Nucleic Acids Research, 1993, Vol. 21, No. 16 3915–3916

Sequencing of cDNA using anchored oligo dT primers

Mark G.Thomas, Sarah A.Hesse, Andrew T.McKie and Farzin Farzaneh
Molecular Medicine Unit, Kings College School of Medicine and Dentistry, The Rayne Institute, 123 Coldharbour Lane, London SE5 9NU, UK

Received May 7, 1993; Revised and Accepted July 6, 1993

Sequencing of cDNAs isolated from phage or plasmid libraries has traditionally been a laborious but reliable process (1). This is because the conventional approach has involved the isolation and purification of insert cDNA followed by its cloning into sequencing vectors such as M13 or Bluescript. Polymerase chain reaction (PCR) mediated direct sequencing avoids many of these time-consuming steps. Using primers complimentary to sequences flanking the cloning sites of various cDNA cloning vectors, the inserts may be amplified and sequenced directly (2). However, a commonly encountered problem with this approach is the difficulty in reading through the poly (A) sequences which are present at the 3′ end of most cDNAs. We have found that after the poly (A/T) sequence, the sequencing ladder often becomes unreadable. We believe this to be due to single base insertions/deletions occuring at high frequency in poly (A/T) stretches during PCR amplification. In order to remedy this, priming of the sequencing reaction would need to occur either on or after the poly (A/T) stretch. Using oligo dT as a primer would be ineffective because oligo dT would prime randomly at different points along the poly (A) tail. Here we describe a simple and highly efficient means of overcoming this problem using oligo dT primers with degenerate 3′ ends as sequencing primers. The presence of the degenerate 3′ position 'fixes' the position of the sequencing primer, thus allowing the direct sequencing of PCR amplified cDNA inserts without the need for further recloning into sequencing vectors. We have employed this approach to sequence cDNA clones isolated from lambda gt10, gt11 or lambda Zap libraries that contain stretches of poly (A/T) at the 3′ end.

In detail, cDNA inserts were amplified directly from isolated clones with vector flanking primers, one of which was biotinylated. The amplification product was bound to streptavidin-coated magnetic particles (Dynabeads, Dynal Ltd) and the two strands separated by denaturation with NaOH (3). Both strands were then sequenced separately using flanking primers as sequencing primers in order to gain information about both ends of the cDNA insert. The presence of poly (A/T) stretches commonly results in the generation of typical ladders shown in Figure 1; lanes A, F and G. In this situation, oligo dT primers with 3 and 4 nucleotide degeneracy at the penultimate and ultimate positions of the 3′ end respectiveley, were used as sequencing primers. Using this approach we were able to obtain clearly readable sequence starting a few bases away from the poly (A) tail (Figure 1; lanes B and C).

The 3′ degenerate oligo dT primer mix [$^{5'}dT_{(18)}$(G/A/C)(A/G/C/T)$^{3'}$] was made by synthesising the following four oligonucleotides which were then mixed in an equimolar ratio:

$^{5'}dT_{(18)}$(G/A/C)A$^{3'}$
$^{5'}dT_{(18)}$(G/A/C)G$^{3'}$
$^{5'}dT_{(18)}$(G/A/C)C$^{3'}$
$^{5'}dT_{(18)}$(G/A/C)T$^{3'}$

It is necessary to make four primers since it is impossible to have degeneracy at the 3′ using normal oligonucleotide synthesis protocols. An alternative and simpler method for the synthesis of the combined mixture of the primers described above is to make a single oligonucleotide incorporating TOPS phosphoramidite (Cambridge Research Biochemicals, Northwich, England) at the penultimate 3′ position, with the last 3′ base being any of the four bases. TOPS phosphoramidite (4) is a bifunctional reagent that may be incorporated at any position during oligonucleotide synthesis. Upon deprotection the TOPS phosphoramidite is both cleaved and cyclised to release an oligonucleotide with a 3′ hydroxyl terminal. Degeneracy can thus be introduced at the two positions directly 5′ to the TOPS phosphoramidite. The loss of the TOPS reagent during deprotection would result in the generation of the same oligonucleotide mixture as above.

The primer mixture was diluted to stock solutions of 5 ng/μl or 20 ng/μl; and 1 μl of each dilution used for each sequencing reaction (Figure 1; lanes 2 and 3). Primer annealing and sequencing was carried out using the standard sequenase method (5).

This effect is not specific to the biotinylated template (compare lane A with F and G), nor is it due to secondary structures as shown by the persistence of the problem despite the inclusion of dITP (see lanes F and G). Comparison of lanes A with B and C demonstrate that this simple procedure allows rapid and direct PCR mediated sequencing of the 3′ ends of cDNAs containing poly (A/T) stretches.

REFERENCES

1. Sanger,F., Niklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
2. Mullis,K.B. and Faloona,F. (1987) *Methods Enzymol.* **155**, 335.
3. Hultman,T., Stahl,S., Hornes,E. and Uhlen,M. (1989) *Nucleic Acids Res.* **17**, 4937–4946.
4. Holland,D., Garman,A.J., Edge,H.D. and Mclean,N.J. (1982) International Patent Application, Publication no. W092–06103.
5. Tabor,S. and Richardson,C.C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4767–4771.

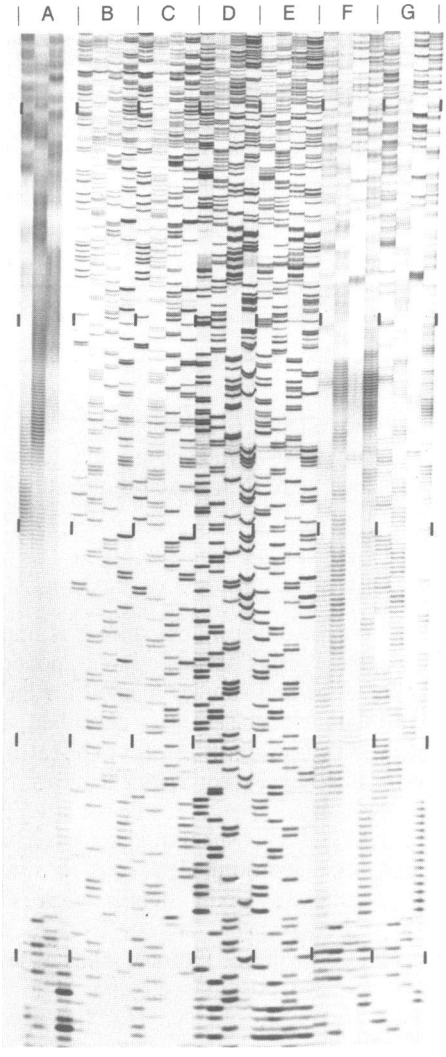

Figure 1. Sequencing gel of samples amplified from cDNA clones selected from a Lambda gt10 human fetal heart library. The cDNA inserts were amplified in 100 μl volumes using 20 μl of phage stock in SM buffer (100 mM NaCl, 5 mM Tris−HCl pH 7.5, 8.1 mM $MgSO_4$, 0.01% gelatin), 0.3 μg of each of the lambda gt10 flanking primers, *gt10for* ($^{5'}$ Biotin-CTTTGAGCAAGTTCAGCC-TGGTTAAG$^{3'}$) and *gt10rev* ($^{5'}$GAGGTGGCTTATGAGTATTTCTTCC-AGGGTA$^{3'}$), a 10×PCR buffer (400 mM KCl, 100 mM Tris−HCl pH8.8 and 1% Triton X-100), designed to be used with SM buffer, and dNTPs to a final concentration of 200 μM. PCR products were precipitated by the addition of 1/3 volume 8 M ammonium acetate and 1 volume isopropanol. Following incubation at room temperature for 10 minutes and centrifugation at 12,000 g, the samples were re-dissolved in 20 μl TE and bound to Dynabeads as described bv Hultman *et al.* (3). The unbiotinylated strands of the amplified cDNA products were recovered by adding 1/10 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ethanol to the NaOH washes. Following incubation at −20°C for 2 hours, the samples were centrifuged at 12,000 g for 10 minutes and the pellets washed twice with 70% ethanol, vacuum-dried and redissolved in 5 μl of water. Primer annealing and sequencing were carried out using the sequenase method (5) and samples were loaded in the gel in the order A, G, C and then T. Lanes A to E show sequences obtained from a lambda gt10 clone of the human ribosomal protein L18 cDNA (Thomas *et al.* manuscript in preparation) using different strands of the cDNA and different primers. Lanes A to C are the biotinylated strand primed with 5 ng of *gt10rev* (A), or either 5 ng (B) or 20 ng (C) of the 3′ degenerate oligo dT primer mix. Lanes D and E are the unbiotinylated (NaOH wash) strand, both primed with *gt10for* using dGTP (D) and dITP (E) sequencing reactions. Lanes F and G show another example of this phenonoma using the unbiotinylated strand of insert cDNA from an uncharacterised lambda gt10 clone. Both were primed with *gt10rev* using dGTP (F) and dITP (G) sequencing reactions.