Molecular cloning of mouse PSP mRNA

Hans O. Madsen and J. Peter Hjorth

Department of Molecular Biology, University of Aarhus, C.F. Møllers Allé 130, 8000 Aarhus C, Denmark

Received 19 November 1984; Accepted 14 December 1984

ABSTRACT

PSP is the most abundant translation product of mouse parotid glands where its production is co-ordinated with that of salivary amylase. The synthesis of these two proteins apparently is restricted to this tissue. In order to enable us to study common regulatory elements in the genes of the two proteins, double stranded cDNA, synthesized for parotid gland poly (A)+ RNA, was cloned. DNA sequencing of three clones complementary to the most abundant messenger indicated overlap and resulted in a total sequence of 867 nucleotides. Translation of this sequence revealed that at one end the amino acid sequence was the same as the N-terminal sequence of PSP. The sequence contains 60 nucleotides coding for part of or the complete signal peptide, 645 nucleotides coding for the PSP protein, and 162 nucleotides that apparently are not translated. Southern blot analysis suggests a simple structure for the PSP gene in mouse and man.

INTRODUCTION

In addition to salivary amylase the parotid salivary glands of mice (Mus musculus) produce and secrete large amounts of PSP (parotid secretory protein) the function of which is as yet unknown. On SDS-acrylamide gels PSP appears as a 20 kd band that slowly takes up the protein stain. Its structural gene, Psp, is located on chromosome 2 and as such is unlinked to the salivary amylase locus, Amy-1, on chromosome 3. However, salivary amylase and PSP show the same tissue distribution in that they are produced only in the parotid glands, not in the other salivary glands nor in the pancreatic gland, and they appear in mouse saliva in a constant ratio. This indicates an interesting tissue co-ordination and a common regulation of amylase and PSP synthesis, either at the translation or at the transcription/processing levels. Here we report the initial
steps taken towards clarifying the mechanisms involved in the regulation by studying the cell-free translation of parotid gland mRNA and isolating cDNA clones for PSP mRNA which can be used to study the transcription and gene structure of Psp.

MATERIALS AND METHODS

(a) Isolation of mouse parotid gland poly(A)⁺ RNA

Parotid glands of 20 mice were dissected out and pulverized in liquid nitrogen. Total RNA was isolated by the guanidinium thiocyanate procedure using sedimentation of the RNA through a cushion of CsCl. Partial purification of poly(A)⁺ RNA was obtained using oligo(dT) cellulose chromatography.

(b) In vitro translation of mouse parotid gland poly(A)⁺ RNA

Poly(A)⁺ RNA was translated in a rabbit reticulocyte lysate system with a 60 min incubation at 30 °C. Hybrid arrested translation was performed using EcoRI linearized recombinant plasmid DNA in hybridization to poly(A)⁺ RNA. Translation products were analysed by electrophoresis in a 12% SDS-polyacrylamide gel containing 8 M urea followed by either fluorography or staining with Coomassie Brilliant blue and autoradiography. Relevant bands were cut out of the gel and counted by liquid scintillation.

(c) Construction of a mouse parotid gland cDNA library

cDNA synthesis was performed using oligo(dT) priming and reverse transcriptase in the first strand synthesis and Kornberg polymerase in the second strand synthesis. Single stranded cDNA was analysed by 3.5% polyacrylamide, 8 M urea gel electrophoresis followed by autoradiography and estimation of the relative proportions between discrete signals by scanning. The solution of double stranded cDNA was adjusted to 0.3 M NaCl, 0.03 M NaAc (pH 4.5), extracted once with phenol and CHCl₃ followed by Sephadex G-50 chromatography. Peak fractions were pooled and adjusted to 3 mM ZnCl₂. Double stranded cDNA at a concentration of 25 μg/ml was flush ended by digestion with 3500 units/ml S₁-nuclease (SIGMA) at 37 °C for 15 min. The digestion was stopped by adjusting the solution to 0.5% SDS, and was followed by ethanol precipitation. The S₁-digested cDNA was fractionated by 3.5% acrylamide gel electrophoresis. cDNA
in the size range of approximately 400-1600 bp was electroeluted, ethanol precipitated, and resuspended in distilled water. The cDNA was oligo(dC) tailed while PstI digested pBR327 DNA was oligo(dG) tailed. Conditions were chosen so as to give dC and dG tails approximately 15 nucleotides long. Equimolar amounts of vector and dC tailed cDNA were co-precipitated with ethanol, resuspended in 0.1 M NaCl, 10 mM tris (pH 8.0), 1 mM EDTA, and annealed. The annealed DNA was used to transform E. coli K803. Recombinants were selected on NZ-agar plates supplemented with 5 µg/ml tetracycline.

Colony hybridizations and plasmid preparations using the alkaline extraction procedure were performed as described. Nick-translation and Southern blot hybridization were performed essentially as described by the authors, except that bovine serum albumin was not included in the prehybridization or hybridization mixtures and that the hybridization mixture contained 10% dextran sulphate.

(d) DNA sequence analysis

cDNA inserts of pMPd12, pMPd16, and pMPd39 were recovered by agarose gel electrophoresis onto dialysis tubing, modifying the method by omitting Whatman 3 MM paper. 10 µg of each insert was self-ligated, sheared by sonication, and end repaired using T4 DNA polymerase. Fragments 200-400 bp long were recovered by preparative electrophoresis, blunt-end ligated into the SmaI site of M13mp9 RF DNA, and transfected into JM101TGl, a restriction minus E. coli strain kindly provided by Dr. M. Biggin, Cambridge. Single stranded phage DNA was prepared from colourless plaques and sequences were determined by using a modification of the dideoxynucleotide chain-termination method and using a 15-mer nucleotide from Biolabs as primer. In this way the sequence for both strands of all three cDNAs were obtained.

(e) Isolation and sequence analysis of PSP

0.5 ml saliva was obtained by isoproterenol stimulation of 10 C3H/As females and adjusted to 1 mM EDTA, 1 mM J-CH2COOH and 1 mM PMSF. Amylase was removed from the saliva by chromatography through Sepharose 4B coupled with cyclohepta-amylose. The peak fraction of unbound material was fractionated by
preparative PAGE in a tris-glycine, pH 8.1 buffer \(^1,2^3\). PSP was recovered from the gel by electrodialysis, then extensively dialysed against 5 mM HAc and lyophilized. The N-terminal sequence was determined automatically by a Beckman 890C sequencer \(^2^4\).

RESULTS AND DISCUSSION

(a) Identification of the \textit{in vitro} translation product of PSP mRNA

The 20 kd PSP protein must be about 190 amino acid residues long. The product of \textit{Psp}\(^b\), from mouse strain C3H, has no methionine residue whereas that of \textit{Psp} \(^a\), from YBR, has one or two \(^1\). In addition 15-20\% leucine has been found in PSP purified from saliva by preparative electrophoresis (data not shown). We can demonstrate the presence of PSP mRNA by using the above characteristics of PSP to identify the cell-free translation product. Parotid gland poly(A)\(^+\) RNA was isolated from four mouse strains which represent all possible combinations of the different \textit{Psp} and \textit{Amy-1} alleles of the strains C3H and YBR, as given in Table 1. The poly(A)\(^+\) RNA was translated in a rabbit reticulocyte lysate system \(^6\) in the presence of either \(^3\)H-L-leucine or \(^35\)S-L-methionine. Electrophoresis of the labelled translation products through SDS-acrylamide gels revealed discrete, labelled polypeptides (Fig.1). \textit{In vivo} labelled parotid gland proteins from the same strains were run in parallel. Amylase from the strains C3H and YBR show a slight allelic difference in mobility \(^2^2\). This difference identifies the \textit{in vitro} translation products that turn out to be 1-2 kd larger than the \textit{in vivo} forms. This size difference is in keeping with the 15 residue signal peptide present in pre-

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Mouse strain} & \textbf{Alleles at} & \textbf{in vivo} & \textbf{in vitro} \\
 & \textbf{Amy-1} & \textbf{Psp} & \textbf{RC} & \textbf{RM} & \textbf{RC} & \textbf{RM} \\
\hline
1 & a & b & .124 & .197 & .145 & .217 \\
2 & a & a & .117 & .185 & .112 & .168 \\
3 & c & a & .285 & .452 & .234 & .351 \\
4 & c & b & .211 & .334 & .239 & .358 \\
\hline
\end{tabular}
\caption{Incorporation of \(^2\)H-L-leucine in amylase relative to PSP, RC, and the calculated molar ratio of their synthesis, RM.}
\end{table}
Figure 1. In vivo and in vitro labelled parotid gland proteins from mouse strains 1-4 separated on a 10% acrylamide gel with 8 M urea and visualized by fluorography. For in vivo labelling, mice were starved overnight, fed for 1 h, and injected in the peritoneal cavity with 100 mCi labelled amino acid. After 1 h the parotid glands were removed and homogenized. In vitro labelled proteins were made by cell free translation of parotid gland poly(A)+ RNA. Both 35S-L-methionine and 3H-L-leucine were used as tracers. For each type of labelling, the same amount of radioactivity was applied on the gel for each of the four strains. A and B represent salivary amylase isolated from in vivo 35S-methionine labelled parotid glands from strains 1 and 4. The insert represents a longer fluorography of the same gel showing in vivo 3H-leucine labelled samples.

Amylase 25. When parotid glands are in vivo labelled with 3H-leucine, PSP is the predominating radioactive protein (Fig. 1). In vitro translation in the presence of 3H-leucine produces a major, labelled polypeptide that is 1-2 kd larger. When 35S-methionine was used for labelling this polypeptide incorporated a higher amount of label in those strains (2 and 3) which produce a methionine containing PSP. This indicates that the polypeptide is the translation product of PSP mRNA and suggests that the pre-PSP has a signal peptide of approximately the same
size as that of amylase.

The synthesis of PSP relative to amylase was determined by counting the radioactivity in the relevant bands of the \textit{in vivo} and \textit{in vitro} \(^3\text{H}\)-leucine labelled polypeptides (Table 1). The ratios of the counts were corrected for the number of leucine residues per molecule to give the molar ratios of their synthesis. For this we used the estimated value of 20\% for PSP, i.e., 38 per molecule and 4 per signal peptide whereas salivary amylase has 24 and 4, respectively. In parotid glands of strains (1 and 2) with the C3H amylase allele, \textit{Amy-1}^a, approximately 5 molecules of PSP are synthesized for each amylase molecule both \textit{in vivo} and \textit{in vitro}. In those mouse strains (3 and 4) with the YBR amylase allele, \textit{Amy-1}^c, which results in an \textit{in vivo} doubling of the rate of amylase synthesis \(^1\) approximately 5 molecules of PSP are produced for every 2 molecules of amylase both \textit{in vivo} and \textit{in vitro} (Table 1). This emphasizes that no special regulatory conditions for translation of amylase and PSP mRNAs are present in mouse parotid glands as opposed to rabbit reticulocytes.

2\% salivary amylase mRNA was found in a mixture of salivary glands from A/J \(^27\), a strain that has the same amylase allele as C3H. We can combine this information with the corresponding protein synthesis ratio presented above to give some idea about the amount of PSP mRNA, because as shown later, the translation efficiencies of amylase and PSP mRNAs are comparable. We know that all the salivary amylase mRNA is derived from the parotid glands and as such its concentration there is much higher than the 2\% stated. There must therefore be an correspondingly five-fold higher concentration of PSP mRNA in parotid glands and we have taken advantage of this abundance in the subsequent cloning experiments.

\textbf{(b) Construction of a parotid gland cDNA library}

\(^32\text{P}\)-Labelled single stranded cDNA was synthesized by reverse transcription of parotid gland poly(A\(^+)\) RNA from mouse strain 4 (\textit{Psp}^b, \textit{Amy-1}^c) in the presence of oligo(dT) primer. Denaturing acrylamide gel electrophoresis and autoradiography of the cDNA showed 3 major, discrete compounds (Fig.2).

The two most intense signals corresponding to compounds of
Figure 2. $^{32}$P-Labelled reverse transcript of parotid gland poly(A)$^+$ RNA (S) and HinfI/HindIII endonuclease digest of pBR327 DNA labelled with $^{32}$P by nick translation (M). Samples were denatured with formamide and analysed on a 3.5% acrylamide gel with 8 M urea.

approximately 1600 and 1000 nucleotides are likely to be full-length cDNA products of amylase and PSP mRNAs and indicate that the PSP mRNA contains, in addition to the expected 600 nucleotides which code for PSP and 60 nucleotides which code for the signal peptide, about 300 nucleotides that are not translated. Scanning of the autoradiogram to find the ratio of the 1600 nucleotide compound to the 1000 nucleotide compound gave a value of 0.6, which compares well to the mass ratio of 0.55 derived from the translation molar ratio of 0.37 by correction for the size of the mRNA molecules (1600/1000). This suggests an equal translation efficiency for the two mRNAs.

The single stranded cDNA was converted to double stranded cDNA using Kornberg polymerase. The double stranded cDNA was then treated with S1 nuclease and fractionated by acrylamide gel electrophoresis in order to isolate compounds in the size range of 400-1600 bp. These were then oligo(dC) tailed, annealed with oligo(dG) tailed pBR327 DNA, and used to trans-
Figure 3. Restriction endonuclease cleavage maps of the cDNA inserts of A: pMPdl2, B: pMPdl6, and C: pMPd39. Plasmid vector DNA is shown as open boxes. A third HinfI site in A, positioned between the two shown HinfI sites, was not detected by this analysis but was deduced from the DNA sequence (D). HindIII (†), PstI (♀), HinfI (♂), Sau3A (♀), and HaeIII (♀).

Form E.coli 803. Seventy recombinant clones were isolated.

(c) Identification of PSP cDNA clones

Because of the abundance and size (ca. 1000 nucleotides) of PSP mRNA, recombinant clones were screened by in situ hybridization, using first ³²P-labelled cDNA synthesized by reverse transcription of parotid gland poly(A)⁺ RNA, and then the 1000 nucleotide compound of the same transcription eluted from an electrophoretic acrylamide gel separation. As a control, hybridization was performed using ³²P-nick translated amylase cDNA plasmid pMSa104. Eight clones were selected for further analysis. They were selected on the basis of giving intense hybridization signals with both the mixture of cDNAs from total poly(A)⁺ RNA and the eluted 1000 nucleotide cDNA, but giving no signal with the amylase cDNA. A Southern blot analysis was performed on the PstI digested DNA from these 8 clones, using the ³²P-labelled 1000 nucleotide cDNA compound as probe in the hybridization. The three clones with the strongest signals, pMPdl2, pMPdl6, and pMPd39, were mapped with 4 different restriction endonucleases. The clones had overlapping parts with identical restriction maps (Fig.3) suggesting that they were derived from the same mRNA.
Identification of the three clones was attempted through hybrid arrest of translation. The clones were individually hybridized to samples of parotid gland poly(A)+ RNA. Each sample was then in vitro translated using $^{35}$S-methionine as tracer. When compared to reference samples with either the amylase cDNA clone pMSa104, or without any added cDNA, the samples showed a quantitative reduction of the product previously identified as PSP by SDS-acrylamide gel electrophoresis and autoradiography (data not shown). No reduction was seen when mRNA hybridized to the three cDNAs was liberated by boiling for a short time before in vitro translation. These results indicate that all three isolated cDNA clones contain sequences complementary to PSP mRNA.

(d) Sequence of PSP cDNA clones

PSP isolated by preparative acrylamide gel electrophoresis of mouse saliva was used in determining the amino acid sequence of its N-terminal part by automatic sequencing: N$_2$-Leu,Leu,Gly,Glu,Leu,Gly,Ser,Ala,Val,Asn,Asn,Leu,Lys,Ile,Leu,Asn,Pro,Pro,Ser,Glu.

The cDNA inserts of pMPd12, pMPd16 and pMPd39 were individually sheered, shotgun cloned into the SmaI site of M13mp9 and sequenced. The three clones had identical overlapping regions and were aligned to give a total sequence of 867 bp as shown in Fig.4. In the deduced translation product of the largest open reading frame of this nucleotide sequence, the amino acid sequence of residues 21-40 is identical to the N-terminal sequence found for PSP. The many non-polar amino acids among residues 1-20 suggest that this region constitutes the signal peptide, with the first residue acting as the initiating methionine. This would be in keeping with the estimated size of the signal peptide previously found from cell-free translation.

The first stop codon in this reading frame at position 706, defines the C-terminal of the protein. The deduced size of the processed PSP is therefore 215 residues or 22.963 kd which compares well with the size estimated by SDS-acrylamide gel electrophoresis. The bases 709-867 must then be from the 3'nontranslated region of the mRNA but contain neither signals for poly A addition nor a poly A tail. The messenger is there-
The combined sequence of the inserts of pMPd12, pMPd16, and pMPd39 corresponding to a portion of PSP mRNA. The amino acid translation shown contains a sequence of 20 residues (boxed) which is identical to the N-terminal sequence of PSP. Restriction endonuclease sites are underlined.

**Figure 4**
Table 2. Amino acid composition of PSP as deduced from cDNA sequence

<table>
<thead>
<tr>
<th>Residues per mol.</th>
<th>Per cent</th>
<th>Residues per mol.</th>
<th>Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>11</td>
<td>5.1</td>
<td>Leu</td>
</tr>
<tr>
<td>Arg</td>
<td>2</td>
<td>0.9</td>
<td>Lys</td>
</tr>
<tr>
<td>Asn</td>
<td>20</td>
<td>9.3</td>
<td>Met</td>
</tr>
<tr>
<td>Asp</td>
<td>8</td>
<td>3.7</td>
<td>Phe</td>
</tr>
<tr>
<td>Cys</td>
<td>2</td>
<td>0.9</td>
<td>Pro</td>
</tr>
<tr>
<td>Gln</td>
<td>10</td>
<td>4.7</td>
<td>Ser</td>
</tr>
<tr>
<td>Glu</td>
<td>6</td>
<td>2.8</td>
<td>Thr</td>
</tr>
<tr>
<td>Gly</td>
<td>14</td>
<td>6.5</td>
<td>Trp</td>
</tr>
<tr>
<td>His</td>
<td>0</td>
<td>0.0</td>
<td>Tyr</td>
</tr>
<tr>
<td>Ile</td>
<td>14</td>
<td>6.5</td>
<td>Val</td>
</tr>
</tbody>
</table>

Moreover likely to extend in both directions beyond the part we have cloned. The amino acid composition deduced from the nucleotide sequence and given in Table 2, shows an unusually high (22.8%) content of leucine for PSP in agreement with preliminary direct determination.

Taken altogether, our results show that the cDNAs of pMPd12, pMPd16, and pMPd39 contain the coding sequence for PSP.

(e) PSP genes in mouse and man

The cDNA insert of pMPd39 was $^{32}$P-labelled by nick-translation and used to detect PSP genes by hybridization to Southern blottings of restriction endonuclease digestions of genomic DNA. BamHI digests of mouse liver and human leucocyte DNA preparations gave 2 and 3 discrete hybridization signals, respectively, whereas EcoRI digests gave 3 and 1 (data not shown). The small number of hybridization signals indicates that the cDNA hybridizes only to the non-repetitive part of genomic DNA and that the organization of PSP genes is rather simple.

The relatively large amount of PSP found in mouse saliva suggests that it has an important function. However, despite the shown presence of a PSP gene in humans, no protein with properties comparable to PSP has been described.

(f) Conclusion

We have isolated cDNA clones of an mRNA coding for a signal peptide of at least 20 amino acids and a protein of 215 amino acids which we have identified as being PSP. Since these clones hybridize strongly to a major cDNA component of approxi-
mately 1000 nucleotides, this must be the size of PSP mRNA, leaving about 300 nucleotides untranslated in the messenger. The 1000 nucleotide compound is the main reverse transcript of parotid gland poly(A)+ RNA which means that PSP mRNA is the most abundant mRNA. The ratio of amylase to PSP cDNA transcripts is similar to the ratio of the synthesis of the two proteins, suggesting that the coordinated level of the proteins in saliva is a consequence of a fixed ratio of the corresponding mRNAs. The coordination in expression level of the two proteins is therefore due to regulation of the transcription rate or the RNA processing.

We are now engaged in cloning and sequencing the mouse PSP gene to determine any possible resemblance to the salivary amylase gene that could explain their coordinated regulation.

ACKNOWLEDGEMENT

We thank T.E. Petersen for operating the sequencer, B. Hother Nielsen for typing and M.E. Clark for correcting the manuscript. This work was supported by the Danish Natural Science Research Council.

REFERENCES

15. Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P.