Novel strategy for synthesis of full-length double-stranded cDNA transcripts without dC-dG tails

Shin-Ichi Fukuoka and George A. Scheele*

Laboratory of Cell and Molecular Biology, Dana Research Institute, DA-554, Thorndike Laboratory and Department of Medicine (Division of Gastroenterology), Harvard Medical School and Beth Israel Hospital, 330 Brookline Avenue, Boston, MA 02215, USA

Submitted June 19, 1991

Previous attempts to obtain full-length double-stranded (ds) cDNA transcripts have relied on deoxy cytosine (dC)-deoxy guanine (dG) tailing methods. In this approach first strand cDNA transcripts are tailed with either dC or dG using terminal deoxynucleotidyl transferase (tdt) and second strand cDNA is primed with oligo-dG or -dC, respectively, and synthesized using DNA polymerase I or Klenow fragment (1). Modifications of this method have used cDNA transcripts tailed with dC and plasmid vectors tailed with dG (2, 3). By use of these tailing methods homopolymeric dC-dG regions (10–30 bp) are inserted upstream of cDNA constructs. We have observed (H. Lütcke and G. A. Scheele, unpublished observations) that such dC-dG regions inhibit (a) *in vitro* transcription of DNA using RNA polymerases (75% inhibition in the case of anionic trypsinogen cDNA), (b) translation of resulting RNA transcripts in wheat germ or reticulocyte lysate protein synthesizing systems (complete inhibition) and (c) sequencing of single stranded (ss) or ds DNA. Synthesis of ds cDNA by self-priming (4) or RNase-H (5) methods avoid the introduction of dC-dG tails but do not preserve the full-length character of cDNA constructs and therefore may not reveal the transcriptional start-site (cap site) or the complete 5' nontranslational sequence.

In the novel method described here, a dC homopolymer tail (~15 bases) is added to first cDNA transcripts using tdt. Synthetic oligo ribo-G (rG, 15 mer, synthesized by Peninsula Laboratories, Belmont, CA) is used to prime the synthesis of full-length second strand cDNA. After completion of synthesis of ds cDNA, rG nucleotides in the RNA-DNA heteroduplex tail are removed with RNase H and the exposed ss dC homopolymer tail is removed with T4 DNA polymerase. The resulting full-length ds cDNA transcripts, devoid of C/G tails, are then inserted into any desirable vector, including expression vectors, using adapter-linker assisted ligation.

To demonstrate the fidelity of our method we employed dog pancreas poly(A)⁺ RNA, which contains three well-defined size classes of high-abundance mRNA (~1.7 kb transcripts encode alpha-amylase [55 kDa] and procarboxypeptidase A and B forms [40–44 kDa], ~1.0 kb transcripts encode serine protease zymogens [trypsinogen, chymotrypsinogen, and proelastase forms, 23–27 kDa], and ~0.6–0.8 kb transcripts encode colipase and phospholipase A₂ forms [10–16 kDa]. Comparison of the abundance of first and second cDNA

transcripts representing each of these three classes of mRNA provides a critical index for efficiency of full-length ds cDNA synthesis. As shown in Figure 1, oligo deoxy-G- and oligo ribo-G-primed synthesis of second strand cDNA (lanes 3 and 4, respectively) were identical. Higher molecular weight bands in lanes 3–5 (e.g. the faint band at 3.2kb) represent dimers resulting from 'snap-back' hybridization during second strand cDNA synthesis. The hair pin nature of these snap-back structures prevent their ligation into functional Lambda phage vectors. Based on incorporation of ³²P-dCTP during synthesis reactions, size classes of second strand transcripts compared favourably with those representing first strand transcripts (not shown). These results may be due to the difficulty in optimizing RNase H digestion for heteroduplex (mRNA-cDNA) transcripts of different lengths. In contrast, enzymatic methods are not available to remove dC-dG tails with precision.

Based on other studies (S.-I. Fukuoka and G. Scheele, unpublished findings), we recommend that deoxy-C tailing and ribo-G priming be employed since (a) oligo ribo-G is insensitive to RNase A digestion, (b) 3' tailing (number of bases added per time, temperature, and concentration of dNTP) may be controlled more easily with dC than dG, and (c) hybridization between oligo-dC and oligo-rG is more stable than that observed with oligo-dG and oligo-rC as judged by insensitivity to digestion with T4 polymerase. Since oligo ribo-G is less stable than oligo deoxy-G, we recommend that aliquots of the synthetic ribonucleotide be stored in sealed vials under nonoxidizing conditions. Once suspended in buffer the ribonucleotide remains active for approximately 1–2 months at –80°C.

We have demonstrated that synthesis of full-length ds cDNA devoid of upstream dC-dG tails improves the *in vitro* expression of cDNA transcripts as judged by both transcription of DNA and translation of RNA. This technique should therefore be the method of choice for the construction of cDNA libraries, particularly those employed for heterologous gene expression (6).

ACKNOWLEDGEMENT

Supported by US Public Health Service research grant DDK-18532.

* To whom correspondence should be addressed

REFERENCES

1. Land, H., Grez, M., Hauser, H., Lindenmaier, W. and Schutz, G. (1981) *Nucleic Acids Res.* **9**, 2251–2266.
2. Okayama, H. and Berg, P. (1982) *Mol. Cell. Biol.* **2**, 161–170.
3. Vieira, J. and Messing, J. (1982) *Gene* **19**, 259–268.
4. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Section 8.14.
5. Gubler, U. and Hoffman, B. J. (1983) *Gene* **25**, 263–269.
6. Fukuoka, S.-I. and Scheele, C. A., Patent filed, January 5, 1991.

Figure 1. Analysis of ss and ds cDNA transcripts separated in alkaline agarose gels. Panel A: Fidelity of second-strand cDNA synthesis in three methods used for construction of cDNA libraries. Lanes 1 and 7, ^{32}P -labeled molecular weight markers (Lambda phage DNA restricted with Hind-III); Lanes 2 and 6, first cDNA strands synthesized with dog pancreas poly(A) $^{+}$ mRNA (100 ng/ μl), AMV reverse transcriptase (1 U/ μl), dNTPs $\times 4$ (0.5 mM each), and tailed with ^{32}P -dCTP in the presence of terminal deoxynucleotidyl transferase (2 U/ μl) (tail length, ~ 15 b); Lane 3, radiolabeled second strand cDNAs synthesized with dC-tailed first-strand cDNA transcripts, priming with synthetic oligo deoxy-G (15-mer, 50 ng/ μl), and synthesis of second cDNA strands with DNA polymerase I (1 U/ μl) in the presence of ^{32}P -dCTP; Lane 4, same as lane 3, except that oligo ribo-G (50 ng/ μl) was used for priming the synthesis of second cDNA strands; Lane 5, radiolabeled second-strand cDNAs using the method of Gubler and Hoffman. Panel B: Removal of dC-rG heteroduplex tails by RNase H and T4 polymerase. Lane 8, ds cDNA transcripts synthesized with ^{32}P -dCTP tailing of first cDNA strands and ribo-G priming for synthesis of second cDNA strands and treated with T4 DNA polymerase (0.25 U/ μl); Lane 9, ds cDNA product, displayed in lane 8, treated with RNase H (0.2 U/ μl) and T4 polymerase (0.25 U/ μl).