Short Technical Reports

SHORT TECHNICAL REPORTS

Manuscripts published in the Short Technical Reports section are shorter and less comprehensive in scope than full Research Reports.

Use of Long RT-PCR to Characterize Splice-Variant mRNAs

BioTechniques 25:224-229 (August 1998)

ABSTRACT

Recent advances in long reverse transcription (RT)-PCR technology allow the copying of full-length coding regions of large mRNAs in one step. Using long RT-PCR, one can be certain that a given cDNA is derived from a single mRNA. In what to our knowledge is a novel application, we can isolate and characterize splice variants for any given mRNA in a systematic manner. We optimized long RT-PCR to copy the full-length coding region of human multidrug resistance (MDR1) mRNA or the major vault protein (MVP) mRNA in one step, so that only one full-length PCR product was synthesized in each case. Such stringent conditions are necessary to ensure that smaller than full-length products derived from total cell RNA are true splice variants. Twenty MDR1 double-stranded (ds)cDNAs, isolated from either the full-length or one prominent splice-variant DNA band, visualized on agarose gels, were cloned and sequenced. Two were full-length, wild-type in sequence as expected, and the rest were splice-variant mRNAs. Fourteen of the clones were identical and encoded a prominent splice-variant mRNA that can be detected in two tumor cell lines. This approach is shown to be generally applicable to the systematic analysis of splice-variant mRNAs derived from any gene.

INTRODUCTION

Traditionally, long reverse transcription polymerase chain reaction (RT-PCR) has not allowed the efficient synthesis of the full-length coding region of long mRNAs in one step. The consequence of this has been that to isolate a full-length cDNA, one must piece together smaller cDNAs that together reconstitute the entire mRNA. The danger in this approach is that the various cDNAs might derive from different but related mRNAs or from various splice-variant mRNAs. The reconstructed

cDNA would not be an accurate copy of the original mRNA. The cloning of splice-variant mRNAs is even more difficult. It is virtually impossible to prove that reconstituted cDNAs correspond to a novel splice-variant mRNA when the splice products cannot be predicted in advance. Because many splice products can be rare in abundance relative to the full-length mRNA, the analysis is more difficult.

Recently, RT-PCR technology has improved dramatically, and it is now possible to copy the full-length coding regions of even long mRNAs (6 kb) in one step in one tube (5,9). Much of the progress has been due to the commercial availability of optimized reverse-transcriptases that can reverse-transcribe RNA at the high temperatures necessary to disrupt secondary structure (12). The addition of a proofreading enzyme to the DNA polymerase mixture during PCR has also greatly increased the efficiency of this reaction by allowing the correction of mismatched nucleotides and subsequent polymerization through such regions of DNA (1). Combined, the availability of these improved reagents allows for effective long RT-PCR.

Our laboratory is interested in the roles that two particular mRNAs might have in conferring the drug-resistant phenotype to cancer cells: (i) The multidrug resistance (MDR1) mRNA, which encodes P-glycoprotein, an ABC transporter and a drug pump (2); and (ii) the major vault protein (MVP) mRNA, which encodes a protein product that is part of a ribonucleoprotein (RNP) complex that is frequently up-regulated in many cancer cells (7,8,14). We have developed RT-PCR technology that allows us to copy the full-length coding region of each of these mRNAs in one step in one tube. Most importantly, through the use of long RT-PCR technology, we can be absolutely sure that the splice-variant mRNAs we identify originate from one original mRNA. To our knowledge, this is a novel application of this technology in that we have isolated and characterized splice-variant mRNAs from these genes in a systematic manner that provides an overall picture of the splicing pattern. We have identified the predominant splice products as well as novel splice

forms of the *MDR1* mRNA not previously identified.

Other laboratories have also identified splice-variant mRNAs, but the methodology used was quite different from ours. Devine et al. screened 1.7×10^4 recombinant colonies from a cDNA library to identify splice-variant mRNAs of *MDR1* (4). In a recent report, opioid receptor splice-variant mRNAs were identified by RT-PCR, but only a small portion of each mRNA was examined (6). Other methodologies that might allow the synthesis of a full-length, double-stranded (ds)cDNA might result in a much higher background of nonspecific DNA amplification because of the use of a nonspecific 5' primer. We believe our methodology offers distinct advantages to previous approaches.

MATERIALS AND METHODS

Cell Lines and RNA Isolation

The vincristine-resistant cell line, HL60/VCR, was selected from the drug-sensitive parent cell line, HL60, by growth in increasing concentrations of the drug (10). The HL60 cell lines were developed by and obtained from Dr. Melvin Center (Kansas State University). The doxorubicin-resistant cell line, U937-30A-50P, was selected from the drug-sensitive parent cell line, U937, by growth in increasing concentrations of the drug (15). All these cell lines were maintained in RPMI 1640 Media with 10% fetal bovine serum (FBS) and 50 $\mu\text{g/mL}$ of gentamicin sulfate (all from Life Technologies, Gaithersburg, MD, USA). Cells were harvested in mid-log growth phase for total RNA isolation. Total cell RNA was isolated from 1.0×10^7 cells of each cell line using TRIzol[®] Reagent as described by the manufacturer (Life Technologies).

Primers for Long and Conventional RT-PCR

Our base numbering system for all cDNAs is based upon the GenBank[®] accession number, with base 1 being the 5' most base of the 5' untranslated region. For example, the ATG start co-

Short Technical Reports

don for *MDR1* begins at base 425. The PCR primer pairs we describe for long RT-PCR always include both the ATG start site and the 3' termination codon. Primers (MDR1)U423 [5'-GGATGGA-TCTTGAAGGGGACC] and (MDR1) L4254 [5'-AGTCAGAGTTCACCTGG-CGCTTTG] were used to amplify the entire coding region of *MDR1* mRNA (GenBank Accession No. M14758). Primers (MVP)U96 [5'-AGGAGTC-ACCATGGCAACTG and (MVP) L2190 [5'-CCAGTCAGAAGCCGA-GAA] were used to amplify the entire coding region of the human *MVP* mRNA (GenBank Accession No. X79-882). Primers (MDR1)U3037 [5'-AGGAGTTGTTGAAATGAAAATG-TTGT] and (MDR1)L3202 [5'-CTT-TCCTCAAAGAGTTTCTGTATGGT] were used to amplify a 191-bp fragment of the *MDR1* mRNA coding region located near its 3' end by conventional short RT-PCR. Finally, primers

(GAPDH)U900 [5'-GGTGGTCTCC-TCTGACTTCAACAG] and (GAP-DH)L1113 [5'-TCTCTTCCTCTTGT-GCTCTTGCTG] were used to amplify a 213-bp fragment of the mRNA, which encodes glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (GenBank Accession No. M33197) using short conventional RT-PCR. All primers were selected with a computer program, OLIGO™ Version 5.0 (National Biosciences, Plymouth, MN, USA).

Long RT-PCR

The Titan™ Long RT-PCR System (Boehringer Mannheim, Indianapolis, IN, USA) was used with modifications, detailed below. RT-PCRs were performed in a final volume of 50 μ L. Although the RT-PCR buffer contains dimethyl sulfoxide (DMSO), we have determined that the addition of a further 4% of DMSO improved product yields dramatically. A typical reaction uses two mixtures, with all the reagents added in the order listed. Mixture No. 1 contained a final concentration of 0.2 mM of each dNTP plus 5 μ Ci [α -³²P]dCTP (NEN Life Science Prod-

ucts, Boston, MA, USA), 5 mM dithiothreitol (DTT), 1 U RNase inhibitor, 500 ng of total cell RNA and 5.0 μ L of a 40% DMSO solution, all in a thin-walled PCR tube that is kept on ice (final volume of 28 μ L). Mixture No. 2 contained 10 μ L of 5 $\times$ RT-PCR buffer, 1 μ L of Titan enzyme mixture and two PCR primers selected to be used in either long RT-PCR or short conventional RT-PCR, each at a 0.4 pmol/ μ L final concentration.

Mixture No. 2 was brought to a final volume of 22 μ L with Nuclease-Free water (Promega, Madison, WI, USA). Mixture No. 1 was brought to 50°C, at which time, Mixture No. 2 was added (hot start). RT was performed at 50°C for 30 min. This was followed by a 1-min denaturation of DNA at 94°C and 40 cycles of PCR amplification using a 30-s 94°C denaturation step, a 30-s 55°C annealing step and a 5-min 68°C primer extension step. To demonstrate the absence of contaminating DNA, Mixtures No. 1 and No. 2 were mixed and kept on ice, and an RT reaction was omitted. A standard PCR was then run to detect any contaminating DNA. PCR products were resolved on 1% agarose

Figure 1. Long RT-PCR analysis of RNA and DNA controls and total RNAs isolated from four different human tumor cell lines. The long RT-PCR DNA product obtained with 10 ng of an in vitro-transcribed *MDR1* RNA transcript that encodes the full-length coding region of the mRNA (lane 1); a negative control in which the same long RT-PCR conditions were used in the absence of an RT step (lane 2); the PCR DNA product obtained following the amplification of 5 ng of a full-length human *MDR1* cDNA (ATCC 65705; ATCC, Rockville, MD, USA) (lane 3); a negative control in which the same PCRs were run in the absence of added DNA template (lane 4); long RT-PCR analysis of *MDR1* mRNA in HL60/parent cells (lane 5), HL60/VCR cells (lane 6), U937 cells (lane 7) and U937/10A cells (lane 8). Long RT-PCR analysis of the *MVP* mRNA from HL60/VCR cells (lane 9); a negative control as detailed in lane 2 but utilizing HL60/VCR total RNA (lane 10). Lane M, a 1-kb DNA size marker (Life Technologies).

Figure 2. Conventional short RT-PCR analysis of total RNAs isolated from four different human tumor cell lines. (A) Conventional RT-PCR analysis of *GAPDH* mRNA in HL60/parent cells (lane 1), HL60/VCR cells (lane 2), U937 cells (lane 3) and U937/10A cells (lane 4). (B) Conventional RT-PCR analysis of *MDR1* mRNA in the same cell lines in the same order as in Panel A. Negative controls in which the same conventional short RT-PCR conditions were used in the absence of an RT step with HL60/VCR RNA (lane 5, Panels A and B). Lane M, a 1-kb DNA size marker.

gels, and DNA bands were detected by staining with 0.5 µg/mL ethidium bromide using standard methods (13). Gels were dried under vacuum at 80°C for 2 h and exposed to X-Omat X-ray film (Scientific Imaging Systems [Eastman Kodak], New Haven, CT, USA) at -80°C overnight. A successful long RT-PCR is not dependent on the particular kit used. We have also successfully used a TaKaRa Long RT-PCR Kit (PanVera, Madison, WI, USA) to copy the mRNAs described in this report.

Isolation and Characterization of Full-Length and Splice-Variant *MDR1* mRNAs

Following long RT-PCR, *MDR1* DNA bands of interest were extracted from agarose gels using the QIAquick™ Gel Extraction Kit (Qiagen, Chatsworth, CA, USA) as described by the manufacturer. Double-stranded cDNAs were then subcloned into pCR® 2.1 (Invitrogen, Carlsbad, CA, USA) and sequenced using the PRISM™ Ready Reaction Dye Terminator Cycle Sequencing Kit and a Model 373A DNA Sequencer (both from PE Applied Biosystems, Foster City, CA, USA). The sequences were analyzed using the BESTFIT and MAP programs of the GCG Version 8.1 software package (Genetic Computer Group, Madison, WI, USA).

RESULTS AND DISCUSSION

We have developed a technique that allows us to detect both full-length and splice-variant mRNAs of *MDR1*, *MVP* or any other mRNA of interest. PCR primers are designed to include or reside 5' to the AUG start codon and to include or reside 3' to the stop codon of the mRNA of interest, and stringent RT-PCR conditions are developed. When a total cell RNA population is analyzed, both full-length and splice-variant mRNAs can be observed following gel electrophoresis of dscDNA products. A few splice-variant dscDNAs can be detected by ethidium bromide staining, while rarer products must be detected by ³³P autoradiography. cDNA clones must be sequenced multiple times to ensure that PCR-gen-

erated errors are detected and to define the prevalence of different splice-variant mRNAs.

To use long RT-PCR technology to isolate and characterize splice-variant mRNAs, it was necessary that the technique meet two clear technical criteria. First, one must be able to copy the full-length coding regions of mRNAs of interest in one step in one tube, thus ensuring that the splice products observed originate from one common mRNA. Secondly, and perhaps not as obvious, the reverse transcriptase and PCR conditions must be very stringent; the use of a pure, full-length RNA transcript control must result in the synthesis of only one full-length dscDNA product. If smaller, artifactual dscDNA products were produced as a result of PCR primer mispairing, the smaller DNAs would appear to be splice-variant products when in fact they were not. We have developed stringent long RT-PCR conditions, that when applied with a pure, full-length *MDR1* mRNA template control, results in the synthesis of only one full-length dscDNA product band (Figure 1, lane 1). Therefore, it is clear that our reverse transcriptase conditions allow us to fully denature the entire mRNA and to efficiently copy through all hairpin stems and loops. Smaller, artifactual DNA bands that might be confused as splice-variant mRNAs, are not produced even when autoradiograms are overexposed for days (data not shown). The use of a full-length *MDR1* cDNA template control also results in the synthesis of only one full-length dsDNA PCR product band (Figure 1, lane 3). This again demonstrates that PCR conditions are sufficiently stringent, that primer mispairing does not occur and DNA secondary structure does not present a technical problem. Therefore, we can be certain that the application of this technology to the analysis of a total cell RNA population that contains both full-length *MDR1* mRNA as well as smaller splice-variant mRNAs will accurately detect all forms of the mRNA, and that the detected splice-variant mRNAs will not be the result of a technical artifact.

Figure 1 shows an application of this technology. In HL60/VCR cells and in U937-10A-50P cells (Figure 1, lanes 6 and 8), both a full-length *MDR1* mRNA

Short Technical Reports

coding region (3854 bases) and a prominent splice-variant mRNA band (ca. 1873 bases) are observed by ethidium bromide staining and by ^{33}P autoradiography (^{33}P autoradiography data not shown). The two prominent PCR product bands present in HL60/VCR cells (Figure 1, lane 6, Nos. 1 and 2) were extracted from gels, and the DNAs were subcloned and sequenced. The full-length *MDR1* mRNA band contained only wild-type *MDR1* mRNA as expected (Figure 1, lane 6, band No. 1). The major splice product DNA band contained four different splice products (Figure 1, lane 6, band No. 2). Fourteen of the 18 clones analyzed coded for an identical predominant splice-variant mRNA, in which 51% of the coding region has been removed. The splice removes bases 764–2744, which includes exons 6–18. Conventional 5' and 3' intron splice sites are utilized (3). As far as we know, this is a novel splice prod-

uct not reported previously. Other splice-variant mRNAs included in this approximately 1873-bp band encoded rarer transcripts, all but one of which also spliced out exons 6–18. In addition, the rarer splice products were smaller and contained 1 additional splice. Many additional splice-variant mRNA bands can be observed on the gel and are in the process of being characterized. Nested PCRs verify that the majority of these mRNAs derive from the *MDR1* mRNA. Note that some splice-variant mRNAs are present in parent HL60 cells that are absent in HL60/VCR cells.

The failure to detect full-length *MDR1* mRNA in the two parent cell lines (HL60/Parent and U937) with long RT-PCR (Figure 1, lanes 5 and 7) is not due to partial or total degradation of the total RNA preparations. A control demonstrates that *GAPDH* mRNA is detectable at approximately equal levels in all cell lines, thus ensuring that

all RNAs are relatively undegraded (Figure 2A). A more sensitive conventional short product RT-PCR assay further demonstrates that the two parent cell lines negative for detection of full-length product do indeed contain a very low but detectable level of *MDR1* mRNA (Figure 2B). In HL60 and U937 parent cells, both a 191-bp *MDR1* band and a second, artifactual, higher molecular weight band are observed in the same gel lanes. We see artifactual PCR bands only when the *MDR1* mRNA concentration is very low, and the PCR substrate is limited. In samples with higher *MDR1* mRNA concentration, these artifactual bands are not observed. The point to be made is that HL60/parent and U937 cells do contain a very low level of *MDR1* mRNA.

That these two cell lines have very low levels of *MDR1* mRNA but undetectable full-length mRNA products, is explained by two factors. First, the effi-

ciency of conversion of a full-length, 3854-base or a splice-form, 1873-base mRNA to a dscDNA is much less efficient than that of a smaller 191-bp conventional RT-PCR product. Secondly, the entire *MDR1* mRNA sequence can be scanned to find optimal conventional RT-PCR primers; whereas, long RT-PCR primers are constrained to include or flank the start AUG and termination codons and, in general, are not as efficient in PCR. One should therefore be cautious in the interpretation of the inability to detect long RT-PCR products; low levels of a mRNA might still be present that are only detectable with a more sensitive conventional RT-PCR assay. In addition, the use of standard quantitative RT-PCR will determine which splice-variant mRNAs are present at levels that are likely to be biologically significant, e.g., at greater than one copy of mRNA per cell (11).

The methodology described here is generally applicable to the analysis of other mRNAs and is not dependent on the long RT-PCR kit used, although each mRNA may require some adjusting of the long RT-PCR conditions. We have used this technique to copy the full-length coding regions of the *MDR1*, *MVP* and prostate-specific antigen (PSA; data not shown) mRNAs. In all cases, splice-variant mRNAs are observed and will be characterized.

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. Chen, C.J., J.E. Chin, K. Ueda, D.P. Clark, I. Pastan, M.M. Gottesman and I.B. Roninson. 1986. Internal duplication and homology with bacterial transport proteins in the *mdr1* (P-glycoprotein) gene from multidrug-resistant human cells. *Cell* 47:381-389.
3. Chen, C.J., D. Clark, K. Ueda, I. Pastan, M.M. Gottesman and I.B. Roninson. 1990. Genomic organization of the human multidrug resistance (*MDR1*) gene and origin of P-glycoproteins. *J. Biol. Chem.* 265:506-514.
4. Devine, S.E., A. Hussain, J.P. David and P.W. Melera. 1991. Full length and alternatively spliced *pgp1* transcripts in multidrug-resistant chinese hamster lung cells. *J. Biol. Chem.* 266:4545-4555.
5. Frey, B. and B. Suppmann. 1995. Demonstration of the Expand™ PCR system's greater fidelity and higher yields with a lacI-based PCR fidelity assay. *Biochemica* 2:34-35.
6. Gaveriaux-Ruff, C., J. Peluso, K. Befort, F. Simmonin and C. Zilliox. 1997. Detection of opioid receptor mRNA by RT-PCR reveals alternative splicing for the j- and k-opioid receptors. *Mol. Brain Res.* 48:298-304.
7. Izquierdo, M.A., G.L. Scheffer, M.J. Flens, G. Giaccone, H.J. Broxterman, C.J. Meijer, P. Van Der Valk and R.J. Scheper. 1996. Broad distribution of the multidrug resistance-related vault lung resistance protein in normal human tissues and tumors. *Am. J. Pathol.* 148:877-887.
8. Kedersha, N.L. and L.H. Rome. 1986. Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA. *J. Cell Sci.* 103:699-709.
9. Kübler, G. and B. Frey. 1995. Amplification of fragments up to 7.7 kb from human dystrophin RNA by RT-PCR with AMV reverse transcriptase and the Expand™ long template PCR system. *Biochemica* 3:15-16.
10. McGrath, T. and M.S. Center. 1987. Adriamycin resistance in HL60 cells in the absence of detectable P-glycoprotein. *Biochem. Biophys. Res. Commun.* 145:1171-1176.
11. Mulder, J., N. McKinney, C. Christopherson, J. Sninsky, L. Greenfield and S. Kwok. 1994. Rapid and simple PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma: application to acute retroviral infection. *J. Clin. Microbiol.* 32:292-300.
12. Myers, T.W. and D.H. Gelfand. 1991. Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* 30:7661-7666.
13. Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed., CSH Laboratory Press, Cold Spring Harbor, NY.
14. Scheffer, G.L., P.L.J. Wijngaard, M.J. Flens, M.A. Izquierdo, M.L. Slovak, H.M. Pinedo, C.J.L.M. Meijer, H.C. Clevers and R.J. Scheper. 1995. The drug resistance-related protein LRP is the human major vault protein. *Nat. Med.* 1:578-582.
15. Slapak, C.A., N. Mizunuma and D.W. Kufe. 1994. Expression of the multidrug resistance associated protein and P-glycoprotein in doxorubicin-selected human myeloid leukemia cells. *Blood* 84:3113-3121.

Address correspondence to Dr. Robert E. Moore, Lilly Research Laboratories, Lilly Corporate Center, DC 0424, Eli Lilly and Company, Indianapolis, IN 46285-0424, USA. Internet: moore_robert_e@lilly.com

Received 29 December 1997; accepted 30 March 1998.

**Y. Hu, L.R. Tanzer, J. Cao,
C.D. Geringer and R.E.
Moore**
*Eli Lilly and Company
Indianapolis, IN, USA*