3796–3797 Nucleic Acids Research, 1995, Vol. 23, No. 18

cRACE: a simple method for identification of the 5′ end of mRNAs

Ichiro N. Maruyama*, Terese L. Rakow and Hiroko I. Maruyama

Department of Cell Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, USA

Received May 19, 1995; Accepted August 18, 1995

We have devised a simple method for identification of the 5′ end of mRNAs in which the first strand cDNA is circularized and/or joined into a concatemeric form by T4 RNA ligase and the resulting single-stranded DNA is subsequently used as a template for amplification of the 5′ end by the polymerase chain reaction (1) (PCR) with gene-specific primers. The technique has been used for determination of the 5′ end of *Caenorhabditis elegans unc-13* and T12A2.4 mRNAs.

Most clones isolated from cDNA libraries lack their 5′ end, due to premature termination of cDNA synthesis by reverse transcriptase. Therefore, 5′-end mapping of mRNA is a crucial step for the analysis of a gene and its promoter. A variety of methods are available for mapping mRNA ends, including RNase protection, S1 mapping and primer extension (2). These methods require relatively large amounts of mRNA and often pose difficulty in identifying the 5′ end of rare mRNAs. An alternative procedure has been described as the rapid amplification of cDNA ends (RACE) (3), in which mRNA is reverse-transcribed with a gene-specific primer and the cDNA is modified at the 5′ end with a homo-oligonucleotide synthesized by terminal deoxynucleotidyl transferase. The tailed cDNA is subjected to PCR amplification with a complimentary oligonucleotide to the tail and a gene-specific primer. By this technique, it is often difficult to detect the 5′ end of rare mRNAs, due to inefficiency of the tailing reaction and/or non-specific priming by a homopolymeric primer. To overcome this problem, anchored RACE has recently been introduced (4–6). To prevent self ligation of cDNA, the method requires an excess amount of 'anchor' oligonucleotide that is linked to first-strand cDNA for subsequent PCR amplification. Furthermore, to prevent self-ligation of the anchors, the 3′ end of the oligonucleotide has to be modified by incorporating an unusual nucleotide such as dideoxynucleotides. Ligation of the modified oligonucleotide anchor to mRNA requires multi-step chemical and enzymatic reactions to remove the 5′-end capping structure. To simplify the multi-step processing of cDNA and mRNA, we have devised a method that we have denoted as cRACE (circular or concatemeric first-strand cDNA-mediated RACE). The method has been used for identification of the 5′ end of the *C.elegans unc-13* mRNA, which encodes a diacylglycerol/phorbol ester receptor and plays a role in neurotransmission (7). The mRNA is expressed in only a subset of neurons at a very low level, ~ 0.1% of abundant mRNAs such as the myosin heavy chain mRNA in the nematode (T.L.R. and I.N.M., unpublished observation).

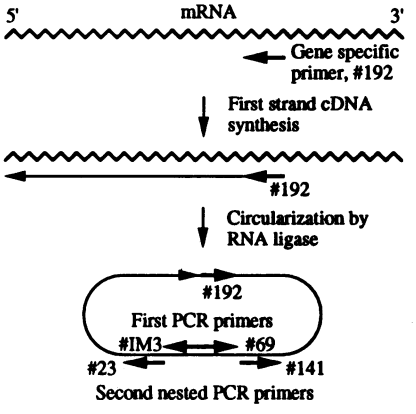

Figure 1. Schematic diagram of cRACE for amplification of the 5′ end of the *unc-13* mRNAs. Pictures are not drawn to scale. Primers for cDNA synthesis and PCR are depicted by short, horizontal arrows with their name (refer to the text). Wavy and straight lines represent the mRNA and first strand cDNA, respectively. When the first strand cDNA derived from an abundant mRNA is ligated at a high concentration, the cDNA may form concatemers instead of a circular structure. However, both circular and concatemeric cDNAs produce the same products by subsequent PCR amplifications.

Total RNA from a mixed culture of *C.elegans*, strain N2, was prepared as previously described (8) and mRNA was purified with oligo(dT) cellulose (Pharmacia). First-strand cDNA of the *unc-13* mRNA was synthesized from 0.5 μg of the mRNA, using a cDNA synthesis kit (Promega) with a gene-specific oligonucleotide #192, 5′-pGTTCCTACAGTTTGACCATTT, which was synthesized by a DNA synthesizer (Applied Biosystems) and phosphorylated by T4 polynucleotide kinase using standard techniques (9). After hydrolysis of the template mRNA with 0.5 N NaOH (final concentration) at 37°C for 10 min, the cDNA was precipitated with ethanol and redissolved in 10 μl of a reaction mixture containing 25% polyethylene glycol (mol. wt ~8000), 1 mM hexamine cobalt chloride, 0.01 mM ATP, 10 mM Tris–HCl (pH 8.0), 10 mM $MgCl_2$, 1 μg/ml bovine serum albumin, and 10 U T4 RNA ligase (New England BioLabs). This

* To whom correspondence should be addressed

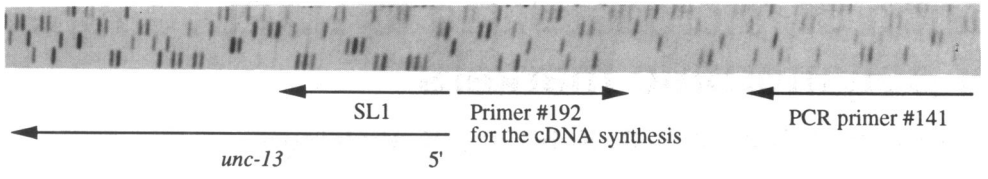

Figure 2. Nucleotide sequence of the junctional region between the 5′ end and the primer used for the cDNA synthesis of *unc-13*. The lanes are T, C, G and A from top to bottom. Note that the *unc-13* mRNA was trans-spliced to a splice leader sequence SL1. This sequence proves that reverse transcriptase synthesized complete cDNA without missing a single base at the 5′ end and the first strand cDNA was circularized without eliminating a single base at either end. The 5′-terminal sequence of the *unc-13* cDNA was confirmed by DNA sequencing of the corresponding genomic region (not shown). The same precise processing of the T12A2.4 mRNA and cDNA was also observed (not shown).

ligation mixture was incubated at 22°C for 16 h and an aliquot was directly used as a template for the first PCR amplification with gene-specific primers #IM3 (5′-AAAACCCTTTGGCCCATA) and #69 (5′-TGGGACAAATTAATAGGT), as illustrated in Figure 1. A PCR reaction was set up in a total volume of 20 µl containing 0.2 µl of the cDNA treated with T4 RNA ligase, 1 µM each of the primers, 0.1 mM each dNTP, 10 mM Tris–HCl (pH 7.0), 50 mM KCl, 2.5 mM $MgCl_2$, and 1 U *Taq* polymerase (Stratagene). The reaction was then carried out for 30 cycles with the following conditions: 94°C for 30 s denaturation, 55°C for 30 s annealing and 72°C for 1 min extension, using an Ericomp EasyCycler (San Diego, CA). The resulting PCR product was diluted 10^3-fold with sterile H_2O, and a 2 µl aliquot was used as a template for second nested PCR amplification with primers #23 (5′-ATCTCGAGCGATTTGTTTCAAAAATGAATTC) and #141 (5′-ATGGTACCTTCAAATGGATCATG) under the same conditions. The product with a size of 360 bp was digested with *Xho*I and *Kpn*I, of which sequences are underlined in the primers, and cloned into a plasmid vector, Bluescript II KS+ (Stratagene). The DNA sequence of the product was determined by the dideoxy chain-termination method (10) and it was verified that the *unc-13* mRNA was trans-spliced to a splice-leader sequence SL1, 5′-GGTTTAATTACCCAAGTTTGAG (11), as shown in Figure 2.

In order to test reproducibility of cRACE, we have also applied the technique to identify the 5′ end of T12A2.4 mRNA, which has been defined by the *C.elegans* genome sequencing project (12). The gene encodes protein kinase C CII regulatory domains but has not previously been analysed at the mRNA or protein level. Primers for the first-strand cDNA synthesis and subsequent PCR reactions were selected from the exon sequences of the gene. We successfully identified a cRACE product and its sequence identified a new first exon that is located 1418 bp upstream of the first exon previously defined by the genome sequencing project. It has also been found that the mRNA product is trans-spliced to the SL1 leader sequence.

In summary, we have successfully used cRACE to identify the 5′ end of the low abundant *unc-13* mRNA and previously unanalysed T12A2.4 mRNA. The technique is similar to a method (13) in which circularization of RNA is followed by PCR amplification. The latter requires a multi-step decapping process of mRNA prior to the circularization. The cRACE method shares a number of advantages of the other RACEs for mapping the 5′ end of cDNA, including precise mapping of the end to the single nucleotide level based on DNA sequencing. However, unlike other RACEs, cRACE requires neither decapping of mRNA nor a modified 'anchor' oligonucleotide. In addition, PCR primers for cRACE are all gene-specific and non-specific PCR products should be less likely to be amplified. These features demonstrate that cRACE may be a simpler, hence more efficient alternative to the conventional RACE.

ACKNOWLEDGEMENTS

The authors are grateful to Sydney Brenner for his support and encouragement, and Ichiro Kuwabara and Gi Mikawa for their comments on the manuscript. This work was supported in part by the Lucille P. Markey Charitable Trust and Taitec Corp.

REFERENCES

1 Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science*, **239**, 487–491.
2 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1988) *Current Protocols in Molecular Biology*. Wiley Interscience, NY.
3 Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8998–9002.
4 Dumas Milne Edwards, J.B., Delort, J. and Mallet, J. (1991) *Nucleic Acids Res.*, **19**, 5227–5232.
5 Troutt, A.B., McHeyzer-Williams, M.G., Pulendran, B. and Nossal, G.J.V. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 9823–9825.
6 Liu, X. and Gorovsky, M.A. (1993) *Nucleic Acids Res.*, **21**, 4954–4960.
7 Maruyama, I.N. and Brenner, S. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 5729–5733.
8 Maruyama, I.N. and Brenner, S. (1992) *Gene*, **120**, 135–141.
9 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
10 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
11 Krause, M. and Hirth, D. (1987) *Cell*, **49**, 753–761.
12 Wilson, R. *et al.* (1994) *Nature*, **368**, 32–38.
13 Mandl, C.W., Heinz, F.X., Puchhammer-Stockl, E. and Kunz, C. (1991) *BioTechniques*, **10**, 485–486.