FEBS 18495 FEBS Letters 407 (1997) 153–158

Molecular cloning of hemocyanin cDNA from *Penaeus vannamei* (Crustacea, Decapoda): structure, evolution and physiological aspects

Daniel Sellos*, Soazig Lemoine, Alain Van Wormhoudt

Station de Biologie Marine, Museum National d'Histoire Naturelle et Collège de France, P.O. Box 225, 29182 Concarneau Cedex, France

Received 17 February 1997; revised version received 19 March 1997

Abstract Hemocyanin is present as 2 subunits in the hemolymph of *Penaeus vannamei*. Isolated from a hepatopancreas cDNA library of this penaeid shrimp, the cDNA chain (2095 bp) corresponds to a full length hemocyanin messenger as determined by Northern hybridization, with a short 5′ untranslated region (17 bp), an open reading frame (1989 bp counting initiation and termination codons) coding for a signal peptide (13 residues) and a mature hemocyanin (648 amino acids), and a 3′ untranslated region (89 bp) followed by the polyadenylated track. It is the first time that the existence of a hydrophobic signal peptide is shown in arthropod hemocyanin. Two primary N-terminal sequences are determined and a 3-fold increase of mRNA content, measured in the hepatopancreas during the premoult stages, is reported. The low level of polymorphism shown by *P. vannamei* hemocyanin, along with its weak percentage identity with counterparts and its similarity with hemocyanin from *Panulirus interruptus*, suggests that this arthropod hemocyanin may be a primitive subunit that has evolved independently, following gene duplication.

Key words: Hemocyanin; cDNA nucleotide sequence; Invertebrate; Crustacea

1. Introduction

Hemocyanin is a copper-containing respiratory pigment assuming O_2 transport, colorless when de-oxygenated, but appearing light blue when oxygenated. Hemocyanins occur in two phyla of invertebrates: the arthropoda [1–6] and the mollusca [7], occurring in several different classes: Crustacea, Myriapoda (Chilopoda), Merostomata (horseshoe crab) and Arachnida. As the main protein component of hemolymph, hemocyanin typically represents up to 95% of the total amount of protein. In crustacean hemolymph, several components differing in size, as estimated by their sedimentation coefficients ranging from 39 S to 16 S (as it is in *Palinurus* or *Penaeus)* were determined [8,9]. The hexamer, with a molecular mass of 4.5×10^5 Da, is the predominant form in the most primitive crustacean Decapoda such as *Penaeus setiferus* [9] or *P. monodon* [10]. For the mean value of the smallest functional component, the polypeptide monomers contain 600–660 amino acid residues depending on species (625 for chelicerate hemocyanins and 660 for crustacean protein). Between three and eight different subunits, classified on immunological data, have been distinguished in crustacean hemocyanin [11]. Alpha type is present in all of them, but beta and gamma types may be absent [12]. The hepatopancreas has been shown to be the site of hemocyanin synthesis [13]. Until now, only several partial N-terminal and some complete amino acid sequences of the different subunits from *Panulirus interruptus* were determined in crustacea [14–16], and no cDNA sequence was further characterized. Hemocyanin cDNAs were isolated from the tarantula *Eurypelma californicum* [17] and from the mollusk *Octopus dofleini* [7]. Surprisingly, no leader sequence coding for a signal peptide was detected in the hemocyanin gene in tarantula [18]. The authors argue that the secretion of this extracellular pigment is achieved by the accumulation of this protein in the cytoplasm released by holocrine secretion.

In this paper, we report on the purification and electrophoretic characterization of *Penaeus vannamei* hemocyanin and on the molecular cloning and complete sequencing of one hemocyanin cDNA from the hepatopancreas. The presence in the deduced protein sequence of an hydrophobic signal peptide, absent in the other invertebrate hemocyanins, is discussed in relation to differences in the mode of secretion of this protein in relation to the penaeid moulting cycle.

2. Materials and methods

Shrimps (*Penaeus vannamei*) were obtained from Ifremer (Tahiti). The hepatopancreas was removed by dissection, immediately frozen in liquid nitrogen and stored at −80°C. Determination of the moulting stages was made according to the method of Drach and Tchernigotzeff [19].

2.1. Hemocyanin and RNA extractions and purifications

Hemocyanin was taken in a 1 ml syringe by puncturing the heart. After coagulation, the clot was disrupted by passage through the needle and the sample was centrifuged 5 min at $14\,000\times g$. Hemocyanin complex was obtained from the serum of one animal by chromatography on a Sephadex G100 column (15×300 mm). Proteins from the main peak eluted in the exclusion volume of the column were characterized by electrophoresis on a 3% stacking gel and 7% SDS-PAGE acrylamide gel [20]. An aliquot of the G100 purified hemocyanin was used for the determination of the N-terminal amino acid sequence [21].

Total RNA extractions were made following the guanidine thiocyanate method [22] after disruption of the digestive glands in liquid nitrogen. Poly-A^+ RNAs were purified by affinity chromatography on oligo dT-cellulose column. Total RNA was isolated from 50 mg fresh weight individual hepatopancreas from 2–6 shrimps for each of the different moulting stages using the rapid RNA extraction kit from Stratagene. A 256 bp *Eco*RI coding fragment of the shrimp hemocyanin cloned cDNA was used to probe RNA, which was fractionated by denaturing electrophoresis on 1.5% agarose gels in phosphate buffer. Transfer to nylon membrane (Hybond N^+, Amersham) was performed as described by the manufacturer. Dots of known RNA amounts of each of the moulting stages, quantified by spectrophotometry, were directly fixed on membranes to estimate variation of the hemocyanin mRNA during the moult cycle. The membranes were treated as for the screening procedure.

*Corresponding author. Fax (33) 02 98 97 81 24.
E-mail: sellos@sb-roscoff.fr

EBI sequence accession number: X82502.

PII S0014-5793(97)00350-5

2.2. cDNA library screening

A lambda ZAP cDNA library for shrimp hepatopancreas was established following the manufacturer's protocol (Stratagene, La Jolla, CA, USA). The amplified cDNA library, containing 8×10^9 phages per ml, was first used to select randomly 15 isolated clones among 500 phage plaques. In a second step, plaques were transferred to Hybond N^+ membranes (Amersham) and screened using a 5′ end 498 bp *Eco*RI fragment of the cloned cDNA we obtained for the first selection as a probe covering nucleotides 1262–1759. Prehybridization of the duplicate membranes was achieved in a 50% formamide solution containing 1% sodium dodecyl sulfate (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 and 8×10^8 cpm/μg) for 16 h at the same temperature. The filters were washed twice in 2×sodium saline citrate (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.

In order to have complete cDNAs, a third screening was performed using a synthetic oligonucleotide (CATGAGTGCCTCATTACGATGCC) based on the 5′ end of the longest previously isolated cDNA hybridizing to nucleotides 275–297, end-labelled with [γ-^{32}P]ATP [23].

2.3. Plasmid subcloning and sequencing

The existence of five restriction sites for *Eco*RI and two for *Bam*HI in the complete cDNA (clone Hemo20) allow us to obtain six subclones in pBluescript SK and KS minus (Stratagene) covering the whole insert. The two *Bam*HI restriction sites in position 762 and 1107 were used to generate a subclone encompassing the portion containing the two close by *Eco*RI sites. All these subclones were sequenced in the both orientations. In the case of the large *Eco*RI fragment (657 bp), two internal primers were used (the one used for the screening on the complementary strand and another one with sequence CCAATTCCTTTGATCCCGTT (hybridizing to nucleotide 143–162, on the coding strand). DNA was sequenced following the dideoxynucleotide method with modification for extended DNA sequencing. Electrophoresis of the extended products was performed with two or three successive loadings on a 5% acrylamide/bisacrylamide (30/0.8) gel.

2.4. Extension of 5′ mRNA end

Specific partial hemocyanin cDNAs were synthesized using a RACE 5′/3′ kit (Boehringer Mannheim) from total RNA with a specific anti-sense oligonucleotide: CATGAGTGCCTCATTACGATGCC hybridizing to nucleotides 275–297. The 5′ end of the hemocyanin cDNA was amplified using an oligo-dT anchor primer hybridizing to the appended poly-A tail and a second specific anti-sense oligonucleotide (HEMO2R): CCATCACTGTAGGTGCCTAAGTTTC, hybridizing to nucleotides 164–188.

3. Results and discussion

3.1. Hemocyanin characterization

Hemocyanin from *Penaeus vannamei* was observed as two close prominent bands of protein with molecular weights around 73 and 75 kDa as determined by SDS-denaturing gel electrophoresis (Fig. 1). These results confirm the work of Chan et al. [24] and similar results have been recorded for *Penaeus setiferus* [8] and *Penaeus semisculatus* [25]. The bands represent the monomeric polypeptides of hemocyanin and suggest subunit heterogeneity.

3.2. Library screening and hemocyanin cDNA characterization

By screening the cDNA library, approximately 8% of the phage plaques were found to be positive. Further screening with the oligonucleotide probe allowed us to isolate incomplete cDNAs. Using the 5′ RACE method with Pfu DNA polymerase (Stratagene) and the reverse oligonucleotide HEMO2R, a 250 bp fragment was obtained and cloned. The analysis of these cloned inserts confirmed the full length sequence of the cDNA.

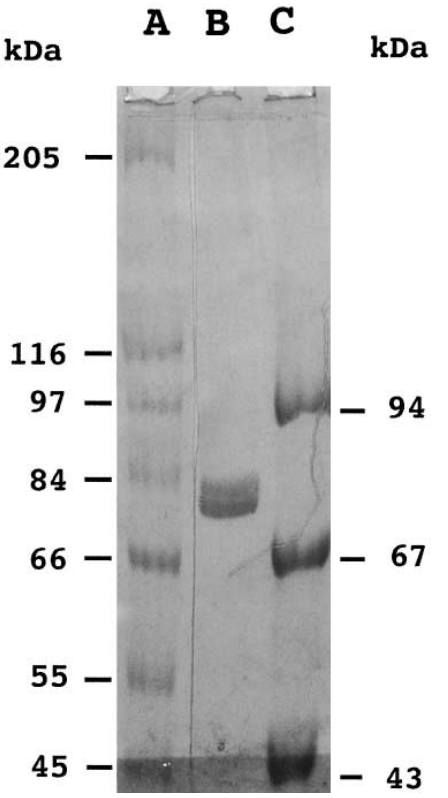

Fig. 1. SDS-denaturing electrophoresis of *Penaeus vannamei* hemocyanin on a 7% acrylamide/bisacrylamide gel. 4 μg of protein was denatured and loaded on the gel. After electrophoresis, proteins were stained with Coomassie blue. A: High range markers (Sigma); myosin (205 000), β-galactosidase (116 000), phosphorylase b (97 000), fructose-6-phosphate kinase (84 000), bovine serum albumin (66 000), glutamic deshydrogenase (55 000) and ovalbumin (43 000). B: Crustacean hemocyanin sample. C: Low molecular weight markers (Pharmacia); phosphorylase b (94 000), albumin (67 000) and ovalbumin (43 000).

The longest, complete cDNA sequence recorded is 2095 bp long (without the poly-A tail), including a 17 bp 5′ untranslated region, a 1989 bp coding sequence (counting initiation and termination codons) and an 89 bp untranslated 3′ end with the polyadenylated track (Fig. 2). The translation initiation site bears no resemblance to a Kozak motif [26] and does not contain an in-frame stop sequence. Moreover the presence of the very short 5′ untranslated extension has been observed for major digestive enzymes such as trypsin [27], cathepsin [28] and amylase [29], synthesized and secreted in the penaeid digestive gland. Such an extension is also found in prophenoloxidase (a member of the hemocyanin family) in *Pacifastacus leniusculus* [30]. The open reading frame encodes 661 amino acids forming a preprotein composed of a hydrophobic signal peptide (13 residues; not observed in tarantula [18]) and a mature protein (648 amino acids). This hydrophobic signal sequence ends with a stretch of 4 alanines corresponding to an Ala-X-Ala motif, a frequent consensus preceding the cleavage site of signal peptides [31]. Given that hemocyanin should be stored in secretory vesicles as has been observed for digestive enzymes in the F cells of the crustacean hepatopancreas [27–29], the presence of this sequence, absent in the other known invertebrate hemocyanins, implies that the crustacean hemocyanin passes through the endoplasmic reticulum membrane.

```
ATG AGG GTC TTA GTG GTT CTT GGG CTT GTC GCT GCT GCC GCC TTT CAG GTG GCC AGT GCA GAT GTT CAG CAG CAG AAA   78
Met R  V  L  V  V  L  G  L  V  A  A  A  A  F  Q  V  A  S  A  D  V  Q  Q  Q  K   12

GAC GTC CTT TAT CTC CTG AAC AAA ATC TAT GGA GAC ATT CAG GAC GGA GAC CTG CTG GCT ACT GCC AAT TCC TTT GAT  156
 D  V  L  Y  L  L  N  K  I  Y  G  D  I  Q  D  G  D  L  L  A  T  A  N  S  F  D   38

CCC GTT GGA AAC TTA GGC AGC TAC AGT GAT GGT GGT GCA GCC GTG CAA AAA CTG GTG CAG GAC CTT AAT GAC GGC AAA  234
 P  V  G  N  L  G  S  Y  S  D  G  G  A  A  V  Q  K  L  V  Q  D  L  N  D  G  K   64

CTC TTG GAG CAG AAA CAC TGG TTC TCC CTT TTC AAT ACA AGG CAT CGT AAT GAG GCA CTC ATG CTT TTC GAC GTC CTC  312
 L  L  E  Q  K  H  W  F  S  L  F  N  T  R  H  R  N  E  A  L  M  L  F  D  V  L   90

ATC CAC TGC AAA GAC TGG GCA TCC TTT GTC GGC AAT GCA GCC TAC TTC CGT CAG AAA ATG AAC GAA GGA GAG TTT GTT  390
 I  H  C  K  D  W  A  S  F  V  G  N  A  A  Y  F  R  Q  K  M  N  E  G  E  F  V  116

TAT GCC TTG TAT GTT GCC GTG ATC CAC TCG TCT TTG GCT GAA CAG GTG GTG CTC CCT CCC CTC TAT GAG GTC ACA CCG  468
 Y  A  L  Y  V  A  V  I  H  S  S  L  A  E  Q  V  V  L  P  P  L  Y  E  V  T  P  142

CAC CTC TTC ACC AAC AGT GAA GTT ATC GAA GAA GCT TAT CGT GCC AAA CAG AAG CAG ACG CCT GGC AAA TTC AAG TCC  546
 H  L  F  T  N  S  E  V  I  E  E  A  Y  R  A  K  Q  K  Q  T  P  G  K  F  K  S  168

TCC TTT ACG GGA ACC AAG AAA AAC CCT GAA CAG AGA GTG GCA TAT TTC GGT GAA GAT ATC GGC TTG AAT ACC CAT CAC  624
 S  F  T  G  T  K  K  N  P  E  Q  R  V  A  Y  F  G  E  D  I  G  L  N  T  H  H  194
                EcoRI
GTT ACC TGG CAT ATG GAA TTC CCC TTC TGG TGG AAT GAT GCT TAC GGC CAT CAT CTG GAT CGC AAA GGA GAA AAC TTC  702
 V  T  W  H  M  E  F  P  F  W  W  N  D  A  Y  G  H  H  L  D  R  K  G  E  N  F  220
                                                   BamHI
TTC TGG ATT CAT CAC CAA CTT ACC GTC CGA TTT GAT GCT GAA CGT CTG TCC AAT TAT CTG GAT CCA GTA GGT GAA CTC  780
 F  W  I  H  H  Q  L  T  V  R  F  D  A  E  R  L  S  N  Y  L  D  P  V  G  E  L  246

CAG TGG AAC AAG CCC ATT GTA GAT GGC TTT GCT CCC CAC ACC ACT TAC AAG TAT GGA GGT CAG TTC CCT GCT CGT CCT  858
 Q  W  N  K  P  I  V  D  G  F  A  P  H  T  T  Y  K  Y  G  G  Q  F  P  A  R  P  272
                                  EcoRI                                EcoRI
GAC AAT GTT AAA TTC GAA GAT GTG GAC GAT GTT GCT CGA ATT CGA GAT ATG GTC ATC GTG GAG AGT CGA ATT CGT GAT  936
 D  N  V  K  F  E  D  V  D  D  V  A  R  I  R  D  M  V  I  V  E  S  R  I  R  D  298

GCC ATT GCC CAT GGC TAT ATA GTT GAC AGT GAG GGC AAA CAC ATT GAC ATC AGT AAT GAG AAA GGT ATT GAC ATT CTT 1014
 A  I  A  H  G  Y  I  V  D  S  E  G  K  H  I  D  I  S  N  E  K  G  I  D  I  L  324

GGT GAT ATC ATC GAA TCC TCA CTA TAC AGT CCC AAC GTG CAG TAC TAT GGA GCT TTA CAT AAC ACT GCC CAT ATT GTA 1092
 G  D  I  I  E  S  S  L  Y  S  P  N  V  Q  Y  Y  G  A  L  H  N  T  A  H  I  V  350
          BamHI
CTA GGC CGT CAA GGG GAT CCT CAT GGA AAG TTT GAT TTA CCA CCT GGT GTG CTG GAA CAC TTC GAA ACT GCC ACC CGT 1170
 L  G  R  Q  G  D  P  H  G  K  F  D  L  P  P  G  V  L  E  H  F  E  T  A  T  R  376

GAT CCC AGC TTC TTC CGG CTT CAC AAG TAT ATG GAT AAC ATT TTC AAA GAA CAC AAG GAC AAC CTA CCC CCA TAC ACC 1248
 D  P  S  F  F  R  L  H  K  Y  M  D  N  I  F  K  E  H  K  D  N  L  P  P  Y  T  402
             EcoRI
AAA GCC GAT TTG GAA TTC TCT GGC GTG TCT GTC ACA GAG CTA GCC GTT GTA GGT GAA CTG GAG ACC TAC TTT GAA GAT 1326
 K  A  D  L  E  F  S  G  V  S  V  T  E  L  A  V  V  G  E  L  E  T  Y  F  E  D  428

TTC GAA TAC AGT CTT ATC AAC GCA GTT GAT GAT GCT GAA GGA ATC CCA GAT GTG GAA ATC AGC ACA TAT GTG CCT CGT 1404
 F  E  Y  S  L  I  N  A  V  D  D  A  E  G  I  P  D  V  E  I  S  T  Y  V  P  R  454

CTT AAC CAC AAA GAG TTC ACT TTT AGG ATT GAT GTA GAG AAT GGA GGT GCT GAG AGA TTG GCT ACA GTT CGT ATC TTT 1482
 L  N  H  K  E  F  T  F  R  I  D  V  E  N  G  G  A  E  R  L  A  T  V  R  I  F  486

GCC TGG CCT CAT AAA GAC AAC AAC GGA ATC GAG TAT ACA TTT GAC GAA GGT CGC TGG AAT GCC ATC GAG TTG GAT AAG 1560
 A  W  P  H  K  D  N  N  G  I  E  Y  T  F  D  E  G  R  W  N  A  I  E  L  D  K  506

TTC TGG GTA TCT TTG AAG GGT GGA AAA ACT TCA ATT GAA CGC AAG TCC ACG GAA TCT TCA GTA ACT GTA CCG GAC GTG 1638
 F  W  V  S  L  K  G  G  K  T  S  I  E  R  K  S  T  E  S  S  V  T  V  P  D  V  532

CCA AGC ATA CAT GAC CTG TTT GCA GAA GCC GAG GCA GGC GGC GCT GGC CTT GCC AAA TTC GAG AGT GCA ACA GGC CTA 1716
 P  S  I  H  D  L  F  A  E  A  E  A  G  G  A  G  L  A  K  F  E  S  A  T  G  L  558
                                                EcoRI
CCA AAC AGG TTC CTT CTC CCC AAG GGC AAC GAT AGA GGC CTG GAA TTC GAC CTT GTG GTG GCG GTG ACT GAT GGT GAT 1794
 P  N  R  F  L  L  P  K  G  N  D  R  G  L  E  F  D  L  V  V  A  V  T  D  G  D  584

GCC GAC TCA GCA GTG CCG AAC CTT CAT GAG AAT ACC GAG TAC AAT CAC TAC GGT TCC CAT GGC GTG TAC CCC GAT AAG 1872
 A  D  S  A  V  P  N  L  H  E  N  T  E  Y  N  H  Y  G  S  H  G  V  Y  P  D  K  610

CGT CCT CAT GGT TAT CCT CTG GAC CGC AAA GTT CCA GAT GAG CGC GTG TTT GAA GAT CTT CCT AAC TTT AAG CAC ATC 1950
 R  P  H  G  Y  P  L  D  R  K  V  P  D  E  R  V  F  E  D  L  P  N  F  K  H  I  636

CAA GTT AAG GTC TTC AAT CAT GGT GAA CAC ATT CAT TAG TGATTCAAGCTTACAACTCTTCATATGAATATGTCGATGAATGAATTGTTT 2040
 Q  V  K  V  F  N  H  G  E  H  I  H Ter                                                              648

CTTTGAATTATTTCCCACAATAAAGAACATATGCATAC(A)n                                                           2078
```

Fig. 2. Nucleotide sequence of shrimp hepatopancreas hemocyanin cDNA and deduced amino acid sequence. The amino acid sequence is numbered sequentially from the first amino acid of the mature protein. The proposed signal peptide is represented by amino acids −13 to −1 (bold letters). The hybridization sites of the three oligonucleotide probes used for the screening, the sequencing and the 5′ extension are underlined. The proposed polyadenylation site is also underlined. Restriction sites for *Eco*RI and *Bam*HI and initiation and termination codons are shown in bold letters. Histidines involved in the two copper binding sites are marked with asterisks.

Hemocyanin typically shows a high degree of polymorphism, based primarily on electrophoretic migration or immunoreactivity (see Markl and Decker [32] or Van Holde and Miller [33] for a review). In the tarantula and in *Panulirus interruptus* several hemocyanin sequences can be established or deduced from cDNAs. However, using our *Penaeus vannamei* cDNA library, during the first random selection and the two further screenings, around 50 clones were isolated and the inserts were characterized and partially sequenced. All these cDNAs showed the same nucleotide alignment with the ex-

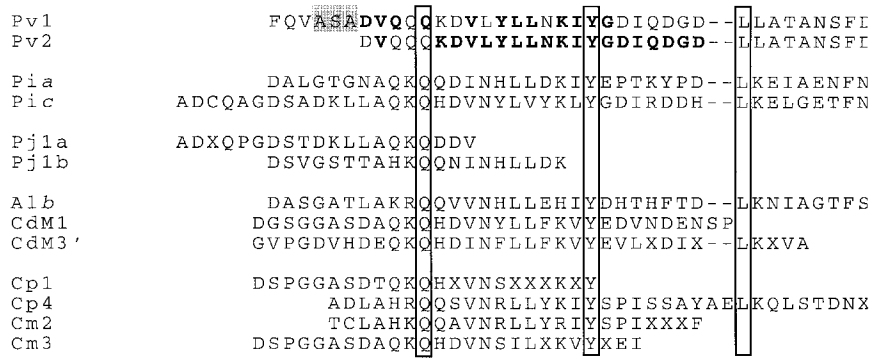

Fig. 3. Alignment of N-terminal partial amino acid sequences of crustacean hemocyanins. The proposed deletions in *P. vannamei* (Pv) compared to the other sequences are indicated by deletion symbols (−). Undetermined residues are indicated by X. Residues in bold characters were confirmed by amino acid sequencing. The putative proteolytic motive is presented in a gray box. Pi: *Panulirus interruptus* [12], Pj: *Palinurus japonicus* [14], Alb: *Astacus leptodactylus* [36], CdM: *Cherax destructor*, Cp: *Cancer pagurus* and Cm: *Carcinus maenas* [14].

ception of two positions: base 435 was found to be G or C; and base 1038, A or G. Even when screening was performed using the short synthetic oligonucleotide probe with low stringent conditions, all the selected positive clones wore inserts with this low level of polymorphism. Nonetheless, these two changes in the nucleotide sequence lead to a single change in the amino acid primary structure of the two subunits, thus showing that residue 131 could be a glutamine or a histidine.

The length recorded for the complete mature protein was 648 amino acids, corresponding to a mass of 73 570 Da and approaching the molecular mass as determined by SDS-gel electrophoresis (73–75 kDa). No putative carbohydrate attachment site was determined. In fact, *N*-type glycosylation is found only in *Panulirus* and *Astacus* hemocyanins [34], although some Asn-X-Ser/Thr sequences have been determined in other crustaceae [35]. Only one cysteine residue is shown, unlike *P. interruptus* (a) subunit, which presents three disulfide bridges, or *P. interruptus* (c) subunit with one disulfide bridge. In *P. vannamei*, the absence of these bridges may cause a certain flexibility in the molecule essential for cooperativity in oxygen binding.

Two overlapping N-terminal sequences were obtained. One was 6 amino acids longer than the other. The first sequence starts with phenylalanine in position 1 of the deduced mature protein sequence. The second protein sequence starts six residues later with the aspartic acid in position 7 of the deduced protein. The alignment of these sequences with other crustacean hemocyanins confirms the existence of subunits with truncated N-terminal portions (Fig. 3), possibly accounting, in part, for the difference in the observed molecular weights. Moreover, the presence of an Ala-X-Ala motive (position 4–6) inside the N-terminal sequence of the longest subunit explains the existence of the shorter subunit by a partial hydrolysis.

3.3. Sequence alignment

The mature protein sequence was aligned with sequences of the two subunits (a) and (c) found in *P. interruptus* [15,16]. The maximum percentage of identity is found when aligning *P. vannamei* and *P. interruptus* variant (c) hemocyanins (Fig. 4). In addition, the alignments showed the same location for several deletions specifically in the C-terminal domain of the two proteins. As expected, the second domain with the oxygen binding site was the most conserved part of the chain. The third domain had the largest number of replacements and resembles more closely the *P. interruptus* (c) subunit, especially in the stretch 583–605. As suggested by Neuteboom et al. [15], hemocyanin (a) and (c) result from gene duplications that occur before the divergence of *Astacus* and *Panulirus*. Our results suggest that this duplication occurred after the appearance of Penaeides, probably in the Paleozoic [38], since the percentage of identity of the *P. vannamei* subunit is more or less the same as that of the two *P. interruptus* subunits, (a) and (c). As such, *P. vannamei* hemocyanin may represent a primitive subunit [39].

3.4. Hemocyanin mRNA expression

Hybridization of total RNA from the hepatopancreas of *P. vannamei* with homologous cDNA probes showed a strong band approximately 2100 bp in length (Fig. 5, insert). This size corresponds to the longest isolated cDNA and fits well with the apparent sizes ranging from 1900 to 2300 bp as determined in *Callinectes sapidus* [40]. Our results show that hemocyanin is synthesized in the hepatopancreas confirming the results obtained in the crayfish using [^{35}S]methionine incorporation [41,42], in the blue crab using hybridization of hepatopancreas mRNA with an oligonucleotide based on the amino acid sequence encompassing the 'copper B' binding site [40], and recently in penaeids [43]. Moreover, we have shown that hemocyanin is synthesized as a preprotein in a very large amount (8% of the cDNA clones). In contrast, the frequency of trypsin or chymotrypsin (the two highly produced digestive enzymes) coding sequences in this cDNA library is only 0.5 and 0.6% respectively [44].

Variations in the amount of hemocyanin mRNA were studied during the moult cycle (Fig. 5). Hemocyanin mRNA amount increased largely from a minimum in stage A to a maximum in stages D1″–D1‴. This pattern has been observed in the expression of chymotrypsin [44], trypsin [27], and other proteins synthesized in the hepatopancreas. In all cases, minima are observed in stages D3–AB and maxima occurred during stages D1″–D1‴. These results corroborate the hypothesis of hemocyanin synthesis in hepatopancreatic cells starting in stage D0, with a coordinate release of cellular constituents under the stimulation by ecdysteroids during the premoult. Biphasic variations are seen for hemocyanin synthesis in *Astacus leptodactylus* [13] hepatopancreas in relation with ecdysteroid levels. The importance of ecdysteroids for the moulting process is emphasized by their stimulative effect on many

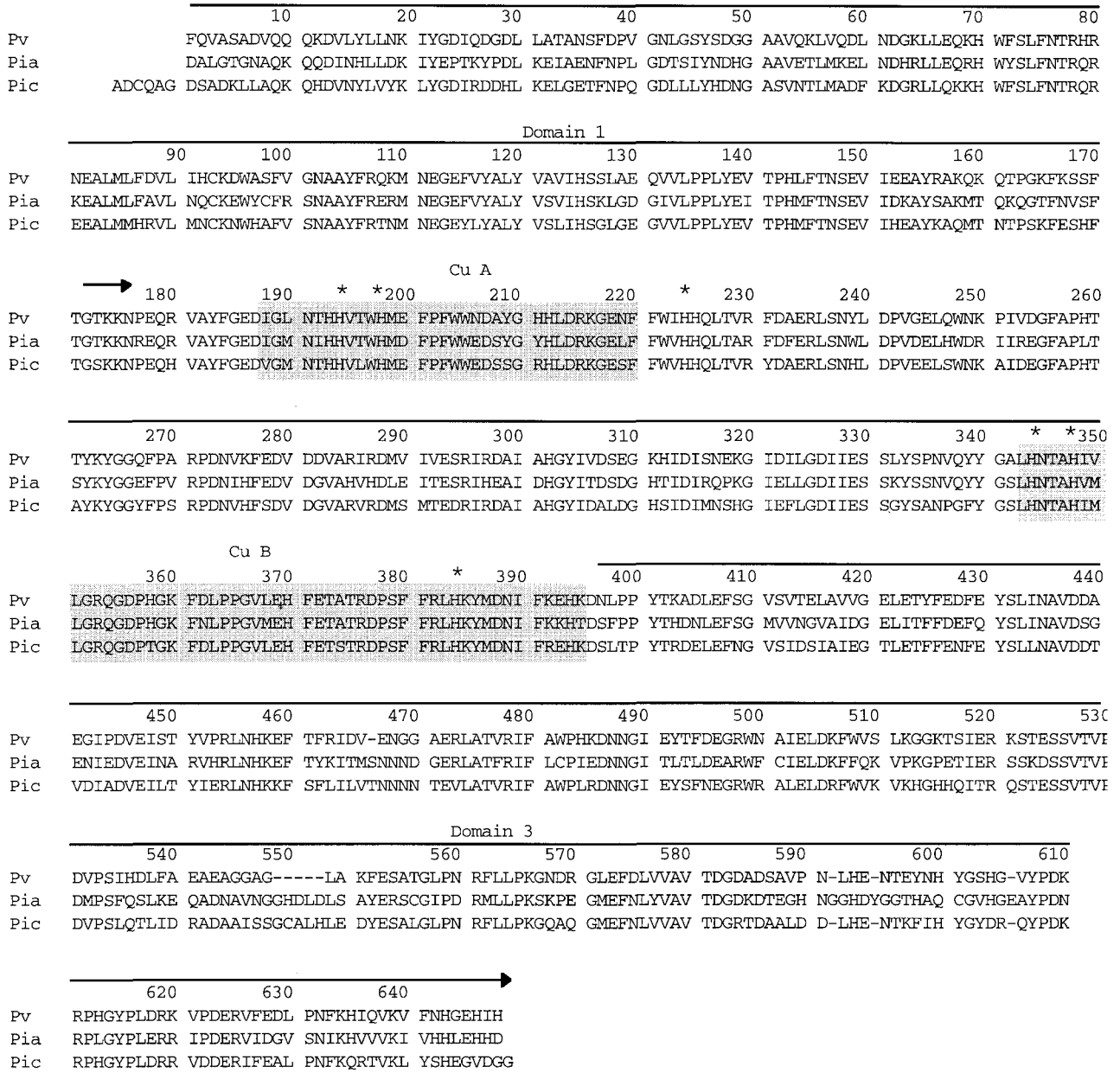

Fig. 4. Alignment of arthropod hemocyanin amino acid sequences [37]. Amino acid sequences are shown for shrimp hemocyanin (Pv), for two of the three subunits of *Panulirus interruptus* (Pia and Pic) [15,16]. Gaps have been introduced in the sequences to facilitate alignments. The numbering above the alignments relates to the shrimp mature hemocyanin as used in Fig. 2. The two copper A and B binding sites are framed. Domain 1 and 2 extents are indicated by arrows.

protein expressions in different tissues, such as actin in crustacean muscle [45]. Actin, which is considered a reference protein in mammals, varies during the moult cycle and, therefore, cannot be used as a control for crustaceans.

Instead, R-cells may be a good candidate for the hemocyanin synthesis location because they display ultrastructural features for the export of metabolites to other organs through hemolymph [46–48]. Certain authors [49], however, consider that the reserve inclusion cells located in small hemal sinuses are responsible for the synthesis and the storage of hemocyanin. Such a hypothesis needs further substantiation since the number of these cells is very low.

In the future, the cDNA we described here will be used as a probe to confirm the hepatopancreas cell type in which hemocyanin is synthesized.

Acknowledgements: The authors are grateful to Miss A. Bullock for the improvement of the English.

References

[1] M.A. Alikhan, Comp Biochem Physiol 39A (1971) 735–741.
[2] N.B. Terwilliger, R.C. Terwilliger, M. Appelstein, C. Bonaventura, J. Bonaventura, Biochemistry 18 (1979) 102–108.
[3] J. Berthet, P. Berthet, Arch Int Physiol Biochem 71 (1963) 124–125.
[4] W. Wieser, Naturwissenschaften 52 (1965) 133–134.

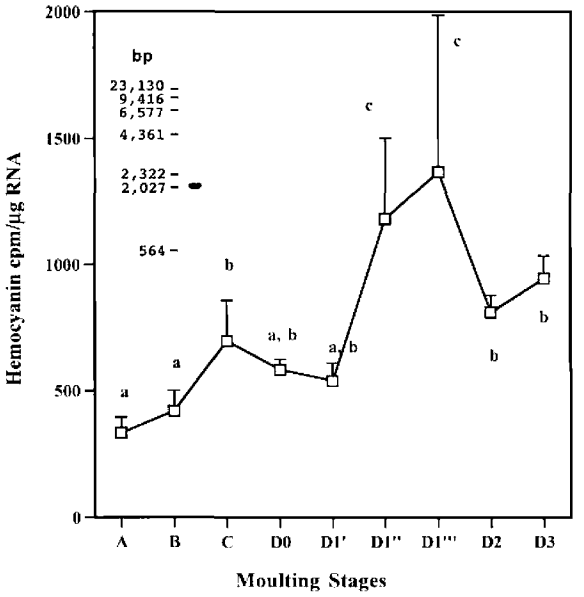

Fig. 5. Variation of the amount of hemocyanin messenger RNA during the intermoult cycle. RNA amounts were quantified by spectrophotometry. 1 µl dots of total RNA extracted from individuals at different moulting stages were fixed on nylon membranes and hybridized as described in Section 2. The curve represents mean values of 2–6 determinations for each moulting stages. Standard errors of mean (S.E.M.) as well as Student analyses are indicated. The different letters correspond to significant values at 0.05 S. Insert: Northern blot of *Penaeus vannamei* total RNA.

[5] R.W. Freel, J Exp Zool 204 (1978) 267–274.
[6] C. Herberts, J. De Frescheville, Comp Biochem Physiol 70B (1981) 65659.
[7] W.H. Lang, K.E. Van Holde, Proc Natl Acad Sci USA 88 (1991) 244–248.
[8] H.A. Kuiper, W. Gaastra, J.J. Beintema, E.F.J. Van Bruggen, A.M.H. Schepman, J. Drenth, J Mol Biol 99 (1975) 619–629.
[9] M. Brouwer, C. Bonaventura, J. Bonaventura, Biochemistry 17 (1978) 2148–2154.
[10] Ellerton HD, Anderson DM. In: Lang J, editor. Invertebrate Oxygen Binding Proteins: Structure, Active Site and Function. New York: Marcel Dekker, 1981:159–170.
[11] J. Markl, Biol Bull 171 (1986) 90–115.
[12] B. Neuteboom, W. Dokter, J. Van Gijsen, H. Rensink, J. De Vries, J.J. Beintema, Comp Biochem Physiol 94B (1989) 593–597.
[13] K.D. Spindler, R. Hennecke, G. Gellissen, Gen Comp Endocrinol 85 (1992) 248–253.
[14] B. Neuteboom, S.J. Sierdsema, J.J. Beintena, Comp Biochem Physiol 49 (1989) 587–592.
[15] B. Neuteboom, P.A. Jekel, J.J. Beintema, Eur J Biochem 206 (1992) 243–249.
[16] P.A. Jekel, H.J. Bak, N.M. Soeter, J.M. Vereijken, J.J. Beintema, Eur J Biochem 178 (1988) 403–412.
[17] R. Voit, G. Feldmaier-Fuchs, J Biol Chem 265 (1990) 19447–19452.
[18] W. Voll, R. Voit, Proc Natl Acad Sci USA 87 (1990) 5312–5316.
[19] P. Drach, C. Tchernigotzeff, Vie Milieu A18 (1967) 595–617.
[20] K. Weber, M. Osborn, J.Biol Chem 244 (1969) 4406–4412.
[21] Humkapiller HW, Grandlund-Moyer K, Whiteley HW. In: Shinley JW, editor. Methods of Protein Microcharacterization. Clifton, NJ: Humana Press, 1986:223–247.
[22] J.J. Chirgwin, A.E. Przbyla, R.J. MacDonald, W.J. Rutter, Biochemistry 18 (1979) 5294.
[23] V.V. Suggs, R.B. Wallace, T. Hirose, E.H. Kawashima, K. Itakura, Proc Natl Acad Sci USA 78 (1981) 6613.
[24] S.M. Chan, S.M. Rankin, L.L. Keeley, Biol Bull 175 (1988) 185–192.
[25] M. Khayat, B. Funkestein, A. Tiez, E. Lubzens, Comp Biochem Physiol 112B (1995) 31–38.
[26] M. Kozak, Cell 44 (1986) 283–292.
[27] B. Klein, G. Le Moullac, D. Sellos, A. Van Wormhoudt, Int J Biochem Cell Biol 28 (1996) 551–563.
[28] Le Boulay C, Van Wormhoudt A, Sellos D. Comp. Biochem. Physiol. 1996 (in press).
[29] A. Van Wormhoudt, D. Sellos, J Mol Evol 42 (1996) 543–551.
[30] A. Aspan, T. Huang, L. Cerenius, K. Soderhall, Proc Natl Acad Sci USA 92 (1995) 939–943.
[31] G. Von Heijne, J Mol Biol 173 (1984) 243–251.
[32] J. Markl, H. Decker, Adv Comp Environ Physiol 13 (1992) 325–376.
[33] Van Holde KE, Miller KI. in: Advances in Protein Chemistry. New York: Academic Press, 1995:1–81.
[34] D. Tseneklidou-Stoeter, G.J. Gerwig, J.P. Kamerling, K. Spindler, Biol Chem Hoppe-Seyler 376 (1995) 531–537.
[35] B. Linzen, M. Soeter, A.F. Riggs, H.J. Schneider, W. Schartau, M.D. Moore, E. Yokota, P.Q. Behrens, H. Nakashima, T. Takagi, T. Nemoto, J.M. Vereijken, H.J. Bak, J.J. Beintema, A. Volbeta, W.P.J. Gaykema, W.G.J. Hol, Science 4713 (1985) 519–5525.
[36] H.J. Schneider, R. Drexel, G. Feldmaier, B. Linzen, F. Lottspeich, A. Henschen, Z Physiol Chem 364 (1983) 1357.
[37] P. Dessen, C. Fondrat, C. Valencien, C. Mugnier, Comp Appl Biosci 6 (1990) 355–356.
[38] Glaessner MF In: Moore RC, editor. Arthropoda 4, part R, II. Boulder, CO: Geological Society of America, and Laurence, KS: University of Kansas Press, 1969:R399–R533.
[39] L.G. Abele, Mem Queensland Museum 31 (1991) 101–108.
[40] J. Rainer, M. Brouwer, Comp Biochem Physiol 104B (1993) 69–73.
[41] Hennecke R, Gellisen G, Spindler-Barth M, Spindler KD. In: Préaux G, Lontie R, editors. Invertebrate Dioxygen Carriers. Leuven: Leuven University Press, 1990.
[42] G. Gellissen, R. Henneke, K.D. Spindler, Experientia 47 (1991) 194–195.
[43] M. Khayat, B. Funkenstein, A. Tiez, E. Lubzens, Comp Biochem Physiol 1 (1995) 31–38.
[44] D. Sellos, A. Van Wormhoudt, FEBS Lett 309 (1992) 219–224.
[45] N.M. Whiteley, E.W. Taylor, A. El Haj, J Exp Biol 167 (1992) 277–284.
[46] G. Vogt, Zoomorphology 114 (1994) 83–101.
[47] T. Caceti, K.F. Neck, D.H. Lewis, R.F. Sis, J Mar Biol Assoc UK 68 (1988) 323–337.
[48] E. Sagrista, M. Durfort, Cell Tissue Res 266 (1991) 585–590.
[49] J.E. Paulus, H. Laufer, Int J Invert Reprod Dev 11 (1987) 29–44.