UAB
Universitat Autònoma de Barcelona

Dipòsit digital de documents de la UAB

This is the **accepted version** of the journal article:

Nadal, Anna; Robertson, Hugh D.; Guardia, Jaime; [et al.]. «Characterization of the structure and variability of an internal region of hepatitis C virus RNA for M1 RNA guide sequence ribozyme targeting». Journal of General Virology, Vol. 84 Núm. 6 (June 2003), p. 1545-1548. DOI 10.1099/vir.0.18898-0

This version is available at https://ddd.uab.cat/record/305677

under the terms of the IN COPYRIGHT license

Anna Nadal; Hugh D. Robertson; Jaime Guardia and Jordi Gómez (2003). The definitive peer reviewed, edited version of this article is published in Journal of General Virology, vol. 84, issue. 6, 2003, https://doi.org/10.1099/vir.0.18898-0

Characterization of the structure andvariability of an internal region of hepatitis C virus RNA for M1 RNA guide sequence ribozyme targeting

Anna Nadal,[1] Hugh D. Robertson,[2] Jaime Guardia[1] and Jordi Gómez[1*]

[1]Servicio de Medicina Interna-Hepatología, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, 08035 Barcelona, Spain

[2]Department of Biochemistry, Weill Medical College of Cornell University, New York, NY 10021, USA

*Corresponding author

ABSTRACT

Accessibility to folded RNA and low potential of variation in the target RNA are crucial requirements for ribozyme therapy against virus infections. In hepatitis C virus (HCV), the sequence of the 5'UTR is conserved but the highly folded RNA structure severely limits the number of accessible sites. To expand investigation of targeting in the HCV genome, we have considered an internal genomic region whose sequence variation has been widely investigated and which has a particularly conserved RNA structure, which makes it accessible to the human RNase P *in vitro*. We have first mapped the accessibility of the genomic RNA to complementary DNAs within this internal genomic region. We performed a kinetic and thermodynamic study. Accordingly, we have designed and assayed four RNase P M1 RNA guide sequence ribozymes targeted to the selected sites.

Considerations of RNA structural accessibility and sequence variation indicate that several target sites should be defined for simultaneous attack.

INTRODUCTION

M1 RNA is the catalytic subunit of RNase P from Escherichia coli, a ribonucleoprotein complex that removes a 5' leader sequence from tRNA precursors (Guerrier-Takada et al., 1983). M1 RNA may be converted to an endoribonuclease that specifically cleaves other RNAs through the covalent attachment of a guide sequence (GS) which is complementary to the target RNA and to the 3' end of the M1 RNA. Reduced levels of virus replication of herpes simplex virus (Trang et al., 2000a) and cytomegalovirus (Trang et al., 2000b) in cell culture have been reported with the use of M1GSs.

Hepatitis C virus (HCV) is a good candidate for ribozyme therapeutic studies because treatments available for chronic hepatitis, especially in HCV genotype 1b-infected patients, yield a small percentage of successful responses. One of the problems associated with this pioneering task is the quasispecies nature of the HCV genome – a wide spectrum of mutant sequences is generated soon after infection and changes continuously in composition over time (Martell et al., 1992; Domingo et al., 1999). The 5'UTR is the most conserved locus within the viral genome and efforts regarding HCV RNA therapeutics have been focused on this locus (Macejack et al., 2000). Nevertheless, this region has a highly stable RNA structure, which severely limits the number of accessible sites for ribozyme targeting.

MATERIALS AND METHODS

DNA templates for in vitro transcription of catalytic RNAs were constructed by PCR with the gene for M1 RNA (Liu & Altman, 1995). The 5' PCR primer was oligo101 (5'-GGAATTCTAATACGACTCACTATAGAAGCTGACCAG-ACAGTCGCC-3'). The 3' primers, which contained the appropriate GSs, were M1HCV2682 (5'-GTGGTGGCTG-GTCCCTGGTATGACCATG-3'), M1HCV2734 (5'-GTG-GTGCTCCTCTTGACATATGACCATG-3'), M1HCV2743 (5'-GTGGTGACATTACCACCATATGACCATG-3') and M1HCV2770 (5'-GTGGTGGACCGGGAGATGTATGACCATG-3'). Sequences in bold correspond to GSs that allow M1GS RNAs to hybridize with HCV target sequences at the corresponding position. For ribozyme cleavage assays of the HCV transcript, non-labelled, re-natured M1GSs (8 nM) were incubated with labelled SI transcript RNA (8 nM) at 37˚C in a volume of 10 ml for 30 min in 50 mM Tris, pH 7?5, 100 mM NH4Cl, 100 mM MgCl2 and 4 % (wt/vol) polyethylene glycol. Reactions were stopped by addition of 7 M urea and the cleavage products were separated on 4 % denaturing polyacrylamide gels and autoradiographed.

RESULTS AND DISCUSSION

To discover vulnerable sites at other loci of the HCV genome, we chose the RNase P M1 RNA model. Because the M1 ribozyme is a large molecule whose recognition capability has steric requirements, especially when dealing with long RNA substrates, we analysed accessibility in an internal viral genomic region comprising the structural/ non-structural junction region (Fig. 1). Accumulated knowledge, which favours analysis of this region, includes the recent demonstration that the relevant RNA structure is accessible to human RNase P in vitro and that accessibility is conserved despite variation in the primary sequence (Nadal et al., 2002). In addition, hundreds of viral sequences from a large cohort of patients suffering from hepatitis C have been determined for this region, thereby providing information on sequence variation.

RNA used as substrate (SI RNA) in the mapping study, corresponding to HCV nt 2486–3040, was obtained by transcription of a plasmid derived from pN(1–4728) Bluescript (Nadal et al., 2002). DNA-directed RNase H cleavage was performed exactly as described by Lyons et al. (2001). For gel-shift analysis, oligonucleotides were labelled at the 5' end with [c-32P]ATP, annealed in hybridization buffer (20 mM HEPES, pH 8?0, 50 mM KCl, 10 mM MgCl2 and 1 mM DTT) to non-labelled SI transcript RNA and subsequently electrophoresed under non-denaturing con- ditions in 4 % polyacrylamide gels. Control reactions without RNA transcripts were also performed.

Structural accessibility analyses with HCV SI RNA trans- cripts involved a 554 base RNA and 10 complementary DNA oligonucleotides 18 bases in length. Fig. 1 shows a representation of the HCV genome, indicating the location of the SI transcript RNA and also each complementary oligonucleotide identified by the number of its first nucleotide residue. A gel-shift mobility assay and RNase H mapping methods were used.

Two sets of annealing reactions were performed, one for gel-shift mobility analysis and the other for RNase H analysis. These assays were evaluated as has been described previously (Lyons et al., 2001) and representative results for these experiments are shown in Fig. 1.

By correlating SI RNA band shifts and the rate of cleavage, we can distinguish regions of relatively open structure (C2745) from tightly folded ones (C2678), and even from others folded in very stable local structures (C2848). Interestingly, the inaccessibility of the region

flanking nt 2848 coincides with a tRNA-like structure (Nadal et al., 2002). Table 1 describes the results from the whole set of oligonucleotides tested. Consistent results were obtained with mobility shift and RNase H experiments.

Preheating increases both the percentage of band shifting and the extent of RNase H cleavage in most locations but has no effect in several others. RNase H cleavage kinetics analysis shows that cleavage increases continuously over time and that it does not reach saturation during the 60 min of the reaction, indicating that the RNA is not folded into alternative structures of very different accessibility.

Based on the hybridization results obtained with com- plementary DNAs described above, four GSs, each 13 nt in length, were designed and incorporated into the 3' end of the M1 ribozyme RNA. Two GSs were designed to target HCV RNA at the most accessible site, another targeted a partially accessible site and the last targeted a completely inaccessible site. GSs were designed according to the substrate requirements for a G base 5' to the scissile bond (Trang et al., 2000a), although small displacements relative to the corresponding DNA oligonucleotide were thus introduced. The results of cleavage assays (Fig. 2) showed significant differences between each M1GS, as predicted from the RNase H cleavage and gel-shift patterns. Two of the four GSs were found to induce M1 cleavage of the target RNA (Fig. 2, lanes 4 and 5). The sizes of the cleavage products were compared to size markers in the gel and found to be in agreement with the expected sizes.

Regarding HCV variability, sequence analysis was per- formed on fragments of 212 nt (nt 2658–2869) encompas- sing the central region of SI RNA. We took advantage of the 540 nucleotide sequences obtained from virus samples (serum and liver) from 39 patients chronically infected with HCV (genotype 1b); these sequences had been obtained previously in our laboratory (Cabot et al., 2000). On average, seven sequences (4–12) were aligned for each sample, serum and liver, and the 'quasispecies consensus sequence' (i.e. after averaging the sequences that comprise the particular quasispecies) was obtained. We evaluated the percentage of conservation in consecutive sequence intervals, 13 nt in length, starting from position 2658 and ending at position 2869 – the optimal length for a GS linked to M1 RNA to direct specific cleavage of the RNA substrate is 13 complementary bases (Liu & Altman, 1995). There was no potential target in which the aligned sequences differed by less than four point mutations. In 68 % of the intervals, there were four to seven point mutations, while the

rest of the intervals differed by more than 50 % of the positions (data not shown). The EMBL database accession numbers for the sequences analysed are AJ247658–AJ248197.

Ribozyme therapeutics against pathogenic RNAs have two main factors. One is the ribozyme: its activity, stability, delivery and co-localization with the target; the other is the RNA substrate, the properties of which, in the case of real virus quasispecies, most likely limit the efficiency of a ribozyme as an effective antiviral candidate. Here we have explored the effect of RNA structural accessibility in vitro and the variation at the target sites from viral RNA sequences from HCV-infected patients.

Conclusions from several studies (reviewed briefly by Lima et al., 1997) suggest a correlation between oligonucleotides binding to a folded RNA structure in vitro and its biological activity. The HCV SI RNA transcript of 554 nt is large enough to expect a folded structure. In fact, we have described recently that a part of this RNA adopts a tRNA- like structure (Nadal et al., 2002). Furthermore, the 5'- and 3'-termini of the RNA transcript correspond to sequence positions internal to the viral RNA. Such artificial termini may lower the stability of regions near them. Therefore, we limited targeting to the central region of the transcript (nt 2658–2869), the region for which sequence variation data were available from previous studies. In comparison with structural studies employing similar methods performed in the HCV 5'UTR (Lyons et al., 2001), we observed both open and closed RNA domains. The open region in S1 does not reach 100 % accessibility, as in some sites of the 5'UTR, but the transition between open and closed domains is not so abrupt. The potential consequence for targeting is that the efficiency of ribozyme cleavage might be reduced but the number of target possibilities is not so restricted.

Cleavage efficiencies obtained by the M1GSs corresponded well with the results of structural accessibility analyses. This may be interpreted in that the folded M1 ribozyme did not impede the binding of GSs to the target and that the type of mapping performed in this work is useful to detect sites vulnerable to M1GS attack.

Nevertheless, the RNase H and M1GS patterns of cleavage were not exactly the same. This was particularly evident at sequence 2745–2761 (Fig. 2), where M1GS cleavage activity was lower than expected. Because cleavage was designed to occur between the optimal 5' guanosine G2743 and 3' pyrimidine U2742 residues (Trang et al., 2000a), the reduced

cleavage activity at this site may be best ascribed to local steric or electrostatic repulsion effects on the M1 ribozyme.

Considering variation, alignments of quasispecies consensus sequences obtained from 39 different patients and from different samples (liver or serum) revealed that, in the virus loci studied, the most structurally 'open' region contains genetic motifs that are 13 bases in length with four to seven mutations. It might be argued that these mutations may be selected easily to evade the action of the M1GS ribozyme by avoiding specific recognition between the GS and the target. Nevertheless, this argument does not invalidate the examination, for targeting purposes, of other regions different from those that are highly conserved. In each case, the extent of sequence variation may not correlate with the relative ability to escape attack with M1GS ribozymes.

The complex behaviour of RNA virus quasispecies in response to selective factors makes such direct correlation difficult. Although there are no reports on virus resistance to M1GS in vivo or in cell culture, examples of fitness recovery of debilitated viruses either in model systems (Domingo et al., 1999) or in multidrug-resistance virus variants in human immunodeficiency virus patients treated with multiple reverse transcriptase inhibitors and protease inhibitors shows that even mutants thought to be improbable may arise (Cabana et al., 1999). The most likely conclusion is that the level of RNA variation that arises during HCV infection will require multiple targeting.

The picture that emerges is that of a spatially limited accessibility and those subtle displacements within the open region lead to important susceptibility changes to cleavage by the ribozyme. It also appears that up to 50 % of the positions within the open regions are variant positions among sequence isolates from different patients. Compiling the efforts carried out by independent groups working with different aspects of HCV RNA structure accessibility and variation would greatly improve the progress on therapeutics with ribozyme combinations.

ACKNOWLEDGEMENTS

We thank Dr Fenyong Liu from University of California, Berkeley for providing the M1 clone pFL117 and for helpful comments. Work was funded by the Ministerio de Ciencia y

Tecnología (SAF1999- 0108 and BIO00-0347) and the Ministerio de Sanidad y Consumo (FISS-01/1351).

REFERENCES

Cabana, M., Clotet, B. & Martinez, M. A. (1999). Emergence and genetic evolution of HIV-1 variants with mutations conferring resistance to multiple reverse transcriptase and protease inhibitors. *J Med Virol* 59, 480–490.

Cabot, B., Martell, M., Esteban, J. I., Sauleda, S., Otero, T., Esteban, R., Guardia, J. & Gó mez, J. (2000). Nucleotide and amino acid complexity of hepatitis C virus quasispecies in serum and liver. *J Virol* 74, 805–811.

Domingo, E., Escarm´ıs, C., Menendez-Arias, L. & Holland, J. J. (1999). *Origin and Evolution of Viruses*, pp. 141–161. Edited by

E. Domingo & R. H. J. Webster. London: Academic Press.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849–857.

Lima, W. F., Brown-Driver, V., Fox, M., Hanecak, R. & Bruice, T. W. (1997). Combinatorial screening and rational optimization for hybridization to folded hepatitis C virus RNA of oligonucleotides with biological antisense activity. *J Biol Chem* 272, 626–638.

Liu, F. & Altman, S. (1995). Inhibition of viral gene expression by the catalytic RNA subunit of RNase P from *Escherichia coli*. *Genes Dev* 9, 471–480.

Lyons, A. J., Lytle, J. R., Gó mez, J. & Robertson, H. D. (2001). Hepatitis C virus internal ribosome entry site RNA contains a tertiary structural element in a functional domain of stem–loop II. *Nucleic Acids Res* 29, 2535–2541.

Macejack, D. G., Jensen, K. L. & Jamison, S. F. & 9 other authors (2000). Inhibition of hepatitis C virus (HCV)-RNA-dependent translation and replication of a chimeric HCV poliovirus using synthetic stabilized ribozymes. *Hepatology* 31, 769–776.

Martell, M., Esteban, J. I., Quer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J. & Gó mez, J. (1992). Hepatitis C virus (HCV) circu- lates as a population of different but closely related genomes: quasi- species nature of HCV genome distribution. *J Virol* 66, 3225–3229.

Nadal, A., Martell, M., Lytle, J. R. & 7 other authors (2002). Specific cleavage of hepatitis C virus RNA genome by human RNase P. *J Biol Chem* 277, 30606–30613.

Trang, P., Kilani, A., Kim, J. & Liu, F. (2000a). A ribozyme derived from the catalytic subunit of RNase P from *Escherichia coli* is highly effective in inhibiting replication of herpes simplex virus 1. *J Mol Biol* 301, 817–826.

Trang, P., Lee, M., Nepomuceno, E., Kim, J., Zhu, H. & Liu, F. (2000b). Effective inhibition of human cytomegalovirus gene expression and replication by a ribozyme derived from the catalytic

RNA subunit of RNase P from *Escherichia coli*. *Proc Natl Acad Sci U S A* 97, 5812–5817.

TABLES

TABLE 1

Table 1. Summary of the gel-shift assay and the RNase H cleavage assay of the whole set of oligonucleotides tested

	Gel shift		RNase H treatment	
DNA	Not preheated	Preheated	Not preheated	Preheated
C2658	+	++	+	++
C2678	++	++	++	++
C2700	–/+	+	++	++
C2724	++	++++	++	+++
C2745	+++	+++++	+++	+++++
C2767	+	++	+	++
C2786	–/+	–/+	–/+	–/+
C2810	–/+	–/+	+	+
C2829	+	+	+	+
C2848	–	–	–	–

Results have been quantified by visual estimation and are expressed in percentage of shifted DNA or RNA cleavage by RNase H: –, 0 %; –/+, <10 %; +, 10–30 %; ++, 30–50 %; +++, 50–70 %; ++++, 70–90 %; +++++, >90 %.

FIGURES

FIGURE 1

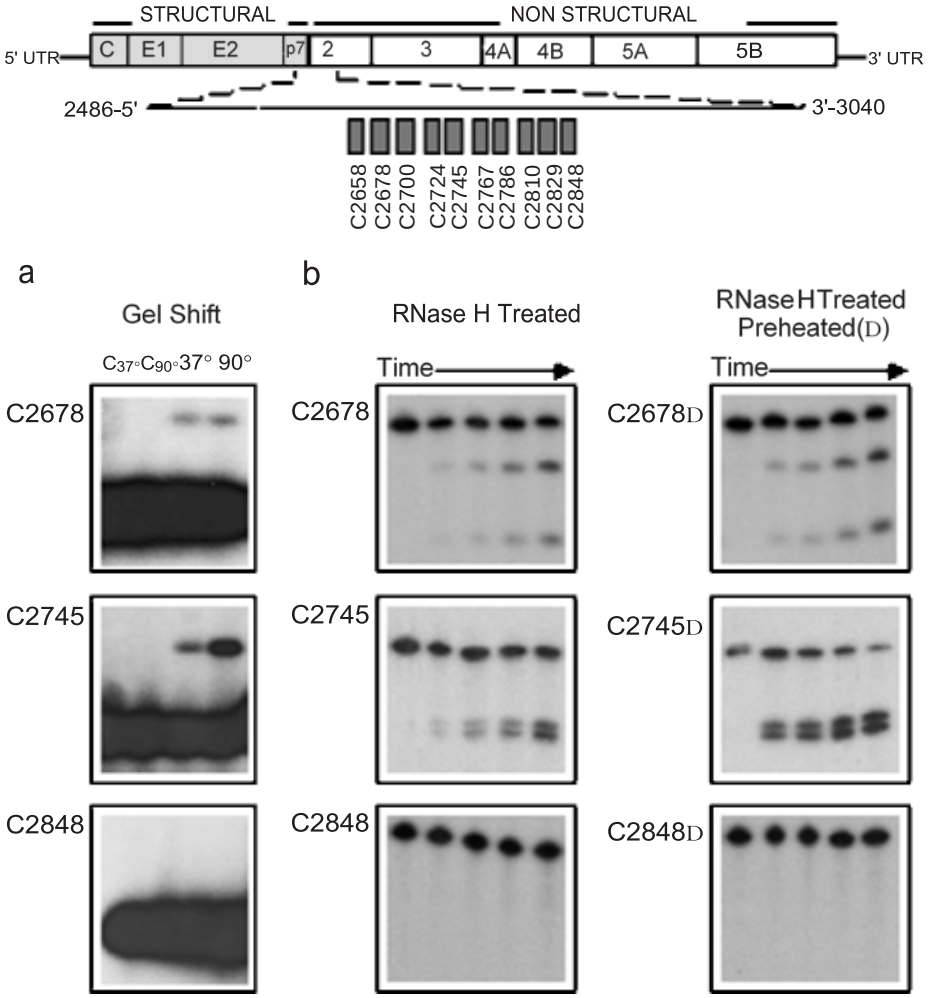

Figure 1. Accessibility of HCV SI RNA transcript to comple- mentary oligonucleotides. Schematic representation of the HCV genome and location of the HCV RNA SI transcript. Gel-shift assays (a) and kinetics analysis of RNase H cleavage (b) for three representative oligonucleotides of the 10 tested are shown. (a) Control reactions (lanes C37◦ and C90◦) and indivi- dual annealing of unlabelled SI transcript with labelled DNAs (lanes 37◦ and 90◦), either preheated or not, are shown. (b) In the left column, the results of mixing RNA substrate, DNA and RNase H are shown. In each individual reaction profile, the left- most lane represents time 0, while each lane to the right repre- sents, sequentially, 5, 15, 30 and 60 min of incubation. The right column represents reactions in which substrate and DNA were preheated to 90 ◦C before annealing.

FIGURE 2

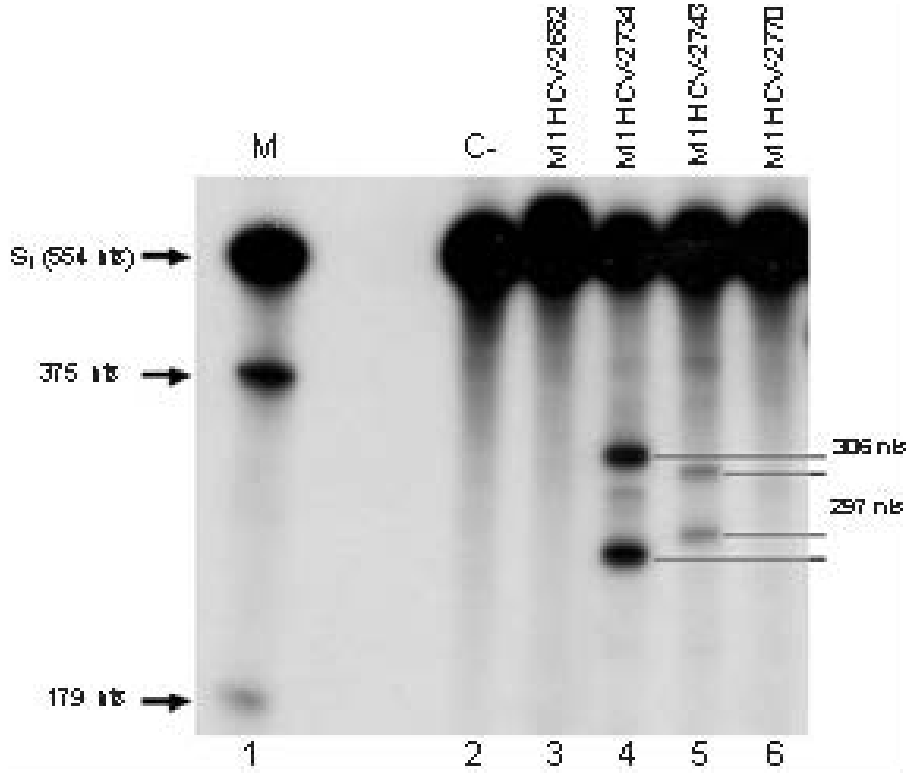

Figure 2. Cleavage of SI transcript RNA by M1GS ribozymes. Lane 1, size markers: incubation of RNA SI, which is 554 bases long, with the human RNase P enzyme produces a band 375 bases long and a band 179 bases long. SI RNA and pro- duct RNA sequences were determined directly by RNA finger- printing (Nadal et al., 2002). Lane 2, no ribozyme was added to the reaction. Lanes 3–6, incubation of each M1GS with SI transcript. The main band at the top of the gel corresponds to the substrate; the other bands are the two specific cleavage products generated by each M1GS ribozyme. The estimated size of cleavage products deduced from the relative mobility in the polyacrylamide/urea gel from two independent experiments does not deviate more than 2 % of the expected length.