Organization and alternative splicing of the murine phospholipase D2 gene

Olga E. REDINA and Michael A. FROHMAN¹

Department of Pharmacological Sciences, and The Institute for Cell and Developmental Biology, State University of New York, Stony Brook, New York 11794-8651, U.S.A.

Phospholipase D (PLD) catalyses the hydrolysis of phosphatidylcholine, generating phosphatidic acid and choline. Mammalian PLD activity derives from a family of membrane-associated enzymes that are activated by a wide variety of signal transduction events. cDNA species encoding human, mouse and rat PLD1 and PLD2 have recently been reported. In this study we undertook to determine the organization of the mouse PLD2

gene. We report that the gene spans 17.1 kb and contains 25 exons. Mouse PLD2 is notable for a relatively GC-rich and large 5′ untranslated region. Proximal promoter sequences upstream of the first exon contain several consensus SP1 sequences (GGGCGG) but lack TATA and CAAT boxes. Finally, alternatively spliced cDNA species identified for PLD1 and PLD2 are discussed in the context of the PLD2 genomic organization.

INTRODUCTION

Phosphatidylcholine-specific phospholipases D (PLDs) are phosphodiesterases that catalyse the hydrolysis of phosphatidylcholine to produce phosphatidic acid and choline (reviewed in [1]). PLD activity is present in both prokaryotes and eukaryotes. Activation of PLD has been observed in a wide variety of mammalian cells and tissues, particularly in the context of changes in physiological processes such as metabolic regulation, neural and cardiac stimulation, secretion, mitogenesis, and actin polymerization (reviewed in [2]). Regulation of PLD at both protein and RNA levels is currently the subject of intense investigation.

The existence of multiple forms of mammalian phospholipase D was clearly established by a number of biochemical studies that characterized enzymological properties of different PLD activities [2,3]. Over the past 2 years, mammalian cDNA species encoding PLD activity and corresponding to two distinct genes, PLD1 and PLD2, have been cloned [4–7] and their chromosomal locations determined [6]. PLD1 and PLD2 are members of a highly conserved gene family [4]; bacterial, plant, yeast and mammalian PLDs encode conserved sequences that mediate catalysis [8,9] and potentially other common properties [10]. Human and mouse PLD1 proteins are 92% identical [4,6]. Human, mouse and rat PLD2 are similarly conserved ([5,7], and GenBank accession number AF033850). However, PLD1 is only approx. 50% identical with PLD2 [5].

Analysis of mouse PLD1 (mPLD1) and mPLD2 RNA distribution has revealed that both are expressed in the many tissues previously reported to exhibit PLD activity [6]. However, hybridization studies *in situ* demonstrated that, whereas PLD1 and PLD2 are simultaneously expressed in some cells, other cells express only one or even neither gene. Moreover, PLD1 and PLD2 mRNA levels become up-regulated on differentiation of HL60 neutrophil-like and C6 glioma cells, and down-regulated on induction of apoptosis [11–13]. Finally, PLD1 exhibits alternative splicing of unknown significance [14]. Taken together, these findings suggest that an important aspect of PLD regulation involves changes in rates of transcription and in mRNA processing.

To investigate the mechanisms underlying transcriptional

control and mRNA processing and to obtain the information required for transgenic approaches such as gene inactivation by homologous recombination, it is necessary to establish genomic organization. In this paper we describe the genomic structures of the mPLD2 gene, including its intron–exon organization and proximal promoter region sequences. We also describe alternatively spliced PLD2 isoforms isolated from cDNA libraries or PCR amplifications, or present in the Genbank database.

EXPERIMENTAL

P1 clones

P1 genomic clones for mPLD2 were obtained from an ES cell library through a screen conducted by Genome Systems (St. Louis, MO, U.S.A.) by using PLD2-specific amplification primers as described below. The primers consisted of the sequences GTAACCCAGAAGAACCTCTTTCCC [nt 7–30 of the open reading frame (ORF)] and CAAGGGGTGTGCTTTCAGAGGC (nt 149–171 of the ORF) and were designed on the basis of the cDNA sequence published previously [5]. PCR was performed with a GeneAmp PCR System 9600 (Perkin-Elmer) with a hot start (98 °C for 5 min) and 30 cycles of amplification (denaturation for 30 s at 94 °C, annealing for 30 s and 64 °C and extension for 1 min at 68 °C). The PCR products were sequenced with a Sequenase PCR Product sequencing kit (Version 2.0; Amersham Life Science, Cleveland, OH, U.S.A.) to confirm their identities.

Southern blot analyses

P1 clones obtained from Genome Systems were digested with a panel of restriction enzymes, electrophoresed on a 0.8% agarose gel and blotted to a nylon membrane (Hybond-N, Amersham Life Science). The membrane was hybridized with mPLD2 cDNA probes labelled with [α -³²P]dCTP (3000 Ci/mmol; NEN Life Science Product, Boston, MA, U.S.A.) by random priming (Random primed DNA labeling kit; Boehringer Mannheim, Germany). Hybridization was performed overnight at 65 °C in 1% (w/v) BSA/0.4 M sodium phosphate (Na₂HPO₄, 7H₂O + H₃PO₄)/15% (v/v) deionized formamide/1 mM

Abbreviations used: mPLD, mouse PLD; PLD, phospholipase D; RACE, rapid amplification of cDNA ends; UTR, untranslated region.

¹ To whom correspondence should be addressed (e-mail Michael@pharm.som.sunysb.edu).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession numbers AF052291, AF052292, AF052293 and AF052294.

Figure 1 Genomic structures of mPLD2 and alternatively spliced isoforms

mPLD2 spans 17.1 kb. The alternatively spliced isoforms mPLD2si-a and mPLD2si-b are shown as isolated from the mouse brain cDNA library. Presumably they represent partial length clones but it is not known whether the transcripts from which they were derived encode wild-type N-termini. Similarly the remainder of the structure of the human PLD2 isoform lacking the 33 bp depicted is not known. mPLD2si-c was isolated through PCR amplification of embryonic cDNA species by using primers directed against the first and last exons. Abbreviation: hPLD2, human PLD2.

EDTA/7% SDS (w/v) and washed twice with $0.2 \times \text{SSC}/0.1\%$ SDS at 65°C .

Subcloning and gene structure detection

The 85 kb P1 genomic clones were restricted with *Bam*HI, *Pst*I or *Eco*RI, and shotgun subcloned into correspondingly digested Bluescript-KS plasmids. The exon-intron structure of mPLD2 was determined by sequence analysis of the clones that hybridized to the full-length PLD2 cDNA probe.

Confirmation of contiguous genomic organization

A linear reconstruction of the genomic locus from individual *Bam*HI fragments was developed. In some instances the *Bam*HI sites occurred in the cDNA, eliminating the possibility of intervening non-hybridizing fragments. To examine the other sites, PCR amplification across putative junctions was performed with the Expand[®] Long Template PCR System kit (Boehringer Mannheim, Germany), using the P1 clone as template. The PCR products were then digested with *Bam*HI and analysed on a 1% (w/v) agarose gel to determine whether additional genomic material was present. Such studies were repeated and confirmed with *Bam*HI restriction of the corresponding and overlapping *Pst*I or *Eco*RI fragments. At one site (between exons 10 and 11), the presence of four additional *Bam*HI fragments was observed. To order the fragments, a 6 kb *Pst*I fragment that contained the

entire region was subjected to partial digestions by *Bam*HI. The *Bam*HI reactions were stopped by the addition of $1\ \mu\text{l}$ of 0.1 M EDTA after 1, 2, 4, 8, 15, 30 or 60 min of digestion. Southern analysis of the partly digested samples with probes from the extreme ends of the fragment was used to order the *Bam*HI sites.

Proximal promoter analysis

Possible consensus sequences important for the regulation of mPLD2 transcription were identified with the MACSIGNALSCAN program to analyse the proximal promoter region sequence. The approximate transcription start point for mPLD2 was detected by using 5' rapid amplification of cDNA ends (RACE) as described previously [15]. The site was confirmed by using staggered PCR as described in the text. The G+C distribution was determined with the program DNASIS with a group size of 150 bp.

Splice isoforms

Alternatively spliced isoforms for mPLD2 were identified through screening a mouse brain cDNA library as previously described [5,6] (ICR outbred strain, Harlan Sprague-Dawley, newborns, Lambda ZAP II vector; Stratagene). Unpublished expressed sequence tags were located by using searches of the GenBank database. An additional alternatively spliced isoform was identified through the amplification of mPLD2 from a variety of

embryonic cDNA first-strand mixtures by using primers located in the first and last exons.

Expression and analysis of mPLD2 expression

Phospholipids were purchased from Avanti polar lipids. PtdIns(4,5) P_2 was isolated as described [4]. L- α -Dipalmitoyl [choline-methyl- 3H]phosphatidylcholine was obtained from American Radiolabeled Chemicals. All other reagents were obtained from previously noted standard sources and were of analytical grade unless otherwise specified [4]. COS-7 cells were maintained in Dulbecco's modified Eagle's medium with 10% (v/v) fetal calf serum. For transfections the cells were grown in 35 mm dishes and then switched into Opti-MEM I media (Gibco BRL). mPLD2 or control cDNA species were transiently over-expressed in COS-7 cells as described previously by using the mammalian expression vector, pCGN [4,5,8]. The transfection efficiency was observed to be approx. 5–20%. mPLD2 protein expression and stability was assessed by using Western blotting as described previously [8]. mPLD2 activity was measured with the head group release assay *in vitro* in the presence or absence of guanosine 5'-[γ -thio]triphosphate-activated ADP-ribosylation factor as described previously [4,5,8].

RESULTS AND DISCUSSION

Isolation of mPLD2 genomic clones

We previously reported the cloning of a second mammalian PLD gene, mPLD2 [5]. To obtain genomic clones for mPLD2 we synthesized a pair of primers corresponding to nt 7–30 and 149–171 of its open reading frame. An approx. 270 nt fragment was amplified from genomic DNA, instead of the expected 164 bp product, suggesting the presence of a small intron between the primers. Sequencing of the PCR product confirmed that the correct product had been amplified and that it contained an

intron of 108 bp. The primers were then used to screen a Genome Systems P1 mouse ES cell library by using a PCR sib-selection strategy.

Three P1 clones were obtained and digested with a panel of restriction enzymes that included *Bam*HI, *Eco*RI, *Hind*III and *Sac*I. The pattern of DNA fragments observed with ethidium bromide staining was similar in most respects but also revealed unique bands for each P1 plasmid (results not shown), indicating that they represented independent clones. Southern blot hybridization with the PLD2 cDNA identified an identical set of fragments in each of the three P1 clones, demonstrating that the differences between them lay significantly outside the PLD2 transcription unit. A single clone was chosen for further analysis. Shotgun subcloning of the P1 plasmid DNA restricted with *Bam*HI, *Eco*RI or *Pst*I into pBluescript was performed and hybridizing plasmids were identified. Subsequently analysis, as described below, determined that the entire mPLD2 gene had been recovered.

Structure of the mPLD2 gene

The exon–intron structure of the mPLD2 gene was determined through sequence analysis, which revealed that mPLD2 is interrupted by 24 introns (Figure 1). Most exons and introns were found to be relatively short (Table 1). Long exons were found only in the 5' and 3' untranslated regions (UTRs). All exons encoding protein sequence and 13 of the 24 introns were less than 200 nt in length. Only one intron was moderately sized (approx. 6.7 kb). Taken as a whole, the approximate length of the mPLD2 transcription units is 17.1 kb. The intron/exon junctions are described in Table 1. The 3' junctions of all of the splice sites exhibited an 11 nt T/C-rich sequence. All of the splice donor and splice acceptor sites were in good agreement with the GT/AG rule [16]. The genomic organization of PLD genes in yeast and in

Table 1 mPLD2 Intron/exon junctions

Capitals indicate exon sequences; lower-case letters indicate intron sequences.

Exon no.	Exon size (bp)	Intron donor	Intron size (bp)	Acceptor
1	≥ 320–449	CCTTAGCTAGgtgcttttttg	231	cctcatatttcctagGATGACTGTA
2	110	GAGGATCCAGgtatagaggg	108	ttttccctccagCTGATCGAAT
3	131	CGGATCCAAGgtcagtcacc	103	tttcataatccagGTGGGAACCT
4	143	CTTTGGCTCGgtgagtagaa	374	totcatotctgtagCTTTGCTGTG
5	106	CAGCAACAGgtatgcagca	80	cngtctcccttcagAAATACTTGG
6	66	CCACGCCATGgtaaggtcct	514	ttcttctctgcagACAGAATTTC
7	59	CCAAAGGACTgtgagtggtg	118	gcttgctccctgtagGGAAGGGGTG
8	93	GGTCCAAGAGgtgagcaagg	89	cccttcgctccagGTGGCTGGTG
9	153	CCTCCCACAGgtgaggtctt	~ 600	ttattttctgcccagGTCCCTGATT
10	150	TGGCCCGGTGgtgagacctt	~ 6700	tggactcccccagGTTTGTGAAT
11	84	CAGACTGGTGgtgagtgagg	122	gcccctggcttctagGTTGAGTCCT
12	79	GAGGAAGGCGgtgaggagtg	~ 980	cccctcactaccagGAAGAAGGTG
13	111	CAACATAAAGgtgacgttca	138	tcaccctactctagGTGATCGCAC
14	171	ACATTTACAGgtaattgcc	454	tgctctctacagATCCACAC
15	127	CCTTTTGAAGgtgaggtccc	581	tcccgaggactccagATTTTCATCGA
16	119	TTTACCAAGgtgctgagtc	114	attccacatggcagACCACCAAGG
17	114	CACTGTGCAGgtgacagtcc	178	tcaccatcccatagGTCTTGAGGT
18	105	CTACATTGAGgtctgacgca	178	gatccctgccccagAATCAGTTCT
19	89	AGGCTCACGagtaaggaagg	100	tatttttctctcagACAGGGGCAG
20	114	TCACCTACAGgtgaccact	166	ttcttctctctcagGACCCTGTGT
21	50	AAAGCAGCCAggtgagtgctg	212	tctcttgggtacagTGGGGACTGC
22	135	GTCATCATTTGgtcagtgctg	177	gcccttgaacccagGTTCTGCGAA
23	154	GATGTTTCAGgttaggagcg	~ 942	ctattctctcnacagTGTGATTCTT
24	115	CTATGAGCAGgtggaagctg	219	ccctttccatccagATCTTCGCT
25	772

Figure 2 mPLD2 5' UTR and proximal promoter region

Upper panel: the sites of initiation of cDNA species recovered from mouse (nt -673, one clone; nt -551, seven clones) and rat (nt -556) are shown by filled circles, relative to the ATG translation initiation codon, which is indicated by double underlining. Potentially important transcriptional regulatory elements [30] are underlined. The hairpin loop structure at -601 to -664 encodes a stem of 11 bp and a loop of 42 bp [$\Delta G = -21$ kcal/mol (87.9 kJ/mol)]. Capital letters, exons 1 and 2; lower-case letters, presumptive proximal promoter sequence and intron 1; rPLD2, rat PLD2. Lower panel: RACE analysis for mPLD2. The final amplification was performed with a gene-specific primer initiating at nt 53. The approx. 430 nt band shown in lane 2 corresponds to cDNA species initiating at nt -551 in the genome DNA (-331 in the cDNA). Lane 3 represents a negative control for which the amplifications were performed in the absence of cDNA template. Molecular masses in kDa are indicated at the left.

the castor bean has been reported previously [17,18]. PLD in yeast exists as a single uninterrupted open reading frame [17]. Castor bean PLD is encoded by four exons [18], but the intron/exon junctions do not correspond to locations of any of the intron/exon junctions determined here for mPLD2 (results not shown).

5' UTR and proximal promoter

RACE PCR [15] was used to acquire full-length cDNA species at the 5' end [5]. A large number of clones were obtained that ended at approximately the same location (Figure 2, and results not shown). The mPLD2 5' UTR sequence is additionally thought likely to be full or nearly full-length, because the composite length of the mPLD2 cDNA is 3.7 kb and Northern blot analysis

previously determined that the mPLD2 mRNA is approx. 3.5 kb [6]. PCR amplifications of embryonic cDNA by using a common 3' primer located in the second exon and a series of increasingly distal primers located in the first exon or presumed 5' flanking sequence were used to confirm that the 5' end of the PLD2 transcript did not extend significantly past the two sites identified by RACE (results not shown). The longest cDNA clone examined initiated 442 nt 5' of the ATG start codon, which was 117 nt further upstream than the site of initiation reported for a rat cDNA PLD2 clone [7], as shown in the corresponding genomic sequence in Figure 2 (upper panel). However, most of the clones initiated 320 nt 5' of the ATG, at a site 5 nt downstream of the site corresponding to the 5' end of a cDNA described for rat PLD2 [7]. The rat and mouse 5' UTR sequences are highly conserved throughout their common regions (87%; Figure 3),

Figure 3 Comparison of mouse and rat 5' UTR sequences

The sequences are 87% identical, and identities are indicated by vertical lines. The initial in-frame stop codon upstream of the ATG is underlined for each cDNA. No additional ATGs are found for either mouse or rat in the 5' UTR upstream of the ATG that initiates the PLD2 open reading frame.

but are not conserved with the human PLD2 5' UTR (results not shown).

Taking into account intron 1, which separates the 5' UTR from the coding region and is 231 nt in length, the transcription start site for mPLD2 lies approx. 551–673 nt upstream of the ATG start codon in the genomic DNA (Figure 2, upper panel). At the extreme 5' end of mPLD2, a potential stem (11 bp) and loop (42 bp) structure is apparent at positions –601 to –664. The free energy for the base-paired structure is –21 kcal/mol (–88 kJ/mol). A similar configuration has been reported for other genes, including for example the hamster 3-hydroxy-3-methylglutaryl-CoA reductase gene [19]. Such structures have been found to mediate specific interactions with RNA-binding proteins [20,21].

The proximal promoter region was examined by sequence analysis of a 6.5 kb *Eco*RI genomic fragment that contained approx. 3 kb of 5' flanking DNA. TATA and CAAT boxes are frequently found near the point of initiation for eukaryotic genes (at approx. –25 and –80 nt relative to the site of initiation respectively). However, neither was identified in the most proximal 1 kb of 5' flanking DNA for mPLD2, suggesting that mPLD2 might instead be regulated by SP1 or ETF transcriptional factors similarly to other genes that lack TATA and CAAT boxes. Such genes characteristically exhibit GC-rich regions immediately 5' of the transcription start site [19,22–25] and encode SP1 or ETF consensus binding sites nearby. Alternatively, RNA polymerase II basal transcription factor TFIID can interact with initiator regions of TATA-less genes and can promote transcription even in the absence of such sites [26–28].

The 5' UTR and flanking region for mPLD2 are relatively GC-rich (Figure 4). This region contains a GC-rich island ($G + C > 50\%$) extending from the ATG start codon to –1050, which is preceded by a GC-poor region ($G + C < 50\%$) from –1050 to –1503. In addition, islands of CpG are found in or near the PLD2 GC-rich region (Figure 4). A similar pattern of GC distribution has been reported for other TATA-less genes, for example the rat carbonic anhydrase II gene [29]. Three SP1 consensus sites (GGGCGG) were found at positions –681, –710 and –726. PLD activity and PLD1 and PLD2 mRNA levels have been reported to be up-regulated by cAMP and retinoic acid [11–13], but corresponding canonical consensus sites for the cAMP and retinoic acid response elements were not found in the 1 kb of sequence immediately upstream of the transcription start site or in the first intron, suggesting as one possibility that the increase in transcription does not represent an

Figure 4 Distribution of G/C basepairs upstream of the ATG translation initiation codon

The percentage ($G + C$) was calculated by using a window of 150 nt (top panel). In the bottom panel the distribution of GC dinucleotide pairs is compared with that of CG dinucleotides (by using a window of 50 nt), revealing conserved islands of CG dinucleotides in the proximal promoter, exon 1, and part of intron 1 but not in more distal portions of the promoter. The distribution of GC and CG dinucleotides in the 3' UTR is shown for comparison. The thin horizontal line indicates the 50% ($G + C$) level in the top panel, and the frequency of GC or CG dinucleotides expected if the distributions were random in the bottom panel. In the bottom panel, the thin experimental line shows the number of GC dinucleotides found in 50 nt window, and the thick line the number of CG dinucleotides.

immediate-early response. Alternatively, such sites could be located outside of this region or the response could be mediated by atypical elements.

Figure 5 Comparison of PLD1 and PLD2 in the region of the PLD1-specific loop sequence and alternative splice site

Human PLD1b (hPLD1b) lacks 38 amino acid residues encoded by hPLD1a [14]. Dashes indicate amino acids not present. The filled circles indicate intron/exon junctions. The 3' site of divergence corresponds to the location of an intron/exon junction for hPLD1a [14]. Mouse and rat PLD1 also encode both isoforms [6,14]. The intron 14/exon 15 junction for mPLD2 occurs within 2 nt of the PLD1 splice site.

3' UTR structure and polyadenylation signal

A classic polyadenylation signal sequence (AAUAAA) was previously reported to be present at 529 nt 3' of the stop codon and 19 nt upstream of the poly(A) tail in the cDNA clones isolated [5]. Three other classic polyadenylation signal sequences were found in introns 7, 9 and 14 (Figure 1), although their use was not observed.

Splice isoforms for mPLD2

Alternative pre-mRNA splicing is a widely used mechanism for gene regulation and the generation of different protein isoforms. Two alternatively spliced PLD1 isoforms that differ in the inclusion or exclusion of a 38-residue exon and are highly conserved in human, rat and mouse have been described, although the biological significance of this alternative processing is unknown [14]. We have identified four alternatively spliced isoforms of PLD2, found either through screening of a mouse brain cDNA library or through GenBank searches. mPLD2 splice isoform a (mPLD2si-a) lacks exons 10–24 and the coding portion of exon 25 (Figure 1). mPLD2si-b lacks portions of exons 24 and 25 (part of the open reading frame from exon 25 is retained but is frame-shifted). mPLD2si-c lacks a portion of exon 7 and exons 8–21 and includes a portion of intron 21. The splice sites used by the isoforms do not correspond to exon/intron junctions defined by the wild-type cDNA reported previously (Figure 1). A composite cDNA for mPLD2si-b (with a wild-type N-terminus) was expressed and assayed; stable protein was generated but in contrast with wild-type mPLD2, which exhibits robust constitutive activity *in vitro* [5,8], no enzymic activity was detected for mPLD2si-b (results not shown). In addition to the two murine alternate isoforms, a human PLD2 expressed sequence tag present in GenBank encodes a sequence in which 11 amino acid residues (33 nt) have been excised from exon 23. The 3' (but not the 5') site of divergence matches an mPLD2 intron/exon junction.

mPLD2 lacks an 116-residue 'loop' region present in the longer form of PLD1 (PLD1a; see [5,14]). Cognate PLD proteins from non-animal species also lack the loop region [4,14], suggesting that PLD1 acquired this sequence through the inclusion of novel DNA (as opposed to PLD2's having lost one or more exons). The site at which the sequence is missing lies 2 nt away from an intron/exon junction for mPLD2 (intron 14/exon 15) and at the location where alternative splicing takes place for PLD1 (Figure 5). This finding raised the possibility that the PLD1 'loop' region might have arisen from the inclusion of the equivalent of part of intron 14 into the final mRNA product. However, translation of intron 14 into protein yielded peptides

with no demonstrable similarity with PLD1, suggesting another origin for the extra sequence or extensive divergence of the intron since the duplication event that created PLD1 and PLD2.

Concluding remarks

The mPLD2 genomic structure is noteworthy in that it is largely composed of very small exons and introns. Such structural arrangements often lead to the presence of numerous alternatively spliced variants, several examples of which are reported here. Because PLD1 has also been reported to exhibit alternative splicing and shares at least one genomic intron/exon boundary in common with PLD2, it is likely to be organized similarly. It remains to be determined whether any of the alternatively spliced products are of biological importance. The results presented here provide the basic information required to undertake promoter analyses and to generate inactivated alleles for mPLD2 by homologous recombination.

We thank S. Tsirka, J. Engebrecht and A. Morris and members of the Frohman lab for their critical reading of the manuscript, and Paul Steed and Bernard Lassegue for sharing unpublished results. This work was supported by a grant from the NIH to M.A.F. (GM54813) and by funding from Onyx. O.R. was supported by an American Heart Association Fellowship.

REFERENCES

- Exton, J. H. (1997) *J. Biol. Chem.* **272**, 15579–15582
- Klein, J., Chalifa, V., Liscovitch, M. and Koffelholz, K. (1995) *J. Neurochem.* **65**, 1445–1455
- Provost, J. J., Fudge, J., Israelit, S., Siddiqi, A. R. and Exton, J. H. (1996) *Biochem. J.* **319**, 285–291
- Hammond, S. M., Altshuler, Y. M., Sung, T. C., Rudge, S. A., Rose, K., Engebrecht, J., Morris, A. J. and Frohman, M. A. (1995) *J. Biol. Chem.* **270**, 29640–29643
- Colley, W., Sung, T. C., Roll, R., Hammond, S. M., Altshuler, Y. M., Bar-Sagi, D., Morris, A. J. and Frohman, M. A. (1997) *Curr. Biol.* **7**, 191–201
- Colley, W. C., Altshuler, Y. M., Sue-Ling, C. K., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Branch, K. D., Bollag, R. J., Bollag, W. B. and Frohman, M. A. (1997) *Biochem. J.* **326**, 745–753
- Kodaki, T. and Yamashita, S. (1997) *J. Biol. Chem.* **272**, 11408–11413
- Sung, T. C., Roper, R., Zhang, Y., Rudge, S. A., Temel, R., Hammond, S. M., Morris, A. J., Moss, B., Engebrecht, J. and Frohman, M. A. (1997) *EMBO J.* **16**, 4519–4530
- Morris, A. J., Engebrecht, J. and Frohman, M. A. (1996) *Trends Pharmacol. Sci.* **5**, 182–185
- Ponting, C. P. (1996) *Protein Sci.* **5**, 2353–2357
- Ohguchi, K., Nakashima, S., Tan, Z., Banno, Y., Dohi, S. and Nozawa, Y. (1997) *J. Biol. Chem.* **272**, 1990–1996
- Nakashima, S., Ohguchi, K., Frohman, M. A. and Nozawa, Y. (1998) *Biochim. Biophys. Acta*, in the press
- Yoshimura, S., Nakashima, S., Ohguchi, K., Sakai, H., Shinoda, J., Sakai, N. and Nozawa, Y. (1996) *Biochem. Biophys. Res. Commun.* **225**, 494–499
- Hammond, S. M., Jenco, J. M., Nakashima, S., Cadwallader, K., Cook, S., Nozawa, Y., Frohman, M. A. and Morris, A. J. (1997) *J. Biol. Chem.* **272**, 3860–3868
- Frohman, M. A. (1994) *PCR Methods Appl.* **4**, S40–S58

-
- 16 Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383
 - 17 Rose, K., Rudge, S. A., Frohman, M. A., Morris, A. J. and Engebrecht, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12151–12155
 - 18 Xu, L. W., Zheng, L., Coughlan, S. J. and Wang, X. M. (1996) *Plant Mol. Biol.* **32**, 767–771
 - 19 Reynold, G. A., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. and Luskey, K. L. (1984) *Cell* **38**, 275–285
 - 20 Shi, H., Hoffman, B. E. and Lis, J. T. (1997) *Mol. Cell. Biol.* **17**, 2649–2657
 - 21 Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F. and Bouvet, P. (1966) *J. Mol. Biol.* **260**, 34–53
 - 22 Ishii, S., Xu, Y., Stratton, R. H., Roe, B. A. and Merlino, G. T. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4920–4924
 - 23 Ishii, S., Merlino, G. T. and Pastan, I. (1985) *Science* **230**, 1378–1381
 - 24 Kageyama, R., Merlino, G. T. and Pastan, I. (1989) *J. Biol. Chem.* **264**, 15508–15514
 - 25 Geiser, A. G., Busam, K. J., Kim, S. J., Lafyatis, R., O'Reilly, M. A., Webbink, R., Roberts, A. B. and Sporn, M. B. (1993) *Gene* **129**, 223–228
 - 26 Purnell, B. A., Emanuel, P. A. and Gilmour, D. S. (1994) *Genes Dev.* **8**, 830–842
 - 27 Martinez, E., Chiang, C. M., Ge, H. and Roeder, R. G. (1994) *EMBO J.* **13**, 3115–3126
 - 28 Kaufmann, J. and Smale, S. T. (1994) *Genes Dev.* **8**, 821–829
 - 29 McGowan, M. H., Neubauer, J. A. and Stolle, C. A. (1997) *Gene* **186**, 181–188
 - 30 Hames, B. D. and Higgins, S. J. (1993) *Gene Transcription: A Practical Approach*, IRL Press, Oxford
-

Received 19 September 1997/5 January 1998; accepted 9 January 1998