Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer

(polymerase chain reaction/5′ and 3′ cDNA ends/cDNA cloning/low-abundance mRNAs/*int-2* gene)

Michael A. Frohman, Michael K. Dush, and Gail R. Martin

Department of Anatomy, University of California, San Francisco, CA 94143

ABSTRACT We have devised a simple and efficient cDNA cloning strategy that overcomes many of the difficulties encountered in obtaining full-length cDNA clones of low-abundance mRNAs. In essence, cDNAs are generated by using the DNA polymerase chain reaction technique to amplify copies of the region between a single point in the transcript and the 3′ or 5′ end. The minimum information required for this amplification is a single short stretch of sequence within the mRNA to be cloned. Since the cDNAs can be produced in one day, examined by Southern blotting the next, and readily cloned, large numbers of full-length cDNA clones of rare transcripts can be rapidly produced. Moreover, separation of amplified cDNAs by gel electrophoresis allows precise selection by size prior to cloning and thus facilitates the isolation of cDNAs representing variant mRNAs, such as those produced by alternative splicing or by the use of alternative promoters. The efficacy of this method was demonstrated by isolating cDNA clones of mRNA from *int-2*, a mouse gene that expresses four different transcripts at low abundance, the longest of which is ≈2.9 kilobases. After <0.05% of the cDNAs produced had been screened, 29 independent *int-2* clones were isolated. Sequence analysis demonstrated that the 3′ and 5′ ends of all four *int-2* mRNAs were accurately represented by these clones.

Despite the development of numerous cDNA cloning strategies (1–5), obtaining full-length cDNA copies of low-abundance mRNAs remains a formidable task. We describe here a simple, rapid, and efficient cDNA cloning strategy that is based on the DNA polymerase chain reaction (PCR) technique developed by Saiki *et al.* (6). PCR employs two oligonucleotide primers, one complementary to a sequence on the (+) strand and the other to a downstream sequence on the (−) strand. Reiterative cycles of denaturation, annealing, and extension are used to generate multiple copies of the DNA that lies between the two primers. PCR has been used for a variety of purposes (7–13), including the detection of allelic polymorphisms and of DNA sequences unique to rare cells in a population, as well as the cross-species isolation of homologous genes. The feature common to all applications of PCR to date has been the use of primers designed to match two known or presumed genomic or cDNA sequences. In contrast, we have devised an application of this technique that achieves amplification and cloning of the region between a single short sequence in a cDNA molecule and its unknown 3′ or 5′ end. We demonstrate here the utility of this strategy, termed "rapid amplification of cDNA ends" (RACE), by using it to obtain clones representing transcripts of a gene, *int-2*, expressed at low abundance in the early mouse embryo.

METHODS

3′-End Amplification of cDNAs. See Fig. 1 for the scheme. *Reverse transcription.* One microgram of poly(A)$^+$ RNA (ref. 14; ref. 15, pp. 91–98) in 16.5 μl of water was heated at 65°C for 3 min, quenched on ice, added to 2 μl of 10× RTC buffer (1× RTC buffer is 50 mM Tris·HCl, pH 8.15 at 41°C/6 mM MgCl$_2$/40 mM KCl/1 mM dithiothreitol/each dNTP at 1.5 mM), 0.25 μl (10 units) of RNasin (Promega Biotec, Madison, WI), 0.5 μl of (dT)$_{17}$-adaptor (1 μg/μl), and 10 units of avian myeloblastosis virus reverse transcriptase (Life Sciences, Saint Petersburg, FL), and incubated for 2 hr at 41°C. The reaction mixture was diluted to 1 ml with TE (10 mM Tris·HCl, pH 7.5/1 mM EDTA) and stored at 4°C.

Amplification. The cDNA pool (1 μl) and amplification (3′amp) and adaptor primers (25 pmol each) in 50 μl of PCR cocktail [10% (vol/vol) dimethyl sulfoxide/1× *Taq* polymerase buffer (New England Biolabs)/each dNTP at 1.5 mM] were denatured (5 min, 95°C) and cooled to 72°C. Then 2.5 units of *Thermus aquaticus* (*Taq*) DNA polymerase (Perkin–Elmer–Cetus) was added and the mixture was overlaid with 30 μl of mineral oil (Sigma) at 72°C and annealed at 50–58°C for 2 min. The cDNA was extended at 72°C for 40 min. Using a DNA Thermal Cycler (Perkin–Elmer–Cetus), we carried out 40 cycles of amplification by using a step program (94°C, 40 sec; 50–58°C, 2 min; 72°C, 3 min), followed by a 15-min final extension at 72°C.

5′-End Amplification of cDNAs. See Fig. 1. One microgram of poly(A)$^+$ RNA was reverse transcribed as described above except for the addition of 20 μCi (1 Ci = 37 GBq) of [^{32}P]dCTP and the substitution of 20 pmol of 5RT primer for (dT)$_{17}$-adaptor. Excess 5RT was removed as follows: the 20-μl cDNA pool was applied to a Bio-Gel A-5m (Bio-Rad) column (in a 2-ml serological pipette plugged with silane-treated glass wool) equilibrated with 0.05× TE. Void volume (0.8 ml) and 30 one-drop fractions were collected. Fractions −4 to +3 relative to the first peak of radioactivity were pooled, concentrated by centrifugation under reduced pressure (Speedvac), and adjusted to 23 μl. For tailing, 1 μl of 6 mM dATP, 6 μl of 5× tailing buffer (Bethesda Research Laboratories), and 15 units of terminal deoxynucleotidyl-transferase (Bethesda Research Laboratories) were added, and the mixture was incubated for 10 min at 37°C and heated for 15 min at 65°C. The reaction mixture was diluted to 500 μl in TE and 1- to 10-μl aliquots were used for amplification as described for the 3′-end procedure, except for the substitution of (dT)$_{17}$-adaptor (10 pmol), adaptor (25 pmol), and amplification (5′amp, 25 pmol) primers.

Southern and RNA Blot Analysis. Ten-microliter aliquots of RACE reaction products were separated by electrophoresis [1% agarose gel containing ethidium bromide (EtdBr) at 0.5

Abbreviations: RACE, rapid amplification of cDNA ends; PCR, (DNA) polymerase chain reaction; 3′amp and 5′amp, gene-specific primers for 3′- and 5′-end amplification; 5RT, gene-specific primer for reverse transcription; EtdBr, ethidium bromide.
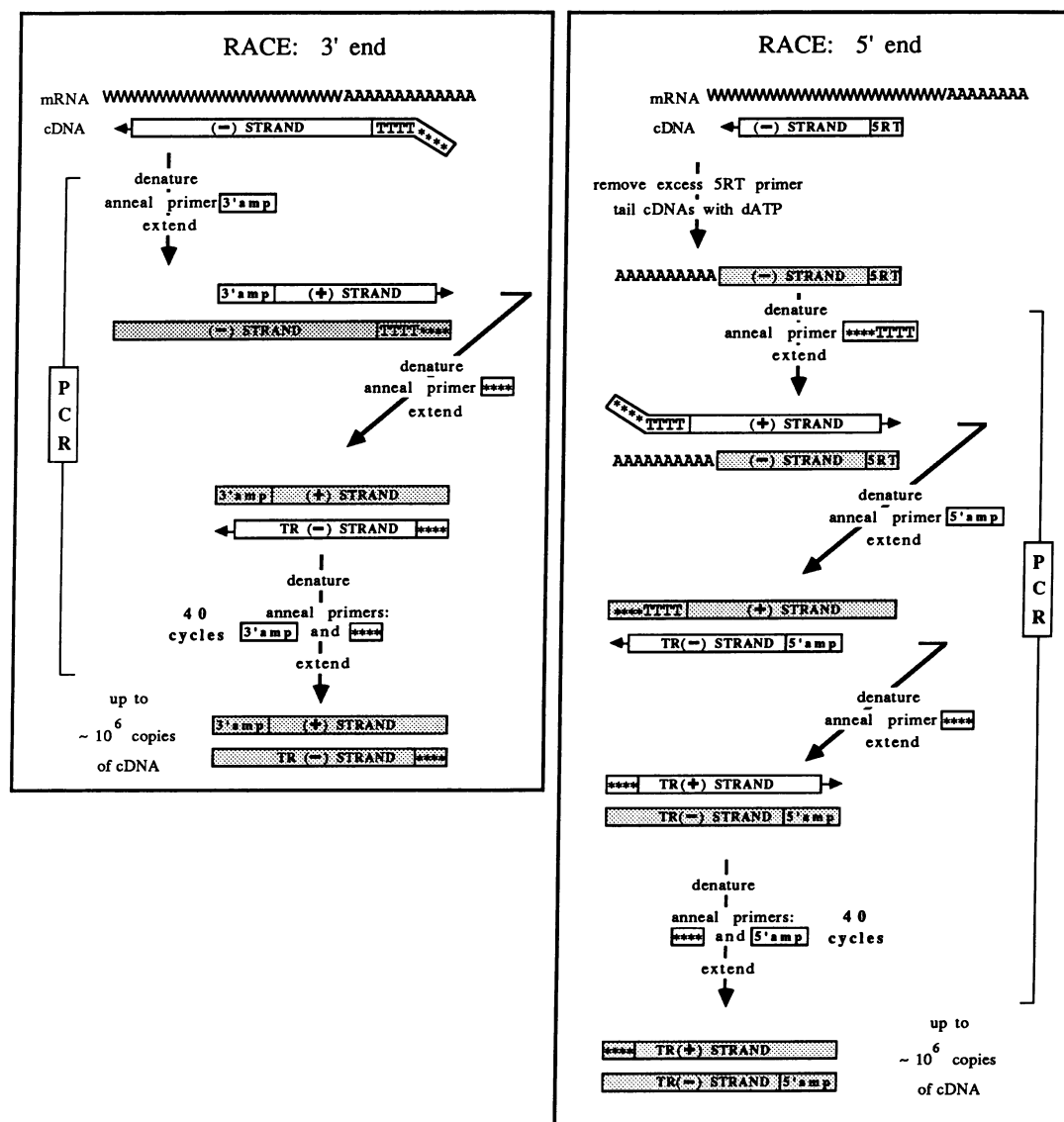
FIG. 1. Schematic representation of the RACE protocol. Primers: ****TTTT (dT)$_{17}$-adaptor, 5'-GACTCGAGTCGACATC-GATTTTTTTTTTTTTTTTTT-3'. This sequence contains the *Xho* I, *Sal* I, and *Cla* I recognition sites. ****, Adaptor, 5'-GACTCGAGTCGACATCG-3'. 3'amp (amp refers to amplification), specific to gene of interest, complementary to (−) strand. 5RT (RT refers to reverse transcription) and 5'amp, specific to gene of interest, complementary to (+) strand. Open rectangles represent DNA strands actively being synthesized; shaded rectangles represent DNA previously synthesized. At each step the diagram is simplified to illustrate only how the new product formed during the previous step is utilized. A (−) or (+) strand is designated as "truncated" (TR) when it is shorter than the original (−) or (+) strand, respectively.

µg/ml], transferred to GeneScreen (New England Nuclear; ref. 16), and hybridized at high stringency (17) with a ^{32}P-labeled probe (Bethesda Research Laboratories nick-translation kit). RNA blot analysis was carried out as previously described (17).

Cloning and Sequencing of cDNAs. RACE products were transferred into TE by using spun column chromatography (ref. 15, pp. 466–467), digested with restriction enzymes that recognize sites in the adaptor (*Cla* I or *Sal* I) or *int-2* sequences, and separated by electrophoresis. Regions of the gel containing specific products were isolated, and the DNA was extracted with Glassmilk (Bio 101, San Diego, CA) and cloned in a Bluescript vector (Stratagene, San Diego, CA). Plasmids with *int-2* cDNA inserts were identified by colony lift hybridization (ref. 15, p. 324). Restriction analyses (ref. 15, p. 104) were carried out on plasmid DNA prepared by the alkaline lysis method (18). Mini-prep plasmid DNA was sequenced with

Sequenase (United States Biochemicals, Cleveland), using the supplier's recommendations.

It should be noted that whereas all 23 clones generated after *Sal* I cleavage of the adaptor contained an intact *Sal* I site, many of the clones generated after cleavage with *Cla* I were altered at the *Cla* I site and could no longer be cleaved by that enzyme (M.A.F., unpublished data). It is unknown whether such alteration was due to the positioning of the *Cla* I site immediately adjacent to the oligo(dT) sequence.

RESULTS

The RACE Protocol. To obtain cDNA 3' ends (Fig. 1), a RNA population is reverse transcribed to create a cDNA (−) strand. The primer used is a 35-base oligonucleotide with 17 dT residues and an adaptor sequence containing three endonuclease recognition sites that are found infrequently in genomic DNA [(dT)$_{17}$-adaptor; see legend to Fig. 1]. The

presence of the adaptor places a unique sequence at the unknown end of the cDNA. Next, a gene-specific amplification primer (3′amp) is annealed with a small portion of the first (−) strand product (generally <1 ng) and extended to generate a complementary second (+) strand. Reiterative PCR cycles using 3′amp and adaptor primers then amplify the double-stranded cDNAs. Specificity of amplification is dependent on base-pairing of 3′amp only to molecules representing the mRNA of interest. The adaptor primer is used instead of the $(dT)_{17}$-adaptor primer because initial experiments suggested that long stretches of dT residues do not base pair well at the temperatures used to prevent mismatching of the specific primers.

A similar strategy is used to obtain cDNA 5′ ends (Fig. 1). The initial specificity is achieved by using a gene-specific primer for reverse transcription (5RT). The primer-extended products are then separated from excess primer and a poly(A) tail is added by using dATP and terminal deoxynucleotidyl-transferase. Second (+) strand synthesis is carried out with the $(dT)_{17}$-adaptor primer. Finally, the products are amplified by using the adaptor primer and a gene-specific amplification primer (5′amp) consisting of a sequence located upstream of the extension primer sequence. This "internal" primer increases the specificity and efficiency of the amplification reaction; it should bind only to the cDNAs of interest, whereas 5RT would bind to all (+) strand cDNAs present, including any generated through mismatched hybridization of 5RT during reverse transcription. Moreover, 5′amp would not bind to any residual 5RT that had been tailed. Such tailed 5RT, because of its small size and relatively large initial concentration, would serve as a potent substrate for a PCR reaction that included the 5RT and $(dT)_{17}$-adaptor primers, and this would greatly decrease the efficiency of the cDNA amplification.

After amplification, the products are analyzed by restriction and Southern blot analyses and cloned. The optimal cloning strategy utilizes one restriction endonuclease that cleaves within the adaptor sequence and a second that cleaves within the amplified region. This adds selectivity to the cloning process, since most of the nonspecific cDNA products will not be cleaved by the second enzyme and thus will not be cloned. Alternatively, the RACE products can be cloned intact by also incorporating a restriction site into the 5′ end of the gene-specific primer. Finally, full-length cDNA clones can be reconstructed from separate but overlapping 3′ and 5′ RACE products, or they can be synthesized by using primers whose sequences were obtained by analysis of the extreme 5′ and 3′ ends of the RACE products for PCR amplification of reverse-transcribed mRNA.

Isolation of *int-2* cDNA Clones by Using the RACE Protocol. We tested this scheme by generating cDNA copies of mRNA from *int-2*, a gene characterized by multiple transcripts expressed at very low abundance (≈2 copies per cell, ref. 19), and whose complete sequence is known (20). Because we had recently isolated *int-2* cDNA clones by conventional methods (21), we had a standard against which we could judge the results of the present study. On the basis of our previous results, we expected to obtain *int-2* cDNA clones with alternative start and polyadenylylation sites (Fig. 2A).

The first 3′amp primer used, 3′amp-D, was specific to the long *int-2* mRNAs (Fig. 2A). The 3′-end cDNA product of the amplification reaction was expected to be ≈1050 bp in length. DNA of the expected size was visible after EtdBr staining (Fig. 2B, lane 3′D), and its identity was confirmed by Southern blot analysis using an *int-2* cDNA probe (Fig. 2C, lane 3′D). In addition, most samples contained *int-2* cDNAs of unexpected size, which were subsequently determined to be single-stranded molecules (see below). In some circumstances other reaction products were visible after EtdBr staining, but they did not hybridize to the *int-2* probe; we found that such nonspecific PCR products could be mini-

mized by optimizing the annealing temperature (M.A.F., unpublished data). EtdBr-staining bands were also visible in the control sample, but they did not hybridize with the *int-2* probe (Fig. 2 B and C, lane 3′C).

The second 3′amp primer used, 3′amp-U, was common to both short and long classes (Fig. 2A) and thus was expected to generate cDNA products 480 and 1580 bp in length, respectively. Of these, only the 480-bp product was detectable by staining with EtdBr (Fig. 2B, lane 3′U), but the presence of both cDNAs was demonstrated by Southern blot analysis (Fig. 2C, lane 3′U). The observation that the 480-bp cDNA is ≈10-fold more abundant than the 1580-bp cDNA is inconsistent with the fact that the short classes of mRNA are only ≈2-fold more abundant than the long ones (Fig. 2A). This finding, that the relative abundance of different cDNA products generated by multiple cycles of PCR may not accurately reflect the composition of the mRNA population, is presumably due to a decrease in PCR efficiency as extension length increases.

When *int-2* cDNA 5′ ends were produced by using the RACE strategy, we expected ≈690- and ≈930-bp products representing the 5′ ends of the class B and the less abundant class A mRNAs, respectively (Fig. 2A). Although both fragments could be detected by Southern blot analysis, they were not always visible after EtdBr staining (Fig. 2 B and C, lane 5′P).

The specific products of the 3′- and 5′-end RACE reactions were isolated from agarose gels and cloned. The possibility afforded by the RACE protocol of separating different classes of cDNAs by gel electrophoresis prior to cloning was particularly useful in obtaining cDNAs representing class A mRNAs, which would otherwise have been difficult to detect in the presence of the much more numerous class B clones. However, we did not further size-select the RACE products in either class, and therefore we were able to examine the range of cDNAs produced by this protocol. Together the clones obtained represented all four transcripts expressed from the *int-2* gene, the longest of which is ≈2.9 kilobases (kb).

The cDNA clones representing the 3′ ends of the long and short classes of *int-2* mRNA were expected to terminate at *int-2* genomic positions 7508 and 6404, respectively. Sequence analysis of four independent cDNA clones (three of the long and one of the short class) confirmed that the products obtained by the RACE method were identical to those obtained by conventional cDNA cloning methods (21). Previous studies suggested that the 5′ ends of class A and B *int-2* mRNAs begin around positions 938 and 1677/78 of the *int-2* genomic sequence, respectively (21); however, single start sites are unlikely since the *int-2* promoters lack TATA boxes, and promoters of this type usually initiate transcription at multiple sites (23–25). Sequence analysis suggested that the majority of the clones studied were full length: 11/16 class B clones and 5/9 class A clones began at or within 21 bp of the presumed start sites (Fig. 2D). A higher frequency of full-length cDNA clones presumably could have been obtained by precise size selection of the RACE products prior to cloning.

In the course of developing and testing the RACE protocol, we discovered the following:

(*i*) For some mRNAs (not *int-2*) not all of the 3′ end cDNAs were full length. We found this to be the case when we characterized cDNA clones of transcripts from the mouse *En-1* gene (17); sequence analysis of 3′-end *En-1* cDNA clones suggested that in some cases the $(dT)_{17}$-adaptor primer had bound to an A-rich region in the coding sequence (AAAGAAGAAAAAGAA) upstream of the poly(A) tail. This could be minimized by decreasing the concentration of the $(dT)_{17}$-adaptor primer to 15 ng/μg of mRNA (M.A.F., unpublished data).

(*ii*) Some of the specific reaction products appeared to be single-stranded cDNAs. These reaction products hybridized
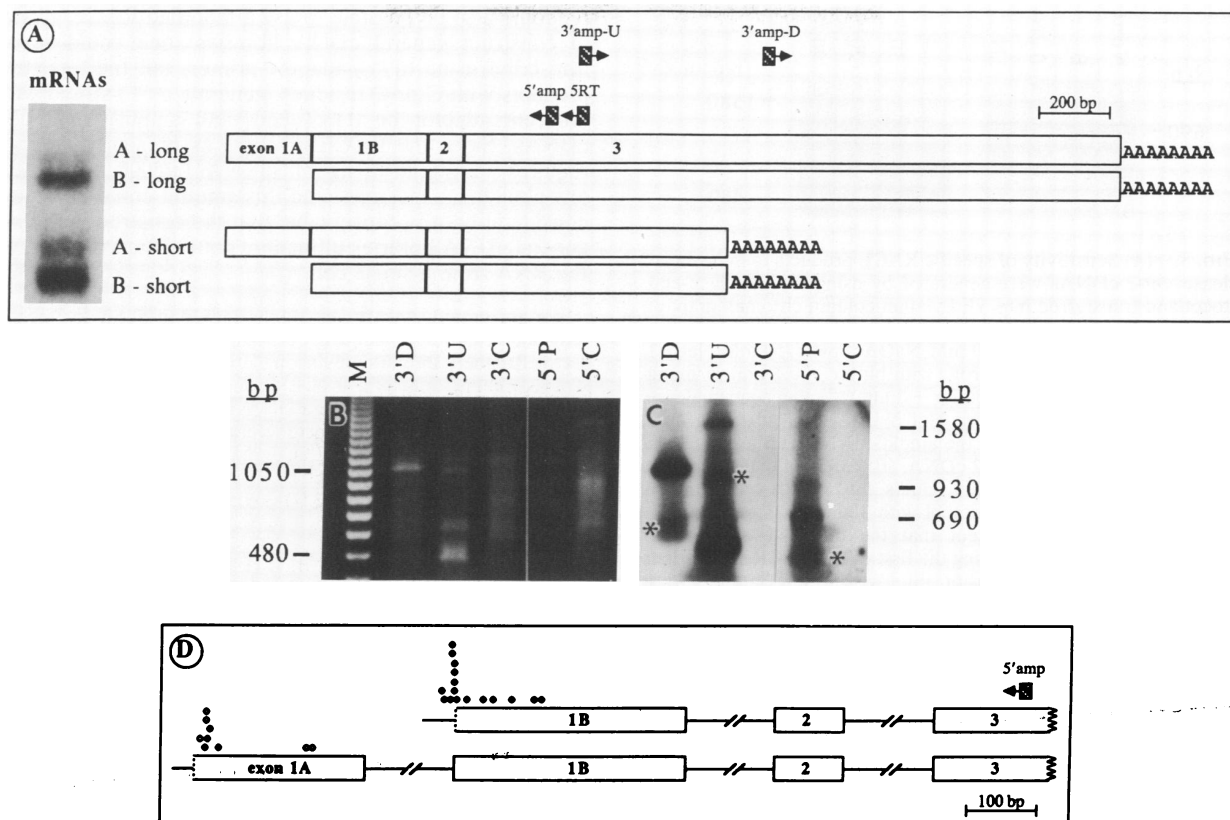
FIG. 2. Isolation of *int-2* cDNA clones by using the RACE protocol. (*A*) *int-2* mRNAs in teratocarcinoma-derived endodermal cells. (*Left*) RNA blot hybridized with an *int-2* exon 3 probe, showing four *int-2* transcripts (21). (*Right*) Depiction of structures of the four classes of *int-2* transcripts found in these cells as previously defined (21). Primers used are depicted as hatched boxes with arrows. Primer sequences: 3'amp primers 3'amp-U (5'-GGCAGAAGAAGCAGAGC-3') and 3'amp-D (5'AGAAGGCAGTGATCTGC-3'), complementary to (−) strand; 5RT (5'-GGCTCTGCTTCTTCTGC-3') and 5'amp primer (5'-CTCATGGTCCTTGTGGC-3'), complementary to (+) strand. bp, Base pairs. (*B* and *C*) Analysis of *int-2* RACE reaction products. Poly(A)$^+$ RNA isolated from endoderm-like cells obtained by treating F9 embryonal carcinoma cells with retinoic acid and dibutyryl-cAMP (22) was subjected to the RACE protocol. The reaction products were analyzed by electrophoresis and EtdBr staining (*B*) and Southern blot hybridization using an *int-2* cDNA probe (*C*). Marker lane (M) contains a 123-bp ladder (Bethesda Research Laboratories); sample lanes are labeled with the name of the gene-specific primer used in the amplification reaction (3'D, 3'amp-D, 3'U, 3'amp-U; 5'P, 5'amp) or as controls in which amplification was carried out in the absence of an *int-2* primer (3'C, 5'C). An * indicates single-stranded *int-2* cDNA. (*D*) Analysis of 5'-end *int-2* cDNA clones. The thin lines represent *int-2* genomic DNA and the open boxes the exons of the class A and B *int-2* mRNAs. The 5' ends of the mRNAs are symbolized by dotted vertical lines to indicate that multiple start sites are used. The jagged lines indicate that only the 5' portion of exon 3 is shown. The filled dots indicate the positions of the 5' ends of RACE clones as determined by sequence analysis, and the open dots the 5' ends as determined by restriction analysis.

to *int-2* probes, but they did not cleave with appropriate restriction endonucleases and were eliminated by digestion with mung bean nuclease (Fig. 3A). The quantity of these single-stranded products was also decreased by altering the buffer conditions (Fig. 3*A*, lane B), suggesting that complementary DNA strands are present and that appropriate binding can occur under suitable conditions. Single-stranded products are not usually observed with PCR, but use of a gene-specific primer in conjunction with a primer that binds to every cDNA present and that contains a long homopolymeric stretch may lead to greater production of one of the strands, or to mismatched pairing of some strands after the final round of PCR. Because these single-stranded products cannot easily be cloned, it is important to determine which of the specific products are double stranded.

(*iii*) The relative amount of longer cDNA products could be dramatically increased in several cases by substantially increasing the extension time of the initial (+) strand synthesis reaction (Fig. 3*B*). We also noted that small changes in the annealing temperatures (2°C) could cause significant changes in the ratio of specific to nonspecific cDNA amplification (M.A.F., unpublished data). In practice, we suggest using a conservative annealing temperature (50°C) for initial studies and raising it until the optimal temperature is reached for obtaining the cDNA product of interest, as judged by Southern blot analysis.

DISCUSSION

The RACE protocol described here provides an alternative to other cDNA cloning methods that is advantageous in many respects. A single gene-specific primer is used to generate 3'-end clones, whose sequences can be used, if necessary, to design the primers required for 5'-end cloning. The protocol involves few steps and the only reagents that must be specially prepared are the oligonucleotide primers. Moreover, steps common to other methods can be much simpler in the RACE protocol. For example, terminal deoxynucleotidyltransferase tailing conditions are generally optimized and monitored, since it is considered important to limit the length of the tail (4, 5). Using the RACE protocol, however, we found that the length of the poly(A) tail on the cDNA products did not exceed 34 nucleotides, even though no precautions were taken to limit tail length.

The RACE protocol is rapid and can be executed in 1 day. Moreover, it is unique in that substantial information about
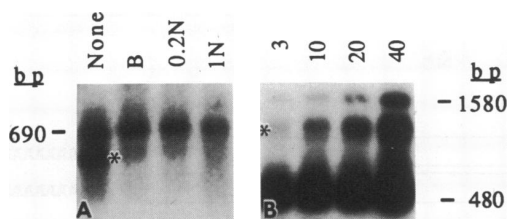
FIG. 3. (*A*) Sensitivity of specific RACE products to mung bean nuclease. *int-2* cDNAs were generated by using the RACE 5'-end protocol and transferred into TE by spun column chromatography. The samples were then incubated at 30°C for 30 min with no additions (None), with buffer only [B: 30 mM NaOAc, pH 5/50 mM NaCl/1 mM ZnCl$_2$/5% (vol/vol) glycerol], with buffer + 0.2 unit of mung bean nuclease (0.2N), or with buffer + 1 unit of mung bean nuclease (1N), and analyzed by Southern blot hybridization using an *int-2* probe. The 690-bp fragment is the amplified class B 5'-end cDNA; * indicates single-stranded *int-2* cDNA. (*B*) Effect of extension time on production of specific RACE products. *int-2* cDNA was prepared by using the RACE 3' end protocol with 3'amp-U as described in *Materials and Methods*, except that the time allowed for the synthesis of the (+) strand during the first cycle of PCR was varied; in subsequent steps the extension time was 3 min. Samples were analyzed by Southern blotting. Lanes are designated according to the length of the extension time (3, 10, 20, and 40 min). The 480- and 1580-bp fragments are 3'-end cDNAs of the short and long *int-2* mRNAs, respectively; * indicates single-stranded *int-2* cDNA.

the cDNAs can be obtained prior to cloning and screening steps. Specifically, the RACE products can be examined by Southern blotting to determine whether the desired cDNAs are present and to evaluate the complexity of the cDNA population. This allows cDNAs representing unusual transcripts, such as those created by alternative splicing or by the use of alternative promoters, to be selected for cloning.

The RACE protocol is more efficient than other methods. This is evident from a comparison of the results of cloning *int-2* cDNAs by using conventional library construction vs. the RACE method. Starting with a total of 15 μg of poly(A)$^+$ RNA, we constructed two cDNA libraries, screened 2.4 × 10^6 recombinant phage, and isolated a total of 9 *int-2* cDNAs (21). In the present study we started with 0.5 μg of poly(A)$^+$ RNA, generated and screened several hundred recombinant plasmids with <0.05% of the RACE products, and isolated 29 independent *int-2* cDNA clones. Sequence analysis demonstrated that the 3' and 5' ends of all of four *int-2* mRNAs were accurately represented by the RACE products. Moreover, all of the 3'-end clones terminated at known polyadenylylation sites and the majority of the 5'-end RACE clones appeared to be full length. Thus, we have used this technique to readily obtain full-length cDNA clones representing transcripts 1.4 to 2.9 kb long. As our data suggest, however, PCR amplification becomes less efficient with increasing fragment length. Therefore, in cases in which the generation of full-length cDNAs would require amplification over a region more than several kilobases in length, production of the desired cDNAs may be difficult. Nevertheless, in most cases it should be possible to devise strategies to overcome this limitation. Another potential problem is that a limited number of errors are introduced by the use of multiple rounds of amplification (ref. 7; ≈0.1%, unpublished data of M.A.F.). Therefore, the sequence of a cDNA clone obtained by the RACE method should be verified by comparison with sequences of several independent cDNA clones or genomic DNA.

Other potential applications of this protocol are evident. For example, primers based on amino acid sequence might be adequate for RACE cDNA cloning, although our preliminary data suggest that the feasibility of this approach may be limited by nonspecific amplification due to primer mismatch.

It should also be possible to use a modified RACE protocol to construct general cDNA libraries: one could reverse transcribe using (dT)$_{17}$-adaptor primer, tail the (−) strand products with G or C residues, generate a (+) strand with a different adaptor on its 5' end, and amplify the pooled cDNAs. Since the RACE protocol as presently described requires less than 1 ng of mRNA (≈5000 copies of a low-abundance message) and this amount could be reduced considerably, it is conceivable that the construction of such general cDNA libraries ultimately could be carried out with the RNA from a single cell.

1. Gubler, U. & Hoffman, B. J. (1983) *Gene* 25, 263–269.
2. Heidecker, G. & Messing, J. (1987) *Methods Enzymol.* 159, 28–40.
3. Efstratiadis, A., Kafatos, F. C. & Maniatis, T. (1977) *Cell* 10, 571–586.
4. Okayama, H., Kawaichi, M., Brownstein, M., Lee, F., Yokota, T. & Arai, K. (1987) *Methods Enzymol.* 159, 3–27.
5. Coleclough, C. (1987) *Methods Enzymol.* 159, 64–83.
6. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* 230, 1350–1354.
7. Saiki, R. K., Gelfand, D. H., Stoffel, B., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* 239, 487–491.
8. Kogan, S. C., Doherty, M. & Gitschier, J. (1987) *N. Engl. J. Med.* 317, 985–990.
9. Embury, S. H., Scharf, S. J., Saiki, R. K., Gholson, M. A., Golbus, M., Arnheim, N. & Erlich, H. A. (1987) *N. Engl. J. Med.* 316, 656–661.
10. Lee, M. S., Chang, K., Cabanillas, F., Freireich, E. J., Trujillo, J. M. & Stass, S. A. (1987) *Science* 237, 175–178.
11. Lee, C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M. & Caskey, C. T. (1988) *Science* 239, 1288–1291.
12. Ou, C.-Y., Kwok, S., Mitchell, S. W., Mack, D. H., Sninsky, J. J., Krebs, J. W., Feorino, P., Warfield, D. & Schochetman, G. (1988) *Science* 239, 295–297.
13. Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J. & Scott, J. (1987) *Cell* 50, 831–840.
14. Cathala, G., Savouret, J., Mendez, B., West, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983) *DNA* 2, 329–335.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1991–1995.
17. Joyner, A. L., Kornberg, T., Coleman, K., Cox, D. & Martin, G. R. (1985) *Cell* 43, 29–37.
18. Birnboim, C. & Doly, J. (1979) *Nucleic Acids Res.* 7, 1513–1523.
19. Jakobovits, A., Shackleford, G. M., Varmus, H. E. & Martin, G. R. (1986) *Proc. Natl. Acad. Sci. USA* 83, 7806–7810.
20. Moore, R., Casey, G., Brookes, S., Dixon, M., Peters, G. & Dickson, C. (1986) *EMBO J.* 5, 919–924.
21. Mansour, S. B. & Martin, G. R. (1988) *EMBO J.* 7, 2035–2041.
22. Strickland, S., Smith, K. K. & Marotti, K. R. (1980) *Cell* 21, 347–355.
23. Melton, D. A., Konecki, D. S., Brennand, J. & Caskey, C. T. (1984) *Proc. Natl. Acad. Sci. USA* 81, 2147–2151.
24. Reynolds, G. A., Baser, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. & Luskey, K. L. (1984) *Cell* 38, 275–285.
25. Ishii, S., Merlino, G. T. & Pastan, I. (1985) *Science* 230, 1378–1381.