Secretory Pancreatic Stone Protein Messenger RNA

Nucleotide Sequence and Expression in Chronic Calcifying Pancreatitis

Dominique Giorgi, Jean-Paul Bernard, Sylvie Rouquier, Juan Iovanna, Henri Sarles, and Jean-Charles Dagorn
Institut National de la Santé et de la Recherche Médicale, U.315 F-13009 Marseille, France

Abstract

The pancreatic stone protein and its secretory form (PSP-S) are inhibitors of $CaCO_3$ crystal growth, possibly involved in the stabilization of pancreatic juice. We have established the structure of PSP-S mRNA and monitored its expression in chronic calcifying pancreatitis (CCP). A cDNA encoding pre-PSP-S has been cloned from a human pancreatic cDNA library. Its nucleotide sequence revealed that it comprised all but the 5' end of PSP-S mRNA, which was obtained by sequencing the first exon of the PSP-S gene. The complete mRNA sequence is 775 nucleotides long, including 5'- and 3'-noncoding regions of 80 and 197 nucleotides, respectively, attached to a poly(A) tail of ~ 125 nucleotides. It encodes a preprotein of 166 amino acids, including a prepeptide of 22 amino acids. No overall sequence homology was found between PSP-S and other pancreatic proteins. Some homology with several serine proteases was observed in the COOH-terminal region, however. The mRNA levels of PSP-S, trypsinogen, chymotrypsinogen, and colipase in CCP and control pancreas were compared. PSP-S mRNA was three times lower in CCP than in control, whereas the others were not altered. It was concluded that PSP-S gene expression is specifically reduced in CCP patients.

Introduction

Pancreatic secretion of vertebrates contains ~ 20 major protein components (1). The structure and function of most of them have been the subject of numerous studies in the last decades. Secretory protein includes zymogens, enzymes, an enzyme cofactor (colipase) and a trypsin inhibitor, all of which are directly involved in the digestive process. More recently, our laboratory reported the presence in human pancreas of a polymorphic glycoprotein without recognized enzyme activity (2). It was named "secretory pancreatic stone protein" (PSP-S)[1] because of its immunological identity with the "pancreatic

stone protein," the major component of the protein matrix present in calculi of patients suffering from chronic calcifying pancreatitis (CCP) (3). In human pancreatic juice collected over an appropriate mixture of protease inhibitors, PSP-S comprised four species with M_r between 16 and 20 kD, named PSP-S2 to PSP-S5 (4). Immunoprecipitation of in vitro translation products of pancreatic RNAs showed that PSP-S was synthesized as a single polypeptide (5). Hence, its molecular weight heterogeneity is probably due to posttranslational processing. An additional PSP-S form (PSP-S1) with a M_r of 15 kD appears in juice upon activation (controlled or not), albeit minimal. PSP-S1 derives from PSP-S2-5 by trypsin-like cleavage of an Arg-Ile bond in the NH_2-terminal part of the PSP-S backbone (6). PSP-S1 was independently discovered by Gross et al. in human (7) and bovine (8) secretions, and named "pancreatic thread protein" (PTP) because of its ability to form fibrils at neutral pH.

The abundance of PSP-S in juice (10–14% of total protein) suggests that it plays an important role in the exocrine pancreatic function. In vitro experiments have shown that PSP-S inhibited $CaCO_3$ crystal growth (9). Evidence that normal pancreatic secretion is supersaturated in $CaCO_3$ (10) prompted us to suggest that the physiological role of PSP-S may be related to its inhibitory properties. The demonstration that PSP-S concentration was diminished in the juice of patients suffering from chronic calcifying pancreatitis supported that hypothesis (11). In this work, we report the cloning and complete sequencing of human PSP-S mRNA, and the alterations of its expression associated with CCP.

Methods

Experimental procedures

Construction of a human pancreatic cDNA library in $\lambda gt10$. Total RNA was prepared from fragments of human pancreas by the technique of Chirgwin et al. (12). The polyadenylated fraction was purified by affinity chromatography on oligo-dT cellulose (13), and used to direct cDNA synthesis in the presence of reverse transcriptase. DNA polymerase was used in conjunction with RNase H to synthesize the complementary cDNA strands. Enzymes and reagents were obtained as a kit from Amersham International (Amersham, UK) and used according to their recommendations. The double-strand (ds)-cDNA was ligated into bacteriophage $\lambda gt10$ DNA with Eco R1 linkers and packaged using the Amersham packaging kit. The library contained a total of 5.5×10^4 independent recombinant clones, and was amplified to a titer of 3×10^{10} plaque-forming units (pfu)/ml.

Biohazards associated with the experiments described in this publication have been examined previously by the French National Control Committee.

Library screening. The cDNA library was screened by the "plaque screening" procedure (14), with a 23-mer oligonucleotide mixture cov-

---

---

1. *Abbreviations used in this paper:* CCP, chronic calcifying pancreatitis; PSP, pancreatic stone protein; PSP-S, secretory forms of the pancreatic stone protein; PTP, pancreatic thread protein.

---

2. DeCaro, A., Z. Adrich, B. Fournet, C. Capon, J. Bonicel, J. De Caro, and M. Rovery, manuscript in preparation.

ering all possible sequences encoding the octapeptide DLYCQNMN present in position 61–68 of pre-PSP-S ("BL-1," kindly synthesized by Dr. J. L. Bergé-Lefranc, INSERM U38, Marseille). The oligonucleotides were labeled with [α-^{32}P]ATP and T4 polynucleotide kinase (15) to a specific activity of 1–2 × 10^9 cpm/μg, and used to screen in duplicate ~ 10^5 pfu. 15 positive clones were amplified and screened again in the same conditions except that they were plated at a much lower density; 25 individual positive plaques were obtained. Screening conditions were the following (16, 17): the filters were prehybridized for 2 h at 37°C in 6× SSC, 1× Denhardt's solution, 0.5% SDS, 0.05% Na pyrophosphate, and 100 μg/ml salmon sperm DNA. Hybridization was conducted at 37°C for 16 h in the same buffer in which salmon sperm DNA was replaced by 20 μg/ml tRNA, in the presence of 0.5–1 × 10^9 cpm of oligonucleotide probe. Filters were then washed twice for 1 h and several times at 37°C, 15 min each in 6× SSC, 0.05% Na pyrophosphate, until < 100 cps could be detected with a portable Geiger counter (Mini Instruments Ltd, Burnham-on-Crouch, UK).

A human genomic library in the bacteriophage λEMBL3 was purchased from Clontech Laboratories (Palo Alto, CA). It was screened for potential PSP-S genomic clones with a PSP-S cDNA fragment (i31) and ^{32}P-labeled by random priming (18) to a specific activity of 10^9 cpm/μg. A first plaque screening was performed in duplicate on 10^5 pfu, and the positivity of the clones was confirmed by rescreening individual plaques, which were eventually purified (14). Hybridization and washing conditions were as described by Southern (19).

DNA and RNA analysis by filter hybridization. DNA fragments were transferred to nitrocellulose filters either from 0.4% agarose gels by the technique of Southern (19), or from 7% acrylamide gels by electrotransfer (20). Hybridization conditions with oligonucleotides or inserts were the same as for plaque screening. RNAs were analyzed by electrophoresis on agarose-methylmercury gels (21), transferred to nitrocellulose, and hybridized to random primer-labeled inserts by the technique of Thomas (22). Quantitative dot-blot analysis of RNAs was performed according to White and Bancroft (23) as already described (24). Human chymotrypsinogen, trypsinogen, and colipase cDNA probes were selected from the cDNA library by oligonucleotide screening or heterologous hybridization using the corresponding rat probes (kindly provided by Dr. R. MacDonald, University of Texas, Dallas TX) and their nature confirmed by sequencing. The autoradiographs were scanned with a Multiscan apparatus (LKB Instruments, Gaithersburg, MD), and the results analyzed as in reference 24.

Primer extension experiments. 5 μg of pancreatic RNA were incubated in 5 mM methyl mercury hydroxide for 10 min at room temperature and then used in the cDNA synthesis procedure described by Maniatis et al. (25), in which oligo-dT was replaced by 100 ng of the oligonucleotide DG3, ^{32}P-labeled by kinasing to 2 × 10^9 cpm/μg.

DNA sequencing. DNA fragments were subcloned in the M13 mp18/mp19 system (26), and the single stranded recombinant phages were sequenced with Sequenase (United States Biochemical Corp., Cleveland, OH) by the method of Sanger (27), following the protocol supplied. Sequencing was initiated either by the universal M13 primer or by appropriate synthetic oligonucleotides as described in figure legends.

Sequence comparison. The complete nucleotide sequence of PSP-S mRNA was compared with the 15,465 sequences of Genbank (28). The search was conducted with the BISANCE system (CITI2, Centre Interuniversitaire d'Informatique, Paris) using the program of Kanehisa (29). The complete amino acid sequence was compared with the 5,415 sequences of the National Biomedical Research Foundation data bank, using the FASTP program (CITIZ) (30).

Patients

Four fragments of pancreas were obtained from cadaver kidney transplant donors (three males, one female, aged 14–40 yr). The pancreases were macroscopically normal but no histological study was performed. The other fragments were obtained from the Service de Gastroentérologie, Hôpital Sainte-Marguerite, Marseille. Three samples were taken at surgery from patients presenting with obstructive pancreatitis due to

endocrine tumors. The diagnoses have been verified on histological sections of the resected pancreas that confirmed the tumors and showed the typical lesions of obstructive pancreatitis and the absence of calcifications. These patients were two males and a female whose ages ranged from 62 to 67. Their average daily alcohol consumption was 0, 50, and 110 ml ethanol per d. The four donors and the three cases of obstructive pancreatitis served as controls. In five patients, the diagnosis of chronic pancreatitis was done on histological sections of the resected pancreas. Four of the patients presented with visible pancreatic calculi (calcifications) on x-ray films of the abdomen. The fifth case was a man who used to drink 200 ml ethanol per d for years, when he started presenting 7 yr ago with recurrent attacks of pancreatic pain. Surgery was indicated because on retrograde pancreatography, the only anomaly was a dilation of a collateral duct in the head of the pancreas. As a small cancer could not be excluded in a patient presenting with neither visible pancreatic calculi nor exocrine or endocrine insufficiency, a cephalic duodenopancreatectomy was performed. The first routine sections only showed a slight degree of periductal fibrosis and no tumor. Systematic histological examination of the pancreatic head was then performed, which revealed typical lesions of chronic pancreatitis in a few groups of lobules. A decreased number of acini, dilation of small ducts with atrophy of their epithelium, and large intraductal plugs were found. We diagnosed a very slight but typical form of CCP. These five patients were males aged 45–54. Their daily alcohol consumption was 30, 60, 200, 300, and 600 ml ethanol per d.

Results

Screening the λgt10 library of human pancreatic cDNA for PSP-S clones. A stretch of seven amino acids with minimum codon degeneracy was chosen in the PSP sequence, and a mixture of 128 23-mer oligonucleotides covering all possible sequences homologous to the corresponding mRNA was synthesized (BL-1, Fig. 1). Approximately 10^5 clones of the library were screened with ^{32}P-labeled BL-1, 25 of which were positive. The DNA of 20 recombinant phages was analyzed. Eco RI restriction generated in most of these two fragments, showing the presence of an Eco R1 recognition sequence internal to the cDNA. The longest inserts yielded two Eco R1 fragments of 540 and 250 bp, respectively, for a total of ~ 790 bp. Southern blot analysis demonstrated hybridization of the oligonucleotides to the 250-bp fragment. The fragments were separately ligated into M13 mp18 DNA, and sequenced as described in Methods.

Sequence determination of the recombinant clones. Fig. 1 demonstrates the sequencing strategy. Complete overlap of the two complementary sequences of the 540-bp fragment could
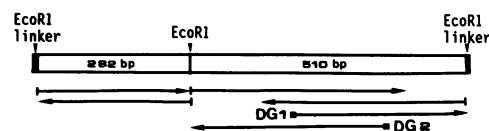
Figure 1. Sequencing strategy of the λDG3-2 cDNA clone. The sequence of the λDG3-2 insert was determined after subcloning in M13 mp19 of the two fragments generated by Eco R1 restriction (282 and 510 bp). *Arrows,* direction and length of sequencing. Complete sequence of both strands of the 510-bp fragment could not be obtained from the universal M13 primer. This was achieved by sequencing the same M13 mp19 clones from synthetic oligonucleotide primers (DG1 and DG2). DG1 was a 19-mer homologous to nucleotides 479–497 of PSP-S mRNA (see Fig. 4) and DG2 was a 18-mer homologous to nucleotides 642–660, in an opposite orientation. DG1 and DG2 sequences were, respectively, 5'-CCTGGCTACTGT-GTGAGCC-3' and 5'-GTTGGAGAGATGGACCGG-3'.

not be obtained when we sequenced from the universal M13 primer only. Additional sequencing was therefore performed, using synthetic oligonucleotide primers (DG1 and DG2) complementary to internal portions of the two cDNA strands (Fig. 1). On the contrary, the 250-bp fragment was short enough to be entirely sequenced from the M13 primer. The typical sequence of the Eco R1 linker used in the construction of the library was found at one end of each fragment, indicating that the fragments were connected by their other end in the cloned cDNA. A polyadenosine stretch specific to the 3′ end of mRNAs was present next to the Eco R1 linker on the 540-bp fragment, allowing orientation of the cDNA. Evidence that the recombinant clone under study actually pertained to PSP-S was obtained from the amino acid sequence deduced from the cDNA sequence. It revealed a single open reading frame, leading to a polypeptide sequence in which the portion of the PSP-S sequence already published could be found (see below).

Primer-extension analysis of the 5′ end of PSP-S mRNA. The 5′ end of a cDNA does not necessarily correspond to its mRNA start point. To estimate the actual length of the PSP-S mRNA, a primer extension experiment was conducted on human pancreatic RNA. The primer (DG3) was a ^{32}P-labeled synthetic oligonucleotide (20-mer) homologous to nucleotides 72–90 from the 5′ end of the cDNA. The size of the extended primer was measured by comparison with a sequencing ladder, and was found 31 nucleotides longer than the cDNA sequence inserted into DG3-2 (Fig. 2 *A*). Hence, at least 31 nucleotides were missing from the 5′ of the cDNA inserted into DG3-2. A "strong stop" (premature termination of the extension) was observed 10 nucleotides downstream.

Sequence of a PSP-S gene fragment including the 5′ end of the RNA (Fig. 3). Rather than look for full-length cDNA clones to determine the nucleotide sequence of the 5′ end of the RNA, we screened a human genomic library constructed in the bacteriophage λEMBL3 with the aim of studying the first exon of the PSP-S gene. A screening of 4×10^5 recombinant phages revealed two positive clones. One of them (λHG-2) was mapped by Eco R1 restriction and the fragments were probed with the oligonucleotide DG3 after Southern blotting. A 235-bp fragment was revealed by the probe, subcloned after purification in M13 mp18/19, and sequenced in both orientations (Fig. 3). The mRNA start site predicted from

the primer extension experiment (Fig. 2 *A*) was a thymidine, 82 nucleotides upstream from the first possible initiation codon. Finding a pyrimidine in position 1 is not usual, but sequences around and upstream from that nucleotide conform to the consensus established for protein-coding genes (31, 32): a typical promoter sequence (TAGAAA) was present at position −27, and pyrimidines were found in positions −1 and 2–6. No other position in the region compatible with our data met any of the requirements for a mRNA start site.

The coding sequence that started at position 82 was interrupted at position 146. A very strong homology was observed when the sequence around nucleotide 146 was compared with the consensus exon-intron boundary sequence (33). The 235-bp gene fragment therefore contains the 3′ end of the PSP-S gene first exon followed by 42 nucleotides of intronic sequence.

Analysis of PSP-S mRNA sequence. The sequence of PSP-S mRNA, obtained by alignment of the cDNA clones DG3-2-282 and DG3-2-510 (Fig. 1), and completed by data from the genomic clone λHG-2 comprises 775 nucleotides, excluding the poly(A) tail (Fig. 4). The size of that mRNA, estimated by Northern blot analysis of total pancreatic RNA, is ∼ 900 nucleotides (Fig. 2 *B*). A poly(A) tail of ∼ 125 nucleotides is therefore attached to mature PSP-S mRNA, only a portion of which was present in the cDNA.

Three methionines at positions 81, 105, and 123 are potential initiation sites for translation. In vertebrates, however, protein synthesis typically starts from the upstream methionine (34) and the initiation codon is surrounded by a well-conserved sequence motif. The most important features of that consensus sequence are an A in position −3 and a G immediately after the ATG codon (35), which are actually present around position 81 only. Hence, synthesis of prePSP-S very probably starts at nucleotide 81.

The 3′ nontranslated sequence of PSP-S mRNA comprises 197 nucleotides. Two canonical polyadenylation signals (AATAAA) (36) are present in tandem, in positions 751 and 757.

Encoded prePSP-S. The single continuous open reading frame prescribes a preprotein of 166 amino acids, with a M_r of 18,690. De Caro et al.[2] recently demonstrated that the NH$_2$-terminal end of mature PSP-S was a glutamine (blocked in the form of 5-oxoproline) in position 23 of the preprotein se-
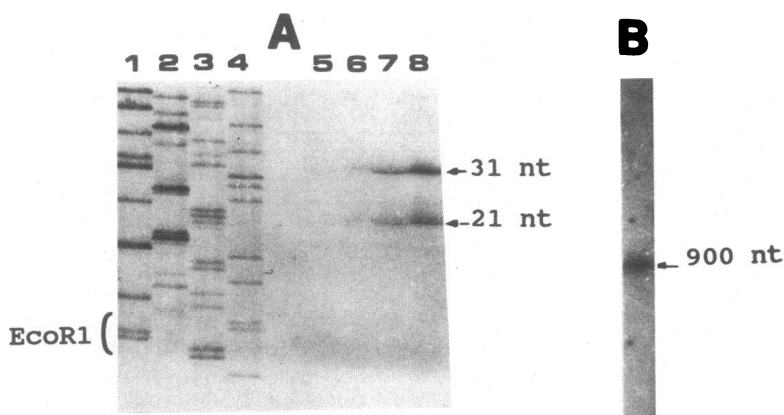

Figure 2. Primer extension and Northern blot analysis of PSP-S mRNA. (*A*) Primer extension: a synthetic 20-mer oligonucleotide (DG3) homologous to nucleotides 72–90 of PSP-S cDNA sequence present in λDG3-2 was used to prime the reverse transcription of the 5′ end of PSP-S mRNA. DG3 sequence was 5′-TCAGGC-AGGAGATCAGACTG-3′. The size of the resulting cDNA was estimated by comparison with a sequencing ladder obtained with a recombinant M13 subclone of the 5′ end of λDG3-2 insert, and DG3 as primer (lanes *1–4*). The Eco R1 linker bordering the 5′ end of λDG3-2 insert is indicated. In lanes *5–8* were loaded, respectively, 0.3, 0.7, 1.5, and 3×10^3 cpm of the primer extension reaction. Two major bands were visible, corresponding to cDNAs, respectively, 21 and 31 nucleotides longer than the cDNA inserted in λDG3-2.

(*B*) Northern blot: 5 μg of pancreatic RNA were submitted to agarose-methylmercury gel electrophoresis and transferred onto nitrocellulose. The inserts of clones DG3-2-294 and DG3-2-510, respectively, homologous to the 5′ and 3′ ends of PSP-S mRNA were ^{32}P-labeled and hybridized either independently or in combination to the transferred RNA. In the three instances, a single band was revealed, whose length was estimated to 900 nucleotides (*arrow*) by comparison with markers run in parallel.

Figure 3 sequence:

```
                                              ▼
gaattcagagagagagatgt[tagaaa]gataacaggtgctctgctctcTTCTTCAAACCCTCTTCCCTG    68

TGTTGTCCTACAGAGATTGCTGATTTGCTCCTTAAGCAAGAGATTCACTGCCGCTAAGC ATG GCT    133
                                                             Met Ala      2

CAG ACC AAC TCG TTC TTC ATG CTG ATC TCC TCC CTG ATG TTC CTG TCT CTG    184
Gln Thr Asn Ser Phe Phe Met Leu Ile Ser Ser Leu Met Phe Leu Ser Leu     19

AGC CAA G gtgagattttcccccacacttcccacaacccca                            224
Ser Gln G                                                                21
```

Figure 3. Sequence and organization of a genomic fragment comprising the first (5') exon of the PSP-S gene. Eco RI-Bam H1 restriction of λHG-2 insert generated six fragments, of which one, 235 bp long, hybridized to the oligonucleotide DG3. That fragment was subcloned in M13 mp19 and sequenced. The promoter sequence (TAGAAA) is boxed, the proposed RNA start site is indicated by an arrow, the intronic sequences are in lower-case letters and the nucleotides of the exon-intron boundary that conform to the consensus sequence are underlined.

quence. The prepeptide therefore comprises 22 amino acids. Its sequence conforms to the characteristics of eukaryotic signal sequences, with an NH$_2$-terminal net charge of +1 and a hydrophobic core (37). Interestingly enough, the 3' boundary of the first exon in the PSP-S gene (Fig. 3) corresponds to the end of the prepeptide (Gly in position 22), showing segregation in that exon of a functional domain of the protein, as described for other secretory proteins (31, 38). Mature PSP-S is a 144-amino acids polypeptide (Fig. 4). The Arg-Ile bond cleaved by trypsin in PSP-S to generate PSP-S1 was found 11 residues downstream from the initial glutamine, and the following 133-amino acid sequence was identical to the PSP-S1 sequence already published. The predicted COOH-terminal end coincided with the end of PSP-S1, excluding posttranslational processing in that region.

Comparison of the COOH-terminal sequences of PSP-S and other serine proteases. A first computer search for homologies between the nucleotide sequence of PSP-S mRNA and the 15,465 sequences indexed in Genbank gave negative re-

Figure 4 sequence:

```
     aqaqaaajaqatattaqaaaaataacaqatactctactctc              -1
                   ↓
ttcttcaaaccctcctcttccctatattctcctacaaaattactaatttactccttaaacaaaagat    68

tcactaccactaaac ATG GCT CAG ACC AAC TCG TTC TTC ATG CTG ATC TCC TCC    122
                Met Ala Gln Thr Asn Ser Phe Phe Met Leu Ile Ser Ser     13

CTG ATG TTC CTG TCT CTG AGC CAA GGC CAA GAG GCC CAG ACA GAG TTG CCC    173
Leu Met Phe Leu Ser Leu Ser Gln Gly Gln Glu Ala Gln Thr Glu Leu Pro     30

CAG GCC CGG ATC AGC TGC CCA GAA GGC ACC AAT GCC TAT CGC TCC TAC TGC    224
Gln Ala Arg Ile Ser Cys Pro Glu Gly Thr Asn Ala Tyr Arg Ser Tyr Cys     47
              ↓
TAC TAC TTT AAT GAA GAC CGT GAG ACC TGG GTT GAT GCA GAT CTC TAT TGC    275
Tyr Tyr Phe Asn Glu Asp Arg Glu Thr Trp Val Asp Ala Asp Leu Tyr Cys     64

CAG AAC ATG AAT TGG GGC AAC CTG GTG TCT GTG CTC ACC CAG GCC GAG GGT    326
Gln Asn Met Asn Ser Gly Asn Leu Val Ser Val Leu Thr Gln Ala Glu Gly     81

GCC TTT GTG GCC TCA CTG ATT AAG GAG AGT GGC ACT GAT GAC TTC AAT GTC    377
Ala Phe Val Ala Ser Leu Ile Lys Glu Ser Gly Thr Asp Asp Phe Asn Val     98

TGG ATT GGC CTC CAT GAC CCC AAA AAG AAC CGC CGC TGG CAC TGG AGC AGT    428
Trp Ile Gly Leu His Asp Pro Lys Lys Asn Arg Arg Trp His Trp Ser Ser     115

GGG TCC CTG GTC TCC TAC AAG TCC TGG GGC ATT GGA GCC CCA AGC AGT GTT    479
Gly Ser Leu Val Ser Tyr Lys Ser Trp Gly Ile Gly Ala Pro Ser Ser Val     132

AAT CCT GGC TAC TGT GTG AGC CTG ACC TCA AGC ACA GGA TTC CAG AAA TGG    530
Asn Pro Gly Tyr Cys Val Ser Leu Thr Ser Ser Thr Gly Phe Gln Lys Trp     149

AAG GAT GTG CCT TGT GAA GAC AAG TTC TCC TTT GTA TGC AAG TTC AAA AAC    581
Lys Asp Val Pro Cys Glu Asp Lys Phe Ser Phe Val Cys Lys Phe Lys Asn     166

tagaggcaactagaaaatacatgtctaaaactaatccagcaattacaacagaatcaaaaattaaaccq    649

aaccatctctccaactcaactcaacctggacactctcttctctgctgaatttgccttgttaatcttca   717

ataatttacctaccccaatctttaaaaccctaaat[aataaaataaa]catattttccactaaaa    782
```

Figure 4. Complete sequence of pre-PSP-S mRNA and deduced sequence of the encoded protein. Sequences obtained from the cloned cDNA (Fig. 1) and from the PSP-S genomic clone (Fig. 3) were aligned. The proposed mRNA cap site is numbered 1. The 5' end of the cDNA (nucleotide 31) is indicated by an arrow. The Eco R1 site internal to the cDNA is boxed. Noncoding sequences are in lower-case letters, and the two potential polyadenylation sites are underlined. In the protein sequence of pre-PSP-S, the prepeptide is underlined and the Arg-Ile bond whose hydrolysis generates the PSP-S1 form is marked by an arrow.

sults. A second search was conducted in the National Biomedical Research Foundation protein bank of 5,415 sequences. As shown on Fig. 5, significant homology was observed in the COOH-terminal regions of PSP-S and several serine proteases: bovine and human trypsinogens, bovine chymotrypsinogen, and plasmatic kallikrein. It is noteworthy that the tripeptide Ser-Trp-Gly, which is involved in the specificity pocket of pancreatic serine proteases (39), is present at positions 123–125 of PSP-S.

Expression of PSP-S, trypsinogen, chymotrypsinogen, and colipase mRNAs in the pancreas of controls and patients suffering from CCP. Quantitative filter hybridization experiments (dot blots) were conducted on total RNA extracted from pancreatic fragments obtained from five patients presenting with CCP and from seven control individuals (Fig. 6). PSP-S mRNA concentration was significantly lower in CCP than in control, with a mean value of 388±70 OD$_{492}$ U/µg RNA compared 1,167±187 OD$_{492}$ U/µg RNA in controls. In the latter group, no significant difference was observed between fragments from pancreatic surgery and from kidney transplantation donors. In contrast to PSP-S, no significant difference was observed between controls and CCP patients in the concentrations of trypsinogen, chymotrypsinogen, and colipase mRNAs. Mean values (±SE) for CCP and controls were, respectively:trypsinogen, 590±157 and 637±94; chymotrypsinogen, 430±78 and 368±36; and colipase, 417±137 and 533±136.

Records of daily alcohol consumption were not available in transplantation donors. In the other cases (three controls, five CCP), values obtained were in the same range for both groups and did not correlate with PSP-S mRNA concentrations (not shown).

Discussion

The proteins of pancreatic secretion in a variety of animal species have been the subject of so many studies, that recent reports suggesting the presence of a new protein entity in pancreatic juice, abundant (up to 14% of total exported protein) and related to a polypeptide found in pancreatic stones, were understandably received with some reservation. Scheele (40), however, already observed by double-dimension gel analysis of human pancreatic proteins the presence of several spots in the 16–19-kD region, with acidic isoelectric points. These spots were eventually attributed to the secretory form of PSP-S. Also, Gross et al. (7) reported the presence of a 15-kD protein forming fibrillar aggregates in the juice of PTP. Sequence comparison between PTP and PSP-S revealed that PTP and PSP-S1, the carboxy-terminal part of PSP-S generated by trypsin-like cleavage of the Arg-Ile bond in positions 11–12

```
PSP-S           S G S  L V  S Y K  S W G  I  G  A P  S  S V N  P G
Bov.Tg1         S G K  L Q  G I V  S W G  S  G  C A Q  K N K  P G
Hum.Tg1         N G Q  L Q  G V V  S W G  N  G  C A Q  K N K  P G
Hum.Plasm.Kal.  M W R  L V  G I T  S W G  E  G  C A R  R E Q  P G
Bov.ChTgA       A W T  L V  G I V  S W G  S  S T C  S  T S T  P G
```

Figure 5. Sequence homology in the COOH-terminal ends of PSP-S and serine proteases. Sequence comparison was obtained by searching the National Biomedical Research Fund database with the FASTP program. Sequence alignment was possible without introducing deletions.

(6), were identical molecules. These results provide additional evidence pointing in the same direction. Clones selected in a human pancreatic cDNA library with an oligonucleotide derived from a portion of PSP-S1 encoded PSP-S. Moreover, no other mRNA species than PSP-S mRNA could be revealed by Northern blot analysis when probing total pancreatic RNA with a cDNA encoding the PSP-S1 portion of PSP-S. The presence of a mRNA encoding PSP-S1, distinct from PSP-S mRNA, is therefore very unlikely. Moreover, transformation of PSP-S into PSP-S1 does not take place in the intracellular compartments where protein maturation normally occurs, as PSP-S1 is absent from human juice collected on a mixture of protease inhibitors (4). Hence, PSP-S transformation takes place in juice. Whether this is a normal process or the consequence of incidental trypsin activity in pancreatic ducts remains to be clarified, however.

The amino acid sequence of PSP-S deduced from the mRNA sequence was in complete agreement with partial sequences already published (41). The carbohydrate moiety of PSP-S is borne by the amino-terminal amino acid stretch, upstream from PSP-S1, whose sequence was not available (4).

Figure 6. mRNA concentrations of trypsinogen, chymotrypsinogen, colipase, and PSP-S in the pancreas of CCP patients and in controls. mRNA concentrations were measured by dot-blot hybridization of equal amounts of RNA from control (Cont.) or CCP tissues (CCP) to ^{32}P-labeled cDNA probes. They were expressed as arbitrary OD_{492} U/μg RNA obtained after scanning the autoradiograms of the blots. o, individual data controls; and □, individual data CCP patients, respectively; ×, mean values (±SE) for each group. ■ in the CCP groups, the patient presenting with a mild form of the disease (see Patients). The significance of the difference in PSP-S mRNA concentration between controls and CCP patients was estimated by Wilcoxon test (52).

These data indicate that the threonine in position 5 seems the most probable O-glycosylation site of the molecule. The carboxy-terminal amino acid predicted by the mRNA sequence is that of PSP-S1. Hence, no carboxy-terminal proteolysis occurs in the posttranslational processing of pre-PSP-S.

A computer search for overall sequence homology confirmed that PSP-S was distinct from the other pancreatic proteins described to date. In particular, no structural evidence could be found supporting the results of Amouric et al. (42) that PSP-S1, called "protein X" by these authors, was in fact a degradation product generated by rapid activation of trypsinogen.

A search conducted by Petersen (43) for partial homologies between PSP-S and sequences from the National Biomedical Research Fund database pointed to an interesting similarity (up to 31% homology) between PSP-S, thrombomodulin and several lectins. Based on that homology, Petersen could predict the positions of two of the three disulfide bridges of PSP-S (Cys 42-Cys 139 and Cys 115-Cys 132), which were eventually confirmed in our laboratory (44). Similar results were reported by Patthy (45). Searching the same database with a slightly different window, we could find in addition a significant homology in the carboxy-terminal regions of PSP-S and several serine proteases. Interestingly enough, the homologous region surrounded a tripeptide involved in the specificity pocket (39) of all tested serine proteases. Studies on PSP-S gene structure may reveal whether it arose from ancestor lectin and serine protease genes by shuffling of exons encoding structural or functional domains from both protein families.

The demonstration that PSP-S was a potent inhibitor of $CaCO_3$ crystal growth in vitro (8) led to the proposition that its function could be to control $CaCO_3$ crystallorhea in juice by selective binding to the sites of crystal growth (46). The presence of such an inhibitor was expected in pancreatic juice that is naturally supersaturated in $CaCO_3$ (10), and similar inhibitors have been described in other supersaturated biological fluids (47–49). Studies showing that PSP-S concentration was lower in the juice of patients presenting with CCP than in controls (11) suggested a relationship between inadequate PSP-S levels and the disease. Immunolocalization studies demonstrating the presence of lesser amounts of PSP-S in the rough endoplasmic reticulum and zymogen granules of CCP acinar cells than in normal tissue (50) indicated that it could be a matter of decreased rate of synthesis. Because changes in protein biosynthesis rates often result from changes in the amounts of the corresponding mRNAs (51), we measured by quantitative filter hybridization the amount of PSP-S mRNA in total pancreatic RNA from controls and CCP patients. It was found three times lower in the latter group, whereas the amounts of trypsinogen, chymotrypsinogen, and colipase mRNA were not altered. Thus, a specific decrease in tissular PSP-S mRNA concentration is associated with CCP. A comparable decrease in the rate of PSP-S biosynthesis is therefore expected, which could account for the decreased PSP-S levels observed in CCP juice.
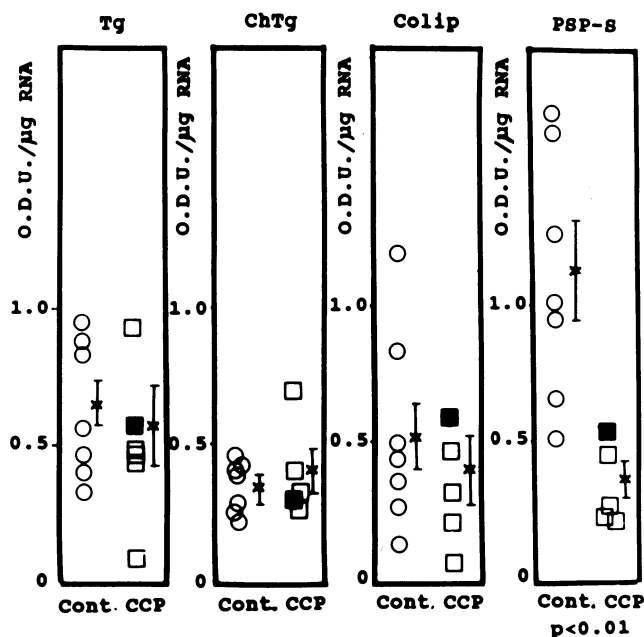
Whether decreased PSP-S gene expression is primary or secondary to the disease is not established. There are, however, indications that the decrease is not a mere consequence of pancreatitis. For instance, PSP-S mRNA levels of patients with obstructive pancreatitis, where the lesions are very similar to those of CCP, were not different from levels of patients without pancreatic disease (Fig. 6). Moreover, in a patient with typical but very mild lesions of CCP and no pancreatic insufficiency, PSP-S mRNA level was still lower than in controls. If this was confirmed, it would suggest that altered PSP-S gene expression confers predisposition to CCP.

Acknowledgments

References

1. Rinderknecht, H. 1986. Pancreatic secretory enzymes. *In* The Exocrine Pancreas: Biology, Pathology, and Diseases. V. L. W. Go., J. D. Gardner, F. P. Brooks, E. Lebenthal, E. P. DiMagno, and G. A. Scheele, editors. Raven Press, New York. 163–183.

2. Montalto, G., J. Bonicel, L. Multigner, M. Rovery, H. Sarles, and A. DeCaro. 1986. Partial amino acid sequence of human pancreatic stone protein, a novel pancreatic secretory protein. *Biochem. J.* 238:227–232.

3. DeCaro, A., J. Lohse, and H. Sarles. 1979. Characterization of a protein isolated from pancreatic calculi of men suffering from chronic calcifying pancreatitis. *Biochem. Biophys. Res. Commun.* 87:1176–1182.

4. DeCaro, A., L. Multigner, J. C. Dagorn, and H. Sarles. 1988. The human pancreatic stone protein. *Biochimie (Paris).* 70:1209–1214.

5. Giorgi, D., J. P. Bernard, A. DeCaro, L. Multigner, R. Lapointe, H. Sarles, and J. C. Dagorn. 1985. Pancreatic stone protein. I. Evidence that it is encoded by a pancreatic messenger ribonucleic acid. *Gastroenterology.* 89:381–386.

6. Rouimi, P., J. Bonicel, M. Rovery and A. DeCaro. 1987. Cleavage of the Arg-Ile bond in the native polypeptide chain of the human pancreatic stone protein. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 216:195–199.

7. Gross, J., R. I. Carlson, A. W. Brauer, M. N. Margolies, A. L. Warshaw, and J. R. Wands. 1985. Isolation, characterization and distribution of an unusual pancreatic human secretory protein. *J. Clin. Invest.* 76:2115–2126.

8. Gross, J., A. W. Brauer, R. F. Bringhurst, C. Corbett, and M. N. Margolies. 1985. An unusual bovine pancreatic protein exhibiting pH-dependent globule-fibril transformation and unique amino acid sequence. *Proc. Natl. Acad. Sci. USA.* 82:5627–5631.

9. Multigner, L., and A. DeCaro. 1987. Pancreatic stone protein. Kinetic studies on calcium carbonate crystal growth inhibition by human pancreatic stone protein. *Digestion.* 38:43–44.

10. Moore, E. W., and H. J. Verine. 1985. Pathogenesis of pancreatic and biliary $CaCO_3$ lithiasis: the solubility product (K'sp) of calcite determined with the Ca^{++} electrode. *J. Lab. Clin. Med.* 106:611–618.

11. Multigner, L., H. Sarles, D. Lombardo, and A. DeCaro. 1985. Pancreatic stone protein. II. Implication in stone formation during the course of chronic calcifying pancreatitis. *Gastroenterology.* 89:387–391.

12. Chirgwin, J. M., A. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 24:5294–5299.

13. Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by affinity chromatography on oligothymidilic acid cellulose. *Proc. Natl. Acad. Sci. USA.* 69:1408–1412.

14. Benton, W. D., and R. W. Davis. 1977. Screening gt recombinant clones by hybridation to single plaques in situ. *Science (Wash. DC).* 196:180–182.

15. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499–560.

16. Woods, D. E., A. F. Markham, A. T. Ricker, G. Goldberger, and H. R. Colten. 1982. Isolation of cDNA clones for the human complement protein factor B, a class III major histocompatibility complex gene product. *Proc. Natl. Acad. Sci. USA.* 79:5661–5665.

17. Whitehead, A. S., G. Goldberger, D. E. Woods, A. F. Markham, and H. R. Colten. 1983. Use of cDNA clones for the fourth component of human complement (C4) for analysis of a genetic deficiency of C4 in guinea-pig. *Proc. Soc. Natl. Acad. Sci. USA.* 80:5387–5389.

18. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6–13.

19. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.

20. Bittner, M., P. Kupferer, and C. F. Morris. 1980. Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzoxyoxymethyl cellulose or nitrocellulose sheets. *Anal. Biochem.* 102:459–471.

21. Bailey, J. M., and N. Davidson. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. *Anal. Biochem.* 70:75–85.

22. Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments to nitrocellulose. *Proc. Soc. Natl. Acad. Sci. USA.* 77:5201–5205.

23. White, B. A., and F. C. Bancroft. 1982. Cytoplasmic dot hybridization. *J. Biol. Chem.* 257:8569–8572.

24. Giorgi, D., J. P. Bernard, R. Lapointe, and J. C. Dagorn. 1984. Regulation of amylase messenger RNA concentrations in rat pancreas by food content. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1521–1524.

25. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *In* Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545.

26. Messing, J., R. Crea, and P. H. Seeberg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* 9:309–332.

27. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

28. Wilbur, W. J., and D. J. Lipman. 1983. Rapid similarity searches of nucleic acids and protein data banks. *Proc. Natl. Acad. Sci. USA.* 80:726–730.

29. Goad W. B., and M. I. Kanehisa. 1982. Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries. *Nucleic Acids Res.* 10:247–263.

30. Lipman D. J., and W. R. Parson. 1985. Rapid and sensitive protein similarity searches. *Science (Wash. DC).* 227:1435–1441.

31. Breathnach, R., and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50:349–384.

32. Bucher, P., and E. N. Trifonov. 1986. Compilation and analysis of eukaryotic POLII promoter sequences. *Nucleic Acids Res.* 14:10009–10026.

33. Padgett, R. A., P. J. Grabowski, M. M. Konarska, S. Seiler, and

P. A. Sharp. 1986. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* 55:1119–1150.

34. Kozak, M. 1984. Comparison and analysis of sequences upstream from the translational start site in eukaryotic mRNA. *Nucleic Acids Res.* 12:857–872.

35. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125–8148.

36. Nevins, J. R. 1983. The pathway of eukaryotic mRNA formation. *Annu. Rev. Biochem.* 52:441–466.

37. von Heijne, G. 1984. Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in procaryotic and eukaryotic cell. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:2315–2318.

38. Scheele, G., and R. Jacoby. 1983. Proteolytic processing of presecretory proteins is required for development of biological activities in pancreatic exocrine proteins. *J. Biol. Chem.* 258:2005–2009.

39. Stroud, R. M., M. Krieger, R. E. Koeppe, A. A. Kossiakoff, and J. L. Chambers. 1975. Structure-function relationships in the serine proteases. *In* Proteases and Biological Control. E. Reich, D. B. Rifkin, and E. Shaw, editors. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 13–32.

40. Scheele, G., D. Bartelt, and W. Bieger. 1981. Characterization of human exocrine pancreatic proteins by two-dimensional isoelectric focusing/sodium dodecylsulphate gel electrophoresis. *Gastroenterology.* 80:461–473.

41. DeCaro, A., J. J. Bonicel, P. Rouimi, J. D. DeCaro, H. Sarles, and M. Rovery. 1987. Complete amino acid sequence of an immunoreactive form of human pancreatic stone protein isolated from pancreatic juice. *Eur. J. Biochem.* 168:201–207.

42. Amouric, M., C. Barthe, C. Kopeyan, C. Figarella, and O. Guy-Crotte. 1987. Protein X, a proteolysis product of human pancreatic juice. Immunological relationship to trypsinogen I. *Biol. Chem. Hoppe-Seyler.* 368:1525–1532.

43. Petersen, T. E. 1988. The amino-terminal domain of thrombomodulin and pancreatic stone protein are homologous with lectins. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 231:51–53.

44. Rouimi, P., J. DeCaro, J. Bonicel, M. Rovery, and A. DeCaro. 1988. The disulfide bridges of the immunoreactive form of human pancreatic stone protein isolated from pancreatic juice. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 229:171–174.

45. Patthy, L. 1988. Homology of human pancreatic stone protein with animal lectins. *Biochem. J.* 253:309–311.

46. DeCaro, A., L. Multigner, H. Lafont, D. Lombardo, and H. Sarles. 1984. The molecular characteristics of a human pancreatic acidic phosphoprotein that inhibits calcium carbonate crystal growth. *Biochem. J.* 222:669–677.

47. Oppenheimer, F. G., D. I. Hay, and C. Franzblau. 1971. Proline-rich proteins from human parotid saliva. Isolation and partial characterization. *Biochemistry.* 10:4233–4238.

48. Hay, D. I., E. C. Moreno, and D. H. Schlesinger. 1979. Phosphoprotein inhibitors of calcium phosphate precipitation from salivary secretions. *Inorg. Perspect. Biol. Med.* 2:271–285.

49. Nakagawa, Y., V. Abram, F. J. Kezdy, E. T. Kaiser, and F. L. Coe. 1983. Purification and characterization of the principal inhibitor of calcium oxalate monohydrate crystal growth in human urine. *J. Biol. Chem.* 258:12594–12600.

50. Lechene de la Porte, P., A. DeCaro, H. Lafont, and H. Sarles. 1986. Immunocytochemical localization of pancreatic stone protein in the human digestive tract. *Pancreas.* 1:301–308.

51. Maniatis, T., S. Goodbourn, and J. A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science (Wash. DC).* 236:1237–1245.

52. Wilcoxon, S. 1947. Probability tables for individual comparisons by ranking methods. *Biometrics.* 3:119–122.