JOURNAL OF CLINICAL MICROBIOLOGY, Dec. 1996, p. 3085–3091 Vol. 34, No. 12
0095-1137/96/$04.00+0

Long PCR and Its Application to Hepatitis Viruses: Amplification of Hepatitis A, Hepatitis B, and Hepatitis C Virus Genomes

RAYMOND TELLIER, JENS BUKH, SUZANNE U. EMERSON, ROGER H. MILLER,† AND ROBERT H. PURCELL*

Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland

Received 10 April 1996/Returned for modification 11 July 1996/Accepted 4 September 1996

In this study we amplified virtually the entire genomes of hepatitis A virus (a member of the *Picornaviridae* family), hepatitis B virus (a member of the *Hepadnaviridae* family), and hepatitis C virus (a member of the *Flaviviridae* family) by using the recently described technique of long PCR. In order to do this, we first demonstrated, using the λ phage, that long PCR can be made highly sensitive and that the sensitivity can be further enhanced by nested long PCR. We also showed, using tobacco mosaic virus as a model, that a reverse transcriptase reaction can be linked to a long PCR, enabling the nearly full-length amplification of the genomes of RNA viruses. We then applied these techniques to serial dilutions of titrated stocks of well-characterized strains of hepatitis A, B, and C viruses. We amplified the nearly full-length sequence of each of these viruses from a small number of viral genomes, demonstrating the sensitivity of the process.

Recent studies (1, 5) have shown that the technique of PCR can be improved to permit the synthesis of amplicons of up to 40 kb. In principle, this would greatly benefit the field of virology by enabling the amplification of greater fractions of viral genomes and in many cases virtually the entire genome.

This technique could be especially useful for the study of viruses such as hepatitis C virus (HCV), a single-stranded RNA virus of positive polarity, which exhibits an extensive genetic heterogeneity (2). However, HCV is generally present in patients' sera in low titers and cannot replicate efficiently in cell culture. Consequently, for long PCR to be useful, it must be optimized to a high degree of sensitivity and efficiency. Furthermore, efficient reverse transcription (RT) of long fragments of RNA followed by long PCR (long RT-PCR) must be achieved in order to amplify the genomes of RNA viruses. In this study we showed that long PCR could be optimized to a high degree of sensitivity (for some templates it was within 1 order of magnitude of that of short PCR protocols), that the strategy of nested PCR was applicable and resulted in higher yields and sensitivity, and that long RT-PCR for the amplification of RNA viruses was feasible and highly sensitive. We applied these techniques to the nearly full-length amplification of the genomes of hepatitis A virus (HAV), hepatitis B virus (HBV), and HCV, using titrated stocks of well-characterized viral strains.

MATERIALS AND METHODS

Virus stocks. Purified DNA of the λ phage was purchased from Promega (Madison, Wis.). Purified RNA from the vulgare strain of tobacco mosaic virus (TMV) was purchased from Boehringer Mannheim (Indianapolis, Ind.). Strain HM175/24a (15), a cytopathogenic mutant of HAV that is well adapted to cell culture, was used in this study. The stock used for the extraction of viral RNA contained 4.4×10^7 radioimmunofocus units (RFU) per ml (29). We used a serum containing the adr subtype of HBV with an infectious titer of 10^8 50% chimpanzee infectivity doses (CID_{50}) per ml (28). Plasma containing strain H-77 (genotype 1a) of HCV (8, 21) was used in this study. It had an infectivity titer of $10^{6.5}$ CID_{50} per ml (8, 24a).

DNA purification. The HBV DNA was purified from serum by treatment with proteinase K and sodium dodecyl sulfate followed by phenol-chloroform extraction and ethanol precipitation, as previously described (14).

RNA purification. The HAV RNA was purified from cell lysate and HCV RNA was purified from plasma or serum with TRIzol (GIBCO BRL, Gaithersburg, Md.), according to the manufacturer's recommendations, except that all mixing steps were performed by repeated inversion or tapping of the tube to prevent RNA shearing. In each case the RNA pellet was resuspended in 10 μl of 10 mM dithiothreitol containing 5% (vol/vol) RNasin (20 to 40 U/μl) (Promega). Aliquots of RNA were stored at −80°C.

Serial dilutions of nucleic acid stocks. Serial 10-fold dilutions were prepared from the purified nucleic acids. The λ phage DNA and the HBV DNA were diluted in double-distilled H_2O, and the aliquots were stored at −20°C. The RNAs of TMV, HAV, and HCV were diluted in 10 mM dithiothreitol containing 5% (vol/vol) RNasin (20 to 40 U/μl) (Promega); aliquots of diluted RNA were stored at −80°C.

Primers. The primers were synthesized in a 392 DNA/RNA synthesizer (Applied Biosystems, Foster City, Calif.) for the λ phage, TMV, HAV, and HCV; they were supplied by an outside contractor for HBV.

Primers specific for λ phage were deduced from the complete λ DNA sequence (GenBank accession number J02459) or taken from reference 1. The sense primer λS17K (5′ CGATGTTATTGAAATCTGCGATGATGACTAT GC 3′) and the antisense primer MBL002 (1) (5′ GCAAGACTCTATGAGAA GCAGATAAGCGATAAG 3′) constituted one pair and bracketed a genome fragment of approximately 11 kb. The sense primer λS19K (5′ GTGGTCCGG CAGTACAATGGATTACCGTAAGA 3′) and the antisense primer λRS24K (5′ GTACAGTCATAGATGGTCGGTGGGAGGTGGTAC 3′) constituted another pair and bracketed a genome fragment of approximately 5 kb internal to the 11-kb fragment mentioned above.

Primers specific for TMV were deduced from the sequence of the vulgare strain of TMV (GenBank accession number J02415) The sense primer TMVS (5′ ACAATTACAATGGCATACACACAGACAGCTACC 3′) and the antisense primer TMVRS (5′ GATCCAAGACACAACCCTTCGATTTAAGTG 3′) bracketed a nearly full-length fragment (approximately 6.24 kb) of the genome.

Primers specific for HAV were deduced from the sequence of HAV strain HM175/24a (15) (GenBank accession number M59810) supplemented, for the termini of the genome, by the sequence of the wild-type strain HM175 (7) (GenBank accession number M14707). The sense primer HAVLA1 (5′ TGCAGGT TCAGGGTTCTTAAATCTGTTTCTC 3′) and the antisense primer HAVLA2 (5′ CAAATCATGAAAGGTCACAAATGAAACACTGG 3′) bracketed a nearly full-length fragment (approximately 7.2 kb) of the genome. The antisense primer HAVRS (5′ TATTTACTGATAAAAGAAATAAAC 3′) was located at the very end of the 3′ terminus of the genome, immediately before the poly(A) tail.

Primers specific for HBV were deduced from the consensus sequence of an alignment of the sequences of 33 strains of HBV (22). The sense primer (relative to the negative strand) HBVLA1 (5′ TTAACCTAATCTCCTCCCCCAACTC CTCCC 3′) and the antisense primer HBVLA2 (5′ ACCTCTGCCTAATCAT CTCATGTTCATGTC 3′) constituted one pair and bracketed a nearly full-length fragment (approximately 3.15 kb) of the genome. The sense primer HBVL1 (5′ CCAACTCCTCCCAGTCTTTAAACAAACAGTC 3′) and the an-

* Corresponding author.

† Present address: Targeted Interventions Branch, Division of AIDS, National Institutes of Health, Bethesda, Md.

tisense primer HBVL2 (5′ AATCATCTCATGTTCATGTCCTACTGTTCAAG 3′) bracketed a genome fragment of approximately 3.12 kb internal to the fragment of 3.15 kb mentioned above.

Primers specific for HCV were deduced from the sequence of HCV strain H (13 [GenBank accession number M67463], 21 [GenBank accession numbers M62381, M62383, and M62385]). Whenever possible we have used the sequence determined by Ogata et al. (21), who directly sequenced the H-77 strain of HCV extracted from the plasma stock used in this study. Several primers were used in different combinations, as detailed in Results. The sense primers were UNI 40 (primer a1 in reference 3) (5′ ACTGTCTTCACGCAGAAAGCGTCTAGC CAT 3′), UNI 49 (primer a3 in reference 3) (5′ ACGCAGAAAGCGTCTAG CCATGGCGTTAGT 3′), HUT 3364S (5′ AGGAGATACTGCTTGGGCCAG CCGACGGAATG 3′), HUT 4326S (5′ ATCGGCACTGTCCTTGACCAAGC AGAGACTGCG 3′), and HUT 5244S (5′ GTCACCCTGACGCACCCAATC ACCAAATACATC 3′). The two primers UNI 40 and UNI 49 are located in a highly conserved domain of the 5′ end of the 5′ noncoding region (5′ NC) of the HCV genome, with UNI 40 being the most external (3, 4). The antisense primers were HUTLA2 (5′ GGGCCGGGCATGAGACACGCTGTGATAAATGTC 3′), HUTLA4 (5′ AAATGTCTCCCCCGCTGTAGCCAGCCGTGAACC 3′), UNI 311R (primer a2 in reference 3) (5′ CGAGACCTCCCGGGGCACTCGC AAGCACCC 3′), and UNI 304R (primer a4 in reference 3) (5′ TCCCGGGG CACTCGCAAGCACCCTATCAGG 3′). The primers HUTLA2 and HUTLA4 are at the 3′ end of the sequence coding for the polyprotein, in a region conserved as shown by an alignment of multiple strains of HCV (2). The primers UNI 311R and UNI 304R are located in a highly conserved domain of the 3′ end of the 5′ NC of the HCV genome (3).

Long PCR with the KlenTaq LA-16 (KLA-16) polymerase mix. The reactions were performed in a total volume of 50 μl with final concentrations of 20 mM Tris-HCl (pH 8.55 at 25°C), 150 μg of bovine serum albumin per ml, 16 mM $(NH_4)_2SO_4$ and 3.5 mM $MgCl_2$ (PC2 buffer [1] and 250 μM each deoxynucleoside triphosphate (dNTP). Each reaction contained 20 pmol of each primer and 0.4 μl of the KlenTaq LA-16 (KLA-16) enzyme mix (1). This enzyme mix was prepared in a proportion of 15 μl of KlenTaq1 (25 U/μl) (Ab Peptides, St. Louis, Mo.) to 1 μl of cloned *Pfu* (2.5 U/μl) (Stratagene, La Jolla, Calif.). Reactions were performed in thin-walled PCR tubes (Stratagene). Each reaction mixture was overlaid with exactly 40 μl of mineral oil (Sigma, St. Louis, Mo.). The PCR was performed in a Robocycler thermal cycler (Stratagene) with the following cycling parameters: denaturation at 99°C for 35 s, annealing at 67°C for 30 s, and elongation at 68°C for durations that depended on the application (see Results). Different numbers of cycles were studied (see Results). Stringent precautions were observed to prevent cross-contamination, as previously described (3).

Long PCR with the Advantage KlenTaq polymerase mix. The reactions were performed essentially as described above, except that the buffer supplied with the kit was used instead of the PC2 buffer, and 1 μl of the Advantage KlenTaq polymerase mix (Clontech, Palo Alto, Calif.) (a mixture of KlenTaq1, Deep Vent, and TaqStart anti-*Taq* antibody) per reaction was used instead of the KLA-16 mix.

Nested PCR. When a second round of PCR was performed, 5 μl (10%) of the first-round PCR mixture was added to 45 μl of the second-round PCR mixture, prepared as described above except for the amount of buffer, which took into account the contribution from the first round. The contributions of dNTPs and primers were considered negligible. The pair of primers used in the second round of long PCR was always internal to the pair used in the first round.

RT. An RNA aliquot was thawed on ice, incubated for 2 min at 65°C, and chilled on ice. To the RNA was added, from a master mix, 0.5 μl of RNasin (20 to 40 U/μl) (Promega), 1 μl of 100 mM dithiothreitol (Promega), 1 μl of a 10 mM stock solution of dNTPs (Pharmacia, Piscataway, N.J.), 2.5 μl of a 10 μM primer stock solution, 1 μl (200 U) of Superscript II reverse transcriptase (GIBCO BRL), and 4 μl of 5× 1st Strand Synthesis Buffer (GIBCO BRL). In initial experiments the reaction mixture was incubated at 50°C for 1 h. In subsequent experiments the reaction mixture was incubated at 42°C for 1 h, 1 μl of RNase H (1 to 4 U/μl) (GIBCO BRL) and 1 μl of RNase T_1 (900 to 3,000 U/μl) (GIBCO BRL) were added, and the reaction mixture was incubated at 37°C for 20 min.

Purification of the cDNA after RT. Different methods of purification of the cDNA prior to the PCR were studied.

(i) Sodium acetate precipitation. One-tenth volume of 3 M sodium acetate (pH 5.2) and 2 volumes of isopropanol were added to the RT mix; mixing steps were performed by repeated inversion or tapping of the tube to prevent cDNA shearing. The solution was left to precipitate overnight at −20°C. The pellet was recovered by centrifugation, washed with 70% ethanol, resuspended in 10 μl of double-distilled H_2O, and transferred to the long PCR reaction mix.

(ii) DNA Clean Sweep kit. A DNA Clean Sweep Kit (Amresco, Solon, Ohio) was used essentially according to the manufacturer's recommendations except that 20 μg of glycogen (Boehringer Mannheim) was used as a carrier instead of the supplied yeast tRNA, and mixing was performed as described above. The cDNA pellet was resuspended in 10 μl of double-distilled H_2O, incubated at 95°C for 3 min, and added to the long PCR reaction mix.

(iii) Transfer method (19). In the transfer method procedure 2 μl (approximately 10%) of the unpurified RT mixture was transferred directly to the long PCR reaction mix.

Nested PCR of the 5′ NC of HCV. We used *Taq* polymerase and the procedure described by Bukh et al. (3), using the primer pairs UNI 40-UNI 311R (outer pair) and UNI 49-UNI 304R (inner pair). To determine the PCR titer of the H-77 stock, the RT was performed within the 5′ NC with the antisense primer UNI 311R, and transfer to the PCR mixture was as previously described (3). To optimize the RT with Superscript II, we used the HUTLA2 primer. After the RT, the cDNA was precipitated with sodium acetate essentially as described above but with the addition of glycogen (20 μg) as a carrier and then transferred into the PCR mixture (see Results).

Analysis of PCRs. After completion of the PCRs, 10 μl of the PCR mix was loaded on an agarose gel, and DNA bands were separated by electrophoresis. The bands were visualized by staining with ethidium bromide and UV transillumination. The specificity of selected amplicons was confirmed by direct sequencing or restriction enzyme digestion.

RESULTS

In preliminary experiments we synthesized amplicons of 28 and 35 kb from 1.0 ng of λ DNA by using exactly the procedure and primers described by Barnes (1). However, since the genomes of the various hepatitis viruses studied in our laboratory are all less than 10 kb, we were interested in optimizing the long PCR for templates of such length. We were especially interested in achieving a high degree of sensitivity, since our ultimate goal was to amplify the entire genomes of these viruses, many of which are available only in low titer. We set the basic parameters of the protocol as outlined in Materials and Methods and optimized the procedure.

We first attempted to increase the sensitivity of the long PCR procedure in amplifying the λ phage template. Using the primer pair λS17K-MBL002, the KLA-16 enzyme mix, and an elongation time of 13 min, we readily obtained the expected amplicon of 11 kb. The minimum amount of λ DNA template required to obtain this amplicon depended on the number of cycles in the PCR round, as follows: 10^{-1} ng with 19 cycles, 10^{-2} ng with 25 cycles, and 10^{-4} ng with 35 cycles (occasionally, a small amount of amplicon could even be synthesized from 10^{-5} ng after 35 cycles). For this reason the PCR rounds consisted of 35 cycles in all other experiments. For experiments with the KLA-16 enzyme mix we used exclusively mixes that would pass the quality test of synthesizing the 11-kb amplicon from 10^{-4} ng of λ phage DNA. Such a KLA-16 mix retained this sensitivity for at least 6 months. However, occasionally we had difficulty in preparing a mix that would perform at the required sensitivity.

Next, we used the primer pair λS19K-λRS24K to synthesize a 5-kb amplicon from the λ DNA. Using the same elongation time of 13 min for 35 cycles, we obtained this amplicon with as little as 10^{-4} ng of template, albeit with significant size heterogeneity. We tried shorter durations for the elongation step, which resulted in a lowered sensitivity. However, we reproducibly obtained the amplicon of 5 kb from 10^{-4} ng by using an elongation time of 6 min 30 s for the first 25 cycles and an elongation time of 13 min for the last 10 cycles.

It is well established that the use of nested PCR with *Taq* polymerase can increase the sensitivity, specificity, and yield. Therefore, we next performed a nested long PCR procedure with the λ DNA. We used the primer pair λS17K-MBL002 for the first round of PCR, transferred 5 μl from the reaction mix to a new PCR mix with the primer pair λS19K-λRS24K (internal to the first pair), and performed a second round of PCR. We obtained an amplicon from as little as 10^{-6} ng of λ DNA following nested PCR. Further, the yield from the higher dilutions of template was much improved.

Finally, we performed long PCR with the Advantage KlenTaq polymerase mix. With the primer pair λS17K-MBL002 and an elongation time of 13 min for 35 cycles, we reproducibly obtained amplicons of 11 kb from as little as 10^{-5} ng of λ DNA, a sensitivity 10-fold greater than that with the KLA-16

TABLE 1. Summary of the most sensitive results of long PCR with DNA viruses

DNA virus	Polymerase mix	Amplicon size (kb) (primer pair)	Duration of elongation step in PCR (no. of cycles)	Endpoint sensitivity (approx no. of genome copies)
λ phage	KLA-16[a]	11 (λS17K-MBL002)	13 min (35)	10^{-4} ng (2,000)
		5 (λS19K-λRS24K)	6 min 30 s (25), 13 min (10)	10^{-4} ng (2,000)
		Final product, 5 (first round, λS17K-MBL002; second round, λS19K-λRS24K)	First round, 13 min (35); second round, 6 min 30 s (25), 13 min (10)	10^{-6} ng (20)
	Advantage KlenTaq polymerase mix[a]	11 (λS17K-MBL002)	13 min (35)	10^{-5} ng (200)
		Final product, 5 (first round, λS17K-MBL002; second round, λS19K-λRS24K)	First round, 13 min (35); second round, 6 min 30 s (25), 13 (10)	10^{-6} ng (20)
HBV	KLA-16	3.15 (HBVLA1-HBVLA2)	6 min (25), 12 min (10)	10^2 CID_{50} (10^3)[b]
		Final product, 3.12 (first round, HBVLA1-HBVLA2; second round, HBVL1-HBVL2)	Each round, 6 min (25), 12 min (10)	10^1 CID_{50} (10^2)
	Advantage KlenTaq polymerase mix	3.15 (HBVLA1-HBVLA2)	6 min (25), 12 min (10)	10^2 CID_{50} (10^3)

[a] See Materials and Methods
[b] 1 $CID_{50} \approx 10$ genome copies.

mix. The use of this mix for a nested PCR, as described above, yielded a sensitivity of 10^{-6} ng, unchanged from that obtained previously. Overall, we could amplify large fragments from a very small number of genome copies of λ phage (Table 1).

On the basis of the results obtained with the λ phage template, we attempted to amplify the genome of HBV. We used a serial dilution of HBV DNA obtained from a serum sample with a known infectious titer (28). With the primer pair HBVL1-HBVL2 and the KLA-16 mix, we could reproducibly obtain an amplicon of the expected size (approximately 3.12 kb), with a sensitivity varying with the different elongation times we tested. Our best results were obtained with an elongation time of 6 min for the first 25 cycles and 12 min for the last 10 cycles, which reproducibly gave a sensitivity of 10^2 CID_{50}. Using the latter parameters, we also obtained an amplicon of the expected size (approximately 3.15 kb) with the primer pair HBVLA1-HBVLA2, with the same sensitivity of 10^2 CID_{50}. Since the primer pair HBVLA1-HBVLA2 is external, we could perform a nested PCR by transferring 5 μl of the PCR mixture into a new PCR mix with the primer pair HBVL1-HBVL2. As expected, we obtained a greater yield; the sensitivity was improved to 10^1 CID_{50} of HBV. To illustrate the general usefulness of such a procedure, we tested DNAs extracted from the sera of eight patients chronically infected with HBV. In all cases, we obtained the expected 3.15-kb amplicon after one round of amplification.

Using the Advantage KlenTaq polymerase mix, with the primer pair HBVLA1-HBVLA2 and with the same parameters as described above, we achieved the same sensitivity of 10^2 CID_{50} after one round of PCR. Thus, we could amplify nearly full-length amplicons from a small number of genome copies of HBV (Table 1).

It is noteworthy that with both λ and HBV templates and with both enzyme mixes, we observed minor amplicons of different molecular sizes in addition to the expected amplicon. We speculate that they constitute a PCR artifact, most probably arising from false priming events. Consistent with this hypothesis is the observation that these amplicons increased in diversity and quantity when the initial amount of template was increased or when the second round of a nested long PCR was performed.

With the exception of HBV, all of the known human hepatitis viruses are single-stranded RNA viruses; long PCR for these agents would therefore initially require an efficient RT step and a method to link the RT reaction and the PCR. As a model to study the feasibility of the long RT-PCR, we used TMV, an RNA virus whose genome is commercially available in great quantity and purity.

In a first series of experiments with TMV, the RT reactions were performed at 50°C for 1 h with the antisense primer TMVRS, after which the cDNA was purified by sodium acetate-isopropanol precipitation. Long PCR was then performed with the KLA-16 enzyme mix and the primer pair TMVS-TMVRS, with an elongation time of 13 min for 35 cycles. We obtained an amplicon of the expected size (6.2 kb) from as little as 10^{-4} ng (and occasionally 10^{-5} ng) of TMV RNA.

Since this result demonstrated a high degree of sensitivity of long RT-PCR, we attempted to amplify long segments of the HCV genome. We used serial dilutions of RNA extracted from a plasma stock of known infectious titer containing isolate H-77 (8, 21) of HCV; its titer in genome equivalents (GE), as measured by nested RT-PCR on the 5′ NC with the primer pairs UNI 40-UNI 311R and UNI 49-UNI 304R, was found to be 10^8 GE/ml.

Using the primer pairs HUT5244S-HUTLA2, HUT4326S-HUTLA2, and HUT3364S-HUTLA2 and several elongation times, we obtained the expected amplicons of 4, 5, and 6 kb, respectively, from 10^5 GE of HCV. However, the quantity of PCR product was low, and artifactual amplicons were present in comparable amounts. Further, these results proved to be difficult to reproduce, and we could not amplify larger fragments of the HCV genome.

A low yield of nearly full-length cDNA from the RT would decrease the sensitivity. Therefore, we studied how to improve the yield of long cDNA by performing the RT, with various modifications, from the 3′ end of the genome (with the HUTLA2 primer) on a serial dilution of HCV RNA. This was followed by sodium acetate precipitation and nested PCR on the 5′ NC, which would yield a signal only from virtually full-length cDNA. The procedure used as described above for the RT yielded a signal with 10^3 GE of HCV. We improved this result to 10^2 GE by incubating the RT reaction mixture at 42°C instead of 50°C. However, additional improvements were not obtained by denaturing the RNA longer or at a higher temperature prior to cDNA synthesis or by performing the cDNA synthesis with a higher concentration of Superscript II or with

TABLE 2. Summary of the most sensitive results obtained with long RT-PCR with RNA viruses[a]

RNA virus	Amplicon size (kb) (primer pair)	Duration of elongation step in PCR (no. of cycles)	Endpoint sensitivity (approx no. of genome copies)
TMV	6.2 (TMVS-TMVRS)	9 min 45 s (15), 11 min (10), 13 min (10)	10^{-5} ng (3,000)
HCV	4 (HUT5244S-HUTLA2)	9 min (35)	10^2 GE (10^3)[b]
	6 (HUT3364S-HUTLA2)	9 min 45 s (15), 11 min (10), 13 min (10)	10^3 GE (10^4)
	9.25 (UNI 40-HUTLA2)	9 min 45 s (15), 11 min (10), 13 min (10)	10^4 GE (10^5)
	Final product, 9.22 (first round, UNI 40-HUTLA2; second round, UNI 49-HUTLA4)	Each round, 9 min 45 s (15), 11 min (10), 13 min (10)	10^4 GE (10^5)[c]
HAV	7.2 (HAVLA1-HAVLA2)	9 min 45 s (15), 11 min (10), 13 min (10)	4.4×10^3 RFU (4.4×10^4–4.4×10^6)[d]

[a] All of these experiments were performed with prior RNA denaturation at 65°C for 2 min, RT at 42°C for 1 h, treatment with RNase H and RNase T_1, and transfer of 2 μl of the unpurified RT mixture to a long PCR mixture with the Advantage KlenTaq polymerase mix.
[b] GE ≈ 10 genome copies.
[c] The yield of the nested PCR is significantly increased.
[d] 1 RFU ≈ 10 to 1,000 genome copies.

a longer incubation time. In subsequent experiments cDNA synthesis was performed at 42°C for 1 h.

Next, we speculated that residual RNA from the RT mixture could interfere with the long PCR by binding to the template. To address this potential problem, we treated the mixture from the completed RT reaction with RNase H and RNase T_1 in all subsequent experiments. Last, we hypothesized that false priming events during RT could generate short cDNAs that would be amplifiable, since they incorporated the antisense primer. We therefore removed the unused primers by purifying the cDNA with the DNA Clean Sweep kit, and we performed the PCR with an antisense primer that was internal to the primer used in the RT. With these two modifications, upon reverse transcribing HCV RNA with HUTLA2 as the primer and performing nested PCR on the 5′ NC, we still obtained a signal with 10^2 GE. After employing this modified RT procedure, we proceeded with long PCR, with HUTLA4 as the new antisense primer. Using the same cycling parameters as described above, we observed a definite improvement of the PCR: amplicons of the expected size (4 or 6 kb) were reproducibly synthesized in greater quantity as the major product. However, the sensitivity remained at 10^5 GE, and we still could not amplify longer fragments of HCV. We then repeated the experiments with the Advantage KlenTaq polymerase mix. Whereas the sensitivity remained 10^5 GE for the 4- and 6-kb amplicons, for the first time we could amplify a larger fragment of the HCV genome: by using the primers UNI 40 and HUTLA4 for PCR, we obtained the expected 9.23-kb amplicon, with a sensitivity of 10^6 GE.

The fact that we could not increase the sensitivity of the long PCR beyond 10^5 GE led us to question whether the cDNA purification method was efficient at low concentrations, and consequently we explored the simple procedure of merely transferring a small amount (2 μl) of unpurified RT mixture into the PCR mixture (19). This modification for TMV with the Advantage KlenTaq polymerase mix and an elongation time of 13 min for 35 cycles produced the expected amplicon of 6.2 kb, albeit with a lower sensitivity of 10^{-3} ng in the RT reaction. However, using a modified elongation time, we reproducibly obtained an amplicon with as little as 10^{-5} ng of TMV RNA in the RT reaction (Table 2). We used this revised procedure for RT-PCR on HCV with HUTLA2 as antisense primer for both the RT and PCR. We obtained the 4-kb amplicon with a sensitivity of 10^2 GE, the 6-kb amplicon with a sensitivity of 10^3 GE, and the 9.25-kb amplicon with a sensitivity of 10^4 GE (Fig. 1A; Table 2). When we used the latter reaction to set up a nested PCR with the primers UNI 49 and HUTLA4 (bracketing a 9.22-kb region), the sensitivity at the end of the second round was the same (10^4 GE), but there was a considerable increase in the yield of the amplicon (Fig. 1B; Table 2). Thus, by simply transferring approximately 10% of the RNase-treated RT mix to the long PCR mix, we significantly improved the sensitivity of the long RT-PCR for HCV. In order to confirm the above-described observations, we performed concurrent experiments with aliquots from a serial dilution of HCV RNA, using reaction components from a master mix. In a first series of experiments, the RT was performed concurrently at 42 and 50°C and was followed by RNase treatment and transfer of 2 μl into a long PCR mix with the Advantage Klentaq polymerase mix and the primer pair UNI 40-HUTLA2 (bracketing the 9.25-kb amplicon). To enhance the signal, nested PCR with the primer pair UNI 49-HUTLA4 (bracketing the 9.22-kb amplicon) was also performed. The sensitivity after performing RT at 42°C was 10^4 GE, as expected, but it was only 10^6 GE if RT was performed at 50°C. A second series of experiments was performed essentially as described above, except that RT was performed at

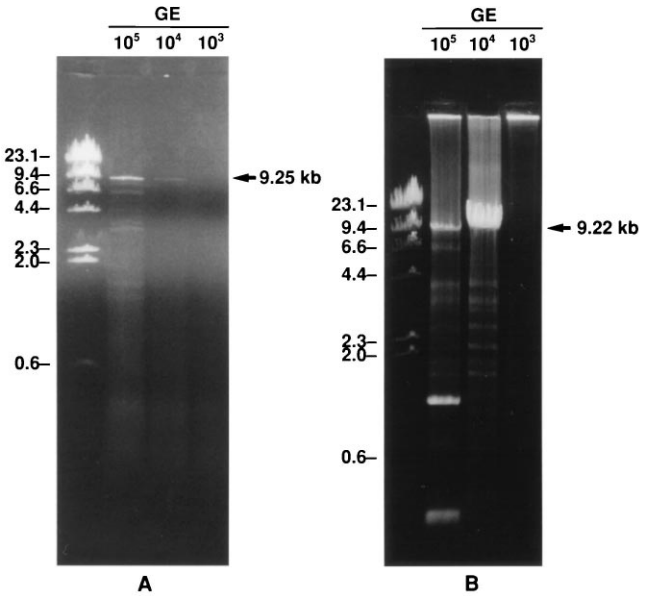

FIG. 1. Long RT-PCR amplification of HCV RNA. (A) The expected 9.25-kb amplicon was obtained from as little as 10^4 GE. (B) The 9.22-kb amplicons were obtained by nested PCR amplification of the PCR product from panel A. In each panel, the first lane contained molecular size markers (λ *Hin*dIII fragments).

42°C with and without RNase treatment. We observed that RNase treatment improved the sensitivity by 1 or 2 orders of magnitude.

We next demonstrated that HCV cDNA could be amplified from samples containing a relatively low viral titer. First, we extracted RNA from 10 μl of a serum sample containing 10^6 GE/ml from a chimpanzee infected with a genotype 5a isolate of HCV. We performed the RT with the primer HUTLA2 and the long PCR with the primers UNI 40 and HUTLA2, as described above, and we obtained the 9.25-kb amplicon as expected. Second, we extracted RNA from 100 μl of a chimpanzee serum containing 10^5 GE of a genotype 3a isolate of HCV per ml. After modifying the sequences of the primers UNI 40 and HUTLA2 to obtain a perfect match with the sequence of genotype 3a (4, 23) and performing the RT and long PCR as described above, we obtained an amplicon of the expected size.

As a final application, we used the long RT-PCR to amplify a nearly full-length fragment of the HAV genome. First, we performed the RT with the primer HAVRS, treated the mixture with RNases, and purified the cDNA with the DNA Clean Sweep kit; the cDNA was then amplified with the primer pair HAVLA1-HAVLA2, using the Advantage KlenTaq Polymerase mix. We obtained the expected band of 7.2 kb, starting with RNA equivalent to 4.4×10^4 RFU. Next, we performed the RT with the HAVLA2 primer, treated the mixture with RNases, and transferred 2 μl to a PCR mixture as described above. We obtained the 7.2-kb amplicon from the RNA equivalent of 4.4×10^3 RFU (Table 2).

The results summarized in Tables 1 and 2 were obtained with finalized protocols and represent our best results in terms of sensitivity and yield of the expected amplicon. The identities of representative amplicons were verified by direct sequencing. Figure 1 illustrates results of the long PCR for HCV and demonstrates the unambiguous synthesis of an amplicon of the expected size, as well as the increased yield conferred by nested PCR.

DISCUSSION

We studied the feasibility of applying the technique of long PCR to amplify nearly complete viral genomes. The procedure proved to be applicable to DNA viruses and, following RT, to RNA viruses. We limited our study to targets of 11 kb or less and endeavored to improve sensitivity and formulate rules of optimization.

We kept the procedure for the long PCR simple and in fact quite close to that originally described by Barnes (1). We were able to amplify diverse targets, spanning a broad range of GC content (49.9, 49.1, 43.3, 58.8, and 37.8% for λ phage, HBV, TMV, HCV, and HAV, respectively), at a high sensitivity by optimizing primarily the elongation time. We did not find it necessary to change the cycling temperatures, the durations of the denaturation or annealing steps, or the concentrations of Mg ion or other components of the mix. In our hands the Advantage KlenTaq polymerase mix was more sensitive and consistent than the KLA-16 mixes that we could prepare from enzymes purchased separately.

We amplified 5- or 11-kb fragments from as little as 200 genome copies of λ DNA in the long PCR. It is noteworthy that, at least for λ DNA in this size range, a greater amplicon length did not adversely affect the sensitivity. However the elongation time had to be carefully adjusted to the size of the desired amplicon: not only must it be long enough for the polymerases to synthesize the desired length of product, but elongation times that were too long were in fact deleterious. We have found, as have other authors (5), that a stepwise increase in the elongation time can be beneficial. It should be emphasized that a serial dilution from a well-quantified stock was essential for optimization, since suboptimal parameters may still enable amplification from a large amount of template. We found that the optimal parameters for PCR with KLA-16 generally worked well with the Advantage KlenTaq polymerase mix.

The feasibility of nested long PCR has not been addressed previously. We have shown that nested long PCR increased the yield and sensitivity, even with inner primers that had a significant overlap with the outer pair. The sensitivity of the nested long PCR for a 5-kb amplicon of λ DNA was 20 genome copies, which is within 1 order of magnitude of the sensitivity of the best PCR regimen. Increased sensitivity was also observed for nested long PCR of the HBV genome, with sensitivities of 10^2 CID_{50} for the first round of PCR (with 35 cycles) and 10 CID_{50} (≈100 genome copies) for the second round (with 35 cycles). This is comparable to the result obtained by Günther et al. (11), who amplified the whole genome of HBV from as little as 300 genome copies, using the Expand High Fidelity enzyme mix (Boehringer Mannheim) and a 40-cycle long PCR round. Remarkably, those authors demonstrated that the full-length amplicons they recovered were infectious. Our approach differed in that we chose primers that spanned a very highly conserved region, near the termini of the negative strand. Thus, while our amplicons are not expected to be infectious, we should be able to amplify almost the entire genome from most HBV-infected patients irrespective of the particular strain involved. Given the relatively high titer of HBV in most patients and the sensitivity of this assay, amplification from most clinical samples should be feasible. In that regard, the positive results obtained for each of the eight cases of chronic HBV infection that we studied are encouraging.

Our initial approach to long RT-PCR was to perform the RT in the buffer optimal for the reverse transcriptase, purify the cDNA, and perform the long PCR in the buffer optimal for the DNA polymerase. As our final results showed, this concern about buffer incompatibility was unfounded, at least with Superscript II and the Advantage KlenTaq polymerase mix. Recently this observation was also made with a different long PCR enzyme mix (eLONGase Enzyme mix; GIBCO BRL) (19). In fact, purification with the DNA Clean Sweep kit resulted in a loss of sensitivity (although it permitted the use of an inner antisense primer for the PCR and therefore in some circumstances may be advantageous), which probably is attributable to a low recovery rate at a low concentration of nucleic acids. The sensitivity of the long RT-PCR for TMV RNA was ≈3,000 genome copies (10^{-5} ng) after one round of PCR. This is 1 order of magnitude lower than the sensitivity of one round of PCR for λ DNA (200 genome copies). Since only 10% of the cDNA is transferred from the RT to the long PCR, this suggests that, at least for some templates, the sensitivity of the combined RT-PCR could be as high as that of PCR from DNA templates if one could efficiently recover all of the synthesized cDNA. So far, however, the other purification methods that we evaluated have not been found to improve on the transfer technique (unpublished data).

An important step in long RT-PCR was the optimization of the RT, which we assessed by reverse transcribing serially diluted HCV RNA with a primer (HUTLA2) almost at the 3′ end of the genome and performing standard nested PCR in the 5′ NC. (We note a recent example [17] of a similar strategy to study RT of the dengue virus genome.) Of the various modifications we studied, the most noticeable improvement came from performing the RT at 42°C rather than at 50°C. Even

then, comparison with the standard RT-PCR in the 5′ NC (3) suggested that at most 1% of the cDNAs synthesized from HUTLA2 are near full length. Nonetheless, it was enough to obtain a considerable improvement in generating long amplicons from HCV with KLA-16. Finally, the synthesis of near-full-length amplicons (9.25 kb) of HCV with the Advantage KlenTaq polymerase mix enabled us to show unequivocally an improvement of 2 orders of magnitude in sensitivity when the RT is performed at 42°C rather than at 50°C. This result is in good agreement with the reported optimal temperature for the Superscript II enzyme (9); it should be pointed out, however, that RT at 50°C may be advantageous for some templates with strong secondary structures. Finally, the treatment with RNase H and RNase T_1 did offer a more modest, but real, improvement; beneficial effects of RNase H treatment have been reported (19). One would expect the benefits of treatment with RNase to vary according to the sequence of the target as well as the initial amounts of template RNA and extraneous RNA.

It is noteworthy that the highest sensitivities of the long PCR procedure obtained in this study were achieved with commercially obtained templates (λ DNA and TMV RNA) with high purity which contained a high percentage of undamaged genomes. While the relatively lower sensitivities obtained with HAV, HBV, and HCV may reflect a lack of optimization, they may also reflect inferior quality of the templates.

Interestingly, we achieved useful sensitivities with long RT-PCR by using the same elongation parameters for three different viruses (Table 2), in the range of 6 to 9.3 kb: therefore, this regimen appears to be a good first approximation for amplifying many other RNA viruses in that size range (note that the three viruses span a broad range of GC content: 43.3% for TMV, 58.8% for HCV, and 37.8% for HAV).

For HAV and HCV, the RNA extraction method that we chose is one of many variations of the original method of Chomczynski and Sacchi (6). These methods have been recommended for use in clinical microbiology laboratories (20, 24). This method can easily be scaled up to extract RNA from 1 ml of serum or plasma. Thus, our long RT-PCR methodology would yield positive results with many clinical samples having a titer of 10^4 to 10^5 GE of HCV per ml. Of course, this sensitivity was achieved with a serial dilution of purified RNA; one could object that inhibitors of RT-PCR which might copurify with the RNA would also be diluted and that a more realistic test would be to extract RNA from lower-titer samples. However, when we extracted RNA from chimpanzee sera with lower viral titers and performed long RT-PCR, we obtained the 9.25-kb amplicon at the expected sensitivity, suggesting that inhibitors were not a problem. It should also be noted that the type of RNA purification method used here has been found to be highly effective in removing RT-PCR inhibitors from sewage samples (26).

For primer design the most obvious difference in the context of long PCR comes from the high annealing temperature, which requires primers with high T_ms but which increases the tolerance for a weak secondary structure. Experimental validation of the suitability of the primers with a serial dilution of the target is, in our opinion, essential. While we have not probed systematically the limits of tolerance for primer mismatch, we suspect that they are not very permissive, since the annealing temperature is quite stringent; in turn, it may be difficult to lower the annealing temperature, since this is likely to result in false priming and the synthesis of artifactual amplicons. This may constitute a fundamental limitation of the technique.

Our primary goal in this study was to achieve a high degree of sensitivity of long PCR and long RT-PCR for the amplification of viral genomes. We reasoned that restricting ourselves to the amplification of nearly full-length sequences, as opposed to full-length sequences, permitted more flexibility in primer optimization and avoided the possible loss of targets by truncation at the termini. Focusing on large subgenomic fragments also enabled us to exploit highly conserved regions; long PCR can be used to bridge large gaps between conserved regions, so that even if standard PCR methods may be preferable for the sensitive detection of pathogens at very low titers, long PCR should facilitate genetic studies of mutants, genotypes, or quasispecies. In that context the positive amplification from cell culture, as we demonstrated with HAV, should prove to be a valuable technique, since many viruses, such as the enteroviruses, can be routinely cultured in the clinical laboratory. There are situations in which further molecular characterization of such isolates is required (see, for example, references 16 and 18) and in which long PCR could prove to be very helpful. As additional examples of the use of long PCR for further molecular characterization, we note two recent studies using long PCR to analyze integrated retroviruses (12, 25) and one study using long PCR to amplify the genomes of human papillomaviruses (27). Clearly, it should be possible to obtain infectious cDNAs of viral genomes by using a single round of long RT-PCR, and in fact we recently succeeded in doing this for a mutant of HAV by using the conditions outlined in this paper (29). As noted above, Günther et al. (11) have demonstrated the infectivity of amplified full-length HBV genomes. Additionally, in a recent study Gritsun and Gould (10) generated, by long RT-PCR, two overlapping fragments of a tick-borne encephalitis virus from which a full-length infectious cDNA was generated by fusion PCR. These examples further illustrate the expected benefits of long PCR to the field of virology.

ACKNOWLEDGMENTS

We thank W. M. Barnes for helpful suggestions. We also thank Yu Mei Wen for providing HBV DNA, Tatiana S. Tsareva for the synthesis of the primers, and Carmean Hutton for help with the manuscript.

R.T. is the recipient of a Fellowship from the Medical Research Council of Canada.

REFERENCES

1. **Barnes, W. M.** 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. Proc. Natl. Acad. Sci. USA **91:**2216–2220.
2. **Bukh, J., R. H. Miller, and R. H. Purcell.** 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. Semin. Liver Dis. **15:**41–63.
3. **Bukh, J., R. H. Purcell, and R. H. Miller.** 1992. Importance of primer selection for the detection of hepatitis C virus RNA with the polymerase chain reaction assay. Proc. Natl. Acad. Sci. USA **89:**187–191.
4. **Bukh, J., R. H. Purcell, and R. H. Miller.** 1992. Sequence analysis of the 5′ noncoding region of hepatitis C virus. Proc. Natl. Acad. Sci. USA **89:**4942–4946.
5. **Cheng, S., C. Fockler, W. M. Barnes, and R. Higuchi.** 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. Proc. Natl. Acad. Sci. USA **91:**5695–5699.
6. **Chomczynski, P., and N. Sacchi.** 1987. Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
7. **Cohen, J. I., J. R. Ticehurst, R. H. Purcell, A. Buckler-White, and B. M. Baroudy.** 1987. Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses. J. Virol. **61:**50–59.
8. **Feinstone, S. M., H. J. Alter, H. P. Dienes, Y. Shimizu, H. Popper, D. Blackmore, D. Sly, W. T. London, and R. H. Purcell.** 1981. Non-A, non-B hepatitis in chimpanzees and marmosets. J. Infect. Dis. **144:**588–598.
9. **GIBCO BRL.** 1994. Technical bulletin 18064-1. GIBCO BRL, Gaithersburg, Md.
10. **Gritsun, T. S., and E. A. Gould.** 1995. Infectious transcripts of tick-borne encephalitis virus, generated in days by RT-PCR. Virology **214:**611–618.
11. **Günther, S., B.-C. Li, S. Miska, D. H. Krüger, H. Meisel, and H. Will.** 1995.

A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients. J. Virol. **69:**5437–5444.

12. **Herchenröder, O., R. Turek, D. Neumann-Haefelin, A. Rethwilm, and J. Schneider.** 1995. Infectious proviral clones of chimpanzee foamy virus (SFVcpz) generated by long PCR reveal close functional relatedness to human foamy virus. Virology **214:**685–689.
13. **Inchauspe, G., S. Zebedee, D.-H. Lee, M. Sugitani, M. Nasoff, and A. M. Prince.** 1991. Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. Proc. Natl. Acad. Sci. USA **88:**10292–10296.
14. **Kaneko, S., R. H. Miller, S. M. Feinstone, M. Unoura, K. Kobayashi, N. Hattori, and R. H. Purcell.** 1989. Detection of serum hepatitis B virus DNA in patients with chronic hepatitis using the polymerase chain reaction assay. Proc. Natl. Acad. Sci. USA **86:**312–316.
15. **Lemon, S. M., P. C. Murphy, P. A. Shields, L.-H. Ping, S. M. Feinstone, T. Cromeans, and R. W. Jansen.** 1991. Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination. J. Virol. **65:**2056–2065.
16. **Lin, K.-H., H.-L. Wang, M.-M. Sheu, W.-L. Huang, C.-W. Chen, C.-S. Yang, N. Takeda, N. Kato, K. Miyamura, and S. Yamazaki.** 1993. Molecular epidemiology of a variant of coxsackievirus A24 in Taiwan: two epidemics caused by phylogenetically distinct viruses from 1985 to 1989. J. Clin. Microbiol. **31:**1160–1166.
17. **Liu, H. S., H.-C. Tzeng, and C.-C. Chen.** 1995. Monitoring the cDNA synthesis of dengue-2 virus by RT-PCR. J. Virol. Methods **51:**55–60.
18. **Mulders, M. N., A. M. Van Loon, H. G. A. M. Van der Avoort, J. H. Reimerink, A. Ras, T. M. Bestebroer, M. A. Drebot, O. M. Kew, and M. P. Koopmans.** 1995. Molecular characterization of a wild poliovirus type 3 epidemic in The Netherlands (1992 and 1993). J. Clin. Microbiol. **33:**3252–3256.
19. **Nathan, M., L. M. Mertz, and D. K. Fox.** 1995. Optimizing long RT-PCR. Focus **17:**78–80.
20. **Nolte, F. S., C. Thurmond, and P. S. Mitchell.** 1994. Isolation of hepatitis C virus RNA from serum for reverse transcription PCR. J. Clin. Microbiol. **32:**519–520.
21. **Ogata, N., H. J. Alter, R. H. Miller, and R. H. Purcell.** 1991. Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. Proc. Natl. Acad. Sci. USA **88:**3392–3396.
22. **Ogata, N., R. H. Miller, K. G. Ishak, and R. H. Purcell.** 1993. The complete nucleotide sequence of a pre-core mutant of hepatitis B virus implicated in fulminant hepatitis and its biological characterization in chimpanzees. Virology **194:**263–276.
23. **Okamoto, H., H. Tokita, M. Sakamoto, M. Horikita, M. Kojima, H. Iizuka, and S. Mishiro.** 1993. Characterization of the genomic sequence of type V (or 3a) hepatitis C virus isolates and PCR primers for specific detection. J. Gen. Virol. **74:**2385–2390.
24. **Podzorski, R. P., and D. H. Persing.** 1995. Molecular detection and identification of microorganisms, p. 130–157. *In* P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. H. Yolken (ed.), Manual of clinical microbiology, 6th ed. ASM Press, Washington, D.C.

24a. **Purcell, R. H.** Unpublished data.

25. **Salminen, M. O., C. Koch, E. Sanders-Buell, P. K. Ehrenberg, N. L. Michael, J. K. Carr, D. S. Burke, and F. E. McCutchan.** 1995. Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction. Virology **213:**80–86.
26. **Shieh, Y.-S. C., D. Wait, L. Tai, and M. D. Sobsey.** 1995. Methods to remove inhibitors in sewage and other fecal wastes for enteroviruses detection by the polymerase chain reaction. J. Virol. Methods **54:**51–66.
27. **Stewart, A.-C. M., P. E. Gravitt, S. Cheng, and C. M. Wheeler.** 1995. Generation of entire papillomavirus genome by long PCR: frequency of errors produced during amplification. Genome Res. **5:**79–88.
28. **Tabor, E., R. H. Purcell, W. T. London, and R. J. Gerety.** 1983. Use of and interpretation of results using inocula of hepatitis B virus with known infectivity titers. J. Infect. Dis. **147:**531–534.
29. **Tellier, R., J. Bukh, S. U. Emerson, and R. H. Purcell.** 1996. Amplification of the full-length hepatitis A virus genome by long reverse transcription-PCR and transcription of infectious RNA directly from the amplicon. Proc. Natl. Acad. Sci. USA **93:**4370–4373.