'Race no more': an alternative approach to cloning the 5' end of transcripts

John H. Weis

Division of Cell Biology and Immunology, Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT 84132, USA

Received June 8, 1994; Revised and Accepted July 24, 1994

The identification and isolation of the sequences at the 5' termini of transcripts can be a frustrating and time consuming experience. Although there are a variety of RACE (rapid amplification of cDNA ends) protocols (1,2) and commercial kits which can all be used with varying degrees of success, the cloning of these sequences still remains a troublesome endeavor. During the full cloning of a novel integrin subunit, the murine α M290 gene (3), it became clear that all of the cDNA clones isolated from a standard cDNA library lacked the final expected 300–400 bp of the mature transcript. In order to isolate these sequences a different PCR approach, compared to the standard RACE technique, was taken.

In our experience faster cycle times, 25 to 30 sec per cycle (4), have been more consistent and efficient in the amplification of rare transcripts than slower cycle times. Because of this greater productivity a fast cycle protocol was utilized to determine if a linear amplification would produce enough product to directly isolate the desired sequences. Accordingly, random primed cDNA was amplified using our previously described RT-PCR protocol (5,6) using only a single 21b oligonucleotide (325) (5'-TCACTCCCTGG TACCTGCCC-3') which, via primer extension, was anticipated to produce a 300–400 bp fragment (see Figure 1). Two identical reactions with high stringency annealing temperature were amplified for either 25 or 50 cycles (59°C-1 sec, 72°C-5 sec, 95°C-1 sec). With linear amplification, any band present in the 25 cycle lane should be doubled in intensity in the 50 cycle lane. However, after gel electrophoresis analysis of the resultant products, no bands were evident in the 25 cycle lane but a broad band of the appropriate size was found in the 50 cycle lane (data not shown). (Decreasing the annealing temperature simply created a more homogenous smear within the lane suggesting inappropriate priming.) This band was cut and eluted, DNA recovered and reamplified, for only 25 cycles, with either the 325 oligo alone, or in combination with the 288 oligo (5'-TGGAGC TGCCTCTGGCTACCA-3') (Figure 1, panel A). The broad band, band a, was again evident with 325 alone, no product with 288 alone (which was expected since only 25 cycles were utilized), but a variety of products with the #288/#325 set. The rapid amplification of these fragments indicated that the reactions were exponential, not linear, suggesting the 325 oligo had somehow become incorporated on both ends of the primary (band a) cDNA product. Reasoning that the longest 325/288 fragment would represent the desired product, band b was cut, eluted, and re-amplified (Figure 1, panel B). The b fragment was apparently contaminated with a small amount of the 325 only fragment (band a), however the bulk of

the fragment preparation clearly possessed a 325 oligo sequence on one end and the 288 sequence on the other (band b). The b fragment was cloned and sequenced. The sequences from this fragment exactly matched those expected for this region of the transcript including 5' untranslated sequences, initiating ATG, and sequences encoding a signal sequence followed by the 288 oligo sequence. Additionally, as expected, the sequence information confirmed that the 325 oligo had been incorporated at the 5' terminus of this fragment.

The use of the 325 oligo alone to isolate the 5' portion of the transcript either suggested that there were two sequences in the gene that were similar to the 325 oligo or that the 325 oligo somehow was incorporated onto the 5' end of the cDNA fragment during the PCR amplification step. The latter model was deemed most likely, however, to prove this would require the generation of the same 5' terminal sequences using a completely unrelated oligo sequence for the gene specific PCR priming. Thus a third oligo, 358 (5'-CAGGTCTGGTTGTTGCTGGGGTC-3'), was used alone in a PCR reaction for 25 and 50 cycles. Again only the 50 cycle run demonstrated an appropriately sized fragment, band c (Figure 2, panel A). That fragment was excised, recovered, and reamplified with 357 (5'-GTACGCGTGCATC-TGCTCCAGCTATG-3') alone, 358 alone, and 357/358 for 25 cycles. As expected, fragment c is regenerated in both of the reactions which contained the 358 oligo. However, the generation of fragment d indicated that the c fragment was derived from the 5' end of the transcript and that its amplification (fragment c) was driven by the 358 oligo being incorporated on both the 3' end and 5' ends, similar to that seen for the 325 oligo in Figure 1. This finding was also confirmed by DNA sequencing (data not shown). Thus this 5' extension protocol is not specific for any oligo sequence.

Figure 3 illustrates our perception of the initial single oligo primed reaction. The ABC oligo can be used by the Taq polymerase to extend the primary strand to the 5' terminus (#1). The enzyme (spotted ball) then remains on this end (#2) until the ABC oligo in solution lines up in the correct orientation to the primary strand (#3). The enzyme then copies the ABC oligo to provide the complementary sequence on the 5' end of the extending strand (#4). Following denaturation the ABC oligo then can be used to prime of the complementary sequence and exponential amplification ensues (#5). An alternative scenario is that the ABC oligo ligates to the primary strand via a contaminating ligase or the Taq polymerase itself. However, the efficiency of the reaction was not been increased by adding exogenous ligase (data not shown). Since 50 cycles are required

Figure 1. Amplification of 5' sequence of the α M290 transcript. Total RNA obtained from TGF- β induced murine T cells was reversed transcribed and amplified using the RT-PCR protocol previously described (5,6). Briefly reverse transcription was carried out using 5 μ g of total RNA in a 50 μ l reaction mixture containing 1 $\times$ RT buffer (Gibco-BRL, Gaithersburg, MD), 10 mM DTT (Gibco-BRL), 0.125 mM each of the four deoxyribonucleotide triphosphates, 0.5 μ g of random primers (New England Biolabs, Beverly, MA), and 400 units M-MLV Reverse Transcriptase (Gibco-BRL). After 60 minutes incubation at 37°C, 2 μ g of DNase-free RNase was added and incubated for 5 minutes at 37°C. Reaction volume was adjusted to 270 μ l to 0.4 M NaCl, followed by a phenol/chloroform, chloroform extraction, and precipitated with ethanol at -20°C overnight. PCR was performed with adaptations for rapid cycling with the air thermocycler (4). Each 10 μ l reaction contained 200 ng cDNA, 70 pmol of primer, 1 $\times$ reaction buffer (10 $\times$ stock consisted of 500 mM Tris pH 8.3, 30 mM MgCl₂, 200 mM KCl, and 5 mg/ml BSA), 0.8 mM dNTPs, 2.5 μ Ci ³²P-dCTP (3000 Ci/mmol: New England Nuclear, Boston, MA), and 0.72 units AmpliTaq DNA Polymerase (5 U/ μ l, Cetus, diluted 1:7). Each 10 μ l reaction was added to glass microcapillary tubes (# 1705, Idaho Technology, Idaho Falls, ID), the ends sealed, and the tubes inserted into the 1605 Air Thermocycler (Idaho Technology). Following PCR, the ends of the microcapillary tubes were scored and samples were removed with a Captrol microaspirator. All 10 μ l of sample was added to an equal volume of stop solution, heated to 95°C for 5 minutes, and electrophoresed in a 6% Acrylamide gel. Radiographic bands were detected by autoradiographic exposure for variable times at -70°C. Sizes of bands were estimated by the migration of ³²P end labeled *Msp*I digest of pBR322. A broad band of approximately 300 bp was generated after 50 cycles of amplification with the total T cell cDNA and oligonucleotide 325 alone using conditions described above. The band was excised and eluted from the gel, re-amplified (A) using either oligonucleotides 325 or 288 alone, or together (RPCR conditions: 59°C-1 sec, 72°C-5 sec, 95°C-1 sec; 25 cycles), and the resulting products resolved by denaturing gel electrophoresis. The band b identified in panel A was recovered and re-amplified (B) with oligonucleotides 325 or 288 alone or together (RPCR conditions: 59°C-1 sec, 72°C-5 sec, 95°C-1 sec; 25 cycles). The line drawing indicates the positions of the primer sequences within the α M290 transcript.

to see the expected product (while in a normal RT-PCR reaction the transcript can be identified in 25 cycles) this must represent a uncommon event. However, given the greater sensitivity provided for by rapid cycling, even a rare event can be amplified to the point of cloning.

This experimental protocol was developed in order to isolate the 5' region of the transcript encoding the murine α M290 protein. The α M290 gene mRNA is found in relatively rare quantities and as such certainly does not qualify as an abundant transcript (3). The fact that two different oligonucleotide sequences could be used to isolate the 5' end of the α M290 transcript suggests this protocol should be useful in the isolation of similar sequences from both rare and abundant transcripts.

Figure 2. Single primer amplification is not oligonucleotide specific. A. The T cell cDNA described in Figure 1 was amplified with oligonucleotide 358 alone for 50 cycles (RPCR conditions: 59°C-1 sec, 72°C-5 sec, 95°C-1 sec). The predominant product, band c, was excised and eluted, and re-amplified with oligonucleotides 357 or 358 alone, or together (B) (RPCR conditions: 59°C-1 sec, 72°C-5 sec, 95°C-1 sec; 25 cycles). The line drawing indicates the positions of the primer sequences within the α M290 transcript.

Figure 3. Schematic representation of the proposed single oligonucleotide amplification process.

ACKNOWLEDGEMENTS

This work was supported by NIH grants AI 24158 and AI 32958 with additional assistance from CA 42014. The author is an Established Investigator of the American Heart Association. Technical assistance was provided by Sue Natali.

REFERENCES

1. Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.
2. Loh, E.Y., et al. (1989) *Science* **243**, 217-220.
3. Smith, T.J., et al. (1994) *Immunity* In press.
4. Wittwer, C.T. and Garling, D.J. (1991) *Biotechniques* **10**, 76-83.
5. Tan, S. and Weis, J.H. (1992) *PCR Methods and Applications*, **2**, 137-143.
6. Weis, J.H., et al. (1992) *Trends Genet.* **8**, 263-264.