Journal of General Virology (1998), **79**, 2359–2366. Printed in Great Britain

Translation efficiencies of the 5′ untranslated region from representatives of the six major genotypes of hepatitis C virus using a novel bicistronic reporter assay system

Adam J. Collier, Shixing Tang and Richard M. Elliott

Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, UK

The 5′ untranslated region (5′UTR) of hepatitis C virus (HCV) contains an internal ribosome entry site (IRES) which directs translation of the viral open reading frame (ORF). The 5′UTR is highly conserved between virus isolates in both primary sequence and predicted secondary structure. We cloned and sequenced the 5′ regions (nt 18 of the 5′UTR to nt 15 of the core coding sequence) of HCV isolates representing the six major genotypes and sub-cloned these into a bicistronic, dual luciferase reporter construct. The relative expression of the two luciferases, one directed by the HCV IRES and the other by cap-dependent ribosome scanning, was used to compare the activities of the different IRES elements in transfected cells. The 5′UTR from a genotype 2b isolate was the most efficient at directing translation in all four cell lines tested: BHK-21, HeLa-T4 , HuH7 and HepG2. In HepG2 cells the 2b 5′UTR was three times as efficient as the type 6a 5′UTR, which was generally the least active IRES tested. These data suggest that HCV isolates are not able to translate their ORF with equal efficiency, and provide a starting point from which further sequence-function studies can be undertaken.

Introduction

Hepatitis C virus (HCV) is the main causative agent of post-transfusion hepatitis (Choo *et al.*, 1989) and chronic infection often leads to liver cirrhosis and hepatocellular carcinoma (Saito *et al.*, 1990). The genome of HCV is a single-stranded, positive-sense RNA of approximately 9500 nucleotides (nt) which contains a single open reading frame (ORF) encoding a polyprotein of approximately 3000 amino acids. This polyprotein is cleaved by cellular and two virus-encoded proteases to yield the mature structural and non-structural proteins (reviewed in Houghton *et al.*, 1994). Based on genomic organization, HCV has been placed in a new genus, *Hepacivirus*, within the family *Flaviviridae*.

The majority of HCV isolates so far identified can be classified into six main groups, designated genotypes 1–6, with subdivisions within each (a, b, c, etc.) (Simmonds *et al.*, 1994; Chamberlain *et al.*, 1997*a*). Genotypes 1, 2 and 3 have a broad distribution in Europe, USA, the Far East and parts of Africa. Genotype 4 is the principal type in North and Central Africa and the Middle East, whilst genotypes 5 and 6 have a more restricted distribution, type 5 being found in South Africa and type 6 in Hong Kong (for review see Simmonds, 1995, and references therein).

Author for correspondence: Richard Elliott.
Fax +44 141 337 2236. e-mail elliott@vir.gla.ac.uk

The 5′-untranslated region (5′UTR) of the HCV genome (341 nt in most isolates) is highly conserved and has the potential to form a stable secondary and tertiary structure (Brown *et al.*, 1992; Honda *et al.*, 1996*a*; Le *et al.*, 1995; Lemon & Honda, 1997; Tsukiyama-Kohara *et al.*, 1992). Tsukiyama-Kohara *et al.* (1992) and Wang *et al.* (1993) showed that the HCV 5′UTR is able to direct translation of the ORF by a cap-independent mechanism in a manner analogous to that seen in picornaviruses (Pelletier & Sonenberg, 1988). In contrast to the picornaviruses, where only a portion of the much longer 5′UTR is involved in internal ribosome entry site (IRES) activity, the HCV IRES represents almost all of the 5′UTR sequence (Rijnbrand *et al.*, 1995). By comparison of predicted RNA secondary structures (Brown *et al.*, 1992; Honda *et al.*, 1996*a*; Lemon & Honda, 1997), it is also clear that the HCV 5′UTR is more similar to that of pestiviruses and GB virus B (GBV-B), other members of the *Flaviviridae*, than to that of picornaviruses.

To date, there has been little work undertaken comparing the efficiency of the IRES elements from different HCV genotypes. Tsukiyama-Kohara *et al.* (1992) investigated the

0001-5660

relative efficiencies of genotypes 1b and 2b to direct translation *in vitro*, work which has since been extended in an *in vivo* system by Kamoshita *et al*. (1997), and Buratti *et al*. (1997) who compared the IRES elements of genotypes 1b, 2a and 3a.

Here we compare the ability of the 5′UTRs from representatives of the six major HCV genotypes to initiate translation in four cell lines to investigate further the effects of the relatively minor differences in sequence and predicted secondary structure.

Methods

■ **Cell culture.** HuH7 cells (Nakabayashi *et al*., 1982) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% newborn calf serum (NCS). HeLa-T4$^+$ (Maddon *et al*., 1988) and HepG2 cells were grown in DMEM supplemented with 10% foetal calf serum (FCS) and 1% non-essential amino acids (NEAA). BHK-21 cells were grown in Glasgow modified Eagle's medium supplemented with 10% NCS, 100 U/ml penicillin, 100 U/ml streptomycin and 10% tryptose phosphate broth. All cells were grown as monolayer cultures.

■ **HCV cDNA clones.** cDNAs corresponding to the HCV 5′UTRs of genotypes 1a, 2b and 3a (established by restriction fragment length polymorphism; Davidson *et al*., 1995) were provided by V. McKechnie, University of Glasgow. Plasmids containing HCV sequences from nt 61 to 5595 of genotype 4a and nt 61 to 725 of genotype 6a were provided by R. Chamberlain (Adams *et al*., 1997; Chamberlain *et al*., 1997*a*). An HCV genotype 1b-positive serum was obtained from a Chinese blood donor (Tang *et al*., 1994) and a genotype 5a-positive serum was provided by P. Simmonds, University of Edinburgh (Chamberlain *et al*., 1997*b*). HCV RNA was extracted from the sera using the Qiamp viral RNA kit (Qiagen), according to the manufacturer's instructions. The purified RNAs were reverse transcribed using Moloney murine leukaemia virus reverse transcriptase (Mo-MLV RT, Promega) and primer AC8 (5′ CCGACGCTGCAGATGTACCCCATGAG) using conditions described previously (Chamberlain *et al*., 1997*a*).

The 5′UTRs of HCV genotypes 1a, 1b, 2b, 3a and 5a were amplified by PCR using the primer pairs listed in Table 1. PCR involved 30 cycles of heating at 94 °C for 45 s, 58 °C for 30 s and 72 °C for 1 min. *Taq* DNA polymerase (Gibco-BRL) was used for amplification. These primers amplified fragments from nt 18 of the 5′UTR to nt 15 of the virus coding sequence.

The 5′UTR of genotype 4a was constructed by inserting a fragment corresponding to nt 18 to 83 of genotype 1b into pTZ18R containing genotype 4a cDNA (nt 64 to 5595; Chamberlain *et al*., 1997*a*) using a unique *Nco*I site within the 5′UTR and an *Eco*RI site within the plasmid's polylinker region. This template was then amplified as described above, using primers 5A and 6-145, to give genotype 4a 5′UTR (nt 18 to 356). Similarly, the genotype 6a 5′UTR (nt 18 to 358) was constructed by inserting a fragment corresponding to nt 18 to 83 of genotype 1b into a plasmid containing nt 64 to 725 of genotype 6a (Adams *et al*., 1997). This template was then amplified as described above using primers 5A-6 and 6-6. It should be noted that nt 18 to 83 of genotypes 1b, 4a and 6a are identical except for positions 29 (A or G) and position 33 (deleted in 6a; Fig. 1A); these genotype-specific residues were maintained in the appropriate primers.

Following purification by agarose gel electrophoresis, PCR products from the six genotypes were digested with *Bam*HI and cloned into pTZ18R. These clones were then sequenced on an ABI Prism 377 DNA sequencer (Perkin-Elmer) using reverse and universal primers. To confirm the genotype of the isolated clones, sequences were matched with sequences in the GenBank and EMBL databases which have previously

Table 1. Primer sequences used to amplify representative HCV 5′UTR sequences

Nucleotides forming the *Bam*HI cloning sites are shown in lower-case letters.

Genotype	Forward primer	Reverse primer
1a	**5A**	**6-145**
	5′ ggatccGGCGACACTCC($^{G}_{A}$)CCA	5′ ggatccAGGATTCGTGCTCATGGTGC
1b	**5A**	**6-145**
	5′ ggatccGGCGACACTCC($^{G}_{A}$)CCA	5′ ggatccAGGATTCGTGCTCATGGTGC
2b	**5A**	**6-2**
	5′ ggatccGGCGACACTCC($^{G}_{A}$)CCA	5′ ggatccAGGATTTGTGCTCATGATGC
3a	**5A**	**6-3**
	5′ ggatccGGCGACACTCC($^{G}_{A}$)CCA	5′ ggatccAGGAAGTGTGCTCATGTTGC
4a	**5A**	**6-145**
	5′ ggatccGGCGACACTCC($^{G}_{A}$)CCA	5′ ggatccAGGATTCGTGCTCATGGTGC
5a	**5A**	**6-145**
	5′ ggatccGGCGACACTCC($^{G}_{A}$)CCA	5′ ggatccAGGATTCGTGCTCATGGTGC
6a	**5A-6**	**6-6**
	5′ ggatccGGCGACACTCCGCCATTATCC	5′ ggatccTGGAAGTGTGCTCATGATGC

A

```
                18                                                      67
1a          GGCGACACUC CaCCAUgaAU CACUCCCCUG UGAGGAACUa CUGUCUUCAC
1b          GGCGACACUC CaCCAUagAU CACUCCCCUG UGAGGAACUa CUGUCUUCAC
2b          GGCGACACUC CgCCAUgaAU CACUCCCCUG UGAGGAACUa CUGUCUUCAC
3a          GGCGACACUC CaCCAUggAU CACUCCCCUG UGAGGAACUu CUGUCUUCAC
4a          GGCGACACUC CaCCAUagAU CACUCCCCUG UGAGGAACUa CUGUCUUCAC
5a          GGCGACACUC CaCCAU.gAU CACUCCCCUG UGAGGAACUa CUGUCUUCAC
6a          GGCGACACUC CgCCAU.uAU CACUCCCCUG UGAGGAACUa CUGUCUUCAC
Consensus   GGCGACACUC C-CCAU--AU CACUCCCCUG UGAGGAACU- CUGUCUUCAC
                                                    IIa
                68                                                     117
1a          GCaGAAAGCG uCUAGCCAUG GCGUUAGUAu GAGUGUcGug CAGCCUCCAG
1b          GCaGAAAGCG uCUAGCCAUG GCGUUAGUAu GAGUGUcGug CAGCCUCCAG
2b          GCaGAAAGCG uCUAGCCAUG GCGUUAGUAu GAGUGUcGua CAGCCUCCAG
3a          GCgGAAAGCG cCUAGCCAUG GCGUUAGUAc GAGUGUcGug CAGCCUCCAG
4a          GCaGAAAGCG uCUAGCCAUG GCGUUAGUAu GAGUGUuGua CAGCCUCCAG
5a          GCaGAAAGCG uCUAGCCAUG GCGUUAGUAu GAGUGUcGaa CAGCCUCCAG
6a          GCaGAAAGCG uCUAGCCAUG GCGUUAGUAu GAGUGUcGua CAGCCUCCAG
Consensus   GC-GAAAGCG -CUAGCCAUG GCGUUAGUA- GAGUGU-G-- CAGCCUCCAG
                        IIb
                118                                                    167
1a          GaCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUgCGGAACC GGUGAGUACA
1b          GaCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUgCGGAACC GGUGAGUACA
2b          GcCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUgCGGAACC GGUGAGUACA
3a          GaCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUgCGGAACC GGUGAGUACA
4a          GaCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUuCGGAACC GGUGAGUACA
5a          GaCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUgCGGAACC GGUGAGUACA
6a          GcCCCCCCCU CCCGGGAGAG CCAUAGUGGU CUgCGGAACC GGUGAGUACA
Consensus   G-CCCCCCCU CCCGGGAGAG CCAUAGUGGU CU-CGGAACC GGUGAGUACA
                                                    IIIa
                168                                                    214
1a          CCGGAAUugC caGgacGACc GGGUCCUUUC ..UUGGAu.a AACCCgCUCa
1b          CCGGAAUugC caGgacGACc GGGUCCUUUC ..UUGGAu.c AACCCgCUCa
2b          CCGGAAUuaC cgGaaaGACu GGGUCCUUUC ..UUGGAu.a AACCCaCUCu
3a          CCGGAAUcgC ugGggaGACc GGGUCCUUUC ..UUGGAg.c AACCCgCUCa
4a          CCGGAAUcgC cgGgauGACc GGGUCCUUUC ..UUGGAu.u AACCCgCUCa
5a          CCGGAAUugC cgGgauGACc GGGUCCUUUC ..UUGGAu.a AACCCgCUCa
6a          CCGGAAUugC caGgauGACc GGGUCCUUUC caUUGGAuca AACCCgCUCa
Consensus   CCGGAAU--C --G---GAC- GGGUCCUUUC --UUGGA--- AACCC-CUC-
                                    IIIb
                215                                                    264
1a          AUgcCugGag AUUUGGGCGU GCCCCCGCaA GAcugCUAGC CGAGUAGuGU
1b          AUgcCugGag AUUUGGGCGU GCCCCCGCgA GAcugCUAGC CGAGUAGuGU
2b          AUguCcgGuc AUUUGGGCGU GCCCCCGCaA GAcugCUAGC CGAGUAGcGU
3a          AUacCcaGaa AUUUGGGCGU GCCCCCGCgA GAucaCUAGC CGAGUAGuGU
4a          AUgcCcgGaa AUUUGGGCGU GCCCCCGCaA GAcugCUAGC CGAGUAGuGU
5a          AUgcCcgGag AUUUGGGCGU GCCCCCGCgA GAcugCUAGC CGAGUAGuGU
6a          AUgcCugGag AUUUGGGCGU GCCCCCGCaA GAcugCUAGC CGAGUAGcGU
Consensus   AU--C--G-- AUUUGGGCGU GCCCCCGC-A GA---CUAGC CGAGUAG-GU
              IIIb        IIIc                          IIId
                265                                                    314
1a          UGGGUcGCGA AAGGCCUUGu GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
1b          UGGGUcGCGA AAGGCCUUGu GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
2b          UGGGUuGCGA AAGGCCUUGu GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
3a          UGGGUcGCGA AAGGCCUUGu GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
4a          UGGGUcGCGA AAGGCCUUGc GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
5a          UGGGUcGCGA AAGGCCUUGu GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
6a          UGGGUuGCGA AAGGCCUUGu GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
Consensus   UGGGU-GCGA AAGGCCUUG- GGUACUGCCU GAUAGGGUGC UUGCGAGUGC
              IIId                        IIIe
                315                                            356
1a          CCCGGGAGGU CUCGUAGACC GUGCAcCAUG AGCACgaaUC Cu
1b          CCCGGGAGGU CUCGUAGACC GUGCAcCAUG AGCACgaaUC Cu
2b          CCCGGGAGGU CUCGUAGACC GUGCAuCAUG AGCACaaaUC Cu
3a          CCCGGGAGGU CUCGUAGACC GUGCAaCAUG AGCACacuUC Cu
4a          CCCGGGAGGU CUCGUAGACC GUGCAcCAUG AGCACgaaUC Cu
5a          CCCGGGAGGU CUCGUAGACC GUGCAcCAUG AGCACgaaUC Cu
6a          CCCGGGAGGU CUCGUAGACC GUGCAuCAUG AGCACacuUC Ca
Consensus   CCCGGGAGGU CUCGUAGACC GUGCA-CAUG AGCAC---UC C-
                                    IV
```

B

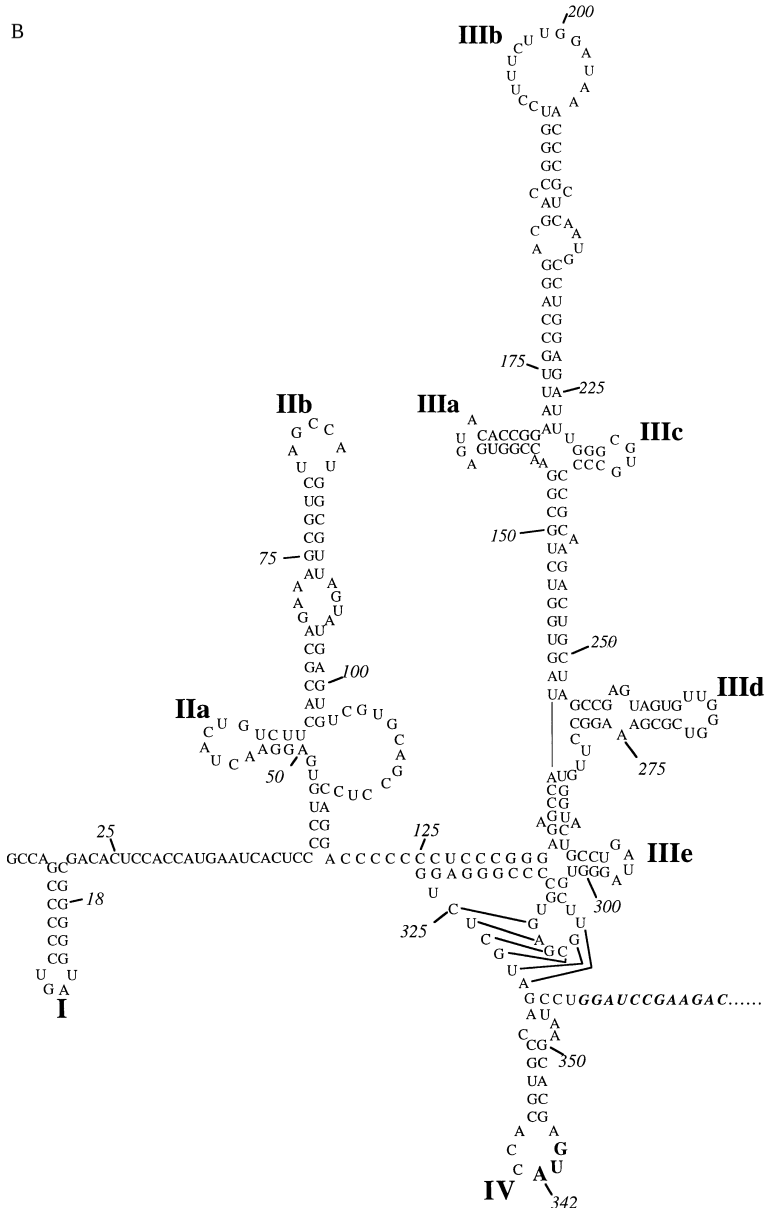

Fig. 1. HCV 5′UTR sequences. (A) Alignment of the HCV 5′UTR sequences used in this study. The sequences are from nt 18 to 356, numbered according to the genotype 1a sequence. The consensus sequence shows the residues that are conserved in all seven sequences, and the predicted stem–loop structures designated in Fig. 1 (B) are underlined. The sequences correspond to those of the prototypes of each genotype and have the following database accession numbers: 1a, M67463; 1b, D00832; 2b, D31606; 3a, D17763; 4a, Y11604; 5a, Y13184; 6a, Y12083. (B) Predicted secondary structure of the HCV 5′UTR. The sequence shown is that of the genotype 1a 5′UTR, HCV-H (Inchauspé *et al.*, 1991), and the structure is based on those proposed by Brown *et al.* (1992), Honda *et al.* (1996*a*) and Kamoshita *et al.* (1997). Stem–loop structures are labelled for reference.

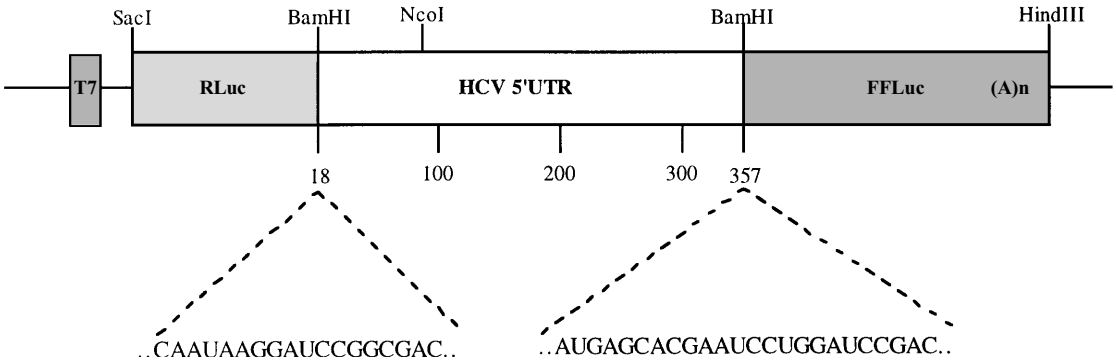

Fig. 2. Structure of the bicistronic, dual luciferase reporter plasmid. The two luciferase reporter genes are indicated by shaded boxes; RLuc, renilla luciferase and FFLuc, firefly luciferase. The HCV RNA sequence is shown as an open box with the junctions between the HCV RNA sequence and the reporter sequences underneath. Restriction enzyme cleavage sites used for DNA manipulation are indicated over the cDNA. T7, T7 promoter sequence.

been assigned to specific genotype groups. An alignment of the sequences used in this work is shown in Fig. 1(A).

The bicistronic reporter vector, pRL, which contains the renilla luciferase and firefly luciferase genes flanking a *Bam*HI cloning site, was constructed as follows. The upstream (control) reporter (renilla luciferase) gene was amplified by PCR using 10 ng of pRL-Null (Promega) as a template and 50 pmol each of primers AC7A (5′ GAGCTCATGACTTCGAAAGTT) and AC7B (5′ GGATCCTTATTGTTCATTTTT) in a final reaction volume of 50 µl. The downstream (assay) reporter (firefly luciferase) gene was amplified using 10 ng of pGEM2:luc (Promega) as a template and 50 pmol each of primers AC3C (5′ GGATCCGAAGACGCCAAAAAC) and AC4 (5′ AAGCTTTTTTTTTTTTTTTTTTTACAATTTGGACTT) in a final reaction volume of 50 µl. PCR involved 25 cycles of heating at 94 °C for 45 s, 55 °C for 30 s and 72 °C for 2 min. These PCR products were digested with *Sac*I and *Bam*HI or *Bam*HI and *Hin*dIII respectively, purified by agarose gel electrophoresis, cloned into pTZ18R and their sequence determined as above. The different 5′UTR cDNAs were then subcloned into pRL at the *Bam*HI site (Fig. 2), and the correct orientation was confirmed by appropriate restriction enzyme digestion and nucleotide sequencing.

■ **Transfection of cultured cells.** Subconfluent monolayers of cells in 24-well plates (15 mm diameter) were infected with vTF7-3, a recombinant vaccinia virus expressing T7 RNA polymerase (Fuerst *et al.*, 1986), at 5 p.f.u. per cell in 300 µl serum-free medium (OptiMEM; Gibco-BRL) for 30 min at 37 °C. The inoculum was removed and the cells washed once with OptiMEM. The cells were then transfected with plasmid DNA in 500 µl OptiMEM containing liposomes (15 µl for HeLa-T4$^+$ and BHK-21, 5 µl for HuH7 and HepG2 cell transfections) prepared as described by Rose *et al.* (1991). Four replicate wells were transfected with each construct, and each experiment was performed at least twice by two independent workers. Following a 2 h incubation at 37 °C, 1 ml of growth medium was added to the wells. The cells were harvested 16 h later and cell lysates assayed for luciferase activity as described below.

■ **Dual reporter assay.** Cell lysates were prepared from transfected cells and dual luciferase assays performed exactly according to the manufacturer's instructions (Promega). Luciferase activities were measured using a Biotrace M3 benchtop luminometer and the ratio of firefly luciferase to renilla luciferase activity was calculated and used as a measure of IRES function. Arbitrarily, the ratio obtained using the genotype 1a IRES was taken as 100% activity and the ratios calculated for the other genotype IRES elements were normalized to this.

Results and Discussion

HCV 5′UTR sequences

In designing the constructs for this study, the paramount consideration was to include the sequences within the HCV RNA which previous studies suggested represent the minimum IRES element. Hence, the constructs employed include the sequence from nt 18 of the 5′UTR to nt 15 of the core protein-coding sequence. This is because the 5′ boundary of the IRES appears to reside downstream of the most distal stem–loop, and indeed removal of domain I (see Fig. 1B) may slightly enhance activity (Honda *et al.*, 1996*b*; Kamoshita *et al.*, 1997; Reynolds *et al.*, 1995; Rijnbrand *et al.*, 1995; Tsukiyama-Kohara *et al.*, 1992; Wang *et al.*, 1993). The exact 3′ boundary of the HCV IRES is still the subject of much debate (reviewed by Lemon & Honda, 1997). Tsukiyama-Kohara *et al.* (1992) and Wang *et al.* (1993) demonstrated efficient expression of reporter sequences fused directly to the initiator AUG, whilst Reynolds *et al.* (1995) presented data which suggested that 12 to 30 nt of the core coding sequence were an absolute requirement for IRES activity. More recently the predicted secondary structure model for the HCV IRES has been refined following comparative sequence analysis with the GBV-B 5′UTR, double-strand specific RNase cleavage analysis and site-directed mutagenesis (Honda *et al.*, 1996*a*; Lemon & Honda, 1997). A conclusion from this work was that the core coding sequence that constitutes part of the domain IV structure (Fig. 1B) was not obligatory for translation but may be necessary to modulate translation during virus replication. In the present study the core coding sequences needed to form the most proximal stem–loop of domain IV were included to reflect current information regarding the predicted secondary structure of the HCV IRES.

Fig. 1(A) shows the nucleotide sequences of the different 5′UTR regions used in this study. A predicted secondary structure for the 5′UTR of HCV-H (Inchauspé *et al.*, 1991)

based on models proposed by Brown *et al*. (1992), Honda *et al*. (1996*a*) and Kamoshita *et al*. (1997) is shown in Fig. 1(B). Within the highly conserved primary sequence of the 5′UTR, there are small regions of sequence polymorphism (reviewed by Smith *et al*., 1995), most notably between nt 175 to 187 and nt 210 to 224. While many sequence differences are covariant maintaining the predicted secondary structure, this is not always the case, resulting in either an increase or decrease in the predicted stability of the double-stranded regions (e.g. in stem–loop IIIc). At positions 175 and 224 genotypes 3a and 4a have U to C and G to A changes, respectively, which maintain non-Watson–Crick base pairing at this site but may destabilize the secondary structure. Genotype 2b has a G to C change at position 224 which has a similar effect. At positions 176 and 223, genotype 2b increases the stability of the stem structure by introducing a Watson–Crick base pair, while at positions 178 to 181 and 218 to 221 there are a series of covariant changes. At positions 183 and 214 there are further differences, which have the potential to introduce new base pair interactions and these alter the predicted secondary structure in this region. One of the most radical differences within the primary sequences is in the apical loop (IIIb) of the type 6a 5′UTR, where there are three insertions. This is not seen in any of the other genotypes, all of which maintain the same number of nucleotides in the loop. Genotype 3a also includes three changes between stem–loops IIIc and IIId (nt 247 to 249); again these changes may alter the stem–loop structure. The final cluster of changes occurs within domain IV, nt 351 to 353, which is within the core protein-coding sequence.

Comparison of the translational efficiencies of 5′UTR in cultured cells

Our assay procedure used a bicistronic mRNA to analyse IRES activity both *in vivo* and *in vitro*. The downstream cistron would ordinarily be accessed inefficiently by ribosomes which have completed translation of the upstream cistron; however, if an IRES sequence is inserted before the downstream ORF, translation is considerably increased (reviewed by Borman *et al*., 1995). In transfection experiments the upstream translation product acts as an internal control to account for differences in transfection efficiency. The dual luciferase system was chosen because it satisfies all of the required criteria in terms of speed, ease of use, accuracy and reproducibility, with typical errors of 2% to 15% between replicates within an assay. Importantly, both reporter enzymes are assayed in the same cell lysate preparation, and the activities are determined in essentially the same way using a benchtop luminometer. In addition, the dose-response curves of the two reporters are linear over at least seven orders of magnitude of protein concentration, so that the ratio of the two luciferase activities from any one construct will essentially remain constant over a large range of input DNA. Plasmids were constructed in which a T7 promoter drives transcription of the bicistronic mRNA comprising renilla luciferase, the HCV 5′UTR sequence and then firefly luciferase (Fig. 2); these plasmids were transfected into cells infected with vTF7-3, a recombinant vaccinia virus which expresses T7 RNA polymerase (Fuerst *et al*., 1985), to allow cytoplasmic transcription of the bicistronic mRNA.

We have compared the IRES activity of representatives of all the major HCV genotypes using the above assay. Four different cell lines were used, BHK-21, HeLa-T4$^+$, HuH7 and HepG2, and each construct was transfected into four replicate monolayers; IRES activity was calculated as the ratio of the two luciferase activities. The results of these experiments are summarized in Fig. 3. Arbitrarily the activity of the prototypic 1a 5′UTR sequence, as found in HCV-H (Inchauspé *et al*., 1991), was taken as 100%, and the activities of the other IRES elements were normalized with respect to this. The genotype 2b 5′UTR was the most efficient at initiating translation in all four cell lines tested, whereas the genotype 6a IRES was the least efficient in BHK and particularly in HepG2 cells. In HuH7 cells there was the least variation between the efficiencies of the different IRES elements. Relative to the 1a IRES, the IRES elements of genotypes 1b, 3a, 4a and 5a were similar in their ability to initiate translation in all the cell lines used. The greatest difference between IRES elements was seen in HepG2 cells, where the 2b IRES was nearly 3-fold more active than the 6a IRES.

By normalizing the renilla luciferase activities obtained in different cell types with respect to each other, and then normalizing the firefly luciferase activities, we estimated that HCV IRES activity was about 2-fold higher in BHK and HepG2 cells than in HeLa-T4$^+$ or HuH7 cells. Our results are in agreement with those of Borman *et al*. (1997) who also found HCV IRES functioned similarly in BHK and HepG2 cells, but lower in HeLa cells. These data suggest that the interaction of the HCV IRES with host factors may differ between different cell types, though it is perhaps surprising that one of the human hepatocyte-derived lines, HuH7, did not support higher levels of IRES activity compared with the non-primate, non-hepatocyte BHK cell line. Recently Pestova *et al*. (1998) reconstituted HCV IRES-mediated internal ribosome entry *in vitro* from purified translation components, and showed that there was no absolute requirement for non-canonical cellular factors for internal initiation of translation. However, these results do not preclude a role for cellular factors in modulating IRES activity.

The vaccinia virus–T7 system using transfected plasmid DNAs is a particularly convenient method to assess IRES activity, but we also investigated alternative approaches. Capped, *in vitro*-transcribed RNAs were made and transfected directly into cells using cationic lipids. The overall luciferase activities were lower than with DNA transfections, but comparison of the activities of the different IRES elements gave results similar to those shown in Fig. 3 (data not shown).

Many authors have examined IRES activity in cell-free systems, and so we also investigated the bicistronic constructs

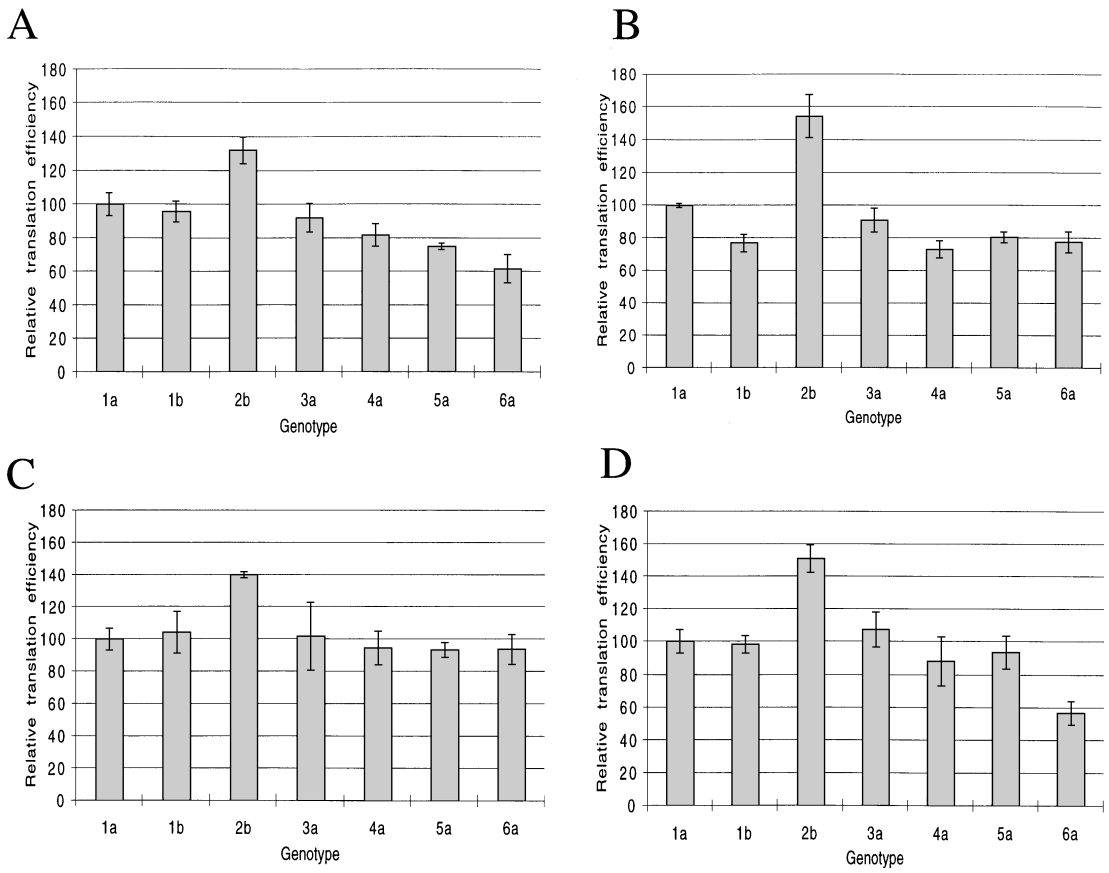

Fig. 3. Relative efficiencies of the IRES elements of different HCV genotypes *in vivo*. IRES activity was assessed by measuring the ratio of renilla and firefly luciferases produced from transfected bicistronic plasmid constructs into vTF7-3-infected cells. The activity of the genotype 1a IRES was arbitrarily taken as 100% with the activities of the other IRES elements expressed relative to this. The results are shown for (A) BHK-21 cells, (B) HeLa-T4$^+$ cells, (C) HuH7 cells and (D) HepG2 cells.

in rabbit reticulocyte lysates, either by translation of *in vitro*-transcribed RNAs or by coupled *in vitro* transcription–translation (Promega TnT Quick Coupled Transcription/Translation system). The genotype 2b IRES was always the most active at directing translation of firefly luciferase (as measured enzymically or by gel electrophoresis of radiolabelled proteins followed by phosphor-imaging). However, the degree of enhancement of the 2b IRES compared to the IRES elements of the other genotypes, and the relative activities of the different IRES elements, was more variable between different experiments (data not shown). Thus we are less confident in using the cell-free system than the *in vivo* approach. A complication of the cell-free system is that the optimal potassium ion concentration differs for cap-dependent and IRES-dependent translation (Borman *et al.*, 1995) such that translations are carried out under conditions which favour either one or the other activity, or else are suboptimal for both, making comparisons difficult.

Conclusions

Previous studies have reported differences in the IRES efficiency when comparing 5′UTR sequences from two or three different HCV types *in vitro* or *in vivo* (Tsukiyama-Kohara *et al.*, 1992; Buratti *et al.*, 1997; Kamoshita *et al.*, 1997). Our results confirm those obtained *in vitro* by Tsukiyama-Kohara *et al.* (1992) and *in vivo* by Kamoshita *et al.* (1997) that HCV genotype 2b has the most active IRES. However, in our experiments the IRES activities of the 1b and 3a genotypes were similar whereas Buratti *et al.* (1997) reported the 3a IRES to be only 50% as active as 1b. However, as acknowledged by the authors, the 3a construct of Buratti *et al.* contained genotype 1 core sequences rather than genotype 3 as in our construct, which may have affected the secondary structure and hence IRES activity. Considering that there are only 17 nucleotide differences between the most and least active IRES elements, genotypes 2b and 6a, it seems likely that relatively subtle effects on secondary structure may have more drastic effects on function. These results provide a starting point for such structure–function analysis, but mapping the important domains may prove difficult, however, as Kamoshita *et al.* (1997) were unable to identify regions influencing activity in chimeric IRES elements derived from genotypes 1b and 2b.

Some reports have suggested biological differences between genotypes 1 and 2, in terms of quantity of virus in serum

or sensitivity to interferon (Kohara *et al*., 1995; Zein & Persing, 1996) and in the distribution of type 1 and type 2 HCV infections, which may correlate with the replication efficiencies of the various genotypes. One factor affecting replication may be initiation of protein translation. Recent reports describing the establishment of infectious clones of HCV which cause disease in chimpanzees (Kolykhalov *et al*., 1997; Yanagi *et al*., 1997) make it feasible to study the effects, if any, of different IRES elements on the replication and disease progression of HCV in these animals.

Finally, it should be noted that the sequences we have used in these experiments correspond to those described for the different genotypes of HCV, but there is some variation in the 5′UTR even between different isolates of the same genotype. Hence caution should be exercised in describing the noted effects as 'genotype-specific' – rather the effects are specific for a particular nucleotide sequence.

We thank R. Chamberlain, M. McElwee and P. Simmonds for providing some of the cDNA clones or HCV-positive sera used in this study. This work was supported by an MRC (ROPA) Project Grant to R.M.E., a Wellcome Trust Travelling Fellowship (047027/2/96/2) to S.T. and a Wellcome Trust Equipment Grant (046745/Z/96).

References

Adams, N., Chamberlain, R. W., Taylor, L., Davidson, F., Lin, C. K., Elliott, R. M. & Simmonds, P. (1997). Complete coding sequence of hepatitis C virus genotype 6a. *Biochemical and Biophysical Research Communications* **234**, 393–396.

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*. *Nucleic Acids Research* **23**, 3656–3663.

Borman, A. M., Le Mercier, P., Girard, M. & Kean, K. M. (1997). Comparison of picornaviral IRES-driven internal initiation of translation in cultured cells of different origins. *Nucleic Acids Research* **25**, 925–932.

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. *Nucleic Acids Research* **20**, 5041–5045.

Buratti, E., Gerotto, M., Pontisso, P., Alberti, A., Tisminetzky, S. G. & Baralle, F. E. (1997). In vivo translational efficiency of different hepatitis C virus 5′-UTRs. *FEBS Letters* **411**, 275–280.

Chamberlain, R. W., Adams, N., Saeed, A. A., Simmonds, P. & Elliott, R. M. (1997*a*). Complete nucleotide sequence of a type 4 hepatitis C virus variant, the predominant genotype in the Middle East. *Journal of General Virology* **78**, 1341–1347.

Chamberlain, R. W., Adams, N., Taylor, L. A., Simmonds, P. & Elliott, R. M. (1997*b*). The complete coding sequence of hepatitis C virus genotype 5a, the predominant genotype in South Africa. *Biochemical and Biophysical Research Communications* **236**, 44–49.

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.

Davidson, F., Simmonds, P., Ferguson, J. C., Jarvis, L. M., Dow, B. C., Follett, E. A. C., Seed, C. R. G., Krusius, T., Lin, C., Medgyesi, G. A., Kiyokawa, H., Olim, G., Duraisamy, G., Cuypers, T., Saeed, A. A., Teo, D., Conradie, J., Kew, M. C., Lin, M., Nuchaprayoon, C., Ndimbie, O. K. & Yap, P. L. (1995). Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5′ non-coding region. *Journal of General Virology* **76**, 1197–1204.

Fuerst, T. R., Niles, E. G., Studier, F. W. & Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesises bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences, USA* **83**, 8122–8126.

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

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

Houghton, M., Selby, M., Weiner, A. & Choo, Q.-L. (1994). Hepatitis C virus. *Current Studies in Hematology and Blood Transfusion* **61**, 1–11.

Inchauspé, G., Zebedee, S., Lee, D.-H., Sugitani, M., Nasoff, M. & Prince, A. M. (1991). Genomic structure of the human prototype strain H of hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 10292–10296.

Kamoshita, N., Tsukiyama-Kohara, K., Kohara, M. & Nomoto, A. (1997). Genetic analysis of internal ribosome entry site on hepatitis C virus RNA: implication for involvement of the highly ordered structure and cell type-specific transacting factors. *Virology* **233**, 9–19.

Kohara, M., Tanaka, T., Tsukiyama-Kohara, K., Tanaka, S., Mizokami, M., Lau, J. Y. N. & Hattori, N. (1995). Hepatitis C virus genotypes 1 and 2 respond to interferon-α with different virologic kinetics. *Journal of Infectious Diseases* **172**, 934–938.

Kolykhalov, A. A., Agapov, E. V., Blight, K. J., Mihalik, K., Feinstone, S. M. & Rice, C. M. (1997). Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* **277**, 570–574.

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

Lemon, S. M. & Honda, M. (1997). Internal ribosome entry sites within the RNA genomes of hepatitis C virus and other flaviviruses. *Seminars in Virology* **8**, 274–288.

Maddon, P. J., McDougal, J. S., Clapham, P. R., Dalgleish, A. G., Jamal, S., Weiss, R. A. & Axel, R. (1988). HIV infection does not require endocytosis of its receptor, CD4. *Cell* **54**, 865–874.

Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T. & Sato, J. (1982). Growth of human hepatoma cell lines with differentiated functions in chemically defined medium. *Cancer Research* **42**, 3858–3863.

Pelletier, J. & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320–325.

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 RNAs. *Genes & Development* **12**, 67–83.

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

Rijnbrand, R., Bredenbeek, P., van der Straaten, T., Whetter, L., Inchauspé, 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.

Rose, J. K., Buonocore, L. & Whitt, M. A. (1991). A new cationic liposome reagent mediating nearly quantitative transfection of animal cells. *Biotechniques* **10**, 520–525.

Saito, I., Miyamura, T., Ohbayashi, A., Harada, H., Katayama, T., Kikuchi, S., Watanabe, Y., Koi, S., Onji, M. Y. O., Choo, Q.-L., Houghton, M. & Kuo, G. (1990). Hepatitis C infection is associated with the development of hepatocellular carcinoma. *Proceedings of the National Academy of Sciences, USA* **87**, 6547–6549.

Simmonds, P. (1995). Variability of hepatitis C virus. *Hepatology* **21**, 570–583.

Simmonds, P., Alberti, A., Alter, H. J., Bonino, F., Bradley, D. W., Brechot, C., Brouwer, J. T., Chan, S. W., Chayama, K., Chen, D. S., Choo, Q. L., Colombo, M., Cuypers, H. T. M., Date, T., Dusheiko, G. M., Esteban, J. L., Fay, O., Hadziyannis, S. J., Han, J., Hatzakis, A., Holmes, E. C., Hotta, H., Houghton, M., Irvine, B., Kohara, M., Kolberg, J. A., Kuo, G., Lau, J. Y. N., Lelie, P. N., Maertens, G., McOmish, F., Miyamura, T., Mizokami, M., Nomoto, A., Prince, A. M., Reesink, H. W., Rice, C., Roggendorf, M., Schalm, S. W., Shikata, T., Shimotohno, K., Stuyver, L., Trepo, C., Weiner, A., Yap, P. L. & Urdea, M. S. (1994). A proposed nomenclature of hepatitis C genotypes. *Hepatology* **19**, 1321–1324.

Smith, D. B., Mellor, J., Jarvis, L. M., Davidson, F., Kolberg, J., Urdea, M., Yap, P.-L. Simmonds, P. & The International HCV Collaborative Study Group (1995). Variation of the hepatitis C virus 5′ non-coding region: implications for secondary structure, virus detection and typing. *Journal of General Virology* **76**, 1749–1761.

Tang, S. X., Meng, Q. H., Ma, X. K., Zhang, X. T. & Jiang, Y. T. (1994). HCV RNA detection and genotyping by polymerase chain reaction in anti-HCV positive professional blood donors in China. *Chinese Journal of Hepatology* **2**, 33–35.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). Internal ribosome entry site with hepatitis C virus RNA. *Journal of Virology* **66**, 1476–1483.

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. *Journal of Virology* **67**, 3338–3344.

Yanagi, M., Purcell, R. H., Emerson, S. U. & Buhk, J. (1997). Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. *Proceedings of the National Academy of Sciences, USA* **94**, 8738–8743.

Zein, N. N. & Persing, D. H. (1996). Hepatitis C genotypes: current trends and future implications. *Mayo Clinic Proceedings* **71**, 458–462.

Received 29 April 1998; Accepted 22 June 1998