2015

Purification, Characterization, and Analysis of the Allergenic Properties of Myosin Light Chain in *Procambarus clarkii*

Yong-Xia Zhang
Jimei University

Heng-Li Chen
Jimei University

Soheila J. Maleki
USDA-ARS, soheila.maleki@ars.usda.gov

Min-Jie Cao
Jimei University

Ling-Jing Zhang
Jimei University

See next page for additional authors

Authors

Yong-Xia Zhang, Heng-Li Chen, Soheila J. Maleki, Min-Jie Cao, Ling-Jing Zhang, Wen-Jin Su, and Guang-Ming Liu

Purification, Characterization, and Analysis of the Allergenic Properties of Myosin Light Chain in *Procambarus clarkii*

Yong-Xia Zhang,[†] Heng-Li Chen,[†] Soheila J. Maleki,[‡] Min-Jie Cao,[†] Ling-Jing Zhang,[†] Wen-Jin Su,[†] and Guang-Ming Liu*,[†]

[†]College of Food and Biological Engineering, Fujian Collaborative Innovation Center for Exploitation and Utilization of Marine Biological Resources, Jimei University, 43 Yindou Road, Xiamen, Fujian 361021, China

[‡]Agriculture Research Service, Southern Regional Research Center, U.S. Department of Agriculture, 1100 Robert E. Lee Boulevard, New Orleans, Louisiana 70124, United States

Ⓢ *Supporting Information*

ABSTRACT: Myosin light chain (MLC) plays a vital role in cell and muscle functions and has been identified as an allergen in shrimp. In this study, MLC with a molecular mass of 18 kDa was purified from crayfish (*Procambarus clarkii*) muscle. Its physicochemical characterization showed that the purified MLC is a glycoprotein with 4.3% carbohydrate, highly stable to heat, acid−alkali, and digestion, and weakly retains IgE-binding activity when its secondary structure was altered. Serological assays suggested that conformational epitopes predominate over linear epitopes in the purified MLC. Two isoforms of the *MLC* gene (*MLC1* and *MLC2*) were cloned, and the purified MLC was identified as MLC1. Analysis of the secondary and tertiary structures of the MLCs indicated that MLC1 has four conformational epitopes and three linear epitopes, whereas MLC2 had a major conformational epitope and three linear epitopes. These results are significant for understanding hypersensitization of humans to crayfish.

KEYWORDS: *Procambarus clarkii, myosin light chain, purification, physicochemical characterization, allergenicity, epitopes*

■ INTRODUCTION

Seafood, including fish, crustaceans, and mollusks, is a widely consumed source of dietary protein because of its delicate flavor and high nutritive value. China, Japan, and North America are the three largest consumers of seafood, but the increasing consumption of seafood has been followed by increases in adverse reactions.[1] Most of these are mediated by immunoglobulin E (IgE) and generate mucocutaneous (95.7%), respiratory (23.9%), gastrointestinal (16.3%), anaphylactic (11.9%), and cardiovascular symptoms (3.3%).[2]

Procambarus clarkii, commonly known as the crayfish or freshwater lobster, is native to south-central United States and northeastern Mexico and is one of the shellfish with the greatest commodity value worldwide. The crayfish is characterized by its high protein, low fat, low calorie, and high nutritional value, as well as its delicious taste and texture. Therefore, it is considered to be a popular food by consumers.[3,4] The increased consumption of crayfish has been accompanied by more frequent allergic reactions.[2] In China, *P. clarkii* is an important economic resource among freshwater shellfish, and the total production of crayfish in 2013 was about 603 520 tons, 8.78% higher than the 554 821 tons produced in 2012.[5] A survey of 182 children suffering from urticaria in China reported that 58.79% of subjects had an allergic reaction to seafood, including crayfish, indicating that crayfish is an important allergenic food.[6] Three well-characterized allergens in crayfish have been reported. The heat-stable myofibrillar protein tropomyosin (TM) is recognized as a major allergen of crayfish.[7−9] Generally, TM is reported to be the pan-allergen of seafood, occurring in shrimp,[10,11] lobster,[7] crab,[12] and mollusks.[13] Arginine kinase (AK), a myosinogen protein

with a molecular mass of 40 kDa, is also a major allergen of crayfish[14] and has been identified as a major cross-reactive allergen in mollusks and crustaceans.[10,13] Another water-soluble allergic myosinogen protein discovered in crayfish is sarcoplasmic calcium-binding protein (SCP), which consists of three subunits and occurs as three isoforms with molecular masses of 22 kDa, and all of its subunits and isoforms are allergenic.[6] SCP is also a shrimp allergen in the white leg Pacific shrimp (*Litopenaeus vannamei*) and the black tiger prawn (*Penaeus monodon*).[15,16] These three allergens share similar physicochemical characteristics with most food allergens in that they are glycoproteins, are relatively stable to heat and acid−alkali, and retain their immunological activity after digestion.[14] Several novel allergens of crustaceans have also been identified, such as hemocyanin,[17] triosephosphate isomerase,[10] and troponin C.[10]

Myosin light chain (MLC) from *L. vannamei*, the North Sea shrimp (*Crangon crangon*), and *P. monodon* is recognized as a novel shrimp allergen.[10,18,19] Myosin, the most important functional muscle protein, is a multimeric protein composed of two heavy chains with a molecular mass of ∼200 kDa and four light chains with a molecular mass of ∼18−20 kDa.[20,21] These proteins are actin-binding molecular motors and play vital roles in a wide range of functions, including cell adhesion, phagocytosis, muscle contraction, and cytokinesis. The light chains are categorized into two classes: the essential light chain

(MLC1) with a molecular mass of ~18 kDa and the regulatory light chain (MLC2) with a molecular mass of ~20 kDa.[20,22] It has been demonstrated that MLC2 is phosphorylated by its kinase, which is regulated the expression of IgE, and is reported to be increased in the smooth muscle of the allergen-sensitized human airway.[23] MLC has been reported to have IgE-binding activity in *L. vannamei*, *C. crangon*, and *P. monodon*, and the immunological effects of recombinant MLC have been analyzed in these organisms. However, it is unclear whether MLC is the major allergen in *P. clarkii*, nor has native MLC been characterized or the isoform identified.

The objectives of this study were to identify and purify MLC from *P. clarkii* and to analyze its physicochemical and immunological characteristics during changes in its secondary structure. We were able to clone the cDNA encoding the MLC proteins and to identify the allergenic MLC isoform. A homology model of the MLC structure was developed and used to predict its epitopes. These data will extend our understanding of MLC and provide valuable references for further research into the molecular nature of hypersensitivity to crayfish.

■ MATERIALS AND METHODS

Crayfish. Crayfish (*P. clarkii*) were purchased from a local market in Xiamen, China. Crayfish were peeled, and the muscle samples were collected and immediately used in the experiments or frozen at −70 °C for later analysis.

Chemicals. Diethylaminoethyl−Sephacel (DEAE−Sephacel) and Sephacryl S-200 HR were from Amersham Biosciences (Uppsala, Sweden). The protein standards for sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) were from Fermentas (Vilnius, Lithuania), and the prestained protein standards for Western blot analysis were from New England BioLabs (Beverly, Massachusetts, USA). Horseradish peroxidase (HRP)-conjugated goat anti-human IgE antibody was from Kirkegaard and Perry Laboratories (Gaithersburg, Maryland, USA), and the enhanced chemiluminescence (ECL) substrate for Western blot analysis was from Pierce (Rockford, Illinois, USA). 3,3′,5,5′-Tetramethylbenzidine (TMB) was from Tiangen (Beijing, China). Ponceau S concentrate and porcine pancreatic enzymes were purchased from Sigma-Aldrich (St. Louis, Missouri, USA). Porcine pepsin was prepared in our laboratory. All reagents were of analytical grade.

Human Sera. Sera were collected from seven crustacean-allergic patients (nos. 749, 514, 974, 067, 785, 980, and 791), who were referred by the Medical Center of Jimei University in Xiamen, China. Serum samples from nonallergic (NA) individuals (nos. 780 and 981) were used as the negative controls. All patients and NA individuals voluntarily provided their sera (nos. U245749, U394514, T800974, U246067, U390785, U405980, U393791, U397780, and V490981). Written informed consent was obtained from each patient. The IgE level in each serum was measured with an ImmunoCAP assay (Phadia AB, Uppsala, Sweden), and the results are shown in Table 1. An ImmunoCAP score >0.35 kUA/L was considered positive. All sera were stored at −30 °C until analysis.

Purification of the Target Protein. The preliminary treatment was carried out as described previously,[24] with slight modifications. All procedures were performed at 0−4 °C unless otherwise stated. The crayfish muscle tissue was minced and homogenized with 10 vol (v/w) of ice-cold 50 mM NaCl buffer containing 2 mM $NaHCO_3$ and 10 mM EDTA. The homogenate was then centrifuged at 8000g for 10 min at 4 °C. The precipitate was resuspended in the same solution and then homogenized and centrifuged again four more times. The last precipitate, which contained the myofibrillar proteins, was washed three times with cold absolute acetone and dried completely. The dried powder was extracted overnight with 10 vol (v/w) of 20 mM Tris-HCl (pH 7.5) containing 1 M NaCl and 1 mM EDTA. The mixture was then centrifuged at 12 000g for 20 min at 4 °C. The supernatant was subjected to isoelectric precipitation at pH 4.0 with 1 N HCl. After centrifugation

Table 1. Clinical and Laboratory Characterizations of Seven Crustacean-Allergic Patients and Two Nonallergic Individuals[a]

serum no.	age	sex	symptoms	sIgE shrimp (f24) by ImmunoCAP (kUA/L)
749	23	F	U, P	11.1
514	19	M	E, N	1.28
974	18	M	N, P	45.1
067	22	F	A, D, R	9.68
785	19	F	GIT, Dr	61.7
980	19	F	U, D, P	1.93
791	24	M	A, D	2.35
780	19	F	ND	0.15
981	23	M	ND	0.06

[a]F, female; M, male; U, urticaria; P, pruritus; E, emesis; N, nausea; A, asthma; D, dyspnea; R, rhinitis; GIT, gastrointestinal problems; Dr, diarrhea; ND, not determined/unknown.

at 12 000g for 20 min at 4 °C, the precipitate was dissolved with 20 mM Tris-HCl (pH 7.5) and brought to pH 7.6 with 1 N NaOH and then subjected to ammonium sulfate fractionation with 41−60% saturated ammonium sulfate. Following centrifugation at 12 000g for 20 min at 4 °C, the precipitate was dissolved in 20 mM Tris-HCl (pH 7.5) containing 1 mM EDTA and then dialyzed in the same solution overnight. After centrifugation at 12 000g for 20 min at 4 °C, the supernatant was applied to a DEAE−Sephacel column (2.5 × 10 cm) and eluted with a linear gradient of 0−0.5 M NaCl at a flow rate of 0.6 mL/min. The fractions containing the target protein were concentrated by the ultrafiltration of a 3 kDa filtering membrane and then applied to a Sephacryl S-200 HR gel column (1.5 × 100 cm) equilibrated with 20 mM Tris-HCl (pH 7.5) containing 0.2 M NaCl. The column was eluted at a flow rate of 0.5 mL/min, and the fractions containing the target protein were collected and detected on a 12% SDS−PAGE gel stained with Coomassie Brilliant Blue R-250 (CBB). The protein concentrations in the eluted solutions were estimated by measuring the absorbance at 280 nm.

Peptide Analysis by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS). After the 18 kDa target protein had migrated on the 12% SDS−PAGE gel and was stained with CBB, the band was cut from the gel and sent to Shanghai Applied Protein Technology Company Limited (Shanghai, China) to digest with trypsin for analysis with a 4800 Plus MALDI-TOF/TOF analyzer from Applied Biosystems (Foster City, California, USA), based on the method of Yan et al.[25] The MS conditions were as follows: laser source, ND; wavelength, 355 nm; accelerating voltage, 2 kV; and substrate, α-cyano-4-hydroxycinnamic acid. The positive ion mode was used, and the data were obtained automatically. The resulting peptide spectra were compared with those in the National Center for Biotechnology Information (NCBI) database.

Carbohydrate Content and Carbohydrate−Peptide Linkage Characterization. The total carbohydrate content of the 18 kDa target protein was determined with the phenol−sulfuric acid method,[26] with some modifications. In brief, 0.3 mL of 6% phenol in water was added to the sample solution (0.5 mL) in a test tube, and 1.5 mL of concentrated sulfuric acid was added slowly down the side of the tube. The tube was then vortexed, covered, and incubated for 20 min at 100 °C. After all the tubes were cooled to room temperature, the absorbance of each was read at 490 nm. All the experiments were performed in duplicate. The β-elimination was used to determine the carbohydrate−peptide linkages in the target protein, according to a previously described method.[6] The target protein sample (0.3 mg/mL, 90 μL) was mixed with 10 μL of 0.5 M NaOH and incubated at 37 °C for 16 h, after which the absorbance at 220−300 nm was measured. The target protein sample was replaced with the same volume of distilled water, as the blank, and NaOH was replaced with the same volume of distilled water, as the negative control. All experiments were performed in duplicate.

Thermal Stability and pH Stability. To determine the thermal stability of the 18 kDa target protein, it was incubated in buffer (10 mM Tris-HCl, pH 7.5) at different temperatures (30, 40, 50, 60, 70, 80, 90, and 100 °C) for 30 min. The protein was incubated at 0 °C as the control. To test its pH stability, the target protein was incubated for 1 h at room temperature in buffer at different pH (1.0, 2.0, 3.0, 4.0, 5.0, 7.0, 9.0, and 11.0). The protein was incubated at pH 7.5 as the control. Then all the samples were analyzed with 15% SDS–PAGE and dot blot.

Stability during Simulated Digestion. The stability of the target protein in vitro in simulated gastric fluid (SGF, pepsin) and simulated intestinal fluid (SIF, pancreatin) was assessed with the method of Chen et al.,[14] with some modifications. The total volume of the reaction solution was 1 mL. For gastric digestion, the ratio of pepsin to the target protein was 1:500 (w/w). An aliquot of 100 μL of the reaction solution was removed after different time intervals (0, 1, 2, 5, 10, 15, 30, or 60 min), and the reaction was immediately terminated by the addition of 30 μL of 200 mM Na$_2$CO$_3$. For intestinal digestion, the ratio of pancreatin to the target protein was 1:500 (w/w). The reactions were performed for 0, 1, 15, 30, 60, 120, 180, or 240 min, after which they were immediately terminated by heating at 95 °C for 5 min. A target protein sample, dissolved in SGF/SIF with no corresponding enzyme, was used as the undigested control. The final concentration of each sample was adjusted to 0.25 mg/mL. The samples were analyzed with 15% SDS–PAGE and a competitive inhibition enzyme-linked immunosorbent assay (ELISA).

Western Blot and Dot Blot Analysis. Western blot analysis was performed as described by Chen et al.,[14] with some modifications. Briefly, protein samples separated with SDS–PAGE were electrophoretically transferred to a nitrocellulose membrane, and then the transfer efficiency was determined by Ponceau S staining of the membrane. The membrane was washed with TBST (20 mM Tris-HCl, pH 8.0, containing 0.15 M NaCl and 0.05% Tween-20) and blocked with 5% skim milk in TBST. They were then incubated for 6 h at room temperature with human serum (diluted 1:4) as the primary antibody and washed with TBST. The membrane was then incubated for 1.5 h with HRP-labeled goat anti-human IgE antibody (diluted 1:20000) as the secondary antibody. After being extensively washed with TBST, the proteins were visualized with ECL. In the dot blot analysis, the samples were blotted onto the nitrocellulose membrane, blocked, and incubated with sera from patients (diluted 1:4) and HRP-labeled goat anti-human IgE antibody (diluted 1:20000) as described above. The signal was detected with ECL.

Indirect and Competitive Inhibition ELISA (ciELISA). Each well of a polystyrene 96-well ELISA plate (Nunc Maxisorb, Denmark) was coated with the 18 kDa target protein (0.2 μg) and incubated at 4 °C overnight or at 37 °C for 2 h. The coated plates were washed with TBST and blocked with 200 μL/well 5% skim milk at 37 °C for 2 h. After being washed and blocked, the coated plates were incubated at 37 °C for 2.5 h with 50 μL of patient serum (diluted 1:5 with 1% skim milk in TBST) as the primary antibody and then at 37 °C for 1.5 h with HRP-conjugated goat anti-human IgE (diluted 1:2000) as the secondary antibody. TMB (100 μL) was then added to each well and incubated at 37 °C for 20 min. After the reaction was terminated by the addition of 50 μL of 2 M sulfuric acid, the absorbance was measured at 450 nm with an automated ELISA plate reader (Benchmark 96; Bio-Rad Laboratories, Hercules, California, USA). The specific IgE-binding activity of the purified protein-derived digestion fragments was confirmed with a ciELISA, based on the method of Li et al.[27] In contrast to indirect ELISA, after the plates were blocked with 5% skim milk, 50 μL of human sera (diluted 1:2 with 1% skim milk in TBST) was mixed with the same volume of inhibitor sample (the purified protein-derived digestion fragments), in the concentration range of 0.025–50 μg/mL in another plate. After incubation at 37 °C for 1.5 h, 50 μL of the mixture (serum plus inhibitor) was transferred to the ELISA plate precoated with the purified protein as the allergen and incubated for 1.5 h at 37 °C. As above, after incubation with the secondary antibody and reaction with TMB, the absorbance at 450 nm was measured. The inhibition rate was calculated to evaluate the percentage of inhibition by each inhibitor sample using the following formula:

$$\text{inhibition rate } (\%) = ([X - Y]/[X - Z]) \times 100$$

where X is the absorbance of the patient's serum not incubated with the inhibitor, and Y and Z are the absorbance of an allergic patient's serum and control NA serum reacted with the inhibitor, respectively. All the samples were tested in duplicate and the mean values calculated.

Circular Dichroism (CD) Spectroscopy. Far-UV CD spectra were measured with a Chirascan circular dichroism spectrometer (Applied Photophysics Ltd., Surrey, UK). The measurements of purified protein in buffer (10 mM Tris-HCl, pH 7.5) were made at 25 °C using a rectangular quartz cuvette (Starna, Essex, UK) with a path length of 0.1 cm at a protein concentration of 0.12 mg/mL. The purified protein was treated at different temperatures (0 and 100 °C) for 30 min or at room temperature in different pH buffers (pH 1.0, 7.5, and 11.0) for 1 h. The samples were scanned three times at 1 nm/s at a bandwidth of 1 nm. The response time was 2 s, and the wavelength range was 190–260 nm. The data were averaged, and the spectrum of the sample-free control was subtracted. The secondary structures of the proteins were determined with the CDNN program (version 2.0) based on the variables of the amino acid sequence and the molecular weight of the target protein.[28]

Cloning and Sequencing the MLC Genes. Total RNA was prepared from crayfish (*P. clarkii*) muscle based on the TRIzol method.[29] The first-strand cDNA was synthesized with the TIANScript RT kit (Tiangen, Beijing, China), according to the manufacturer's instructions, and oligo(dT)$_{18}$ primers were designed according to the manufacturer's instructions. Based on the MALDI-TOF-MS results and the alignment of the MLC cDNA sequences of *C. crangon* and *P. monodon*, degenerate oligonucleotide primers were designed as the sense primers for PCR. Using these primers and the synthesized cDNA, a 300 bp fragment of the *MLC* gene was amplified by PCR in the GeneAmp 9700 thermal cycler (Applied Biosystems, Foster City, California, USA). The PCR thermal cycling program was 5 min at 94 °C, followed by 30 cycles of 45 s at 94 °C, 30 s at 57 °C, and 45 s at 72 °C, and a final extension at 72 °C for 7 min. The PCR product was purified by agarose gel electrophoresis and cloned into the pGEM-T Easy vector (Promega), and the DNA sequence was analyzed. 5′-RACE primers and 3′-RACE primers were designed based on the sequence information from the RT–PCR product. RACE and RACE–PCR were performed with the SMART RACE amplification kit (Clontech, Palo Alto, California, USA). The PCR amplification product was purified, cloned, and sequenced. The full-length gene sequences of the *MLCs* (*MLC1*, 1096 bp; *MLC2*, 1362 bp) were obtained by overlapping the RT–PCR fragment, the 3′-RACE fragment, and the 5′-RACE fragment. DNA nucleotide sequencing was performed with the ABI PRISM 3730 DNA analyzer (Applied Biosystems, Foster City, California, USA), and the sequences were compared with the nucleotide sequences in the GenBank database. The sequences were submitted to GenBank under accession numbers JX173714 (gil401606250) for *P. clarkii MLC1* cDNA and JX013934 (gil392499120) for *P. clarkii MLC2* cDNA.

Structural Modeling and Epitope Prediction. The three-dimensional structure of the target protein was simulated at the Web site http://swissmodel.expasy.org/ based on the homology modeling by the amino acid sequence of the target protein to search the suitable template. The continuous linear epitopes were predicted by the DNASTAR Protean software using the method of Kyte–Doolittle for hydrophilicity plot prediction, the method of Karplus–Schulz for flexible regions prediction, the method of Jameson–Wolf for antigenic index prediction, and the method of Emini for surface probability plot prediction. The peptide fragments were selected to satisfy the conditions: (i) highly hydrophilic; (ii) appropriate antigenic index; and (iii) situated in a flexible and surface-exposed region. Discontinuous conformational epitopes were predicted at the DiscoTope 2.0 server (http://www.cbs.dtu.dk/services/DiscoTope/) by the calculation of surface accessibility, a novel epitope propensity, and the propensity of residues in spatial proximity. The three-dimensional structure and epitopes were represented on the molecular surface of the protein with PyMOL.
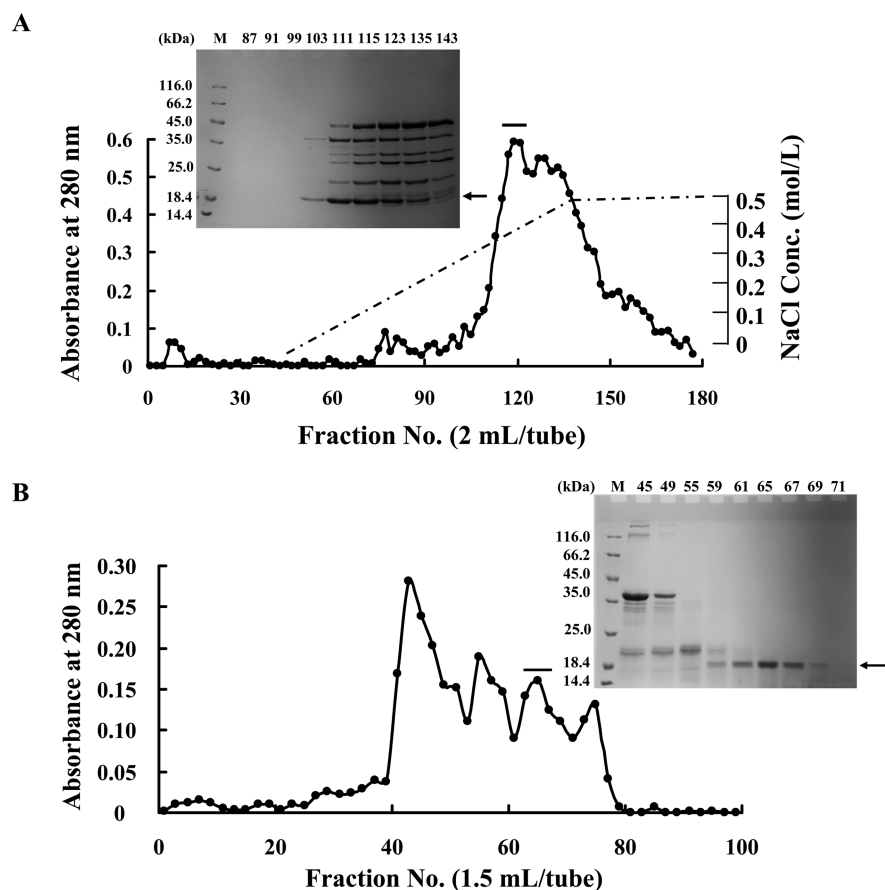
Figure 1. Chromatographic purification of the 18 kDa protein in crayfish. (A) DEAE−Sephacel chromatography. (B) Sephacryl S-200 HR gel chromatography. The numbers above the lanes correspond to the fraction numbers. Target protein fractions under the bar were pooled.

RESULTS

Purification of the 18 kDa Protein. The 18 kDa protein was purified from crayfish muscle with ammonium sulfate precipitation and sequential column chromatography on DEAE−Sephacel and Sephacryl S-200 HR. The fractions from the DEAE−Sephacel column were eluted with a linear gradient of 0−0.5 M NaCl (Figure 1A). The fractions that contained relatively high concentrations of the 18 kDa target protein were then pooled, concentrated, and separated on a Sephacryl S-200 HR gel-filtration column, which removed the contaminating proteins. After Sephacryl S-200 HR chromatography, the highly purified protein migrated on SDS−PAGE as a single band with a molecular weight of about 18 kDa (Figure 1B). Approximately 5.6 mg of the purified protein was obtained from 140 g of crayfish muscle.

Identification by MALDI-TOF-MS. The purified 18 kDa protein band was excised from the SDS−PAGE gel, destained, and digested in-gel with trypsin. The resulting peptides were analyzed with MALDI-TOF-MS. Peptide mass fingerprinting of the purified protein showed multiple peaks ranging from 800 to 4000 Da. The peaks with signal-to-noise ratios >50 were analyzed with MS/MS, as shown in Figure 2A. The effective peaks were compared with the NCBInr database using the Mascot 2.2 search engine. One peptide fragment containing 13 amino acid residues was 100% identical to MLC of *Drosophila yakuba* (gi| 195503715) and 92.3% identical to that of *C. crangon* (gi| 238477331) and *Marsupenaeus japonicas* (gi|290908988) (Figure 2B). These results confirmed the purified 18 kDa protein as MLC.

Glycoprotein Characterization. The results of the phenol−sulfuric acid analysis indicated that the total carbohydrates in MLC accounted for about 4.3 ± 0.05% (w/w) of the whole protein. The β-elimination was used to determine the presence of O-linked glycans in the MLC structure. As shown in Figure 3A, when MLC was treated with alkali, the UV absorbance at 240 nm increased significantly compared with the absorbance before alkali treatment. This result indicates that one or more O-glycans may be present in the structure of MLC.

Thermal and pH Stability. Thermal and pH stabilities are accepted characteristics of most allergens. The thermal stability of purified MLC was analyzed with SDS−PAGE and dot blot. As illustrated in Figure 3B, MLC was stable after exposure to different temperatures, even up to 100 °C. MLC IgE-binding activity at different temperatures was determined with dot blot using the patients' sera. The IgE-binding activity of crayfish MLC was detected in this assay and did not increase or decrease obviously in the temperature range of 30−100 °C (Figure 3C).

The pH stability of purified crayfish MLC was also examined with SDS−PAGE and dot blot. As illustrated in Figure 3D, MLC remained stable when exposed to acidic (pH 1.0−5.0) and alkaline (pH 9.0, 11.0) conditions. Dot blot was also used to analyze the IgE-binding activity of MLC. Compared with the control, no obvious change was detected in the IgE-binding activity of MLC after exposure to different pH values (Figure 3E).

Stability during Digestion. It was determined whether MLC was stable under in vitro conditions simulating human gastrointestinal digestion. The stability of purified MLC in

A   b ions: 89.08, 185.13, 299.18, 412.24, 526.31, 623.35, 724.43, 837.52, 908.56, 1021.63, 1134.69, 1263.73

y ions: 141.11, 274.17, 389.24, 502.33, 573.37, 686.47, 784.48, 884.57, 998.61, 1111.71, 1225.70, 1337.86

m/z: 1409.82, Charge: +1

y_{12} y_{11} y_{10} y_9 y_8 y_7 y_6 y_5 y_4 y_3 y_2 y_1

Ala–Leu–Asn–Leu–Asn–Pro–Thr–Leu–Ala–Leu–Ile–Glu–Lys

b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8 b_9 b_{10} b_{11} b_{12}

Figure 2. Mass spectrometric analysis of the 18 kDa protein in crayfish. (A) Map of MS/MS. (B) Protein sequence alignments of MLCs from *Procambarus clarkii*, *Drosophila yakuba*, *Crangon crangon*, and *Marsupenaeus japonicus*.

pepsin digestion was analyzed with SDS−PAGE. The original band (18 kDa) was gradually digested over time. Proteolytic fragments smaller than 14.4 kDa appeared after prolonged peptic digestion. The original band was digested into fragments smaller than can be resolved by SDS−PAGE by pepsin within 1 h (Figure 3F). In contrast, a single proteolytic fragment smaller than 14 kDa was observed after the pancreatin digestion of MLC. The ~14.4 kDa fragment appeared after 15 min and remained undigested up to 240 min (Figure 3G). Competitive inhibition ELISA was used to analyze the IgE-binding activity of the MLC-derived digestion fragments. As shown in Figure 3H, the IgE-binding activity of the digested fragments was reduced by pancreatin digestion but not significantly with pepsin digestion. When the concentration of the inhibitor was 2.5 μg/mL, the inhibition rate of the pepsin digestion fragments was 80%, whereas that of the undigested MLC was 85%, and the inhibition

rate of the pancreatin digestion fragments was only 57%. The pepsin digestion fragments at a concentration of 0.095 μg/mL were capable of binding 50% of the IgE epitopes bound by undigested MLC at 0.075 μg/mL, whereas the pancreatic digestion fragments had a 50% inhibition rate at concentrations up to 0.224 μg/mL. This indicates that pepsin digestion had little effect on the IgE-binding activity of MLC. However, the IgE-binding activity of MLC that had been digested by pancreatin over a period of 4 h was reduced compared with that of the undigested MLC sample.

Secondary Structure of Purified MLC. Far-UV CD spectra were examined to confirm the predicted structure of purified MLC and to evaluate the effects of different physicochemical conditions on the protein's secondary structure. As shown in Figure 4, the control curves (0 °C and pH 7.5) displayed two minima, around 222 and 208 nm, and one maximum around 190
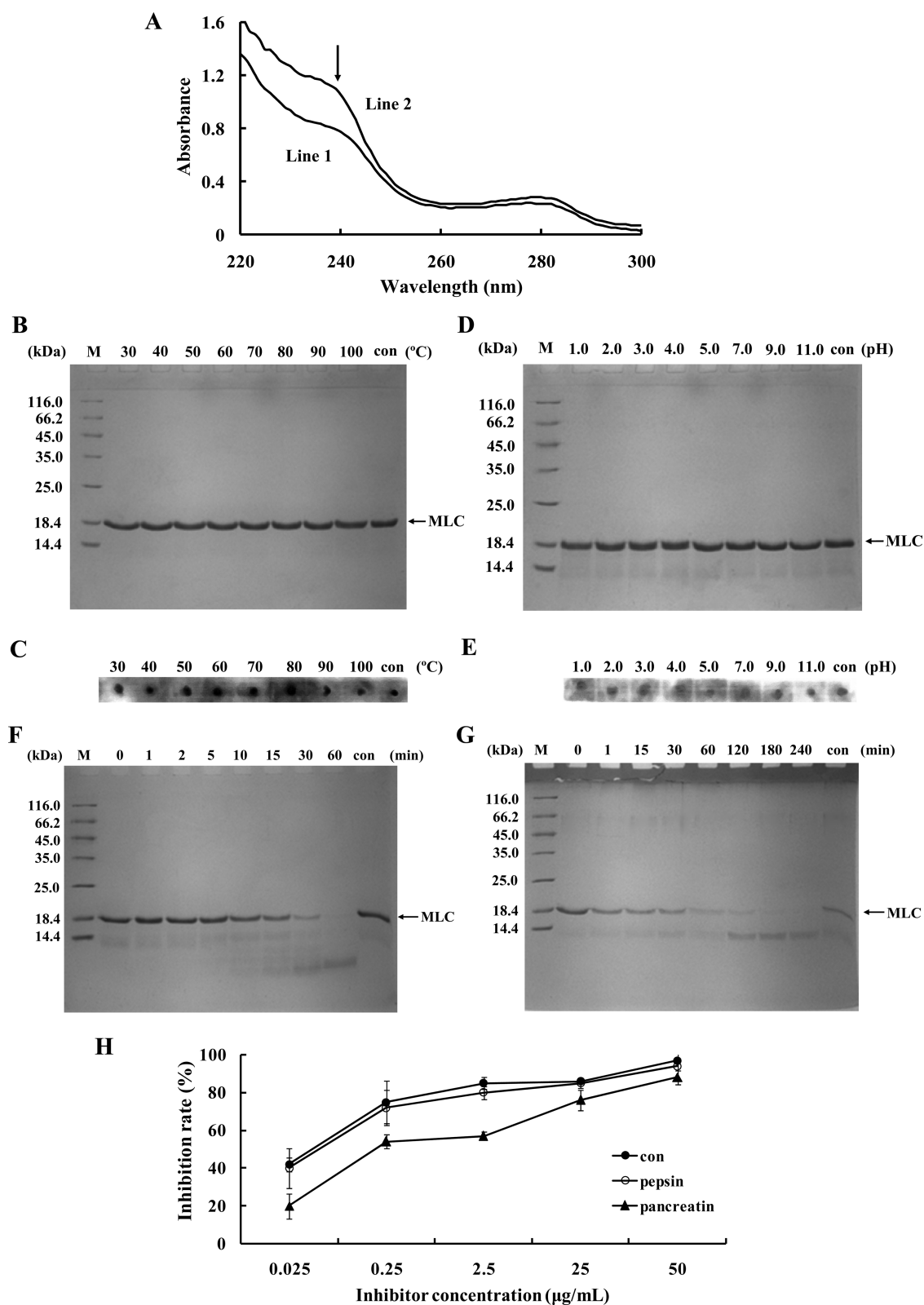
Figure 3. Physicochemical characterization of purified crayfish MLC. (A) UV spectral analysis of the purified MLC. Line 1: untreated MLC. Line 2: MLC treated with low-concentration alkali. The change at 240 nm is marked by an arrow. (B) SDS−PAGE analysis of the thermal stability of MLC. con: MLC was incubated at 0 °C. (C) Dot blot of thermally treated MLC. con: MLC was incubated at 0 °C. (D) SDS−PAGE analysis of the pH stability of MLC. con: MLC was incubated at pH 7.5. (E) Dot blot of MLC treated with different pH values. con: MLC was incubated at pH 7.5. (F) SDS−PAGE analysis of MLC stability to pepsin digestion. con: MLC was untreated with pepsin. (G) SDS−PAGE analysis of MLC stability to pancreatin digestion. con: MLC was untreated with pancreatin. (H) ciELISA of the digested fragments of MLC. MLC digested by pepsin (○, 1 h); pancreatin (▲, 4 h); and undigested (●).
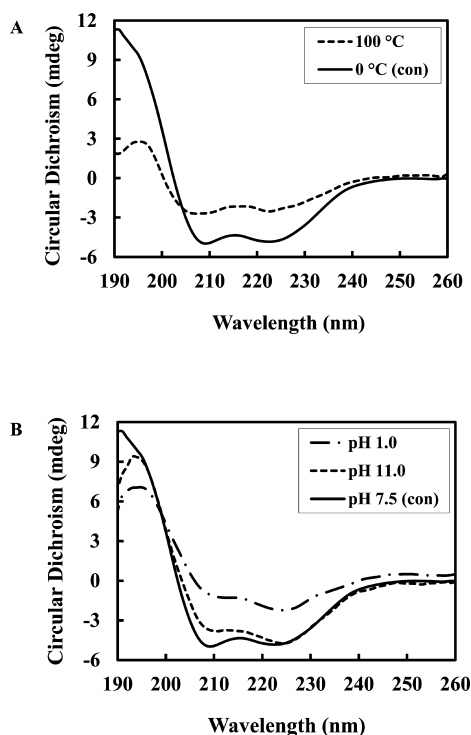
Figure 4. Effects of different experimental conditions on the stability of purified crayfish MLC1. (A) CD spectral analysis of MLC1 subjected to high temperature (100 °C, - - -) or 0 °C (—) as the control. (B) CD spectral analysis of MLC1 subjected to strong acid (pH 1.0, ·····); strong alkali (pH 11.0, - - -), or pH 7.5 (—) as the control.

nm, which are typical of α-helical secondary structures, although there was no fully canonical shapes that reflect the presence of other secondary structures, such as the extended strands, the β-turns, and/or the random coils. At 0 °C, the secondary structure was 33.5% α-helices, 16.2% extended strands, and 34.2% random coils. However, high-temperature treatments had a noticeable effect on the secondary structure, significantly increasing the extended strands (23.5%), which are characterized by a maximum at around 195 nm, and partially increasing the β-turns and/or the random coils (Figure 4A). Extended strands (25.6%) and random coils (39.3%) were also markedly increased after strong acid treatment, whereas strong alkali did not significantly increase the extended strands (19.2%) (Figure 4B). Among the different treatments tested, the secondary structure of purified crayfish MLC was most effectively altered by high temperature and strong acid.

Determination of the Predominant Epitope Type of Purified MLC. The IgE-binding activity of the purified crayfish MLC in folded form (with indirect ELISA and dot blot analysis, Figure 5A,B) was compared with MLC in its denatured, linear form (Western blot analysis, Figure 5C), with seven individual serum samples (nos. 749, 514, 974, 067, 785, 980, and 791) from crustacean-allergic patients and two sera from NA individuals that were pooled (nos. 780 and 981). With indirect ELISA (Figure 5A), the IgE from each of the seven sera bound to the native folded MLC, at least 2-fold higher than that of the negative control sera. In the dot blot analysis, all seven sera reacted with the native folded MLC, but no IgE-binding activity was detected in the negative control sera (Figure 5B). However, as shown in Figure 5C, only three sera (nos. 067, 785, and 791) reacted weakly to the denatured, linear MLC in the Western blot analysis.

This indicates that the native folded MLC has higher IgE-binding activity than the denatured linear MLC.

Cloning and Sequencing MLC1. To identify the isoform of the purified MLC and to predict of the MLC structures, the *MLC1* gene was cloned. As shown in Figure 6A, the full-length gene sequence of crayfish *MLC1* is 1096 bp, including the ATG start codon at nucleotides (nt) 63−65, the TAA stop codon at nt 523−525, and the poly-A tail. The open reading frame of crayfish *MLC1* is 462 bp, and the predicted protein consisted of 153 amino acid residues, with a predicted molecular weight of 17.59 kDa by DNASTAR EditSeq software, which is equivalent to the ∼18 kDa molecular weight estimated with SDS−PAGE (Figure 1B). The protein has a predicted isoelectric point of 4.58 by DNASTAR EditSeq software. There are two potential N-glycosylation sites in the amino acid sequence of crayfish MLC1 (marked with a box in Figure 6A), which were identified by the conserved motif Asn−X−Ser/Thr/Cys. A protein sequence alignment showed that the MLC proteins from different crustacean species share high homology (Figure 6B). The amino acid sequence of crayfish MLC1 was 92.81% identical to the amino acid sequence of *C. crangon* MLC (gi|238477331) and *M. japonicus* MLC (gi|290908988), 84.97% to *Palaemon varians* MLC (gi|238683675), 64.05% to *Periplaneta americana* MLC (gi|359326577), and 53.80% to *Artemia franciscana* MLC (gi|152013723).

Cloning and Sequencing MLC2. The *MLC2* gene was also cloned to distinguish between the isoforms of MLC. The full-length gene sequence of crayfish *MLC2* is 1362 bp, including the ATG start codon at nt 63−65, the TAA stop codon at nt 589−591, and the poly-A tail (Figure 6C). The open reading frame of crayfish *MLC2* is 528 bp, and the predicted protein consists of 175 amino acid residues, with a predicted molecular weight of 19.07 kDa and a predicted isoelectric point of 4.21 by DNASTAR EditSeq software. There is only one potential N-glycosylation site in the amino acid sequence of crayfish MLC2 (marked with a box in Figure 6C). A protein sequence alignment showed that the sequence homology of MLC1 and MLC2 is only 14.2%, whereas the MLC2 proteins of different crustacean species share relatively high homology (Figure 6D). The amino acid sequence of crayfish MLC2 is 62.92% identical to the amino acid sequence of *P. monodon* MLC2 (gi|357640499) and *L. vannamei* MLC (gi|184198734).

Structural Modeling and Epitope Prediction in Crayfish MLC. The three-dimensional structure of crayfish MLC1 was determined based on its amino acid sequence by a homology model (template: 4QBD, 56.95% identity, Figure 7A). Crayfish MLC1 is mainly composed of α-helices (54.25%) and random coils (22.22%), with a few β-turns or extended strands analyzed by SOPMA secondary structure prediction. This is roughly consistent with the results of the CD spectral analysis. The conformational epitopes were identified with the DiscoTope 2.0 server, and the most likely regions included 10 discontinuous amino acid residues (25, 57−60, 96−99, and 136). The 10 predicted discontinuous epitope residues were molecularly mapped onto the three-dimensional structure of crayfish MLC1 and formed four conformational epitopes (labeled with slate in Figure 7B). The linear (continuous) epitopes of crayfish MLC1 were predicted with the DNASTAR Protean software based on its secondary structure. Three continuous regions (amino acid residues 53−58, 76−83, and 95−99) of linear epitopes occurred in the most hydrophilic, flexible, and surface-exposed regions and were shown in blue on the model of crayfish MLC1 (Figure 7C). Crayfish MLC2 is also mainly composed of
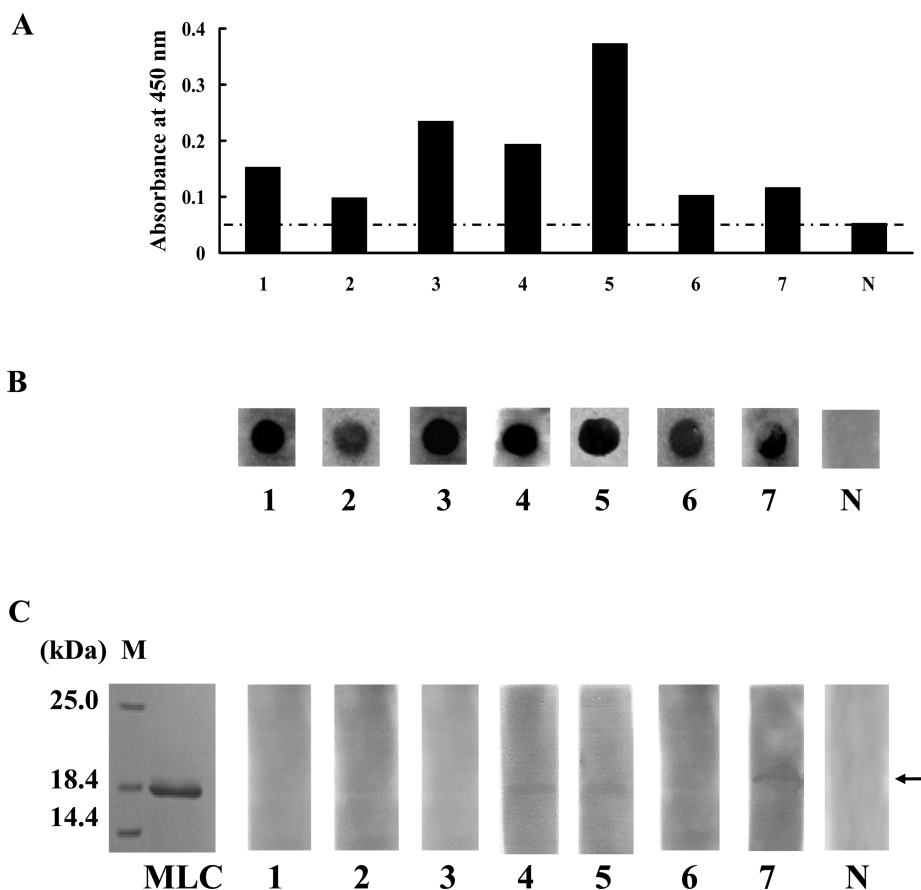
Figure 5. Comparison of IgE reactivity to native and denatured MLC from crayfish. 1−7: sera of seven crustacean-allergic patients (corresponding to serum nos. 749, 514, 974, 067, 785, 980, and 791); N: pooled sera from nonallergic individuals (nos. 780 and 981) as the negative control. (A) ELISA. (B) Dot blot. (C) Western blot. M, protein marker; MLC, SDS−PAGE analysis of the purified MLC.

α-helices (49.14%) and random coils (32.57%), with a few β-turns or extended strands. The most likely regions of conformational epitopes were 47 discontinuous residues (amino acid residues 31−32, 34−36, 38−39, 41−47, 49−71, 73−79, and 81−83), which composed a major conformational epitope. Three continuous regions (amino acid residues 5−11, 92−99, and 169−172) of linear epitopes were predicted (data not shown).

DISCUSSION

The growing international trade in crayfish has increased its popularity and the frequency of its consumption, which has been accompanied by an increase in the incidence of allergic diseases.[1,2] MLC was identified as a novel allergen in the white leg Pacific shrimp (*L. vannamei*),[18] the North Sea Shrimp (*C. crangon*),[10] and the black tiger prawn (*P. monodon*).[19] A previous study of crayfish TM demonstrated that an undefined myofibrillar protein, with a molecular mass about 18 kDa, had IgE-binding activity in the sera from crustacean-allergic patients.[9] It was inferred to be a new allergen in *P. clarkii*, although it is unclear whether this *P. clarkii* protein was MLC and, if so, which isoform it might be. This study was undertaken to determine whether this 18 kDa protein was MLC.

Few studies have purified native MLC for further analysis. In this study, the 18 kDa protein was purified by a combination of isoelectric precipitation, ammonium sulfate precipitation, ion exchange, and size exclusion chromatography and identified as MLC with MALDI-TOF-MS. The MLC content in muscle tissue is very low compared with the content of crayfish AK.[14] Only 11 sera from 137 sera collected from crayfish-allergic patients (which also had IgE-binding activity to other crustaceans) reacted with *P. clarkii* MLC in a serological assay. This indicates that MLC is a minor allergen in *P. clarkii* and may cross-react with the myofibrillar proteins of other crustaceans.

The molecular mass of crayfish MLC is 18 kDa, which is similar to that of *C. crangon* MLC, smaller than that of *L. vannamei* MLC (20 kDa), and larger than that of *P. monodon* MLC (16 kDa). It is interesting that the cloned sequence of the purified MLC was highly homologous to *C. crangon* MLC (Figure 6B), whereas *L. vannamei* MLC (gi|184198734) was only 13.11% identical and *P. monodon* MLC (gi|357640499) only 13.97% identical to the purified MLC, whereas the *L. vannamei* and *P. monodon* MLCs were 62.92% identical to MLC2 from *P. clarkii* (gi|392499121). This indicates that the purified MLCs from *P. clarkii* and MLC from *C. crangon* are MLC1, and that the MLCs from *L. vannamei* and *P. monodon* are MLC2. Because MLC1 is the essential light chain and MLC2 is the regulatory light chain, further studies should be performed to determine whether MLC2 in *P. clarkii* is also allergenic and to compare the allergenic properties of MLC1 and MLC2.

Most food allergens share similar physicochemical characteristics in that they are glycoproteins, relatively stable to heat and acid−alkali, and may remain allergenic after digestion.[14] In this study, the total carbohydrate content of MLC1 was about 4.3 ± 0.05% (w/w) of the whole protein. Two identical potential N-glycosylated sites (Asn−Gly−Thr) were identified by analyzing

A

```
1     ACGCGGGGACTTCAGGTCTGCCT TTTGCCCCGT TAACGTGGTGGTGCTAGAGGCTAAGTCA
61    AGATGGCCGCGGA TCTCAGTGCTCGTGATG TAGAGAGGGTGAAAT TCGCCT TCTCTATCT
1      M  A  A  D  L  S  A  R  D  V  E  R  V  K  F  A  F  S  I
121   ATGACTTCGAAGGCAACGGCACCATGGATGCCT TCTACA TT GGCGAT TGCCTTCGCGCCC
20     Y  D  F  E  G [N  G  T] M  D  A  F  Y  I  G  D  C  L  R  A
181   TGAACCTCAACCCCACCT TAGCTATCA T TGAGAAGGTCGGAGGCAAGGAGAAGAGGAAGG
40     L  N  L  N  P  T  L  A  I  I  E  K  V  G  G  K  E  K  R  K
241   AGAAGAT TATTAAGCTCGAGGAAT TCATACCCATCT TCGCT CAGG T TA AGAAGGACAAGG
60     E  K  I  I  K  L  E  E  F  I  P  I  F  A  Q  V  K  K  D  K
301   ACGCCGGCTCCTACGAGGACT TCATGGAAGTGCTGAAGCTCTACGAC AA ATCTGAA TG
80     D  A  G  S  Y  E  D  F  M  E  V  L  K  L  Y  D  K  S  E [N]
361   GCACTA TGATGTACGCCGAGCT TGAACACATCCTCCTG TCCCT TGGTGAGCGCCTCGAA AA
100   [G  T] M  M  Y  A  E  L  E  H  I  L  L  S  L  G  E  R  L  E
421   AGTCCGAATTGGAGCCAGTGCTG AAGGAGTGCTGCAACGAAGAGGATGAAGACGGC T TCA
120    K  S  E  L  E  P  V  L  K  E  C  C  N  E  E  D  E  D  G  F
481   T TCCCTACGAGCCT TTCCT TAAGAAGATGACTCAAC TACTCT A A ATACGTCCGCCA TCT T
140    I  P  Y  E  P  F  L  K  K  M  T  Q  L  L  *
541   CT T T TGTGACGTCATCACAGAGTGCGCATGCGTGCA ACACCACCGCCA T TCGGAGTGTCA
601   GTGCGCAGTGTACA ATCAT TCTGTGGGCTCGACCA A ACCACAGACAT T T AT TAGCTAAAT
661   TAT ACACATACACCA A AAGTATATAAAGAT TACCAA GATGAAAGTGAATA T AAGTGGAAA
721   AAT TGTTATA TT TGACA A TGTGAGCCACCCGTCACA TGCTCA TCACT TCCT GTCCGT ATG
781   TGTCT TGTGTCCGTGT AGCGGTGAAGGTCCCATGGAT CTGTGGA ACGTCTCCT T T TTCTG
841   TCTCA TCCACGCGGGGTGAAGGAGGGTCAAGGCAGGCGCGACTCGCGT TGCCT TGTAGA
901   CCTGAAGTGACCACT GA AT TGA TCCGCACAT AGT TGACCACCCGACATCACCTAGAGGTG
961   GGAGACTATAGACGCC AT TATC A AGACTCCCT CCCCGCA ATAGCCAGCCGGCCCAGCTGA
1021  CCATCGCCGCCATCT T GCGGGCCCCGCGCGCGCGCGGGAGGGACAAGAAAACTCAAAAAA
1081  AAAAAAAAAAAAAAAA
```

B

```
Procambarus clarkii      MAADLSARDVERVKFAFSIYDFEGNGTMDAFYIGDCLRALNLNPTLAIIEKVGGKEKRKE
Crangon crangon          MAADLSARDVERVKFAFSIYDFEGNGIDAFYIGDCLRALNLNPTLAVIEKVGGKEKRKE
Marsupenaeus japonicus   MAADLSARDVERVKFAFSIYDFEGNGTMDAYYIGDCLRALNLNPTLAVIEKVGGKEKRKE
Palaemon varians         MAADLSARDVERVKFAFSIYDFEGNGTMDAFYIGDCLRALNLNPTLAVIEKVGGKEKRKE
Periplaneta americana    MAGDLSARDIERANAFSIYDFEGTGTVDAIYLGDILRGLNLNPTQATVEKMGGTKKKNE
Artemia franciscana      MA............FSFTILMETRQLMAFYLGDCLRALNLNPTLAQIEKLGGTKKKNE

Procambarus clarkii      KIIKLEEFIPIFAQVKKDKDAGSYEDFMEVLKLYDKSENGTMMYAELEHILLSLGERLEK
Crangon crangon          KMIKLDEFMPLFAQVKKDKDAGSYEDFIEVLKLYDKAENGTMMYAELEHILLSLGERLDK
Marsupenaeus japonicus   KMIKLDEEMPIFAQVKKDKDASFEDFMEVLKLYDKAENGTMMYAELEHILLSLGERLEK
Palaemon varians         KMIKLDEFMPLFAQVKKDKDAGSYEDFIEVLKLYGKSENGTMMYAELEHIPSLGERLEK
Periplaneta americana    KKLKVEEFLPIFSGVKKEKDVKCYEDFLEGLKLYDKAEDGKNLLAELTHTLLSLGERLSD
Artemia franciscana      KKLKVEEFLPIFSQVKKEKDHGTLEDFIECMKLYDKSENGTMMLAELNHILLALGEKLED

Procambarus clarkii      SELEPVLKECCNEEDEDGFIPYEPFLKKMTQLL....
Crangon crangon          AELEPILRECGPPEDDEGLIPFEPFVKKLTQLL....
Marsupenaeus japonicus   AELEPVLKECCPEEDEEGCIPYEPFLKKITQLL....
Palaemon varians         AELEPILKECGPPEDDEEGCIPYEPFVKKLTTLL....
Periplaneta americana    REVDTIVKDCAEPEDEDGFIPYAQFLKNVMS......
Artemia franciscana      SQVDSIFKECMDPEDDEGNIPYTPELKRLLAGPYPDA
```

C

```
1     ACGCGGGAGAGGAGGTCGAGGAACCTGACTATGGGAGTGATAAGGTTGATGTATCGCG
61    ATCATGCCTGCCTCAAAGGGTTCCCGTTCTTCTTCCAAGAAGGCGAAGAAGGGAGGAAGC
1      M  P  A  S  K  G  S  R  S  S  S  K  K  A  A  K  K  G  G  S
121   AATGTCT TCGATATGTTCACCCAGAAGCAGGTGGCGGAGTTCAAGGAGGGCTTCCAGGTG
20     N  V  F  D  M  F  T  Q  K  Q  V  A  E  F  K  E  G  F  Q  V
181   ATGGACCGTGATCGTGATGGCATCATCAGCAAGGACGACCTCCGTGGTGTCTACGACGAA
40     M  D  R  D  R  D  G  I  I  S  K  D  D  L  R  G  V  Y  D  E
241   ATCGGCCGTATTGTCACCGACAAGGACCTGACGACA TGATTGCTGACGCACCTGGCCCA
60     I  G  R  I  V  T  D  K  D  L  D  D  M  I  A  D  A  P  G  P
301   ATTAACT TCACCACCCTCCCGCACA TGT TCGCCTCACGCTCCTCTGGAGAGTCCGACGAC
80     I [N  F  T] T  L  P  H  M  F  A  S  R  S  S  G  E  S  D  D
361   GACGATGTTGTTGCCAAGTCTTTCCGT GCCTTCGAGAAGGAAGCCGGTCAGATCGACTCC
100    D  D  V  V  A  K  S  F  R  A  F  E  K  E  A  G  Q  I  D  S
421   GAGCAGTTCCGTGCAATGCTT ATGGCTTTCGGTGACAAGTTCACGCCCGCCGGAAGTCGAC
120    E  Q  F  R  A  M  L  M  A  F  G  D  K  F  T  P  A  E  V  D
481   GATGCCTTTGAGCAGATGGAAATCGACGAGGACACCGGCA TCATCGACTCCAACGCCCTC
140    D  A  F  E  Q  M  E  I  D  E  D  T  G  I  I  D  S  N  A  L
541   ATCGCCAT GTTGTGCGCCCGGTGGAGCAGAAAAGGGTGAGGGTGAGGCTTAAATACCCAAA
160    I  A  M  L  C  A  G  G  A  E  K  G  E  G  E  A  *
601   TTCCCGTCCATCGTCCGTGCCCGAAGGAAATGTGGGTGAGCAGTAGCGGAAGCCATCTGC
661   AGGGGTTGCTACCTGTAGCTGT GACACTCAGGCGCTTC TTGGCTAT CTTCTGGCCCCTGC
721   TGCTGCT ACTTCCTCTCTCA TCAATGCTCGCCACT TCTT CCA TACG TCACCACCACCCA T
781   ATT TCT TACGGGCGAAGTCACAAGCAGGATAGACAATTGAGGCTGGACCCTTCGCCCTCT
841   CCCTCTTCTGT AATCCTGCCGTCA TTATC TCCATCTACGATA GACCCTACAAGCGGCGGC
901   TTTCACCAGACCTCAAGATCGGCGTCCGCGAACTCGCAGTGTCTTTGTGAAAAGAAAAAA
961   GCCAAGCCGGGAGCGGAGTGACAGTGACAGTGAATATGAATGTAAACATTGCCGCCGTGA
1021  TCCAACAACCAGAGGGCCTACCGCTCCC TTGCTAGTGTTACAT CCGAACAGAACT AAAACG
1081  AGACAT CTGTGAC TAAAGTAAAGTACGAGGTCGTAGTGTGAGGGAGCTCAGTGTCATCTG
1141  CCTCAT CTTTAAACCAGTGTAGTGACGACTTTTACGAGCT GGAGAAGCCACCGCGGTGAC
1201  CGCTACGCAGTTGTCGTCACTGATGCTGGACCAGAGTCGGGTGCCGGTGTACTGCAGCTC
1261  TAGCACTTGT AAATGTTTT CAGTTGTGAAGCA TTGTAT TAACA TCGGCAAGATA TA TTCAA
1321  TAAAAT GTTTT ACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

D

```
Procambarus clarkii    MPASKGSRSSSKKAKK..GGSNVFDMFTQKQVAEFKEGFQVMDRDRDGIISKDDLRGVYD   58
Litopenaeus vannamei   MSRSKGSRSSSKRSKKSGGGSNVFDMFTQRQVAEFKEGFQLMDRDKDGVIGKTDLRGTFD   60
Penaeus monodon        MSRSKGSRSSSKRSKKSGGGSNVFDMFTQRQVAEFKEGFQLMDRDKDGVIGKTDLRGTFD   60

Procambarus clarkii    EIGRIVTDKDLDDMIADAPGPINFTTLPHMFASRSSGESDDDDVVAKSFRAFEKEAGQID   118
Litopenaeus vannamei   EIGRIATDQELDEMLADAPAPINFTMLLNMFAERQTGESDDDDVVAKAFLAFADEEGNID   120
Penaeus monodon        EIGRIATDQELDEMLADAPAPINFTMLLNMFAERQTGESDDDDVVAKAFLAFADEGGNID   120

Procambarus clarkii    SEQFRAMLMAFGDKFTPAEVDDAFEQMEIDEDTGIIDSNALIAMLCAGGAEKGEG.EA   175
Litopenaeus vannamei   CDTFRHALMTWGDKFSSQEADDALDQMDIDDG.GKIDVQGVIQMLTAGGGDDAAAEEA   177
Penaeus monodon        CDTFRHALMTWGDKFSSQEADDALDQMDIDDA.GKIDVQSVIQMLTAGGGDDAAAEEA   177
```

Figure 6. Cloning *MLC* genes and sequence alignment of crayfish MLC proteins. (A) Nucleotide and deduced amino acid sequences of crayfish MLC1 (the N-glycosylation site is marked with a box). (B) Sequence alignment of crayfish MLC1 and the MLC proteins of other crustaceans. (C) Nucleotide and deduced amino acid sequences of crayfish MLC2 (the N-glycosylation site is marked with a box). (D) Sequence alignment of crayfish MLC2 and the MLC proteins of other crustaceans.

the primary structure and mass spectra of the protein. The carbohydrates in crayfish SCP are reported to constitute about 4.9% of the protein, with one N-glycosylation site. MLC and SCP have reportedly similar characteristics, and data in this report support this conclusion.[6,18] An O-linked glycan in the MLC1 structure was detected with β-elimination. The O-glycan, which occurs in clusters along short regions of the peptide chain that contain repeating units of Ser and Thr, is unstable under alkaline conditions at low temperature. After NaOH treatment, Ser and Thr are transformed into α-amino acids. Both of these unsaturated amino acids display characteristic UV absorption at 240 nm.[6] According to the typical characterization of glycoproteins, MLC1 is a glycoprotein. N- and O-linked allergenic glycoproteins, which have common core structures, are associated with the induction of an IgE response in allergic patients.[30] The N-glycosylated sites of MLC1 were partially at the predicted epitope regions (amino acid residues 25 and 99). Therefore, the relationship between the allergenicity and glycosylation of crayfish MLC1 warrants further research.

In this study, MLC1 was stable at 30−100 °C and under strongly acidic and alkaline conditions and retained its IgE-binding activity at different temperatures and pH value. In contrast, a study by Kamath et al.[19] demonstrated that IgE binding to MLC of *P. monodon* was only detected in heated prawn extract and not in raw prawn extract. It was inferred that

the thermal stability of MLC1 differs from that of MLC2. The thermal stability and pH stability of crayfish MLC1 were similar to those of TM, another myofibrillar protein and the major allergen of shellfish, which is stable under acidic and alkaline conditions and retains its IgE-binding activity after high-temperature and high-pressure processing.[31] However, the IgE-binding activities of crayfish myosinogens AK and SCP are reduced at high temperatures or under different acidic or alkaline conditions.[6,14] It has been reported that, under different physiological and experimental conditions, some allergens may undergo conformational changes that cause the exposure of different combinations of epitopes recognized by the IgE antibodies of allergic subjects.[18] In this study, the changes in the secondary structure of crayfish MLC1 under severe conditions, including high temperature (100 °C), strong acid (pH 1.0), and strong alkali (pH 11.0), were assessed to observe the changes in its secondary structure. High temperature and strong acid treatments had significant effects on the secondary structure of MLC1, but its IgE-binding activity remained intact despite the changes in its structure. A future study should focus on effective processing methods to mask or disrupt the IgE epitopes. Food allergens generally remain intact during gastro-intestinal digestion. Crayfish MLC1 appeared to be more resistant to gastric digestion than to intestinal digestion and retained its immunological activity (detected with ciELISA).
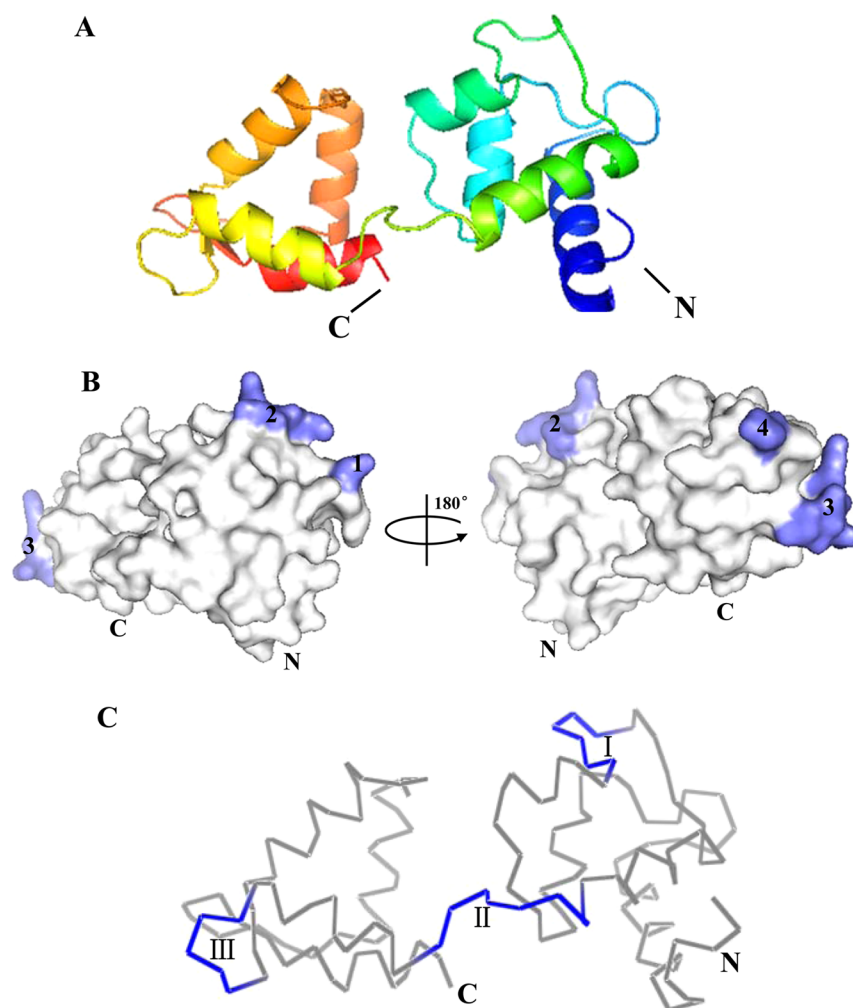
Figure 7. Structural modeling and epitope prediction of crayfish MLC1. (A) Three-dimensional structure of crayfish MLC1. N and C indicate the N- and C-termini, respectively. (B) Molecular mapping of the conformational epitopes of crayfish MLC1 onto its three-dimensional structure. The areas occupied by the conformational epitopes are shown in slate. Epitopes 1–4 are Asn^{25}, Lys^{57}–Glu^{60}, Lys^{96}–Asn^{99}, and Glu^{136}, respectively. (C) Molecular mapping of the linear epitopes of crayfish MLC1 onto its three-dimensional structure. The areas occupied by the linear epitopes are shown in blue. Epitopes I–III, Gly^{53}–Arg^{58}, Lys^{76}–Ser^{83}, and Asp^{95}–Asn^{99}, respectively.

Interestingly, MLC1 was degraded after gastrointestinal digestion, but the degraded fragments had IgE-binding activity almost equivalent to that of the undigested MLC1 sample. In a previous study, crayfish AK and SCP were also degraded to varying degrees by gastrointestinal digestion but remained immunologically active.[6,14] Certain IgE-binding sites on crayfish MLC1 are thought to be left intact by digestive enzymes, and the major IgE-binding sites of allergens in crayfish are stable after digestion. It was reported that Ca^{2+} binding was essential for the formation of the conformation structure of calcium-binding proteins like parvalbumin, and the depletion of Ca^{2+} resulted in the reduction of IgE-binding activity.[32,33] MLC belongs to the EF-hand superfamily, and the influences of Ca^{2+} on the structure of MLC and its allergenicity need further research.

The epitopes can be classified as continuous or discontinuous. Continuous epitopes can be identified in denatured proteins and peptide fragments, whereas discontinuous epitopes depend on the higher-order structures of the native protein.[34] The study of epitopes is critical to understanding the nature of allergens and the improvement of diagnosis, detection, and therapies. To determine the predominant IgE epitope type in crayfish MLC1, indirect ELISA and dot blot analysis were performed on folded MLC1 and Western blot analysis was performed on the unfolded MLC1 with sera from allergic individuals. In both the indirect ELISA and the dot blot analysis, the native MLC1 was immobilized on the plate or a nitrocellulose membrane, respectively. In both of these assays, the folded MLC1 was found to maintain IgE binding under high heat and different pH conditions for all allergic sera tested. In the Western blot, where MLC1 was in denatured or linear form, only the serum IgE from three allergic individuals displayed weakly positive binding. These results imply that the conformational epitopes of MLC1 are predominantly recognized by IgE, rather than the linear epitopes. Considering that the IgE binding by unfolded MLC1 is significantly reduced compared to the folded MLC1, it is possible that the digested fragments retain some structural features that are able to compete with intact MLC1 for IgE binding.

The conformational and linear epitopes of the crayfish MLCs were also predicted in silico. MLC1 was predicted to contain four conformation epitopes and three linear epitopes, and the regions formed by amino acids 57–60 and 96–99 contained both conformational and linear epitopes. MLC2 also showed a major predicted conformational epitope and three linear epitopes. MLC2 clearly was predicted to have more epitopes than MLC1,

but it was not recognized by the IgE in sera from crustacean-allergic patients in a previous study. However, in future studies, it will be important to investigate IgE binding to MLC2 with the same patient sera used in this study for MLC1. Fundamental studies of allergens focus on IgE-binding epitopes, which are used as biomarkers to predict the clinical reactivity of subjects who are sensitized to a specific food, and provide alternative immunotherapeutic strategies.[35] If some parallels can be found between predicted and the experimental IgE-binding patterns, perhaps efforts to reduce the allergenicity of food and diagnostic tools may be improved via epitope prediction tools. Also, recombinant MLC1 and MLC 2 can be used to aid in clinical diagnoses and in further research to provide basic information for the hypersensitization to crayfish and the evaluation of food allergies.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Additional experimental details. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.5b01318.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +86-592-6183383. Fax: +86-592-6180470. E-mail: gmliu@jmu.edu.cn.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Lee, A. J.; Gerez, I.; Shek, L. P.-C.; Lee, B. W. Shellfish allergy—an Asia-Pacific perspective. *Asian Pac. J. Allergy Immunol.* **2012**, *30*, 3−10.

(2) Pedrosa, M.; Boyano-Martínez, T.; García-Ara, C.; Quirce, S. Shellfish Allergy: a Comprehensive Review. *Clin. Rev. Allergy Immunol.* **2014**, DOI: 10.1007/s12016-014-8429-8.

(3) Mona, M. H.; Geasa, N.; Sharshar, K.; Morsy, E. Chemical composition of freshwater crayfish (*Procambarus clarkii*) and its nutritive value. *Egyptian J. Aquatic Biol. Fisheries* **2000**, *4*, 19−34.

(4) Fahmy, S. R.; Hamdi, S. A.; Abdel-Salam, H. A. Curative effect of dietary freshwater and marine crustacean extracts on carbon tetrachloride-induced nephrotoxicity. *Aust. J. Basic Applied Sci.* **2009**, *3*, 2118−2129.

(5) *China Fisheries Yearbook*; China Agricultural Press: Beijing, China, 2014 (in Chinese).

(6) Chen, H. L.; Cao, M. J.; Cai, Q. F.; Su, W. J.; Mao, H. Y.; Liu, G. M. Purification and characterisation of sarcoplasmic calcium-binding protein, a novel allergen of red swamp crayfish (*Procambarus clarkii*). *Food Chem.* **2013**, *139*, 213−223.

(7) Leung, P. S.; Chen, Y. C.; Mykles, D. L.; Chow, W. K.; Li, C. P.; Chu, K. H. Molecular identification of the lobster muscle protein tropomyosin as a seafood allergen. *Mol. Mar. Biol. Biotechnol.* **1998**, *7*, 12−20.

(8) Reese, G.; Ayuso, R.; Lehrer, S. B. Tropomyosin: an invertebrate pan-allergen. *Int. Arch. Allergy Immunol.* **1999**, *119*, 247−258.

(9) Lin, J. W.; You, H. Y.; Shen, H. W.; Cao, M. J.; Cai, Q. F.; Liu, G. M. Purification and Allergenic Analysis of Procambarus clarkii Tropomyosin. *J. Jimei Uni. (Natural Sci.)* **2012**, *17*, 167−174.

(10) Bauermeister, K.; Wangorsch, A.; Garoffo, L. P.; Reuter, A.; Conti, A.; Taylor, S. L.; Lidholm, J.; DeWitt, Å. M.; Enrique, E.; Vieths, S. Generation of a comprehensive panel of crustacean allergens from the North Sea Shrimp Crangon crangon. *Mol. Immunol.* **2011**, *48*, 1983−1992.

(11) Abdel Rahman, A. M.; Kamath, S. D.; Gagné, S. b.; Lopata, A. L.; Helleur, R. Comprehensive proteomics approach in characterizing and quantifying allergenic proteins from northern shrimp: toward better occupational asthma prevention. *J. Proteome Res.* **2013**, *12*, 647−656.

(12) Rosmilah, M.; Shahnaz, M.; Zailatul, H.; Noormalin, A.; Normilah, I. Identification of tropomyosin and arginine kinase as major allergens of *Portunus pelagicus* (blue swimming crab). *Trop. Biomed.* **2012**, *29*, 467−478.

(13) Shen, H. W.; Cao, M. J.; Cai, Q. F.; Ruan, M. M.; Mao, H. Y.; Su, W. J.; Liu, G. M. Purification, cloning, and immunological character-ization of arginine kinase, a novel allergen of *Octopus fangsiao*. *J. Agric. Food Chem.* **2012**, *60*, 2190−2199.

(14) Chen, H. L.; Mao, H. Y.; Cao, M. J.; Cai, Q. F.; Su, W. J.; Zhang, Y. X.; Liu, G. M. Purification, physicochemical and immunological characterization of arginine kinase, an allergen of crayfish (*Procambarus clarkii*). *Food Chem. Toxicol.* **2013**, *62*, 475−484.

(15) Shiomi, K.; Sato, Y.; Hamamoto, S.; Mita, H.; Shimakura, K. Sarcoplasmic calcium-binding protein: identification as a new allergen of the black tiger shrimp Penaeus monodon. *Int. Arch. Allergy Immunol.* **2008**, *146*, 91−98.

(16) Ayuso, R.; Grishina, G.; Ibáñez, M. D.; Blanco, C.; Carrillo, T.; Bencharitiwong, R.; Sánchez, S.; Nowak-Wegrzyn, A.; Sampson, H. A. Sarcoplasmic calcium-binding protein is an EF-hand-type protein identified as a new shrimp allergen. *J. Allergy Clin. Immunol.* **2009**, *124*, 114−120.

(17) Piboonpocanun, S.; Jirapongsananuruk, O.; Tipayanon, T.; Boonchoo, S.; Goodman, R. E. Identification of hemocyanin as a novel non-cross-reactive allergen from the giant freshwater shrimp *Macro-brachium rosenbergii*. *Mol. Nutr. Food Res.* **2011**, *55*, 1492−1498.

(18) Ayuso, R.; Grishina, G.; Bardina, L.; Carrillo, T.; Blanco, C.; Ibáñez, M. D.; Sampson, H. A.; Beyer, K. Myosin light chain is a novel shrimp allergen, Lit v 3. *J. Allergy Clin. Immunol.* **2008**, *122*, 795−802.

(19) Kamath, S. D.; Rahman, A. M. A.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. Effect of heat processing on antibody reactivity to allergen variants and fragments of black tiger prawn: a comprehensive allergenomic approach. *Mol. Nutr. Food Res.* **2014**, *58*, 1144−1155.

(20) Funkenstein, B.; Skopal, T.; Rapoport, B.; Rebhan, Y.; Du, S.; Radaelli, G. Characterization and functional analysis of the 5′ flanking region of myosin light chain-2 gene expressed in white muscle of the gilthead sea bream (Sparus aurata). *Comp. Biochem. Physiol., Part D: Genomics Proteomics* **2007**, *2*, 187−199.

(21) Han, F.; Wang, Z. Y.; Wang, X. Q. Characterization of myosin light chain in shrimp hemocytic phagocytosis. *Fish Shellfish Immunol.* **2010**, *29*, 875−883.

(22) Liu, J.; Han, L. N.; Zhang, Q.; Wang, Q. L.; Chang, Q.; Zhuang, H.; Liu, J.; Li, M.; Yu, D.; Kang, Z. S. Cloning and molecular characterization of a myosin light chain gene from Puccinia striiformis f. sp. tritici. *World J. Microbiol. Biotechnol.* **2014**, *30*, 631−637.

(23) Balhara, J.; Redhu, N. S.; Shan, L.; Gounni, A. S. IgE Regulates the Expression of smMLCK in Human Airway Smooth Muscle Cells. *PLoS One* **2014**, *9*, e93946.

(24) Huang, M. C.; Ochiai, Y. Fish fast skeletal muscle tropomyosins show species-specific thermal stability. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **2005**, *141*, 461−471.

(25) Yan, L. J.; Zhan, C. L.; Cai, Q. F.; Weng, L.; Du, C. H.; Liu, G. M.; Su, W. J.; Cao, M. J. Purification, characterization, cDNA cloning and in vitro expression of a serine proteinase from the intestinal tract of sea cucumber (*Stichopus japonicus*) with collagen degradation activity. *J. Agric. Food Chem.* **2014**, *62*, 4769−4777.

(26) Dubois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P.; Smith, F. Colorimetric method for determination of sugars and related substances. *Anal. Chem.* **1956**, *28*, 350−356.

(27) Li, Z.; You, J.; Luo, Y. K.; Wu, J. P. Purification and characterization of parvalbumin isotypes from grass carp (*Ctenophar-yngodon idella*). *J. Agric. Food Chem.* **2014**, *62*, 6212−6218.

(28) Böhm, G.; Muhr, R.; Jaenicke, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng., Des. Sel.* **1992**, *5*, 191−195.

(29) Guo, C.; Cao, M. J.; Liu, G. M.; Lin, X. S.; Hara, K.; Su, W. J. Purification, characterization, and cDNA cloning of a myofibril-bound serine proteinase from the skeletal muscle of crucian carp (*Carassius auratus*). *J. Agric. Food Chem.* **2007**, *55*, 1510−1516.

(30) Fötisch, K.; Vieths, S. N-and O-linked oligosaccharides of allergenic glycoproteins. *Glycoconjugate J.* **2001**, *18*, 373−390.

(31) Yu, H. L.; Cao, M. J.; Cai, Q. F.; Weng, W. Y.; Su, W. J.; Liu, G. M. Effects of different processing methods on digestibility of *Scylla paramamosain* allergen (tropomyosin). *Food Chem. Toxicol.* **2011**, *49*, 791−798.

(32) Moraes, A.; Ackerbauer, D.; Kostadinova, M.; Bublin, M.; Oliveira, G. A.; Ferreira, F.; Almeida, F.; Breiteneder, H.; Valente1, A. P. Solution and high-pressure NMR studies of the structure, dynamics, and stability of the cross-reactive allergenic cod parvalbumin Gad m 1. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 3032−3042.

(33) Swoboda, I.; Bugasjska-Schretter, A.; Linhart, B.; Verdino, P.; Keller, W.; Schulmeister, U.; Sperr, W. R.; Valent, P.; Peltre, G.; Quirce, S.; Douladiris, N.; Papadopoulos, N. G.; Valenta, R.; Spitzauer, S. A recombinant hypoallergenic parvalbumin mutant for immunotherapy of IgE-mediated fish allergy. *J. Immunol.* **2007**, *178*, 6290−6296.

(34) Barlow, D.; Edwards, M.; Thornton, J. Continuous and discontinuous protein antigenic determinants. *Nature* **1986**, *322*, 747−748.

(35) Mao, H. Y.; Cao, M. J.; Maleki, S. J.; Cai, Q. F.; Su, W. J.; Yang, Y.; Liu, G. M. Structural characterization and IgE epitope analysis of arginine kinase from *Scylla paramamosain*. *Mol. Immunol.* **2013**, *56*, 463−470.

Supporting Information

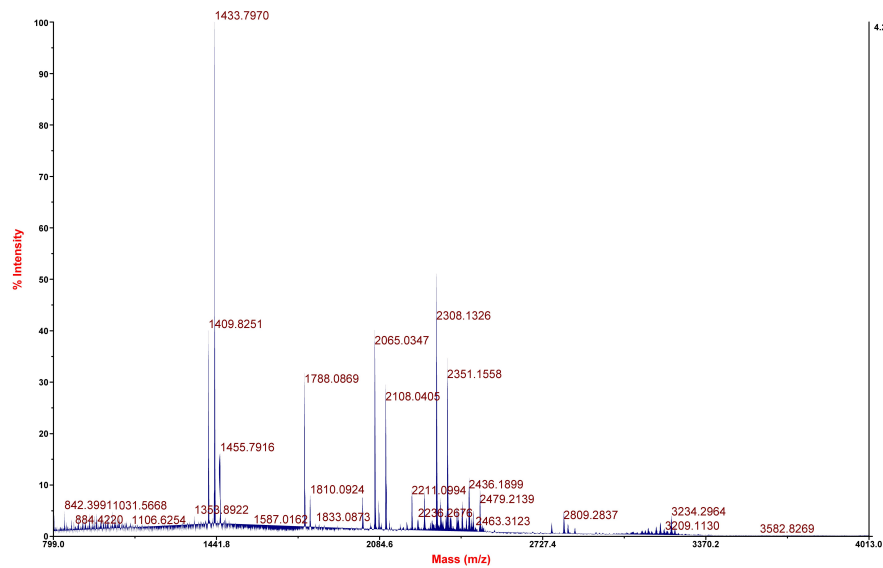**Figure 1.** Map of mass spectrometry of the 18 kDa protein in crayfish.

Table 1. Primers Used in Cloning *MLC1* from *Procambarus clarkii*.

Name	Sequence (5'-3')
MLC1-R1	ATCTTCTTGAGGAAAGGTTCGTAGG
MLC1-R2	GTTCGTAGGGAATGCAACCTTCTTC
MLC1-F1	GGATCTCAGTGCTCGTGATGTCG
MLC1-F2	TTCGCCTTCTCCATCTATGACTT
MLC1-3R1	GCTATCATTGAGAAGGTCGGAGG
MLC1-3R2	TCAGGTTAAGAAGGACAAGGACG
MLC1-5F1	TCAAGCTCGGCGTACATCATAGT
MLC1-5F2	GGCGTCCTTGTCCTTCTTAACCT

Table 2. Primers Used in Cloning *MLC2* from *Procambarus clarkii*.

Name	Sequence (5'-3')
MLC2-R1	ATGTCCATCTGGTCGAGGGCGTC
MLC2-R2	TCCTCTTGAACAGCGGGGTCCAGTA
MLC2-F	GAGTTCAAGGAGGGCTTCCAGCTGA
MLC2-3R1	CCGACAAGGACCTGGACGACA
MLC2-3R2	GATGTTGTTGCCAAGTCTTTCCG
MLC2-5F1	GGAAAGACTTGGCAACAACATCG
MLC2-5F2	AGCAATCATGTCGTCCAGGTCC

The Principles of Structural Modeling and Epitope Prediction.

The structural modeling was analyzed by homology modeling using SWISS-MODEL. Submitting the amino acid sequence of the target protein to the workspace, a library of experimental protein structures was searched to identify suitable templates. On the basis of a sequence alignment between the target protein and the template structure, a three-dimensional model for the target protein is generated.

The linear epitopes prediction was analyzed by four parameters: hydrophilicity plot predicted using the method of Kyte–Doolittle, flexible regions predicted using the method of Karplus–Schulz, antigenic index predicted using the method of Jameson–Wolf, surface probability plot predicted using the method of Emini, and then select the peptide fragments which were satisfied the four conditions simultaneously.

The conformational epitopes prediction was analyzed by DiscoTope server. The method utilizes calculation of surface accessibility (estimated in terms of contact numbers) and a novel epitope propensity amino acid score. The final scores are calculated by combining the propensity scores of residues in spatial proximity and the contact numbers.