Published online September 7, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 16 ***4715–4724***
doi:10.1093/nar/gkh805

NMR studies of the structure and Mg^{2+} binding properties of a conserved RNA motif of EMCV picornavirus IRES element

Marie Phelan, Ryan J. Banks, Graeme Conn[1] and Vasudevan Ramesh*

Department of Chemistry and [1]Department of Biomolecular Sciences, University of Manchester Institute of Science and Technology, PO Box 88, Manchester M60 1QD, UK

Received June 15, 2004; Revised and Accepted August 17, 2004

ABSTRACT

The structure and Mg^{2+} binding properties of a conserved 75mer RNA motif of the internal ribosome entry site (IRES) element of encephalomyocarditis virus picornavirus have been investigated by ^{1}H-NMR and UV melting experiments. The assignment of the imino proton resonances with characteristic chemical shift dispersion for canonical and non-canonical base pairs confirmed the predicted secondary structure of the 75mer and its fragments. Addition of Mg^{2+} resulted in a dramatic increase in apparent melting temperature, with the 75mer RNA registering the biggest increase, from 63 to 80°C, thus providing evidence for enhanced stability arising from Mg^{2+} binding. Similarly, addition of Mg^{2+} induced selective changes to the chemical shifts of the imino protons of a GCGA tetraloop in the 75mer, that is essential for IRES activity, thereby highlighting a possible structural role for Mg^{2+} in the folding of the 75mer. Significantly, the same protons show retarded exchange to water solvent, even at elevated temperature, which suggest that Mg^{2+} induces a conformational rearrangement of the 75mer. Thus, we propose that Mg^{2+} serves two important roles: (i) enhancing thermodynamic stability of the 75mer RNA (and its submotifs) via non-specific interactions with the phosphate backbone and (ii) promoting the folding of the 75mer RNA by binding to the GCGA tetraloop.

INTRODUCTION

Translation initiation in eukaryotes is a complex process (1) and on the vast majority of eukaryotic cell mRNAs, the initiation of protein synthesis is achieved following the recognition of the 5′ terminal cap-structure (m7GpppN..) of mRNAs by the cap-binding complex eIF4F (2). This translation initiation factor is a heterotrimer consisting of eIF4E (which binds to the cap), eIF4A (which has RNA helicase activity) and eIF4G (a scaffold protein). The eIF4F complex acts as a bridge between the mRNA and the small ribosomal subunit. The eIF4G subunit has binding sites for a number of proteins, as well as eIF4E and eIF4A (two independent sites); these include eIF3 (associated with the 40S ribosomal subunit), poly(A) binding protein and Mnk-1 (an eIF4E-kinase).

Picornavirus RNAs function like mRNAs within mammalian cells, but are uncapped and the initiation of translation on picornavirus RNA occurs up to 1300 nt away from the 5′ terminus (3,4). This process is directed by an internal ribosome entry site (IRES) (5,6). The picornavirus IRES elements are generally ~450 nt in length. Two major classes of picornavirus IRES elements have been defined (7,8); the entero-/rhinovirus IRES elements constitute one class while the cardio-/aphthovirus IRESs [including the encephalomyocarditis virus (EMCV) (9) and foot-and-mouth disease virus (FMDV) (10) elements] make up a second class of elements. The cardio-/aphthovirus IRES elements function efficiently in the rabbit reticulocyte lysate cell-free translation system. In contrast, the entero-/rhinovirus IRES elements function rather poorly in this system, unless they are supplemented with HeLa cell proteins. Both classes of elements function well within certain mammalian cell lines. All picornavirus IRES elements contain extensive secondary structure (11), but these predicted structures are completely different between the two classes of IRES element, and there is no significant sequence similarity between them (12). This observation raises the important question of just exactly how each of the picornaviridae viruses direct internal ribosomal entry when their IRES elements show no sequence similarity (5,6). The answer to this question most probably is that the specific secondary and/or tertiary structures of the IRES element (which are not sequence dependent) are more important than the primary structure for initiating translation.

The EMCV IRES is perhaps the best characterized picornavirus IRES (13–16). It consists of multiple domains (labelled H to L) (13,15). Certain regions within the EMCV IRES are highly conserved and have stable predicted secondary structures. These include a complex 'hammerhead' type structure at the apex of the I-domain (13,15), and also the J-K domains which bind the initiation factors eIF4G and eIF4A (16). The hammerhead region includes a four-way junction and three short stem–loop elements. At the apex of one of the stem–loops is a GNRA tetraloop motif (13,15), which is an essential feature of the IRES (14). Modification of the GCGA sequence to GCGC completely inactivated the IRES (17);

*To whom correspondence should be addressed. Tel: +44 161 200 4539; Fax: +44 161 200 4559; Email: vasudevan.ramesh@umist.ac.uk

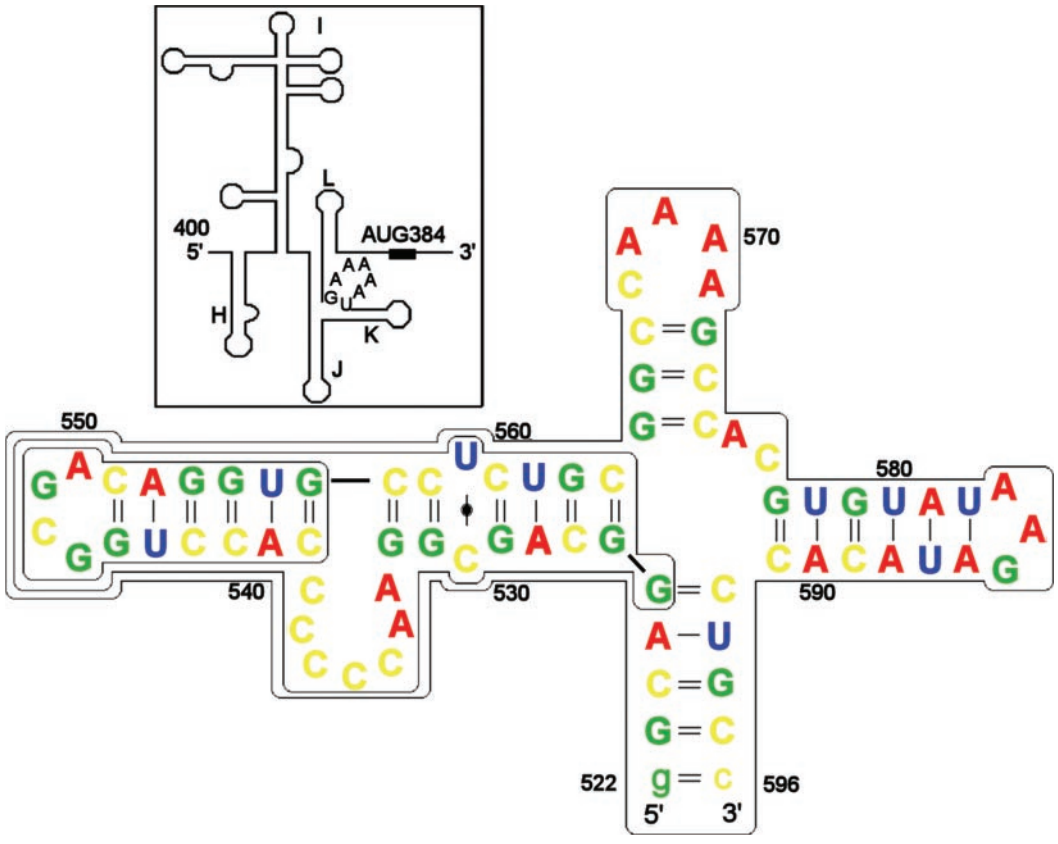


Figure 1. Schematic diagram showing the colour-coded nucleotide sequence and predicted 'hammerhead' type secondary structure of the single-stranded 75mer RNA motif, spanning residues 5′-G522 to C596-3′, of the IRES element of EMCV RNA. The secondary structure depicted is based on the nomenclature proposed previously (46). The intact 75mer motif is endowed with several smaller secondary structural elements such as Watson–Crick base pair helical stems, hairpin loops, internal loop (mismatch base pair between C531 and U559), internal cytosine-rich, heptanucleotide bulge (A534–C540) and the four-way junction. The four residues spanning G547–A550 are predicted to form the well-known GNRA tetraloop motif (33,34) and have been previously demonstrated to be important for activity (14). Similarly, mutation of C→A544 causes severe reduction in activity (13). The scheme also shows a normal base pair g522–c596 (in small letters) added to the beginning of the sequence to increase the transcription promoter efficiency. The single-stranded 16mer (C541 to G556) and 38mer (G526 to C563) RNA submotifs, derived from the intact 75mer RNA motif, also used for NMR and other investigation are shown in separate inner boxes. The inset shows an outline of the secondary structure of the various domains within the intact IRES of EMCV, including the above cited 75mer motif as part of the domain I.

indeed selection from a pool of up to 256 variants of this tetraloop indicated that only sequences that fitted a RNRA consensus displayed high-IRES activity (14). However, the molecular basis for this critical role in IRES function is not known and no proteins are known to interact with this hammerhead region (15,16). The effect of mutation on IRES function may then be attributed, at least in part, to disrupting the three-dimensional (3D) structural integrity and considerable stability conferred by this motif, including its dependence on Mg^{2+} ions, and hence warrants a structural investigation.

The determination of the structure of the whole EMCV IRES (434 nt) is beyond the reach of NMR methods at present; however, studies on the hepatitis C virus IRES have shown the utility of determining the structures of specific domains from within the whole structure (18,19). Thus, we have selected a highly conserved 75mer motif containing the GNRA tetraloop (Figure 1) that is critical for the activity within the EMCV IRES for NMR structural investigation. Furthermore, we have probed the role of Mg^{2+} in the folding of the RNA by monitoring changes to the imino proton resonances of the RNA.

MATERIALS AND METHODS

Chemical and enzymatic synthesis of RNA

The 16mer RNA (Figure 1) was synthesized chemically (Dharmacon) and used without further purification. DNAs coding the 38mer and 75mer fragments (Figure 1) were cloned into a plasmid for *in vitro* run-off transcription of each RNA (with a 3′-self cleaving ribozyme) and were used to prepare RNAs as described previously (20). The purity of each RNA was confirmed by denaturing PAGE which showed a single band corresponding to RNA of the expected length (data not shown). Transcription and purification of the 38mer and 75mer RNAs were performed on a scale such as to enable the measurement of the 1D proton NMR spectra of each RNA and to independently verify their purity. The RNA samples were annealed by heating to 80°C for 3 min and slowly cooling them to 4°C.

Empirical free energy calculation

The empirical free energy calculation (21) of the 16mer, 38mer and 75mer RNA motifs of EMCV IRES element

(Figure 1) was carried out using the Mfold program (22) available on the web server (http://www.bioinfo.rpi.edu/applications/mfold).

Unconstrained molecular modelling and molecular dynamics simulation

The starting structures for the 16mer and 38mer stem–loop models were constructed from the A-form double-helical stems using the Nucleic Acid Database within HyperChem with default atom types. The modelled structures were geometry optimized using the Polak–Ribiere conjugate gradient algorithm. Molecular dynamics simulation of the aqueous construct was carried out for up to 100 ps characterizing the converged 10 ps structures observed throughout the simulation by geometry and energetic analysis (23–25).

UV absorbance melting (T_m) experiments

UV melting curves were collected on a Cary 400 UV/Vis spectrophotometer running in dual beam mode, and with a heating rate of 1.0°C/min (alternative heating rates, e.g. 0.5°C/min, or cooling rather than heating, gave essentially identical curves). Samples contained ~25 μg of RNA in buffer containing 10 mM MOPS, pH 6.8 and 50 mM NaCl. Melting curves were also collected in the same buffer but with additional 2 mM $MgCl_2$. Up to five curves were collected at one time (measuring absorbance at both 260 and 280 nm), with the sixth position used for a temperature probe immersed in the buffer. The melting curves were converted into plots of the first derivative ('melting profiles') using the program 'ODderiv', and deconvolutions of these melting profiles into sequential unfolding transitions were performed using 'Global Melt Fit' (http://www.jhu.edu/~chem/draper) (26,27).

NMR spectroscopy

One- and two-dimensional proton NMR experiments (TOCSY and NOESY) (28,29) were carried out using a Varian unity 500 spectrometer, equipped with a triple resonance inverse geometry gradient probe and electronic variable temperature unit. To afford higher sensitivity, selected 2D experiments were also carried out using Bruker 600 MHz, equipped with a modern cryoprobe, and Varian 800 MHz NMR spectrometers. The spectra were acquired using Vnmr 6.1C software hosted by a Sun UNIX workstation (Varian) and XWINNMR software (Bruker) interfaced to the spectrometers, and the manufacturer supplied pulse programs with modifications where necessary. The spectral width was set to 22 p.p.m. for experiments in ^1H_2O and the transmitter/carrier was positioned on the water signal to minimize any artefacts. The large resonance due to the water protons was suppressed by the WATERGATE pulse sequence (30). Typically, 32–64 scans were averaged for each FID before apodization and then performing the Fourier transformation. ^1H chemical shifts are reported with respect to dilute concentration of internal dioxane (0.01%, v/v) used as an internal reference standard.

Most of the 1D and 2D NMR data were initially processed using Varian Vnmr 6.1C software or Bruker XWINNMR software. For display and analysis purposes, the data were later reprocessed using MestRe-C 3.4 software (http://www.mestrec.com) running on Microsoft Windows XP. Usually, each time domain (t_1, t_2) data was zero-filled twice before apodization by a shifted sine bell function and then Fourier transformation to yield frequency domain (δ_1, δ_2) spectra.

The NMR samples, as a 0.6 ml solution each, were prepared by dissolving the RNA (0.28–1.3 mM) in 90% ^1H_2O + 10% ^2H_2O solvent mixture containing 20 mM PO_4^{3-}, pH 6.0. Titrations with Mg^{2+} were carried out by adding microlitre volumes of a stock solution of $MgCl_2$ to RNA samples.

RESULTS

Empirical free energy of EMCV IRES RNA motifs

In order to assess and compare the overall thermodynamic stability of the 16mer, 38mer and 75mer RNAs, the free energy of folding (21) (ΔG, kcal mol^{-1}) of each RNA was calculated using the empirical free energy prediction protocol provided within the Mfold program (22). The predicted free energy of folding for wild-type and mutant tetraloop analogues of all three RNAs are summarized in Table 1.

Inspection of Table 1 shows that the consensus 16mer RNA stem–loop consisting of the GCGA tetra loop (Figure 1) has a low free energy (−9.7 kcal mol^{-1}, column 2) and hence is predicted to be endowed with good stability. The corresponding data for the wild-type 38mer RNA motif shows that, despite an internal 7 nt asymmetrical bulge and a single C–U mismatch (Figure 1), the predicted net free energy (−14.0 kcal mol^{-1}, column 3) is still lower than the free energy of the 16mer within it. The predicted free energy of folding for the wild-type 75mer motif (Figure 1), however, shows the lowest free energy (−29.1 kcal mol^{-1}, column 4), as compared to those of 16mer and 38mer submotifs, which suggests a large contribution of the secondary structure of the four-way junction (Figure 1) to the enhanced stability of this intact motif.

Table 1 also shows that nucleotide variation of the consensus GCGA tetraloop at the fourth position (A) to any other nucleotide (G, U or C) is accompanied by a net increase in the free energy of the corresponding mutant analogue of all three RNAs, thus lowering their overall stability. This highlights a critical structural role for the fourth nucleotide of the GCGA tetraloop within the tertiary structure of the RNAs.

To allow a comparison of the thermodynamic stability of the 75mer EMCV IRES RNA with other RNAs of similar size and disposed with a range of secondary structural elements, the

Table 1. The empirical free energy (ΔG, kcal mol^{-1}) due to folding of single-stranded 16mer, 38mer and 75mer RNA motifs of EMCV IRES calculated by the Mfold program (http://www.bioinfo.rpi.edu/applications/mfold)

RNA tetraloop	16mer ΔG kcal mol^{-1}	38mer ΔG kcal mol^{-1}	75mer ΔG kcal mol^{-1}
GCG**A**	**−9.7**	**−14.0**	**−29.1**
GCG**G**	−6.8	−11.9	−26.2
GCG**U**	−6.9	−11.2	−26.3
GCG**C**	−6.9	−11.2	−26.3
		$tRNA^{Phe}$ 76mer	−20.5

The calculated free energy of mutant RNAs, carrying single nucleotide mutation at the fourth nucleotide position (in boldface) of GCG**A** tetraloop of each RNA is shown. Also tabulated is the ΔG of the single-stranded 76mer $tRNA^{Phe}$, whose 3D has been determined before using X-ray crystallography (31) and hence is chosen as a reference for the empirical calculation (22).

empirical free energy of $tRNA^{phe}$ (76mer) (31) was calculated, and the predicted value (−20.5 kcal mol^{-1}) is shown separately at the bottom of Table 1. The data show that the single-stranded 75mer EMCV RNA is predicted to be considerably more stable than the single-stranded 76mer cloverleaf motif of $tRNA^{phe}$ whose 3D X-ray structure has already been determined (31). This suggests that the proposed structure (Figure 1) is extremely likely to be that undertaken by this sequence of nucleotides in the EMCV picornavirus.

Unconstrained molecular modelling and molecular dynamics simulation of EMCV IRES RNA motifs

We have carried out molecular dynamics simulation of 16mer and 38mer RNAs and analysis of the output parameters in terms of potential energy of base stacking and base pairing, and puckering of sugar along the backbone for both RNA motifs (32). The overlay of the different molecular dynamics simulations (up to 100 ps) showed good agreement with greater deviation at the bulge, the tetraloop and the free ends of the stem region. This implied a dynamic process of constant reorganization in the bulge and tetraloop regions. The sugar pucker alteration in the second nucleotide is like that observed before by Heus and Pardi (33,34). The stacking also behaved in the expected manner with the CGA bases of the GCGA tetraloop stacked on top of the 3′-strand while G alone was stacked on the 5′-strand. This change in direction of the hairpin between nucleotides 1 and 2 is characteristic for GNRA tetraloops. The first and fourth bases in the tetraloop formed an unusual G–A base pair (35).

UV absorbance melting studies of EMCV IRES RNA motifs and the effect of Mg^{2+}

Melting profiles (see Materials and Methods) for each RNA in the absence or in the presence of 2 mM Mg^{2+} are shown in Figure 2. In the absence of Mg^{2+}, the 16mer hairpin unfolds in a single transition with a T_m of 72°C. As expected, the 38mer and 75mer RNA profiles are more complex, with apparent transitions at T_m's of 46 and 66°C, and 49°C (a broad 'shoulder') and 63°C, respectively. The stability of all three RNAs is dramatically enhanced by the addition of Mg^{2+} to the melting buffer (Table 2). The Mfold program was used to compute T_m and other thermodynamic parameters for each RNA; these values (data not shown) were consistently higher than the experimental apparent values from the melting profiles. However, we note that the order of stability, i.e. the decrease in the highest apparent T_m with increasing RNA size/complexity, was reproduced. The unfolding pathways of RNAs with a number of secondary (and possibly tertiary) structural elements may be complex, and may consist of a number of sequential transitions (i.e. where there is an obligatory order to the unfolding of these elements). The program 'Global Melt Fit' was used to fit sequential two-state transitions to each melting profile (simultaneously fitting 260 and 280 nm profiles), with one, two and three transitions used for the 16mer, 38mer and 75mer, respectively. Each profile is fit with values for ΔH, absorbance and T_m (data not shown). The fitted values for both T_m and ΔH were both substantially lower than that calculated by Mfold (see Discussion). Despite this, however, the melting profiles indicate that each RNA is stable and well folded under conditions suitable for NMR structural analysis.

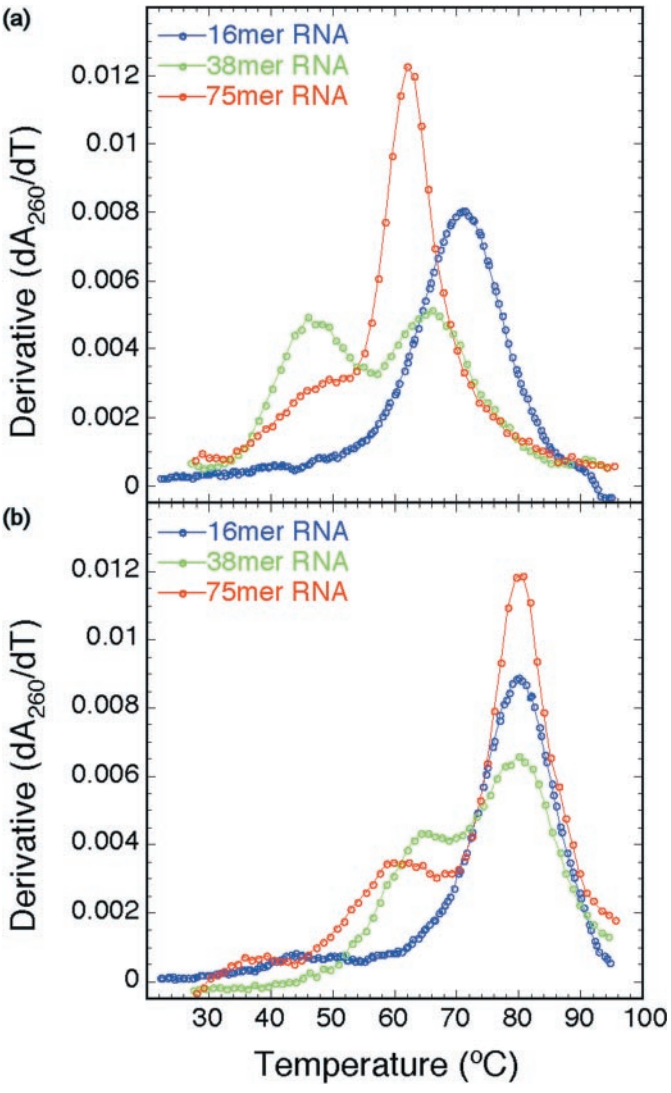


Figure 2. First derivative plots of the UV absorbance melting curves, d*A*/d*T* versus *T* (°C), of 16mer, 38mer and 75mer RNAs (10 mM MOPS and 50 mM NaCl, pH 6.8), without Mg^{2+} (**a**) and with 2mM Mg^{2+} (**b**). The individual traces are colour coded to distinguish the melting profile of each RNA.

Table 2. The UV absorbance melting temperature T_m (°C) of 16mer, 38mer and 75mer RNA motifs of EMCV IRES (10 mM MOPS and 50 mM NaCl, pH 6.8), measured without and with 2 mM Mg^{2+}

RNA	T_m (°C) No Mg^{2+}		With Mg^{2+}		ΔT_m	
16mer	—	72	—	80	—	8
38mer	46	66	64	80	18	14
75mer	49	63	60	80	11	17

The T_m values reported for the first and second transitions of each RNA are based on the melting profiles shown in Figure 2.

Sequence-specific NMR resonance assignment of EMCV IRES RNA motifs

The 1D proton NMR spectra of 16mer, 38mer and 75mer RNA samples were measured in ^1H_2O which showed resonances, with good intensity characteristic of well-folded A-form RNA,

such as the low-field imino protons arising from stacked base pairs of helical stems. There were very few resonances high field of 3.5 p.p.m. (data not shown). Owing to its larger molecular size, the spectrum of the 75mer (~24 kDa) showed increased line widths and greater resonance overlap, in contrast to the resolved and narrower resonances observed in the spectra of the 16mer (~5 kDa) and 38mer (~12 kDa) RNAs. Comparison of the spectra of the three RNAs showed similar chemical shifts of some of the resonances, hinting at the possibility of conserved tertiary structural interactions among them. The NOESY spectra of all the three RNA samples were measured, without and with three equivalents of Mg^{2+}, and the sequence-specific assignment (36) of the observed imino proton resonances was carried out separately for each RNA and its Mg^{2+} complex. Owing to a slightly better quality of the Mg^{2+} bound RNA spectra in general, the results of sequence-specific assignment are described below for each RNA when it is in complex with Mg^{2+}, but they are equally applicable to free RNA.

Figure 3 shows the NOESY spectrum of 16mer RNA–Mg^{2+} complex, where sequential connectivities have been traced for imino protons in close proximity of nucleotides adjacent to each other in the sequence of the stem–loop (bottom panel). A striking feature of the spectrum is the identification of the high-field shifted and well-resolved unpaired imino protons of G547/G549 nucleotides of the GCGA tetraloop, which provided a reliable starting point for tracing the sequential connectivity of the lower stem region (Figure 1). All of the cross peaks have been unambiguously identified and their connectivities were assigned to sequence-specific nucleotides of the stem–loop (Figure 1). Despite the fraying of the G556–C541 terminal base pair [a common occurrence (36)], resulting in rapid exchange of G556 imino proton with protons of solvent ^1H_2O, a distinct NOE cross peak due to U555–G556 correlation can be observed. This confirms both the general stability of the RNA and the high sensitivity afforded by the cryoprobe used to detect weak cross peaks.

The successful assignment of the G and U imino protons of the 16mer RNA enabled the identification and assignment of the exocyclic amino and non-exchangeable H2 protons of canonical base pairing C and A, respectively, via NOE correlations (Figure 3, upper panel). Thus, a number of cross peaks observed due to G imino and C amino protons in close proximity, and similarly between U imino and A (H2) protons, have been labelled with their specific assignments (upper panel). The connectivity traced across the bottom and upper panels due to the coupling between base paired G546 imino and C551 amino protons, and similarly between unpaired G547/G549 imino and C551 amino protons is especially significant from the point of both confirmation of assignment and structure of the GCGA tetraloop.

The NOESY spectrum of 38mer RNA–Mg^{2+} complex was analysed similarly to carry out its sequence-specific resonance assignment (Supplementary Material 1). Many of the cross peaks due to protons in close proximity as identified in the NOESY spectrum of the 16mer (Figure 3) are conserved in the spectrum of the 38mer, including the unpaired imino protons G547/G549 of the tetraloop, with minor differences in chemical shifts, which suggest that the conformation of the 16mer stem–loop motif is not changed significantly as an integral part of the larger 38mer motif (Figure 1). Owing to low sensitivity

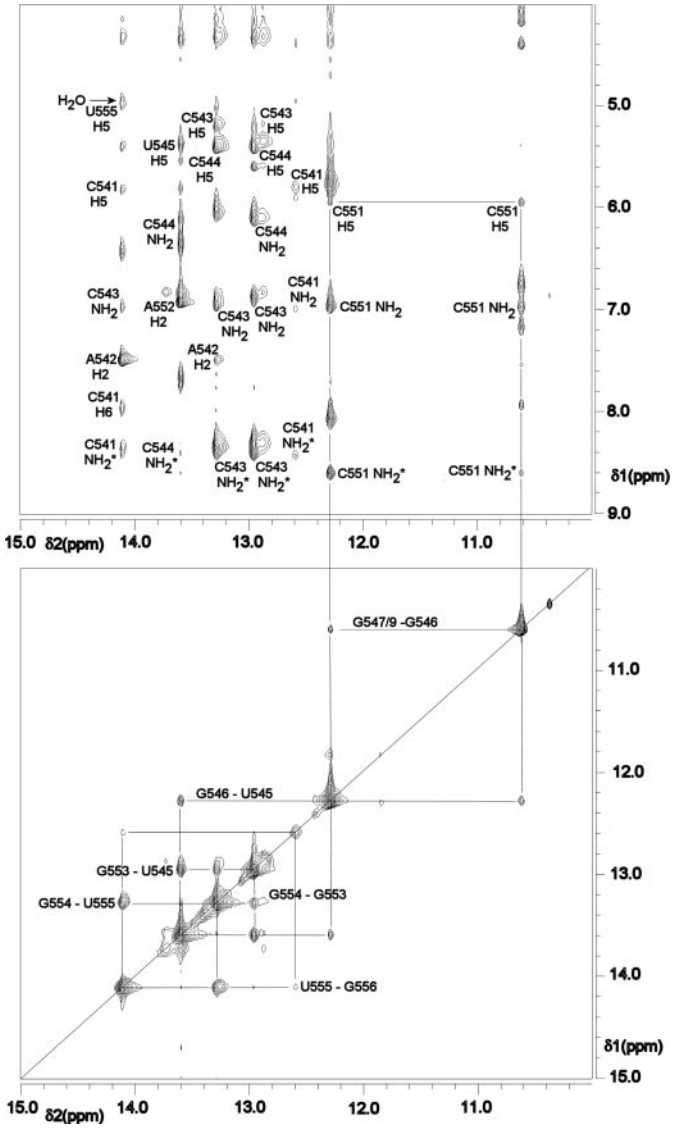
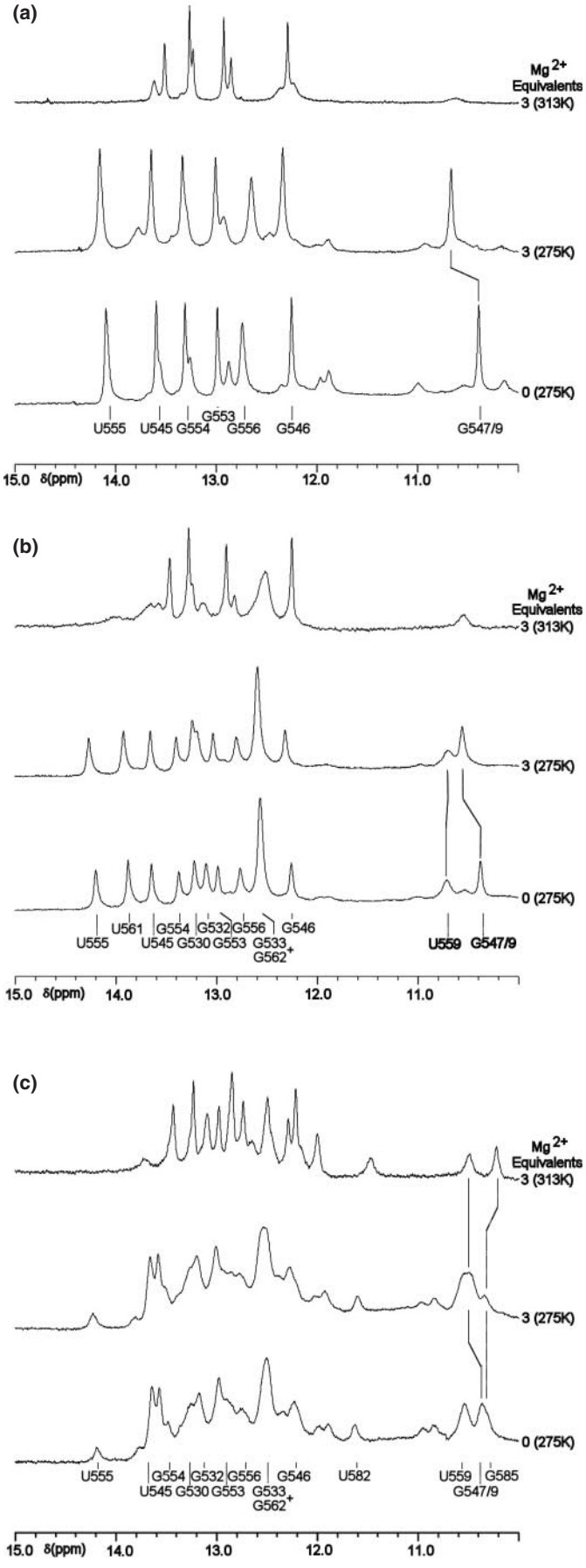


Figure 3. A 600 MHz ^1H-^1H correlated, phase-sensitive NOESY spectrum of 16mer RNA (1.27 mM) dissolved in 90% ^1H_2O + 10% ^2H_2O, 20 mM PO_4^{3-}, pH 6.0 and containing three molar equivalents of Mg^{2+} at 275 K. Mixing time used for cross relaxation between protons in close proximity was 250 ms. (Bottom panel) Sequential connectivities tracing cross peaks due to cross relaxation between adjacent imino protons (NH_i–NH_{i+1}) of each base pair and are labelled sequence specifically. (Upper panel) Identification and assignment of selected exchangeable amino protons (hydrogen bonded NH_2^* and non-hydrogen bonded NH_2, and non-exchangeable aromatic H5 and H2 protons of various cytosine, uracil and adenine residues based on observed NOE cross peaks arising from cross relaxation by neighbouring imino protons. Thus, the NOE connectivities shown by the imino protons of G547/G549 and G546 to the exocyclic amino (NH_2) and aromatic (H5) protons of C551 have been traced with vertical lines across both bottom and upper panels.

of the sample (0.28 mM) and increased resonance line widths, the 2D NOESY spectrum of the 75mer RNA–Mg^{2+} complex did not exhibit enough cross peaks for tracing NOE connectivities to render full sequential assignment (Supplementary Material 2). However, the corresponding 1D proton spectra of free and Mg^{2+} bound RNA (described in the next section, Figure 4) showed many resolved imino proton resonances and they could be identified from the prior assignment of 16mer and 38mer RNAs noted above.



The chemical shifts of the assigned imino protons of 16mer, 38mer and 75mer RNAs, without and with Mg^{2+}, are listed in Table 3.

NMR spectra of the effect of Mg^{2+} on EMCV RNA motifs

The effect of Mg^{2+} on 16mer, 38mer and 75mer RNA motifs was monitored by measuring the 1D proton NMR spectra of each RNA before and after the addition of three equivalents of Mg^{2+}; the results are shown in Figure 4a–c. Mg^{2+}-induced selective chemical shift changes to the assigned imino proton resonances of the RNAs are indicated by vertical lines (bottom and middle traces, Figure 4a–c). Inspection of Table 3 reveals that the unpaired G547/G549 imino protons of the GCGA tetraloop of 16mer, 38mer and 75mer RNAs show significant change in the chemical shift ($\Delta\delta$) of 0.26, 0.15 and 0.12 p.p.m., respectively. The remainder of the chemical shift changes noted, including the U559 imino proton of the C–U mismatch, are minor and relatively small in magnitude (<0.1 p.p.m.). These results indicate that Mg^{2+} has a greater and selective affinity for the GCGA tetraloop and may be involved in non-specific electrostatic interactions with the negatively charged phosphate backbone.

As well as the chemical shift changes noted above (Figure 4), the addition of Mg^{2+} to the 75mer RNA was also accompanied by an increase in line widths to the assigned imino proton resonances of both base-paired stem and unpaired loop nucleotides of the RNA (bottom and middle traces, Figure 4c). In contrast, there were no such changes to the line widths of 16mer and 38mer RNAs after the addition of Mg^{2+} (bottom and middle traces, Figure 4a and b). This indicates that Mg^{2+} induces a conformational rearrangement of the 75mer, in addition to the electrostatic interaction with the GCGA tetraloop within the ion complex.

The effect of Mg^{2+} on the stability of RNAs was probed by measuring the spectra of each complex at an elevated temperature of 313 K (upper trace, Figure 4a–c). At this higher temperature, the intensities of resonances due to exchange labile imino protons of 16mer and 38mer were sharply attenuated, with the unpaired loop nucleotides G547/G549 (10.52 p.p.m.) and base paired U555, U545 and U561

Figure 4. (**a**) Stack plot of the 500 MHz ^{1}H-NMR spectra (10.0–15.0 p.p.m.) of 16mer RNA (1.27 mM) dissolved in 90% ^1H_2O + 10% ^2H_2O and 20 mM PO_4^{3-}, pH 6.0 at 275 K measured before (bottom trace) and after the addition of three molar equivalents of Mg^{2+} at 275 K (middle trace) and 313 K (upper trace). The spectra show Mg^{2+}-induced changes in chemical shifts of the imino proton resonances of G547/G549 indicated by a slanted vertical line. (**b**) Stack plot of the 500 MHz ^{1}H-NMR spectra (10.0–15.0 p.p.m.) of 38mer RNA (0.78 mM) dissolved in 90% ^1H_2O + 10% ^2H_2O and 20 mM PO_4^{3-}, pH 6.0 at 275 K measured before (bottom trace) and after the addition of three molar equivalents of Mg^{2+} at 275 K (middle trace) and 313 K (upper trace). The spectra show Mg^{2+}-induced changes in chemical shifts of the imino proton resonances of G547/G549 and U559 indicated by slanted vertical lines. (**c**) Stack plot of the 500 MHz ^{1}H-NMR spectra (10.0–15.0 p.p.m.) of 75mer RNA (0.28 mM) dissolved in 90% ^1H_2O + 10% ^2H_2O and 20 mM PO_4^{3-}, pH 6.0 at 275 K measured before (bottom trace) and after the addition of three molar equivalents of Mg^{2+} at 275 K (middle trace) and 313 K (upper trace). The spectra show Mg^{2+}-induced changes in chemical shifts of the imino proton resonances of G547/G549 and U559 indicated by slanted vertical lines. Further, these resonances exhibit retarded exchange with protons in bulk water solvent at the elevated temperature of 313 K, unlike those that did not retard in (a) and (b). The sequence-specific assignment of the various imino protons of 16mer, 38mer and 75mer RNAs is shown underneath the bottom trace of each stack plot (a), (b) and (c) above.

Table 3. The ^{1}H-NMR chemical shifts (p.p.m.) of the assigned imino protons of 16mer, 38mer and 75mer RNAs at 275 K measured without and with three molar equivalents of Mg^{2+}

Residue	16mer RNA Chemical shift δ (p.p.m.) No Mg^{2+}	With Mg^{2+}	Δδ (p.p.m.)	38mer RNA Chemical shift δ (p.p.m) No Mg^{2+}	With Mg^{2+}	Δδ (p.p.m.)	75mer RNA Chemical shift δ (p.p.m) No Mg^{2+}	With Mg^{2+}	Δδ (p.p.m.)
$G^{547/9}$	10.42	10.68	0.26	10.37	10.52	0.15	10.38	10.50	0.12
G^{546}	12.28	12.35	0.07	12.25	12.28	0.03	12.24	12.27	0.03
U^{545}	13.62	13.66	0.04	13.64	13.69	0.05	13.64	13.65	0.01
G^{553}	13.02	13.02	0.00	12.98	13.01	0.03	12.98	13.00	0.02
G^{554}	13.34	13.35	0.01	13.37	13.37	0.00	—	—	—
U^{555}	14.12	14.17	0.05	14.19	14.25	0.06	14.19	14.22	0.03
G^{556}	12.77	12.66	−0.11	12.76	12.78	0.02	12.75	12.78	0.03
G^{533} (+G^{562})				12.56	12.56	0.00	12.51	12.53	0.02
G^{532}				13.10	13.19	0.09	13.18	13.19	0.01
U^{559}				10.71	10.67	−0.04	10.83	10.83	0.00
G^{530}				13.22	13.23	0.01	—	—	—
U^{561}				13.87	13.91	0.04	—	13.80	—
G^{562} (+G^{533})				12.56	12.56	0.00	12.51	12.53	0.02
G^{527}				—	—	—	—	—	—
G^{526}				—	—	—	—	—	—

The values reported are based on the 600 MHz NMR data (Figure 3) and 500 MHz, 800 MHz NMR data (Supplementary Material).

disappearing almost completely (upper traces, Figure 4a and b). This shows that Mg^{2+} does not retard the rate of exchange of these labile imino protons of 16mer and 38mer RNAs with those of bulk water solvent. Upon lowering the temperature back to 275 K, the same labile imino proton resonances were recovered and restored to their original intensities for both the RNAs (data not shown). The effect of Mg^{2+} on the stability of 75mer RNA at the elevated temperature of 313 K was, however, different in many respects. The imino proton resonance line widths showed a uniform increase, due to the shorter correlation time of the molecule accompanied by a reduction in the intensity of base paired U555 and U545 protons (middle and upper traces, Figure 4c). The unpaired imino protons of the assigned nucleotides G547/G549 within the GCGA tetraloop showed a dramatic retardation of solvent exchange even at 313 K (upper trace, Figure 4c). In contrast, no retardation was noted above for the same protons of the tetraloop, when they are part of the 16mer and 38mer RNAs (upper traces, Figure 4a and b). This suggests that Mg^{2+} may play a distinct role in the structure of the 75mer in addition to stabilizing electrostatic interactions with the phosphate backbone of the GCGA tetraloop (37–40).

The results of the effect of Mg^{2+} binding and increase in temperature on the imino proton resonances are shown schematically in the form of a histogram, taking the stem–loop secondary structure of the 16mer RNA as an example (Figure 5). The histogram clearly demonstrates that Mg^{2+} induces maximum effect (largest Δδ) on the imino protons of the tetraloop and, in comparison, much smaller changes to the stem protons (upper panel). The histogram also shows that Mg^{2+} binding does not significantly affect the small chemical shift variation with increase in temperature observed for the imino protons of the tetraloop of apo RNA (bottom panel). In contrast, the stem protons of apo RNA that show large chemical shift variation with increase in temperature (Δδ ∼0.2 p.p.m.) are considerably retarded after Mg^{2+} binding (bottom panel). This suggests that Mg^{2+} binding renders the RNA to adopt as homogeneous a population of conformation as possible and limiting the chances for dynamic fluctuation.

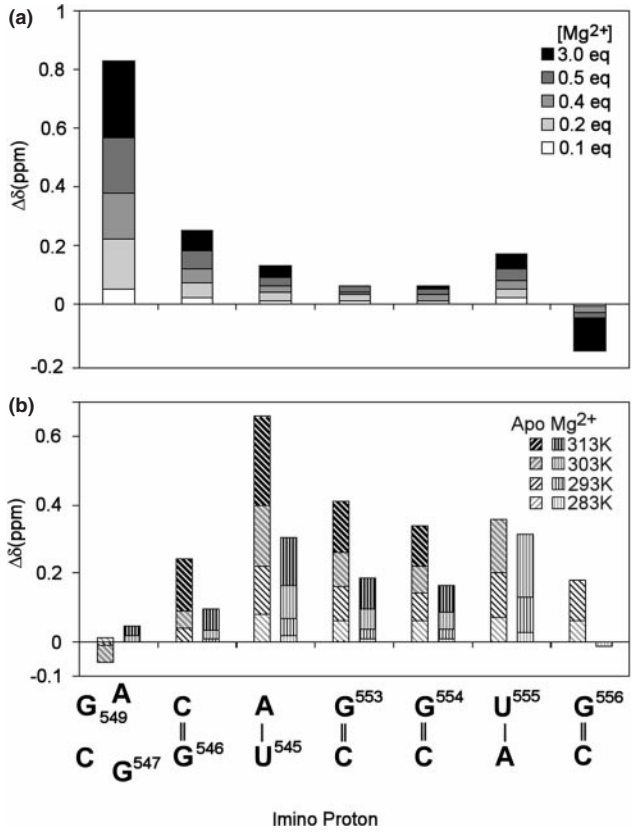


Figure 5. A histogram depicting (**a**) the variation of change in imino proton chemical shift (Δδ) along the sequence of the stem–loop secondary structure of 16mer RNA (5′-C541-G556-3′) after incremental addition of molar equivalents of Mg^{2+} (upper panel) and (**b**) the variation of change in imino proton chemical shift (Δδ) along the sequence of the stem–loop secondary structure of the same RNA, in the apo and Mg^{2+} bound states, as a function of increasing temperature, 283–313 K (bottom panel).

Furthermore, we have also characterized the changes to the non-exchangeable, aromatic proton resonances of 75mer RNA upon variation in temperature at regular intervals between 2 and 65°C (Supplementary Material 3). Although these

resonances (5.5–7.0 p.p.m.) have not been specifically assigned yet, distinct and sharp changes to these can be observed from 50°C upwards with disappearance of selected few resonances and narrowing of some others. As a number of these resonances can be reliably identified with H5 protons of cytosine residues based on chemical shifts, we speculate that the observed changes may correspond to the unfolding of C-rich, heptanucleotide bulge (the second melting transition at lower T_m, Table 2).

DISCUSSION

The calculated empirical free energies of the 16mer, 38mer and 75mer RNAs of EMCV IRES indicate that the RNA motifs are endowed with a globally well-folded structures and thermodynamic stability. The data also compare favourably with the calculated free energy of the structurally well-characterized 76mer $tRNA^{Phe}$ clover leaf motif, used as a reference (31). The critical role played by the conserved GCGA tetraloop of each RNA was probed by a single substitution of the fourth nucleotide A550 successively with G, U and C, and recalculating the empirical free energy for each of the resulting mutant RNAs. The results showed a uniform increase in the free energy of the RNA as a consequence of the mutation, thereby highlighting a possible structural role for the fourth nucleotide A of the GCGA tetraloop, in agreement with the previous biochemical studies on the same tetraloop (17). The results also correlate with the recently published data on the IRES from foot-and-mouth disease, another picornavirus, which shares a similar secondary structural motif also consisting of a GNRA tetraloop, where mutation of the fourth nucleotide of the tetraloop loop (A to G) caused strong RNA re-organization affecting several distal nucleotides in the domain (41).

The low-energy 3D structures, generated by unconstrained molecular dynamics simulation of the single-stranded 16mer and 38mer motifs, are consistent across the different secondary structural elements constituting the respective motifs, and thus validate their secondary structure prediction. This simulation also generated the cytosine-rich heptanucleotide bulge structure for the nucleotides spanning A534–C540 (Figure 1), with noticeably high energy as expected. Overall, the results of simulation were mutually consistent for the 16mer and 38mer motifs, and aided the experimental NMR investigation. It may be noted that the molecular dynamics simulations were run only in the absence of Mg^{2+}, and hence the role of the latter is not known (42).

The experimental UV melting data show that, in the absence of Mg^{2+}, the three RNAs have different maximum stabilities. The 16mer has a slightly higher T_m (72°C) in comparison to the 38mer and 75mer RNAs, which have lower values (apparent T_m's of 66 and 63°C, respectively) despite the fact they incorporate this same 16mer motif. The Mfold program predicted substantially higher values than that were experimentally observed. Clearly, the presence of additional structural elements such as the cytosine-rich, heptanucleotide bulge in the 38mer and 75mer RNAs have an overall destabilizing effect on the RNA structure. After the addition of Mg^{2+}, these differences were removed and all three RNAs melt with an experimental apparent T_m of ~80°C. The 75mer RNA thus showed the largest apparent T_m increase of 17°C, demonstrating the increased dependence of the 75mer on Mg^{2+}, or a greater effect of Mg^{2+} on the 75mer to achieve a globally well-folded tertiary structure. Such a large increase in T_m hints at the possibility of a specific role for Mg^{2+}, in addition to non-specific interaction with the negatively charged backbone (38,43,44).

The melting profiles of RNA molecules with complex structures, such as the 38mer and 75mer, are often complex (45). Such RNAs are assumed to unfold in a number of transitions, though these may not necessarily correspond to specific structural features of each RNA. The melting of RNAs with tertiary structure is expected to be hierarchical, with tertiary structure unfolding preceding that of the secondary structural elements that comprise it (as it is in the case of tRNA, for example). The experimental melting profiles for the two larger RNA fragments do have additional apparent transitions at lower T_m in addition to that discussed above. To gain better understanding of the unfolding of these RNAs, we used the Global Melt Fit (GMF) program to fit each melting profile to a number of sequential unfolding transitions. Ideally, the total enthalpy from these transitions should be close to the enthalpy for unfolding from nearest neighbour calculations, as calculated in Mfold. This was not the case here, as the enthalpies of fitted transitions were all substantially lower. A plausible explanation is that (i) too few transitions were included in the fit and/or (ii) it may be that these RNAs substantially deviate from the implicit assumption that each unfolds in a series of sequential transitions. In the absence of other data (beyond the reasonable scope of this investigation), such as from differential scanning calorimetry (DSC), it is not possible to reliably calculate enthalpies and free energies.

However, for the purposes of this study, we wished to gain some insight to the origin of these lower temperature transition(s) to correlate with the direct information on RNA structure available from the NMR data. The low-temperature transition for the 38mer and 75mer RNAs can be fit with a single two-state transition. What might this correspond to in the RNA domain? From separate UV-melt analysis of 16mer, 38mer and 75mer RNAs containing specifically modified nucleotides (5-fluoro uridine), we deduced that the low-temperature apparent transition most likely corresponds to the C-rich heptanucleotide bulge (unpublished data). Although the detailed effects of the modification are unknown, only the lower temperature transition for these RNAs is unaffected by the modification; the 16mer melting profile and the higher T_m apparent transitions for the other RNAs are all changed. The heptanucleotide bulge is the only region lacking a uridine residue (Figure 1).

The 1D spectra of 16mer, 38mer and 75mer RNAs (without and with Mg^{2+}) show many resolved peaks with conserved chemical shifts among the three RNAs, for the low-field shifted, exchangeable imino proton resonances. This helped in carrying out the sequence-specific resonance assignment of the imino protons of 16mer and 38mer RNAs in their respective NOESY spectra. Owing to increased peak overlap and broader line widths, the NOESY spectrum of the 75mer produced only a smaller number of cross peaks and hence its assignment is incomplete and tentative in some cases. However, the important sequential connectivity, due to NOE between stacked imino protons of nucleotides G547/G549 within the tetraloop, and G546, allowed their unambiguous assignment. The successful assignment of the 16mer, 38mer and 75mer RNAs

simultaneously confirmed the secondary structural elements endowed on them, especially the GCGA tetraloop.

The effect of Mg^{2+} on the spectra of the RNA motifs is significant with selective changes induced to the chemical shifts, line widths and intensities of the imino proton resonances. The imino protons of G547/G549 showed distinctly larger changes in comparison to the resonances of other nucleotides, thus indicating a greater effect of Mg^{2+} binding to the GCGA tetraloop within the RNAs. The same imino protons of the tetraloop also showed, in the presence of Mg^{2+}, the maximum retardation to exchange with protons of solvent water for the 75mer, vis-à-vis a complete loss in intensity for the 16mer and only a slight retardation for the 38mer. The protection against solvent water exchange conferred using Mg^{2+} on the imino protons of GCGA tetraloop of the 75mer suggests a conformational change caused by the metal ion, in order to achieve greater thermodynamic stability for the otherwise labile tetraloop. These NMR results allow us to speculate that Mg^{2+} ions serve two major roles: (i) involvement in chelation to the GCGA tetraloop and serving as an anchor to promote the folding of the 75mer RNA and (ii) enhancing thermodynamic stability of the 75mer RNA via non-specific interactions with the phosphate backbone. These results are also consistent with the UV melting and molecular dynamics simulation data discussed above, and support the published biochemical data.

In summary, we have provided the first NMR experimental evidence for the folding and stability of a functionally important RNA motif of the EMCV IRES element. We have further demonstrated that Mg^{2+} plays a major role in the tertiary structure organization of the motif, including the critical GCGA tetraloop.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Miss Nicola Carden-Nowell for assistance with RNA preparation and the staff of MRC Biomedical NMR Centre (National Institute of Medical Research, Mill Hill, London, UK) for the allocation of high-field NMR time and valuable support. We are grateful to Dr G. Belsham (Institute of Animal Health, Pirbright, Surrey, UK) for valuable discussions and helpful comments on the manuscript. We thank the EPSRC and UMIST for postgraduate studentship support to M.P. and R.J.B., respectively. G.C. thanks The Wellcome Trust for a research career development fellowship (Ref. no. 061444). We are grateful to the NESBiC for a pump-priming grant support for the project.

REFERENCES

1. Sonenberg,N. and Dever,T. (2003) Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.*, **13**, 56–63.
2. Gingras,A.C., Raught,B. and Sonenberg,N. (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.*, **68**, 913–963.
3. Belsham,G. and Sonenberg,N. (1996) RNA–protein interactions in regulation of picornavirus RNA translation. *Microbiol. Rev.*, **60**, 499–511.
4. Roberts,L., Seamons,R. and Belsham,G. (1998) Recognition of picornavirus internal ribosome entry sites within cells; influence of cellular and viral proteins. *RNA*, **4**, 520–529.
5. Jackson,R., Howell,M. and Kaminski,A. (1990) The novel mechanism of initiation of picovirus RNA translation. *Trends Biochem. Sci.*, **15**, 477–483.
6. Jackson,R. and Kaminski,A. (1995) Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA*, **1**, 985–1000.
7. Pilipenko,E., Blinov,V., Romanova,L., Sinyakov,A., Maslova,S. and Agol,V. (1989) Conserved structural domains in the 5′ untranslated region of picornaviral genomes: an analysis of the segment controlling translation and neurovirulence. *Virology*, **168**, 201–209.
8. Pilipenko,E., Blinov,V. and Agol,V. (1990) Gross rearrangements within the 5′-untranslated region of the picornaviral genomes. *Nucleic Acids Res.*, **18**, 3371–3375.
9. Jang,S.K., Krausslich,H.-G., Nicklin,M.J.H., Duke,G.M., Palmenberg,A.C. and 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.*, 2636–2643.
10. Lopez de Quinto,S., Lafuente,E. and Martinez-Salas,E. (2001) IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. *RNA*, **7**, 1213–1226.
11. Palmenberg,A.C. and Sgro,J.-Y. (1997) Topological organization of picornaviral genomes: statistical prediction of RNA structural signals. *Semin. Virol.*, **8**, 231–241.
12. Witwer,C., Rauscher,S., Hofacker,I. and Stadler,P. (2001) Conserved RNA secondary structures in picornaviridae genomes. *Nucleic Acids Res.*, **29**, 5079–5089.
13. Van Der Velden,A., Kaminski,A., Jackson,R. and Belsham,G. (1995) Defective point mutants of the encephalomyocarditis virus internal ribosome entry site can be complemented in trans. *Virology*, **214**, 82–90.
14. Robertson,M., Seamons,R. and Belsham,G. (1999) A selection system for functional internal ribosome entry site (IRES) elements: analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES. *RNA*, **5**, 1167–1179.
15. Kolupaeva,V., Pestova,T., Hellen,C. and Shatsky,I. (1998) Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *J. Biol. Chem.*, **273**, 18599–18604.
16. Clark,A.T., Robertson,M.E.M., Conn,G.L. and Belsham,G.J. (2003) Conserved nucleotides within the J domain of the encephalomyocarditis virus internal ribosome entry site are required for activity and for interaction with eIF4G. *J. Virol.*, **77**, 12441–12449.
17. Roberts,L.O. and Belsham,G. (1997) Complementation of defective picornavirus internal ribosome entry site (IRES) elements by the co-expression of fragments of the IRES. *Virology*, **227**, 53–62.
18. Sarnow,P. (2003) Viral internal ribosome entry elements: novel ribosome–RNA complexes and roles in viral pathogenesis. *J. Virol.*, **77**, 2801–2806.
19. Lukavsky,P., Kim,I., Otto,G. and Puglisi,J. (2003) Structure of HCV IRES domain II determined by NMR. *Nature Struct. Biol.*, **10**, 1033–1038.
20. Walker,S.C., Avis,J.M. and Conn,G.L. (2003) General plasmids for producing RNA *in vitro* transcripts with homogeneous ends. *Nucleic Acids Res.*, **31**, e82.
21. Turner,D.H., Sugimoto,N. and Freier,S.M. (1988) RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 167–192.
22. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
23. Cheatham,T.E., Miller,J.L., Fox,T., Darden,T.A. and Kollman,P.A. (1995) Molecular dynamics simulations on solvated biomolecular systems: the particle mesh ewald method leads to stable trajectories of DNA, RNA and Proteins. *J. Am. Chem. Soc.*, **117**, 4193–4194.
24. Miller,J.L., Cheatham,T.E.,III and Kollman,P.A. (1999) Simulation of nucleic acid structure. In Neidle,S. (ed.), *Oxford Handbook of*

Nucleic Acid Structure. Oxford University Press, Chapter 4, pp. 95–115.
25. Cheatham,T.E. and Kollman,P. (1997) Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA and DNA:RNA duplexes. *J. Am. Chem. Soc*., **119**, 4805–4825.
26. Puglisi,J. and Tinoco,I. (1989) Absorbancy melting curves of RNA. *Methods Enzymol*., **180**, 304–325.
27. Conn,G.L., Gutell,R.R. and Draper,D.E. (1998) A functional ribosomal RNA tertiary structure involves a base triple interaction. *Biochemistry*, **37**, 11980–11988.
28. Bax,A. and Davis,D.G. (1985) MLEV-17 based 2D homonuclear magnetisation transfer spectroscopy. *J. Magn. Reson*., **65**, 355–360.
29. Bodenhausen,G., Kogler,H. and Ernst,R.R. (1984) Selection of coherence transfer pathways in NMR pulse experiments. *J. Magn. Reson*., **58**, 370–388.
30. Piotto,M., Saudek,V. and Sklenar,V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR*, **2**, 661–665.
31. Robertus,J.D., Ladner,J.E.., Finch,J.T., Rhodes,D., Brown,R.S., Clark,B.F.C. and Klug,A. (1974) Structure of yeast phenylalanine tRNA at 3Å resolution. *Nature*, **250**, 546–551.
32. Banks,R. (2003) Structure and energetics of a conserved RNA motif. MSc (Cheminformatics) Thesis, 'UMIST', Manchester, UK.
33. Heus,H. and Pardi,A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science*, **253**, 191–194.
34. Jucker,F., Heus,H., Yip,P., Moors,E. and Pardi,A. (1996) A network of heterogeneous H bonds on GNRA tetraloops. *J. Mol. Biol*., **264**, 968–980.
35. Perbrandt,M., Nolte,A., Lorenz,S., Bald,R., Betzel,C. and Erdmann,V.A. (1998) Crystal structure of domain E of *Therma flavus* 5S rRNA: a helical RNA structure including a hairpin loop. *FEBS Lett*., **429**, 211–245.
36. Wuthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, NY.
37. Laing,L., Gluick,T. and Draper,D. (1994) Stabilization of RNA structure by Mg ions. Specific and non-specific effects. *J. Mol. Biol*., **237**, 577–587.
38. Pyle,A.M. (2002) Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem*., **7**, 679–690.
39. Draper,D.E. (2004) A guide to ions and RNA structure. *RNA*, **10**, 335–343.
40. Gonzalez,R.L. and Tinoco,I. (2001) Identification and characterisation of metal ion binding sites in RNA. *Methods Enzymol*., **338**, 4421–4443.
41. Fernandez-Miragall,O. and Martinez-Salas,E. (2003) Structural organization of a viral IRES depends on the integrity of the GNRA motif. *RNA*, **9**, 1333–1344.
42. Auffinger,P., Bielecki,L. and Westhof,E. (2003) The Mg^{2+} binding sites of the 5S rRNA loop E motif as investigated by molecular dynamics simulations. *Chem. Biol*., **10**, 551–561.
43. Tinoco,I. and Bustamante,C. (1999) How RNA folds. *J. Mol. Biol*., **293**, 271–281.
44. Serra,M.J., Baird,J.D., Dale,T., Fey,B.L., Retatagos,K. and Westhof,E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. *RNA*, **8**, 307–323.
45. Cole,P.E., Yang,S.K. and Crothers,D.M. (1972) Conformational changes of transfer ribonucleic acid. Equilibrium phase diagrams. *Biochemistry*, **11**, 4358.
46. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.