Cloning, expression and chromosomal localization of the rat pancreatitis-associated protein III gene

Nelson J. DUSETTI,* Jean-Marc FRIGERIO,* Claude SZPIRER,† Jean-Charles DAGORN* and Juan L. IOVANNA*‡

*U.315 Institut National de la Santé et de la Recherche Médicale, 46 Bd de la Gaye, F-13009 Marseille, France and †Université Libre de Bruxelles, Département de Biologie Moléculaire, Rue des Chevaux, 67, B-1640 Rhode-St-Genese, Belgium

PAP III belongs to the family of pancreatitis-associated proteins, recently characterized as pancreatic secretory proteins structurally related to C-type lectins, and whose expression is induced during the acute phase of pancreatitis. In this paper, we describe the cloning, characterization and chromosomal localization of the rat PAP III gene. The gene was isolated from a genomic library using a PCR-based method and characterized over 2.5 kb of gene sequence and 1.7 kb of 5′-flanking sequence. The 5′ end of the coding sequence was determined by primer extension of the PAP III transcript. The PAP III coding sequence spanned over six exons. We found striking similarities between PAP III and PAP I and II genes, in genomic organization as well as in promoter sequences. Moreover, the rat PAP III gene was mapped to chromosome 4 using mouse–rat hybrid cells, a localization which coincides with that of the PAP I and II genes. The three genes could therefore derive from the same ancestral gene by duplication. Expression of the PAP III gene was compared with that of PAPs I and II. Expression levels were similar in pancreas, where PAP III mRNA concentration increased within 6 h following induction of pancreatitis, reached maximal levels (> 200 times control values) at 24–48 h, and decreased thereafter. In the intestinal tract, where PAP II is not expressed, the pattern of PAP III expression was comparable with that of PAP I; fasting induced a decrease in its mRNA concentration by more than 80%, which could be reversed within 6 h upon feeding. PAP III is therefore a new member of the PAP gene family, more closely related to the PAP I gene.

INTRODUCTION

Acute pancreatitis induces important changes in the expression of a number of genes in the rat pancreas (Iovanna et al., 1991a, 1992). In contrast to an overall decrease found for most pancreatic enzymes, mRNA level and protein synthesis of the rat 'pancreatitis-associated protein' (PAP) increased dramatically (Iovanna et al., 1991b). In fact, PAP was not detectable in the pancreas of control animals. It could be detected in juice 6 h after induction of an experimental acute pancreatitis, reached a maximum during the acute phase (24–48 h) and disappeared during recovery (Keim et al., 1991). On the other hand, the human homologue of PAP was isolated from pancreatic juice (Keim et al., 1992), and subsequently cloned (Orelle et al., 1992) and characterized as a novel serum indicator of the course of acute pancreatitis (Iovanna et al., 1994).

Cloning of the PAP mRNA from a rat pancreatic cDNA library provided the primary structure of the encoded preprotein, whose 175-amino-acid sequence, including an N-terminal signal peptide of 26 amino acids, was typical of secretory proteins (Iovanna et al., 1991b). Whereas PAP was expressed by the pancreas during the acute phase of pancreatitis only, the same molecule was constitutively expressed by the epithelial cells of the small intestine (Kamimura et al., 1992; Dusetti et al., 1993a; Iovanna et al., 1993; Itoh and Teraoka, 1993). On the other hand, Lasserre et al. (1992) have shown that PAP mRNA, called HIP by these authors, was expressed at a high level in some hepatocellular carcinomas.

Sequence similarity between PAP and other pancreatic proteins was significant with lithostathine/*reg* I only (Terazono et al., 1988; Rouquier et al., 1989,1991). A computer-assisted sequence comparison revealed a striking sequence similarity between the C-terminal region of PAP and the carbohydrate-recognition domain of calcium-dependent lectins (Drickamer, 1988; Iovanna et al., 1991b), suggesting a common ancestor gene and providing new insight into the phylogeny of that class of lectins (Dusetti et al., 1993a). This is supported by the recent finding that PAP could bind lactose (Christa et al., 1994).

A second rat PAP was identified from a pancreatic cDNA library, and called PAP II (Frigerio et al., 1993a), the original PAP being named PAP I. It was also overexpressed by the pancreas during the course of acute pancreatitis but could not be detected in the intestinal tract nor in other tissues (Frigerio et al., 1993a). More recently, the family extended to a third protein, PAP III, whose sequence similarity with PAP I and II amounted to 66 and 63% of amino acids, respectively (Frigerio et al., 1993b). Analysis of the PAP III gene was undertaken to complete the positioning of the PAP family on the lectin phylogenic tree and to obtain the structural and functional information required for eventual studies on the PAP promoter.

EXPERIMENTAL

Isolation of a PAP III genomic clone

Preliminary stage

A PCR-based method was used for the rapid isolation of a PAP III genomic clone based on gene tracking in serial dilutions of the

Abbreviations used: CRD, carbohydrate recognition domain; IL-6, interleukin 6; PAP, pancreatitis-associated protein; p.f.u., plaque forming units.

‡ To whom correspondence should be addressed.

The nucleotide sequence reported in this paper has been submitted to the GenBank Data Bank with accession number U09193.

Table 1 Mapping of the PAP I, II and III genes

Note: The rat chromosome content of each of the rat–mouse somatic cell hybrids is indicated, and the presence of rat-specific PAP I, II and III bands in each hybrid is also shown. The number of discordant clones is shown along the botton row. Symbols: +, chromosome present in more than 55% of the cells; (+), chromosome present in 25–55% of the cells; (–), chromosome present in less than 25% of the cells; –, chromosome absent. The cell hybrids used in this study, derived from the fusion of mouse hepatoma cells (BWTG3) with adult rat hepatocytes, have been described previously [44], except for LB260TG4 which is a new clone.

	PAP			Rat chromosome																				
Hybrid	I	II	III	X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
LB150.1	+	+	+	+	–	–	+	+	–	–	+	–	+	(–)	+	+	+	–	–	(+)	(+)	+	(+)	–
LB251	+	+	+	+	+	+	–	+	–	(+)	+	–	–	+	–	+	+	–	–	–	+	–	+	–
LB1040TG3	+	+	+	–	–	–	–	+	–	+	(–)	–	–	–	+	–	–	–	+	+	–	+	–	–
LB780	+	+	+	+	–	+	+	+	+	–	+	–	–	+	+	–	+	–	–	–	+	+	–	(+)
BS511	+	+	+	+	–	+	+	+	(–)	+	+	+	–	–	+	–	+	–	+	+	+	–	+	–
LB260TG4	+	+	+	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–
LB210.I	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–	+	–	+	–	–
LB210B	+	+	+	+	–	–	–	+	–	–	–	–	–	–	–	–	+	+	–	+	–	+	–	–
LB161	+	+	+	+	–	+	+	+	+	+	+	–	+	+	–	(–)	+	+	+	+	+	+	+	(+)
LB780-8	–	–	–	+	–	+	–	–	–	–	+	–	–	+	+	–	–	–	–	–	–	+	–	–
LB330TG3	+	+	+	–	–	–	–	+	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–
Discordant				5	8	6	5	0	7	5	5	8	7	7	6	6	4	8	6	4	4	6	5	7

library by primer-specific PCR amplification. To test the specificity for the PAP III gene of the selected set of oligonucleotides, 50 ng of rat genomic DNA was amplified by PCR, utilizing sense (5′-GAAGATGCCAAGGAAGATGTG-3′) and antisense (5′-CTAGGCCTTGAACTTGCAGAC-3′) primers (Frigerio et al., 1993b). Amplification was performed in 1 × PCR buffer (50 mM KCl, 10 mM Tris/HCl, pH 8.3, 2 mM $MgCl_2$ and 0.01 % gelatin) containing 125 μM dNTP, 1 % DMSO, 25 pmol of each primer and 2.5 units of *Taq* polymerase in a final volume of 50 μl. Reaction times were as follows: first cycle, denaturation at 94 °C for 2 min, addition of *Taq* polymerase, annealing at 55 °C for 2 min and extension at 74 °C for 2 min. Subsequent 28 cycles, denaturation at 94 °C for 10 s, annealing at 55 °C for 2 min and extension at 74 °C for 2 min. Last cycle, denaturation at 94 °C for 10 s, annealing at 55 °C for 2 min and extension at 74 °C for 10 min. The PCR product was eventually sequenced after asymmetric PCR amplification, as previously described (Orelle et al., 1992). A rat genomic library constructed in λEMBL3 SP6/T3 with inserts generated by partial digestion with *Sau*3A of rat liver DNA (Clontech) was screened as described below.

First step

The presence of at least one copy of the PAP III gene in the amplified library was confirmed as follows: about 10^6 recombinant clones of the library were amplified in liquid medium to 10^{10} plaque forming units (p.f.u.)/ml. DNA from an aliquot was purified by standard methods (Maniatis et al., 1982), of which 5 ng were PCR amplified with PAP III-specific oligonucleotides, as described above.

Second step

Ten aliquots of the amplified library, each containing about 10^5 p.f.u. (named λB10^{5}1 to λB10^{5}10 pools), were again amplified in liquid medium to 10^{10} p.f.u./ml. DNA purification and PCR reactions were performed as in the first step. One pool showing adequate PCR amplification (λB10^{5}3) was selected for use in the next step.

Third step

Ten aliquots of the amplified λB10^{5}3 pool, each containing about 10^4 p.f.u. (named λC10^{4}1 to λC10^{4}10 pools) were amplified on Luria–Bertani plates. DNA was purified from each pool and PCR reactions were performed as in the first step. Again one pool showing adequate PCR amplification (λC10^{4}1) was selected.

Fourth step

In the last step of cloning, about 1.5×10^4 p.f.u. of the amplified λC10^{4}1 pool were used for plaque hybridization screening (Benton and Davis, 1977) with PAP III cDNA insert (Pat 3) as probe (Frigerio et al., 1993b). Hybridization was carried out at 65 °C in a buffer containing 6 × SSC (SSC is 150 mM NaCl/15 mM sodium citrate), 0.1 % BSA, 0.1 % Ficoll, 0.1 % polyvinylpyrrolidone, 0.5 % SDS and 200 μg/ml denatured herring sperm DNA. Clones showing positive hybridization were plaque-purified by two rounds of screening and characterized by restriction enzyme mapping.

Nucleotide sequencing

Appropriate fragments were subcloned and purified in the plasmid vector pBluescript (KS^+/SK^+)(Stratagene). Nucleotide sequence was determined by the dideoxy chain termination method (Sanger et al., 1977) with Sequenase 2.0 (United States Biochemical Corp) according to the manufacturer's specifications and using synthetic oligonucleotides as primers.

Determination of the transcription initiation site by primer extension

Primer extension analyses were performed as described (Dusetti et al., 1993a). Briefly, an antisense oligonucleotide (5′-AGGC-TGGACAGCAGCATCCAGGAC-3′) was synthesized with its 5′-end at position +54 of the PAP III mRNA. Approximately 10 pmol of oligonucleotide was end-labelled using [γ-^{32}P]ATP and T4 polynucleotide kinase for 40 min at 37 °C. Free [γ-

[^{32}P]ATP was removed from the labelled oligonucleotides by two successive precipitations with ammonium acetate and ethanol and radioactivity incorporation was quantified. The end-labelled oligonucleotide was hybridized to 10 μg of total pancreatic or intestinal RNA for 10 min at 80 °C followed by 12 h at 30 °C in 30 μl of formamide buffer (40 mM Pipes, 1 mM EDTA, 400 mM NaCl and 80 % deionized formamide, pH 6.4). As a negative control, 10 μg of yeast tRNA was incubated in the same conditions with the radiolabelled oligonucleotide. The heteroduplex was precipitated, washed with 70 % ethanol and dissolved in 18 μl of 50 mM Tris (pH 8.2) containing 140 mM KCl, 10 mM $MgCl_2$, 4 mM dithiothreitol, 40 units of placental RNase inhibitor and 2 mM of deoxynucleotide triphosphates. Fifty units of avian myeloblastosis virus reverse transcriptase were added to each sample and incubation was conducted at 42 °C for 60 min. The samples were ethanol-precipitated and the extended primers were analysed by electrophoresis on 6 % acrylamide/urea gels. An M13 mp18 sequencing ladder was used as a size marker.

Northern-blot and dot-blot analyses

For Northern-blot hybridization, 20 μg of denatured RNA was loaded on to each gel lane and submitted to electrophoresis on a 1 % agarose gel (Maniatis et al., 1982). Gels were stained with ethidium bromide to check the quality of the RNA. The RNA was transferred to nylon filters (Hybond N, Amersham, U.K.) which were baked and prehybridized for 4 h at 42 °C in a buffer containing 50 % deionized formamide, 5 × SSPE (SSPE is 180 mM NaCl, 1 mM EDTA, 10 mM NaH_2PO_4, pH 7.5), 0.1 % BSA, 0.1 % Ficoll, 0.1 % polyvinylpyrrolidone, 0.5 % SDS and 200 μg/ml denatured herring sperm DNA. Hybridization was conducted at 42 °C for 16–20 h in the same buffer in the presence of the ^{32}P-labelled probe (Feinberg and Vogelstein, 1983). Then the filters were washed four times for 5 min at room temperature in 2 × SSC/0.1 % SDS, twice for 15 min at 50 °C in 0.1 × SSC/0.1 % SDS, and once for 30 min in 0.1 × SSC. Filters were dried and exposed to Hyperfilm TM (Amersham, U.K.).

Quantitative analysis of RNAs (dot-blots) was performed according to White and Bancroft (1982), as described previously (Calvo et al., 1991). RNA samples (5, 2.5, 1.25 and 0.625 μg) were denatured and dotted on to prewashed nitrocellulose filters using a Minifold apparatus (Schleicher & Schuell, Inc.). The filters were washed and baked. Dot blots were prehybridized, hybridized, and washed as described for Northern blots. Filters were exposed to Hyperfilm TM (Amershan, U.K.). The autoradiograms of the blots were scanned with a Multiscan apparatus (LKB instruments) and the results analysed as described previously (Calvo et al., 1991), the slopes of the regression lines giving an estimate of the mass of the probed mRNA per μg of total RNA.

Animal experiments

This study was conducted in male Sprague–Dawley rats weighing 180–200 g. Animals were fed standard rat chow (UAR, Villemoisson sur Orge), and were kept in individual cages in a room with controlled temperature and light.

Rat RNA samples

Total RNA was prepared from normal pancreas, from pancreas with acute pancreatitis, stomach, duodenum, jejunum, ileum, caecum, colon, brain, liver, kidney, salivary gland, heart, testes, spleen, skeletal muscle, lung, prostate and thymus.

Acute pancreatitis

Experimental acute pancreatitis was induced by retrograde injection of 200 μl of 0.4 % sodium taurocholate in the main pancreatic duct, as described by Lankisch et al. (1974). Animals were killed after 1.5, 3, 6, 12, 24, 48, 96, 168 and 240 h. RNA from pancreas was prepared as described previously (Calvo et al., 1991).

Nutritional manipulations

Access to food and water was *ad libitum* before starting the experiments. In a first experiment, 110 rats were separated into two groups. One group had free access to food, whereas the other group was fasted. Five rats per group were killed after 0, 3, 6, 12, 24 and 48 h and the ileums were removed for RNA extraction. In a second experiment, rats starved for 24 h were given food and killed after 3, 6, 12 and 24 h, whereas control animals remained fasted. Both protocols started at 09.00.

Analysis of somatic cell hybrid DNA by PCR

PAP III chromosomal localization

The PCR was used to specifically amplify rat PAP III sequences in DNA from a panel of somatic cell hybrids. The 11 hybrids used are listed in Table 1. DNA was prepared from somatic cell hybrids and PCR was carried out using 50 ng of DNA. The same set of oligonucleotides, 1 × PCR buffer and reaction times as for PCR gene cloning (see above) were used for DNA amplification, except that the annealing temperature was 65 °C. Aliquots (5 μl) were removed and analysed on a 1.2 % agarose gel. DNA products were visualized with ethidium bromide under UV light.

PAP I and PAP II chromosomal localization

The same strategy was used for chromosomal assignment of PAP I and PAP II. The oligonucleotides used for PAP I (Dusetti et al., 1993a) DNA amplification were 5′-GAAGACTCTCCGAAGAAAATACCC-3′ and 5′-ACCTGTAAATTTGCAGACGTAGGG-3′, and for PAP II (Frigerio et al., 1993a) were 5′-ATCCCAGATCACTGCAAGG-3′ and 5′-CTGCGGGTCTACTGCTTGAAC-3′.

Computer analysis

The FASTA program (Lipman and Pearson, 1985) was used to search the Genbank database for sequence similarities. Sequence comparisons and alignments were performed using GAP (Needleman and Wunsch, 1970), CLUSTAL (Higgins and Sharp, 1988) and TREE (Feng and Doolitle, 1987) analyses.

RESULTS

Isolation of the rat PAP III gene

Before screening, an amplification by PCR of rat genomic DNA was performed with our selected set of oligonucleotides (see the Experimental section). This yielded a single product of about 1400 bp. The partial sequence of that product was identical to the sequence reported for rat PAP III mRNA (Frigerio et al., 1993b). In the first step of screening, a single product of 1400 bp, similar to that obtained from genomic DNA, was obtained from the library after PCR amplification. In the second step, the amplified product was observed in λB10^{5}3, λB10^{5}4 and λB10^{5}9

Figure 1 Nucleotide sequence of the rat PAP III gene

The nucleotide sequence of the rat PAP III gene and 1705 nucleotides of the 5′-flanking sequence is given. Exonic sequences are in bold letters. Introns 1, 2, 3, 4 and 5 are 276, 546, 197, 579 and 179 bp long respectively. Numbering begins at the 5′-end of the first exon of the gene. The TATAA (position −31) and CAAT (position −40) sequences are boxed. A putative acute-phase reaction signal (in reverse orientation) is underlined.

pools. In the third step, amplification was observed in λC10^{4}1 and λC10^{4}4 pools. After that step, and with the evidence that the gene was present in a population of only 10^4 recombinant phages, the cloning (fourth step) was carried out after two rounds of plaque hybridization screening with a cDNA coding for PAP III (Frigerio et al., 1993b) as probe. This yielded the λGPAP III clone containing the PAP III gene.

Organization of the PAP III gene

Insert DNA from λGPAP III was restriction-digested, subcloned and the sequence in both strands was obtained by the dideoxy method using a combination of strategies of staggered deletions and sequence 'walking' with multiple synthetic oligonucleotide primers. The entire sequence of the rat PAP III gene was determined, including 1705 bp of the 5′-flanking region (Figure 1). These data and comparison with the PAP III cDNA sequence revealed that the PAP III gene spanned a region of 2551 bp, containing six exons interrupted by five introns. Exons ranged in size from 21 to 291 bp. Exons averaged 128.28 ± 33.47 bp (mean ± S.E.M.) consistent with the reported average exon size of 137 bp (Hawkins, 1988). As shown in Figure 1, intron/exon boundaries conformed to the known GT/AG donor/acceptor site rule, essentially maintaining the consensus sequence proposed by Mount (1982).

Determination of the transcription initiation sites

To map the rat PAP III gene transcription initiation site(s), primer extension analysis was performed using total RNA from

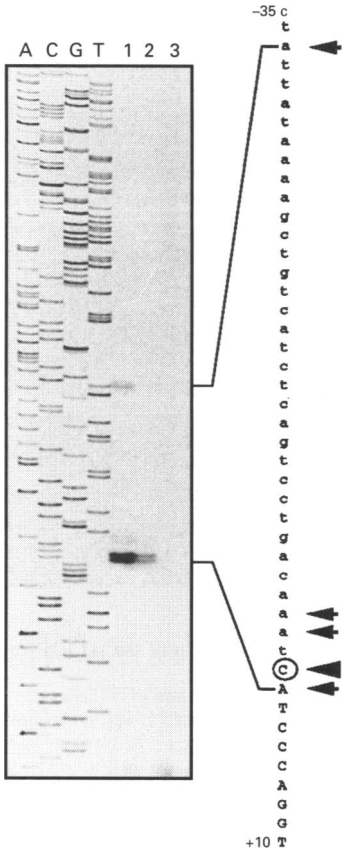

Figure 2 Determination of the 5′-end of the rat PAP III transcript by primer extension analysis

An end-labelled oligonucleotide (5′-AGGCTGGACAGCAGCATCCAGGAC-3′) complementary to the PAP III mRNA was annealed to 10 μg of RNA from pancreas with acute pancreatitis (1), ileum (2) or yeast tRNA (3) and extended with reverse transcriptase. An M13 mp18 ladder was used as a size marker. The nucleotide sequence of the region encompassing the initiation sites is presented on the right-hand side with the initiations indicated by arrows.

pancreas with acute pancreatitis and from intestine (see the Experimental section). Five putative transcriptional initiation sites were identified in pancreatic RNA and four in intestinal RNA. Based on the intensity of the signal, the major transcriptional site was found in both tissues 26 bases upstream of the first translated ATG codon (Figure 2), and the corresponding nucleotide was numbered +1. No primer extension product was obtained in the control experiment involving yeast tRNA.

Northern-blot analyses of RNAs from several rat tissues

We examined PAP III mRNA expression in several tissues from normal rat and in pancreas with acute pancreatitis by Northern-blot analysis using rat PAP III cDNA as probe. As shown in Figure 3, PAP III was expressed strongly in pancreas with pancreatitis and ileum, weakly in duodenum and jejunum, and no signal was obtained on normal pancreas, stomach, caecum, colon, brain, liver, kidney, salivary gland, heart, testes, spleen, skeletal muscle, lung, prostate or thymus. On longer exposure a signal was observed in heart and thymus (data not shown). RNAs from the same tissues obtained from rats 2 days after induction of acute pancreatitis showed a comparable pattern of PAP III mRNA expression. In addition, no signal was obtained after probing with PAP III RNAs extracted from liver after induction of an acute phase or from kidney with acute experimental failure (data not shown).

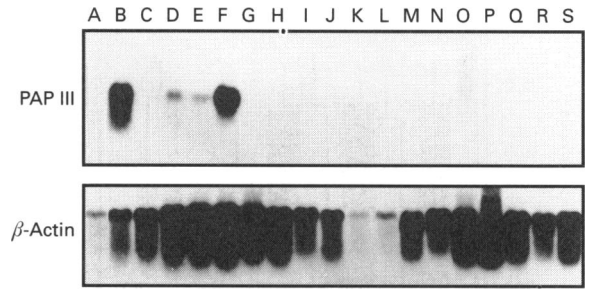

Figure 3 Northern-blot analysis of RNAs from various rat tissues using rat PAP III cDNA

RNA samples were prepared from the following tissues: lane A, normal pancreas; lane B, pancreas with acute pancreatitis; lane C, stomach; lane D, duodenum; lane E, jejunum; lane F, ileum; lane G, caecum; lane H, colon; lane I, brain; lane J, kidney; lane K, liver; lane L, salivary gland; lane M, heart; lane N, testes; lane O, spleen; lane P, skeletal muscle; lane Q, lung; lane R, prostate; lane S, thymus. Ten micrograms of RNA was separated on an agarose gel and blotted to a nylon membrane. The blot was probed with the random-primer-labelled PAP III cDNA (Pat 3) or β-actin cDNA.

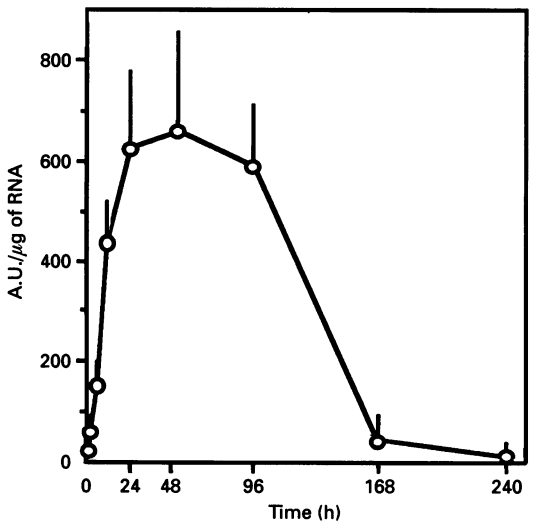

Figure 4 Kinetics of PAP III mRNA concentration changes following induction of acute pancreatitis

Values were obtained by dot-blot hybridization of equal amounts of pancreatic RNA to ^{32}P-labelled PAP III cDNA probe. They were estimated by scanning the autoradiograms of the blots and expressed as absorbance units (A.U./μg of RNA). Values are expressed as means ± S.E.M.

Expression of PAP III mRNA during the course of acute pancreatitis in the pancreas

Quantification of the increase in PAP III mRNA concentration during the course of an experimental acute pancreatitis was performed by dot-blot hybridization (Figure 4). PAP III mRNA concentration increased within 6 h of induction of pancreatitis, reached maximal levels at 24–48 h, and decreased thereafter. Levels at 7–10 days were similar to control values.

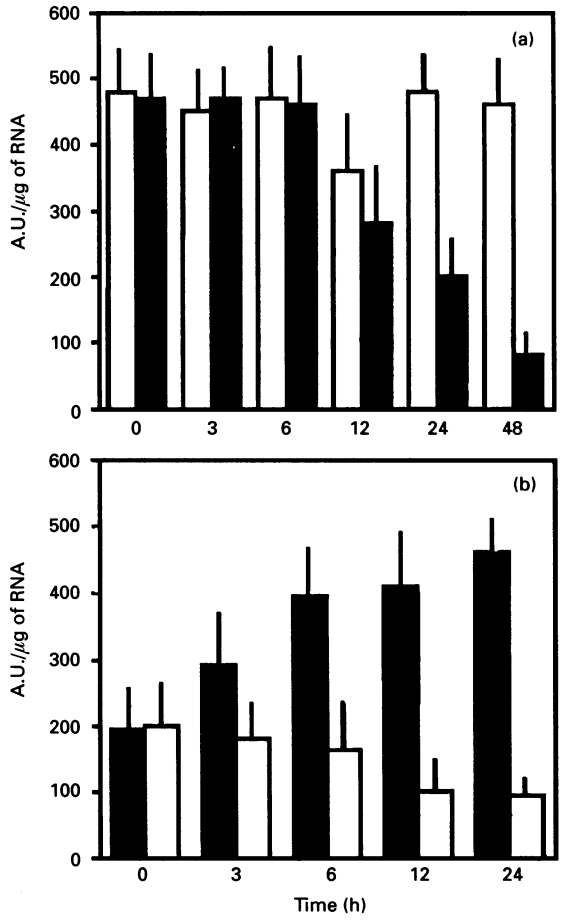

Figure 5 Effect of fasting and feeding on PAP III mRNA concentration in ileum

(**a**) Groups of 5 rats were killed at differents times after they were denied access to food (filled columns). Rats with free access to food were used as controls (empty columns). (**b**) Rats starved for 24 h were given food *ad libitum* and killed after 0, 3, 6, 12 or 24 h (filled columns). Starving was maintained in control rats (empty columns). PAP III mRNA concentration in their ileum was measured as described in Figure 4. Values are expressed as means ± S.E.M.

Influence of fasting and feeding on PAP III mRNA expression in the ileum

We selected ileum, the section of the intestine where constitutive expression of PAP III mRNA is highest, to establish the kinetics of PAP III mRNA expression after access to food was denied, starting at 09:00. Figure 5 shows that PAP III mRNA concentration decreased to 58% of the initial value after 12 h of fasting and to 42% and 16% after 24 and 48 h respectively. In the matching group, with free access to food, a transient decrease to 77% of control was observed after 12 h. The influence of feeding was also analysed in rats fasted for 24 h (Figure 5). PAP III mRNA concentration increased after food ingestion. Plateau values of animals fed *ad libitum* were reached after 6 h. Actin and villin cDNA were used in control experiments performed on the same RNAs. No change was observed during fasting or refeeding (data not shown).

Mapping of rat PAP III gene to chromosome 4

The results of the PCR amplification of rat genomic DNA, and DNA from a panel of somatic cell hybrids and their parental cell line using the PAP III primers are illustrated in Figure 6. Rat DNA sequences were specifically amplified in the hybrid cells, as a product of the expected size of 1400 bp, and no band was seen using DNA from the mouse parent cell line. The DNA of appropriate size was obtained only in the hybrid cell lines which contained an intact chromosome 4 (Table 1) and was absent from the other cell lines, clearly indicating assignment of the PAP III gene to chromosome 4.

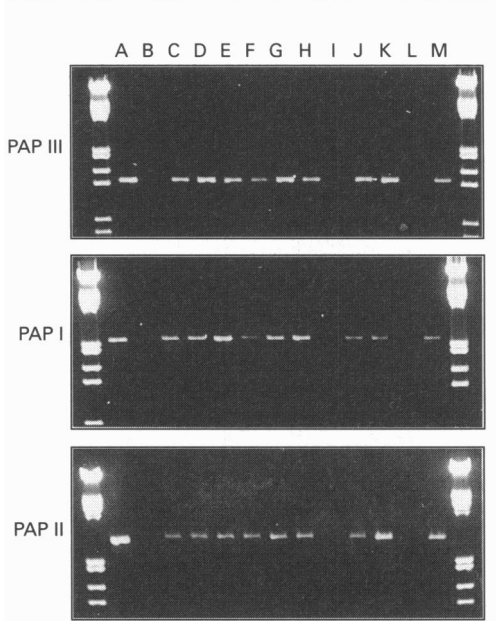

Figure 6 Chromosomal localization of the rat PAPs genes

DNA samples from somatic cell hybrids and parental cell lines were subjected to PCR analysis using specific primers for PAP III, PAP I and PAP II (see the Experimental section). Lane A is DNA from the parental adult rat hepatocytes (HRSD) and lane B is DNA from the parental mouse hepatoma cell line (BWTG3). Lanes C to M are DNA from the somatic cell hybrid panel: LB150.1, LB251, LB1040TG3, LB780, BS511, LB260TG4, LB210.I, LB210B, LB161, LB780-8, LB330TG3 (Szpirer et al., 1984). Size markers are *Eco*RI + *Hin*dIII digestion products of λDNA.

Assignment to chromosome 4 of the PAP I and PAP II genes

We have established the chromosomal location of the genes encoding PAP I and PAP II in the rat, using PCR amplification of DNA from the same panel of mouse–rat somatic cell hybrids. PAP I and PAP II were specifically amplified as products of 2100 and 2500 bp respectively. As for PAP III, appropriate DNA amplifications were obtained in hybrid cell lines which contained the intact chromosome 4 (Figure 6 and Table 1).

DISCUSSION

A third PAP cDNA was recently isolated and characterized in the rat and was shown to be the product of a separate gene (Frigerio et al., 1993b). We report here the structural and functional characteristics of the rat PAP III gene, and describe the implications of these findings for the knowledge of the PAP gene family.

The PAP III gene

A PCR-based method (Dusetti et al., 1993b, 1994) was used to isolate a genomic clone containing the entire PAP III gene, which is 2.5 kb in length and contains six exons and five introns. The cap site, as mapped by primer extension analysis, was a C residue

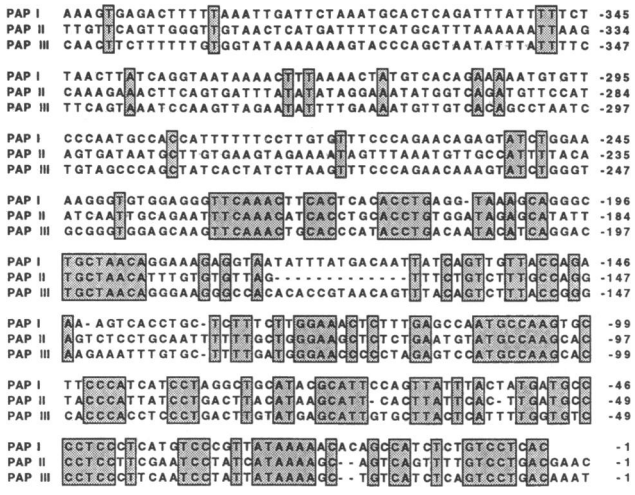

Figure 7 Sequence comparison between the 5′-flanking regions of rat PAP I, II and III genes

The numbering for each gene sequence starts at the position of the capped nucleotide (determined by primer extension analysis). Gaps are introduced to optimize sequence similarity. Identical sequences are shaded.

26 bp upstream from the translation start codon (ATG) (Figure 1). Examination of 1705 bp of sequence upstream from the cap site showed the presence of several consensus sequences corresponding to transcription regulatory elements. The classical TATAA box (Breathnach and Chambon, 1981) and CAAT box (McKnight and Tjian, 1986) sequences were found at −31 and −40 respectively (Figure 1). Three additional TATAA boxes were located at positions −380, −489 and −972 respectively (Figure 1). Finally, a putative acute-phase signal (CTGGGAA), described as 'interleukin-6 (IL-6)-response element' (Hattori et al., 1987), was observed in reverse orientation at position −269. This IL-6-response element is present in many acute-phase genes, in which it is required for full induction (Kunz et al., 1989). An acute-phase response element in the promoter region of the PAP III gene might therefore participate in the increased PAP expression which occurs during the acute phase of pancreatitis.

PAP III and the PAP gene family

The organization of the PAP III gene is very similar to that of the genes encoding PAP I (Dusetti et al., 1993a) and PAP II (Frigerio et al., 1993a). The number of exons is identical. They have the same size, except for exon 5 which is one codon shorter in the PAP III gene. In addition, the interruption of the reading frame by introns occurs inside codons at the same position in the three genes. Comparison of the promoter regions of these genes confirmed their relationship. In the proximal region of the 5′-flanking sequences of the three genes, we found 53% identity over 232 nucleotides (Figure 7). A higher degree of similarity with PAP III was, however, observed for PAP I than for PAP II. Conservation amounted to 65% of nucleotides between PAP III and PAP I, whereas a 13 nucleotide deletion in position −164 of the PAP II promoter had to be introduced in order to reach 63% conservation with PAP III (Figure 7). In addition, upstream from position −232, promoters from PAPs III and I showed 78% identity over a 43-nucleotide stretch in which homology with PAP II was not significant (17% nucleotide identity). That stretch includes the IL-6-responsive element (Dusetti et al., 1993a; Frigerio et al., 1993a). It is noteworthy that among the encoded proteins, PAPs III and I show 66% amino acid identity compared with 63% between PAPs III and II and 58% between PAPs I and II.

Expression of the three PAP genes during the acute phase of pancreatitis is unique among pancreatic proteins and co-ordinated, which is in agreement with the presence of homologous regions in their promoters. In addition, the closer similarity of the PAP III and PAP I promoters, compared with PAP II, might account for their tissue distribution. Whereas PAP I and PAP III are constitutively expressed in the intestine, and induced by feeding (Iovanna et al., 1993; and Figure 5), PAP II is not expressed (Frigerio et al., 1993a). However, assignment of specific sequences of the PAP promoters to characteristic features of their expression must await functional studies involving deletions and mutant reporter gene fusion constructs.

The experiments using PCR-amplification with DNA from a panel of somatic cell hybrids and specific sets of oligonucleotides for PAP I, II and III genes have shown that the PAP locus is on chromosome 4 (Table 1). The significant sequence similarity between PAP I, PAP II and PAP III, their similar gene organization and their localization on the same chromosome suggest strongly that they derived from the same ancestral gene by gene duplication.

Comparative mapping of homologous genes from a number of species provides information regarding the evolution of mammalian chromosomes. The human PAP I gene maps to the chromosome 2p12 (Dusetti et al., 1994). This finding confirms the existence of conservation between the 2p12 region of human and the rat chromosome 4.

Evolutionary tree of the PAPs and PAP-related genes

We have established an evolutionary tree for the PAPs and PAP-related proteins as described in the Experimental section. The sequences of rat PAP I (Iovanna et al., 1991b), II (Frigerio et al., 1993a) and III (Frigerio et al., 1993b), human PAP (Orelle et al., 1992), mouse PAP I (Itoh and Teraoka, 1993), bovine PTP (de la Monte et al., 1990), rat lithostathine/*reg* I (Rouquier et al., 1991), human lithostathine/*reg* I (Giorgi et al., 1989) and II (Bartoli et al., 1993) and mouse lithostathine/*reg* I and II (Unno et al., 1993) were obtained from the literature. The dendrogram shows a close relationship between all these sequences (Figure 8). Among them, the PAPs and bPTP (probably the bovine homologue of the PAP) form a subgroup distinct from the lithostathine/*reg* sequences. This is further supported by the presence in all proteins from the PAP subgroup of a 5-amino-acid stretch in position 84–89, absent in proteins from the lithostathine/*reg* subgroup (Rouquier et al., 1991; Giorgi et al., 1989; Bartoli et al., 1993; Unno et al., 1993).

PAPs and PAP-related sequences constitute a new protein family that shares with the C-type lectins the structural features of their Carbohydrate-Recognition Domain (CRD). Based on sequence similarities, functional properties and exon–intron organization of the genes, Bezouska et al. (1991) proposed a classification of the C-type lectins into four groups: lectins with CRD in proteoglycans (group I); lectins with CRD joined to N-terminal membrane anchors, constituting endocytosis receptors for serum glycoproteins (group II); lectins with CRD associated with collagenous domains (group III); and lectins with CRD in the LEC-CAMs (group IV). The PAP family does not fit into any of these groups. Differences in structural features between the PAP family and each of the four groups of lectins were as important as differences observed among these groups. In addition, the pattern of introns within the CRD-coding region was different for the PAP group and for the others (see below).

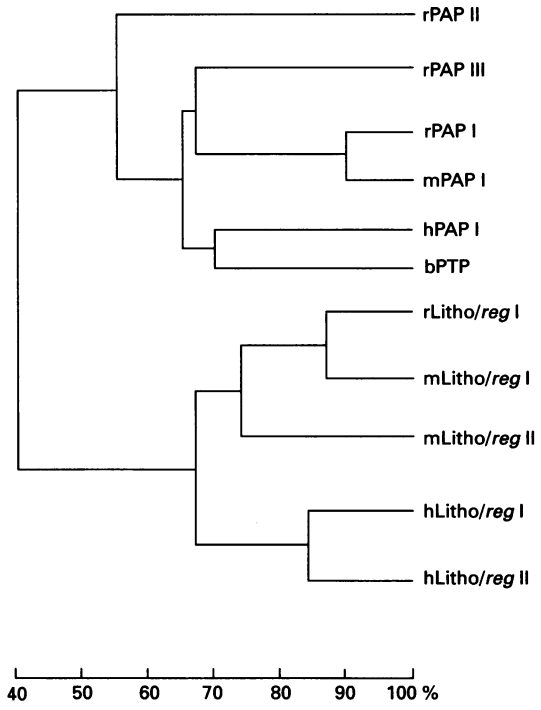

Figure 8 Phylogenetic tree showing the relationship of PAPs and PAP-related proteins

A dendrogram was constructed by comparison of the amino acid sequences of rat PAP I (Iovanna et al., 1991b), II (Frigerio et al., 1993a) and III (Frigerio et al., 1993b), human PAP (Orelle et al., 1992), mouse PAP I (Itoh and Teraoka, 1993), bovine PTP (de la Monte et al., 1990), rat lithostathine/*reg* I (Rouquier et al., 1991), human lithostathine/*reg* I (Giorgi et al., 1989) and II (Bartoli et al., 1993) and mouse lithostathine/*reg* I and II (Unno et al., 1993) as described in the Experimental section. The horizontal scale indicates the percentage sequence similarity.

Hence, we propose that the PAPs and PAP-related proteins are considered as a fifth group of C-type animal lectins.

The intron–exon organization of the CRD-containing genes was also used as a criterion to explore the phylogeny of lectins (Dusetti et al., 1993a). Earlier compilations of lectin gene structures showed that their CRD was either encoded by a single exon, or by three exons with conserved intron–exon boundaries. Description of the CRD of PAP I (Dusetti et al., 1993a) showed the existence of a third intron–exon pattern, which led to the conclusion that PAP I had diverged very early from the other CRD-containing genes (Dusetti et al., 1993a). The present results on PAP III and recent information on several genes related to PAP (Frigerio et al., 1993a; Dusetti et al., 1993b, 1994; Giorgi et al., 1989; Bartoli et al., 1993; Unno et al., 1993) showing that all members in that family have the same intron–exon organization of their CRDs provide further support for that conclusion.

N.J.D. was supported by a grant from the Fondation pour la Recherche Médicale (Paris), J.-M.F. was supported, in part, by a fellowship from Eurorga-Jouveinal. C.S. is Research Director of the National Fund for Scientific Research (FRSM, Belgium). Work in Brussels was supported by the Belgian program on interuniversity attraction poles initiated by the Belgian State Prime Minister's Office Science Policy Programming, and the Fund for Medical Scientific Research (FRSM, Belgium). Technical assistance of P. Garrido, M. Rivière and R. Grimaud is gratefully acknowledged.

Received 19 August 1994/17 November 1994; accepted 23 November 1994

REFERENCES

Bartoli, C., Gharib, B., Giorgi, D., Sansonetti, A., Dagorn, J.-C. and Bergé Lefranc, J.-L. (1993) FEBS Lett. **327**, 289–293
Benton, W. D. and Davis, R. W. (1977) Science **196**, 180–182
Bezouska, K., Crichlow, G. V., Rose, J. M., Taylor, M. E. and Drickamer, K. (1991) J. Biol. Chem. **266**, 11604–11609
Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. **50**, 349–383
Calvo, E., Dusetti, N., Cadenas, B., Dagorn, J.-C. and Iovanna, J. L. (1991) Pancreas **6**, 150–156
Christa, L., Felin, M., Morali, O., Simon, M., Lasserre, C., Brechot, C. and Seve, A. (1994) FEBS Lett. **337**, 114–118
de la Monte, S., Ozturk, M. and Wands, J. R. (1990) J. Clin. Invest. **86**, 1004–1013
Drickamer, K. (1988) J. Biol. Chem. **263**, 9577–9560
Dusetti, N., Frigerio, J.-M., Keim, V., Dagorn, J.-C. and Iovanna, J. L. (1993a) J. Biol. Chem. **268**, 14470–14475
Dusetti, N., Frigerio, J.-M., Dagorn, J.-C. and Iovanna, J. L. (1993b) Biochim. Biophys. Acta **1174**, 99–102
Dusetti, N., Frigerio, J.-M., Fox, M., Swallow, D. M., Dagorn, J.-C. and Iovanna, J. L. (1994) Genomics **19**, 108–114
Feinberg, A. P. and Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13
Feng, D. F. and Doolitle, R. F. (1987) J. Mol. Evol. **25**, 351–360
Frigerio, J.-M., Dusetti, N., Keim, V., Dagorn, J.-C. and Iovanna, J. L. (1993a) Biochemistry **32**, 9236–9241
Frigerio, J.-M., Dusetti, N., Garrido, P., Dagorn, J.-C. and Iovanna, J. L. (1993b) Biochim. Biophys. Acta **1216**, 329-331
Giorgi, D., Bernard, J.-P., Rouquier, S., Iovanna, J., Sarles, H. and Dagorn, J.-C. (1989) J. Clin. Invest. **84**, 100–106
Hattori, M., Abraham, L. J., Northemann, W. and Fey, G. H. (1987) Proc. Natl. Acad. Sci. U.S.A. **87**, 2364–2368
Hawkins, M. D. (1988) Nucleic Acids Res. **16**, 9893–9903
Higgins, D. G. and Sharp, P. M. (1988) Gene **73**, 237–244
Iovanna, J. L., Keim, V., Michel, R. and Dagorn, J.-C. (1991a) Am. J. Physiol. **261**, G485–G489
Iovanna, J. L., Orelle, B., Keim, V. and Dagorn, J.-C. (1991b) J. Biol. Chem. **266**, 24664–24669
Iovanna, J. L., Lechene de la Porte, P. and Dagorn, J.-C. (1992) Pancreas **7**, 712–718
Iovanna, J., Keim, V., Bosshard, A., Orelle, B., Frigerio, J.-M., Dusetti, N. and Dagorn, J.-C. (1993) Am. J. Physiol. **265**, 611–618
Iovanna, J. L., Keim, V., Nordback, I., Montalto, G., Letoublon, C., Camarena, J., Lévy, P., Berthézène, P. and Dagorn, J.-C. (1994) Gastroenterology **106**, 728–734
Itoh, T. and Teraoka, H. (1993) Biochim. Biophys. Acta **1172**, 184–186
Kamimura, T., West, C. and Beutler, E. (1992) Gene **118**, 299–300
Keim, V., Iovanna, J. L., Rohr, G., Usadel, K. and Dagorn, J.-C. (1991) Gastroenterology **100**, 775–782
Keim, V., Iovanna, J. L., Orelle, B., Verdier, J.-M., Büsing, M., Hopt, U. and Dagorn, J.-C. (1992) Gastroenterology **103**, 248–254
Kunz, D., Zimmermann, R., Heisig, M. and Heinrich, P. C. (1989) Nucleic Acids Res. **17**, 1121–1138
Lankisch, P. G., Winckler, K., Bokermann, M., Schmidt, H. and Creutzfeldt, W. (1974) Scand. J. Gastroenterol. **9**, 725–729
Lasserre, C., Christa, L., Simon, M., Vernier, P. and Bréchot, C. (1992) Cancer Res. **52**, 5089–5095
Lipman, D. J. and Pearson, W. R. (1985) Science **227**, 1435–1441
Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 1–545, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
McKnight, S. and Tjian, R. (1986) Cell **46**, 795–805
Mount, S. M. (1982) Nucleic Acids Res. **10**, 459–472
Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. **48**, 443–453
Orelle, B., Keim, V., Masciotra, L., Dagorn, J.-C. and Iovanna, J. L. (1992) J. Clin. Invest. **90**, 2284–2291
Rouquier, S., Giorgi, D., Iovanna, J. L. and Dagorn, J.-C. (1989) Biochem. J. **264**, 621–622
Rouquier, S., Verdier, J.-M., Iovanna, J. L., Dagorn, J.-C. and Giorgi, D. (1991) J. Biol. Chem. **266**, 786–791
Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Szpirer, J., Levan, G., Thorn, M. and Szpirer, C. (1984) Cytogenet. Cell Genet. **38**, 142–149
Terazono, K., Yamamoto, H., Takasawa, S., Shiga, K., Yonemura, Y., Tochino, Y. and Okamoto, H. (1988) J. Biol. Chem. **263**, 2111–2114
Unno, M., Yonekura, H., Nakagawara, K., Watanabe, T., Miyashita, H., Moriizumi S., Okamoto H., Itoh T. and Teraoka, H. (1993) J. Biol. Chem. **268**, 15974–15982
White, B. A. and Bancroft, F. C. (1982) J. Biol. Chem. **257**, 8569–8572