Identification and Quantification of Allergenic Tropomyosin from Shellfish

Thesis submitted by

Martina Koeberl

(Magistra rer. nat.)

For the degree of Doctor of Philosophy

College of Public Health, Medical and Veterinary Sciences

James Cook University

Townsville, Australia

July 2015

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Andreas for giving me the opportunity to join James Cook University to obtain a doctoral degree. Thank you Andreas for guidance, support and help throughout the whole project as well as sharing some privacy to make everyone welcome in Townsville. Moreover, to present me such a new and interesting research project and make my project reality with the help from NMI and JCU.

Secondly, I would like to thank the National Measurement Institute. Especially I would like to thank James from NMI for his personal interest in allergy research and pushing the boundaries to apply new techniques for allergen analytics and therefore organising funding and support to realise my PhD research project. Moreover, I would like to thank Saman from NMI, who always coped with a smile to the instrumental difficulties during my PhD. I also would like to thank Dean for his support, Louise for organising time on the mass spectrometers and Jenny for the support with finances.

Thirdly, I would like to thank Anas, my co-supervisor, for the support and guidance during my experiments and providing his professional advice. Moreover, I would like to thank Waters Cooperation, especially Chris, for helping with the qTOF and the thousand phone calls to solve instrumental difficulties. Furthermore, I would like to thank the team of the Molecular Immunology Group at JCU: Sandip, Michael, Shanshan, Nik, Susan, Thimo, Roni and Fiona - thanks for your support, help and friendship.

Last, but definitely not least, I would thank the most important people in my life - my parents, my family and my friends. These are the people who supported me all my life and made me who I am today.

STATEMENT AND CONTRIBUTION OF OTHERS

Statement and contribution of others in the thesis		
Nature of Assistance	**Contribution**	**Name and Affiliation**
Intellectual support	Project plan and development	• A/Prof. Andreas Lopata, James Cook University • James Roberts, National Measurement Institute • Dr. Saman Buddhadasa, National Measurement Institute
	Editorial Assistance	• A/Prof. Andreas Lopata, James Cook University
Financial support	Research	• ARC Future Fellowship (A/Prof. Andreas Lopata) • National Measurement Institute • NHMRC Project grant (A/Prof. Andreas Lopata)
	Stipend	• National Measurement Institute
Data collection	Blood sample collection	• Prof. Jennifer Rolland, Monash University • Prof. Robyn O'Hehir, Monash University and The Alfred Hospital
Technical support	Recombinant protein expression	• Dr. Michael Smout, James Cook University
	Mass spectrometry	• Dr. Anas Abdel Rahman • Chris Buck, Waters Cooperation • Pat Sacchetta, Waters Cooperation

Statement and contribution of others in the chapters	
Chapter 1: Introduction to mass spectrometry for allergen detection	
Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems. **Journal of Proteome Research.** 13, (8), 3499-3509.	
Dean Clarke (DC), National Measurement Institute A/Prof. Andreas Lopata (AL), James Cook University	MK and AL co-developed the aim of the review article. The figures and tables were generated by MK with editorial inputs from AL. Editorial inputs of the manuscript were obtained from AL and DC.
Chapter 2: Molecular and immunological analysis of tropomyosin from King prawn (*Melicertus latisulcatus*)	
Koeberl, M.; Kamath, S. D.; Saptarshi, S. R.; Smout, M. J.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). Auto-Induction for High Yield Expression of Recombinant Novel Isoallergen Tropomyosin from King Prawn (*Melicertus Latisulcatus*) for Improved Diagnostics and Immunotherapeutics. **Journal of Immunological Methods.** 415, (0), 6-16.	
Dr. Sandip Kamath (SK), James Cook University Dr. Shruti Saptarshi (SS), James Cook University Dr. Michael Smout (MS), James Cook University Prof. Jennifer Rolland (JR), Monash University and The Alfred Hospital, Australia Prof. Robyn O'Hehir (ROH), Monash University and The Alfred Hospital, Australia A/Prof. Andreas Lopata (AL), James Cook University	MK and AL co-developed the research question. MK performed the experimental work and data analysis. SK provided recombinant tropomyosin from Black Tiger prawn, primers for tropomyosin and myosin light chain, the structural modelling of tropomyosins and myosin light chains. SS performed the CD spectroscopic analysis. MS provided knowledge about the auto-induction system. JR and ROH provided the patient sera. The figures and tables were generated by MK with editorial input from AL. Editorial inputs of the manuscript were obtained from AL, SK, MS, JR and ROH.
Chapter 3: Identifying signature peptides for allergenic tropomyosin to distinguish crustacean from molluscs	
Dr. Anas Abdel Rahman (ARA), King Faisal Specialist Hospital and Research Center (KFSHRC) A/Prof. Andreas Lopata (AL), James Cook University	MK and AL co-developed the research question. MK performed the experimental work and data analysis. ARA provided assistance with the selection of signature peptides. The figures and tables were generated by MK with editorial input from AL.
Chapter 4: Identifying possible allergen-candidates in crustacean and molluscs	
A/Prof. Andreas Lopata (AL), James Cook University	MK and AL co-developed the research question. MK performed the experimental work and data analysis. The figures and tables were generated by MK with editorial input from AL.
Chapter 5: Quantification of allergenic tropomyosin applying signature peptides	
Dr. Anas Abdel Rahman (ARA), King Faisal Specialist Hospital and Research Center (KFSHRC) A/Prof. Andreas Lopata (AL), James Cook University	MK and AL co-developed the research question. MK performed the experimental work and data analysis. ARA provided assistance with the validation of the LC/MRM method. The figures and tables were generated by MK with editorial input from AL.
Chapter 6: Analysis of food samples for allergenic tropomyosin	
A/Prof. Andreas Lopata (AL), James Cook University	MK and AL co-developed the research question. MK performed the experimental work and data analysis. The figures and tables were generated by MK with editorial input from AL.

PUBLICATIONS ARISING FROM THESIS

At the time of this thesis submission, two journal manuscripts describing the research findings from Chapters 1 and 2 were published.

- **Koeberl, M**.; Clarke, D.; Lopata, A. L. (2014). Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems. Journal of Proteome Research. 13, (8), 3499-3509. (Chapter 1)

- **Koeberl, M.;** Kamath, S. D.; Saptarshi, S. R.; Smout, M. J.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). Auto-Induction for High Yield Expression of Recombinant Novel Isoallergen Tropomyosin from King Prawn (*Melicertus Latisulcatus*) for Improved Diagnostics and Immunotherapeutics. Journal of Immunological Methods. 415, (0), 6-16. (Chapter 2)

ABSTRACT

Shellfish allergy belongs to the “Big 8” food allergies and can cause anaphylaxis in allergenic individuals. In the gastronomy setting, shellfish includes two groups, namely crustacean and molluscs. Both groups include large numbers of different species, whereas all species have tropomyosin as one of the major allergens. However, the amino acid sequence of tropomyosin in different shellfish species varies greatly. The amino acid homology within crustacean is very high, whereas the amino acid homology within molluscs is much lower, followed by the amino acid homology between crustacean and molluscs. The legislation in Canada and the European Union therefore requires a different label for crustacean and molluscs on food products, to protect allergic consumers. However, currently there is no analytical method available to distinguish crustacean and molluscs allergens.

Current methods for shellfish allergen detection and quantification are mainly based on antibodies utilised enzyme-linked immunosorbent assay (ELISA) techniques. All experimentally developed or commercially available ELISAs for shellfish allergen quantification are based on tropomyosin recognition. However, due to the amino acid homology between different tropomyosins, ELISAs cannot distinguish between crustacean and molluscs. Recently mass spectrometry techniques have been applied for allergen detection, allergen identification and allergen quantification to overcome the disadvantages of ELISAs. A detailed review on the current status of food allergy detection using mass spectrometry is provided **in chapter 1**. The main aim of this PhD thesis was to develop and validate a novel quantitative mass spectrometry method (LC/MRM) to be able to distinguish between crustacean species and mollusc species.

To develop the LC/MRM based method for the quantification of tropomyosin from crustacean or molluscs, a positive tropomyosin control was generated **in chapter 2**, namely recombinant tropomyosin from King prawn (*Melicertus latisulcatus*). King prawn is a commonly consumed prawn species in Australia and the amino acid sequence of tropomyosin was analysed for the first time in this thesis. The investigated tropomyosin from King prawn was registered as the novel allergen Mel l 1 and compared to the well investigated prawn species, Black Tiger prawn

(*Penaeus monodon).* The results demonstrated that the two tropomyosins had different amino acid sequences, resulting in different IgE binding of allergenic patient's sera. The expressed recombinant tropomyosin was in the following chapters used as positive control for the development of the LC/MRM method.

To develop the LC/MRM method for the quantification of tropomyosin from crustacean or molluscs **in chapter 3 and chapter 4**, twenty-two commonly consumed shellfish species in Australia have been analysed by mass spectrometry (LC/qTOF). Overall, 32 different proteins were identified, when analysing raw and whole heated shellfish extracts, whereas 16 identified proteins have been previously reported as being allergenic. The main protein identified in raw crustacean was arginine kinase, whereas in whole heated crustacean it was tropomyosin. In contrast to mollusc species where three major proteins were identified in raw and whole heated molluscs extracts, namely actin, arginine kinase and tropomyosin.

Fourteen tryptic peptides derive from tropomyosin have been identified analysing extracts of 22 different shellfish species by LC/qTOF. These 14 tryptic peptide were *in silico* analysed using 106 different tropomyosins to identified signature peptides to distinguish crustacean and molluscs. Based on the *in silico* data four peptides were selected to distinguish shellfish subgroups based on tropomyosin. In detail, peptide 1 is unique for crustacean and peptide 3 is unique for molluscs. Peptide 2 is present in crustacean and cephalopods, but not in bivalves and peptide 4 is present in prawns and lobster, but not in krill and crabs, whereas the latter includes exceptions.

To develop the LC/MRM method for the quantification of tropomyosin from crustacean or molluscs **in chapter 5** the four identified peptides were chemically synthesised. Applying these four chemical synthesised peptides, the LC/MRM method was successfully developed and validated for different shellfish species. As predicted *in silico*, the four peptides were quantified in 22 shellfish species utilising raw and whole heated extracts, with the exception of four species. Overall, the concentration of tropomyosin is higher in whole heated extracts compared to raw extracts. Moreover, in whole heated crustacean the concentration of tropomyosin is higher compared to whole heated molluscs.

The validated LC/MRM method was applied **in chapter 6** for the quantification of allergenic tropomyosin in highly processed food samples. Moreover, the LC/MRM method was compared with two commercial available ELISA kits, to confirm that both chemical and antibody based methods can quantify highly processed tropomyosin. Overall, both methods can detect allergenic TM, whereas the ELISAs had difficulties quantifying mollusc tropomyosin. Moreover, the ELISAs can certainly not distinguish crustacean from mollusc in food samples. The concentrations quantified for the food samples varied for the LC/MRM method compared to the two ELISA kits, whereas the results of one ELISA kit were similar to the concentrations quantified by LC/MRM. Thus one can assume the other ELISA kit overestimates the concentration of allergenic tropomyosin in food samples.

Overall in this PhD thesis a novel quantitative LC/MRM method was developed and validated to distinguish allergenic tropomyosin from crustacean species and from mollusc species. The validated LC/MRM method was successfully applied for the quantification of allergenic crustacean tropomyosin and allergenic mollusc tropomyosin in 22 different shellfish species and for thirteen highly processed food samples. Therefore it was demonstrated that the quantification of tropomyosin by LC/MRM is a suitable, specific and sensitive alternative to currently existing antibody based methods, such as ELISAs. The work presented in this thesis provides an important contribution towards the detection and quantification of allergenic tropomyosin from crustacean and molluscs, to fulfil the international legislation requirements for processed food.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

°C	Degrees centigrade
µg	Microgram
µl	Microlitre
µm	Micrometer
µM	Micromolar
ACN	Acetonitrile
AK	Arginine kinase
BLAST	Basic Local Alignment Search Tool
BMV	Department of Health Human Services Food and Drug Administration using the Guidance for Industry for Bioanalytical Method Validation
B.Q.	Below quantification
BTP	Black Tiger prawn (*Penaeus monodon*)
BSA	Bovine serum albumin
CD	Circular dichroism (CD) spectroscopy
cDNA	Complementary deoxynucleic acid
CV	Coefficient of variation
Da	Daltons
DDA	Data dependent acquisition
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ELISA	Enzyme-linked immunosorbent assay
EU	European Union
ESI	Electrospray ionisation

FA	Formic acid
FBA	Fructose-1,6-bisphospahtase
G	Gravity
GAPDH	Glyveraldehyde-3-phospate dehydrogenase
GO	Gene Ontology (GO) Annotation Database
H	Hour(s)
H_2O	Water
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
IUIS	International Union of Immunological Societies
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IS	Internal standard
IT	Ion trap
ITPG	Isopropyl-β-D-thiogalactoside
kDa	Kilo daltons
KP	King prawn (*Melicertus latisulcatus*)
L	Litre
LB	Luria–Bertani medium
LC	Liquid chromatography
LC/MS	Liquid chromatography coupled with mass spectrometer
LC/qTOF	Liquid chromatography coupled with a quantitative time of flight mass spectrometer
LC/MRM	Liquid chromatography coupled with a multiple reaction monitoring mass spectrometer
LLOD	Lower limit of detection
LOD	Limit of detection

LOQ	Limit of quantification
MALDI	Matrix-assisted laser desorption/ionization
Mascot	Mascot software, the benchmark for identification, characterisation and quantitation of proteins using mass spectrometry data.
MHC	Myosin heavy chain
Min	Minute(s)
MLC	Myosin light chain
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS^E	MS^E records data without discrimination or preselection. In a low energy channel the data about precursor ion is acquired and in a second channel it elevates the energy for fragmentation.
MSMS	MS/MS: precursor ion scan, product ion scan
MWCO	Molecular weight cut-off
m/z	Mass-to-charge ratio
N.D.	Not detected
nTM	Native tropomyosin
nM	Nanomolar
nm	Nanometer
OD600	Optical density measured at a wavelength of 600 nm
PBS	Phosphate-buffered saline
PLGS	ProteinLynx Global Server
PCR	Polymerase chain reaction
ppm	Parts per million
PTM	Post-translational modification

PVDF	Polyvinylidene difluoride
QC	Quality control
qTOF	Quantitative time of flight
RNA	Ribonucleic acid
RT	Room temperature
rTM	Recombinant tropomyosin
SCBP	Sarcoplasmic calcium binding protein
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
Sec	Second(s)
SCOP	Structural classification protein database
SOC	Super optimal broth with catabolite repression
TEMED	Tetramethylethylenediamine
TIC	Total ion chromatogram
TIM	Triose phosphate isomerase
TM	Tropomyosin
TOF	Time of flight
UPLC	Ultra performance liquid chromatography
V	Voltage
WC	Work concentration

CHAPTER 1

INTRODUCTION TO MASS SPECTROMETRY FOR ALLERGEN DETECTION

Data of this chapter are published in:

Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). *Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems.* ***Journal of Proteome Research.*** 13, (8), 3499-3509.

1.1 General introduction

Food allergies are increasing worldwide and becoming a public health concern. Therefore food legislation requires detailed declarations of potential allergens in food products, leading to an increased capability to analyse for the presence of food allergens. Shellfish allergy belongs to “The Big 8” food allergies and therefore needs to be declared on food products. However, the shellfish group is very heterogeneous, including various species of crustacean and molluscs in the human diet. The major allergen in shellfish is tropomyosin and tropomyosin shows high amino acid sequence homology within the shellfish group, but even higher similarity between crustacean, insects, arachnids and nematodes. The high amino acid sequence identity causes a high cross-reactivity within different species of the shellfish group, but also between other invertebrates from different phyla. Due to amino acid sequence homology, current methods cannot distinguish between crustacean and molluscs allergens. However, the food legislation in Canada and the European Union require different food labels for crustacean and molluscs.

Currently antibody based methods are mainly utilised to quantify allergens, however, these methods have several disadvantages. Moreover, the legislation for shellfish allergens leads to a technical challenge for detection methods, due to cross-reactivity between tropomyosin from different species. Recently mass spectrometry (MS) techniques have been developed and applied for food allergen analysis. Nonetheless, quantification of allergens by MS is not routinely employed. This chapter compares the different aspects of food allergen identification and quantification applying advanced MS techniques, including multiple reaction monitoring (MRM) for allergen quantification. The latter provides lower limits of quantification (LOQ) for multiple allergens in simple or complex food matrices, while being robust and reproducible. Moreover, this chapter summarises the current research for food allergen analysis by MS, showing that the quantification of food allergens is technical feasible by MRM. Therefore in this thesis it is demonstrated that quantification and differentiation of crustacean and molluscs, utilising tropomyosin and MRM, is achievable, as required by the legislation.

1.2 Food allergy

Food allergies are caused by proteins, also termed allergens, which are generally not considered harmful to the human body. Hence, food allergy is a hypersensitive reaction of the human immune system. Currently, sensitisation rates to one or more allergen among children are globally 40-50%.[1] Worldwide an estimated 220-250 million people suffer from food allergy.[1] Typical allergic symptoms include mild to severe reactions, such as urticaria, vomiting, rhinitis, asthma and life threatening anaphylaxis.[2-7]

Ninety percent of all food allergies are caused by eight food groups. These eight groups, often referred to as “The Big 8” food allergies, include egg, fish, milk, peanut, soy, tree nuts, wheat and shellfish. Currently more than 600 food allergens are known on the molecular level. Of these, 206 are officially registered by the allergen nomenclature subcommittee established by the International Union of Immunological Societies (IUIS) (http://www.allergen.org). To complicate the analysis, different food sources are known to have more than one allergenic protein. The Allergome database (http://www.allergome.org) reports 6969 allergens derived from 2454 species. In this database, for example, celery has six known allergens, molluscs three, peanut 13, lupine only one and mustard five. In soybean at least 212 proteins are reported to be possibly allergenic, of which approximately 20 proteins are confirmed to be allergenic.

The prevalence of food allergies is increasing worldwide, as is the reported number of identified food allergens. Children have a higher prevalence of food allergies with approximately 4-8% compared to adults with approximately 1-5%, based on large population studies.[2, 4, 7-9] The overall prevalence of allergenic sensitisation to “The Big 8” food allergies for adults and children is shown in figure 1.1. Over 75% of allergic children are allergic to milk, peanut, egg and tree nuts. In contrast, the vast majority of adults are allergic to shellfish, peanut, tree nuts and fish. This different distribution of food allergy prevalence can be explained by the fact that children have the ability to outgrow milk and egg allergy. Contrary to shellfish and peanut allergy, which cannot be outgrown, hence are lifelong allergies. Unfortunately, shellfish, peanut and tree nut allergies can cause severe reactions, in worst case anaphylaxis which can potentially lead to death.

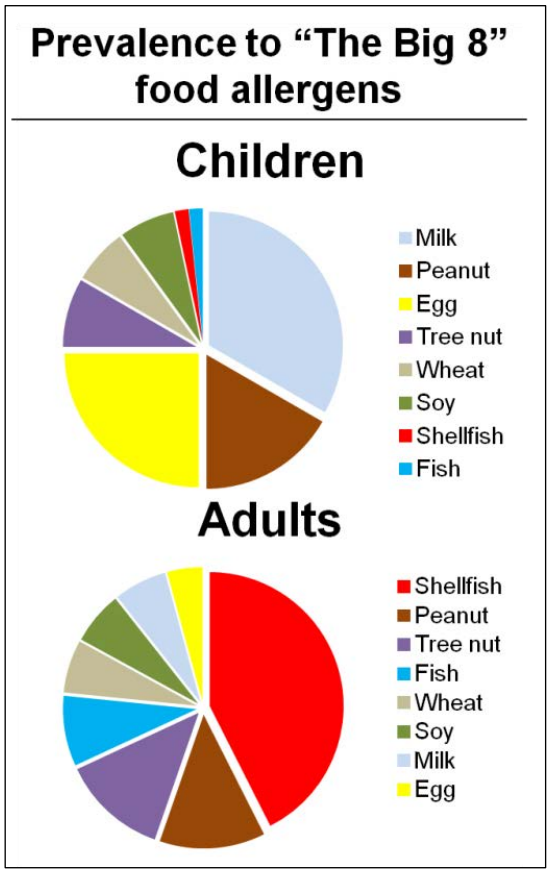


Figure 1.1: Prevalence profile of allergic sensitisation to "The Big 8" food allergens among children and adults**.** The highest prevalence data are listed in decreasing order.[10]

1.3 Classification of shellfish

In human diet shellfish is classified as species from crustacean and molluscs, although these species are originated from different phyla (Figure 1.2). Unfortunately, a large variety of shellfish species are reported to cause server allergic reactions. Over 50,000 different crustacean species and over 100,000 different mollusc species can be found worldwide and are consumed by humans.[6, 11] However, molluscs are not as often consumed as crustacean.

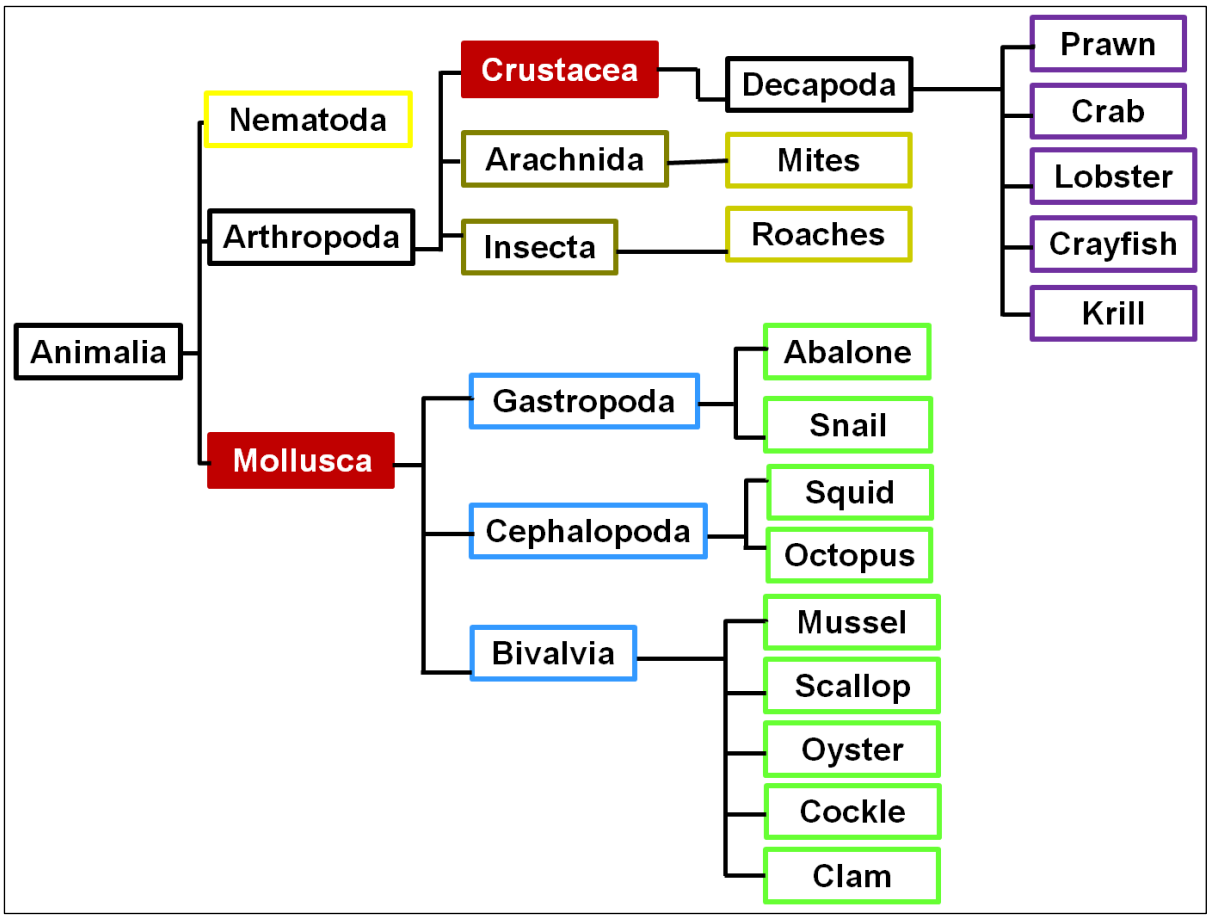


Figure 1.2: Schematic taxonomy tree view of the shellfish group and related invertebrate species, potentially causing allergic reactions. The crustacean group is encircled in purple, whereas the classes of the molluscs are encircled in blue and subgroups green.

Crustaceans are a subphylum of the arthropod phylum. However, the classes of insects and arachnids also belong to the phylum of the arthropods. Moreover, the phylum of nematodes is closely related to the arthropods phylum. In contrast to the molluscs phylum, which is more distanced to the arthropods. The classes of insects, arachnids and nematodes include species such as house dust mites, cockroaches and Anisakis, which are also known to cause allergies.

The edible crustaceans belong mainly to the order of decapods. Decapods are classified into further groups, which play an important role in human diet. These groups are prawns/shrimps, crabs, lobsters, crayfish and krill. The term “prawn” and “shrimp” are often used interchangeably in research as well as industry. The term prawn is more commonly used in Australia and other commonwealth countries, whereas the term shrimp is most commonly used in Europe and the

United States. In this PhD thesis the edible decapods will be divided into three main groups, (1) prawns (including prawns/shrimps), (2) crabs and (3) lobsters (including lobsters and crayfish). The lobsters and crayfish will be grouped together, because they are very closely related, sharing the same infra-order (Astacidea). However, every main group includes various species.

The phylum of molluscs is divided into 11 classes, however only three play a role in human diet (Figure 1.2). These three classes within the molluscs are gastropods, bivalves and cephalopods. The most commonly consumed molluscs species belong to the class of bivalves, including the main groups of mussels, oysters, scallops, clams and cockles, followed by the class of cephalopods, including the main groups of squid and octopus, and the class of gastropods, including the main groups of abalone and snail. However, every main group includes various species, playing a role in human consumption.

Overall, the shellfish group is a very heterogeneous group and includes many different species. Due to the diversity of the shellfish group the detection and quantification of shellfish allergens is technical challenging. Moreover, only few species have been investigated for their allergenicity yet.

1.4 What is a food allergen?

Food allergens are proteins that mostly originate from plant or animal sources. Most allergens are water soluble proteins in the range between 3-160 kDa, mostly between 20-70 kDa.[12] These proteins can be functional proteins, enzymes or structural proteins.[13-15] Allergens are very stable, considering chemical or physical treatments. Moreover, allergens show a high resistance to pH, denaturing chemicals, heat and degradation by proteases and proteolysis.[13] Structural elements to enhance stability of food allergens are, for example, disulphide bounds and N-glycosylation.[13] The glycosylation of proteins can increases the ability to become absorbed by respiratory or gastrointestinal mucosa.[13]

Food allergens are often highly evolutionary conserved proteins, thus they have a very specific function. To fulfil these function proteins need to have a certain structure. Although there are many different food allergens, it was reported by Jenkins et al.[14] that animal food allergens can be classified in three major protein

families and 14 minor protein families. The major protein families are tropomyosins, EF-hand proteins and caseins. The minor families only contain one to three reported allergens and are: κ-casein, lipocalin, serum albumin, c-type lysozyme, transferrin, immunoglobulin, arginine kinase, serpin, ovomucin, vitogellin N, Kazal inhibitor, very low-density lipoprotein, Kunitz bovine pancreatic trypsin inhibitor and myosin tail.[14]

Almost all animal food allergens have homologues protein in the human proteome. Nevertheless, it seems that proteins are not allergens when they share more than 62% amino acid sequence identity with the human homolog protein.[7, 16, 17] If the amino acid sequence identity of proteins is less than 54% to the human genome all proteins can become potential allergens.[16] For cross-reactivity between allergens more than 35-40% amino acid sequence identity is necessary.[3, 18] However, the structure of the allergen is important for the possibility of being cross-reactive. It was reported, when proteins share the same tertiary structure, the amino acid identity can be as low as 20-30% and still be cross-reactive.[17, 19]

1.5 Shellfish allergens

Various species of crustacean and a few mollusc species have been investigated for their allergenic properties in the literature. The Allergome database (http://www.allergome.org) reports 297 entries for crustacean allergens, with 14 different allergens. Whereas for molluscs allergens only 156 entries are reported, including three different allergens, which are tropomyosin, paramyosin and arginine kinase. Thirty-one crustacean allergens and three mollusc allergens are registered with the IUIS, as summarised in table 1.1 and table 1.2.

Table 1.1: Allergenic tropomyosins from shellfish registered with the IUIS. The table is divided into the main groups of shellfish, summarising the allergen name, the common name and the scientific name.

Shellfish Species					
		Allergen Name	**Common name**	**Scientific name**	**Allergen**
Crustacea	Prawns	Cra c 1	North Sea shrimp	*Crangon crangon*	Tropomyosin
		Lit v 1	Vannamei prawn	*Litopenaeus vannamei*	Tropomyosin
		Mac r 1	Giant freshwater prawn	*Macrobrachium rosenbergii*	Tropomyosin
		Met e 1	Greasyback shrimp	*Metapenaeus ensis*	Tropomyosin
		Pan b 1	Northern shrimp	*Pandalus borealis*	Tropomyosin
		Pen a 1	Northern brown shrimp	*Penaeus aztecus*	Tropomyosin
		Pen i 1	Indian prawn	*Penaeus indicus*	Tropomyosin
		Pen m 1	Black Tiger prawn	*Penaeus monodon*	Tropomyosin
	Crabs	Cha f 1	Swimmer crab	*Charybdis feriatus*	Tropomyosin
		Por p 1	Blue Swimmer crab	*Portunus pelagicus*	Tropomyosin
	Lobsters	Hom a 1	American lobster	*Homarus americanus*	Tropomyosin
		Pan s 1	Spiny lobster	*Panulirus stimpsoni*	Tropomyosin
Mollusca	Gastropods	Hel as 1	Brown garden snail	*Helix aspersa*	Tropomyosin
	Cephalopods	Tod p 1	Squid	*Todarodes pacificus*	Tropomyosin

1.5.1 The major shellfish allergen - tropomyosin

Tropomyosin (TM) is a highly conversed protein in the animal kingdom and found in muscle and non muscle tissues.[20, 21] However, only invertebrate TM is allergenic

to humans.[20] TM plays a major role in the contraction of muscles; hence it occurs in large amounts in muscle tissue, such as skeletal, cardiac and smooth muscles. Wrapped around actin filaments, the coiled-coil α-helices of TM regulate the muscle contrition by cooperatively blocking or exposing the myosin-binding sites on actin. This blocking also involves the troponin complex, which is activated by changes of the calcium concentration.[20, 21]

Allergenic tropomyosins in shellfish and other related invertebrates consist out of 284 amino acids and have a molecular weight of approximately 34-39 kDa.[7, 20, 22] The two linear parallel α-helical TM molecules are held together by various hydrophobic residues on the interface between helices as well as salt bindings, flanking the two α-helical chains. These structures make TM very stable towards temperature, chemical and physical treatment and highly water soluble.[13, 23, 24]

1.5.1.1 Allergenic tropomyosin cross-reactivity

Tropomyosin, as the major allergen in shellfish, is a well established pan-allergen within the shellfish group but also with related invertebrates. The cross-reactivity of the pan-allergen TM is due to the amino acid identity and the similar linear epitope regions for TM. The amino acid sequence of TM is very similar for the different shellfish groups. In detail, within crustacean the amino acid identity is 88-100%, sharing the highest sequence identity with 98-100% within the prawns.[6, 20, 25] In comparison to crustacean, the amino acid identity within molluscs is only 60-90%.[20, 22] The amino acid identity between crustacean and molluscs is only 55-80%.[7] However, the amino acid identity between crustacean and other invertebrates is as high as 74-100%,[20] with Anisakis sharing the lowest sequence identity with 74%.[7] Allergenic tropomyosin is only 50-55% identical with the non allergenic human tropomyosin.[7] As described in section 1.4 a protein is more likely to be an allergen, if it shares less than 62% amino acid sequence identity with the human analogue.[7, 16, 17]

An epitope is the region of an allergen that is recognised by the human immune system, mainly by IgE antibodies. Due to amino acid sequence homology of tropomyosin in different shellfish species, epitopes can be similar or even identical. Currently four crustacean species,[26-30] two related invertebrates,[27] and three

mollusc species[31-33] have been investigated for TM epitopes. Figure 1.3 shows that the linear epitopes in different species can be found throughout the whole TM sequence, however, some regions are commonly reported for being epitopes. For example the amino acid region from 91-101 is reported as epitope for all species investigated, interestingly this region shares high amino acid sequence identity within crustacean, molluscs and other related invertebrates. In contrast to amino acid region from 50-60, where epitopes have been reported for most species investigated, however, the amino acid sequence homology in this area is diverse for shellfish. Nonetheless, even if the same species was investigated, different epitopes have been reported.[26, 28, 29]

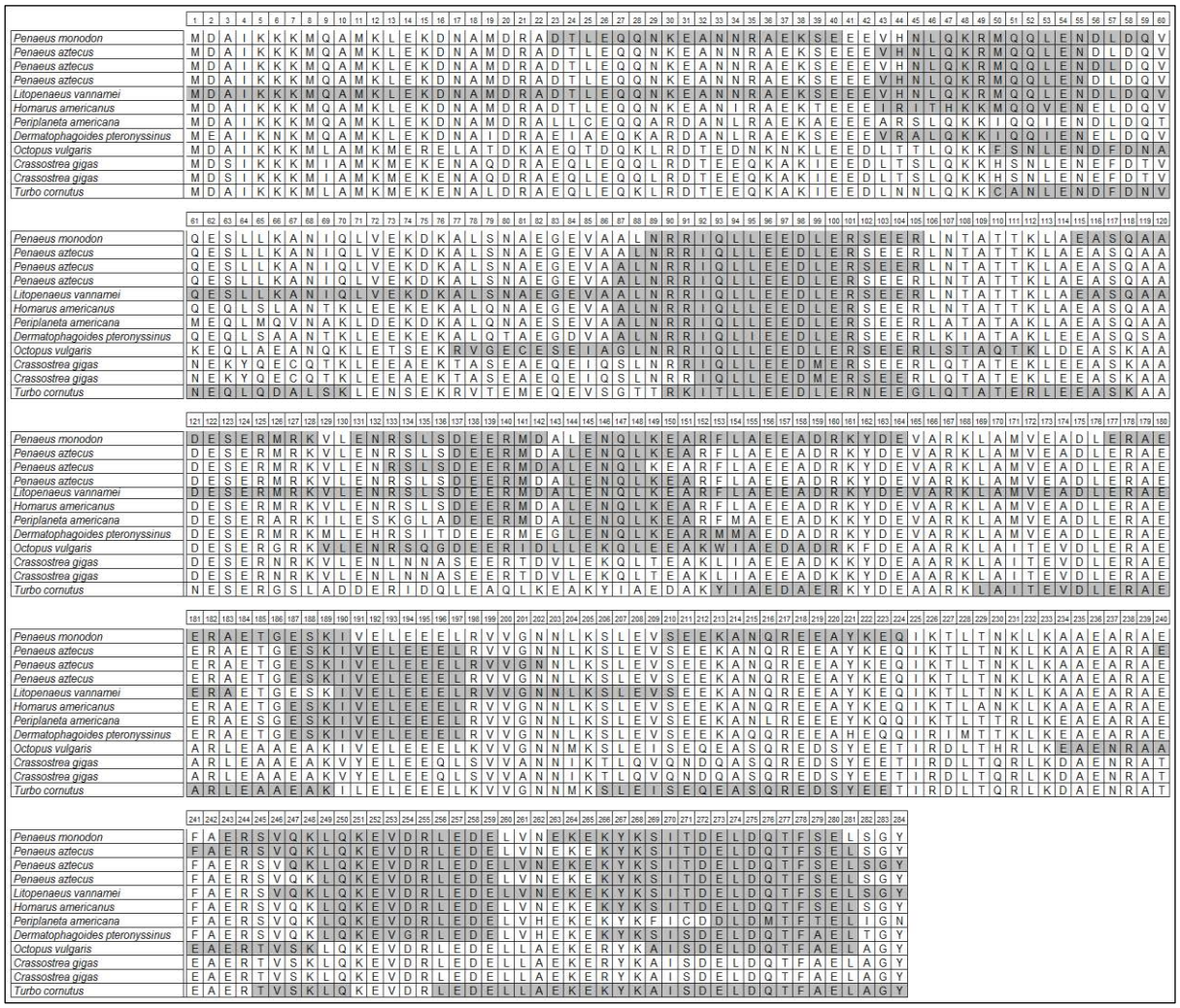


Figure 1.3: Tropomyosin sequences aligned with reported epitopes highlighted in gray for investigated species in the literature.

Overall, epitopes of tropomyosin vary, suggesting the presence of both common and more or less species specific epitopes. Studies at the molecular level, demonstrating that mollusc tropomyosins do not always have IgE-binding epitopes in the same amino acid regions, support the notion that cross-reactivity between molluscs is limited compared to cross-reactivity between crustaceans.[11, 27]

1.5.2 Other shellfish allergens

There are several other allergens found in shellfish. However, there are not as well characterised and investigated as the major allergen tropomyosin. Additionally, more research has been carried out using species of the crustacean group, compared to the group of molluscs. Hence, these other reported allergens could be major allergens once they are better investigated.[11] Thus, an allergen is classified as major allergen when more than 50% of the patient's sera demonstrate IgE reactivity binding. If less than 50% of patient's sera show IgE binding it is considered as minor allergen.

Other established allergens in crustacean are arginine kinase, myosin light chain and sarcoplasmic calcium binding protein, whereas arginine kinase and paramyosin are established allergens in molluscs. Excluding TM, other allergens registered with the IUIS in crustacean are arginine kinase, myosin light chain 1, myosin light chain 2, triose phosphate isomerase, troponin C and troponin I (Table 1.2). Other allergens that have been reported for crustacean are actin,[34] enolase,[35] fructose-1,6-bisphosphate,[36] glyceraldehyde-3-phosphate dehydrogenase,[34, 35] hemocyanain,[35, 37] myosin heavy chain,[34, 35] and titin.[36] Table 1.1 and table 1.2 summarise that for molluscs only TM and one unknown protein are registered with the IUIS. Moreover, only two other allergens are reported in literature for molluscs, being arginine kinase[38] and paramyosin.[38, 39]

Table 1.2: Other shellfish allergens registered with the IUIS. The table is divided into the main groups of shellfish, summarising the allergen name, the common name, the scientific name and the registered allergen.

Shellfish Species					
		Allergen Name	**Common name**	**Scientific name**	**Allergen**
Crustacea	Prawns	Arc s 8	Crustacean species	*Archaeopotamobius sibiriensis*	Triosephosphate isomerase
		Art fr 5	Brine shrimp	*Artemia franciscana*	Myosin, light chain 1
		Cra c 2	North Sea shrimp	*Crangon crangon*	Arginine Kinase
		Cra c 4			Sarcoplasmic calcium-binding protein
		Cra c 5			Myosin, light chain 1
		Cra c 5			Troponin C
		Cra c 8			Triosephosphate isomerase
		Lit v 2	Vannamei prawn	*Litopenaeus vannamei*	Arginine kinase
		Lit v 3			Myosin, light chain 2
		Lit v 4			Sarcoplasmic calcium-binding protein
		Pen m 2	Black Tiger prawn	*Penaeus monodon*	Arginine kinase
		Pen m 3			Myosin light chain- 2
		Pen m 4			Sarcoplasmic calcium-binding protein
		Pen m 6			Troponin C
	Lobsters	Hom a 3	American lobster	*Homarus americanus*	Myosin light chain 2
		Hom a 6			Troponin C
		Pon l 4	Narrow-clawed crayfish	*Pontastacus leptodactylus*	Sarcoplasmic calcium-binding protein
		Pon l 6			Troponin I
Mollusca	Gastropods	Hal m 1	South African abalone	*Haliotis midae*	unknown

1.6 Legislation

The prevalence of food allergies and severity of allergenic reactions are increasing as well as the number of different allergens. In consequence governments try to protect allergic individuals. Many countries have enacted laws to ensure health of allergic individuals based on the Codex Alimentarius Commission (a joint committee with delegates from both the Food and Agriculture Organization of the United Nations and the World Health Organization).[1] However, as summarised in table 1.3, different countries mandate a different selection of allergens for food labelling. Already 14 different food groups are required for allergen labelling in the European Union followed by Canada (11 allergens) and Australia/New Zealand with nine allergens, compared with Japan mandating just four allergens.[40-42]

In terms of shellfish allergens, table 1.3 shows that crustacean (shellfish) allergens need to be labelled on food products in all countries. Moreover, in Canada and the European Union, molluscs allergens need be declared separately Thus, as explained in section 1.5.1.1, the major allergen in shellfish shares lower amino acid homology within shellfish and linear epitopes can vary, leading to lower cross-reactivity between crustacean and molluscs.[11] Due to these reasons crustacean and molluscs need to be labelled differently in these countries.

Table 1.3: Allergens requiring labelling on food products. "The Big 8" food allergens are ordered alphabetically (shaded in purple) and additional allergens are below them as currently required by legislation. "✔" indicates that the allergen needs to be labelled on every food product.

Source/ Allergen	Codex Alimentarius	Australia/ New Zealand	European Union	United States	Canada	China	Hong Kong	Japan	Korea	Mexico
Crustacean	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Egg	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Fish	✔	✔	✔	✔	✔	✔	✔		✔	✔
Milk	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Soy	✔	✔	✔	✔	✔	✔	✔		✔	✔
Peanut	✔	✔	✔	✔	✔	✔	✔			✔
Tree Nuts	✔	✔	✔	✔	✔	✔	✔			✔
Wheat/Cereals	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Celery			✔							
Lupine			✔							
Mustard			✔		✔					
Sesame		✔	✔		✔					
Shellfish/ Molluscs			✔		✔					
Sulphur oxide and sulphites			✔							

1.7 Analytical difficulties for food allergen detection

The detection and the quantification of food allergens is challenging. Hence, there are numerous different allergens which can be found in different food products. Therefore the food matrix can be very different. However, if an allergen is directly analysed from the unprocessed food product, they can be extracted due to the allergen solubility. For example, allergens from egg and milk are spread evenly

throughout food product and are easily extracted.[43] Shellfish and fish tissues are uncomplicated to disperse. On the other hand, peanut, tree nuts, sesame and soy contain more oily components and therefore need more attention in allergen extraction.[43] For example, quite complex food matrices to analyse are chocolate, due to their high content of fat, carbohydrates and polyphenoles.[43]

Nevertheless, food products are often processed during food production. These include chemical and physical treatments, mainly to increase shelf life. These processes can significantly alter the physicochemical and the structural properties of allergens, thereby increasing or attenuating the allergenicity.[44] Moreover, the structure and the solubility of an allergen can change. During the food processing allergens can undergo modifications, such as unfolding, aggregation, glycosylation, oxidation, denaturation or interference with other compounds of the food matirix.[44, 45] Glycosilation and oxidation are examples for post-translational modifications (PTMs). These PTMs can result in a possible change to the secondary and tertiary structure of an allergen. All modification can influence the allergen or the epitopes, especially conformational epitopes, changing the antibody recognition site.[46] In terms of food safety, processing can also lead to a pronounced reduction of analytical sensitivity when using biological methods such as enzyme-linked immunosorbent assay (ELISA).[44] This way false-negative results might be obtained that discriminate modified but allergenic proteins.

The Maillard reaction is a commonly investigated non enzymatic glycation reaction that occurs during food processing.[24] Protein modifications via the Maillard reaction are known to modulate the allergenicity of food proteins.[44] For example, it was reported that the peanut allergen Ara h 1 increased IgE-binding capacity when being roasted.[47] In the case of tropomyosin it was also described that the Maillard reaction increases allergenicity.[48]

1.8 Current biological methods for food allergen quantification

Legislation demands accurate and robust and sensitive methods, however, current methods have some drawbacks. Technical methods to detect allergens are either not validated and the limit of detection and quantification is not sensitive enough or

just not available. Moreover, there is a lack of well characterised standards or reference materials. Therefore right or wrong declarations cannot be proven.

Current methods for food allergen analysis are biological methods based on (1) antibodies or on (2) DNA. Antibody based methods are mainly 'Enzyme Linked Immunosorbent Assays' (ELISA) or immunoblotting. ELISA is a quantitative food allergen detection method, whereas immunoblotting is only semi-quantitative. DNA base methods use PCR techniques. The most common method for food allergen detection and quantification are ELISAs. For most of "The Big 8" food allergens ELISA kits and PCR methods are available. The big exception is fish, thus only one PCR method is available.[2] Most ELISA kits, however, only target a single allergen from a single food product. For example, for the analysis of milk allergens, ELISA kits are commercially available to detect casein, β-lactoglobulin and total allergen content (casein and β-lactoglobulin). None of the other known milk allergens including α-lactoglobulin, which is also considered as a major allergen in milk, is targeted.[49] For other allergens, which need to be labelled in some countries (Table 1.3) only PCR methods are available, e.g. for lupin and celery.

However, DNA based methods and immunoblotting are not commonly used for food allergen quantification. Therefore these methods will not be discussed in detail. Briefly, the advantages of DNA based methods are that they are sensitive and highly specific. The disadvantages of DNA based methods are that they only detect the possibility of allergen being present. Moreover, the sequence of DNA from the allergen is necessary to generate primers of the detection. Immunoblotting, as mentioned above, is only semi-quantitative.

1.8.1 ELISAs for food allergen quantification – biological method

The advantages of ELISAs are their sensitivity and selectivity, thus antibodies are highly specific to the allergen. Many antibodies are commercially available, however, often antibodies are poorly characterised.[50] Cross-reactivity can occur when using antibody based methods, leading to potential false positive results.[23, 36, 50, 51] Complex food matrices can comprise interfering components, e.g. polyphenols or tannins, which interact with or bind to proteins and antibodies.[52, 53] Additionally, food processing or sample preparation can modify allergens, which

subsequently are not recognised by the target antibody, leading to potential false negative results.[12, 54]

A comparative study by Heick et al.[55] of two commercially available ELISA kits for soy noted that the detection for spiked flour samples varied by a factor of 10. When they examined hazelnut in spiked processed bread they observed that results between ELISA kits varied by a factor of 3. Quantifying hazelnut in both unprocessed and processed samples using two different ELISA kits resulted in significant differences of up to 40%.[55] A similar comparative study of commercial ELISA kits for hazelnut detection by Cucu et al.[56] demonstrated that all kits evaluated produced false-positive and false-negative results. In some kits the actual hazelnut protein concentration was 17-49% underestimated and another kit overestimated the concentration by 27%.

Johnson et al.[50] performed a multi-laboratory evaluation of egg and milk allergens and demonstrated that all kits underestimated the concentration of egg. Only one kit quantified the milk protein content with acceptable accuracy at 6 and 15 mg/kg. All milk and egg ELISA kits were able to detect the lowest spiked concentration (3 mg/kg, however, the limit of quantification (LOQ) was for egg about 10 mg/kg and for milk 30 mg/kg. This highlights that the current methods would have difficulties to detect allergens in certain types of food consumed in larger quantities. Furthermore, some ELISA kits have a very low dynamic range and the generated results are difficult to compare.[50, 57]

However, the sensitivity of ELISA is very high. LOQ for ELISA reported in literature range from 0.3-1.5 ppm and limit of detection (LOD) 0.2-2 ppm, respectively.[58] Whereas other found LOD of ELISA kits vary from 1-5 ppm.[12, 59] The LOQ and LOD of allergen analysis are dependent on complexity of allergen and food matrices analysed. Moreover, it is dependent on the standard used in the commercial available ELISA kit. Unfortunately, standards provided are poorly characterised and therefore generated results using ELISA kits are not comparable at all.[52, 57, 60]

1.8.2 Biological methods for shellfish allergen analysis

Both, ELISA kits and PCR methods are available for the detection of shellfish allergens. All developed methods, when specified, state that they use tropomyosin (TM) for the detection of shellfish allergens. In Australia, two commercial available ELISA kits are available for the shellfish allergen detection. One company specifies to detect TM and using anti-tropomyosin antibodies, whereas the other company specifies to use crustacean proteins as standards and anti-crustacean-proteins antibodies. Moreover, five ELISA systems have been experimentally developed worldwide by various research groups.[61-65] All of these ELISAs used TM for detection and quantification of shellfish allergens. Out of five ELISAs only one is designed to detect crustacean and molluscs allergens,[65] however, it cannot distinguish crustacean and molluscs, as it would be required by legislation in Canada and the European Union. The other four ELISAs are designed to detect crustacean TM, however, only one has been tested not to be cross-reactive with molluscs TM.[61] Out of five experimental developed ELISAs three were tested positive for cross-reactivity with other invertebrates.[62-64] In a compassion study using two commercial available crustacean allergen ELISA kits in Japan, it was reported that both ELISAs are suitable to detect crustacean TM in highly processed food products.[51]

In summary, for many food allergens ELISA kits are commercially available. ELISAs are sensitive and specific. However, these methods are restricted to single well known allergens and less well characterised allergens are excluded. The lack of general consistent standards and reference material for allergen analysis make a comparison of results between different ELISA kits very difficult. Moreover, for shellfish allergens, there is no method available that can distinguish between crustacean and molluscs allergens, as it is required by food legislation. These insufficiencies clearly highlight that alternative methods for the quantification of food allergens are urgently needed. New technologies using chemical methods, such as mass spectrometry (MS) therefore become more relevant for allergen analysis, to support allergen labelling and certification process for allergen analysis.[12]

1.9 Mass spectrometry for food allergen detection and quantification - chemical method

The disadvantages of current established methods for allergen analyses are numerous; hence alternatives have been investigated in recent years. In particular, non biological methods, such as mass spectrometry (MS) systems have been investigated and developed to overcome the drawbacks of ELISAs. Different MS systems are commonly used in proteomics, including allergen analysis. Generally, a mass spectrometer is composed of three different parts: ion source, mass analyser and detector. Most commonly used as ion sources are matrix-assisted laser desorption/ionisation (MALDI) or electrospray ionisation (ESI). For mass analyser (quantitative) time of flight ((q)TOF) and ion trap (IT) are used in proteomics. Combining different ion sources and mass analysers leads to hybrid MS systems, such as ESI- qTOF (referred as qTOF in this thesis), ESI-IT (referred as IT in this thesis), or MALDI-TOF (referred as MALDI in this thesis). Those different MS systems are applied in proteomics to identify proteins utilising peptides.[66, 67] Moreover, qTOF and IT systems can quantify peptides by applying higher voltages in the collision cell, resulting in typical peptide fragments, which then can be quantified.[42] To maximise the amount of peptides that can be identified, MS systems are regulary coupled with liquid chromatography (LC), to separate proteins or peptides.[68] Stationary phases most commonly utilised by LC, to separate protein and peptide, are reversed phase columns. However, it is also possible to use other stationary phases to achieve better separation accordingly to the allergen analyses.[12, 68-70] Figure 1.4 compares and summarises the major differences for biological methods and chemicals methods, applying antibodies or MS systems for allergen detection and quantification. Overall, MS systems in the field of proteomics are relatively new technique, however, utilising MS systems for allergen identification and quantification is a new approach and therefore currently not applied for routine analysis.[42, 57, 71, 72]

Comparison Between Biological and Chemical Methods for Allergen Analysis

Food Matrix Containing Allergens

Biological Methods			**Chemical Methods**		
Antibody-based			**MS-based**		
	ELISA	**Immunoblot**	**MALDI**	**qTOF/IT**	**MRM**
Extraction Buffer	Strong Detergents	Strong Detergents	Weak Buffers	Weak Buffers	Weak Buffers
Treatment	No	No	Yes/No	Digestion	Digestion
Detection Method	UV Spectrometer	Visual	Mass Spectrometer	Mass Spectrometer	Mass Spectrometer
Allergens can be Analysed	One	One	Multiple	Multiple	Multiple
Standard Required?	Yes (internal kit standard)	No	No	Yes/No	Yes
Cross-Reactive?	Yes	Yes	No	No	No
Species Specific?	No	No	Yes	Yes	Yes
Can be Standardised?	No	No	Yes	Yes	Yes
Results Comparable?	No	No	Yes	Yes	Yes
Time	Time Consuming	Time Consuming	Fast Analysis	Fast Analysis	Fast Analysis
Cost	Medium	Low	Medium	Medium	Medium
Quantification	LOD (lowest 0.2 ppm) LOQ (lowest 0.3 ppm)	Semi - quantitative	Semi - quantitative	LOD (lowest 0.06 ppm) LOQ (lowest 3.7 ppm)	LOD (lowest 0.001 ppm) LOQ (lowest 0.01 ppm)

Figure 1.4: Comparison between biological methods (antibody based methods) and chemical methods (MS based methods) for the detection and quantification of food allergens.

1.9.1 Comparison between biological methods and chemical methods for food allergen detection

Several authors in the literature compared detection methods, such as ELISA, PCR and MS systems for food allergen analysis. In detail, Weber et al.[49] found that the results of ELISA and MS were comparable when analysing milk allergens in orange sherbet, lacto-free ice cream, milk powder, oatmeal cereal and cookies

extract. However, for processed products containing milk and soy allergens the MS detection was outstanding in comparison to ELISAs. Heick et al.[73] found that when analysing peanut, hazelnut, walnut and almond by ELISA and MS that both methods were capable to detect the allergens. Lee and Kim[74] found in their comparison study that PCR is not suitable when analysing egg allergens, thus it cannot distinguish between egg and chicken proteins, whereas only egg proteins are allergens. Moreover, the ELISA technique was suitable to detect egg allergen in trace amounts whereas MS was not able to detect ovomucoid.[74]

1.9.2 Advantages of mass spectrometry

The advantages of MS are ease of sample preparation, fast analysis and analysis of more than one allergen at a time. Moreover, MS is robust and stable and can easily be automated and standardised, with potential low LOD and LOQ.[12, 55, 73] Another advantage is to have better defined standards, which make the comparison of results between methods and laboratories much easier. High resolving power and sensitivity coupled with this independence from structural changes allows MS to detect allergens in trace amounts.[53]

Post-translational modifications (PTMs), e.g. occurring during heat treatment, can result in a possible change to the secondary and tertiary structure of an allergen. The sensitivity and the specificity of ELISAs can depend on the 3-D structure of allergens, whereas MS is based on the structurally independent amino acid sequence. PTMs can therefore have a huge impact on allergenicity of the protein, which can lead to false positive or false negative results by ELISAs.[42, 67] Nevertheless, applying LC/MS systems PTMs can be detected, providing additional information on primary and secondary protein structure.[57, 75]

The main disadvantage of the MS systems is that the equipment is very costly. Moreover, protein/peptide analysis can be effected by a range of chemical properties (size, charge, hydorphobicity) of the allergen. Some proteins/peptides show poor ionisation and therefore not sensitive in detection. The salt concentration in sample matrix can also influence ionisation, whereas higher salt content in the sample leads to higher noise of the analysis. Therefore it is required to have a standard for MS analysis.[12, 67, 76, 77]

1.9.3 Operating mass spectrometry for allergen analysis

To analyse allergens by MS system a set of methodological steps has to be followed. A schematic workflow for the identification and quantification of food allergens is shown in figure 1.5. Briefly, the allergens need to be extracted from the food matrix, followed by digestion into peptides utilising enzymes. Nevertheless, MS systems also can be applied to analyse intact allergens, but for detection and quantification of food allergens this approach is not used. The digested peptides are analysed using MS systems, such as MALDI, LC/qTOF or LC/IT. The obtained data generated from the LC/MS systems is then processed applying bioinformatics tools and protein databases. The softwares most commonly used for this purpose are Mascot and SEQUEST,[42, 55] providing detailed information about peptides in the food samples. The peptides are identified according to their mass-to-charge ratio and the allergenic protein identified by comparing the derived amino acid sequences with known proteins.[42, 55, 78] The peptide sequences that are uniquely specific for a particular allergen are termed ‘signature peptides’.

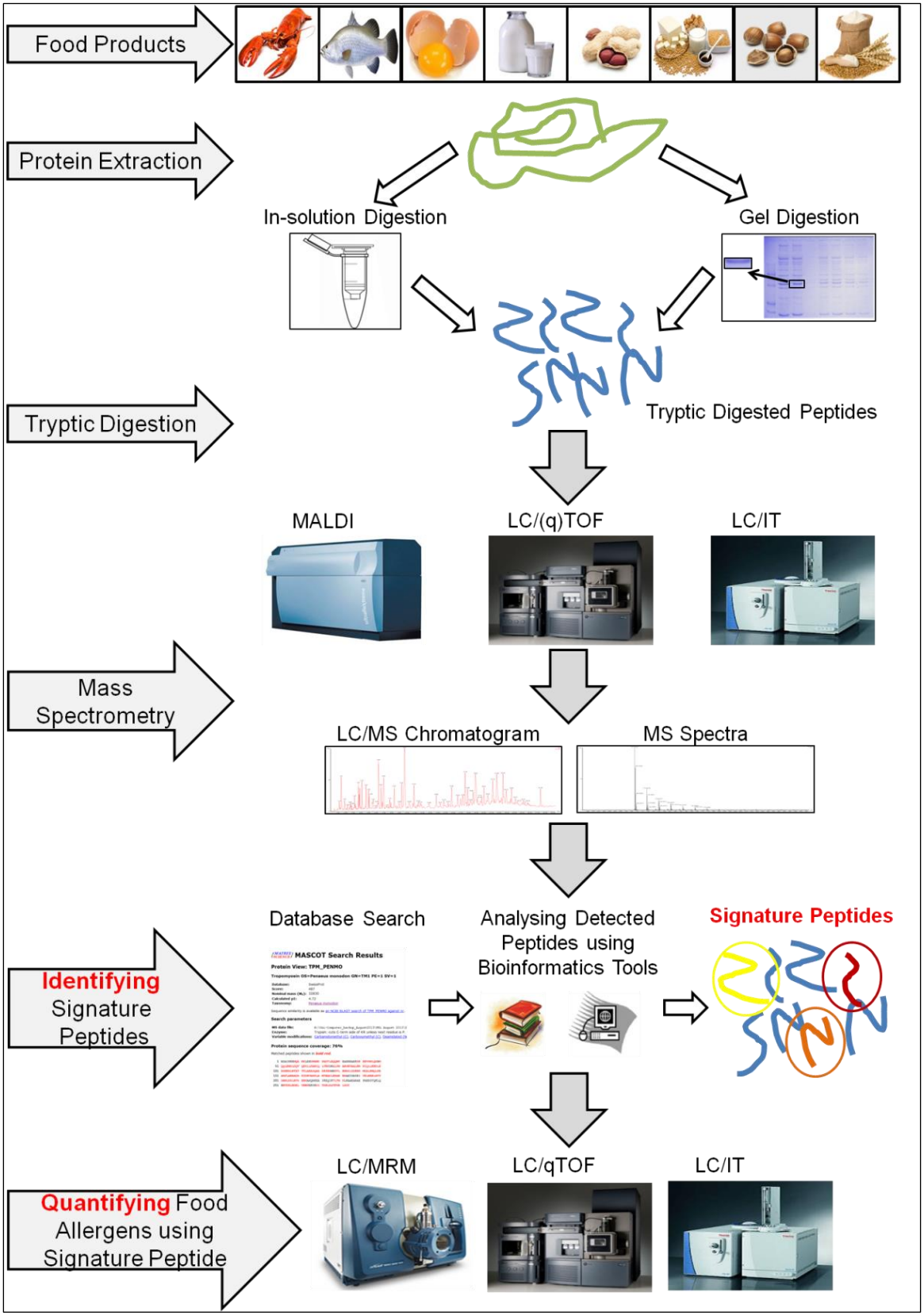


Figure 1.5: Schematic workflow for the detection, the identification and the quantification of food allergens by mass spectrometry. The figure shows the basic steps on how to identify signature peptides and apply these for food allergen quantification.[10]

1.9.3.1 Sample digestion for allergen analysis by MS systems

Digestion of the allergenic proteins is commonly performed prior to MS analysis, following the bottom-up strategy, as explained in detail by Monaci and Visconti.[42] The digestion cleaves large proteins into smaller peptides, thus potentially matrix interferences and associated interactions with other proteins are reduced. These reductions remove complicating factors and make the analysis by LC/MS more reproducible.[79] Ideally a complete digestion is achieved in a very short time with a maximum of peptides generated and without missed cleavage sites.[12, 73] The peptides generated should be stable over time and easily detected by MS. In some food samples, enzymes are naturally present and unwanted degradation of proteins occurs. Thus, the digestion process depends on the individual allergen structure with careful consideration of disulphide bridges, structural folding, solubility and glycosylation. Comparing methods adequately requires a total, robust and reproducible digestion method.

Various enzymes are available, with specific cleavage sites. However, the most commonly used enzyme is trypsin due to the well known cleavage sites between the amino acid arginine (R) and lysine (K). R and K occur in proteins common enough to derive sufficient peptide fragments from most proteins.[42] As more information about trypsin peptides is available, protein identification is more likely using database searches. Trypsin is also preferred, as it occurs naturally in the stomach and therefore is representative *in vivo* cleavage of all proteins.

Abdel Rahman, et al.[75] compared different enzymes and demonstrated that V8 enzymes have poor efficiency for in-gel digestion due to their enzyme size. Many peptides digested with V8 had missed cleavage sites, with a maximum of five missed cleavage sites. Carrera et al.[80] showed that more peptides could be identified when fish species were digested with trypsin compared to endoproteinase GluC. In contrast, Sealey-Voyksner et al.[68] generated the highest yield of peptides utilising pepsin compared with trypsin and chymotrypsin when analysing wheat gluten. Surprisingly, this study reported that a higher sample concentration, with the optimal enzyme to protein ratio, did not lead to increased yield of generated peptides due to incomplete digestion.

In summary, when identifying and detecting allergens by MS systems, the sample preparation is a critical step. The proteins are extracted from the food matrices and digested with enzymes generating peptides. The tryptic digestion approach is most commonly referenced in the literature and has the most data publically available. Trypsin digestion also occurs naturally in the stomach when allergens are ingested.

1.9.3.2 Identifying signature peptides by MS

Specific criteria must be fulfilled for the designation of a 'signature peptide' as summarised in figure 1.6. A signature peptide is defined as a theoretical tryptic peptide that is exclusively present in one group, but not in any other group. It is critical that signature peptides are unique to the target protein and detectable by the MS systems of choice.[12, 81] The selected signature peptides do not need to be the most intense signals found in the MS spectra, but they do need to be sufficiently intense to allow clear separation from other peptides or the MS background, even if only present in low quantities.[57] Other criteria are that peptides with amino acids prone to PTMs[12, 57, 71, 81] or peptides with missed cleavage sites after digestion should not be selected as signature peptides.[12, 57, 66] Preferably, the signature peptide should be between six and twenty amino acids long.[57, 66, 71, 82]

Critical Steps for Selecting Signature Peptide and Transition	
Signature Peptide Selection	• Unique • Easy Detected with MS Systems • No Modifications •6-20 Amino Acid Long • Stable • Abundant in Food Matrix • Easily Extracted • Reproducible (Digestion)
Transition Selection	• Easy Detected with MS Systems • Abundant in Food Matrix • No Modifications • Reproducible Transition Time • Reproducible Retention Time • Precursor Ion: Double or Triple Charged • Product Ion: Reproducible Transitions Specific to Peptide

Figure 1.6: Selection criteria for signature peptides and transitions for the quantification of food allergens.

1.9.3.2.1 Protein identification versus species identification

Basically there are two different approaches on how to select a signature peptide for the allergen analysis. Depending on the investigation, one option is to choose a signature peptide which is unique for the allergen selected. A more advanced approach is, which is not always possible due to allergen amino acid sequence similarity, to select a signature peptide that is unique for the allergen as well as for the species.

The species identifications of related species with similar allergens are currently almost exclusively carried out by DNA based methods.[83, 84] However, DNA does not represent the allergen nor the phenotype or any PTMs.[83, 84] Ortea et al.[83] were able to distinguish six different prawn species by MALDI analysing the minor allergen arginine kinase (Table 1.4 and table B1.1). Therefore MS systems have potential for species differentiation.

However, as seen in supplementary table B1.1, Carrera et al.[85] tried to identify different signature peptides for the fish allergen β-parvalbumin in different fish species. Although combining different peptides only five species out of 19 were unique. The remaining 14 species have identical or similar peptides identified. For β-parvalbumin, the sequence identity of different species is between 60-80%.[2, 4] Houston et al.[86] identified ten allergens in soy and were aiming to report two signature peptides per allergen (total of 20), but due to signature peptides selection criteria could only find 15 ideal tryptic peptides. For other allergens, such as the major allergen in shellfish, tropomyosin, the amino acid homology is as high as 100% for some species.[6]

In summary, the signature peptides should be carefully selected. The more information available, the higher is the certainty that the peptide represents the allergen and species of interest. It was shown that the MS systems have the ability for species specific signature peptides. Unfortunately, due to amino acid sequence homology of some allergens in different species it is not always possible to select species specific signature peptides.

1.9.4 Signature peptides identified by MS systems

Several signature peptides have been identified for the most common food allergens and are reported in the literature. Figure 1.7 shows “The Big 8” most common food allergens, where allergens analysed by different MS systems are highlighted in yellow. Overall, currently 46 different allergens have been identified from “The Big 8” food allergen group, with 32 signature peptides being reported. Moreover, allergens were analysed from 11 different food products. Table 1.4 summarises, as example, the well known pan-allergens in crustacean that have been investigated by MS. A complete summary of all allergens, species and food sources investigated by various MS systems are presented in table B1.1. Currently allergenic food proteins are identified by MS for crustacean (8 allergens), egg (5 allergens), fish (1 allergen), milk (7 allergens), peanut (3 allergens), soy (7 allergens), tree nuts (5 allergens) and wheat (10 allergens). The allergens and signature peptides reported will be described according to “The Big 8” food groups in section 1.9.4.1-1 9.4.8.

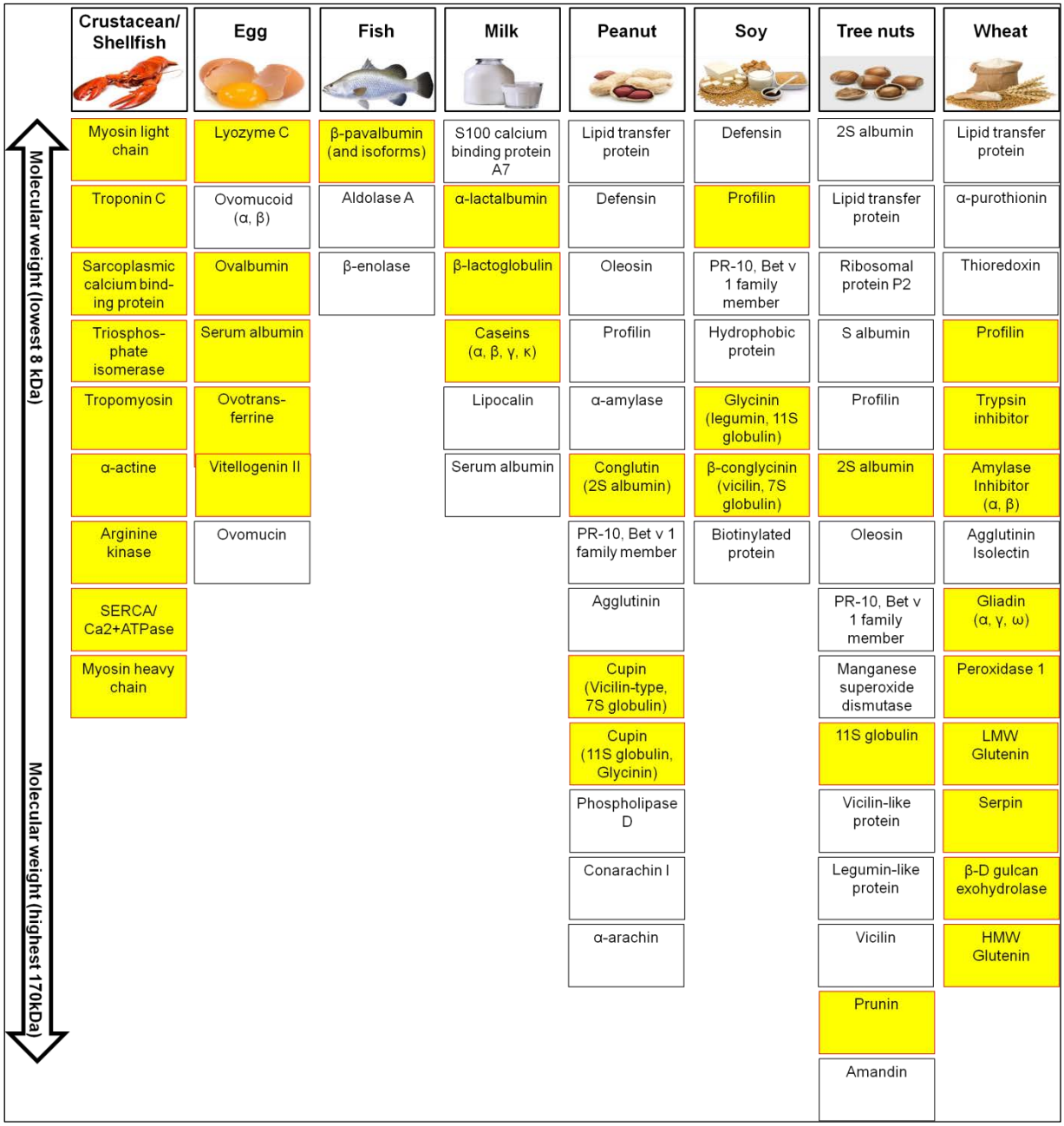


Figure 1.7: Common food allergens from "The Big 8" food allergen groups. The allergens are ordered from top to bottom by increasing molecular weight. Yellow highlights indicate allergens investigated by MS systems.[10]

1.9.4.1 Crustacean

Eight different crustacean allergens have been investigated by three different research groups (Table 1.4 and table B1.1).[34, 82-84, 87-91] Arginine kinase being the most investigated allergen in this allergen group, although only being a minor allergen in the crustacean group. Ortea et al.[84] attempted to identify different

signature peptides for seven different prawn species, using arginine kinase and two different MS systems. However, only for two species unique peptides were identified, whereas the other species peptides identified were identical. Abdel Rahman, et al.[75, 90] was successful in identifying different signature peptides for tropomyosin from Black Tiger prawn, Northern prawn and Snow crab. For mollusc allergens, as belonging to together with the crustacean to the shellfish group, only one species has been investigated. Several peptides have been reported for arginine kinase from *Octopus fangsiaorginine*.[38]

1.9.4.2 Egg

Five different allergens are identified in egg (Table B1.1) by five different research groups.[53, 55, 73, 74, 92, 93] Interestingly, identical peptides are reported for the same allergens investigated (Gal d 3). For Gal d 2 only Azarnia et al.[92] identified a signature peptide, however, this signature peptide is different to the other peptides detected for Gla d 2.[53, 74]

1.9.4.3 Fish

The major allergen in fish is parvalbumin and was investigated by Carrera et al.[80, 85] (Table B1.1). In 2010 Carrera et al.[80] could fully *de novo* sequence 25 new parvalbumin isoforms. Additional 16 new isoforms were partially sequenced investigating 13 species of the Merlucciidae family, which includes cod-like fish and many hakes. Outcomes of the study in 2012 are described in section 1.9.3.2.1.

Table 1.4: Alphabetical summary of the well characterised pan-allergens in crustacean analysed by different MS systems and their peptides published in the literature. "*"indicates recommendation use as a signature peptide. Allergen name in brackets confirms the registration with the IUIS. LOD given as published. (N.D.= Not Determined)

Allergen (registered allergen)	Peptides identified (*recommended signature peptides)	Species/ Allergen source	LOD	MS system used	Ref.
Crustacean					
Arginine kinase (Pen m 2)	• FLQAANACR • GTRGEHTEAEGGIYDISNK	*(Penaeus monodon)*	N.D.	MALDI and IT	83
Arginine kinase (Lit v 2)	• FLQAANACR • GTRGEHTEAEGGIYDISNK	*(Litopenaeus vannamei)*	N.D.	MALDI and IT	
Arginine kinase (Pen m 2)	• *AVFDQLKEK • *VSSTLSSLEGELK • *TFLVWVNEEDHLR • *LEEVAGKYNLQVR	*(Penaeus monodon)*	N.D.	LC/IT	84
Arginine kinase (Lit v 2)	• *VSSTLSSLEGELK • *TFLVWVNEEDHLR • *LEEVAGKYNLQVR	*(Litopenaeus vannamei)*	N.D.	LC/IT	
Arginine kinase	• *LVSAVNEIEK	*(Chionoecetes opilio)*	N.D.	MALDI, LC/qTOF and LC/MRM	82
Arginine kinase	• AVFDQLKEK • VSSTLSSLEGELK • GTYYPLTGMSK • LIDDHFLFK	*(Penaeus monodon)*	N.D.	MALDI and LC/qTOF	90
Myosin light chain (Lit v 3)	• KGGXNVFDMFTQK • SSGESDDDDVVAASIR	*(Litopenaeus vannamei)*	N.D.	LC/MALDI	87
Myosin light chain	• EGFQLMDR	*(Pandalus borealis)*	N.D	MALDI, LC/qTOF and LC/MRM	34
Sarcoplasmic Ca-binding protein (Lit v 4)	• YMYDIDDDGFLDK • NDFECLAVR • GEFSAADYANNQK • NLWNEIAELADFNKDG	*(Litopenaeus vannamei)*	N.D.	LC/IT	88
Sarcoplasmic calcium binding protein	• VATVSLPR	*(Chionoecetes opilio)*	N.D.	MALDI, LC/qTOF and LC/MRM	75
Tropomyosin (Pen m 1)	• *ANIQLVEK	*(Penaeus monodon)*	N.D.	LC/qTOF	90
Tropomyosin	• *SQLVENELDHAQEQLSAATHK	*(Chionoecetes opilio)*	3 nM	LC/MRM	91
Tropomyosin (Pen b 1)	• *SEEEVFGLQK	*(Pandalus borealis)*	0.25 nM	MALDI, LC/qTOF and LC/MRM	34

1.9.4.4 Milk

Seven allergens are identified in milk (Table B1.1).[49, 53, 73, 94-98] The identified signature peptides for α-S1 casein and β-casein are identical between all four research groups.[49, 53, 73, 95] One peptide is consistent with two of the four research groups for β-casein, which may be due to different MS systems applied. κ-casein was reported with two different peptides, due to the fact that Molle and Leonil[94] were trying to identify and quantify different glycosylated and non glycosylated forms of the macro peptide κ-casein.

1.9.4.5 Peanut

The three major peanut allergens (Table B1.1)[19, 52, 59, 69, 70, 73, 79, 97, 99-109] include various isoforms of cupin and conglutin. Chassaigne et al.[102] detected unique peptides for all three peanut allergens, including two different isoforms. However, only the Ara h 1 signature peptide from the 2007 study were reported again in 2009.[59, 102] All research groups investigating Ara h 1 reported similar peptides, in contrast to Shefcheck and Musser [52] and Helbing et al.,[106] who chose the most abundant peptides, opposed the identified criteria set out in section 1.9.3.2. Ara h 2 was analysed by three groups with no corresponding peptides reported. Heick et al.[73] identified different peptides for Ara h 3/4 than Chassaigne et al., [59, 102] which may again be explained by the different MS systems used.

1.9.4.6 Soy

Four research groups investigated seven different allergens in soy (Table B1.1).[69, 73, 86, 103] Cucu et al.[103] reported one signature peptide each for Gly m 5 and Gly m 6, which are not modified during food processing, such as the Maillard reaction. Heick et al. [73] reported one peptide for glycinin consistent with Cucu et al., [103] and a different peptide consistent with Houston et al.[86] None of the reported peptides are identical between Houston et al. [86] and Cucu et al. [103]

1.9.4.7 Tree Nut

The group of tree nuts include a variety of different real nuts, fruits and legumes. The Brazil nut was investigated by Moreno et al.,[104] who reported up to six isoforms for the 2S albumin (Table B1.1). Walnut, hazelnut and almond (Table B1.1) were investigated by Heick et al. [73]

1.9.4.8 Wheat

Ten different allergens have been reported for wheat allergens by five research groups (Table B1.1). [68-70, 108, 109] Wheat allergens are therefore the most investigated food allergens. α-amylase inhibitor with similar peptides, was identified by two research groups. The investigated β-amylase resulted in only one common peptide by MALDI and LC/qTOF.[70]

Overall, the current literature demonstrates that the analysis of allergens by different MS systems is successful. MS analysis overcomes the major drawbacks of established methods, such as non specific antibody allergen reaction and cross-reactivity. Moreover, it was demonstrated that many allergens can be analysed in one single LC/MS run. MS systems make it possible to generate information about amino acid sequences of allergens as well as identifying PTMs and isoforms. Many different peptides for the different allergens have been reported. Similar or identical peptides, including signature peptides are reported for the specific allergens investigated, despite the different MS systems applied. These findings highlight that MS is an outstanding approach in detecting and quantifying specific food allergens. However, consistent signature peptides which can be used as standards and reference materials need to be further evaluated.

1.10 MRM systems for absolute allergen quantification

The use of multiple reaction monitoring (MRM) for peptide analysis has been a more recent approach for clinical applications and for allergen analysis. New MRM systems are now available utilising triple quadruples with extended mass ranges.[66] However, MRM does not provide accurate mass determinations when compared to other MS systems, such as MALDI, qTOF and IT, used for

peptide/protein/allergen analysis. In the MRM approaches, the time window of a selected signature peptide is scanned for a defined time period, accordingly to its elution or retention time in the chromatographic run.[12, 55] More recent MRM systems are able to narrow the scanning window of precursor ions to 0.2 m/z, which aids selectivity and accuracy.[105] These narrow windows make it possible to analyse many different signature peptides in one single LC/MRM run[66] as demonstrated by Picotti et al.[93]

A great advantage of MRM is that it allows the precise quantitative determination of the target peptides in complex samples.[12, 57, 93] MRM also offers a broad dynamic range (up to five magnitudes), which is essential for the quantification of allergens with highly variable concentrations.[93] Moreover, the high sensitivity of MRM allows the precise quantification of individual proteins, as well as different isoforms.[55] This is an advantage, because allergens, e.g. parvalbumin, have different isoforms and cannot be easily detected by antibody based methods. Since the quantification of proteins by MRM is a relatively new technique, the development of methods has to be carefully investigated, designed and validated[12, 57] as illustrated in figure 1.5. Selecting an appropriate tryptic signature peptide for MRM is the most critical step.[12, 57,72] The ionisation of the selected peptides must be complete and distinguishable from the matrix.[57, 73] The quantification can be based on a single peptide, if specifically unique to the target protein, however, two to three peptides per proteins are preferred to achieve better specificity.[12, 57,72]

1.10.1 Standards used for absolute allergen quantification by MRM

Well characterised standards are required to ensure that the correct peptide, matching the target allergen, is quantified. With an appropriate standard, LC/MRM parameters can be optimised to achieve sensitivity of the target peptide.[66, 93] These signature peptides can be generated from a recombinant allergen, synthetic peptides or identical isotopic labelled peptides.[19, 66, 71, 72] MRM approaches quantifying small molecules have most commonly used isotopic labelled standards.[12, 66, 71, 72, 93] General characteristics that an internal standard for

LC/MRM quantifications should fulfil are summarised in figure 1.6. In detail, these standards should be (1) unique to relevant analyte (2) present in matrix (3) stable and robust (4) easy to be detected by LC/MRM (5) have good ionisation qualities and fragmentation behaviour (6) should be available in databases (7) should be easily digested, thus not missed tryptic cleavage sites (for instance sequences like KRXX or XXRK should be avoided) (8) should not interfere with sample matrix (9) should not interfere with naturally occurring products (10) be absent of chemical modifications such as cystein (thiol group) methionine (oxidable) asparagine and glutamine (deamination and N-terminal glutamine pyroglutamate) (11) and be absent of PTMs.[42, 67, 75, 82]

Different standards are used for the MRM quantification in literature and will be described in the following paragraphs. Johnson et al.[78] suggest that isotopic labelled peptides are a realistic alternative for the development of reference methods,[78] although the best standards are isotopic labelled proteins. However, these proteins are prohibitively expensive. The isotopic labelled peptides have been used successfully for protein quantification over the last two decades.[67, 90] For example, Kitteringham et al.[66] and Meng and Veenstra[71] suggest using isotopic labelled standards, as they match characteristics of the original peptide.[67, 91] These characteristics include the physiochemical and chromatographic performance as well as ionisation efficiancy.[57, 71, 91] The isotopic labelled peptides have a slightly different mass-to-charge ratio, generated from the elemental isotope label used, and can therefore be distinguished from the natural peptide.[81, 91] The mass difference between labelled and natural peptide should be between 5-6 Daltons (Da). At least three to four transitions should be selected per peptide.[57, 71] The major disadvantage of isotopic labelled standards remains their cost.

Alternatives to isotopic labelled standards have been investigated to reduce the expense of the method development. Synthetic peptides without an isotopic label are one option. Rauh[57] suggests optimising the MRM parameters with synthetic peptides or crude peptides, as all LC and MS parameters will match the native protein, as explained above. As the synthetic peptide is not distinguishable from the native peptide, the absolute quantification is still successful.[57, 81, 93, 107]

Sealey-Voyksner et al.[68] considered the use of isotopic label standards in their study. The recovery, the accuracy and the injection to injection precision of various

control samples demonstrated that the isotopic labelled standards were not necessary. Pak et al.[107] also found the accuracy and the reproducibly via spiking samples with the synthetic standards was suitable. This study also found an acceptable dynamic range by qTOF and IT. However, the LC/MRM methods using unlabeled spikes need to be more consistent and reliable to be able to achieve absolute quantification.[55]

Another alternative is the use of recombinant proteins as they are generally well defined and better characterised, molecularly and chemically. The recombinant proteins can be produced in large amounts and are easier to purify, compared to traditional native protein extract purifications.[12, 19] The recombinant allergen proteins are also useful to investigate digestion efficiency, losses during the sample preparation, the peptides generation, the LC retention time and the transitions times.[12, 71]

Parameters optimised with recombinant proteins are equivalent to the native allergens.[57, 71] The recombinant proteins can be generated from different species, which may differ only by a few amino acids, and therefore may assist in characterising species specific allergens.[71] As recombinant proteins are usually expressed in *Escherichia coli*, they are, however, not suitable for the determination of PTMs.[71] While recombinant proteins are ideal for the development of precise and accurate methods by LC/MRM,[57, 67, 72] only few highly purified recombinant proteins are currently available.[12]

In summary, the MRM quantification requires standards, to confirm the reproducibility and the accuracy of the target molecules. The standards used most commonly are isotopic labelled peptides. When establishing the MRM methods for absolute quantification of allergens the use of both, recombinant proteins and isotopic labelled standards, would be best. Once the method is established, optimised and validated synthetic peptides without isotopic labels can replace the labelled standards.

1.10.2 Operating MRM for absolute allergen quantification

Many biomarkers are published in the allergy field, but only one per year is actually well characterised and methods fully validated.[71] To establish, evaluate

and validate new MRM methods several critical steps should be considered. A brief workflow is shown in figure 1.5. Firstly, signature peptides best fitting the selection criteria should be identified (Figure 1.6). Optimisation of the LC/MRM parameters should be performed as described above. Each MRM transition should be optimised by MS to achieve maximum sensitivity.[55, 57, 72] Secondly, samples should be spiked with synthetic or isotopic labelled peptides to demonstrate the intra-laboratory and inter-laboratory method the reproducibility and the precision. All sample preparation steps, precursor and product ion generation must be proven to be stable and reproducible. Synthetic or isotopic labelled standards should be spiked into the matrix before and after sample preparation to identify losses in the extraction process and obtain recovery information. The absence of peptides in blank samples should also be demonstrated. Thirdly, using recombinant protein, repeated digestions should be carried out to demonstrate that the same signature peptides can be identified consistently and losses during the sample preparation evaluated. Finally, the methods reproducibility, precision, linearity, accuracy and recoveries should be calculated.[68, 72, 91]

In MRM mode, peptides will be fragmented, generating mainly y-ions and b-ions. Sealey-Voyksner et al.[68] also demonstrated that cleavage sites between amino acid F (phenylalanine) and P (proline) exhibited different product ion fragment patterns. The fragment spectra generated provides essential information about the amino acid sequence of the peptide and the protein source.[57] Since fragmentation can vary from instrument to instrument, it is important to optimise each peptide individually to determine the best balance of signal-to-noise ratio, transition time and collision energy.[71, 72]

At least one, preferably three signature peptides, should be selected per protein. These peptides should not interfere with other peptides or matrix compounds.[57, 71, 72] For each of these three peptides, the precursor ions need to be selected. For each precursor ion at least two, preferable three product ions should be selected.[57, 71] In total per protein at least six, preferably nine MRM transition will be chosen. The more transitions are chosen the more specific the MRM quantification will be for the target protein. However, too many transitions will lead to a loss of sensitivity.[57, 71]

If the amino acid sequence of the detected peptide is specific to the target protein, then the MRM transition is also specific. The precursor peptides chosen should not have too many charge states and should be double or at the most triple charge.[73, 78] The higher the charges are on peptides the more transitions should be chosen to quantify precursor ion accurately.[66] The product ions chosen for MRM transition time should also have a higher m/z than the precursor ion to guarantee peptide specificity.[73] Sealey-Voyksner et al.[68] reported that a longer dwell time improved the signal-to-noise ratio by two to three fold. Generally, the larger the investigated allergen, the easier false positive transition time and fragments can occur.[66] Peptides used as standards must therefore be reproducible after digestion and have very small LC retention time windows.[66, 68]

In summary, many factors have to be considered when establishing and validating a new MRM method for the absolute allergen quantification. Most importantly the signature peptide to be used as standard in the LC/MRM method must be carefully selected. With accurate standards the LC/MRM methods are, in contrast to antibody based methods, reproducible, reliable, robust, accurate and sensitive for different food allergen quantification. Most importantly the results acquired by LC/MRM are comparable between methods and laboratories.

1.11 LOD/LOQ identified by MS systems

The first study to use synthetic peptide for allergen analysis by MRM system was in 2005 by Molle and Leonil.[94] However, LOD and LOQ data was not published. Monaci et al.[95] used milk allergen standards and demonstrated a LOD of 1 ppm using LC/qTOF. The LOD for spiked wine samples was 5 ppm, thus Monaci et al.[95] recommend that MRM sensitivity can be increased and warrants further investigation. Nevertheless, LOD and LOQ derived by qTOF and IT are comparable with reported ELISA values.

Monaci and van Hengel[96] demonstrated MRM to be almost ten times more sensitive compared to a UV detector in a LC method for milk allergen detection. When comparing identical preparation and digestion by LC/qTOF and LC/MRM Shefcheck et al.[79] showed that sensitivity of the latter was 10 fold higher compared to LC/qTOF. Molle and Leonil.[94] identified and quantified by MRM total casein

macropeptide variant A, variant B and aglyco- casein macropeptide in different dairy samples with good sensitivity and accuracy. These peptides originated from κ-casein with a LOQ of 10 picomolar. Subsequently in 2008 Monaci and van Hengel[96] analysed milk allergen spiked samples, achieving an LOD and LOQ of 1 ppm and 4 ppm, respectively. Careri et al.[100] calculated LOD and LOQ for Ara h 2 to be 5 ppm and 13 ppm. In the same study LOD and LOQ for Ara h 3/4 was about 1 ppm and 3 ppm. Abdel Rahman, et al.[34, 91] reported a crustacean tropomyosin and argine kinase LOQ of 3 and 0.25 nanomolar. Sealey-Voyksner et al.[68] was able to detect trace levels of cereal allergens by LC/MRM with an LOD and LOQ of 0.01-0.03 ppm and 0.01-0.1 ppm, respectively. These studies demonstrate that MRM is comparable and can be more sensitive than allergen detection by ELISA.

Studies quantifying allergens in the past five years using MS systems, especially MRM, have confirmed that low LOD and LOQ can be easily achieved. MRM can detect peptides in a low femtomolar and attomolar range equivalent to low ppm and ppb range, calculated by ELISA. Moreover, LC/MRM methods are easy to compare and standardise and therefore validated LC/MRM methods should be used for absolute allergen quantification.

1.12 Concluding remarks

The prevalence of food allergies are increasing worldwide and therefore represents a growing public health concern. Governments protect allergic consumers by regulating the labelling of food products containing potential allergens. Currently more than 600 different food allergens are known which demonstrates the variety of existing allergens and the difficultly in allergen analysis and subsequent food labelling. However, detecting and quantifying food allergens remains problematic. The analysis is complicated by complex food matrices, multiple allergens and different food sources. Moreover, a lack of standardised analytical methods means legislation cannot be properly policed. To date the most common quantitative methods for allergen analysis is the ELISA. However, ELISA methods have several drawbacks. Furthermore, for shellfish allergen detection, currently no method is available to distinguish crustacean and molluscs as required by legislation in numerous countries. Therefore the development of new

methods for the quantification of food allergens is suggested, which are robust, reliable, comparable, stable, sensitive and easy to standardise. MRM fulfils all these factors and in addition demonstrates similar LOD and LOQ compared to ELISA, which makes MRM the method of choice for absolute allergen quantification. Well characterised standards are required for all MS based methods, which can be used as reference materials for intra-laboratory and inter-laboratory comparison.

1.13 Summary and research aims of thesis

As explained in this **chapter 1**, shellfish involve species of crustacean and molluscs. The shellfish consumption and exposure is increasing worldwide and therefore prevalence of allergy towards shellfish and allergic reactions for individuals are becoming public health concern. The food labelling legislation in Canada and the European Union requires a different labelling for allergens from crustacean and molluscs. However, current available methods for the detection of shellfish allergens cannot distinguish between those two groups. The existing commercial and non commercial detection methods for shellfish allergens are based on tropomyosin, the major allergen in shellfish. Mass spectrometry, as explained in detail in this **chapter 1**, in comparison to biological methods, can successfully identify and quantify allergens, while being selective, sensitive, robust and reliable.

To establish a detection method for shellfish allergen quantification, the **aim of chapter 2** is to investigate a commercially important Australian crustacean species, King prawn (*Melicertus latisulcatus*). The novel investigated tropomyosin sequence from King prawn will be compared to the well investigated related prawn species, Black Tiger prawn. Moreover, recombinant tropomyosin will be expressed, and two different expression systems compared for higher yield of expressed recombinant allergens.

In **chapter 3**, the chemical method LC/qTOF (liquid chromatography coupled with a quantitative time of flight mass spectrometer) will be applied to identify a signature peptides that are unique for crustacean or unique for molluscs, hence these two groups can be distinguished. The signature peptides will be identified

using the protein tropomyosin, the major allergen in shellfish. These signature peptides will be further utilised in **chapter 5** for the development and validation of a sensitive robust chemical method (liquid chromatography coupled with a multiple reaction monitoring mass spectrometer (LC/MRM)) which is able to quantify allergenic crustacean tropomyosin as well as allergenic mollusc tropomyosin.

In **chapter 4**, the chemical method (LC/qTOF) will be used to identify possible other shellfish allergen candidates in twenty-two shellfish species commonly consumed in Australia. Moreover, the impact of temperature (heating/cooking) on the protein profile of crustacean and molluscs will be investigated.

Finally, the chemical methods (LC/MRM and LC/qTOF) will be compared with two commercial ELISA kits available in Australia in **chapter 6**. Food samples will be analysed and compared using these two chemical and one biological method.

The work presented in this thesis provides an important contribution towards the detection and quantification of allergenic tropomyosin from crustacean and molluscs, to fulfil the international legislation requirements.

The validated method for the qualitative and quantitative detection of TM in food products will lead to a standardised analytical detection. Therefore food labelling can be controlled, thus leading to a better protection for allergic individuals. The chemical and molecular knowledge of more shellfish species will help to understand cross-reactivity between shellfish and related allergens better. Thus, diagnosis and treatment of patients will be improved.

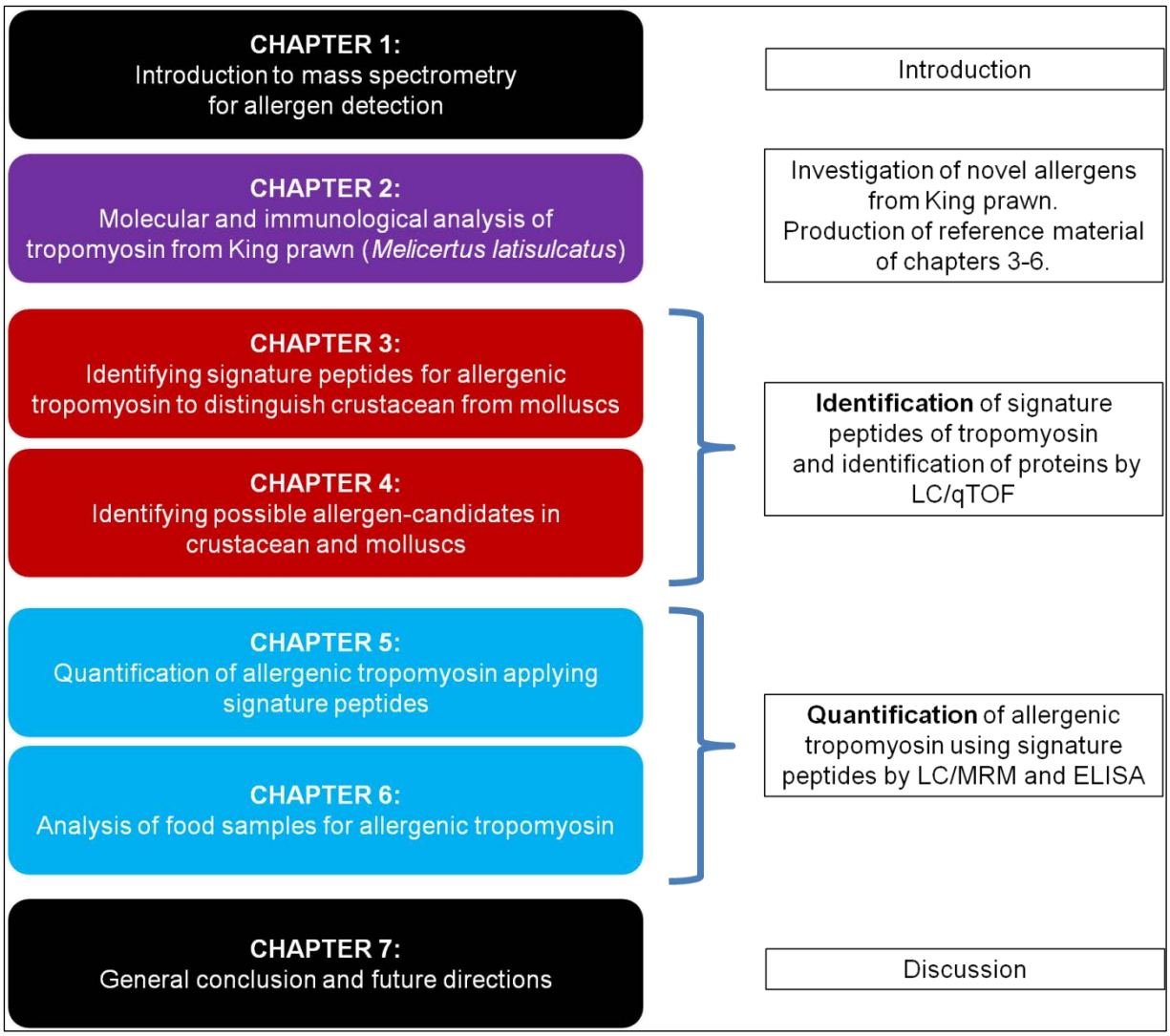


Figure 1.8: Flowchart of thesis chapters.

1.14 References

1. Pawankar, R.; Canonica, G. W.; Holgate, S. T.; Lockey, R. F., *Wao White Book on Allergy*. 2011.

2. Lopata, A. L.; Lehrer, S. B. (2009). *New Insights into Seafood Allergy.* Current Opinion in Allergy and Clinical Immunology. 9, (3), 270-277.

3. Ortolani, C.; Pastorello, E. A. (2006). *Food Allergies and Food Intolerances.* Best Practice & Research in Clinical Gastroenterology. 20, (3), 467-483.

4. Sharp, M.; Lopata, A. (2013). *Fish Allergy: In Review.* Clinical Reviews in Allergy & Immunology. 1-14.

5. Untersmayr, E.; Jensen-Jarolim, E. (2006). *Mechanisms of Type I Food Allergy.* Pharmacology & Therapeutics. 112, (3), 787-798.

6. Lopata, A. L.; O'Hehir, R. E.; Lehrer, S. B. (2010). *Shellfish Allergy.* Clinical and Experimental Allergy. 40, (6), 850-858.

7. Sicherer, S. H.; Sampson, H. A. (2010). *Food Allergy.* Journal of Allergy and Clinical Immunology. 125, (2), S116-S125.

8. Sicherer, S. H. (2011). *Epidemiology of Food Allergy.* Journal of Allergy and Clinical Immunology. 127, (3), 594-602.

9. Hill, D. J.; Hosking, C. S.; Zhie, C. Y.; Leung, R.; Baratwidjaja, K.; Iikura, Y.; Iyngkaran, N.; Gonzalez-Andaya, A.; Wah, L. B.; Hsieh, K. H. (1997). *The Frequency of Food Allergy in Australia and Asia.* Environmental Toxicology and Pharmacology. 4, (1-2), 101-110.

10. Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). *Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems.* Journal of Proteome Research. 13, (8), 3499-3509.

11. Wulf Becker, F. B., Daniel Brasseur, Jean-Louis Bresson, Albert Flynn, Alan A. Jackson, Pagona Lagiou, Martinus Løvik, Geltrude Mingrone, Bevan Moseley, Andreu Palou, Hildegard Przyrembel, Seppo Salminen, Stephan Strobel, Henk van den Berg, and Hendrik van Loveren. (2007). *Opinion of the Scientific Panel on Dietetic Products, Nutrition and Allergies (Nda) Related to the Evaluation of Molluscs for Labelling Purposes.* EFSA Journal. 4 [2006], (3), 1-25.

12. Picariello, G.; Mamone, G.; Addeo, F.; Ferranti, P. (2011). *The Frontiers of Mass Spectrometry-Based Techniques in Food Allergenomics.* Journal of Chromatography A. 1218, (42), 7386-7398.

13. Breiteneder, H.; Mills, E. N. C. (2005). *Molecular Properties of Food Allergens.* Journal of Allergy and Clinical Immunology. 115, (1), 14-23.

14. Jenkins, J. A.; Breiteneder, H.; Mills, E. N. C. (2007). *Evolutionary Distance from Human Homologs Reflects Allergenicity of Animal Food Proteins.* Journal of Allergy and Clinical Immunology. 120, (6), 1399-1405.

15. Radauer, C.; Bublin, M.; Wagner, S.; Mari, A.; Breiteneder, H. (2008). *Allergens Are Distributed into Few Protein Families and Possess a Restricted Number of Biochemical Functions.* Journal of Allergy and Clinical Immunology. 121, (4), 847-852.

16. Breiteneder, H. (2008). *Can Any Protein Become an Allergen?* Revue Francaise D Allergologie Et D Immunologie Clinique. 48, (3), 135-138.

17. Pomes, A. (2008). *Common Structures of Allergens.* Revue Francaise D Allergologie Et D Immunologie Clinique. 48, (3), 139-142.

18. Breiteneder, H.; Mills, E. N. C. (2005). *Plant Food Allergens—Structural and Functional Aspects of Allergenicity.* Biotechnology Advances. 23, (6), 395-399.

19. Chapman, M. D.; Smith, A. M.; Vailes, L. D.; Arruda, L. K.; Dhanaraj, V.; Pomes, A. (2001). *Recombinant Allergens for Diagnosis and Therapy of Allergic Disease.* Clinical and Experimental Allergy. 31, (1), 164-164.

20. Reese, G.; Ayuso, R.; Lehrer, S. B. (1999). *Tropomyosin: An Invertebrate Pan–Allergen.* International Archives of Allergy and Immunology. 119, (4), 247-258.

21. Li, X.; Holmes, K. C.; Lehman, W.; Jung, H.; Fischer, S. (2010). *The Shape and Flexibility of Tropomyosin Coiled Coils: Implications for Actin Filament Assembly and Regulation.* Journal of Molecular Biology. 395, (2), 327-339.

22. Emoto, A.; Ishizaki, S.; Shiomi, K. (2009). *Tropomyosins in Gastropods and Bivalves: Identification as Major Allergens and Amino Acid Sequence Features.* Food Chemistry. 114, (2), 634-641.

23. Kamath, S. D.; Abdel Rahman, A. M.; Komoda, T.; Lopata, A. L. (2013). *Impact of Heat Processing on the Detection of the Major Shellfish Allergen Tropomyosin in Crustaceans and Molluscs Using Specific Monoclonal Antibodies.* Food Chemistry. 141, (4), 4031-9.

24. Mills, E. N. C.; Sancho, A. I.; Rigby, N. M.; Jenkins, J. A.; Mackie, A. R. (2009). *Impact of Food Processing on the Structural and Allergenic Properties of Food Allergens.* Molecular Nutrition & Food Research. 53, (8), 963-969.

25. Motoyama, K.; Suma, Y.; Ishizaki, S.; Nagashima, Y.; Lu, Y.; Ushio, H.; Shiomi, K. (2008). *Identification of Tropomyosins as Major Allergens in Antarctic Krill and Mantis Shrimp and Their Amino Acid Sequence Characteristics.* Marine Biotechnology. 10, (6), 709-718.

26. Ayuso, R.; Lehrer, S. B.; Reese, G. (2002). *Identification of Continuous, Allergenic Regions of the Major Shrimp Allergen Pen a 1 (Tropomyosin).* International Archives of Allergy and Immunology. 127, (1), 27-37.

27. Ayuso, R.; Reese, G.; Leong-Kee, S.; Plante, M.; Lehrer, S. B. (2002). *Molecular Basis of Arthropod Cross-Reactivity: Ige-Binding Cross-Reactive Epitopes of Shrimp, House Dust Mite and Cockroach Tropomyosins.* International Archives of Allergy and Immunology. 129, (1), 38-48.

28. Ayuso, R.; Sanchez-Garcia, S.; Pascal, M.; Lin, J.; Grishina, G.; Fu, Z.; Ibanez, M. D.; Sastre, J.; Sampson, H. A. (2012). *Is Epitope Recognition of Shrimp Allergens Useful to Predict Clinical Reactivity?* Clinical and Experimental Allergy. 42, (2), 293-304.

29. Reese, G.; Viebranz, J.; Leong-Kee, S. M.; Plante, M.; Lauer, I.; Randow, S.; Moncin, M. S. M.; Ayuso, R.; Lehrer, S. B.; Vieths, S. (2005). *Reduced Allergenic Potency of Vr9-1, a Mutant of the Major Shrimp Allergen Pen a 1 (Tropomyosin).* Journal of Immunology. 175, (12), 8354-8364.

30. Zheng, L.-N.; Lin, H.; Pawar, R.; Li, Z.-X.; Li, M.-H. (2011). *Mapping Ige Binding Epitopes of Major Shrimp (Penaeus Monodon) Allergen with Immunoinformatics Tools.* Food and Chemical Toxicology. 49, (11), 2954-2960.

31. Ishikawa, M.; Ishida, M.; Shimakura, K.; Nagashima, Y.; Shiomi, K. (1998). *Purification and Ige-Binding Epitopes of a Major Allergen in the Gastropod Turbo Cornutus.* Biosci Biotechnol Biochem. 62, (7), 1337-43.

32. Ishikawa, M.; Ishida, M.; Shimakura, K.; Nagashima, Y.; Shiomi, K. (1998). *Tropomyosin, the Major Oyster Crassostrea Gigas Allergen and Its Ige-Binding Epitopes.* Journal of Food Science. 63, (1), 44-47.

33. Ishikawa, M.; Suzuki, F.; Ishida, M.; Nagashima, Y.; Shiomi, K. (2001). *Identification of Tropomyosin as a Major Allergen in the Octopus Octopus Vulgaris and Elucidation of Its Ige-Binding Epitopes.* Fisheries Science. 67, (5), 934-942.

34. Abdel Rahman, A. M.; Kamath, S. D.; Gagne, S.; Lopata, A. L.; Helleur, R. (2013). *Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention.* Journal of Proteome Research. 12, (2), 647-656.

35. Khanaruksombat, S.; Srisomsap, C.; Chokchaichamnankit, D.; Punyarit, P.; Phiriyangkul, P. (2014). *Identification of a Novel Allergen from Muscle and Various Organs in Banana Shrimp (Fenneropenaeus Merguiensis).* Annals of Allergy, Asthma & Immunology. 113, (3), 301-306.

36. Kamath, S. D.; Abdel Rahman, A. M.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. 58, (5), 1144-1155.

37. Piboonpocanun, S.; Jirapongsananuruk, O.; Tipayanon, T.; Boonchoo, S.; Goodman, R. E. (2011). *Identification of Hemocyanin as a Novel Non-Cross-Reactive Allergen from the Giant Freshwater Shrimp Macrobrachium Rosenbergii.* Molecular Nutrition & Food Research. 55, (10), 1492-1498.

38. Shen, H. W.; Cao, M. J.; Cai, Q. F.; Ruan, M. M.; Mao, H. Y.; Su, W. J.; Liu, G. M. (2012). *Purification, Cloning, and Immunological Characterization of Arginine Kinase, a Novel Allergen of Octopus Fangsiao.* Journal of Agricultural and Food Chemistry. 60, (9), 2190-2199.

39. Suzuki, M.; Kobayashi, Y.; Hiraki, Y.; Nakata, H.; Shiomi, K. (2011). *Paramyosin of the Disc Abalone Haliotis Discus Discus: Identification as a New Allergen and Cross-Reactivity with Tropomyosin.* Food Chemistry. 124, (3), 921-926.

40. Cheftel, J. C. (2005). *Food and Nutrition Labelling in the European Union.* Food Chemistry. 93, (3), 531-550.

41. Gendel, S. M. (2012). *Comparison of International Food Allergen Labeling Regulations.* Regulatory toxicology and pharmacology : RTP. 63, (2), 279-85.

42. Monaci, L.; Visconti, A. (2009). *Mass Spectrometry-Based Proteomics Methods for Analysis of Food Allergens.* Trac-Trends in Analytical Chemistry. 28, (5), 581-591.

43. Goodwin, P. R. (2004). *Food Allergen Detection Methods: A Coordinated Approach.* Journal of Aoac International. 87, (6), 1383-1390.

44. Lepski, S.; Brockmeyer, J. (2013). *Impact of Dietary Factors and Food Processing on Food Allergy.* Molecular Nutrition & Food Research. 57, (1), 145-152.

45. Wal, J. M. (2003). *Thermal Processing and Allergenicity of Foods.* Allergy. 58, (8), 727-729.

46. Thomas, K.; Herouet-Guicheney, C.; Ladics, G.; Bannon, G.; Cockburn, A.; Crevel, R.; Fitzpatrick, J.; Mills, C.; Privalle, L.; Vieths, S. (2007). *Evaluating the Effect of Food Processing on the Potential Human Allergenicity of Novel Proteins: International Workshop Report.* Food and Chemical Toxicology. 45, (7), 1116-1122.

47. Maleki, S. J.; Chung, S. Y.; Champagne, E. T.; Raufman, J. P. (2000). *The Effects of Roasting on the Allergenic Properties of Peanut Proteins.* Journal of Allergy and Clinical Immunology. 106, (4), 763-768.

48. Nakamura, A.; Watanabe, K.; Ojima, T.; Ahn, D.-H.; Saeki, H. (2005). *Effect of Maillard Reaction on Allergenicity of Scallop Tropomyosin.* Journal of Agricultural and Food Chemistry. 53, (19), 7559-7564.

49. Weber, D.; Raymond, P.; Ben-Rejeb, S.; Lau, B. (2006). *Development of a Liquid Chromatography-Tandem Mass Spectrometry Method Using Capillary Liquid Chromatography and Nanoelectrospray Ionization-Quadrupole Time-of-Flight Hybrid Mass Spectrometer for the Detection of Milk Allergens.* Journal of Agricultural and Food Chemistry. 54, (5), 1604-1610.

50. Johnson, P. E.; Rigby, N. M.; Dainty, J. R.; Mackie, A. R.; Immer, U. U.; Rogers, A.; Titchener, P.; Shoji, M.; Ryan, A.; Mata, L.; Brown, H.; Holzhauser, T.; Dumont, V.; Wykes, J. A.; Walker, M.; Griffin, J.; White, J.; Taylor, G.; Popping, B.; Crevel, R.; Miguel, S.; Lutter, P.; Gaskin, F.; Koerner, T. B.; Clarke, D.; Sherlock, R.; Flanagan, A.; Chan, C.-H.; Mills, E. N. C. (2014). *A Multi-Laboratory Evaluation of a Clinically-Validated Incurred Quality Control Material for Analysis of Allergens in Food.* Food Chemistry. 148, (0), 30-36.

51. Sakai, S.; Matsuda, R.; Adachi, R.; Akiyama, H.; Maitani, T.; Ohno, Y.; Oka, M.; Abe, A.; Seiki, K.; Oda, H.; Shiomi, K.; Urisu, A. (2008). *Interlaboratory Evaluation of Two Enzyme-Linked Immunosorbent Assay Kits for the Determination of Crustacean Protein in Processed Foods.* Journal of Aoac International. 91, (1), 123-129.

52. Shefcheck, K. J.; Musser, S. M. (2004). *Confirmation of the Allergenic Peanut Protein, Ara H 1, in a Model Food Matrix Using Liquid Chromatography/Tandem Mass Spectrometry (Lc/Ms/Ms).* Journal of Agricultural and Food Chemistry. 52, (10), 2785-2790.

53. Tolin, S.; Pasini, G.; Simonato, B.; Mainente, F.; Arrigoni, G. (2012). *Analysis of Commercial Wines by Lc-Ms/Ms Reveals the Presence of Residual Milk and Egg White Allergens.* Food Control. 28, (2), 321-326.

54. Cucu, T.; De Meulenaer, B.; Kerkaert, B.; Vandenberghe, I.; Devreese, B. (2012). *Maldi Based Identification of Whey Protein Derived Tryptic Marker Peptides That Resist Protein Glycation.* Food Research International. 47, (1), 23-30.

55. Heick, J.; Fischer, M.; Kerbach, S.; Tamm, U.; Popping, B. (2011). *Application of a Liquid Chromatography Tandem Mass Spectrometry Method for the Simultaneous Detection of Seven Allergenic Foods in Flour and Bread and Comparison of the Method with Commercially Available Elisa Test Kits.* Journal of Aoac International. 94, (4), 1060-1068.

56. Cucu, T.; Platteau, C.; Taverniers, I.; Devreese, B.; de Loose, M.; de Meulenaer, B. (2011). *Elisa Detection of Hazelnut Proteins: Effect of Protein Glycation in the Presence or Absence of Wheat Proteins.* Food Additives and Contaminants Part a-Chemistry Analysis Control Exposure & Risk Assessment. 28, (1), 1-10.

57. Rauh, M. (2012). *Lc-Ms/Ms for Protein and Peptide Quantification in Clinical Chemistry.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences. 883, 59-67.

58. Monaci, L.; Visconti, A. (2010). *Immunochemical and DNA-Based Methods in Food Allergen Analysis and Quality Assurance Perspectives.* Trends in Food Science & Technology. 21, (6), 272-283.

59. Chassaigne, H.; Norgaard, J. V.; van Hengel, A. J. (2007). *Proteomics-Based Approach to Detect and Identify Major Allergens in Processed Peanuts by Capillary Lc-Q-Tof (Ms/Ms).* Journal of Agricultural and Food Chemistry. 55, (11), 4461-4473.

60. Careri, M.; Elviri, L.; Lagos, J. B.; Mangia, A.; Speroni, F.; Terenghi, M. (2008). *Selective and Rapid Immunomagnetic Bead-Based Sample Treatment for the Liquid Chromatography-Electrospray Ion-Trap Mass Spectrometry Detection of Ara H3/4 Peanut Protein in Foods.* Journal of Chromatography A. 1206, (2), 89-94.

61. Fuller, H. R.; Goodwin, P. R.; Morris, G. E. (2006). *An Enzyme-Linked Immunosorbent Assay (Elisa) for the Major Crustacean Allergen, Tropomyosin, in Food.* Food and Agricultural Immunology. 17, (1), 43-52.

62. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M. X.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* International Journal of Hygiene and Environmental Health. 217, (7), 740-750.

63. Seiki, K.; Oda, H.; Yoshioka, H.; Sakai, S.; Urisu, A.; Akiyama, H.; Ohno, Y. (2007). *A Reliable and Sensitive Immunoassay for the Determination of Crustacean Protein in Processed Foods.* Journal of Agricultural and Food Chemistry. 55, (23), 9345-9350.

64. Werner, M. T.; Faeste, C. K.; Egaas, E. (2007). *Quantitative Sandwich Elisa for the Determination of Tropomyosin from Crustaceans in Foods.* Journal of Agricultural and Food Chemistry. 55, (20), 8025-8032.

65. Zhang, H.; Lu, Y.; Ushio, H.; Shiomi, K. (2014). *Development of Sandwich Elisa for Detection and Quantification of Invertebrate Major Allergen Tropomyosin by a Monoclonal Antibody.* Food Chemistry. 150, (0), 151-157.

66. Kitteringham, N. R.; Jenkins, R. E.; Lane, C. S.; Elliott, V. L.; Park, B. K. (2009). *Multiple Reaction Monitoring for Quantitative Biomarker Analysis in Proteomics and Metabolomics.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences. 877, (13), 1229-1239.

67. Sancho, A. I.; Mills, E. N. C. (2010). *Proteomic Approaches for Qualitative and Quantitative Characterisation of Food Allergens.* Regulatory Toxicology and Pharmacology. 58, (3), S42-S46.

68. Sealey-Voyksner, J. A.; Khosla, C.; Voyksner, R. D.; Jorgenson, J. W. (2010). *Novel Aspects of Quantitation of Immunogenic Wheat Gluten Peptides by Liquid Chromatography-Mass Spectrometry/Mass Spectrometry.* Journal of Chromatography A. 1217, (25), 4167-4183.

69. Akagawa, M.; Handoyo, T.; Ishii, T.; Kumazawa, S.; Morita, N.; Suyama, K. (2007). *Proteomic Analysis of Wheat Flour Allergens.* Journal of Agricultural and Food Chemistry. 55, (17), 6863-6870.

70. Sotkovsky, P.; Hubalek, M.; Hernychova, L.; Novak, P.; Havranova, M.; Setinova, I.; Kitanovicova, A.; Fuchs, M.; Stulik, J.; Tuckova, L. (2008). *Proteomic Analysis of Wheat Proteins Recognized by Ige Antibodies of Allergic Patients.* Proteomics. 8, (8), 1677-1691.

71. Meng, Z. J.; Veenstra, T. D. (2011). *Targeted Mass Spectrometry Approaches for Protein Biomarker Verification.* Journal of Proteomics. 74, (12), 2650-2659.

72. Simon, R.; Jubeaux, G.; Chaumot, A.; Lemoine, J.; Geffard, O.; Salvador, A. (2010). *Mass Spectrometry Assay as an Alternative to the Enzyme-Linked Immunosorbent Assay Test for Biomarker Quantitation in Ecotoxicology: Application to Vitellogenin in Crustacea (Gammarus Fossarum).* Journal of Chromatography A. 1217, (31), 5109-5115.

73. Heick, J.; Fischer, M.; Popping, B. (2011). *First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry.* Journal of Chromatography A. 1218, (7), 938-943.

74. Lee, J. Y.; Kim, C. J. (2010). *Determination of Allergenic Egg Proteins in Food by Protein-, Mass Spectrometry-, and DNA-Based Methods.* Journal of Aoac International. 93, (2), 462-477.

75. Abdel Rahman, A. M.; Lopata, A. L.; O'Hehir, R. E.; Robinson, J. J.; Banoub, J. H.; Helleur, R. J. (2010). *Characterization and De Novo Sequencing of Snow Crab Tropomyosin Enzymatic Peptides by Both Electrospary Ionization and Matrix-Assisted Laser Desorption Ionization Qqtof Tandem Mass Spectrometry.* Journal of Mass Spectrometry. 45, (4), 372-381.

76. Contreras, M. d. M.; Lopez-Exposito, I.; Hernandez-Ledesma, B.; Ramos, M.; Recio, I. (2008). *Application of Mass Spectrometry to the Characterization and Quantification of Food-Derived Bioactive Peptides.* Journal of Aoac International. 91, (4), 981-94.

77. Lee, W. C.; Lee, K. H. (2004). *Applications of Affinity Chromatography in Proteomics.* Analytical Biochemistry. 324, (1), 1-10.

78. Johnson, P. E.; Baumgartner, S.; Aldick, T.; Bessant, C.; Giosafatto, V.; Heick, J.; Mamone, G.; O'Connor, G.; Poms, R.; Popping, B.; Reuter, A.; Ulberth, F.; Watson, A.; Monaci, L.; Mills, E. N. C. (2011). *Current Perspectives and Recommendations for the Development of Mass Spectrometry Methods for the Determination of Allergens in Foods.* Journal of Aoac International. 94, (4), 1026-1033.

79. Shefcheck, K. J.; Callahan, J. H.; Musser, S. M. (2006). *Confirmation of Peanut Protein Using Peptide Markers in Dark Chocolate Using Liquid Chromatography-Tandem Mass Spectrometry (Lc-Ms/Ms).* Journal of Agricultural and Food Chemistry. 54, (21), 7953-7959.

80. Carrera, M.; Canas, B.; Vazquez, J.; Gallardo, J. M. (2010). *Extensive De Novo Sequencing of New Parvalbumin Isoforms Using a Novel Combination of Bottom-up Proteomics, Accurate Molecular Mass Measurement by Fticr-Ms, and Selected Ms/Ms Ion Monitoring.* Journal of Proteome Research. 9, (9), 4393-4406.

81. DeSouza, L. V.; Siu, K. W. M. (2013). *Mass Spectrometry-Based Quantification.* Clinical Biochemistry. 46, (6), 421-431.

82. Abdel Rahman, A. M.; Kamath, S. D.; Lopata, A. L.; Robinson, J. J.; Helleur, R. J. (2011). *Biomolecular Characterization of Allergenic Proteins in Snow Crab (Chionoecetes Opilio) and De Novo Sequencing of the Second Allergen Arginine Kinase Using Tandem Mass Spectrometry.* Journal of Proteomics. 74, (2), 231-241.

83. Ortea, I.; Canas, B.; Calo-Mata, P.; Barros-Velazquez, J.; Gallardo, J. M. (2009). *Arginine Kinase Peptide Mass Fingerprinting as a Proteomic Approach for Species Identification and Taxonomic Analysis of Commercially Relevant Shrimp Species.* Journal of Agricultural and Food Chemistry. 57, (13), 5665-5672.

84. Ortea, I.; Canas, B.; Gallardo, J. M. (2011). *Selected Tandem Mass Spectrometry Ion Monitoring for the Fast Identification of Seafood Species.* Journal of Chromatography A. 1218, (28), 4445-4451.

85. Carrera, M.; Canas, B.; Gallardo, J. M. (2012). *Rapid Direct Detection of the Major Fish Allergen, Parvalbumin, by Selected Ms/Ms Ion Monitoring Mass Spectrometry.* Journal of Proteomics. 75, (11), 3211-3220.

86. Houston, N. L.; Lee, D.-G.; Stevenson, S. E.; Ladics, G. S.; Bannon, G. A.; McClain, S.; Privalle, L.; Stagg, N.; Herouet-Guicheney, C.; MacIntosh, S. C.; Thelen, J. J. (2010). *Quantitation of Soybean Allergens Using Tandem Mass Spectrometry.* Journal of Proteome Research. 10, (2), 763-773.

87. Ayuso, R.; Grishina, G.; Bardina, L.; Carrillo, T.; Blanco, C.; Ibanez, M. D.; Sampson, H. A.; Beyer, K. (2008). *Myosin Light Chain Is a Novel Shrimp Allergen, Lit V 3.* Journal of Allergy and Clinical Immunology. 122, (4), 795-802.

88. Ayuso, R.; Grishina, G.; Ibanez, M. D.; Blanco, C.; Carrillo, T.; Bencharitiwong, R.; Sanchez, S.; Nowak-Wegrzyn, A.; Sampson, H. A. (2009). *Sarcoplasmic Calcium-Binding Protein Is an Ef-Hand-Type Protein Identified as a New Shrimp Allergen.* Journal of Allergy and Clinical Immunology. 124, (1), 114-120.

89. Abdel Rahman, A. M.; Helleur, R. J.; Jeebhay, M. F.; Lopata, A. L., Characterization of Seafood Proteins Causing Allergic Diseases. In *Allergic Diseases - Highlights in the Clinic, Mechanisms and Treatment* Pereira, P. C., Ed. 2012; pp 107-140.

90. Abdel Rahman, A. M.; Kamath, S.; Lopata, A. L.; Helleur, R. J. (2010). *Analysis of the Allergenic Proteins in Black Tiger Prawn (Penaeus Monodon) and Characterization of the Major Allergen Tropomyosin Using Mass Spectrometry.* Rapid Communications in Mass Spectrometry. 24, (24), 3624-3624.

91. Abdel Rahman, A. M.; Lopata, A. L.; Randell, E. W.; Helleur, R. J. (2010). *Absolute Quantification Method and Validation of Airborne Snow Crab Allergen Tropomyosin Using Tandem Mass Spectrometry.* Analytica Chimica Acta. 681, (1-2), 49-55.

92. Azarnia, S.; Boye, J. I.; Mongeon, V.; Sabik, H. (2013). *Detection of Ovalbumin in Egg White, Whole Egg and Incurred Pasta Using Lc–Esi-Ms/Ms and Elisa.* Food Research International. 52, (2), 526-534.

93. Picotti, P.; Rinner, O.; Stallmach, R.; Dautel, F.; Farrah, T.; Domon, B.; Wenschuh, H.; Aebersold, R. (2010). *High-Throughput Generation of Selected Reaction-Monitoring Assays for Proteins and Proteomes.* Nature Methods. 7, (1), 43-U5.

94. Molle, D.; Leonil, J. (2005). *Quantitative Determination of Bovine Kappa-Casein Macropeptide in Dairy Products by Liquid Chromatography/Electrospray Coupled to Mass*

Spectrometry (Lc-Esi/Ms) and Liquid Chromatography/Electrospray Coupled to Tamdem Mass Spectrometry (Lc-Esi/Ms/Ms). International Dairy Journal. 15, (5), 419-428.

95. Monaci, L.; Losito, I.; Palmisano, F.; Visconti, A. (2010). *Identification of Allergenic Milk Proteins Markers in Fined White Wines by Capillary Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry.* Journal of Chromatography A. 1217, (26), 4300-4305.

96. Monaci, L.; van Hengel, A. J. (2008). *Development of a Method for the Quantification of Whey Allergen Traces in Mixed-Fruit Juices Based on Liquid Chromatography with Mass Spectrometric Detection.* Journal of Chromatography A. 1192, (1), 113-120.

97. Newsome, G. A.; Scholl, P. F. (2012). *Quantification of Allergenic Bovine Milk As1-Casein in Baked Goods Using an Intact 15n-Labeled Protein Internal Standard.* Journal of Agricultural and Food Chemistry. 61, (24), 5659-5668.

98. Zhang, J.; Lai, S.; Zhang, Y.; Huang, B.; Li, D.; Ren, Y. (2012). *Multiple Reaction Monitoring-Based Determination of Bovine A-Lactalbumin in Infant Formulas and Whey Protein Concentrates by Ultra-High Performance Liquid Chromatography–Tandem Mass Spectrometry Using Tryptic Signature Peptides and Synthetic Peptide Standards.* Analytica Chimica Acta. 727, (0), 47-53.

99. Amnuaycheewa, P.; de Mejia, E. G. (2010). *Purification, Characterisation, and Quantification of the Soy Allergen Profilin (Gly M 3) in Soy Products.* Food Chemistry. 119, (4), 1671-1680.

100. Careri, M.; Costa, A.; Elviri, L.; Lagos, J. B.; Mangia, A.; Terenghi, M.; Cereti, A.; Garoffo, L. P. (2007). *Use of Specific Peptide Biomarkers for Quantitative Confirmation of Hidden Allergenic Peanut Proteins Ara H 2 and Ara Ill 3/4 for Food Control by Liquid Chromatography-Tandem Mass Spectrometry.* Analytical and Bioanalytical Chemistry. 389, (6), 1901-1907.

101. Careri, M.; Elviri, L.; Maffini, M.; Mangia, A.; Mucchino, C.; Terenghi, M. (2008). *Determination of Peanut Allergens in Cereal-Chocolate-Based Snacks: Metal-Tag Inductively Coupled Plasma Mass Spectrometry Immunoassay Versus Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry.* Rapid Communications in Mass Spectrometry. 22, (6), 807-811.

102. Chassaigne, H.; Tregoat, V.; Norgaard, J. V.; Maleki, S. J.; van Hengel, A. J. (2009). *Resolution and Identification of Major Peanut Allergens Using a Combination of Fluorescence Two-Dimensional Differential Gel Electrophoresis, Western Blotting and Q-Tof Mass Spectrometry.* Journal of Proteomics. 72, (3), 511-526.

103. Cucu, T.; De Meulenaer, B.; Devreese, B. (2012). *Maldi Based Identification of Soybean Protein Markers - Possible Analytical Targets for Allergen Detection in Processed Foods.* Peptides. 33, (2), 187-196.

104. Moreno, F. J.; Jenkins, J. A.; Mellon, F. A.; Rigby, N. M.; Robertson, J. A.; Wellner, N.; Clare Mills, E. N. (2004). *Mass Spectrometry and Structural Characterization of 2s Albumin Isoforms from Brazil Nuts (Bertholletia Excelsa).* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 1698, (2), 175-186.

105. Gallien, S.; Duriez, E.; Demeure, K.; Domon, B. (2013). *Selectivity of Lc-Ms/Ms Analysis: Implication for Proteomics Experiments.* Journal of Proteomics. 81, 148-158.

106. Hebling, C. M.; McFarland, M. A.; Callahan, J. H.; Ross, M. M. (2012). *Global Proteomic Screening of Protein Allergens and Advanced Glycation Endproducts in Thermally Processed Peanuts.* Journal of Agricultural and Food Chemistry. 61, (24), 5638-5648.

107. Pak, H.; Pasquarello, C.; Scherl, A. (2011). *Label-Free Protein Quantifcation on Tandem Mass Spectra in an Ion Trapping Device.* Journal of Integrated Omics. 1, (2), 211-215.

108. Prandi, B.; Faccini, A.; Tedeschi, T.; Galaverna, G.; Sforza, S. (2013). *Lc/Ms Analysis of Proteolytic Peptides in Wheat Extracts for Determining the Content of the Allergen Amylase/Trypsin Inhibitor Cm3: Influence of Growing Area and Variety.* Food Chemistry. 140, (1–2), 141-146.

109. Uvackova, L.; Skultety, L.; Bekesova, S.; McClain, S.; Hajduch, M. (2013). *Mse Based Multiplex Protein Analysis Quantified Important Allergenic Proteins and Detected Relevant Peptides Carrying Known Epitopes in Wheat Grain Extracts.* Journal of Proteome Research.

CHAPTER 2

MOLECULAR AND IMMUNOLOGICAL ANALYSIS OF TROPOMYOSIN FROM KING PRAWN (*Melicertus latisulcatus*)

Data of this chapter are published in:

Koeberl, M.; Kamath, S. D.; Saptarshi, S. R.; Smout, M. J.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Auto-Induction for High Yield Expression of Recombinant Novel Isoallergen Tropomyosin from King Prawn (Melicertus Latisulcatus) for Improved Diagnostics and Immunotherapeutics.* ***Journal of Immunological Methods****.* 415, 6-16.

2.1 Introduction

Food allergies are increasing worldwide and therefore represent a growing public health concern. Consequently, the diagnosis of food allergy and the detection of allergens in food products are of increasing importance to protect consumers. Subsequently, the development of better assays to quantify allergens, the standardisation of the allergen preparations is of great importance. Recombinant allergenic proteins become therefore an essential tool for the detection of allergens,[1-3] the diagnosis of allergic sensitisation[4-9] and immunotherapeutic applications.[6, 10, 11] Moreover, the production of recombinant pan-allergens allows species specific investigation of allergens. Although similar species might have a similar genome, different allergens can be expressed as demonstrated for fish, house dust mite and shellfish.[6, 12, 13] Pan-allergens can cause different IgE binding patterns in sensitised patients. Therefore the investigation and expression of pan-allergens is necessary for better diagnostics and the development of immunotherapeutics.

Shellfish allergy is one of the major food allergies. Allergic reaction towards shellfish can range from mild reactions to life threatening anaphylaxis, affecting 2% of the world population.[14] Tropomyosin (TM) is the major allergen in shellfish and therefore frequently used in the detection and quantification of shellfish allergens in food products as well as in the diagnosis of allergic sensitisation towards shellfish. However, TM is a pan-allergen within shellfish and also cross-reactive with mites, cockroaches and nematodes.[14, 15] Minor allergens that have been reported for prawns are for example, myosin light chain, arginine kinase and sarcoplasmic calcium binding protein.[16] Different species from the shellfish group have been investigated for their allergenic properties. However, King prawn (*Melicertus latisulcatus*) has not been investigated yet. Recombinant tropomyosins from different species have been reported in the literature. Allergic tropomyosins investigated belong to five different arthropod groups. These groups include (1) prawns,[2, 17-21] (2) crabs,[13, 22, 23] (3) house dust mites,[24-27] (4) cockroaches[28-30] and (5) nematodes.[31-34] In most studies TM was expressed in *Escherichia coli* (*E. coli*) with the exception of German cockroach TM[28] and chicken TM,[35] being expressed in yeast.

Escherichia coli (*E. coli*) is a widely used expression system for the production of recombinant proteins, thus it has a well characterised genome, established techniques for genetic manipulation and it is easy and inexpensive to cultivate.[36-40] Historically, *E. coli* are divided into two major strains, the K-strain and the B-strain.[41] Both strains are commonly used in research and industry, while both strains seem to have advantages and disadvantages. Both strains grow similarly when cultured, however, the production of by-products (such as acetic acid, formic acid, and lactic acid) seems to differ[42, 43] and in addition the B-strains seem to grow faster in comparison to K-strains. This can lead to a higher biomass and therefore higher protein production.[42, 43] Consequently this would mean less unwanted by-products and possible inhibition of growth and expression of heterologous proteins. The disadvantages of B-strains is, however, that they can lose the plasmids.[42]

E. coli are mainly cultured in Luria–Bertani (LB) media at 37°C and expression is chemically induced using isopropyl-β-D-thiogalactoside (IPTG). In 2005, a different concept, so called auto-induction, was introduced by Studier.[44] Auto-induction uses a media in which glucose, glycerol and lactose are present at the same time. *E.coli* prefers glucose over lactose for growth, however, the consumption of glycerol and lactose follows when glucose is depleted. The auto-induction system is activated through the suppression of the *lac* operon by glucose. In turn, when glucose is metabolised, *E. coli* will utilise lactose and the *lac* operon will be activated. Therefore, if glucose and lactose are present in the media simultaneously, the induction will be delayed.[44]

The efficiency of expressing recombinant TM is only reported in a few publications. TM yields achieved are published for house dust mites,[24, 26] cockroach[28] and chicken.[35] Despite the different sources of TM and expression systems, the purified recombinant TM yields reported are relatively low and vary between 7.2 mg/l[28] and 26 mg/l.[26] However, larger quantities are needed as standards for the detection of allergens, diagnostics for allergic sensitisation and development of immunotherapeutics. Studier[44] reported that auto-induction can increase the protein yield by 10-fold over the IPTG induction, with up to 50 mg/l. To achieve this higher protein yield for allergens this chapter compares for the first time this novel auto-induction system with conventional IPTG induction to generate larger

quantities of the major shellfish allergen, tropomyosin. In this chapter the novel amino acid sequence of tropomyosin from King prawn is reported. This novel TM of King prawn was used to compare the expression in *E. coli* K-strains (NM522, TOP10) and B-strain (BL21(DE3)RIPL) as well as auto-induction versus IPTG induction. It is established for the first time a high cell density culture using ZY-5052 auto-inducing medium for allergen expression. Moreover, the molecular and immunological properties of recombinant tropomyosin from King prawn was compared to the recombinant tropomyosin of Black Tiger prawn (Pen m 1).

2.1.1 Aims

The aims of this chapter are:

- Purifying natural tropomyosin from King prawn
- cDNA sequencing of tropomyosin from King prawn
- Expressing recombinant tropomyosin from King prawn and comparing ITPG induction versus auto-induction as expression systems
- Comparing molecular and immunological characteristics from King prawn with the well characterised allergens from Black Tiger prawn

2.2 Materials and methods

2.2.1 Protein extraction from King prawn and Black Tiger prawn

Green headless King prawns (*Melicertus latisulcatus* (KP)) and raw Black Tiger prawns (*Penaeus monodon* (BTP)) were obtained from the local market (Townsville, Australia) and transported to the laboratory on ice. Two different extracts were made per species, named (1) “raw-extract” and (2) “whole-heated-extract”. For the raw extract the outer shell of the fresh prawn was removed and the abdominal muscles shredded into small pieces and homogenised, using an Ultra Turrax blender (IKA, Germany), in 200 ml of phosphate buffered saline (PBS) (Thermo Fisher Scientific, Australia). The total homogenised extract was kept at 4°C overnight, while continuously shaking. The protein extracts were centrifuged (3000 g for 15 minutes), followed by sterile filtration (0.22 µm) (Millipore, USA). For the whole heated extract the whole prawn was homogenised, with its outer shell, the whole heated extracted was heated in PBS at 100°C for 15 minutes and extracted over night. Both generated extracts were stored at -80°C until further use.

2.2.2 Protein estimation

Protein concentrations of different extracts and purified proteins were estimated using Pierce® 660 nm Protein Assay (Thermo Fisher Scientific, Australia). The protein concentration of protein extracts and purified proteins were estimated following the manufacturer’s instructions using bovine serum albumin (BSA) (Bio-Rad, Australia) as standard.

2.2.3 Purifying natural tropomyosin from King prawn

The tropomyosin (TM) from KP was purified from the whole heated extract using a strong anion-exchange chromatographic column (Bio-Rad, Australia) on a Biologic Duoflow system (Bio-Rad, USA). The mobile phase consisted of buffer A (30 mM sodium acetate buffer, pH 5.5) and buffer B (30 mM sodium acetate, 1 M NaCl, pH 5.5). Before loading the proteins onto the column, the protein extract were exchanged into the chromatographic starting buffer A using Amikon spin filters of 3

kDa molecular weight cut off (MWCO). After equilibrating the column with the buffer A, five mg of protein extract was loaded onto the column. The flow rate was set to 1.5 ml/min and the tropomyosin was purified using following gradient: 0-7 min 0% B; 7-17 min 0-35%B; 17-23 min 35-60%B; 23-33 min 60-100%B; 33-35 min 100%B; 35-40 min 0%B. Elution profile was generated using UV280 UV206 UV212 with Bio-Rad QuadTech UV spectrophotometer (Bio-Rad, USA). The collected fractions were exchanged and concentrated in PBS using a spin filters with a 3 kDa MWCO. The final purified natural tropomyosin from KP was stored at -80°C until further use.

Dr. Sandip Kamath[18] kindly provided natural purified tropomyosin from Black Tiger prawn for further studies.

2.2.4 CD spectroscopy for natural and recombinant tropomyosins

To compare the structures of tropomyosin from KP and BTP circular dichroism (CD) spectroscopy was performed. The natural purified and the generated recombinant tropomyosin (rTM) were prepared in PBS, pH 7.2 and adjusted to a final concentration of 100 µg/ml. CD spectroscopy was performed on a J715Spectropolarimeter (Jasco, USA) with continuous nitrogen flushing at 25°C. All measurements were performed using a 10 mm quartz-cuvette over a wavelength range of 190–260 nm. For wavelength analysis, the tropomyosin samples were scanned with a step width of 0.2 nm and bandwidth of 1 nm at 100 nm/min averaging over eight scans. Final data were expressed as mean residual ellipticity (θ) after subtracting the PBS blank spectrum.

Dr. Sandip Kamath[16] kindly provided recombinant tropomyosin from Black Tiger prawn for further studies.

2.2.5 SDS–PAGE analysis

Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) was performed to separate proteins on a 12% polyacrylamide gel. Proteins were diluted in 5x SDS–PAGE loading buffer containing dithiothreitol (DTT), heated for 5 min (100°C) and loaded onto a gel. The amount of protein loaded was for figure

2.2 two µg and for figure 2.10 twenty µg, respectively. Precision Plus protein standards (Bio-Rad, Australia) were used to estimate the molecular weights of individual proteins, using the Mini-PROTEAN® Tetra Cell (Bio-Rad, USA) system at 200 V. Proteins were visualised by Coomassie Brilliant Blue R-250 (Bio-Rad, Australia) staining. To estimate the expression level of TM in cell lysate, densitometric analysis was performed using the Quantity One 1-D analytical software (Bio-Rad, USA).

2.2.6 cDNA sequencing of King prawn tropomyosin

To investigate the amino acid sequence from TM of King prawn (*Melicertus latisulcatus* (KP)) cDNA was amplified and analysed. The green headless KP were obtained from the local market (Townsville, Australia) and transported to the laboratory on ice. The total RNA was isolated from KP meat using TRIzol® reagent (Life Technologies, Australia), with 0.1 g of muscle tissue in 1 ml of reagent. RNA extracted was quantified using the NanoDrop ND 1000 spectrophotometer (Thermo Fisher Scientific, USA) and the first strand cDNA synthesised using cDNA Synthesis kit (Bioline, Australia). The generated cDNA was used as a template to amplify the coding region of TM using forward (5`-GCGGATCCGACGCCATCAAGAAGAAGATGC-3´) and reverse (5`-GCGAATTCTTAGTAGCCAGACAGTTCGCTG-3´) primers. The PCR was run for 35 cycles, denaturation at 95°C for 0.45 minutes, annealing 55°C for 0.45 minutes and elongation at 72°C for 1 minute. The generated PCR products were purified on a low melting agarose gel. The amplified PCR product was cloned into the sequencing vector pCR2.1 using the TOPO® Cloning Reaction (Invitrogen, USA) protocol and transformed into chemical competent *E. coli.* cells. Positive colonies were screened by colony PCR using the gene specific oligonucleotide primers for the presence of inserts. The plasmid DNA was purified from overnight cultures using the QIAprep SPIN reaction KIT (Qiagen, Germany) and products analysed to generate the nucleotide sequence of tropomyosin from King prawn by Macrogen sequence analysis (Macrogen, South Korea). *E. coli.* cells were stored in glycerol stock at -80°C until further use.

The investigated cDNA sequence was converted into the amino acid sequence and published as Genbank accession number JX171685. To examine differences in amino acid sequence of TM from KP, the amino acid sequence of TM from KP was aligned with TM from Black Tiger prawn (*Penaeus monodon*), Genbank accession number HM486525.

2.2.7 cDNA sequencing of King prawn myosin light chain

To investigate the amino acid sequence from myosin light chain (MLC) of KP (*Melicertus latisulcatus*) cDNA was amplified and analysed. The same procedure was applied as described for the cDNA sequencing of TM (Section 2.2.6), with different primers and different PCR settings. The generated cDNA was used as a template to amplify the coding region of MLC using forward (5`-GCGGATCCTCCCGCAAGTCAGGCTCTCG-3´) and reverse (5`-GCGAATTCTTAGGCTTCCTCGGCGGCC-3´) primers. The PCR was run for 35 cycles, denaturation 94°C for 2 minutes, annealing 52.3°C for 0.45 minutes and elongation at 72°C for 1 minute. Different annealing temperatures were investigated with 52.2°C being chosen for further PCR amplifications. The plasmid DNA was analysed to generate the nucleotide sequence of MLC from King prawn by Macrogen sequence analysis (Macrogen, South Korea). *E. coli.* cells were stored in glycerol stock at -80°C.

To examine differences in amino acid sequence of MLC from KP, the amino acid sequence of MLC from KP was aligned with MLC from Black Tiger prawn (*Penaeus monodon*), Genbank accession number AET87131.1.

2.2.8 Expression of recombinant tropomyosin of King prawn

To express recombinant tropomyosin (rTM) the coding region for tropomyosin was cloned into the expression vector pProEX HTb (kindly provided by James Burnell), using the restriction enzymes, BamH1 and EcoR1 (Promega, USA). Ligation of the coding region into the expression vector was conducted using T4 DNA Ligase (Invitrogen, USA). The cloned vector was transformed into chemically competent *E. coli* cells; (1) NM522 (kindly provided by James Burnell), (2) BL21(DE3)RIPL

(kindly provided by Patrick Schaeffer) or (3) TOP10 (Invitrogen, USA), using heat shock and incubation in Super optimal broth with catabolite repression (SOC) medium at 37°C for 1 hour. The cells were grown overnight on LB agar with 100 µg/ml of ampicillin at 37°C. The colonies were tested for the presence of the insert by PCR. Plasmid and insert were confirmed by Macrogen sequence analysis (Macrogen, South Korea), using QIAprep SPIN reaction KIT (Qiagen, Germany). Cells were stored in glycerol stock at -80°C.

2.2.8.1 Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction

To investigate and optimise the expression of TM from KP in LB media with IPTG induction *E. coli* strains were grown over various hours (0 -24 hours) and soluble and insoluble fractions of bacteria culture analysed. In detail, *E. coli* cells were grown overnight in LB media with 100 µg/ml of ampicillin at 37°C (0.09 g). An aliquot of the overnight culture was transferred into LB broth containing 100 µg/ml ampicillin and grown until culture reached middle to late log phase. The expression of TM was inducted by adding 1 mM IPTG, subsequent 1 ml aliquots of cultures were analysed over 24 hours. These 1 ml aliquots were tested for the solubility of TM according to Studier.[44] Briefly, for testing TM in soluble fraction the 1 ml aliquots were centrifuged (5 min, 20000 g) and supernatant analysed by SDS–PAGE (10 µl loaded). The pellet of the insoluble fraction was resuspended in 5x SDS–PAGE loading buffer containing DTT (30 µl) and separated by SDS–PAGE (20 µl loaded). As results of the tested 1 ml aliquots, the optimised method was that cells were harvested after 5 hours of IPTG induction. The cells were harvested by centrifugation (20 min, 17000 g) and resuspended in 10 ml extraction buffer (25 mM Tris–HCl, 300 mM NaCl, pH 8) prior to applying French pressure cell to generate cell lysis. Cell lysate was again centrifuged (20 min, 25000 g) and supernatant collected (referred as first supernatant). Cell lysate was resuspended in extraction buffer (5 ml) and centrifuged, thereby generating the second supernatant.

2.2.8.2 Auto-induction

The method used for the auto-induction is based on a previous study by Studier.[44] In detail, the *E. coli* strains were grown in ZYP-0.8G medium with 100 µg/ml of ampicillin at 37°C for 4 hours. An aliquot was transferred into ZY-5052 media, containing 100 µg/ml ampicillin. Similar to the optimisation of IPTG induction, different time point (0-24 hours) of growing bacteria culture were analysed using 1 ml aliquots. The 1 ml aliquots were tested for soluble and insoluble fractions as described above. Three different approaches were investigated to optimise the expression of TM in auto-induction media; (1) antifoam was added to the media, (2) metal content was diminished, (3) glucose concentration was increased by 10-fold, with increase expression time of 48 hours. In the optimised method, the culture was grown over night at 37°C, for 18 hours with 0.13 g rotation and the media was composed of ZY-5052 broth containing 100 µg/ml ampicillin, 100 µM $MgSO_4$ (Sigma-Aldrich, USA), 100 µM $Fe(III)Cl_3$ (Sigma-Aldrich, USA) and antifoam (10 µl) (Sigma-Aldrich, USA).

The auto-induction experiments were performed in two different volumes (250 ml and 1000 ml). For the expression in 250 ml volume, the harvested and lysated cells were resuspended in 10 ml extraction buffer (25 mM Tris–HCl, 300 mM NaCl, pH 8) prior to applying French pressure cell. Subsequently the cell lysate was centrifuged (20 min, 25000 g), supernatant collected and cell lysate resuspended again in the extraction buffer (5 ml) and centrifuged. This generates the second supernatant. The pellet was resuspended in 10 ml extraction buffer, leading to the third supernatant. For the expression in 1000 ml the volume of the extraction buffer was increased to 30 ml prior to applying French pressure cell. All subsequent supernatants collected (12 supernatants in total) were resuspended in 30 ml extraction buffer.

2.2.9 Purification of recombinant tropomyosin of King prawn

To purify the expressed rTM from the supernatants affinity chromatography was performed, using as a nickel charged metal chelate column (GE Healthcare, USA).The HisTrap purification was performed on a Biologic Duoflow system (Bio-Rad, USA) using 250 mM imidazole as elution buffer. The fractions collected from

the Biologic Duoflow system containing the purified recombinant tropomyosin were dialysed in PBS and stored at -80°C until further use.

2.2.10 Immunoblotting

2.2.10.1 Immunoblotting applying antibodies

To confirm the presence of TM various antibodies were used. After proteins separation by SDS–PAGE, proteins were transferred to an activated Polyvinylidene fluoride (PVDF) membrane (Bio-Rad, Australia) using the Semi-dry TransBlot electrophoretic transfer system (Bio-Rad, USA) for 25 min at 15 V. After blocking with 5% skim milk powder in PBS with 0.05% Tween-20 (Sigma-Aldrich, USA) for 1 hour at room temperature, the membrane was incubated for 1 hour at room temperature applying primary antibodies, followed by incubation for 1 hour at room temperature utilising secondary antibodies. For TM detection in various samples, the proteins were exposed to (1) monoclonal anti-tropomyosin antibody (Abcam, Cambridge, MA, USA) diluted 1:6000 and subsequently rabbit anti-mouse IgG horseradish peroxidase labelled antibody (Sigma-Aldrich, USA) diluted 1:40000; (2) in-house[45] polyclonal anti-crustacean antibody diluted 1:30000 and subsequently anti-rabbit IgG HRP conjugate (Promega, USA) diluted 1:40000; (3) anti-His antibody (GE Healthcare, UK) diluted 1:3000 and subsequently rabbit anti-mouse IgG horseradish peroxidase labelled antibody (Sigma-Aldrich, USA) diluted 1:40000. Antibody binding was visualised using the enhanced chemiluminescence (ECL) kit (Bio-Rad, Australia).

2.2.10.2 Patient sera IgE immunoblotting

IgE immunoblotting was performed to analyse patient IgE antibody reactivity to rTM. Ten subjects with a confirmed clinical history of allergic reactivity to crustacean and one nonatopic subject were recruited by The Alfred Hospital, Allergy Clinic, Melbourne Victoria, Australia.[19, 46] Twenty-five µg of the recombinant TM were loaded on 12% SDS–PAGE and resolved at 170 V until the tracker dye reached the bottom. The proteins were transferred to an activated PVDF membrane and blocked with 5% skim milk powder in PBS with 0.05% Tween-20 (Sigma-Aldrich, USA) and subsequently incubated overnight with 1:10 diluted

patient sera using a slot blot apparatus (Idea Scientific, USA). IgE binding was detected using rabbit anti-human IgE polyclonal antibody (1:10000) (DAKO Corporation, USA) and subsequently goat-anti-rabbit IgG HRP conjugate (Promega, USA) (1:10000).

Ethics approval for this chapter was granted by James Cook University's Ethics committee (Project number H4313) in collaboration with The Alfred Hospital (Project number 192/07) and Monash University's Ethics Committees (MUHREC CF08/0225).[19]

2.3 Results

2.3.1 Purifying natural tropomyosin from King prawn

The raw extract and the whole heated extract from King prawn and Black Tiger prawn were made as described in the material and methods section. Moreover, the protein concentration of various extracts was estimated. In order to generate pure tropomyosin for following chapters as reference protein tropomyosin from KP was purified from the whole heated extract using a strong anion-exchange chromatographic column. Figure 2.1 shows the ion exchange purification profile of King prawn tropomyosin, using an increasing NaCl concentration and measuring the absorbance at different nanometers (nm) with the maximum absorbance for TM at 260 nm. The purified natural TM was compared with the raw extract and the whole heated extract by SDS-PAGE (Figure 2.2). For the raw extracts more bands are visible in comparison to the whole heated extract. Moreover, there are minor differences between KP and BTP in the extracts and the natural TM. To confirm that the natural TM of KP has the same molecular shape as the natural TM of BTP, CD spectroscopy analysis was performed (Figure 2.3). The CD spectrometric analysis demonstrated identical secondary structures for both natural tropomyosins. TM exhibited a distinct negative signal at 208 and 222 nm, typical for an alpha helical protein.

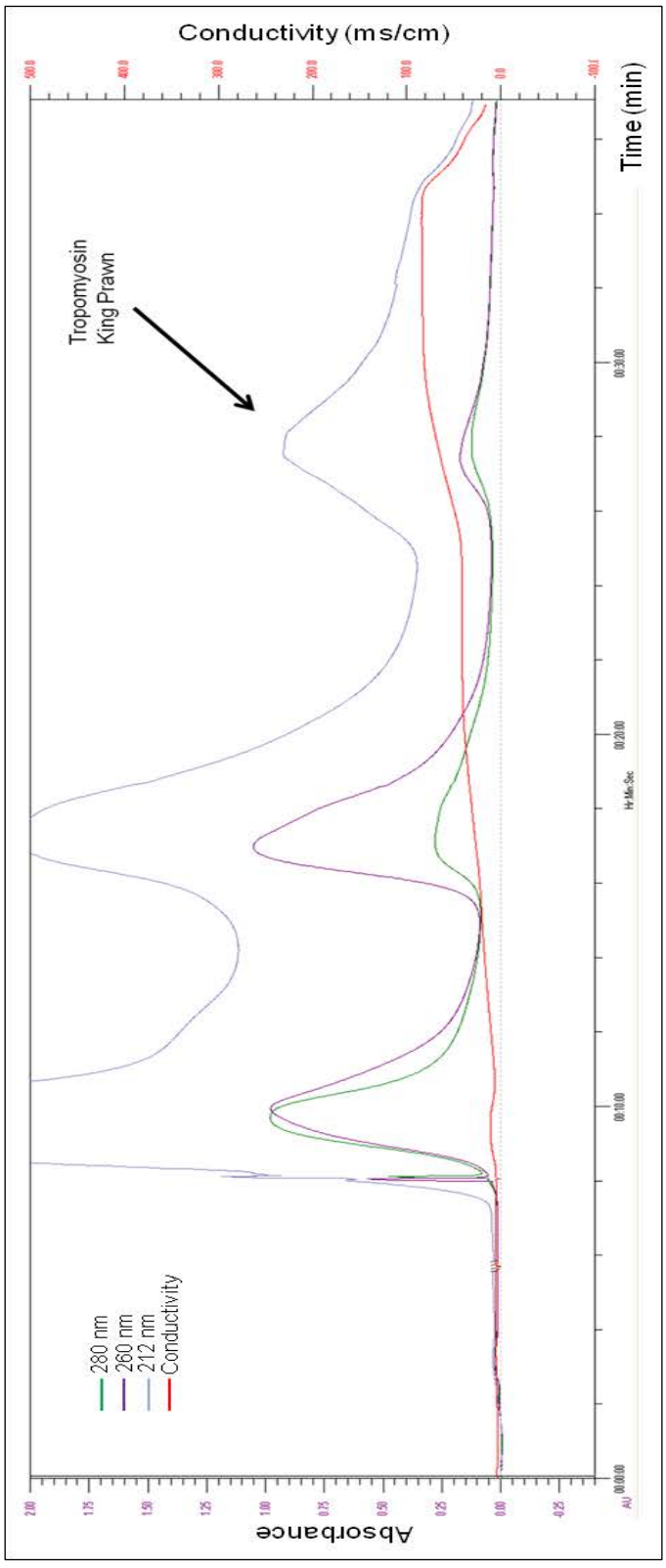


Figure 2.1: Ion exchange purification profile of King prawn tropomyosin, using an increasing NaCl concentration and measuring the absorbance at different nm.

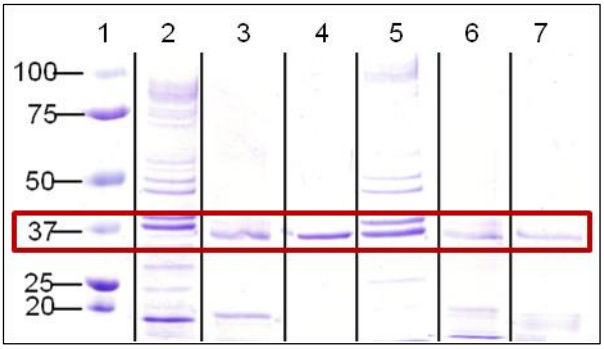


Figure 2.2: SDS-PAGE protein analysis of prawn extracts and natural purified tropomyosins. The red box highlights the molecular weight where TM is expected. Lane numbers represent: 1. Marker; 2. KP raw extract; 3. KP whole heated extract; 4. KP natural TM; 5. BTP raw extract; 6. BTP whole heated extract; and 7. BTP natural TM.

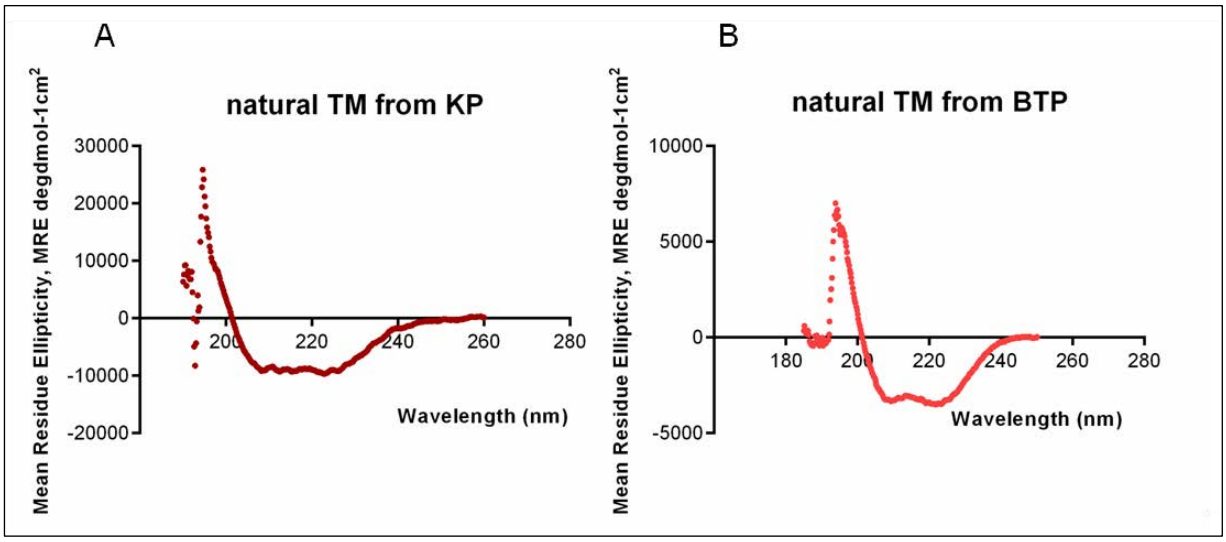


Figure 2.3: CD spectrometric analysis of natural TM from (A) KP and natural TM from (B) BTP.

2.3.2 Comparison of tropomyosin amino acid sequences derived by cDNA

Tropomyosin (TM) is the major allergen in crustaceans, a well established pan-allergen. To investigate the amino acid sequence of TM of King prawn (*Melicertus latisulcatus* (KP)) cDNA was amplified and analysed. The amino acid sequence of tropomyosin from King prawn was deduced from the cDNA generated in this chapter and published as Genbank accession number JX171685.

Figure 2.4 displays the alignment of amino acid sequences of TM from KP and is compared to TM from Black Tiger prawn (BTP). The amino acids highlighted in red show the differences between KP and BTP. Moreover, as indicated by the black boxes (Figure 2.4), amino acids differ in one of the IgE antibody binding regions that were identified by Reese et al.[47] for the allergen Pen a 1. Although species are genetically related, they can express different allergens. Interestingly it was found that the amino acid sequence homology between KP and BTP is only 95% (Figure 2.4). A total of 11 amino acids are different for KP and BTP. Moreover, all variances are between amino acid position 34-71, which overlaps with a predicted epitope region for prawn tropomyosin.

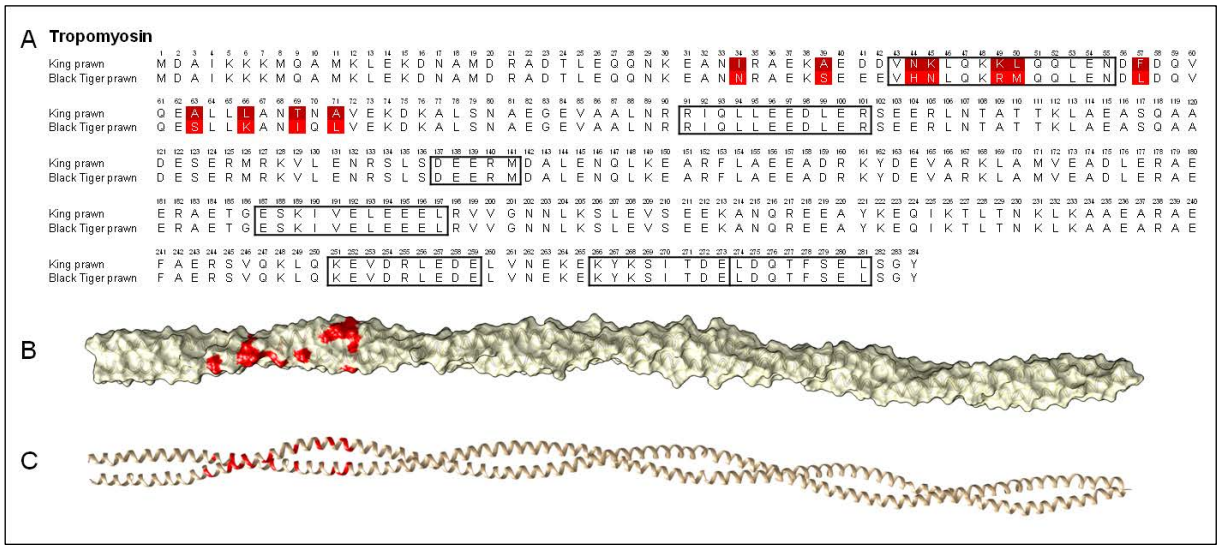


Figure 2.4: Comparison of TM from KP and BTP (A) Amino acid sequence alignment of tropomyosin from KP and BTP. The amino acids substitutions are marked in red for differences between KP and BTP Black boxes indicate IgE binding regions previously identified by Reese et al[47] for the allergen Pen a 1. (B) Spatial model of TM with differences from KP and BTP highlighted in red. (C) Ribbon model of TM with differences from KP and BTP highlighted in red.

2.3.3 Comparison of myosin light chain amino acid sequences derived by cDNA

To investigate if other known allergens in prawns show low amino acid sequence identity, a minor allergen, myosin light chain (MLC) was chosen. The amino acid sequence of MLC from King prawn was deduced from the cDNA.

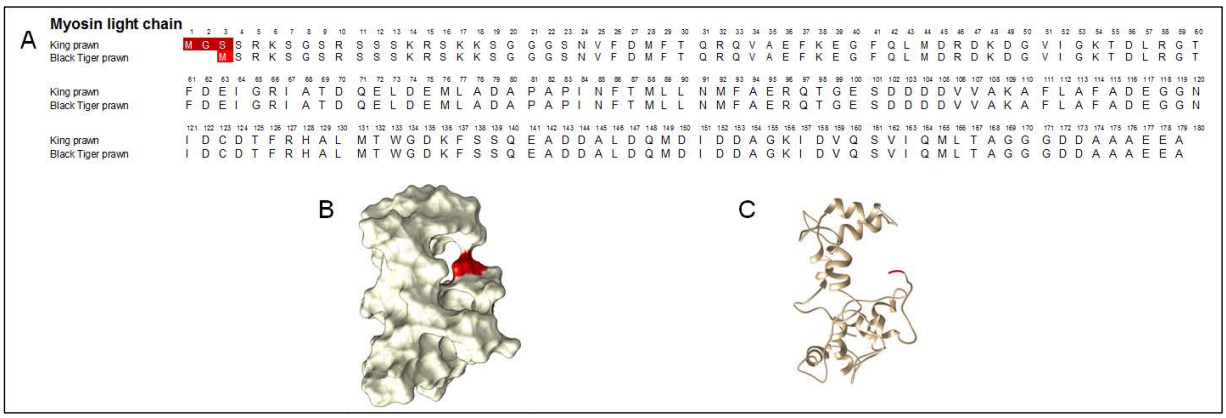


Figure 2.5: Comparison of MLC from KP and BTP (A) Amino acid sequence alignment of MLC from KP and BTP. The amino acids substitutions are marked in red for differences between KP and BTP. (B) Spatial model of MLC with differences from KP and BTP highlighted in red. (C) Ribbon model of MLC with differences from KP and BTP highlighted in red.

Figure 2.5 displays the alignment of amino acid sequences of MLC from KP and is compared to MLC from Black Tiger prawn (BTP). The amino acids highlighted in red show the differences between KP and BTP. Only three amino acids are different for KP in comparison to BTP, all of them being at the beginning of the N-terminal end of the protein. Therefore MLC from KP and BTP are 99.44% identical. Moreover, as it can be seen in the spatial and ribbon model the structure of MLC does not seem to change with the additional amino acid at the N-terminal end for the KP in comparison to BTP. Although the major allergen TM from KP shows low amino acid identity with BTP, the minor allergen MLC shows high amino acid identity between KP and BTP.

2.3.4 Expression of recombinant tropomyosin of King prawn

In order to investigate if tropomyosin is an allergen, TM from KP was expressed as recombinant protein. Moreover, the recombinant tropomyosin (rTM) from KP was used in the following chapters as reference protein. However, large amounts of rTM will be necessary for method development and immunotherapeutics applications. Therefore two expression systems for the expression of allergenic tropomyosin from KP were investigated and compared (Section 2.3.4.1 – 2.3.4.4).

2.3.4.1 IPTG induced expression of tropomyosin

The first aim was to optimise the expression time of TM in *E. coli* cells. NM522 and BL21(DE3)RIPL cells were grown in 25 ml LB media and 1 ml aliquots were tested for the expression of TM. To visualise the expression 1 ml aliquots were separated by SDS–PAGE, before and after IPTG induction. Moreover, to ensure presence of soluble tropomyosin, both soluble and insoluble fractions were tested. Figure 2.6 visualises the growing conditions of NM522 cells over a period of 21 hours. The BL21(DE3)RIPL performed very similar based on measured OD600 values for the soluble fractions. Three hours after IPTG was introduced, the concentration of TM was high enough to be visualised by SDS-PAGE. In the following two hours of growing the culture the amount of expressed TM increased, however, expression did not increase considerably between 5 and 24 hours. Therefore it was focused on a more repeat production of recombinant tropomyosin after 5 hours instead.

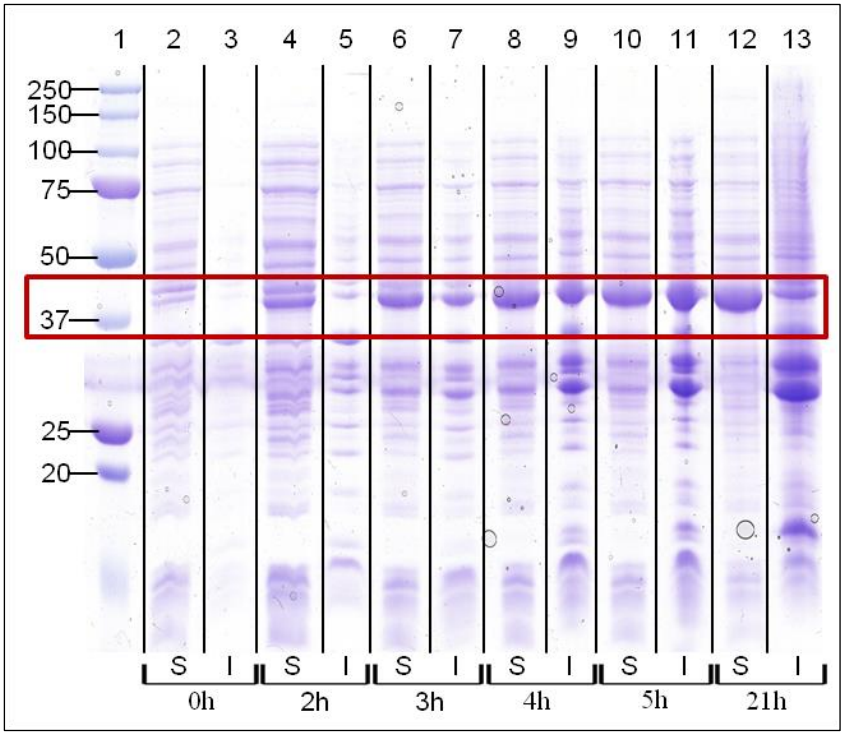


Figure 2.6: SDS–PAGE protein analysis of bacterial fractions for NM522 cells using LB media and IPTG induction. Fractions were analysed for soluble (S) and insoluble (I) TM at different time points of bacteria culture growth. The red box highlights the molecular weight where TM is expected. Lane numbers represent: 1. Marker; 2. 0 hours soluble fraction; 3. 0 hours insoluble fraction; 4. 2 hours soluble fraction; 5. 2 hours insoluble fraction; 6. 3 hours since soluble fraction; 7. 3 hours insoluble fraction; 8. 4 hours soluble fraction; 9. 4 hours insoluble fraction; 10. 5 hours soluble fraction; 11. 5 hours insoluble fraction; 12. 21 hours soluble fraction; and 13. 21 hours insoluble fraction.

2.3.4.2 Auto-induction system for tropomyosin

The first auto-induction experiments were similar to the IPTG optimisation. The changes made were that culture medium had a higher OD600, in comparison to the IPTG induced cells. Accordingly, 1 ml aliquots of cell cultures were analysed for TM in the soluble and insoluble fractions. Briefly, TM was mainly present in the soluble fractions, however, some was formed in inclusion bodies. Although more protein was expressed, the amount of insoluble protein did not increase in comparison to the IPTG induction (Table 2.1).

Table 2.1: Summary of different growing conditions and TM yield generated for different *E. coli* strains and 250 ml LB media (IPTG induction) or ZY-5052 media (auto-induction) investigated.

	NM522 (250 ml LB media)	BL21(DE3)RIPL (250 ml LB media)	BL21(DE3)RIPL (250 ml ZY-5052 media)
Optimal harvesting time	5 hours after ITPG induction	5 hours after ITPG induction	18 hours after growing bacteria culture
OD 600 nm at optimal harvesting time	0.934	1.002	13.601
g bacteria wet pellet	0.21	1.92	2.15
% of TM of total protein (densitometric analysis based on figure 2.8a)	43%	54.46%	68.55%
Ratio of insoluble TM to total protein (densitometric analysis based on figure 2.8a)	10.23	8.06	7.7
Ratio of yield mg/g bacteria wet pellet	1.9	2.29	2.51
Yield mg purified TM	**0.4**	**4.4**	**5.4**

To further optimise the auto-induction system three different approaches were investigated: (1) the addition of antifoam increased the cell density and cells could be harvested at higher OD600, leading to higher expression of TM; (2) the ZY-5052 media contains different metals. However, chelate bindings between the His-tag of recombinant protein and metals in the ZY-5052 media could be possible. Therefore an expression experiment without the addition of metals salts was performed, however, adding $MgSO_4$ and $Fe(III)Cl_3$ seem to be essential to generate expression of TM. Moreover, no chelate binding with the addition of metals was observed by SDS-PAGE analysis before and after purification; (3) glucose concentration was increased 10-fold, to investigate if cell density and therefore expression can be further increased. Increasing the glucose concentration in ZY-5052 media led to higher cell density, but not to increased TM expression. The increase of glucose concentration probably inhibited the *lac* operon, even if the culture was grown for longer time (48 hours).

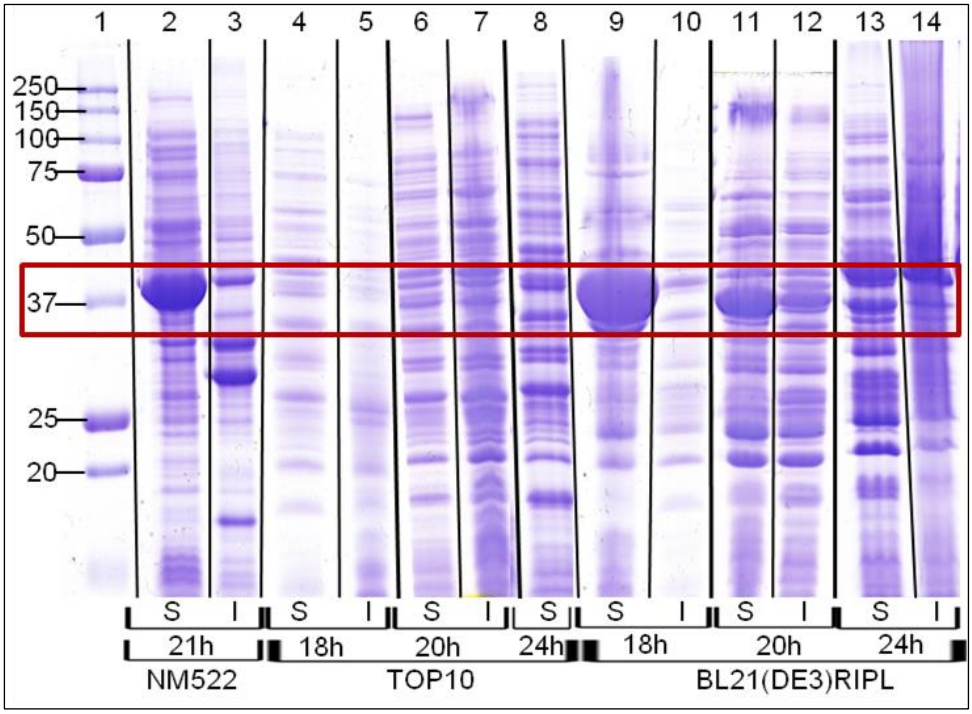


Figure 2.7: SDS-PAGE protein analysis of bacterial cultures after IPTG induction (NM522) or auto-induction (TOP10 and BL21(DE3)RIPL). Fractions were tested for soluble (S) and insoluble (I) TM at different time points of culture growth. The red box highlights the molecular weight where TM is expected. Lane numbers represent: 1. Marker; 2. NM522 21 h soluble; 3. NM522 21 h insoluble; 4. TOP10 ZY-5052 18 h soluble; 5. TOP10 18 h insoluble; 6. TOP10 20 h soluble; 7. TOP10 20 h insoluble; 8. TOP10 24 h soluble; 9. BL21(DE3)RIPL ZY-5052 media 18 h soluble; 10. BL21(DE3)RIPL 18 h insoluble; 11. BL21(DE3)RIPL 20 h soluble; 12. BL21(DE3)RIPL 20 h insoluble; 13. BL21(DE3)RIPL 24 h soluble; and 14. BL21(DE3)RIPL 24 h insoluble.

NM522 cells demonstrated good expression when IPTG induced, however, did not express TM when being auto-induced. NM522 cells are Δ(*lac-pro*AB) and therefore not suitable for auto-induction. To be able to compare *E. coli* K-strain with *E. coli* B-strain TOP10 cells were used. Figure 2.7 shows that TOP10 cells can be used for auto-induction. However, in comparison to IPTG induced NM522 cells the expression is very low, even after 24 hours. The expression of TM in TOP10 cells was confirmed by immunoblotting using anti-His antibody and polyclonal anti-crustacean antibody. Densitometric analysis showed that TM comprised only about 5.3% of the total protein as demonstrated by SDS–PAGE. Figure 2.7 demonstrates that BL21(DE3)RIPL cells are expressing considerable

more TM. Therefore all further auto-induction experiments were harvested after 18 hours and grown under the conditions explained in the material and method section.

2.3.4.3 IPTG induction versus auto-induction of tropomyosin of King prawn

To compare the same experiments, the IPTG (NM522 and BL21(DE3)RIPL) and auto-induction system (BL21(DE3)RIPL) using the same volume (250 ml) were performed. Table 2.1 summarises the most important findings for the different inductions and bacteria strains investigated. The saturation in the auto-inducing media was considerably higher than in the LB media under similar conditions, resulting in an increase in the number of producing cells and therefore, in the final concentration of TM. The different fractions of the supernatants were visualised by SDS–PAGE and the presence of His-TM (rTM) confirmed using a specific anti-His antibody (Figure 2.8). In summary, the expression of recombinant TM was similar between the NM522 and the BL21(DE3)RIPL cells using LB media. In figure 2.8b one additional antibody binding band can be seen at 75 kDa, representing the dimer of recombinant tropomyosin.

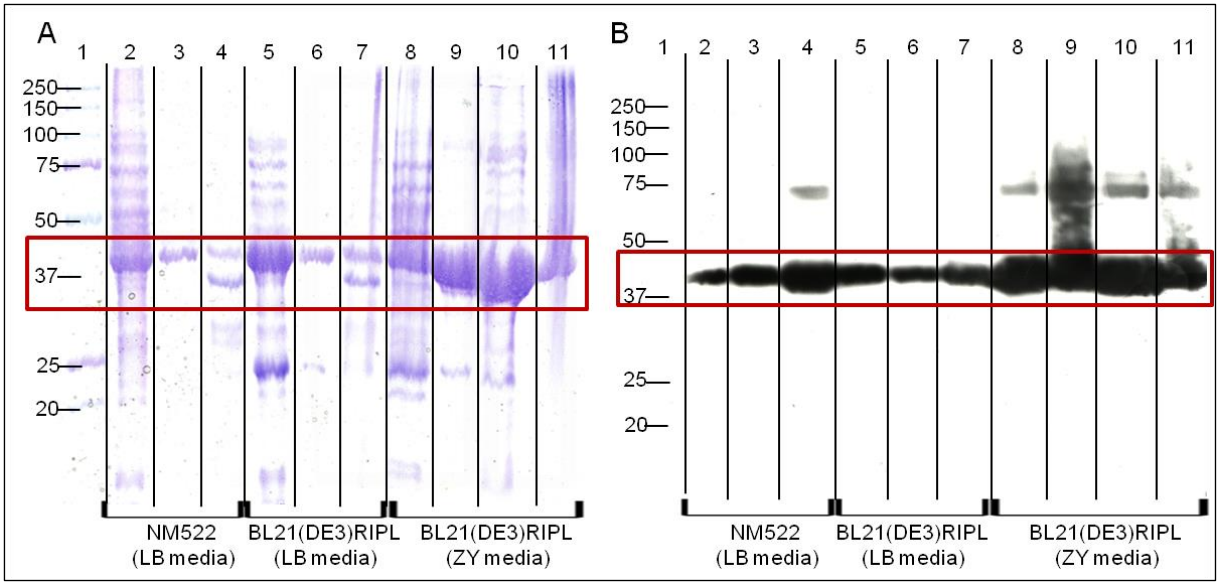


Figure 2.8: Molecular and immunological comparison of different *E. coli* strains and induction systems. Proteins in generated fractions were separated by (A) SDS–PAGE and confirmation of expressed recombinant protein by (B) anti-His antibody. The lane number represent analysed supernatants of NM522 IPTG induction (Lane 2-4), BL21(DE3)RIPL cells for IPTG induction (Lane 5-7) and BL21(DE3)RIPL cells auto-induction (Lane 8-11). The red boxes highlight the molecular weight where TM is expected. Lane numbers represent: 1. Marker; 2. 1st supernatant; 3. 2nd supernatant; 4. pellet; 5. 1st supernatant; 6. 2nd supernatant; 7. pellet; 8. 1st supernatant; 9. 2nd supernatant; 10. 3rd supernatant; and 11. Pellet.

2.3.4.4 Up-scaling expression of tropomyosin by auto-induction

The results in table 2.1 show clearly that the auto-induction using BL21(DE3)RIPL cells result in the highest TM yield and was used to upscale this system. The TM expressed was in a total volume of 1 litre of ZY-5052 media using the optimised conditions. The separation of different supernatants (12 supernatants in total) by SDS-PAGE can be seen in figure 2.9. Using 1 litre of ZY-5052 media a total of 62 mg/l TM could be purified. The higher yield achieved in the 1 litre expression, compared to the 250 ml volume can be explained with higher cell density. Thus, the cell density is directly related to the culture conditions and the medium.

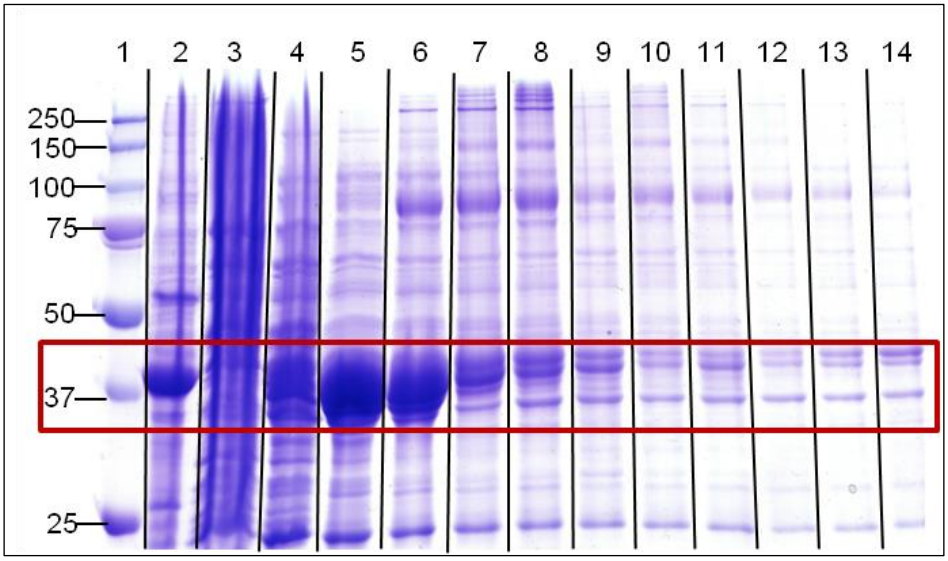


Figure 2.9: Proteins in supernatants generated in 1 litre of auto-induction media were separated by SDS-PAGE. The red box highlights the molecular weight where TM is expected. Lane numbers represent: 1. Marker; 2. Pellet; 3. 1st supernatant; 4. 2nd supernatant; 5. 3rd supernatant; 6. 4th supernatant; 7. 5th supernatant; 8. 6th supernatant; 9. 7th supernatant; 10. 8th supernatant; 11. 9th supernatant; 12. 10th supernatant; 13. 11th supernatant; and 14. 12th supernatant.

2.3.5 Molecular and immunological comparison of King prawn and Black Tiger prawn

To investigate the molecular and immunological properties of the generated recombinant TM (rTM) from KP, this protein was compared to previously generated rTM from Black Tiger prawn. Therefore protein extracts and rTM from KP and BTP were separated by SDS-PAGE. The difference in the protein separation profile from figure 2.2 and figure 2.10a can be explained that for figure 2.10 the amount of protein separated was higher. The separated proteins were detected by three different antibodies to investigate TM binding captivities. Figure 2.10 shows the separated protein extracts from KP and BTP by SDS–PAGE and antibody binding using three different antibodies; (1) monoclonal anti-tropomyosin antibody; (2) polyclonal anti-crustacean antibody; (3) anti-His antibody. The SDS–PAGE profiles and the antibody reactivity of the two recombinant tropomyosins are very similar, with the exception of slightly stronger monoclonal anti-tropomyosin antibody binding to BTP. As expected, the anti-His antibody only bound to the

recombinant TM and not to the naturally occurring proteins (Figure 2.10b). The rTM of KP and BTP share the same molecular weight by SDS–PAGE (Figure 2.10a), however, slightly higher than calculated from the amino acid sequence with 36055.94 Da for KP and 36978.91 Da for BTP, respectively.

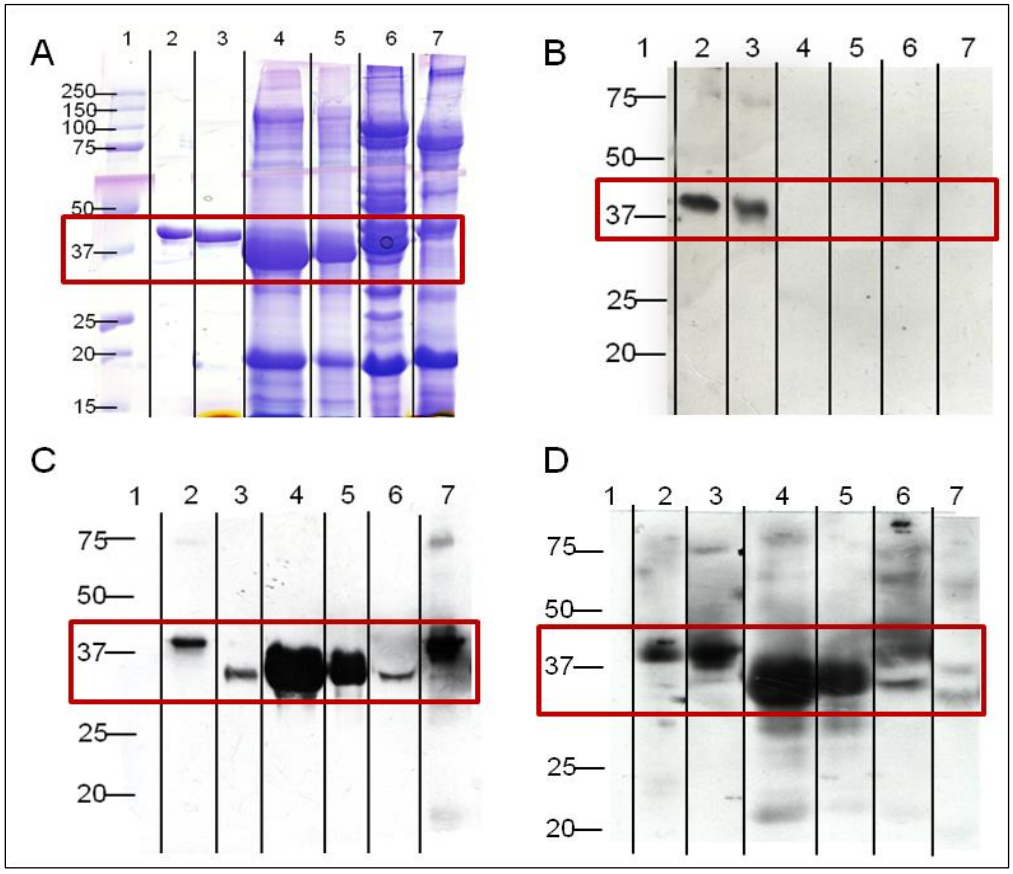


Figure 2.10: SDS-PAGE and immunoblotting analysis of protein extracts and recombinant TM from King prawn and Black Tiger prawn. Different protein extracts of KP and BTP are separated by (A) SDS–PAGE and the presence of recombinant protein demonstrated by (B) anti-His antibody (C) monoclonal anti-tropomyosin antibody and (D) polyclonal anti-crustacean antibody. The red boxes highlight the molecular weight where TM is expected. Lane numbers represent: 1. Marker; 2. rTM KP; 3. rTM BTP; 4. KP whole heated extract; 5. BTP whole heated extract; 6. KP raw extract; and 7. BTP raw extract.

The secondary structure of the rTM from KP was compared to the rTM from BTP using CD spectroscopy. Figure 2.11 demonstrates identical secondary structures. Very similar to the natural purified tropomyosins, the recombinant tropomyosins

exhibited a distinct negative signal at 208 and 222 nm, typical for an alpha helical protein.

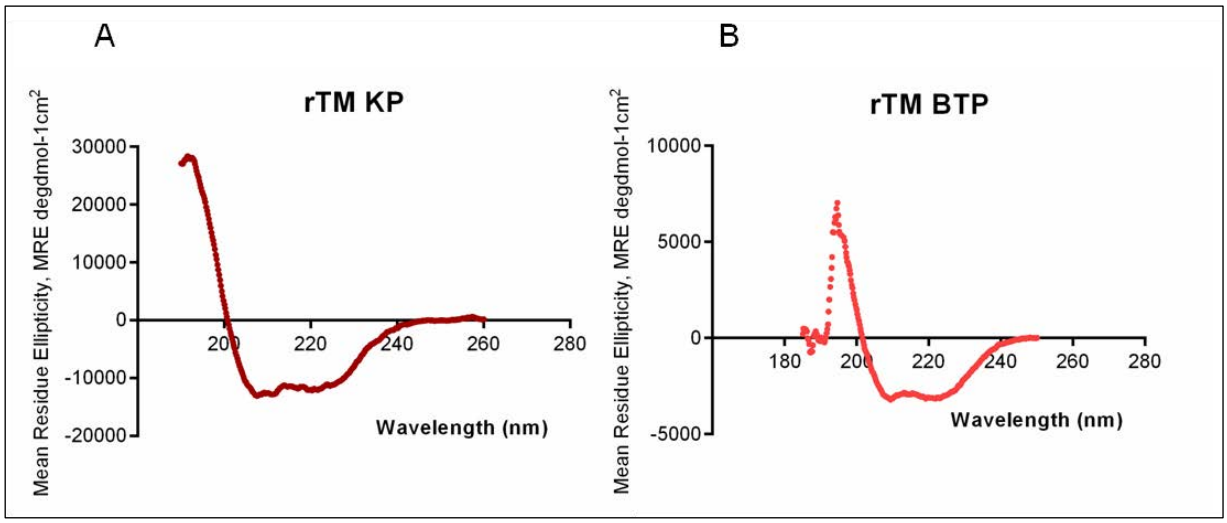


Figure 2.11: CD spectrometric analysis of recombinant TM (rTM) from (A) KP and rTM from (B) BTP.

Figure 2.12 illustrates the IgE antibody reactivity of ten crustacean allergic patients to rTM KP and rTM BTP. Out of ten patients analysed six patients show a positive IgE binding to rTM KP. Therefore TM from KP was registered as allergen, with the International Union of Immunological Societies (IUIS), following their nomenclature and named Mel I 1. Overall, the crustacean allergic patients demonstrate very similar IgE binding to rTM KP and rTM BTP. However, all patients show a stronger IgE binding to rTM BTP. Interestingly, patient A and patient C only recognise rTM from BTP. As presented in figure 2.4, the only region where the two tropomyosins differ are between amino acid positions 34-71. Reese et al.[47] identified one IgE binding epitope within this region at amino acid position 42–55 for Pen a 1. In this chapter two patients only recognise rTM BTP and it is therefore expect very specific IgE binding to an epitope in this region, possible the same as identified by Reese et al.[47]

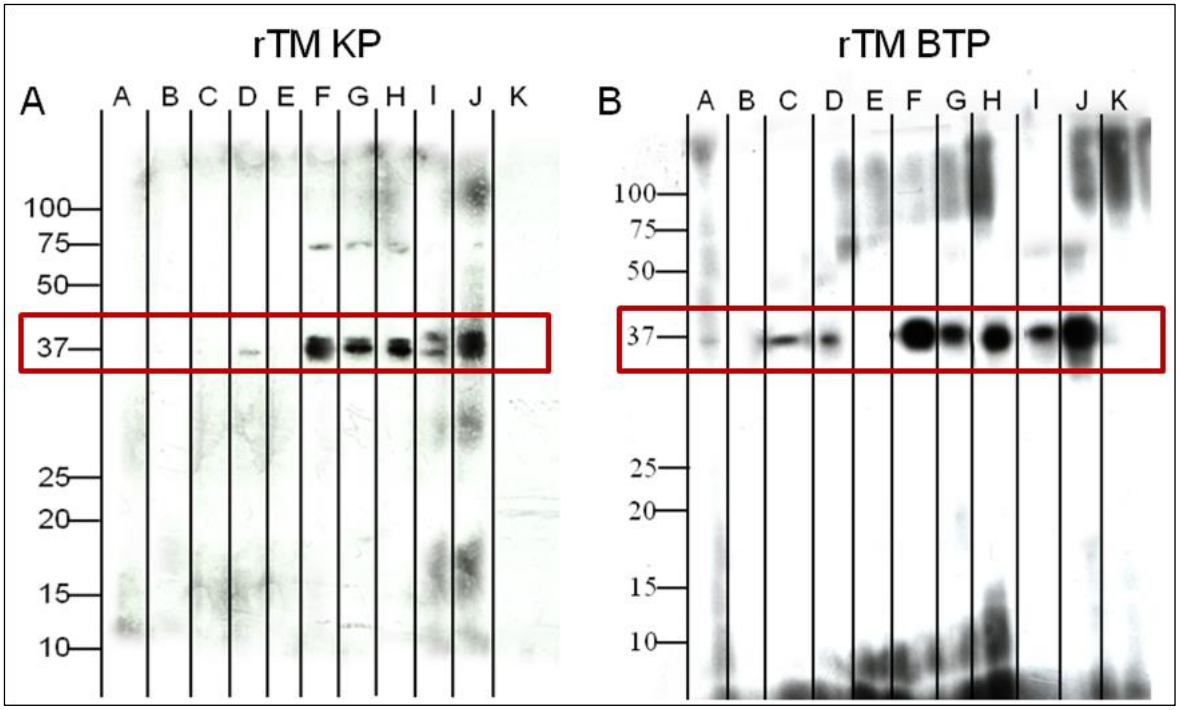


Figure 2.12: Immunoblotting analysis applying sera from ten crustacean allergic patients and one negative control for (A) rTM KP and (B) rTM BTP. The red boxes highlight the molecular weight where TM is expected. Lane numbers represent: A. Patient 1; B. Patient 2; C. Patient 3; D. Patient 4; E. Patient 5; F. Patient 6; G. Patient 7; H. Patient 8; I. Patient 9; J. Patient and 10. K negative control serum.

2.4 Discussion

The current diagnostic tools for the detection and quantification of allergens in food products need to be improved. Therefore, recombinant allergenic proteins become an essential tool for method developments and reference materials. In this chapter, natural tropomyosin from KP was purified and recombinant tropomyosin from King prawn was expressed. The natural TM and rTM from KP will be used as reference material in the following chapters. The yield of expressed functional recombinant TM was increased using B-Strain of *E. coli* and a specific auto-induction system. This is the first reported experiment to demonstrate that high yield recombinant allergen can be expressed using the auto-induction system. The yield achieved with the auto-induction systems for TM (62 mg/l) is higher than reported by Studier[44] with 30-50 mg/l. Moreover, much lower quantities were reported for cockroach (Bal g 7) with 7.2 mg/l,[28] 20 mg/l for chicken[35] expressed in yeast. Anisakis TM with 15 mg/l[32] and house dust mite (Der p 10) with 26 mg/l [26] expressed in *E. coli. W*ith 62 mg/l of King prawn TM in this chapter the auto-induction systems generated considerable more recombinant protein as previously reported using other systems for TM expressions.

TM is the major allergen in prawns, a well established pan-allergen. The amino acid sequence for tropomyosin from King prawn was investigated and registered as a novel pan-allergen, Mel I 1. Mel I 1 was compared with Pen m 1, a well characterised allergenic tropomyosin from Black Tiger prawn.[2, 13, 18, 19] In this chapter it was investigated that the amino acid identity for KP and BTP prawn is only 95%. The amino acid identity within prawns is usually as high as 98-100%,[14] thus 95% between KP and BTP shows relatively low amino acid sequence identity. However, as it was demonstrated, only because one pan-allergen is different within two prawn species, does not imply that other (pan-)allergens are different, thus, the amino acid sequence identify for myosin light chain was 99.44% for KP and BTP.

Overall, the molecular and immunological characteristics between the allergenic TM of KP and BTP are very similar. Nevertheless, species specific IgE reactivity was demonstrated for Mel I 1 and Pen m 1. Species specific reactivity to prawns has previously been demonstrated for Black Tiger prawn compared with Fresh water shrimp[48] and White shrimp compared with Brown shrimp.[49] The latter study

analysed 31 patients by radioallergosorbent testing (RAST) and skin prick testing and found three patients reacting only to one prawn species by RAST and one patient by skin prick testing, respectively.[49] Jirapongsananuruk et al.[48] reported that out of 60 prawn allergic children, 12 reacted only positive to Black Tiger prawn, whereas 16 children only reacted positive to Fresh water shrimp, by skin prick testing and oral food challenge. Both studies used prawn extracts for their studies. Moreover, in the amino acid region where KP and BTP are different (amino acid sequence 34-71), one epitope was identified for Pen a 1 (amino acid 43-55 or 45-57),[47, 50-52] whereas two epitope were identified for Pen m 1 (amino acid 23-40 and 45-59).[53] Therefore it is suggested that there is species specific IgE reactivity for KP and BTP and can be explained by the amino acid difference for KP and BTP.

Due to the comparison of KP extracts and recombinant TM and natural TM with the BTP analogues, it is suggested that TM is a major allergen in KP. While the molecular and immunological properties of both allergens are very similar the species specific reactivity is of great importance for component resolved diagnosis.

2.5 References

1. Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). *Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems.* Journal of Proteome Research. 13, (8), 3499-3509.

2. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M. X.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* International Journal of Hygiene and Environmental Health. (0).

3. Abdel Rahman, A. M.; Lopata, A. L.; O'Hehir, R. E.; Robinson, J. J.; Banoub, J. H.; Helleur, R. J. (2010). *Characterization and De Novo Sequencing of Snow Crab Tropomyosin Enzymatic Peptides by Both Electrospary Ionization and Matrix-Assisted Laser Desorption Ionization Qqtof Tandem Mass Spectrometry.* Journal of Mass Spectrometry. 45, (4), 372-381.

4. Jeong, H.; Barbe, V.; Lee, C. H.; Vallenet, D.; Yu, D. S.; Choi, S.-H.; Couloux, A.; Lee, S.-W.; Yoon, S. H.; Cattolico, L.; Hur, C.-G.; Park, H.-S.; Segurens, B.; Kim, S. C.; Oh, T. K.; Lenski, R. E.; Studier, F. W.; Daegelen, P.; Kim, J. F. (2009). *Genome Sequences of Escherichia Coli B Strains Rel606 and Bl21(De3).* Journal of Molecular Biology. 394, (4), 644-652.

5. Lopata, A. L.; Kamath, S. (2012). *Shellfish Allergy Diagnosis –Gaps and Needs.* Current Allergy & Clinical Immunology. 25, (2), 60-66.

6. Vrtala, S.; Huber, H.; Thomas, W. R. (2014). *Recombinant House Dust Mite Allergens.* Methods. 66, (1), 67-74.

7. Bohle, B.; Vieths, S. (2004). *Improving Diagnostic Tests for Food Allergy with Recombinant Allergens.* Methods. 32, (3), 292-299.

8. Shreffler, W. G. (2011). *Microarrayed Recombinant Allergens for Diagnostic Testing.* Journal of Allergy and Clinical Immunology. 127, (4), 843-849.

9. van der Ventel, M. L.; Nieuwenhuizen, N. E.; Kirstein, F.; Hikuam, C.; Jeebhay, M. F.; Swoboda, I.; Brombacher, F.; Lopata, A. L. (2011). *Differential Responses to Natural and Recombinant Allergens in a Murine Model of Fish Allergy.* Molecular Immunology. 48, (4), 637-646.

10. Burks, A. W.; Calderon, M. A.; Casale, T.; Cox, L.; Demoly, P.; Jutel, M.; Nelson, H.; Akdis, C. A. (2013). *Update on Allergy Immunotherapy: American Academy of Allergy, Asthma &Amp; Immunology/European Academy of Allergy and Clinical Immunology/Practall Consensus Report.* Journal of Allergy and Clinical Immunology. 131, (5), 1288-1296.e3.

11. Nowak-Węgrzyn, A.; Sampson, H. A. (2011). *Future Therapies for Food Allergies.* Journal of Allergy and Clinical Immunology. 127, (3), 558-573.

12. Sharp, M. F.; Kamath, S. D.; Koeberl, M.; Jerry, D. R.; O'Hehir, R. E.; Campbell, D. E.; Lopata, A. L. (2014). *Differential Ige Binding to Isoallergens from Asian Seabass (Lates Calcarifer) in Children and Adults.* Molecular Immunology. 62, (1), 77-85.

13. Abramovitch, J. B.; Kamath, S.; Varese, N.; Zubrinich, C.; Lopata, A. L.; O'Hehir, R. E.; Rolland, J. M. (2013). *Ige Reactivity of Blue Swimmer Crab (Portunus Pelagicus) Tropomyosin, Por P 1, and Other Allergens; Cross-Reactivity with Black Tiger Prawn and Effects of Heating.* Plos One. 8, (6), e67487.

14. Lopata, A. L.; O'Hehir, R. E.; Lehrer, S. B. (2010). *Shellfish Allergy.* Clinical and Experimental Allergy. 40, (6), 850-858.

15. Lopata, A. L.; Lehrer, S. B. (2009). *New Insights into Seafood Allergy.* Current Opinion in Allergy and Clinical Immunology. 9, (3), 270-277.

16. Kamath, S. D.; Abdel Rahman, A. M.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. 58, (5), 1144-1155.

17. Bauermeister, K.; Wangorsch, A.; Garoffo, L. P.; Reuter, A.; Conti, A.; Taylor, S. L.; Lidholm, J.; DeWitt, Å. M.; Enrique, E.; Vieths, S.; Holzhauser, T.; Ballmer-Weber, B.; Reese, G. (2011). *Generation of a Comprehensive Panel of Crustacean Allergens from the North Sea Shrimp Crangon Crangon.* Molecular Immunology. 48, (15–16), 1983-1992.

18. Kamath, S. D.; Abdel Rahman, A. M.; Komoda, T.; Lopata, A. L. (2013). *Impact of Heat Processing on the Detection of the Major Shellfish Allergen Tropomyosin in Crustaceans and Molluscs Using Specific Monoclonal Antibodies.* Food Chemistry. 141, (4), 4031-9.

19. Kamath, S. D.; Rahman, A. M. A.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. n/a-n/a.

20. Leung, P. S. C.; Chow, W. K.; Duffey, S.; Kwan, H. S.; Gershwin, M. E.; Chu, K. H. (1996). *Ige Reactivity against a Cross-Reactive Allergen in Crustacea and Mollusca: Evidence for Tropomyosin as the Common Allergen.* Journal of Allergy and Clinical Immunology. 98, (5, Part 1), 954-961.

21. Leung, P. S. C.; Chu, K. H.; Chow, W. K.; Ansari, A.; Bandea, C. I.; Kwan, H. S.; Nagy, S. M.; Gershwin, M. E. (1994). *Cloning, Expression, and Primary Structure of Metapenaeus-Ensis Tropomyosin, the Major Heat-Stable Shrimp Allergen.* Journal of Allergy and Clinical Immunology. 94, (5), 882-890.

22. Leung, P. S. C.; Chen, Y. C.; Gershwin, M. E.; Wong, H.; Kwan, H. S.; Chu, K. H. (1998). *Identification and Molecular Characterization of Charybdis Feriatus Tropomyosin, the Major Crab Allergen.* Journal of Allergy and Clinical Immunology. 102, (5), 847-852.

23. Liang, Y. L.; Cao, M. J.; Su, W. J.; Zhang, L. J.; Huang, Y. Y.; Liu, G. M. (2008). *Identification and Characterisation of the Major Allergen of Chinese Mitten Crab (Eriocheir Sinensis).* Food Chemistry. 111, (4), 998-1003.

24. Aki, T.; Fujikawa, A.; Wada, T.; Jyo, T.; Shigeta, S.; Murooka, Y.; Oka, S.; Ono, K. (1994). *Cloning and Expression of Cdna Coding for a New Allergen from the House-Dust Mmite, Dermatophagoides-Farinae-Homology with Human Heat-Shock Cognate Proteins in the Heat-Shock Protein-70 Family* Journal of Biochemistry. 115, (3), 435-440.

25. Aki, T.; Kodama, T.; Fujikawa, A.; Miura, K.; Shigeta, S.; Wada, T.; Jyo, T.; Murooka, Y.; Oka, S.; Ono, K. (1995). *Immunochemical Characterization of Recombinant and Native Tropomyosins as a New Allergen from the House Dust Mite, Dermatophagoides Farinae.* Journal of Allergy and Clinical Immunology. 96, (1), 74-83.

26. Asturias, J. A.; Arilla, M. C.; Gomez-Bayon, N.; Martinez, A.; Martinez, J.; Palacios, R. (1998). *Sequencing and High Level Expression in Escherichia Coli of the Tropomyosin Allergen (Der P 10) from Dermatophagoides Pteronyssinus.* Biochimica Et Biophysica Acta-Gene Structure and Expression. 1397, (1), 27-30.

27. Pittner, G.; Vrtala, S.; Thomas, W. R.; Weghofer, M.; Kundi, M.; Horak, F.; Kraft, D.; Valenta, R. (2004). *Component-Resolved Diagnosis of House-Dust Mite Allergy with Purified Natural and Recombinant Mite Allergens.* Clinical & Experimental Allergy. 34, (4), 597-603.

28. Jeong, K. Y.; Lee, J.; Lee, I. Y.; Hong, C. S.; Ree, H. I.; Yong, T. S. (2004). *Expression of Tropomyosin from Blattella Germanica as a Recombinant Non-Fusion Protein in Pichia Pastoris and Comparison of Its Ige Reactivity with Its Native Counterpart.* Protein Expression and Purification. 37, (2), 273-278.

29. Satinover, S. M.; Reefer, A. J.; Pomes, A.; Chapman, M. D.; Platts-Mills, T. A. E.; Woodfolk, J. A. (2005). *Specific Ige and Igg Antibody-Binding Patterns to Recombinant Cockroach Allergens.* Journal of Allergy and Clinical Immunology. 115, (4), 803-809.

30. Santos, A. B. R.; Chapman, M. D.; Aalberse, R. C.; Vailes, L. D.; Ferriani, V. P. L.; Oliver, C.; Rizzo, M. C.; Naspitz, C. K.; Arruda, L. K. (1999). *Cockroach Allergens and Asthma in Brazil: Identification of Tropomyosin as a Major Allergen with Potential Cross-Reactivity with Mite and Shrimp Allergens.* Journal of Allergy and Clinical Immunology. 104, (2), 329-337.

31. Alvite, G.; Esteves, A. (2009). *Echinococcus Granulosus Tropomyosin Isoforms: From Gene Structure to Expression Analysis.* Gene. 433, (1–2), 40-49.

32. Asturias, J. A.; Eraso, E.; Martinez, A. (2000). *Cloning and High Level Expression in Escherichia Coli of an Anisakis Simplex Tropomyosin Isoform.* Molecular and Biochemical Parasitology. 108, (2), 263-266.

33. Hartmann, S.; Sereda, M. J.; Sollwedel, A.; Kalinna, B.; Lucius, R. (2006). *A Nematode Allergen Elicits Protection against Challenge Infection under Specific Conditions.* Vaccine. 24, (17), 3581-3590.

34. Sereda, M. J.; Hartmann, S.; Büttner, D. W.; Volkmer, R.; Hovestädt, M.; Brattig, N.; Lucius, R. (2010). *Characterization of the Allergen Filarial Tropomyosin with an Invertebrate Specific Monoclonal Antibody.* Acta Tropica. 116, (1), 61-67.

35. Hilario, E.; Lataro, R. C.; Alegria, M. C.; Lavarda, S. C. S.; Ferro, J. A.; Bertolini, M. C. (2001). *High-Level Production of Functional Muscle A-Tropomyosin in Pichia Pastoris.* Biochemical and Biophysical Research Communications. 284, (4), 955-960.

36. Graeslund, S.; Sagemark, J.; Berglund, H.; Dahlgren, L.-G.; Flores, A.; Hammarstroem, M.; Johansson, I.; Kotenyova, T.; Nilsson, M.; Nordlund, P.; Weigelt, J. (2008). *The Use of Systematic N- and C-Terminal Deletions to Promote Production and Structural Studies of Recombinant Proteins.* Protein Expression and Purification. 58, (2), 210-221.

37. Graeslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; dhe-Paganon, S.; Park, H. W.; Savchenko, A.; Yee, A.; Edwards, A.; Vincentelli, R.; Cambillau, C.; Kim, R.; Kim, S. H.; Rao, Z.; Shi, Y.; Terwilliger, T. C.; Kim, C. Y.; Hung, L. W.; Waldo, G. S.; Pelef, Y.; Albeck, S.; Unger, T.; Dym, O.; Prilusky, J.; Sussman, J. L.; Stevens, R. C.; Lesley, S. A.; Wilson, I. A.; Jochimiak, A.; Collart, F.; Dementieva, I.; Donnelly, M. I.; Eschendeldt, W. H.; Kim, Y.; Stols, L.; Wu, R.; Zhou, M.; Burley, S. K.; Emtage, J. S.; Sauder, J. M.; Thompson, D.; Bain, K.; Luz, J.; Gheyi, T.; Zhang, F.; Atwell, S.; Almo, S. C.; Bonanno, J. B.; Fiser, A.; Swaminathan, S.; Studier, F. W.; Chance, M. R.; Sail, A.; Acton, T. B.; Xiao, R.; Zhao, L.; Ma, L. C.; Hunt, J. F.; Tong, L.; Cunningham, K.; Inouye, M.; Anderson, S.; Janjua, H.; Shastry, R.; Ho, C. K.; Wang, D.; Wang, H.; Jiang, M.; Montelione, G. T.; Stuart, D. I.; Owens, R. J.; Daenke, S.; Schutz, A.; Heinemann, U.; Yokoyama, S.; Bussow, K.; Gunsalus, K. C. (2008). *Protein Production and Purification.* Nature Methods. 5, (2), 135-146.

38. Papaneophytou, C. P.; Kontopidis, G. (2014). *Statistical Approaches to Maximize Recombinant Protein Expression in Escherichia Coli: A General Review.* Protein Expression and Purification. 94, (0), 22-32.

39. Papaneophytou, C. P.; Rinotas, V.; Douni, E.; Kontopidis, G. (2013). *A Statistical Approach for Optimization of Rankl Overexpression in Escherichia Coli: Purification and Characterization of the Protein.* Protein Expression and Purification. 90, (1), 9-19.

40. Xu, J. L.; Banerjee, A.; Pan, S. H.; Li, Z. J. (2012). *Galactose Can Be an Inducer for Production of Therapeutic Proteins by Auto-Induction Using E. Coli Bl21 Strains.* Protein Expression and Purification. 83, (1), 30-36.

41. Daegelen, P.; Studier, F. W.; Lenski, R. E.; Cure, S.; Kim, J. F. (2009). *Tracing Ancestors and Relatives of Escherichia Coli B, and the Derivation of B Strains Rel606 and Bl21(De3).* Journal of Molecular Biology. 394, (4), 634-643.

42. Waegeman, H.; De Lausnay, S.; Beauprez, J.; Maertens, J.; De Mey, M.; Soetaert, W. (2013). *Increasing Recombinant Protein Production in Escherichia Coli K12 through Metabolic Engineering.* New Biotechnology. 30, (2), 255-261.

43. Yoon, S. H.; Han, M.-J.; Jeong, H.; Lee, C. H.; Xia, X.-X.; Lee, D.-H.; Shim, J. H.; Lee, S. Y.; Oh, T. K.; Kim, J. F. (2012). *Comparative Multi-Omics Systems Analysis of Escherichia Coli Strains B and K-12.* Genome Biology. 13, (5).

44. Studier, F. W. (2005). *Protein Production by Auto-Induction in High Density Shaking Cultures.* Protein Expression and Purification. 41, 207-234.

45. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* Int J Hyg Environ Health. 217, (7), 740-50.

46. Johnston, E.; Kamath, S.; Lopata, A.; Schaeffer, P. (2014). *Tus-Ter-Lock Immuno-Pcr Assays for the Sensitive Detection of Tropomyosin-Specific Ige Antibodies.* Bioanalysis. 6, (4), 465-476.

47. Reese, G.; Viebranz, J.; Leong-Kee, S. M.; Plante, M.; Lauer, I.; Randow, S.; Moncin, M. S. M.; Ayuso, R.; Lehrer, S. B.; Vieths, S. (2005). *Reduced Allergenic Potency of Vr9-1, a Mutant of the Major Shrimp Allergen Pen a 1 (Tropomyosin).* Journal of Immunology. 175, (12), 8354-8364.

48. Jirapongsananuruk, O.; Sripramong, C.; Pacharn, P.; Udompunturak, S.; Chinratanapisit, S.; Piboonpocanun, S.; Visitsunthorn, N.; Vichyanond, P. (2008). *Specific Allergy to Penaeus Monodon (Seawater Shrimp) or Macrobrachium Rosenbergii (Freshwater Shrimp) in Shrimp-Allergic Children.* Clinical & Experimental Allergy. 38, (6), 1038-1047.

49. Morgan, J. E.; O'Neil, C. E.; Daul, C. B.; Lehrer, S. B. (1989). *Species-Specific Shrimp Allergens: Rast and Rast-Inhibition Studies.* Journal of Allergy and Clinical Immunology. 83, (6), 1112-1117.

50. Ayuso, R.; Lehrer, S. B.; Reese, G. (2002). *Identification of Continuous, Allergenic Regions of the Major Shrimp Allergen Pen a 1 (Tropomyosin).* International Archives of Allergy and Immunology. 127, (1), 27-37.

51. Ayuso, R.; Reese, G.; Leong-Kee, S.; Plante, M.; Lehrer, S. B. (2002). *Molecular Basis of Arthropod Cross-Reactivity: Ige-Binding Cross-Reactive Epitopes of Shrimp, House Dust Mite and Cockroach Tropomyosins.* International Archives of Allergy and Immunology. 129, (1), 38-48.

52. Ayuso, R.; Sanchez-Garcia, S.; Pascal, M.; Lin, J.; Grishina, G.; Fu, Z.; Ibanez, M. D.; Sastre, J.; Sampson, H. A. (2012). *Is Epitope Recognition of Shrimp Allergens Useful to Predict Clinical Reactivity?* Clinical and Experimental Allergy. 42, (2), 293-304.

53. Zheng, L.-N.; Lin, H.; Pawar, R.; Li, Z.-X.; Li, M.-H. (2011). *Mapping Ige Binding Epitopes of Major Shrimp (Penaeus Monodon) Allergen with Immunoinformatics Tools.* Food and Chemical Toxicology. 49, (11), 2954-2960.

2.6 Summary chapter 2:

Molecular and Immunological Analysis of Tropomyosin from King Prawn *(Melicertus latisulcatus)*

- ✔ The novel investigated amino acid sequence of tropomyosin from King prawn is only 95% identical to Black Tiger prawn, whereas for myosin light chain it is 99.44% identical
- ✔ Two expression systems were compared for the expression of allergenic proteins using recombinant tropomyosin from King prawn
- ✔ With auto-induction the yield of expressed tropomyosin could be increased to 62 mg/l
- ✔ Different IgE binding was observed with crustacean allergic patients for recombinant tropomyosin from King prawn and Black Tiger prawn
- ✔ Tropomyosin from King prawn was registered as allergen, Mel l 1

The purified tropomyosin from King prawn will be used in the next chapter for the development of the LC/qTOF method. Recombinant tropomyosin will be used in the next chapters as positive control.

CHAPTER 3

IDENTIFYING SIGNATURE PEPTIDES FOR ALLERGENIC TROPOMYOSIN TO DISTINGUISH CRUSTACEAN FROM MOLLUSCS

3.1 Introduction

Allergic reactions to shellfish are a severe type of food allergy and can lead to anaphylaxis, which can result in death.[1] Two invertebrate groups, crustacean and molluscs are generally referred to as shellfish and are common constituents in the diet of many populations.[1] To protect allergic consumers food labelling regulations are enforced in various countries. The legislation in the European Union and Canada require a different label each for crustacean and molluscs.[2] Crustacean and molluscs incorporate a large number of species and for detection purposes it is necessary to decide whether the intent is to identify one specific species or distinguish crustacean from molluscs as a group. From the legislation and the food industries perspective the differentiation between crustacean and molluscs as groups seems to be presently more important. However, current methods are not able to distinguish between these two groups. Most commonly for shellfish allergen detection and quantification ELISAs are used. Presently experimental developed ELISAs as well as commercial available ELISA kits for shellfish allergen detection are using tropomyosin (TM) for the detection and quantification of shellfish allergens.[3-7] TM is the major allergen in the shellfish group, however, it is also a pan-allergen that is cross-reactive with mites, cockroaches and nematodes.[1, 8] The amino acid sequence of TM is highly conserved throughout the invertebrates. In detail the amino acid sequence identity within crustacean is 88-100%[9], within the molluscs 60-90%[10] and between both shellfish groups 50-70%.[1, 8] Due to amino acid sequence identity and resulting cross-reactivity, ELISAs cannot distinguish easily between crustacean and molluscs. Recently mass spectrometry has been applied for food allergen detection and quantification to overcome the drawbacks of ELISAs.[2] Abdel Rahman investigated in different studies[11-14] three crustacean species to identify species specific signature peptides using TM. In detail, he reported the signature peptide ANIQLVEK for Black Tiger prawn,[11] SEEEVFGLQK for Northern prawn[12] and SQLVENELDHAQEQLSAATHK for Snow crab.[13, 14] Hence, it was shown that, despite the amino acid homology for shellfish and especially for crustacean, it is possible to apply mass spectrometry for species specific tropomyosin detection. However, for the TM detection using mass spectrometry, only three crustacean species have presently been investigated. To increase the knowledge on different crustacean and molluscs

species twenty-two different shellfish species will be analysed in this chapter by mass spectrometry, in particular liquid chromatography (LC) coupled with a time of flight (qTOF) mass spectrometer. This analysis will identify and show which peptides are easily detected using mass spectrometry. The aim of this chapter is to identify signature peptides of TM unique for crustacean and unique for molluscs, hence the two subgroups forming the shellfish group causing allergic reactions and can be used to distinguish these groups, fulfilling the demand from legislation and industry.

3.1.1 Aims

The aims of this chapter are:

- Analysing 11 different crustacean species and 11 different molluscs species (total of 22 shellfish species) by mass spectrometry (LC/qTOF)
- Evaluating peptides derived of tropomyosin that are easily detected and identified by LC/qTOF
- Aligning evaluated peptides with more tropomyosin sequences from crustacean and molluscs species to identify possible signature peptides
- Identifying signature peptides for tropomyosin for the differentiation of crustacean and molluscs

3.2 Materials and methods

The equipment used in this chapter is liquid chromatography (LC) coupled with a quadruple time of flight (qTOF) mass spectrometer. The LC/qTOF system is an Acquity ultra performance liquid chromatography (UPLC) coupled with an electrospray ionisation (ESI) interface to a Xevo G2 QTof mass spectrometer (Waters Corporation, USA).

The tryptic digested peptides were separated by an Acquity UPLC BEH 300, C18 column (1.7 µm 2.1 x 1500 mm (Waters Corporation, Australia)), protected by a guard column containing the same stationary material. The flow rate was 0.4 ml/min and injection volume was 20 µl using full loop setting. The temperature for the column oven was 45°C and for the autosampler 10°C. Mobile phase A was 0.1% formic acid in water (Thermo Fisher Scientific, Australia) and solvent B was 0.1% formic acid in acetonitrile (Thermo Fisher Scientific, Australia).

The qTOF instrument was acquiring in positive resolution mode. For calibration of the qTOF sodium formate (Sigma-Aldrich, USA) (1 ppm) was used. Leucine Enkephalin (Waters Corporation, Australia) (2 ng/µl) was used as lock spray and constantly injected and acquired in the third mass channel.

The Xevo G2 QTof mass spectrometer can acquire data as “data independent acquisition” (MS^E) or as “data dependent acquisition” (DDA). The DDA method is the more traditional way of acquiring MS data. Briefly, peptides become ionised and a signal can be detected. When the signal rises above the noise in full scan mode, the peptide will be selected for fragmentation, generating MS/MS data. This approach is extremely powerful; however, peptides are randomly selected depending on their signal intensity. Therefore the dynamic range depends on the detection system and the sample complexity. Overall, DDA is a serial biased discontinuous process. In contrast, MS^E is a parallel unbiased continuous process; this increases the number of peptides that can be detected and the reproducibility, accuracy and sensitivity of peptide signals, leading to a potentially enhanced peptide and protein identification, including low abundance peptides. However, MS^E data requires more data analysis, thus data is more complex and includes more information. However, currently there is a lack of complete software solutions to fully exploit the MS^E data[48]. Nevertheless, the raw MS^E data becomes

processed by three different algorithms in the ProteinLynx Global Server, possible increasing statistical errors. The dynamic range is independent from the sample complexity, but dependent on the detection system and the chromatography separation of the precursor. Applying the mixed precursor integration minimises the inherent biases associated with DDA, leading to higher resolution and better suitability for label free quantification.

3.2.1 LC/qTOF method development

The preliminary tune page parameters were: Capillary voltage 5 kV; Sampling cone voltage 40 kV; Extraction cone voltage 5 kV; Source temperature 150°C; Desolvation temperature 500°C; Cone gas flow 30 l/h and Desolvation gas flow 900 l/h, respectively. Moreover, the lock spray was infused at a flow rate of 20 µl/min.

The preliminary LC gradient was: 0-1 min 2%B; 2.5-35 min 10-50%B; 35-47 min 50-90%B; 47-52 min 90%B; 52-60 min 2%B.

3.2.1.1 Preliminary MS^E method

In MS^E mode a mass range of 100-2000 m/z was chosen for the acquisition of the data. The data was acquired in continuum mode, with 3 scans to average and a mass window of +/- 0.5 m/z. The scan time was 1 sec, with an interval of 30 sec. The collision energy was set to 6 eV.

3.2.1.2 Preliminary DDA method

In Data Dependent Acquisition (DDA) mode a mass range of 400-1950 m/z was chosen, with an intensity threshold of 1000. The data was acquired in continuum mode. The survey scan time was 0.5 sec and survey interscan time was 0.01 sec. The two highest compounds were selected for tandem mass spectrometry mode (MS/MS mode).

For MS/MS mode a mass range of 100-1950 m/z was chosen. As precursor the two highest (intensity) peaks were chosen for the occurring time window. The

MS/MS mode switched back to MS mode when the total ion chromatogram (TIC) was falling below threshold (100) or after 2.5 sec. The MS/MS scan time was 1 sec. The charge state tolerance window and the charge state extraction window was set to 3 and 2, respectively. For low masses (100 Da) and high masses (1950 Da) the collision energy was ramped from 25-50 eV, with a survey cone voltage of 40 V.

3.2.2 Final LC/qTOF method

All digested proteins were analysed four times, using the combination of final LC and final MS methods. These four final methods are: 15 minutes MS^E, 60 minutes MS^E, 15 minutes DDA and 60 minutes DDA.

The final tune page parameters were: Capillary voltage 2 kV; Sampling cone voltage 20 kV; Extraction cone voltage 2 kV; Source temperature 120°C; Desolvation temperature 500°C; Cone gas flow 30 l/h and Desolvation gas flow 900 l/h, respectively. Moreover, the lock spray was infused with at a flow rate of 5 µl/min.

3.2.2.1 Final LC methods

Method A (15 min): The tryptic digested peptides were separated by following gradient: 0-1 min 5%B; 1-13.2 min 5-50%B; 13.2-14 min 50-90%B; 14-14.3 min 90%B; 14.3-14.5 min 90-5%B; 14.5-15 min 5%B. The flow from the LC was diverted into waste from 0-0.5 min and from 14.3-15 min.

Method B (60 min): The tryptic digested peptides were separated by following gradient: 0-1 min 5%B; 2.5-35 min 10-50%B; 35-47 min 50-90%B; 47-52 min 90%B; 52-60 min 5%B. The flow from the LC was diverted into waste from 0-0.5 min and from 54-60 min.

3.2.2.2 Final MS^E method

In the MS^E mode a mass range of 100-1950 m/z was chosen, for the detection of tryptic digested peptides. The data was acquired in continuum mode, with 3 scans

to average and a mass window of +/- 0.5 m/z. The scan time was 0.5 sec, with an interval of 30 sec. The collision energy was set to 6 eV over the time period of the acquisition.

3.2.2.3 Final DDA method

In the DDA mode a mass range of 300-1950 m/z was chosen, with an intensity threshold of 1000. The data was acquired in continuum mode. The survey scan time was 0.1 sec and survey interscan time was 0.01 sec. The three highest compounds were selected for MS/MS mode.

For MS/MS mode a mass range of 100-1950 m/z was chosen. As precursor the two highest (intensity) peaks were chosen for the occurring time window. The MS/MS mode switched back to MS mode when the TIC was falling below threshold (100) or after 0.6 sec. The MS/MS scan time was 1 sec. The charge state tolerance window and the charge state extraction window was set to 3 and 2, respectively. For low masses (100 Da) the collision energy was ramped from 20-40 eV, for high masses (1950 Da) the collision energy was ramped from 25-50 eV, with a survey cone voltage of 40 V.

3.2.3 Data processing

All raw generated and uninterpreted data was processed using ProteinLynx Global Server (PLGS) v2.3 (Waters Corporation, Australia) and converted into pkl files. The converted pkl were searched with Mascot daemon search engine (version 2.4) and Swissprot database (549,832 entries – November 2015) with a 0.1 Da tolerance against the database generated theoretical peptide ion masses and a minimum of one matched peptide. The preliminary Mascot settings used are variable modifications of carbamidomethyl-C and N-terminus and oxidation of M (methionine). Up to 20 missed cleavage sites were allowed. As final Mascot search parameters variable modifications of carbamidomethyl-C and N-terminus, deamidation of N (asparagine) and Q (glutamine) and oxidation of M were chosen. Up to seven missed cleavage sites were allowed.

3.2.4 Preliminary protein digestion methods

3.2.4.1 Protein digestion using trypsin spin columns

Soluble protein extract or purified TM from KP was reduced with dithiothreitol (Sigma-Aldrich, USA) (10 mM) at 60°C for 30 min and alkylated in the dark with iodoacetamide (Sigma-Aldrich, USA) (10 mM) at 37°C for 30 min. The solution was loaded on a trypsin spin column (Sigma-Aldrich, USA), which has been washed, equilibrated and prepared according to the manufacturer's instructions. The samples were incubated for various time points, varying from 15 minutes up to 22 hours, at room temperature (RT) or 37°C. After incubation the tryptic digested peptide were eluted twice with 100 µl of 0.1% formic acid (Thermo Fisher Scientific, Australia).

3.2.4.2 Protein digestion using immobilised TPCK trypsin

Protein extracts were reduced and alkylated as mentioned above. Immobilised TPCK Trypsin (Thermo Fisher Scientific, Australia) was added to the extracts in and enzyme to substrate ratio 1:5 prior to digestion. The manufactures instructions requested that the digestion will take place at 37°C, therefore no incubate experiment was carried out at RT. Moreover, the digestion time was increased to two hours, being the shortest digestion time recommended by the manufacturer. After incubation the resin was removed by centrifugation (1 minute, 800 g) and 50 µl formic acid was added to adjust the pH.

3.2.4.3 Reagents tested to possible improve digestion efficiency

To potential improve digestion and reduced missed cleavage sites two reagents were added before samples were reduced and alkylated: (1) Rapigest (Waters Corporation, USA) was added to protein extract in a ratio of 1:10 (v/v) and (2) urea (Sigma-Aldrich, USA) was added to protein in a concentration of 800 mM.

3.2.5 Final protein digestion method

A total of 200 µl of soluble whole heated protein extract was reduced with dithiothreitol (10 mM) at 60°C for 30 min and alkylated in the dark with iodoacetamide (10 mM) at 37°C for 30 min. The solution was loaded on a trypsin spin column (Sigma-Aldrich, USA), which has been washed, equilibrated and prepared according to the manufacturer's instructions. The samples were incubated for 15 minutes at room temperature and eluted twice with 100 µl of 0.1% formic acid. The eluted peptides were analysed by LC/qTOF.

3.2.6 SDS-PAGE and immunoblotting

SDS-PAGE and immunoblotting was performed as explained in detail in chapter 2. Briefly, digested protein extracts were separated on a 12% polyacrylamide gel. The protein components were resolved at 200 V until the tracker dye reached the base.[15, 16] To investigate if digested proteins can be still be detected by antibodies, monoclonal anti-tropomyosin antibody and in-house[17] polyclonal anti-crustacean antibody were applied.

3.2.7 Recombinant TM and purified TM

The recombinant TM from King prawn was expressed using auto-induction as explained in detail in chapter 2 and as published in Koeberl et al.[15] Recombinant TM from Black Tiger prawn was expressed as described in Kamath et al.[3, 18, 19] Tropomyosin was purified from KP, using the same strong anion-exchange purification method as described in chapter 2. The same purification was performed using whole heated extract and raw extract from KP, referred as "purified whole heated TM" and "purified raw TM".

3.2.8 Shellfish species analysed by mass spectrometry

Table 3.10 summaries the different shellfish species that have been analysed by the final LC/qTOF methods, grouped into their phyla and subgroups of edible shellfish species, including the GenBank Accession number, when available. For analysing the shellfish species whole heated protein extracts were made as

described in chapter 2. Briefly, the shellfish species were obtained for a local market (Queensland, Australia) and transported to the laboratory on ice. The crustacean species were heated including their outer shell, whereas for mollusc species only the edible muscle parts were heated. The detailed preparation of sample extracts, including the estimated protein concentration by Pierce® 660 nm Protein Assay using BSA as standard, is summarised in table 3.10. The samples were heated at 100°C in PBS for 20 minutes, after cooling only the edible muscles pieces were shredded into small pieces and further homogenised. The total homogenised extract was kept at 4°C overnight while continuously shaking. The protein extracts were centrifuged, followed by sterile filtration and stored at -80°C until further use. The species Jade Hybrid Tiger abalone (*Haliotis laevigata x Haliotis rubra*) will be referred as Hybrid abalone in this thesis.

3.2.9 Tropomyosin analysis

The phylogenetic analysis was performed for tropomyosin to analyse the evolutionary relationship of the allergenic and non allergenic tropomyosin using 106 tropomyosin sequences. The tree was constructed using MEGA6 software and applying the Neighbour-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches. The evolutionary distances were computed using the Poisson correction method and were in the units of the number of amino acid substitutions per site. The Pair wise deletion was applied to all positions containing gaps and missing data. The analysis involved 106 amino acid sequences for tropomyosin recruited from the Genbank. There were a total of 289 positions in the final dataset for tropomyosin.

Table 3.1: The 106 tropomyosin sequences from the NCBI database with the GenBank accession numbers, including the common name, scientific name are sorted alphabetically by their subgroup and the allergen name is given in brackets, when registered with IUIS.

Common name	Scientific name	GenBank accession number	Allergen name
Prawns:			
Brown shrimp, Common shrimp, Bay shrimp	*Crangon crangon*	ACR43473.1	(Cra c 1);
Northern Brown shrimp	*Penaeus aztecus*	AAZ76743.1;	(Pen a 1)
Fleshy prawn, Chinese White shrimp, Oriental shrimp	*Fenneropenaeus chinensis*	ADA70137.1	
Banana prawn	*Fenneropenaeus merguiensis*	ADC55381.4	
White leg shrimp, Pacific White shrimp	*Litopenaeus vannamei*	ACB38288.1	(Lit v 1)
Oriental River shrimp	*Macrobrachium nipponense*	AHJ10947.1	
Giant Freshwater prawn	*Macrobrachium rosenbergii*	ADC55380.1;	
Greasyback shrimp	*Metapenaeus ensis*	Q25456.1	(Met e 1)
Kuruma prawn	*Marsupenaeus japonicas*	BAF47263.1	
King prawn	*Melicertus latisulcatus*	AGF86397.1	(Mel l 1)
Japanese Mantis Shrimp	*Oratosquilla oratoria*	BAF95206.1	
Cold Water prawn	*Pandalus eous*	BAF47264.1	
Northern shrimp.	*Pandalus borealis*	CBY17558.1	
Indian prawn	*Penaeus indicus*	[20]	
Black Tiger prawn	*Penaeus monodon*	ADM34184.1	(Pen m 1)
Mantis shrimp	*Squilla aculeata*	AAR87379.1	
Mantis shrimp	*Squilla oratoria*	ABQ57495.1	
Crabs:			
Crucifix crab	*Charybdis feriata*	Q9N2R3.1	(Cha f 1);
Snow crab	*Chionoecetes opilio*	A2V735.1	
Snow crab	*Chionoecetes opilio*	BAF47267.1	
Hair crab.	*Erimacrus isenbeckii*	BAF47269.1	
Hair crab.	*Erimacrus isenbeckii*	BAF47268.1	
Chinese Mitten crab	*Eriocheir sinensis*	ABO71783.1	
Red King crab	*Paralithodes camtschaticus*	BAF47266.1	
Red King crab	*Paralithodes camtschaticus*	BAF47265.1	
Blue Swimmer crab	*Portunus pelagicus*	AGE44125.1	
Blood-spotted Swimming crab	*Portunus sanguinolentus*	ABL89183.1	
Swimming crab	*Portunus trituberculatus*	ABS12234.1	
Orange mud crab	*Scylla olivacea*	AAX37289.1	
Serrated Swimming Crab	*Scylla serrata*	ABS12233.1	
Lobsters and crayfish:			
Yabby	*Cherax destructor*	AGW22428.1	
American lobster	*Homarus americanus*	O44119.1	(Hom a 1)
Southern Rock lobster, Red Rock lobster, Spiny Rock lobster	*Jasus edwardsii*	AGW22426.1	
South African Cape Rock lobster	*Jasus lalandii*	AFY98827.1	
Red Swamp Crayfish	*Procambarus clarkii*	ACN87223.1	
Chinese Spiny lobster	*Panulirus stimpsoni*	O61379.1;	(Pan s 1)
Chinese Spiny lobster	*Panulirus stimpsoni*	AAC38996.1	
Slipper lobster	*Thenus orientalis*	AGW22427.1;	
Krill:			

Pacific krill	*Euphausia pacifica*	BAF76431.1;	
Krill	*Euphausia superba*	BAF76430.1	
Insects:			
German cockroach	*Blattella germanica*	AAF72534.1	(Bla g 7)
American cockroach	*Periplaneta Americana*	ACS14052.1	(Per a 7)
Red flour beetle	*Tribolium castaneum*	XP 967128.1	
Mites:			
Flour mite	*Acarus siro*	ABL09305.1	(Aca s 10)
Brown Legged Grain mite	*Aleuroglyphus ovatus*	AAX37287.1	
Storage mite	*Blomia tropicalis*	ABU97466.1	(Blo t 10)
Furniture mite	*Chortoglyphus arcuatus*	AEX31649.1	
American house dust mite	*Dermatophagoides farina*	Q23939.2	(Der f 10)
European house dust mite	*Dermatophagoides pteronyssinus*	O18416.1	(Der p 10)
Storage mite, Furniture mite, Food mite	*Glycyphagus domesticus*	AAQ54614.1	
Hay mite	*Lepidoglyphus destructor*	Q9NFZ4.1	(Lep d 10)
Itch mite	*Sarcoptes scabiei*	AFH08744.1	
Nematodes:			
Herring worm	*Anisakis simplex*	AEQ28167.1	(Ani s 3)
Roundworm	*Caenorhabditis elegans*	NP_001263595.1	
Soybean Cyst nematode	*Heterodera glycines*	AAQ12016.1	
Gastropods:			
Freshwater snail	*Biomphalaria glabrata*	AAA27817.1	
Ass's-ear abalone	*Haliotis asinine*	**AAP85231.1**	
Japanese abalone	*Haliotis discus discus*	BAH10148.1	
Coloured abalone	*Haliotis diversicolor*	Q9GZ71.1	
Red abalone	*Haliotis rufescens*	CAA53028.1	
Garden snail	*Helix aspersa*	O97192.1	(Hel a 1)
Neptunea polycostata	*Neptunea polycostata*	BAH10150.1	
Horned turban	*Turbo cornutus*	BAH10149	
Cephalopods:			
Hawaiian Bobtail squid	*Euprymna scolopes*	AAN35188.1	
Spear squid	*Loligo bleekeri*	ADD64463.1	
Common octopus	*Octopus vulgaris*	BAE54433.1	
Red Flying squid	*Ommastrephes bartramii*	BAE54432.1	
Golden cuttlefish	*Sepia esculenta*	BAE54429.1	
Common cuttlefish	*Sepia officinalis*	AEE87268.1	
Bigfin Reef squid	*Sepioteuthis lessoniana*	BAE54430.1	
Japanese Flying squid	*Todarodes pacificus*	BAE54431.1	(Tod p 1)
Bivalves:			
Bay scallop	*Argopecten irradians*	AAX37290.1	
Akazara scallop	*Chlamys nipponensis*	BAA36219.1	
Tasmanian oyster	*Crassostrea gigas*	BAH10152.1	
Eastern oyster	*Crassostrea virginica*	AAC61869.1	
Mangrove oyster	*Crassostrea rhizophorae*	AAR17060.1	
Egg cockle	*Fulvia mutica*	AEV23866.1	
Asiatic hard clam	*Meretrix lyrata*	AEV23864.1	
Common Orient clam	*Meretrix meretrix*	AEV23865.1	
Noble scallop	*Mimachlamys nobilis*	Q9GZ69.1	
Yesso scallop	*Mizuhopecten yessoensis*	BAB17858.1	
Blue mussel	*Mytilus edulis*	Q25457.1	
Mediterranean mussel	*Mytilus galloprovincialis*	P91958.1	
Asian Green mussel	*Perna viridis*	Q9GZ70.1	
Sakhalin surf-clam	*Pseudocardium sachalinensis*	BAH10154.1	
Japanese Carpet shell	*Ruditapes philippinarum*	BAH10157.1	
Blood cockle	*Scapharca broughtonii*	BAH10151.1	
Razor clam	*Sinonovacula constricta*	ABU53681.1	

Gould's Razor shell	*Solen strictus*	BAH10156.1	
Malaysian cockle	*Tegillarca granosa*	AEV23866.1	
Horse clam	*Tresus keenae*	BAH10155.1	
Mammals:			
Cow	*Bos taurus*	NP_001013608.1	
Horse	*Equus caballus*	P02561.2	
Human	*Homo sapiens*	AAB59509.1	
Mouse	*Mus musculus*	NP_001157723.1	
Pig	*Sus scrofa*	NP_001090952.1	
Chicken:			
Chicken	*Gallus gallus*	AAA49112.1	
Fish:			
Alaska pollock	*Gadus chalcogrammus*	BAC44994.1	
Channel catfish	*Ictalurus punctatus*	NP_001187627.1	
Golden Grey mullet	*Liza aurata*	P84335.1	
Rainbow smelt	*Osmerus mordax*	ACO08849.1	
Salmon	*Salmo salar*	ACN10871.1	
Tuna	*Thunnus thynnus*	BAJ11924.1	
Zebrafish	*Danio rerio*	NP_001019638.1	
Zebrafish	*Danio rerio*	NP_998323.1	
Frogs:			
African clawed frog	*Xenopus laevis*	CAA36488.1	

3.2.10 Peptide analysis

To be able to identify signature peptides for crustacean or molluscs 106 different tropomyosin sequences were used. The same 106 tropomyosin sequences were used from NCBI with the GenBank accession numbers as listed above. Identified tryptic peptides, based on the sequence of Black Tiger prawn (Pen m 1), were aligned. Two more alignments were performed using tryptic peptides based on squid (*Todarodes pacificus*) (Tod p 1).

3.3 Results

3.3.1 Investigation of digestion method of TM using purified whole heated TM from KP and preliminary LC/qTOF settings

To investigate the LC/qTOF analysis as well as the digestion methods, firstly purified TM from whole heated KP extract was studied. In the first step different digestion duration and digestion temperatures were investigated, followed by two considerations of the digestion efficiency of the applied trypsin spin columns. Thus, as first step, 50 µg purified TM from whole heated extracts was loaded on trypsin spin column and incubated for four different time points, 15 minutes, 30 minutes, 1 hour and 2 hours, respectively. This temperature range was chosen, as the manufactures instructions specify that digestion at room temperature (RT) for 15 minutes is suitable. The LC/qTOF data were converted into pkl files and searched with Mascot. To examine the quality of analysis the mascot score and the number of missed cleavage sites are reported in table 3.1, as example for identified peptides, the peptide score and the e-value see figure B2.6.

As first step the digestion time and the digestion temperature were evaluated. Summarised in table 3.1, tropomyosin could be detected in all digested samples and different LC/qTOF methods applied. Moreover, the Mascot score as well as the missed cleavage sites are similar for different digestion time and digestion temperature investigated. The minor differences in the Mascot score and the missed cleavage sites can be explained by measurement uncertainties. TM is highly heat stable, thus Mascot scores are not obviously decreasing with increased digestion time and temperature. However, the digestion efficiency does not improve when digestion temperature and duration are increased. Therefore, it was conclude that digestion at RT for 15 minutes, as recommended by the manufacturer, is suitable for digestion and detection of TM.

Table 3.2: Mascot scores and missed cleavage sites for preliminary LC/qTOF results for different digestion times and temperature using purified whole heated TM from KP.

Purified tropomyosin from King prawn (whole heated extract)				
Time of digestion	**Temperature**	**MS method**	**TM Mascot score**	**Missed cleavage sites**
15 min	RT	MSE	62	6
		DDA	893	0
	37°C	MSE	61	6
		DDA	425	0
30 min	RT	MSE	62	5
		DDA	798	0
	37°C	MSE	131	3
		DDA	1038	0
1 hour	RT	MSE	50	5
		DDA	879	0
	37°C	MSE	86	0
		DDA	1105	3
2 hours	RT	MSE	59	2
		DDA	369	0
	37°C	MSE	65	1
		DDA	435	0

Table 3.3: Repeated analysis of purified whole heated TM from KP using trypsin spin columns form different production lines, showing Mascot scores and missed cleavage sites.

Purified tropomyosin from King prawn (whole heated extract)			
Replica	**MS method**	**TM Mascot score**	**Missed cleavage sites**
Replica 1	MSE	54	0
	DDA	311	1
Replica 2	MSE	54	0
	DDA	277	0
Replica 3	MSE	48	0
	DDA	393	1
Replica 4	MSE	50	0
	DDA	317	1
Replica 5	MSE	54	0
	DDA	262	1
Replica 6	MSE	48	0
	DDA	275	1

As next step uncertainties in the digestion method were considered, due to applied trypsin spin columns. Therefore, 25 µg of tropomyosin were digested for 15 minutes at room temperature using six different trypsin spin columns. These columns investigated were supplied by Sigma-Aldrich in three different kits. The Mascot score and missed cleavage sites are shown in table 3.2, demonstrating that the trypsin digestion is reproducible throughout different production lines from Sigma-Aldrich.

Table 3.4: Testing for trypsin digestion reproducibility applying the same trypsin spin column and purified whole heated TM from KP, summarising Mascot scores and missed cleavage sites.

Purified tropomyosin from King prawn (whole heated extract)			
Column used	**MS method**	**TM Mascot score**	**Missed cleavage sites**
1st time	MSE	187	5
	DDA	619	0
2nd time	MSE	217	4
	DDA	610	10
3rd time	MSE	180	2
	DDA	619	13
4th time	MSE	141	1
	DDA	735	16
5th time	MSE	155	2
	DDA	601	3
6th time	MSE	191	2
	DDA	603	14
7th time	MSE	117	1
	DDA	427	10
8th time	MSE	228	0
	DDA	532	9

As a final step, the performance of trypsin in one column was investigated, thus the manufactures specify that the column can be reused at least four times. To confirm this, 50 µg of tropomyosin was digested using the same trypsin spin column eight times. The column was equilibrated and re-equilibrated according to the manufacturer's instructions and digestion time was 15 minutes at RT. As summarised in table 3.3 the column seems to give reproducible results reusing the column up to eight times. However, the number of missed cleavage sites seems to

increase when trypsin spin columns were reused. Therefore for the final experiments for each shellfish species extract a new trypsin spin column was used for protein digestion.

Overall, in this section 3.3.1 it was shown that trypsin spins columns can digest purified whole heated TM from KP very rapidly. The digestion time can be reduced to 15 minutes at RT. The different Mascot scores for the MS^E and the DDA methods which can be observed in tables 3.1 to 3.3 can be explained with the different amount of purified whole heated TM used for digestion. Moreover, the settings of MS^E and DDA methods were changed between set up experiments, to improve detection and identification simultaneously. All investigated aspects demonstrate that the digestion method is suitable and does not need to be improved. However, these results generated in this section were obtained using purified whole heated TM from KP, therefore as next step the raw and whole heated extract from KP will be investigated to demonstrate that digestion is still suitable in more complex matrices.

3.3.2 Investigation of digestion method of TM using raw and whole heated extract from KP and preliminary LC/qTOF settings

In this section it was investigated if the digestion time and temperature needs to be adjusted when using a protein extract in comparison to purified TM, hence the matrix is more complex. Therefore whole heated extract and raw extract from KP were applied. The first extract investigated was the whole heated extract from King prawn, thus it was shown in chapter 2, due to heat treatment less proteins are present, leading to enhanced TM in this extract. Therefore, 200 µg of King prawn whole heated extract were digested, using trypsin spin columns at room temperature and 37°C. Digestion time was 15 minutes, 30 minutes, 2 hours, 4 hours, 6 hours and 22 hours. As shown in table 3.4 with whole heated extract the Mascot scores decreases as the incubation time and temperature increases. Arginine kinase (AK) is summarised in this table as well, because it was identified in whole heated extract, moreover, as shown below in raw KP extract AK was the major identified protein, not TM. However, AK was not always identified in whole heated extract, due to temperature sensitivity of arginine kinase. AK is an

acceptable protein for prawn identification, thus it is a known minor allergen in crustacean.

Table 3.5: Mascot scores and missed cleavage sites obtained from whole heated King prawn extract for TM and AK for different digestion time and temperature applied.

King prawn - whole heated extract					
Time of digestion	**Temperature**	**MS method**	**TM Mascot score**	**Missed cleavage sites**	**AK Mascot score**
15 min	RT	MSE	503	0	not identified
		DDA	1241	0	84
	37°C	MSE	255	0	not identified
		DDA	676	0	not identified
30 min	RT	MSE	288	1	not identified
		DDA	1264	1	63
	37°C	MSE	108	1	not identified
		DDA	140	1	not identified
2 hours	RT	MSE	394	7	not identified
		DDA	604	2	66
	37°C	MSE	369	3	not identified
		DDA	556	2	62
4 hours	RT	MSE	401	4	not identified
		DDA	629	1	not identified
	37°C	MSE	345	3	45
		DDA	500	0	not identified
6 hours	RT	MSE	303	0	not identified
		DDA	648	9	not identified
	37°C	MSE	372	2	not identified
		DDA	639	14	45
22 hours	RT	MSE	401	0	not identified
		DDA	444	5	not identified
	37°C	MSE	377	1	42
		DDA	444	5	not identified

Overall, table 3.4 shows, that the digestion of 15 min at RT results in high Mascot scores and no missed cleavage sites applying whole heated protein extract from King prawn, a more complex matrix compared to purified TM. Moreover, longer digestion time leads to lower Mascot scores and more missed cleavage sites, which is undesirable. Therefore for whole heated KP extract the digestion of 15 min at RT is advisable.

The next step was to see the digestion and identification performance of LC/qTOF method using raw extracts. Therefore 200 µg of King prawn raw extract were digested using the same digestion durations and temperatures as for the whole heated extract. However, as shown in table 3.5, tropomyosin could not be identified for any settings tested. Nevertheless, arginine kinase could be detected in all samples. AK is temperature sensitive, hence with increasing digestion time and temperature the Mascot score for AK identification decreases. Therefore for raw KP extract the digestion of 15 min at RT is advisable.

Nevertheless, it was unexpected that TM could not be identified in KP raw extract. Therefore, to confirm that tropomyosin is present in the raw KP extract, the digested whole heated and raw extract were analysed by SDS-PAGE and immunoblotting. As displayed in figure 3.1 the antibodies detect the tropomyosin in raw as well as in whole heated extract, when the extracts are undigested, with the visible band around 37 kDa. When raw and whole heated extracts are digested at room temperature for various time points, the digestion profile does not visible change on the SDS-PAGE gel and immunoblotting. This confirms that digestion at room temperature for 15 minutes is adequate for the digestion of the extracts. The monoclonal anti-tropomyosin antibody is not binding to tropomyosin after digestion; therefore it is assumed that the epitope region on TM for the monoclonal anti-tropomyosin antibody was successfully affected by digestion. However, the in-house anti-crustacean antibody still recognises digested fragments of tropomyosin, with the main binding region below 20 kDa, possible indication its residual allergenicity. The digested protein separation was achieved on a 12% SDS-PAGE, however, this percentage is not suitable for peptide separation, thus peptides have a lower molecular weight. Therefore it is assumed that bands visible below the 20 kDa are fully digested peptides converged at the edge of the polyacrylamide gel. Overall, the anti-crustacean antibody shows that TM is successfully digested using the trypsin spin columns, thus the TM band at approximately 37 kDa is not detected.

Table 3.6: Mascot scores and missed cleavage sites obtained from raw King prawn extract for TM and AK for different digestion time and temperature applied.

King prawn - raw extract					
Time of digestion	**Temperature**	**MS method**	**TM Mascot score**	**AK Mascot score**	**Missed cleavage sites**
15 min	RT	MSE	not identified	239	1
		DDA	not identified	476	1
	37°C	MSE	not identified	147	0
		DDA	not identified	365	1
30 min	RT	MSE	not identified	34	0
		DDA	not identified	231	1
	37°C	MSE	not identified	61	0
		DDA	not identified	227	1
2 hours	RT	MSE	not identified	189	1
		DDA	not identified	567	2
	37°C	MSE	not identified	86	1
		DDA	not identified	442	1
4 hours	RT	MSE	not identified	104	0
		DDA	not identified	369	1
	37°C	MSE	not identified	180	0
		DDA	not identified	624	2
6 hours	RT	MSE	not identified	115	0
		DDA	not identified	324	3
	37°C	MSE	not identified	80	0
		DDA	not identified	225	1
22 hours	RT	MSE	not identified	98	0
		DDA	not identified	381	3
	37°C	MSE	not identified	66	0
		DDA	not identified	45	0

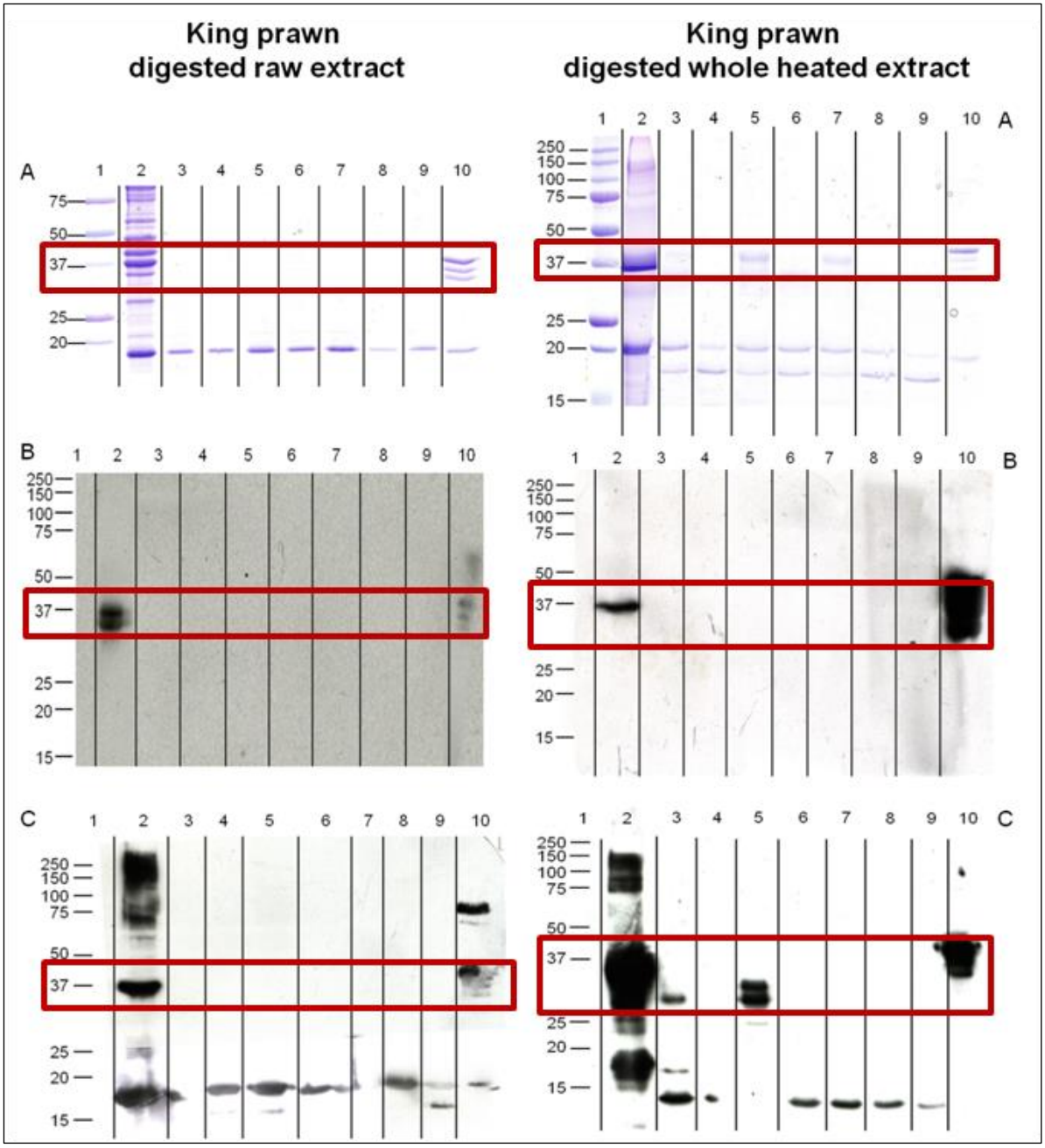


Figure 3.1: SDS-PAGE and immunoblotting analysis of digested King prawn protein extracts. Different protein extracts were digested at RT for various time points, followed by (A) SDS-PAGE separation. Immunoblotting was performed by (B) monoclonal anti-tropomyosin antibody and (C) polyclonal anti-crustacean antibody. The red boxes highlight the molecular weight where undigested TM is expected. Lane numbers represent: 1. Marker; 2. Undigested extract; 3. Extract digested for 15 min; 4. Extract digested for 30 min; 5. Extract digested for 1 hour; 6. Extract digested for 2 hours; 7. Extract digested for 4 hours; 8. Extract digested for 6 hours; 9. Extract digested for 22 hours; and 10. rTM KP.

Nevertheless, some bands are visible for digestion at two time points (digestion time of 1 hour and 4 hours) by SDS-PAGE gel for digested whole heated extracts in figure 3.1. Moreover, applying the anti-crustacean antibody, bands are also visible for digestion time 15 minutes and 1 hour. However, the monoclonal anti-tropomyosin antibody does not detect TM for any time point. Thus, it is hypothesised that these bands are (1) not generated of TM, but potential other proteins present in the whole heated extract or (2) not fully digested peptides of TM for these time points, which confirms the results of reported Mascot scores, that increased digestion duration does not improve the digestion efficiency. Overall, the protein analysis by SDS-PAGE and immunoblotting show that TM is present in raw KP extract. Moreover, that the digestion using trypsin spin columns for 15 min at RT is suitable to digest most proteins as visualised by SDS-PAGE separation applying protein staining and antibody detection.

As it was shown by SDS-PAGE and immunoblotting, raw KP extract contains TM. Therefore as next step different reagents were added to the protein extracts to investigate if TM can be identified by LC/qTOF in raw King prawn extracts. RapiGest or urea was added before digestion, hence these reagents are suitable for LC/qTOF applications and both should improve digestion efficiency applying enzymes. Moreover, a different trypsin (Immobilised TPCK trypsin) was tested for digestion efficiency. The digestion time was two hours for this comparison, thus it is the shortest digestion duration as recommended by the immobilised trypsin manufacturer. The Mascot scores and missed cleavage sites for different digestions are summarised in table 3.6, using total of 200 µg protein for digestion. Interestingly, using the immobilised trypsin resulted in poor identification for arginine kinase. However, applying immobilised trypsin, with both reagent and the DDA method, TM was identified, although the Mascot score is low in comparison to whole heated extract. Using the trypsin spin column, TM was identified in one experiment by the DDA method. However, generally the addition of detergents did not result in successful identification of TM in raw KP extract. The TM identification extract based on Mascot score, using whole heated KP extract, seems very similar, in absence or presence of reagents. Moreover, results in table 3.6 demonstrate that AK is temperature sensitive.

Table 3.7: Digestion of KP raw and whole heated extracts using reagents for possible better digestion, potential resulting in an improved identification of TM. Mascot scores and missed cleavage sites obtained from King prawn extracts for TM and AK for different digestion time, temperature and reagents are summarised in this table.

King prawn - raw extract						
Time of digestion	**Temperature**	**Trypsin + Reagent added**	**MS method**	**TM Mascot score**	**AK Mascot score**	**Missed cleavage sites**
2 hours	RT	Trypsin spin column + RapiGest	MSE	not identified	285	2
			DDA	not identified	535	4
	37°C	Trypsin spin column + RapiGest	MSE	not identified	132	0
			DDA	68	664	3
2 hours	RT	Trypsin spin column + Urea	MSE	not identified	182	0
			DDA	not identified	735	5
	37°C	Trypsin spin column + Urea	MSE	not identified	282	0
			DDA	not identified	640	4
2 hours	37°C	Immobilised trypsin	MSE	not identified	not identified	
			DDA	not identified	210	1
2 hours	37°C	Immobilised trypsin + RapiGest	MSE	not identified	219	3
			DDA	41	601	3
2 hours	37°C	Immobilised trypsin + Urea	MSE	not identified	not identified	
			DDA	41	171	0
King prawn - whole heated extract						
2 hours	RT	Trypsin spin column + RapiGest	MSE	486	not identified	9
			DDA	1184	60	18
2 hours	RT	Trypsin spin column + Urea	MSE	500	not identified	10
			DDA	1066	54	16

Table 3.8: Digestion of KP raw extracts using reagents for possible better digestion, potential resulting in an improved identification of TM. Mascot scores and missed cleavage sites obtained from King prawn extracts for TM and AK for different digestion time and reagents using trypsin spin columns are summarised in this table.

King prawn - raw extract						
Time of digestion	**Temperature**	**Trypsin + Reagent added**	**MS method**	**TM Mascot score**	**AK Mascot score**	**Missed cleavage sites**
15 min	RT	Trypsin spin column + RapiGest	MSE	not identified	405	2
			DDA	not identified	427	2
15 min	RT	Trypsin spin column + Urea	MSE	not identified	254	4
			DDA	not identified	370	3
30 min	RT	Trypsin spin column + RapiGest	MSE	not identified	160	2
			DDA	not identified	193	2
30 min	RT	Trypsin spin column + Urea	MSE	not identified	254	2
			DDA	not identified	370	4
1 hour	RT	Trypsin spin column + RapiGest	MSE	not identified	192	2
			DDA	not identified	195	2
1 hour	RT	Trypsin spin column + Urea	MSE	not identified	141	2
			DDA	not identified	198	3
2 hour	RT	Trypsin spin column + RapiGest	MSE	not identified	177	2
			DDA	not identified	325	4
2 hour	RT	Trypsin spin column + Urea	MSE	not identified	279	2
			DDA	not identified	307	4
4 hour	RT	Trypsin spin column + RapiGest	MSE	not identified	459	5
			DDA	not identified	599	2
4 hour	RT	Trypsin spin column + Urea	MSE	not identified	556	3
			DDA	not identified	601	4
6 hour	RT	Trypsin spin column + RapiGest	MSE	not identified	not identified	
			DDA	not identified	not identified	
6 hour	RT	Trypsin spin column + Urea	MSE	not identified	294	2
			DDA	not identified	327	1
22 hour	RT	Trypsin spin column + RapiGest	MSE	not identified	not identified	
			DDA	not identified	not identified	
22 hour	RT	Trypsin spin column + Urea	MSE	not identified	not identified	
			DDA	not identified	not identified	

Overall, table 3.6 summarises that the immobilised trypsin seems to have less digestion efficiency. Moreover, the addition of reagents possible improves the

identification of AK and TM, but it is not clear from the Mascot results, which reagent performs better. Therefore, the raw KP extract was digested in precents of reagents for various time points (Table 3.7) at room temperature, thus it was generally shown in table 3.5 that digestion at 37°C resulted in lower Mascot scores for TM and AK is heat sensitive.

Summarised in table 3.7, even with the addition of reagents, TM cannot be identified in KP raw extract. Moreover, it is not clear if RapiGest or urea provides higher digestion efficiency. However, increasing the digestion time to 22 hours and 6 hours, when using RapiGest, resulted in a loss of identification for arginine kinase. The highest scores were obtained when digestion time was four hours, followed by 15 minutes digestion time. However, the number of missed cleavage sites is higher at four hours compared to 15 minutes.

Nevertheless, it was shown by SDS-PAGE and immunoblotting analysis that raw King prawn extract contains tropomyosin. However, by the LC/qTOF method it could not be identified. Thus, to examine why TM was not identified in raw extract due to (1) the natural of TM structure in raw extract and therefore resulting insufficient digestion or (2) if the concentration of TM is too low in raw KP extract, TM from raw KP extract was purified. The purification was performed identically to the purification process using the whole heated extract, as described in detail in chapter 2 and briefly in section 3.2.7.

The results in table 3.8 show that purified tropomyosin from raw King prawn extract can be digested, detected and identified using the final digestion method by LC/qTOF method and applying Mascot search. Therefore it is assumed that the concentration of TM in raw KP is too low to be identified, however the digestion methods applying trypsin spin column at 15 minutes at RT is adequate for the digestion of natural TM. It is suggested when analysing raw prawn extracts by mass spectrometry, using TM for prawn identification might not be suitable, but other proteins/allergens, such as arginine kinase are suitable to identify prawn proteins in samples.

Table 3.9: Mascot scores and missed cleavage sites obtained of digesting purified TM from raw KP extract using trypsin spin column.

Purified tropomyosin from King prawn (raw extract)			
Time of digestion	**Temperature**	**MS method**	**TM Mascot score**
15 min	RT	MSE	119
		DDA	96

As it can be seen in table 3.1 to 3.8 the Mascot scores obtained using the MS^E and DDA methods vary. As already mentioned in section 3.3.1, the qTOF settings for, the tune page, the MS^E and DDA methods were modified accordingly to the results after each preliminary experiment set, including changes in the LC method. Thus, the Mascot scores obtained from DDA were approximately 15 fold higher compared to MS^E methods in table 3.1. However, the aim was to have comparable and reproducible LC/qTOF methods, thus changes in qTOF methods were applied to obtain higher Mascot scores for both methods. As it can be seen in table 3.8 and 3.9, the Mascot scores between MS^E and DDA methods are similar. With the final LC/qTOF methods Mascot scores and better protein identification can be achieved.

Overall, preliminary results summarised in section 3.3.1 and 3.3.2 show that the detection and identification of TM in purified and whole heated extract can be easily achieved using the very short and uncomplicated digestion method. Moreover, TM can be identified by MS^E and DDA methods. Thus, the following experiment, analysing 22 different whole heated shellfish extracts, will be applied using new trypsin spin columns for each digestion and digesting the extracts at room temperature for 15 minutes. Moreover, each digested species will be analysed by four LC/qTOF methods as described in section 3.2.2.

3.3.3 Protein extracts analysed by LC/qTOF

Table 3.9 summaries the protein extracts from KP analysed in section 3.3.1 and 3.2.2. The amounts of prawn muscle tissue used to make 100 ml of protein extracts are given. Moreover, it summaries the protein concentration for raw and

whole heated extract, purified raw TM and purified whole heated TM, using Pierce® 660 nm Protein Assay.

Table 3.10: Protein extracts and purified TM from King Prawn generated and the protein concentration estimated by Pierce® 660 nm Protein Assay.

King prawn extracts				
		Common name	**g edible muscle in 100 ml PBS**	**Protein concentration in mg/ml**
Crustacea	Prawns	King prawn whole heated extract	18	1.3
		King prawn raw extract	19.5	0.8
		King prawn purified whole heated extract	18	0.25
		King prawn purified raw extract	19.5	0.125

Table 3.10 summarises the 22 shellfish species which will be analysed in the following sections, with the common name, the scientific name and the GenBank accession number when available. Moreover, the amount of edible muscle extract used in 100 ml PBS used for generation the different whole heated shellfish extracts and their protein concentration are given. Interestingly, although the amount of muscle tissue used for various whole heated extracts are similar, the estimated protein concentrations by Pierce® 660 nm Protein Assay seem to vary.

Table 3.11: Shellfish species analysed by LC/qTOF, arranged by their major subgroups, including the common name, scientific name and the GenBank accession number. Moreover, information about the whole heated protein extracts made for various shellfish species and their protein concentration estimated by Pierce® 660 nm Protein Assay are given.

Shellfish Species						
		Common name	Scientific name	Accession numbers (GenBank)	g edible muscle in 100 ml PBS	Protein concentration in mg/ml
Crustacea	Prawns	King prawn	*Melicertus latisulcatus*	AGF86397.1	18	1.3
		Black Tiger prawn	*Penaeus monodon*	ADM34184.1	16.5	6.5
		Vannamei prawn	*Litopenaeus vannamei*	ACB38288.1	16.5	2
		Banana prawn	*Fenneropenaeus merguiensis*	ADC55381.4	16.5	2
		Green Tiger prawn	*Penaeus semisulcatus*	Not available	16.5	1.3
	Crabs	Blue Swimmer crab	*Portunus pelagicus*	AGE44125.1	16.5	1.5
		Sand crab	*Ovalipes australiensis*	Not available	16.5	4
		Mud crab	*Scylla serrata*	ABS12233.1	43.5	3
	Lobsters	Rock lobster	*Jasus edwardsii*	AGW22426.1	16.5	2.5
		Slipper lobster	*Thenus orientalis*	AGW22427.1	16.5	1
		Yabby	*Cherax destructor*	AGW22428.1	16.5	0.8
Mollusca	Gastropods	Jade Hybrid Tiger abalone	*Haliotis laevigata x Haliotis rubra*	Not available	37	1
		Sea snail	*Turbo cornutus*	BAH10149.1	16.5	0.25
	Cephalopods	Squid (Calamari)	*Sepioteuthis lessoniana*	BAE54430.1	16.5	3
		Octopus	*Octopus vulgaris*	BAE54433.1	16.5	3.5
	Bivalves	Blue mussel	*Mytilus edulis*	AAA82259.1	16.5	0.25
		Green mussel	*Perna viridis*	AAG08988.1	16.5	0.25
		Scallop (*fumatus)*	*Pecten fumatus*	Not available	16.5	1.7
		Scallop (*yessonensis)*	*Patinopecten yessonensis = Mizuhopecten yessoensis*	BAB17858.1	36	0.5
		Tasmanian oyster	*Crassostrea gigas*	BAH10152.1	33.5	0.55
		Sydney Rock oyster	*Saccostrea glomerata*	Not available	23	0.4
		Tuatua cockle	*Paphies subtriangulata*	Not available	21.5	0.7

3.3.4 TM identification in shellfish species by LC/qTOF

To identify signature peptides based on tropomyosin to distinguish crustacean from molluscs, 22 shellfish species (11 crustacean, 11 molluscs) have been analysed by the four LC/qTOF methods using the whole heated extracts. Moreover, rTM KP and rTM BTP have been analysed. Examples of the generated LC/qTOF chromatograms are shown in figure B2.1 and figure B2.2. The 22 shellfish species chosen are representing species for each subgroup of crustacean and molluscs and are commonly consumed in Australia. Moreover, some of the species have not been investigated yet' therefore the TM amino acid sequence is unknown. The overall most important finding is that tropomyosin was identified in all species, with the exception of Sydney Rock oyster.

The analysis of recombinant TM from KP and BTP showed that similar peptides and sequence coverage is achieved compared with the whole heated extract as shown in figure 3.2. Moreover, the LC/qTOF chromatograms are comparable with peaks detected for recombinant TM and whole heated extracts (Figure B2.1 and figure B2.2). These results demonstrate that recombinant TM perform very similar to whole heated extracts, thus are suitable to be used as standard proteins for future allergen identifications.

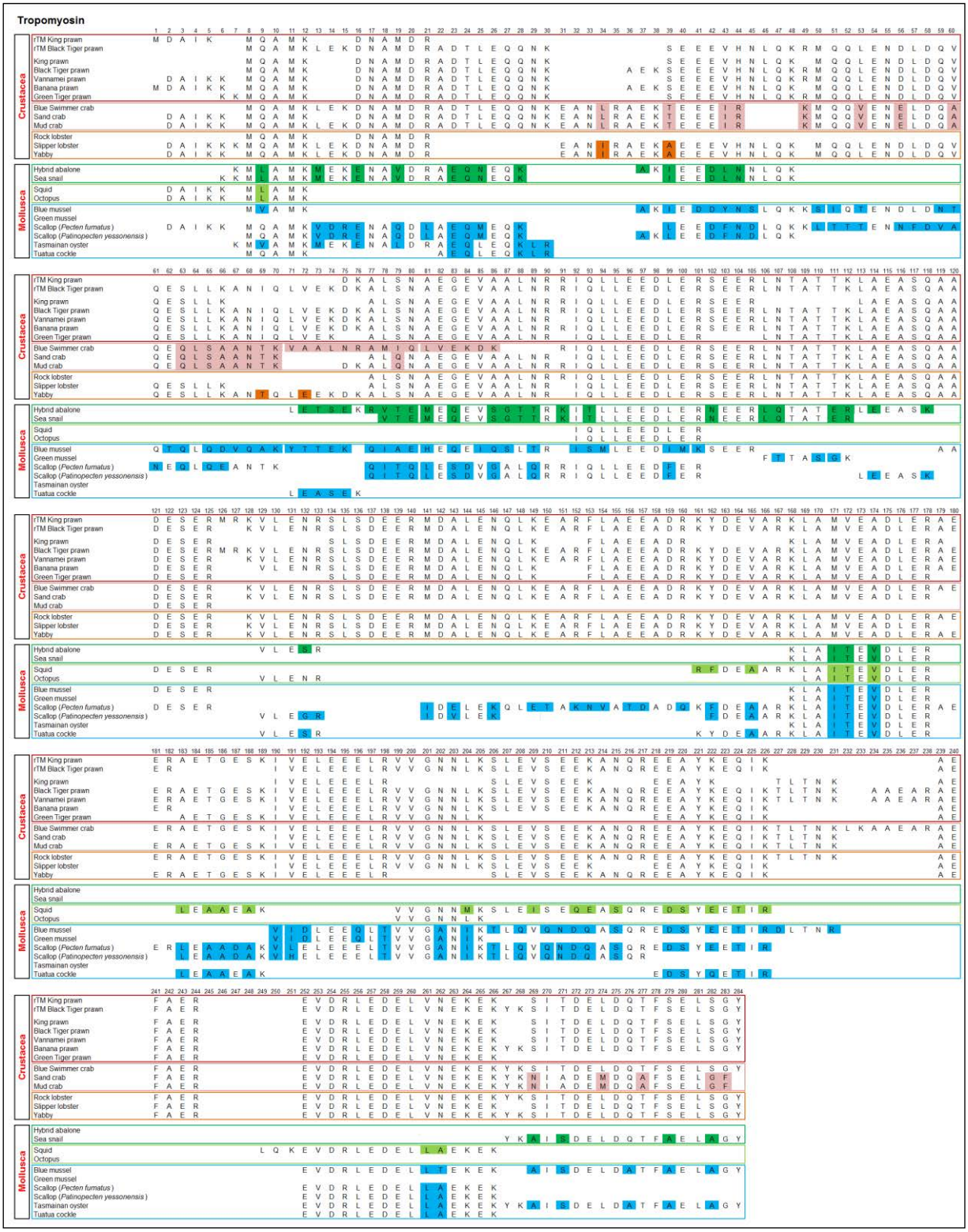


Figure 3.2: All peptides aligned for the various species analysed by LC/qTOF and identified for TM using Mascot. The amino acids highlighted in different colours vary of the TM sequence of Pen m 1.

Figure 3.2 aligns all peptides for the various species identified using the Mascot search. The amino acids highlighted in different colours vary of the TM sequence of Pen m 1. Generally, more peptides were identified for crustacean in comparison to mollusc species. For crustacean the TM sequence coverage by the four LC/qTOF methods is very high, with approximately 80%, whereas the sequence coverage for molluscs varies between 9-51%. The least TM sequence coverage, representing the number of identified peptides, is for gastropods and cephalopods. According to the estimated protein concentration by Pierce® 660 nm Protein Assay (Table 3.10) the protein concentration of Sea snail is relatively low compared to Hybrid abalone, thus more peptide have been identified for Hybrid abalone. Nevertheless, for cephalopods the highest protein content was estimated for all shellfish species investigated, with the exception of BTP and Sand crab. However, little peptides have been identified for cephalopods. Moreover, more peptides have been identified for squid in comparison to octopus, whereas the latter has the higher estimated protein concentration. Therefore it is assumed that less peptides were identified for molluscs in comparison to crustacean, might be due to the fact that mollusc species are less investigated and the amino acid sequence of tropomyosin varies more, thus cannot be identified using Mascot. Moreover, it shows that mollusc tropomyosins vary more in the amino acid sequence than crustacean. Especially for gastropods and cephalopods only a few TM sequences are available. Interestingly, for molluscs the identification of peptides is generally low in the amino acid region of 101-161 whereas between amino acid 232-248 no peptides have been identified. This could be that the peptides of these species (1) vary too much from known sequences and (2) peptides are not suitable for MS detection. The region that varies most for crustacean is between amino acid 33-86, followed by 269-284, where some crabs were identified with different peptides.

3.3.5 Signature peptide analysis

As illustrated in figure 3.2, overall many peptides for the different species have been identified for TM. However, figure 3.2 also shows that the crustacean peptides are highly identical, therefore it is challenging to find crustacean and mollusc specific signature peptides. To increase the panel of TM sequences, 106

tropomyosin sequences have been aligned in figure 3.3. These 106 species include many different crustacean and currently all available molluscs TM sequences. Moreover, sequences from related invertebrates that are know pan-allergens or potential cross-reactive have been added. Fish, commonly eaten “seafood”, and unrelated vertebrates (frogs) and were also analysed. Finally, human, chicken and other mammals have been added to the peptide analysis. The TM alignment shows that non allergenic TM sequences are far distanced from allergenic TM. Moreover, crustacean and molluscs species cluster, however, other invertebrates, such as mites, insects and nematodes are closer related to crustacean, than crustacean to molluscs. The subgroups of crustacean are closely related and cluster, with some exceptions. Interestingly, the subgroups of cephalopods and gastropods cluster as well, whereas the bivalves are separated into two clusters.

To be able to identify signature peptides to distinguish crustacean or molluscs the same 106 different tropomyosin sequences were aligned (Figure 3.4, figure 3.5 and figure B 2.3) with tryptic digested peptides based on Pen m 1 and Tod p 1, respectively. The addition of various non allergenic TM sequences is important, hence peptides that might be in food samples due to other food components or contamination should not be selected. The tryptic peptides that have been commonly identified in various species were analysed for potential being signature peptides to distinguish crustacean and molluscs, however, peptides shorter than six amino acids were excluded. Hence, 14 tryptic peptides were suitable for the alignment with 106 different tropomyosin sequences. As reference for the peptide alignment the tropomyosin/peptide sequence of Pen m 1 was chosen.

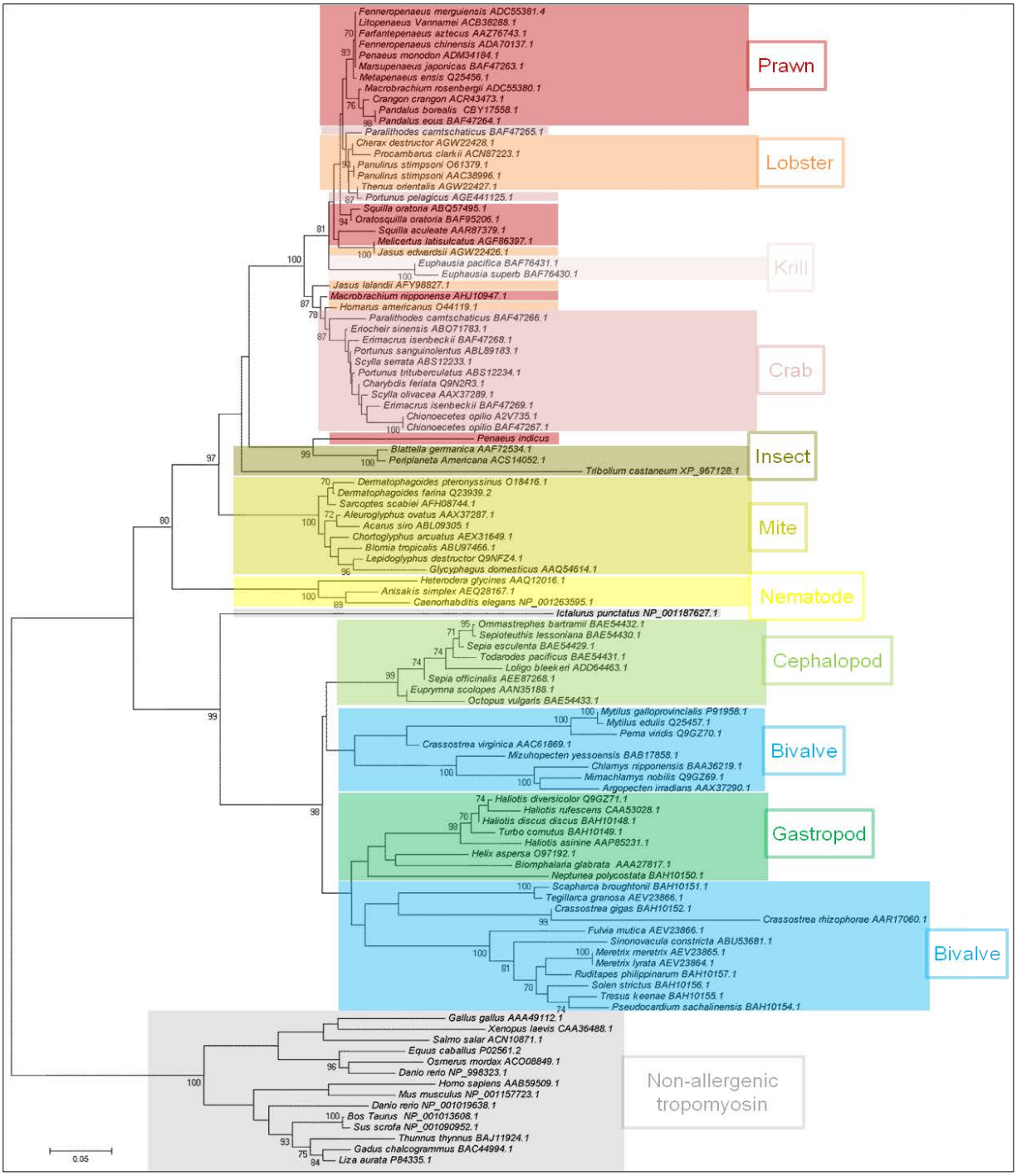


Figure 3.3: The phylogenetic tree for allergenic and non allergenic tropomyosin using 106 tropomyosin sequences. The subgroups of shellfish, the related invertebrates and non allergenic tropomyosin are shown in different colours for better visualisation. The phylogenetic tree was constructed using MEGA6 software. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches.

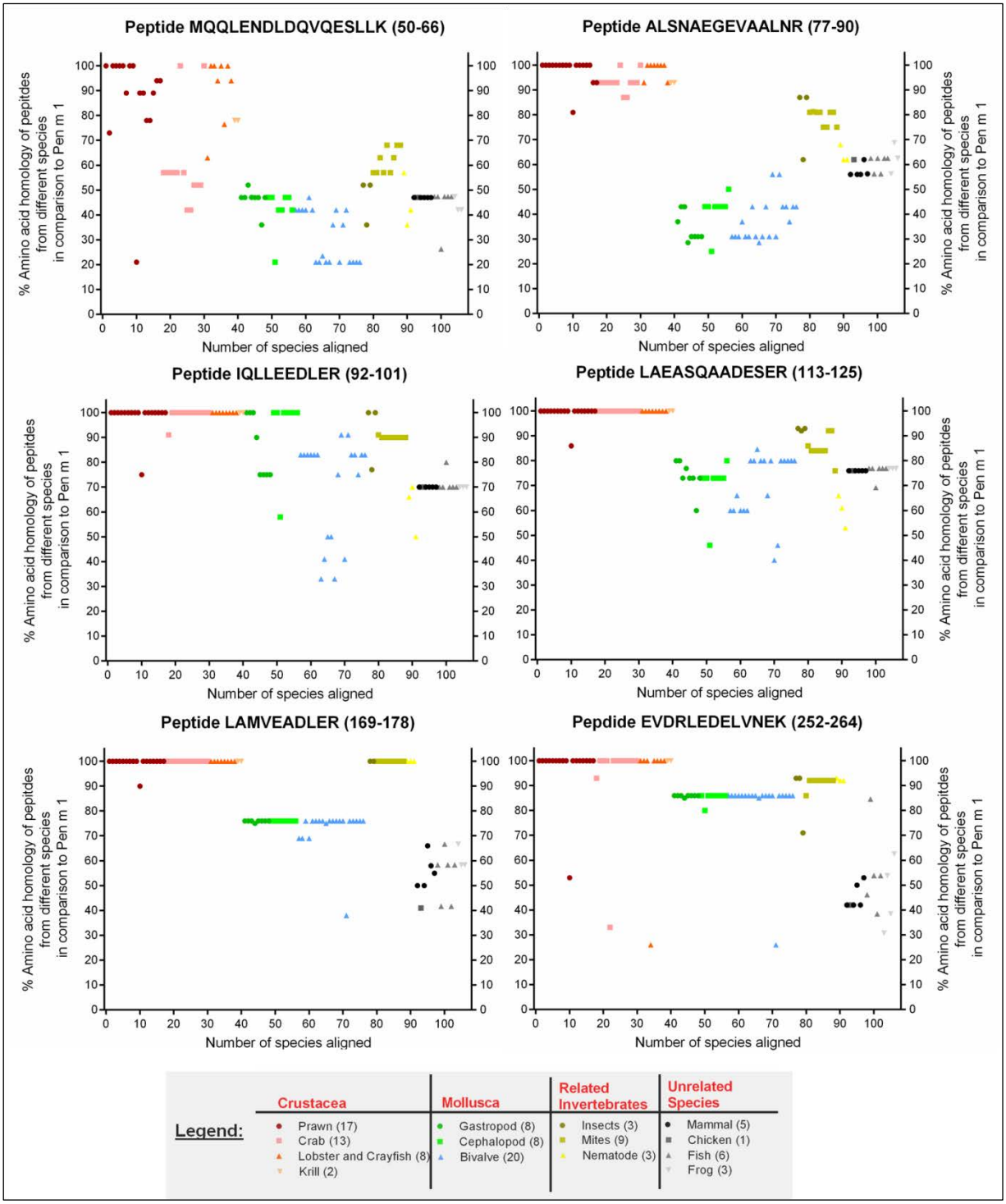


Figure 3.4: Amino acid alignment using commonly identified tryptic identified peptides of TM analysing 22 shellfish species by LC/qTOF. Peptide sequence used as reference for alignment is originated from Pen m 1. Amino acid homology is shown in percentage for 106 species and colour coded by their shellfish subgroups.

Figure 3.4 and figure B2.3 show the peptide alignment for the 14 tryptic peptides. These 14 tryptic peptides are in amino acid regions: 22-30, 50-60, 67-74, 77-90, 92-101, 113-125, 141-149, 153-160, 161-167, 169-178, 190-198, 206-213, 252-264 and 269-284, respectively. The alignment visualises that the non allergenic TM sequences do not share 100% amino acid homology for any of the investigated tryptic peptide. Unfortunately, none of the tryptic peptides includes all crustacean or all mollusc species, commonly one species does not show 100% amino acid sequence identity with the selected tryptic peptides. For prawns the big exception is Indian prawn (*Penaeus indicus)*, however, for this species there is no GenBank accession number available. Moreover, as displayed in the phylogenetic tree, Indian prawn is distanced compared to the other crustacean. The exception for crabs is Swimming crab (Cha f 1) (*Charybdis feriata*). The mollusc species are too diverse, when being aligned with peptides from Pen m 1, to name an exception for the subgroups.

Overall, as visualised in figure 3.2, figure 3.4 and figure B2.3, for crustacean the N-terminal region of TM is the most diverse and therefore not suitable to select a signature peptide that is unique for crustacean. Nevertheless, from amino acid 87 onwards, the crustaceans are highly identical. Peptides 113-125, 153-160, 252-264 and 269-284 are the only peptides for crustacean that are not identical with other related invertebrates. Peptide 113-125 is the only peptide, which has only one exception (Indian prawn) not sharing 100% amino acid identity. Therefore this peptide, 113-125 LAEASQAADESER, is suitable and was selected as signature peptide for crustacean and will be referred as “peptide 1” in this thesis.

Surprisingly, only peptide 92-101 shares 100% amino acid identity between crustacean and molluscs. All crustacean show 100% amino acid identity, with the exceptions of Indian prawn and Swimming crab as mentioned above. None of the bivalves share 100% amino acid identity, however, all cephalopods do, with the exception of Hawaiian bobtail squid (*Euprymna scolopes*). When analysing the gastropods, it is hypothesised that snails share 100% amino acid homology, but abalone do not. Thus, four analysed abalone species have 75% amino acid sequence identity, whereas three snails share 100% and one snail 90% (*Turbo cornutus*). Interestingly, this amino acid region was reported to be an IgE antibody binding epitope in various species,[21-26] thus possible explaining the immunological

cross-reactivity between crustacean and molluscs. Due do these interesting finding peptide 92-101 IQLLEEDLER was selected and will be referred as “peptide 2” in this thesis.

Figure 3.4 and figure B2.3 show peptides where mollusc species seem less diverse are amino acid positions 22-30, 67-74, 113-125, 153-160, 161-167, 169-178, 206-231 and 252-264. However, peptides 22-30, 67-74, 113-125 and 206-213 have different cleavage sites compared to crustacean and therefore different peptides would be observed. Moreover, peptide 153-160 could not be identified for any mollusc species. Furthermore, peptides 161-167 and 206-213 were not identified for gastropods. Due to this analyse peptides 169-178 and 252-264 were aligned using a molluscs species as reference protein, namely *Todarodes pacificus*, thus Tod p 1 is an official allergen registered with the IUIS. Figure 3.5 clearly shows that there are fewer exceptions for peptide 169-178, namely Sea snail (*Neptunea polycostata*) and Mangrove oyster (*Crassostrea rhizophorae*). However, for Mangrove oyster the partial tropomyosin like protein sequence is available, thus this might be that the amino acid sequence for peptide 3 is different. Moreover, peptide 252-264 was frequently detected as peptide 252-266, which means it includes a missed cleavage site and therefore is not suitable as signature peptide. Therefore peptide 169-178 LAITEVDLER was selected as signature peptide for molluscs and will be referred as “peptide 3” in this thesis.

It was further investigated if there would be a peptide suitable to distinguish the mollusc subgroups. However, as explained above, molluscs peptides are very diverse and cleavage sites do not match the crustacean peptides. Hence, peptide 2 is with the current knowledge to only tryptic peptide that can roughly distinguish the subgroups of molluscs. To find better tryptic peptides to distinguish the mollusc subgroups, more mollusc species need to be analysed for their TM amino acid sequence as well as for peptides that can be easily detected and identified by mass spectrometry.

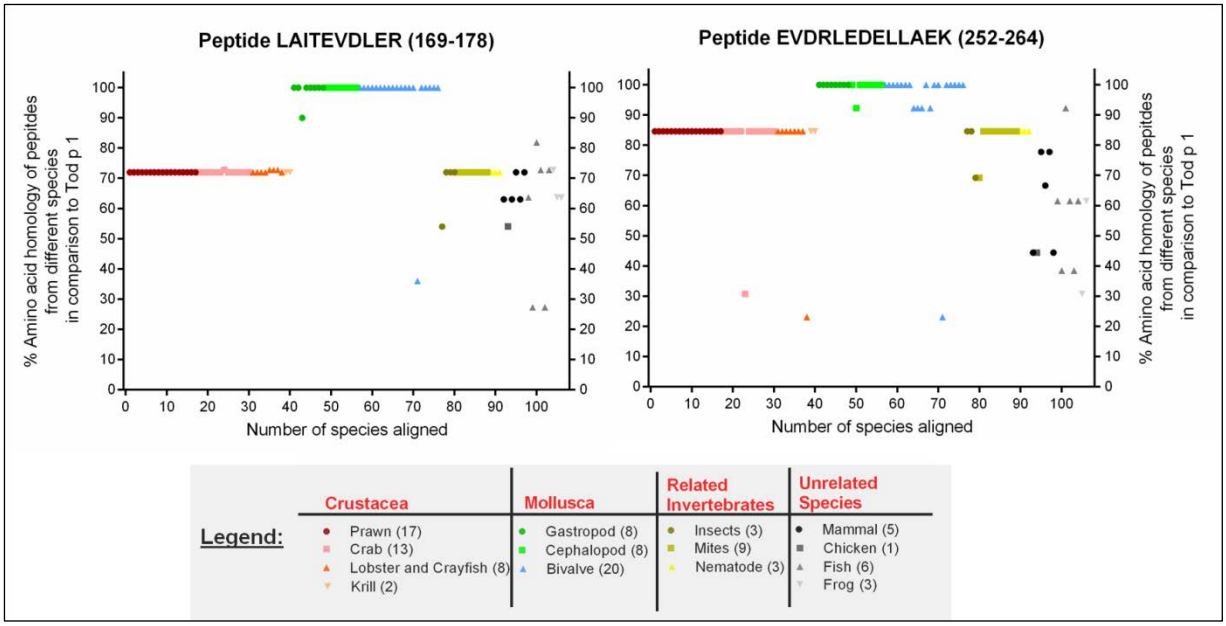


Figure 3.5: Amino acid alignment using commonly identified tryptic peptides of TM analysing 22 shellfish species by LC/qTOF. Peptide sequence used as reference for alignment is originated from Tod p 1. Amino acid homology is shown in percent for 106 species and colour coded by their shellfish subgroups.

To possible distinguish crustacean allergenic TM by their subgroups, peptides in the region from amino acid region 0- 87 were analysed. On the one hand peptide 22-30 is too similar, whereas on the other hand peptide 50-66 is too diverse. Peptide 67-74 shows different tryptic cleavage sites for crabs in comparison to prawns and lobsters. Nevertheless, peptide 77-90 is only present in prawns and lobsters, not in crabs or krill. The exceptions not sharing 100% amino acid identity for prawns are, Indian prawn, *Squilla oratoria* and *Oratosquilla oratoria* and for lobsters, American lobster (*Homarus americanus*) and Red Swamp crayfish (*Procambarus clarkii*). In contrast, crabs sharing 100% amino acid identity are Blue Swimmer crab (*Portunus pelagicus*) and Red King crab (*Paralithodes camtschaticus*), whereas for Red King crab two TM isoforms are known and only the slow-tonic isoform shares 100% amino acid homology. Overall, peptide 77-90 is the only tryptic peptide analysed that can be potentially distinguish between prawns and lobsters compared to crabs and krill. Thus, out of 40 crustacean species analysed only seven species do not fulfil these criteria for this selected peptide. Therefore, peptide 77-90 ALSNAEGEVAALNR was selected as peptide to

distinguish subgroups of crustacean and will be referred as “peptide 4” in this thesis.

3.3.6 Selected signature peptides

The four selected peptides were searched with Standard Protein BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using the default settings. All peptides were identified for TM. For Peptide 1 and 2 all 100 sequences analysed had a 100% query cover and 100% identity, whereas for peptide 3 only 64 and for peptide 4 only 19 species fulfilled these criteria. These findings are expected, thus less TM sequences are available for mollusc (peptide 3) and peptide 4 is more specific for the crustacean subgroups.

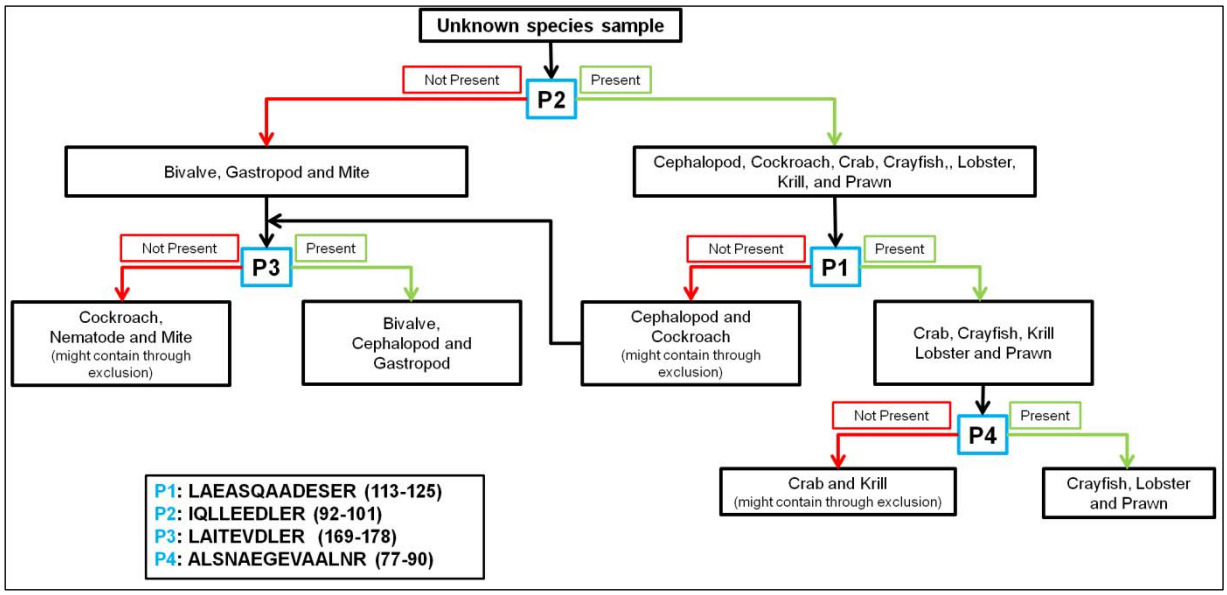


Figure 3.6: Flowchart for the identification of allergenic TM by subgroups of shellfish species using the four selected peptide, analysing unknown shellfish samples. (P= peptide)

To summarise and visualise the four selected peptides and their specificity for subgroup differentiation from shellfish and related invertebrates species a flowchart was generated (Figure 3.6). As it can be seen with peptide 1 and 3 the differentiation between crustacean and molluscs is possible. Moreover, selecting

peptide 2 can differentiate between cephalopods and bivalves and peptide 4 can potentially distinguish between prawns and lobsters versus crabs and krill.

Overall, this chapter shows that the identification of TM in various shellfish species is possible by LC/qTOF. Moreover, two signature peptides have been selected to be able to distinguish between crustacean and molluscs. Additionally, two peptides have been selected for the differentiation of the subgroups from crustacean and molluscs.

3.4 Discussion

Food legislation and industry have a demand for an analytical method that can distinguish between crustacean and mollusc species. Due to antibody cross-reactivity of the major allergen tropomyosin (TM) ELISAs are not suitable to fulfil these criteria. Therefore in this chapter, mass spectrometry (LC/qTOF) was applied to identify signature peptides, which are can distinguish crustacean from molluscs.

The analysis of purified whole heated TM and whole heated extract from KP showed that good digestion can be achieved in a shorter time using trypsin as reported in literature.[27] However, one trypsin performed better than another trypsin investigated. The efficiency of trypsins can vary as was also reported by Burkhart et al.[28] The Mascot scores achieved and peptides identified for TM by LC/qTOF for TM detection and identification are suitable. Nevertheless, when analysing raw King prawn (KP) extract, TM could not be identified, although the digestion, based on arginine kinase is suitable. Improving the identification of TM in raw KP extract with increasing digestion time and adding reagents (RapiGest or urea) could not be achieved.

Although it was reported by Yu et al.[29] that boiling accelerated the digestion of TM, it is not assumed that this is the reason why TM was not identified in raw KP extract. The SDS-PAGE and immunoblotting profile for whole heated extracts appears to be similar with reported results.[29-31] However, in this chapter it was shown that the digestion efficiency is shorter that observed for carb[29, 30] and prawns.[31] Overall, for whole heated KP extract, and especially for raw KP extract, less fragments were observed in comparison to Yu et al.,[29] but similar to Liu et al.[30, 31] It is suspected that TM could not be identified in KP raw extract is due to the low quantity of TM, thus it was shown that purified TM from raw KP extract can be identified using short digestion and the LC/qTOF methods detailed in this chapter.

Food allergen detection and quantification by mass spectrometry is a rather new application in the flied of allergen detection. Therefore, there are currently no guidelines available on how many peptides per protein should be chosen. Generally, the more peptides per protein become selected, the lower the chances

of identifying false positive proteins. Moreover, in terms of quantification, more transition per peptide reduces the possibility for false positive proteins. Two peptides per protein seems to be the approximated suggestion,[32, 33] as well as being the applied research. Nevertheless, there are various publications using only one peptide and few publications using more than two peptides per protein. In detail, a maximum of six peptides were selected by Sealey-Voyksner[34] for wheat allergens. Three studies did not explain how the signature peptides were selected, thus using two to four peptides for milk allergens.[35-37] Seven studies selected one to six peptides based on their signal intensity found by MS, after applying signature selection criteria for peanut allergens allergen.[34, 38-44] Eight studies used all signature peptides they were able to identify for various allergens, ranging from one peptide per allergen up to four peptides per allergen.[11-14, 38, 39, 45, 46] This demonstrates, that sometimes it is not possible to select more than one signature peptide per allergen. Four studies mentioned that the selected signature peptides overlapping with predicted IgE binding epitopes of peanut and milk allergens investigated.[36, 37, 40, 41]

The signature peptides identified and applied for the quantification of allergenic TM are all based on one single peptide, due to the analysed amino acid sequence homology.[11-14] Overall, in this chapter four peptides were identified for TM, however, a maximum of three peptides and a minimum of one peptide for different species can be detected applying these four peptides. These four peptides are: peptide 1 LAEASQAADESER (113-125), peptide 2 IQLLEEDLER (92-101), peptide 3 LAITEVDLER (169-178) and peptide 4 ALSNAEGEVAALNR (77-90).

Peptide 1 is unique for crustacean with one exception, whereas the region of 115-128 has been reported as epitope region for Pen m 1.[26] Peptide 3 is unique for molluscs with two exceptions and region 169-189 has been reported as epitope for *Turbo cornutus*.[23]

The amino acid sequence of peptide 2 is included in crustacean and cephalopods, but exclude bivalves. Interestingly, the amino acid area of peptide 2 has been reported as IgE binding region in various species. In detail, for prawns amino acid region 85-105[21] and 87-101[22, 25] was reported as epitope for Pen a 1 and 89-105[26] for Pen m 1. Amino acid region 87-101 was also found to be an epitope for lobster (Hom a 1) and house dust mite (Per a 7), whereas house dust mite (Der p 10) only

showed little IgE reactivity in this region.[22] This confirms the peptide alignment, thus Per a 7 shares 100% amino acid sequence identify for peptide 2, whereas Der p 10 only shows 80% amino acid identity. All other crustacean species that have been investigated for IgE binding regions, and have confirmed epitopse in the mentioned region share 100% amino acid identity for selected peptide 2. For molluscs amino acid region 92-105 for oyster (*Crassostrea gigas*)[47] and octopus for 77-112 (*Octopus vulgaris*)[24] were found to be IgE reactive, whereas region 92-118 was not an epitope for Sea snail (*Turbo cornutus).*[23] The amino acid sequence for octopus is 100% with peptide 2, however for oyster and Sea snail it was 91% and 90%, respectively, having one amino acid substituted for peptide 2. The amino acid substituted for oyster has the same property (leucine substituted for methionine), whereas the substitution for Sea snail does not share the same characteristics (glutamine substituted for threonine). The different amino acid substitutions might be the reason that an epitope was reported for oyster, but not for Sea snail.[23, 47] Moreover, other species shown to have 100% amino acid identity with peptide 2 are American and German cockroaches.

Peptide 4 is present in prawns and lobsters, but not in crabs and krill. The species specific signature peptides selected for TM by Abdel Rahman[11-14] are in the region where the amino acid sequence of TM varies more (39-76). However, peptide 4 was detected for Northern prawn[12] and Snow Crab,[13] whereas the latter showed to have a different amino acid sequence, as demonstrated in the peptide analysis. This shows that selected peptides can be detected with different digestion methods and mass spectrometer applied, thus peptides can be used as standards with different digestion methods, mass spectrometer and laboratories used.

Using more than the four selected peptides in this chapter for TM differentiation and detection was investigated, however, TM is too similar and too diverse at the same time to identify more suitable peptides with current knowledge. The majority of different shellfish species, with the exception of bivalves will be detected with at least two peptides, fulfilling the suggestion in the literature. Moreover, species of crustacean and molluscs can now be distinguished using the identified peptides, in this study and fulfil the requirements of the legislation and the food processing industry.

3.5 References

1. Lopata, A. L.; O'Hehir, R. E.; Lehrer, S. B. (2010). *Shellfish Allergy.* Clinical and Experimental Allergy. 40, (6), 850-858.

2. Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). *Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems.* Journal of Proteome Research. 13, (8), 3499-3509.

3. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* Int J Hyg Environ Health. 217, (7), 740-50.

4. Fuller, H. R.; Goodwin, P. R.; Morris, G. E. (2006). *An Enzyme-Linked Immunosorbent Assay (Elisa) for the Major Crustacean Allergen, Tropomyosin, in Food.* Food and Agricultural Immunology. 17, (1), 43-52.

5. Seiki, K.; Oda, H.; Yoshioka, H.; Sakai, S.; Urisu, A.; Akiyama, H.; Ohno, Y. (2007). *A Reliable and Sensitive Immunoassay for the Determination of Crustacean Protein in Processed Foods.* Journal of Agricultural and Food Chemistry. 55, (23), 9345-9350.

6. Werner, M. T.; Faeste, C. K.; Egaas, E. (2007). *Quantitative Sandwich Elisa for the Determination of Tropomyosin from Crustaceans in Foods.* Journal of Agricultural and Food Chemistry. 55, (20), 8025-8032.

7. Zhang, H.; Lu, Y.; Ushio, H.; Shiomi, K. (2014). *Development of Sandwich Elisa for Detection and Quantification of Invertebrate Major Allergen Tropomyosin by a Monoclonal Antibody.* Food Chemistry. 150, (0), 151-157.

8. Lopata, A. L.; Lehrer, S. B. (2009). *New Insights into Seafood Allergy.* Current Opinion in Allergy and Clinical Immunology. 9, (3), 270-277.

9. Sicherer, S. H.; Sampson, H. A. (2010). *Food Allergy.* Journal of Allergy and Clinical Immunology. 125, (2), S116-S125.

10. Emoto, A.; Ishizaki, S.; Shiomi, K. (2009). *Tropomyosins in Gastropods and Bivalves: Identification as Major Allergens and Amino Acid Sequence Features.* Food Chemistry. 114, (2), 634-641.

11. Abdel Rahman, A. M.; Kamath, S.; Lopata, A. L.; Helleur, R. J. (2010). *Analysis of the Allergenic Proteins in Black Tiger Prawn (Penaeus Monodon) and Characterization of the Major Allergen Tropomyosin Using Mass Spectrometry.* Rapid Communications in Mass Spectrometry. 24, (24), 3624-3624.

12. Abdel Rahman, A. M.; Kamath, S. D.; Gagne, S.; Lopata, A. L.; Helleur, R. (2013). *Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention.* Journal of Proteome Research. 12, (2), 647-656.

13. Abdel Rahman, A. M.; Lopata, A. L.; O'Hehir, R. E.; Robinson, J. J.; Banoub, J. H.; Helleur, R. J. (2010). *Characterization and De Novo Sequencing of Snow Crab Tropomyosin Enzymatic Peptides by Both Electrospary Ionization and Matrix-Assisted Laser Desorption Ionization Qqtof Tandem Mass Spectrometry.* Journal of Mass Spectrometry. 45, (4), 372-381.

14. Abdel Rahman, A. M.; Lopata, A. L.; Randell, E. W.; Helleur, R. J. (2010). *Absolute Quantification Method and Validation of Airborne Snow Crab Allergen Tropomyosin Using Tandem Mass Spectrometry.* Analytica Chimica Acta. 681, (1-2), 49-55.

15. Koeberl, M.; Kamath, S. D.; Saptarshi, S. R.; Smout, M. J.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Auto-Induction for High Yield Expression of Recombinant Novel Isoallergen Tropomyosin from King Prawn (Melicertus Latisulcatus) for Improved Diagnostics and Immunotherapeutics.* J Immunol Methods. 415, (0), 6-16.

16. Kamath, S. D.; Rahman, A. M. A.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. n/a-n/a.

17. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M. X.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* International Journal of Hygiene and Environmental Health. (0).

18. Kamath, S. D.; Abdel Rahman, A. M.; Komoda, T.; Lopata, A. L. (2013). *Impact of Heat Processing on the Detection of the Major Shellfish Allergen Tropomyosin in Crustaceans and Molluscs Using Specific Monoclonal Antibodies.* Food Chemistry. 141, (4), 4031-9.

19. Kamath, S. D.; Abdel Rahman, A. M.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. 58, (5), 1144-1155.

20. Shanti, K. N.; Martin, B. M.; Nagpal, S.; Metcalfe, D. D.; Rao, P. V. (1993). *Identification of Tropomyosin as the Major Shrimp Allergen and Characterization of Its Ige-Binding Epitopes.* The Journal of Immunology. 151, (10), 5354-63.

21. Ayuso, R.; Lehrer, S. B.; Reese, G. (2002). *Identification of Continuous, Allergenic Regions of the Major Shrimp Allergen Pen a 1 (Tropomyosin).* International Archives of Allergy and Immunology. 127, (1), 27-37.

22. Ayuso, R.; Reese, G.; Leong-Kee, S.; Plante, M.; Lehrer, S. B. (2002). *Molecular Basis of Arthropod Cross-Reactivity: Ige-Binding Cross-Reactive Epitopes of Shrimp, House Dust Mite and Cockroach Tropomyosins.* International Archives of Allergy and Immunology. 129, (1), 38-48.

23. Ishikawa, M.; Ishida, M.; Shimakura, K.; Nagashima, Y.; Shiomi, K. (1998). *Purification and Ige-Binding Epitopes of a Major Allergen in the Gastropod Turbo Cornutus.* Biosci Biotechnol Biochem. 62, (7), 1337-43.

24. Ishikawa, M.; Suzuki, F.; Ishida, M.; Nagashima, Y.; Shiomi, K. (2001). *Identification of Tropomyosin as a Major Allergen in the Octopus Octopus Vulgaris and Elucidation of Its Ige-Binding Epitopes.* Fisheries Science. 67, (5), 934-942.

25. Reese, G.; Viebranz, J.; Leong-Kee, S. M.; Plante, M.; Lauer, I.; Randow, S.; Moncin, M. S. M.; Ayuso, R.; Lehrer, S. B.; Vieths, S. (2005). *Reduced Allergenic Potency of Vr9-1, a Mutant of the Major Shrimp Allergen Pen a 1 (Tropomyosin).* Journal of Immunology. 175, (12), 8354-8364.

26. Zheng, L.-N.; Lin, H.; Pawar, R.; Li, Z.-X.; Li, M.-H. (2011). *Mapping Ige Binding Epitopes of Major Shrimp (Penaeus Monodon) Allergen with Immunoinformatics Tools.* Food and Chemical Toxicology. 49, (11), 2954-2960.

27. Lopez-Ferrer, D.; Capelo, J. L.; Vazquez, J. (2005). *Ultra Fast Trypsin Digestion of Proteins by High Intensity Focused Ultrasound.* Journal of Proteome Research. 4, (5), 1569-1574.

28. Burkhart, J. M.; Schumbrutzki, C.; Wortelkamp, S.; Sickmann, A.; Zahedi, R. P. (2012). *Systematic and Quantitative Comparison of Digest Efficiency and Specificity Reveals the Impact of Trypsin Quality on Ms-Based Proteomics.* Journal of Proteomics. 75, (4), 1454-1462.

29. Yu, H.-L.; Cao, M.-J.; Cai, Q.-F.; Weng, W.-Y.; Su, W.-J.; Liu, G.-M. (2011). *Effects of Different Processing Methods on Digestibility of Scylla Paramamosain Allergen (Tropomyosin).* Food and Chemical Toxicology. 49, (4), 791-798.

30. Liu, G.-M.; Cao, M.-J.; Huang, Y.-Y.; Cai, Q.-F.; Weng, W.-Y.; Su, W.-J. (2010). *Comparative Study of in Vitro Digestibility of Major Allergen Tropomyosin and Other Food Proteins of Chinese Mitten Crab (Eriocheir Sinensis).* Journal of the Science of Food and Agriculture. 90, (10), 1614-1620.

31. Liu, G.-M.; Huang, Y.-Y.; Cai, Q.-F.; Weng, W.-Y.; Su, W.-J.; Cao, M.-J. (2011). *Comparative Study of in Vitro Digestibility of Major Allergen, Tropomyosin and Other Proteins between Grass Prawn (Penaeus Monodon) and Pacific White Shrimp (Litopenaeus Vannamei).* Journal of the Science of Food and Agriculture. 91, (1), 163-170.

32. Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. (2008). *Selected Reaction Monitoring for Quantitative Proteomics: A Tutorial.* Molecular Systems Biology. 4, (1), n/a-n/a.

33. Picariello, G.; Mamone, G.; Addeo, F.; Ferranti, P. (2011). *The Frontiers of Mass Spectrometry-Based Techniques in Food Allergenomics.* Journal of Chromatography A. 1218, (42), 7386-7398.

34. Sealey-Voyksner, J. A.; Khosla, C.; Voyksner, R. D.; Jorgenson, J. W. (2010). *Novel Aspects of Quantitation of Immunogenic Wheat Gluten Peptides by Liquid Chromatography-Mass Spectrometry/Mass Spectrometry.* Journal of Chromatography A. 1217, (25), 4167-4183.

35. Monaci, L.; van Hengel, A. J. (2008). *Development of a Method for the Quantification of Whey Allergen Traces in Mixed-Fruit Juices Based on Liquid Chromatography with Mass Spectrometric Detection.* Journal of Chromatography A. 1192, (1), 113-120.

36. Newsome, G. A.; Scholl, P. F. (2012). *Quantification of Allergenic Bovine Milk As1-Casein in Baked Goods Using an Intact 15n-Labeled Protein Internal Standard.* Journal of Agricultural and Food Chemistry. 61, (24), 5659-5668.

37. Zhang, J.; Lai, S.; Zhang, Y.; Huang, B.; Li, D.; Ren, Y. (2012). *Multiple Reaction Monitoring-Based Determination of Bovine A-Lactalbumin in Infant Formulas and Whey Protein Concentrates by Ultra-High Performance Liquid Chromatography–Tandem Mass Spectrometry Using Tryptic Signature Peptides and Synthetic Peptide Standards.* Analytica Chimica Acta. 727, (0), 47-53.

38. Heick, J.; Fischer, M.; Kerbach, S.; Tamm, U.; Popping, B. (2011). *Application of a Liquid Chromatography Tandem Mass Spectrometry Method for the Simultaneous Detection of Seven Allergenic Foods in Flour and Bread and Comparison of the Method with Commercially Available Elisa Test Kits.* Journal of Aoac International. 94, (4), 1060-1068.

39. Heick, J.; Fischer, M.; Popping, B. (2011). *First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry.* Journal of Chromatography A. 1218, (7), 938-943.

40. Chassaigne, H.; Norgaard, J. V.; van Hengel, A. J. (2007). *Proteomics-Based Approach to Detect and Identify Major Allergens in Processed Peanuts by Capillary Lc-Q-Tof (Ms/Ms).* Journal of Agricultural and Food Chemistry. 55, (11), 4461-4473.

41. Shefcheck, K. J.; Callahan, J. H.; Musser, S. M. (2006). *Confirmation of Peanut Protein Using Peptide Markers in Dark Chocolate Using Liquid Chromatography-Tandem Mass Spectrometry (Lc-Ms/Ms).* Journal of Agricultural and Food Chemistry. 54, (21), 7953-7959.

42. Shefcheck, K. J.; Musser, S. M. (2004). *Confirmation of the Allergenic Peanut Protein, Ara H 1, in a Model Food Matrix Using Liquid Chromatography/Tandem Mass Spectrometry (Lc/Ms/Ms).* Journal of Agricultural and Food Chemistry. 52, (10), 2785-2790.

43. Careri, M.; Costa, A.; Elviri, L.; Lagos, J. B.; Mangia, A.; Terenghi, M.; Cereti, A.; Garoffo, L. P. (2007). *Use of Specific Peptide Biomarkers for Quantitative Confirmation of Hidden Allergenic Peanut Proteins Ara H 2 and Ara Ill 3/4 for Food Control by Liquid Chromatography-Tandem Mass Spectrometry.* Analytical and Bioanalytical Chemistry. 389, (6), 1901-1907.

44. Careri, M.; Elviri, L.; Maffini, M.; Mangia, A.; Mucchino, C.; Terenghi, M. (2008). *Determination of Peanut Allergens in Cereal-Chocolate-Based Snacks: Metal-Tag Inductively Coupled Plasma Mass Spectrometry Immunoassay Versus Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry.* Rapid Communications in Mass Spectrometry. 22, (6), 807-811.

45. Monaci, L.; Losito, I.; Palmisano, F.; Visconti, A. (2010). *Identification of Allergenic Milk Proteins Markers in Fined White Wines by Capillary Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry.* Journal of Chromatography A. 1217, (26), 4300-4305.

46. Weber, D.; Raymond, P.; Ben-Rejeb, S.; Lau, B. (2006). *Development of a Liquid Chromatography-Tandem Mass Spectrometry Method Using Capillary Liquid Chromatography and Nanoelectrospray Ionization-Quadrupole Time-of-Flight Hybrid Mass Spectrometer for the Detection of Milk Allergens.* Journal of Agricultural and Food Chemistry. 54, (5), 1604-1610.

47. Ishikawa, M.; Ishida, M.; Shimakura, K.; Nagashima, Y.; Shiomi, K. (1998). *Tropomyosin, the Major Oyster Crassostrea Gigas Allergen and Its Ige-Binding Epitopes.* Journal of Food Science. 63, (1), 44-47.

48. Bilbao, A.; Varesio, E.; Luban, J.; Strambio-De-Castillia, C.; Hopfgartner, G.; Müller, M.; Lisacek, F. (2015). *Processing strategies and software solutions for data-independent acquisition in mass spectrometry.* Proteomics. 15 (5-6), 964-980.

3.6 Summary chapter 3:

Identifying Signature Peptides for Allergenic Tropomyosin to Distinguish Crustacean from Molluscs

- ✔ Four liquid chromatography (LC) coupled with quadruple time of flight (qTOF) mass spectrometer methods were developed and rapid protein digestion was evaluated
- ✔ Analysis of 11 different crustacean species and 11 different mollusc species (total of 22 shellfish species) by LC/qTOF for their identified peptides based on allergenic tropomyosin.
- ✔ Aligning 14 commonly detected tryptic peptides of tropomyosin with 106 other tropomyosin sequences to identify signature peptides
- ✔ Four signature peptides have been identified. Peptide 1 is unique to crustaceans and peptide 3 is unique to molluscs. Peptide 2 occurs in crustacean and cephalopods, not in bivalves. Peptide 4 can be found in crayfish, lobsters and prawns, not in crabs or krill, however there are some exceptions

The four identified peptides derived of tropomyosin to distinguish crustacean and mollusc species will be chemical synthesised and applied in chapter 5 for the development and validation of a quantitative liquid chromatography coupled with multiple reaction monitoring (LC/MRM) method.

In chapter 4 the same 22 shellfish species will be analysed for all proteins identified in raw and whole heated extracts, by the four LC/qTOF methods developed in this chapter.

CHAPTER 4

IDENTIFYING POSSIBLE ALLERGEN-CANDIDATES IN CRUSTACEAN AND MOLLUSCS

4.1 Introduction

Every protein can potentially be an allergen, but it was summarised in several studies that most allergens belong to certain protein families, have a certain structure and a particular function.[1-4] Hundred thirty-eight allergen families have been analysed for their structure using the Structural Classification Proteins (SCOP) database.[2] It was found that 125 families include either an α-helical or β-sheet structures or a combination of both. Only 13 allergen families had other structures, such as membrane proteins or cell surface proteins.[2] In the same study the allergens were also analysed according to their function, using the Gene Ontology (GO) Annotation Database. When classified by the molecular function and the GO number, it was demonstrated that 238 allergens have a binding function and 175 allergens are catalytic active, whereas only 53 had nutrient reservoir activity as a molecular function.[2] Animal food allergens, which include shellfish allergens, could be classified into three major protein families and 14 minor protein families.[3] Moreover, proteins that share more than 62% amino acid homology with the human analogue seem not to be allergenic.[1, 2, 5-9] Whereas, if the amino acid sequence identity to the human analogue is less than 54% all proteins are potential allergens.[10]

Shellfish, including crustaceans and molluscs, are important food sources throughout the world. A total of sixteen proteins from shellfish have been reported as allergens in the literature, however, not all 16 allergens are reported in every shellfish species. The major allergen in shellfish, namely tropomyosin (TM), has been investigated in the previous chapters. Other established allergens in crustacean are arginine kinase, myosin light chain and sarcoplasmic calcium binding protein,[7] whereas arginine kinase and paramyosin are established allergens in molluscs. Excluding TM, other allergens registered with the International Union of Immunological Societies (IUIS) in crustacean are arginine kinase, myosin light chain 1, myosin light chain 2, triose phosphate isomerase, troponin C and troponin I. Other allergens that have been reported for crustacean are actin,[11] enolase,[12] fructose-1,6-bisphosphate,[13] glyceraldehyde-3-phosphate dehydrogenase,[11, 12] hemocyanain,[12, 14] myosin heavy chain,[11, 12] and titin.[13] For molluscs only TM is registered with the IUIS. Moreover, only two other allergens are reported in literature for molluscs, being arginine kinase[15] and paramyosin.[15, 16]

Nevertheless, shellfish species include more proteins than the 16 reported allergens. Therefore in this chapter 22 shellfish species will be investigated by LC/qTOF and all identified proteins will be analysed for their potential allergenicity.

4.1.1 Aims

The aims of this chapter are:

- Analysing 22 shellfish species by mass spectrometry (LC/qTOF) utilising raw and whole heated protein extracts
- Identifying all proteins in 22 shellfish species applying Mascot database
- Investigating species specific and shellfish subgroups specific protein profiles utilising raw and whole heated extracts
- Analysing all identified proteins for their potential allergenicity

4.2 Materials and methods

4.2.1 Shellfish species analysed by LC/qTOF

The same 22 shellfish species that have been analysed in chapter 3 will be analysed in this chapter. However, in chapter 3 only the whole heated extracts were analysed with focus on tropomyosin. In this chapter the raw and the whole heated extracts will be utilised by LC/qTOF and all the identified proteins will be analysed. Both extracts for various species were made as described in chapter 2 and chapter 3. Briefly, for the raw extracts the shells of shellfish species were removed and edible muscle parts were shredded into small pieces and homogenised in PBS. For the whole heated extracts, the crustacean species were heated including their outer shell, whereas for the mollusc species only the edible muscle parts were heated. The detailed preparation of sample extracts, including the estimated protein concentration by Pierce® 660 nm Protein Assay using BSA as standard, is summarised in table 3.10 and table 4.1. The total homogenised extract was kept at 4°C overnight while continuously shaking. The protein extracts were centrifuged, followed by sterile filtration and stored at -80°C until further use. The species Jade Hybrid Tiger abalone (*Haliotis laevigata x Haliotis rubra*) will be referred as Hybrid abalone in this thesis.

4.2.2 LC/qTOF method and sample digestion

The same digestion method and LC/qTOF analysis was performed for each extract and species as described in detail in chapter 3. Briefly, a total of 200 µl of soluble protein extract was reduced and alkylated. The solution was loaded onto a trypsin spin column, which has been washed, equilibrated and prepared according to the manufacturer's instructions. The samples were incubated for 15 minutes at room temperature and eluted twice with 100 µl of 0.1% formic acid. The eluted solution was analysed with a UPLC coupled with an ESI interface to a Xevo G2 QTof MS mass spectrometer (Waters Corporation, USA). All the digested proteins were analysed four times, using the combination of the final LC and the final MS methods. These four final methods are: 15 minutes MS^E, 60 minutes MS^E, 15 minutes DDA and 60 minutes DDA.

4.2.3 LC/qTOF data and protein analysis

All raw and uninterpreted data generated by LC/qTOF was processed using ProteinLynx Global Server (PLGS) v2.3 (Waters Corporation, Australia) and converted into pkl files. The converted pkl files were searched with Mascot daemon search engine (version 2.4) and Swissprot database (549,832 entries – November 2015) with a 0.1 Da tolerance against the database generated theoretical peptide ion masses and a minimum of one matched peptide. The Mascot search parameters included variable modifications of carbamidomethyl-C and N-terminus, deamidation of N (asparagine) and Q (glutamine) and oxidation of M (methionine). Up to seven missed cleavage sites were allowed.

All proteins identified by Mascot were considered, however, some proteins were excluded from the analysis after manually verifying the protein and peptide scores. Proteins were exclude when (1) the Mascot score was below 20[28] (2) the identified protein was trypsin (3) the source and/or function of the identified protein is unknown (4) the protein identified matched human analogues, thus were considered as contamination (5) the protein were only derived from bacteria, thus were considered as contamination (6) only one peptide was matched to the identified protein. Furthermore, when more than one species belonging to the shellfish group was identified for a specific protein, the highest Mascot score was used. If other species were identified, including shellfish species, the highest shellfish Mascot score was used. However, if none of the identified species belonged to the shellfish group, the highest Mascot score was used. [29]

The proteins identified were individually analysed for each species, each extract and each of the four LC/qTOF methods. The Mascot scores[29] for all the proteins identified per species and extract were summed up and this value was set as 100%. The relative percentage of the individual proteins was calculated based on the total Mascot score. All percentages for the four individual runs were added and divided by four, giving the final species specific percentages shown in figure 4.4 to figure 4.9. Furthermore, the percentage for each species and the identified proteins was added and divided by the amount of species analysed in the shellfish

subgroup. This number is shown in percent in the pie charts (Figure 4.2 and figure 4.3).

To investigate if the identified proteins are possible allergens, the proteins were analysed according to their function, structure, protein family, heat stability and amino acid identity to the human analogue. These criteria were selected, thus it is reported in the literature, that if proteins meet these criteria the protein is more likely to be an allergen.[2, 3, 10, 17] The identified proteins were classified according to their function using the Gene Ontology (GO) Annotation Database (http://www.ebi.ac.uk/GOA).[2] Moreover, the structures of proteins were classified using the Structural Classification Proteins (SCOP) database (http://scop.mrc-lmb.cam.ac.uk/scop).[2]

4.3 Results

4.3.1 Shellfish species analysed by LC/qTOF

Table 3.10 in chapter 3 summarises the details about the whole heated shellfish extracts, whereas table 4.1 summaries the details of the raw shellfish extracts. Both tables report the summary of the raw and the whole heated extracts and their protein concentration of all the shellfish species analysed in this chapter. Moreover, the amount of edible shellfish muscle tissue homogenised in 100 ml PBS is reported. The same effect can be observed in table 4.1 as reported in chapter 3, that although the amount of muscle tissue used for the various extracts is similar, the estimated protein concentrations by Pierce® 660 nm Protein Assay seemed to vary.

Figure B2.4 and figure B2.5 display as examples LC/qTOF chromatograms generated for the shellfish species using the raw and the whole heated extracts. Figure B2.4 represents the crustacean species displaying the LC/qTOF chromatogram of Mud Crab, whereas figure B2.5 represents the mollusc species showing the LC/qTOF chromatogram of Hybrid abalone. Both figures visualise that the chromatograms differ for the four LC/qTOF methods applied, the species and the extract analysed, however, overall the chromatograms include various peaks, which could be identified as many peptides originated from various proteins. All the identified proteins, applying Mascot, will be described and analysed in the following sections.

Table 4.1: Shellfish species analysed by LC/qTOF, arranged by their major subgroups, including the common name, the scientific name and the GenBank accession number. Moreover, information about the raw protein extracts made for the various shellfish species and their protein concentration estimated using the Pierce® 660 nm Protein Assay are given.

Shellfish Species						
		Common name	**Scientific name**	**Accession numbers (GenBank)**	**g edible muscle in 100 ml PBS**	**Protein concentration in mg/ml**
Crustacea	Prawns	King prawn	*Melicertus latisulcatus*	AGF86397.1	19.5	0.8
		Black Tiger prawn	*Penaeus monodon*	ADM34184.1	16.5	7
		Vannamei prawn	*Litopenaeus vannamei*	ACB38288.1	16.5	12
		Banana prawn	*Fenneropenaeus merguiensis*	ADC55381.4	16.5	12
		Green Tiger prawn	*Penaeus semisulcatus*	Not available	16.5	8.5
	Crabs	Blue Swimmer crab	*Portunus pelagicus*	AGE44125.1	16.5	3
		Sand crab	*Ovalipes australiensis*	Not available	16.5	8
		Mud crab	*Scylla serrata*	ABS12233.1	46.5	5
	Lobsters	Rock lobster	*Jasus edwardsii*	AGW22426.1	16.5	2
		Slipper lobster	*Thenus orientalis*	AGW22427.1	16.5	3
		Yabby	*Cherax destructor*	AGW22428.1	16.5	1.5
Mollusca	Gastropods	Jade Hybrid Tiger abalone	*Haliotis laevigata x Haliotis rubra*	Not available	41	5
		Sea snail	*Turbo cornutus*	BAH10149.1	16.5	3
	Cephalopods	Squid (Calamari)	*Sepioteuthis lessoniana*	BAE54430.1	16.5	10
		Octopus	*Octopus vulgaris*	BAE54433.1	16.5	5
	Bivalves	Blue mussel	*Mytilus edulis*	AAA82259.1	16.5	5
		Green mussel	*Perna viridis*	AAG08988.1	16.5	4
		Scallop (*fumatus)*	*Pecten fumatus*	Not available	16.5	3.3
		Scallop (*yessonensis)*	*Patinopecten yessonensis = Mizuhopecten yessoensis*	BAB17858.1	32	2
		Tasmanian oyster	*Crassostrea gigas*	BAH10152.1	25	3.3
		Sydney Rock oyster	*Saccostrea glomerata*	Not available	22	4
		Tuatua cockle	*Paphies subtriangulata*	Not available	16	0.15

4.3.2 All the proteins identified for analysed shellfish species utilising Mascot

Overall, a total of 32 different proteins have been identified for the raw and the whole heated extracts analysing the 22 shellfish species. Out of these 32 proteins 16 proteins have been reported as being allergenic in shellfish. Interestingly, when the identified proteins are sorted according to their GO number and function, it can be seen that with the exception of transmembrane protein 2, all proteins have a specific molecular function (Figure 4.1). Approximately half of the proteins are either binding proteins or catalytic active proteins. Moreover, eleven identified proteins play a role in the muscle contraction, whereas ten proteins are enzymes.

For the identification of the proteins the Mascot database was used. The vast majority of the proteins were identified with high certainty, having a high Mascot sore and good sequence coverage with the identified peptides. However, some proteins had more than one matched protein and/or more than one species matched. The proteins that could be potential different proteins and/or contamination are explained below.

Actin, glyeraldehyde-3-phopsphate dehydrogenase (GAPDH), triose phosphate isomerase (TIM) and fructose-1,6-bisphosphatase (FBA) were commonly identified with different species, therefore these proteins were subdivide in two groups, assigning if the identified species included or exclude a shellfish species. Moreover, myosin heavy chain (MHC) was identified as (1) MHC, muscle and the species *Drosophila melanogaster* or as (2) MHC straight muscle and the species *Argopecten irradians*. Interestingly, these two myosin heavy chains only share 60% amino acid homology, therefore they were analysed separately and will be referred as MHC muscle or MHC straight muscle.

The proteins that have been identified, which might be possible contaminations (due to sample processing) are TIM, enolase, serine/threonine protein kinase, thymosin beta and actin. Whereas the latter is a highly conserved protein throughout the whole animal kingdom and therefore shares high amino acid identity with humans, shellfish species and any other species from the animal kingdom. However, serine/threonine protein kinase and FBA are potential bacterial contaminations, but still will be analysed in this chapter.

Molecular Function	Binding Proteins	Protein binding	Cytoskeletal protein binding	Tropomyosin (TM)
				Myosin light chain (MLC)
				Troponin C
				Troponin I
				Titin
			Actin binding	Actin
				Thymosin beta
			Calmodulin binding	Calmodulin
		Ion binding	Metal ion binding	Sarcoplasmic calcium binding protein (SCBP)
				Enolase
				Heavy metal binding protein
				Cytochrome C
		Organic cyclic compound binding	Nucleic acid binding	Heat shock factor protein 2
				Histone H2A
				Ribosomal S30 protein
		Small molecule binding	Nucleotide binding	Tubulin alpha chain
				Protein ycf2
	Catalytic active protein	Transferase activity	Transferase activity, transferring phosphorus - c	Arginine kinase (AK)
				Serine/Threonine protein kinase
				Nucleoside diphosphate kinase
		Oxidoreductase activity	Oxidoreductase activity, acting on CH-OH group	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)
				Glucose-6-phosphate-1-dehydrogenase
			Oxidoreductase activity on superoxide	Superoxide dismutase [Cu-Zn]
		Isomerase activity	Intramolecular oxidoreductase activity	Triose phosphate isomerase (TIM)
		Lyase activity	Carbon-carbon lyaseactivity	Fructose-1,6-bisphosphate aldolase (FBA)
		Hydrolase activity	Motor activity	Paramyosin
				Myosin heavy chain (MHC), Muscle (*Drosophila melanogaster*)
				Myosin heavy chain (MHC), Straight muscle (*Argopecten irradians*)
			ATPase activity	Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type
	Transporter activity	Substrate specific transporter activity	Oxygen transporter activity	Hemocyanin C chain
	Structural molecule activity	Structural constituent for ribosome	Substrate specific transporter activity	40S ribosomal protein S28
Celluular compoment	Membrane	Membrane part	Instrinsic component of membrane	Transmembrane protein 2

Figure 4.1: All the proteins identified using Mascot for the 22 shellfish species analysed by LC/qTOF for the raw and the whole heated extracts. The identified proteins are sorted by their function (GO number) showing common abbreviation of the protein in brackets. Proteins highlighted in red are allergens registered with the IUIS and proteins highlighted in green are reported allergens in the literature.

Myosin light chain (MLC) was identified as myosin light chain (Fragment) and as myosin regulatory light chain. The latter is referred as myosin light chain (MLC) in this chapter and the identification is based on peptides matching areas of the whole myosin regulatory light chain, including various shellfish species, however, mainly mollusc species. The MLC (Fragment) consist in the Mascot database only as two peptides, matching the MLC of Black Tiger prawn (*Penaeus monodon*). Thus, MLC is a protein and MLC (Fragment) is only a tryptic dipeptide, these two identified proteins matches will be analysed separately.

The peptide matching the identified protein sarcoplasmic calcium binding protein (Fragment) ((SCBP (Fragment)) is also present in trypsin originated from pig (*Sus scrofa*). Therefore, it is assumed that this protein is not sarcoplasmic calcium binding protein (SCBP), but trypsin originated form pig, thus two categories were used for SCBP, using (1) sarcoplasmic calcium binding protein, alpha-B and A chains (SCBP), which is a known allergen[7, 18] and (2) sarcoplasmic calcium binding protein (Fragment).

Titin was identified in four mollusc species, namely octopus, scallop (*yessonensis*), Tasmanian oyster and Tuatua cockle. A maximum of two peptides have been identified for titin per species, however, the peptides are different for the four species. Moreover, the identified peptides are very short, thus, with one exception, only five amino acids long. Additionally, these short peptides can originate from many other proteins and species. Therefore, in this chapter it cannot be confirmed that titin was present in any protein extract. Similar, the peptides that have been identified for paramyosin matched many other proteins, thus it is not certain if paramyosin was present in the extracts. However, titin[13] and paramyosin[15, 16] have been described as shellfish allergens, thus these proteins are included in the protein analysis.

Protein ycf2 is not found in shellfish, hence, it is only a plant protein. The exact function of protein ycf2 is unknown, although the literature indicates that it has an essential function, which is probably not related to photosynthesis. However, this protein was identified in many shellfish species analysed, therefore it is included protein ycf2 in the analysis, although it is uncertain where protein ycf2 is originated from.

4.3.3 Analysis of the identified proteins from shellfish species utilising Mascot

All the proteins that have been identified using Mascot for various shellfish species extracts analysed by LC/qTOF are listed in figure 4.1 and will be described in the following sections. Figure 4.2 and figure 4.3 display all the proteins identified as pie chart using percentage for the raw and the whole heated extracts, divided by the subgroups of shellfish. Moreover, figure 4.4 to 4.9 visualise all the protein identified as bar graph separately for each shellfish species analysed. The percentages shown in figure 4.2 to 4.9 are based on the Mascot scores and calculations as explained in the material and method section.

The overall protein profile for the raw crustacean extracts is quite similar, with arginine kinase (AK) being the most identified protein in the raw crustacean, with 20-35% (Figure 4.2). As the detailed bar graph analysis shows, TM was only identified in one raw prawn species, representing only 5% of all proteins identified (Figure 4.5). Actin is overall identified with 5% in the raw crustacean species (Figure 4.2). For all the whole heated crustacean extracts TM is clearly the most identified protein with about 45%, followed by actin with about 10%. Nevertheless, the overall identified profile pattern for the raw crustacean and the whole heated crustacean is highly similar.

Interestingly, the identified raw and whole heated profile from the gastropods are similar to the crustacean, whereas the overall protein profile of the cephalopods and the bivalves are different to the crustacean and the gastropods, but similar to each other (Figure 4.3). In detail, in the raw gastropods the main protein identified is AK (25%) followed by actin (20%) and TM with 5%. In comparison to the gastropods and the crustacean, AK is only 2-5% of all the identified proteins in the cephalopods and the bivalves. For the whole heated gastropod extracts the main protein is TM (35%), followed by actin with 15%. The overall protein profile for the raw cephalopods and the bivalves is actin with 30%, followed by TM with 5-10%. The main protein in the whole heated extracts is actin for the cephalopods with 25% and 5% for the bivalves. The TM percentage is 25% for the bivalves and 10% for the cephalopods, respectively.

4.3.3.1 Main proteins identified for the shellfish species utilising Mascot

A total of 32 different proteins have been identified (Figure 4.1). However, analysing all the shellfish species utilising the raw and the whole heated extracts, three proteins are more commonly identified, resulting from higher Mascot scores compared to other proteins. The three main proteins are tropomyosin (TM), arginine kinase (AK) and actin (Figure 4.4 to 4.9).

TM is the major allergen in shellfish[7] and was identified in every whole heated extract, with the exception of Sydney Rock oyster. The detailed analysis for the identification of TM in the whole heated extracts can be found in chapter 3. In the raw extracts, TM was more commonly identified in molluscs compared to crustacean, thus oysters are the only mollusc species where TM was not identified. In contrast to the crustacean, where TM was only identified in the crabs, Rock lobster, Slipper lobster and Green Tiger prawn. These results confirm the finding from chapter 3, that the identification of TM in the raw extracts is achievable with the applied digestion method. Moreover, that TM is not identified in the raw prawns (with the exception of Green Tiger prawn), thus it is hypothesised that the concentration of TM in the raw prawns is too low to become identified by the LC/qTOF analysis. Especially TM being an important muscle protein and therefore must be present from a biological point of view.

Interestingly, in the raw shellfish species the protein with the overall highest Mascot score is AK, not TM. AK is a known allergen in crustacean[7] and was also reported to be a mollusc allergen.[15] Overall, AK was identified in all raw crustacean, gastropods, cephalopods and two bivalve species. The overall Mascot score for AK and the different prawn species are comparable with the ones reported for King prawn in chapter 3, confirming the findings that in the raw prawns AK is the main protein and TM cannot be identified. However, one can assume that TM is present in all the raw extracts (and all the whole heated extracts), including the species that were not identified, thus it was shown in chapter 3 by immunoblotting and with the purified TM from King prawn raw extract, that TM is present.

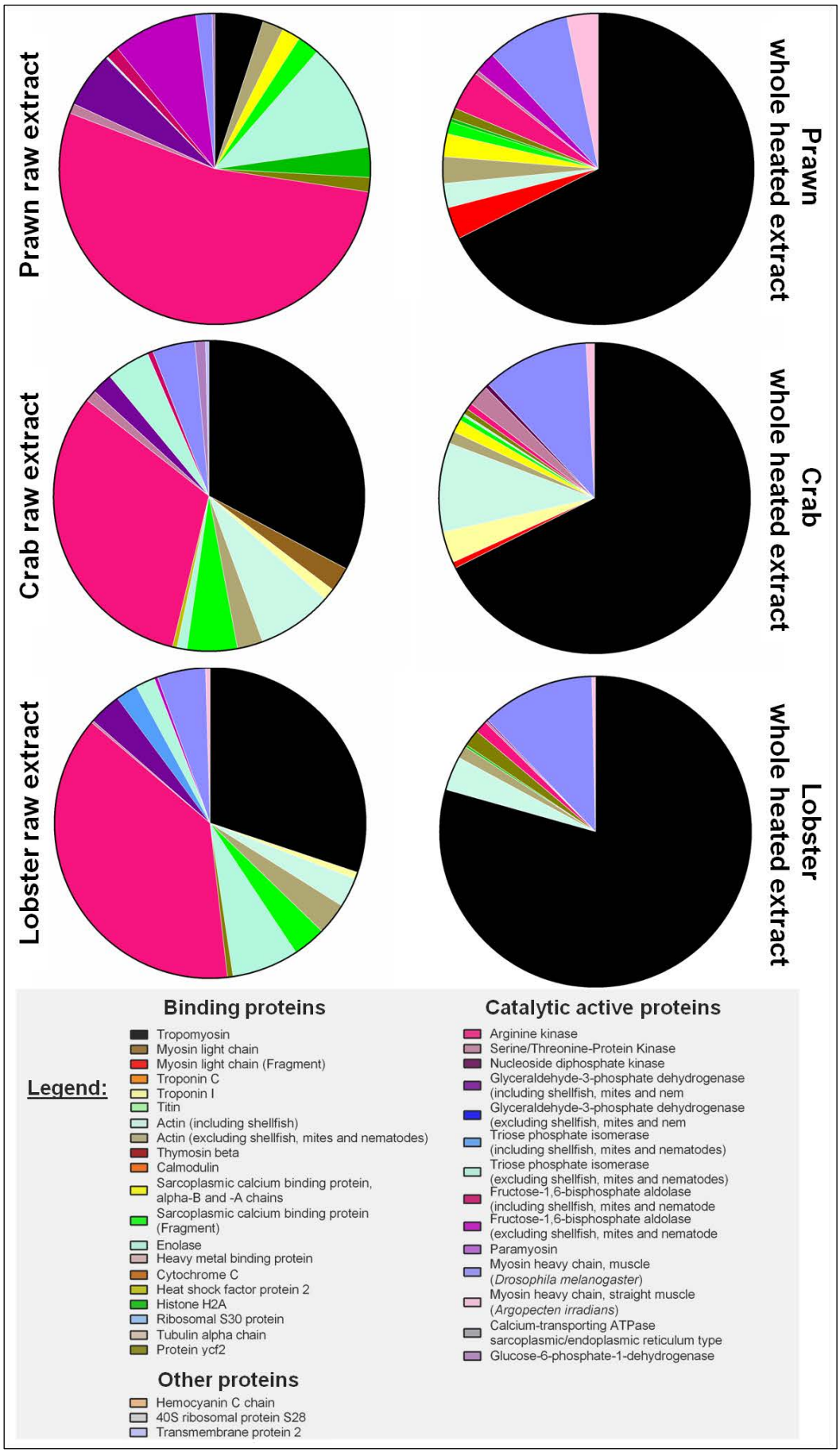


Figure 4.2: Comparison of all the proteins identified for all the crustacean subgroups in the raw and the whole heated protein extracts. Percentages for the different species were summed, followed by dividing with the number of species analysed in the subgroup.

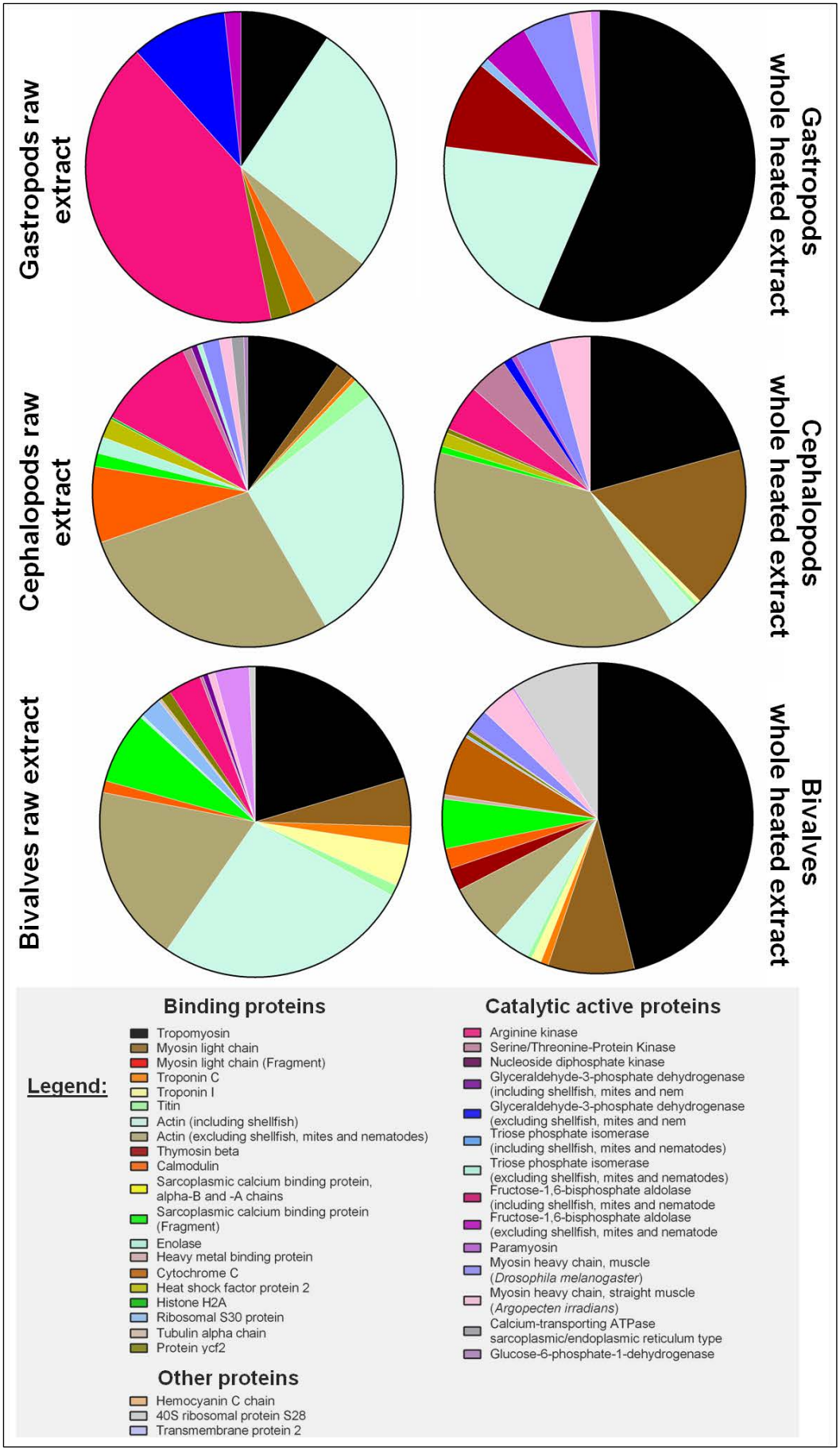


Figure 4.3: Comparison of all the proteins identified for all the mollusc subgroups in the raw and the whole heated protein extracts. Percentages for the different species were summed, followed by dividing with the number of species analysed in the subgroup.

Actin is the third protein that was identified in many shellfish species extracts. In contrast to TM and AK, actin is not commonly reported to be an allergen in any shellfish species.[11] In the raw extracts actin was identified for all species, except three prawn species. In the whole heated extracts, actin was identified in all shellfish species, except three bivalves and one prawn. Combining the raw and the whole heated extracts actin was identified in every shellfish species analysed. Actin was identified with higher percentage in the whole heated crustacean compared to the raw crustacean. In contrast to the molluscs where actin was identified more in the raw extracts compared to the whole heated molluscs extracts. However, when actin was identified using Mascot, not all samples matched shellfish species.

4.3.3.2 Other proteins identified for the shellfish species utilising Mascot

Overall, 32 different proteins have been identified, with the tree main proteins detailed above. In the following section the 29 other identified proteins will be analysed.

Figure 4.4 to 4.9 display the overall proteins identified in the crustacean and the molluscs subgroups in the raw and the whole heated extracts. The list of all proteins identified is summarised in figure 4.1. Some proteins were only identified in crustacean, others only in molluscs. The proteins only identified in crustacean are MLC (Fragment), SCBP, TIM, hemocyanin C chain and transmembrane protein 2. The proteins only identified in molluscs are troponin C, titin, thymosin beta, calmodulin, heavy metal binding protein, cytochrome C, 30S ribosomal protein, superoxide dismutase, calcium transporting ATPase sarcoplasmic/endoplasmic reticulum type and 40S ribosomal protein S28.

Transmembrane protein 2, as mentioned in section 4.3.2, is the only protein identified that is a cellular component. However, it was only identified in Blue Swimmer crab raw extract. Moreover, hemocyanin C chain was only identified in Rock lobster and yabby raw extracts. Therefore 40S ribosomal protein S28 is the only protein that is not a binding or catalytic active protein that was identified in

different extracts and different species. Nevertheless, 40S ribosomal protein S28 was only identified in molluscs, not in crustacean.

The six proteins that have only been identified in one species are (1) heavy metal binding protein in Blue mussel whole heated extract (2) cytochrome C in Sydney Rock oyster whole heated extract, although cytochrome C is a heat labile protein (3) tubulin alpha chain in Blue mussel raw extract (4) nucleoside diphosphate kinase in Blue Swimmer crab whole heated extract (5) paramyosin in squid whole heated extract and (6) calcium transporting ATPase sarcoplasmic/endoplasmic reticulum type in squid raw extract. Moreover, titin was only identified in scallop (*yessonensis*) and octopus raw extracts. Whereas glucose-6-phosphate-1-dehydrogenase was only identified in Sand crab and octopus raw extract, potentially confirming that glucose-6-phosphate-1-dehydrogenase is a heat labile protein.[11, 12]

Myosin regulatory light chain (MLC) is only identified in the cephalopods and the bivalves in the raw and the whole heated extracts and Blue Swimmer crab raw extract, although MLC is a known allergen in crustacean,[7] but has not been reported as allergen in molluscs. However, myosin light chain (Fragment) was detected in the prawns and the crabs whole heated extracts, but not in the molluscs. This different pattern for crustacean and molluscs might be due to the fact that MLC (Fragment) is only matched to the sequence of Black Tiger prawn (as explained in section 4.3.2) and the amino acid sequence of the mollusc species for myosin light chain might not include those two peptides. However, it is hypothesised that myosin light chain from crustacean was not identified as myosin regulatory light chain due to sequence unavailability.

The protein troponin (Troponin C and troponin I) has been reported as crustacean allergen,[19] but not as mollusc allergens. However, troponins were identified in three crustacean species and also in three mollusc species. In detail, for the crustacean the troponins were identified in two raw and two whole heated extracts, whereas they were identified in four raw and two whole heated mollusc species, thus overall similar to MLC, the troponins were more frequently identified in the molluscs when compared to the crustacean.

Thymosin beta was mainly identified in oysters, namely in Tasmanian oyster whole heated extract and Sydney Rock oyster in the raw and the whole heated extracts. Moreover, thymosin beta was also identified in Sea snail whole heated extract.

Calmodulin was identified in all the raw gastropods and the cephalopod extracts, with the exception of Sea snail. Additionally, it was identified in Blue Mussel whole heated extract.

Sarcoplasmic calcium binding protein (SCBP) is a known allergen in crustacean,[7] but has not been reported as allergen in molluscs. SCBP (Fragment) was identified in many shellfish species including the raw and the whole heated extracts. Nevertheless, as explained in section 4.3.2 it is hypothesised that SCBP (Fragment) is trypsin, not SCBP. Sarcoplasmic calcium binding protein, alpha-B and A chains (SCBP) was mainly identified in the prawns and Blue Swimmer crab, but not in the lobsters. Moreover, with the exception of King prawn raw extract SCBP was only identified in the whole heated extracts. However, the Mascot scores for the identification of SCBP were relatively low compared to other known allergens.

Enolase[12] is a heat labile protein and therefore was possible only detected in the raw extracts. It was mainly identified in the crustacean, but also in squid and Blue mussel. Furthermore, heat shock factor protein 2 could only be identified in octopus raw and whole heated extracts as well as whole heated Mud crab extract.

Histone H2A was only identified in the raw prawn extracts, with the exception of Blue Swimmer crab raw extract. Moreover, it was identified in Banana prawn whole heated extract. In contrast to ribosomal S30 protein, which was only identified in molluscs, namely Tuatua cockle raw and whole heated extracts and Hybrid abalone whole heated extract.

Protein ycf2 was identified in various raw and whole heated extracts. Unfortunately, there is no clear pattern in which species or extract it is more frequently identified. Nevertheless, as mentioned in section 4.3.2, protein ycf2 is only present in plants, not in the animal kingdom. Therefore it is unclear if this protein is a contamination or if the protein is really present in all these different shellfish species. The identification of this protein could be a possible contamination, thus the Mascot scores achieved are relatively low in comparison

to other proteins identified. Moreover, the Mascot scores in the different extracts and the shellfish species did not differ much. The same minimal variation of the Mascot scores and overall distribution of various extracts and the shellfish species analysed was observed for serine/theroine protein kinase. Moreover, when serine/theroine protein kinase was identified it was matched to different bacteria strains, never to shellfish species or shellfish related species. Therefore it is assumed that serine/theroine protein kinase was not identified in the shellfish species, but is more likely a contamination.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) has been reported to be a heat stable crustacean allergen.[11, 12] However, GAPDH was only identified in one whole heated extract. Nevertheless, in the raw extracts it has been identified in the crustacean and the molluscs. With two exceptions, GAPDH has been identified with species that match species from shellfish, mites or nematodes. Moreover, it was identified more often in the crustacean species compared to the molluscs.

Superoxide dismutase was mainly identified in the bivalves and Hybrid abalone whole heated extract. However, the Mascot score was relatively low for all the species including superoxide dismutase.

Triose phosphate isomerase (TIM) has been reported as heat labile crustacean allergen,[13] hence it was not identified in the whole heated extracts. Similar fructose-1,6-bisphosphate aldolase (FBA) is a known crustacean allergen,[13] but has not been reported as allergen for molluscs. Unfortunately, TIM and FBA were mainly matched to species that did not included shellfish species, mites or nematodes. This might be due to the sequence unavailability or diverse amino acid sequences of shellfish for TIM and FBA. With the exception of squid TIM was only identified in the crustacean. Similar, FBA was only identified in the crustacean, with the exception of Sea snail and squid.

Myosin heavy chain (MHC) was identified throughout all extracts and species analysed. However, the majority of MHC was found in the heated extracts. Moreover, MHC muscle was identified more often compared to MHC straight muscle, especially in the whole heated extracts. The only whole heated crustacean extract where MHC muscle was not identified is yabby. Interestingly, it appears that the identification of MHC is higher in the whole heated extracts compared to the raw extracts. It is hypothesised, this might be due the

temperature treatment, consequently denaturing MHC and leading to better digestion efficacy. Overall, the identification of both myosin heavy chains was good, although the Mascot scores are relatively low. This can be explained with the size of MHC, thus although several peptides were matched to MHC, the sequence coverage was low considering the whole protein.

4.3.4 Species specific analysis of the identified proteins

4.3.4.1 Proteins identified in prawn species

Surprisingly, in the raw prawn extracts the major shellfish allergen, tropomyosin (TM), was only identified in Green Tiger prawn out of five different prawn species investigated (Figure 4.4). The main protein identified in the raw prawn extracts is obviously arginine kinase (AK) ranging from 23-75% based on the Mascot scores. Other proteins that have been identified in at least in one species with over 20% are enolase, GAPDH and FBA as compared to the whole heated extract, where the main protein is definitely TM. With the exception of MHC muscle all the other proteins were below 10%. Furthermore, all the identified proteins are either binding proteins or catalytic active proteins.

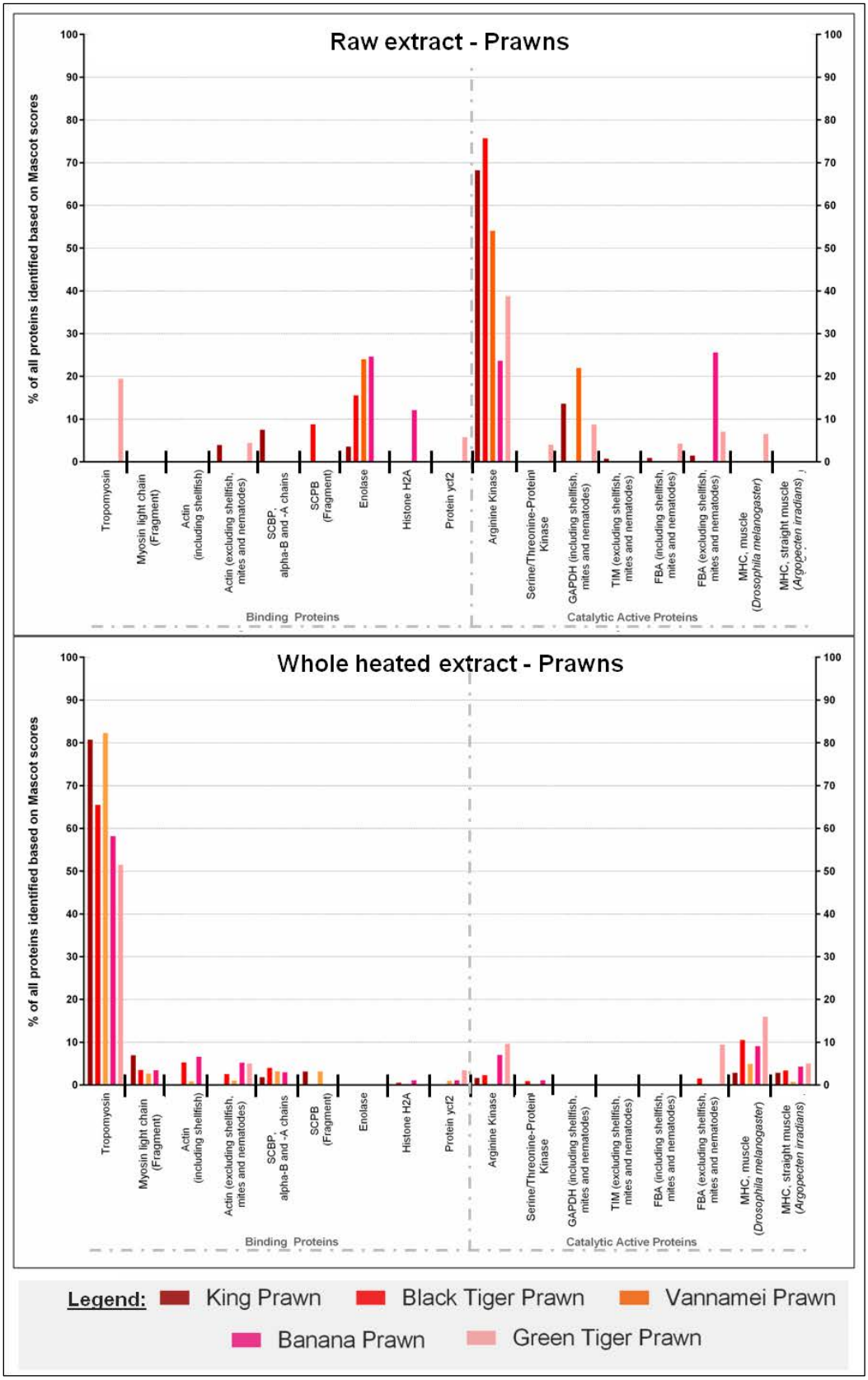


Figure 4.4: All the proteins identified for the analysed prawn species by LC/qTOF utilising raw and whole heated extracts. The identified proteins are shown in percent, calculated based on the Mascot scores for the identified proteins.

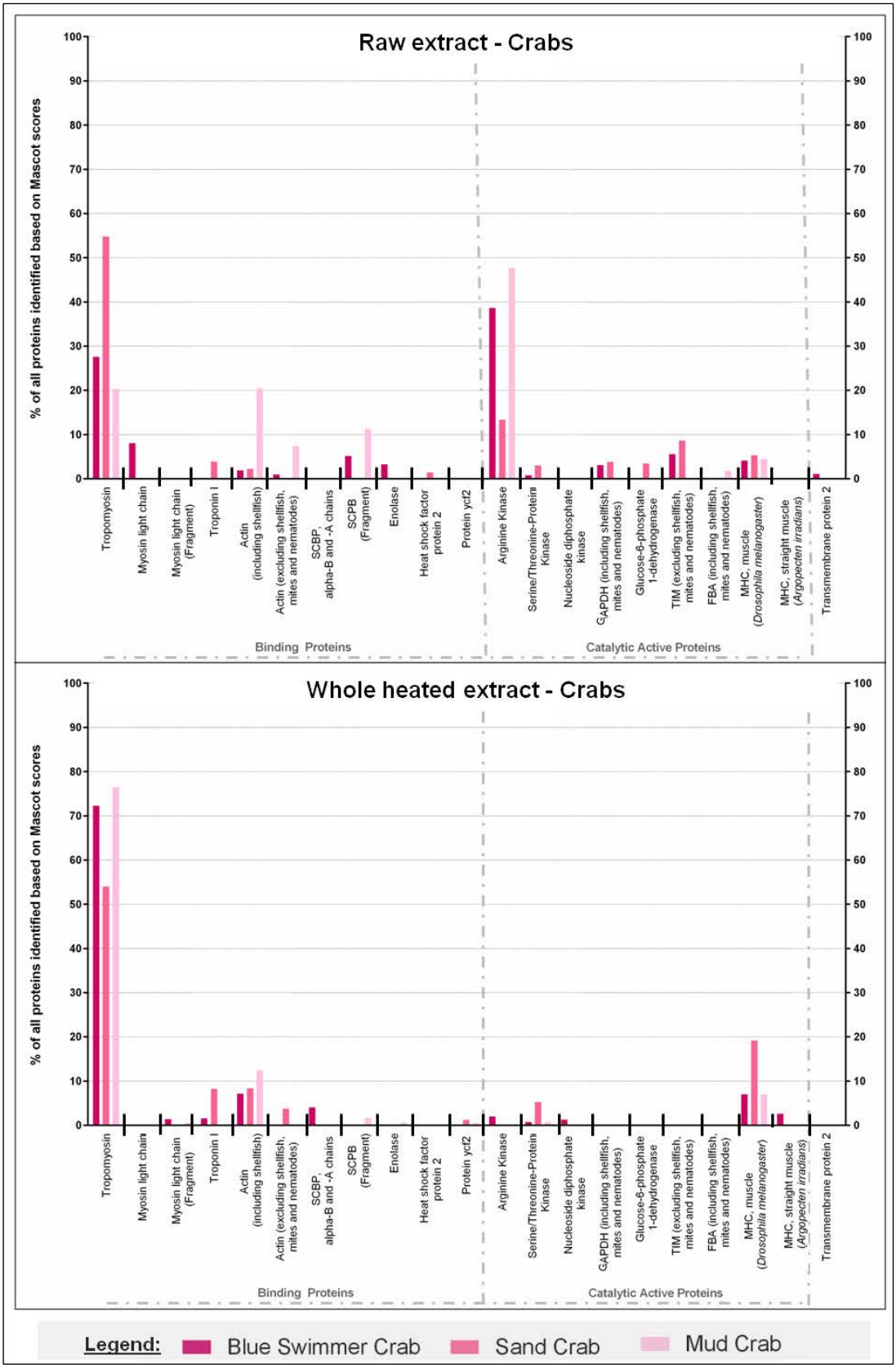


Figure 4.5: All the proteins identified for the analysed crab species by LC/qTOF utilising raw and whole heated extracts. The identified proteins are shown in percent, calculated based on the Mascot scores for the identified proteins.

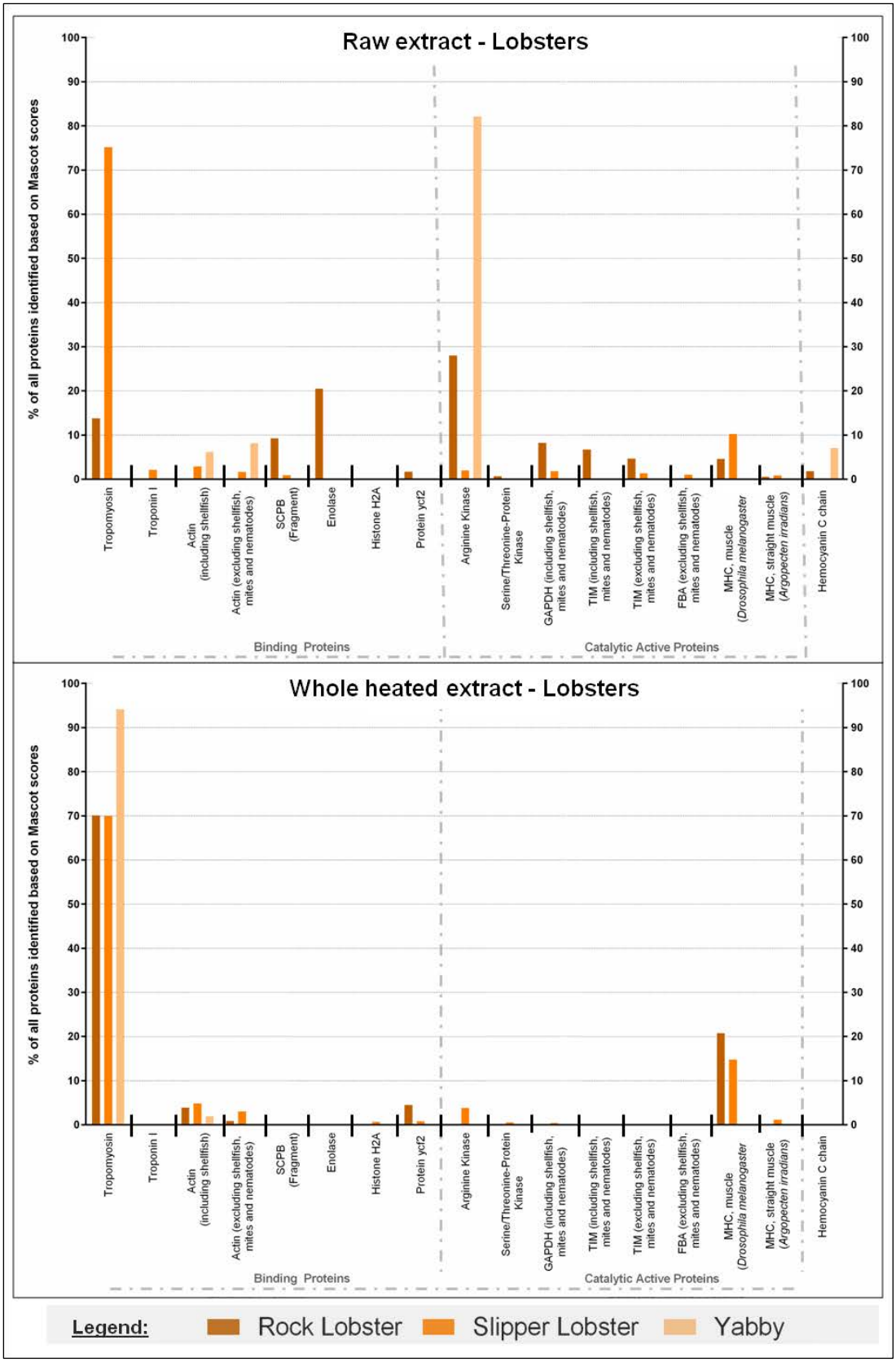


Figure 4.6: All the proteins identified for the analysed lobster species by LC/qTOF utilising raw and whole heated extracts. The identified proteins are shown in percent, calculated based on the Mascot scores for the identified proteins.

4.3.4.2 Proteins identified in crab species

Figure 4.5 displays that in the crab raw extracts TM could be identified. TM and AK are the two main proteins having higher Mascot scores, with 65% of the proteins identified. Interestingly, when the Mascot score was higher for AK, TM score was lower and vice versa. Other proteins identified were less than 10% based on the Mascot scores, with the exception of actin and SCBP (Fragment), identified in Mud crab. All the proteins identified in the raw extract are either binding proteins or catalytic active proteins, with one protein not belonging to these two groups in Blue Swimmer crab. However, for the whole heated extracts, only binding proteins and catalytic active proteins have been identified. In the whole heated extract TM is clearly the protein with the highest Mascot scores for all the species investigated. Similar to the raw and the whole heated extracts, only for Mud crab actin was identified with more than 10%, whereas for Sand crab MHC muscle was identified with 19%. Overall, for the raw extract TM and AK are the main proteins, whereas for the whole heated extracts it is TM.

4.3.4.3 Proteins identified in lobster species

Very similar to the raw crabs, the two main proteins in the raw lobsters are TM and AK (Figure 4.6). However, the Mascot scores added for TM and AK lead to more than 75% for Slipper lobster and Yabby, but only 40% for Rock lobster. All the other proteins identified are below 10%, with the exception of enolase for Rock lobster. Similar to the crabs, for the raw and the whole heated lobster extracts, all the proteins identified are either binding proteins or catalytic active proteins, with the exception of hemocyanin C chain, which was identified in raw Rock lobster and raw yabby extracts. In the whole heated lobster extracts the main protein is TM, followed by MHC muscle with over 10% for Rock lobster and Slipper lobster. All the other proteins identified in the whole heated lobster extracts are below 5%.

Overall, in the raw crustacean and the whole heated crustacean the two main proteins identified are AK and TM, whereas the in whole heated extract it is definitely TM. Moreover, MHC muscle seems to be more present in the crustacean extracts compared to MHC straight muscle. Furthermore, the percentage for MHC

muscle seems to increase slightly for the whole heated extracts compared to the raw extracts.

4.3.4.4 Proteins identified in gastropod species

The least amount of proteins was identified for the gastropods (Figure 4.7), with only 14 proteins in comparison to the bivalves with the maximum of 24 identified proteins. Nevertheless, in comparison to the crustacean, in the raw gastropod extracts three proteins are identified with higher Mascot scores, namely AK, TM and actin, whereas the Mascot scores were higher for actin in comparison to TM. All the other proteins have been identified with less than 5%, with the exception of GAPHD in raw Sea snail extract. Similar to the prawns, all the proteins identified are either binding or catalytic active proteins in the raw or the whole extracts. The two main proteins in the whole heated extracts are TM, followed by actin. All other proteins identified are below 10%, with the exception of thymosin beta in Sea snail.

4.3.4.5 Proteins identified in cephalopod species

Figure 4.8 visualises that for the raw cephalopod extracts the main protein identified is actin. AK, being the main protein for all the other raw species investigated, is only present with 20% for squid and below 5% for octopus. TM is the only other protein identified in squid having more than 10%. Identical to the gastropods, only binding proteins and catalytic active proteins have been identified for the raw and the whole heated cephalopod extracts. The three main proteins identified in the whole heated cephalopod extracts are actin, TM and myosin light chain (MLC). For squid the order of these proteins are TM and actin, followed by MLC, whereas for octopus the order is actin, MLC and then TM. All the other proteins identified were below 10%.

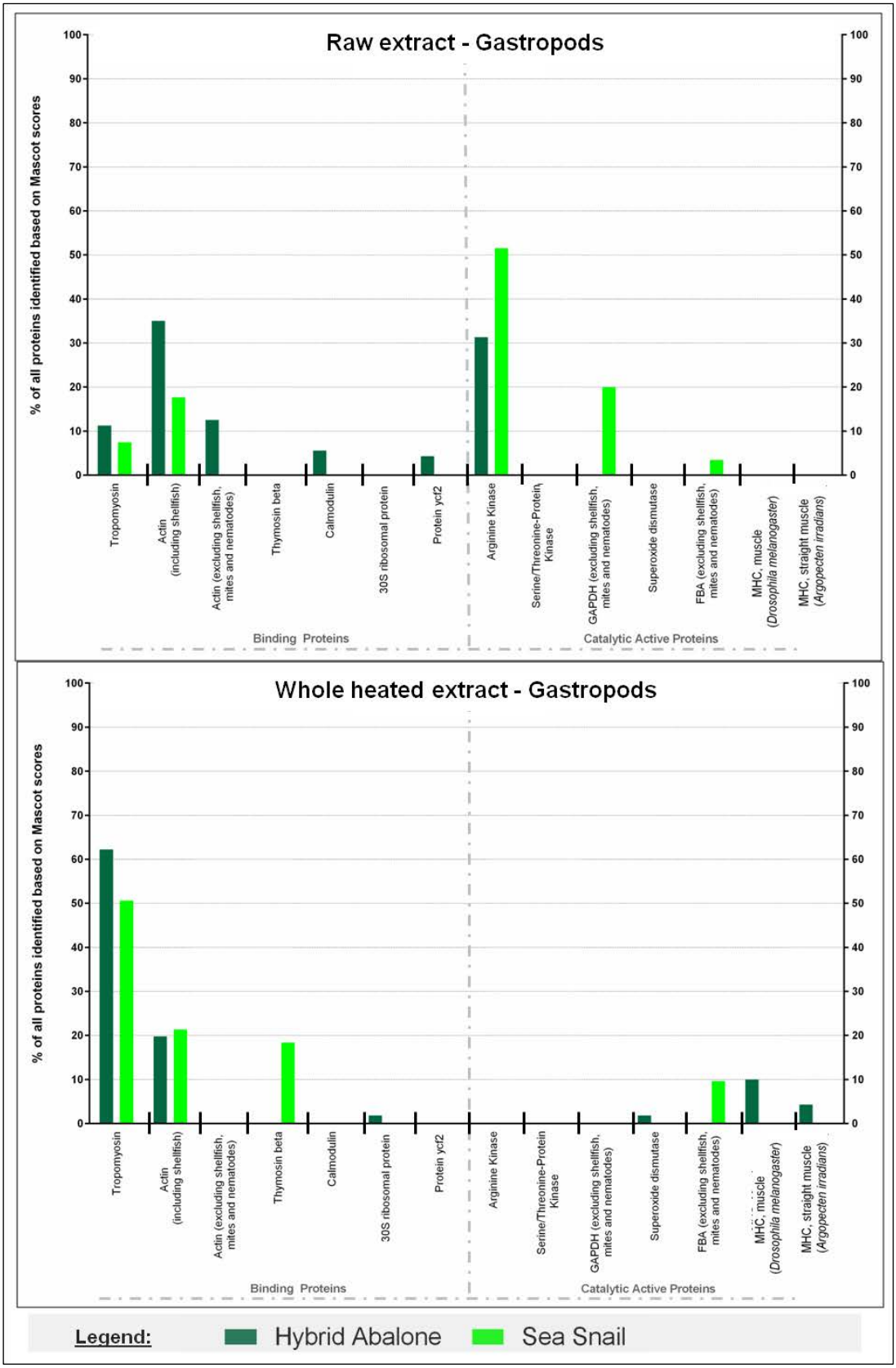


Figure 4.7: All the proteins identified for the analysed gastropod species by LC/qTOF utilising raw and whole heated extracts. The identified proteins are shown in percent, calculated based on the Mascot scores for the identified proteins.

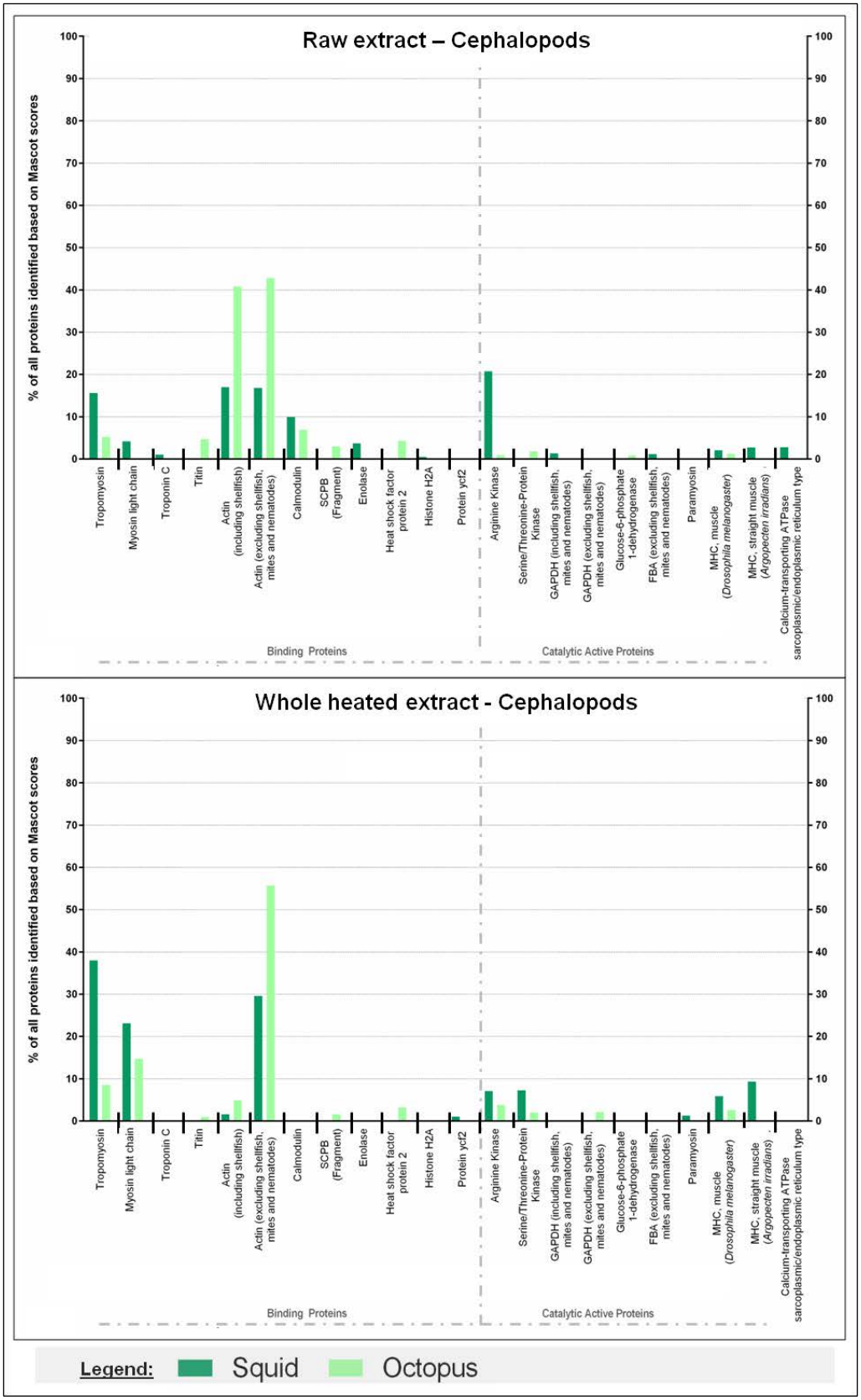


Figure 4.8: All the proteins identified for the analysed cephalopod species by LC/qTOF utilising raw and whole heated extracts. The identified proteins are shown in percent, calculated based on the Mascot scores for the identified proteins.

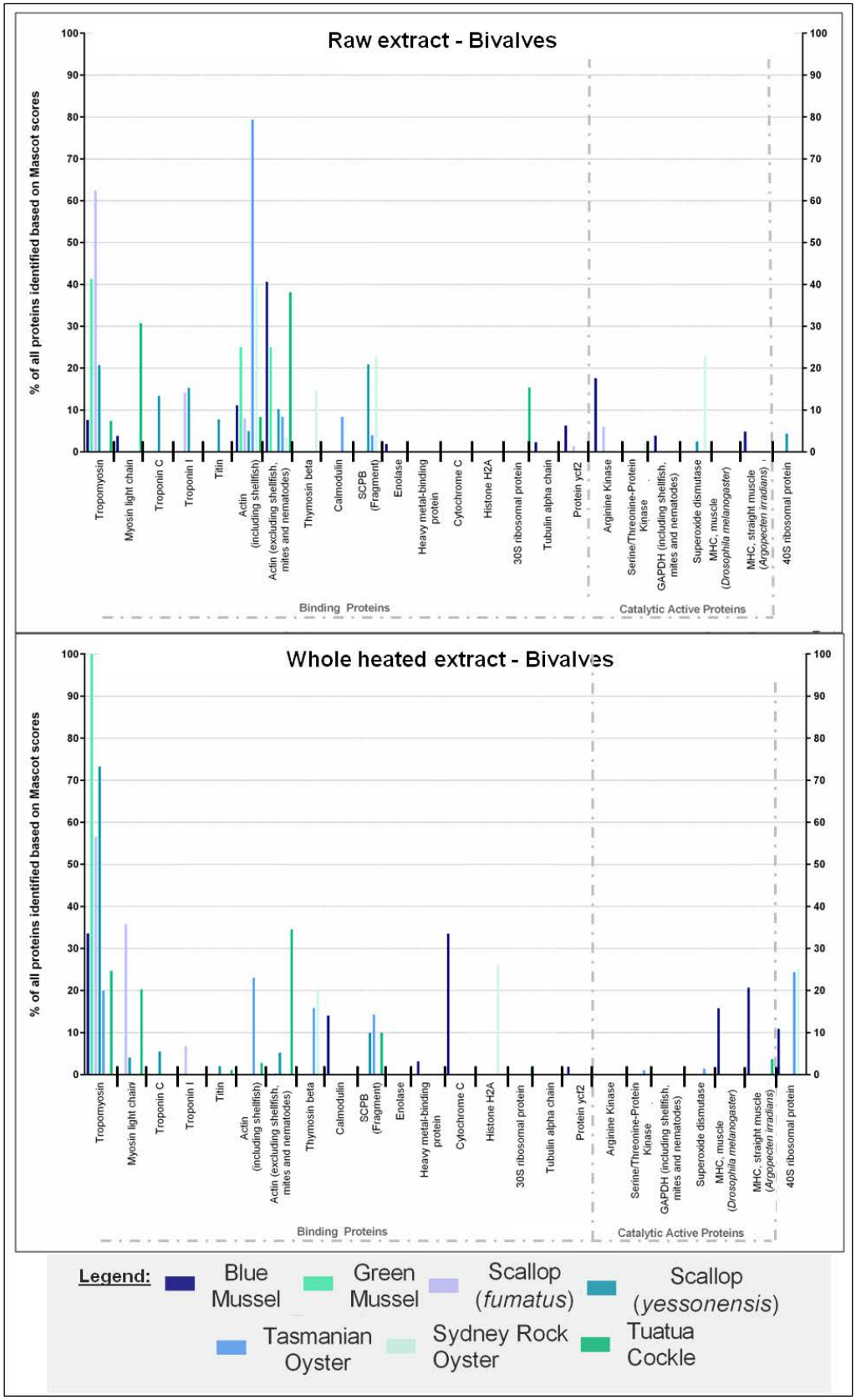


Figure 4.9: All the proteins identified for the analysed bivalve species by LC/qTOF utilising raw and whole heated extracts. The identified proteins are shown in percent, calculated based on the Mascot scores for the identified proteins.

4.3.4.6 Proteins identified in bivalve species

Interestingly, more binding proteins, in comparison to the other subgroups, were identified for the bivalves (Figure 4.9). However, similar to the other mollusc species the main protein identified is clearly actin, followed by TM. There are several other proteins that have been identified with more than 10%, thus it seems the protein pattern is different for the bivalves compared to the other shellfish species. Surprisingly, compared to all the other species AK was only identified for two species in the raw extracts and not in the whole heated extracts. For the whole heated bivalve extracts the main protein identified is TM, for Green mussel is was the only protein being identified. Similar to the other species, in the whole heated extracts more MHC is identified, although the relation of MHC muscle and MHC straight muscle does not seem to change. Actin is still present in three bivalve species after the heat treatment.

Overall, in the raw molluscs and the whole heated molluscs the three main proteins are actin, AK and TM. Moreover, the percentage for MHC muscle seems to increase slightly for the whole heated extracts compared to the raw extracts, similar as observed for the crustacean.

4.3.5 Possible allergen-candidate analysis

Every protein can be potentially an allergen, however, certain biochemical characteristics increase the chances of a protein being allergenic. A protein is more likely to be an allergen if the protein (1) shares less than 62% amino acid sequence homology with the human analogue, (2) belongs to specific protein families, (3) shows a certain structure, and (4) is heat stable. Therefore all 32 identified proteins are analysed for being potential allergen-candidates applying these four criteria (Figure 4.10). Interestingly, when all the identified proteins are classified according their structure using the Structural Classification Proteins (SCOP) database, it was observed that all the identified proteins have potential allergenic structures, with the exception of calmodulin and protein ycf2. However, only ten identified proteins belong to the 17 important food allergen protein families, with eight being already reported allergens. Only four proteins are heat

labile, however, three heat labile proteins already have been reported as being allergenic.

Figure 4.10 summarises that seven identified proteins are registered with the IUIS as allergens. Moreover, nine proteins have been reported as being allergens. With the exception of actin,[11] enolase,[12] and MHC,[11, 12] one can conclude that all the reported allergens are definitely allergens. However, MHC and enolase have a high amino acid sequence identity with the human homologue (50-70% and 60-65%), thus it is suggested MHC and enolase might be potential allergen. Furthermore, the amino acid identified of actin with the human homologue is 92-99%, thus it is hypothesised that actin from shellfish is not allergenic to humans.

Sixteen proteins have been identified, which have not previously been reported as shellfish allergens, whereas nine proteins are not considered as potential allergens, as the amino acid sequence homology of the shellfish proteins is similar to the human analogues. Moreover, one protein is heat labile and one does not share the same SCOP structure. Nevertheless, eight identified proteins have low amino acid sequence identity with the human analogue, share similar SCOP structure and are heat stable. The identification of two proteins, protein ycf2 and serine/theronine protein kinase, using Mascot was low and matched to non shellfish species. Therefore it is hypothesised that these two proteins could be allergens, but not shellfish allergens.

Overall, six identified proteins could be potential novel shellfish allergens, which have not been identified in the literature. However, further work and experiments are necessary to confirm that these six proteins are actual allergens.

Protein	Sequence Identity to humans	Belonging to 17 important protein families	heat stable	Function	Structure (SCOP)	Reported allergen	Is it an allergen?
Tropomyosin	~55%	Tropomyosin	yes	Muscle protein	Coiled coil protein	Registered with IUIS	**yes**
Myosin light chain	~10-40%	EF hand domain	yes	Muscle protein	α protein	Registered with IUIS	**yes**
Troponin C	~20-45%	EF hand domain	yes	Muscle protein	α protein	Registered with IUIS	**yes**
Troponin I	~20-30%	no	yes	Muscle protein	Coiled coil protein	Registered with IUIS	**yes**
Titin	~ 30%	no	yes	Muscle protein	β protein	Reported allergen	**yes**
Actin	~92-99%	no	yes	Muscle protein	α protein	Reported allergen	no
Thymosin beta	~10-70%	EF hand domain	yes	Muscle protein	Peptide	no	possible
Calmodulin	~50-98%	EF hand domain	yes	Muscle protein	α protein	no	no
Sarcoplasmic calcium binding protein	~5-12%	EF hand domain	yes	Muscle protein	α protein	Registered with IUIS	**yes**
Enolase	~60-65%	no	no	Glycolysis	α and β protein	Reported allergen	possible
Heavy metal binding protein	~5-25%	no	yes	Detoxification	α and β protein	no	possible
Cytochrome C	~70-75%	no	no	ATP synthesis	α protein	no	no
Heat shock factor protein 2	~60-80%	no	yes	Chaperone	α protein	no	no
Histone H2A	~80-90%	no	yes	DNA folding	α protein	no	no
Ribosomal S30 protein	~3%	no	yes	Ribosomal protein	α protein	no	possible
Tubulin alpha chain	~70-75%	no	yes	Globular protein	α and β protein	no	no
Protein ycf2	no human analogue	no	yes	Cell survival	unknown	no	possible
Arginine kinase	~35-55%	Arginine kinase	yes	ATP synthesis	α protein	Registered with IUIS	**yes**
Serine/Threonine protein kinase	no human analogue	no	yes	ATP synthesis	α protein	no	possible
Nucleoside diphosphate kinase	~20-50%	no	yes	ATP synthesis	α and β protein	no	possible
Glyceraldehyde-3-phosphate dehydrogenase	~55-75%	no	yes	Glycolysis	α and β protein	Reported allergen	**yes**
Glucose-6-phosphate-1-dehydrogenase	~60-70%	no	no	NADPH synthesis	α and β protein	no	no
Superoxide dismutase [Cu-Zn]	~10-44%	no	yes	Oxygen binding	β protein	no	possible
Triose phosphate isomerase	~35-48%	no	no	Glycolysis	α and β protein	Registered with IUIS	**yes**
Fructose-1,6-bisphosphate aldolase	~60-80%	no	yes	Glycolysis	α and β protein	Reported allergen	**yes**
Paramyosin	~30-40	Myosin tail	yes	Muscle protein	Coiled coil protein	Reported allergen	**yes**
Myosin heavy chain, muscle	~50-70%	Myosin tail	yes	Muscle protein	Coiled coil protein	Reported allergen	possible
Myosin heavy chain, straight muscle	~50-70%	Myosin tail	yes	Muscle protein	Coiled coil protein	Reported allergen	possible
Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type	~72%	no	yes	ATP synthesis	β protein	no	no
Hemocyanin C chain	no human analogue	no	yes	Oxygen binding	α protein	Reported allergen	**yes**
40S ribosomal protein S28	~84%	no	yes	Ribosomal protein	α protein	no	no
Transmembrane protein 2	~5-30%	no	yes	Membrane protein	α and β protein	no	possible

Figure 4.10: Possible allergen-candidate analysis for all the identified proteins in various shellfish extracts analysed by LC/qTOF. The figure summarises for the identified proteins their amino acid sequence homology to the human analogue, their function, and their molecular structure, if the protein was previously reported as allergen and if the protein could be a potential allergen based on specific molecular characteristics.

4.4 Discussion

Twenty-two shellfish species have been analysed using the raw and the whole heated protein extracts by LC/qTOF. The LC/qTOF data was analysed with Mascot to identify various proteins present in the different shellfish species extracts. Overall, 32 different proteins were identified applying these methods. However, the identification of proteins using Mascot needs to be carefully investigated, thus some protein matches are based on protein fragments and not the whole protein sequence. Moreover, it is important to confirm which species the identified protein matches, thus the species of interests are not always included. This could be due to unknown sequences or potential false positive peptide matches and therefore needs to be carefully considered.

Several protein that have been identified in this chapter were identified in previous studies analysing seven prawns,[11, 18-23] one crab[11, 24] and one octopus[15] species. The identification of tropomyosin in various shellfish species by mass spectrometry was explained and discussed in chapter 3, therefore will not be discussed in this chapter. Myosin light chain (MLC) was identified for three prawns, matching the Pacific White prawn,[11] *Artemia franciscana*[19] or German cockroach.[20] Neither publication mentioned that MLC was identified as MLC (Fragment) or myosin regulatory light chain, as MLC was identified in this chapter. Similar, sarcoplasmic calcium binding protein (SCBP) was identified in Pacific White prawn[11] in one study, but as sarcoplasmic calcium binding protein, alpha-B and A chains[18] in another study. The latter identification is detailed in this chapter, however, neither study report the identification of SCBP (Fragment) where one can assume that it is originated form trypsin, not from SCBP. Nevertheless, it was reported that MLC and SCBP seem to be cross-reactive allergens within crustacean and some other invertebrates, where SCBP shows high amino acid sequence identity, in contrast to MLC having relatively low amino acid identity within the shellfish group.[18, 20, 25]

Arginine kinase (AK) is reported to be an allergen in crustacean and molluscs. In comparison to TM, AK shares less amino acid identity within the shellfish group, however, it was still reported to be a cross-reactive allergen in the crustacean and insects.[15, 25, 26] Moreover, although AK is an enzyme it was reported to be heat stable,[13, 26] whereas crustacean AK seems to be more stable compared to octopus.[15] Although AK is reported as minor allergen in shellfish, this chapter

demonstrated that in many raw shellfish extracts AK was the main identified protein, not tropomyosin. The identification of AK in the literature was achieved with good Mascot scores and related species, as demonstrated by four studies[11, 15, 21, 27] and nine shellfish species analysed. AK was identified in this chapter with very similar Mascot scores and identified peptides as reported in the literature analysing the raw and the whole heated extracts. It was reported that AK is still immune reactive after tryptic difestion,[15] thus it is suggest that AK, as a shellfish allergen, needs further investigations, because it was mainly present in the raw extracts and therefore believed that AK deserves more attention as cross-reactive shellfish allergen.

Other proteins identified for crustacean in the literature have been reported by Abdel Rahman[11, 13, 27] and are actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), myosin heavy chain (MHC), titin, troponin C and SERCA/ smooth endoplasmic reticulum Ca2+ATPase, whereas the latter was the only protein not identified in this chapter. The identification of actin, GAPDH and troponin C in this chapter was similar achieved by Abdel Rahman,[11] matching shellfish species, whereas MHC was matched to an insect. Nevertheless, titin was identified with two peptides,[13] whereas in this chapter titin was identified with different peptides, which are rather short and could be obtained from different proteins and species. Therefore one can assume titin was not present in the shellfish species analysed in this chapter. Moreover, Abdel Rahman[11] found actin, GAPDH and MHC to be allergenic. According to the protein analysis, GAPDH and MHC can be possible allergens, however, it is hypothesised that actin is not an allergen due to high amino acid homology with the human analogue.

The 32 proteins identified in this chapter have been analysed for their potential allergenicity. Therefore the structure and the protein families were analysed according to Jenkins et al.[3] and Radauer et al.[2] With the exception of thymosin beta, all the identified proteins obtain structures that have higher prevalence of being allergens.[2] However, only ten identified proteins belong to the 17 main protein families that were described as including more allergens compared to other protein families.[3] Interestingly, when the identified proteins were analysed for their function using the GO Annotation Database[2] it was found that 29 proteins are either binding proteins or catalytic active proteins. Thus, 17 binding proteins and

13 catalytic active proteins have been identified. It was also reported by Radauer et al.[2] that most allergenic proteins are binding protein or catalytic active proteins.

Sixteen identified proteins already have been reported as being shellfish allergens. The analysis of the identified proteins in this study confirmed that 15 proteins are allergenic proteins, however, as explained above, it cannot be confirm that actin[11] is a possible allergen. Furthermore, eight other identified proteins could be potential allergens, whereas for two proteins the identification was poor.

Overall, in this study 32 proteins were identified in various shellfish extracts. Twenty-one proteins are possible allergens. Moreover, enolase, FBA, MHC, MLC, SCBP, troponin C, troponin I and titin were identified in most mollusc species, although presently only reported as crustacean allergens. The main protein identified in the raw crustacean extracts is AK, not TM, thus it is hypothesised that AK as a cross-reactive shellfish allergen might have a bigger impact than currently known. The main proteins identified in the molluscs raw and the whole heated extracts, are actin, followed by TM. Actin was identified throughout all the mollusc extracts analysed. Therefore it seems that the crustacean might be overall potentially more allergenic than the molluscs, thus the main proteins identified in the raw and the whole heated crustacean extracts are well established allergens, whereas it is hypothesised that actin is not an allergen.

4.5 References

1. Breiteneder, H.; Mills, E. N. C. (2005). *Molecular Properties of Food Allergens.* Journal of Allergy and Clinical Immunology. 115, (1), 14-23.

2. Radauer, C.; Bublin, M.; Wagner, S.; Mari, A.; Breiteneder, H. (2008). *Allergens Are Distributed into Few Protein Families and Possess a Restricted Number of Biochemical Functions.* Journal of Allergy and Clinical Immunology. 121, (4), 847-852.

3. Jenkins, J. A.; Breiteneder, H.; Mills, E. N. C. (2007). *Evolutionary Distance from Human Homologs Reflects Allergenicity of Animal Food Proteins.* Journal of Allergy and Clinical Immunology. 120, (6), 1399-1405.

4. Jenkins, J. A.; Griffiths-Jones, S.; Shewry, P. R.; Breiteneder, H.; Mills, E. N. C. (2005). *Structural Relatedness of Plant Food Allergens with Specific Reference to Cross-Reactive Allergens: An in Silico Analysis.* Journal of Allergy and Clinical Immunology. 115, (1), 163-170.

5. Breiteneder, H.; Mills, E. N. C. (2005). *Plant Food Allergens—Structural and Functional Aspects of Allergenicity.* Biotechnology Advances. 23, (6), 395-399.

6. Ivanciuc, O.; Garcia, T.; Torres, M.; Schein, C. H.; Braun, W. (2009). *Characteristic Motifs for Families of Allergenic Proteins.* Molecular Immunology. 46, (4), 559-568.

7. Lopata, A. L.; O'Hehir, R. E.; Lehrer, S. B. (2010). *Shellfish Allergy.* Clinical and Experimental Allergy. 40, (6), 850-858.

8. Pomes, A. (2008). *Common Structures of Allergens.* Revue Francaise D Allergologie Et D Immunologie Clinique. 48, (3), 139-142.

9. Sicherer, S. H.; Sampson, H. A. (2010). *Food Allergy.* Journal of Allergy and Clinical Immunology. 125, (2), S116-S125.

10. Breiteneder, H. (2008). *Can Any Protein Become an Allergen?* Revue Francaise D Allergologie Et D Immunologie Clinique. 48, (3), 135-138.

11. Abdel Rahman, A. M.; Kamath, S. D.; Gagne, S.; Lopata, A. L.; Helleur, R. (2013). *Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention.* Journal of Proteome Research. 12, (2), 647-656.

12. Khanaruksombat, S.; Srisomsap, C.; Chokchaichamnankit, D.; Punyarit, P.; Phiriyangkul, P. (2014). *Identification of a Novel Allergen from Muscle and Various Organs in Banana Shrimp (Fenneropenaeus Merguiensis).* Annals of Allergy, Asthma & Immunology. 113, (3), 301-306.

13. Kamath, S. D.; Abdel Rahman, A. M.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. 58, (5), 1144-1155.

14. Piboonpocanun, S.; Jirapongsananuruk, O.; Tipayanon, T.; Boonchoo, S.; Goodman, R. E. (2011). *Identification of Hemocyanin as a Novel Non-Cross-Reactive Allergen from the Giant Freshwater Shrimp Macrobrachium Rosenbergii.* Molecular Nutrition & Food Research. 55, (10), 1492-1498.

15. Shen, H. W.; Cao, M. J.; Cai, Q. F.; Ruan, M. M.; Mao, H. Y.; Su, W. J.; Liu, G. M. (2012). *Purification, Cloning, and Immunological Characterization of Arginine Kinase, a Novel Allergen of Octopus Fangsiao.* Journal of Agricultural and Food Chemistry. 60, (9), 2190-2199.

16. Suzuki, M.; Kobayashi, Y.; Hiraki, Y.; Nakata, H.; Shiomi, K. (2011). *Paramyosin of the Disc Abalone Haliotis Discus Discus: Identification as a New Allergen and Cross-Reactivity with Tropomyosin.* Food Chemistry. 124, (3), 921-926.

17. Radauer, C.; Breiteneder, H. *Evolutionary Biology of Plant Food Allergens.* Journal of Allergy and Clinical Immunology. 120, (3), 518-525.

18. Ayuso, R.; Grishina, G.; Ibanez, M. D.; Blanco, C.; Carrillo, T.; Bencharitiwong, R.; Sanchez, S.; Nowak-Wegrzyn, A.; Sampson, H. A. (2009). *Sarcoplasmic Calcium-Binding Protein Is an Ef-Hand-Type Protein Identified as a New Shrimp Allergen.* Journal of Allergy and Clinical Immunology. 124, (1), 114-120.

19. Bauermeister, K.; Wangorsch, A.; Garoffo, L. P.; Reuter, A.; Conti, A.; Taylor, S. L.; Lidholm, J.; DeWitt, Å. M.; Enrique, E.; Vieths, S.; Holzhauser, T.; Ballmer-Weber, B.; Reese, G. (2011). *Generation of a Comprehensive Panel of Crustacean Allergens from the North Sea Shrimp Crangon Crangon.* Molecular Immunology. 48, (15–16), 1983-1992.

20. Ayuso, R.; Grishina, G.; Bardina, L.; Carrillo, T.; Blanco, C.; Ibanez, M. D.; Sampson, H. A.; Beyer, K. (2008). *Myosin Light Chain Is a Novel Shrimp Allergen, Lit V 3.* Journal of Allergy and Clinical Immunology. 122, (4), 795-802.

21. Ortea, I.; Canas, B.; Calo-Mata, P.; Barros-Velazquez, J.; Gallardo, J. M. (2009). *Arginine Kinase Peptide Mass Fingerprinting as a Proteomic Approach for Species Identification and Taxonomic Analysis of Commercially Relevant Shrimp Species.* Journal of Agricultural and Food Chemistry. 57, (13), 5665-5672.

22. Ortea, I.; Canas, B.; Gallardo, J. M. (2011). *Selected Tandem Mass Spectrometry Ion Monitoring for the Fast Identification of Seafood Species.* Journal of Chromatography A. 1218, (28), 4445-4451.

23. Abdel Rahman, A. M.; Kamath, S.; Lopata, A. L.; Helleur, R. J. (2010). *Analysis of the Allergenic Proteins in Black Tiger Prawn (Penaeus Monodon) and Characterization of the Major Allergen Tropomyosin Using Mass Spectrometry.* Rapid Communications in Mass Spectrometry. 24, (24), 3624-3624.

24. Abdel Rahman, A. M.; Lopata, A. L.; O'Hehir, R. E.; Robinson, J. J.; Banoub, J. H.; Helleur, R. J. (2010). *Characterization and De Novo Sequencing of Snow Crab Tropomyosin Enzymatic Peptides by Both Electrospary Ionization and Matrix-Assisted Laser Desorption Ionization Qqtof Tandem Mass Spectrometry.* Journal of Mass Spectrometry. 45, (4), 372-381.

25. Ayuso, R. (2011). *Update on the Diagnosis and Treatment of Shellfish Allergy.* Current Allergy and Asthma Reports. 11, (4), 309-316.

26. Yadzir, Z. H. M.; Misnan, R.; Abdullah, N.; Bakhtiar, F.; Arip, M.; Murad, S. (2010). *Identification of Ige-Binding Proteins of Raw and Cooked Extracts of Loligo Edulis (White Squid).* Southeast Asian Journal of Tropical Medicine and Public Health. 41, (3), 653-659.

27. Abdel Rahman, A. M.; Kamath, S. D.; Lopata, A. L.; Robinson, J. J.; Helleur, R. J. (2011). *Biomolecular Characterization of Allergenic Proteins in Snow Crab (Chionoecetes Opilio) and De Novo Sequencing of the Second Allergen Arginine Kinase Using Tandem Mass Spectrometry*. Journal of Proteomics. 74, (2), 231-241.

28. Lasonder, E.; Ishihama, Y.; Andersen, J.; Vermunt, A. M. W.; Pain, A.; Sauerwein, R. W.; Eling, W. M. C.; Hall, N.; Waters, A. P.; Stunnenberg, H. G.; Mann, M. (2015). *Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry*. Nature. 419, (6906), 537-542

29. Ishihama, Y.; Oda, Y.;Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. (2005) *Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein.* Molecular & Cellular Proteomics.4, (9), 1265-1272.

4.6 Summary chapter 4:

Identifying Possible Allergen-Candidates in Crustacean and Molluscs

- ✔ 22 shellfish species were analysed by mass spectrometry (LC/qTOF), utilising raw and whole heated protein extracts
- ✔ 32 different proteins were identified, mainly being binding proteins or catalytic active proteins
- ✔ The main protein identified in the crustacean raw extracts is arginine kinase, and in the whole heated extracts it is tropomyosin
- ✔ The main protein identified in the molluscs raw extracts is actin, and in the whole heated extracts it is actin and tropomyosin
- ✔ 21 identified proteins are possible shellfish allergens, whereas six have not previously been reported
- ✔ 11 reported shellfish allergens and two possible novel allergens have been identified in crustacean
- ✔ 12 reported shellfish allergens and five possible novel allergens have been identified in molluscs

In the previous chapters tropomyosin was identified by LC/qTOF and in this chapter all known and possible novel allergen were identified. In chapter 5 the amount of tropomyosin will be quantified using the same 22 shellfish species and the identified signature peptide from chapter 3. Therefore in chapter 5 a novel quantitative LC/MRM method will be established and validated for the absolute quantification of tropomyosin, confirming the low concentrations of TM found in the raw shellfish extracts in this chapter.

CHAPTER 5

QUANTIFICATION OF ALLERGENIC TROPOMYOSIN APPLYING SIGNATURE PEPTIDES

5.1 Introduction

Mass spectrometry (MS) has recently been applied as an alternative method for food allergen identification, detection and quantification. Presently 46 different food allergens have been investigated using different MS systems, however, only 28 have been quantified.[1] Furthermore, only 19 have been quantified by liquid chromatography (LC) and mulitiple reaction monitoring (MRM) for the sensitive quantification in nine publications.[1] Overall, the development as well as the validation for the LC/MRM methods for absolute allergen quantification was only explained in detail in four studies.[2-5] Moreover, three studies used isotopic labelled peptides for the validation of the LC/MRM methods,[2, 4, 5] whereas Sealey-Voyksner et al.[3] considered using isotopic labelled peptides but came to the conclusion that they were not necessary.

Shellfish allergens are, using current methods, detected and quantified based on tropomyosin (TM), the major heat stable allergen in shellfish.[6-10] Species belonging to the shellfish allergen group include the group of crustacean and molluscs. Presently, the legislation in the European Union and Canada are demanding a different food allergen label for crustacean and molluscs,[1] however, current available methods cannot distinguish between those two groups. Mass spectrometry has been successfully applied to identify species specific signature peptides obtained from TM for Black Tiger prawn,[11] Northern prawn[12] and Snow crab.[2, 5, 13] One LC/MRM method was successfully developed and validated using the identified signature peptide for Snow crab, using the chemically synthesised analytical and isotopic labelled signature peptide.[2, 5]

Nevertheless, from the perspective of the legislation and the food industry, the differentiation of crustacean and molluscs is currently more important than species specific identification and quantification of TM. Therefore in chapter 3 four signature peptides were identified originated from TM to distinguish crustacean and molluscs. These four peptides will be chemically synthesised as analytical mass spectrometry standards and isotopic labelled to be used as internal standards for the development and validation of an absolute quantification method by LC/MRM in this chapter. Applying these four peptides, TM was quantified in different crustacean and mollusc species with a minimum of one peptide per species and a maximum of three peptides per species.

5.1.1 Aims

The aims of this chapter are:

- Developing a liquid chromatography (LC) method coupled with multiple reaction monitoring (MRM) using four signature peptides derived from tropomyosin, identified in the previous chapter to distinguish crustacean species and mollusc species
- Validating the developed LC/MRM method for the four selected peptides to distinguish crustacean and molluscs
- Analysing 22 shellfish species by the validated LC/MRM method using raw and whole heated extracts
- Confirming that the four selected peptides can distinguish between crustacean and molluscs with the predicted selection of peptides

5.2 Materials and methods

5.2.1 Peptide synthesis

The peptides which have been identified in chapter 3 were chemically synthesised by Sigma-Aldrich, USA as analytical mass spectrometry (MS) standards and internal standards in the unlabelled and the isotopic labelled form (Table 5.1). The isotopic label was produced using Leu-OH-d_{10}. The purity of the peptides were <98%, as specified by the manufacture. Table 5.1 summaries the abbreviation of the peptides with the selection criteria for distinguishing the shellfish subgroups. Moreover, the sequences of the four peptides with the position of the isotopic label and their monoisotopic mass are given.

Table 5.1: Summary of the chemically synthesised peptides, including the lot number as named by the manufacture, the peptide specificity, the amino acid sequence, the monoisotopic mass and the position of the isotopic label for the internal standards (IS).

Peptide	Lot #	Peptide specificity	Label position	Amino acid sequence	Mass
Peptide 1	89811	Unique to crustacean		LAEASQAADESER	1376.8
IS 1	89810		B1 = Y13	**L**AEASQAADESER	1386.69
Peptide 2	89813	Include crustacean, cephalopods, but not bivalves		IQLLEEDLER	1257.4
IS 2	89812		B3 = Y7	IQ**L**LEEDLER	1267.73
Peptide 3	89815	Unique to crustacean		LAITEVDLER	1158.3
IS 3	89814		B8 = Y3	**L**AITEVDLER	1168.69
Peptide 4	59817	Include prawn and lobster, but not crabs and krill		ALSNAEGEVAALNR	1414.9
IS 4	89816		B12 = Y3	ALSNAEGEVAA**L**NR	1424.79

The peptides were dissolved in water at a concentration of 5 mg/ml (stock concentration) and the working concentration (WC) was 100 micromolar (µM) in 0.1% formic acid (FA). The stock concentration and the working concentration were stored at -80° C until further use.

5.2.2 MRM preliminary settings

The equipment used in this chapter is liquid chromatography (LC) coupled with a triple quadruple (MRM) mass spectrometer. The LC/MRM system is an Acquity ultra performance liquid chromatography (UPLC) coupled with an electrospray ionisation (ESI) interface to a Xevo TQ mass spectrometer (Waters Corporation, USA). Data processing was performed using Mass Lynx 4.1 software (Waters Corporation, USA).

The peptides were firstly analysed via direct infusion in a concentration of 500 nanomolar (nM) in 20:80 acetonitrile (ACN): water (H_2O) + 0.1% FA. As preliminary settings the four highest transitions were chosen, to investigate if the intensity changes with different collision energies and cone voltages applied. The cone voltage was evaluated from 10-50 eV and the collision energy applied ranged from 0-100 eV. Four transitions were monitored for the LC method development. Moreover, the optimisation of the desolvation temperature in the range of 200-600°C was assessed, using the 15 minutes LC method A from chapter 3 as preliminary LC method.

5.2.3 LC preliminary settings

To test the separation of the peptides according to the stationary phase, different LC columns were tested. The columns have all been designed by the manufacture (Waters Corporation, Australia) to separated peptides and/or small molecules. The columns tested were (1) ACQUITY UPLC BEH C8 1.7 µm; 2.1 x 100 mm (designed for small molecule separation); (2) ACCQTAG ULTRA 2.5 µm; 2.1 x 100 mm (designed for amino acid separation); (3) ACQUITY UPLC BEH300 C4 1.7 µm; 2.1 x 50 mm (designed for intact protein separation); (4) ACQUITY UPLC BEH300 C18 1.7 µm; 1.0 x 100 mm (designed for peptide separation); (5) ACQUITY UPLC BEH130 C18 1.7 µm; 2.1 x 100 mm (designed for peptide separation); (6) XSELECT CSH C18 2.5 µm; 2.1 x 50 mm Column XP (designed for small molecule separation).

The different stationary phases were evaluated with the unlabelled peptides and two different LC gradients. The gradient of method I was shallower, to ensure that peptides will bind to stationary phase, when utilising different materials. Method II

was the shorter version of the 15 minutes LC method A (Chapter 3). In detail, for the method I the gradient was: 0-7 min 5-35%B; 7-7.5 min 35-70%B; 7.5-7.9 min 70-70%B; 7.9-8 min 70- 5%B; 8-10 min 5%B; for the method II the gradient was: 0-7 min 5-50%B; 7-7.5 min 50-90%B; 7.5-7.9 min 90%B; 7.9-8 min 90-5%B; 8-10 min 5%B. The mobile phase A was 0.1% formic acid in water (Thermo Fisher Scientific, Australia) and the mobile phase B was 0.1% formic acid in acetonitrile (Thermo Fisher Scientific, Australia), with a flow rate of 0.4 ml/min. The injection volume was 5 µl, the column oven temperature was set at 45°C and autosampler temperature at 10°C. The concentration of the peptides for this experiment was 50 nM.

To further improve the separation of peptides, the concentration of the mobile phase B was assessed using a linear gradient for 7 minutes from 5-25%B to 5-40%B. Moreover, the flow rate was evaluated in a range of 0.3–0.5 ml/min and the column oven temperature was investigated between 25-50°C. The concentration of the peptides for this experiment was 500 nM.

5.2.4 LC/MRM final method

The following settings and parameters were chosen for the absolute quantification of TM applying the four peptides and the four internal standards (IS). The mobile phase A was 0.1% formic acid in water and the mobile phase B was 0.1% formic acid in ACN. The separation of the peptides was performed on the ACQUITY UPLC BEH 300, C18 column (1.7 µm 2.1 x 100 mm (Waters Corporation, Australia)) with a gradient of: 0-7 min 5-25%B; 7.0-7.5 min 50-90%B; 7.5-8.7 min 90%B; 8.7-8.8 min 90-5%B; 8.8-10 min 5%B. To protect the mass spectrometer from contamination the first 1.5 minutes and the last 2.5 minutes from the LC were diverted into waste. The LC was run at a flow rate of 0.4 ml/min and the column temperature was 40°C. The injection volume was 5 µl of sample using the full loop mode. The final tune page settings were: Capillary voltage 1.5 eV; Cone voltage 28 eV; Desolvation temperature 500°C; Desolvation gas flow 1000 l/h; Cone gas flow 50 l/h; Collision gas flow 0.4 l/h; Collision energy 2 eV and Source Temperature 150°C, respectively. Table 5.2 summarises all settings chosen for the final LC/MRM method, including the precursor and the transitions ions with the

variable cone voltage and the collision energy applied. Moreover, the LC retention times and the set MRM windows are reported.

Table 5.2: Final settings for the developed LC/MRM method, including the precursor and the transitions information for the peptides 1-4 and the IS 1-4, the LC retention times and the MRM windows.

Peptide	$[M + 2H]^{2+}$	Transition	Ion	Collision energy (eV)	Cone voltage (eV)	Fragment intensity	Target ion ratio	MRM time window (min)	LC-MRM retention time (min)
Peptide 1	689.0	777.3	Y7	25	28	Quantification fragment		0-3	2.17
		706.3	Y6	24	28	Medium fragment	1:1.2		
		314.1	B3	25	28	Lowest Fragment	1:2		
IS 1	694.0	520.2	Y4	25	30	Medium fragment	1:2	0-3	2.16
		324.2	B3	35	30	Lowest Fragment	1:2.5		
		195.1	B2	31	30	Quantification fragment			
Peptide 2	629.5	1016.5	Y8	21	28	Medium fragment	1:1.5	6-10	6.57
		355.2	B3	20	28	Lowest Fragment	1:3		
		242.1	B2	22	28	Quantification fragment			
IS 2	634.5	1026.5	Y8	21	28	Medium fragment	1:2	6-10	6.54
		365.2	B3	20	28	Lowest Fragment	1:3		
		242.1	B2	22	28	Quantification fragment			
Peptide 3	580.0	861.4	Y7	20	24	Quantification fragment		5.6-6.8	6.26
		417.2	Y3	29	24	Medium fragment	1:1.4		
		185.1	B2	19	24	Lowest fragment	1:1.7		
IS 3	584.9	871.4	Y7	22	24	Quantification fragment		5.6-6.8	6.22
		770.4	Y6	22	24	Lowest fragment	1:2		
		185.1	B2	23	24	Medium fragment	1:1.2		
Peptide 4	708.3	958.5	Y9	26	28	Lowest fragment	1:1.7	3-5.6	4.93
		829.5	Y8	28	28	Quantification fragment			
		544.3	Y5	23	28	Medium fragment	1:1.3		
IS 4	713.3	968.5	Y9	25	28	Lowest fragment	1:1.7	3-5.6	4.91
		839.5	Y8	27	28	Quantification fragment			
		554.3	Y5	23	28	Medium fragment	1:1.4		

However, it was observed during the LC/MRM method optimisation that the peptide 2 and the peptide 3 are “sticky”. Therefore another LC/MRM method was developed to avoid carry-over of peptides between the samples and the injections. In detail, the LC method gradient is 0-30 min 50%B; 30-35 min 5%B, utilising the same LC settings and mobile phases. The MRM precursor and the transitions were monitored for all the peptides and the internal standards over 35 minutes. This LC/MRM program was applied after each standard curve and quality controls (QC), furthermore, six different samples were measured.

5.2.5 LC/MRM method validation

The method validation was followed according to the Department of Health Human Services Food and Drug Administration using the Guidance for Industry for Bioanalytical Method Validation,[14] (BMV) as suggested by Abdel Rahman.[2, 5]

All standards, quality controls (QC) and samples were spiked using the four isotopic labelled peptides as internal standards (IS 1-4) in a concentration of 200 nM, equalling the concentration of the quality control mid (QC-M). Each MS data point given in the calibration curves and the sample analysis represent triplicate analysis by LC/MRM. The points are measured with the peak area-ratio of the selected fragment ion (quantification ion) of both, the unlabelled and the labelled peptides. Therefore all the final results for the measured samples were corrected by the internal standards for each peptide and the calculations and the results are based on the area-ratio, as shown below:

$$Area - ratio = \frac{Area\ of\ unlablled\ peptide}{Area\ of\ lablled\ peptide}$$

5.2.5.1 Linearity, precision and accuracy

The linearity was measured generating the independent calibration curves on three consecutive days, using ten different concentrations of the standard solution (Table 5.3) along with blank sample (blank processed without the internal standards (IS)) and zero blank samples (blank processed with the IS 1-4). Every

standard curve and QC was prepared freshly on the day of use, diluting the working concentration (WC) as shown in table 5.3. The standard curve was measured in the range from 0.5 nM – 1000 nM. The regression analysis was used to evaluate the linearity of the method and calculated by the least squares regression (R^2).

The precision was calculated using three quality controls (QC), with 20 nM for the quality control low (QC-L), 200 nM for the quality control mid (QC-M) and 800 nM for the quality control high (QC-H). The repeatability (intra-day), the intermediate precision (inter-day) and the reproducibility (mean of the inter-project trails) were calculated and evaluated for the calibration standard curves and the QCs.

Table 5.3: Workflow of the dilution series for the calibration standard curve and the quality controls measured, summarising the final concentrations applied for the method validation and the sample quantification.

Dilution series		
Final peptide concentration	**Work concentrations**	**Work flow of dilution series**
1 µM (1000 nM)	WC2	10 µl (WC) + 990 µl of 0.1% FA
0.5 µM (500 nM)		5 µl (WC) + 995 µl of 0.1% FA
100 nM		100 µl (WC 2) + 900 µl of 0.1% FA
50 nM		50 µl (WC2) + 950 µl of 0.1% FA
25 nM		25 µl (WC2) + 975 µl of 0.1% FA
10 nM		10 µl (WC2) + 990 µl of 0.1% FA
5 nM		5 µl (WC2) + 995 µl of 0.1% FA
2.5 nM		2.5 µl (WC2) + 997.5 µl of 0.1% FA
1 nM		5 µl (WC2) +4995 µl of 0.1% FA
0.5 nM		2.5 µl (WC2) + 4997.5 µl of 0.1% FA
800 nM	QC-H	8 µl (WC) + 992 µl of 0.1% FA
200 nM	QC-M	200 µl (WC2) + 800 µl of 0.1% FA
20 nM	QC-L	20 µl (WC2) + 980 µl of 0.1% FA

The lower limit of detection (LLOD) was determined by calculating the standard deviation (SD) of the response and the slope of the regression equation. The limit of detection (LOD) and the limit of quantification (LOQ) were calculated using 3.3 σ/k and 10 σ/k, respectively, where σ is the SD of the intercept (d) and k is the slope of the curve.

Overall, the method acceptance criteria of the BMV guidelines state that the precision of the calibration curve and the QC samples are considered to be acceptable if the coefficient of variation (CV) is ≤20% for the intra-day and the inter-day precision. Furthermore, the accuracy compared with the nominal value needs to be within ±20% and the precision at the LOQ to be ±20%. Finally, the calibration curves must meet the above criteria and have a correlation coefficient (R^2) of at least 0.99.

5.2.5.2 Selectivity and recovery

To confirm the selectivity of chosen peptides and transition, digested blank matrices (PBS, vegetable soup, chicken soup and chicken muscle) were analysed by the LC/MRM method. Three replicas were made and injected in triplicate and the selectivity was based on the observed area of the total ion chromatogram (TIC) for the peptides 1-4 and the IS 1-4.

The chicken muscle extract has been generated similar to the raw shellfish extracts as described in chapter 4. Briefly, 17 g of chicken muscle (Local supermarket, Townsville, Australia) were homogenised in 100 ml PBS. For vegetable soup and chicken soup, stock soups (Campbell's Real Stock, Australia) were purchased from the local supermarket (Townsville, Australia) and 125 g were extracted in 100 ml PBS. All the total homogenised extracts were kept at 4°C overnight, while continuously shaking, followed by centrifugation and sterile filtration and finally stored at -80°C until further use.

For the recovery experiment the same blank matrices were (1) spiked with the peptides 1-4 and the IS 1-4 in a concentration of 200 nM (concentration of the QC-M), before digestion of the samples or (2) spiked with the peptides 1-4 in a concentration of the QC-M before digestion of the samples. Whereas the IS 1-4 (QC-M) were added after the digestion, prior to LC/MRM analysis. The experiment was performed three times (digestion), where each sample was injected three times. The recovery was calculated according to the area-ratio in percent as shown below:

$$Recovery\ (\%) = \frac{(Blank\ matrices + peptides\ 1{-}4\ ) + IS\ 1{-}4\ added\ before\ digestion\ (1)}{(Blank\ matices + peptides\ 1{-}4\ ) + IS\ 1{-}4\ added\ after\ digestion\ (2)} * 100$$

5.2.5.3 Stability

To confirm the stability of the peptides 1-4 and the IS 1-4, 200 nM (concentration of the QC-M) underwent different temperature treatments (1) samples were stored at -80°C as reference sample for 10 days (2) samples were treated for two hours at 60°C (3) samples were treated for two hours at 37°C (4) samples were treated for two hours at room temperature (5) samples were treated for two days at 4°C (6) samples were treated for one week at -20°C and (7) samples underwent three freeze thaw cycles, where each freeze cycle was 24 hours (-80°C). After the treatments all samples were stored at -80°C until the LC/MRM analysis. The stability experiment was performed in triplicates, where each triplicate was injected three times. The results were calculated as shown below:

$$Stability\ (\%) = \frac{Treated\ sample}{Untreated\ sample\ (-80°C)} * 100$$

5.2.6 Shellfish species protein extracts

Twenty-two different shellfish species were analysed using raw and whole heated protein extracts. The information of the generated extracts and the shellfish species are explained in detail in chapter 3 (Table 3.10) and chapter 4 (Table 4.1). Moreover, the recombinant TM from King prawn (KP) was expressed using auto-induction as explained in detail in chapter 2 and published in Koeberl et al.[15] Tropomyosin was purified from KP raw extract, using the same strong anion-exchange purification method as described in chapter 2.

Each digested protein extract was spiked with the IS 1-4 in a concentration of 200 nM after digestion, prior to the LC/MRM analysis. The standard curve was prepared fresh every day and measured prior the quantity control (QC) samples, followed by the digested shellfish extracts. Between the QC samples and the shellfish species analysis the LC/MRM washing method was performed. The QC

samples were reinjected after six different species were analysed, to confirm the reproducibility of analytical setting.

The digestion method was already investigated in chapter 3 using purified TM from KP raw extract and whole heated extract. Therefore the digestion method was not investigated in this chapter. Briefly, 200 µl of soluble protein extract was reduced and alkylated, followed by digestion using trypsin spin columns for 15 minutes at room temperature. The tryptic digested peptides were eluted twice with 100 µl of 0.1% formic acid followed by the LC/MRM analysis. Each sample was injected in triplicate and the concentrations of the samples were calculated based on the freshly prepared standard curve using the area-ratio values, whereas each standard curve and QC met the method acceptance criteria of the BMV guidelines. The accuracy and the precision were calculated based on the standard deviation (SD) of the triplicate (injection) and was considered acceptable if the SD was ±20%.

When the concentration of the sample was outside the standard curve the samples were diluted after the digestion, to fall in the linear range of the standard curve. For the final results shown in section 5.3.4 the dilution factor was recalculated.

5.3 Results

The aim of this chapter is to develop and validate a LC/MRM method for the absolute quantification of allergenic tropomyosin from shellfish using the four identified peptides from chapter 3. Therefore these four peptides were chemically synthesised as analytical mass spectrometry standards and internal standards (IS) in their unlabelled and isotopic labelled form. Using the chemical synthesised standard solutions the MRM mode was optimised and the precursor and the fragment ions were selected. Moreover, the LC method was optimised for the four selected peptides. The development, the optimisation and the validation will be explained in the following sections.

5.3.1 Preliminary LC/MRM results

In order to develop a quantitative analytical procedure, the LC/MRM was optimised for all four peptides, including the IS 1-4. The precursor and the fragment ions for all the chemical synthesised peptides were firstly analysed via direct infusion in a concentration of 500 nM. All the peptides were found to be double charged precursor ions $[M+2H]^{2+}$ and the product ions were only chosen based on being b-ions or y-ions. However, less specific b 1-ions and y 1-ions were excluded. To achieve the best fragmentation with the most intense fragment ions, the cone voltage and the collision energy applied ranged from 10-50 eV and 0-100 eV, respectively. After optimising the cone voltage and the collision energy for each peptide the highest four transitions were chosen to be able to monitor changes in the fragment patterns when changing other LC/MRM parameters. As next step the desolvation temperature was optimised, using the 15 minutes LC method A (Chapter 3), as preliminary LC method. Figure 5.1 shows the response for the peptides at different desolvation temperatures, ranging from 200-600°C. Although the peptide 4 and the peptide 3 have higher response at 600°C, figure 5.1 clearly shows that there is a decrease for peptide 2 at 600°C. Therefore the desolvation temperature was set at 500°C for all further experiments, thus, resulting in the overall best response for all four peptides.

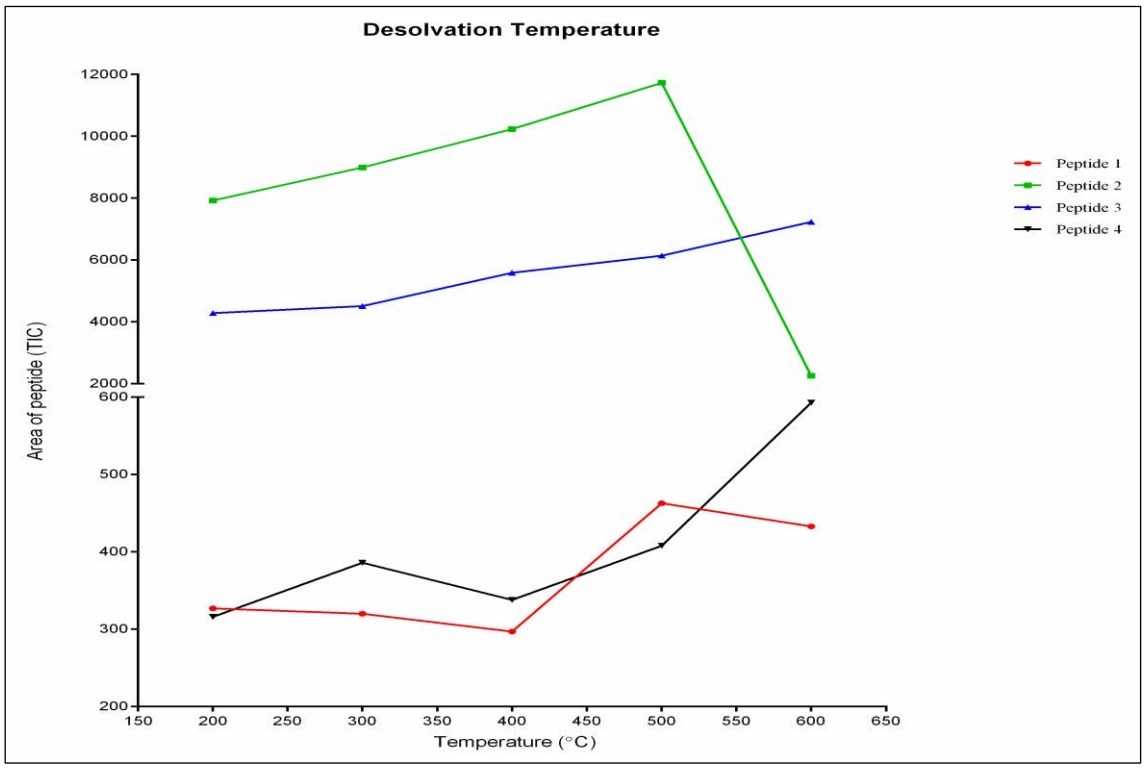


Figure 5.1: Desolvation temperature evaluation in the range of 200-600°C, applying the four chemical synthesised peptides.

To evaluate the best achievable separation of the peptides six different stationary materials have been investigated. Generally, as displayed in figure 5.2, the peptides generated a higher response (area based on the total ion chromatogram (TIC)) for column material designed for peptide separation. In contrast, the peptides were not detected with column 3 (ACQUITY UPLC BEH300 C4), however, this material is designed for protein separation. Column 1 (ACQUITY UPLC BEH C8) performed similar to column 2 (ACCQTAG ULTRA), with the method I giving slightly higher response. Column 5 (ACQUITY UPLC BEH130 C18) and column 6 (XSELECT CSH C18) seem to perform similar, however, only either LC method gave a good response for the peptides. Overall, column 4 (ACQUITY UPLC BEH300 C18) applying the method II showed the best overall response for the four peptides. Therefore this column was selected for all further analysis. Moreover, this column material was also used for the LC/qTOF analysis (Chapter 3 and chapter 4). Furthermore, comparing method I with method II, it was observed that the higher ACN concentration in the gradient reduced the signal in

the blank injections and carry-overs caused by the peptide 2, the IS 2, the peptide 3 and the IS 3.

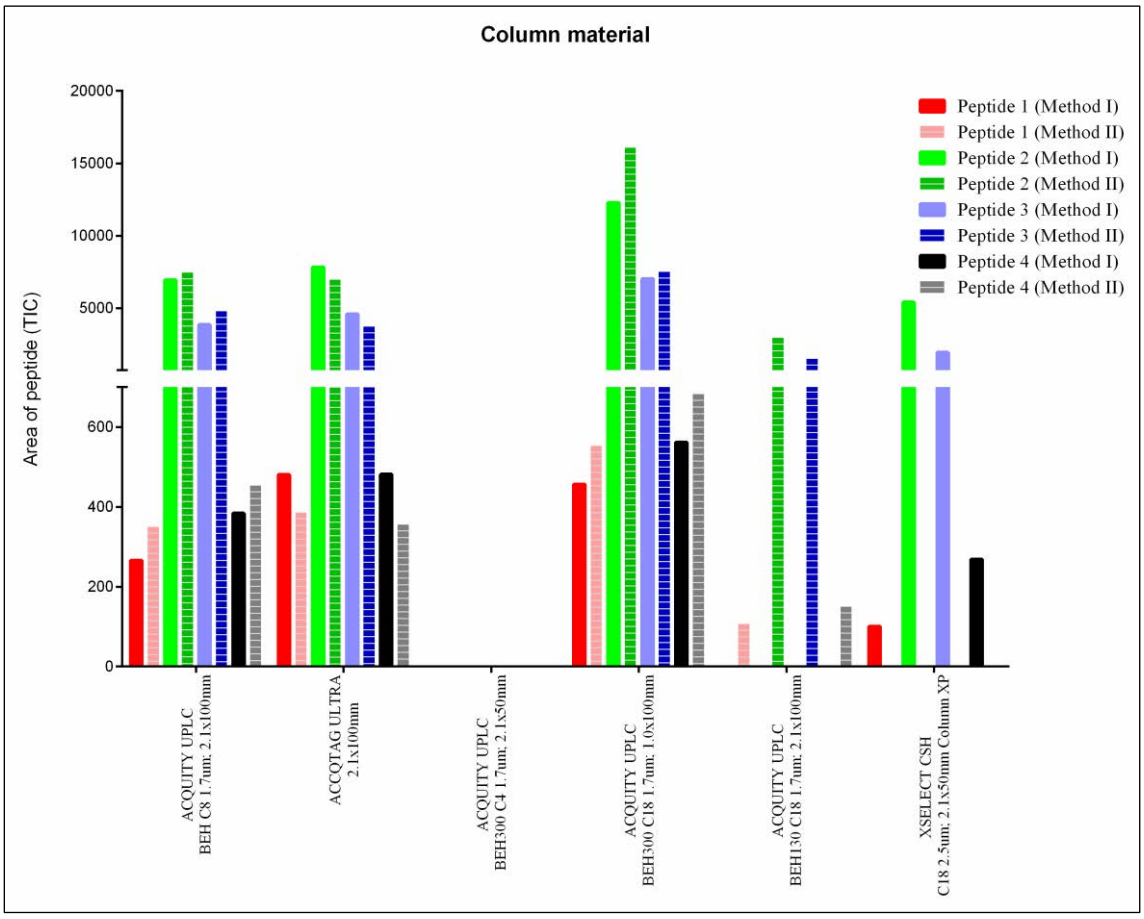


Figure 5.2: Six different stationary materials were investigated with two different LC methods utilising the four chemical synthesised peptides.

The column selected for all further experiment was the column 4 (ACQUITY UPLC BEH300 C18), as shown in figure 5.2. To optimise the separation of the chemical synthesised peptides and the IS 1-4, shallower LC gradients were evaluated (Figure 5.3), ranging from 20-40% ACN as final concentration. It was observed that all four peptides and the IS 1-4 eluted with less than 25% ACN in the mobile phase. Therefore the ACN concentration had to be at least 25%, however, higher ACN percentage lead to steeper LC gradient and the separation of the peptides was lower. Hence, a shallow linear ACN gradient with 25% ACN as maximum concentration was chosen as the final LC method, leading to the best achievable separation for all the peptides. However, the ACN percentage was still risen to

90% after the linear elution gradient, to avoid “stickiness” of the peptides, as observed in the comparison between method I and method II (see above).

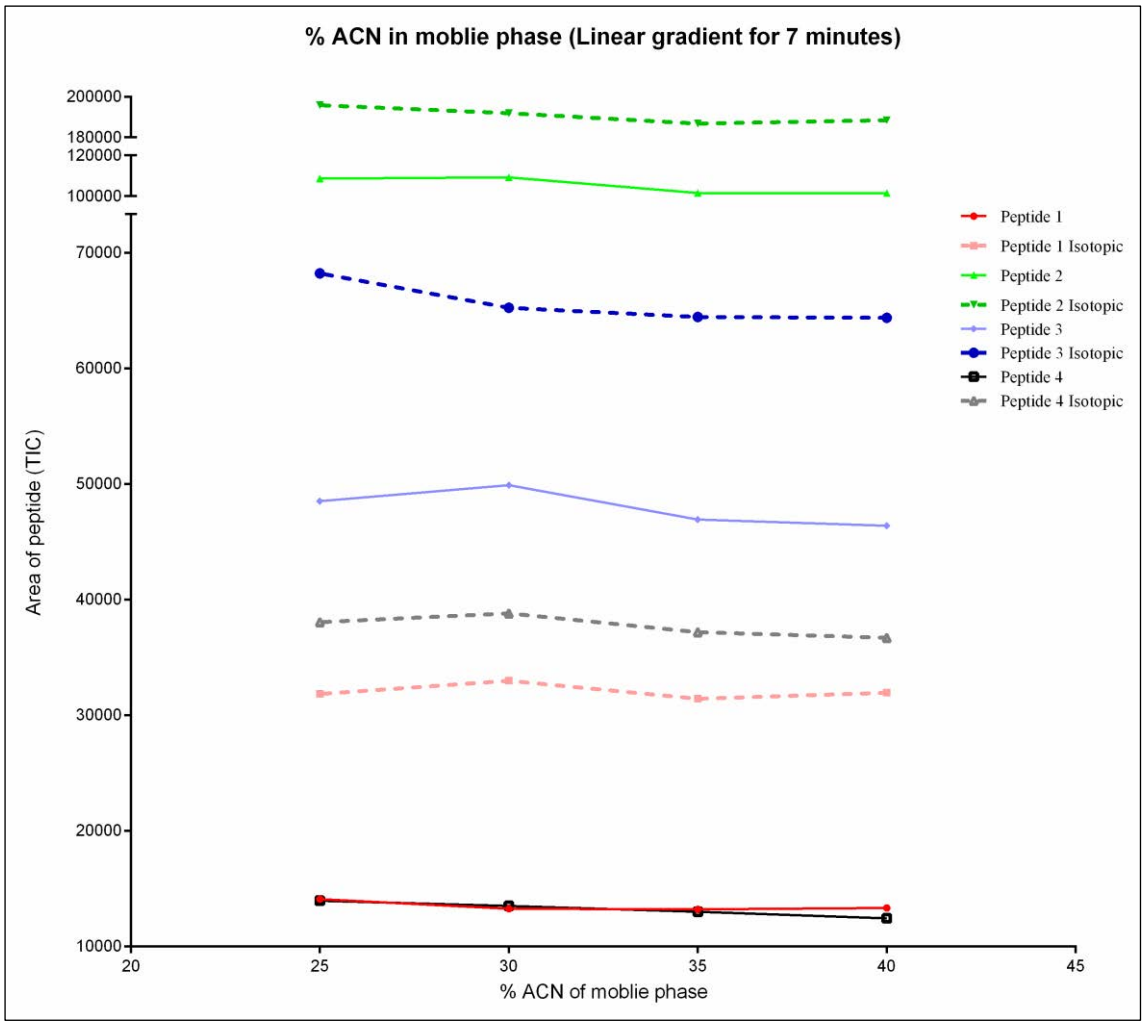


Figure 5.3: Different percentages of acetonitrile in the linear elution gradient were investigated with a maximum of 20-40% ACN, for the chemical synthesised peptides and the internal standards (Isotopic peptides).

Furthermore, different LC flow rates were applied to investigate if the separation of the four peptides and the IS 1-4 can be increased. The results in figure 5.4 visualise that the response is higher when the flow rate is lower. This can be explained with the lower dilution factor of the sample. However, it was observed that the separation of the peptides with the lower flow rate could not be improved. Moreover, the peak-shape of the peptides decreased and double-peaks occurred.

Applying higher flow rates lead to lower separation compared to 0.4 ml/min, thus, 0.4 ml/min was chosen as final flow rate.

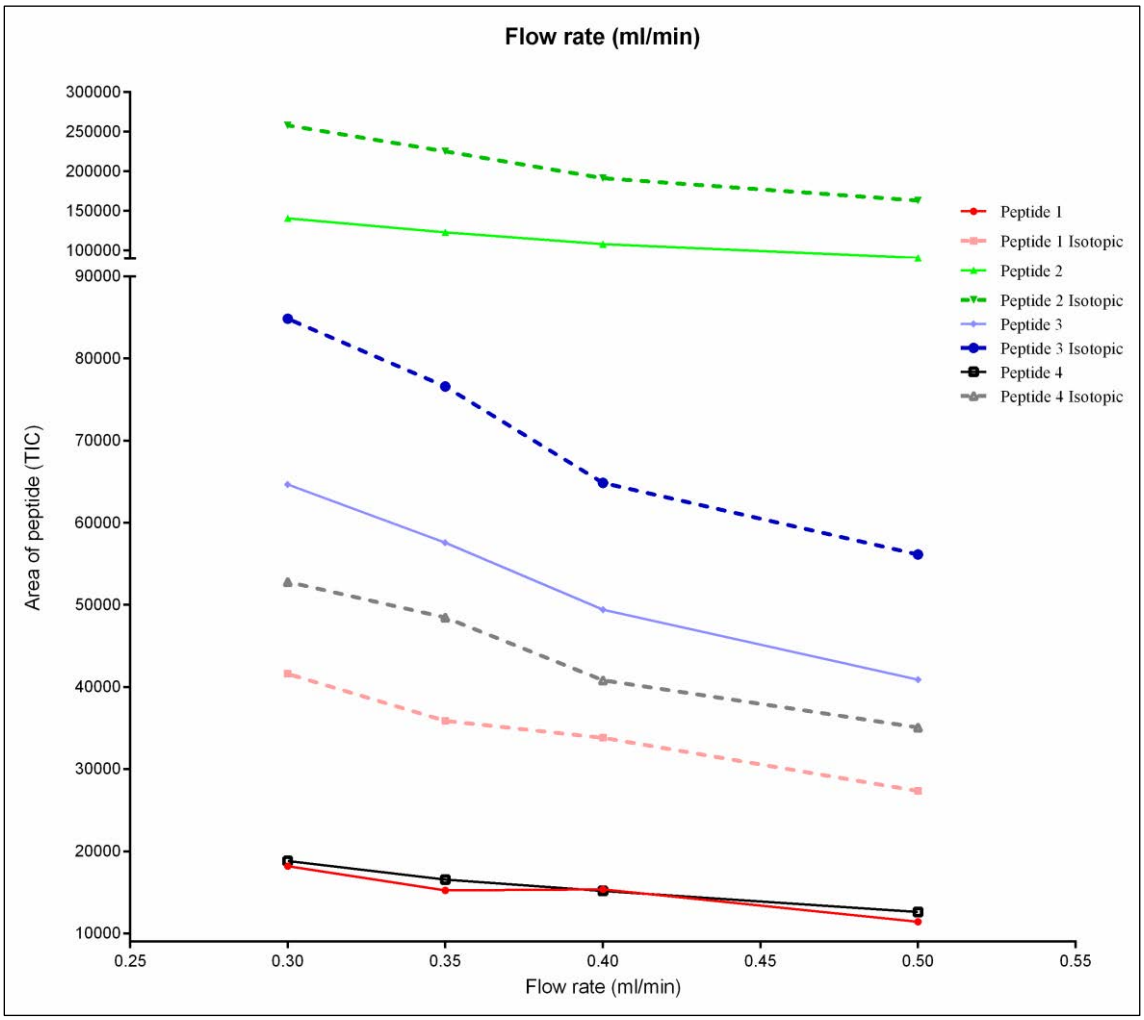


Figure 5.4: The flow rate of the LC method was investigated ranging from 0.3-0.5 ml/min for the chemical synthesised peptides and the internal standards (Isotopic peptides).

As a final step in the LC/MRM development, different column oven temperatures were evaluated. Figure 5.5 summarises the response of all three final transitions chosen per peptide and per IS and different column temperatures applied. Overall, the transitions of peptides respond differently to temperature, for some transitions the impact of different temperature investigated is quite obvious. Most of the transitions showed increased response with increased column oven temperature. However, the signal was dropping for all the transitions at 50°C, with the exception

of the quantification transition of the IS 3. In summary, the overall best column temperature was 40°C and therefore will be used for all further experiments.

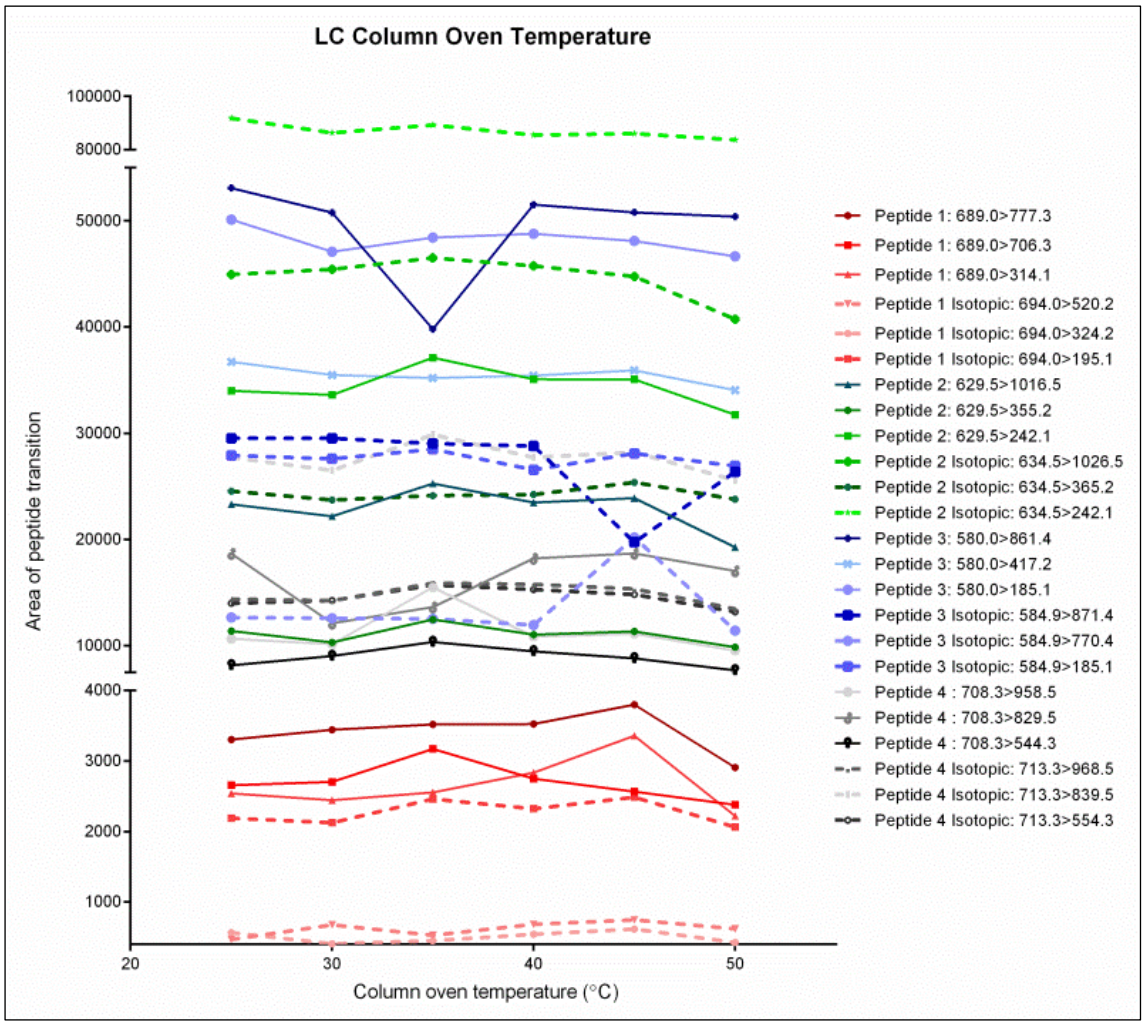


Figure 5.5: The column oven temperature was investigated in a range of 25-50°C for all the chemical synthesised peptides and the isotopic peptides (IS 1-4), displaying all the final three transitions chosen per peptide.

5.3.2 Final LC/MRM method

After the development and optimisation of the LC/MRM method (Section 5.3.1) the final LC/MRM method was chosen as summarised in table 5.2. Moreover, the LC retention time was observed to be stable over time resulting in good peak-shape and low noise for all the transitions and all the chemical synthesised peptides. The peptide precursor and the MRM transitions chosen in combination with the

associated LC retention time assure the specificity and the unambiguous identification of each target peptide and the IS 1-4.

The developed and optimised LC/MRM method was used for the validation of the method (Section 5.3.3), followed by the analysis of the 22 different shellfish species using the raw and the whole cooked extracts (Section 5.3.4). The chemical unlabelled synthesised peptides were used for generating standard curves and quantification, whereas the chemical synthesised isotopic labelled peptides were used as internal standards (IS 1-4) to calculate the peak area-ratio and confirm the accuracy and the precision of the method for the absolute quantification of tropomyosin from the different subgroups of shellfish.

5.3.3 Method validation

The method validation was based on the Guidance for Industry for Bioanalytical Method Validation[14] (BMV) for linearity, accuracy, precision and sensitivity, as explained in detail in section 5.2.5. With the optimised LC/MRM conditions the peptide eluted within 7 min as seen in figure 5.6. The four peptides can be easily separated with the developed LC method. However, the IS have very similar retention time with the unlabelled peptides, due to the high similarity of chemical properties between the labelled and the unlabelled peptides. Nevertheless, due to the isotopic label the mass difference between the peptides and the IS for the double charged precursor ion $[M + 2H]^{2+}$ is five (Figure 5.6 and table 5.2).

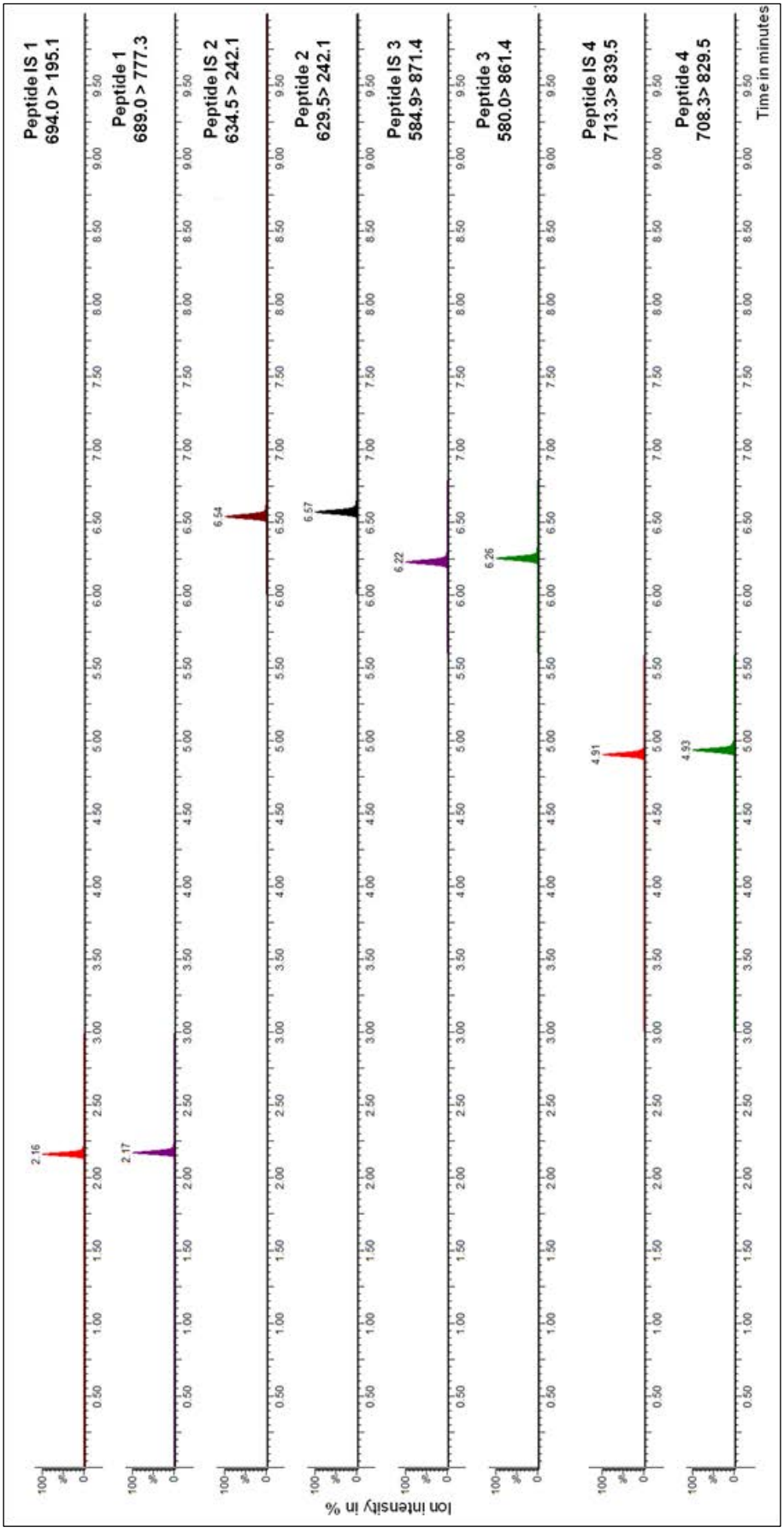


Figure 5.6: Representative LC/MRM chromatogram of the four tropomyosin peptides 1-4 and the internal standards 1-4 at a concentration of 100 nM. The quantification transition is shown for peptides 1-4 and the IS 1-4, including the LC retention time.

5.3.3.1 Linearity, precision and accuracy

The calibrations curve was established after injection of the standard solutions and the internal standards (IS 1-4) at ten different concentrations. The concentration of the ten different standards was recalculated using the equation of the regression curve. Figure 5.7 shows a representative calibration curves for the four peptides, and summarises that a good linear calibration curve was obtained for the peptides in the range of 0.5–1000 nM. The linearity of three different calibration curves was evaluated and the curve fit was calculated by least square regression (Figure 5.7). The linearity of the calibration curve met the criteria set out in the materials and method section, according to the Department of Health Human Services Food and Drug Administration using the Guidance for Industry for Bioanalytical Method Validation[14] (BMV).

For the individual calibration points the precision and the accuracy were calculated using the correlation coefficient, the d-intercept and the slope (k) of the regression line for the three consecutive standards curve generated are summarised in table 5.4. The estimated LLOQ for the peptides 2-4 is 0.5 nM, the LOD is 1 nM and the LOQ is 2.5 nM, whereas for peptide 1 the LLOQ and the LOD are 0.5 nM and the LOQ is 1 nM. The LOQ is the lowest calibration point and was 1 nM (peptide 1) and 2.5 nM (peptide 2-4), respectively, with a consistent precision and accuracy of ≤20% of the nominal value determined at a LOQ >10:1 signal-to-noise-ratio and a LOD >3:1 signal-to-noise-ratio, respectively. Therefore the precision, the accuracy, the LLOQ, the LOD and the LOQ met the criteria of the BMV guidelines.

Table 5.4: Summary of the linearity validation of the calibration curves (Figure 5.7). Average of three different curves over three days (n = 3).

	Peptide 1	Peptide 2	Peptide 3	Peptide 4
Range (nM)	1-1000	2.5-1000	2.5-1000	2.5-1000
Slope (k)	0.008533	1.08E-02	0.017767	0.006067
SD on slope (S_k)	0.000351	5.80E-05	0.000961	0.000153
Intercept (d)	-0.00287	-0.00307	-0.00563	-0.00683
Correlation coefficient (R^2)	1	0.99963	0.9998	0.999967

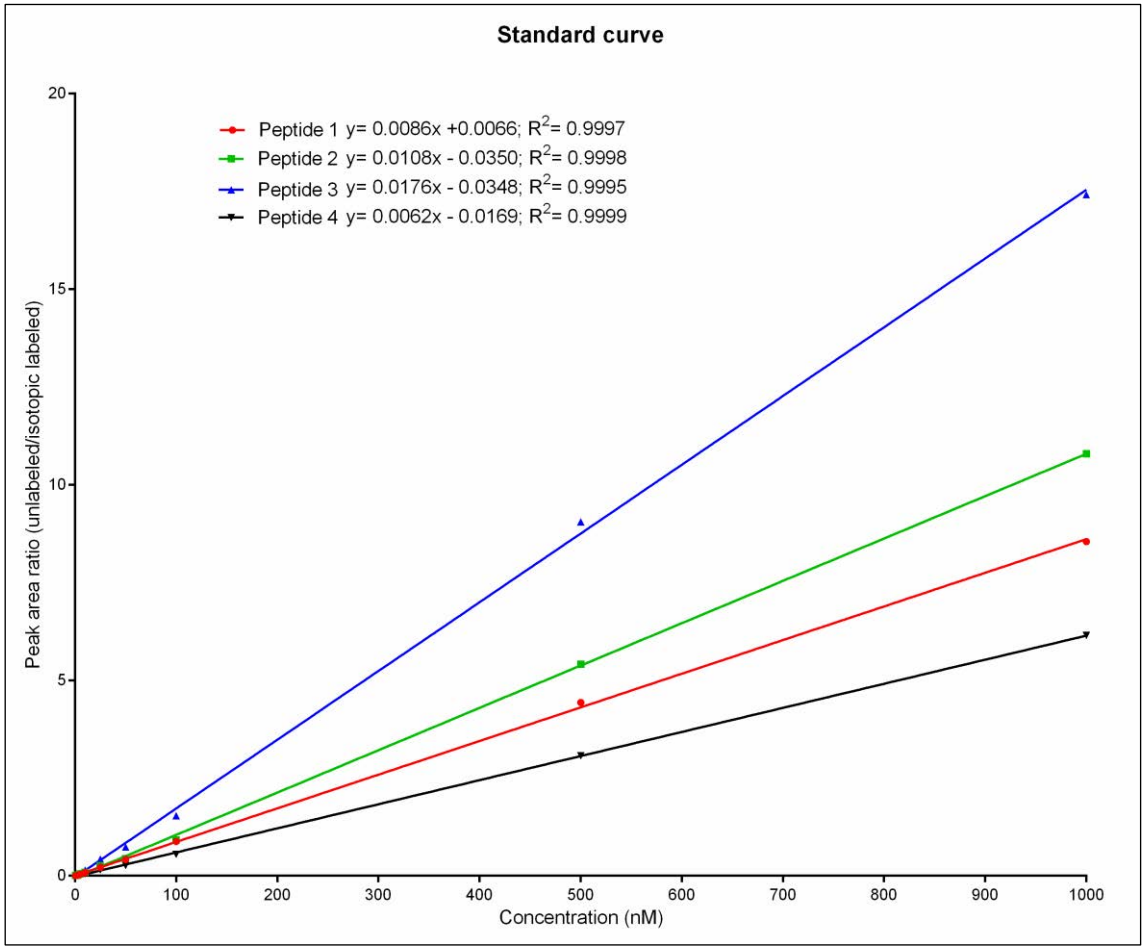


Figure 5.7: Calibration curve for the peptides concentration range of 0.5–1000 nM, where the response is the peak area-ratio of the unlabelled peptides to the isotopic labelled peptides.

Moreover, the accuracy and the precision were evaluated and established over three consecutive days using the QCs. At the intra-day interval, three replica of each QC were injected three times (n = 9 per QC), after running a valid calibration curve. As mentioned above, the concentration of the QCs was recalculated using the equation of the regression curve. The deviation of the mean from the nominal values served as the measure of method accuracy and the precision as summarised in table 5.5. In the inter-day validation study, a total of 27 samples were used to evaluate the method intermediate precision and the reproducibility by including nine of the intra-day samples with the other two days (nine samples each) of each QC concentration. Table 5.5 summarises the results, proving that all the investigated points for the LC/MRM validation meet the BMV guideline criteria.

Table 5.5: Summary of the inter-day and the intra-day validation for the method sensitivity, the precision and the accuracy based on the concentration calculated from the area-ratio, shown for all four peptides and applying three different quality controls.

		Peptide 1			Peptide 2			Peptide 3			Peptide 4		
Statistical parameters		QC-L	QC-M	QC-H	QC-L	QC-M	QC-H	QC-L	QC-M	QC-H	QC-L	QC-M	QC-H
Intra-day (1)	Mean (n=9)	19.31	192.54	776.58	18.99	168.60	742.80	18.24	162.89	702.41	18.43	172.61	750.28
	SD	0.95	4.36	11.64	0.78	2.75	5.58	0.65	2.23	8.06	1.35	6.13	12.10
	Precision CV%	3.43	3.73	2.93	5.04	15.70	7.34	8.79	18.56	12.20	7.83	13.70	6.21
	Accuracy %	96.57	96.27	97.07	94.96	84.30	92.66	91.21	81.44	87.80	92.17	86.30	93.79
Intra-day (2)	Mean (n=9)	20.03	209.04	798.28	16.19	189.44	783.00	17.38	190.40	813.23	16.60	178.53	758.48
	SD	1.13	6.47	10.69	0.44	1.80	8.00	1.45	3.07	8.82	1.84	5.84	18.40
	Precision CV%	0.13	4.52	0.21	19.06	5.28	2.12	13.10	4.80	1.65	17.00	10.73	5.19
	Accuracy %	100.13	104.52	99.79	80.94	94.72	97.88	86.90	95.20	101.65	83.00	89.27	94.81
Intra-day (3)	Mean (n=9)	20.57	191.18	726.62	18.50	193.03	827.67	18.89	197.92	871.69	19.97	161.89	700.89
	SD	1.45	4.21	8.36	0.27	3.07	15.09	1.06	3.73	8.73	2.50	6.26	24.60
	Precision CV%	2.84	4.41	9.17	7.52	3.48	3.46	5.54	1.04	8.96	0.15	19.06	12.39
	Accuracy %	102.84	95.59	90.83	92.48	96.52	103.46	94.46	98.96	108.96	99.85	80.94	87.61
Inter-day	Mean (n=27)	19.96	197.84	768.12	18.00	181.12	789.15	18.16	183.08	792.68	18.34	170.94	736.20
	SD	1.24	9.64	32.36	1.36	11.35	37.58	1.22	15.76	72.68	2.32	9.10	31.70
	Precision CV%	0.19	1.08	3.98	10.52	9.44	1.36	9.18	8.46	0.92	8.29	14.53	7.97
	Accuracy %	99.81	98.92	96.02	89.48	90.56	98.64	90.82	91.54	99.08	91.71	85.47	92.03

Overall, the linearity, the precision and the accuracy for the developed LC/MRM method meet all the specification for the method validation as recommended by the Department of Health Human Services Food and Drug Administration using the Guidance for Industry for Bioanalytical Method Validation.

5.3.3.2 Selectivity and recovery

To confirm the LC/MRM method is selective for the chosen peptides four different blank matrices were digested and analysed by the LC/MRM method. All peaks observed at the correct retention time and the transitions detected were integrated using the total ion chromatogram (TIC). Figure 5.8 summarises the areas for blank matrices and shows that the selected peptides are not detected, thus areas are within the noise and no signal was observed. Nevertheless, it can be seen that more complex matrices (chicken muscle) lead to higher noise, thus the integrated area is higher compared to soups or buffers investigated. However, as visualised in figure 5.8b, it can be seen that the lowest standard used for the calibration curves and representing the LLOQ show higher area compared to all blank matrices analysed.

A

Matrix	Peptide 1	IS 1	Peptide 2	IS 2	Peptide 3	IS 3	Peptide 4	IS 4
PBS	0.00	0.00	0.00	0.00	0.06	0.16	0.00	0.03
Vegetable soup	0.00	0.00	6.56	37.54	1.86	0.86	0.27	0.72
Chicken soup	0.83	0.00	24.32	40.75	3.11	4.31	0.71	0.97
Chicken muscle	1.14	0.00	30.34	14.87	3.20	1.58	0.43	1.52
Lowest standard (0.5 nM)	10.67	2.18	100.54	44.69	66.66	61.70	4.31	4.48

B

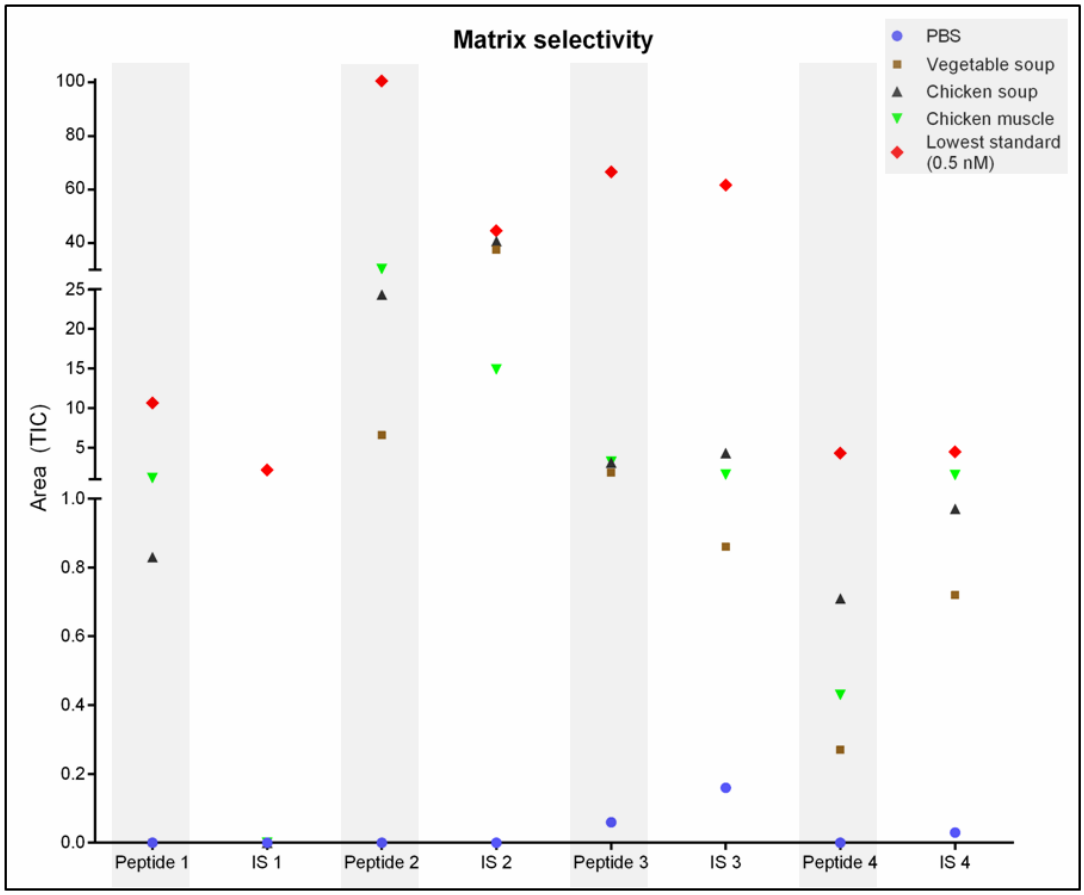


Figure 5.8: Summary of the area from the TIC found for four blank matrices and LLOQ for each of the peptides 1-4 and the IS 1-4, values shown in (A) table and (B) as figure for better visualisation. The digestion was performed in triplicate and each sample was injected three times (n=9).

To investigate if there is a loss of the peptides 1-4 and the IS 1-4 during digestion a recovery experiment was performed. The chemical synthesised peptides were blank digested in 0.1% FA and PBS at the concentration of the QC-M. The concentration of the blank digests was recalculated using the equation of the regression curve and the results were within ±20% of the nominal concentration. Moreover, four blank matrices were investigated and were spiked with the chemical synthesised peptides prior to digestion. The IS 1-4 were either added before or after the digestion and the recovery was calculated. As it can be seen in table 5.6, the recovery is high, being less than ±20%, with the exception of the peptide 2 in chicken muscle. Interestingly, the peptide 4 seems to have a lower recovery compared to the other peptides. Similar to the selectivity experiment, the recovery decreases slightly when matrices are more complex. However, the overall recovery for all the matrices investigated is high for all the four chemically synthesised peptides.

Table 5.6: Recovery values in percent for the spiked blank matrices for each chemically synthesised peptide. The digestion was performed in triplicate and each sample was injected three times (n =9).

Matrix	Peptide 1		Peptide 2		Peptide 3		Peptide 4	
	Recovery in %	SD in %	Recovery in %	SD in %	Recovery in %	SD in %	Recovery in %	SD in %
PBS	104.10	5.68	110.30	5.71	102.00	10.22	96.74	7.1
Vegetable Soup	107.60	15.62	99.23	16.97	103.80	9.88	102.50	10.06
Chicken Soup	119.40	8.8	119.50	18.27	97.65	18.2	95.35	7.54
Chicken Muscle	96.82	4.02	139.50	18.93	100.50	10.25	86.43	11.87

5.3.3.3 Stability

The stability of the chemical synthesised peptides and the IS 1-4 was investigated. The different temperature ranges and durations were chosen, thus, these will be applied during sample preparation for absolute TM quantification. Generally, all the peptides are quite stable during various treatments as summarised in table 5.7. The most stable peptide is peptide 1, followed by the IS 1, the peptide 3 and the IS 3. However, it was observed that the peptide 2, the IS 2, the peptide 4 and the IS 4

seem to be less stable for various treatments when compared to other peptides. Peptide 2, peptide 4, IS 2 and IS 4 were the least stable when heat treated, followed by the freeze and thaw cycles. Nevertheless, it is more important that peptide 1, the IS 1, the peptide 3 and the IS 3 are more stable, thus either of these two peptides will be detected in every shellfish species analysed. Hence, these two peptides are the ones being unique to crustacean or molluscs, respectively. The peptide 2, the IS 2, the peptide 4 and the IS 4 are additional peptides to subdivide the crustacean and the mollusc species further, representing confirming peptides.

Table 5.7: The stability of the chemical synthesised peptides and the IS 1-4 are given in percent for six different temperature treatments. The experiment was performed in triplicate and each sample was injected three times (n = 9). (RT= room temperature, FTC= freeze thaw cycle).

Temperature treatment	4°C		RT		37° C		60°C		-20°C		FTC	
Period of time	2 days		2 hours		2 hours		2 hours		8 days		3 cycles	
	Area in %	SD in %	Area in %	SD in %	Area in %	SD in %	Area in %	SD in %	Area in %	SD in %	Area in %	SD in %
Peptide 1	108.57	2.05	106.26	3.85	103.50	4.44	103.98	4.50	96.34	4.67	99.34	4.41
Peptide 2	107.91	4.20	98.08	8.46	66.54	15.64	63.76	11.77	83.35	8.92	68.80	13.97
Peptide 3	123.62	13.61	104.84	16.82	95.72	9.11	101.72	9.56	86.82	6.53	83.53	10.29
Peptide 4	95.19	5.77	90.38	5.22	78.32	5.54	78.06	4.81	85.75	6.20	80.04	12.06
IS 1	105.65	1.78	103.74	3.47	104.57	2.95	100.05	3.13	95.06	3.85	97.03	2.53
IS 2	107.06	3.88	99.51	8.30	68.89	13.49	64.87	11.54	87.52	8.08	70.38	13.78
IS 3	100.71	9.07	101.12	9.76	76.06	4.80	97.81	11.77	76.62	6.66	115.52	18.28
IS 4	97.89	2.29	90.27	5.11	79.06	6.87	80.37	4.52	87.76	6.32	81.33	10.36

Overall, the developed LC/MRM met all the criteria for the validation using the Guidance for Industry for Bioanalytical Method Validation published by the Department of Health Human Services Food and Drug Administration. Moreover, for the whole LC/MRM development and the validation experiments, it was confirmed that the chemical synthesised peptides and the isotopic labelled peptides show the same properties, such as LC retention time, matrix selectivity and stability. The validated LC/MRM method will be applied in the following section to quantify TM in 22 different shellfish species using the raw and the whole heated

protein extracts. Moreover, it will be observed if the predicted peptides are specific as suggested in chapter 3 and summarised in table 5.1.

5.3.4 Shellfish species analysed by LC/MRM

The four peptides that have been chosen in chapter 3 were chemically synthesised and applied to validate a quantitative LC/MRM method for the absolute quantification of allergenic tropomyosin from shellfish (Section 5.3.1-5.3.3). Briefly, the peptides chosen are unique for crustacean (peptide 1), unique for molluscs (peptide 3), unique for crustacean and cephalopods (peptide 2) and unique for prawns and lobsters, with some exceptions for crabs (peptide 4). To confirm the specificity of the predicted peptides the same 22 shellfish species will be analysed by the validated LC/MRM method, as analysed in chapter 3 and chapter 4 by LC/qTOF.

Table 5.8, table 5.9 and table 5.10 summarise the amount of TM quantified for the different shellfish species, showing all the detected and quantified peptides. The amount of peptides (nM/mg protein) varies for the four selected peptides. However, out of 46 shellfish extracts, for only ten extracts the amount between the different peptides quantified varies more than a factor of five. Therefore it is hypothesised that the minor variations between the quantified peptides are due to stability and sample processing, such as enzyme activity variation between the trypsin spin columns applied. Nevertheless, for major differences observed in the quantified peptides, one can assume this might be to incomplete digestion for some peptides in some protein extracts. Thus it is suggested that the digestion method applied should be analysed and potential optimised for complete protein digestion. However, out of ten extracts where the concentration of different peptides varies more than a factor of five, three belong to the species where one peptide was detected, although based on in silico data analysis (Chapter 3), the peptide should be absence. Potential causes for detection these peptides, including isoforms and undiscovered proteins, are explained in section 5.4.

Analysing the overall detection of the peptides according to their selection criteria explained in detail in chapter 3 and mentioned above, it can be seen in table 5.8 to

table 5.10, that peptide 1 was detected in every crustacean species. However, peptide 1 was also detected in octopus raw and whole heated extracts, as well as in Hybrid abalone whole heated extract. The peptide 2 was detected in every predicted species, except Banana prawn raw extract (Genbank accession number ADC55381.4). It is hypothesised, that the peptide 2 was not detected in Banana prawn due low concentration and incomplete digestion of tropomyosin. However, the peptide 2 was also detected in Hybrid abalone whole heated extract and Blue mussel raw and whole heated extracts. Peptide 3 was only detected in mollusc species and the peptide 4 was only detected in predicted species, with the exception of Hybrid abalone whole heated extract.

It is uncertain why peptide 1 was detected in raw and whole heated octopus extracts, however, the calculated concentration of peptide 1 in octopus extracts is significantly lower compared to the concentration of peptide 2 and peptide 3, which should be present in the octopus species (Genbank accession number for octopus is BAE54433.1). In contrast to Blue mussel raw and whole heated extracts where the peptide 2 was detected in a similar concentration to the peptide 3, although only the latter should be present according to the amino acid sequence, with the Genbank accession number AAA82259.1. In raw Hybrid abalone extract only peptide 3 was detected, however, in whole heated Hybrid abalone extract all four peptides were detected. The tropomyosin amino acid sequence of Hybrid abalone is unknown, therefore it is potential possible that peptides 1, 2 and 4 are present, however, based on known abalone sequences only the mollusc peptide (peptide 3) should be present.

The amino acid sequence of different tropomyosin proteins was analysed for possible other tryptic peptides that would match the masses of the precursor and the fragments of the peptides 1-4 and the IS 1-4, however, no tryptic peptides possible generated from TM would explain the false positive detected peptides in the four mentioned species. Therefore the search was expanded to all possible generated peptides or degradation products of TM, but nothing matched the precursor masses and the transitions selected. Moreover, the masses of the precursor and the transitions chosen for the peptides 1-4 and the IS 1-4 were searched with http://prospector.ucsf.edu to identity possible false negative peptides generated from other proteins, but it showed that the selection of the

masses of the precursor and the transitions for the peptides 1-4 and the IS 1-4 are highly specific for all the selected peptides. Overall, it is no explainable why for octopus, Blue mussel and the Hybrid abalone unpredicted peptides were detected and why peptide 2 was not detected for Banana prawn. Nevertheless, peptide 1 and peptide 3 are the two main peptides to distinguish the crustacean and the molluscs, therefore the overall tropomyosin concentration calculated in various species will be compared using peptide 1 and peptide 3.

Table 5.8 reports the quantified amount of the peptides in the raw shellfish species. Overall, the results of chapter 4 can be confirmed, the concentration of TM in the prawn raw extracts is lower compared to the crabs and the lobsters. Moreover, the TM concentration in Green Tiger prawn is higher compared to the other prawn species, thus, TM from Green Tiger prawn was the only TM identified in the raw prawn extracts in chapter 4. As explained above it is unsure why peptide 2 was not detected in Banana prawn, whereas the TM concentration of Vannamei prawn is lower compared to Banana prawn and the peptide 2 could be detected, although being below the quantification limit. However, TM in raw King prawn could still not be detected, it is still hypothesised that might be due really low concentration in raw King prawn extract, thus table 5.10 shows that the purified TM from raw King prawn extract can be quantified.

Confirming the data of chapter 4, it can be seen in table 5.8, that the TM concentration of the raw crabs is similar. The concentration of raw yabby extract is lower, followed by Rock lobster and Slipper lobster raw extracts. For the gastropods the quantified TM is similar to the identification by LC/qTOF. However, for the cephalopods and the bivalves, the identification profile by LC/qTOF is different compared to the quantified TM amount by LC/MRM. This might be due to the different protein profile pattern observed in the cephalopods and the bivalves in comparison to other shellfish species. However, the results in table 5.8 clearly show that the validated LC/MRM method is sensitive and accurate in the quantification of tropomyosin in various shellfish species, thus TM could be quantified in all species with the exception of King prawn raw extract.

Table 5.8: Summary of the amount of the quantified peptides found in various raw shellfish species extracts, divided into the different subgroups. The concentration is shown in nM peptide per mg protein extract. (N.D.= Not detected; B.Q.= below quantification)

		Raw extracts	nM peptide/mg protein extract			
		Sample name	Peptide 1	Peptide 2	Peptide 3	Peptide 4
Crustacea	Prawns	King prawn	N.D.	N.D.	N.D.	N.D.
		Black Tiger prawn	40.89	33.05	N.D.	60.06
		Vannamei prawn	0.95	B.Q.	N.D.	1.11
		Banana prawn	48.02	N.D.	N.D.	36.98
		Green Tiger prawn	141.72	56.13	N.D.	162.11
	Crabs	Blue Swimmer crab	935.98	502.41	N.D.	6.48
		Sand crab	823.38	337.98	N.D.	N.D.
		Mud crab	658.37	50.03	N.D.	N.D.
	Lobsters	Rock lobster	1,049.61	14.73	N.D.	743.98
		Slipper lobster	7,912.20	1,622.93	N.D..	8,197.64
		Yabby	66.82	40.56	N.D.	19.51
Mollusca	Gastropods	Jade Hybrid Tiger abalone	N.D.	N.D.	85.46	N.D.
		Sea snail	N.D.	N.D.	24.51	N.D.
	Cephalopods	Squid (Calamari)	N.D.	179.18	555.76	N.D.
		Octopus	6.48	931.49	197.40	N.D.
	Bivalves	Blue mussel	N.D.	36.93	68.16	N.D.
		Green mussel	N.D.	N.D.	29.81	N.D.
		Scallop (*fumatus)*	N.D.	N.D..	445.09	N.D.
		Scallop (*yessonensis)*	N.D.	N.D.	1,222.79	N.D.
		Tasmanian oyster	N.D.	N.D.	21.39	N.D.
		Sydney Rock oyster	N.D.	N.D.	18.93	N.D.
		Tuatua cockle	N.D.	N.D.	3,031.16	N.D.

Table 5.9: Summary of the amount of the quantified peptides found in various whole heated shellfish species extracts, divided into the different subgroups. The concentration is shown in nM peptide per mg protein extract. (N.D.= Not detected)

		Whole heated extract	nM peptide/mg protein extract			
		Sample name	Peptide 1	Peptide 2	Peptide 3	Peptide 4
Crustacea	Prawns	King prawn	14,282.26	7,125.41	N.D.	12,583.03
		Black Tiger prawn	2,394.66	1,091.59	N.D.	2,951.93
		Vannamei prawn	9,908.35	6,438.03	N.D.	11,452.92
		Banana prawn	24,722.21	22,348.19	N.D.	36,437.32
		Green Tiger prawn	5,846.16	3,779.99	N.D.	6,280.26
	Crabs	Blue Swimmer crab	89,922.33	53,892.45	N.D.	1,945.08
		Sand crab	11,673.37	9,746.55	N.D.	N.D.
		Mud crab	22,679.74	20,260.37	N.D.	N.D.
	Lobsters	Rock lobster	16,104.78	7,779.21	N.D.	8,658.45
		Slipper lobster	18,841.11	627.78	N.D.	10,441.59
		Yabby	34,507.87	22,075.15	N.D.	18,613.24
Mollusca	Gastropods	Jade Hybrid Tiger abalone	27.25	34.96	204.75	23.58
		Sea snail	N.D.	N.D.	679.33	N.D.
	Cephalopods	Squid (Calamari)	N.D.	8,028.31	647.59	N.D.
		Octopus	16.12	3,136.80	415.13	N.D.
	Bivalves	Blue mussel	N.D.	269.90	879.01	N.D.
		Green mussel	N.D.	N.D.	1,024.76	N.D.
		Scallop (*fumatus*)	N.D.	N.D.	2,869.51	N.D.
		Scallop (*yessonensis*)	N.D.	N.D.	2,599.27	N.D.
		Tasmanian oyster	N.D.	N.D.	1,035.68	N.D.
		Sydney Rock oyster	N.D.	N.D.	298.04	N.D.
		Tuatua cockle	N.D.	N.D.	5,863.82	N.D.

Table 5.9 summarises that TM could be quantified in all whole heated species analysed, with high amount of TM present in all the shellfish species. Many of the samples have a higher concentration than for the optimised linearity of the validated LC/MRM method. Thus, all the whole heated extracts were diluted after digestion to fit the linear range of the validated standard calibration curve. In chapter 3 the only species where TM could not be identified was Sydney Rock oyster, where table 5.9 shows that the concentration of TM in Sydney Rock oyster is generally very low. However, the only subgroups where the identified TM is similar to the quantified TM concentration are the lobster and the cephalopod extracts. It is hypothesised this might be due the lower concentration of TM in these two subgroups in comparison to the prawns and the crabs. Moreover, it is suggested that the identification of TM was less accurate in chapter 4, compared to the quantified TM concentration in this chapter, due to limited sequence availability for molluscs.

Overall, it is clearly demonstrated in table 5.8 and table 5.9 that the concentration of TM is higher in the whole heated extracts compared to the raw extracts. Moreover, in the whole heated crustacean extracts the concentration of TM is higher compared to the whole heated molluscs, potential confirming that the overall main protein in molluscs is actin. However, the TM concentration in the raw crabs, the lobsters and the molluscs is similar, compared to the lower TM concentration in the raw prawns.

TM could not be identified in raw KP extract by the sensitive LC/MRM method. Therefore raw purified TM and rTM KP were analysed by the validated LC/MRM method. Results in table 5.10 show that the concentration of the purified and the recombinant TM is very similar. Moreover, the differences of the calculated concentrations are lower for the four peptide detected. Although TM could not be detected in KP raw extract by validated LC/MRM method, identical to LC/qTOF data (Chapter 3), it was shown that it can be detected in the purified raw KP extract. Therefore one can assume that TM is present, just at very low concentration.

Table 5.10: Summary of the amount of the quantified peptides for the recombinant King prawn and the purified raw King prawn extract. The concentration is shown in nM peptide per mg protein extract. (N.D.= Not detected)

		Sample name	nM peptide/mg protein extract			
		Recombinant and purified TM	Peptide 1	Peptide 2	Peptide 3	Peptide 4
Crustacea	Prawns	King prawn rTM	2,167.31	1,756.14	N.D.	3,863.39
		King prawn raw purified TM	3,535.38	2,517.39	N.D.	4,009.35

The validated LC/MRM method detected TM in all the shellfish extracts, with the exception of KP raw extract. TM could not be identified by LC/qTOF methods for three prawns, yabby, Tasmanian oyster and Sydney Rock oyster, but can be quantified by the more sensitive LC/MRM method. However, the values quantified for species, which were not detected by LC/qTOF, are relatively low for LC/MRM, confirming the hypothesis, that the concentration of TM is lower in these shellfish extracts. Unfortunately, it is uncertain why the concentration of TM in some raw samples is low. Moreover, it is hypothesised that the TM identification applying LC/qTOF underestimates the TM content, especially in the molluscs, due to unknown TM sequences. Thus, for example the identification of TM in Tuatua cockle extracts was low by LC/qTOF, but the quantified TM concentration by LC/MRM in raw and whole heated extract is the highest TM concentration for all mollusc species investigated. However, the digestion method prior to the LC/MRM method should be optimised, thus it seems that some digestions are incomplete, especially the peptide 2 and possible the peptide 3, when comparing the concentrations calculated for different samples in the same extract.

Overall, the developed and validated LC/MRM method is suitable to distinguish the crustacean and the mollusc species, while sensitive and absolute quantification of TM can be achieved. The selected peptides were, with four exceptions, detected in the species as predicted in chapter 3. Even with the crabs it was shown that the peptide 4 could be detected in the species that were exempt from the rule as not being present in the crabs as shown in chapter 3.

5.4 Discussion

A LC/MRM method was developed and validated using the Guidance for Industry for Bioanalytical Method Validation.[14] The linearity, the accuracy, the precision and the sensitivity met all criteria required for the LC/MRM method, using the four peptides for the absolute quantification of allergenic tropomyosin from shellfish. Moreover, with the four peptides applied, the crustacean and the molluscs can be distinguished from each other, as required by the legislation and the food industry.

The four peptides chosen are based on TM, however, for the different species a minimum of one peptide and a maximum of three peptides were detected, although the literature generally recommends to used at least two peptides per protein.[16-18] The more peptides per protein and the more transitions per protein are chosen, the chances of quantifying false positive or false negative proteins is decreasing. However, the more peptides and transitions are selected, the less sensitive the LC/MRM method becomes.[17-19] Most commonly two transitions per peptide are chosen in the literature,[2-4, 20-22] with the exception of Shefcheck et al.[23] and Houston et al.,[24] who chose three transitions per peptide and Molle and Leonil[25] and Careri et al.,[26, 27] who only used one transition per peptide. Nevertheless, the quantification of proteins is only based on one single petide.[18, 19] Therefore the quantification of the crustacean was based on peptide 1 and the quantification of the molluscs was based on the peptide 3. All four peptides are quantified based on the highest fragment ion as quantification ion, but observing three transitions per peptide.

Seven publications[2, 4, 5, 12, 22, 24, 25] used molar concentration (M) for the generation of the calibration standard curve, and five publications[3, 21, 26-28] used grams. The most sensitive calibration curve was reported by Newsome and Scholl[22] ranging from 1.3-150 femtomolar. The calibration curves generated in this chapter are in the range of 1-1000 nanomolar (nM) and 2.5-1000 nM, respectively. These ranges are higher than reported by Newsome and Scholl[22] and Molle and Leonil,[25] however, in the range of Zhang et al.[4] and all the crustacean species quantified by mass spectrometry reported by Abdel Rahamn.[2, 5, 12] Moreover, the LOD (0.5 nM and 1 nM) and the LOQ (1 nM and 2.5 nM) found for the four tropomyosin peptides are similar to the values reported for tropomyosin from Snow crab (LOD 3nM and LOQ 10 nM)[2, 5] and Northern prawn (LOD 0.25 nM).[12]

This is the first study reporting and applying signature peptides for the quantification of a highly similar pan-allergen (TM), thus all other reported studies focus on one specific allergen and species specific signature peptides,[1] therefore results cannot be compared with the literature. However, with the exception of peptide 1 for octopus, the peptides 1, 2 and 4 for Hybrid abalone and the peptide 2 for Blue mussel and Banana prawn, the peptides were detected in all other 18 shellfish species as predicted in chapter 3. Unfortunately, it is not clear why the peptides where detected in those four species, although based on *in silico* data they should be absence. TM and other proteins were searched to fit the precursor and the transitions for the selected peptides using http://prospector.ucsf.edu, however, the precursors and the three transitions chosen are highly specific for the selected peptides, thus no other real peptide could be identified for causing the false positive identified peptides. Nevertheless, for Hybrid abalone the amino acid sequence of TM is unknown, thus the detected peptides could be potential present in this species. Nevertheless, it was reported that different species include TM isoforms with up to three isoforms for crustacean[29] and two isoforms for bivalves,[30] therefore the detected peptides in those species could be potential isoforms, but to date not reported in the Genbank. To investigate why these peptides were detected, more mollusc species would need to be analysed and the availability of more shellfish tropomyosin species would be beneficial.

Interestingly, it was observed that the calculated concentration for the peptides differ, although they are all generated from tropomyosin and therefore the same concentration should be observed. Newsome and Scholl[22] quantified two peptides from the same protein and the concentration reported where similar for those two peptides. However, other studies that quantified more than one protein and therefore analysing more than one peptide demonstrated also that the peptides have a different linear range[24] and the intensity of the peptides detected varies, thus the concentration varies as well.[23, 27] Overall, for most species the factor between the quantified peptides was lower than a factor of 5. Therefore one can assume that this might be due to the stability of peptides and sample preparation, especially difference in enzyme activity and possible incomplete digestion of the peptides. It is suggested that the digestion method should be investigated and

improved to achieve complete digestion of all peptides, as demonstrated by Zhang et al.[4]

The concentration of TM calculated in the different shellfish species by the validated LC/MRM method was similar to the TM identification by LC/qTOF. However, one can assume that the TM identification by LC/qTOF underestimated the concentration in several mollusc species due to sequence unavailability. Furthermore, TM could be quantified by the more sensitive LC/MRM method in more different species than by LC/qTOF identification. Generally, the concentration of TM was higher in all whole heated extracts, compared to the raw extracts, especially for the prawns by LC/MRM. Moreover, the whole heated crustacean had a higher TM content compared to the whole heated molluscs, as previously demonstrated by SDS-PAGE and immunoblotting using a monoclonal anti-tropomyosin antibody.[31]

Overall, a novel LC/MRM method was developed and validated for the quantification of TM. The results of the validation fulfilled the criteria of the Guidance for Industry for Bioanalytical Method Validation.[14] Moreover, the linearity, the LOD and the LOQ observed for the four selected peptides from tropomyosin are similar to values reported for other quantified allergens in the literature. The four peptides were detected in all the shellfish species analysed as predicted, with the exception of four species, therefore they are suitable to distinguish allergenic tropomyosin from crustacean and molluscs, as required by the legislation in the European Union and Canada.

5.5 References

1. Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). *Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems.* Journal of Proteome Research. 13, (8), 3499-3509.
2. Abdel Rahman, A. M.; Lopata, A. L.; Randell, E. W.; Helleur, R. J. (2010). *Absolute Quantification Method and Validation of Airborne Snow Crab Allergen Tropomyosin Using Tandem Mass Spectrometry.* Analytica Chimica Acta. 681, (1-2), 49-55.
3. Sealey-Voyksner, J. A.; Khosla, C.; Voyksner, R. D.; Jorgenson, J. W. (2010). *Novel Aspects of Quantitation of Immunogenic Wheat Gluten Peptides by Liquid Chromatography-Mass Spectrometry/Mass Spectrometry.* Journal of Chromatography A. 1217, (25), 4167-4183.
4. Zhang, J.; Lai, S.; Zhang, Y.; Huang, B.; Li, D.; Ren, Y. (2012). *Multiple Reaction Monitoring-Based Determination of Bovine A-Lactalbumin in Infant Formulas and Whey Protein Concentrates by Ultra-High Performance Liquid Chromatography–Tandem Mass Spectrometry Using Tryptic Signature Peptides and Synthetic Peptide Standards.* Analytica Chimica Acta. 727, (0), 47-53.
5. Abdel Rahman, A.; Gagné, S.; Helleur, R. (2012). *Simultaneous Determination of Two Major Snow Crab Aeroallergens in Processing Plants by Use of Isotopic Dilution Tandem Mass Spectrometry.* Analytical and Bioanalytical Chemistry. 403, (3), 821-831.
6. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* Int J Hyg Environ Health. 217, (7), 740-50.
7. Fuller, H. R.; Goodwin, P. R.; Morris, G. E. (2006). *An Enzyme-Linked Immunosorbent Assay (Elisa) for the Major Crustacean Allergen, Tropomyosin, in Food.* Food and Agricultural Immunology. 17, (1), 43-52.
8. Seiki, K.; Oda, H.; Yoshioka, H.; Sakai, S.; Urisu, A.; Akiyama, H.; Ohno, Y. (2007). *A Reliable and Sensitive Immunoassay for the Determination of Crustacean Protein in Processed Foods.* Journal of Agricultural and Food Chemistry. 55, (23), 9345-9350.
9. Werner, M. T.; Faeste, C. K.; Egaas, E. (2007). *Quantitative Sandwich Elisa for the Determination of Tropomyosin from Crustaceans in Foods.* Journal of Agricultural and Food Chemistry. 55, (20), 8025-8032.
10. Zhang, H.; Lu, Y.; Ushio, H.; Shiomi, K. (2014). *Development of Sandwich Elisa for Detection and Quantification of Invertebrate Major Allergen Tropomyosin by a Monoclonal Antibody.* Food Chemistry. 150, (0), 151-157.
11. Abdel Rahman, A. M.; Kamath, S.; Lopata, A. L.; Helleur, R. J. (2010). *Analysis of the Allergenic Proteins in Black Tiger Prawn (Penaeus Monodon) and Characterization of the Major Allergen Tropomyosin Using Mass Spectrometry.* Rapid Communications in Mass Spectrometry. 24, (24), 3624-3624.
12. Abdel Rahman, A. M.; Kamath, S. D.; Gagne, S.; Lopata, A. L.; Helleur, R. (2013). *Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention.* Journal of Proteome Research. 12, (2), 647-656.
13. Abdel Rahman, A. M.; Lopata, A. L.; O'Hehir, R. E.; Robinson, J. J.; Banoub, J. H.; Helleur, R. J. (2010). *Characterization and De Novo Sequencing of Snow Crab Tropomyosin Enzymatic Peptides by Both Electrospary Ionization and Matrix-Assisted Laser Desorption Ionization Qqtof Tandem Mass Spectrometry.* Journal of Mass Spectrometry. 45, (4), 372-381.
14. Department of Health Human Services Food and Drug Administration, C. f. D. E. a. R. C., Center for Biologics Evaluation and Research (CBER). (2001). *Guidance for Industry Bioanalytical Method Validation.*
15. Koeberl, M.; Kamath, S. D.; Saptarshi, S. R.; Smout, M. J.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Auto-Induction for High Yield Expression of Recombinant Novel Isoallergen Tropomyosin from King Prawn (Melicertus Latisulcatus) for Improved Diagnostics and Immunotherapeutics.* J Immunol Methods. 415, (0), 6-16.
16. Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. (2008). *Selected Reaction Monitoring for Quantitative Proteomics: A Tutorial.* Molecular Systems Biology. 4, (1), n/a-n/a.

17. Picariello, G.; Mamone, G.; Addeo, F.; Ferranti, P. (2011). *The Frontiers of Mass Spectrometry-Based Techniques in Food Allergenomics.* Journal of Chromatography A. 1218, (42), 7386-7398.

18. Rauh, M. (2012). *Lc-Ms/Ms for Protein and Peptide Quantification in Clinical Chemistry.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences. 883, 59-67.

19. Simon, R.; Jubeaux, G.; Chaumot, A.; Lemoine, J.; Geffard, O.; Salvador, A. (2010). *Mass Spectrometry Assay as an Alternative to the Enzyme-Linked Immunosorbent Assay Test for Biomarker Quantitation in Ecotoxicology: Application to Vitellogenin in Crustacea (Gammarus Fossarum).* Journal of Chromatography A. 1217, (31), 5109-5115.

20. Heick, J.; Fischer, M.; Kerbach, S.; Tamm, U.; Popping, B. (2011). *Application of a Liquid Chromatography Tandem Mass Spectrometry Method for the Simultaneous Detection of Seven Allergenic Foods in Flour and Bread and Comparison of the Method with Commercially Available Elisa Test Kits.* Journal of Aoac International. 94, (4), 1060-1068.

21. Heick, J.; Fischer, M.; Popping, B. (2011). *First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry.* Journal of Chromatography A. 1218, (7), 938-943.

22. Newsome, G. A.; Scholl, P. F. (2012). *Quantification of Allergenic Bovine Milk As1-Casein in Baked Goods Using an Intact 15n-Labeled Protein Internal Standard.* Journal of Agricultural and Food Chemistry. 61, (24), 5659-5668.

23. Shefcheck, K. J.; Callahan, J. H.; Musser, S. M. (2006). *Confirmation of Peanut Protein Using Peptide Markers in Dark Chocolate Using Liquid Chromatography-Tandem Mass Spectrometry (Lc-Ms/Ms).* Journal of Agricultural and Food Chemistry. 54, (21), 7953-7959.

24. Houston, N. L.; Lee, D.-G.; Stevenson, S. E.; Ladics, G. S.; Bannon, G. A.; McClain, S.; Privalle, L.; Stagg, N.; Herouet-Guicheney, C.; MacIntosh, S. C.; Thelen, J. J. (2010). *Quantitation of Soybean Allergens Using Tandem Mass Spectrometry.* Journal of Proteome Research. 10, (2), 763-773.

25. Molle, D.; Leonil, J. (2005). *Quantitative Determination of Bovine Kappa-Casein Macropeptide in Dairy Products by Liquid Chromatography/Electrospray Coupled to Mass Spectrometry (Lc-Esi/Ms) and Liquid Chromatography/Electrospray Coupled to Tamdem Mass Spectrometry (Lc-Esi/Ms/Ms).* International Dairy Journal. 15, (5), 419-428.

26. Careri, M.; Costa, A.; Elviri, L.; Lagos, J. B.; Mangia, A.; Terenghi, M.; Cereti, A.; Garoffo, L. P. (2007). *Use of Specific Peptide Biomarkers for Quantitative Confirmation of Hidden Allergenic Peanut Proteins Ara H 2 and Ara Ill 3/4 for Food Control by Liquid Chromatography-Tandem Mass Spectrometry.* Analytical and Bioanalytical Chemistry. 389, (6), 1901-1907.

27. Careri, M.; Elviri, L.; Lagos, J. B.; Mangia, A.; Speroni, F.; Terenghi, M. (2008). *Selective and Rapid Immunomagnetic Bead-Based Sample Treatment for the Liquid Chromatography-Electrospray Ion-Trap Mass Spectrometry Detection of Ara H3/4 Peanut Protein in Foods.* Journal of Chromatography A. 1206, (2), 89-94.

28. Monaci, L.; van Hengel, A. J. (2008). *Development of a Method for the Quantification of Whey Allergen Traces in Mixed-Fruit Juices Based on Liquid Chromatography with Mass Spectrometric Detection.* Journal of Chromatography A. 1192, (1), 113-120.

29. Lopata, A. L.; O'Hehir, R. E.; Lehrer, S. B. (2010). *Shellfish Allergy.* Clinical and Experimental Allergy. 40, (6), 850-858.

30. Fujinoki, M.; Ueda, M.; Inoue, T.; Yasukawa, N.; Inoue, R.; Ishimoda-Takagi, T. (2006). *Heterogeneity and Tissue Specificity of Tropomyosin Isoforms from Four Species of Bivalves.* Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 143, (4), 500-506.

31. Kamath, S. D.; Abdel Rahman, A. M.; Komoda, T.; Lopata, A. L. (2013). *Impact of Heat Processing on the Detection of the Major Shellfish Allergen Tropomyosin in Crustaceans and Molluscs Using Specific Monoclonal Antibodies.* Food Chemistry. 141, (4), 4031-9.

5.6 Summary chapter 5:

Quantification of Allergenic Tropomyosin using Signature Peptides

- ✔ A liquid chromatography (LC) method coupled with multiple reaction monitoring (LC/MRM) was developed, using the four signature peptide derived from tropomyosin, identified in the previous chapter to distinguish crustacean and mollusc species
- ✔ The developed LC/MRM method was validated, fulfilling all the criteria necessary for the validation according to the Department of Health Human Services Food and Drug Administration using the Guidance for Industry for Bioanalytical Method Validation
- ✔ 22 shellfish species were analysed using the validated LC/MRM method employing the raw and whole heated extracts
- ✔ Two peptides (peptide 1 and peptide 3) can distinguish between the different crustacean and mollusc species analysed in this study, with the exception of peptide 1 and the species Hybrid abalone and octopus. One peptide (peptide 2) can distinguish between the crustacean, the cephalopods and the bivalves, with the exception of Blue mussel and Banana prawn. One peptide (peptide 4) can distinguish between the prawns, the lobsters and the crabs as predicted by their amino acid sequence.

The validated LC/MRM method will be applied in chapter 6 for the quantification of TM from crustacean and molluscs using food samples. The results of the validated LC/MRM method will be compared with two commercial available ELISA kits.

CHAPTER 6

ANALYSIS OF FOOD SAMPLES FOR ALLERGENIC TROPOMYOSIN

6.1 Introduction

Shellfish allergens belong to “The “Big 8” food allergens.[1, 2] Shellfish is a generic term and is in general separated into two categories in human consumption, crustacean and molluscs. Either of these two categories can cause severe allergic reactions. Therefore the legislation in Canada and the European Union implemented guidelines to distinguish crustacean from molluscs allergens in food products.[1] Currently available methods for shellfish allergen detection and quantification, such as ELISA, are based on tropomyosin (TM), the major allergen in shellfish. However, not all commercial ELISA kits specify the selection of standards, components, applied antibody specificity or the binding sites of the antibodies. Nevertheless, many ELISA kit manufactures mention that the shellfish allergen quantification is based on TM or other shellfish/crustacean/prawn proteins. Most experimental developed ELSIA kits are based on TM. In detail, different research groups have developed various ELISAs for the detection and quantification of shellfish allergens.[3-7] Four out of five developed ELISAs are targeting exclusively allergenic TM from crustacean, whereas only two were tested and confirmed not to detect mollusc allergens.[5, 6] Only one ELISA targeted crustacean and molluscs TM, but cannot distinguish them.[7] Moreover, three research groups investigated and reported the existing cross-reactivity to other arthropod species.[3, 6, 7] In summary, there is no ELISA system available that would be able to distinguish between crustacean and molluscs allergens.

Food samples are very often highly processed, hence proteins and allergens can become modified.[8] It was reported that heat treatment enhances allergenicity of TM,[9] in contrast to food processing reducing the allergenicity of TM,[10-12] leading to different antibody binding. Moreover, TM can interact with other food components during food processing. For example, the Maillard reaction can easily occur, thus TM is a lysine rich protein and can easily react with reducing sugars, especially if food components have been dried such as noodles.[12, 13] Thus, it was reported that the Maillard reaction increases the allergenicity of TM,[14] but decrease the digestion efficiency with different enzymes investigated.[12, 13] Overall, it was reported, that food processing can have a considerable impact for the detection of allergens using ELISA. Especially heat treatment decreased the detection ability of ELISA.[8] Moreover, other food compounds, such as higher protein or lipid content

decreased the detection of allergens by ELISA.[8] The only current publication comparing different commercial available ELISAs for shellfish allergen quantification was performed by Sakai et al.[15] In this study two available ELISAs in Japan for crustacean allergen detection were compared in an inter-laboratory approach. The results showed that both ELISAs are useful to detect crustacean proteins and are validated appropriately. However, one kit produced more accurate results for spiked crustacean protein in food samples, whilst the other kit performed better in repeatability and reproducibility, due to higher precision of the assay performance of the kit.[15]

ELISAs are currently still the most common technique to analyse food allergens. The aim of this chapter is to compare current antibody based methods with mass spectrometry using highly processed food samples. Therefore in this chapter thirteen food products will be analysed with two commercial available ELISA kits and the mass spectrometric methods developed in chapter 3 and 5. The food samples were selected according to their potential shellfish allergen content, including noodles and oily food products.[8, 12, 13] In detail, three samples do not contain shellfish, four samples definitely contain shellfish and six samples have the possibility to contain shellfish allergens. This analysis of various methods will provide better knowledge about the quantification of highly processed food samples.

6.1.1 Aims

The aims of this chapter are:

- Analysing highly processed food samples utilising specific antibodies
- Analysing highly processed food samples utilising commercial ELISA kits
- Analysing highly processed food samples utilising the developed LC/qTOF and validated LC/MRM methods
- Comparing all methods for the analysis of allergenic tropomyosin
- Comparing all methods for the quantification of allergenic tropomyosin

6.2 Materials and methods

6.2.1 Food protein extraction

Thirteen different food samples have been purchased from the local supermarket (Townsville, Australia). The selected samples either possibly contain highly processed tropomyosin (TM) or are selected as negative controls. Moreover, six samples are unknown for the allergenic tropomyosin content, however, the name of the food product includes a shellfish name. Figure 6.1 displays the expected allergenic TM content, the sample names and pictures as details on the selected food products.

Protein extracts for canned prawn, canned crab, smoked canned mussel, smoked canned oyster and chicken muscle have been generated similar to the raw shellfish extracts as described in chapter 4, thus these food samples include edible muscle tissue. Briefly, for canned food products 40 g and 17 g for chicken muscle were homogenised in 100 ml PBS, respectively. Moreover, for shrimp crackers 16 g were homogenised in 100 ml PBS.

The instant noodles (prawn and chicken noodles, shrimp noodles and crab noodles) were prepared as stated on the cooking instructions, using PBS instead of water. After preparing the noodles, 36 g prawn and chicken noodles, 20 g shrimp noodles and 15 g crab noodles were homogenised in 100 ml PBS. For the liquid food samples, 125 g vegetable soup, 125 g chicken soup, 240 g oyster sauce and 120 g red Thai curry paste were extracted in 100 ml PBS.

All total homogenised extracts were kept at 4°C overnight while continuously shaking, followed by centrifugation and sterile filtration and finally stored at -80°C until further use.

The recombinant TM from King prawn was expressed using auto-induction as explained in detail in chapter 2 and published in Koeberl et al.[16]

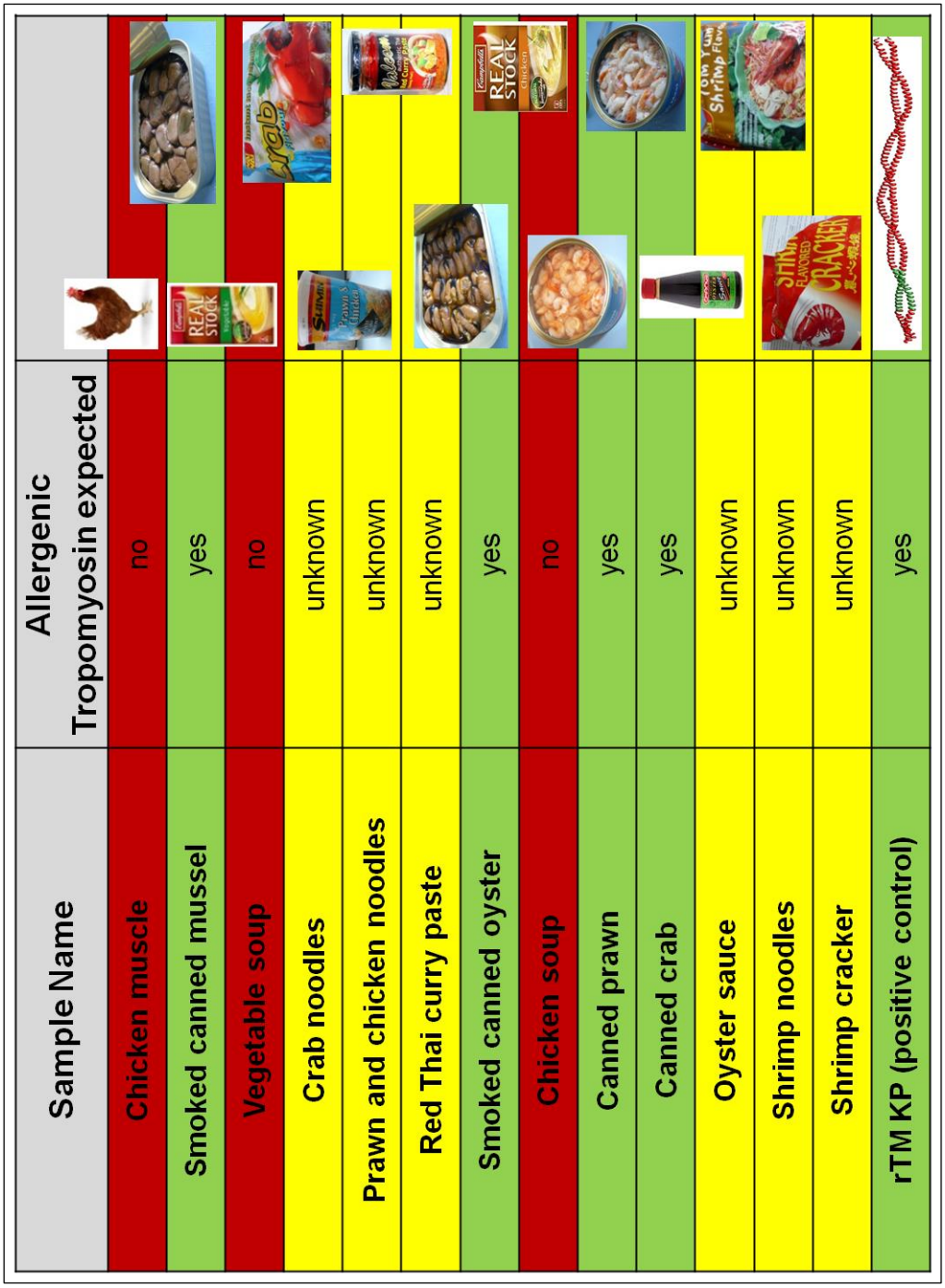


Figure 6.1: Summary of the selected food samples, including pictures of the selected food product of the manufacturer. The samples highlighted in green are expected to contain highly processed allergenic tropomyosin (TM), whereas the samples highlighted in red are not expected to contain any allergenic TM (negative controls). Moreover, the samples highlighted in yellow are unknown to contain allergenic TM.

6.2.2 SDS-PAGE and immunoblotting

SDS-PAGE and immunoblotting was performed as explained in detail in chapter 2. Protein extracts were separated on a 12% polyacrylamide gel, using 20 µl for SDS-PAGE separation and 10 µl for immunoblotting. The protein components were resolved at 200 V until the tracker dye reached the base. To investigate if proteins can be detected by antibodies, monoclonal anti-tropomyosin antibody and in-house[17] generated polyclonal anti-crustacean antibody were applied.

6.2.3 ELISA

Two commercial available ELISA kits for the analysis of shellfish allergens have been purchased from two different companies in Australia, referred as company A and company B. The company A specifies to detect TM and using anti-tropomyosin antibodies, whereas the company B specifies to use crustacean proteins as standards and anti-crustacean-proteins antibodies. The food samples have been extracted as described above. The extracts have been diluted according to the manufactures instructions using the provided extraction buffers, whereas the company B used a greater dilution factor comparing to the company A. The food samples have been applied and analysed following the manufactures instructions. The food samples and the provided standards were analysed in duplicates. Dilutions and the standards were prepared freshly on the day of use and the standard curve was measured with the samples each time. The absorbance of the standards leads to the generation of the calibration curve in ppm. The concentration of the samples was calculated for the various food samples depending on this standard curve (ppm). The concentration of samples has been recalculated to the dilution factor used in the sample preparation and as stated by the manufacture.

6.2.4 Dot blotting

The concentration of various food samples was equally to the ELISA kit manufactures instructions. Samples were applied on PVDF membrane, which has been activated and incubated in sodium bicarbonate buffer (15 mM Na_2CO_3 and 85 mM $NaHCO_3$). The samples were applied and absorbed onto the membrane

while incubating at room temperature for 30 min. The membrane was washed with PBS-T and blocked with 5% skim milk powder in PBS-T, respectively. For the detection of tropomyosin the same primary and secondary antibodies were applied, following the same steps as described in chapter 2, using the monoclonal anti-tropomyosin antibody and the in-house[17] polyclonal anti-crustacean antibody.

6.2.5 Mass spectrometry

All food samples were analysed by both developed mass spectrometry systems, namely the four developed LC/qTOF methods and the validated LC/MRM method. The extracted proteins from the food samples are digested as described in the previous chapters (Chapter 3, 4 and 5), using trypsin spin columns and a digestion time of 15 minutes at room temperature.

The digested food samples were analysed by all four final LC/qTOF methods (Chapter 3) and the proteins were identified using Mascot with the same final search parameters. The sample specific proteins and species summarised in table 6.2 represent the proteins identified with the highest Mascot score.

The digested food samples were analysed with the validated LC/MRM method (Chapter 5). Briefly, each digested protein extract was spiked with internal standard 1-4 in a concentration of 200 nM after digestion, prior to LC/MRM analysis. The standard curve was prepared freshly every day and measured prior to the quality control (QC) samples, followed by the digested food sample extracts. Between the QC samples and the food sample analysis the LC/MRM washing method was performed. The QC samples were reinjected after six different food samples were analysed, to confirm the reproducibility of analytical setting. Each sample was injected in triplicate and the concentrations of the samples were calculated based on the freshly prepared standard curve using the area-ratio values, whereas each standard curve and QC met the method acceptance criteria of the validation. The accuracy and the precision were calculated based on the standard deviation (SD) of the triplicate (injection) and were considered acceptable if the SD was ±20%. When the concentration of the digested food sample was outside the linear range of the standard calibration curve the sample was diluted after digestion.

6.2.6 Calculations for unit conversion

To be able to compare the results of the food samples obtained from two different ELISA kits and the LC/MRM method, the various units were converted into ppm per gram food sample.

The concentration generated by ELISA is in ppm, however, the amount of analysed food sample is different for the two ELISA kits. The quantified amounts of TM by LC/MRM are obtained in nanomolar (nM). Therefore the LC/MRM results were converted into ppm, as final result the peptide 1 and the peptide 3 were used for the quantified TM of crustacean or molluscs, respectively. As next step all the results were recalculated to ppm per one gram of food sample used for the different techniques applied.

6.3 Results

To compare the developed LC/qTOF and the validated LC/MRM methods with the commonly used method for allergen detection (ELISA), thirteen different food samples were analysed. The 13 food samples were chosen to either possible contain highly processed tropomyosin (TM) or were selected as negative controls matrixes, which should not contain and TM (Figure 6.1). In detail, three food samples are negative controls, four samples should contain allergenic TM and six samples could be possible containing allergenic TM, as the product name includes a name related to shellfish. In addition, the recombinant tropomyosin (rTM) from King prawn (KP) was analysed as a positive control. Analysing these 13 different food samples using the two commercial available ELISA kits and the two different mass spectrometry systems will lead to confirmation of the selectively and the sensitivity of mass spectrometry compared to currently commonly used methods for allergen quantification. The results of the method comparison will be explained in the following sections.

6.3.1 Food sample analysis

6.3.1.1 Analysis of food samples by SDS-PAGE and immunoblotting

All thirteen food samples were analysed by SDS-PAGE and immunoblotting (Figure 6.2), to obtain information about the food protein extracts and the possible allergenic TM content. Figure 6.2a shows clearly that the food samples are highly processed, thus for most samples only a smear is visible by SDS-PAGE. Interestingly, the food products containing more oily compounds, such as smoked canned mussel, smoked canned oyster, oyster sauce and red Thai curry paste, show less intense smears as compared to canned prawn and canned crab. Surprisingly, for prawn and chicken noodles bands were visible by SDS-PAGE separation, especially in the 37 kDa range, which is the size of TM.

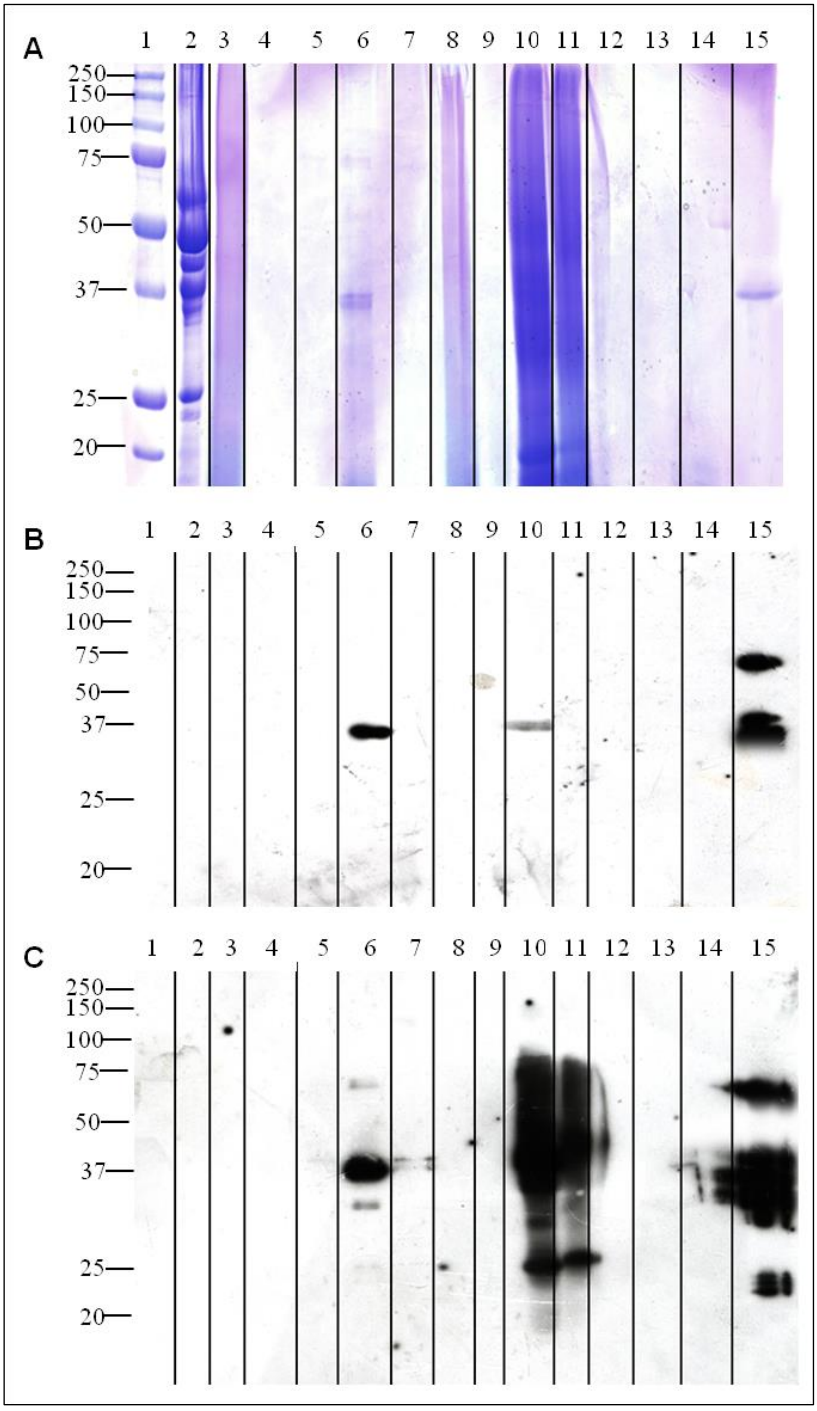


Figure 6.2: Analysis of the 13 food protein extracts by (A) SDS-PAGE (B) monoclonal anti-tropomyosin antibody and (C) polyclonal anti-crustacean antibody. Lane numbers represent: 1. Marker; 2. Chicken muscle; 3. Smoked canned mussel; 4. Vegetable soup; 5. Crab noodles; 6. Prawn and chicken noodles; 7. Red Thai curry paste; 8. Smoked canned oyster; 9. Chicken soup; 10. Canned prawn; 11. Canned crab; 12. Oyster sauce; 13. Shrimp noodles; 14. Shrimp crackers; and 15. rTM KP.

To confirm if the visible bands in the prawn and chicken noodles are TM, the food samples were analysed by immunoblotting applying the monoclonal anti-tropomyosin and the polyclonal anti-crustacean antibodies. Figure 6.2b and figure 6.2c show that TM was identified in prawn and chicken noodles. However, with the monoclonal antibody, TM could only be detected in canned prawn. Applying the polyclonal anti-crustacean antibody more food samples were detected, including samples, which could be detected by the monoclonal anti-tropomyosin antibody. Moreover, an intense band is visible for canned crab. Less intense bindings can be observed for red Thai curry paste and shrimp crackers. The samples chosen as negative controls were not detected by the antibodies, and rTM KP showed strong antibody binding. However, the mollusc food samples were not detected by either antibody applied.

Overall, the SDS-PAGE and immunoblotting analysis showed that the food samples are highly processed. Moreover, some samples were tested positive by the antibodies applied. Nevertheless, due to high processing factor of food samples these two methods for allergenic TM detections are not the best option.

6.3.1.2 Analysis of food samples by two commercial available ELISA kits

The same 13 food samples have been analysed by two commercial available ELISA kits in Australia (company A and company B). The quantification of TM for the food samples was performed as explained in the material and method section. The ppm of the quantified TM calculated according to the manufacturer instructions for the company A and the company B are summarised in table 6.1.

A total of nine food samples were quantified using the ELISA kit from company A. The concentration (ppm) calculated for the food samples utilising the ELISA kit from company A were generally lower for all samples investigated. The samples originated from crustacean, including allergenic TM, were detected correctly, whereas for the food samples including allergenic tropomyosin from molluscs, only smoked canned oyster was quantified. Moreover, this ELISA kit detected one negative control, namely vegetable soup. Potentially the selected negative controls can be contaminated and therefore TM might be present. The analysis of the food

samples by MS and dot blotting technique in the following sections will confirm or contradict the possible contamination of the vegetable soup.

Table 6.1: Summary of the ELISA kits results for TM quantification for the 13 different food samples. The results are shown in ppm and calculated according to the manufactures instructions. The samples highlighted in green are expected to contain highly processed allergenic tropomyosin (TM), whereas the samples highlighted in red are not expected to contain allergenic TM (negative controls). Moreover, the samples highlighted in yellow are unknown to contain TM.

Sample number	Sample name	Company A (Concentration in ppm)	Company B (Concentration in ppm)
2	Chicken muscle	0.00	0.00
3	Smoked canned mussel	0.00	0.00
4	Vegetable soup	**0.05**	0.00
5	Crab noodles	0.00	0.00
6	Prawn and chicken noodles	**0.68**	**15.20**
7	Red Thai curry paste	**0.08**	0.00
8	Smoked canned oyster	**0.11**	0.00
9	Chicken soup	0.00	0.00
10	Canned prawn	**0.57**	**15.00**
11	Canned crab	**0.17**	**7.99**
12	Oyster sauce	0.00	0.00
13	Shrimp noodles	**0.11**	**0.70**
14	Shrimp cracker	**0.11**	**2.02**
15	rTM KP (Positive control)	19.54	170.00

The ELISA kit from company B only quantified six food samples out of ten possible samples containing TM. The negative controls were not detected and all food samples expected to contain TM from crustacean were quantified. Nevertheless, none of the food samples, expected to contain TM from molluscs, were quantified. The manufacturer states that this ELISA kit can be used for crustacean protein quantification, utilising crustacean protein standards and anti-crustacean-protein antibodies. Moreover, the ELISA kit shows minor cross-reactivity to mussels and other arthropods, but not other molluscs. Although the ELISA kit was specified to have slight cross-reactivity to mussels, the smoked canned mussel sample was not quantified. Nevertheless, it was observed that some of the food samples listed

as zero ppm in table 6.1 gave a positive reading, e.g. smoked canned mussel or vegetable soup, however, values were below the LOQ and LOD the manufacturer stated and therefore not considered as positive quantified sample.

Overall, the quantified TM in the food samples is quite different between the two investigated ELISA kits. Moreover, the analysis of food samples by SDS-PAGE and immunoblotting showed different results compared to the ELISAs. The major difference in the instructions from the company A and the company B are sample preparation, extraction buffer and dilution factor. However, in this chapter the sample extracts were the same for the food samples analysis. Therefore, to investigate if there is a different antibody binding with the dilution factor and buffers, the food sample extracts were prepared according to the manufacturer instruction and tested for the binding by the monoclonal anti-tropomyosin antibody and the polyclonal anti-crustacean antibody. Figure 6.3 summaries the results of different buffers and dilution factors analysed by dot-blot technique, thus as it was shown in figure 6.2, the food samples are highly processed and therefore the protein separation by SDS-PAGE was poor.

Figure 6.3 visualises that there is no real difference for the different dilutions factors and buffers systems used in the ELISA kits. The monoclonal anti-tropomyosin antibody detected smoked canned mussel, prawn and chicken noodles, canned smoked oyster, canned prawn, shrimp noodles and the positive control. The polyclonal anti-crustacean antibody detected prawn and chicken noodles, canned prawn, canned crab, shrimp noodles, shrimp crackers and the positive control. Interestingly, the monoclonal anti-tropomyosin antibody detected both mollusc samples, whereas the polyclonal anti-crustacean antibody did not detect any mollusc sample, thus being specific to crustacean samples. The binding site of the monoclonal anti-tropomyosin antibody was predicted[18] and will be explained in the discussion. None of the antibodies detected any of the selected negative controls, including the vegetable soup.

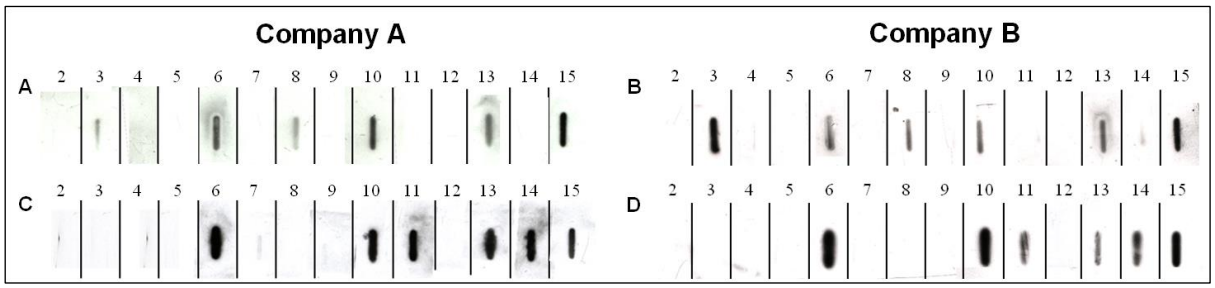


Figure 6.3: Dot blotting analysis of the food samples using two different dilution factors and buffers according to the ELISA kits from the company A and the company B. The food samples have been diluted using extraction buffer and dilution factor for the company A applying (A) the monoclonal anti-tropomyosin antibody and (C) the polyclonal anti-crustacean antibody. The food samples have been diluted using extraction buffer and dilution factor for the company B applying (B) the monoclonal anti-tropomyosin antibody and (D) the polyclonal anti-crustacean antibody. The numbers represent different food samples analysed: 2. Chicken muscle; 3. Smoked canned mussel; 4. Vegetable soup; 5. Crab noodles; 6. Prawn and chicken noodles; 7. Red Thai curry paste; 8. Smoked canned oyster; 9. Chicken soup; 10. Canned prawn; 11. Canned crab; 12. Oyster sauce; 13. Shrimp noodles; 14. Shrimp crackers; and 15. rTM KP.

The monoclonal anti-tropomyosin antibody detected more samples using the dot-blot technique compared to SDS-PAGE separation, whereas the polyclonal anti-crustacean antibody detected different samples using dot-blot technique, but is still specific to crustacean. Therefore one can assume that the highly processed food samples and resulting poor separation of proteins by SDS-PAGE can lead to an underestimated detection using the monoclonal anti-tropomyosin antibody.

6.3.1.3 Analysis of food samples by LC/qTOF

The 13 different food samples were also analysed using the four developed LC/qTOF methods (Chapter 3). Proteins identified in any of the four LC/qTOF methods were considered as results. Table 6.2 shows that for eight food samples analysed specific proteins were identified for the sample, whereas the identified protein with the highest obtained Mascot score is given. Interestingly, for chicken muscle the protein with the highest Mascot sore was myosin heavy chain, followed by tropomyosin, both matching the species *Gallus gallus domesticus* and not any

shellfish species or other species that might be shellfish related. Interestingly, the identified proteins in chicken muscle are similar to the proteins identified in the crustacean species in chapter 4. Nevertheless, allergenic tropomyosin, in detail crustacean TM was only identified for canned prawn, canned crab and the positive control. It is hypothesised that for some samples no sample specific proteins were identified due to high food processing, such as chicken soup or oyster sauce. Therefore it can be concluded that food processing denatures the proteins and protein specific peptides cannot be identified by LC/qTOF.

Table 6.2: Summary of the LC/qTOF results for the 13 different food samples. The results show if allergenic TM was identified and/or other food sample specific proteins were identified with the matched species by Mascot given in brackets. The samples highlighted in green are expected to contain highly processed allergenic tropomyosin (TM), whereas the samples highlighted in red are not expected to contain allergenic TM (negative controls). Moreover, the samples highlighted in yellow are unknown to contain TM.

Sample number	Sample name	Allergenic TM detected	Sample specific proteins identified
2	Chicken muscle	no	Myosin heavy chain (*Gallus gallus domesticus*)
3	Smoked canned mussel	no	Actin (including shellfish species)
4	Vegetable soup	no	no
5	Crab noodles	no	Glutelin type-A 1 (*Oryza sativa subsp. Japonica*)
6	Prawn and chicken noodles	no	2S albumin(*Glycine max*)
7	Red Thai curry paste	no	Defensin (*Capsicum annuum*)
8	Smoked canned oyster	no	Actin (including shellfish species)
9	Chicken soup	no	no
10	Canned prawn	TM (*Penaeus monodon*)	Actin (including shellfish species)
11	Canned crab	TM (*Chionoecetes opilio*)	Actin (excluding shellfish species)
12	Oyster sauce	no	no
13	Shrimp noodles	no	no
14	Shrimp cracker	no	no
15	rTM KP (Positive control)	TM (*Penaeus monodon*)	TM (*Penaeus monodon*)

6.3.1.4 Analysis of food samples by LC/MRM

Analysing the food samples using the validated LC/MRM method shows that allergenic TM is not detected in either of the negative controls. Moreover, table 6.3 reports that all samples which were expected to include allergenic TM have been quantified, this includes both mollusc samples, where the antibodies and the ELISA kits were not able to detected or quantify them. However, the TM concentration in the mollusc samples is rather low compared to the crustacean samples. Moreover, three samples that were unknown for the allergenic TM content were quantified, whereas three other samples were tested negative for allergenic TM.

The four selected peptides have been detected in the different samples as predicted in chapter 3 and confirmed in chapter 5. However, there is still a difference in the concentration between the four different peptides (Table 6.3). Nevertheless, the variation between concentrations is less than a factor of 2.5, which is lower than observed in chapter 5. One can assume this is due to the digestion method and therefore it is suggested that the digestion method should be investigated and optimised. The concentration calculated for the comparison of the methods will be based on peptide 1 for crustacean and peptide 3 for molluscs, as explained in detail in chapter 5.

Table 6.3: Summary of the LC/MRM results for TM quantification in the 13 different food samples. The results are shown in nanomolar (nM). The samples number and the sample names highlighted in green are expected to contain highly processed allergenic tropomyosin (TM), whereas the sample numbers and the names highlighted in red are not expected to contain allergenic TM (negative controls). Moreover, the samples number and the sample names highlighted in yellow are unknown to contain TM. (N.D. = not detected)

		Concentration calculated (nM)			
Sample number	**Sample name**	**Peptide 1**	**Peptide 2**	**Peptide 3**	**Peptide 4**
2	Chicken muscle	N.D.	N.D.	N.D.	N.D.
3	Smoked canned mussel	N.D.	N.D.	**22.98**	N.D.
4	Vegetable soup	N.D.	N.D.	N.D.	N.D.
5	Crab noodles	N.D.	N.D.	N.D.	N.D.
6	Prawn and chicken noodles	**233.28**	**110.88**	N.D.	**233.82**
7	Red Thai curry paste	N.D.	N.D.	N.D.	N.D.
8	Smoked canned oyster	N.D.	N.D.	**38.10**	N.D.
9	Chicken soup	N.D.	N.D.	N.D.	N.D.
10	Canned prawn	**1485.57**	**1051.29**	N.D.	**787.60**
11	Canned crab	**373.47**	**342.89**	N.D.	N.D.
12	Oyster sauce	N.D.	N.D.	N.D.	N.D.
13	Shrimp noodles	**9.84**	**13.81**	N.D.	**21.03**
14	Shrimp cracker	**7.82**	**6.84**	N.D.	**12.41**
15	rTM KP (Positive control)	**2167.31**	**1756.10**	N.D.	**3863.40**

6.3.2 Comparison of all methods applied for food sample analysis

The 13 different food samples and the positive control (rTM KP) have been analysed in the previous section applying different methods, however, not all of the methods can quantify the allergenic TM concentration. Table 6.4 summarises the results for the different samples analysed by the different methods. Overall, all methods were able to detect the positive control with strong antibody binding or high concentration. Moreover, with the exception of the company A and sample

vegetable soup, all known negative samples were not detected by any of these methods.

Table 6.4: Summary of the food samples detected by various methods, where black indicates strong antibody binding or high concentration by ELISA and MS, grey indicates low binding or low concentration and white indicated no binding or no concentration by ELISA and MS. The red boxes show false positive results, whereas the blue boxes show false negative results. The sample numbers and the names highlighted in green are expected to contain highly processed allergenic tropomyosin (TM), whereas the sample numbers and the names highlighted in red are not expected to contain allergenic TM (negative controls). Moreover, the samples and the names highlighted in yellow are unknown to contain TM. (mAb= monoclonal anti-tropomyosin antibody; pAb = polyclonal anti-crustacean antibody)

Detection of TM in food samples using different methods									
Sample number	Sample name	SDS-PAGE mAb	Dot-blot mAb	SDS-PAGE pAb	Dot-blot pAb	Company A	Company B	LC/qTOF	LC/MRM
2	Chicken muscle								
3	Smoked canned mussel								
4	Vegetable soup								
5	Crab noodles								
6	Prawn and chicken noodles								
7	Red Thai curry paste								
8	Smoked canned oyster								
9	Chicken soup								
10	Canned prawn								
11	Canned crab								
12	Oyster sauce								
13	Shrimp noodles								
14	Shrimp cracker								
15	rTM KP (Positive control)								

The food samples that were expected to contain allergenic crustacean tropomyosin (canned prawn and canned crab) were, with the exception of the monoclonal anti-tropomyosin and canned crab, detected by all methods applied (Table 6.4). However, the food samples that were expected to contain allergenic mollusc tropomyosin (smoked canned mussel and smoked canned oyster), were not identified by all methods investigated. In detail, smoked canned mussel was only detected by the monoclonal anti-tropomyosin antibody and the validated LC/MRM method. In contrast to smoked canned oyster, this was detected using the LC/MRM, the ELISA kit from company A and the monoclonal anti-tropomyosin antibody. It is suggested that in the smoked canned food samples allergenic tropomyosin is present, however, that the investigated antibodies and the antibodies used in the two ELISA kits are more specific to crustacean, not to molluscs, and therefore the mollusc samples were not detected by all methods investigated.

Table 6.4 shows that the samples that were unknown to contain TM were identified differentially by the various methods analysed. However, crab noodles and oyster sauce are negative for TM content by all the methods investigated. Therefore it is concluded that these two samples are not containing any allergenic TM, although the title of the food sample would imply this, in contrast to shrimp noodles and shrimp crackers, in which allergenic TM has been detected by most methods investigated. The LC/qTOF method did not identify allergenic TM for either of these two food samples, hence, LC/qTOF only indentified allergenic TM in two food samples. Shrimp noodles were detected by the monoclonal anti-tropomyosin antibody, but shrimp crackers were not. Overall, it is concluded that shrimp crackers and shrimp noodles include allergenic TM, thus most methods applied could detect allergenic TM.

Red Thai curry paste (Table 6.4) was only identified to contain TM by the ELISA kit from company A and the polyclonal anti-crustacean antibody. However, the polyclonal anti-crustacean antibody detected red Thai curry paste only with less accurate immunoblotting, not by dot blotting. Potentially it is possible that vegetable soup and red Thai curry paste contain allergenic TM, however, since all food samples were analysed by four different antibodies and two different mass spectrometry methods and been only detected by one/two antibodies, it is

concluded that these two food samples do not include allergenic tropomyosin. Moreover, the concentration calculated, using the ELISA kit from company A, for vegetable soup is the lowest point of the linear range and the value obtained for red Thai curry paste is at the end of the linear range, thus, it is concluded that the positive detection of red Thai curry paste and vegetable soup are false positive results and do not include allergenic TM.

The major differences between the company A and the company B are the range of the calibration curve and different values for the LOQ, which could explain why the company A seem to have false positives and the company B did not, thus the cut-off value for positive readings is higher for the company B. Overall, the absorbance for the different food samples measured are similar, however, the different dilution factors and the different ranges of the calibration curves lead to different results for the investigated food samples. Therefore, the ppm calculated for the various food samples are higher for the company B, ranging from factor 6 up to factor 26.

Overall, the LC/qTOF methods detected the least positive samples and the ELISA kit of company A the most, however, this includes two false positive. Interestingly, the detection of the ELISA kit from company B was identical to the polyclonal anti-crustacean antibody. The positive control (rTM KP) was detected for every method applied, and the negative controls were also correctly identified, with the exception of vegetable soup by company A. Overall, out of six unknown food samples, three are positive for allergenic tropomyosin content, namely prawn and chicken noodles, shrimp noodles and shrimp crackers. Whereas crab noodles, red Thai curry paste and oyster sauce do not contain allergenic TM. However, all four antibody based methods had difficulties to detect allergenic mollusc TM. This could be explained with the antibodies specificity, thus especially the polyclonal antibodies investigated are designed to be specific for crustacean and/or prawn TM and/or bind to whole unprocessed TM. Moreover, it is suggested that the LC/qTOF is not a very sensitive method for the identification of TM in highly processes food samples. The validated LC/MRM method was the only method investigated and compared that can distinguish and quantify allergenic crustacean TM or allergenic mollusc TM in all the analysed samples.

6.3.3 Comparison of quantitative methods applied for food sample analysis

With all methods tested and compared, only the ELISA kits and the LC/MRM methods are suitable to quantify shellfish allergens. To compare the different results obtained by the different methods, all values have been calculated into ppm per gram food sample. Table 6.5 and figure 6.4 summarise these results. The concentration calculated by all three methods for the positive control (rTM KP) is the highest for all samples analysed, thus rTM only includes allergenic TM from King prawn, and not any other proteins or food components. Based on calculated rTM KP concentration, the factor between company A and company B is a factor of 9, between company A and LC/MRM factor 650 and between company B and LC/MRM a factor of 5700. Therefore it is uncertain which method provides the most accurate results. However, as seen in figure 6.4, the results of company A and the LC/MRM methods are more similar than compared to company B. Therefore it is assumed that company A and the LC/MRM generate more accurate results, whereas company B overestimates the TM concentration. Nevertheless, more experiments including spiking experiments need to be carried out to determine which method generated the most accurate results.

However, besides the positive control (rTM KP) and canned prawn all the values calculated in ppm per gram sample seem relatively low for all methods investigated. This can be explained with the conversion of the units, thus in the previous sections it can be seen, especially for the LC/MRM results, that the quantification of samples based on the method specification is suitable. Generally, the samples containing most allergenic TM are canned prawn, canned crab, followed by prawn and chicken noodles. Interestingly, although sample canned prawn and canned crab only contained edible muscle tissue in the food protein extract, whereas prawn and chicken noodles contained manly noodles (starch), the concentration of prawn and chicken noodles is high. Hence, for both the ELISA kits the concentration is higher for prawn and chicken noodles compared to canned crab. A possible explanation for this could be that the ELISA kit antibodies are generated for prawn proteins and not for crab proteins, thus crabs and prawns share only around 90% amino acid homology, resulting in potential decreased

antibody binding. In contrast to the concentration calculated by LC/MRM which is higher for canned crab compared to prawn and chicken noodles.

Table 6.5: The concentration of TM calculated as ppm per gram of sample, applying the different quantification methods for shellfish allergens analysing the 13 different food samples. The samples highlighted in green are expected to contain highly processed allergenic tropomyosin (TM), whereas the samples highlighted in red are not expected to contain allergenic TM (negative controls). Moreover, the samples highlighted in yellow are unknown to contain TM.

		ppm per gram sample		
Sample Number	Sample Name	Company A	Company B	LC/ MRM
2	Chicken muscle	0.0	0.0	0.0
3	Smoked canned mussel	0.0	0.0	0.3
4	Vegetable soup	0.9	0.0	0.0
5	Crab noodles	0.0	0.0	0.0
6	Prawn and chicken noodles	30.8	1375.2	4.4
7	Red Thai curry paste	3.3	0.0	0.0
8	Smoked canned oyster	4.5	0.0	0.5
9	Chicken soup	0.0	0.0	0.0
10	Canned prawn	23.7	1249.2	25.6
11	Canned crab	7.1	668.6	6.5
12	Oyster sauce	0.0	0.0	0.0
13	Shrimp noodles	4.6	58.3	0.3
14	Shrimp cracker	6.5	184.4	0.4
15	rTM KP (Positive control)	19,540,000.0	170,000,000.0	29,852.7

Figure 6.4 displays the same results as summarised in table 6.5, for better visualisation of the different concentration calculated for the 13 food samples and rTM KP. Comparing the absolute quantified values, figure 6.4 visualise that the values generated by the company B are higher than the concentrations calculated by the company A and the validated LC/MRM method. Thus, one can assume that the company B potentially overestimates the concentration of allergenic TM present in the positive food samples.

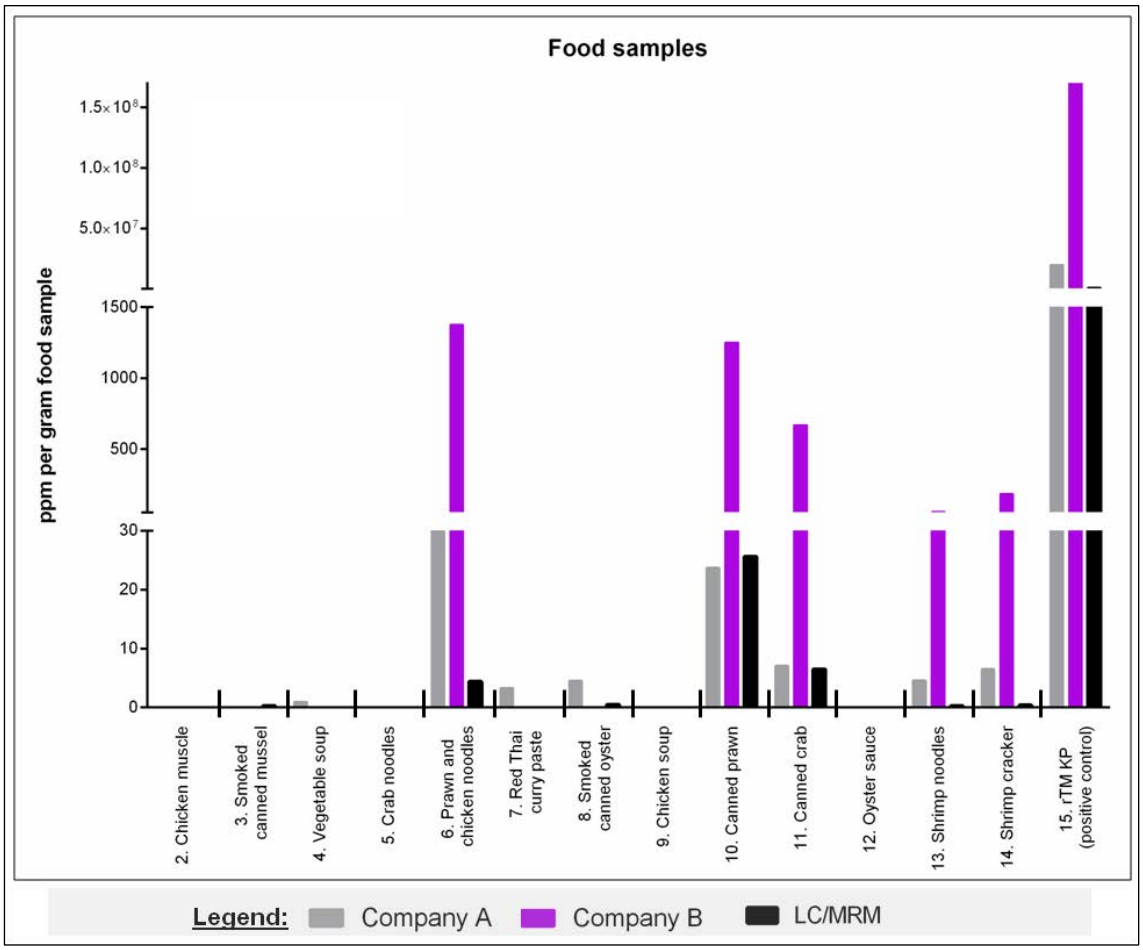


Figure 6.4: The concentration of TM calculated applying different quantification methods for TM analysing various food samples (ppm per g of food samples).

Chapter 6

Overall, the investigated antibodies, the ELISAs and the mass spectrometry methods are suitable to detect highly processed food samples. Therefore, one can assume that the antibody binding sites as well as the chosen quantification peptides are stable towards all food sample processing and the different influences of food matrices, thus TM can be detected and quantified. However, the antibody detection and the LC/qTOF identification are not fully suitable for the quantification of the major shellfish allergen TM.

The concentration calculated for the 13 food samples using the two ELISA kits are different, due to the different calibration curves and the dilution factors. Both the ELISA kits were able to quantify crustacean allergens in various samples. However, the company B did not detect any mollusc allergens and the company A

only detected one mollusc sample out of two analysed. This could be explained with the antibody specificity used in the different ELISA kits. Unfortunately, company A seemed to generate two false positive samples, namely vegetable soup and red Thai curry paste. The validated LC/MRM method was the only method investigated which was able to detect all allergenic tropomyosin from crustacean and mollusc samples and most importantly can distinguish between crustacean and molluscs. The concentrations calculated using the LC/MRM are similar to the values generated by the ELISA kit from the company A. Thus, one can assume that the company B overestimates the concentration of TM.

6.4 Discussion

Thirteen different food samples were analysed utilising antibody based methods (ELISA and immunoblotting) and mass spectrometry methods (LC/qTOF and LC/MRM). Three food samples were chose as negative controls, four samples should contain allergenic TM (two crustacean and two mollusc food samples) and six samples could be possible containing allergenic TM, based on the product name which included a relevant name to shellfish species.

Overall, the ability to detect shellfish allergens in the processed food samples is independent from the food processing and methods investigated, similar to Zhang et al.[7] Factors that can influence allergen detection and quantification mentioned in the literature, such as temperature treatment,[8] pressure treatment,[8, 12] possible occurring Maillard reaction[12, 13] (noodles), and lipids[8] (smoked canned samples) had no obvious effect for the detection of allergenic TM by the investigated methods. Although it is suggested that SDS-PAGE separation and immunoblotting are not suitable to detect allergen in highly processed food samples. As demonstrated by Yu et al., [10] the separation of highly processed food samples is insufficient by SDS-PAGE. Although, it was reported in the literature that food processing influences the stability of TM and decreases allergenicity,[8, 10, 11, 19] in this chapter it was demonstrated that the antibodies still can detect highly processed TM. Moreover, the peptides chosen for the LC/MRM quantification are stable towards food processing, as they still can be detected and quantified and the digestion efficiency was not decreased as reported by Nakamura et al.[12] Nevertheless, it is recommend that the applied digestion method in this thesis should be improved.

Comparing the results of the positive crustacean food samples by ELISA and LC/MRM it was found that the concentration for canned crab is considerably lower compared to canned prawn. Interestingly, although sample canned prawn and canned crab only contained edible muscle tissue in the food protein extract, whereas prawn and chicken noodles contained manly noodles (starch), the concentration of prawn and chicken noodles is higher compared to canned crab for both ELISA kits. A possible explanation for this could be that the ELISA kit antibodies are generated for prawn proteins and not for crab proteins, thus crabs (Cha f 1, *Portunus trituberculatus* (Genbank accession number ABS12234.1),

Scylla serrata (Genbank accession number ABS12233.1), *Chionoecetes opilio* (Genbank accession number A2V735.1) and *Portunus pelagicus* (Genbank accession number AGE44125.1)) and prawns (Mel I 1, Pen m 1 and Lit v 1) share only around 90% amino acid homology, potential leading to decreased antibody binding. In contrast to the concentration calculated by LC/MRM, which is higher for canned crab compared to prawn and chicken noodles. Surprisingly, the monoclonal anti-tropomyosin antibody was not able to detect canned crab, although it was shown by immunoblotting that the monoclonal antibody detects crab species.[18] Moreover, the amino acid identity between crabs and prawns in the predicted monoclonal antibody binding region (amino acid region from 9-19)[18] is 100%, with the exception of *Portunus pelagicus*, only sharing 90% amino acid homology.

However, the polyclonal anti-crustacean antibody and the commercial available ELISA kits investigated have difficulties to detect mollusc tropomyosin. This could be explained due to the specificity of the ELISA antibodies and the polyclonal anti-crustacean antibody[17] for crustaceans/prawns. As demonstrated by immunoblotting using different shellfish species and the monoclonal anti-tropomyosin antibody,[18] both mollusc food samples were detected. In comparison to prawns/crustacean, mussels and oysters have 63.64% amino acid homology in the predicted monoclonal anti-tropomyosin binding region,[18] whereas the overall TM amino acid homology between prawns (Mel I 1, Pen m 1 and Lit v 1) and mussels (*Perna viridis* (Genbank accession number Q9GZ70.1), *Mytilus galloprovincialis* (Genbank accession number P91958.1) and *Mytilus edulis* (Genbank accession number Genbank accession number Q25457.1)) is ~55% and ~60% for oysters (*Crassostrea gigas* (Genbank accession number BAH10152.1) and *Crassostrea virginica* (AAC61869.1)), respectively. This similar amino acid homology in this region can explain why the monoclonal anti-tropomyosin antibody can detect both food mollusc samples. However, considering the other antibodies (ELISAs and the polyclonal anti-crustacean antibody), it is not quite clear why the company A detected smoked canned oyster but not smoked canned mussel. The amino acid region where mussels and oysters differ the most for TM are from amino acid position 50-80, thus possible the antibody of the company A has a binding region to tropomyosin in this region.

Nevertheless, one can assume that the company A and the company B applied polyclonal antibodies in their ELISA kits (this information in not stated by the manufacturers), thus there should be more than one binding site for tropomyosin.

Overall, the investigated antibodies, the ELISAs and the mass spectrometry methods are suitable to detect highly processed food samples. Moreover, the positive control was detected by all methods investigated. The LC/qTOF methods detected the least allergenic tropomyosin, but could identify sample specific proteins in eight food samples. It is concluded that out of six unknown food samples, three contain allergenic tropomyosin, whereas the other three do not include allergenic tropomyosin. Moreover, the negative controls are negative and the food samples chosen to contain allergenic tropomyosin are positive. Nevertheless, the monoclonal anti-tropomyosin antibody, the polyclonal anti-crustacean antibody and the LC/qTOF method are not suitable for the general quantification of the shellfish allergens. For allergenic TM quantification the two commercial ELISA kits and the validated LC/MRM method can be utilised.

The ELISA kit from company A quantified shellfish allergens in 8 out of 13 food samples, including smoked canned oyster, one of the selected molluscs samples. However, the ELISA kit from company A quantified two false positives, namely vegetable soup and red Thai curry paste. The ELISA kit from company B quantified allergenic tropomyosin in 5 out of 13 food samples and could not detect allergenic tropomyosin from mollusc samples. Interestingly, the detection of the ELISA kit from company B was identical to the polyclonal anti-crustacean antibody. One can assume that the difficulties of the ELISAs and the polyclonal anti-crustacean antibody to detect allergenic mollusc TM can be explained by the fact, that the antibody was designed specifically to crustacean and/or prawn TM.[17, 18] The validated LC/MRM method was the only method investigated and compared that can distinguish and quantify allergenic crustacean tropomyosin or allergenic mollusc tropomyosin. In all seven food samples containing allergenic tropomyosin, tropomyosin could be quantified for crustacean or molluscs by LC/MRM.

The concentrations calculated are diverse for seven food samples containing allergenic tropomyosin and the positive control using the two ELISA kits. These findings are different to the results reported by Sakai et al.,[15] when comparing two commercial available ELISA kits in Japan. It is hypothesised that the different

values calculated for the company A and the company B are due to the different calibration curves and the dilution factors. The concentrations calculated by LC/MRM are similar to the values generated by the ELISA kit from company A. Thus, one can assume that the company B overestimates the concentration of TM. Therefore it is suggested that more food samples and recovery experiment should be carried out to better understand the different concentration generated for the different food samples. Nevertheless, it was shown in this chapter that highly processed food samples can be distinguished and quantified for allergenic crustacean or allergenic molluscs tropomyosin using the novel validated LC/MRM method developed in chapter 5. Therefore it is concluded that mass spectrometry, especially LC/MRM is a good alternative for ELISAs for the absolute quantification of food allergens.

6.5 References

1. Koeberl, M.; Clarke, D.; Lopata, A. L. (2014). *Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems.* Journal of Proteome Research. 13, (8), 3499-3509.

2. Lopata, A. L.; O'Hehir, R. E.; Lehrer, S. B. (2010). *Shellfish Allergy.* Clinical and Experimental Allergy. 40, (6), 850-858.

3. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* Int J Hyg Environ Health. 217, (7), 740-50.

4. Fuller, H. R.; Goodwin, P. R.; Morris, G. E. (2006). *An Enzyme-Linked Immunosorbent Assay (Elisa) for the Major Crustacean Allergen, Tropomyosin, in Food.* Food and Agricultural Immunology. 17, (1), 43-52.

5. Seiki, K.; Oda, H.; Yoshioka, H.; Sakai, S.; Urisu, A.; Akiyama, H.; Ohno, Y. (2007). *A Reliable and Sensitive Immunoassay for the Determination of Crustacean Protein in Processed Foods.* Journal of Agricultural and Food Chemistry. 55, (23), 9345-9350.

6. Werner, M. T.; Faeste, C. K.; Egaas, E. (2007). *Quantitative Sandwich Elisa for the Determination of Tropomyosin from Crustaceans in Foods.* Journal of Agricultural and Food Chemistry. 55, (20), 8025-8032.

7. Zhang, H.; Lu, Y.; Ushio, H.; Shiomi, K. (2014). *Development of Sandwich Elisa for Detection and Quantification of Invertebrate Major Allergen Tropomyosin by a Monoclonal Antibody.* Food Chemistry. 150, (0), 151-157.

8. Török, K.; Hajas, L.; Horváth, V.; Schall, E.; Bugyi, Z.; Kemény, S.; Tömösközi, S. (2015). *Identification of the Factors Affecting the Analytical Results of Food Allergen Elisa Methods.* European Food Research and Technology. 1-10.

9. Kamath, S. D.; Abdel Rahman, A. M.; Voskamp, A.; Komoda, T.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Effect of Heat Processing on Antibody Reactivity to Allergen Variants and Fragments of Black Tiger Prawn: A Comprehensive Allergenomic Approach.* Molecular Nutrition & Food Research. 58, (5), 1144-1155.

10. Yu, H.-L.; Cao, M.-J.; Cai, Q.-F.; Weng, W.-Y.; Su, W.-J.; Liu, G.-M. (2011). *Effects of Different Processing Methods on Digestibility of Scylla Paramamosain Allergen (Tropomyosin).* Food and Chemical Toxicology. 49, (4), 791-798.

11. Long, F.; Yang, X.; Wang, R.; Hu, X.; Chen, F. (2015). *Effects of Combined High Pressure and Thermal Treatment on the Allergenic Potential of Shrimp (Litopenaeus Vannamei) Tropomyosin in a Mouse Model of Allergy.* Innovative Food Science & Emerging Technologies. 29, (0), 119-124.

12. Nakamura, A.; Watanabe, K.; Ojima, T.; Ahn, D.-H.; Saeki, H. (2005). *Effect of Maillard Reaction on Allergenicity of Scallop Tropomyosin.* Journal of Agricultural and Food Chemistry. 53, (19), 7559-7564.

13. Nakamura, A.; Sasaki, F.; Watanabe, K.; Ojima, T.; Ahn, D.-H.; Saeki, H. (2006). *Changes in Allergenicity and Digestibility of Squid Tropomyosin During the Maillard Reaction with Ribose.* Journal of Agricultural and Food Chemistry. 54, (25), 9529-9534.

14. Mills, E. N. C.; Sancho, A. I.; Rigby, N. M.; Jenkins, J. A.; Mackie, A. R. (2009). *Impact of Food Processing on the Structural and Allergenic Properties of Food Allergens.* Molecular Nutrition & Food Research. 53, (8), 963-969.

15. Sakai, S.; Matsuda, R.; Adachi, R.; Akiyama, H.; Maitani, T.; Ohno, Y.; Oka, M.; Abe, A.; Seiki, K.; Oda, H.; Shiomi, K.; Urisu, A. (2008). *Interlaboratory Evaluation of Two Enzyme-Linked Immunosorbent Assay Kits for the Determination of Crustacean Protein in Processed Foods.* Journal of Aoac International. 91, (1), 123-129.

16. Koeberl, M.; Kamath, S. D.; Saptarshi, S. R.; Smout, M. J.; Rolland, J. M.; O'Hehir, R. E.; Lopata, A. L. (2014). *Auto-Induction for High Yield Expression of Recombinant Novel Isoallergen Tropomyosin from King Prawn (Melicertus Latisulcatus) for Improved Diagnostics and Immunotherapeutics.* J Immunol Methods. 415, (0), 6-16.

17. Kamath, S. D.; Thomassen, M. R.; Saptarshi, S. R.; Nguyen, H. M. X.; Aasmoe, L.; Bang, B. E.; Lopata, A. L. (2014). *Molecular and Immunological Approaches in Quantifying the Air-Borne Food Allergen Tropomyosin in Crab Processing Facilities.* International Journal of Hygiene and Environmental Health. (0).

18. Kamath, S. D.; Abdel Rahman, A. M.; Komoda, T.; Lopata, A. L. (2013). *Impact of Heat Processing on the Detection of the Major Shellfish Allergen Tropomyosin in Crustaceans and Molluscs Using Specific Monoclonal Antibodies.* Food Chemistry. 141, (4), 4031-9.

19. Huang, Y.-Y.; Liu, G.-M.; Cai, Q.-F.; Weng, W.-Y.; Maleki, S. J.; Su, W.-J.; Cao, M.-J. (2010). *Stability of Major Allergen Tropomyosin and Other Food Proteins of Mud Crab (Scylla Serrata) by in Vitro Gastrointestinal Digestion.* Food and Chemical Toxicology. 48, (5), 1196-1201.

6.6 Summary chapter 6:

Analysis of Food Samples for Allergenic Tropomyosin

- ✔ Thirteen different food samples (three negative controls, four positive controls and six unknown samples) were analysed by monoclonal anti-tropomyosin antibody, one in-house polyclonal anti-crustacean antibody, two commercial available ELISA kits, LC/qTOF and LC/MRM methods
- ✔ All methods can detect tropomyosin in highly processed food samples. The recombinant tropomyosin from King prawn was identified correctly by all methods investigated
- ✔ The monoclonal anti-tropomyosin antibody detected the allergenic crustacean tropomyosin and mollusc tropomyosin, whereas the polyclonal anti-crustacean antibody only detected allergenic crustacean tropomyosin
- ✔ The ELISA kit from company A quantified two false positive, but could quantify one mollusc sample. The ELISA kit from company B only quantified crustacean samples
- ✔ The LC/qTOF method identified eight samples with sample specific proteins, including two allergenic tropomyosins from crustacean samples
- ✔ The validated LC/MRM method quantified all seven positive food samples and could successfully distinguish between crustacean and mollusc samples
- ✔ The calculated concentration for the two ELISA kits and the LC/MRM method are different. The ELISA kit from company A and the LC/MRM generated similar concentrations, whereas the ELISA kit from company B seemed to overestimate the allergen concentration

Highly processed food samples can be quantified for their allergenic tropomyosin content applying the ELISAs or the validated LC/MRM method, whereas the latter is the only method that can distinguish correctly between crustacean and molluscs.

CHAPTER 7

GENERAL CONCLUSION AND FUTURE DIRECTIONS

The shellfish consumption and exposure is increasing worldwide and therefore shellfish allergy is becoming a public health concern. Shellfish involves the crustacean group and the mollusc group, whereas both groups can be further subdivided and include a large number of different species. The food labelling legislation in Canada and the European Union requires a different labelling for allergens from crustacean and molluscs, to protect individual allergic consumers. However, current available methods for the detection of shellfish allergens cannot distinguish between those two groups. The existing commercial and non commercial detection methods for shellfish allergens are based on tropomyosin, the major allergen in shellfish. Moreover, current methods are mainly based on antibody detection, applying enzyme-linked immunosorbent assay (ELISA) technique. However, due to amino acid similarity of tropomyosin from shellfish and resulting cross-reactivity of antibodies, ELISAs cannot distinguish between crustacean and molluscs. Mass spectrometry has been recently applied for allergen food identification, detection and quantification to overcome the disadvantages of ELISA techniques.

For the absolute quantification of tropomyosin from crustacean or molluscs, fulfilling the demand of legislation and industry, a sensitive mass spectrometry (LC/MRM) method was developed and validated in this PhD thesis. To develop the LC/MRM method 22 different shellfish species, commonly consumed in Australia, were analysed by mass spectrometry (LC/qTOF) in chapter 3 and 4. The positive control, recombinant tropomyosin from the King prawn (*Melicertus latisulcatus*) was generated in chapter 2. The quantification method by LC/MRM was validated in chapter 5 and the validated LC/MRM method was applied in chapter 6 analysing 13 different food samples for tropomyosin.

King prawn (*Melicertus latisulcatus*) is a commercially important Australian prawn species, however, was not previously investigated. Therefore in chapter 2 the novel tropomyosin sequence and myosin light chain sequence from King prawn were investigated and compared to the well investigated related prawn species, Black Tiger prawn (*Penaeus monodon).* The results showed that the two tropomyosins had different amino acid sequences, in contrast to the myosin light chains which are almost identical. The tropomyosins have identical structure, but the IgE binding capability of allergenic patient's sera is different. Overall, it was

demonstrated that two related prawn species can express different allergens, thus species specific investigation of various allergens is necessary to better understand cross-reactivity and allergic reactions.

Crustacean species are more frequently investigated in the literature compared to mollusc species. This is reflected in the different shellfish species analysed by mass spectrometry in the literature with eight crustacean species and one mollusc species. To increase the panel of analysed shellfish species, twenty-two commonly consumed shellfish species in Australia were analysed by mass spectrometry (Chapter 3-5). In detail, eleven crustacean species and eleven mollusc species were analysed by LC/qTOF, where Black Tiger prawn, Vannamei prawn and Banana prawn are the only species that was previously investigated by mass spectrometry. Moreover, raw and whole heated extracts have been analysed for the 22 shellfish species, thus it was previously shown and confirmed in chapter 2, that heat treatment can influence the allergenicity of allergens.

The 22 shellfish species have been analysed by LC/qTOF in chapter 3 and 4. Whereas in chapter 3 the focus was on tropomyosin, the major allergen in shellfish, thus for the development of the absolute quantification of tropomyosin to distinguish crustacean and molluscs, signature peptides are required. Overall, with one exception, tropomyosin was identified in all whole heated extracts by LC/qTOF. However, the identification based on Mascot and matched peptides for tropomyosin varied for the different species. Overall, crustaceans were identified with higher certainty, originating from higher Mascot scores and more matched tryptic peptides for the different tropomyosin sequences as in comparison to molluscs. It is suggested that this might be due to the fact that mollusc species are less investigated and the amino acid sequence of tropomyosin varies more, thus cannot be identified using Mascot. Moreover, the amino acid sequences of mollusc tropomyosins vary more within the mollusc group as compared to the amino acid sequences within the crustacean group.

Fourteen tryptic peptides were identified within various species and therefore *in silico* aligned with 106 published tropomyosin sequences (GenBank) to be able to identify the best possible peptides for the absolute quantification of crustacean tropomyosin or mollusc tropomyosin. Overall, four peptides were identified for tropomyosin to distinguish shellfish subgroups. Based on *in silico* data peptide 1 is

unique for crustacean and peptide 3 is unique for molluscs. Peptide 2 is present in crustacean and cephalopods, but not in bivalves and peptide 4 is present in prawns and lobster, but not in krill and crabs, whereas the latter shows to have exceptions.

The analysis of 22 shellfish species utilising raw and whole heated extract by LC/qTOF demonstrated that the identified profile pattern for crustacean and molluscs as well as untreated and heat treated samples are different. Overall, 32 different proteins were identified. The main protein identified in raw crustacean was arginine kinase, whereas in whole heated extracts it was tropomyosin. In raw and whole heated mollusc extracts three main proteins were identified, namely actin, arginine kinase and tropomyosin. Moreover, the 32 different identified proteins have been analysed for their potential allergenicity, whereas 16 identified proteins have been previously reported as being shellfish allergens. The analysis of 32 identified proteins in chapter 4 confirmed that 15 proteins out of 16 reported proteins are allergenic proteins, however, it could not be confirmed that actin is a possible allergen. Furthermore, eight other identified proteins could be potential allergens, which have not been previously reported.

The four identified peptides from chapter 3 were applied for the development and validation of the LC/MRM method, following the validation criteria of the Department of Health Human Services Food and Drug Administration using the Guidance for Industry for Bioanalytical Method Validation. As predicted *in silico*, the four peptides were quantified in 22 shellfish species utilising raw and whole heated extracts, with four species exempted. Therefore a maximum of three peptides and a minimum of one peptide for different species were quantified in the validated LC/MRM method. The concentration of tropomyosin in the different shellfish species calculated by the validated LC/MRM method was similar to the tropomyosin identification by LC/qTOF. Overall, the concentration of tropomyosin is higher in whole heated extracts compared to raw extracts. Moreover, in whole heated crustacean the concentration of tropomyosin is higher compared to whole heated molluscs, potential confirming that the overall main protein in molluscs is actin. However, the tropomyosin concentration in raw extracts from crabs, lobsters and molluscs is similar, compared to the lower tropomyosin concentration in raw prawns.

The overall results of protein and allergen identification by LC/qTOF showed that the identification of proteins must be carefully investigated. Hence, it was observed that some protein matches are based on protein fragments and not the whole protein sequence. Moreover, it is important to confirm which species the identified protein matches, thus the species of interests are not always included. This could be due to unknown sequences or potential false positive peptide matches and needs to be carefully considered. Moreover, it is assumed that the tropomyosin identification by LC/qTOF underestimated the concentration in several mollusc species due to sequence unavailability. More tropomyosin originating from molluscs need to be sequenced and included into the database to allow better and more specific quantification of molluscs allergens using the detailed mass spectrometry approach.

The validated LC/MRM method was applied in chapter 6 for the quantification of allergenic tropomyosin in highly processed food samples. Moreover, the LC/MRM method was compared with two commercial available ELISA kits, to confirm that both methods can detect and quantify highly processed tropomyosin. Overall, both methods can detect allergenic tropomyosin, whereas the ELISAs had difficulties in quantifying mollusc samples. Moreover, the ELISAs can certainly not distinguish crustacean and mollusc in food samples. The concentrations quantified for the food samples varied for the LC/MRM methods and the two ELISA kits. However, the results of one ELISA kit were similar to the concentrations quantified by LC/MRM, thus it can be assumed that the other ELISA kit overestimated the concentration of allergenic tropomyosin in food samples.

Overall, in this PhD thesis a novel quantitative LC/MRM method was developed and validated to distinguish crustacean and mollusc species. The validated LC/MRM method was successfully applied for the quantification of allergenic crustacean tropomyosin and allergenic mollusc tropomyosin in 22 different shellfish species and for 13 highly processed food samples. Therefore it was demonstrated that the quantification of tropomyosin by LC/MRM is a suitable alternative to currently existing antibody based methods, such as ELISAs. The work presented in this thesis provides an important contribution towards the detection and quantification of allergenic tropomyosin from crustacean and molluscs, to fulfil the international legislation requirements.

Future directions

A) Species and allergens

It is strongly recommended that more crustacean species and mollusc species should be investigated for all possible proteins and especially reported allergens. It was also demonstrated in this PhD thesis that known allergens for closely related species can result in different allergenicity for allergic patients. Moreover, although some allergens might be very similar in related species, other allergens can be different. Furthermore, many proteins, which have been reported as allergens in crustacean, can also be identified in mollusc species, although not being reported as allergens. Moreover, some proteins which have not been previously reported as allergens could be potential allergen as demonstrated in this thesis. Therefore, further work and experiments are necessary to confirm that identified potential allergens are actual allergens. Moreover, actin should be investigated for possible allergenicity, although having high amino acid homology with the human analogue.

Furthermore, it is proposed that the reported cross-reactive allergen arginine kinase should be investigated more carefully, as it was shown that it is the main protein in raw crustacean and one of the three main proteins identified in raw and whole heated molluscs. It is recommended to investigate the different allergenicity of crustacean and molluscs more carefully, thus crustacean allergic patients could potentially be able to consume molluscs and vice versa. Moreover, it is suggested that the raw prawn extracts should be evaluated for the tropomyosin content as it was observed that the concentration of tropomyosin in raw prawns was very low and in raw King prawn extract tropomyosin could not be detected at all by mass spectrometry.

B) Mass spectrometry

For the identification of allergens and proteins by mass spectrometry it is strongly recommend that identified proteins, using thorough database searches, need to be carefully investigated. Moreover, it is suggested that some guidelines should be developed and established for the identification of allergens by mass spectrometry. This should include the amount of peptides and the quality of peptides that need to be identified for suitable protein identification. The amount of

peptides detected per protein might vary with the size of the protein. Furthermore it is proposed that selection criteria should be developed for peptides that match more than one protein and/or more than one species. A list of proteins that are contaminations for different food allergens analysed would be beneficial.

For the quantification of allergens and proteins by mass spectrometry it is strongly recommend, that there should be criteria on how many peptides per allergen need to be selected and how many transitions per peptide should be monitored, possible stating that less specific transitions such as b 1-ions, y 1-ions should not be selected. For the LC/MRM method validation a standard is necessary. It is suggested having a cheaper internal standard than isotopic labelled peptides. For example, using tryptic peptides from BSA as internal standards could be possible. These peptides could be monitored by every mass spectrometer, every application and every laboratory. However, using peptides derive from BSA as reference material for method validation and accuracy would need to be intensively investigated and reviewed. For allergen quantification the amount of peptides to be selected could vary with the size of the allergen and amino acid homology between allergens. If the aim of the research is to find species specific signature peptides and one signature peptide would be sufficient, however, more peptides from this allergen should be selected. These selection criteria should be universal applied for all allergen quantifications. Nevertheless, this would make the LC/MRM allergen quantification more expensive.

In terms of the validated quantitative LC/MRM method in this thesis, it is advised that the digestion method for the validated LC/MRM method should be optimised, thus it was observed that some samples showed insufficient digestion with the current method. Moreover, species that included or excluded unpredicted peptides should be investigated and reanalysed to find the origin of theses incorccrectly indentified peptides. Furthermore, all species and food samples analysed are based on one extract. Therefore it is recommended analysing different samples and extracts of the same species to investigate variation in tropomyosin content and/or expression. Furthermore, more shellfish species and food samples should be analysed by the validated LC/MRM method and different ELISA kits, to confirm the selectivity, reproducibility and accuracy of LC/MRM method as compared to ELISAs.

The experimental work and data analysis presented in this thesis confirms that mass spectrometry can be used as an alternative method to existing methods for food allergen detection and quantification. Moreover, the validated LC/MRM method can quantify allergenic tropomyosin from crustacean and tropomyosin from molluscs. Hence, the results of this PhD thesis provide an important contribution towards the detection and quantification of allergenic tropomyosin from crustacean and molluscs, to fulfil the international legislation requirements.

APPENDIX A

BUFFERS AND SOLUTIONS

All buffers were sterile filtered (0.45 µm) (Millipore, USA) after preparation. As water ddH_2O was used, which has been sterile filtered (0.45 µm) and autoclaved before making up buffers and solutions.

A1.1 General buffers

Phosphate buffered saline (PBS), volume 1 litre:

NaCl – 8 g (Thermo Fisher Scientific, Australia)

KCl – 0.2 g (Thermo Fisher Scientific, Australia)

Na_2HPO_4 – 1.44 g (Thermo Fisher Scientific, Australia)

KH_2PO_4 – 0.24 g (Thermo Fisher Scientific, Australia)

Adjust pH to 7.4 (Mettler-Toledo AG, Germany)

Store at 4°C until use.

PBS-T, volume 1 litre:

PBS - 1 l

Tween-20 – 0.5 ml (Sigma-Aldrich, USA)

A1.2 SDS-PAGE solutions

Acrylamide (Solution A)

Acrylamide Bis-acrylamide, 29:1, 40% Solution (Merck, USA)

Store at 4°C until use.

Solution B, volume 100 ml:

2 M Tris-HCl, pH 8.8 – 75 ml (Thermo Fisher Scientific, Australia)

10% Sodium dodecyl sulfate (SDS) in H_2O – 4 ml (Sigma-Aldrich, Australia)

H_2O – 21 ml

Store at 4°C until use.

Solution C, volume 100 ml:

1 M Tris-HCl, pH 6.8 – 50 ml

10% SDS in H_2O – 4 ml

H_2O – 46 ml

Store at 4°C until use.

12% SDS-PAGE gel recipe:

Resolving gel, volume 20 ml (4 minigels):

Solution A – 6 ml

Solution B – 5 ml

H_2O – 8.9 ml

10% Ammonium persulphate – 100 µl (Sigma-Aldrich-Aldrich, Australia)

Tetramethylethylenediamine (TEMED) – 10 µl (Sigma-Aldrich, Australia)

Stacking gel, volume 10ml (4 minigels):

Solution A – 0.93 ml

Solution C – 2.5 ml

H_2O – 6.5 ml

10% Ammonium persulphate – 100 µl

TEMED – 10 µl

5X Protein sample loading buffer, volume 10 ml:

1 M Tris-HCl, pH 6.8 – 0.6 ml

50% Glycerol – 5 ml (Thermo Fisher Scientific, Australia)

10% SDS – 2 ml

1 M Dithiothreitol (DDT) – 500 µl (Sigma-Aldrich, Australia)

1% Bromophenol blue – 1 ml (Bio-Rad, Australia)

Stored at -80°C until use.

1X Gel Electrophoresis running buffer, volume 1 litre:

Tris – 3 g

Glycine – 14.4 g (Sigma-Aldrich, Australia)

SDS – 1 g

Adjust to pH 8.3

Store at 4°C until use.

Coomassie staining solution, volume 100 ml:

Coomassie blue (R-250) – 1 g (Bio-Rad, Australia)

Methanol (AR grade) (MeOH) – 20 ml (Thermo Fisher Scientific, Australia)

Glacial acetic acid – 10 ml (Thermo Fisher Scientific, Australia)

H_2O – 70 ml

Store at room temperature until use.

SDS-PAGE gel destaining solution, volume 1 litre:

MeOH – 500 ml

Glacial acetic acid – 100 ml

H_2O – 400 ml

Store at room temperature until use.

A1.3 Immunoblotting buffers

Transfer buffer, volume 200 ml:

Tris – 1.164 g

Glycine – 0.58 g

10% SDS – 750 µl

MeOH – 40 ml

Blocking buffer, volume 100 ml:

5% Skimmed milk powder in PBS-T– 5 g (Local Supermarket, Australia)

Antibody dilution buffer, volume 100 ml:

1% Skimmed milk powder in PBS-T – 1 g

Dot-blot coating buffer, pH 9.6, volume 30 ml:

30 mM Na_2CO_3 – 15 ml (Sigma-Aldrich-Aldrich, Australia)

170 mM $NaHCO_3$ – 15 ml (Sigma-Aldrich-Aldrich, Australia)

A1.4 Protein purification buffers

Natural tropomyosin purification:

30 mM sodium acetate buffer, pH 5.5 (Buffer A), volume 1litre:

Glacial acetic acid – 0.27 g

Sodium acetate (trihydrate) – 3.46 g (Sigma-Aldrich, Australia)

30 mM sodium acetate, 1 M NaCl, pH 5.5 (Buffer B), volume 1 litre:

Glacial acetic acid – 0.27 g

Sodium acetate (trihydrate) – 3.46 g

NaCl – 58 g

Recombinant tropomyosin purification:

25 mM Tris–HCl, 300 mM NaCl, pH 8 (Buffer A), volume 1 litre:

Tris – 3 g

NaCl – 17.5 g

25 mM Tris–HCl, 300 mM NaCl, 500 mM Imidazole, pH 8 (Buffer B), volume 1 litre:

Tris – 3 g

NaCl – 17.5 g

Imidazole – 34 g (Sigma-Aldrich, Australia)

A1.5 Molecular biology solutions

Luria-Bertani (LB) medium, volume 1 litre:

Tryptone – 10 g (Sigma-Aldrich, Australia)

NaCl – 10 g

Yeast extract – 5 g (Thermo Fisher Scientific, Australia)

Adjust to pH 7.4 and autoclave at 121°C for 15 min. Store at 4°C until use.

LB agar plates, volume 1 litre:

Tryptone – 10 g

NaCl – 10 g

Yeast extract – 5 g

Agar – 15 g (Sigma-Aldrich, Australia)

Adjust to pH 7.4 and autoclave at 121°C for 15 min, cool to 50°C and aseptically add ampicillin stock (1 ml) before pouring into Petri-plates. Plates are stored at 4°C until use.

Ampicillin stock, volume 1 ml:

Ampicillin – 100 mg (Amresco, USA)

Store at -20°C until use.

Glycerol stocks, volume 1 ml:

Fresh overnight culture – 700 µl

50% sterile glycerol – 300 µl (Thermo Fisher Scientific, Australia)

Stored at -80°C until use.

Cell lysate extraction buffer, volume 1 litre:

25 mM Tris–HCl, 300 mM NaCl, pH 8

Tris – 3 g

NaCl – 17.5 g

ZY MEDIA, volume 1 litre:

Tryptone – 10 g

Yeast extract – 5 g

Adjust to pH 7.4 and autoclave at 121°C for 15 min. Store at 4°C until use.

20X NPS, volume 1 litre:

$(NH_4)_2SO_4$ – 66 g (Sigma-Aldrich, Australia)

KH_2PO_4 – 136 g

Na_2HPO_4 – 142 g

Stir the components well until dissolved. Adjust to pH 6.8 and autoclave at 121°C for 15 min. Store at 4°C until use.

50X5052, volume 1 litre:

Glycerol – 250 g

Glucose – 25 g (Sigma-Aldrich, Australia)

Alpha-lactose monohydrate – 100 g (Sigma-Aldrich, Australia)

Stir the components well until dissolved. Adjust to pH 6.8 and autoclave at 121°C for 15 min. Store at 4°C until use.

1 M $MgSO_4$, volume100 ml:

$MgSO_4 \cdot 7H_2O$ – 24.65 g (Sigma-Aldrich, Australia)

Stir the components well until dissolved. Sterile filter and store at 4°C until use.

0.1 M FeCl3 • 6H2O in 0.1 M HCl 100 ml, volume 100 ml:

$FeCl_3 \cdot 6H_2O$ – 2.7 g (Sigma-Aldrich, Australia)

0.1 M HCl – 0.36 g (Sigma-Aldrich, Australia)

Stir the components well until dissolved. Sterile filter and store at 4°C until use.

40% Glucose, volume 100 ml:

Glucose – 40g

Stir the components well until dissolved. Adjust to pH 6.8 and autoclave at 121°C for 15 min. Store at 4°C until use.

Antifoam in MeOH, volume 10 ml:

Antifoam – 1ml (Sigma-Aldrich, Australia)

MeOH – 9 ml

Store at 4°C until use.

ZY–0.8G, volume 1 litre:

ZY – 930 ml

1 M $MgSO_4$ – 1 ml

0.1 M $FeCl_3 \cdot 6H_2O$ in 0.1 M HCl – 0.5 ml

40% Glucose – 20 ml

20X NPS – 50 ml

Ampicillin (100 mg/ml) – 0.5 ml

Antifoam in MeOH – 10 µl

Stored at 4°C until use.

ZY–5052, volume 1 litre:

ZY media – 928 ml

1 M $MgSO_4$ – 1 ml

0.1 M $FeCl_3 \cdot 6H_2O$ in 0.1 M HCl – 0.5 ml

50X 5052 – 20 ml

20X NPS – 50 ml

Ampicillin (100 mg/ml) – 0.5 ml

Antifoam in MeOH – 10 µl

Store at 4°C until use.

A1.6 Mass spectrometry buffers

0.1% formic acid, volume 1 litre:

Formic acid – 1 ml (Thermo Fisher Scientific, Australia)

Store at room temperature until use.

0.1% acetonitrile, volume 1 litre:

Acetonitrile (ACN) – 1 ml (Thermo Fisher Scientific, Australia)

Store at room temperature until use.

1 M DDT, volume 1 ml:

DTT – 154.25 mg (Sigma-Aldrich-Aldrich, Australia)

Store at -20°C until use.

2 M iodoacetamide, volume 1 ml:

Iodoacetamide – 369.92 mg (Sigma-Aldrich, Australia)

Store at -20°C until use.

100 mM Sodium format:

Sodium format – 6.8 g (Sigma-Aldrich, Australia)

Store at 4°C until use.

Leucine Enkephalin, volume 30 ml:

0.1% formic acid – 15 ml

ACN + 0.1% formic acid – 15 ml

Leucine Enkephalin – 30 µl (Waters Corporation, Australia)

8 M urea

Urea – 480.48 mg (Sigma-Aldrich, Australia)

Store at -20°C until use.

RapiGest, volume 1 ml:

RapiGest – 1 mg (Waters Corporation, Australia)

Store at -20°C until use.

APPENDIX B

SUPPLEMENTARY TABLES AND FIGURES

B.1 Supplementary tables

Table B1.1: Alphabetical summary of all food allergens currently analysed by different MS systems and their peptides published in the literature. "*" indicates recommendation use as a signature peptide. Allergen name in brackets confirms the registration with the International Union of Immunological Societies (IUIS). LOD and LOQ given are provided in ppm or as published. (C# carbamidomethylated cysteine; N.D= Not Determined; Ref.= References)

Allergen (registered allergen)	Peptides identified (*recommended signature peptides)	Species/ Allergen source	LOD (LOQ)	MS system used	Ref.
Crustacean					
α-Actin	• AGFAGDDAP • AVFPSIVGRPR • DAYVGDEAQSKR • RGILTLK • IAPEESPVLLTEAPLNPK • TTGIVLDTGDGVTHTVPIYEGYC LPHAILR • LDLAGRDLTAYLTK • GYSFTTTAEREIVR • SYELPDGQVITIGNER • CDIDIRK • KDLFANNVLSGGTTMYPGIADR • EITALAPPTIK • IKIIAPPER • IIAPPERK • EEYDESGPGIVHR	*(Chionoec etes opilio)*	N.D.	MALDI, LC/qTo F and LC/MR M	Abdel Rahman et al, (2011). Biomolecul ar Characteriz ation of Allergenic Proteins in Snow Crab (Chionoece tes Opilio) and De Novo Sequencin g of the Second Allergen Arginine Kinase Using Tandem Mass Spectromet ry. Journal of Proteomics . 74, (2), 231-241..
Actin (isoforms)	• VAPEEHPVLLWEAPLNPK • DITNYLGK • SYELPDGQVITISNER • GYSFTTTAER • EEYDESGPGIVHRK • SYELPDGQVITISNER • AVFPSIVGR • EGYSFTTTAER • SYELPDGQVITIGNER • EITGLAPSSIK • EITALAPSSIK	*(Pandalus borealis)* Crude extract		MALDI, LC/qTo F and LC/MR M	Abdel Rahman et al.. (2013). Comprehe nsive Proteomics Approach in Characteriz ing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupation al Asthma Prevention. Journal of Proteome Research. 12, (2), 647-656..
Arginine kinase (Pen m 2)	• FLQAANAC#R • GTRGEHTEAEGGIYDISNK	*(Penaeus monodon)*	N.D.	MALDI and IT	Ortea et al. (2009). Arginine Kinase Peptide Mass Fingerprinti ng as a Proteomic Approach for Species Identificatio n and Taxonomic Analysis of Commercia
Arginine kinase (Lit v 2)	• FLQAANAC#R • GTRGEHTEAEGGIYDISNK	*(Litopena eus vannamei)*	N.D.	MALDI and IT	
Arginine kinase	• FLQAANAC#R • GTRGEHTEAEGGIYDISNK	*(Fennero penaeu s*	N.D.	MALDI and IT	

		indicus)			lly Relevant Shrimp Species. Journal of Agricultural and Food Chemistry. 57, (13), 5665-5672.
	• FLQAANAC#R • GTRGEHTEAEGGIYDISNK	*(Farfante penaeus notialis)*	N.D.	MALDI and IT	
	• FLQAANAC#R • GTRGEHTEAEGGIYDISNK	*(Pleoticus muelleri)*	N.D.	MALDI and IT	
	• FLQAANAC#R • LVDDHFLFVSGDR	*(Pandalus borealis)*	N.D.	MALDI and IT	
Arginine kinase (Pen m 2)	• *AVFDQLKEK • *VSSTLSSLEGELK • *TFLVWVNEEDHLR • *LEEVAGKYNLQVR	*(Penaeus monodon)*	N.D.	LC/IT	Ortea et al. (2011). Selected Tandem Mass Spectrometry Ion Monitoring for the Fast Identification of Seafood Species. Journal of Chromatography A. 1218, (28), 4445-4451.
Arginine kinase (Lit v 2)	• *VSSTLSSLEGELK • *TFLVWVNEEDHLR • *LEEVAGKYNLQVR	*(Litopenaeus vannamei)*	N.D.	LC/IT	
Arginine kinase	• *LTNAVNEIEKR	*(Pleoticus muelleri)*	N.D.	LC/IT	
	• *ALFDQLKDKK • *TFLVWVNEEDHLR • *LEEVAGKYNLQVR	*(Fennero penaeus merguiensis)*	N.D.	LC/IT	
	• *SFLVWVNEEDQLR	*(Pandalus borealis)*	N.D.	LC/IT	
	• *TFLVWVNEEDHLR • *LEEVAGKYNLQVR	*(Fennero penaeus indicus)*	N.D.	LC/IT	
	• *VSSTLSSLEGELK • *TFLVWVNEEDHLR	*(Farfante penaeus notialis)*	N.D.	LC/IT	
Arginine kinase	• *LVSAVNEIEK	*(Chionoecetes opilio)* This signature peptide can be used for arginine kinase from other crustacean species	N.D.	MALDI, LC/qToF and LC/MRM	Abdel Rahman et al. (2011). Biomolecular Characterization of Allergenic Proteins in Snow Crab (Chionoecetes Opilio) and De Novo Sequencing of the Second Allergen Arginine Kinase Using Tandem Mass Spectrometry. Journal of Proteomics. 74, (2), 231-241.

Arginine kinase	• AVFDQLKEK • VSSTLSSLEGELK • GTYYPLTGMSK • LIDDHFLFK • IISMQMGGDLGQVFRR • LTSAVNEIEKR • IPFSHHDR • GTRGEHTEAEGGIYDISNK	*(Penaeus monodon)*	N.D.	MALDI and LC/qTo F	Abdel Rahman et al. (2010). Analysis of the Allergenic Proteins in Black Tiger Prawn (Penaeus Monodon) and Characterization of the Major Allergen Tropomyosin Using Mass Spectrometry. Rapid Communications in Mass Spectrometry. 24, (24), 3624-3624.
Arginine kinase (isoforms)	• *QQLVDDHFLFVSGDR • SIDGFGLSPGITK • VGVENLMK • KQLVDDHFLFMSGDR • NLQVAGMERDWPEGR • GIFHNAEK • TFLVWVVEEDQLR • TFLVWLVEEDQLR • TFLVWTVEEDHLR • AVGDSVK • YGYVHSCPTNLGTGMRASVHVD LPGTWK • CEELKVQPR • KHRLGYSEVELVQCMIDGVNTL YAEDVALQKK	*(Pandalus borealis)* Synthetic and isotopcically labeled synthetic signature peptide	0.25 nM	MALDI, LC/qTo F and LC/MR M	Abdel Rahman et al. (2013). Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention. Journal of Proteome Research. 12, (2), 647-656
Glyceraldehyde-3-phosphate dehydrogenase	• GIDGFGR • HVYNEMKPENIPWSK • GAGQNIIPSSTGAAK • AGAHMKGGAK • AGAEYIVESTGVFTTIEK • AGAHMKGGAK • LTGMAFR • VPTPDVSVVDLTVR • AGIQLSK	*(Pandalus borealis)* Crude extract		MALDI, LC/qTo F and LC/MR M	
Myosin light chain (Lit v 3)	• KGGXNVFDMFTQK • SSGESDDDDVVAASIR	*(Litopenaeus vannamei)*	N.D.	LC/MALDI	Ayuso et al. (2008). Myosin Light Chain Is a Novel Shrimp Allergen, Lit V 3. Journal of Allergy and Clinical Immunology. 122, (4), 795-802.
Myosin light chain	• EGFQLMDR	*(Pandalus borealis)*	N.D	MALDI, LC/qTo F and LC/MR M	Abdel Rahman et al. (2013). Comprehensive

Myosin heavy chain	• LTQEAVADLER • ELQARIEEL • LDEAGGATSAQIELNK • DEAGGATSAQIELNKKR • DLKLTQEAV • DLLRQLEEA • ELQARIEEL • ELSQVRQEI • LTQEAVADLER • QIEEAEEIAALNLAK	*(Pandalus borealis)* Crude extract	N.D	MALDI, LC/qTo F and LC/MR M	Proteomics Approach in Characteriz ing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupation al Asthma Prevention. Journal of Proteome Research. 12, (2), 647-656
Sarcoplasm ic Ca-binding protein (Lit v 4)	• YMYDIDDDGFLDK • NDFECLAVR • GEFSAADYANNQK • NLWNEIAELADFNKDG • DGEVTVDEFK • VFIANQFKAIDVNGDGK • AIDVNGDGK • VGLDEYR • SAFAEVKEIDDAYNK • EIDDAYDK • LTTEDDRK • KAGGLTLER • AGGLTLER • YQELYAQFISNEDEK	*(Litopena eus vannamei)*	N.D.	LC/IT	Ayuso et al. (2009). Sarcoplas mic Calcium-Binding Protein Is an Ef-Hand-Type Protein Identified as a New Shrimp Allergen. Journal of Allergy and Clinical Immunolog y. 124, (1), 114-120.
Sarcoplasm ic calcium binding protein	• VATVSLPR	*(Chionoec etes opilio)*	N.D.	MALDI, LC/qTo F and LC/MR M	Abdel Rahman et al. (2011). Biomolecul ar Characteriz ation of Allergenic Proteins in Snow Crab (Chionoece tes Opilio) and De Novo Sequencin g of the Second Allergen Arginine Kinase Using Tandem Mass Spectromet ry. Journal of Proteomics . 74, (2), 231-241..

Sarcoplasmic calcium binding protein (isoforms)	• DNRVKYVVRYMYDIDN • NGFLDKNDFECLAVKNTLI • DGFLDKNDFECLALRNTLI • ECRGEWSAEKYAANQK • IMSNLWNEIAELADFNK • TFUANQFK • TIDVNGDGL • TVDVNGDGL • VGVDEYRLDCITR • AGGINIARYQELYAQFISNPDEK • CNAVYLFGPLKEVV	*(Pandalus borealis)* Crude extract	N.D	MALDI, LC/qToF and LC/MRM	Abdel Rahman et al. (2013). Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention. Journal of Proteome Research. 12, (2), 647-656.
SERCA/ smooth endoplasmic reticulum Ca2+ATPase	• YGPNELPAEEGK • NAESAIEALKEYEPEMGK • EIVPGDLVEISVGDKIPADLR • IDQSILTGESVSVIK • NILFSGTNVAAGK • TQMAETEEIKTPLQQK • VGEATETALIVLGEK • EFTLEFSR • VIVITGDNK • KAEIGIAMGSGTAVAK	*(Chionoecetes opilio)*	N.D.	MALDI, LC/qToF and LC/MRM	Abdel Rahman et al.. (2011). Biomolecular Characterization of Allergenic Proteins in Snow Crab (Chionoecetes Opilio) and De Novo Sequencing of the Second Allergen Arginine Kinase Using Tandem Mass Spectrometry. Journal of Proteomics. 74, (2), 231-241.
Tropomyosin (Pen m 1) **Tropomyosin (Pen m 1)**	• *ANIQLVEK • AIKKKMQAMKLE • LAEASQAADESER	*(Penaeus monodon)* *(Penaeus monodon)*	N.D. N.D.	LC/qToF MALDI	Abdel Rahman et al. (2010). Analysis of the Allergenic Proteins in Black Tiger Prawn (Penaeus Monodon) and Characterization of the Major Allergen Tropomyosin Using Mass Spectrometry. Rapid Communications in Mass Spectrometry. 24, (24), 3624-3624.

Tropomyosin	• *SQLVENELDHAQEQLSAATHK	*(Chionoecetes opilio)*	N.D.	MALDI and LC/qTOF	Abdel Rahman et al. (2010). Characterization and De Novo Sequencing of Snow Crab Tropomyosin Enzymatic Peptides by Both Electrospary Ionization and Matrix-Assisted Laser Desorption Ionization Qqtof Tandem Mass Spectrometry. Journal of Mass Spectrometry. 45, (4), 372-381.
Tropomyosin	• *SQLVENELDHAQEQLSAATHK	*(Chionoecetes opilio)*	3 nM	LC/MRM	Abdel Rahman et al. (2010). Absolute Quantification Method and Validation of Airborne Snow Crab Allergen Tropomyosin Using Tandem Mass Spectrometry. Analytica Chimica Acta. 681, (1-2), 49-55.
Tropomyosin (Pen b 1) (full sequence coverage)	• *SEEEVFGLQK	*(Pandalus borealis)* Synthetic and isotopcically labeled synthetic peptide	0.25 nM	MALDI, LC/qTOF and LC/MRM	Abdel Rahman et al. (2013). Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention. Journal of Proteome Research. 12, (2), 647-656

Troponin C	• KGFMTPER • AAEFNFR	*(Chionoecetes opilio)*	N.D.	MALDI, LC/qTOF and LC/MRM	Abdel Rahman, et al. (2011). Biomolecular Characterization of Allergenic Proteins in Snow Crab (Chionoecetes Opilio) and De Novo Sequencing of the Second Allergen Arginine Kinase Using Tandem Mass Spectrometry. Journal of Proteomics. 74, (2), 231-241.
Troponin C	• DYEINELNIQVNDLR • DKKKLFEGGW • FLIEEDEEALKTELR • DEEALKTELR • GLDPEALTGKHPPK	*(Pandalus borealis)* Crude extract	N.D	MALDI, LC/qTOF and LC/MRM	Abdel Rahman et al. (2013). Comprehensive Proteomics Approach in Characterizing and Quantifying Allergenic Proteins from Northern Shrimp: Toward Better Occupational Asthma Prevention. Journal of Proteome Research. 12, (2), 647-656
Molluscs					
Arginine kinase	• TKFGGTLDACIR • VQHPVPDFGDVNNLNIGDLDPS GSLIVSTR • SHDSFGFPPVLK • QLTDDHFLFFNDSDR • TFLCWVNEEDHLRLISMQK • GIHGEHTESVGGVYDISNK • MGLTEYEAVTEMMR • GVNEIIR	*Octopus fangsiao*	N.D.	MALDI	Shen et al. (2012). Purification, Cloning, and Immunological Characterization of Arginine Kinase, a Novel Allergen of Octopus Fangsiao. Journal of Agricultural and Food Chemistry. 60, (9), 2190-2199.
Egg					
Lysozyme C (Gal d 4)	• NTDGSTDYGILQINSR • FESNFNTQATNR	Detected in wine from	N.D.	LC/qTOF	Tolin et al. (2012). Analysis of Commercial Wines by

		species *Gallus gallus*			Lc-Ms/Ms Reveals the Presence of Residual Milk and Egg White Allergens. Food Control. 28, (2), 321-326.
Ovalbumin	• HIATNAVLFFGR • YPILPEYLQCVK • DILNQITKPNDVYSFSLASR • ELINSWVESQTNGIIR	Synthetic standards	0.58 ppm standard; 42 ppm in spiked bread	LC/IT	Heick et al. (2011). Application of a Liquid Chromatography Tandem Mass Spectrometry Method for the Simultaneous Detection of Seven Allergenic Foods in Flour and Bread and Comparison of the Method with Commercially Available Elisa Test Kits. Journal of Aoac International. 94, (4), 1060-1068. AND Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943.
Ovalbumin (Gal d 2)	• ISQAVHAAHAEINEAGR	Detected in whole egg, egg biscuit and mayonnaise from species *Gallus gallus*	N.D.	LC/qTOF	Lee and Kim (2010). Determination of Allergenic Egg Proteins in Food by Protein-, Mass Spectrometry-, and DNA-Based Methods. Journal of Aoac International. 93, (2), 462-477.
Ovalbumin (Gal d 2)	• HIATNAVLFFGR • LTEWTSSNVMEER • GGLEPINFQTAADQAR • VASMASEK • AFKDEDTQAMPFR • ISQAVHAAHAEINEAGR	Detected In wine from species *Gallus gallus*	0.1 ppm	LC/qTOF	Tolin et al (2012). Analysis of Commercial Wines by Lc-Ms/Ms Reveals the Presence of Residual Milk and

	• ELYRGGLEPINFQTAADQAR • VTEQESKPVQMMYQIGLFR • NVLQPSSVDSQTAMVLVNAIVFK • LYAEER • ILELPFASGTMSMLVLLPDEVSG LEQLESIINFEK • DILNQITKPNDVYSFSLASR				Egg White Allergens. Food Control. 28, (2), 321-326.
Ovalbumin (Gal d 2)	• *EDTQAMPFRV	Ovalbumin, egg white and whole egg, Raw pasta incurred with 1000 ppm whole egg	N.D.	LC/qTOF	Azarnia et al. (2013). Detection of Ovalbumin in Egg White, Whole Egg and Incurred Pasta Using Lc–Esi-Ms/Ms and Elisa. Food Research International. 52, (2), 526-534.
Ovotransferrin (Gal d 3)	• GAIEWEGIESGSVEQAVAK	Detected in whole egg and mayonnaise from species *Gallus gallus*	N.D.	LC/qTOF	Lee and Kim (2010). Determination of Allergenic Egg Proteins in Food by Protein-, Mass Spectrometry-, and DNA-Based Methods. Journal of Aoac International. 93, (2), 462-477.
Ovotransferrin (Gal d 3)	• GTEFTVNDLQGK • TDERPASYFAVAVAR • GAIEWEGIESGSVEQAVAK • KGTEFTVNDLQGK • DQLTPSPR • AQSDFGVDTK	Detected In wine from species *Gallus gallus*	N.D.	LC/qTOF	Tolin et al. (2012). Analysis of Commercial Wines by Lc-Ms/Ms Reveals the Presence of Residual Milk and Egg White Allergens. Food Control. 28, (2), 321-326.
Serum albumin precursor (Gal d 5)	• MPQVPTDLLLETGKK	Detected in mayonnaise *Gallus gallus*	N.D.	LC/qTOF	Lee and Kim (2010). Determination of Allergenic Egg Proteins in Food by Protein-, Mass Spectrometry-, and DNA-Based Methods. Journal of Aoac International. 93, (2), 462-477.
Vitellogenin II	• SAVSASGTTETL • RFPAVLPQMPL	Detected in whole egg, egg biscuit and mayonnaise from species *Gallus gallus*	N.D.	LC/qTOF	
Fish					
β-parvalbumi	• AEGTFK • VGLTGK	11 different	N.D.	LC/IT and	Carrera et al. (2010). Extensive

n (25 different β-parvalbumin isoforms)		species from the Merlucciidae family		FTICR	De Novo Sequencing of New Parvalbumin Isoforms Using a Novel Combination of Bottom-up Proteomics, Accurate Molecular Mass Measurement by Fticr-Ms, and Selected Ms/Ms Ion Monitoring. Journal of Proteome Research. 9, (9), 4393-4406.
β-parvalbumin	• *SGFIEEDELK • *SGFIEEEELK • *LFLQNFSASAR • *AVGAFSAAESFNYK	(*Gallus gallus*)	N.D.	LC/IT	Carrera et al.(2012). Rapid Direct Detection of the Major Fish Allergen, Parvalbumin, by Selected Ms/Ms Ion Monitoring Mass Spectrometry. Journal of Proteomics. 75, (11), 3211-3220
	• *SGFIEEDELK • *SGFIEEEELK • *AVGAFSAAESFNYK	*(Meleagris gallopavo)*	N.D.	LC/IT	
	• *SGFIEEDELK • *SGFIEEEELK • *AVGAFSAAESFNYK	*(Taeniopygia guttata)*	N.D.	LC/IT	
	• *SGFIEEDELK	*(Ornithorhunchus anaticus)*	N.D.	LC/IT	
	• *SGFIEEDELK	*(Anolis carolinensis)*	N.D.	LC/IT	
	• *SGFIEEDELK • *SGFIEEEELK • *LFLQNFSAGAR • *LFLQNFSASAR • *IGVEEFQALVK	*(Xenopus laevis)*	N.D.	LC/IT	
	• *SGFIEEDELK • *SGFIEEEELK • *LFLQNFSAGAR • *LFLQNFSASAR	*(Xenopus tropicalis)*	N.D.	LC/IT	
	• *SGFIEEEELK • *IGVEEFQALVK	*(Rana catesbeiana)*	N.D.	LC/IT	
	• *SGFIEEEELK	*(Cavia porcellus)*	N.D.	LC/IT	
	• *SGFIEEEELK	*(Triakis semifasciata)*	N.D.	LC/IT	
	• *LFLQNFSAGAR • *SGYIEEEELK	*(Latimeria chalumnae)*	N.D.	LC/IT	
	• *IGVEEFQALVK	*(Rana temporaria)*	N.D.	LC/IT	
	• *IGVEEFQALVK	*(Rana esculenta*	N.D.	LC/IT	

	• *IGVEEFQALVK	) (*Limnonectes macrodon*)	N.D.	LC/IT	
Milk					
α-S1 casein	• YLGYLEQLLR • FFVAPFPEVFGK	Synthetic standards	0.11 ppm standard; 5 ppm in spiked bread	LC/IT	Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943.
α-S1 casein	• *YLGYLEQLLR • *FFVAPFPEVFGK	Detected in white wine	1 ppm	LC/qTOF	Monaci et al. (2010). Identification of Allergenic Milk Proteins Markers in Fined White Wines by Capillary Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry. Journal of Chromatography A. 1217, (26), 4300-4305
α-S1 casein	• YLGYLEQLLR • HQGLPQEVLNENLLR • FFVAPFPEVFGKEK	Detected in wine from species *Bos taurus*	0.06 ppm	LC-qTOF	Tolin et al. (2012). Analysis of Commercial Wines by Lc-Ms/Ms Reveals the Presence of Residual Milk and Egg White Allergens. Food Control. 28, (2), 321-326.
α-S1 casein	• *YLGYLEQLLR • *FFVAPFPEVFGK	Detected in chocolates, cookies, baby foods, frozen desserts	1.25 ppm	LC/qTOF	Weber et al. (2006). Development of a Liquid Chromatography-Tandem Mass Spectrometry Method Using Capillary Liquid Chromatography and Nanoelectrospray Ionization-Quadrupole Time-of-Flight Hybrid

					Mass Spectrometer for the Detection of Milk Allergens. Journal of Agricultural and Food Chemistry. 54, (5), 1604-1610.
α-S1 casein	• HQGLPQEVLNENLLR ($^{13}C^{15}N$-labeled) • YLGYLEQLLR ($^{13}C^{15}N$-labeled) • FFVAPFPEVFGK ($^{13}C^{15}N$-labeled)	CFSAN cookies and FAPAS Biscuits	CFSAN cookies 0.4 0.6 fmol (LOQ: 1.3 fmol) FAPAS Biscuits 0.5 and 0.7 fmol (LOQ: 6 fmol)	LC/qIT	Newsome and Scholl (2012). Quantification of Allergenic Bovine Milk As1-Casein in Baked Goods Using an Intact 15n-Labeled Protein Internal Standard. Journal of Agricultural and Food Chemistry. 61, (24), 5659-5668.
α-S2 casein	• NAVPITPTLNR • FALPQYLK	Synthetic standards	0.11 ppm standard; 5 ppm in spiked bread	LC/IT	Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943..
α-S2 casein	• *YIPIQYVLSR • *NAVPITPTLNR	Detected in white wine	1 ppm	LC/qTOF	Monaci et al. (2010). Identification of Allergenic Milk Proteins Markers in Fined White Wines by Capillary Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry. Journal of Chromatography A. 1217, (26), 4300-4305113-
β-casein	• *DMPIQAFLLYQEPVLGPVR • *GPFPIIV	Detected in white wine	1 ppm	LC/qTOF	

					120.
β-casein	• DMPIQAFLLYQEPVLGPVR • VLPVPQK • AVPYPQR	Detected in wine from species *Bos taurus*	0.06 ppm	LC-qTOF	Tolin et al. (2012). Analysis of Commercial Wines by Lc-Ms/Ms Reveals the Presence of Residual Milk and Egg White Allergens. Food Control. 28, (2), 321-326.
κ-casein	• *SPAQILQWQVLSNTVPAK	Detected in white wine	1 ppm	LC/qTOF	Monaci et al. (2010). Identification of Allergenic Milk Proteins Markers in Fined White Wines by Capillary Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry. Journal of Chromatography A. 1217, (26), 4300-4305
κ-casein	• VQVTSTAV	Synthetic peptide	N.D.	LC/UV and LC/MRM	Molle and Leonil (2005). Quantitative Determination of Bovine Kappa-Casein Macropeptide in Dairy Products by Liquid Chromatography/Electrospray Coupled to Mass Spectrometry (Lc-Esi/Ms) and Liquid Chromatography/Electrospray Coupled to Tamdem Mass Spectrometry (Lc-Esi/Ms/Ms). International Dairy Journal. 15, (5), 419-428.
α-lactalbumin	• m/z = 1419.7 • m/z = 1577.3	Spiked and detected in mixed-fruit juices	1 ppm (LOQ: 4 ppm)	LC/MRM	Monaci and van Hengel, (2008). Development of a Method for the Quantification of Whey Allergen Traces in Mixed-Fruit Juices Based on Liquid

					Chromatography with Mass Spectrometric Detection. Journal of Chromatography A. 1192, (1), 113-120.
α-lactalbumin	• *VGINYWLAHK • KILDKVGINNYWLAHKALCSE	Synthetic standard spiked in infant formula	(LOQ: 100 ppm)	LC/MRM	Zhang et al. (2012). Multiple Reaction Monitoring-Based Determination of Bovine A-Lactalbumin in Infant Formulas and Whey Protein Concentrates by Ultra-High Performance Liquid Chromatography–Tandem Mass Spectrometry Using Tryptic Signature Peptides and Synthetic Peptide Standards. Analytica Chimica Acta. 727, (0), 47-53.
α-lactoglobulin	• m/z = 1082.2 • m/z = 1149.7 • m/z = 1226.2 • m/z = 1313.6	Spiked and detected in mixed-fruit juices	1 ppm (LOQ: 4 ppm)	LC/MRM	Monaci and van Hengel, (2008). Development of a Method for the Quantification of Whey Allergen Traces in Mixed-Fruit Juices Based on Liquid Chromatography with Mass Spectrometric Detection. Journal of Chromatography A. 1192, (1), 113-120.
β-lactoglobulin	• m/z = 1077.1 • m/z = 1144.4 • m/z = 1220.5 • m/z = 1307.5	Spiked and detected in mixed-fruit juices	1 ppm (LOQ: 4 ppm)	LC/MRM	
Peanut					
Cupin: Vicillin-type, 7S globulin (Ara h 1)	• *VLLEENAGGEGEER • *DLAFPGSEQVEK	Raw, mild roasted and strong roasted	7-40 ng/protein	LC/qTOF	Chassaigneet et al. (2007). Proteomics-Based Approach to Detect and Identify

		peanut *(Arachis hypogaea)*			Major Allergens in Processed Peanuts by Capillary Lc-Q-Tof (Ms/Ms). Journal of Agricultural and Food Chemistry. 55, (11), 4461-4473
Cupin: Vicillin-type, 7S globulin (Ara h 1)	• DLAFPGSGEQVE • GTGNLELVAVR	Synthetic standards	0.2 ppm standard; 11 ppm in spiked bread	LC/qIT	Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943..
Cupin: Vicillin-type, 7S globulin (Ara h 1)	• *SFNLDEGHALR • *NNPFYFPSR • *IFLAGDKDNVIDQIEK • *NTLEAAFNAEFNEIR	Spiked in ice cream	10 ppm	LC/qTOF	Shefcheck and Musser (2004). Confirmation of the Allergenic Peanut Protein, Ara H 1, in a Model Food Matrix Using Liquid Chromatography/Tandem Mass Spectrometry (Lc/Ms/Ms). Journal of Agricultural and Food Chemistry. 52, (10), 2785-2790.
Cupin: Vicillin-type, 7S globulin (Ara h 1)	• *VLLEENAGGEGEER • *DLAFPGSEQVEK	Spiked in dark chocolate	2 ppm	LC/MRM	Shefcheck et al. (2006). Confirmation of Peanut Protein Using Peptide Markers in Dark Chocolate Using Liquid Chromatography-Tandem Mass Spectrometry (Lc-Ms/Ms). Journal of Agricultural and Food Chemistry. 54, (21), 7953-7959
Cupin: Vicillin-type, 7S globulin (Ara h 1)	• NNPFYFPSR • IFLAGDKDNVIDQIEK	Raw and roasted peanut flour	N.D.	LC/IT	Hebling et al. (2012). Global Proteomic Screening of Protein Allergens and Advanced Glycation Endproduct

					s in Thermally Processed Peanuts. Journal of Agricultural and Food Chemistry. 61, (24), 5638-5648.
Cupin: Vicillin-type, 7S globulin isoform 1 (Ara h 1)	• VASISATHAKSSPYQKKTENPC • ASVSATQKSPYRKT • WGPAEPE • EGEQEWGTPGSEVR • IVQIEAR • SSEMMEGVIVK • KGSEEGDOTNPINLR • EEEEDEDEEEEGSNR	Raw peanut extract *(Arachis hypogaea)*	N.D.	LC/qTOF	Chassaigne et al. (2009). Resolution and Identification of Major Peanut Allergens Using a Combination of Fluorescence Two-Dimensional Differential Gel Electrophoresis, Western Blotting and Q-Tof Mass Spectrometry. Journal of Proteomics. 72, (3), 511-526.
Cupin: Vicillin-type, 7S globulin isoform 2 (Ara h 1)	• ASVSATQKSPYRKT • WGPAEPE • EGEQEWGTPGSEVR • IVQIEAR;EHVQELTK • GSEEDITNPINLR		N.D.	LC/qTOF	
Conglutin (2S albumin) (Ara h 2)	• *CCNELNEFENNQR • *CMCEALQQIMENQSDR	In rice crispy and chocolate based snacks	5 ppm (LOQ: 14 ppm)	LC/qTOF and LC/MRM	Careri et al (2007). Use of Specific Peptide Biomarkers for Quantitative Confirmation of Hidden Allergenic Peanut Proteins Ara H 2 and Ara III 3/4 for Food Control by Liquid Chromatography-Tandem Mass Spectrometry. Analytical and Bioanalytical Chemistry. 389, (6), 1901-1907.
Conglutin (2S albumin) (Ara h 2)	• *RQQWELQGDR	Raw, mild roasted and strong roasted peanut (*Arachis hypogaea*)	7-40 ng/protein	LC/qTOF	Chassaigneet al. (2007). Proteomics-Based Approach to Detect and Identify Major Allergens in Processed Peanuts by Capillary Lc-Q-Tof (Ms/Ms). Journal of Agricultural and Food Chemistry. 55, (11), 4461-4473
Conglutin (2S	• CCNELNEFENNQR • APQRCDLEVESGGR	Raw and roasted	N.D.	LC/IT	Hebling et al. (2012). Global Proteomic

albumin) (Ara h 2)	• CMCEALQQIMENQSDR • CDLEVESGGR	peanut flour			Screening of Protein Allergens and Advanced Glycation Endproducts in Thermally Processed Peanuts. Journal of Agricultural and Food Chemistry. 61, (24), 5638-5648.
Conglutin (2S albumin) isoform 1 (Ara h 2)	• DLEVESGGRD	Raw peanut extract (*Arachis hypogaea*)	N.D.	LC/qTOF	Chassaigne et al. (2009). Resolution and Identification of Major Peanut Allergens Using a Combination of Fluorescence Two-Dimensional Differential Gel Electrophoresis, Western Blotting and Q-Tof Mass Spectrometry. Journal of Proteomics. 72, (3), 511-526.
Conglutin (2S albumin) isoform 2 (Ara h 2)	• DEDSYERDPY • HPSQDPD	Raw peanut extract (*Arachis hypogaea*)	N.D.	LC/qTOF	
Cupin: Legumin-type, 11S globulin, Glycinin (Ara h 3/4)	• *AHYQVVDSNGDT • *SPDIYNPQAGSLK	In rice crispy and chocolate based snacks	1 ppm (LOQ: 3.7 ppm)	LC/qTOF and LC/MRM	Careri et al (2007). Use of Specific Peptide Biomarkers for Quantitative Confirmation of Hidden Allergenic Peanut Proteins Ara H 2 and Ara Ill 3/4 for Food Control by Liquid Chromatography-Tandem Mass Spectrometry. Analytical and Bioanalytical Chemistry. 389, (6), 1901-1907.
Cupin: Legumin-type, 11S globulin, Glycinin (Ara h 3/4)	• AHYQVVDSNGDR • SPDIYNPQAGSLK		3 ppm (LOQ: 10 ppm)	LC/IT	Careri et al. (2008). Selective and Rapid Immunomagnetic Bead-Based Sample Treatment for the Liquid Chromatography-Electrospray Ion-Trap Mass Spectrometry Detection

					of Ara H3/4 Peanut Protein in Foods. Journal of Chromatography A. 1206, (2), 89-94.
Cupin: Legumin-type, 11S globulin, Glycinin (Ara h 3/4)	• *AHYQVVDSNGDT • *SPDIYNPQAGSLK • *AQSENYEYLAFK	Raw, mild roasted and strong roasted peanut *(Arachis hypogaea)*	7-40 ng/protein	LC/qTOF	Chassaigneet al. (2007). Proteomics-Based Approach to Detect and Identify Major Allergens in Processed Peanuts by Capillary Lc-Q-Tof (Ms/Ms). Journal of Agricultural and Food Chemistry. 55, (11), 4461-4473
Cupin: Legumin-type, 11S globulin, Glycinin (Ara h 3/4)	• RPFYSNAPQEIFIQQGR • WLGLSAEYGNLYR	Synthetic standards	0.2 ppm standard; 11 ppm in spiked bread	LC/IT	Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943..
Cupin: Legumin-type, 11S globulin, Glycinin (Ara h 3/4)	• SQSENFEYVAFK • SPDIYNPQAGSLKTANDLNLLILR	Raw and roasted peanut flour	N.D.	LC/IT	Hebling et al. (2012). Global Proteomic Screening of Protein Allergens and Advanced Glycation Endproducts in Thermally Processed Peanuts. Journal of Agricultural and Food Chemistry. 61, (24), 5638-5648.
Cupin: Legumin-type, 11S globulin, Glycinin Isoform 1 (Ara h 3/4)	• GETESEEGAIVTVR	Raw peanut extract *(Arachis hypogaea)*	N.D.	LC/qTOF	Chassaigne et al. (2009). Resolution and Identification of Major Peanut Allergens Using a Combination of Fluorescence Two-Dimensional Differential Gel Electrophoresis, Western Blotting and Q-Tof Mass Spectromet
Cupin: Legumin-type, 11S globulin, Glycinin Isoform 4	• FQGQDQSQQQQDSHQK • AHVQVVDSNGDR	Raw peanut extract *(Arachis hypogaea)*	N.D.	LC/qTOF	

(Ara h 3/4)					ry. Journal of Proteomics. 72, (3), 511-526.
Soy					
α-chain of the conglycinin (Gly m 5)	• *(R)QQQEEQPLEVRK(Y)	*(Glycine max)*	N.D.	MALDI/ qTOF	Cucu et al.. (2012). Maldi Based Identification of Soybean Protein Markers - Possible Analytical Targets for Allergen Detection in Processed Foods. Peptides. 33, (2), 187-196.
β-conglcinin α-subunit (Gly m 5)	• *LITLAIPVNKPGR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	Houston et al. (2010). Quantitation of Soybean Allergens Using Tandem Mass Spectrometry. Journal of Proteome Research. 10, (2), 763-773.
G1 glycinin (Gly m 6)	• *(R)VFDGELQEGR(V)	*(Glycine max)*	N.D.	MALDI/ qTOF	Cucu et al.. (2012). Maldi Based Identification of Soybean Protein Markers - Possible Analytical Targets for Allergen Detection in Processed Foods. Peptides. 33, (2), 187-196.
G1 glycinin (Gly m 6)	• *VLIVPQNFVVAAR • LSAEFGSLR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	Houston et al. (2010). Quantitation of Soybean Allergens Using Tandem Mass Spectrometry. Journal of Proteome Research. 10, (2), 763-773.
G2 glycinin (Gly m 6)	• *NLQGENEEEDSGAIVTVK • LSAQYGSLR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	
G3 glycinin (Gly m 6)	• *FYLAGNQEQEFLQYQPQK • LSAQFGSLR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	
G4 glycinin (Gly m 6)	• *VESEGGLIQTWNSQHPELK	20 varieties of soybean	N.D.	LC/IT and LC/MRM	

		(Glycine max)			
Glycinin precursor (Gly m 6)	• *NGLHLPSYSPYPR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	
Glycinin	• NLQGENEGEDKGAIVTVK • VFDGELQEGR • SQSDNFEYVSFK • EAFGVNMQIVR	Synthetic standards	0.24 ppm standard; 24 ppm in spiked bread	LC/IT	Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943..
Gly m Bd 28K	• DGPLEFFGFSTSAR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	Houston et al. (2010). Quantitation of Soybean Allergens Using Tandem Mass Spectrometry. Journal of Proteome Research. 10, (2), 763-773.
Gly m Bd 30K	• EESETLVSAR	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	
Kunitz trypsin inhibitor 1	• *DTVDGWFNIER	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	
Kunitz trypsin inhibitor 2	• *FIAEGHPLSLK • *VSDDEFNNYK	20 varieties of soybean *(Glycine max)*	N.D.	LC/IT and LC/MRM	
Profilin (Gly m 3)	• YMVIQGEPGAVIR • KGPGGVTVK • GPGGVTVKK	In soymilk and soybean meal	N.D.	LC/qTOF	Amnuaycheewa and de Mejia (2010). Purification, Characterisation, and Quantification of the Soy Allergen Profilin (Gly M 3) in Soy Products. Food Chemistry. 119, (4), 1671-1680.
Tree Nut					
2S albumin	• QDCPEQMQR	Brazil nut	N.D.	MALDI	Moreno et

(six different isoforms)	• EQMQRQQMLSHCR • QQMLSHCR • MYMRQMMK • CEGLRMMMR • MMRMMQQK • MQQEEMQPR • LAENIPSR • CNLSPQR • MMMRMAENLPSR	*(Bertholletia excels)* purified extract			al. (2004). Mass Spectrometry and Structural Characterization of 2s Albumin Isoforms from Brazil Nuts (Bertholletia Excelsa). Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 1698, (2), 175-186.
2S sulfur-rich seed storage albumin (Ber e 1)	• QEECREQMQR • EQMQRQQMLSHCR • GRSEQQCR • QQMLSHCR • QQQLNHCR • GEEPHLDECCEQLER • MQQEEMQPR • QREEMELQGEQMQR • KAENLLSR • LAENIPSR • MAENLPSR • LAENIPSRCNLSPMR • MAENLPSRCNLSPMR • CNLSPMR	Brazil nut *(Bertholletia excels)* purified extract	N.D.	MALDI	
2S albumin (Jug r 1)	• DLPNECGISSQR • QCCQQLSQMDEQCQCEGLR • GEEMEEMVQSAR	Walnut Synthetic standards	10 ppm standard; 70 ppm in spiked bread	LC/IT	Heick et al. (2011). First Screening Method for the Simultaneous Detection of Seven Allergens by Liquid Chromatography Mass Spectrometry. Journal of Chromatography A. 1218, (7), 938-943..
11S globulin (Cor a 9)	• ADIYTEQVGR • INTVNSNTLPVLR • QGQVLTIPQNFAVAK • ALPDDVLANAFQISR	Hazelnut Synthetic standards	0.32 ppm standard; 5 ppm in spiked bread	LC/IT	
Prunin	• GNLDFVQPPR • GVLGAFSGCPETFEESQQSSQQGR • ALPDEVLANAYQISR • NGLHLPSYSNAPQLIYIVQGR	Almond Synthetic standards	0.13 ppm standard; 3 ppm in spiked bread	LC/IT	
Wheat					
α-amylase inhibitor	• LQCNGSQVPEAVVRDCCQQLANISEWCR • DCCQQLANISEWCRCDALYNMLDSMYK • CDALYNMLDSMYKEHGAQEGQAGTGAFPR • EHGAQEGQAGTGAFPRCR	In wheat flour *(Triticum aestivum* L.*)*	N.D.	MALDI-TOF/TOF	Akagawa et al. (2007). Proteomic Analysis of Wheat Flour Allergens. Journal of Agricultural and Food Chemistry. 55, (17), 6863-6870.

α-amylase inhibitor	• SGPWMCYPGQAFQVPALPACR • EHGAQEGQAGTGAFPR • LATIHNVR • DCCQQLAHISEWCR • QEGQAGTGAFPR	In wheat flour *(Triticum aestivum and Triticum spelta)*	N.D.	MALDI/TOF and LC/qTOF	Sotkovsky et al. (2008). Proteomic Analysis of Wheat Proteins Recognized by Ige Antibodies of Allergic Patients. Proteomics. 8, (8), 1677-1691.
β-amylase inhibitor	• FFVDNGTYLTEQGR • ASLNFTCAEMR • YDPTAYNTILR • SAPEELVQQVLSAGWR • DAGQYNDAPQR • SFPGIGEFICYDK • AAAAMVGHPEWEFPR	In wheat flour *(Triticum aestivum and Triticum spelta)*	N.D.	MALDI/TOF and LC/qTOF	
CM3 α-amylase/trypsin inhibitor (Tri a 30)	• IALPVPSQPVDPR • FIALPVPSQPVDPR	18 samples analysed belonged to *Triticum turgidum* spp durum species; two different varieties (D240, Levante and Svevo) and cultivation areas	N.D.	LC/IT	Prandi et al (2013). Lc/Ms Analysis of Proteolytic Peptides in Wheat Extracts for Determining the Content of the Allergen Amylase/Trypsin Inhibitor Cm3: Influence of Growing Area and Variety. Food Chemistry. 140, (1–2), 141-146.
β-D-Glucanexohydrolase	• TAGTTILSAIK	In wheat flour *(Triticum aestivum and Triticum spelta)*	N.D.	MALDI/TOF and LC/qTOF	Sotkovsky et al. (2008). Proteomic Analysis of Wheat Proteins Recognized by Ige Antibodies of Allergic Patients. Proteomics. 8, (8), 1677-1691.
γ-gliadin	• NFLLQQCNHVSLVSSLVSIILPR • TLPTMCNVYVPPDCSTINIPYANI DAGIGGQ	In wheat flour *(Triticum aestivum* L.*)*	N.D.	MALDI-TOF/TOF	Akagawa et al. (2007). Proteomic Analysis of Wheat Flour Allergens. Journal of Agricultural and Food Chemistry. 55, (17), 6863-6870.

γ-gliadin (isoforms)	• QPQQPFP • QTQQPQQPFP • LALQTLPAMC • YIPPHCSTTI	In wheat grain	N.D	LC/qTOF	Uvackova et al. (2013). Mse Based Multiplex Protein Analysis Quantified Important Allergenic Proteins and Detected Relevant Peptides Carrying Known Epitopes in Wheat Grain Extracts. Journal of Proteome Research.
LMW glutenin	• VFLQQQCSPVAMPQSLAR • SQMLQQSSCHVMQQQCCQQLP QIPQQSR • TLPTMCNVNVSLYR • VPFGVGTGVGGY	In wheat flour *(Triticum aestivum L.)*	N.D.	MALDI-TOF/TOF	Akagawa et al. (2007). Proteomic Analysis of Wheat Flour Allergens. Journal of Agricultural and Food Chemistry. 55, (17), 6863-6870.
Peroxidase 1	• GAVVSCADILALAAR • DSVVVSGGPDYR	In wheat flour *(Triticum aestivum and Triticum spelta)*	N.D.	MALDI/TOF and LC/qTOF	Sotkovsky et al. (2008). Proteomic Analysis of Wheat Proteins Recognized by Ige Antibodies of Allergic Patients. Proteomics. 8, (8), 1677-1691.
Profilin (Tri a12)	• YMVIQGEPGVVIR • YMVIQGEPGVVIR • DFEEPGHLAPTGLFLGGTK	In wheat flour *(Triticum aestivum and Triticum spelta)*	N.D.	MALDI/TOD and LC/qTOF	
Serpin (Tri a 33)	• LSIAHQTR • VAFANGVFVDASLQLKPSFQELA VCK • QFSMYILLPEAPGGLSNLAEK • VTTGLIK • LSAEPEFLEQHIPR • LASTISSNPK • SAASNAAFSPVSLHSALSLLAAG AGSATR • EDTSGVVLFIGHVVNPLLSS • AAEVTTQVNSWVEKVTSGR • NILPSGSVDNTTKLVLANALYFK • LVLANALYFK • VLKLPYK • LSAEPDFLER • QGGDNRQFSMYILLPEAPGGLS SLAEK • CLGLQLPFSDEADFSEMVDSPM PQGLR	In wheat flour *(Triticum aestivum L.)*	N.D.	MALDI-TOF/TOF	Akagawa et al. (2007). Proteomic Analysis of Wheat Flour Allergens. Journal of Agricultural and Food Chemistry. 55, (17), 6863-6870.

wheat gluten	• LQPQNPSQQQPQEQVPL • TQQPQQPFPQQPQQPFPQ • VPVPQLQPQNPSQQQPQEQVPL • RPQQPYPQPQPQY • QPQQPFPQTQQPQQPFPQ • PQQSPF	Synthetic standards and various different flours and different food products	0.001–0.03 ppm (LOQ: 0.01-0.1 ppm)	LC/MRM	Sealey-Voyksner et al. (2010). Novel Aspects of Quantitation of Immunogenic Wheat Gluten Peptides by Liquid Chromatography-Mass Spectrometry/Mass Spectrometry. Journal of Chromatography A. 1217, (25), 4167-4183.

B.2 Supplementary figures

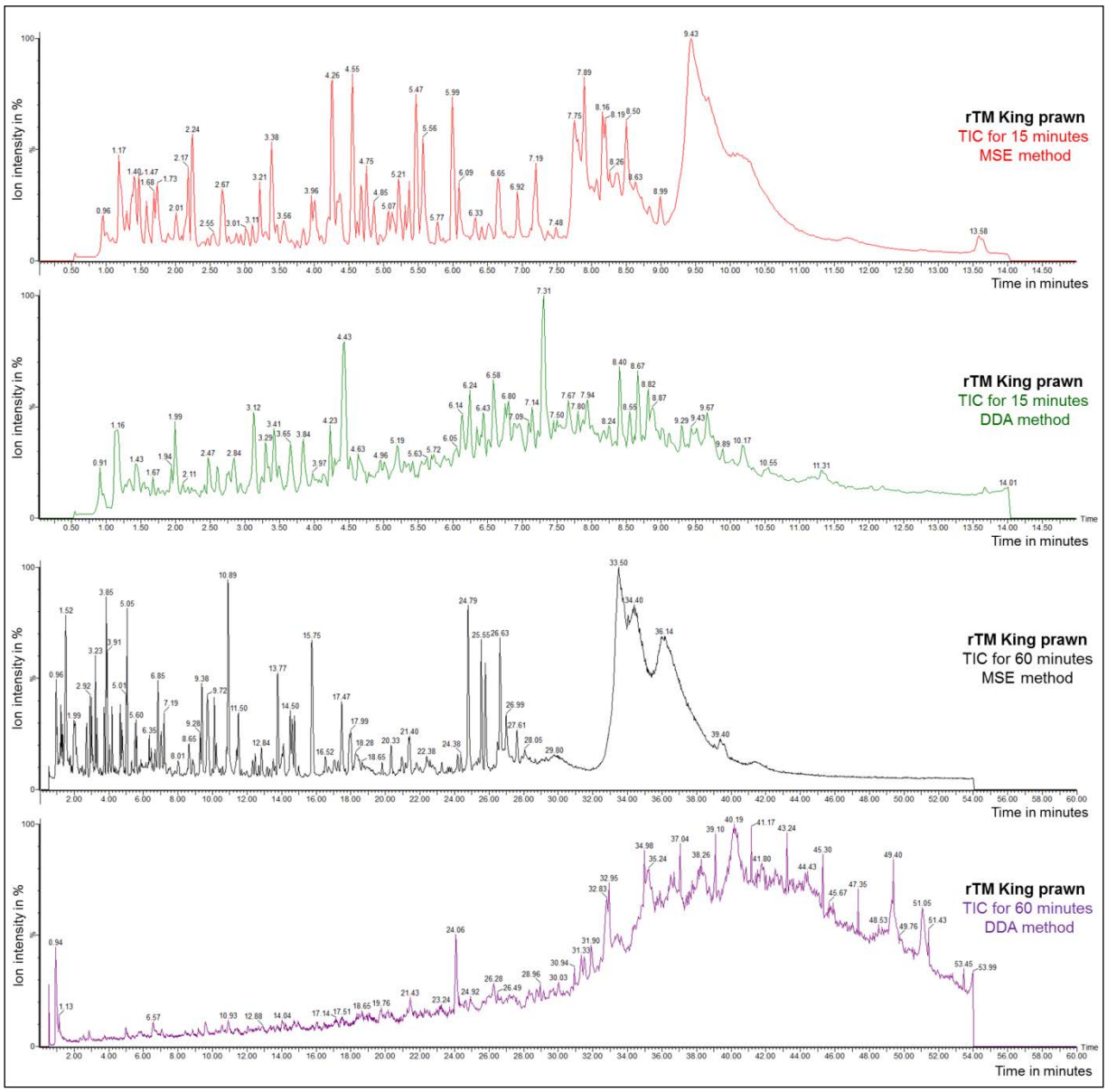


Figure B2.1: LC/qTOF chromatograms for four LC/qTOF methods applied for digested recombinant tropomyosin from King prawn showing the total ion chromatogram (TIC).

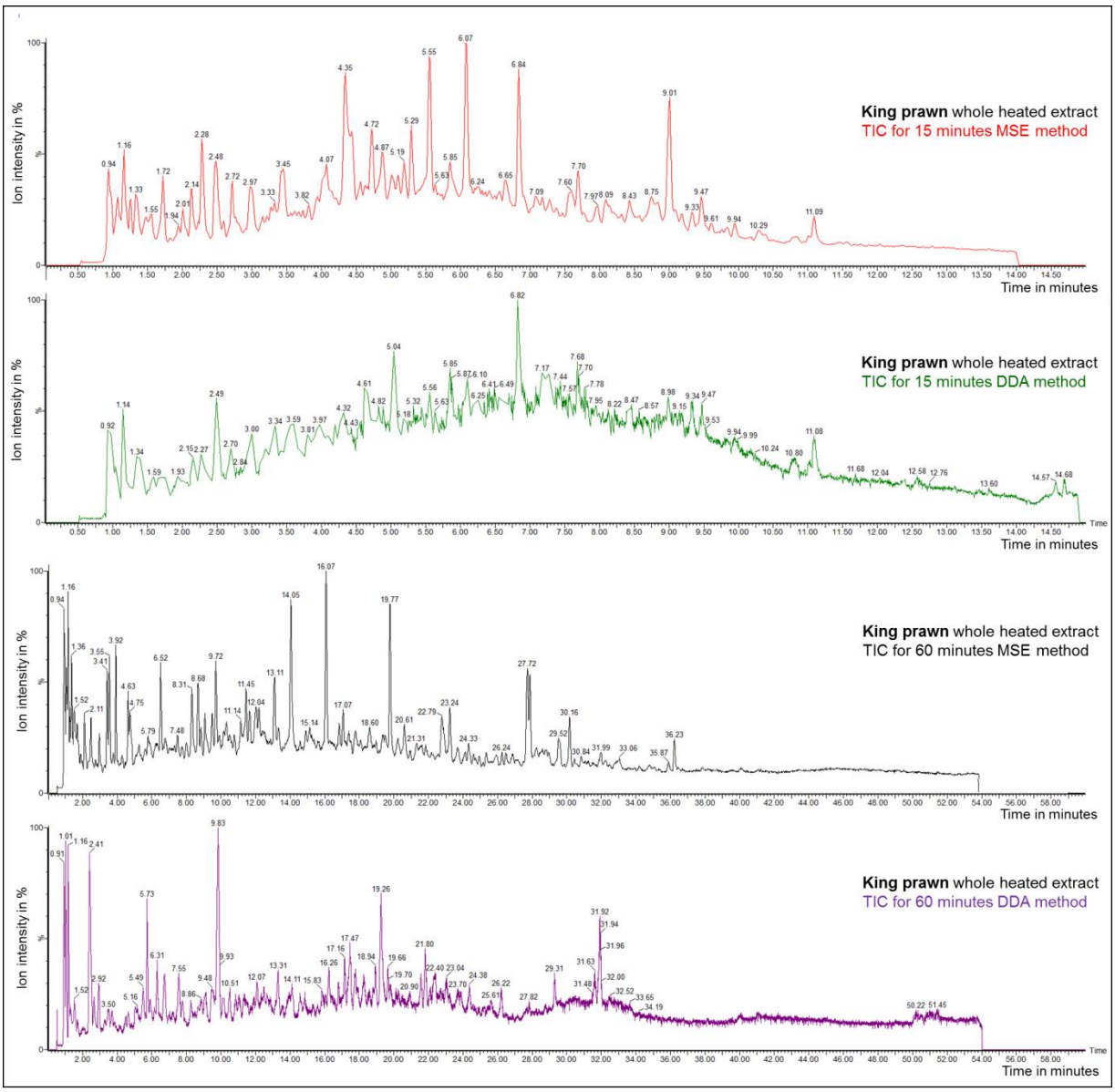


Figure B2.2: LC/qTOF chromatograms for four LC/qTOF methods applied for digested whole heated extract from King prawn showing the total ion chromatogram (TIC).

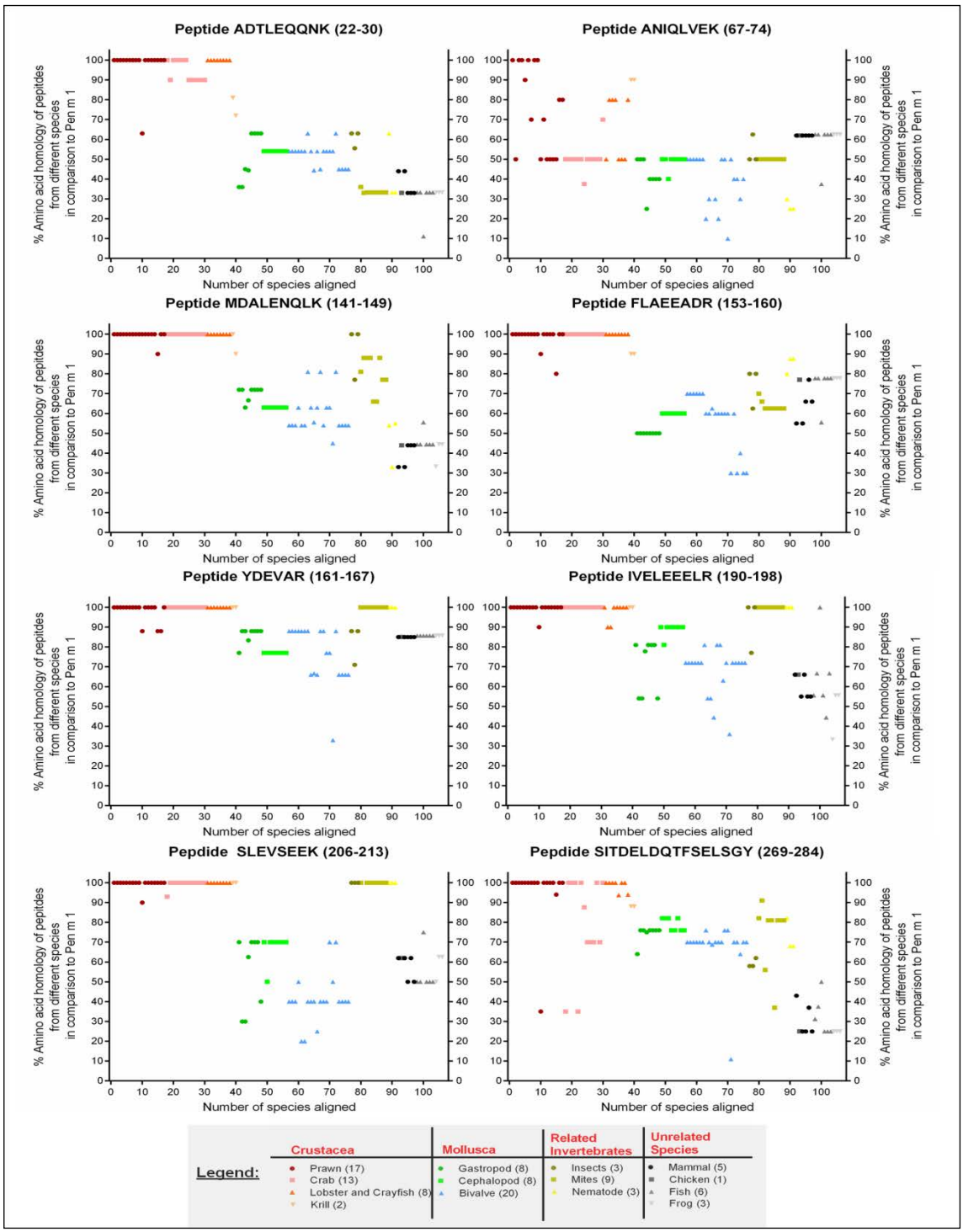


Figure B2.3: Amino acid alignment using commonly identified tryptic peptides of TM analysing 22 shellfish species by LC/qTOF. Peptide sequence used as reference for alignment is originated from Pen m 1. Amino acid homology is shown in percent for 106 species and colour coded by their shellfish subgroups.

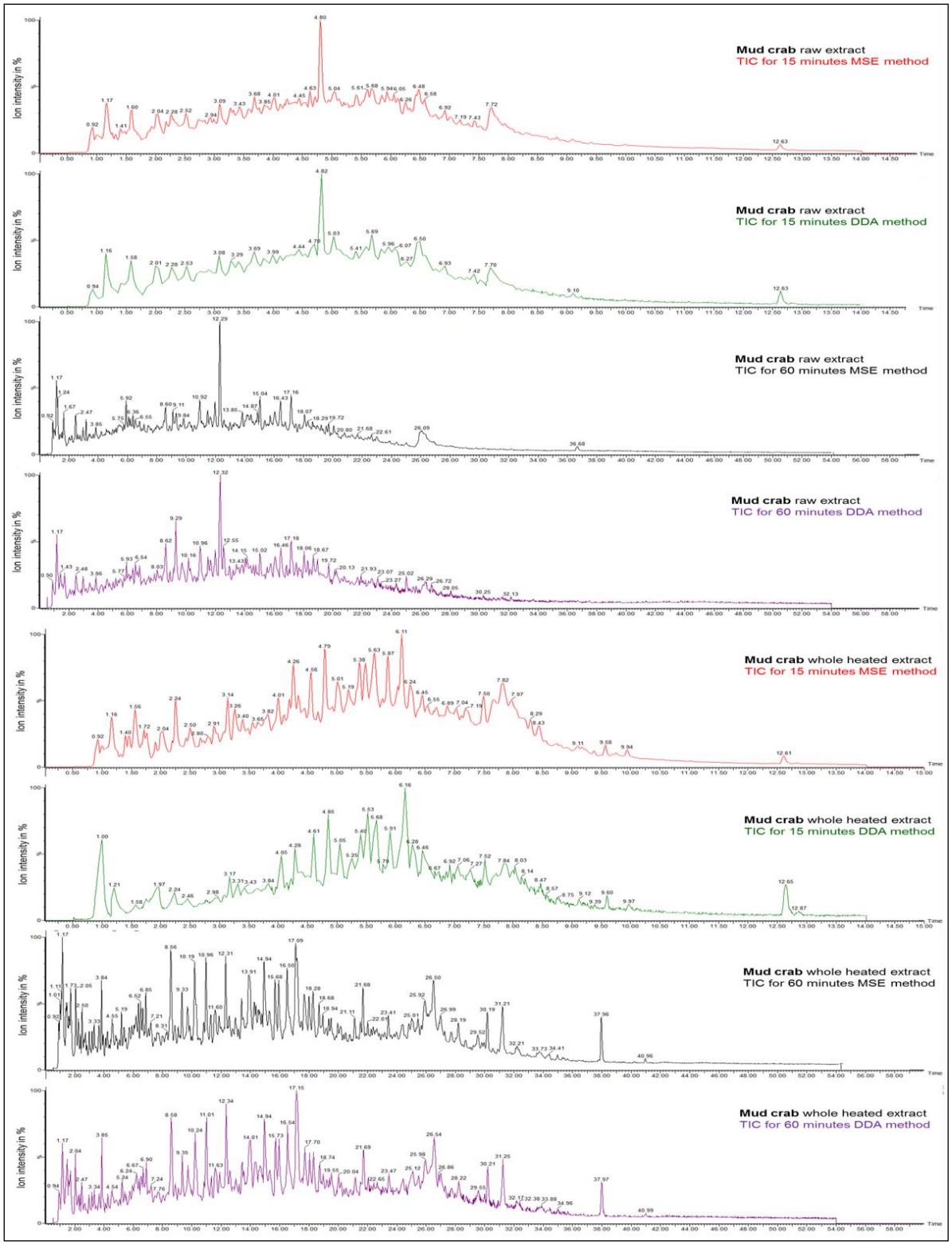


Figure B2.4: LC/qTOF chromatograms for four LC/qTOF methods applied for digested raw and whole heated extracts from Mud crab showing the total ion chromatogram (TIC).

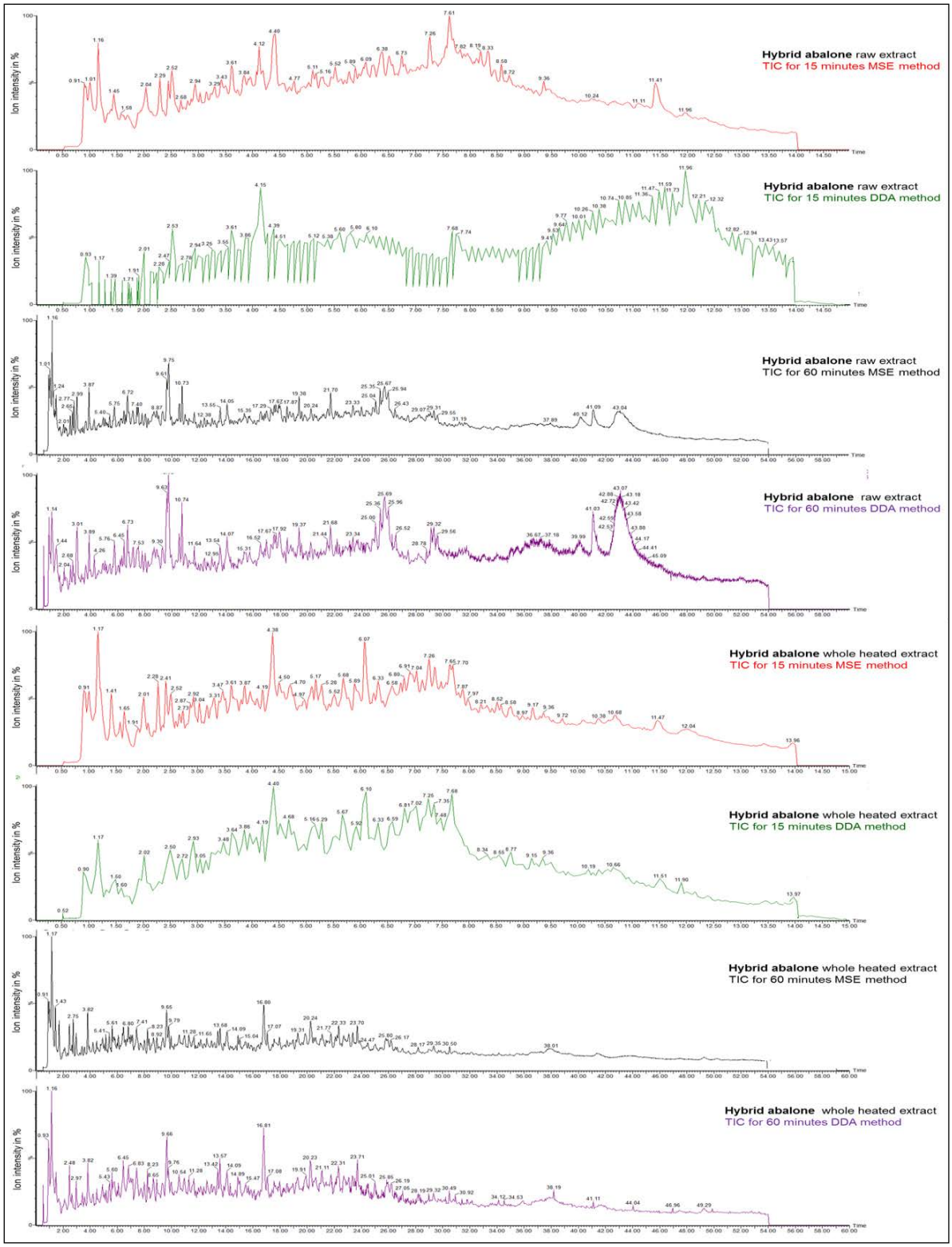
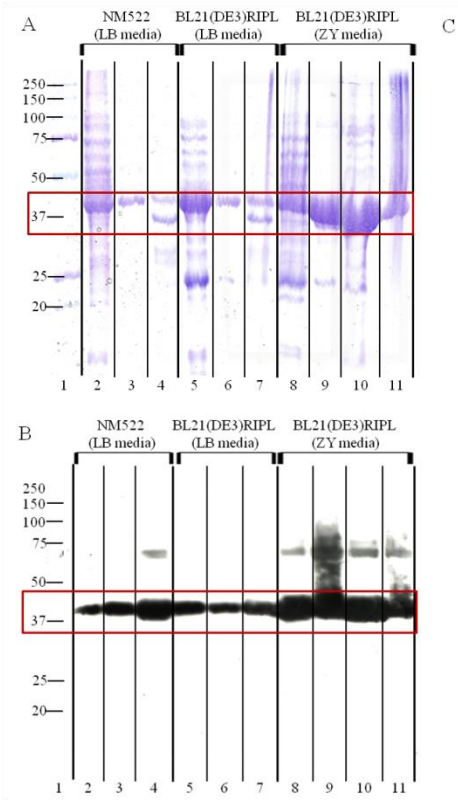


Figure B2.5: LC/qTOF chromatograms for four LC/qTOF methods applied for digested raw and whole heated extracts from Hybrid abalone showing the total ion chromatogram (TIC).



C

Summary of Mascot results for the identification of recombinant TM, expressed using different expression systems using different strains

		NM522 LB media		BL21DE3RIPL LB media		BL21DE3RIPL ZY media	
Tropo-mysoin	Score	537		763		589	
	Sequence coverage	71%		70%		77%	
Peptide	Peptide number in sequence	Peptide sequence	Score	Expect	Score	Expect	Score	Expect
	1-5	.MDAIK.K	25	1.50E-01	30	0.14	31	1.00E-01
	8-12	K.MQAMK.L	30	2.80E-01	29	0.47	31	2.70E-01
	16-21	K.DNAMDR.A	43	8.00E-02	41	0.12	42	1.10E-01
	22-30	R.ADTLEQQNK.E	47	1.30E-02	41	0.11	48	1.40E-02
	22-35	R.ADTLEQQNKEANIR.A			42	0.04		
	31-35	K.EANIR.A			25	0.69		
	75-90	K.DKALSNAEGEVAALNR.R	15	2.40E-01	33	0.064	32	2.00E-01
	77-90	K.ALSNAEGEVAALNR.R	105	8.20E-09	120	2.50E-09	122	1.70E-09
	92-101	R.IQLLEEDLER.S	80	2.40E-05	76	7.10E-05	76	7.20E-05
	92-105	R.IQLLEEDLERSEER.L	85	8.90E-06	93	1.60E-06	79	4.00E-05
	92-112	R.IQLLEEDLERSEERLNTATTK.L	60	3.70E-03	81	6.00E-04	45	2.80E-03
	106-112	R.LNTATTK.L	63	1.20E-03	60	2.50E-02	63	1.20E-03
	113-125	K.LAEASQAADESER.M	116	4.40E-09	115	6.70E-09	119	2.40E-09
	128-133	R.KVLENR.S	51	1.90E-02	52	1.60E-02	51	1.70E-02
	134-140	R.SLSDEER.M	16	3.00E+00	28	1.40E+00	22	1.40E+00
	141-149	R.MDALENQLK.E	46	4.30E-02	56	2.10E-02	45	4.90E-02
	141-152	R.MDALENQLKEAR.F			51	2.40E-02	39	4.90E-01
	153-160	R.FLAEEADR.K	48	6.20E-03	50	2.50E-02		
	153-167	R.FLAEEADRKYDEVAR.K			100	2.80E-08	108	5.20E-09
	161-167	R.KYDEVAR.K	27	6.60E-02	27	7.10E-01	23	4.80E-01
	168-178	R.KLAMVEADLER.A	111	2.30E-07	111	2.20E-08	94	9.90E-07
	168-182	R.KLAMVEADLERAEER.A	55	3.70E-03	88	5.70E-07	72	3.40E-05
	169-178	K.LAMVEADLER.A	73	1.30E-04	77	5.10E-05	75	9.00E-05
	169-182	K.LAMVEADLERAEER.A	49	4.00E-02	68	1.00E-04	46	4.60E-02
	190-198	K.IVELEEELR.V	77	5.10E-05	74	9.30E-05	77	4.80E-05
	190-250	K.IVELEEELRVVGNNLK.S	51	3.80E-03	91	3.70E-07	54	9.60E-04
	206-217	K.SLEVSEEK.A	22	5.80E-01	34	2.10E-02	19	5.70E-01
	218-226	R.EEAYKEQIK.T			42	1.00E-02	26	7.40E-01
	239-244	R.AEFAER.S	15	9.60E-01	35	2.20E-01	37	3.80E-01
	252-264	K.EVDRLEDELVNEK.E	43	9.70E-03	64	3.00E-04	43	3.90E-02
	252-266	K.EVDRLEDELVNEKEK.Y	24	9.50E-02	43	1.50E-01	37	7.20E-02
	269-284	K.SITDELDQTFSELSGY.-	9	2.20E-01			22	1.10E-02

Figure B2.6: Molecular and immunological comparison of different E. coli strains and induction systems. Proteins in generated fractions were separated by A) SDS-PAGE and confirmation of expressed recombinant proteinwith B) anti-His tag antibody. The lane number represent analyzed supernatants ofNM522 IPTG induction (Lane 2–4), BL21(DE3)RIPL cells for IPTG induction (Lane 5–7) and BL21(DE3)RIPL cells auto-induction (Lane 8–11). Lane numbers represent: 1. Marker; 2. 1st supernatant; 3. 2nd supernatant; 4. pellet; 5. 1st supernatant; 6. 2nd supernatant; 7. pellet; 8. 1st supernatant; 9. 2nd supernatant; 10. 3rd supernatant; and 11. Pellet. The expressed proteins were further confirmed using mass spectrometry C). The summarized results of the mass spectrometry analysis, including sequence coverage, peptides identified and scores of purified TM of the different strain and induction system are displayed.