

## SHORT COMMUNICATION

## Reverse Transcription-Polymerase Chain Reaction Construction of Plasmid-based, Full-length cDNA Libraries from *Leishmania infantum* for in Vitro Expression Screening

Bernard Couvreur/<sup>+</sup>, Alex Bollen, Dominique Le Ray\*, Jean-Claude Dujardin\*

Laboratory of Applied Genetics, IBMM, Free University of Brussels, Gosselies, Belgium \*Laboratory for Protozoology, Department of Parasitology, Prince Leopold Institute of Tropical Medicine, Antwerp, Belgium

*We describe a streamlined reverse transcription-polymerase chain reaction methodology for constructing full-length cDNA libraries of trypanosomatids on the basis of conserved sequences located at the 5' and 3' ends of trans-spliced mRNAs. The amplified cDNA corresponded to full-length messengers and was amenable to in vitro expression. Fractionated libraries could be rapidly constructed in a plasmid vector by the TA cloning method (Invitrogen). We believe this is useful when there are concerns over the use of restriction enzymes and phage technology as well as in cases where expression of proteins in their native conformation is desired.*

Key words: kinetoplastid - cDNA libraries - spliced leader

All kinetoplastids and euglenids mRNAs and a significant proportion of them in helminths are flanked by known, conserved sequences: the spliced leader at the 5' end and a polyA tail at the 3' end (Bonen 1993, Nilsen 1995). It is therefore possible to amplify by the polymerase chain reaction (PCR), after reverse transcription (RT), only those cDNA molecules that correspond to full-length messengers, i.e. extending from the 5' spliced leader or mini-exon to the polyA<sup>+</sup> tail. PCR amplification can also produce conveniently large quantities from small amounts of ss cDNA. Indeed, researchers have used this concept to prepare libraries or select full-length transcripts in helminths (Devaney et al. 1996, Blaxter et al. 1996, Gregory et al. 1997, Fernandez et al. 2002) and in trypanosomatids (El Sayed et al. 1995, Ajioka et al. 1995, Levick et al. 1996).

Here we describe a streamlined procedure to construct full-length cDNA libraries from small amount of *Leishmania* promastigote mRNA. As compared to published methods, the following changes have been made: (i) a 3' PCR primer colinear to the adaptor part of the modified oligodT primer used for reverse transcription, (ii) the 5' end of PCR primers were chosen to favour addition of non-template A's by the Taq polymerase (Magnusson et al. 1996), (iii) size fractionation of the amplified material, (iv) direct insertion of amplicons into a linearised plasmid vector by the TA cloning method (Invitrogen), (v) in vitro expression of single or pooled clones in a combined (one-tube) eukaryotic transcription/translation system.

This has several advantages. First, the 3' PCR primer can be designed to have a melting temperature near that of the mini-exon-derived 5' primer so as to increase the annealing temperature thus reducing chances of non-specific priming; non-oligodT 3' primers also avoid spurious priming at internal A-rich stretches (advised by Devaney et al. 1996). Secondly, direct insertion of PCR products into a plasmid vector decreases the amount of work and the risks of losing interesting clones entailed by restriction-and-ligation approaches and by  $\lambda$  phage derivatives technology. Thirdly, fractionation of the material is possible either before or after the RT-PCR step. Finally, given that full-length cDNA's possess signals necessary for translation, such libraries are amenable to screening by in vitro expression of pooled plasmid DNA. It is therefore possible to obtain rapidly the full-length sequence (as well as a small amount of the recombinant protein in the native conformation) of proteins for which a specific ligand is available. Finally, the methodology could be particularly relevant when searching for protective antigens by plasmid DNA immunization and challenge in a process called "expression library immunisation" (Barry et al. 1995).

*RT of polyA<sup>+</sup> RNA and PCR amplification* - Poly A<sup>+</sup> RNA was extracted from promastigotes of *Leishmania infantum* by standard methodology (Sambrook et al. 1989). One hundred ng was reverse transcribed using the Superscript II enzyme (Life Technologies) and 100 pmole NotI-Oligod(T)18 (Pharmacia) as primer plus 40 units Human Placenta Rnase Inhibitor (Roche Molecular Biochemicals) in 20  $\mu$ l final volume, followed by digestion of the template RNA with Rnase H (Life Technologies). Two  $\mu$ l of reverse transcription product was amplified by PCR on a DNA thermal cycler 480 (Perkin Elmer) using the Expand High Fidelity PCR System (Roche Molecular Biochemicals).

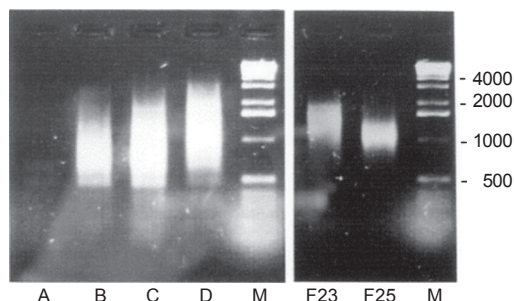
The 5' primer (5'-AACTAACGCTATATAAGT ATCAGTTT-3') was colinear to the 5' end of the mini-exon and the 3' primer (5'-AGAATTCGCGGCCGC

<sup>+</sup>Corresponding author. Present address: Veterinary and Agrochemical Research Centre, Groeselenberg 99, B-1180 Bruxelles, Belgium. Fax: +32-2-379.0670. E-mail: becou@var.fgov.be  
Received 18 September 2002  
Accepted 14 March 2003

AGGAAT-3') was colinear to the arbitrary part of the NotI-Oligod(T)18 used for reverse transcription. Primers were used at 50 pmole (400 nM) along with 0.5 mM each dNTP's and 1.5 mM MgCl<sub>2</sub> in 50 µl of final volume. The following PCR cycling conditions were applied to ss cDNA: 16 cycles of (94°C X 1 min, 60°C X 1 min, 68°C X 10 min) with elongation time increased 10 sec per cycle starting at cycle 11. Cycling was preceded by a 2-min denaturation at 94°C and terminated with a final elongation of 7 min at 72°C. These conditions yielded ~ 6 µg ds cDNA per reaction, ranging from 700 bp to over 4000 bp.

Elongation time and the number of cycles appeared essential. Increasing elongation times from 1 min to 5 and 10 min and decreasing the number of cycles from 25 to 16 increased the amount of material produced as well as the size distribution of amplicons (Figure, left panel). Finally, material from "mock" reverse transcription reactions, performed in the absence of polyA+ RNA or in the absence of primers, did not yield any amplification products (not shown).

*Size fractionation of RNA and amplified cDNA* - Two methods were used in parallel for enriching libraries with fragments of a given size. Firstly, polyA+ RNA was fractionated by isokinetic centrifugation on a 10-to-30% sucrose gradient after Ricca et al. (1981). Fractions N° 23 (F23) and N° 25 (F25) (out of 38) contained RNA ranging from ~ 800 bp to > 1900 bp and 750 to < 1900 bp, respectively. They coded for proteins in the 20 to 35 kDa range mainly. F23 and F25 were reverse transcribed and submitted to our PCR (Figure, right panel). As expected, material amplified from the fractions showed a restricted size distribution as compared to material from unfractionated RNA and the average size distribution accommodated the estimated sizes of the messengers (F23: 800-2000 bp vs F25: 800-1400 bp). Second, 40 µl of PCR reaction product from unfractionated ss cDNA (~ 4 µg amplified ds cDNA) was loaded onto a 6.5% acrylamide gel in 89 mmol l<sup>-1</sup> Tris-Borate, 2 mol l<sup>-1</sup> EDTA pH 8.0 buffer cast in a miniprotean apparatus (Biorad). The electrophoresis was conducted at 100 V for 135 min. The system was then disassembled



Polymerase chain reaction amplification of full-length cDNA's. Left panel: four sets of amplification conditions were tested on single-stranded cDNA from unfractionated polyA+ RNA. A: 30 cycles of (94°C X 1 min, 60°C X 1 min, 72°C X 1 min); B: 25 cycles of (94°C X 1 min, 60°C X 1 min, 68°C X 5 min) with elongation time increased 10 sec per cycle starting at cycle 11; C: 25 cycles as in B but with 10 min elongation time; D: 16 cycles as in C. Right panel: set D conditions were applied to ss cDNA obtained from fractions F23 and F25 of sucrose gradient size-fractionated polyA+ RNA.

and the size markers lanes cut off the gel for ethidium bromide staining. From the unstained sample lane, a piece of gel corresponding to 1000-1500 bp was cut and the DNA material was recovered in 1.5 ml 10 mM Tris buffer pH 8.0 by isotachopheresis after Ofverstedt et al. (1984). The material was then concentrated on a centricon-30 miniconcentrator (Millipore). A total of 250 ng cDNA was recovered, it ranged from 900 to 1600, bp the modal size being ~1300 bp (not shown).

*Insertion of amplified cDNA into pCRII plasmid vector* - Twenty ng of size fractionated or unfractionated amplified cDNA was inserted into the pCRII vector by the TA cloning method (Invitrogen). Competent bacteria were transformed with 2 µl ligation mixture and spread onto Luria-Bertani agar plates containing 150 µg/ml ampicillin and 80 µg/ml Xgal (Duchefa). From all libraries, over 95% white colonies were recovered, of which 90 to 100% actually contained inserted DNA. On average cloning efficiencies of 1.4X10<sup>5</sup> were observed: a typical ligation experiment (i.e. 50 ng vector + 20 ng amplified ds cDNA) yielding a total of 10<sup>4</sup> colonies. The average size of the inserts was estimated roughly in each library. In the unfractionated one, it ranged from 500 to 800 bp. The library constructed from PAGE-fractionated ds cDNA had inserts in the 800-1500 bp, consistent with the size distribution of the input material. Inserts in libraries constructed from sucrose-gradient fractionated polyA+ RNA ranged from 700 to 1700 bp (F23) and 700 to 1200 (F25).

*Evaluation of the libraries by sequencing and in vitro expression* - Inserts of a few clones from the three libraries were partially sequenced at both ends. In all cases the mini-exon sequence at one end and a polyA stretch followed by the adaptor primer sequence at the other end were found, indicating that full-length mRNAs had indeed been amplified and cloned. A few of them (mostly short ones for rapidity) were then fully sequenced and compared to databases: some could be identified on the basis of homology with previously described genes or the presence of typical motifs on the putative peptide sequences, e.g. proteasome subunit iota with A-type subunit signature, Eif-5 with hypusine signature, ribosomal proteins (Table). Long polypyrimidine stretches in untranslated regions, a feature of kinetoplastid mRNA's, were also detected.

RNase-free plasmid DNA was then extracted from individual cultures or from pools of 12 to 48 micro-wells cultures using the Wizard miniprep method (Promega). Plasmid DNA's were submitted to in vitro expression in 25 µl reticulocyte lysate of the TNT coupled SP6 transcription-translation system (Promega) in the presence of 20 units Human Placenta RNase Inhibitor (Roche Molecular Biochemicals) and <sup>35</sup>S-methionine (Amersham). A strong radioactive band was observed for every plasmid with the insert in the direct orientation vis-à-vis the SP6 promoter. No expression was observed from inserts in the reverse orientation or from control plasmid devoid of insert. This is consistent with the observation that, in all sequences examined, nucleotides surrounding the first methionine codon conformed to the consensus for efficient initiation of eukaryotic translation (Kozak 1989). Also, the sizes of expression products were consistent

TABLE  
Sequence analysis of clones from the full-length cDNA libraries

Clone	Accession number	Insert size (bp), orientation	Homology and motifs	Calc MW (Da)	Obs MW (kDa)
89 <sup>U</sup>	<b>Y13060</b>	2000, R	Proteasome subunit Iota, possesses alpha-type signature	27220	29
31 <sup>P</sup>	<b>AJ278538</b>	1300, D	Possesses prenylation motif	21748	26
32 <sup>P</sup>	5' end: AJ297399 3' end: AJ297400	1600, R	NS NS		ND
33 <sup>P</sup>	5' end: AJ297401	1500, D	NS		ND
313 <sup>P</sup>	5' end: AJ278628 3' end: AJ297398	1900, R	Cytochr. C oxydase subunit 10 NS		ND
41 <sup>P</sup>	<b>AJ404610</b>	1300, D	Nascent polypeptide associated complex: alpha subunit.	18138	27.5
46 <sup>P</sup>	<b>AJ278539</b>	600, D	Eukaryotic initiation factor 5a, possesses hypusine signature	17880	24.5
54 <sup>P</sup>	<b>AJ278540</b>	475, D	Ribosomal protein S33	9746	15
65 <sup>P</sup>	<b>AJ278541</b>	600, D	Ribosomal protein S25	10870	18
34B1 <sup>P</sup>	<b>AJ278542</b>	1200, D	Ribosomal protein L10 aka QM protein homologue	24585	25.5
34B6 <sup>P</sup>	5' end: AJ276157 3' end: AJ276158	1350, D	NAC, alpha subunit NS		ND
34B8 <sup>P</sup>	5' end: AJ276159 3' end: AJ276160	1250, D	NS NS		ND
106 <sup>F23</sup>	5' end: AJ276151 3' end: AJ276152	1600, D	NS NS		ND
109 <sup>F23</sup>	5' end: AJ276153 3' end: AJ276154	880, D	Ribosomal protein L18		ND
91 <sup>F25</sup>	5' end: AJ276145 3' end: AJ276146	1100, D	Ribosomal protein L7 NS		ND
95 <sup>F25</sup>	5' end: AJ276147 3' end: AJ276148	1250, R	Ribosomal protein L2		ND
910 <sup>F25</sup>	5' end: AJ276149 3' end: AJ276150	1200, D	NS		ND
91E2 <sup>F25</sup>	5' end: AJ276155 3' end: AJ276156	1600, D	NS NS		19
91E4 <sup>F25</sup>	<b>AJ276107</b>	800, D	Ribosomal protein L15	24436	25

U: from unfractionated polyA+ RNA; P: from Page-fractionated cDNA; F23 and F25, from sucrose gradient RNA fraction no. 23 and no. 25 respectively; D: direct; R: reverse; NS: no significant homology (blast score > 0.01); Calc: calculated; Obs: observed; ND: not determined

with the calculated size from the corresponding *ORF*'s (Table) The overestimation of relative molecular weights of lower size proteins (< 14 kDa) is a common artefact of SDS/PAGE (Hames & Rickwood 1990).

Finally, all clones possessed the native stop codon and possible chimerical or spurious protein starting at plasmid-encoded ATG's upstream of the insert would be aborted in all reading frames by the presence of stop codons (TAA's) at position -1, +4 and +14 relative to the first nucleotide of the mini-exon.

In conclusion, using a streamlined RT-PCR approach, µg quantities of ds cDNA ranging from 800 bp to over 4000 bp could be produced in a single reaction from one-tenth of the product of a typical reverse transcription re-

action. The material could be fractionated by size and usable amounts were recovered. The cDNA was readily inserted into a plasmid vector by the TA cloning method (Invitrogen) and the ligation bias towards smaller fragments was corrected by size selection of the inserted material. Finally, sequencing and in vitro expression experiments indicated that the cloned cDNA's indeed corresponded to functional full-length messengers.

#### REFERENCES

- Ajioka JW, Melville SE, Coulson RMR, Espinoza JR, Khan SM, Wan KL 1995. Techniques associated with protozoan parasite genome analysis. In KW Adolph, *Methods in Molecular Genetics*, Vol. 6, Academic Press, London, p. 30-47.  
Barry MA, Lai WC, Johnston SA 1995. Protection against

- mycoplasma infection using expression-library immunization. *Nature* 377: 632-635.
- Blaxter ML, Raghavan N, Ghosh I, Guiliano D, Lu W, Williams SA, Slatko B, Scott AL 1996. Genes expressed in *Brugia malayi* infective third stage larvae. *Mol Biochem Parasitol* 77: 77-93.
- Bonen L 1993. Trans-splicing of pre-mRNA in plants, animals, and protists. *FASEB J* 7: 40-46.
- Devaney E, Martin SAM, Thompson FJ 1996. Stage-specific gene expression in lymphatic filarial nematodes. *Parasitol Today* 12: 418-424.
- El-Sayed NMA, Alarcon CM, Beck JC, Sheffield VC, Donelson J 1995. cDNA expressed sequence tags of *Trypanosoma brucei rhodesiense* provide new insights into the biology of the parasite. *Mol Biochem Parasitol* 73: 75-90.
- Fernandez C, Gregory WF, Loke P, Maizels RM 2002. Full-length-enriched cDNA libraries from *Echinococcus granulosus* contain separate populations of oligo-capped and trans-spliced transcripts and a high level of predicted signal peptide sequences. *Mol Biochem Parasitol* 122: 171-180.
- Gregory WF, Blaxter ML, Maizels RM 1997. Differentially expressed, abundant trans-spliced cDNAs from larval *Brugia malayi*. *Mol Biochem Parasitol* 87: 85-95.
- Hames BD, Rickwood D 1990. *Gel Electrophoresis of Proteins: a Practical Approach*, 2nd ed., Oxford University Press, New York.
- Kozak M 1989. The scanning model for translation: an update. *J Cell Biol* 108: 229.
- Levick MP, Blackwell J, Connor V, Coulson RMR, Miles A, Smith HE, Wan KL, Ajioka JW 1996. An expressed sequence tag analysis of a full-length, spliced-leader cDNA library from *Leishmania major* promastigotes. *Mol Biochem Parasitol* 76: 345-348.
- Magnusson VL, Ally DS, Nylund SJ, Karanjawala ZE, Rayman JB, Knap JI, Lowe AL, Ghosh S, Collins FS 1996. Substrate nucleotide-determined non-templated addition of adenine by *Taq* DNA polymerase: implications for PCR-based genotyping and cloning. *Biotechniques* 21: 700-709.
- Nilsen TW 1995. Transplicing: an update. *Mol Biochem Parasitol* 73: 1-6.
- Ofverstedt LG, Hammarstrom K, Balgobin N, Hjerten S, Petterson U, Chattopadhyaya J 1984. Rapid and quantitative recovery of DNA fragments from gels by displacement electrophoresis (isotachopheresis). *Biochem Biophys Acta* 782: 120-126.
- Ricca GA, Hamilton RW, McLean JW, Conn A, Kalinyak JE, Taylor JM 1981. Rat alpha 1-acid glycoprotein mRNA. Cloning of double-stranded cDNA and kinetics of induction of mRNA levels following acute inflammation. *J Biol Chem* 256: 10362-10368.
- Sambrook J, Fritsch EF, Maniatis T 1989. *Molecular Cloning: a Laboratory Manual* 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.