Benchmarks

5'-RACEing Across a Bridging Oligonucleotide

BioTechniques 32:480-482 (March 2002)

Rapid amplification of cDNA 5' ends (5'-RACE) is a convenient way to retrieve information about the 5' termini of mRNAs because entire open reading frames are difficult to predict from genomic information and expressed sequence tag libraries generally lack full 5' termini. To achieve this, numerous methods have been developed, each of which has been successful in specific situations (3-5,11-14,17). However, two major generic drawbacks are associated with these procedures. First, a general non-gene-specific primer is introduced for PCR amplification that can lead to nonspecific amplification products (3,5,11-14,17). Second, the methods used to introduce the non-gene-specific primer sequence into the cDNA can be problematic. Some protocols involve ligation of ssDNA by T4 RNA ligase, the efficiency of this reaction being quite low (4,11). Others use terminal deoxynucleotidyl transferase (TdT) to tail the first-strand cDNA (5,11,12) and require additional steps, such as cDNA purification before and after tailing, which reduces the recovery of cDNA. Furthermore, the TdT tailing itself can be inefficient and difficult to control. To circumvent these drawbacks, we have devised a novel and effective method to amplify cDNA 5' ends, termed BO-5'RACE, for 5'-RACEing across a bridging oligonucleotide (BO).

During an examination of the structural proteins residing within the peritrophic matrix, a structure lining the insect midgut of *Mamestra configurata* (Bertha armyworm), we isolated a 1.8-kb clone from a midgut cDNA library that bore a high degree of similarity to the *Trichoplusia ni* mucin (16). Northern blot studies revealed that the full-length transcript was approximately 3.5 kb. To obtain the complete mucin cDNA sequence, we developed the BO-5'RACE protocol described below. Total RNA was isolated from midgut tissues of fourth instar larvae using TRIzol® Reagent (Invitrogen, Carlsbad, CA, USA). The partial sequence

of the mucin cDNA was used to design oligonucleotides (Table 1) for reverse transcription (RT), the BO, and PCR (F1, F2, R1, and R2) to test the BO-5'RACE protocol (Figure 1). Notably, the RT primer is 5'-phosphorylated to allow for the head-to-tail ligation of the first-strand cDNA. A restriction endonuclease recognition site can also be introduced into the 5' end of the RT primer if cloning of the cDNA 5' terminus is required. The standard first strand cDNA synthesis reaction con-

sisted of 5 µg total RNA, 2 pmol RT primer, first-strand buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂), 10 mM DTT, 6 mM MgCl₂, 0.5 mM each dNTP, and 200 U SUPERSCRIPT II™ (Invitrogen) in a 20-µL reaction. This mixture was incubated at 42°C for 60 min, and the reverse transcriptase was inactivated by incubation at 70°C for 15 min. The remaining mRNA was eliminated with the addition of 1 µL RNase H (Invitrogen) and incubation at 37°C for 1 h to pre-

Figure 1. BO-5'RACE protocol. Flowchart outlining the events involved in BO-5'RACE. An internal RT primer is used to catalyze first-strand cDNA synthesis. At elevated dGTP concentrations, SUPERSCRIPT II adds non-template-encoded C residues to the 3' termini of first-strand cDNA molecules derived from mRNA possessing intact 7-methylguanosine cap structures. The poly-C tail anneals to the three Gs of the BO that tether the 5' and 3' termini of the first-strand cDNA. The RT primer for cDNA synthesis is phosphorylated to facilitate inter- or intramolecular ligation of the first-strand cDNA, as directed by the BO. Two gene-specific primers are then used to amplify the unknown 5' terminus.

Table 1. Primers Used for BO-5'RACE of *M. configurata* Mucin cDNA

Primer	Sequence (5'→3')
RT ^a	CTTGTCGCAGTCAGATTC
BO	CATGAATCTGACTGCGACAAGGGG
F1	CTCCTACTACCGAAGCAGACAGCG
F2	ATACCTCTCCCAATGACCCCTGAC
R1	CCTCTCGATCTTCTTACCAGCG
R2	TCGTGGGGGAGGTGGTGAATG

^aRT primer must be 5'-phosphorylated.

vent possible interference with the subsequent BO-assisted ligation. Annealing of the BO to the first-strand cDNA was achieved by adding 10 pmol BO to the cDNA reaction, incubating at 95°C for 5 min, and then cooling to room temperature for 30 min. A 1-μL aliquot of the BO-annealed cDNA was then ligated in a 20-μL reaction containing ligation buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1 mM rATP, 25 μg/mL BSA) and 0.4 U T4 DNA ligase (New England Biolabs, Beverly, MA, USA) at 16°C overnight. If the target mRNA is present in low quantities, then the entire BO-annealed

first-strand cDNA can be precipitated with ethanol before ligation.

To examine the effects of elevated dCTP or dGTP relative to the other dNTPs on BO-5'RACE efficiency, a skewed ratio of dNTPs was used. The dNTP concentration was lowered to 5–50 μM, this being sufficient for cDNA synthesis, and the concentration of dGTP (or dCTP) was increased to 5 mM, resulting in a 100–1000 times excess. PCR amplification was performed using 1 μL ligated cDNA with R1 and F1 primers under the following reaction conditions: 5 min of incubation at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 63°C, and 2 min at 70°C, and a final 10-min extension at 72°C. A 100-μL PCR consisted of 2.5 U *Pfu* DNA polymerase (Stratagene, La Jolla, CA, USA), buffer (20 mM Tris-HCl, pH 8.8, 2 mM MgSO₄, 10 mM KCl, 10 mM (NH₄)₂SO₄, 0.1% Triton® X-100, 0.1 mg/mL nuclease-free BSA), 0.1 μM each primer, 0.2 mM each dNTP, and 1 μL ligated cDNA. Amplified DNA was separated by agarose gel electrophoresis; when used as the DNA template, an internal region of the 1.2-kbp DNA band was amplified by a second set of gene-specific primers, R2 and F2, which confirmed the derivation from the mucin cDNA. Sequencing of the 1.2-kbp fragments from each reaction revealed a common 5' termini with the characteristic three Gs associated with the BO, which suggested that we had obtained the 5' end of the mucin cDNA. The full-length *M. configurata* mucin cDNA was determined to be 2960 bp long (not including the poly A tail), closely matching the size estimated by Northern analysis, and encoding a

protein of 811 amino acids. The sequence of *M. configurata* mucin has been deposited in GenBank® (accession no. AY057052).

BO-5'RACE is based on the TdT-like activity of reverse transcriptases (i.e., the ability to add non-template-encoded nucleotides to the 3' termini of their products). The normal preference of nucleotides for this addition is A > G >> T ≈ C (2). However, when 5' 7-methylguanosine-capped mRNA was used as the substrate, C was added preferentially (6,10,13,15,17) because reverse transcriptases reverse-transcribe the 7-methylguanosine cap nucleotide into a 3'-terminal C residue (6,15). It has been demonstrated that reverse transcribing capped mRNA with avian myeloblastosis virus (AMV) reverse transcriptase resulted in 88% of the first-strand cDNA possessing a single additional C, while 6% had two Cs and 6% had an additional A residue (10). Normally, the addition of a second or third non-template-encoded nucleotide is less common (1,7,10,13); however, under specific cDNA synthesis conditions, such as elevated Mg²⁺ levels (17) or the addition of Mn²⁺ (13), longer stretches of extra nucleotides are incorporated (13,17).

In BO-5'RACE, a BO is introduced that possesses three Gs at its 3' end. The three-G tail is designed to hybridize with the non-template-encoded Cs added to the first-strand cDNA—three Gs being chosen to increase hybrid stability. Another property of the BO is that it encompasses the region complementary to the RT primer used for cDNA synthesis. As such, the BO tethers the “head” and “tail” of the first-strand cDNA and coordinates the intra- and intermolecular ligation events, although the frequency of the latter will be relatively low because of the limited abundance of full-length mRNAs. Two gene-specific primers are then used to amplify the 5' terminus of the cDNA. The results showed that the BO-5'RACE worked well when cDNA synthesis was carried out with elevated Mg²⁺ levels and an equal molar ratio of dNTPs (0.5 mM each) (Figure 2, lane 1). Other studies have shown that under biased ratios of dNTPs, reverse transcriptases preferentially incorporate the nucleotide whose complementary

Figure 2. Amplification of *M. configurata* mucin 5' terminus using primers F1 and R1 by BO-5'RACE. The amplification conditions were as described in the text. The arrowhead indicates the position of the expected DNA fragment. Lane M, DNA molecular weight markers (GeneRuler™; MBI Fermentas, Hanover, MD, USA); lane 1, products amplified from cDNA synthesized using equal molar amounts of dNTPs (0.5 mM each); lane 2, products amplified from cDNA synthesized using excess dGTP (5 mM dGTP and 5 μM each dNTP); lane 3, amplification with primer F1 alone; and lane 4, amplification with primer R1 alone.

Benchmarks

dNTP is present in higher relative concentrations in the reaction (8,9). Indeed, when cDNA was synthesized using higher levels of dGTP relative to the other dNTPs, the BO-5'RACE often generated stronger and more specific products (Figure 2, lane 2). This is likely due to a higher proportion of cDNAs with elongated non-template-encoded C tails able to hybridize to the BO. As predicted, synthesis of cDNA under higher dCTP conditions reduced the yield of BO-5'RACE product because this will reduce the length of the C-tail addition (8,9). To determine whether increasing dGTP relative to other dNTPs leads to the introduction of mutations in the first-strand cDNA, we sequenced the fragments generated under normal cDNA synthesis conditions, at elevated dCTP and dGTP levels. Analysis of the sequencing results showed no base substitutions, insertions, or deletions, which suggested that, in the presence of sufficient dNTPs, increasing one dNTP relative to other three does not lead to elevated mutation rates.

In comparison to other 5'-RACE methods, the BO-5'RACE protocol preferentially amplifies full-length cDNAs in a simple manner with no need for complicated cap-capturing methods. Also, specificity and efficiency are increased because of certain inherent properties. The BO-assisted ligation occurs only with the cDNA molecules with 3'-protruding C termini, reducing possible contamination from genomic DNA. The PCR amplification is carried out with two gene-specific primers rather than the single primer normally used in other 5'-RACE methods. Also, the two gene-specific primers in BO-5'RACE oppose one another, making amplification from genomic DNA impossible. Finally, degraded mRNA does not possess the 5' 7-methylguanosine cap structure and is not a suitable substrate for BO-assisted first-strand cDNA ligation.

While no 5'-RACE method is technically generic enough to allow for high-throughput isolation of full-length cDNAs, the BO-5'RACE method is proving to be relatively robust, as we have been able to amplify other difficult mRNA templates containing highly repetitive regions, such as those found

in mucin. Simplicity and reliability become especially important when a rare mRNA is the subject of investigation.

REFERENCES

1. **Brownstein, M.J., J.D. Carpten, and J.R. Smith.** 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques* 20:1004-1010.
2. **Chen, D. and J.T. Patton.** 2001. Reverse transcriptase adds nontemplated nucleotides to cDNAs during 5'-RACE and primer extension. *BioTechniques* 30:574-582.
3. **Chenchik, A., L. Diachenko, F. Moqadam, V. Tarabykin, S. Lukyanov, and P.D. Siebert.** 1996. Full length cDNA cloning and determination of mRNA 5'- and 3'-ends by amplification of adaptor-ligated cDNA. *BioTechniques* 21:526-534.
4. **Eyal, Y., H. Neumann, E. Or, and A. Frydman.** 1999. Inverse single-strand RACE: an adaptor-independent method of 5' RACE. *BioTechniques* 27:656-658.
5. **Fehr, C., M. Fickova, C. Hiemke, and N. Dahmen.** 1999. Rapid cloning of cDNA ends polymerase chain reaction of G-protein-coupled receptor kinase 6: an improved method to determine 5'- and 3'-cDNA ends. *Brain Res. Protocols* 3:242-251.
6. **Hirzmann, J., D. Luo, J. Hahnen, and G. Hobom.** 1993. Determination of messenger RNA 5'-ends by reverse transcription of the cap structure. *Nucleic Acids Res.* 21:3597-3598.
7. **Magnuson, V.L., D.S. Ally, S.J. Nylund, Z.E. Karanjawala, J.B. Rayman, J.I. Knapp, A.L. Lowe, S. Ghosh, and F.S. Collins.** 1996. Substrate nucleotide-determined non-templated addition of adenine by *Taq* DNA polymerase: implications for PCR-based genotyping and cloning. *BioTechniques* 21:700-709.
8. **Martinez, M.A., M. Sala, J.-P. Vartanian, and S. Wain-Hobson.** 1994. Hypermutagenesis of RNA using human immunodeficiency virus type 1 reverse transcriptase and biased dNTP concentrations. *Proc. Natl. Acad. Sci. USA* 91:11787-11791.
9. **Martinez, M.A., M. Sala, J.-P. Vartanian, and S. Wain-Hobson.** 1995. Reverse transcriptase and substrate dependence of the RNA hypermutagenesis reaction. *Nucleic Acids Res.* 23:2573-2578.
10. **Mules, E.H., O. Uzun, and A. Gabriel.** 1998. *In vitro* Tyl reverse transcription can generate replication intermediates with untidy ends. *J. Virol.* 72:6490-6503.
11. **Schafer, B.C.** 1995. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Anal. Biochem.* 227:255-273.
12. **Schmidt, W.M. and M.W. Mueller.** 1996. Controlled ribonucleotide tailing of cDNA ends (CRTC) by terminal deoxynucleotidyl transferase: a new approach in PCR-mediated analysis of mRNA sequences. *Nucleic Acids Res.* 24:1789-1791.
13. **Schmidt, W.M. and M.W. Mueller.** 1999. CapSelect: a highly sensitive method for 5' cap-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Res.* 27:e31.
14. **Shi, X. and S.G.W. Kaminskyj.** 2000. 5' RACE by tailing a general template-switching oligonucleotide. *BioTechniques* 29:1192-1195.
15. **Volloch, V.Z., B. Schweitzer, and S. Rits.** 1995. Transcription of the 5'-terminal cap nucleotide by RNA-dependent DNA polymerase: possible involvement in retroviral reverse transcription. *DNA Cell Biol.* 14:991-996.
16. **Wang, P. and R.R. Granados.** 1997. Molecular cloning and sequencing of a novel invertebrate intestinal mucin cDNA. *J. Biol. Chem.* 272:16663-16669.
17. **Zhu, Y., A. Chenchik, and P.D. Siebert.** 1996. Synthesis of high-quality cDNA from nanograms of total polyA+ RNA with the CapFinder PCR cDNA library construction kit. *CLONTECHniques* 11:12-13.

This work is supported by an NSERC Strategic Projects grant no. 215742-98 to S.M.H. and D.D.H. We are grateful to M. O'Grady and S. Shah, who did the fundamental work on mucin. Technical assistance and advice from Dr. Janet Hill and Brenda Haug is also appreciated. Sequencing and oligonucleotide synthesis were performed by the DNA Technology Unit, NRCC-PBI, Saskatoon. Address correspondence to Dr. Dwayne H. Hegedus, Molecular Plant Protection, Saskatoon Research Center, Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, SK, S7N 0X2, Canada. e-mail: hegedusd@em.agr.ca

Received 1 October 2001; accepted 29 November 2001.

X. Shi^{1,2}, T. Karkut^{1,2,3}, M. Chahmanhkah^{1,3}, M. Alting-Mees², S.M. Hemmingsen^{2,3}, and D. Hegedus^{1,3}

¹*Agriculture and Agri-Food Canada*

²*National Research Council*

³*University of Saskatchewan Saskatoon, SK, Canada*

For reprints of this or any other article, contact
Reprints@BioTechniques.com