<table>
<thead>
<tr>
<th>Title</th>
<th>Primary structure of human pancreatic α-amylase gene: its comparison with human salivary α-amylase gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Horii, Akira</td>
</tr>
<tr>
<td>Citation</td>
<td></td>
</tr>
<tr>
<td>Issue Date</td>
<td></td>
</tr>
<tr>
<td>Text Version</td>
<td>ETD</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/11094/36665">http://hdl.handle.net/11094/36665</a></td>
</tr>
<tr>
<td>DOI</td>
<td></td>
</tr>
<tr>
<td>rights</td>
<td></td>
</tr>
</tbody>
</table>
Primary structure of human pancreatic α-amylase gene: its comparison with human salivary α-amylase gene

(Recombinant DNA; gene conversion; cosmid; evolutionary mouse and human divergence; exon-intron junctions)

Akira Horii, Mitsuru Emi *, Naohiro Tomita, Takahiro Nishide, Michio Ogawa *, Takesada Mori * and Kenichi Matsubara

Institute for Molecular and Cellular Biology, Osaka University, Yamadaoka, Suita 565 (Japan) and " Second Department of Surgery, Osaka University Medical School, Fukushima-ku, Osaka 553 (Japan) Tel. (06)431-0031

Received 18 June 1987
Revised 11 August 1987
Accepted 18 August 1987

SUMMARY

We have determined the entire structure of the human pancreatic α-amylase (Amy2) gene. It is approx. 9 kb long and is separated into ten exons. This gene (amy2) has a structure very similar to that of human salivary α-amylase (Amyl) gene [Nishide et al. Gene 41 (1986a) 299–304] in the nucleotide sequence and the size and location of the exons. The major difference lies in the fact that amyl has one extra exon on the 5' side. Other differences are at the 5' border of exon 1 and the 3' border of exon 10. The close similarity of these two genes, as compared with mouse pancreatic and salivary amylase genes, suggests that during evolution, the divergence into the two amylase genes may have occurred after the divergence of mice and man.

INTRODUCTION

Human α-amylases which hydrolyze α-1,4 glucosidic bonds, consist of two major isozymes, i.e., salivary type and pancreatic type. These two isozymes differ in molecular size (Matsuura et al., 1978; Stiefel et al., 1973), isoelectric point (Matsuura et al., 1978) and antigenic properties (Boehm-Truitt et al., 1978). Our previous studies on cDNAs of salivary and pancreatic α-amylases (Nakamura et al., 1984; Nishide et al., 1986b) have shown that at least two different, but closely related genes are transcribed in tissue-specific fashion. In addition, salivary α-amylase gene (amyl), but not pancreatic α-amylase gene

Abbreviations: aa, amino acid(s); Amyl, human salivary α-amylase; amyl, gene, cDNA or mRNA coding for Amyl; Amy2, human pancreatic α-amylase; amy2, gene, cDNA or mRNA coding for Amy2; bp, base pair(s); cDNA, DNA complementary to RNA; cos, cosmid; kb, kilobase(s) or 1000 bp; N, any nucleoside; nt, nucleotide(s); p, plasmid; PolI, Klenow (large) fragment of E. coli DNA polymerase I.
(amy2), is expressed in some cancers (Y. Nakamura, N.T., T.N., M.E., A.H., M.O., T.M., G. Kosaki, T. Okabe, M. Fujisawa, N. Osawa, T. Kameya and K.M., manuscript in prep.; N.T., N. Matsuura, A.H., M.E., T.N., M.O., T.M., O. Doi, and K.M., ms. in prep.). Thus, it seems to be of interest to clarify the mechanisms of the specific expression control, and therefore, we started to analyze these genes. This paper reports the structure of human amy2 and compares it with the structure of amyl1 (Nishide et al., 1986a). The results show that these two genes are remarkably similar, albeit they express differently, suggesting that they separated very recently, or that there has been a very efficient gene conversion process to conserve the sequence during evolution.

MATERIALS AND METHODS

(a) Enzymes and reagents

Restriction endonucleases were purchased from Takara Shuzo (Kyoto, Japan) and Toyobo (Osaka Japan). PolIk and Escherichia coli T4 ligase were from Takara Shuzo.

(b) Library screening

The human genomic cosmid library carrying approx. 40–45-kb inserts was kindly provided by Dr. Y. Nakamura (Lau et al., 1983). Recombinant colonies carrying human amyl1 and amy2 were screened using the mixture of 32P-labeled human amyl1 and amy2 cDNAs (Nishide et al., 1986b) as the probe. Selection of a clone (cosP2) that carries the amy2 is described in the text.

(c) Mapping of cosmid clones

Each cosmid clone was digested with some restriction endonucleases, electrophoresed in 0.5% or 0.35% agarose gels (Maniatis et al., 1982), transferred and hybridized to 32P-labeled cDNA, or to 32P-labeled fragment of the cosmid vector DNA, according to Southern (1975). We also used double digestion or partial digestion for fine mapping.

(d) Sequence analysis

DNA fragments were extracted from the cosmid clone (cosP2), subcloned in pUC8 (Vieira et al., 1982) or in Bluescribe (Vector Cloning Systems), and in M13 vectors mp10, mp11 (Messing, 1983), mp18 or mp19 (Yanisch-Perron et al., 1985), and sequenced by the dideoxy method (Sanger et al., 1980; Hattori et al., 1986).

(e) Poly(A)+RNA preparation

A resected human pancreas was frozen in liquid nitrogen and stored at -70°C. RNA was isolated as described (Nakamura et al., 1984). Poly(A)+ RNA was purified by oligo(dT)-cellulose column chromatography.

(f) Primer extension analysis

Primer extension was carried out using pancreatic poly(A)+ RNA as a template and a 5' end-labeled synthetic oligodeoxynucleotide primer (5'-CCAG-CAGAACCCAATGGT-3'). This primer is complementary to a sequence of pancreatic α-amylase cDNA covering codons 9 through 14. Analysis was carried out in a 6% polyacrylamide–8 M urea gel along with appropriate dideoxy sequencing samples as size markers.

RESULTS AND DISCUSSION

(a) A cosmid carrying the human pancreatic α-amylase gene

The human genomic cosmid library was screened with human amyl1 and amy2 cDNA probes (Nishide et al., 1986b). Out of 1.2 × 105 colonies screened, nine clones carrying full-length DNA covering an amylase gene were obtained. Eight of them gave an identical restriction cleavage pattern with that of salivary α-amylase gene (Nishide et al., 1986a). One, which we named cosP2, gave a different cleavage map. As will be shown below, the cosP2 carries the DNA that covers the entire human pancreatic α-amylase gene.
Fig. 1. Restriction map of human genes \textit{amy2} and \textit{amy1}. Relative positions of exons, deduced from the nucleotide sequence, are indicated by solid boxes. B, \textit{Bam} III; E, \textit{Eco} RI; G, \textit{Bgl} II; H, \textit{Hind} III; Hpa, HpaI; N, \textit{Bam} III; P, \textit{Pst} I; S, \textit{Sac} I; X, \textit{Xba} I. Only relevant restriction sites are displayed. A few errors in the restriction map previously published (Nishide et al., 1986a) have been corrected. The numbering of the exons in \textit{amy1} has been revised so as to unify the nomenclatures for the two genes. The primer used for extension analysis is also shown, with its 5'-3' orientation, above the exon 1 of \textit{amy2}.

(b) The structure of \textit{amy2}

The insert in the cosP2 was analyzed by restriction endonucleases. The fragments were monitored by their ability to hybridize to \textit{amy1} and \textit{amy2} cDNA regional probes (Nakamura et al., 1984; Nishide et al., 1986b). The fragments were subcloned for fine restriction mapping and nucleotide sequencing. The restriction map and location of exons (see below) are shown in Fig. 1, and the nucleotide sequence, along with the predicted amino acid sequence, is shown in Fig. 2. The nucleotide sequence of the amino acid coding region, along with its 5'- and 3'-flanking regions which were regarded to be exons, completely agrees with that of human pancreatic \(\alpha\)-amylase cDNA (Nishide et al., 1986b). These sequences also agree with the cDNA sequence reported by Wise et al. (1984), except for 1 nt (thymine instead of cytosine) in exon 8 at aa position 370. These results indicate that the gene we analyzed in cosP2 is actually the human pancreatic \(\alpha\)-amylase gene.

Fig. 1 also shows the map and location of exons in \textit{amy1}. The two genes are very similar in structure including the locations and lengths of exons as well as introns. One of the major differences lies in the fact that \textit{amy1} has an extra exon, referred to as exon S*, which is lacking from \textit{amy2} (see below). The 5' border of exon 1 and the 3' border of exon 10, each differ to some extent between \textit{amy2} and \textit{amy1}. The difference in the 5' border of exon 1 (an extra 32 nt is transcribed in the exon 1 of the \textit{amy1}) can be used to monitor the presence of a small amount of \textit{amy2} mRNA among large amounts of \textit{amy1} mRNA (Y. Nakamura, N.T., T.N., M.E., A.H., M.O., T.M., G. Kosaki, T. Okabe, M. Fujisawa, N. Osawa, T. Kameya and K.M., ms. in prep.)

In addition to the overall homologies between transcribed sequences in \textit{amy2} and \textit{amy1}, we see striking similarities in the size of introns and the sequences that cover the 5' and 3' flanking regions of the two genes. The 5' untranscribed region of \textit{amy2} which corresponds to the exon S of \textit{amy1} is 93% homologous to the exon S. This raised the question as to whether \textit{amy2} is transcribed in this region. Accordingly, we determined the transcription start point of \textit{amy2} by primer extension analysis, as shown in Fig. 3, using the mRNA from human pancreas. The result showed unambiguously that the transcription starts at 14 bp upstream from the start codon in the exon 1. S1 nuclease mapping confirmed this result (not shown). Therefore, the homologous sequence to the exon S in \textit{amy2} is not transcribed.

The TATA box of \textit{amy1} lies 28-28 nt upstream from the transcription start point which lies in the exon 1 in \textit{amy1}. The TATA box of \textit{amy2} lies farther upstream from the exon S (34-28 nt upstream from the cap site of \textit{amy1}).

* The exons in \textit{amy1} were renumbered so as to unify the nomenclature of corresponding exons in the two genes (viz., \textit{amy1} and \textit{amy2}). Thus, \textit{amy1} consists of exons S, 1, 2, 3,... 9, 10 and \textit{amy2} consists of exons 1, 2, 3,... 9, 10.
Fig. 2. The nucleotide and predicted amino acid sequences of human amy2 and amy1. The entire sequence for amy2 is shown. Only the differing nucleotides or amino acids in amy1 are presented beneath amy2 sequence. Exons are shaded and boxed with dashed lines, and the putative TATA boxes and poly(A) signals are boxed with solid lines. A few errors in the sequences of nucleotides and predicted amino acids in amy1 previously published (Nishide et al., 1986a,b) have been corrected. Dashes in the nucleotide sequences indicate deletions.
Fig. 3. Primer extension analysis. Lanes A, G, C and T show the known nucleotide sequence ladders employed as size standards (30–70 nt). Lane P shows the cDNA primed with a 32P-5’-end-labeled 18-nt synthetic oligodeoxynucleotide (see MATERIALS AND METHODS, section f, and Fig. 1) using pancreatic poly(A)+ RNA as a template. The arrow indicates the primer extension product, 56 nt in size, containing a 42 nt coding region and a 14-nt 5’ nontranslated region.

(c) Comparison of amylase genes between man and mouse

Schibler et al. (1982) published structures of α-amylase genes in mouse. They showed that in this animal the amy2 gene consists of ten exons and the amyl gene of eleven. In addition to the pancreas and salivary gland, α-amylase can be produced in the liver, whose gene consists of eleven exons. The structures of mouse and human α-amylase genes are similar from exon 1 through exon 10, which is equivalent to amy2 (Fig. 4). These four genes show close similarity in sizes of exons 1 through 10. In this region, the nucleotide sequences of mRNAs representing amy2 exons are 85% homologous between man and mice, and 81% homologous, with respect to amyl exons (Hagenbüchle et al., 1980). Fig. 4 shows that the sizes of introns in human and mouse amylase genes differ to some extent. To our surprise, introns in human pancreatic and salivary amylase genes also show very close similarities. As far as we sequenced, the nucleotide sequences are more than 90% homologous. The extra exon (exon S) in the salivary amylase genes in both man and mice have no sequence homology. Moreover, their size and location show no similarities. Thus, the size of exon S in man is 154 bp and is located 357 bp upstream from the 5’ end of exon 1, whereas in mice, the size is 50 bp and the location is 7.77 kb upstream from the 5’ end of exon 1 (Schibler et al., 1982). It thus seems likely that at some stage of evolution, the exon S in both species was created by mutation in the upstream region to the basal ten exons which may represent the primordial amylase gene and is similar to amy2. The expression of amylase in mouse liver (Schibler et al., 1980; 1982; 1983; Young et al., 1981) is also mediated through an additional exon, called exon L, that is located at the 5’ side of the basal ten exons. Acquisition of such an extra exon for expression in the salivary gland (or in the liver) in both mammalian species is interesting for the tissue-specific expression of a gene, and suggests that similar mechanisms may be found in some other genes, particularly those coding isozymes that are expressed in tissue-specific fashion.

The nucleotide sequence homology between the basal ten exons in amyl and amy2 is 98% in humans. Not only are the lengths and locations of the exons similar, but also the introns and flanking regions. Close similarity between the basal ten exons is also observed in amyl and amy2 in mice (89% homologous). On the other hand, the overall nucleotide sequence homology between corresponding exons from the two mammals is 81% (salivary type) and 85% (pancreas type). These data suggest that during evolution, the divergence into the two amylase genes (viz. amyl1 and amy2) occurred after the divergence of mice and man. Alternatively, the highly conserved sequence homologies within a species, may be accounted for by gene conversion among the members of the amylase gene family.

Swift et al. (1984), Boulet et al. (1986) and Osborn et al. (1987) found a conserved sequence, possibly a part of the enhancer sequence that plays a role in pancreas-specific gene expression. The corresponding 5′ region of human amy2 has the almost identical sequence (AGGTCAATTAGATGATTT-CCATGAGAGCTT) at 153–183 bp upstream from the cap site of amy2. Interestingly, the human
Fig. 4. Comparison of human and mouse pancreatic and salivary α-amylase genes. Exons are represented by blackened boxes and introns by open boxes. Numbers denote the size (in kb) of exons or introns.
amy1 has this sequence, too, except for substitution of 9th T by G in intron S. The 5'-flanking region of amy2 is 94% homologous to the exon S of amy1, as shown in Fig. 2. Our unpublished data show that up to at least 0.6 kb from the cap site of human amy1, the sequences for amy1 and amy2 are very similar. Thus, from sequencing studies alone, it is not possible to point out the cis-acting control regions for tissue-specific expression. Further analyses are needed to clarify these problems.

ACKNOWLEDGEMENTS

We thank Dr. Y. Nakamura for providing us with the human DNA cosmid library, Drs. T. Yamamoto and Y. Takahashi for their advice and discussions, and Mr. T. Umekage for his technical assistance.

REFERENCES


Communicated by H. Yoshikawa.