Cloning of the shark Po promoter using a genomic walking technique based on the polymerase chain reaction

L.Fors, R.A.Saavedra*⁺ and L.Hood

Biology Division (147 – 75), California Institute of Technology, Pasadena, CA 91125, USA

Received September 20, 1989; Revised and Accepted February 8, 1990

ABSTRACT

We have cloned the putative shark promoter of protein zero (Po) using a novel application of the ligation mediated single-sided polymerase chain reaction (PCR) method. This method uses linker ligation and subsequent amplifications with a linker primer and multiple specific primers to generate specificity. The method allowed us to amplify approximately 305 base pairs of shark genomic DNA sequence immediately upstream from the 5' end of our full-length Po cDNA. The Po promoter was shown to be directly linked to its first exon, contain a transcription initiation start site and sequences commonly found in eukaryotic promoters. This genomic walking technique will be useful for cloning promoters, insertion sites, and other sequences of interest without the need for constructing and screening genomic libraries.

INTRODUCTION

Myelin of the vertebrate nervous system is a dynamic structure that provides electric insulation and facilitates the rapid conduction of nerve impulses along axons. The major myelin proteins in the mammalian central nervous system (CNS) are proteolipid protein (PLP) and myelin basic protein (MBP); whereas, in the peripheral nervous system (PNS) they are protein zero (Po) and MBP. Sharks diverged from other vertebrates about 400 million years ago and are the most primitive living vertebrates that show myelin organized as a concentric multilamellar structure¹. We have previously isolated and sequenced the major CNS myelin proteins and their corresponding cDNAs from the shark *Heterodontus francisci*¹. The sequence data showed that the major proteins of shark CNS myelin are similar to mammalian PNS proteins Po and MBP. This differential distribution of myelin proteins between sharks and mammals, the fact that Po expression is restricted to the mammalian PNS and the unconventional promoter sequence found in mammalian Po¹⁻² led us to clone the shark Po promoter.

Our inability to clone this promoter by conventional techniques combined with the large size of the shark genome (about two times the size of the mammalian genome³) and the time and cost

associated with constructing and screening of genomic libraries motivated us to find an alternative strategy to clone the shark Po promoter. The polymerase chain reaction (PCR)⁴ allows the amplification of a region of DNA between two sequence-specific oligonucleotides used as primers, and appropriately modified could provide the needed alternative. The PCR technique takes advantage of a DNA-polymerase from *Thermus aquaticus* (Taq DNA-polymerase), which is functional at high temperatures. This thermostability permits one to denature the template, hybridize primers and carry out the polymerization reaction at high temperatures. The high reaction temperature makes possible artifacts due to spurious hybridization less likely. Conventional PCR requires one to know the sequences complementary to both oligonucleotide primers. A modification called 'inverted PCR' (IPCR) allows the amplification of unknown sequences located outside of the known DNA region⁵. This procedure involves cutting the DNA with restriction endonucleases and ligating the resulting free ends under conditions that favor circularization of the DNA. As a result of this manipulation, oligonucleotides complementary to the 5' and 3' ends of the known sequences can now be used to amplify the unknown sequence of interest located at the 5' and/or 3' ends.

Recently a ligation mediated single-sided PCR strategy has been developed for *in vivo* footprinting⁶. This procedure randomly cleaves DNA with dimethyl sulfate/piperidine, uses a 5' specific primer in an extension reaction to synthesize the sequence of interest, attaches a linker at the newly created 3' blunt end and amplifies that region through 16 cycles of PCR using a linker primer and a second specific 5' oligonucleotide.

In this article, we describe an adaptation of the ligation mediated single-sided PCR method that allowed us to amplify and clone the putative shark Po promoter.

MATERIALS AND METHODS

Oligonucleotides A set of four oligonucleotides was synthesized in a 380B automatic DNA synthesizer (Applied Biosystems Inc., Foster City, CA), using the chemistry of Caruthers and coworkers (1981)⁷⁻⁸. Three of these oligonucleotides, the specific primers, were used for the required extensions and amplification reactions.

* To whom correspondence should be addressed

⁺ Present address: Laboratory for The Study of Skeletal Disorders, The Children's Hospital and Harvard Medical School, 300 Longwood Avenue, Boston, MA 02115, USA

The specific primers were: a 20-mer (specific primer Po1; 5'-TAAGAAGTTGAGTTACTCCC-3'), a 31-mer (specific primer Po2; 5'-AGGAATGGCTCTGACGGCTGGCACTGTTTAA-3'), and a 25-mer (specific primer Po3; 5'-CACTGTTT-AAGCTGCACCTCTCTCT-3'). The fourth oligonucleotide (probe) was synthesized from the 5' most sequence of exon I and used to identify the specific amplification products at each stage of the procedure. This probe was a 26-mer (Po6; 5'-GTCT-CTCTCCTTTACCAGCGCCGAA-3').

Restriction enzyme cleavage Fifteen μ g of liver DNA from the shark *Heterodontus francisci* were cleaved with a mixture of 40 units each of BglII, EcoRI, HindIII for two hours at 37°C. The DNA mixture was then placed at 68°C for 10 min, extracted with phenol/chloroform/isoamyl alcohol (50:48:2) mix, ethanol precipitated and washed. The resulting pellet, which contained genomic fragments with protruding or recessed ends and with a mean size of 1 to 2 Kb, was resuspended in 20 μ l of 10 mM Tris, pH 8.0, 0.1 mM EDTA (TE buffer) and used in the primer extension.

Primer extension The extension was done by mixing 3 μ g of restricted shark liver DNA, 0.15 pM of specific primer 1 in a modified SEQUENASE buffer (40 mM Tris, pH 7.7, 50 mM NaCl) to a final volume of 15 μ l. Denaturation at 95°C for 3 min was followed by annealing at $T_m - 5^\circ\text{C}$ (47°C) for 30 min. The mixture was then placed on ice and 7.5 μ l of 20 mM MgCl_2 , 20 mM dithiothreitol (DTT) and 0.1 mM of each deoxynucleotide triphosphate (dNTP) was added followed by 1.5 μ l of freshly diluted (1:4 in ice-cold TE buffer) SEQUENASE (modified T7 DNA-polymerase from United States Biochemical Co., Cleveland, OH). The extension reaction was incubated at 45°C for 5 min. The reaction tube was then transferred to ice and 6 μ l of ice-cold 310 mM Tris, pH 7.7 was added followed by 15 min incubation at 67°C to inactivate the enzyme.

Linker Two complementary oligonucleotides to serve as a linker were synthesized using the procedure described above⁷⁻⁸. One was a 25-mer (5'-GCGGTGACCCGGGAGATCTGAATTC-3') the other was an 11-mer (5'-GAATTCAGATC-3')⁶. These two oligonucleotides were hybridized by boiling and slowly cooling them to room temperature over a period of several hours. The result is a partially double stranded linker with one blunt and one recessed or protruding end.

Ligation reaction The blunt end of the linker was ligated to the blunt end of the specific primer 1 extended DNA. The 25-mer linker oligonucleotide (linker primer) can then be used with specific primer 2 in the subsequent PCR amplification. To the extension reaction 20 μ l of ice-cold mix (117.5 mM MgCl_2 , 43 mM DTT and 125 μ g/ml bovine serum albumin, BSA) and 25 μ l of ice-cold ligation mix were added to yield a final concentration of 50 mM Tris, pH 7.7, 10 mM MgCl_2 , 20 mM DTT, 3 mM ATP, 50 μ g/ml BSA, 100 pM linker and 3 units of T4 DNA-ligase. The ligation reaction was incubated at 22°C for 2 hr followed by 10 min at 70°C to inactivate the enzyme. The ligation reaction was precipitated after adding 0.1 volume sodium acetate, 10 μ g tRNA and 2.5 volumes of 100% ethanol, washed with 75% ethanol and resuspended in 50 μ l of water.

Polymerase chain reactions PCR was performed in a Cetus-Perkins Thermocycler using a standard protocol (Gene AmpTM, Perkin-Elmer Cetus Instruments catalogue, Norwalk, CT). Each reaction mixture (a total volume of 100 μ l) contained the reaction

buffer (100 mM Tris-HCl, pH 8.3 at room temperature, 500 mM KCl, 15 mM MgCl_2 and 0.1% gelatin), 200 μ M each dNTP, template DNA (for the first amplification: 1 μ g total shark DNA, which had been restricted, extended and linker ligated; for the second amplification: approximately 1 ng obtained from a region excised from a low melting point agarose, LMP-agarose, gel), 825 ng of each primer, and 2.5 units of Taq DNA-polymerase (Perkin-Elmer Cetus Instruments, Norwalk, CT). The samples were denatured for 7 min at 94°C, followed by 25 or 30 amplification cycles (1 min denaturation at 94°C, 2 min annealing at 66°C, and 3 min extension at 76°C).

Agarose gel electrophoresis and Southern blots The PCR products were analyzed by agarose gel electrophoresis and transferred onto Zeta-Probe Blotting membranes (Bio-Rad Laboratories, Richmond, CA) as described by Southern (1977)⁹. The membranes were hybridized for 18 hr in 5 $\times$ saline sodium citrate (SSC), 10 mM EDTA, 1% sodium dodecyl sulfate (SDS) with the oligonucleotide probe at $T_m - 10^\circ\text{C}$ (62°C). The membranes were washed in 5 $\times$ SSC and 0.1% SDS, two times for 10 min at 22°C, and two times for 30 min at 37°C, and exposed for 18 hr to identify the bands of interest. When indicated LMP-agarose was used to facilitate the subsequent step of the procedure.

Cloning of amplified sequences It is probable that the amplified products contain incompletely polymerized ends. Therefore, the amplified fragments were first treated with T4 DNA-polymerase in ligation buffer, 10 μ M of each dNTP, and 5% polyethylene glycol for 20 min at 37°C, to create blunt ends, followed by 5 min at 65°C to inactivate the enzyme. The resulting insert was ligated directly into the SmaI restriction sequence of M13mp18 and pUC-18 using a 3:1 molar excess of insert for 18 hr at room temperature and transformed into competent DH-5-alpha cells (Bethesda Research Laboratories, Bethesda, MD).

DNA sequencing DNAs from cloned plasmid preparations or amplified PCR fragments were sequenced using the dideoxynucleotide method of Sanger *et al.* (1977)¹⁰; SEQUENASE) in combination with appropriate oligonucleotides used as primers⁷⁻⁸. The gels were autoradiographed using XR-5 film from Kodak (Rochester, NY).

Linkage of the amplified sequences to exon I A PCR experiment was performed using shark genomic DNA as the template, and an oligonucleotide complementary to the promoter and the 3'-most sequence amplified from exon I as primers. In this experiment, specific primer Po2 and a 28-mer (PoC; 5'-GACCAGTAGGCATGTTGCATCTGCAAAC-3') were used.

Determination of the transcription initiation sites Extension experiments were done using 10 μ g of total shark brain RNA¹¹. The RNA was hybridized with 3 $\times 10^5$ cpm of specific oligonucleotide Po2 (purified by gel electrophoresis after end-labeling with ³²P-gamma-adenosine triphosphate and T4 polynucleotide-kinase), in 10 mM Tris, pH 8.0, 250 mM KCl and 10 mM EDTA for 45 min at 62°C. Samples were allowed to cool to room temperature and 2.5 volumes of 50 mM Tris, pH 8.0, 10 mM MgCl_2 , 5 mM DTT, 325 μ g each dNTP, 10 μ g/ml actinomycin and 8 to 10 units of reverse transcriptase (AMV 007 from Life Science, Plainview, NY) were added. The samples were incubated for 45 min at 42°C. The reactions were stopped by putting them on ice, and 2 volumes of a 400 mM

Figure 1. The chromosomal walking procedure. A. The specific oligonucleotides, the probe and the linker (including restriction sites⁶). B. A schematic diagram of the genome walking procedure.

ammonium acetate, 10 mM EDTA mix were added. The samples were extracted with phenol/chloroform/isoamyl alcohol, ethanol precipitated, resuspended in dye buffer, and loaded on a sequencing gel adjacent to a M13mp18 sequence ladder that served as a size marker.

Sequence similarities The sequence similarities between the shark and mammalian promoters were determined by computing the number of identical base pairs (bp) after manual alignment of the corresponding sequences.

RESULTS

The shark Po cDNA clone that we previously isolated contained a 5' untranslated region, a coding region, a relatively long 3' untranslated region, a putative polyadenylation signal and a

polyadenosine tail similar to its mammalian counterpart¹. RNA blot analysis showed that this cDNA clone was approximately the same size as its respective transcript. Thus, we considered this clone to be a full-length cDNA. Next, we attempted to isolate genomic clones containing the shark Po promoter by screening shark genomic libraries generated by *Sau*III A partial digestions (constructed by us or kindly provided by G. Litman, Florida) using our shark Po cDNA clone as a probe. After constructing three and unsuccessfully screening a total of five shark genomic libraries, we decided to attempt to clone the shark Po promoter using alternative methods. We chose the IPCR⁵ and the single-sided PCR methods⁶. For reasons still unclear to us, we were unable to amplify the region of interest using the IPCR method. Therefore, we focused on modifying the single-sided PCR method developed for *in vivo* footprinting⁶ into a technique for walking from a known region of a genome (i.e. from exon I) toward an unknown region (i.e. the promoter).

Figure 2. Results from specific steps of the procedure. A. Ethidium bromide picture of an agarose gel after the first round of amplification. Lane 1, 1 Kb ladder marker; lane 2, 123 bp ladder marker; lane 3, PCR products (faint smear marked by a bracket symbol) from 25 cycles of amplification with linker primer and specific oligonucleotide Po2. B. Southern blot of the gel shown in A. Lane 1 and 2, 1 Kb and 123 bp ladder markers, respectively (not visible here); lane 3, Po band (at about 400 bp, marked by an arrow head) detected by probe Po6. C. Ethidium bromide picture of an agarose gel after the second round of amplification. Lane 1, 123 bp ladder marker; lane 2, PCR product (a band of about 385 bp, marked by an arrow head) from 25 cycles of amplification with linker primer and specific oligonucleotide Po3.

Specific primers, probes and linker A set of three specific primers was synthesized as described in Materials and Methods (Figure 1A). Specific primer 1 is separated by 10 bases from specific primer 2, which in turn overlaps 10 bases with specific primer 3. An additional oligonucleotide (probe) was also synthesized from the 5' most sequence of exon I. The specific primers and the probe were all complementary to sequences previously identified as the 5' most end of exon I on the shark Po cDNA¹.

The genome walking scheme Figure 1B shows our genome walking scheme. High molecular weight shark liver DNA (greater than 50 kilobases [Kb], as determined by agarose gel electrophoresis; data not shown) was digested with a mix of three

different restriction endonucleases with six bp recognition sites (see Materials and Methods), all of which restrict the DNA leaving recessed or protruding ends. The restricted DNA was then denatured by heating it to 95°C for 2 min and specific primer 1 was allowed to anneal by lowering the temperature to $T_m - 5^\circ\text{C}$ (47°C) for 30 min. SEQUENASE was used to extend for 5 min at 45°C from specific primer 1 up to the nucleotide sequence where the restriction enzymes had cut (unknown up to this step of the procedure). Theoretically, only those fragments which are primed could produce a partially double stranded DNA molecule blunted at the unknown end. A DNA linker with a single blunt end was ligated directionally onto the double stranded blunt end of the extension product using T4 DNA-ligase. This linker has

no 5' phosphates and is staggered to avoid self-ligation and provide directionality. Also, the duplex between the 25-mer and 11-mer is stable at the ligation temperature, but denatures easily during subsequent PCR.

The resulting ligation product was then denatured by heating at 94°C for 1 min and specific primer 2 was allowed to anneal by lowering the temperature to 66°C for 2 min. Specific primer 2 was extended using Taq DNA-polymerase, and a first PCR (25 cycles) was carried out using specific primer 2 and the linker primer (25-mer), both of which have a comparable T_m . The amplification products were electrophoresed on an agarose gel. Ethidium bromide staining of this gel revealed a faint smear instead of a discrete band (Figure 2A). Two possible reasons for this smear are: i) specific primer 1 (a 20-mer) in combination with the 45°C primer extension temperature could not provide sufficient specificity and/or ii) unwanted blunt ends capable of accepting the linker were produced by contaminants in the restriction enzymes used. The DNA was then transferred onto a membrane, hybridized with a ^{32}P -labeled oligonucleotide probe, and autoradiographed (Figure 2B). The probe hybridized with a single band located on the smear. A band of about 400 bp was obtained. The amplification products were then run on a LMP-agarose gel electrophoresis and the regions from 325 to 445 bp was excised, melted and used as template for a new round of PCR (25 cycles). This time, specific primer 3 and the linker primer were used. The product of this second round of amplification was a discrete band of 385 bp (clone Po385) with little or no background (Figure 2C). The identity of this bands was determined by hybridization to a ^{32}P -labeled oligonucleotide probe complementary to the 5' most end of exon I (data not shown).

The sequence of the amplification products The amplification product Po385 was sequenced (see Materials and Methods), and the clone shown to contain three regions: specific primer 3 with the remaining 5' sequence of exon I at the 3' end¹, an unknown sequence representing the putative Po promoter in the middle, and the linker primer sequence at the 5' end. In addition, the 5' most end of the unknown region contained one of the expected restriction endonuclease recognition sites (EcoRI; Figure 3).

The known and the unknown regions of clone Po385 are located adjacent to each other in the shark genome We next tested whether the known and unknown regions of the amplification product are linked in the shark genome, using PCR (30 cycles). In this experiment, an oligonucleotide complementary to the known region and the unknown region, facing each other, served as primers with shark genomic DNA as template. The oligonucleotides used were specific primer Po2 and PoC. The expected size of the amplification product was 315 bp. The observed amplification product was a band of about 315 bp (Figure 4A). This result shows that the known and unknown regions which are adjacent to each other in the clone Po385 are also adjacent to each other in the shark genome.

Determination of the transcription initiation site To determine the transcription initiation site for shark Po, a primer extension experiment was performed using total shark brain RNA, in combination with primer Po2 (see Materials and Methods). The primer was annealed and extended, producing discrete products whose end points identify potential transcription initiation sites. The initiation of transcription for Po lies around the sequence 5'-GCTCGACT-3', where C is the most probable site (see

Figure 3. Nucleotide sequence of the shark Po promoter. In upper-left box (light line), the linker; in bottom-right box (heavy line), the 5' end of shark Po cDNA¹; underlined (light), putative shark Po TATA and CAT sequences; underlined (heavy), a sequence of alternated Py/G (with Z-DNA formation potential); +1 (>), indicates the transcription initiation site.

site (see Figures 3 and 4B). This site was located 15 bp upstream from the 5' end of our original Po cDNA clone¹. These data also support the idea that our original cDNA clone is a full-length copy of its corresponding mRNA.

DISCUSSION

The sequence analysis of clone Po385 identified a previously unknown DNA region which we called the shark Po promoter for the following reasons. First, this DNA sequence was shown to be adjacent to the 5' end of its corresponding cDNA in the shark genome. Second, the previously unknown region contains a transcription initiation site, as determined by primer extension experiments. A weak consensus sequence (5'-PyPyCAPyPy-PyPy-3'; Py = pyrimidine) for the initiation of transcription has been identified in mammalian genes, with transcription starting at A¹². Our proposed initiation site for Po (5'-CTCGACTG-3') is 60% similar to this weak consensus sequence, and initiates at C. Third, the previously unknown region shows other sequence features that are often found in eukaryotic promoters. It shows a TATA box (5'-TTAAAA-3'), a CAT box (5'-CAATTT-3'), and a sequence of 26 alternated Py/G (5'-TGTGCGTGTGTGCG-TGCGTGTGTGTG-3') which could potentially form left handed Z-DNA¹³⁻¹⁴. These sequences are located -26 to -31, -113 to -118 and -250 to -276 bp from the transcription initiation site, respectively. The proposed shark Po TATA and CAT boxes are 67% and 83% similar to their mammalian counterparts², respectively. Additionally, we found two regions which are similar in the shark and mouse promoters, and they are separated from each other by 121 and 123 bp, respectively. These regions are located -78 to -115 and -238 to -262 from the transcription initiation site, and they are 66% and 72% similar between shark and mouse, respectively. We also found a third region (-61 to -85) of similarity in the shark and mouse promoters, which in shark, overlaps by 8 bp with the 3' most end of the downstream most region. The third region is 72% similar between shark and mouse. This positional information and sequence similarity taken against an overall similarity of 40% between the shark and mammalian² promoter sequences suggests that these are functionally important features of true promoters.

The large size (6×10^9 bp) of the shark haploid genome³ could explain why a single copy gene like Po¹ was absent from out

Figure 4. The promoter sequence is linked to the 5' end of the shark Po cDNA¹, and it contains a transcription initiation site. A. Lane 1, 1 Kb ladder marker; lane 2, the 315 bp band from Po using Po2 (complementary to the 5' end of Po cDNA¹) and PoC (complementary to the promoter region; from -214 to -241 in Figure 3) as primers; lane 3, 123 bp ladder marker. B. Lanes 1-4, the A, C, G and T sequencing reactions of M13mp18 using the universal primer to serve as size markers; lane 5, total RNA extended with Po2 (77 bp, marked by a bracket symbol).

genomic libraries. Therefore, the shark provides an excellent system to test this genome walking strategy. A major focus in the design of any amplification strategy is specificity. We have solved this problem by adapting, as follows, the ligation mediated PCR technique of Mueller and Wold (1989)⁶. First, the linker is ligated to the only blunt ended DNA molecule that should be available after the first DNA denaturation and primer extension. Since the DNA was cut with enzymes which produce only protruding or recessed ends, then only those molecules specifically extended should have a single blunt end which can directionally accept the single blunt end of the linker. Second, each of the subsequent amplification steps has been carried out using the linker primer and a new specific primer that is fully contained within the DNA sequence of the product of the previous

step. Third, the specific amplification products from the first round of PCR were significantly enriched by size fractionation on LMP-agarose before the final round of amplification. Fourth, the desired amplification products were properly identified at each step of the procedure by using a specific oligonucleotide probe.

We have shown that the above strategy can be used to walk from a known region towards an unknown region in a large genome. There are several possible technical improvements which could increase its effectiveness. First, the length of the cloned products could be increased by adjusting the number and nature of restriction endonucleases used to digest the genomic DNA. In our experiments the length of the Po promoter region we amplified was limited by the EcoRI site found at its 5' end (see Figure 3). Second, the first specific primer could be longer

(a 25-mer or more), and Taq DNA-polymerase could be used instead of SEQUENASE in order to perform the initial extension step at a temperature higher than 45°C. Third, the length of each cycle or the number of cycles could be altered. Modifications of this kind however, must take into consideration the size limit of the PCR products (now up to about 10 Kb) and the probability of obtaining amplification artifacts. The appropriate manipulation of some or all of these variables should result in a higher yield of products of the desired length.

The advent of new biotechnologies has made the cloning and sequencing of large genomes a manageable enterprise. The large size and sequence organization of eukaryotic genomes makes it possible that some regions of these genomes will not yield to present cloning techniques. It would also be advantageous to clone smaller regions of genomic DNA without having to construct and screen a genomic library. Our genome walking technique is independent of conventional cloning steps, and therefore provides a method of choice to fill in genomic library gaps or to clone any other region of interest.

ACKNOWLEDGEMENTS

We wish to thank Paul Mueller and Barbara Wold for providing us with the linker oligonucleotides, essential theoretical insights and practical suggestions. We also wish to thank Teresa Geffroy for outstanding technical support and Bernard Vernooy for critically reading the manuscript. The oligonucleotides used in this work were synthesized at the Caltech microchemical facilities by Suzanna Horvath and coworkers. This work was supported by grants from the National Institutes of Health and the National Multiple Sclerosis Society (NMSS) to LH. The data contained in this article are a part of LF's thesis, in partial fulfillment for the degree of doctor of philosophy from Caltech; he also acknowledges the support from NIH grant NS14069. RAS was a post-doctoral fellow of the NMSS.

REFERENCES

1. Saavedra, R.A., Fors, L., Aebersold, R.H., Arden, B., Horvath, S., Sanders, J. and Hood, L. (1989) *J. Mol. Evol.* **29**, 149–156.
2. Lemke, G., Lamar, E. and Patterson, J. (1989) *Neuron* **1**, 73–83.
3. Hinegardner, R. (1976) *Comp. Biochem. Physiol.* **55B**, 367–370.
4. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.T., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science* **239**, 487–491.
5. Triglia, T., Peterson, M.G. and Kemp, D.J. (1988) *Nucl. Acids Res.* **16**, 8186.
6. Mueller, P.R. and Wold, B. (1989) *Science* **246**, 780–786.
7. Beaucage, S.L. and Caruthers, M.H. (1981) *Tetrahedron Lett.* **22**, 1859–1862.
8. Horvath, S.J., Firca, J.R., Hunkapiller, T., Hunkapiller, M.W. and Hood, L. (1987) *Meth. Enzymol.* **154**, 314–326.
9. Southern, E. (1977) *J. Mol. Biol.* **98**, 503–518.
10. Sanger, F., Nicklen, S. and Coulson, A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
11. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. Smale, S.T. and Baltimore, D. (1989) *Cell* **57**, 103–113.
13. Rich, A., Nordheim, A. and Wang, A. (1984) *Annu. Rev. Biochem.* **53**, 791–846.
14. Bourbon, H.-M., Prudhomme, M. and Amalric, F. (1988) *Gene* **68**, 73–84.