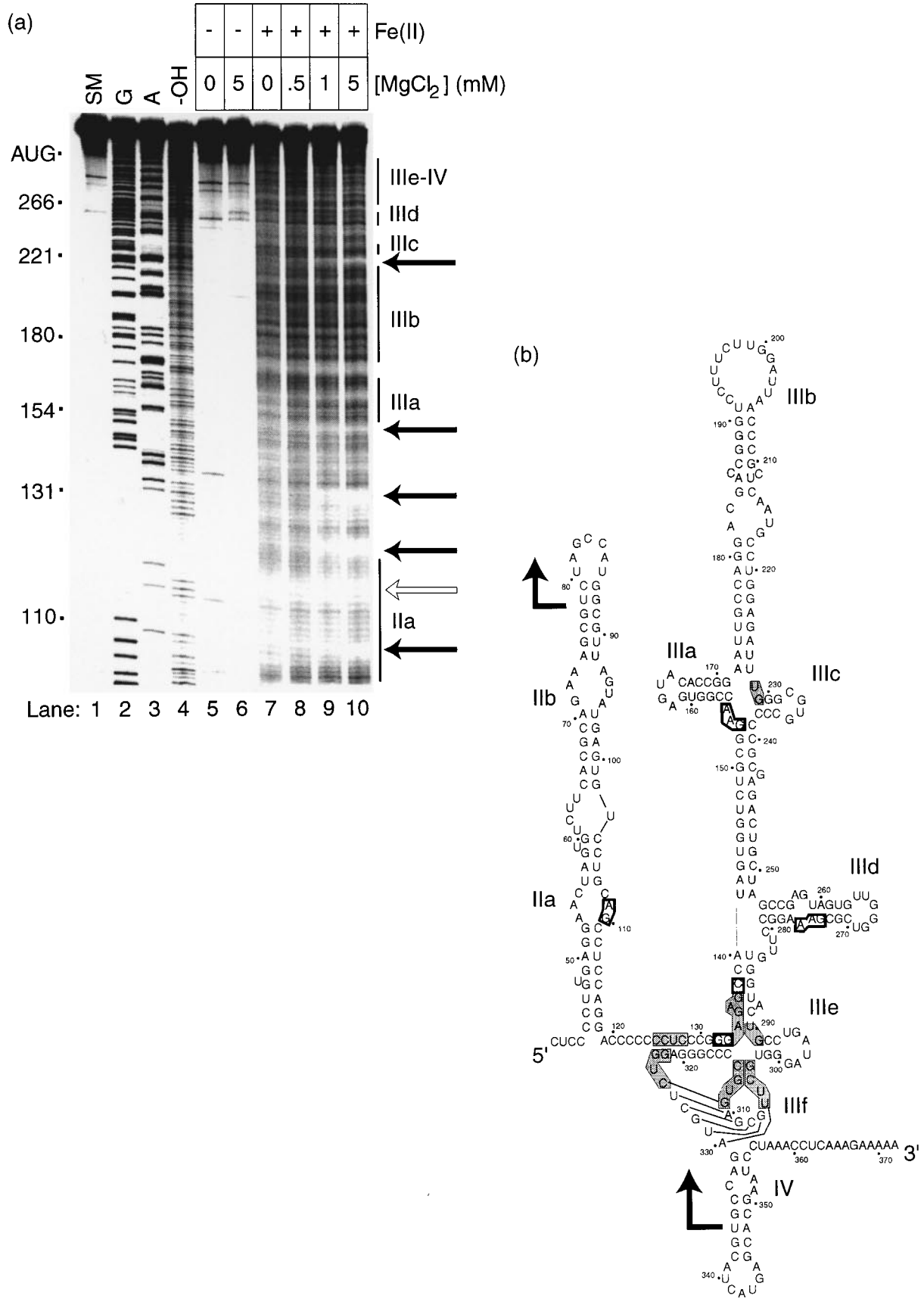


Figure 2 (*legend opposite*)

(Figure 3(a)). Upon addition of $MgCl_2$ (or other ions, see below), most of the RNA backbone becomes protected from RNase T_1 digestion, consistent with duplex formation. Several regions of the sequence remain accessible to the enzyme (Figure 3), including strong cleavage sites in hairpin loops IIb, IIIa, IIIb, IIId, and IV, and weaker cleavage sites in loop IIIc. This pattern is consistent with the proposed secondary structure. Interestingly, several of the regions predicted to be single-stranded are not cleaved by RNase T_1: loop IIIe, the sequence that bridges domain III and domain IV, an internal loop in IIa, and smaller internal loops in the IIb and IIIb helices. The protection of the bridging sequence and loop IIIf is consistent with the formation of a proposed pseudoknot (Wang *et al*., 1995), but none of the other protected regions (loop IIIe, aforementioned internal loops) have hitherto been proposed to be involved in any secondary or tertiary structure. Figure 3(b) contains a summary of all the observed protections superimposed on the HCV IRES secondary structure. It is interesting to note that the G residues between nucleotides 120 and 136 are not cleaved by the enzyme even under denaturing conditions (Figure 3(a), lane 2), implying that residual structure exists even in the unfolded state.

To determine the concentration of $MgCl_2$ required to fully fold the HCV IRES, we quantified the observed protections from Fe(II)-EDTA cleavage at nucleotide U228 and nucleotide G135, two widely separated regions of the sequence, using 0-10 mM $MgCl_2$. Folding is complete at the 10 mM magnesium ion concentration; cleavage experiments at magnesium ion concentrations up to 100 mM showed no additional protection. Uptake of $MgCl_2$ upon folding is cooperative: the Hill coefficients for the protection at U228 and G136 are 2.9 and 2.2, respectively (Figure 4(a)). The magnesium concentration at which the folding transition occurs is between 0.5 and 1 mM, with an apparent dissociation constant of 0.7 mM for the U228 protection and 0.8 mM for the G136 protection. The cooperative nature of magnesium uptake was verified using RNase T_1 (Figure 4(b)). Addition of 0.1 mM Mg^{2+} results in only slight changes in the protection pattern (Figure 4(b), lane 5) but addition of 0.25 mM Mg^{2+} results in nearly complete protection (Figure 4(b), lane 6). No further changes are observed in the presence of additional ion (Figure 4(b), lanes 7-12). Thus the changes in structure that are detectable by RNase T_1 probing occur at less than 0.25 mM $MgCl_2$, a lower ion concentration than detected by Fe(II)-EDTA. This is probably because RNase T_1 is sensitive to secondary structure formation and to global effects that require less ionic strength than tight local collapse.

To determine if folding of the HCV IRES RNA specifically requires magnesium, we conducted Fe(II)-EDTA and RNase T_1 probing experiments in the presence of other cations, including monovalent ions and polyamines. RNase T_1 protection patterns identical with those obtained in the presence of $MgCl_2$ were obtained with both spermidine and calcium ion (Figure 5(a), lanes 7 and 8). The monovalent potassium ion at 50 mM does not fully fold the RNA, although clearly some protection is induced (Figure 5(a), compare lanes 4 and 6). With increasing concentrations of KCl, the IRES folds in several steps until at 500 mM KCl, most of the protection from RNase T_1 seen with magnesium is also observed (Figure 5(b)). However, even at 500 mM KCl, there are two cleavage sites that indicate the RNA is not completely folded in the IIIe/IIIf region. We assign these bands to G295 and G303. Addition of 2.5 mM magnesium to the HCV IRES in 500 mM KCl results in a final complete folding (Figure 5(b), lane 10). Since RNase T_1 is unable to probe the formation of subtle specific structures formed in the junction regions, we used Fe(II)-EDTA to probe the RNA in the presence of polyamine. Figure 5(c) shows that 500 μM spermidine induces a protection pattern identical with that obtained with magnesium ion, suggesting that charge neutralization, rather than specific magnesium ion binding, is the primary driving force in tertiary structure formation and close backbone packing.

Point mutations disrupt HCV IRES function and structure

If the HCV IRES RNA tertiary structure is critical for translation initiation, we expect the structure to be sensitive to point mutations that disrupt function. To test this, we first identified two sets of point mutants that affect the ability of the IRES to drive internal initiation of translation in rabbit reti-

Figure 2. Fe(II)-ETDA cleavage: $MgCl_2$ induces the HCV IRES RNA to fold. (a) Fe(II)-EDTA probing of 5′ end-labeled HCV IRES RNA. Black arrows indicate where folding protects the RNA backbone (lanes 7-10). Additional protections are observed in the compressed portion of the gel corresponding to domains IIId and IIIe-IV. These protections are observable if the gel run time is lengthened (data not shown). The open arrow indicates a region where the backbone is protected in the absence of added ions, hence it is a ''background'' protection. Lanes 5 and 6 are controls without Fe(II) ion. Lanes 1, 2, 3, and 4 contain starting material (unmodified) RNA, an RNase T_1 sequencing ladder (denaturing conditions), an RNase U_2 sequencing ladder (denaturing conditions), and a hydrolysis ladder, respectively. (b) Folding-induced backbone protection from Fe(II)-EDTA is superimposed on the secondary structure of the IRES. Two black arrows indicate the limit of backbone readily analyzed using this gel system. Shaded regions indicated strong protections (2 to 2.5-fold), and open boxes delineate more subtle protections (less than twofold). The region of most extensive protection is centered on four-way junction IIIef.

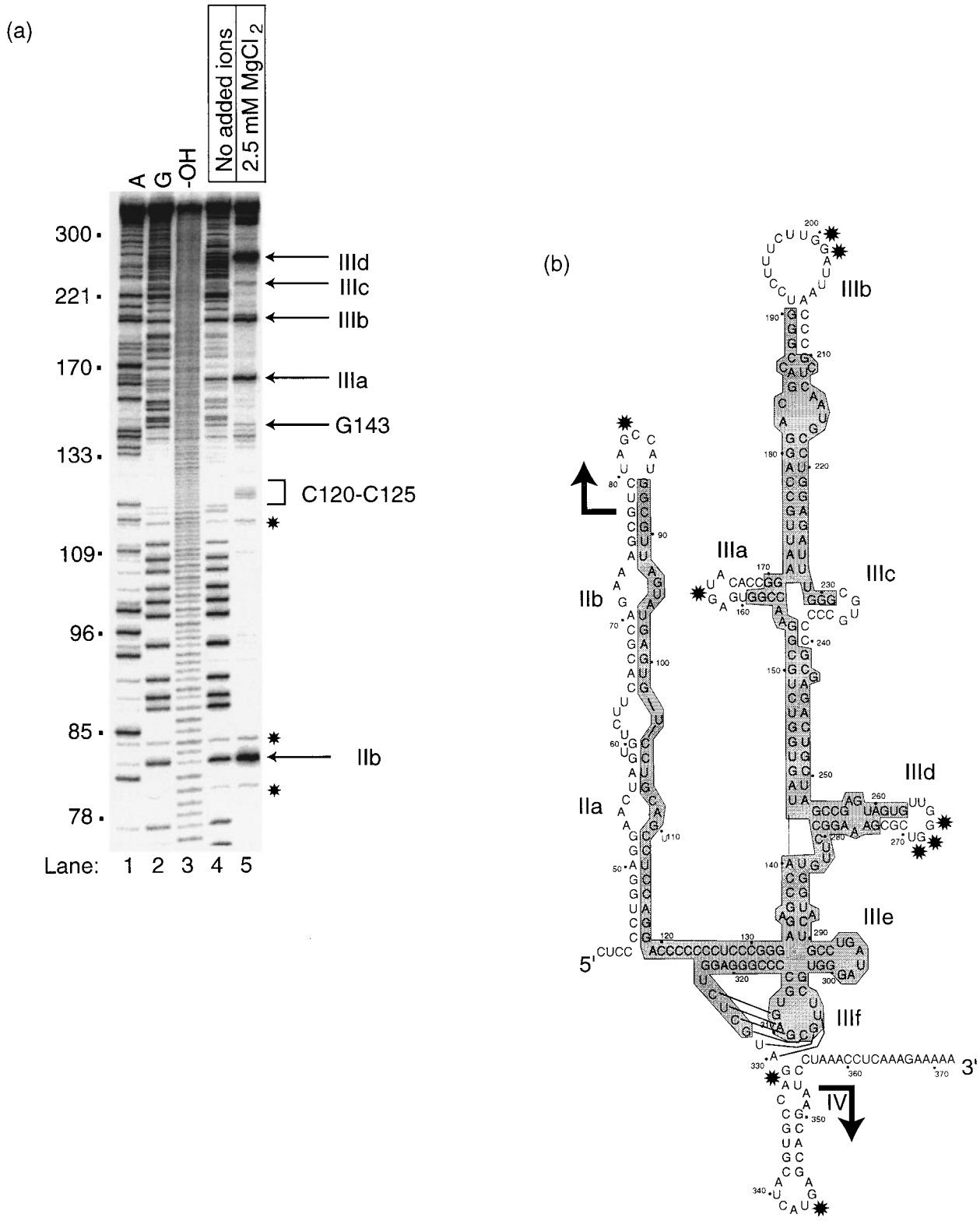


Figure 3. Ion-induced protection from RNase T_1 cleavage. (a) Partial RNase T_1 digestion of 5′ end-labeled IRES RNA shows that ion-induced folding induces protection from RNase T_1 cleavage. Lane four contains IRES RNA subjected to partial RNase T_1 digestion in the absence of added cation, lane 5 contains 2.5 mM $MgCl_2$. The locations of the apical loops IIId, IIIc, IIIb, IIIa, and IIb are shown with arrows. Small stars indicate several locations of autohydrolysis. Note that backbone region C120-C125 is prone to background hydrolysis, indicating possible flexibility. (b) Secondary structure of HCV IRES with RNase T_1 cleavage protection data overlaid. Arrows indicate the limits of our data, the shaded regions are protected from cleavage in the presence of magnesium, stars indicate locations of enhanced cleavage upon folding. Upon addition of ion, most of the backbone is not cleaved, a result consistent with formation of the proposed secondary structure (Brown *et al.*, 1992).

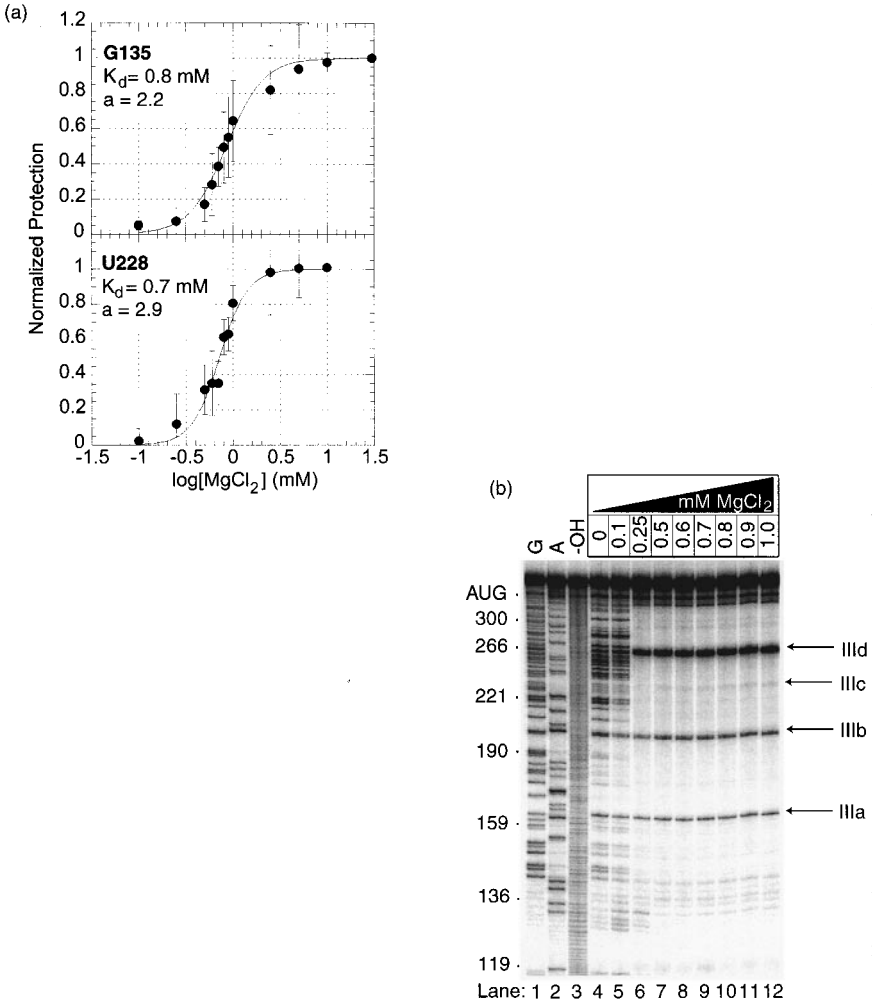


Figure 4. Ionic requirement of the HCV IRES RNA fold. (a) Normalized protection *versus* log[$MgCl_2$] plots of the folding induced protection from Fe(II)-EDTA at G136 and at U228. Normalized protection can be interpreted as the fraction of RNA folded. The data were fit to a Hill equation: $f = (L^a/K_d^a)/(1 + L^a/K_d^a)$; where f is the fraction of RNA folded, L is the free ligand concentration, K_d is the apparent dissociation constant for all sites, and a is the Hill coefficient. Error bars represent standard errors of the mean from three independent experiments. Inset: The apparent dissociation constant and Hill coefficient obtained from a non-linear least squares fit. (b) Protection from RNase T_1 shows the cooperative uptake of magnesium that accompanies folding. Probing experiments with magnesium concentrations of 0.1 mM and below (lane 4 and 5) do not produce the final protection pattern but at 0.25 mM $MgCl_2$ (lane 6), the cleavage pattern of the folded IRES is obtained. Addition of more magnesium does not result in any further changes in the cleavage pattern (lanes 7-12).

culocyte lysate (Table 1). The first set contains two clinical isolates: JL15.36 and JL15.26. JL15.36 differs from wild-type genotype 1b in two positions (U204A and U340C) but retains full activity. JL15.26 contains both of the mutations found in JL15.36, plus one additional mutation, U228C, in the IIIabc junction. This mutation results in a >95 % inhibition of HCV IRES activity. The second set of mutants contained point mutations in loop IIId, a highly conserved loop which has enhanced susceptibility to RNase T_1 cleavage upon folding. Mutants G266C, G267C, G268C, and G(266-268)C exhibited nearly a complete loss of activity, while mutants U264A, U265A, and U269A exhibit an intermediate loss of activity (Table 1).

To determine if point mutations that affect function also affect IRES tertiary structure, we conducted probing experiments on the two junction IIIabc mutants and five of the loop IIId mutants. Fe(II)-EDTA and RNase T_1 protection patterns of fully functional mutant JL15.36 were indistinguishable from wild-type IRES. However, introduction of the U228C mutation (JL15.26) results in the loss of magnesium-dependent protection from Fe(II)-EDTA of the IIIabc junction, including the regions from nucleotides 153-155 and nucleotides 228-229. Protections in other regions of the backbone, such as those around nucleotide 136 and nucleotide 126 are not diminished (Figure 6(a)). The U228C mutation also changes the RNase T_1 protection pattern in several significant ways (Figure 6(b)). First, there are changes in the cleavage pattern obtained in the absence of magnesium ion. Most notable are the changes in the region between nucleotides 210 and 224, which is cleaved less in the construct containing U228C. In the presence of 2.5 mM $MgCl_2$, the cleavage patterns are very similar, with the exception of nucleotides G158, G159, and G161. RNase T_1 cleaves these in the presence of magnesium. Furthermore, the JL15.36 and JL15.26 RNAs have differing mobilities in a native polyacrylamide electrophoretic gel containing 2.5 mM magnesium chloride, implying that the mutation has altered the three-dimensional shape of the molecule (Crothers & Drak, 1992; Batey & Williamson, 1998). The U228C mutation present in mutant JL15.26 results in decreased mobility and a smearing of the band compared to JL15.36 (Figure 7).

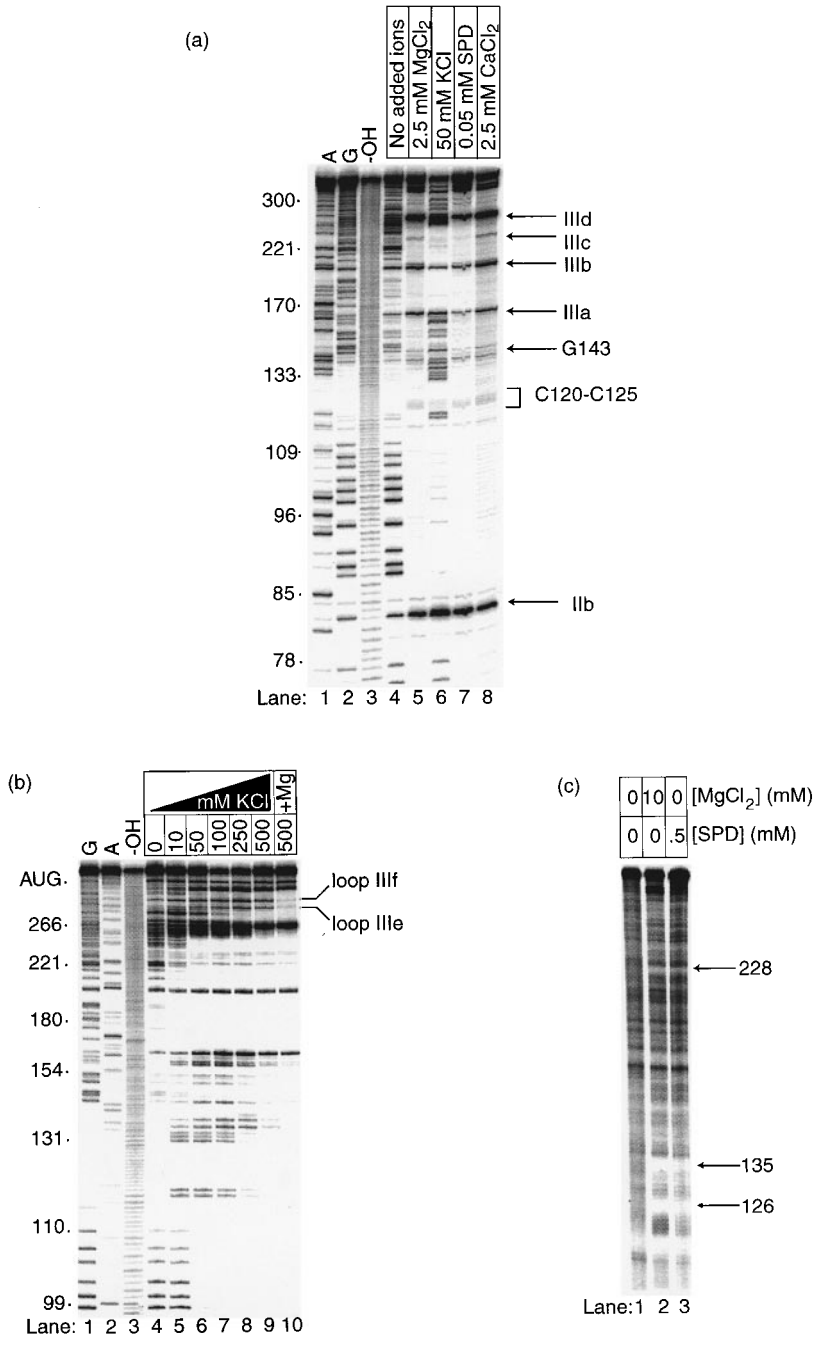


Figure 5. Ability of various cations to fold the HCV IRES. (a) Partial RNase T_1 digestion of 5′ end-labeled HCV IRES RNA in the absence of added cation (lane 4), with 2.5 mM magnesium (lane 5), 50 mM KCl (lane 6), 0.05 mM spermidine (lane 7), and 2.5 mM $CaCl_2$ (lane 8). Spermidine and calcium are able to produce cleavage patterns nearly identical to that obtained with magnesium. (b) Folding of the RNA by increasing amounts of KCl occurs along a much broader concentration range than with magnesium. The RNA goes through several secondary structure changes. For example, between nucleotides 99-110, a transition occurs between 10 and 50 mM KCl (lane 5 and 6), and between nucleotides 131-154, a transition is seen at 250-500 mM KCl (lanes 8 and 9). At 500 mM KCl (lane 9), there is cleavage in the loop IIIe and IIIf regions that indicates incomplete folding. These bands disappear upon addition of $MgCl_2$ (lane 10). (c). Folding of the HCV IRES in the presence of spermidine induces protection from Fe(II)-EDTA cleavage. The protection in the presence of 0.5 mM spermidine (SPD, lane 3) is virtually identical with that with 10 mM $MgCl_2$ (lane 2). Locations of protection are shown with arrows.

Loop IIId is critical for function (Table 1). To determine if loop IIId is important for folding, we subjected loop IIId mutants U265A, G266C, G267C, G268C, and G(266-268)C to RNase T_1 cleavage experiments, again probing both in the presence and absence of $MgCl_2$. The loop IIId mutants exhibited protection pattern changes that differed greatly depending on the positions of the mutations (Table 1). Mutant G266C had the most extensive changes (Figure 8(a), compare lanes 5 and 7). Without $MgCl_2$, the mutant RNA is cleaved in a pattern similar to wild-type IRES. However, introduction of the mutation causes a profound effect on the RNase T_1 cleavage pattern in the presence of $MgCl_2$. Specifically, nucleotides in loop IIId, IIIe, and the intervening sequence are still cleaved by the enzyme in the presence of magnesium. Thus, the sequence of loop IIId is affecting the structure of downstream elements of the IRES, including loop IIIe. The structural effect of mutating a neighboring nucleotide, G267C, is much more subtle. This mutant does not display dramatic changes in the protection pattern in either the folded or the unfolded states (Figure 8(a), lanes 8 and 9), indicating that this nucleotide does not play a critical role in promoting proper structure. Yet another loop IIId mutant, G268C, has a cleavage pattern different from either G266C or G267C

Table 1. Activities (relative to wild-type) of mutants used in this study

Mutant	Activity (%)
Wild-type	100
U264A	13
U265A	47
G266C	3
G267C	2
G268C	3
U269A	40
G(266-268)C	2
JL15.36 (U204A, U340C)	>100[a]
JL15.26 (U204A, U340C, U228C)	5[a]

[a] Determined with bicistronic assay. All other mutants were analyzed using a monocistronic assay (see Materials and Methods). The use of different reporter systems may nominally affect the measured activity values, but does not affect the overall result.

(Figure 8(a), lanes 10 and 11). Unlike G266C or G267C, this mutant has a different cleavage pattern in the absence of added magnesium from that of wild-type IRES. Specifically, nucleotides G253, G256, G258, and G271 are more susceptible to cleavage than other G residues in the sequence. Upon addition of $MgCl_2$, the cleavage pattern becomes similar to wild-type, with only very subtle differences. Clearly, loop IIId mutants perturb the structure well beyond this stem-loop (Figure 8(b)).

SAXS of the HCV IRES reveals an extended structure

To determine how the addition of magnesium affects the global size and shape of the wild-type HCV IRES, we conducted small-angle X-ray scattering experiments in buffers containing 0 and 2.5 mM $MgCl_2$. These two conditions represent the folded and unfolded states of the IRES RNA. As a control, we also analyzed the P4-P6 domain of the *Tetrahymena* group I intron under identical buffer conditions (Murphy & Cech, 1993). The P4-P6 domain folds in a magnesium ion-dependent manner into a compact structure that has been deter-

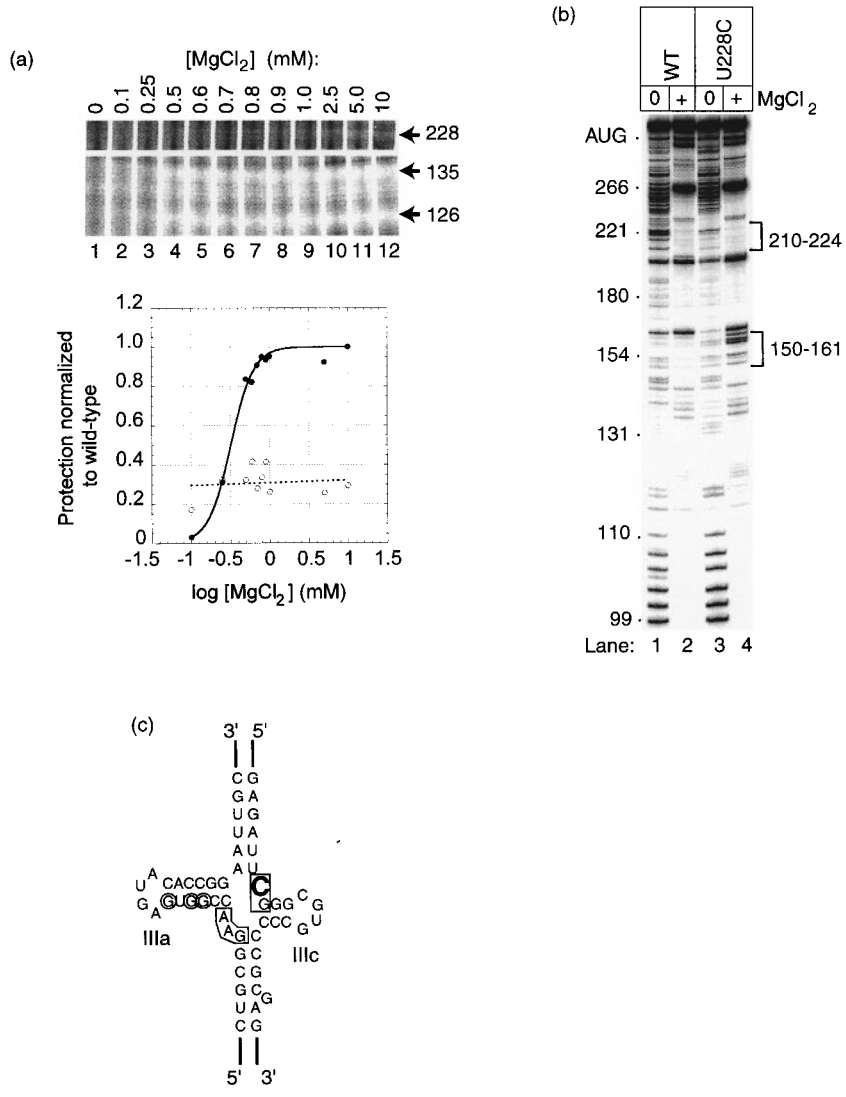


Figure 6. Effect of mutation U228C on HCV IRES structure. (a) Fe(II)-EDTA cleavage of U228C in the presence of increasing magnesium shows a visible loss of protection at nucleotide 229, but the protections at nucleotide 135 and 126 are not affected. A plot of protection *versus* log[$MgCl_2$] at nucleotides 228 (open circles) and 136 (filled circles) in mutant JL15.26 for this gel is shown. The degree of protection of the mutants was normalized against the degree of protection of wild-type HCV IRES, and the data was fit to a Hill equation. It was not possible to fit nucleotide 228 to the equation. (b) RNase T_1 cleavage patterns of wild-type and mutant RNA reveal changes both without $MgCl_2$ (lanes 1 and 3), and with 2.5 mM $MgCl_2$ (lanes 2 and 4). Introduction of the U228C mutation changes the secondary and tertiary structure of the RNA. (c) The protection changes of the folded RNA observed upon introduction of the U228C mutation (shown as a large outlined C residue) are overlaid on a portion of the HCV IRES secondary structure. Boxed residues lost protection from Fe(II)-EDTA upon introduction of the mutation and circled residues became accessible to RNase T_1.

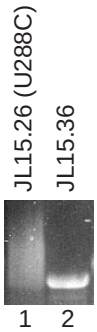


Figure 7. Native gel electrophoresis of mutants JL15.26 and JL15.36. JL15.36 migrates as a tight single band, while JL15.26 (U228C) migrates slower and is smeared.

mined by X-ray crystallography (Cate *et al*., 1996). Figure 9 shows a Guinier plot of P4-P6 scattering with and without magnesium. In the absence of magnesium ion, the curve possesses an increasing slope as Q^2 approaches zero (zero scattering angle), indicating possible structural heterogeneity. Upon addition of magnesium, the Guinier plot becomes more linear, as observed in the Guinier plots of other globular RNAs (Pilz & Kratky, 1970; Wilhelm & Pilz, 1982). The change in the small-angle scattering of the P4-P6 domain upon addition of magnesium is interpreted as a collapse of the molecule from a heterogeneous, extended structure to a folded, compact state. This collapse is accompanied by a decrease in the measured radius of gyration (R_g) from 40 Å to 30 Å and a decrease in the measured maximum interatomic distance (D_{max}) from 140 Å to 100 Å. These values agree with those determined from the crystal structure ($R_g = 33$ Å, $D_{max} = 110$ Å). Thus, magnesium-induced folding is accompanied by a radical global conformational change that is readily detectable by small-angle scattering.

We conducted an identical analysis of scattering data obtained from wild-type HCV IRES RNA in the absence and presence of magnesium. In the absence of the magnesium ion, the measured scattering of the HCV IRES yields a Guinier plot in which the slope increases dramatically as the zero angle limit is approached. Upon addition of magnesium, the scattering changes, but does not become linear. Rather, the curve is similar to the curve obtained for the unfolded IRES in that the slope of the curve is increasing as zero scattering angle is approached. The steep slope near the zero scattering angle might be interpreted as aggregation; however, native gel electrophoresis under identical conditions does not show any higher molecular mass species (data not shown). We were unable to derive values for R_g or D_{max} from this data, possibly due to the inability to accurately measure intensity at the zero scattering angle limit, or the fact that extended structures are poor scattering particles. Despite these limitations, it is clear that addition of magnesium is inducing a global change in the IRES which is different in character from that induced in the P4-P6 RNA. The P4-P6 domain folds into a globular shape with a clearly defined cross-section and tightly packed helices, whereas the IRES does not.

Discussion

The HCV IRES RNA forms an ion-dependent tertiary fold in solution

Like other viral IRES elements, the HCV IRES contains sequence and secondary structural elements that are highly conserved among different strains. There are at least two components of the preinitiation complex that bind directly to the HCV IRES RNA: the 40 S ribosomal subunit and initiation factor 3 (eIF3). It has been unclear how the IRES RNA recognizes these components. One possibility is that the IRES RNA is unstructured when not bound. Alternatively, the IRES RNA could fold into a stable tertiary structure that presents recognition sites to the translational apparatus.

To test these possibilities, we performed chemical and enzymatic probing of the IRES RNA in solution under various ionic conditions. Our results show that the RNA has a unique and stable three-dimensional structure at physiological salt concentrations. Furthermore, the concentration of magnesium ion required to fold the IRES is similar to the concentration needed for maximum IRES-mediated translation *in vitro* (Borman *et al*., 1995). Our data indicate that specific divalent metal ions are not required, since we observe identical protection patterns in the RNA backbone with either magnesium or spermidine, and protection of G residues from RNase T_1 cleavage is observed with a variety of polyvalent cations. Interestingly, high concentrations of potassium ion were insufficient to induce complete folding of the RNA, as assayed by RNase T_1 cleavage. The loop IIIe stem-loop, one of the regions that emanates from the tightly folded junction IIIe/f (Figure 1), became protected only upon the addition of magnesium ions (Figure 5(b)).

It is instructive to compare the HCV IRES to the P4-P6 domain of the *Tetrahymena* group I intron, an RNA whose magnesium-dependent fold has been well documented. In the P4-P6 domain, five magnesium ions coordinate directly to the RNA to form a metal ion ''core'' (Cate *et al*., 1997). Furthermore, monovalent ions bind specifically within a structural motif called the adenosine platform (Basu *et al*., 1998). In contrast, our solution studies of the HCV IRES provide no evidence for such highly specific RNA-metal ion coordination. Rather, the IRES three-dimensional fold requires non-specific charge neutralization in physiological

concentrations of magnesium (Alberts *et al*., 1994) or polyamines.

HCV IRES structure is not globular

The localized protection of the HCV IRES RNA from Fe(II)-EDTA cleavage demonstrates that while there are regions of close backbone packing, the IRES architecture is not globular. The tertiary fold of the IRES is different in character from the solution structure of the *Tetrahymena* group I intron, although these two RNAs have similar size and extent of secondary structure. In the intron, RNA folding is accompanied by close packing of helical regions such that approximately 40% of the backbone is protected from cleavage by Fe(II)-EDTA (Latham & Cech, 1989; Celander & Cech, 1991; Cate *et al*., 1996; Golden *et al*., 1998). In contrast, only 11% of the HCV IRES backbone becomes protected upon folding, and the areas of protection cluster in or near two four-way helical junctions (IIIabc and IIIef). Furthermore, RNase T_1 cleavage in loops IIb, IIIa, IIIb, and IIId is consistent with exposure of these loops to solution, rather than burial within a globular structure.

Small-angle X-ray scattering (SAXS) reveals the size and shape of a macromolecule in solution (reviewed by Moore, 1982). SAXS has been used to determine the size and shape of RNA molecules such as tRNA and ribosomal RNAs (Pilz & Kratky, 1970; Leontis & Moore, 1984), and to investigate the effect of metal ions on RNA structure (Wilhelm & Pilz, 1982). In general, smaller globular or rod-like RNAs have scattering profiles amenable to rigorous size and shape determination. SAXS of the HCV IRES RNA reveals a scattering profile that is not consistent with a globular molecule. Since native gel electrophoresis under identical conditions did not reveal any higher molecular mass species, we interpret the curve as indicative of an extended structure with some flexibility. This is consistent with the lack of extensive backbone protection from Fe(II)-EDTA cleavage.

Two independent domains within the fold

We identify two independently folded domains within the extended IRES structure, based on probing of point mutants that affect IRES function. The first two mutants (JL15.36 and JL15.26) were selected to explore the effect of a mutation in junction IIIabc on the global fold of the IRES. This junction contains five protections from Fe(II)-EDTA cleavage. Introduction of mutation U228C decreases IRES activity to 5% of wild-type and causes a loss of protection from both hydroxyl radicals and RNase T_1 in and directly adjacent to junction IIIabc, but nowhere else. Hence, junction IIIabc comprises an independently folded region of the RNA.

Nucleic acid junctions are known to fold in the presence of metal ions and arrange the emanating helices at specific angles (Mollegaard *et al*., 1994; Bassi *et al*., 1995, 1996; Duckett *et al*., 1995; Batey & Williamson, 1998; Walter *et al*., 1998). Misfolding of junction IIIabc could cause the interhelical angles of the emanating stems to change, detectable by native gel electrophoresis (Crothers & Drak, 1996; Batey & Williamson, 1998). Consistent with this possibility, mutation U228C caused retardation and smearing of the HCV IRES RNA on a native gel. Smearing and mobility retardation suggest that the mutant RNA adopts a ground-state conformation different from wild-type, and is exchanging between different conformers. Such structural changes could inhibit the ability of the IRES to bind the translational machinery, explaining the loss of translation initiation activity by this mutant.

While initial RNase T_1 probing of wild-type HCV IRES suggested that loop IIId is not part of a direct tertiary contact, point mutations in the loop perturb the IRES secondary and tertiary structure. Changes in protection from RNase T_1 appear in the IIId stem, and in the IIIe and IIIf loops, but there are no protection changes in more distal regions of the RNA. This implies the presence of a second independently folded domain, consisting of stem-loop IIId, loop IIIe, loop IIIf, and the proposed pseudoknot. This domain is more extensively protected from solvent than junction IIIabc.

The effects of loop IIId mutations on the protection of loops IIIe and IIIf demonstrate that this stem-loop is involved in the folding of the IRES RNA. The specific role of loop IIId is not yet clear. In the mutants, local secondary structural changes or destabilization around loop IIId might propagate to affect RNase T_1 protection in loops IIIe and IIIf. It is also possible that these loops interact directly via a long-range tertiary contact.

Functional implications of the IRES global architecture

The global architecture of the unliganded HCV IRES RNA that we have elucidated provides a structural explanation for much of the biochemical, functional and biophysical data previously reported. The translation efficiency and integrity of the tertiary structure are strongly correlated, since mutants that assume the correct tertiary fold have wild-type translation initiation activity, while mutants with perturbed structures have reduced activity. Furthermore, the locations of independently folded regions within the overall extended structure of the IRES correspond to binding sites for the 40 S subunit and eIF3 (Figure 10). Toeprinting of these complexes places eIF3 on the independently folded region that includes IIIb and junction IIIabc, and the 40 S subunit on the region surrounding junction IIIef (Pestova *et al*., 1998). Our probing experiments show that the stem-loop that contains the start AUG codon is highly susceptible to RNase T_1 cleavage and autohydrolysis in the presence of $MgCl_2$ (data not shown). This suggests that the stem-loop is extended away from the folded core, and that its secondary structure is

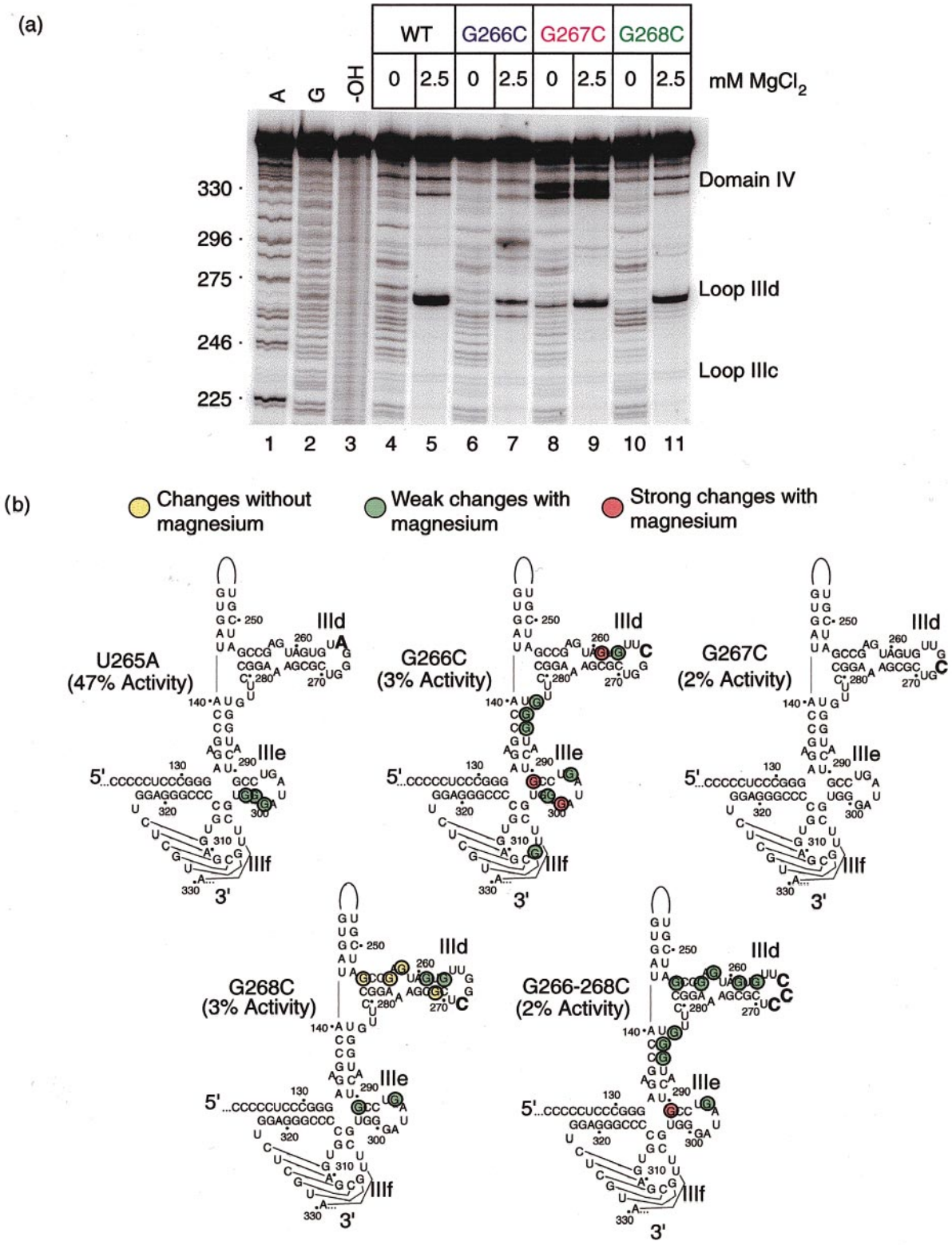


Figure 8 (*legend opposite*)

unstable, perhaps making the start codon readily available for recognition by the decoding site of the ribosome. This correlates well with the fact that destabilization of this stem-loop leads to enhanced IRES activity (Honda *et al.*, 1996a).

Our results show that the HCV IRES RNA folds into a specific structure which is required for recognition by the host translational machinery. Assembly of the preinitiation complex may involve simple docking of this structure with the 40 S ribo-

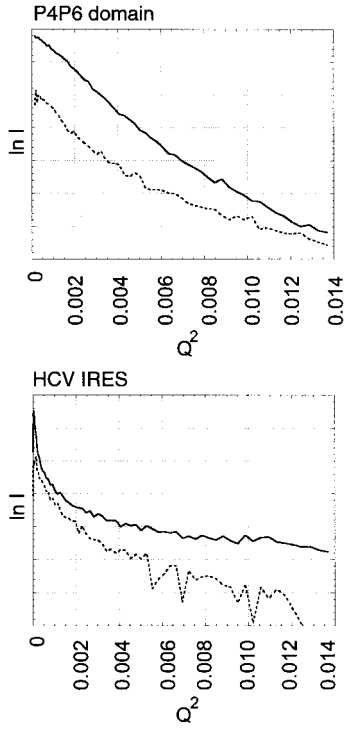


Figure 9. Guinier plots of RNA with and without magnesium. Guinier plots of the solution small-angle X-ray scattering of the P4-P6 domain of the *Tetrahymena* group I intron (left) and of nucleotides 40-372 of the HCV IRES RNA (right) in the absence (broken line) and presence (continuous line) of $MgCl_2$.

somal subunit and eIF3. Alternatively, binding of these components may induce additional folding, as has been demonstrated for the interaction between CBP2 and the group I intron (Weeks & Cech, 1996) and ribosomal protein S15 and its binding site within 16 S ribosomal RNA (Batey & Williamson, 1998). Recognition of the IRES structure may also involve other, as of yet unidentified, cellular proteins. Our results provide a structural framework for experiments to explore these possible binding strategies, and ultimately to understand the mechanism of internal ribosome entry.

Materials and Methods

Plasmid construction for translation assays

All plasmids (except JL15.36 and JL15.26) contain HCV IRES derived from plasmid pK1b, which contained a HCV cDNA (genotype 1b) fragment spanning nucleotides 1-779. To facilitate cloning into additional vectors, an *Ehe* 1 site at was introduced at nucleotide 1 of the HCV IRES. The resultant plasmid (pK1b(e)) served as the template for construction of loop IIId mutants, which were constructed by insertion of oligonucleotides with compatible termini to the *Nhe*I and *Stu*I sites which span IIId. Wild-type (genotype 1b) and loop IIId mutant constructs retained the HCV core protein as the translation reporter. Plasmids JL15.26 and JL15.36 HCV IRES elements were identified from a patient chronically infected with HCV genotype 1b infection. Briefly, total RNA is extracted from 100 µl of serum according to a described protocol (Lau *et al.*, 1993). Reverse transcription and PCR were performed with appropriate primers (sense: GACACTCCACCATGAATC; antisense: GGGAATTCAACTTGACGTC) to generate an amplicon that flanked nucleotides 20-426, which was cloned into pCR3.1 (Invitrogen). For activity determination, the IRES element was cloned into a bicistronic construct driven by phage T7 in which translation of the first reporter (chloramphenicol acetyltransferase) is IRES independent and the second reporter (luciferase) is IRES dependent as detailed by Feller & Lau (1998).

Transcription reactions for translation assays

For wild-type and loop IIId mutants, purified plasmid DNA was linearized with *Avr*II. Linearized plasmids were repurified prior to transcription with GeneClean II kit (BIO 101) and resuspended in sterile RNase-free water. Transcription reactions were carried out using T7 Ribomax kits (Promega) according to manufacturer's protocol. Following treatment with DNase to remove the template DNA, the RNA was purified using RNeasy Midi kits (Qiagen). Purified RNA was suspended in sterile RNase-free water, quantified by absorbance at 260/280 nm, analyzed by agarose gel electrophoresis to examine quantity and integrity of RNA transcripts, and was diluted in sterile water to a final concentration of 0.1 mg/ml.

Translation assays

Purified RNA (0.5 µg) was added to rabbit reticulocyte lysate (Promega) with 20 µCi of [^{35}S] methionine. Translation mixes were incubated at 30 °C for 90 minutes. Following incubation, 4 µl of the translation reac-

Figure 8. RNase T_1 cleavage patterns of loop IIId mutants compared to wild-type IRES RNA. (a) Portion of a denaturing gel containing partial RNase T_1 digestions of wild-type IRES RNA and three loop IIId mutants. Lanes 4 and 5 contains wild-type IRES in the absence of added ion and in the presence of 2.5 mM $MgCl_2$, respectively. Lanes 6 and 7 contain mutant G266C RNA cleaved under identical conditions. The strong cleavage location in domain IV are degradation products that were not observed in repeated experiments. Changes in the cleavage pattern are summarized on a portion of the secondary structure (see (b)). Mutants G267C and G268C are shown in lanes 8 and 9, and lanes 19 and 11, respectively. (b) The cleavage pattern changes of all loop IIId mutants that were probed with RNase T_1 are summarized on a portion of the secondary structure. Yellow circles denote changes in the accessibility of nucleotides in the absence of $MgCl_2$, green circles indicate nucleotides affected in the presence of the ion to only a very subtle degree (just visible on the gel in (a)). Red circles are strong, clearly visible changes in the cleavage pattern. Mutations are denoted with large bold type.

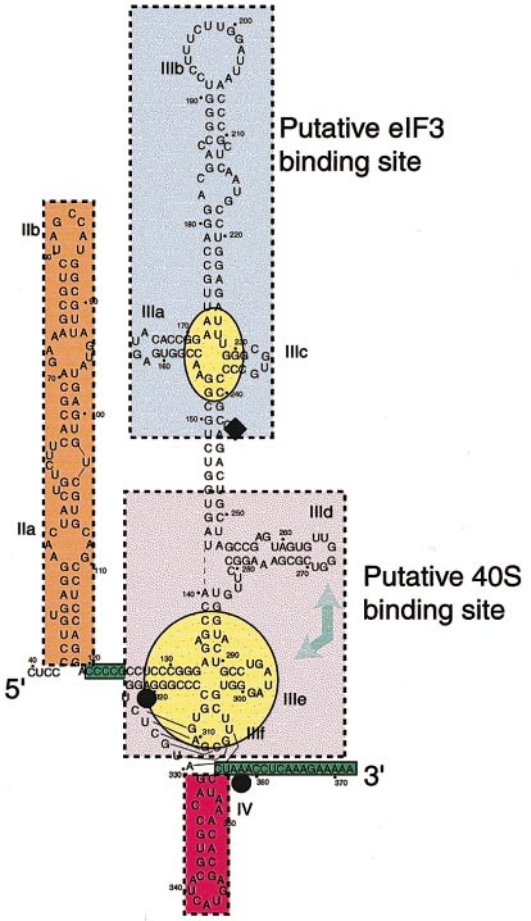


Figure 10. The global architecture of the HCV IRES correlates with function. Putative eIF3 (blue) and 40 S subunit (pink) binding sites correspond to independently folded domains that contain regions of tightly packed RNA backbone. These tightly packed cores are shown in yellow. Primer extension stops for bound 40 S subunit (black circles) and eIF3 (black diamonds) are included (Pestova *et al*., 1999). Alteration of sequence within loop IIId perturbs structures in downstream regions, indicating a possible tertiary contact (shown as a green arrow). The stem-loop that contains the start codon (domain IV) is shown in red; this secondary structure is unstable, consistent with earlier studies linking destabilization of this loop to increased IRES activity (Honda *et al*., 1996a). Sequences shaded in green indicate flexible single-stranded regions. The flexible sequence coding region does not participate in folding; the single-stranded region that bridges domain II (orange) and domain III may provide a flexible linker.

tion was removed and transferred to a tube containing 35 µl of Laemmli sample buffer. Samples were boiled for five minutes and transferred to ice to cool. A 20 µl aliquot of the translation samples was analyzed by SDS-PAGE (20 %) at constant current (10 mA) overnight. Dried gels were visualized on a phosphorimager. Protein bands corresponding to the appropriate reporter protein were selected and quantitated using ImageQuant software. Percentage activity/inhibition studies were repeated three times and values averaged.

Plasmid construction for all other experiments

In order to generate RNA with homogeneous 5′ and 3′ ends, *cis*-acting ribozymes were included in all constructs (Ferré-D'Amaré & Doudna, 1996). A hammerhead ribozyme was used to cleave the 5′ end and a hepatitis δ ribozyme was used to cleave the 3′ end. To generate the plasmids used in *in vitro* transcriptions, the above described parent plasmids were used as templates in PCR reactions to generate inserts with the T7 polymerase promoter, *cis*-acting ribozymes, and *Eco*RI and *Bam*HI restriction sites. These cDNAs were cloned pUC-19 and amplified in either HB101 (Bio-Rad) or DH5α (Gibco-BRL) cell lines. Individual colonies were picked and the plasmid DNAs were sequenced.

RNA transcription and purification

Plasmids were linearized with *Bam*HI and used as templates in *in vitro* transcription reactions. RNA sequences self-cleaved at both the 5′ and 3′ ends of the IRES during the transcription reaction. Reactions were ethanol precipitated and purified on denaturing 6 % polyacrylamide gels. The product band was located by UV shadowing, excised, and passively eluted overnight into DEPC-treated water. RNA was concentrated and washed using ultrafiltration (Amicon) and stored at 4 °C.

Chemical and enzymatic probing

RNA was 5′ end-labeled and purified on a denaturing 10 % polyacrylamide gel, visualized by autoradiography, excised, and passively eluted into 0.5 M sodium acetate (pH 5.2), 0.1 % (w/v) SDS. Labeled RNA was ethanol precipitated and resuspended in 1 × TE buffer for use in subsequent reactions. For Fe(II)-EDTA experiments, at least 100,000 CPM labeled RNA was annealed by heating to 65 °C for one minute then cooled to room temperature. RNA was then added to a tube containing buffer (final concentration of 30 mM Hepes (pH 7.5)), 1 µg tRNA, and desired salt to a final volume of 7 µl, then incubated at 37 °C for five minutes to achieve folding equilibrium. The use of other annealing protocols or temperatures did not effect the probing results. To initiate the cleavage reaction, 1 µl of Fe(II)-EDTA solution (20 mM sodium EDTA, 10 mM ferrous ammonium sulfate), 1 µl of 50 mM sodium ascorbate, and 1 µl of 1 % hydrogen peroxide were added sequentially to the reaction tube and the reaction was incubated at 37 °C for two minutes. Reaction was quenched with 1 µl of 100 mM thiourea and 10 µl of 9 M urea/1 mM EDTA and resolved on 10 % polyacrylamide sequencing gels. They were dried, visualized, and quantified on a phosphorimager. For RNase T_1 probing, RNA was annealed, added to a buffer/salt/tRNA solution, followed by incubation at 37 °C for five minutes. Cleavage reactions were initiated by addition of 1 µl of 0.1 units/µl RNase T_1 (Boehringer Mannheim), incubated at 37 °C for seven minutes, then quenched with addition of 10 µl of 9 M Urea/1 mM EDTA and placed on ice. Reactions were resolved on 10 % polyacrylamide sequencing gels which were dried and visualized as described above.

Native gel electrophoresis

Non-denaturing gels consisted of 6 % polyacrylamide, 1:19 acrylamide:N-N′-methylene-bis-acrylamide ratio, with a 66 mM Hepes, 33 mM Tris-HCl buffer and

2.5 mM $MgCl_2$. For each reaction, 10 μg of IRES RNA was suspended in 30 mM Hepes (pH 7.5), 2.5 mM $MgCl_2$, 5% (v/v) glycerol and heated to 65°C for one minute. Reactions were cooled to room temperature and loaded directly onto the gel. The gel was run at constant power (10 W) for 24 hours, followed by staining with ethidium bromide and visualization on a transilluminator.

Small angle X-ray scattering

RNA samples for use with SAXS were concentrated and the buffer was exchanged using Amicon Centricon concentrators. The buffer was 30 mM Hepes (pH 7.5) with any additional salts added, and the final RNA concentration was approximately 1 mg/ml. Concentrator flow-through was collected and used as the buffer blank during data collection. A detailed description of the SAXS apparatus, collection, and data reduction is found elsewhere (Bu *et al*., 1998). Briefly, buffer scattering data was subtracted from sample data and the resulting profile was circularly averaged. Guinier plots of the scattering profiles were constructed by plotting ln I *versus* $Q^2 \cdot Q = 4\pi \sin(\Theta)/\lambda$, where Θ is the scattering angle and λ is wavelength of the incident X-rays. I is the measured scattering intensity as a function of the scattering angle.

Acknowledgements

We are grateful to Professor A. Nomoto for the gift of plasmid pKIb, to Professor D. Engelman and Dr Z. Bu for assistance with SAXS data collection and interpretation, and to Dr A. Ferré D'Amaré, Dr R. Batey, and L. Doherty for useful discussions and comments.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, L. D. (1994). Editors of *Molecular Biology of the Cell*, 3rd edit., Garland Publishing, New York.

Ali, N. & Siddiqui, A. (1995). Interaction of polypyrimidine tract-binding protein with the 5′ noncoding region of the hepatitis C virus RNA genome and its functional requirement in internal initation of translation. *J. Virol.* **69**, 6367-6375.

Ali, N. & Siddiqui, A. (1997). The La antigen binds 5′ noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation. *Proc. Natl Acad. Sci. USA,* **94**, 2249-2254.

Alter, H. J., Purcell, R. H., Shih, J. W., Melpolder, J. C., Houghton, M., Choo, Q.-L. & Kuo, G. (1989). Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. *New Engl. J. Med.* **321**, 1494-1500.

Bassi, G. S., Mollegaard, N. E., Murchie, A. I., von Kitzing, E. & Lilley, D. M. J. (1995). Ionic interactions and the global conformations of the hammerhead ribozyme. *Nature Struct. Biol.* **2**, 45-55.

Bassi, G. S., Murchie, A. I. & Lilley, D. M. J. (1996). The ion-induced folding of the hammerhead ribozyme: core sequence changes that perturb folding into the active conformation. *RNA,* **2**, 756-768.

Basu, S., Rambo, R. P., Strauss-Soukap, J., Cate, J. H., Ferré-D′, Amaré A. R., Strobel, S. A. & Doudna, J. A. (1998). A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nature Struct. Biol.* **5**, 986-992.

Batey, R. T. & Williamson, J. R. (1998). Effects of polyvalent cations on the folding of an rRNA three-way junction and binding of ribosomal protein S15. *RNA,* **5**, 984-998.

Bernstein, J., Sella, O., Le, S.-Y. & Elroy-Stein, O. (1997). PDGF2/c-sis mRNA leader contains a differentiation-linked internal ribosomal entry site (D-IRES). *J. Biol. Chem.* **272**, 9356-9362.

Borman, A. M., Bailly, J.-L., Girard, M. & Kean, K. M. (1995). Picornavirus internal ribosome entry segments: comparison of translation efficiency and the requirements for optimal internal initiation of translation *in vitro*. *Nucl. Acids Res.* **23**, 3656-3663.

Brown, E. A., Zhang, H., Ping, L.-H. & Lemon, S. M. (1992). Secondary structure of the 5′ nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucl. Acids Res.* **20**, 5041-5045.

Bu, Z., Perlo, A., Johnson, G. E., Olack, G., Engelman, D. M. & Wyckoff, H. W. (1998). A small-angle X-ray scattering apparatus for studying biological macromolecules in solution. *J. Appl. Crystallog.* **31**, 533-543.

Bukh, J., Purcell, R. H. & Miller, R. H. (1992). Sequence analysis of the 5′ noncoding region of hepatitis C virus. *Proc. Natl Acad. Sci. USA,* **89**, 4942-4946.

Buratti, E., Tisminetzky, S., Zotti, M. & Baralle, F. E. (1998). Functional analysis of the interaction between HCV 5′ UTR and putative subunits of eukaryotic initiation factor eIF3. *Nucl. Acids Res.* **26**, 3179-3187.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science,* **273**, 1678-1685.

Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997). A magnesium ion core at the heart of a ribozyme domain. *Nature Struct. Biol.* **4**, 553-558.

Celander, D. W. & Cech, T. R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science,* **251**, 401-407.

Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science,* **244**, 359-362.

Choo, Q.-L., Weiner, A. J., Overby, L. R., Kuo, G., Houghton, M. & Bradley, D. W. (1990). Hepatitis C virus: the major causative agent of viral non-A, non-B hepatitis. *Brit. Med. Bull.* **46**, 423-441.

Clarke, B. (1997). Molecular virology of hepatitis C virus. *J. Gen. Virol.* **78**, 2397-2410.

Crothers, D. M. & Drak, J. (1992). Global features of DNA structure by comparative gel electrophoresis. *Methods Enzymol.* **212**, 46-71.

Duckett, D. R., Murchie, A. I. H. & Lilley, D. M. J. (1995). The global folding of four-way helical junctions in RNA, including that in U1 snRNA. *Cell,* **83**, 1027-1036.

Feller, J. A. & Lau, J. Y. N. (1998). Use of dicistronic vector for the quantitation of HCV IRES activity. In *Hepatitis C Protocols* (Lau, J. Y. N., ed.), pp. 315-324, Humana Press, Totowa, New Jersey.

Ferré-, D′. Amaré A. R. & Doudna, J. A. (1996). Use of *cis*- and *trans*-ribozymes to remove 5′ and 3′ hetero-

geneities from milligrams of *in vitro* transcribed RNA. *Nucl. Acids Res.* **24**, 977-978.

Fukushi, S., Katayama, K., Kurihara, C., Ishiyama, N., Hoshino, F. B., Ando, T. & Oya, A. (1994). Complete 5′ noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. *Biochem. Biophys. Res. Commun.* **199**, 425-432.

Fukushi, S., Kurihara, C., Ishiyama, N., Hoshino, F. B., Oya, A. & Katayama, K. (1997). The sequence element of the internal ribosome entry site and a 25-kilodalton cellular protein contribute to efficient internal initiation of translation of hepatitis C virus RNA. *J. Virol.* **71**, 1662-1666.

Gan, W. & Rhoads, R. E. (1996). Internal initiation of translation directed by the 5′-untranslated region of the mRNA for eIF4G, a factor involved in the picornavirus-induced switch from cap-dependent to internal initiation. *J. Biol. Chem.* **271**, 623-626.

Golden, B. L., Gooding, A. R., Podell, E. L. & Cech, T. R. (1998). A preorganized active site in the crystal structure of the tetrahymena ribozyme. *Science,* **282**, 259-264.

Hahm, B., Kim, Y. K., Kim, J. H., Kim, T. Y. & Jang, S. K. (1998). Heterogeneous nuclear ribonucleoprotein L interacts with the 3′ border of the internal ribosomal entry site of hepatitis C virus. *J. Virol.* **72**, 8782-8788.

Han, J. H., Shyamala, V., Richman, K. H., Brauer, M. J., Irvine, B., Urdea, M. S., Tekamp-Olson, P., Kuo, G., Choo, Q.-L. & Houghton, M. (1991). Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5′ untranslated region and poly(A) tails at the 3′ end. *Proc. Natl Acad. Sci. USA,* **88**, 1711-1715.

Honda, M., Brown, E. A. & Lemon, S. M. (1996a). Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation of hepatitis C virus RNA. *RNA,* **2**, 955-968.

Honda, M., Ping, L.-H., Rijnbrand, R. C. A., Amphlett, E., Clarke, B., Rowlands, D. & Lemon, S. M. (1996b). Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. *Virology,* **222**, 31-42.

Honda, M., Beard, M. R., Ping, L.-H. & Lemon, S. H. (1999). A phylogenetically conserved stem-loop structure at the 5′ border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J. Virol.* **73**, 1165-1174.

Jackson, R. J. & Kaminski, A. (1995). Internal initiation of translation in eukaryotes: The picornavirus paradigm and beyond. *RNA,* **1**, 985-1000.

Jang, S. K., Krausslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C. & Wimmer, E. (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.

Jang, S. K., Davies, M. V., Kaufman, R. J. & Wimmer, E. (1989). Initiation of protein synthesis by internal entry of ribosomes into the 5′ nontranslated region of encephalomyocarditis virus RNA *in vivo*. *J. Virol.* **63**, 1651-1660.

Kamoshita, N., Tsukiyama-Kohara, K., Kohara, M. & Nomoto, A. (1997). Genetic analysis of internal ribosomal entry site on hepatitis C virus RNA: implication for involvement of the highly ordered structure and cell type-specific transacting factors. *Virology,* **233**, 9-18.

Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W.-S. & Kuo, C., *et al*. (1989). An assay for circulating antibodies to major etiologic virus of human non-A, non-B hepatitis. *Science,* **244**, 362-364.

Latham, J. A. & Cech, T. R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science,* **245**, 276-282.

Lau, J. Y. N., Davis, G. L., Kniffen, J., Qian, K. P., Urdea, M. S., Chan, C. S., Mizokami, M., Neuwald, P. D. & Wilber, J. C. (1993). Significance of serum hepatitis C virus RNA levels in chronic hepatitis C. *Lancet,* **341**, 1501-1504.

Le, S.-Y., Sonenberg, N. & Maizel, J. V., Jr (1995). Unusual folded regions and ribosome landing pad within hepatitis C virus and pestivirus RNAs. *Gene,* **154**, 137-143.

Leontis, N. B. & Moore, P. B. (1984). A small angle X-ray scattering study of a fragment derived from *E. coli* 5S RNA. *Nucl. Acids Res.* **12**, 2193-2203.

Macejak, D. G. & Sarnow, P. (1991). Internal initiation of translation mediated by the 5′ leader of a cellular mRNA. *Nature,* **353**, 90-94.

Mollegaard, N. E., Murchie, A. I., Lilley, D. M. J. & Nielsen, P. E. (1994). Uranyl photoprobing of a four-way DNA junction: evidence for specific metal ion binding. *EMBO J.* **13**, 1508-1513.

Moore, P. B. (1982). Small-angle scattering techniques for the study of biological macromolecules and macromolecular aggregates. *Methods Exp. Phys.* **20**, 337-390.

Murphy, F. L. & Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry,* **32**, 5291-5300.

Oh, S.-K., Scott, M. P. & Sarnow, P. (1992). Homeotic gene Antennapedia mRNA contains 5′-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev.* **6**, 1643-1653.

Pan, T. (1995). Higher order folding and domain analysis of the ribozyme from *Bacillus subtilis* ribonuclease P. *Biochemistry,* **34**, 902-909.

Pan, T., Long, D. M. & Uhlenbeck, O. C. (1993). Divalent metal ions in RNA folding and catalysis. In *The RNA World* (Gesteland, R. F. & Atkins, J. F., eds), pp. 271-302, Cold Spring Harbor Laboratory Press, Plainview, NY.

Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. & Hellen, C. U. T. (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.

Pilz, I. & Kratky, O. (1970). On the conformation of phenylalanine specific transfer RNA. Studies on size and shape of the molecule by X-ray small angle scattering. *J. Biochem.* **15**, 401-409.

Poole, T. L., Wang, C., Popp, R. A., Potgieter, L. N. D., Siddiqui, A. & Collett, M. S. (1995). Pestivirus translation initiation occurs by internal ribosomal entry. *Virology,* **206**, 750-754.

Reynolds, J. E., Kaminski, A., Kettinen, H. J., Grace, K., Clarke, B. E., Carroll, A. R., Rowlands, D. J. & Jackson, R. J. (1995). Unique features of internal initiation of hepatitis C virus RNA translation. *EMBO J.* **14**, 6010-6020.

Reynolds, J. E., Kaminski, A., Carroll, A. R., Clarke, B. E., Rowlands, D. J. & Jackson, R. J. (1996). Internal initiation of translation of hepatitis C virus RNA: The ribosome entry site is at the authentic initiation codon. *RNA,* **2**, 867-878.

Rijnbrand, R., Bredenbeek, P., van der Straaten, T., Whetter, L., Inchauspe, G., Lemon, S. & Spaan, W. (1995). Almost the entire 5′ non-translated region of hepatitis C virus is required for cap-independent translation. *FEBS Letters,* **365**, 115-119.

Sizova, D. V., Kolupaeva, V. G., Pestova, T. V., Shatsky, I. N. & Hellen, C. U. T. (1998). Specific interaction of eukaryotic translation initiation factor 3 with the 5′ nontranslationed regions of hepatitis C virus and classical swine fever virus RNAs. *J. Virol.* **72**, 4775-4782.

Tang, S., Collier, A. J. & Elliott, R. M. (1999). Alterations to both primary and predicted secondary structure of stem-loop IIIc of the hepatitis C virus 1b 5′ untranslated region (5′ UTR) lead to mutants severely defective in translation which cannot be complemented *in trans* by the wild-type 5′ UTR sequence. *J. Virol.* **73**, 2359-2364.

Teerink, H., Voorma, H. O. & Thomas, A. A. M. (1995). The human insulin-like growth factor II leader 1 contains an internal ribosomal entry site. *Biochim. Biophys. Acta,* **1264**, 403-408.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* **66**, 1476-1483.

Vagner, S., Gensac, M.-C., Maret, A., Bayard, F., Amalric, F., Prats, H. & Prats, A.-C. (1995). Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. *Mol. Cel. Biol.* **15**, 35-44.

Varaklioti, A., Georgopoulou, U., Kakkanas, A., Psaridi, L., Serwe, M., Caselmann, W. H. & Mavromara, P. (1999). Mutational analysis of two unstructured domains of the 5′ untranslated region of HCV RNA. *Biochem. Biophys. Res. Commun.* **199**, 425-432.

Walter, F., Murchie, A. I. H., Duckett, D. R. & Lilley, D. M. J. (1998). Global structure of four-way RNA junctions studied using fluorescence resonance energy transfer. *RNA,* **4**, 719-728.

Wang, C., Sarnow, P. & Siddiqui, A. (1993). Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J. Virol.* **67**, 3338-3344.

Wang, C., Sarnow, P. & Siddiqui, A. (1994). A conserved helical element is essential for internal initiation of translation of hepatitis C virus RNA. *J. Virol.* **68**, 7301-7307.

Wang, C., Le, S.-Y., Ali, N. & Siddiqui, A. (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.

Weeks, K. M. & Cech, T. R. (1996). Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science,* **271**, 345-348.

Wilhelm, P. & Pilz, I. (1982). Small- and large-angle X-ray scattering studies of counter ion influence on tRNA conformation. *Z. Naturforsch.* **37c**, 1293-1296.

Yen, J.-H., Chang, S. C., Hu, C.-R., Chu, S.-C., Lin, S.-S., Hsieh, Y.-S. & Chang, M.-F. (1995). Cellular proteins specifically bind to the 5′-noncoding region of hepatitis C virus RNA. *Virology,* **208**, 723-732.

Edited by D. E. Draper

(Received 14 May 1999; received in revised form 3 August 1999; accepted 3 August 1999)