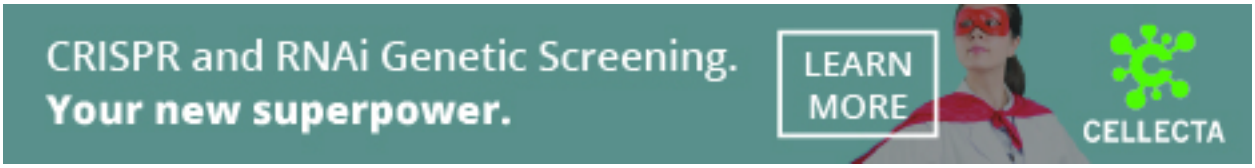
Characterization of the extreme 5' ends of RNA molecules by RNA ligation-PCR.

R Sallie

Genome Res. 1993 3: 54-56



Characterization of the Extreme 5′ Ends of RNA Molecules by RNA Ligation-PCR

Richard Sallie[1]

The Liver Diseases Section, NIDDK, The National Institutes of Health, Bethesda, Maryland 20892

Despite major recent advances in PCR-based methodology, characterization of the extreme terminal sequences of RNA molecules may be difficult and time consuming, particularly when the 3′ end is not polyadenylated.[1] The usual approach to defining the 5′ end involves cDNA synthesis using gene-specific downstream priming of the reverse transcription (RT) reaction, followed by the addition of a homopolymer tail by terminal transferase and single-sided PCR using a homopolymeric dN-based oligonucleotide to prime the 5′ PCR reaction (the RACE protocols). Verification of the extreme 5′ end of the RNA molecule requires two separate reactions in which different homopolymer tails are added by terminal deoxytransferase.[2] However, premature termination of the cDNA reaction cannot be excluded using this technique and would result in undetected and foreshortened cDNA clones, even if dual tailing reactions are used. Additionally, the presence or absence of a 5′-terminal structure, such as a methylated cap or a covalently bound VpG protein, cannot be deduced by homopolymer tailing methodology. Both 5′ hairpin loops[3] and terminal cap structures[4] are known to exist at the 5′ end of many RNA viruses. Recently, head-to-tail ligation of viral molecules with PCR amplification across the 3′ to 5′ junction has been advocated as a method of defining the terminal sequences and/or structure of RNA molecules.[5,6] However, this method requires knowledge of the complete sequences of one end or the other before the junctional sequence can be defined fully. Furthermore, variable-length homopolymer tails on the 3′ end of many RNA molecules prevent application of direct PCR-based sequence strategies to the resulting PCR products.[6]

The 3′ of nonpolyadenylated RNA molecules or molecules of unknown sequence is more problematic. Using cobalt-based buffers, RNA is a suitable substrate for the addition of a homopolymeric tail by terminal transferase. It is conceivable that a strategy based on tailing followed by 3′ single-sided PCR using a 5′ gene-specific primer and an oligo(dN)-based 3′ primer would work. However, definition of the extreme 3′ nucleotide would require two separate tailing experiments.

MATERIALS AND METHODS

A strategy of heminested PCR following end-to-end ligation of a synthetic RNA oligonucleotide was devised and used to characterize the extreme 5′-terminal sequences of both the positive and negative strands of hepatitis C virus (HCV). The overall design of the ligation amplification experiments is described in Figure 1.

RNA Purification and Decapping

HCV RNA was purified from 1 ml of serum obtained from patients known to have high-titer HCV using a proprietary modification of the method of Chomczynski,[7] RNAzol B, according to the manufacturer's instructions. The RNA pellet so obtained was washed three times in 70% alcohol, resuspended in SEPC-treated water, and divided into aliquots. On the assumption that a 5′-methylated cap existed, aliquots were treated with tobacco acid pyrophosphatase (TAP) by digestion in 20 μl of 1× TAP buffer containing 2 mM ATP, 10 units of TAP (Epicentre Technologies, Madison, WI), and 40 units of cloned RNasin (Promega) at 37°C for 1 hr; the decapped RNA was then purified by repeated phenol–chloroform extraction and ethanol precipitation.

End-to-end ligation

The HCV RNA was ligated end to end with a synthetic RNA molecule, RSHEVRNA, 5′-HO-UCUACUUCAACUUCAAGCCGAAUUC-OH-3′ (National Biosciences, Plymouth, MN), designed to be homologous to a hepatitis E virus PCR primer[8] and to contain an *Eco*RI site at the 3′ end for cloning purposes. HCV RNA and RSHEVRNA were ligated head to tail with T4 RNA ligase (New England Biolabs), according to a modification of the methods of Tessier et al.[9] Ligation was carried out in 25 μl of a mixture containing 25% polyethylene glycol, 100 μM ATP, 10 units of T4 RNA ligase (New England Biolabs), 40 units of RNasin (Promega), 500 ng of Gp 32 gene protein (Boehringer Mannheim), and 2 mM hexamine cobalt chloride incubated in a thermal cycler (Biometra trioblock), cycling between 4°C (20 min) and 37°C (20 min) for 16 hr to promote mixing.

Following ligation, the RNA was again purified by phenol–chloroform extractions, ethanol precipitated, and washed,

[1]Permanent address: P.O. Box 677, Nedlands, Western Australia, 6009.

3:54–56©1993 by Cold Spring Harbor Laboratory ISSN 1054-9803/93 $5.00

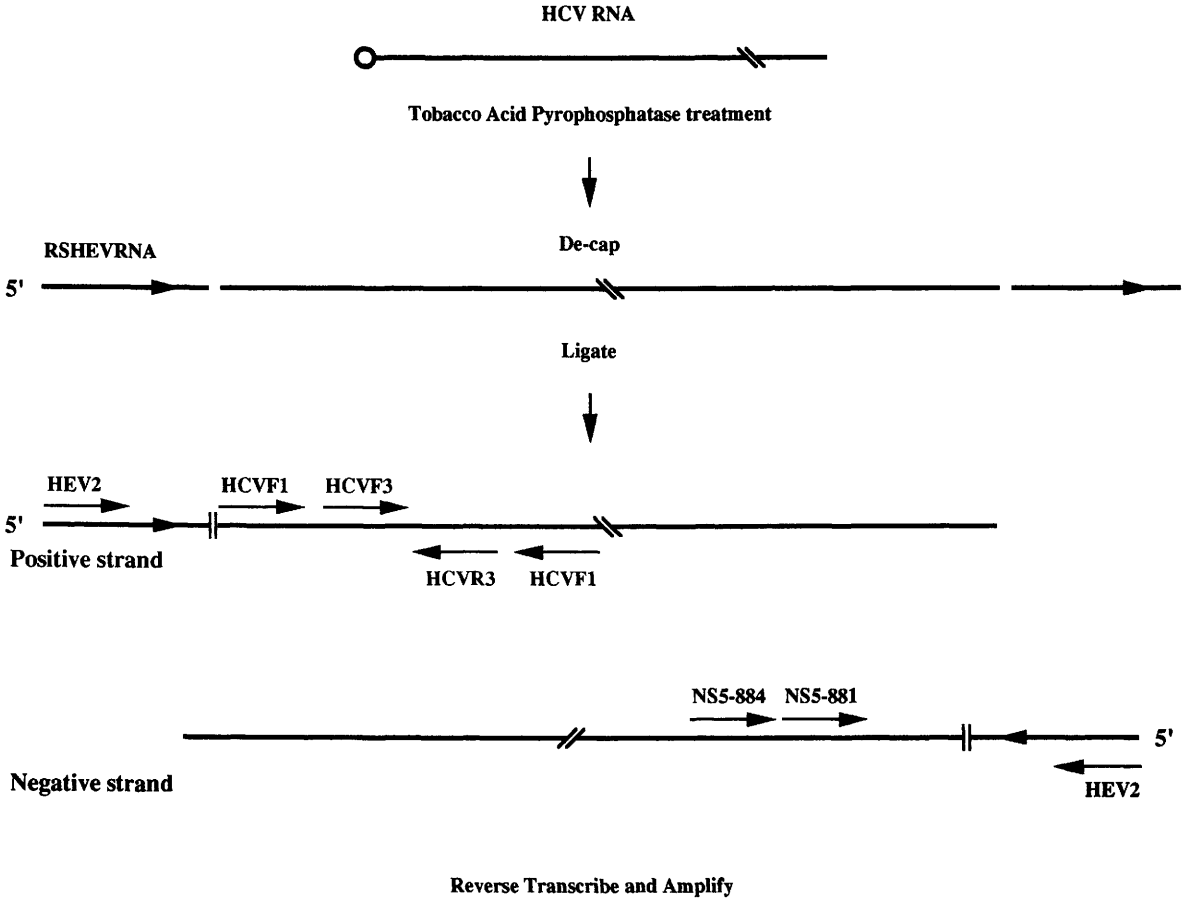


FIGURE 1 Overall design of end-to-end ligation experiment with relative positions of the primers used.

and resuspended in 20 μl of DEPC-treated water.

RT PCR

RT was performed using an AMV RT system (Riboclone, Promega) in a 20-μl volume reaction using 10 ng of either an antisense primer homologous to the 5′-untranslated region of the positive strand of HCV (HCVR1 for the 5′ end), or a sense (antisense with respect to the negative strand) primer complementary to the negative strand (NS5-884); 4 μl of the RNA solution was added as a substrate. Incubation was at 42°C for 30 min.

PCR was then performed using a hemi-nested strategy and hot-start methodology.[10] The DNA oligonucleotide HEV2 (5′-TCTACCTTCAACTTCAAGCC-3′ was used as the upstream primer, and either HCVR1 (HEV2 × HCVR1, outer amplicon) or HCVR3 (HEV2 × HCVR3, inner amplicon) was used as the downstream pair for the 5′ end and either NS5-884 (NS5-884 × HEV2, outer amplicon) or NS5-881 (NS5-881 × HEV2, inner amplicon) was used as the upstream primer and HEV2 as the downstream primer for the 5′ end of the negative strand. Four microliters of the RT reaction was used as substrate in a 100-μl first-round PCR reaction containing 100 ng of each of the primers, 2.5 units of *Thermus aquaticus* DNA polymerase (*Taq* polymerase, Biotech International, Bentley, Western Australia), and dNTPs at 200 μM in 1× PCR buffer. A 2-μl aliquot of the first-round reaction was the used as a substrate for the second-round reaction, which was identical to the first except for change or primers. Cycling parameters for both inner and outer reactions were (95°C for 1 min, 55°C for 1 min, and 72°C for 1 min) × 30 for the 5′ end of the positive strand while (95°C for 1 min, 55°C for 5 min, and 72°C for 1 min) × 30 was used for the negative strand.

RESULTS

Following TAP digestion, the junctional sequence of the synthetic RNA–HCV chimeric molecule could readily be amplified. However, multiple attempts at RNA–RNA ligation without prior incubation in TAP resulted in failure of amplification of chimeric molecules, suggesting the presence of a methylated 5′ cap on the positive strand (TAP should not remove covalently bound VpG proteins). Furthermore, amplification of the 5′ chimera resulted in single bands that were readily amenable to direct PCR-based sequencing (R. Sallie, S.-C. Chia, J.H. Hoofnagle, A.M. DiBiseglie, and S.M. Fernstone, in prep.). In contrast, amplification of the 5′ negative-strand chimera was possible without prior TAP digestion (Fig. 2). This resulted in multiple (four to eight) discrete bands, suggesting the presence of subgenomic transcripts. Again, following excision and purification from the gel, these bands were suitable substrates for direct sequencing procedures.

DISCUSSION

In this paper we describe the use of end-

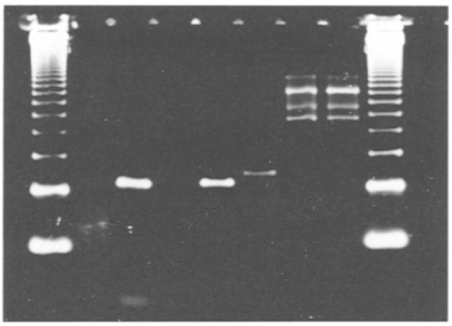


FIGURE 2 Gel electrophoresis (2% Nusieve, 1% agarose composite gel). (Lane *1,9*) 123-bp molecular weight markers; (lane *2*) negative reagent control; (lanes *3,5*) standard HCV PCR (F3 × R3, M. W. ~250 bp; (lane *4*) 5′ chimeric PCR without prior TAP treatment; (lane *6*) chimeric PCR with prior TAP treatment (HEV2 × R3, M. W. ~308 bp); (lanes *7,8*) 3′-end chimera PCR with and without prior TAP treatment (NS5-881 × HEV2). Note multiple bands ranging from ~1200 to ~300 bp (*top* to *bottom*).

to-end ligation of synthetic RNA to RNAs of interest to define the extreme termini of RNA molecules. Although the specific use in this instance was for characterization of terminal sequences of hepatitis C, the methodology is applicable to the extreme 5′ end of any RNA. It is particularly useful for the characterization of the 5′ sequences of both genomic and antigenomic viral RNAs. Determination of the 3′ sequences of the positive strand, not demonstrated here, would theoretically be possible if the RNA oligonucleotide was synthesized with 5′-terminal phosphate rather than hydroxyl groups. However, as the synthetic RNA is in vast molar excess, synthesis of the RNA with phosphate groups on both ends (or blocking the 3′ reactive groups by some other means) would be essential to prevent formation of synthetic RNA concatamers as the preferred reaction. Although determination of the presence or absence of 5′ caps can be deduced from the ability to amplify the junctional sequence with and without prior TAP digestion, this represents a negative experiment, and it is entirely possible that uncapped RNA is not head-to-tail ligated (and therefore not detected by this method) due to the low efficiency of the single-stranded ligation reaction.

REFERENCES

1. Frohman, M., M. Dush, and G. Martin. 1988. Rapid production of full length cDNAs from rare transcripts: Amplification using gene specific oligonucleotide primer. *Proc. Natl. Acad. Sci.* **85:** 8998–9002.
2. Okamoto, H., S. Okada, Y. Sugiyama, S. Yotsumoto, T. Tanaka, H. Yoshizawa, F. Tsuda, Y. Miyakana, and M. Mayumi. 1990. The 5′-terminal sequence of the hepatitis C virus genome. *Jpn. J. Exp. Med.* **60:** 167–177.
3. Han, J.H., V. Shyamala, K.H. Richman, M.J. Brauer, B. Irvine, M.S. Urdea, O.P. Tekamp, G. Kuo, Q.L. Choo, and M. Houghton. 1991. Characterization of the terminal regions of hepatitis C viral RNA: Identification of conserved sequences in the 5′ untranslated region and poly(A) tails at the 3′ end. *Proc. Natl. Acad. Sci.* **88:** 1711–1715.
4. Sagripanti, J.L., R.O. Zandomeni, and R. Weinmann. 1986. The cap structure of simian hemorrhagic fever virion RNA. *Virology* **151:** 146–150.
5. Brock, K.V., R. Deng, and S.M. Riblet. 1992. Nucleotide sequencing of 5′ and 3′ termini of bovine viral diarrhea virus by RNA ligation and PCR. *J. Virol. Methods.* **38:** 39–46.
6. Mandl, C.W., F.X. Heinz, S. Puchhammer, and C. Kunz. 1991. Sequencing the termini of capped viral RNA by 5′-3′ ligation and PCR. *BioTechniques* **10:** 75–76.
7. Chomczynski, P. and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162:** 156–159.
8. Sallie, R., A.E. Silva, H. Smith, C. Tibbs, B. Portmann, A.L.W.F. Eddelston, D. Bradley, and R. Williams. 1991. Detection of Hepatitis E but not "C" in sera of patients with fulminant NANB hepatitis. *Hepatology* (abstr. 81) **14:** 68A.
9. Tessier, D., R. Brousseau, and T. Vernet. 1986. Ligation of single stranded oligodeoxyribonucleosides by T4 RNA ligase. *Anal. Biochem.* **158:** 171–178.
10. Chou, Q., M. Russell, D. Birch, J. Raymond, and W. Bloch. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy number amplifications. *Nucleic Acids Res.* **20:** 1717–1723.

Received April 29, 1993; accepted in revised form May 4, 1993.