J Mol Evol (1996) 42:543–551

JOURNAL OF MOLECULAR EVOLUTION

Cloning and Sequencing Analysis of Three Amylase cDNAs in the Shrimp *Penaeus vannamei* (Crustacea decapoda): Evolutionary Aspects

Alain Van Wormhoudt, Daniel Sellos

Laboratoire de Biologie Marine du Collège de France, BP 225, 29182 Concarneau, France

Received: 22 August 1995 / Accepted: 25 October 1995

Abstract. In *Penaeus vannamei,* α-amylase is the most important glucosidase and is present as at least two major isoenzymes which have been purified. In order to obtain information on their structure, a hepatopancreas cDNA library constructed in phage lambda-Zap II (Strategene) was screened using a synthetic oligonucleotide based on the amino acid sequence of a V8 staphylococcal protease peptide of *P. vannamei* α-amylase. Three clones were selected: AMY SK 37 (EMBL sequence accession number: X 77318) is the most complete of the analyzed clones and was completely sequenced. It contains the complete cDNA sequence coding for one of the major isoenzymes of shrimp amylase. The deduced amino acid sequence shows the existence of a 511-residue-long pre-enzyme containing a highly hydrophobic signal peptide of 16 amino acids. Northern hybridization of total RNA with the amylase cDNA confirms the size of the messenger at around 1,600 bases. AMY SK 28, which contains the complete mature sequence of amylase, belonged to the same family characterized by a common 3′ terminus and presented four amino acid changes. Some other variants of this family were also partially sequenced. AMY SK 20 was found to encode a minor variant of the protein with a different 3′ terminus and 57 amino acid changes.

Phylogenetic analysis established with the conserved amino acid regions of the (β/α) eight-barrel domain and with the total sequence of *P. vannamei* showed close evolutionary relationships with mammals (59–63% identity) and with insect α-amylase (52–62% identity). The use of conserved sequences increased the level of similarity but it did not alter the ordering of the groupings. Location of the secondary structure elements confirmed the high level of sequence similarity of shrimp α-amylase with pig α-amylase.

Key words: α-Amylase isoenzymes — cDNA nucleotide sequence — Invertebrate — Crustacea

Introduction

Alpha amylase (α-1,4 glucan-4-gluconohydrolase, EC 3.2.1.1) catalyzes the hydrolysis of α-D (1,4) glucosidic linkage of starch components, glycogen, and related oligosaccharides. Since the first primary structure deduced from the rat pancreatic cDNA (Pictet et al. 1981), more than 50 complete amino acid sequences from different sources have been determined (Janecek 1994a). Known sequences deal mainly with prokaryotic amylases, vegetals, and some vertebrate enzymes including porcine, rat, and human. In contrast, studies of α-amylase from invertebrates remain limited and only fruit fly and mosquito amylase sequences have been determined (Boer and Hickey 1986; Grossman and James 1993). Although α-amylases have the same function, sequences present a high level of variability, which accounts for less than 10% of identity between bacteria and mammals (Nakajima et al. 1986). Alpha amylases are, however, characterized by a constant (β/α) eight-barrel domain which characterizes also starch hydrolases and related enzymes

Correspondence to: A. Van Wormhoudt

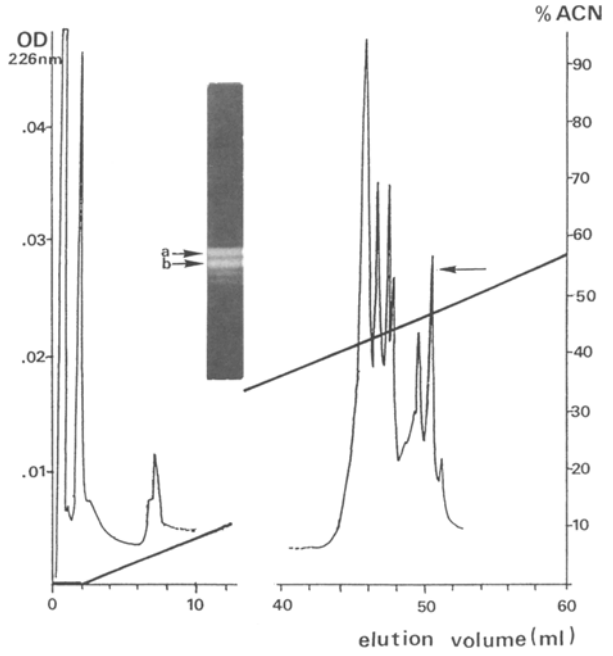

Fig. 1. High performance liquid chromatography elution profile of *Penaeus vannamei* α-amylase hydrolysis products obtained by staphylococcal protease digestion. Separation was achieved on a Vydac column (Cluzeau) with a linear gradient of acetonitrile in the presence of 0.1% TFA. The absorbance was monitored at 226 nm. One peptide (*arrow*) was selected and sequenced (CNRS-Vernaison, France). **Insert:** Revelation of purified amylase activity (two major isoforms were indicated). No significative absorbance was shown between 10 and 40 ml.

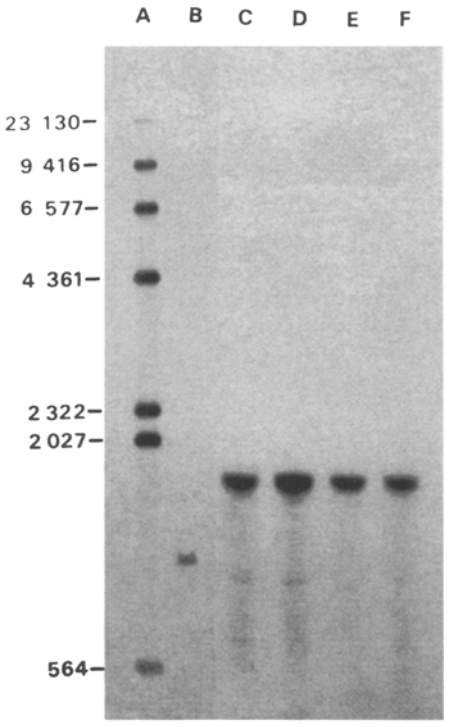

Fig. 2. Autoradiography of a Northern blot of shrimp hepatopancreas total RNA following hybridization with chymotrypsin and amylase cDNA radioactive probes. *Lane A* is lambda DNA *Hin*dIII fragments end-labeled using ^{32}P dNTP and the Klenow fragment of the DNA polymerase I. *Lane B* is shrimp hepatopancreas total RNA hybridized with a ^{32}P-labeled 520-bp-amplified fragment coding for the chymotrypsin (1,068 bp: Sellos and Van Wormhoudt 1992). *Lanes C–F* are shrimp hepatopancreas total RNAs extracted from individual animals and hybridized with a ^{32}P amylase cDNA probe.

(Jespersen et al. 1993). It consists of adjacent regions which may vary either at slow or at more rapid speed, providing sequences suitable for short-range as well as long-range comparisons.

Arthropods are largely polyphyletic (Cisne 1974), and there is still no agreement concerning the origin of crustaceans and their relationships to the other major arthropod groups. During the last decade, mitochondrial DNA and 18S ribosomal RNA have been widely used as genetic markers to relate evolution and population studies but reports on Crustacea are still scarce and cannot alone provide a definitive phylogenetic answer (Palumbi and Benzie 1991; Kim and Abele 1990; Ovendeen et al. 1992). The use of α-amylase sequences will give new insights not only into evolution studies but also into the relationships between their structures and their functions.

In the crustacean Decapoda, a high degree of polymorphism was recently reported for α-amylase in 40 different species (Van Wormhoudt et al. 1995), which included the penaeides, which represent the world's largest number of crustacea of economic importance. Penaeides, which belong to the Dendrobranchiata suborder (Burkenroad 1963; Felgenhauer and Abele 1983; Abele and Felgenhauer 1986; Kim and Abele 1990), are regarded, in terms of morphological and molecular phylogeny, as the most primitive group of the Decapoda (Abele 1991). Moreover, they present one of the smallest isoform numbers (Van Wormhoudt et al.1995). Schram (1982) places the origin of Penaeides in the Carboniferous, but the first fossils were found in the lower Jurassic. *Penaeus* probably arose in the Paleozoic (Glaessner 1969) and is the only genus that has survived through the Cenozoic.

In the shrimp *Penaeus vannamei,* α-amylase is one of the major products of the hepatopancreas. It accounts for approximately 1% of the total protein extract and is associated with the digestive function of this tissue. Shrimp hepatopancreatic α-amylase is a single-chain polypeptide whose amino-terminal end is blocked. It is present as two major equally active forms (a and b) and it has minor components with the same apparent molecular weights on denaturating gel electrophoresis of about 30 kDa (Van Wormhoudt et al. 1995) compared to 55 kDa for amylases from other sources (Keller et al. 1971). These isoenzymes showed important variations during intermolt stages and nutrition experiments (Le Moullac 1995). A repression of the expression of the isoform (a) was reported in dietary experiments while casein increased (Van Wormhoudt et al. 1996).

In this paper we report for the first time on the mo-

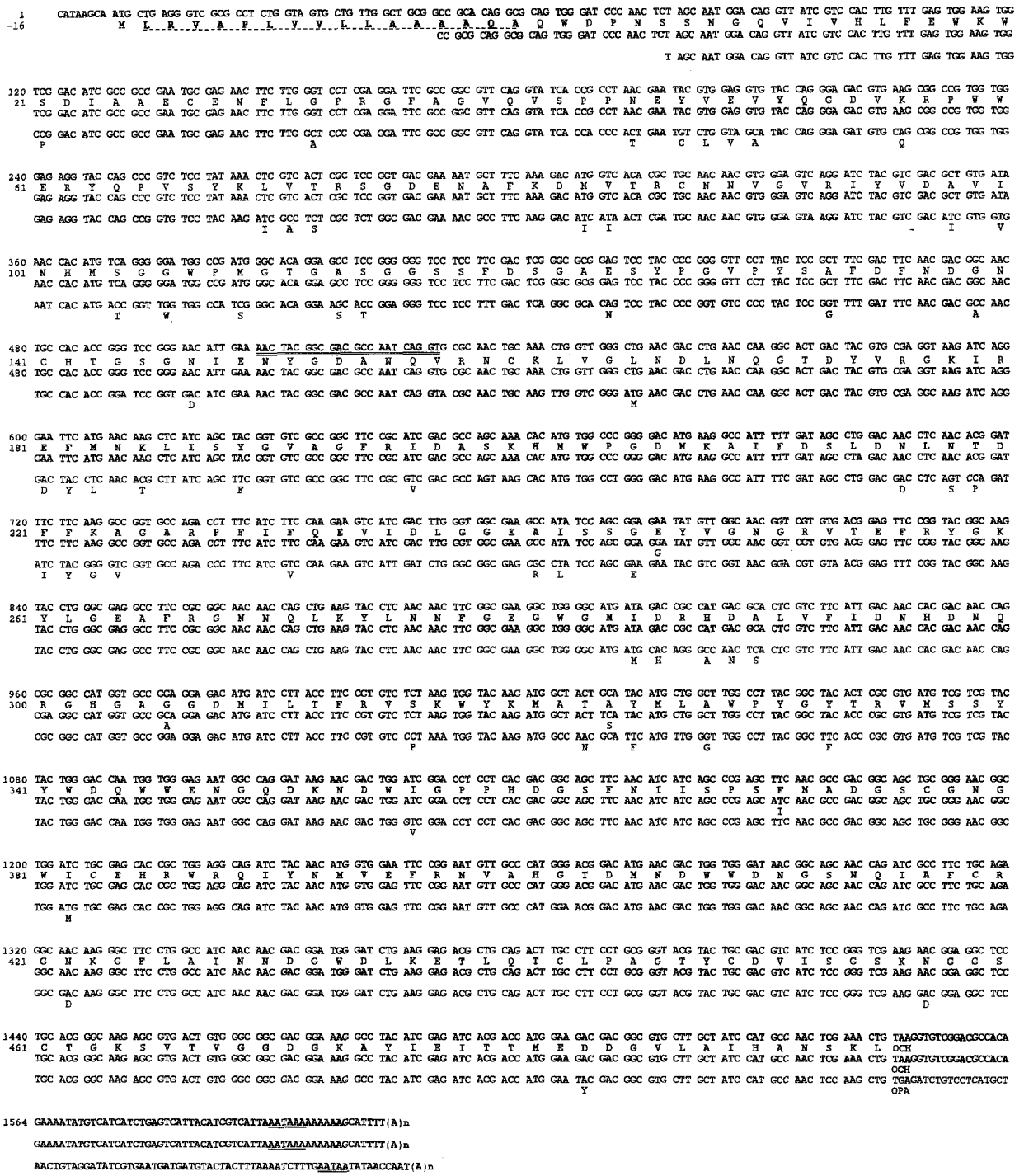

Fig. 3. Nucleotide sequences of shrimp hepatopancreas amylase cDNAs and deduced amino acid sequences. The first sequence is the largest cDNA (AMY SK 37: accession No. EMBL X77318). Next is the sequence of AMY SK 28. The last sequence is AMY SK 20. The amino acid sequence is *numbered* sequentially from the first amino acid of the mature protein. The proposed signal peptide is in *bold and underlined with a dotted line.* When a nucleotide modification in the two other cDNAs results in an amino acid change, the residue is indicated. The nucleotide sequence corresponding to the synthetic oligonucleotide used in the screening procedure and sequencing experiment is *double underlined.* The proposed polyadenylation site is *underlined.*

lecular cloning of the three amylase cDNAs of a marine invertebrate: the shrimp *P. vannamei.* The structure of the encoded proteins and an analysis of phylogenetic relationships (mainly those with insects and mammals) are also given.

Materials and Methods

Shrimp. Shrimp (*Penaeus vannamei*) were obtained from IFREMER (Brest). The hepatopancreas was removed by dissection, immediately frozen in liquid nitrogen, and stored at −80°C until used.

Table 1. Conserved stretches in some amylases enzymes[a]

	I (β_1)	II (β_2)	III (β_3)	IV (loop3)	V (loop3)
S37	16 FEW	35 GFAGVQVSP	95 DAVINH	139 NYGD	154 LNDLN
S28	16 FEW	35 GFAGVQVSP	95 DAVINH	139 NYGD	154 LNDLN
S20	16 FEW	35 GFAGVQVSP	95 DIVVNH	139 NYGD	154 MNDLN
DME	17 FEW	34 GYAGVQVSP	97 DVVFNh	150 NYND	165 LADLL
DSE	17 FEW	35 GFAGVQVSP	97 DVIFNH	143 NYND	158 LRDLN
DER	17 FEW	28 GFAGVQVSP	83 DVVFNH	126 NYND	142 LRDLN
DTE	17 FEW	34 GFAGVQVSP	97 DVVFNH	150 NYSD	165 LRDLN
MSQ	17 FEW	34 GYGGVQLSP	97 DIIINH	150 DWGN	165 LHDLN
AED	17 FEW	35 GYAGVQVSP	96 DIVINH	150 NYND	165 LPDLA
MPA	17 FEW	35 GFAGVQVSP	96 DAVINH	146 NYQD	161 LLDLA
HPA	17 FZW	35 GFGGVQVSP	95 DAVINH	146 NYND	161 LLDLA
PPA	17 FEW	35 GFGGVQVSP	95 DAVINH	136 SYND	151 LLDLA
RPA	17 FEW	35 GFGGVQVSP	95 DAVINH	139 NYND	154 LLDLA
THE	17 FQW	35 GFGAVQVSP	96 DAVINH	150 NWND	165 LADLK
SHY	17 FER	35 GYGYVEVSP	96 DAVVNH	150 DYTN	165 LADLG
SLI	17 FEW	35 GYGYVQVSP	96 DSVINH	150 DYGN	165 LADLD
ALT	16 FEW	35 GYAAVQVSP	96 DTLINH	150 DYGN	165 LADLD
	*	* * **	* **		**

[a] The second line denotes the elements of secondary structure determined by HCA. The enzymes are numbered from the N-terminal end of mature protein. The numbers represent the start of the consensus sequences. The asterisks signify invariable amino acid residues. S37 = *Penaeus vannamei* (hepatopancreas isoform): X77319 EMBL; S28 = *P. vannamei* (hepatopancreas isoform); S20 = *P. vannamei* (hepatopancreas isoform); DME = *Drosophila melanogaster,* P08144 SWISSPROT; DSE = *D. sechellia,* DROAMYSECP GENBANK; DER = *D. erecta,* DROAMYERED; DTE = *D. teissieri,* DROAMYTEID; MSQ = *Anopheles gambiae,* L14753; AED = *Aedes aegyti,* P13080; MPA = *Mus musculus* (mouse pancreas), P00688; HPA = *Homo sapiens* (human pancreas), P04746; PPA = *Sus scrofa* (pig pancreas), P00690; RPA = *Rattus norvegicus* (rat pancreas), P00689; THE = *Thermomonospora curvata,* P29750; SHY = *Streptomyces hygroscopicus,* P08486; SLI = S. limosus, P09794; ALT = *Alteromonas haloplanctis,* Ps9957

Protein Hydrolysis and Peptide Purification. Crude extracts were purified by affinity chromatography on a reticulated starch gel column (Van-Wormhoudt 1983). Pure enzyme was obtained by FPLC. The column (MonoQ, Pharmacia) was eluted with a NaCl gradient in 10 mM Tris-HCl pH = 7.6. Samples were examined by polyacrylamide slab gel electrophoresis (Davis 1964). Amylase activity of isoenzymes was determined on an amylopectine (1%)-agar (1%) gel as described previously (Van Wormhoudt and Favrel 1988). After boiling 50 μg of the protein in the presence of 1% SDS, during 5 min, the hydrolysis of amylase was performed overnight at 37°C with 10 μg of V8 staphylococcal protease (Sigma).

Separation of the obtained peptides was achieved on a Vydac HPLC column (Cluzeau), with a linear gradient of acetonitrile in the presence of 0.1% trifluoroacetic acid. A LKB-Pharmacia HPLC system was used.

RNA Isolation and Northern Blot Characterization. Total RNA extractions were made following the guanidine thiocyanate method (Chirgwin et al. 1979). Digestive glands were disrupted in liquid nitrogen with a grinder. The powdered tissue was immediately dissolved in a 4 M guanidine thiocyanate solution in 12.5 mM Tris HCl pH 7.6 with additional 12.5 mM EDTA and 0.1 M mercaptoethanol (5 ml per g of tissue). After complete dissolution, the solution was centrifuged at 12,000*g* for 10 min and one-tenth of the volume of 20% Sarkosyl was added to the supernatant. After heating the solution to 65°C for 2 min, 0.1 g of cesium chloride was added per ml of solution. This was then layered onto a cushion of 5.7 M cesium chloride in 0.1 M EDTA, pH 8, and centrifuged at 20,000 rpm for 22 h at 20°C in a Beckman centrifuge with a SW 25 rotor. The pellet obtained was dissolved overnight at 4°C in a solution of 5 mM EDTA, pH 8 with additional 5% Sarkosyl and 5% mercaptoethanol (1 ml/g). This solution was then extracted with 1 vol of phenol/chloroform/isoamylol (25/24/1) followed by chloroform and then precipitated in the presence of 0.3 M sodium acetate and ethanol (3 vol). Poly-A RNAs were purified by affinity chromatography on oligo dT-cellulose column. They were denatured with deionized glyoxal and DMSO and characterized by electrophoresis on a 1.5% agarose gel. The transfer of RNA from the gel to the membrane (Hybond N^+, Amersham) was performed passively in the presence of 10 × SSC overnight.

cDNA Library and Screening Procedures. A lambda ZAP c-DNA library for shrimp hepatopancreas was established following the manufacturer's protocol (Stratagene, La Jolla, CA, USA). The unamplified cDNA library, containing 5.8×10^6 independent phages, was amplified to 1.5×10^{10} phages per ml and screened first with a synthetic oligonucleotide

(AAT/CTAT/CGGI GAT/C GCN AAT/C CAA/G GT)

based on the N-terminus sequence of peptide (NYGDANQV) obtained after V8 staphylococcal hydrolysis of the purified shrimp amylase. Plaques were transferred to hybond N^+ membranes (Amersham) and screened with the oligonucleotide end-labeled with gamma ^{32}P ATP and the polynucleotide kinase. In order to have complete cDNAs, a second screening was performed using a 236-bp 5′ fragment of a cloned cDNA probe we obtained from the first screening and covering nucleotides 109–345. This DNA fragment was labeled with the random priming kit from Biolabs using ^{32}P dATP. Prehybridization of the duplicate membranes was achieved in this case in a 50% formamide solution containing 1% SDS, 1 M NaCl, and 100 μg yeast RNA per ml, for 4 h at 42°C. For hybridization, the denatured probe was added (10^6 cpm/ml) for 16 h at the same temperature. The filters were washed twice in 2 × SSC for 5 min at room temperature, then twice in 2 × SSC containing 1% SDS at 65°C for 30 min, and finally twice in 0.1 × SSC for 30 min at room temperature.

Table 1. Continued

VI (β_4)	VII (β_5)	VIII (β_7)	IX (β_8)
182 GFRIDASKH	223 EVID	283 FIDNHD	323 GYTRVMSSY
182 GFRIDASKH	223 EVID	283 FIDNHD	323 GYTRVMSSY
182 GFRVDASKH	223 EVID	283 FIDNHD	323 GFTRVMSSY
193 RDLNDASRH	233 EVID	293 FVDNHD	333 GTPRVMSSF
186 GFRVDAAKH	224 EVID	279 FVDNHD	314 GTPRVMSSF
170 GFRVDAAKH	200 EVID	259 FVDNHD	298 GTPRVMSSF
193 GFRVDAAKH	233 EVID	293 FVDNHD	333 GTPRVMSSF
193 GFRVDAAKH	233 EVID	293 FVDNHD	333 GQLRIMSSF
193 GFRVDAVKH	233 EVID	295 FVDNHD	335 GIVRIMSSF
189 GFRLDASKH	230 EVID	291 FVDNHD	327 GFTRVMSSF
189 GFRIDASKH	230 EVID	291 FVDNHD	327 GFTRVMSSY
179 GFRIDASKH	220 EVID	280 FVDNHD	320 GFTRVMSSY
182 GFRLDAAKH	223 EVID	283 FVDNHD	323 GFTRVMSSY
193 GFRIDAAKH	233 EVIA	295 FVVNHD	335 GTPKVMSSY
193 GFRIDAAKH	233 EVIY	295 FVDNWD	335 GSPNVYSGY
193 GFRIDAAKH	233 EAIH	295 FVDNHD	335 GSPDVHSGY
193 GFRFDASKH	233 EVID	295 FVDNHD	335 GYPKVMSSY
** **	* *	* **	* *

Plasmid Subcloning and DNA Sequencing. The recombinant clones which hybridized with the specific probes were selected and isolated with successive cycles of purification. The phagemids contained in the selected phages were excised following the manufacturer's protocol. The complete cDNA insert (clone AMY SK 37) and the inserts that encode the two other variants of amylase (clones AMY SK 28 and AMY SK 20) extracted from low-melting-point agarose gel and purified with Geneclean kit (Bio Rad) were deleted using *Sac*II and *Bgl*II for AMY SK 37) and AMY SK 20 or *Sac*II and *Pst*I for AMY SK 28 and subcloned in Bluescript SK and KS minus (Stratagene) in order to allow sequencing on both strands. To confirm the nucleotide order in the middle of the antisense strand, sequencing was performed using the synthetic oligonucleotide already used for the screening. Single-stranded DNA was sequenced following the dideoxynucleotide method with modification for extended DNA sequencing with the large fragment of DNA polymerase I.

Structure Analysis. Alpha-amylase sequence was converted into hydrophobic cluster analysis (HCA plot) and compared to pig pancreatic enzyme where the position of the structure elements forming a (β/α) eight-barrel were deduced from crystallography (Raimbaud et al.1989).

Sequence Alignment. The amino acid multialignment between the various sequences presenting the highest degree of homology was done with the Treealign program (Hein 1990). Most sequences were extracted from the SWISSPROT protein sequence data bank. Two unrooted distance trees were calculated on total mature protein sequences and on the 55 amino acid sequences resulting from alignment of the nine consensus domains.

Results and Discussion

Purification and Partial Sequencing of α-Amylases

In crude extracts and after purification, α-amylase is present as two major isofoms, named a and b, which are equally represented. Minor bands are also observed (inset Fig. 1). The pattern of the digestion products is given in Fig. 1. The sequence of the most hydrophobic peptide, generated by staphylococcal hydrolysis of the three variants, is (N-Y-G-D-A-N-Q-V-R-N-C-K-L-V-G-L-N-D-L-N-Q-G-T-D-Y-V-R).

cDNA Library Screening

An oligonucleotide derived from the first eight amino acids of this peptide was used for the primary screening of the library as described in Materials and Methods. Only 0.2% of the phage plaques were positive. Using the *Eco*RI/*Sal*I cDNA fragment as a probe, for the secondary screening, the obtained inserts, ranging from 1,400 to 1,620 base pairs, showed three different types of organization. The longest cDNA, named cDNA1 and represented by clone AMY SK 37, showed an *Xho*I site at position 157 and two *Eco*RI sites at positions 600 and 1242. The second cDNA, referred to as cDNA2 and represented by clone AMY SK 28, has only the *Xho*I site, and the first *Eco*RI site is located at position 600. The third cDNA, referred to as cDNA3 (clone AMY SK 20), lacks the *Xho*I and the two *Eco*RI sites. All these clones were sequenced to confirm this organization.

Northern Analysis

The autoradiographic pattern is given in Fig. 2. One major band, 1,600 bp in length, hybridizes with the ^{32}P-labeled amylase cDNA probe. This size is in accordance with the size of the longest cDNA cloned and confirms the presence of a very short 5′ untranslated region for all the amylase mRNA.

cDNA Analysis

The complete cDNA 1 sequence is 1,620 base pairs long (without the poly A tail). The two other longest cDNAs

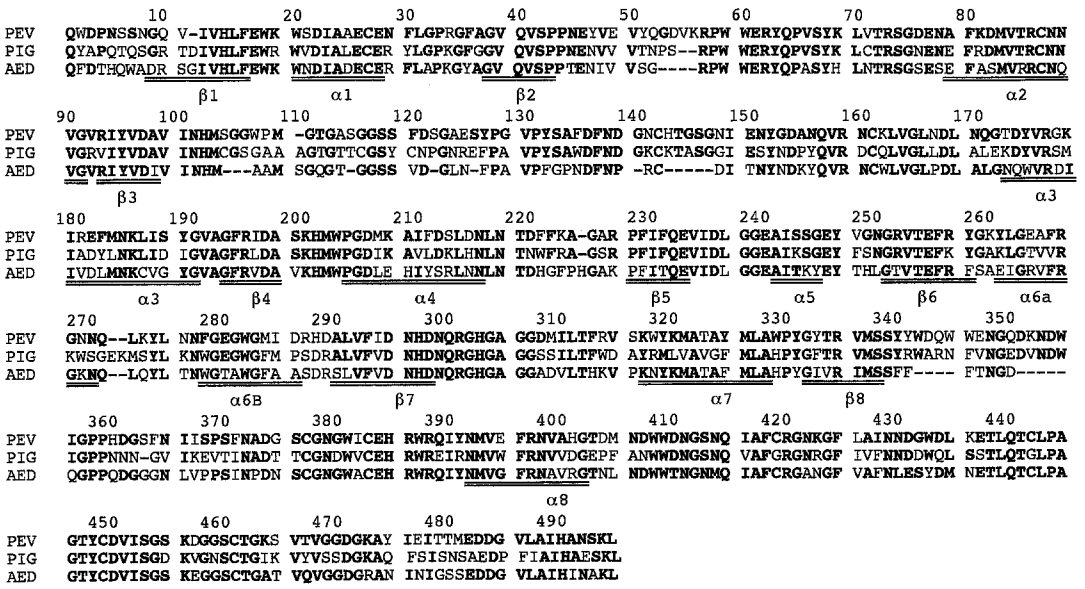

Fig. 4. Location of secondary-structure elements of the (β/α) eight-barrel in *Penaeus vannamei,* pig, and *Aedes* α-amylases. *Numbering* is that of the mature proteins. *PEV = P. vannamei* (X77319), *PIG = Sus scrofa* (POO690). *AED = Aedes aegypti* (P13080) α-amylases. The secondary-structure elements are *double underlined.*

(cDNA 2 and 3) were also entirely sequenced. These three cDNA sequences were aligned and are shown in Fig. 3.

The protein (AMY 37) deduced from the cDNA1 coding sequence is 511 amino acids long. This protein is composed of a highly hydrophobic peptide covering residues −16 to −1 (Kyte and Doolittle 1982). The amino-terminal amino acid, represented by a glutamine, corresponds to the blocked end of the protein (Van-Wormhoudt et al. 1995) and is also confirmed by the alignment of known amylases. The preprotein should be processed to give a mature enzyme of 495 residues. Only ten nucleotide changes are observed when comparing cDNA2 to cDNA1. These ten nucleotide changes produce the substitution of four amino acids (Fig. 3). This amylase variant (AMY 28) contains one less acidic residue (pHi = 5.56) than amylase 37 (pHi = 5.48). Referring to α-amylase purification, AMY 37 may correspond to isoform (a) and AMY 28 to isoform (b) observed after Davis electrophoresis.

In the third cDNA, 145 nucleotide changes are observed when comparing to the major cDNA; 77 of these nucleotide changes give way to 57 changes in the amino acid sequence of the cDNA3 coded protein (AMY 20). Among these 57 changes, 20 are conservative changes. This third variant appears to be more hydrophobic (+2 residues) and less basic (−4), (pHi = 5.35). The stop codon was found to be different and the untranslated 3′ ends present very little homology between cDNA 1 and 2 and cDNA 3. It was also the case for the two variants of the shrimp chymotrypsin we have already cloned and sequenced (Sellos and Van Wormhoudt 1992). The calculated molecular weight of shrimp amylase does not agree with previously determined values based on SDS electrophoresis (Van Wormhoudt et al. 1995) but is comparable to that of other amylases.

The five consensus regions defined by Nakajima et al. (1986) and Janececk (1992), found in almost all the other amylases from different sources, can also be found in the shrimp amylase. These highly conserved regions are very likely either active or substrate binding sites in α-amylases (Table 1). They comprised the catalytic residues Asp 197, Glu 233, and Asp 297 and the Ca^{++} binding site Asp 167 (Fig. 4). More recently, four other regions were established (Janecek 1994a).

Analyses of partial clones have shown the existence of at least two other variants of clone 28, suggesting the existence of at least five different genes. High numbers of genes have also been reported in insects (Levy et al. 1985), in rats (McDonald et al. 1980), in human (Nakamura et al. 1984), and in rice (Huang et al. 1992). The purpose of maintaining multiple isoenzymes for amylases is not known and may be related to food conditions or growth (Le Moullac 1995). It is noteworthy that the expression of one of the major isoforms can be controlled by food nutrients (Van Wormhoudt et al. 1996).

Structure Analysis

Using an HCA plot, α-amylase is characterized by a (β/α) eight-barrel supersecondary structure that constitutes an A-domain. The segments, labeled α 1–8 and β 1–8, correspond to the α-helical and β-sheet structures (Fig. 4) and are similar to those of porcine α-amylase. These segments are characteristic of the "animal" amylase subfamily previously determined (Raimbaud et al. 1989). The C-terminal domain, comprising the amino

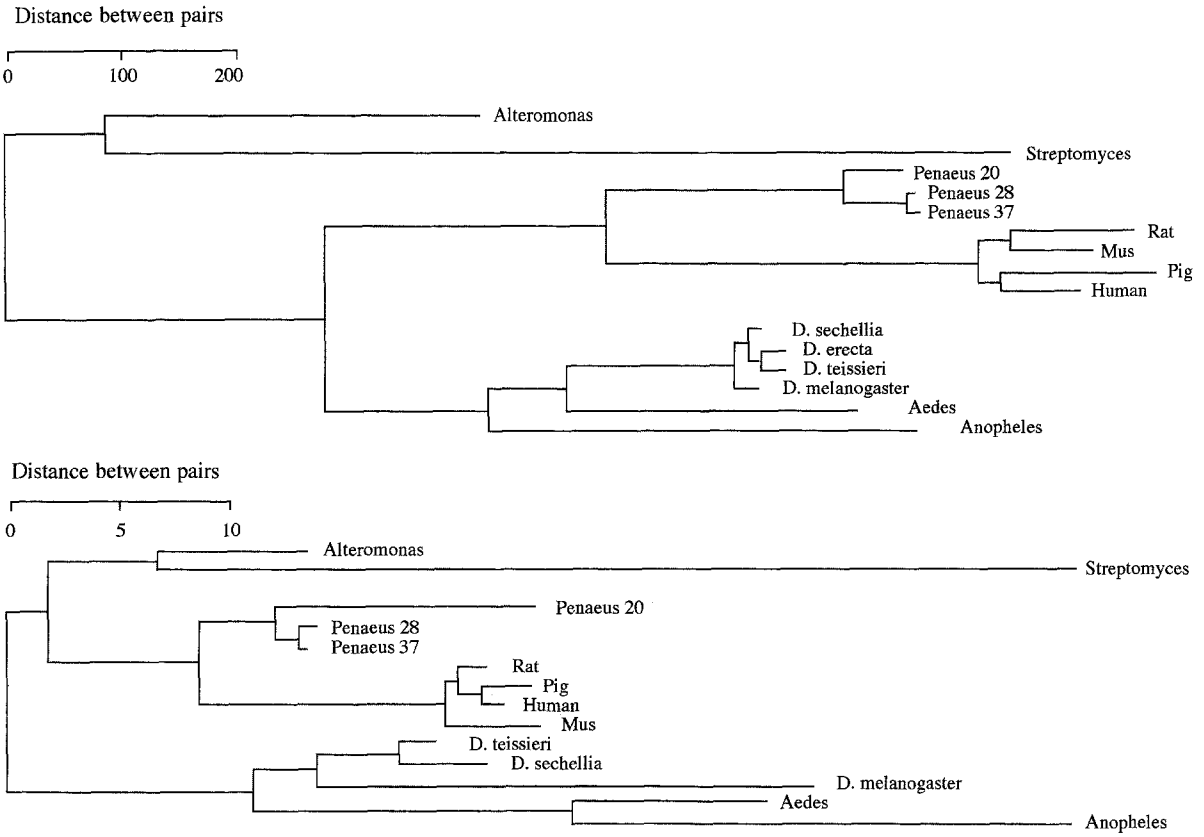

Fig. 5. Phylogenetic relationships within the α-amylase family. Unrooted phylogenetic trees were constructed using parsimony analysis (TREE-ALIGN method of Hein et al. 1990). Trees were constructed with all the sequences (**A**) and compared with those constructed with conserved domains (**B**) (Nakajima et al. 1986); Janecek 1992, 1994b). For ref. see Table 1.

acid residues located after the α8 region, is also folded, with an eight-stranded antiparallel β-sheet (data not shown). The same feature was detected in mammals and in crustaceans, but no significant homology was reported between mammals and other α-amylases except those of *Streptomyces* (Raimbaud et al. 1989).

Other structural features were also examined such as the occurrence of cysteine residues. Ten residues were observed in amylase 37 and 28 and 11 in amylase 20. In pig amylase, 12 cysteines were characterized, giving only 5 disulfide bridges (Pasero et al. 1986). All these bridges are present only in mammal α-amylases (Janecek 1994a). One bridge, holding cysteine 70 to cysteine 115 in pig, cannot be present in insects and in crustaceans because of the absence of cysteine at these positions. Two cysteines exist also in *P. vannamei* in positions 419 and 442 (Fig. 4), as well as in *Aedes* amylase (Grossman and James 1993), which could result in a cysteine bridge. The presence of this putative bridge may be related, in crustaceans, to differences in activity, especially concerning the Michaelis constant displayed by the enzymes during temperature adaptation (Van Wormhoudt 1980).

Sequence Alignment

Raimbaud et al. (1989) have shown that α-amylases have evolved from a common precursor and described six subfamilies, based upon specific similarities occurring in the variable regions of α-amylases from bacterial (three subfamilies), cereal, fungal, and animal origins. In order to evaluate the divergence of amylase sequences over long evolutionary periods, entire amylase amino acid sequences from *P. vannamei* were compared with mammalian (Pasero et al. 1986; Hagenbuchle et al. 1989) or insect sequences (Boer and Hickey 1986) retrieved from data banks. From these examples, Hickey et al. (1987) have shown that animal α-amylase family proteins are highly conserved. The amylases from *Streptomyces* and *Aspergillus* along with those from *Thermo-monospora curvata* and *Alteromonas haloplanctis* (Janecek 1994b) appear to be more similar to mammalian α-amylase than to any other bacterial or cereal amylases and were used for aligments (Fig. 5A). On searching the data bank (Dessen et al. 1990) these proteins contained the best homology scores with shrimp amylase: 38% for *Streptomyces* to 44% for *Alteromonas,* respectively.

From these aligments, *P. vannamei* α-amylases appeared to belong to two families confirming the analysis of the cDNA. This result appeared also on phylogenetic analysis of conserved stretches of (β/α)-barrel domain (Fig. 5B). The percentage of amino acid similarity (52–62%) between shrimp and insect α-amylases is more or less the same as that observed between shrimp and vertebrate α-amylases (59–63%). The use of the four well-

accepted conserved regions, stretches III, V, VI, and VIII in Table 1 (Nakajima et al. 1986), the fifth conserved region (stretch IV in Fig. 5) (Janecek 1992), and the rest of sequence similarities (Janecek 1994a) increased the levels of similarity but did not alter the ordering of the groups (Fig. 5b), suggesting that conservation of the β/α-barrel domain structure of amylase is as important as active sites or conserved sequences (McKay et al. 1985). For other proteins, differences are much more important between active and other sites, and differences in topology of phylogenetic trees between kringle and protease domains, for example, have been reported for serine proteases (Ikeo et al. 1995).

The alignment of amylases provides information not only on evolutionary history but also on the function of these molecules by the knowledge of structural features important for enzyme activity (McGregor 1993). Amylases which are proteins showing a slow degree of evolution (Hickey et al. 1987) may be a good model in which to establish, in the future, phylogenetic relationships among Arthropoda and among other crustaceans. Because of the same degree of homology between vertebrate and insect amylase sequences, it is probable that the divergence of crustacea and insects occurred before the apparition of Burgess fauna, 550 million years ago (Hickey et al. 1987). Moreover, the regulation of the expression of the different isoforms will be studied.

Acknowledgments. This research was partially supported by grants from the "Conseil Régional de Bretagne, France" and from IFREMER (URM 14). We thank Professor G. Marchis Mouren for critical revue of the manuscript, Professor B. Henrissat for determination of the secondary structure, and Miss Andrea Bullock for improving the English text.

References

Abele LG, Felgenhauer BE (1986) Phylogenetic and phenetic relationships among the lower Decapods. J Crust Biol 6:385–400

Abele LG (1991) Comparison of morphological and molecular phylogeny of the Decapoda. Memoirs of the Queensland Museum 31: 101–108

Boer PM, Hickey DA (1986) The α-amylase gene in *Drosophila melanogaster:* nucleotide sequence, gene structure and expression. Nucleic Acid Res 14:8399–8411

Burkenroad MD (1963) The evolution of the Eucarida (Crustacea, Eumalacostaca) in relation to fossil record. Tulane Studies Geol 2:3–16

Cisne JL (1974) Trilobites and the origin of arthropods. Science 186: 13–18

Chirgwin JJ, Przbyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologicaly active ribonucleic acid from sources enriched in ribonucleases. Biochemistry 18:5294–5299

Davis BT (1964) Disc electrophoresis. II—Method and application to human serum proteins. Ann NY Acad Sci 321:404–428

Dessen P, Fondrat C, Valencien C, Mugnier C (1990) BISANCE: French service for access to biomolecular sequence databases. Comp Appl Biosci 6:355–356

Felgenhauer BE, Abele GL (1983) Phylogenetic relationships among shrimp-like Decapods. In: Schram FR (ed) Crustaceans issue 1: Crustacean phylogeny. AA Balkema, Rotterdam, Netherlands pp 291–311

Glaessner MF (1969) Decapoda. In: Moore RC (ed) Treatise on invertebrate Paleontology. part R, Arthropoda 4, II R399–R533. Geological Society of America, Boulder, Colorado and the University of Kansas Press, Lawrence

Grossman GL, James AA (1993) The salivary gland of the vector mosquito *Aedes aegypti* expresses a novel member of the amylase gene family. Insect Mol Biol 1:223–232

Hagenbuchle O, Bovery R, Young RA (1980) Tissue specific expression of mouse α-amylase genes: nucleotide sequence of isoenzyme mRNAs from pancreatic and salivary glands. Cell 21:179–187

Hein J (1990) Unified approach to aligment and phylogenies. In: Doolittle RF (ed) Methods in Enzymology Acad Press NY, 183: 626–644

Hickey DA, Benkel BF, Boer PH, Genest Y, Abukashwa S, Ben David G (1987) Enzyme encoding genes as molecular clocks: the molecular evolution of animal alpha-amylases. J Mol Evol 26:252–256

Huang N, Stebbins GL, Rodriguez RL (1992) Classification and evolution of α-amylases in plants. Proc Natl Acad Sci USA 89:7526–7530

Ikeo K, Takahashi K, Gojoborit J (1995) Different evolution histories of Kringle and protease domains in serine proteases: a typical example of domain evolution. J Mol Evol 40:331–334

Janecek S (1992) New conserved amino acid region of α-amylase in the third loop of their (β/α)8-barrel domain. Biochem J 288:1069–1075

Janecek S (1994a) Sequence similarities and evolutionary relationships of microbial, plant and α-amylases. Eur J Biochem 224:519–524

Janecek S (1994b) Parallel β/α-barrels of α-amylase, cyclodextrin, glycosyltransferase and oligo-1-6 glucosidase versus the barrel of β-amylase: evolutionary distance is a reflection of unrelated sequence. Febs Lett 353:119–123

Jespersen H, McGregor E, Henrissat B, Sierks MR, Svensson B (1993) Starch and glycogen-debranching and branching enzymes: prediction of structural features of the (β/α)8-barrel domain and evolutionary relationships to other amylolytic enzymes. J Prot Chem 12:791–805

Keller P, Kaufman DL, Allan BJ, Williams BL (1971) Isoenzymes of human parotid α-amylases. Biochemistry 10:4867–4874

Kim W, Abele LG (1990) Molecular phylogeny of selected decapod crustaceans based on 18S rRNA nucleotide sequences. J Crust Biol 10:1–13

Kyte J, Doolytle RF (1982) A single method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Le Moullac G (1995) Adaptation des enzymes digestives à l'alimentation chez la crevette *Penaeus vannamei* (Crustacea Decapoda). Thèse EPHE, 120 p

Levy JN, Gemmil RM, Doane WW (1985) Molecular cloning of alpha-amylase from *Drosophila melanogaster.* II Clone organisation and verification. Genetics 11:313–324

McDonald RJ, GreRar MM, Swain WF, Pictet RL, Thomas G, Rutter WJ (1980) Structure of a family of rat amylase genes. Nature 287: 117–122

McGregor EA (1993) Relationships between structure and activity in the α-amylase family of starch metabolising enzymes. Starch 45: 232–237

McKay RM, Baird S, Dove MJ, Errat JA, Gines M, Moranelli F, Nasim A, Willick GE, Yaguchi M (1985) Glucanase gene diversity in prokaryotic and eukaryotic organisms. Biosystems 18:279–292

Nakajima R, Imanaka T, Aiba S (1986) Comparison of amino acid

sequences of eleven α-amylases. Appl Microbiol Biotech 23:353–360

Nakamura Y, Ogama M, Nishida T, Emi M, Kosaki G, Himeno S, Matsubara K (1984) Sequence of cDNAs for human salivary and pancreatic α-amylases. Gene 28:263–270

Ovendeen JR, Brasher DJ, White RWG (1992) Mitochondrial analysis of the red rock lobster *Jasus edwardsii* supports an apparent absence of population subdivision throughout Australasia. Mar Biol 112:319–326

Palumbi SR, Benzie J (1991) Large mitochondrial DNA differences between morphologically similar Penaeid shrimps. Mol Mar Biol Biotech 1:27–34

Pasero L, Mazzei-Pierron Y, Abadie B, Chicheportiche Y, Marchis-Mouren G (1986) Complete amino-acid sequence and location of the five disulfide bridges in porcine b pancreatic α-amylases. Biochim Biophys Acta 869:147–157

Pictet R, McDonald RJ, Swain WF, Grebar MM, Hobart PM, Crawform R, Shen LD, Bell G, Rutter WJ (1981) Differentiation of the pancreas: an analysis of the structure of the amylase and insulin genes. Fortsh Zool B 26:227–245

Raimbaud E, Buleon A, Perez S, Henrissat B (1989) Hydrophobic cluster analysis of the primary sequences of α-amylases. Int J Biol Macromol 2:217–225

Schram FR (1982) The fossil record and evolution of crustacea. In: Abele LG (ed) The biology of crustacea. Acad Press NY, London, 1:241–304

Sellos D, Van-Wormhoudt A (1992) Molecular cloning of a c-DNA that encodes a serine protease with chymotryptic and collagenolytic activities in the hepatopancreas of the shrimp *Penaeus vannamei* (Crustacea, Decapoda). FEBS Lett 309:219–224

Van Wormhoudt A (1980) Régulation d'activité de l'α-amylase à différentes températures d'adaptation et en fonction de l'ablation des pédoncules oculaires et du stade de mue chez *Palaemon serratus*. Biochem System Ecol 8:193–203

Van Wormhoudt A (1983) Variations immunoquantitatives de l' α-amylase au cours du cycle d'intermue à différentes saisons chez Palaemon serratus (Crustacea Decapoda). Mar Biol 74:127–132

Van Wormhoudt A, Favrel P (1988) Electrophoretic characterization of *Palaemon elegans* (Crustacea Decapoda) α-amylase system: study of amylase polymorphism during the intermolt cycle. Comp Biochem Physiol 89B:201–207

Van Wormhoudt A, Bourreau G, Le Moullac G (1995) Amylase polymorphism in Crustacea Decapoda: electrophoretic and immunological studies. Biochem System Ecol 23:139–149

Van Wormhoudt A, Le Moullac G, Klein B, Sellos D (1996) Adjustment of the expression of digestive enzymes to casein level in food in *Penaeus vannamei*. Br J Nutr 55 (in press)