To investigate the coagulation system in crustacean decapoda, a homodimeric glycoprotein of 380 kDa was purified from the hemolymph of tiger shrimp (*Penaeus monodon*) by sequential DEAE anion exchange chromatography. The purified protein was coagulated by the shrimp hemocyte transglutaminase in the presence of Ca\(^{2+}\). The clottable protein contains 44% \(\alpha\) helices and 26% \(\beta\) sheets as determined by circular dichroism spectra. Its conformation is stable in buffer of pH 4–9. To solve its primary structure, partial sequences of the purified polypeptides from cyanogen bromide cleavage and endopeptidase digestion were also determined. A shrimp cDNA expression library was constructed. By combination with antibody screening, reverse transcriptase PCR using degenerate primers from determined amino acid sequences and cDNA library screening with digoxigenin-labeled DNA probes, the entire cDNA of 6124 bp was obtained. This cDNA encodes a protein of 1670 amino acids, including a 14-amino acid signal peptide. With four potential N-glycosylation sites, the clottable protein was found to contain 3.8% high-mannose glycan; and Man\(_9\)GlcNAc and Man\(_8\)GlcNAc were released upon endo-\(\beta\)-N-acetylglucosaminidase hydrolysis. Upon conducting a protein sequence database survey, the shrimp clottable protein shows 36% identities to the crayfish clotting protein and lower similarities to members of insect vitellogenins, apolipoprotein B and mammalian von Willebrand factor. Notably, a region rich in Gln residues, a polyGln motif and five Ser-Lys-Thr-Ser repeats are present in the shrimp protein, suggesting this protein might be a transglutaminase substrate. Northern blot analysis revealed that the clottable protein is expressed in most of the shrimp tissues but not in the mature hemocytes.

**Keywords:** clottable protein; shrimp; cDNA cloning; carbohydrate analysis; tissue distribution; *Penaeus monodon.

Efficient immune systems and clotting reactions are of vital importance to both vertebrate and invertebrate animals. The invertebrate models of plasma coagulation are highly diverse and less studied. Investigation of them should help our understanding of the general principles of hemostasis in animals. Tait [1] described three types of hemolymph coagulation in crustaceans as follows: type A is characterized by rapid agglutination of hemocytes without clotting of the plasma; type B involves cell aggregation coupled with limited clotting of the plasma, and type C shows limited cell lysis leading to plasma clotting and little cell aggregation. These three types are likely variations of the basic mechanism involving both hemocyte and hemolymph [2–4]. Shrimp and other decapoda undergo the type C coagulation.

In crustaceans, clotting is mediated through clottable proteins (coagulogens) present in the plasma and cell factors compartmentalized within the circulating cells. The clottable protein is converted to covalently linked polymers by Ca\(^{2+}\)-dependent transglutaminase which may be released from hemocytes during coagulation [5–7]. In the wound area, the clottable protein oligomerizes to prevent hemolymph loss through breaks in the exoskeleton and dissemination of bacteria throughout the body. On the other hand, cell factors in the hemocytes have been reported to coagulate by serine proteases which may be activated by lipopolysaccharide or \(\beta\)-1,3-glucans [2,8]. This is also linked to the prophenoloxidase activating system [9]. Thus the coagulation is part of humoral immune response in crustaceans [10,11].

The tiger shrimp, *Penaeus monodon*, is an economically important species cultured in Taiwan and south-eastern Asia. We have purified and characterized the clottable proteins of tiger shrimp and several crustaceans [12]. Being homodimeric glycoproteins of about 380–400 kDa, the clottable proteins of Penaid shrimps [12], lobster [13], crayfish [14], freshwater giant prawn [12], and sand crayfish [15] appear to have similar amino acid compositions and N-terminal sequences. The purified clottable proteins formed stable clots by the transglutaminase from the shrimp hemocyte lysate in the presence of Ca\(^{2+}\). We report herein the complete nucleotide and protein sequences, carbohydrate structure and tissue distribution of the tiger shrimp clottable protein. Its structural features and motifs are studied in detail and compared with those of the crayfish clotting protein whose sequence was published very recently [16].
**MATERIALS AND METHODS**

**Purification of clottable protein from hemolymph**

Collection of hemolymph and hemocytes from tiger shrimps (*P. monodon*), and purification of the clottable protein were as described [12]. The molecular mass of the protein was characterized by matrix-assisted laser desorption/ionization time-of-flight MS and SDS/PAGE [12]. Protein concentrations were determined by the Bradford method [17], using BSA as standard.

**Circular dichroism**

The clottable protein stock solution was mixed with buffers to prepare samples of constant protein concentration but with varied salt concentrations or pH. CD measurements were carried out on a J720 spectropolarimeter (Jasco) under constant flushing of nitrogen at 25 °C. Each sample was scanned from 200 nm to 250 nm, and two independent experiments for each sample were performed and found to give identical spectra. The mean residue ellipticity [6] was calculated from the mean residue weight. The spectra were used for analyses of secondary structure of the protein with a computer program provided by Jasco and based on the method of Yang et al. [18].

**Fragmentation and partial amino acid sequencing**

The purified protein (2 mg) was dissolved in 100 μL of 6 M guanidine hydrochloride, 250 mM Tris/HCl, pH 8.5 and reduced by dithiothreitol (0.09 M) at 50 °C for 1 h. After adding 30 μL of 1 M iodoacetic acid, the mixture was incubated at room temperature for 30 min. The desalted and lyophilized S-carboxymethyl protein was re-dissolved in 100 μL of 70% formic acid with 0.2 mg of cyanogen bromide (CNBr) at room temperature for 24 h. On the other hand, the native clottable protein (4 mg·mL⁻¹) was digested at 37 °C for 18–24 h, either with Lys-C endopeptidase (Promega) in 50 mM Tris, pH 8.0, or with V8 protease (i.e. Glu-C, Promega) and 50 mM sodium phosphate buffer, pH 7.8, at an enzyme to substrate ratio of 1 : 50 (w/w). The reaction was stopped by adding dithiothreitol and heated at 95 °C for 5 min. The resultant polypeptides were fractionated by reversed-phase HPLC using a Chemosorb C18 column. The amino acid sequences of the purified peptides were determined with a on-line phenylthiohydantoin-derivative analyzer.

**General methods in molecular biology**

Standard procedures in molecular biology were used for preparation of plasmid DNA, restriction enzyme digestion, DNA agarose gel electrophoresis, DNA ligation, and the transformation of bacteria [19].

**RNA isolation and cDNA library construction**

Total RNA was isolated from the tiger shrimp by extraction in acid guanidinium thiocyanate as described [20]. For library construction a total of approximately 1 mg of total RNA was isolated from the whole shrimps. A cDNA library prepared from poly(A)-enriched RNA by unidirectional insertion of cDNA into λ-ZAP II [21] was constructed using a kit from Stratagene.

**cDNA synthesis**

The total RNA was purified using the RNAzol B kit (Biotex) and the mRNA was purified using QuickPrep Micro mRNA purification kit with oligo(dT)-cellulose chromatography (Pharmacia). The first strand cDNA synthesis was primed with a hybrid oligo(dT) linker-primer and random primers, and was transcribed using moloney murine leukemia virus reverse transcriptase (Gibco BRL). The synthesized cDNA was used as a template in subsequent PCR [22].

**PCR and screening of cDNA library**

Degenerate primers were designed from N-terminal sequence QPGLEYQY (forward primer) and some known peptide sequences of the clottable protein (reverse primers). A PCR product of 500 bp was obtained from degenerate primers, the design being based on the sequences QPGLEYQY and AEEENVQ. Amplified DNA fragments were then purified and ligated into pGEM-T (Promega). Each clone was sequenced by the dideoxy chain termination method [23] with Sequenase (US Biochemical) according to the manufacturer’s instructions. Among the seven clones sequenced, three contained other determined peptide sequences. Thus this 500-bp DNA was used as a probe to screen the shrimp cDNA library by using a digoxigenin (DIG) DNA Labeling kit (Boehringer Mannheim). Approximately 1 × 10⁶ amplified clones were plated at a density of 5 × 10⁵ plaque-forming units per 150-mm petri dish. Hybridization and washing were carried out as previously describe [19]. In brief, nitrocellulose (Drassel) lifts of the phage plates were hybridized at 42 °C overnight in 50% (v/v) formamide containing 5 × NaCl/Cit (NaCl/Cit = 150 mM NaCl, 15 mM sodium citrate, pH 7.5), 0.1% (w/v) SDS, 0.5% (w/v) N-laurylsarcosine, 2% blocking reagent (Boehringer Mannheim), and a DIG-labeled probe. Following hybridization, filters were washed in 2 × NaCl/Cit, 0.1% SDS at 25 °C for 30 min, then in 0.1 × NaCl/Cit, 0.1% SDS at 65 °C for 30 min. Signals were detected using the DIG luminescent detection kit for nucleic acids (Boehringer Mannheim). Clones of interest were further purified by three more screen cycles. The clones were subcloned into the EcoRI site of pBluescript II KS(+) phagemid (Stratagene).

**Immunoscreening of the shrimp cDNA library**

The shrimp expression cDNA library was screened by anti-clottable protein antibodies [19]. The λ-ZAP phages were plated at a density of 5 × 10⁵ plaques per agar plate. After incubation for 3.5 h at 42 °C, the plates were overlaid with nitrocellulose filters (0.45 μm; Micron) that had been impregnated with 10 mM isopropyl-1-thio-D-galactopyranoside. Incubation was continued for 5 h at 37 °C. The filters were then removed, washed with NaCl/P₀ (0.12 M NaCl, 10 mM phosphate) at room temperature, and blocked with 1% polyvinyl-pyrrolidone in NaCl/Pₐ for 16 h at 4 °C. Following blocking, filters were probed with a polyclonal antiserum specific for shrimp clottable protein [12] at a 1 : 500 dilution in NaCl/Pₐ, containing 1% poly(vinyl pyrrolidone), 1 mM EDTA, and 0.4% Triton X-100 at 25 °C for 1 h. The filters were then washed three times with NaCl/Pₐ and incubated with horseradish peroxidase-conjugated anti-rabbit IgG (Sigma) for 1 h at room temperature. The immune complexes were then incubated in 10 mL of NaCl/Pₐ containing 0.2 mg mL⁻¹ dianisobenzamidine and viewed after adding 10 μL of 30% H₂O₂. Phages displaying strong signals were isolated for secondary and third
screening. One positive clone (Cp 5) was isolated after screening 3 × 10^6 plaques, and the clone was subcloned into the EcoRI site of pBluescript II KS(+) phagemid (Stratagene).

Sequence homology and hydrophatic profile

A search for related sequences using data from GenBank, and SWISS-PROT was carried out [24]. Amino acid sequence alignment of the shrimp clotting proteins with crayfish clotting protein [16] was accomplished with the CLUSTAL W multiple alignment program (version 1.64B) [25]. The hydrophatic profile of the shrimp clotting protein sequence was plotted following the method of Kyte and Doolittle [26]. Windows of seven amino acid residues were used to calculate the average hydropophicity of the central residue.

Northern blot analysis

Total RNA was isolated from various shrimp tissues using the RNAzol B kit (Biotex). Total RNA (20 μg) from each tissue was fractionated on 1% formaldehyde/agarose gel in 4-morpholinepropanesulfonic acid buffer and transferred onto a Hybond-N membrane (Amersham). Following prehybridization for 3 h in 50% formamide, 5 × NaCl/Cit, 2% blocking reagent, 0.1% N-laurylsarcosine and 0.02% SDS at 42 °C, the blots were hybridized with a PCR-generated DIG-labeled cDNA probe for 20 h under identical conditions. The blots were washed twice for 5 min with 2 × NaCl/Cit containing 0.1% SDS at room temperature and twice for 15 min at 65 °C with 0.1 × NaCl/Cit containing 0.1% SDS. Detection of DIG signals was accomplished using the DIG luminescent detection kit. The cDNA probe used for Northern blotting was synthesized by a PCR DIG probe synthesis kit (Boehringer Mannheim) with two opposing primers. It was used in the Northern blotting experiments as an internal control of the RNAzol B kit (see Materials and methods) whereas Cp 5 was isolated by immunoscreening. Other clones, Cp 2, Cp 3, Cp 4, were isolated by using both Cp 1 and Cp 5 as probes. Abbreviations of restriction enzyme sites are denoted as: B, BglII; E, EcoRI; H, HindIII; X, XhoI.

CLONING OF TIGER SHRIMP β-ACTIN GENE

Degenerate primers were designed according to the amino acid sequences that are highly conserved in β-actin. The amino acid sequences for designing the two opposing primers are QIMFETF and MKCDVDI. Using these primers and the sequences for designing the two opposing primers are CATC (i.e. forward primer, nucleotides 3915±3934) and GAGGCGATGTGAGAACACATC (i.e. reverse primer, nucleotides 4384±4403).

Carbohydrate analysis

The lyophilized clotting protein (20 μg) was dissolved in 100 μL of water and added to an equal volume of either 4 M trifluoroacetic acid or 8 M hydrochloric acid in a 1.5-mL screw-capped vial; each vial was heated at 100 °C for 4 h or 6 h [27], respectively. The hydrolyzates were evaporated to dryness under vacuum using a Speedvac evaporator. Monosaccharide contents were analyzed with high-performance anion exchange chromatography (HPAEC) consisting of a Bio-LC system (Dionex) equipped with a CarboPac PA-1 column (9 × 250 mm) and a pulsed amperometric detector (PAD-II) [27]. The chromatographic data were managed with AI-450 chromatography software (Dionex). Seven monosaccharides including glucose, mannose, galactose, 6-methylgalactose, fucose, N-acetylglucosamine and N-galactosamine were used as standards. On the other hand, the clotting protein (40 μg) dissolved in 50 μL of water was denatured by heating at 100 °C for 5 min. After cooling down it was added 5 μL of endo-β-N-acetylglucosaminidase and 15 μL of sodium acetate buffer (pH 6.5) and incubated at 37 °C for 18 h. The released oligosaccharides were analyzed by a Dionex HPAEC [28], using the established conditions [29]. Ribonuclease B (Sigma) was hydrolyzed by the same enzyme to give Man_gN CafAc~Man_gN CafAc standards.

Binding test of biotinylated clotting protein

Following the protocol of an immunoprobe biotinylation kit (Sigma), 0.75 mg of purified clotting protein in 200 μL of 0.1 M sodium phosphate buffer (pH 7.4) was mixed with 10 μL of biotinamidocaproate-N-hydroxy-sulfosuccinimide ester (10 mg·mL⁻¹) and incubated with gentle stirring for 30 min at room temperature. The biotinylated protein was purified by a Sephadex G-25 column (1 × 10 cm) equilibrated with NaCl/Pi and stored at −20 °C until use; it was confirmed by Western blot with ExtraAvidin conjugated with alkaline phosphatase (Sigma). The hemolymph was freshly withdrawn from the shrimp and fixed with 10% formalin in 2.25 × NaCl/ P for 15 min at room temperature. The hemocytes were spun down at 500 g for 10 min washed twice with NaCl/P, and then smeared on slides. The slides were overlaid with biotinylated clotting protein (50 μg·mL⁻¹) for 1 h at room temperature and then washed with NaCl/P. Finally, the slides were incubated with 100 μL of 50 times-diluted fluorescence isothiocyanate-conjugated UltraAvadin (Leinco) for 1 h. After washing with NaCl/P, the slides were mounted in 50% glycerol in NaCl/P for observation under a fluorescence microscope (Olympus AHBS3).

RESULTS

Partial amino acid sequence

The N-terminal amino acid sequence of the native protein was determined up to the 30th residue [12]. The clotting protein

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>XH</td>
<td>1</td>
<td>1000</td>
<td>2000</td>
<td>3000</td>
</tr>
</tbody>
</table>

Fig. 1. Restriction enzyme map (A) and sequencing strategy (B) of the clotting protein cDNA. The clone Cp 1 was obtained by PCR (see Materials and methods) whereas Cp 5 was isolated by immunoscreening. Other clones, Cp 2, Cp 3, Cp 4, were isolated by using both Cp 1 and Cp 5 as probes. Abbreviations of restriction enzyme sites are denoted as: B, BglII; E, EcoRI; H, HindIII; X, XhoI.
Table 1. Partial amino acid sequences of tiger shrimp clottable protein. X denotes unidentified residues.

<table>
<thead>
<tr>
<th>Cleaving agent</th>
<th>Fragment</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNBr</td>
<td>CN1</td>
<td>AADYSVVQFSNIEVGDLNKLVDL</td>
</tr>
<tr>
<td>CNBr</td>
<td>CN2</td>
<td>KLPVNLAEEVNVQREH</td>
</tr>
<tr>
<td>Glu-C</td>
<td>E1</td>
<td>QAOQKTQQQVQGTQWEEXFP</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K1</td>
<td>GNFVINRVLFWSTEL</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K2</td>
<td>LNYGAVEEVLGQRX</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K3</td>
<td>TSPQQLANPVHPIDTLLWSVR</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K4</td>
<td>IMHSLINGEGEL</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K5</td>
<td>HLKPKPASAPILSTNFHHF</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K6</td>
<td>RPQSSQAEISVDMWEELKE</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K7</td>
<td>IEAEIKGDXSTYPGIDIT</td>
</tr>
</tbody>
</table>

was also digested directly with Lys-C and Glu-C endoproteinases since its solubility was greatly decreased when the clottable protein was reduced by dithiothreitol. We obtained a total of ten purified polypeptides from reversed-phase HPLC, including two from CNBr cleavage, one from Lys-C digest, and seven from Glu-C digest. Table 1 shows their amino acid sequences.

cDNA cloning of the clottable protein

Degenerate primers were designed based on the partial amino acid sequences obtained from Table 1. After PCR amplification, we isolated a 500-bp clone (Cp 1, see Fig. 1) which encodes the N-terminal sequence of the clottable protein. On the other hand, we got the Cp 5 clone by immunoscreening. The cDNA library was re-screened using both Cp 1 and Cp 5 as probes, and Cp 2, Cp 3 and Cp 4 were thus cloned. These five clones were isolated and sequenced to complete the entire nucleotide sequence of the clottable protein. The deduced amino acid sequence is shown in Fig. 2B. The cDNA has a total length of 6124 bp, including 502 bp of the 5′-untranslated region, an open reading frame of 5010 bp, and 612 bp of the 3′-untranslated region. The putative initiating ATG codon, which agrees with Kozak’s rule [30], is at nucleotide 503. The open reading frame is predicted to encode a protein of 1670 amino acids, including a 14-amino acid signal peptide, two RGD (Arg-Gly-Asp) motifs and four potential N-glycosylation sites (Fig. 2B). The authenticity of the cDNA sequence is confirmed by its accordance with the amino acid sequences of the fragments obtained from CNBr cleavage and proteolytic digests of the clottable protein. The results are consistent with its molecular mass determined by matrix-assisted laser desorption/ionization time-of-flight MS [12].

Table 1. Partial amino acid sequences of tiger shrimp clottable protein. X denotes unidentified residues.

<table>
<thead>
<tr>
<th>Cleaving agent</th>
<th>Fragment</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNBr</td>
<td>CN1</td>
<td>AADYSVVQFSNIEVGDLNKLVDL</td>
</tr>
<tr>
<td>CNBr</td>
<td>CN2</td>
<td>KLPVNLAEEVNVQREH</td>
</tr>
<tr>
<td>Glu-C</td>
<td>E1</td>
<td>QAOQKTQQQVQGTQWEEXFP</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K1</td>
<td>GNFVINRVLFWSTEL</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K2</td>
<td>LNYGAVEEVLGQRX</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K3</td>
<td>TSPQQLANPVHPIDTLLWSVR</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K4</td>
<td>IMHSLINGEGEL</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K5</td>
<td>HLKPKPASAPILSTNFHHF</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K6</td>
<td>RPQSSQAEISVDMWEELKE</td>
</tr>
<tr>
<td>Lys-C</td>
<td>K7</td>
<td>IEAEIKGDXSTYPGIDIT</td>
</tr>
</tbody>
</table>

Fig. 2. Deduced amino acid sequence and 5′ and 3′ untranslated regions of the shrimp clottable protein. (A) Nucleotides 1–600 of the shrimp clottable protein cDNA, containing the 5′ untranslated region (1–502) and the beginning of the coding region (503–600). The deduced amino acid sequence up to residue 33 is also shown. Repeated nucleotide sequences are underlined. (B) Deduced amino acid sequence of the clottable protein. The sequences confirmed by protein sequencing are underlined. The putative signal sequence, N-glycosylation sites, and RGD motifs are doubly underlined. (C) Nucleotides 5401–6124 of the shrimp clottable protein cDNA, containing the stop codon (5512–5514) and the 3′ untranslated region (5515–6124). Three types of repeated sequences are underlined. The two polyadenylation signals are doubly underlined.
Circular dichroism and fluorescence spectra

Figure 3A shows the CD spectra of the purified clottable protein in neutral buffer. The clottable protein is predicted to contain 44% α helix, 26% β sheet, 16% β turn. The tiger shrimps are normally grown in seawater of 0.5±4.5% salinity (i.e. 0.09±0.77 m NaCl), but its CD spectra were not affected by the addition of 0.1±0.5 m NaCl salt (Fig. 3A) or divalent metal ions (e.g. 0.1±5 m CaCl2, 0.1±5 m MgCl2 or 0.1 m ZnCl2) to the protein solution. The protein was stable in buffers of pH between 4 and 9 but denatured at pH $\geq 10$ (Fig. 3B).

The intrinsic fluorescence spectrum of the clottable protein has an emission maximum at 336 nm when excited at 295 nm, typical of tryptophan fluorescence (data not shown). The addition of 5 mM CaCl2, 5 mM MgCl2, 0.1 mM ZnCl2 or 5 mM EDTA did not change the fluorescence spectra. Moreover, the absence of Ca$^{2+}$ or Zn$^{2+}$ in the protein was confirmed by atomic absorption analysis.

Glycan analyses

Sugar composition analyses confirmed that the tiger shrimp clottable protein contains mannose (2.6%) and N-acetylgalactosamine (1.2%) (Fig. 4). Although glucose was also found, its value varied widely from analysis to analysis, and thus it was regarded as a contaminant and not an integral sugar constituent. The sugar composition was also analyzed by a precolumn labeling method [31] with the same result. After heat denaturation and endo-β-N-acetylgalactosaminidase digestion, the glycans released from the clottable protein were subjected to HPAEC analysis. The high-mannose oligosaccharides of ribonuclease B released by endo-β-N-acetylgalactosaminidase were used as standards (Fig. 5A). It was found that the oligosaccharides released from the clottable protein were mainly Man9GlcNAc and Man8GlcNAc (Fig. 5B).

Northern blot

The presence of mRNA of the clottable protein in different shrimp tissues was examined by Northern blotting using its partial cDNA as the probe (Fig. 6A). The same tissue blot was also examined with a tiger shrimp β-actin probe (Fig. 6B) as an internal control. The DIG-labeled cDNA probe for the clottable protein hybridized with a single 6.2-kb band in all the samples except the hemocytes, and the highest levels were found in gill and heart (Fig. 6A). The size of the mRNA identified agrees with that expected for the clottable protein clone (6124 bp). The size of the β-actin mRNA identified is about 2 kb, also similar to that of brine shrimp actin [32].

Hemocytes binding test

The clottable protein sequence contains two Arg-Gly-Asp motifs that are potential sites for cell adhesion or binding [33]. The binding of the biotinylated clottable protein to fixed...
hemocytes was examined by immunofluorescence but no significant response was detected (data not shown).

**DISCUSSION**

Among invertebrate coagulation, the clotting system in horseshoe crab has been well-studied as a model [34,35]. From sequence homology search, the 380 kDa tiger shrimp clottable protein is most similar to the 420 kDa clotting protein of a freshwater crayfish recently published [16]. So far only these two homologous clotting proteins from crustacean have been completely sequenced (Fig. 7); they are 36% identical or about 57% similar. Their low sequence identity is consistent with the fact that the anti-(tiger shrimp clottable protein) serum cannot recognize the crayfish (*Procambarus clarki*) hemolymph clotting protein [12]. They also show even lower sequence similarities to insect vitellogenins [36], apolipoprotein B [37] and to the D domain of mammalian von Willebrand factor [38]. Thus crustacean clottable proteins are evolutionarily related to vitellogenins but apparently play different functions. They do not resemble any protein in the coagulation cascade of horseshoe crab [39] or the vertebrate fibrinogens [13]. The shrimp as well as crayfish clotting proteins are plasma proteins in contrast to the horseshoe crab coagulogen which is much smaller and localized in the blood cells [39]. Apparently, coagulation in crustaceans is rather different from that in horseshoe crab and many types of clotting mechanism exist among the vast number of existing invertebrate species.

The hydrophobic profile (Fig. 8) reveals that the shrimp clottable protein contains a relatively hydrophobic region between amino acid residues 4 and 14, which fits the known feature of signal sequence. No large hydrophobic regions or membrane spanning domain can be identified. Major hydrophilic regions in the protein are located at regions 179±234, 306±323, 771±786, 1077±1090, 1585±1604, and 1625±1656. Three potential N-glycosylation sites at positions 106, 319, and 1301 are located at hydrophilic regions (Fig. 8).

In contrast to other members of the vitellogenin family, the most prominent structural features of the shrimp protein include a Lys-rich domain containing five Ser-Lys-Thr-Ser repeats at positions 203±231, another Ser/Thr-rich region at 1038±1292, a Gln-rich region at 1088±1132, and a polyglutamine run (QQQAQQQQQQQQQQQQQQQQQQQQQQQ) at the C-terminal region. Except for the polyQ, these features are also conserved in the crayfish clotting protein (Fig. 7), but crayfish clotting protein contains less SKTS repeats but more TKTTG repeats [16] than the shrimp counterparts. A potential clue to the functional roles of the polyQ and Ser-Lys-Thr-Ser repeats of tiger shrimp is that the protein is a substrate for the hemocyte transglutaminase. Structural homology of arthropod transglutaminase to that of the mammalian enzymes was reported [40] and previous study...
revealed that the shrimp enzyme has similar properties to the vertebrate platelet transglutaminase or the α-subunit of coagulation Factor XIII [41]. The transglutaminase-catalyzed reaction involves a Ca2+-dependent acyl transfer [6,42]. When a Gln-containing acyl donor polypeptide binds to the enzyme, the nucleophilic attack of the enzyme active site on the γ-carboxamide group of a Gln residue leads to formation of an acyl-enzyme intermediate with release of ammonia. The Lys-containing polypeptide then binds, and the acyl group is transferred to its ε-amino group, resulting in formation of an isopeptide bond. The shrimp clottable protein presumably also contains specific donor (Gln) and acceptor (Lys) residues to bind to the enzyme active site.

According to previous results on the substrate specificities of mammalian transglutaminase [43–45], only certain amino acid residues frequently precede the acceptor lysine residues, e.g. Ser, Val, Gln, or Thr. We notice that there are in total ten Ser-Lys, five Val-Lys, five Gln-Lys, and eight Thr-Lys arrangements in the clottable protein sequence. For example, KSK (residues 205–207) and KQK (residues 772–774) of the shrimp protein bear identical sequences to those of the active acceptors in another transglutaminase substrate [44]. The C-terminal polyQ and the Gln-rich region also have high potential for isopeptide formation [46,47]. As the hydrophobic profile of the protein (Fig. 8) revealed, these Lys-rich or Gln-rich regions and polyQ are located in major hydrophilic region and easily accessible surface of the clottable protein. Thus some of these Gln and Lys residues of the protein are probably utilized in the clotting process catalyzed by the hemocyte derived transglutaminase.

Ten cysteine residues and a GICG or GLCG motif (at position 1511–1514 of the shrimp protein) are conserved in the C-terminal regions of insect vitellogenins and the crustacean clottable proteins (Fig. 7). Previously, a similar motif TCGLCG was found to be conserved in various invertebrate and vertebrate vitellogenins, and in the domains D1 and D2 of human von Willebrand factor and the domain D3 of human mucin 2 [36,48]. The D domains play essential roles in the polymerization of mucin 2 and von Willebrand factor into large polymers, and are involved in the formation of disulfide bonds between subunits [49,50]. Therefore the cysteine-rich C-terminal domain in the clottable protein probably also involves intersubunit interaction.

Functional RGD-recognizing receptors have been found in the granular hemocytes of crayfish [51] but the crayfish clotting protein does not contain any RGD motif [16]. Although two RGD motifs exist in the shrimp clottable protein, both are not flanked by disulfide bonds or other types of conformational constraint [52] and their binding to the hemocytes was not detected. Whether the RGD motif of this clottable protein is functional toward other cell types is not known.

The carbohydrate content (3.8%) of the shrimp clottable protein (Fig. 4) is lower than that reported for the crayfish clotting protein (about 20%, w/w) [53]. The crayfish protein
contains six potential N-glycosylation sites [16] in contrast to four in the shrimp protein; two of the sites are at identical positions in both proteins (Fig. 7). The N-glycan structure of the shrimp clottable protein belongs to the high mannose type very often found in glycoproteins of invertebrates (Fig. 5). However, results of the monosaccharide composition analysis of the shrimp protein gave a ratio of GlcNAc/Man = 2.0 : 5.4 (Fig. 4). It is likely that other side chains, e.g., GlcNAc-GlcNAc-Asn or GlcNAc-Asn which could not be cleaved by endo-β-N-acetylglucosaminidase or O-linked GlcNAc [54], are also present. Alternatively, there could have been disproportionate decomposition of Man over GlcN, which is often observed during acid hydrolysis of glycoproteins of low carbohydrate contents.

Northern analyses (Fig. 6A) revealed that the clottable proteins are expressed in most of the shrimp tissues. Similar to that found in the crayfish [16], the clottable protein is also expressed in hepatopancreas of the shrimp (Fig. 6A, lane 4). The expression of the clottable protein is at its highest levels in gill and heart of the shrimp (Fig. 6A, with a lower level in hepatopancreas, lymphoid organ and muscle, but not in the mature hemocytes (Fig. 6A, lane 3). In order to confirm the Northern results, the blot is also examined with a tiger shrimp β-actin probe (Fig. 6B). In mature hemocytes, the actin gene transcript we obtained was intact and the corresponding band on the gel was as strong as those from the other tissues (Fig. 6B). Therefore the degradation of the clottable protein messenger during total RNA isolation from the mature hemocytes should not occur (Fig. 6B, lane 3). Using an isotonic anti-coagulant improved since previous experiments [59], we could prepare washed and intact shrimp hemocytes, but could not detect the protein in the hemocyte lysate in immunocytochemical experiments (data not shown). From this evidence, we conclude that the mature shrimp hemocytes apparently do not express, synthesize or contain the clottable protein. Previous studies also showed that clottable protein was not found in the hemocytes of other decapods, e.g., Sicyonia ingentis [55] and Penaeus japonicus [56]. This seems logical since active transglutaminase causing polymerization of the protein is localized in the hemocytes [41]. However, the clottable protein is expressed in the shrimp lymphoid organ (Fig. 6A, lane 5) which is part of the hematopoietic tissues capable of generating mature hemocytes [57,58]. Whether or not the hemocytes express the clottable protein before they become mature and released is not clear.

We have previously used the rocket immunoelectrophoretic analyses to quantitate the hemolymph clottable protein of tiger shrimp [59]. The basal level of tiger shrimp clottable protein is about 3 mg·mL−1 of hemolymph in normal intermolt tiger shrimp; it increased by twofold after molting and decreased after molting to the normal level. It could be elevated about 4-fold by repetitive bleeding or surgical ablation of tail or eyestalks of the shrimp [59]. Results of Northern analysis indicate that the highest levels of expression of the clottable protein are at gill and heart of the shrimp (Fig. 6A). These organs play a major role in the hemolymph circulation. Presumably these organs are sensitive to bleeding conditions, and may express clottable protein rapidly to compensate for its loss and to help wound healing. By analogy, in vertebrate blood the expression of fibrinogen increases many fold under acute-phase condition.

In the cDNA of crayfish clotting protein, repeated sequences are found in the 3′ and 5′ untranslated regions and are suggested to be involved in regulating translation and/or stabilizing the mRNA [16]. Similar repeats in the 3′ and 5′ untranslated regions were also identified in the shrimp gene (Fig. 2). The genomic structure, especially the promoter region of the clottable protein gene, should be investigated in more detail to understand the transcriptional regulation of coagulation in crustacea.

**ACKNOWLEDGMENTS**

We thank Shun-Wen Chen and Chen-Sien Chang for amino acid sequencing, Chih-Ming Chou and The-Li Su for DNA sequencing, Mei Tang for carbohydrate analyses, Kuan-Fu Liu and Wen-Ten Cheng for collecting shrimp hemolymph, and Kay-Hooi Khoo for reading of the manuscript. Materials in this paper form part of the dissertation submitted by M.-S. Y. in partial fulfillment for the requirement of the degree of Doctor of Science at the National Taiwan University.

**REFERENCES**


43. Kao, L.R. (1987) Studies on shrimp (Penaeus monodon) hemolymph clotting system and its transglutaminase. Masters Thesis, National Taiwan University, Taiwan.