A PROTEOMIC INVESTIGATION OF THE HEAT STRESS RESPONSE OF THE SOUTH AFRICAN ABALONE *HALIOTIS MIDAE*

By Bridget Calder

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Molecular and Cell Biology,
Faculty of Science

University of Cape Town,
South Africa.

July 2014

Supervised by Prof. V. Coyne and Dr. S. Rafudeen

Declaration

I know the meaning of plagiarism and declare that all of the work in this thesis, save for that which is properly acknowledged, is my own.

Signed:
B. Calder

_____

Acknowledgements

I would like to acknowledge the support and assistance offered to me by my supervisor, Professor Vernon. E. Coyne, and co-supervisor Dr Suhail Rafudeen during my time at the University of Cape Town. The marine biotechnology laboratory has been a source of constant comfort and much needed advice during my time in the molecular and cell biology department. Thank you also to the department as a whole for providing me with so much stimulating tea-time conversation and helpful guidance. I would also like to thank Dr Zac McDonald for his assitance with mass spectrometry in the molecular and cell biology department, without which this work would not have been possible.

Then, I would like to thank my friends and loved ones who have been so patient with me while I completed this work. This would not have been possible without a loving and understanding support structure.

I would also like to formally acknowledge the financial assistance offered to me by the postrgraduate funding office at the University of Cape Town, and the National Research foundation for providing bursaries during my postgraduate studies.

TABLE OF CONTENTS

A PROTEOMIC INVESTIGATION OF THE HEAT STRESS RESPONSE OF THE SOUTH AFRICAN ABALONE *HALIOTIS MIDAE*

By Bridget Calder

2014

ABSTRACT

The abalone *Haliotis midae* has been fished to near-extinction on the South African coastline, primarily to satisfy a growing international market. In order to meet demands, *H. midae* has been produced in South Africa by aquaculture for several decades, and the South African abalone aquaculture industry continues to expand. Internationally, abalone aquaculture has been negatively affected by the outbreak of bacterial and viral diseases, which spread rapidly and lead to high abalone mortality. There is evidence that environmental stresses on abalone farms may lead to immunosuppression, and thereby increase the severity of disease outbreaks. The water temperature on abalone farms fluctuates seasonally, and increased abalone mortality has been associated with warmer water during the summer months. However, the molecular mechanisms affecting the abalone during exposure to stress remain unclear. With advances in proteomics technology, it is possible to identify and quantify the expression of several hundred proteins simultaneously. This study therefore aimed to gain insight into the *H. midae* stress response by using proteomic tools to identify proteins that are differentially regulated in haemocytes during exposure to acute heat stress. Identifying which biochemical pathways are involved in the abalone stress response will give some insight into the molecular mechanism by which *H. midae* responds to heat stress.

Proteins that were significantly regulated under heat shock conditions were identified using two different proteomic techniques: 2D PAGE analysis and iTRAQ. Several proteins with cytoskeletal function were identified by 2D PAGE, indicating that cytoskeletal remodelling may be induced by acute heat stress. As the abalone immune response is predicated on the ability of the haemocytes to engulf foreign particles by phagocytosis, it was therefore hypothesised that damage to the cytoskeleton could lead to immunosuppression. iTRAQ analysis of heat stressed abalone haemocyte total protein was then carried out and additional heat-regulated proteins were identified. A statistically stringent method was employed, and 11 proteins with significant heat-induced differential expression were identified. Additional differentially regulated proteins were identified using statistical seriation to cluster proteins with similar heat-induced expression profiles. As many of the identified proteins lack functional characterisation in abalone, Blast2GO® analysis was performed to associate gene ontology keywords with the protein identities that had the most consistent response to heat stress. The GO terms were then simplified using the tool

REViGO to identify biochemical pathways that were implicated by these protein identities. Based on these data, the cellular functions that appear to be affected by heat stress in *H. midae* are cell signalling, response to stress and metabolic regulation.

As several important proteins identified by iTRAQ remain uncharacterised in abalone, one was selected for further characterisation in order to determine its suitability as a potential molecular heat stress indicator in *H. midae*. Thus, further expression analysis of the molecular chaperone calreticulin (CRT) was carried out using alternate molecular techniques. The transcriptional regulation of CRT in *H. midae* during exposure to heat stress was measured using qPCR, which detected a significant early decrease in the amount of CRT mRNA followed by a return to basal levels. The translational regulation of CRT was then measured by western blot analysis, which detected a significant increase in the amount of CRT in the haemocytes after 120 minutes of exposure to 23 °C. This data also confirmed the expression profile determined by iTRAQ for CRT.

The findings of this study indicate that the *H. midae* molecular heat stress response is complex and interconnected, and capable of rapidly responding to acute thermal stress. However, the metabolic cost to the abalone and the potential damage caused to the haemocytes should be of concern to abalone farmers. This information is also of particular concern when considering the effects of anthropogenic climate change on the survival in the wild of this already threatened species. This study presents a set of molecular tools that may be of use in future molluscan studies, and with the growing popularity of proteomics as a diagnostic tool in aquaculture may serve as a platform for the investigation of the effect of environmental stress on other molluscan species.

Chapter 1
LITERATURE REVIEW

1.1. Introduction

The global human population is predicted to surpass 9 billion by 2050 (United-Nations, 2004). This rate of population growth is predicted to increase particularly in sub-Saharan Africa, an area already suffering from food shortages. With increased urbanisation there will also be a loss of farm land. These predictions would require global food production to increase by 60% (Alexandratos and Bruinsma, 2012) between now and 2050 to meet the demands of the increasing human population. There is cause for concern as to the sustainability and achievability of these goals. Aquaculture, the farming of water-based organisms including fish, crustaceans, molluscs and aquatic plants, is thought to be a possible means through which global food demands may be met (Cunningham, 2005). The consumption of fish and seafood is growing and in 2008 was measured at 17.2 kg/capita/year, contributing 15.7% to the global animal protein intake (De Silva, 2012; Merino et al., 2012). If the current trends continue, the aquaculture sector will need to exceed current production by 30-40 million tonnes per year in order to meet the projected food requirements of 2050 (De Silva, 2012). Aquaculture in sub-Saharan Africa contributes less than 1% of the total global aquaculture production by weight (Cunningham, 2005), despite the suitability of approximately 42 million ha of land for small-scale fish farming (Moehl and Machena, 2000). There is therefore a need to establish good aquaculture production facilities in Southern Africa that are capable of generating economically viable, sustainable products.

1.2. The abalone

Abalone are marine gastropod molluscs belonging to the family haliotidae, genus *Haliotis* (Linnaeus, 1758). Gastropods, literally meaning 'stomach foot', include the snails and slugs, and are characterised by their muscular 'feet' upon which they move. The molluscs are divided into three primary classes – the bivalves such as clams and oysters, cephalopods including squids and octopi, and the gastropods. All three groups are highly diverse, both in body plan and physiology, and many occur in marine environments.

The *Haliotis* genus is identified by their flat, ear-shaped shell with respiratory holes running down the anterior margin. There are six species of *Haliotis* occurring in the coastal waters of South Africa (Sales and Britz, 2001). The commercially viable South African abalone species, *Haliotis midae*, is under pressure from extensive poaching in the wild and is currently farmed (Troell et al., 2006). Globally, abalone are sold at a market value of $15 -

$30/kg, a price that has recently been depressed due to the global financial crisis and an increase in farm production (Cook and Gordon, 2010). The flesh of the animal is considered a delicacy, and the shell is a source of iridescent mother-of-pearl.

1.3. Naturally occurring *H. midae* populations in South Africa

The South African coastline is defined by two currents, the warmer Agulhas on the east coast and the cooler Benguela on the west. This has the effect of causing a temperature gradient, which affects the geographic distribution of marine organisms along the coast. The naturally occurring abalone on the west coast are acclimatised to cold water with an average monthly temperature of 12-13 °C (Vosloo and Vosloo, 2010). The optimum temperature for juvenile abalone, based on maximum growth and feed conversion rates determined by Britz et al. (1997) is between 12 and 20 °C. The thermal history of marine invertebrates has been shown to have an effect on their temperature tolerance limits (Osovitz and Hofmann, 2005).

In the wild, *H. midae* is a generalist, opportunistic herbivore (Troell et al., 2006) which typically grazes nocturnally on commonly occurring species of marine algae such as kelp (*Ecklonia maxima*), *Plocamium spp.* and *Ulva spp.* (Sales and Britz, 2001). Gut bacteria isolated from abalone are capable of hydrolysing a range of complex polysaccharides common to their food of choice ((Erasmus et al., 1997; Sales and Britz, 2001; Ten Doeschate and Coyne, 2008), implying a symbiotic relationship with the animal whereby additional digestive enzymes are produced by the gut microflora.

The reproductive biology of *H. midae* has been studied since 1967 (Newman, 1967) and varying reports of the age of sexual maturity have since been published (Tarr, 1995; Wood, 1993), establishing the age of 100% sexual maturity at approximately 7.2 years. Fertilization occurs externally, when the female releases one of two egg groups and these are fertilised by nearby males. The next group of eggs will be released thereafter, in order to increase the probability of fertilization. Many millions of eggs are released, and survivability of the spawn is low.

Wild *H. midae* may grow to approximately 200 mm shell length after 30 years of age (Newman, 1968); however there is significant variability in growth rate between animals. Larger individual *H. midae* are found in greater numbers in the colder waters of the coast of Southern Africa (between 12-13°C), although they can be found in water up to 20 °C -

despite evidence that elevated water temperature can restrict growth rate and reproduction (Sales and Britz, 2001).

Wild stocks of abalone in South Africa are under intense pressure from illegal poaching due to their high value on the black market (Raemaekers et al., 2011) which has led to severe depletion of naturally occurring *H. midae*. As a resource, abalone have proved notoriously difficult to manage (Branch and Clark, 2006), and have been poached to commercial extinction in at least one area in the western cape (Hauck and Sweijd, 1999). Although a ban was placed on the commercial abalone fishery in 2007 (Raemaekers et al., 2011), international poaching syndicates continue to exploit this natural resource (Sundström, 2012).

1.4. The South African abalone aquaculture industry

Commercial aquaculture of *H. midae* began in 1949, with successful cultivation of spawn only occurring from 1981. The practice has been the subject of substantial research and development since the 1990s, when the University of Cape Town, Rhodes University and the Council for Scientific and Industrial Research entered into partnership with fishing companies (Sales and Britz, 2001).

Internationally, total abalone production reached 22 677 metric tonnes in 2002, of which 1731 metric tonnes was produced by South Africa (Gordon and Cook, 2004). This figure includes illegal production, defined as any harvest beyond the total allowable annual landing quotas – South Africa remains one of the biggest sources of poached abalone internationally (Gordon and Cook, 2004; Cook and Gordon, 2010). In 2008, farm production alone was about 30 760 metric tonnes (Cook and Gordon, 2010), representing a 350% growth in the production of abalone farms over the preceding 6 years. The rapid decline in wild abalone stocks and increasing international demand have been the major driving force for the prevalence of commercial abalone farms, and in South Africa favourable water quality and availability of kelp has led to the success of the industry (Troell et al., 2006) – in the international market, South Africa produced the third highest metric tonnage of cultured abalone in 2002. There were at least twelve farms commercially farming the animal in South Africa in 2008 (Ten Doeschate and Coyne, 2008), employing approximately 1360 people (Troell et al., 2006).

Generally the farming practice involves land based tanks into which seawater is pumped and/or recirculated. Water temperature in these tanks, which are often exposed to direct sunlight, can fluctuate by as much as 4.1 °C daily in the summer months (Vosloo and Vosloo, 2010). Juvenile animals are either produced in on-site hatchery facilities or purchased from other farms, although most farms have their own hatcheries in an effort to remain self-sufficient and minimise the risk of infection. Farmed animals are fed the kelp *E. maxima* as well as commercial, formulated feeds Abfeed™ and Midae Meal™, and take approximately 4 years to reach market size (80 g) (Sales and Britz, 2001).

1.5. The molluscan immune system

Throughout evolutionary history, it has been vitally important for organisms to possess defence systems against the varied and virulent array of pathogenic organisms that have evolved alongside them. In vertebrate animals, the immune function is carried out through two different yet connected systems – acquired and innate immunity. The acquired or 'adaptive' system is mediated by cells known as B and T lymphocytes. This system allows the body to mount increasingly effective responses against repeat infections due to acquired cellular 'memory'. The innate immune system lacks B and T lymphocytes, and is mediated by both humoral and cellular responses.

In invertebrate animals, the immune response is carried out solely through the innate system, which is thought to be evolutionarily older than acquired immunity. Innate immunity involves humoral responses such as the production of anti-microbial peptides and reactive oxygen species (Marmaras and Lampropoulou, 2009), as well as cellular responses by circulating haemocytes. Like all molluscs, abalone possess an innate immune system, although little is understood about the abalone immune system and specifically its response to stress. The preponderance of knowledge in this case is based largely on studies conducted on other invertebrates, and the innate immune system of vertebrates (Hooper et al., 2007a) although the degree of comparability between vertebrates and abalone is not fully comprehended. There is overlap between both the mechanisms and the molecules involved in the invertebrate stress response and that of vertebrates (Ottaviani and Franceschi, 1997). Molecules that are produced by a number of organs during the mammalian stress response (including cytokines, biogenic amines, CRH and ACTH) are all produced by invertebrate haemocytes (Hooper *et. al.*, 2007a).

Within the body cavity of molluscs is a fluid analogous in function to blood in vertebrates known as haemolymph, which is responsible for transporting nutrients and waste products, as well as haemocytes (Marmaras and Lampropoulou, 2009). Haemocytes have the capacity to perform a wide variety of functions including digestion, metabolite transport, wound and shell repair and, notably, the immune response. Not all types of haemocytes are found in all species of animals, and the classification of molluscan haemocytes is a field which requires a large volume of additional research (Martin et al., 2007). Haemocyte classification, carried out by means of microscopy, differential centrifugation, flow cytometry, lectin and antibody binding and functional studies (Hine, 1999; Martin et al., 2007; Travers et al., 2008), has determined that the primary molluscan haemocytes are hyalinocytes and granulocytes (Wang et al., 2008). Travers et al. (2008) however, showed that the abalone *Haliotis tuberculata* haemocytes were comprised of only large and small hyalinocytes, whereas classification of the haemocytes of *H. midae* has not yet been reported in the literature.

In their immune capacity, both hyalinocytes and granulocytes can form pseudopodia, aggregate to a wound and produce reactive oxygen species, chemokines and anti-microbial peptides (Travers et al., 2008; Wang et al., 2008). Hyalinocytes have a high nuclear to cytoplasmic ratio and are less able to phagocytose foreign materials, whereas granulocytes have a low nuclear to cytoplasmic ratio and are highly effective at phagocytosis of foreign materials (Martin et al., 2007). Phagocytosis is the primary means by which apoptotic cells and foreign material are removed from the body of the animal. When an apoptotic cell or pathogen (known as the 'target cell') is recognised as such, the macrophage will engulf and degrade the particle when the phagosome fuses with the lysosome (Travers et al., 2008). Recognition is mediated by surface receptors on the phagocyte, which are activated by the target cells (Marmaras and Lampropoulou, 2009). The process of phagocytosis is strictly controlled and highly complex, requiring a succession of interactions between the target cell and the phagocyte. The act of phagocytosis is carried out by re-arranging the cytoskeleton of the cell, and it has been shown that blocking actin polymerisation leads to a loss of phagocytic ability (May and Machesky, 2001).

1.6. The link between stress and the immune response

Stress and disease are major influences acting on the biological processes of an animal's metabolism, as they disrupt the homeostasis of the body and lead to impaired growth,

which is of clear economic importance to farmers (Lacoste et al., 2002). The term 'stressor' refers to the cause of the 'stress' which is the response in the animal (Hooper et al., 2007a). The vertebrate stress response is mediated primarily through neuroendocrine messengers, such as glucocorticoids and adrenocorticotropic hormone (Lacoste et al., 2002). During the stress response, these messengers will typically divert energy away from non-vital functions such as growth and reproduction in order to provide energy for functions which will allow for the proper physiological response to the stressor. For example, in a fight or flight scenario, respiration is enhanced to facilitate locomotion (Lacoste et al., 2002).

There is less information available concerning the effect of stress on the immune system of invertebrates. Bacterial diseases, such as *Vibrioses*, and Schistosomes (Lacoste et al., 2002; Travers et al., 2007) have become problematic and ultimately costly to the abalone farming industry. Lacoste et al. (2001) showed that the principle stress hormone noradrenaline elicits an inhibitory effect on the phagocytic ability of oyster (*Crassostrea gigas*) haemocytes. As stated previously, phagocytosis is the means by which pathogens are removed from the molluscan body and as such this inhibitory effect would allow for opportunistic infection. The response to stressors such as oxygen limitation, shaking, altered salinity and elevated temperature (Hooper et al., 2007a) in abalone has been linked to the onset of disease (Lacoste et al., 2002; Malham et al., 2003) due to decreased immune capacity and therefore must be closely monitored.

Globally, aquaculture of invertebrates such as molluscs and crustaceans is of vital concern, because it contributes significantly to the economy of a number of developing countries. The major obstacle to overcome in such aquaculture situations is the prevalence of diseases that may have a devastating effect on the cultured species (Bachère, 2003). *Vibrio* species are of particular concern, as they are thought to cause many diseases that are harmful to larval and adult molluscs (Bachère, 2003). Currently, as little is known regarding the exact physiological interactions between pathogen and host in the aquaculture environment, there are few options available for the detection of stress or infection in aquacultured invertebrates.

Common *in vitro* tests used to indicate stress and/or disease in aquacultured invertebrates include the haemocyte count, phenoloxidase measurement, superoxide production, phagocytosis assay and bacterial clearing assays (Hooper et al., 2007a). Although the above

mentioned tests have been developed and optimised for a wide range of animals, most have not been specifically optimised for invertebrates and therefore there have been inconsistencies in their results when applied to abalone in the literature (Hooper et al., 2007a). These tests are based on biochemical pathways that have been well documented in other species, but may still require further research in abalone systems.

1.7. The heat-shock response

Thermal stress disrupts many cellular processes, because different cellular enzymatic reactions have different optimal thermal conditions. Heat shock is therefore potentially extremely disruptive of cellular homeostasis (Suzuki and Mittler, 2006). The uncoupling of enzymatic pathways results in an imbalance in the flux of electrons, which can be transferred to molecular oxygen (O_2) to form reactive oxygen species (ROS) such as O_2- and H_2O_2 (Suzuki and Mittler, 2006). ROS can be extremely damaging to organisms undergoing stress as they can disrupt attempts to return to homeostasis by damaging lipids, proteins and DNA, which can lead to cellular apoptosis (Lesser, 2006).

The cellular heat-shock response is well characterised in some species, and numerous genes and proteins exhibit stress mediated differential expression or activity (Kültz, 2005). It was first shown that a protein responded to sub-lethal heat stress in the 1960s in *Drosophila* (Santoro, 2000), and by the 1980s scientists had demonstrated that there were specific proteins that were induced by heat stress (Feder and Hofmann, 1999). This family of proteins became known as the heat shock proteins (Kregel, 2002), and is categorised according to the molecular weight of the protein (Clark and Peck, 2009). In mammalian cells, the heat shock proteins vary in size from 10 to 110 kDa, including the well-known Hsp70 and 90 groups. Some of these proteins, such as Hsp70, have been characterised extensively and are well understood (Clark and Peck, 2009) – however as their varied biological functions are discovered it is becoming clearer that they are not merely 'heat' stress proteins, but assist in a diverse and vital range of stress responses (Morimoto, 1998; Samali and Orrenius, 1998). Heat shock proteins are involved in cellular metabolism, both under stress conditions and in normal functioning of the cell (Morimoto et al., 1996), and heat inducible proteins have been shown to be induced by other abiotic factors such as O_2 depletion and toxic substances (Sørensen et al., 2003; Tkácová and Angelovicová, 2012).

In order to respond adequately to environmental stresses, the heat-shock response must be under the control of rapid, stress inducible regulation (Morimoto et al., 1996). The

underlying stress signal thought to activate the transcription of heat-shock response genes is the accumulation of non-native or misfolded proteins (Santoro, 2000). Heat-shock response genes code for molecular chaperones, which assist proteins in returning to their native state before extensive cellular damage is caused. The cytoprotective role of heat-shock proteins may confer a transient resistance to diseases and other stresses in the stressed organism (Feder and Hofmann, 1999). However, long-term expression of stress response proteins could have negative effects on the overall health of the organism (Morimoto et al., 1996).

It has been proposed that the characterisation of molecular stress response proteins could facilitate biomonitoring of the status of aquacultured animals (Dimitriadis et al., 2011). Heat shock proteins are an attractive candidate for biomonitoring studies (Feder and Hofmann, 1999) because of the sensitivity of their induction (Dimitriadis et al., 2011), especially in non-model organisms because they are highly conserved between species which facilitates characterisation (Clark and Peck, 2009). Heat shock proteins, however, are only one aspect of the cellular response to heat stress. Other families of proteins with varying functions have also been shown to respond to heat stress (Abele et al., 2001). A meta-study carried out by Wang et al. (2009) of frequently identified stress response proteins across 66 studies identified 44 proteins that were expressed during cellular stress regardless of species. These proteins include representatives from 28 generally detected protein families, with functional classification ranging from cytoskeletal organisation and energy metabolism to molecular chaperones.

1.8. The molluscan response to thermal stress

The cellular response to stress is complex and diverse, and characterising the molluscan response to thermal stress is therefore a challenging task. However, it is important to understand how economically valuable animals such as *H. midae* respond to stress in order to minimise stock losses on abalone farms. Since no annotated complete genomic sequence exists for *H. midae*, what limited information is available is deduced based on work that has been carried out on model invertebrates and other molluscs (Hooper et al., 2007a). A brief overview of the available literature describing the molluscan thermal stress response will follow, emphasising the research undertaken on mussels, oysters and abalone.

1.8.1. Mussel

The genus *Mytilus* is fairly well represented in the literature as it is a North American coastal resident which occurs on both the Atlantic and Pacific Ocean coasts at a range of temperatures and salinities. *Mytilus* is intertidal and sessile, and has a dark shell which absorbs sunlight and therefore increases the heat exposure of the animals, which means that it must be able to withstand thermal stress in its natural environment (Hofmann and Somero, 1995). Studies characterising the mussel response to stress tend to focus on the haemocytes (Winston et al., 1996; Canesi et al., 2006; Donaghy and Volety, 2011; Wang et al., 2012), as bivalve haemocytes respond to external stimuli and reflect the physiological status of the animal (Donaghy and Volety, 2011). Since there are several species within the genus which occur at different coastal temperatures, these closely related organisms, or congeners, are ideal candidates for studying the molecular response to acclimatisation dependent thermal stress. Fields et al. (2012b) compared the proteomic response of the cold acclimatised indigenous species *Mytilus trossulus* to that of the invasive Mediterranean species *Mytilus galloprovincialis* under chronic and acute thermal stress both at low and high temperatures. Since the indigenous species is being displaced by the invader along the Californian coast, Fields et al. (2012b) used 2-dimensional polyacrylamide gel electrophoresis (2D PAGE) to determine whether the invasive species had a physiological advantage in warmer habitats. Clustering analysis revealed differential regulation of cytoskeletal proteins, electron transport chain proteins and molecular chaperones including Hsp70 in both species of mussel. *M. galloprovincialis* appeared to be the more robust species in response to heat shock, which may serve as an advantage in its spread and integration into new environments (Yao and Somero, 2012). This work parallels the work of Hofmann and Somero (1995), who had demonstrated that field-collected *M. trossulus* exhibited increased Hsp70 expression and increased ubiquitination of proteins during the summer months, when their body temperature increased from 10° to 30°C during low tide. Ubiquitination is a post-translational modification that occurs when proteins are irreversibly damaged and must be targeted for degradation, so it is often used as a measure of the ability of an organism to cope with its current environmental state. Fields et al. (2012a) investigated the protein expression patterns of a species of mussel (*Geukensia demissa*) with a very broad natural temperature range during exposure to acute heat stress, and identified similar stress response proteins – cytoskeletal proteins, chaperones, oxidative stress proteins, regulatory proteins and a transcription initiation factor. They identified seven Hsp70 isoforms that were significantly regulated in response to acute heat stress.

Mytilus edulus was used to characterise the oxidative stress response of the mussel to the pro-oxidant menadione (McDonagh and Sheehan, 2007). Since this mussel species is an intertidal filter feeder, it has been identified as a 'sentinel species' which could serve as an early warning for toxic pollution in an aquatic system. The group identified post-translational modifications affecting actin, amongst other proteins, which could be part of a molecular system for sensing the redox state of the cell.

1.8.2. Oysters

The oysters are intertidal bivalve molluscs, which are filter feeders and therefore under constant pressure from pathogen infection and a highly changeable environment (Tirard et al., 1995). The calcified shell of the oyster acts as a physical barrier to harm, but acidification of the oceans can prevent shell formation in larval oysters (Waldbusser et al., 2013), and this ancient defence mechanism may no longer afford sufficient protection from environmental stresses (Zhang et al., 2012a). The loss of shell integrity due to decalcification has consequences for any shelled oceanic mollusc, including the abalone (Byrne et al., 2011).

The annual production of the oyster *Crassostrea gigas* exceeds any other aquatic organism (Fleury et al., 2009), and the oyster has therefore been the subject of more intensive study than other molluscs. *C. gigas* recently became the first mollusc to have its genome fully sequenced (Zhang et al., 2012a). The thermal stress response capabilities of the oyster that can be deduced from its genome are of interest when studying the abalone, as it is the nearest evolutionary relative for which this information is available. The predicted oyster genome was found to contain 88 Hsp70 genes, compared with ~17 in humans and 39 in sea urchins, which may relate to the oyster's natural exposure to fluctuating temperatures as a result of their location in the intertidal zone. Zhang et al. (2012a) found 5844 genes that were differentially expressed in the transcriptome of oysters subjected to 9 stresses including temperature and air exposure. Genes involved in the unfolded protein response (UPR) and the endoplasmic response to stress including chaperones like calreticulin were up-regulated, indicating that the return to cellular homeostasis was a priority in stressed oysters. Although oysters and abalone may have diverged as much as 500 million years ago (Hooper *et. al.*, 2007), the molecular basis for the stress response may be similar between these classes of mollusca, and methodologies that have proved successful in classifying the oyster stress and immune response may also be applied to abalone.

Like the abalone, the haemocytes of the oyster are multifunctional, mobile cells which are capable of mounting cellular and humoral defence against biotic and abiotic stresses. The stress response and characterisation of oyster haemocytes is therefore a focus for research (Tirard et al., 1995; Hégaret et al., 2003b; Cao et al., 2009; Thompson et al., 2011). Techniques employed to characterise the stress response of the oyster have included 2D PAGE, which has been employed to detect proteomic differences in disease resistant vs. susceptible oyster species (Cao et al., 2009) and the effect of metal contamination on the oyster immune response (Thompson et al., 2011). A number of oyster gene sequences have been determined independently to the oyster genome sequencing project, and several stress response pathways have been suggested using evidence obtained by quantitative PCR (Meistertzheim et al., 2007; Nikapitiya et al., 2010). Due to the lack of information available for molecular methodologies applicable to abalone specifically, the body of research available for mussel, oysters and other marine molluscs is extremely useful in the pursuit of abalone research.

1.8.3. Abalone

Some studies have been carried out to characterise the response of abalone to environmental perturbation, although there has not been extensive progress in this field. Investigations into the abalone stress response have focussed primarily on the effect of exposure to toxic compounds and pathogenic infection, with research on the effect of thermal stress focussed mainly on the response of heat shock proteins. Heat shock elicits behavioural responses in abalone including uplifted shell gaping, epipodial spreading (Diaz et al., 2006) and movement (Robinson et al., 2012); and physiological reactions such as changes in oxygen consumption, nitrogen excretion and total haemocyte count (Vosloo and Vosloo, 2010; Dang et al., 2012). Since different species of the *Haliotis* genus occur naturally over a wide range of temperatures, some acclimation dependent heat shock responses have been observed (Dahlhoff and Somero, 1993; Diaz et al., 2006; Li et al., 2012). *H. midae* appears to have a critical thermal maximum of about 28 °C (Hecht 1994), although temperatures of 12-20°C have been shown to be physiologically optimal (Sales and Britz, 2001). When *H. midae* are exposed to temperatures exceeding 20 °C for long periods of time, their growth rate is decreased and mortality increased compared to animals maintained at lower temperatures (Sales and Britz, 2001). Although this has been established through scientific enquiry, there do not appear to be measures in place at abalone farms to prevent water temperatures from exceeding 20 °C during the summer

months (Vosloo and Vosloo, 2010). Martello et al. (2000) showed that stressful conditions (exposure to increased salinity and the aquatic toxin pentachlorophenol) decreased phagocytic ability and impeded cytoskeletal integrity in both the red and black abalone. As phagocytosis is the basis by which abalone haemocyte cells respond to pathogenic infection, they conclude that environmental stress has an immunosuppressive effect in abalone.

Real-time qPCR has been used to determine gene expression profiles of candidate stress genes in several abalone species. Wan et al. (2009) characterised a novel omega glutathione S transferase (GST) in *Haliotis discus discus* and documented increased GST transcription in response to heat shock, heavy metal and endocrine disrupting chemical (EDC) exposure. The glutathione S transferase (GST) family of proteins are involved in detoxification, as they are responsible for the elimination of harmful reactive oxygen species. Hsp70 and Hsp90 have been characterised in *Haliotis tuberculata* (Farcy et al., 2007), and respond to elevated temperature with increased transcription of both genes. Proteomic approaches to characterising the stress response of abalone have so far been limited to 2D PAGE analysis of proteins that are differentially expressed in response to EDCs (Zhou et al., 2010). Aldehyde dehydrogenase and glutathione S transferase, which are both involved in detoxification, along with superoxide dismutase, an oxidative stress response protein, and β tubulin, a cytoskeletal element, responded to treatment with bisphenol A and diallyl phthalate. While this study concluded that proteomics was a powerful platform for the study of aquatic ecotoxicology, Zhou et al. (2010) is currently the only published study that has used proteomics to characterise the abalone stress response. The field of proteomics is rapidly gaining favour in marine biotechnology and aquaculture (Slattery et al., 2012) due to its capacity to characterise transcriptional, translational and post-translational regulatory mechanisms that respond to the effects of environmental changes simultaneously.

1.9. Proteomics and the characterisation of the stress response

It is feasible to quantify the effects of an environmental stimulus on the regulation of a gene using technology such as real-time qPCR, because of the assumption that regulation of mRNA transcription is directly related to the biological processes occurring in a cell. However, there is mounting evidence that there is disjunction between the amount of mRNA in a cell and the biological effect of the protein translated from it (Liebler, 2002; Rodrigues et al., 2012). This could be due to a number of factors, not least the complex post-transcriptional modifications applied to many gene products, and thus there may be functional dissonance between mRNA expression and protein abundance (Rodrigues et al., 2012). However, since proteins are the bioactive molecules, measuring the varying concentration of a particular protein in a cell exposed to altered environmental states represents a more direct measure of the molecular stress response mechanisms (Kennedy, 2002; Tomanek, 2011). A protein or group of proteins which respond predictably to environmental stress may therefore provide a more meaningful quantifier of the effect of that stress (Kennedy, 2002).

The term 'biomarker' has been defined by Hawkridge and Muddiman (2009) as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention", but in aquaculture biomarkers have potential for use in environmental monitoring and risk assessment (Rodrigues et al., 2012). This has proven especially useful in the field of aquatic toxicology (Sanchez et al., 2011), where biomarkers can be used to identify aquatic pollutants (Nesatyy and Suter, 2007) and also improve our understanding of how aquatic organisms respond to pollution (Tomanek, 2011). Technology, like proteomics, that is capable of identifying several hundred biomarkers simultaneously therefore has great potential as a powerful tool in aquaculture research. Proteomics in aquaculture is a rapidly growing field which is primarily concerned with the welfare, health, quality, safety and nutrition of aquacultured organisms (Rodrigues et al., 2012). Despite the urgent lack of molecular information for several economically valuable aquacultured animals, there are still relatively few publications describing the application of proteomics techniques in the field of aquaculture. This is due, in part, to difficulties inherent in the methodology and technology (Slattery et al., 2012; Zhou et al., 2012). To properly evaluate the potential of proteomics in aquaculture studies, it is first necessary to understand some of the technical advantages and limitations inherent to mass spectrometry-based analyses.

1.9.1. An introduction to proteomics technology

High throughput methods for studying the proteome have only recently become available (Nesatyy and Suter, 2007). The genomic revolution paved the way for proteomics by providing reliable databases of genomic sequences, bioinformatics tools capable of performing user-driven searches on these sequence databases, and particularly, the algorithms necessary for analysing high throughput genomic datasets such as microarrays (Liebler, 2002; McDonald and Yates, 2002). Traditional genomics is now moving into a functional phase where gene function can be classified rather than merely accumulating a list of gene names found in a sample (Harry et al., 2000). This means that it is now possible to perform complex clustering and mapping studies of a large number of gene identities, as well as any associated quantitative information. This technology has in turn made the study of proteomics - that is, the identification of all expressed proteins in a cell including post-translational modifications and protein-protein interactions - possible (Tomanek, 2011; Zhou et al., 2012). Proteomics is challenging because there are no analytical tools analogous to the core genomic protocols of PCR and DNA-DNA hybridisation available in the proteomic toolkit (Liebler, 2002). In order to quantify proteins in a cell it is necessary to measure the protein directly, sometimes at extremely low relative concentrations, as it is impossible to amplify more copies of a target protein as it is with a gene of interest using PCR (Harry et al., 2000; Lottspeich, 2009). Additionally, some proteins naturally make up the majority of the total protein in the cell, and might conceal the response of less abundant proteins. Therefore it is necessary for proteomic methods to achieve the sensitivity required to detect low abundance proteins. The difficulty in interpretation of proteomics data is further compounded by another feature of the proteome, which is the presence of post-translational modifications (Gevaert and Vandekerckhove, 2011). Multiple isoforms of the same protein can be present in a cell at any time, which adds further complexity to the proteome (Liebler, 2002). These complications have led to the development of incredibly sensitive and complex technology with which the proteome may be studied.

1.9.1.1. The mass spectrometer

Mass spectrometry is a powerful tool that works by first ionising, then fragmenting and measuring the mass/charge ratio of molecules in order to create spectra that can be used to identify the molecules. The molecule that is being analysed is referred to as the analyte.

All mass spectrometers must consist of an ion source, which is necessary to ionise the sample (Liebler, 2002), a mass analyser which resolves the ions based on their mass to charge ratio (m/z) and a detector (Aebersold and Mann, 2003). In order to measure the mass of incredibly small molecules, mass spectrometer measurements must be carried out on charged analytes that are in the gas phase, which necessitates ionisation of the sample prior to analysis (Yates, 1998). Mass spectrometry was used extensively in non-biological fields for many years before it was applied to the biological sciences (Mann et al., 2001). The invention of so called 'soft' ionisation techniques made it possible to fragment and ionise large, polar biomolecules (Mann et al., 2001; Liebler, 2002). The field of proteomics is generally interested in the qualitative and quantitative assessment of proteins in a tissue or sample, however in recent years lipids and other metabolites have also become popular areas of research (Scalbert et al., 2009). Due to the necessity for ionisation there is an inherent loss of sensitivity with increasing molecule size, which becomes problematic since proteins can be up to several hundred kilo-Daltons in size. In proteomics, it is therefore common practice to first digest the protein sample into smaller fragments using a protease enzyme with a known recognition site. This so called 'peptide centric' or 'bottom up' proteomic workflow has become the standard for the analysis of the proteome (Gevaert and Vandekerckhove, 2011). The most commonly used enzyme is a serine protease called trypsin, which cleaves protein between lysine and arginine residues, unless either is followed by a proline in the C terminal direction (Finehout et al., 2005). This also confers an important feature to all resulting tryptic peptides – they all possess lysine or arginine residues at their C termini and N terminal amino groups (Liebler, 2002). In addition, the frequency at which lysine and arginine occur in the proteome is such that the peptides resulting from tryptic digestion are often of an appropriate length for ionisation and analysis by mass spectrometry (Liebler, 2002).

Two different types of ionisation sources are typically used for proteomic studies; namely MALDI and electrospray ionisation (ESI) (Liebler, 2002). One of the earliest methods used to ionise biological molecules for mass spectrometry analysis was matrix assisted laser desorption ionisation (MALDI) (Mann et al., 2001), which was invented in 1987 (Liebler, 2002). This technology remains in use to this day and is utilised by some of the most sensitive and widely used mass spectrometers. MALDI-TOF analysis is generally acceptable for less complex peptide mixes, for example the identification of a protein from a single spot on a 2D PAGE gel (Aebersold and Mann, 2003), whereas liquid chromatography (LC) coupled with ESI is generally used for more complex samples.

ESI requires that the analytes are resuspended in an aqueous phase, which is called the mobile phase (Liebler, 2002). The analytes in solution are sprayed through a very fine needle into a charged field (Gevaert and Vandekerckhove, 2011) to electrostatically disperse the tiny droplets (Mann et al., 2001). At the orifice of the needle the spray forms a "Taylor cone", releasing charged solvent droplets containing peptides (Gevaert and Vandekerckhove, 2011). The electrical field directs the droplets into the mass analyser. A drying gas is then passed over the droplets, causing evaporation of the solvent until the repulsion of the like-charged ions overcomes the surface tension of the droplet and a Coulomb explosion occurs, shattering the droplet. This occurs until only the peptides remain and pick up the charge before passing into the mass spectrometer.

The pH of the solution will control which amino acid functional groups carry charge, as different groups are protonated at different pHs depending on their physical properties. Amine groups will be protonated under acidic conditions which will confer an overall positive charge to the peptides. Under basic conditions, amines and carboxyl groups are deprotonated and the peptide becomes more negatively charged. A positively charged peptide can be fragmented more easily (Liebler, 2002), but unlike MALDI, the peptide ion products derived by ESI can carry single or multiple charges. Since tryptic peptides have a lysine or arginine residue at their C termini and N terminal amino groups, which are the sites of protonation, ESI of tryptic peptides generally produces multiply charged positive peptide ions. One of the risks of using ESI is that, unlike MALDI, there is no opportunity to re-analyse a sample – once it has been subjected to ESI it is consumed and cannot be re-examined (Wang and Li, 2008). In order to maximise the number of peptide identities obtained in ESI MS/MS, a nano-HPLC column is often included in the machine setup before ESI occurs, so that the incoming peptides have been pre-fractionated before mass analysis occurs (Mann et al., 2001). Sample fractionation will be discussed in some detail below.

ESI sources are generally paired with one of four types of mass analysers: quadrupoles, ion traps, Fourier transform ion-cyclotron or orbitrap (Aebersold and Mann, 2003). Although the technology for each type is different, the role of the mass analyser is always the same in each mass spectrometer (Liebler, 2002) – that is, the resolution of peptides or peptide fragments, and the accurate measurement of their mass to charge ratio (m/z) (Aebersold and Mann, 2003). For the initial MS analysis, the ionised analyte travels from the ion source into a charged chamber, which uses positive charges to direct and filter the

movement of the ionised analyte towards the detector. The detector is usually at the base of a tube, so that incoming ions can be reflected upwards into the tube where they travel in an arc pattern before hitting the detector. The measurement of each ion's m/z is plotted on a graph, creating a spectrum of peaks corresponding to each ionised peptide in the sample. The so called 'mass spectrum' plots the m/z ratio for each parent ion vs. the relative intensity of the signal, which is loosely correlated to the abundance of the peptide ion in the sample. In order to then obtain amino acid sequence information from the parent ions, a second round of mass analysis is then carried out, known as MS/MS or MS2.

After the ionised sample has been passed through the mass spectrometer once, producing a m/z spectrum, the representative peptide/s for a single m/z peak is selected by the tandem MS analyser from the mixture of ionised peptides. This 'parent ion' is channelled into the mass analyser, and then fragmented in a collision cell by collision induced dissociation (CID), or other methods of fragmentation (Jones and Hubbard, 2010), into 'fragment ions' which pass through the mass spectrometer where their m/z is once again measured, producing a second m/z spectrum. The resultant spectrum, which is a record of the m/z ratios for all the fragment ions produced for a specific precursor, is the basis for the deduction of the amino acid sequence of the peptide (Aebersold and Mann, 2003), using technology which will be discussed below. The different types of mass analyser are made distinct by the method by which the ionised analytes are channelled towards the detector, and retained for MS/MS analysis. Part of the complexity of the mass spectrometer machinery is the need to perform MS and MS/MS analysis in tandem, while accurately retaining information for each parent/fragment ion. The rate at which the machine can direct, fragment, and measure m/z for each ion will directly influence the number of parent ions that can be fragmented for MS/MS analysis. Loss of sensitivity, and therefore the resolution at which the sample can be processed, is of great concern and therefore new innovations are constantly being introduced in order to improve the accuracy and sensitivity of the mass analyser.

1.9.1.2. Reduction in proteome complexity

The proteome of a cell can be very complex (Gevaert and Vandekerckhove, 2011), and in a "bottom-up" proteomics workflow the digestion of the proteome into tryptic peptides increases the complexity of the mixture (Ernoult and Guette, 2011). The detection capabilities of any mass spectrometer are limited by the amount of data that can be

captured by the machine (Hubner et al., 2008), because the chances of identifying a peptide decrease as the total number of peptides increases (Liebler, 2002). Therefore, the reduction of sample complexity prior to analysis by mass spectrometry is a fundamental element of any proteomics workflow (Liebler, 2002). This can either occur alongside the analysis in the mass spectrometer, or separately without the use of the mass spectrometer. Methods that do not make use of the mass spectrometer include both gel-based and gel-free technology. One of the most common gel-based methods for separation is two-dimensional polyacrylamide gel electrophoresis (2D PAGE) which separates whole proteins on a polyacrylamide gel first by isoelectric point and then by molecular mass (Pietrogrande et al., 2003). The protein spots can be excised from the polyacrylamide gel, digested with trypsin and analysed by the mass spectrometer individually in order to obtain the protein identity (Kennedy, 2002). Although gel-based methods are popular, there are limitations to the methodology which will be discussed in the following chapters. As an alternative, gel-free separation can be used to achieve separation of peptides in the liquid phase (Hubner et al., 2008). Separating proteins or peptides while in the liquid phase can be beneficial because the sample can then be applied directly to the mass spectrometer for analysis, which reduces the number of intermediary steps and can help to minimise sample loss.

In order to achieve the best possible separation of peptides prior to analysis by mass spectrometry, multidimensional separation strategies can be employed (Kubota et al., 2009), such as a combination between 2D PAGE and LC MS/MS, sometimes called geLC MS/MS (Hubner et al., 2008). After analysing a peptide sample by mass spectrometer in this manner, a raw data file is produced containing spectra for each peptide detected by the mass spectrometer. In order to identify the amino acid sequences represented by each spectrum, and thereby identify the protein that the peptide originated from, it is necessary to use computational analysis to match mass spectra to amino acid sequences.

1.9.1.3. Software for proteomic investigations

The m/z spectrum produced by the fractionation of a peptide in MS/MS analysis corresponds directly to the amino acid sequence of the peptide, because the physical properties of the peptide determine how it fragments in the collision chamber. Different methods are available when computationally determining an amino acid sequence based on a mass spectrum, usually referred to as MS/MS database searching or *de novo* approaches.

De novo sequence determination makes use of the mass spectrum generated by MS/MS directly in order to elucidate the amino acid sequence of the peptide. In *de novo* sequencing, collision induced dissociation (CID) is typically used to fragment the parent ions into fragment ions, and this method of fragmentation breaks the peptide backbone at predictable sites (Seidler et al., 2010). CID often produces complementary charged fragments of peptides known as b and y ions, which overlap with each other. Therefore it is possible to computationally determine which amino acids made up the peptide in order to produce the corresponding mass spectrum based on the 'ladder' of b/y ions detected. One of the disadvantages of the *de novo* sequencing methodology is the computational intensity required to assemble amino acid sequences from a complex sample. Furthermore, the amino acid sequence of a peptide does not verify which proteins are present in the sample, which is a requirement for many proteomics studies (Nesvizhskii, 2007). An alternative to *de novo* sequencing is database searching, which can overcome these disadvantages.

With the advent of high-throughput genomic tools and protein sequence prediction software, several species' entire proteome sequences are stored in databases. It is possible for the user to input a database of proteins that are expected to be found in the sample to be analysed by mass spectrometry, for example the complete proteome of the species in question. The collection of protein sequences in the database is then subjected to *in silico* 'digestion' by denoting the protease used on the protein sample undergoing analysis. The resultant predicted peptides can then be filtered to correct for post-translational modifications, before *in silico* fragmentation, according to pre-set fragmentation rules. The resultant theoretical fragmentation patterns are then matched to the mass spectra determined by MS/MS analysis of the sample. The match between a peptide and a spectrum is known as a peptide spectrum match (PSM). The software then generates a list of potential PSMs, usually with a score for each denoting the confidence of the match (Liebler, 2002). This score is determined in a manner that is unique to the software used to establish the PSMs, and therefore will vary accordingly. The best PSMs are then used to match a peptide to a protein identity from the database, thereby identifying the proteins present in the sample.

Database searching is carried out using sophisticated proteomics software (Xu and Ma, 2006), which is also required to perform quantitative analysis of protein expression in MS based studies, and *de novo* peptide sequence determination (Jones and Hubbard, 2010).

Many software packages exist, including both proprietary and freely available packages (Schulze and Usadel, 2010), which are capable of database directed searching or *de novo* sequencing, or a hybrid of both (Jones and Hubbard, 2010). Peptide sequence determination by these methods can incorrectly match MS/MS spectra with a peptide sequence, and it is therefore necessary for the software to statistically validate results in order to minimise the chance of a false positive identity. When the need for quality control became topical in the field of proteomics, there were varying theories proposed as to the correct means to minimise false positives in a set of PSMs (Tabb, 2007). Although each PSM is associated with a score denoting the confidence of the match, a threshold for the confidence still needs to be established below which the match should be considered a false positive and disregarded. Differing mathematical models were devised to determine the score threshold, but modern software generally makes use of the false discovery rate (FDR). FDR is defined as the expected proportion of false PSMs within the total number of PSMs (Gevaert and Vandekerckhove, 2011). The software adjusts the score threshold in order to minimise the FDR, thereby ensuring that the maximum number of false peptide identities are removed while removing as few correct identities as possible (Gevaert and Vandekerckhove, 2011). As the goal in any proteomics study is to obtain as many protein identities as possible within acceptable error boundaries, the FDR can therefore be seen as a direct trade-off between sensitivity and error rate (Jones and Hubbard, 2010).

The proteome is a complex and dynamic entity which should be the focus of much future research. As the development of proteomics technology advances we will have increased capacity to identify more peptides from more complex biological samples. Proteomics based research is essentially a multi-disciplinary field, and involves complex technology and intricate data analysis. Care must therefore always be taken when designing a study that makes use of proteomics technology, especially in non-model organisms. In the field of aquaculture, proteomics technology has great potential as a tool for enabling better comprehension of the molecular and cellular mechanisms at work in economically valuable species.

1.10. Aims of this study

Abalone are commercially valuable marine molluscs, which are farmed globally and are an important economical contributor to the aquaculture industry in South Africa (Sales and Britz, 2001; Troell et al., 2006), where *Haliotis midae* is a commonly farmed species. Due to high stocking density and intensive farming practices, farmed abalone experience high rates of disease and mortality which have led to catastrophic losses in the past (Malham et al., 2003; Hooper et al., 2007a). Despite their value and the high rate of growth of the global abalone farming industry, there remains a lack of information describing the immune system of this animal. Abalone possess an innate immune system which is largely facilitated by the activity of mobile haemocytes in their circulatory system. The onset of disease in abalone has been linked to environmental factors (Malham et al., 2003) such as temperature, salinity, oxygen availability and water pollution (Girón-Pérez, 2010; Tomanek, 2011). Biomonitoring of the molecular stress response in abalone haemocytes could provide a valuable insight into the link between environmental stress and the onset of disease in abalone.

On the west coast of South Africa, *H. midae* are acclimatised to the natural water temperature of approximately 13°C. However, on abalone farms during the summer months the day to day average temperature can fluctuate by 4.1°C (Vosloo and Vosloo, 2010). It is therefore necessary to understand how the abalone responds to acute heat shock within their thermal tolerance limits (Hecht, 1994). The aquaculture industry in developing countries has the potential to alleviate food shortages, and provide much needed economic growth. The South African abalone *H. midae* was therefore selected as a subject for this study, due to its economic value to abalone farmers in South Africa. Furthermore, although some progress has been made towards understanding the immune system of other species of abalone, the immune system of *H. midae* remains poorly characterised.

Previous studies investigating the effect of environmental stimulus on the immune response of abalone and other molluscs have tended to focus on a single candidate stress gene, or stress gene family. However, the stress response is a complex interconnected system of biochemical pathways, so characterising a single gene or protein can lead to an incomplete understanding of the system. With the recent development of proteomics technology, it is possible to identify and quantify the response of the entire proteome,

thereby providing a more holistic basis for the characterisation of the abalone stress response. Proteomics can identify proteins that respond to stress stimuli, and also identify groups of proteins belonging to specific biochemical pathways, making it a powerful tool (Campos et al., 2012; Slattery et al., 2012). However, proteomics has had limited applications in marine biomonitoring studies, especially in studies regarding the mollusca, due to the lack of fully sequenced genomic databases (Sanchez et al., 2011).

The aim of this study is to gain insight into the abalone molecular heat stress response using proteomics technology. In order to do so, the main objectives of this study were to:

i) Use a discovery-based proteomic approach to identify proteins that are differentially regulated in *H. midae* during acute exposure to thermal stress
ii) Identify biochemical pathways that are involved in the abalone stress response
iii) To corroborate proteomics results using independent technologies such as qPCR and western blots

Chapter 2

TWO DIMENSIONAL POLYACRYLAMIDE GEL ELECTROPHORESIS ANALYSIS OF THE *H. MIDAE* HEAT STRESS RESPONSE

2.1. Introduction

Aquaculture is a rapidly growing industry which has the capability to provide much needed food and economic security (Rodrigues et al., 2012). By improving our understanding of the biological and molecular mechanisms by which aquacultured species respond to an artificial environment we can maximise production. Although disease outbreaks are common in farmed marine animals, there is still a limited understanding of the mechanisms that direct these organisms' immune response to stress and disease (Rodrigues et al., 2012). Farmers who produce marine molluscs and crustaceans have become particularly concerned by recent disease outbreaks, which threaten the sustainability of aquaculture (Bachère, 2003).

Farmed molluscs, including abalone, are highly susceptible to a number of diseases (Sales and Britz, 2001). *Haliotis midae* is an economically valuable species of abalone that is produced by aquaculture in South Africa where, presently, the incidence of disease has been relatively low (Macey et al., 2011) compared to the high infection and mortality rates in Australia and China (Troell et al., 2006; Hooper et al., 2007b). Although a link between stress and increased susceptibility to disease has been established in abalone (Malham et al., 2003), the molecular mechanisms underpinning this connection are poorly understood. Characterisation of the molecular response of abalone to abiotic stresses they are exposed to on the farm will lead to improved productivity.

Farmed abalone are subjected to daily fluctuations in water temperature, which vary in severity depending on the season (Vosloo and Vosloo, 2010). Rapid changes in water temperature affect animals by altering their metabolic rates (Pallares et al., 2012), which may have an effect on their ability to respond to bacterial or viral infection. Higher temperatures have been linked with increased susceptibility to *Vibrio* infection in the abalone *Haliotis diversicolor supertexta* (Lee et al., 2001). The escalation of disease outbreaks in marine molluscs known as 'summer mortality' is caused by changes in water temperature between summer and winter months (Dang et al., 2012). Previous studies in abalone have used total haemocyte counts as an indicator of immune status, as the number of circulating haemocytes is linked to the immune capacity of the abalone (Cheng et al., 2004a; Hooper et al., 2011, 2014; Dang et al., 2012). Other metrics that have been used to measure the effect of heat stress in abalone include the activity of oxidative stress enzymes (Hooper et al., 2014), haemocyte phagocytic activity (Cheng et al., 2004a) and the expression of the

heat inducible protein Hsp70 (Vosloo and Vosloo, 2010; Li et al., 2012). Since the abalone immune system is poorly characterised in the literature, the precise nature of the effect of acute heat stress on the abalone immune system is unknown. Therefore, this study was undertaken to investigate the effects that rapid, short-term increases in temperature (acute heat stress) have on proteins expressed in abalone haemocyte cells.

Historically, the characterisation of genes and proteins has been a slow, laborious and expensive procedure. With the advent of techniques that allow for more comprehensive and rapid studies of proteins, it is now possible to identify and quantify the response of the proteome to environmental stress (McDonald and Yates, 2002). The separation of a protein sample by electrophoresis on an immobilised pH gradient (IPG) was described by Görg et al. (1988). The technique makes use of the physical properties of proteins to separate a complex protein mixture into a tightly resolved cluster of spots on a polyacrylamide gel. Two dimensional polyacrylamide gel electrophoresis (2D PAGE) combined with protein identification by liquid chromatography/mass spectrometry (sometimes referred to as geLC/MS) was one of the earliest methods by which protein identification and quantitative expression profiling of total cellular protein could be carried out (Görg et al., 2004). 2D PAGE remains a popular method in quantitative proteomic investigations, due to its reproducibility and sensitivity.

The separation of the protein sample in the first dimension by 2D PAGE takes place on a gel strip exposed to electrical current, where carrier ampholytes resolve into a gradient over a pH range on the strip. The proteins in the sample move along the gradient according to their charge (Liebler, 2002). The isoelectric point, or pI, of a protein is the pH at which it is no longer affected by negative or positive charge in an electric field. When a protein reaches the point on the strip which corresponds to its pI, it remains stationary. The strip can then be applied to a polyacrylamide SDS gel for electrophoresis in the second dimension, which separates the proteins by molecular weight. The resulting gel contains a reproducible pattern of protein spots, which allows for the determination of the relative quantity of each protein. 2D PAGE remains unique amongst proteomic methods in its ability to detect thousands of proteins simultaneously (Boguth et al., 2000). The use of 2D PAGE technology in aquaculture studies has shown recent promising growth (Rodrigues et al., 2012), as it is capable of identifying and quantifying a large number of proteins. However, this technology remains underutilised in the characterisation of the effects of environmental stress on marine aquacultured species (López et al., 2002). Currently there

are no 2D PAGE based studies describing the molecular mechanisms by which the South African abalone, *Haliotis midae,* responds to acute heat stress.

2.1.1. Aim of this study

The onset of disease in abalone may be linked to environmental stresses such as oxygen depletion, increased salinity, mechanical shaking (simulating forced movement) and changes in temperature (Hooper et al., 2007a). *Haliotis* spp. are considered to be non-model organisms since the full genomic sequence of any member of this genus is not currently available. As such, the molecular basis of the abalone stress response is poorly understood. The aim of this study was therefore to characterise the response of the *H. midae* proteome to an acute increase in temperature, and furthermore to identify specific proteins that can be characterised as components of the molecular stress response of aquacultured abalone.

In order to effectively carry out the aims of this study, an appropriate heat shock temperature had to be selected. Since west coast *H. midae* are naturally acclimatised to 12-13°C (Vosloo and Vosloo, 2010), the suitability of a heat-shock temperature of 23°C was investigated. As determined by Britz et al. (1997), this temperature is higher than the average range for optimal abalone growth, but below the 50% critical thermal maximum of 27.9°C (Hecht, 1994). The ideal heat shock temperature should illicit a physiological response to stress in the animals without permanently weakening the overall health of the animal or causing mortality. Total haemocyte counts have previously been used to assess immune function in abalone, and a rapid increase in the number of circulating haemocytes in abalone responding to acute heat stress has previously been reported (Cheng et al., 2004a; Dang et al., 2012). In order to confirm that the chosen temperature elicited a physiological response, haemocyte counts were carried out at intervals for the duration of a heat shock experiment at 23°C. In order to simplify the analysis, it was assumed that the overall haemocyte cell type composition did not change during exposure to stress.

The second objective of this study was to determine the duration of the cellular response in *H. midae* haemocytes exposed to acute heat stress. This was executed by 2D PAGE analysis of abalone haemocyte protein sampled at timed intervals during exposure to 23 °C water. In addition, a recovery sample was also obtained from animals that were heat shocked for 3 hours and then returned to 13 °C water for 24 hours. The results of this experiment allowed for the selection of appropriate sampling time points for the

identification of proteins that respond to heat shock by 2D PAGE coupled with MS analysis.

The final objective of this study was the use of the available literature to contextualise the identified differentially expressed proteins in this study, and thus suggest biochemical pathways that are affected by acute heat stress in *H. midae* haemocytes.

2.2. Materials and methods

Solutions used in this chapter are listed in Appendix A.
For the purposes of this study, a biological replicate is defined as samples obtained from distinct groups of abalone under the same environmental conditions. A technical replicate is the replication of the same experimental technique on the same sample.

2.2.1. Animal maintenance and sample collection

Cocktail size abalone (6 cm) were maintained at the Department of Agriculture, Forestry and Fisheries (DAFF) research aquarium at Sea Point, Cape Town, Western Cape, South Africa. Baskets of about 30 animals were kept in darkened 100 L tanks with circulating filtered seawater taken from the nearby ocean at 13°C. Oxygen was bubbled into each tank via an air hose. Animals were fed weekly on a diet of fresh *Eklonia maxima* harvested from the beach opposite the aquarium. Prior to each heat shock experiment the abalone were divided into separate baskets and acclimatised overnight in water maintained at 13°C (control temperature). The experimental tank was heated to 23 °C using a standard submersible fish tank heater. Once the water reached the correct temperature, the baskets containing the experimental animals were placed in the heated tank. Abalone were initially observed for known behavioural responses to heat shock, in addition to quantifying the effect of the heat shock by haemocyte counts.

Haemolymph was removed via the pedal sinus of the abalone with a sterile 22 G needle and 2 ml syringe at each sampling time point during timed heat shock treatments. As the animals were too small to individually provide sufficient haemolymph for applications requiring large amounts of protein, 500 µl was sampled from each of 12 individual animals to generate a pooled biological replicate consisting of a total of 6 ml. The pooled haemolymph was then divided into 1 ml aliquots in 1.5 ml eppendorf tubes to facilitate further processing of the samples. The aliquots were immediately centrifuged at 10000 rpm for 5 minutes in order to pellet the haemocytes. The haemolymph was removed and the pellets were immediately frozen in liquid nitrogen. The pellets were transported back to the University of Cape Town in liquid nitrogen, and stored at -80 °C.

2.2.2. Haemocyte count

In order to quantify the physiological effects of heat shock at 23 °C on *H. midae*, 100 µl of haemolymph sampled from heat shocked animals was transferred to an eppendorf containing 200µl of anti-coagulation buffer (Appendix A) on ice. Since the duration of the heat shock response in *H. midae* was initially unknown, haemolymph samples were obtained from the abalone at 0, 15, 30, 45, 60 and 120 minutes. The samples were mixed on a vortex for 15 seconds before the count was performed to ensure homogeneity of the haemocytes in solution. The count was performed using a haemocytometer (Neubauer) on a light microscope at a magnification of 400 X. A 10 µl aliquot of the haemolymph was loaded onto the haemocytometer, and covered with a glass coverslip. Haemocytes on four 0.04 mm^2 squares (0.16 mm^2 total area) were counted in triplicate for each time point, and the experiment was repeated independently four times (n=4). The number of haemocytes per millilitre was calculated using the following equation:

$$Y = [(A/0.1)/X] * 1000 * Z \quad (1)$$

Where:

Y = the number of haemocytes/millilitre

A = average number of haemocytes (technical replicates)

X = the area of the square counted (0.16 mm^2)

Z = the dilution factor

The average number of haemocytes/ml haemolymph sampled at each time point across four biological replicates was calculated and plotted on a bar graph with the standard error of the mean (SEM) indicated. The data were analysed by one way ANOVA using SigmaStat software (Version 3.10 © 2004 Systat software, Inc) with $p < 0.05$ and a Tukey post hoc test.

2.2.3. Heat shock experiments

Two sets of experiments were performed in order to assess the effect of heat stress on the proteome of *H. midae* haemocytes. A preliminary experiment was carried out to determine the duration of the heat shock response, with two biological replicates analysed per time point. In order to determine whether the effects of heat stress persist after cessation of the stress, a recovery sampling point was included for 24 hours post-stress. Based on this information, an appropriate time course could be determined for further characterisation of the abalone molecular heat shock response. A second round of experiments was then

carried out to compare the protein expression profiles of haemocytes sampled from stressed and unstressed *H. midae* after short (30 minutes) and long (180 minutes) exposure to heat shock. This second experiment was carried out in triplicate, with three technical replicates averaged for each biological replicate.

2.2.4. Preparation of total soluble protein

The haemocyte pellets were removed from -80 °C and 200 µl of protein lysis buffer (Appendix A), containing protease inhibitors, was added to each pellet. The haemocytes were thawed in the lysis buffer, and then the pellets were disrupted and homogenised by vortexing for up to 30 minutes at room temperature. Once the pellets were completely disrupted in the lysis buffer the cellular debris was pelleted by centrifugation at 6000 rpm for 30 minutes at 4 °C. The supernatant, containing the cell lysate, was transferred to a fresh eppendorf tube.

The supernatant was diluted up to 500 µl with sterile distilled water immediately before the addition of an equal volume of phenol pH 8.0. The cell lysate and phenol phases were mixed by inverting the tubes and the phases were then separated by centrifugation for 10 minutes at 14000 rpm at 4 °C. The liquid supernatant layer was removed, leaving the interface and phenol layers intact. This was repeated twice; with 500 µl of lysis buffer added to the phenol layer the second time. The protein was precipitated out of the phenol layer by incubating at -20 °C overnight in 0.1 M ammonium acetate in methanol. The precipitated protein was pelleted by centrifugation and washed in 80% acetone. The pellets were then air dried and resuspended in urea lysis buffer (Appendix A), which is a suitable buffer for downstream 2D PAGE analysis. Protein concentration was then determined using Bradford's reagent (Bio-Rad) and a bovine serum albumen standard curve. Protein was stored in 150 µg aliquots (three per time point) at -80 °C.

2.2.5. 2-Dimensional gel electrophoresis

The frozen aliquots of protein were removed from -80 °C and thawed at room temperature. Each heat shock experiment comprised of three biological replicates at each time point, with three technical replicates of each biological replicate. The separation of the protein sample in the first dimension was carried out on an immobilized pH gradient (IPG) according to the manufacturer's instructions (Bio-Rad). Briefly, rehydration buffer including 150 µg of protein sample (Appendix A) was mixed, and 0.01 g of bromophenyl

blue was added. This solution was then carefully pipetted into a clean rehydration tray, and any bubbles were removed. Bio-Rad ReadyStrip™ IPG strips (7 cm, pH 4-7) were placed gel side down into the rehydration solution and allowed to rehydrate for 12-16 hours. The focussing tray for the Bio-Rad Protean® Isoelectric focussing (IEF) system for 7 cm strips was prepared by covering the electrode filaments with Whatmanns filter paper wicks, and the rehydrated strips containing the protein sample were placed into the focussing tray with anode/cathode aligned correctly onto the electrodes. The strips were covered with mineral oil, and the focussing tray was inserted into the isoelectric focussing system. The protein was focussed in the first dimension under the following conditions: 250 V for 20 minutes, 4000 V for 2 hours, 4000 V for 20000 Volt-hours, 500 V for 24 hours (holding step). Once the IEF was complete strips could be removed and stored at -80 °C overnight.

Electrophoretic separation of the protein sample in the second dimension was performed using the Bio-Rad Protean Plus Dodeca cell system, which allows for the simultaneous electrophoresis of up to 12 polyacrylamide gels, thereby minimising any gel to gel variation. The Dodeca cell casting tray was used to cast 12 X 10% polyacrylamide gels (Appendix A) simultaneously. Once cast and polymerised, the gels were removed from the casting tray and excess acrylamide was cut off the edges of the gel assembly using a scalpel blade. The gels were rinsed in distilled H_2O immediately prior to electrophoresis, and blotted dry with whatmanns paper. The strips were removed from -80 °C and incubated in equilibration buffer one (6 M urea, 2% SDS, 0.375 M Tris-HCl pH 8.8, 20% glycerol containing 1% w/v DTT) and then equilibration buffer two (6 M urea, 2% SDS, 0.375 M Tris-HCl pH 8.8, 20% glycerol containing 4.8% w/v iodoacetamide) for 10 minutes each. After equilibration the strips were removed from the equilibration buffer, dried by briefly applying tissue paper, and applied onto 10% acrylamide gels. The strips were held in place by the addition of 1 ml of melted 0.5% agarose in 1x SDS PAGE running buffer. A molecular weight marker was included for estimation of the size of the spots (Fermentas PAGE-ruler). The protein was separated by electrophoresis at 100 V for 2-3 hours at 4 °C. The gels were then stained in Coomassie blue R250, incubated in destaining solution overnight and stored in 7% v/v acetic acid. The gels were scanned at 300 dpi using a Canon CanoScan LiDE110 colour image scanner.

2.2.5.1. Spot analysis

The scanned images of the 2D PAGE gels were analysed using Melanie 7 (Swiss institute of Bioinformatics). Images from each replicate were uploaded into Melanie, and processed concurrently. The molecular weight marker was cropped out of the images so that the software would not detect the ruler bands. A hierarchical match structure was designed for each of the two experimental sets so that the correct structure of biological and technical replicates was contained within each match group. A diagram showing the experimental design and match hierarchy for the preliminary experiment to determine the duration of the haemocyte heat stress response is shown in Figure 2-1-A. A similar diagram showing the experimental design used to investigate the proteomic profiling of an early and late response to heat stress is shown in Figure 2-1-B.

Spots were detected automatically, and noise in the form of dust contamination or scratches on the gel was reduced using the 'smooth', 'min area' and 'saliency' parameters. The smooth function was set to 2, as at this value most of the overlapping spots were properly resolved. Min area and saliency were altered on an image specific basis so that only 'real' spots were detected. Reference images (the image with the greatest number of detected spots) were selected for each match set, and 2 landmark spots were chosen to facilitate proper matching. The gels were then matched, which allowed for spots to be compared between images, grouping each individual spot across all biological/technical replicates. Normalised spot intensity values were calculated by Melanie using the equation:

$$\frac{\text{Spot value} - \text{Central tendency}}{\text{Dispersion}},$$

for each individual gel, which is sensitive to both low and high spot intensities and simplifies the comparison between matches. Spot matches were manually inspected to ensure that the software had aligned the gels correctly. Statistical analysis for each matched set of spots was carried out by one way ANOVA using Melanie to identify spots that were differentially expressed in the heat shocked samples versus the control.

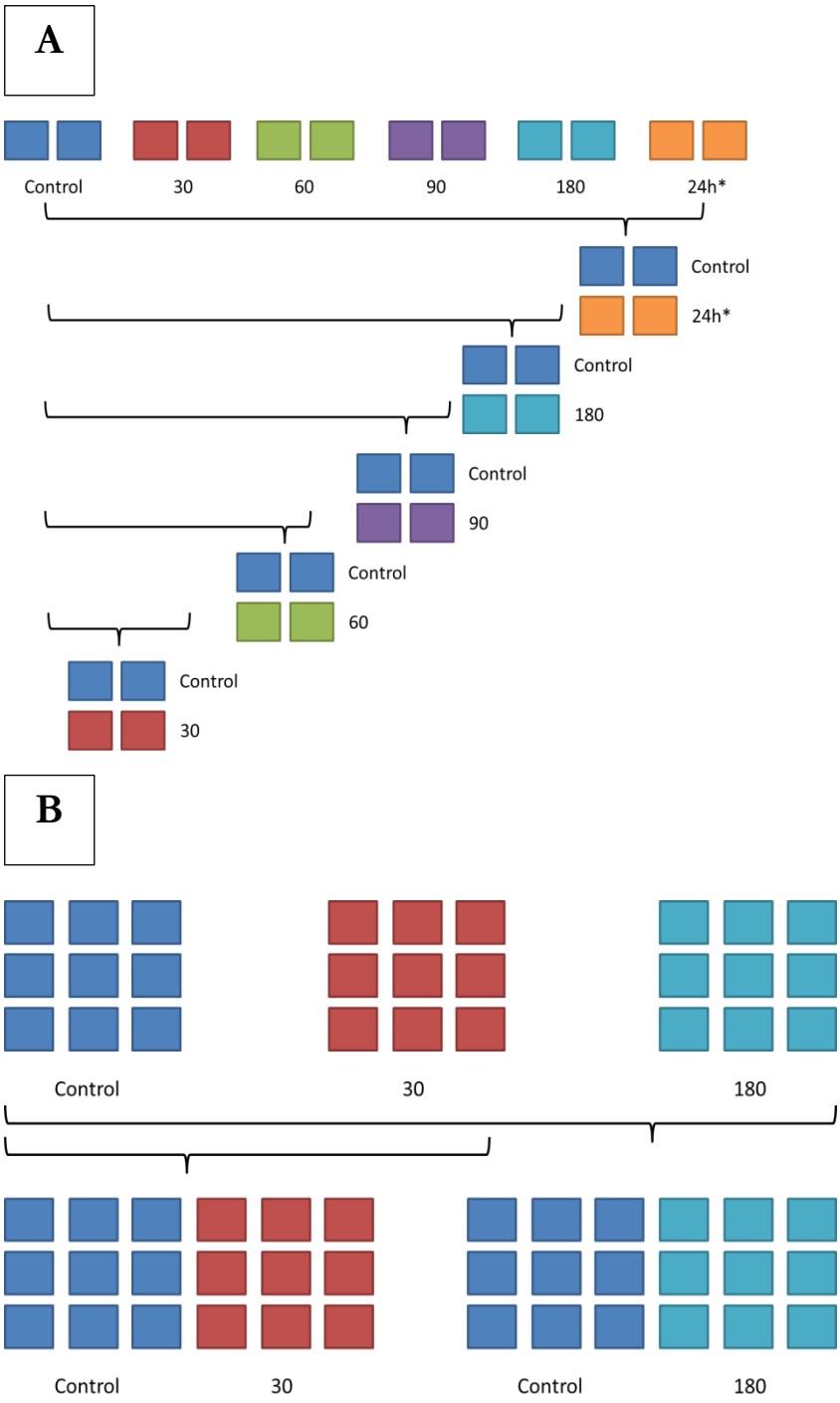


Figure 2-1: Diagram showing the experimental design and match hierarchy used to determine the duration of heat stress in *H. midae* haemocytes, and identify differentially expressed proteins in response to heat stress. Time points are shown in minutes, with the exception of the sample marked by an asterisk (*), which represents a sample obtained from animals post-stress that had recovered in ambient temperature water for 24 hours. Replicates (representing individual 2D PAGE gels) are indicated by squares. Brackets indicate groups that were compared in the analysis. (A) Experimental design of a preliminary investigation to determine the duration of the proteomic heat stress response in the abalone *H. midae*. (B) Experimental design used to compare control (T0) 2D PAGE gel images to experimental 2D PAGE gel images in order to locate differentially expressed proteins in haemocytes from *H. midae* exposed to acute heat shock.

2.2.5.2. Protein excision and analysis by mass spectrometry

In order to identify the protein in the selected statistically significant spots by MS analysis, a new set of 2D PAGE gels were prepared as described in 2.2.5. However, due to the sensitivity of MS analysis it was important to minimise contaminants such as human keratin in all the reagents and solutions used by filter sterilising. The gels were prepared, and then soaked in sterile milli-Q water for 2 hours prior to excising the spots in order to dilute any interfering substances that may persist in the gel. In order to obtain the maximum amount of protein from each spot, three pieces of gel were excised from each of the selected significant spots using the Bio-Rad ExQuest™ spot cutter. The gel pieces were placed into sterile 1.5 ml eppendorf tubes and dehydrated in a vacuum centrifuge until no liquid remained. The pieces were then stored at -80 °C, and sent to the Keck Institute at Yale University for analysis by mass spectrometry on an OT-Orbitrap mass spectrometer. Since the full genomic sequence for abalone was unavailable, identification of differentially expressed proteins necessitated the use of homologous database searching. The Keck Institute recommended the use of MS-BLAST, which matches spectra to peptides generated from the non-redundant BLAST database in order to determine the best peptide spectrum match. The proprietary proteomics software package Mascot (Matrix science) was used by the Keck Institute for peptide spectrum matching and protein identification of the selected spots.

Mascot searching was performed by the Keck Institute with the variable modifications oxidation (M) and propionamide (C) selected, and with a peptide monoisotopic mass tolerance of 20 ppm and fragment mass tolerance of 0.6 Da. The spots were identified based on the proteins with the highest Mascot scores, which were determined by Mascot software. In each case the predicted molecular weight and pI of the protein in the database with the best score was compared to the actual molecular weight and pI of the protein on the 2D gel for verification of the protein identity.

2.3. Results

2.3.1. Haemocyte count and behavioural response to heat stress.

Initially, the suitability of 23 °C as the routine heat shock temperature was established. Abalone that had been exposed to 23 °C water exhibited known behavioural responses to heat stress, including mantle gaping, epipodial lifting and a rapid rotation of their shells through 180°. After two hours of heat stress, the animals were no longer able to maintain attachment to the basket surface, and were producing a large amount of mucus. Haemocyte counts were used as a physiological parameter to assess molluscan immune capacity. The number of haemocytes in haemolymph obtained from *H. midae* via the pedal sinus during a heat shock trial was determined. There was an initial haemocyte count of approximately $9.39x10^6$ cells/ml at time zero, followed by a statistically significant ($p<0.05$) increase in the number of circulating haemocytes after fifteen minutes of heat shock (approximately $1.75x10^7$ cells/ml) (Figure 2-2). Haemocyte numbers remained elevated for the duration of the heat shock trial without any further significant changes.

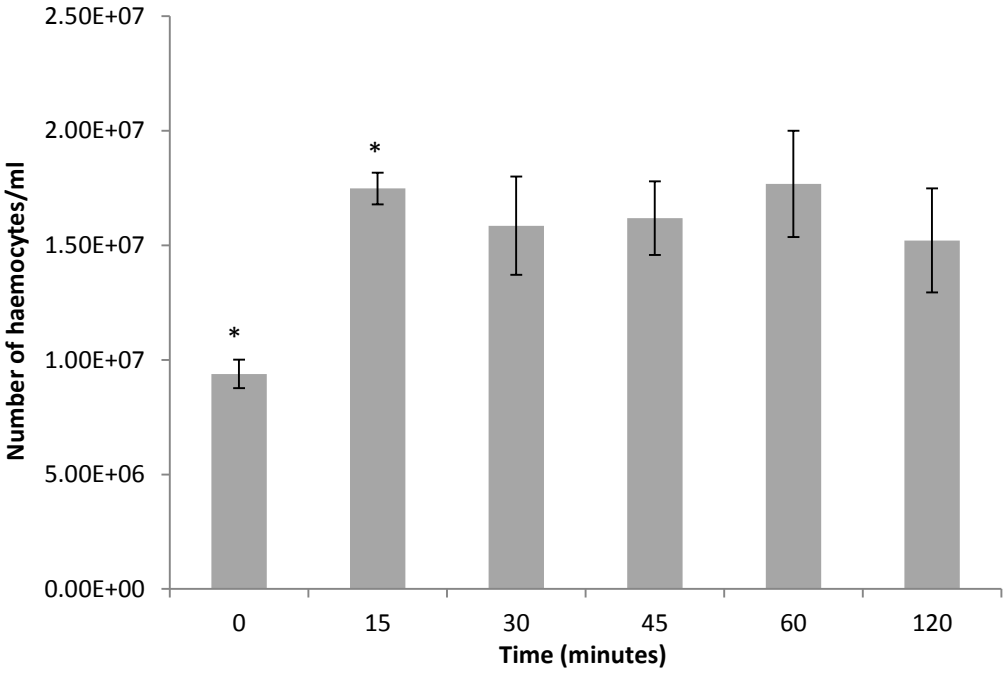


Figure 2-2: The number of circulating haemocytes counted for the duration of a timed heat shock experiment. Data are represented as mean number of haemocytes/ml haemolymph ± standard error for pooled haemocytes sampled from five animals at each time point and across four biological replicates. Statistical analysis was performed using a one way analysis of variance (ANOVA) with a Tukey post-hoc test using SigmaStat 3.10.0. Asterisks (*) indicate time points that are significantly different, ($p<0.05$).

2.3.2. Determination of the duration of the abalone heat shock response

A preliminary investigation was carried out to determine the duration of the abalone heat shock response. The proteomic response of the haemocytes at time intervals during a heat shock experiment was investigated using 2D PAGE and analysed using Melanie software. To assess proteomic changes in the haemocytes at each time interval, the spot patterns were matched between control and experimental samples. Each spot detected by Melanie in the control (T0) sample was matched to its counterpart in the experimental samples. Melanie then calculated intensity values for each spot, which indicates the amount of that protein present in the sample. Spots in these matched sets were then sorted according to the stipulated criteria that the experimental spots should show more than a 2 fold change in intensity compared to the control, with the statistical significance determined by one way ANOVA ($p < 0.01$). The number of spots that satisfied this criterion for each of the time intervals sampled, referred to as verified spots, are presented in Table 2-1.

The prevalence of verified spots was used to signify the duration of the abalone proteomic response to acute heat stress. The sample obtained from abalone that had recovered from a 3 hour heat shock in ambient temperature water for 24 hours had 142 verified spots, indicating the persistence of a physiological response to the stress even after it was removed.

Table 2-1: The duration of the molecular heat stress response in the haemocytes of *H. midae* was determined using 2D PAGE analysis to identify differentially regulated proteins in control vs. heat stressed haemocyte total protein samples. The number of verified spots for each time interval compared to the control (T0) is shown. Spots were assigned as verified if the fold difference of their density compared to the control was more than 2 fold, and statistically significant by one way ANOVA ($p < 0.01$) All spot density determination, fold change and one way ANOVA calculations were performed using Melanie software. The asterisk (*) indicates data obtained from abalone 24 hours post-heat shock (recovery).

Comparison	# of verified spots
0vs30	130
0vs60	105
0vs120	155
0vs180	157
0vs24hours*	142

2.3.3. 2D PAGE analysis of an early (30 minute) and late (180 minute) *H. midae* haemocyte response to acute heat shock

In order to gain further understanding of the molecular mechanisms of *H. midae's* molecular heat stress response, the effect of acute heat stress on the haemocyte proteome was investigated. 2D PAGE was employed to compare the total haemocyte protein expression profiles from untreated (T0) abalone, and abalone which had been exposed to 30 or 180 minutes of heat shock. The 2D PAGE gel images were processed using Melanie software. One-way ANOVA was used to test for differences between spots in control and experimental samples, with significance established if $p < 0.05$.

A total of 900 spot matches were made when comparing the control sample with the sample obtained after 30 minutes of heat stress, while comparison between the control sample and the sample obtained at 180 minutes of heat stress detected 772 spot matches. Of these matches, spots which differed between control and experiment significantly by one way ANOVA ($p < 0.05$) were identified using Melanie. A total of 13 spots were selected that reproducibly showed the most statistically significant changes in intensity between the experimental sample and the control – five spots at 30 minutes (Figure 2-3) and 8 spots at 180 minutes (Figure 2-4). After 30 minutes of heat stress, four of the five chosen significant spots had increased compared to T0 (Spot ID numbers 188, 99, 226 and 19), and after 180 minutes 7 out of the 8 chosen significant spots had significant decreases in expression compared to T0 (Spot ID numbers 87, 168, 103, 76, 31, 38 and 17).

Due to financial constraints, only a small number of differentially expressed spots were chosen for identification by MS analysis. Only 7 of the chosen 13 spots could be reliably located and excised from a Coomassie blue stained gel due to the extremely small size of some of the protein spots. The identities of the excised spots that were determined following MS analysis at the Keck Institute at Yale University are presented in Table 2-2. In order to verify the putative protein identities, the pI and MW from the NCBI database were compared to the pI and MW measured for each spot on the 2D polyacrylamide gel.

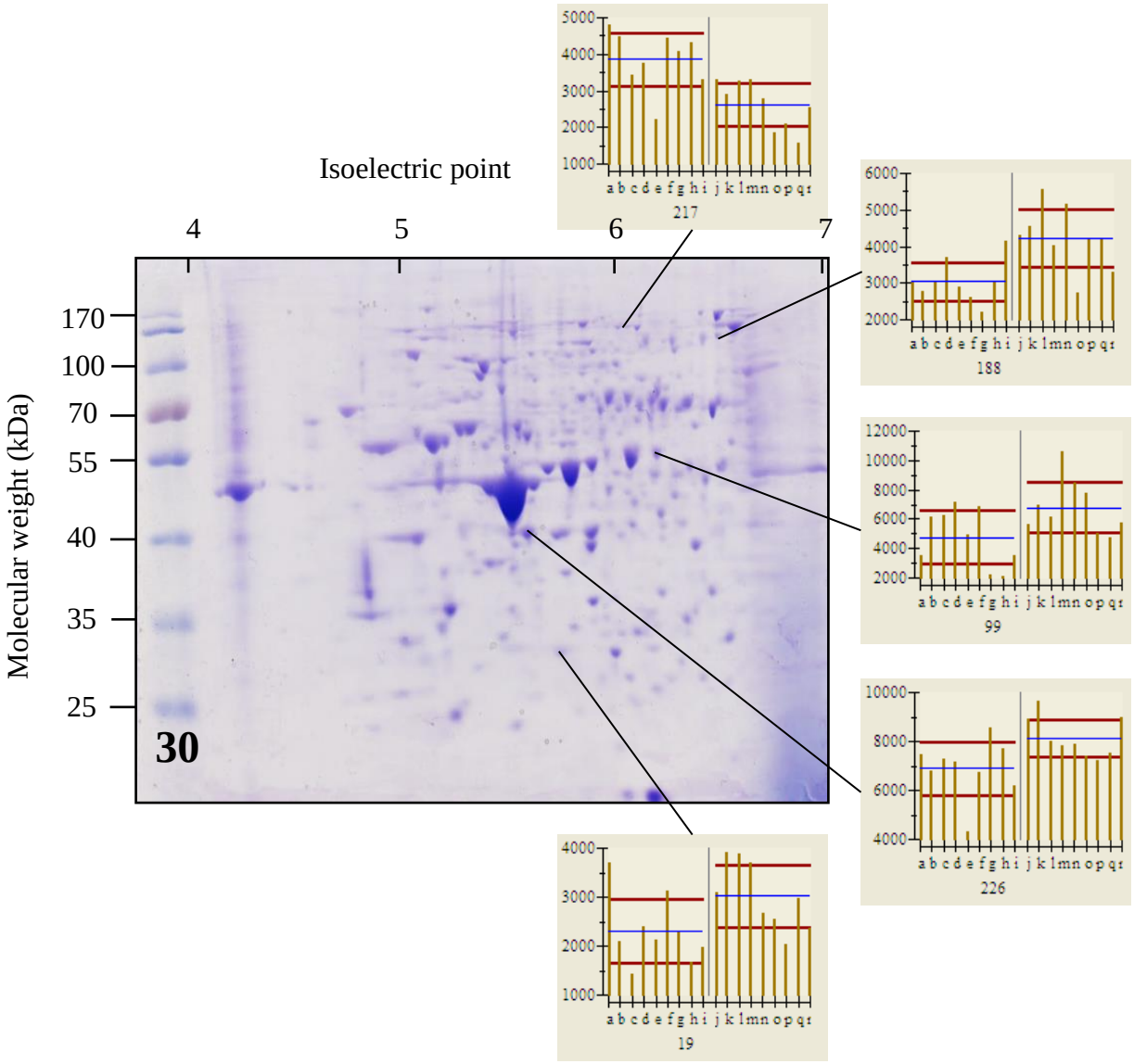


Figure 2-3: Representative 2D PAGE gel showing the proteome of haemocytes from *H. midae* after exposure to 30 minutes of heat shock at 23°C. The molecular weight marker included is the Fermentas PAGE ruler pre-stained marker. Bar graphs were generated by Melanie software and indicate intensity values for each spot across all replicate gels. The bar graphs show the intensity values for that spot for each of the 3 technical repeats performed in the control (T0) biological replicate 1 (a-c), 2 (d-f) and 3 (g-i), followed by T30 biological replicate 1 (j-l), 2 (m-o) and 3 (p-r). Spot identity numbers (below the bar graphs) were assigned by Melanie software. The y axis represents intensity in arbitrary units.

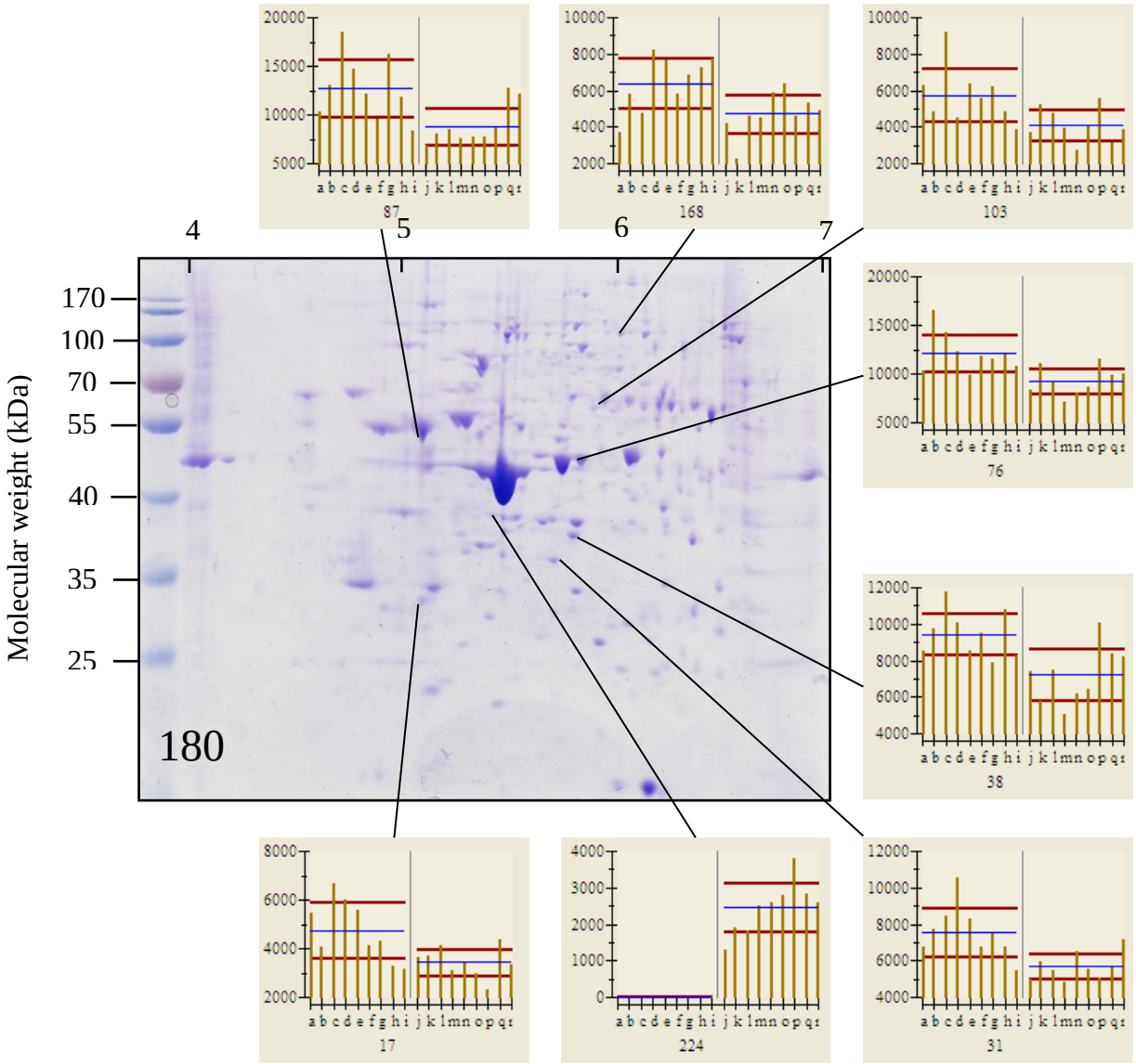


Figure 2-4: Representative 2D PAGE gel showing the proteome of haemocytes from *H. midae* after exposure to 180 minutes of heat shock at 23°C. The molecular weight marker included is the Fermentas PAGE ruler pre-stained marker. Bar graphs were generated by Melanie software and indicate intensity values for each spot across all replicate gels. The bar graphs show the intensity values for that spot for each of the 3 technical replicates performed in the control (T0) biological replicate 1 (a-c), 2 (d-f) and 3 (g-i), followed by T180 biological replicate 1 (j-l), 2 (m-o) and 3 (p-r). Spot identity numbers (below the bar graphs) were assigned by Melanie software. The y axis represents intensity in arbitrary units.

Table 2-2: Protein identities determined by MS analysis of spots excised from 2D PAGE gels of *H. midae* haemocyte total protein. Spots were selected for identification based on statistical significance as determined by analysing three biological replicates of 2D PAGE gel images comprising three technical replicates each for the control (T0) vs. samples obtained after 30 or 180 minutes exposure to acute heat stress. The excised spots were sent to the Keck Institute at Yale University for analysis and identification by MS. The Keck Institute used the Mascot search engine to assign peptide spectrum matches (PSMs) for each of the proteins, and thereby assign a protein identity for each spot. The confidence of the identity is supported by the 'score' value, which is assigned by Mascot and indicates the confidence of the PSM. The % coverage indicates the total coverage of peptides determined by MS analysis of the full length protein sequence in the database. Furthermore, the pI and MW of the protein identified from the NCBI database is presented alongside the pI and MW measured on the 2D PAGE gel.

Match ID	GenBank Accession	Description	Species	Score	Expectation	% coverage	Predicted pI/MW (kDa)	Gel pI/MW (kDa)
17	P45885	Actin-2, muscle specific	*Bactrocera dorsalis*	548	1.70E-48	25.3	5.23/41.8	5.5/32
31	ACB05681	G Protein beta subunit	*Euprymna scolopes*	397	2.60E-33	18.5	5.8/37	6.3/37
38	No identity	-	-	-	-	-	-	-
76	NP_989638	Actin related protein 3	*Gallus gallus*	433	6.10E-37	14.4	5.61/47	6.5/50
87	P02554	B tubulin	*Sus scrofa*	487	2.40E-42	18.9	4.78/50	5.5/50
217	AAH71378	Aars protein, partial	*Danio rerio*	158	2.00E-09	4.5	5.28/107	6.5/120
224	CAA34718	Actin	*Caenorhabditis elegans*	434	5.20E-37	21.5	5.3/42	5.9/40

2.4. Discussion

Aquacultured abalone are produced worldwide and exported for high prices. Under aquaculture conditions, bacterial or viral pathogens of abalone often cause mass mortality (Hooper et al., 2007b). A link between abiotic stresses and increased susceptibility to disease has been reported in abalone, although the mechanisms underpinning this connection remain unclear (Hooper et al., 2007a). Malham et al. (2003) hypothesise that the re-appropriation of energy resources in stressed abalone by metabolic systems necessary to mount a stress response reduces the amount of energy available for immune function and thus leads to immunodepression. There is a need to investigate how abalone respond to common environmental stresses on the farm in order to elucidate the link between stress and the immune system in this economically valuable animal.

The interactions between abalone, their pathogens and the environment leading to disease outbreaks are complex and poorly understood, although increased temperature has been linked to higher incidences of disease and mortality in abalone (Dang et al., 2012). Tank water temperature on abalone farms fluctuates regularly, exposing the abalone to a wide range of temperatures over a short period of time. With current proteomics technology it is possible to investigate the cell-wide mechanisms that are involved in the molecular response to stimulus in marine molluscs (López et al., 2002). 2D PAGE is a technique that has been used to detect proteomic changes in marine molluscs exposed to toxic compounds (Dowling et al., 2006), geographical variation (López et al., 2002) and disease (Cao et al., 2009). However, few studies exist that report the use of proteomic techniques to investigate the cellular response to thermal stress (Serafini et al., 2011). In this study, the effect of acute heat stress on the South African abalone *H. midae* was examined by analysing the proteomic expression profile of abalone haemocytes.

The first objective of this study was to determine an appropriate temperature that would elicit a physiological response in *H. midae* that had been acclimatised to water at 13 °C, without causing unnecessary mortality in the exposed animals. There is a lack of published data reporting temperature fluctuations in abalone farm tanks; however Vosloo and Vosloo (2010) report daily water temperatures that range up to 10 °C. For this study, a temperature of 10 °C higher than ambient was selected for investigation. Although several commonly used physiological assays have been used to determine immune modulation in

abalone, changes in the number of circulating haemocytes is considered to be an indicator of abalone immune capacity (Hooper et al., 2007a). Haemocyte counts and behavioural observation were therefore conducted in order to determine if exposure to water at 23 °C was sufficient to illicit a physiological response in *H. midae* that had been acclimatised at 13 °C.

In the current study, there was a significant increase in the number of haemocytes counted in the haemolymph collected from the pedal sinus of *H. midae* after 15 minutes of exposure to 23 °C water. The number of haemocytes had risen to almost double the initial count, and this elevated number persisted for the duration of the 120 minute heat shock. It has been suggested that stress may affect the association between the endothelium and the haemocytes, causing the haemocytes to leave the connective tissue and enter the haemolymph where they would presumably enter the pedal sinus (Hooper *et al.,* 2007). The research carried out by Malham *et al.,* (2003) reported a similar phenomenon in the abalone *H. tuberculata* during a shaking stress trial, confirming that a change in the number of circulating haemocytes is related to the stress response in abalone. An increase in the number of total circulating haemocytes in abalone exposed to elevated water temperature has been reported previously for *H. rubra* (Dang et al., 2012), *H. diversicolor supertexta* (Cheng et al., 2004a) and a hybrid Australian abalone (Hooper et al., 2014). In addition to elevated haemocyte counts, heat shocked *H. midae* displayed documented abalone heat stress behaviours including uplifted shell gaping and epipodial spreading (Hecht, 1994). These responses are thought to improve evaporative cooling and have been reported in a number of abalone species (Diaz et al., 2006). No mortality was observed in treated abalone in the days following exposure to 23 °C water for 120 minutes. These observations, as well as elevated haemocyte counts for the duration of the exposure to 23 °C water, indicate that the applied stress is sufficient to illicit a physiological response in *H. midae* without causing unnecessary mortality. Therefore, 23 °C was selected as an appropriate temperature for further experimentation in order to characterise the heat stress response in *H. midae*.

Although some investigations report the transcriptional regulation of heat shock proteins in abalone responding to chronic heat stress (De Zoysa et al., 2009) or recovery following acute heat stress for 1 hour (Cheng et al., 2007), there is currently a lack of information regarding the duration of the abalone molecular response while exposed to acute heat stress. Therefore, the next objective of this study was to use 2D PAGE in a preliminary experiment to determine the duration of the proteomic response of *H. midae* while exposed

to water at 23 °C. Abalone were exposed to heated water for 180 minutes, and haemolymph was sampled at intervals for the duration of the trial. In addition, abalone that had been exposed to 23 °C water for 3 hours and then returned to water at 13 °C to recover were sampled 24 hours post-stress. Analysis of the resultant 2D PAGE gels revealed that the proteomic response to heat shock was sustained for the duration of the exposure to the stress, and persisted in the animals that had been allowed to recover for 24 hours. Since the duration of exposure to heated water was sufficient to cause a response, early (30 minutes) and late (180 minutes) sampling points were chosen for the succeeding experiments.

A second 2D PAGE experiment was carried out in order to identify proteins that respond differentially to acute heat stress, and thereby illuminate the molecular stress response mechanisms contributing to the heat stress response in *H. midae*. In this study, the proteomic response of *H. midae* to exposure to water at 23 °C for 30 or 180 minutes was investigated. Although 2D PAGE technology coupled with MS analysis is capable of identifying hundreds of proteins, financial constraints necessitated the selection of only a few spots for identification by MS analysis in this study. Therefore, based on the 2D PAGE analysis, the proteins with the most reproducible expression and statistically significant differences between control and experimental groups were chosen. In order to select these, the differentially expressed spots reported by Melanie were manually examined to eliminate those spots that were overlapping or indistinct, or had artefacts such as bubbles which altered the quantitative data. Thereafter, a total of 13 spots (5 from the 30 minute sample, 8 from the 180 minute sample) were chosen as candidates for MS analysis. However, due to the small size of some of these chosen spots, only 7 could be reliably located and excised from the Coomassie gel. Only spot 217 was successfully excised from the sample taken after 30 minutes of heat shock, and the remaining six spots were excised from the sample taken after 180 minutes. These 7 spots were analysed using LC MS/MS at the Keck Institute at Yale University School of Medicine in order to identify the proteins that were responding reproducibly and significantly to thermal stress in *H. midae*. The putative roles of these proteins in the heat shock response of *H. midae* are discussed below.

One of the chosen spots (spot 38) could not be identified following MS analysis, possibly due to a large amount of keratin contamination in the sample. Out of the remaining 6 proteins that were putatively identified, five identities were related to the cytoskeleton, including isoforms of actin (spot 17, 76 and 224). The putative identity of spot 31 is a G

protein beta subunit which is a known actin binding protein, which performs regulatory functions within the cytoskeleton (Peracino et al., 1998). The cell cytoskeleton is a supportive structure of fibres that gives a cell its shape and rigidity, but is also involved in locomotion and dynamic response to external stimuli (Small et al., 1999). It is comprised of many different isoforms of the proteins actin and tropomyosin (Gunning et al., 2005), which are regulated regionally in the cell by a diverse array of actin binding proteins (May and Machesky, 2001). Actin is common and abundant in all metazoan cells (Small et al., 1999), and in non-muscle cells comprises up to 5% of the total cellular protein by weight (Lodish and Zipursky, 2001). Differential expression of actin has been reported in a number of studies describing the molecular effects of stress stimuli in molluscs, including the response to pollutants (Gómez-Mendikute et al., 2002; Rodriguez-Ortega et al., 2003), heavy metals (Muralidharan et al., 2012) and oxidative stress (McDonagh and Sheehan, 2007). As the cellular basis for the abalone haemocyte immune response is phagocytosis (Travers et al., 2008), which is an actin dependent process, disruption of the actin cytoskeleton may affect the ability of the haemocytes to respond to an opportunistic pathogen infection (Thomson *et al.*, 2012).

In the present study, after 180 minutes of exposure to heat stress, the expression of five cytoskeletal proteins decreased compared to the control. The exception was spot 224 (Actin), which was present in the heat shocked haemocytes but not in the control sample. As these were not differentially regulated in the sample taken after 30 minutes of heat stress, it would seem that disruption of the cytoskeleton may be a late occurrence in response to acute heat stress. The apparent decrease in the expression of cytoskeletal proteins may be attributed to the proteolytic breakdown of cytoskeletal elements during exposure to high temperature. Serafini et al. (2011) determined that the sea squirt (genus *Ciona*) response to high heat stress conditions included significant proteolytic breakdown of actin monomers, and conclude that cytoskeletal elements may be thermally unstable (Tomanek and Zuzow, 2010). The precise reason for this instability is unspecified in the literature, however Rodriguez-Ortega et al. (2003) suggest that stresses causing increased production of reactive oxygen species will alter Ca^{2+} homeostasis of the cell, thereby depolymerising filamentous actin. It is also possible that increased production of reactive oxygen species (ROS) has a direct regulatory effect on the expression of actin genes (Gómez-Mendikute et al., 2002). Thermal stress causes the uncoupling of electrons from metabolic pathways, increasing the amount of ROS in the cell (Huot et al., 1997; Dalle-Donne et al., 2001; Tomanek, 2011).

Advances in proteomics technology have allowed for the identification and quantitative analysis of a large number of proteins. 2D PAGE is a reliable and reasonably high-throughput technique that is suitable for the characterisation of the proteomic response to environmental stimuli, such as heat shock. However, some of the weaknesses of 2D PAGE include the loss of resolution of proteins that are present in very small amounts, highly basic or acidic, or small in size. This study revealed a number of proteins that were differentially expressed in *H. midae* haemocytes after exposure to acute heat stress for 30 or 180 minutes. Seven proteins were chosen for identification by MS analysis, and six of these proteins were assigned putative identities. As many of these proteins performed cytoskeletal roles in the cell, it was hypothesised that cytoskeletal disruption as a result of heat stress was a contributing factor to the immunodepression associated with increased temperature. Although this study has provided some insight into the effect that acute heat stress has on proteins expressed by *H. midae* haemocytes, further research is required to better understand the molecular mechanisms that cause this cytoskeletal perturbation.

Chapter 3

AN iTRAQ BASED PROTEOMIC PROFILING OF THE *H. MIDAE* HAEMOCYTE RESPONSE TO ACUTE THERMAL STRESS

3.1. Introduction

The coupling of 2D PAGE with LC MS/MS has enabled scientists to identify hundreds of proteins simultaneously. This has been beneficial for studying the effects of environmental toxicity on marine invertebrates (López et al., 2002; Silvestre et al., 2010; Sanchez et al., 2011; Thompson et al., 2011). One of the benefits of utilising high-throughput technology is that we are now able to study how groups of proteins respond together in order to generate a more holistic impression of the molecular stress response (Tomanek, 2011; Muralidharan et al., 2012). Despite successful application in the mollusca (McDonagh and Sheehan, 2007; Cao et al., 2009; Fields et al., 2012b), this technology remains underutilised in studies profiling the abalone stress response. Investigations focussing on abalone have previously measured the expression of a specific known stress response gene or protein under stress conditions (Cheng et al., 2007; Wan et al., 2012). In order to address the lack of information, 2D PAGE analysis was previously carried out to identify proteins in *H. midae* that are differentially expressed during exposure to acute thermal stress (Chapter 2).

Although many differentially expressed protein spots were detected using 2D PAGE, financial constraints prevented the identification of all of these proteins. In practice it is expensive and tedious to excise and analyse every spot individually, and therefore only seven spots could be excised and identified. 2D PAGE methodology is also limited in its scope to detect proteins outside the investigated pI range, or those that occur in low abundance. It is also difficult to resolve proteins with very high or low molecular weights on a polyacrylamide gel. In order to identify additional stress response proteins in *H. midae*, and thereby further our understanding of the abalone cellular stress response, a complementary proteomic technique was employed. Advances in MS-based high throughput quantitative proteomic technology have provided a tool that addresses the limitations of 2D PAGE by making use of peptide labelling to quantitatively describe the proteome.

LC MS has emerged as a powerful technology for the identification and quantitative analysis of proteins from complex mixtures. The first LC MS based gel-free isobaric labelling technique was known as ICAT (isotope coded affinity tag) (Linscheid et al., 2009). This and subsequent labelling technology makes use of chemical 'tags' (Zieske, 2006) to label peptides or proteins from different biological samples (Unwin et al., 2012) so that

their expression profiles can be determined following MS analysis (Eng et al., 2011). iTRAQ (isobaric tags for absolute and relative quantitation) is a newer technology that binds isobaric tags of a known m/z value to peptides (Unwin, 2010), and relies on MS/MS fragmentation of the peptides before they can be quantified (Linscheid et al., 2009). A generalised schematic of the iTRAQ workflow is represented in Figure 3-1. Protein samples are digested with trypsin, and then each sample is labelled with a different isobaric tag before being combined (Timms and Cutillas, 2010). The tag binds to the N-terminal amine groups, and/or the lysine side chains of the peptides via an N – hydroxysuccinimide moiety (Unwin et al., 2010). The isobaric tag comprises a balance and reporter group, with lighter balance groups on the heavier reporter groups so that during MS1 the same peptides from different samples will appear at the same mass to charge ratio on the mass spectrum (Linscheid et al., 2009; Unwin et al., 2010; Burkhart et al., 2011). Following fragmentation of the MS1 precursor, the reporter group is cleaved and released along with fragment ions which give a specific m/z signal for each label (Timms and Cutillas, 2010). The height of the m/z peaks for the reporter ions represent the quantity of that peptide in the sample, which can be used to compare protein expression in treated groups to a control sample (Unwin, 2010).

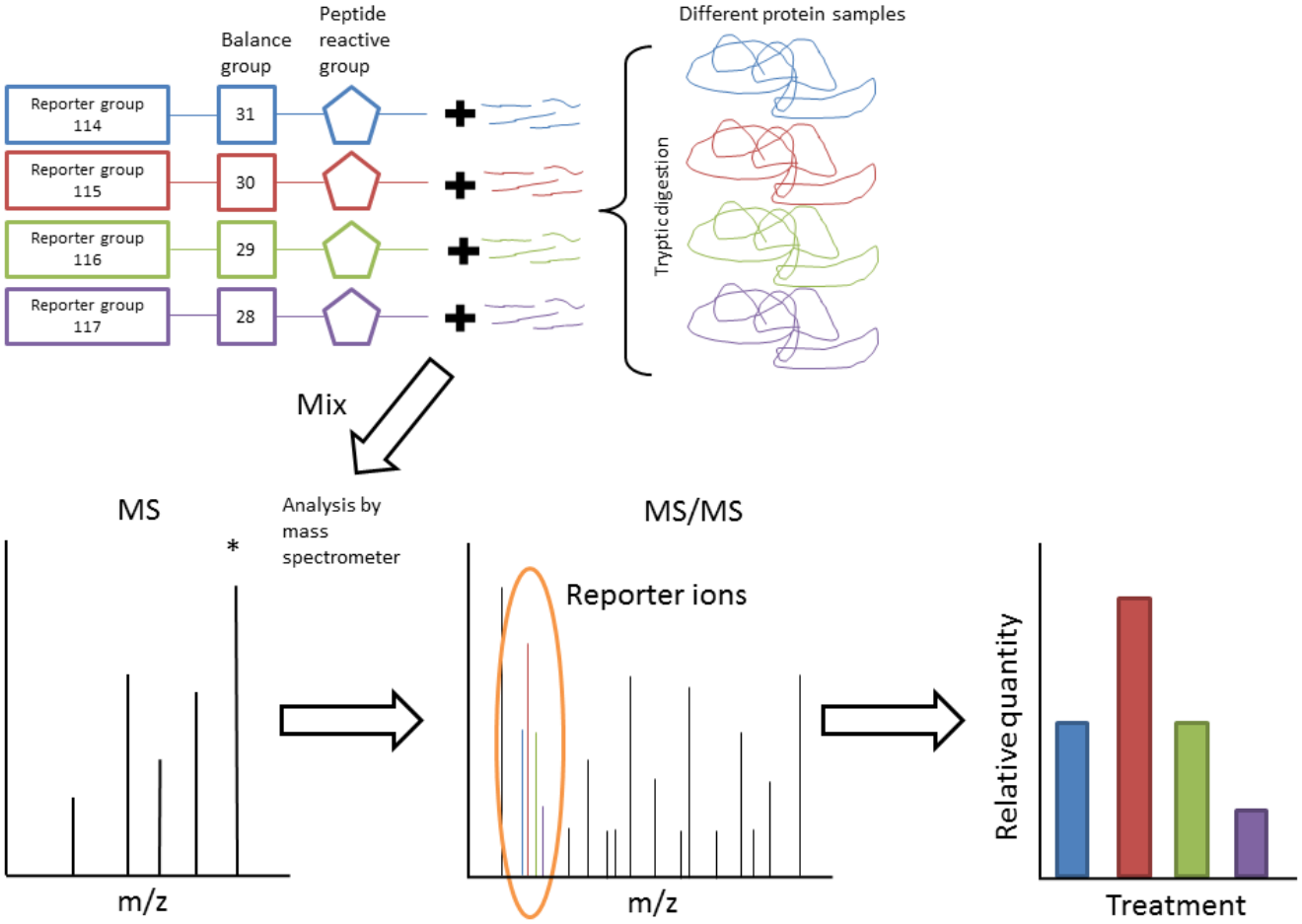


Figure 3-1: A schematic diagram of the iTRAQ workflow. For the sake of brevity 4-plex iTRAQ has been depicted, as 8-plex follows the same basic concept with an additional 4 reporter groups. Reporter groups with different isobaric masses are attached to balance groups in order to offset their mass differences so that the same peptide tagged with different labels will have the same mass. Different samples are labelled with different iTRAQ labels, and then combined. The sample is analysed by mass spectrometry. During MS1, since each peptide has the same mass regardless of sample they will have the same mass on the mass spectrum. A specific peptide is selected (indicated by an asterisk) and fragmented by CID for MS2 analysis, upon which the reporter group is cleaved from the peptide. Therefore the intensity of each isobaric tag in MS/MS indicates the relative abundance of that peptide in its respective sample. MS/MS spectra are then used to identify the peptide in question, either by database searching or by *de novo* sequence determination. Based on Unwin (2010).

iTRAQ tags are currently available through AB SCIEX™ and can be purchased in a 4-plex or 8-plex set. The 8-plex set has reporter ion masses at 113-119 and 121 m/z and allows for the comparison of up to 8 different conditions in a single experiment (Unwin et al., 2012). The ion mass 120 m/z is omitted because it corresponds to the mass of the phenylalanine immonium ion (Unwin, 2010). 8-plex iTRAQ is particularly useful for time-

course studies where several experimental samples are to be compared to the same control (Unwin et al., 2010) because there are 8 available channels. Furthermore, since total protein is first digested into peptides, the pI and molecular weight of the original protein has less of an effect on protein detection than is the case with 2D PAGE analysis. iTRAQ analyses of *Escherichia coli* confirmed that the percentage of identified proteins with high and low pIs and molecular weights was consistent with their abundance in the proteome (Aggarwal et al., 2006).

iTRAQ has been applied in studies with human cell lines, blood plasma and serum, and disease tissue (Song et al., 2008). There is, however, currently a lack of literature supporting high-throughput quantitative proteomic studies of aquacultured marine animals. A recent study described the use of iTRAQ to analyse proteomic changes in the barnacle *Balanus amphitrite* in response to an anti-fouling agent designed to prevent barnacle growth on submerged substrates (Han et al., 2013). This model crustacean has been utilised extensively to study the effects of anti-fouling agents. To our knowledge, this study is the only other one on marine invertebrates in which iTRAQ was used. This is partially due to the lack of adequate genomic databases for many aquacultured species. As with any mass spectrometry based analysis, the identification of a peptide is based on matching its m/z spectrum to that generated from a user-determined database of protein sequences. Although the genome of the oyster *C. gigas* is now fully sequenced and available in the UniProt database, protein sequences for most other marine species are not available, and thus, proteomic studies of non-model organisms must employ alternative strategies.

3.1.1. The implications of database choice in a proteomics investigation

The choice of database to be used in any 'bottom-up' proteomics based analysis is fundamental to the success of the study (Liebler, 2002). Although there are many model organisms with fully sequenced and annotated genomes, their numbers represent a small percentage of the total number of organisms which may be subjected to proteomic analysis. If a complete genomic database for the target species does not exist, it is possible to use sequence information from closely related organisms, which should have a high degree of genetic relatedness. However, genetic differences exist even between closely related organisms, and this may lead to reduced numbers of successfully identified proteins.

Until the advent of high-throughput next-generation sequencing technology the most commonly studied model invertebrates were *Drosophila melanogaster*, *Caenohybditis elegans* and *Aplysia californica*. The use of homologous database searching when the target species is non-model has been reported, and used to identify hundreds of proteins that would otherwise not be identifiable by MS analysis (Habermann et al., 2004; Liska et al., 2004). There is a concern that using a small database, or subset of a whole genome, may lead to the loss of information caused by a high number of false negatives due to missing entries, or false positives due to overestimation of the significance of a peptide match (Jones and Hubbard, 2010). The current scope of the literature lacks definitive guidelines when determining which database will best suit a proteomic study of a non-model organism. It has been proposed that a better strategy may be to use all known sequences for a phylogenetic subset that is closely related to the target (Forné et al., 2010), and software that is capable of concurrent database-based PSMs and *de novo* peptide sequencing (Shevchenko et al., 2009). Although computationally costly, the incorporation of *de novo* peptide sequencing is a powerful tool to corroborate the PSMs determined by database searching and to provide additional peptide sequence information where none exist in the database (Habermann et al., 2004; Zhang et al., 2012b).

3.1.2. Aims

There is an established link between environmental stress and the onset of disease in abalone (Malham et al., 2003; Malagoli et al., 2007; Adamo, 2012). Stresses encountered during exposure to aquaculture conditions could lead to immunosuppression and thus cause increased susceptibility to disease (Adamo, 2012). Previous work carried out on *H. midae* (Chapter 2) suggested that reactive oxygen species (ROS) production under heat shock conditions may lead to the disruption of the cytoskeleton of the haemocytes, which could impair the animal's ability to respond to bacterial infection. However, the detection capability of 2D PAGE technology is limited to a subset of the total proteins, and therefore an alternative molecular strategy was employed.

The aim of this study was to identify stress response proteins in the South African abalone *H. midae* by employing the high-throughput quantitative proteomic technique, iTRAQ. Although iTRAQ has been used in biomarker discovery in cancer and other human diseases, there is not yet any available literature describing the use of iTRAQ in environmental proteomics studies of molluscs. Therefore, available methodology was optimised for use with abalone haemocyte protein samples in order to obtain as many protein identities as possible without compromising accuracy. The objectives of this study included designing a workflow that satisfies the need for proper cell lysis and tryptic digest, as well as allowing efficient iTRAQ labelling followed by off-gel fractionation of the peptide sample. This workflow was then applied to haemocyte samples obtained from heat shocked abalone in order to identify proteins that are differentially regulated as a result of heat stress.

Since protein identification is based on homologous database searching, and many proteins have not been characterised in *Haliotis* it is often difficult to assign functional roles to the proteins. Gene ontology (GO) analysis allows the assignation of species neutral function-based keywords to protein identities. These keywords fall under the domains cellular component, molecular function or biological process. This enables some determination of function, at least superficially, for homology-based protein identities. GO analysis was therefore used to identify putative cellular processes occurring in *H. midae* haemocytes during exposure to thermal stress.

To assess the accuracy of the quantitation of individual proteins in the iTRAQ analysis, the expression profile of selected proteins was investigated using a complementary molecular technique. Western blot hybridisation was used to quantify the expression of selected proteins in *H. midae* haemocytes during exposure to the same experimental conditions as the animals sampled for iTRAQ analysis.

3.2. Materials and Methods

Solutions in this chapter are referenced in Appendix A.

For the purposes of this study, a biological replicate is defined as samples obtained from distinct groups of abalone under the same environmental conditions. A technical replicate is the replication of the same experimental technique on the same sample.

Although the word 'expression' implies directed control by cellular mechanisms, in order to simplify discussion it is used herewith to indicate the increased or decreased concentration of a particular protein in the cell as a result of translational regulation, targeted degradation, post-translational modification or any other cellular mechanism that can affect the quantity of a protein in a cell.

All reagents used were the highest quality and/or proteomics grade.

3.2.1. Animal maintenance

Cocktail sized abalone were maintained at the DAFF research aquarium in Cape Town, South Africa. Animals were housed in 100 l tanks, each containing 2 separate plastic mesh baskets supplied with a continuous flow of filtered seawater at 13 °C and oxygen, and fed fresh kelp (*Eklonia maxima*) once a week. In all cases animals were exposed to experimental conditions no more than once in a 2 month period.

3.2.2. Haemocyte sampling and heat shock

The water in a 100 l tank was heated to 23 °C for the heat shock experiment. At the outset, baskets containing 6-7 animals per time point were placed into the heated water. Since the animals were cocktail sized, equal contributions of haemolymph from several animals were pooled to make up sufficient sample volume. Approximately 500 µl of haemolymph was removed from each animal via the pedal sinus with a 21 G x 1.5" needle at each sample point during timed exposure to heated water. A control haemolymph sample was obtained from unstressed animals. The haemolymph was placed in sterile 1.5 ml eppendorf tubes, and the haemocytes were pelleted by centrifugation. The

haemolymph was removed, and the haemocyte pellets were snap frozen in liquid nitrogen, and placed at -80 °C for storage.

3.2.3. Experimental design for the identification of heat response proteins

Heat shock experiments were carried out to identify abalone stress proteins that are differentially regulated during heat shock using iTRAQ. Haemolymph was sampled from abalone after 30, 90 and 180 minutes of heat shock as described above. 8-plex iTRAQ labels (AbSciEx) were used, and the available channels were split into two sets of four time points in order to incorporate two biological replicates within the same iTRAQ analysis. The entire experiment was then repeated for a total of four biological replicates. A schematic overview of this experimental design is depicted in Figure 3-2.

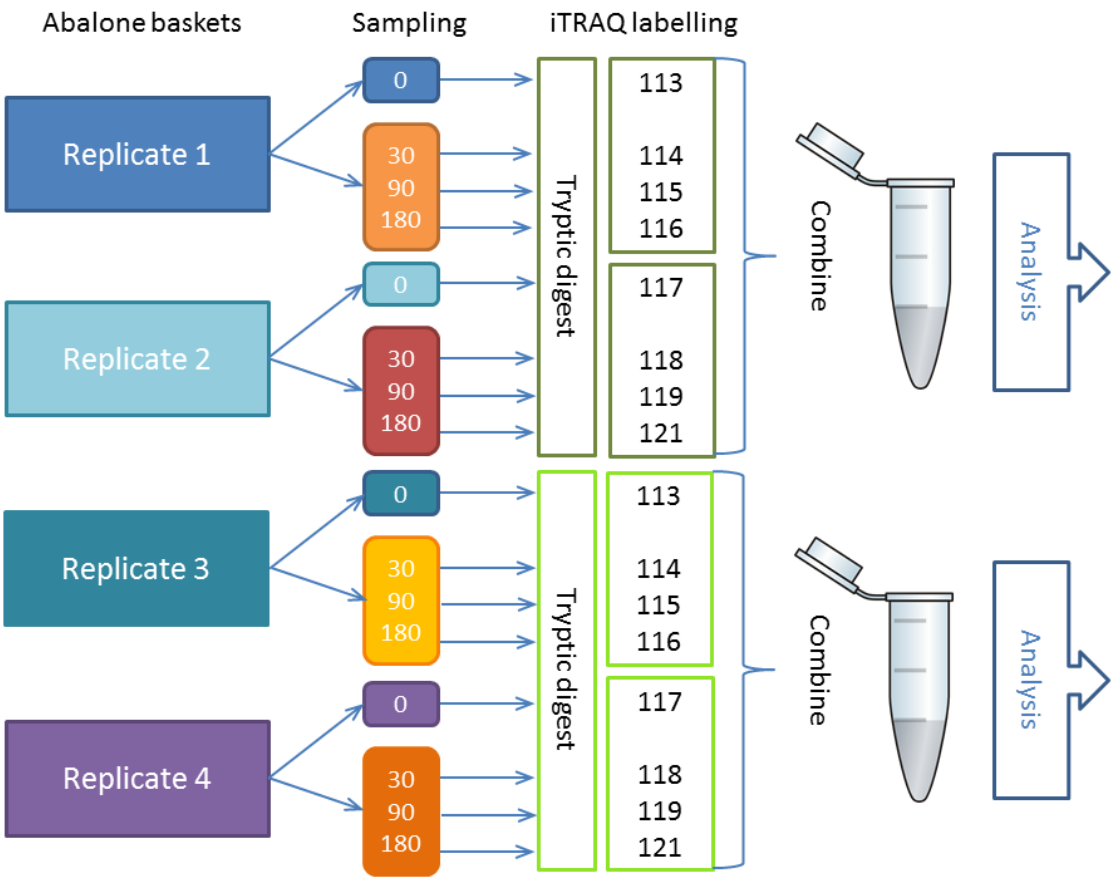


Figure 3-2: iTRAQ experimental design workflow. In each replicate, animals were sampled before heat shock (control sample), and then at 30, 90 and 180 minutes during exposure to heat shock at 23°C. An iTRAQ 8-plex was then split into two groups of 4, each representing a biological replicate. The entire procedure was repeated twice, for a total of four biological replicates made up of two iTRAQ 8-plex label kits.

3.2.4. Sample preparation for iTRAQ analysis

As there is currently no published work describing the appropriate methodology for iTRAQ protein expression profiling of molluscan cells, a workflow was optimised for abalone haemocytes. The workflow began with cell lysis, followed by tryptic digest, iTRAQ labelling, off-gel fractionation and finally MS analysis. At each point in the workflow it was necessary to remove potentially interfering agents in the sample that may have affected the efficiency of the reaction.

3.2.4.1. Cell lysis conditions

In order to determine the best cell lysis buffer for downstream applications, different cell lysis and protein solubilising buffers were investigated. The AB SCIEX iTRAQ and in-solution tryptic digest methodology recommends that the sample be free from thiols such as DTT, active proteases and primary amines, and low in detergents (e.g. SDS) and denaturants (e.g. urea). If any of the above were present in the sample, acetone precipitation was recommended in order to then dissolve the protein in a more suitable buffer. Furthermore, it was recommended that the volume of the tryptic digest should be less than 50 µl, and the pH of the labelling reaction should be > 7.5 in order to maximise labelling efficiency. This meant that any interfering substances needed to be diluted to their appropriate concentrations using a volume less than 50 µl. During previous protein isolations with abalone haemocytes there were concerns that the potentially high concentrations of proteases in the granulocyte haemocytes would degrade the protein sample. The cell lysis buffer previously optimised for abalone haemocytes (used in 2.2.4) contained the chelator EDTA and protease inhibitors, as well as Tris. As these are inappropriate for downstream iTRAQ labelling, buffer exchange was initially attempted by acetone precipitation. In order to satisfy the sample buffer requirements for high efficiency tryptic digest and iTRAQ labelling the secondary buffer was comprised of 1M TEAB with 0.1% w/v SDS.

Buffer exchange by acetone precipitation was performed according to the Applied Biosystems iTRAQ® reagents chemistry reference guide (2004). Briefly, 6 volumes of chilled 100% acetone was added to the sample and incubated at -20 °C overnight. The precipitate was pelleted by centrifugation at 5000 rpm for 5 minutes at 4 °C, and the

acetone was removed from the pellet. The pellet was then resuspended directly in the secondary buffer (1M TEAB with 0.1% SDS). However, the resulting protein pellet was only partially soluble in the secondary buffer. To fully solubilise the pellet, the chaotropic agent urea (8 M) was added to the buffer. The protein pellet was successfully resuspended in 8 M Urea, however in order to reduce the concentration of urea sufficiently for subsequent tryptic digest the sample volume would exceed 50 µl. A buffer exchange protocol that provided both sufficient solubilisation of the protein sample, as well as appropriate buffer compositions for iTRAQ and mass spectrometry analysis workflows was therefore devised.

The protocol was based on a modified filter aided sample preparation (FASP) method reported by Wisniewski et. al. (2009). The haemocyte pellets were lysed in 1% SDS in 0.5M TEAB (Sigma-Aldrich) by vortexing for 30 minutes at room temperature. The cellular debris was then pelleted by centrifugation at 14000 rpm for 30 minutes at 4 °C. The supernatant containing the total cellular protein was removed to a clean 1.5 ml tube, and then quantified using the Pierce™ BCA protein assay kit (Thermo Scientific) using the microplate method, with the provided BSA solution to generate a standard curve. The samples were diluted in water to minimise interference between the BCA reagent and the SDS in the lysis buffer. Each standard and sample (20 µl) was pipetted into a 96 well microplate in triplicate, and 200 µl of working reagent was added to each well. The plate was then incubated for 30 minutes at 37 °C before the absorbency was read at 590 nm on a Thermo Scientific Multiskan plate reader. Protein for tryptic digestion (300 µg) was then placed in a low bind 1.5 ml centrifuge tube (Protein loBind tube, Lasec), and 0.1 volumes of 50 mM TCEP (Sigma-Aldrich) was added. Cysteine disulphide bonds were then reduced by incubation at 60 °C for 1 hour. The reduced protein sample was transferred into a 30 kDa molecular weight cut off centrifugal filter (Merck), inserted into the supplied collection tube, and the volume was reduced to 30 µl by centrifugation at 10000 g at room temperature. The sample was then incubated for 15 minutes at room temperature with 100 µl of 8 M urea in 0.5 M TEAB containing 15 mM MMTS (Sigma-Aldrich) to block cysteine residues. Two washes with 8 M urea in 0.5 M TEAB followed to reduce the concentration of SDS in the sample. In each wash the minimum volume of retentate was left in the filter before the next wash commenced. In a similar fashion, washes with 0.5 M TEAB were then carried out to reduce the concentration of urea in the sample to an acceptable level (approximately 1 M).

3.2.4.2. Tryptic digest

The final 30 µl retentate, containing no more than 0.1% w/v SDS or 1 M urea, was removed from the centrifugal filter by inverting the filter in the provided collection tube and centrifuging for 5 minutes at 10000 g. The sample was removed to a new low bind 1.5 ml tube and proteomics-grade modified trypsin (Trypsin Gold, Promega®) was added to a final protein:trypsin ratio of 100:1 (w/w). Optimal trypsin activity occurs at an alkaline pH so the pH was tested using pH strips, and adjusted to approximately pH 9 with 0.5 M TEAB if necessary. The tryptic digest was allowed to proceed overnight at 37 °C in an air incubator in eppendorf tubes sealed with Parafilm® (Sigma-Aldrich).

3.2.4.3. iTRAQ labelling

The digested protein was labelled using the AbSciEx 8-plex iTRAQ system. The labels were reconstituted with proteomics grade isopropanol (Sigma-Aldrich) and added to each sample respectively. The label and sample were then mixed by vortexing, brought to the bottom of the tube by pulse centrifugation and then incubated at room temperature for 2 hours. After labelling the samples were combined and carefully mixed by vortexing.

3.2.4.4. Off-gel fractionation and preparation of labelled samples for analysis by mass spectrometer

Any contaminating factors were removed from the labelled tryptic peptides through a C-18 resin spin column (Thermo Scientific Pierce) with the modifications to the method suggested by the manufacturer specifically for ESI MS/MS. In order to load the maximum amount of peptide onto the IPG strip for off-gel fractionation, eight columns were therefore used to clean the labelled peptide mix for a maximum possible total of 240 µg of peptide. All C-18 column equilibration, sample binding, washing and eluting was facilitated by centrifugation at 9000 rpm for 1 minute at 15 °C. The samples were prepared by the addition of sample buffer (2% TFA (Sigma-Aldrich) in 20% acetonitrile (Sigma-Aldrich)), at a ratio of 3 parts sample to 1 part sample buffer. The columns were activated and equilibrated with 50% acetonitrile, and 0.5% TFA in 5% acetonitrile respectively, and the peptides were bound to the column. Contaminants were rinsed out by washing with 0.5% TFA in 5% acetonitrile, and the peptides were then eluted into clean low-bind 1.5 ml

centrifuge tubes (Lasec). The cleaned samples in elution buffer (70% acetonitrile with 0.1% formic acid (Sigma-Aldrich)) were dehydrated completely by vacuum centrifugation so that they could be resuspended in 1x peptide rehydration solution (50% glycerol, Bio-Lyte ampholytes (GE Healthcare)) prior to off-gel fractionation.. The cleaned, labelled peptides were separated into 12 fractions across a pI gradient of 4-7 using the 12 well Agilent Off-gel fractionation system. The peptides were separated along a 13 cm IPG strip (pH 4-7, GE healthcare) using the standard pre-loaded protocol for peptide fractionation.

After fractionation it was necessary to remove glycerol and other contaminating substances prior to analysis by ESI MS/MS mass spectrometry. Therefore each fraction was once again applied to a C-18 spin column (Thermo Scientific) and cleaned as described above, and the cleaned eluent was once again dehydrated by vacuum centrifugation before analysis by mass spectrometry.

3.2.4.5. Mass spectrometry settings

Mass spectrometry analysis was carried out on each of the 12 cleaned peptide fractions on an Agilent 6530A Accurate-Mass QTOF LC/MS system coupled to a 1200 series HPLC-chip/MS system. The HPLC-Chip had a 75 μm x 150 mm analytical column packed with ZORBAX 300SB C18 (5 μm; 300 A), and a 160 nL enrichment column. The injection volume was 8 μl, with a 2 μl/min flow on the loading pump and 0.4 μl/min flow on the analytical pump. The mobile phases were comprised of solvent A (0.1% formic acid in water) and solvent B (90% acetonitrile in water with 0.1% formic acid). The gradient was formed with 2% B at 0 minutes, 7% B at 1 minute, 16% B at 14 minutes 30 % B at 44 minutes, 40 % B at 56 minutes, 70% B at 66 minutes and 70% B at 70 minutes, with the stop time at 70 minutes.

The spectra were acquired in positive ion mode with a drying gas flow of 4.0 l/min and a gas temperature of 355 °C. The MS scan range m/z was 250-1800, and the MS scan rate was 6.00 /sec. The MS/MS scan range m/z was 90-1800 and the scan rate was 2.50/sec. The isolation width for MS/MS was set to medium (~4 amu). The ramped collision energy was 4.200 with an offset of -4.2. The capillary voltage was set to 1960.

3.2.4.6. Database compilation

For the purposes of this study, the UniProt SwissProt/TREMBL database (uniprot.org) was accessed on 20 August 2013 in order to download all known molluscan protein sequences. The database was downloaded in FASTA format, and then verified in Peaks Studio and used in all subsequent searches. The taxonomic composition of this database was analysed using the taxonomy tool on the UniProt website in order to show the proportion of protein sequences represented by each molluscan taxonomic group.

3.2.5. iTRAQ data processing

Data files for each of the 12 fractions per iTRAQ experiment were acquired from the Agilent MassHunter software in .d file format, and converted to .mzxml using the free tool MSConvert (ProteoWizard, version 3.0.3662) to facilitate downstream analysis. The .mzxml files were loaded into Peaks Studio software (Bioinformatics Solutions Inc., version 6.0) and the quantitation workflow was selected. SPIDER, PTM Finder and *de novo* searches were all performed in order to maximise the number of peptide identities obtained. The *de novo* search was set up using the default settings for total local confidence and average local confidence. The precursor mass tolerance for the database search was set at 20 ppm, and the fragment mass tolerance was set at 0.1 Da. Two missed enzyme cleavages were allowed per peptide. The fixed modifications included were beta-methylthiolation and iTRAQ labelling at the N terminus, and the variable modifications were set as oxidation of Methionine and iTRAQ labelling at the Y terminus. The false discovery rate was determined by Peaks Studio using a concatenated decoy database which was automatically generated from the search database. The Peaks Quantitation results were then filtered so that the false discovery rate was <1%, which set the -10 logp score at a threshold below which peptides would not be considered in the analysis. Furthermore, only proteins with 2 or more unique peptides were considered. Raw intensity data that satisfied the above criteria was then auto-normalised by Peaks Studio to correct for total channel intensity differences. This data was then exported in a .csv format.

Proteins that respond reproducibly during exposure to heat shock were identified by carrying out two heat shock experiments for analysis by iTRAQ. Proteins in the haemocytes of *H. midae* exposed to thermal stress for 30, 90 and 180 minutes sampled

during these experiments were identified and their expression profiles were determined by iTRAQ analysis. Each iTRAQ experiment was processed individually in Peaks Studio, according to 3.2.5, so that both Peaks search and *de novo* peptide identities were combined into the quantitative analysis. The resulting lists of protein identities and quantitative data were exported into MS Excel, and the two independent iTRAQ experiments (each representing two biological replicates) were combined. This list was then used to identify proteins with quantitative data which were common to both iTRAQ experiments, thereby ensuring that only high quality protein identities were carried forward. The common protein list was manually edited so that it contained only one example of each protein identity, as the homologous database contained multiple entries of the same protein in some cases. In all cases this was because the protein sequence had been reported in more than one species. In cases with ambiguous protein identity the protein identity from the species with the highest -10logp score and/or the most unique peptides between both iTRAQ experiments was selected as the 'correct' identity. Therefore, after manual editing, a list of proteins that were expressed in all four biological replicates with greater than 2 unique peptides and only high quality PSMs was generated.

3.2.5.1. Identification of differentially regulated proteins

The list of protein identities and associated quantitation values was imported into CLC workbench for processing before statistical analysis. The data were quantile normalised (box plots showing the dispersion of expression values for raw and normalised data are included in Appendix B) and then log2 transformed before being exported into MS Excel where the values were normalised to their respective T0 values for each biological repeat and each protein. The differences between the control and experimental groups' relative expression values were then determined by one-way ANOVA using the MS Excel add-on Analysis ToolPak. The proteins with a statistically significant difference between control and heat shocked abalone samples (p value <0.05) were selected. Once the first round of one-way ANOVA ascertained the proteins that were significantly differentially expressed in all four biological replicates, a second round of one-way ANOVAs was performed on those proteins with a similar expression pattern in 3 out of the 4 biological replicates.

The average expression across 4 biological replicates was calculated for each protein identity in MS Excel and analysed using the tool PermutMatrix (version 1.9.3) (Caraux and

Pinloche, 2005), which uses seriation and clustering to graphically represent the expression patterns for genes or proteins in a heat map. The data were ordinated based on Euclidean distance, with hierarchical clustering performed by average distance based on the unweighted means of all pairwise distances (UPGMA) and seriation performed using a multiple-fragment heuristic in order to provide a 'reasonably good' solution. The data matrix was presented in a heat map with an associated dendrogram using a green colour to represent a decrease in amount vs. the control, and a red colour to indicate an increase in amount vs. the control. This matrix was exported into the image manipulation software Gimp (version 2.8.10) in order to resize the image and divide the output .jpeg produced by PermutMatrix over 2 pages for presentation here.

3.2.5.2. Gene ontology analysis of differentially regulated proteins

Blast2GO® is a JAVA based tool used for the functional annotation of sequences, particularly those for which no GO terms are associated. Blast2GO® uses the BLAST algorithm to find homologous sequences to the input database, and then maps relevant GO terms based on homology with proteins from other organisms.

Although Blast2GO® is effective at finding GO terms for non-model protein identities, the terms can often be very general and therefore of limited value in mapping the functional classification of proteins. Therefore, the tool REViGO was used to simplify and visualise the GO terms more meaningfully, so that the important biological processes and molecular functions implicated by the protein identities obtained by iTRAQ could be identified. GO identities obtained by Blast2GO analysis were exported into REViGO, and redundant terms were removed in order to simplify the list within an allowed similarity of 0.5 (for a 'medium' sized output file). The identities were then arranged in treemaps based on similarity, using the UniProt database as a reference. The treemaps clustered GO term clusters into 'superclusters' with similar function, thereby identifying general cellular activities that were affected by heat stress. In order to simplify the analysis and ensure that the data output was meaningful, only proteins showing extremes of expression were analysed in REViGO. Therefore, the dendrogram generated by cluster analysis was used to identify the groups of proteins that responded to heat shock at every sampled time point, either by increasing or a decreasing in expression.

3.2.6. Validation of iTRAQ data by western blot hybridisation

Due to the large amount of information generated by quantitative iTRAQ analysis, it is necessary to validate the data using an alternate molecular technique. Western blot hybridisation analysis was used to validate the expression of specific *H. midae* proteins that were found to respond to thermal stress. Since *H. midae* has no species-specific commercially available antibodies, the antibodies selected were specific to proteins that had the highest homology to specific *H. midae* proteins that were identified by iTRAQ.

Despite the large number of antibodies available commercially, only three of the significantly expressed proteins could be matched to appropriate antibodies. Antibodies for GSTO1 (HPA037604 Sigma) and G6P1D (HPA000247 Sigma) were purchased from the Sigma Prestige antibodies® range**.** The antibody to heat shock protein 70 (Hsp70) had already been used in this laboratory, and was obtained from Abcam (ab79852).

3.2.6.1. Heat shock and total protein isolation

A set of heat shock experiments were carried out as described in 3.2.2 to generate four new biological replicates. Three animals per time point were sampled pre-stress as a control (T0), and then at 30, 90 and 180 minutes exposure to water at 23 °C. As western blotting does not have the sample buffer requirements that iTRAQ has, the original protein lysis buffer (2.2.4) could again be used for cell lysis. Haemocyte total protein was extracted by vortexing cells in 200 µl of protein lysis buffer (Appendix A). The lysate was incubated at room temperature for 5 minutes, and then the cellular debris was pelleted by centrifugation at 7000 rpm for 30 minutes at 4 °C. The supernatant containing the total soluble haemocyte protein was transferred to a clean 1.5 ml eppendorf tube.

3.2.6.2. Determination of total protein concentration

Total protein concentration was determined using the BCA protein assay (Thermo Scientific) using the microplate method as described in 3.2.4.1.

3.2.6.3. Western blot optimisation

To determine the optimum concentrations of primary and secondary antibody, 20 μg protein samples were dotted onto nitrocellulose membrane and allowed to dry at room temperature. Membranes were blocked in blocking buffer (5% w/v skim milk in 1 X TBS (Appendix A)) for 1 hour at room temperature with shaking. Primary antibody diluted in blocking buffer was applied according to the manufacturer's suggested concentrations, which were antibody-dependent and ranged from 1:10000 (Hsp70) to 1:500 (GSTO1). Additional concentrations were selected for the eventuality that the suggested concentration was too dilute. Membranes were incubated with primary antibody overnight at 4°C with shaking. The membrane was then washed three times in 1x TBS for 10 minutes at room temperature with shaking, and then the secondary antibody was applied. Secondary antibodies were applied at either 1:10000 or 1:20000 in blocking buffer and allowed to hybridise for 2 hours at room temperature with shaking. Residual secondary antibody was removed by three washes with 1x TBS for 10 minutes at room temperature with shaking. The dot blots were visualised using the Advansta WesternBright Chemiluminescent ECL HRP substrate according to the manufacturer's instructions, and the Bio-Rad ChemiDoc XR system for detection. The lowest possible concentrations of both primary and secondary antibody to give a signal were selected for each antibody.

The concentration of antigen required to give sufficient signal was determined by performing western blot hybridisation on three different amounts of antigen for each antibody. In each case, a western blot of 5, 10 and 15 μg of total *H. midae* haemocyte protein was carried out using the method described below. The primary and secondary antibodies were applied in the concentrations determined above and the blot was visualised using the detection method described above. The lowest amount of antigen that gave a detectable signal without any saturated pixels was selected for subsequent analyses.

3.2.6.4. Expression analysis by western blot hybridisation

The expression of the selected proteins in the haemocytes of *H. midae* during exposure to heat shock was measured by western blot analysis. Each protein sample was boiled at 95 °C in a 1.5 ml eppendorf with 1x sample application buffer (Appendix A) in order to denature the protein. The samples were then pipetted into gel lanes in a 4% v/v stacking

gel and separated electrophoretically through a one dimensional 10% v/v polyacrylamide separating gel (Appendix A) (Ausubel *et al.,* 1989) under denaturing conditions using the Bio-Rad mini PROTEAN tetra electrophoresis system. The PAGE ruler protein molecular weight marker (Thermo Scientific) was included to determine the molecular weight of the band/s detected. The gels were electrophoresed at a constant voltage of 100 V in 1x SDS electrophoresis buffer (Appendix A) until the bromophenol blue front of the sample application buffer had reached the bottom of the glass plate. The gels were then removed from the electrophoresis tank and prepared for western blotting.

The gels were transferred to a plastic container and incubated in cold Towbin buffer (Appendix A) at 4 °C for approximately 5 minutes to remove residual SDS. A nitrocellulose membrane (Whatman® Protran® 0.2 μM pore size) cut to fit the gel was also placed in Towbin buffer to equilibrate. The western transfer apparatus used was the Bio-Rad mini Trans-Blot cell which was set up according to the manufacturer's instructions. The membrane and gel were placed between two pieces of pre-soaked Whatmann 3 MM filter paper and sandwiched together in the transfer cassette. This was then loaded vertically into the transfer tank, which was filled with Towbin buffer and 100 V was passed through the apparatus for one hour at 4 °C until the protein was completely transferred from the gel to the membrane.

Efficient transfer and equal loading was confirmed by Ponceau S reversible total protein stain, the images of which were captured using the Bio-Rad ChemiDoc XR system with 'colorimetric' settings, optimised for faint bands. The membrane was then immersed in blocking buffer (5% w/v skim milk powder in TBS) on a shaker for 1 hour at room temperature. After blocking, the membrane was incubated with the primary antibody in 5% (w/v) skim milk in TBS overnight at 4 °C on a shaker. The antigen, primary and secondary antibody concentration parameters for each protein in the western blot analysis are presented in (Table 3-1). Thereafter, the membrane was washed in TBS with shaking, before being incubated with goat anti-rabbit peroxidase conjugated secondary antibody (Kirkegaard & Perry Laboratories) in 5% (w/v) skim milk in TBS for 2 hours at 4 °C on a shaker. After incubation with the secondary antibody the membrane was again washed with TBS before visualisation with the WesternBright ECL HRP substrate (Advansta) chemiluminescent detection kit according to the manufacturer's instructions. The substrate was applied directly to the membrane which was placed on the Bio-Rad

Molecular Imager® ChemiDoc™ XRS imager. The apparatus was set up for chemiluminescent imaging using ImageLab software (version 4.1), and the image was acquired with sequential exposures for 100 seconds, with ten seconds between exposures. The image with the fewest saturated pixels was selected for analysis.

The size of the bands detected for each antibody was determined using the Fermentas PAGE ruler prestained molecular weight marker and the molecular weight analysis tool provided in Bio-Rad's ImageLab software. Volume analysis using ImageLab was carried out using a user-defined rectangular area for each western blot. The background subtraction method was set to 'local', and an area containing no bands was defined for background subtraction. The adjusted intensity values for each band were exported into MS Excel in .csv format. Replicate values were used to calculate average relative expression values for each protein. A one-way ANOVA was performed using the MS Excel analysis ToolPak add-on, with statistical significance established if $p < 0.05$.

Table 3-1: The detection parameters for each antibody used in quantitative western blot analysis of the thermal response to heat stress in *H. midae* haemocytes.

Protein target	Antigen concentration (μg)	Primary antibody concentration	Secondary antibody concentration	Exposure time (s)
Hsp70	15	1:10000	1:10000	100
G6P1D	15	1:1000	1:10000	100
GSTO1	15	1:500	1:10000	100

3.3. Results

3.3.1. Database compilation and analysis

In order to successfully identify proteins from *H. midae* using mass spectrometry, a database comprising molluscan protein sequences was accessed and downloaded from the UniPROT repository (www.uniprot.org). At the time of access (20 August 2013), 68821 molluscan sequences were available in the UniPROT/TrEMBL database (Figure 3-3)– of which only 1754 were 'reviewed', and the rest were predicted proteins based on genomic sequences. The *C. gigas* complete proteome accounts for 25978 of these sequences. At the time of access 1282 of the sequences originate from the genus *Haliotis* (Figure 3-4).

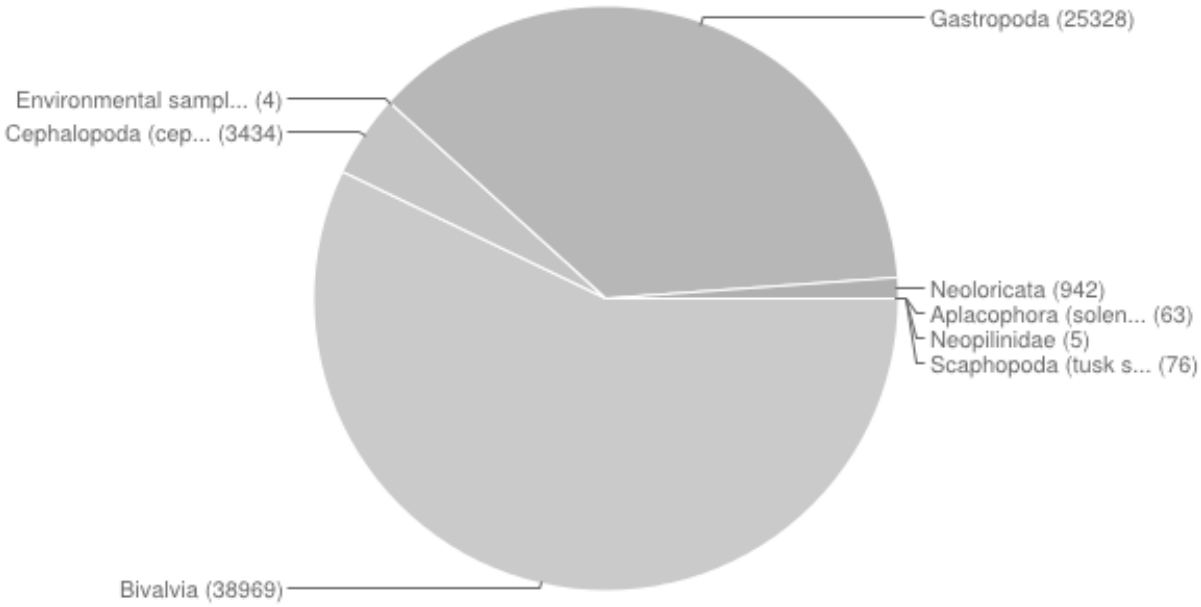


Figure 3-3: The distribution of molluscan classes present in the protein database used for peptide spectrum matching in this study. Pie chart downloaded from www.uniprot.org on 20 August 2013, showing taxonomic makeup of the database. The biggest single species contributing to the database is *C. gigas*, the full deduced proteome of which is included in class Bivalvia. Although the limpet *Lottia gigantea* has a fully sequenced genomic database, it is not included in the UniProt database.

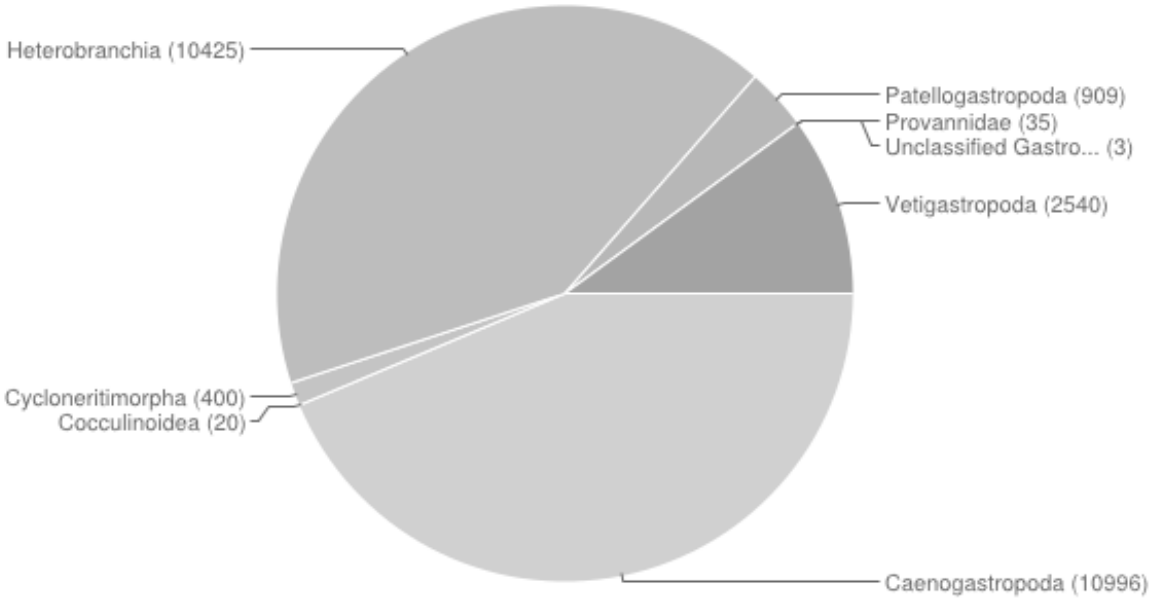


Figure 3-4: The superorder distribution of the major subclass Orthogastropoda of phylum Gastropoda, included in the protein database used for peptide spectrum matching in this study. Superorder Vetigastropoda contains the superfamily Halliotoidea, of which *H. midae* is a member, is represented by 2540 sequences. The majority of Gastropod sequences in this database originate from the Heterobranchia (snails and slugs) and Caenogastropoda (sea snails). Numbers of sequences correct at time of access (20 August 2013).

3.3.2. Identification and expression profiling of heat shock responsive proteins

iTRAQ analysis was used to obtain a list of differentially expressed proteins that are present in the haemocytes of *H. midae* during exposure to acute heat stress. The multiplexing capabilities of iTRAQ analysis make it possible to determine protein expression profiles in up to 8 samples simultaneously. This enabled the analysis of two distinct biological replicates with four sampling points per replicate within one iTRAQ experiment. The entire iTRAQ experiment was then repeated for a total of four biological replicates.

A total of 388 protein identities from iTRAQ #1 and 510 from iTRAQ #2 with more than 2 unique peptides and quantitation data were used to compile a list of 116 proteins that were common to both experiments (Appendix B).

In order to identify proteins which respond to heat stress with statistical significance ($p < 0.05$), one-way ANOVA was used to compare control and experimental expression values. In this manner, 8 proteins that were significant ($p < 0.05$) in all four biological replicates were identified, and an additional 3 proteins that were significant in 3 out of the 4 biological replicates were identified, and are presented in Table 3-2. One of these proteins (B3TK33_HALDV) originates in *H. diversicolor* but has not yet been characterised. The rest of the significant proteins originate in abalone (*H. diversicolor, H. discus discus* and *H. discus hannai*), sea slug (*Aplysia californica*), arrow squid (*Ommastrephes sloanei*), oyster (*C. gigas*) and clam (*Chlamys farreri*). Omega crystallin (CROM_OMMSL) had the largest fold change increase during exposure to heat shock, and haemocyanin isoform 1 (C7FEG7_HALDV) had the largest fold change decrease at a single sample point.

The graphical analysis program PermutMatrix (Caraux and Pinloche, 2005) was used to perform statistical seriation on the dataset. The resulting heat map and dendrogram are shown in Figure 3-5, with up-regulation indicated by red colour and down-regulation indicated by green colour, according to the scale provided. Proteins that cluster in the same branches of the dendrogram have similar heat-induced expression profiles. The dendrogram generated by seriation of the columns (time points) indicates that samples obtained at 30 and 90 minutes group together, with the sample obtained at 180 minutes exhibiting a distinct expression pattern to proteins analysed at the earlier time points. Based on this clustering analysis, groups of proteins were identified that respond to heat shock with increased and decreased expression levels at all the sampling points (Figure 3-5). Thus, 16 up-regulated proteins and 8 proteins that were down-regulated were grouped at the top and bottom of the dendrogram respectively. The expression data for these proteins is presented in Table 3-3.

Table 3-2: Proteins determined to be differentially expressed at various time points (min) during exposure to heat shock. Accessions refer to the UniProt database. P values were determined by one way ANOVA of four biological replicates, except where three biological replicates are indicated by an asterisk. Expression values are average, log2 transformed, quantile normalised iTRAQ label intensity values, normalised to the control (0).

			Expression			
Accession	**Description**	**p-value**	**0**	**30**	**90**	**180**
B3TK33_HALDV	Putative uncharacterized protein (Fragment) OS=*Haliotis diversicolor*	0.035024	0	-0.02	0.07	-0.30
B3TK66_HALDV	40S ribosomal protein S8 OS=*Haliotis diversicolor*	0.019792*	0	0.16	0.17	0.26
B6RAZ8_HALDI	Omega class glutathione-s-transferase 1 OS=*Haliotis discus discus*	0.044*	0	-0.23	0.09	-0.07
B6RB43_HALDI	ATP synthase_ H+ transporting_ mitochondrial F1 complex_ o subunit OS=*Haliotis discus discus*	0.020411	0	-0.27	-0.38	0.09
Q2MJK5_HALDH	Heat shock protein 70 OS=*Haliotis discus hannai*	0.019873*	0	0.17	0.16	0.03
C7EAA2_HALAI	PL10-like protein OS=*Haliotis asinina*	0.024847	0	0.24	0.28	0.24
C7FEG7_HALDV	Hemocyanin isoform 1 (Fragment) OS=*Haliotis diversicolor*	0.024317	0	-0.57	-0.54	-0.31
RHO_APLCA	Ras-like GTP-binding protein RHO OS=*Aplysia californica*	0.04526	0	-0.14	-0.36	-0.31
CROM_OMMSL	Omega-crystallin OS=*Ommastrephes sloanei*	0.018064	0	0.05	0.08	0.32
Q0KHB8_CRAGI	Glucose-6-phosphate 1-dehydrogenase OS=*Crassostrea gigas*	0.052043	0	0.09	0.06	0.25
Q8MUE3_9BIVA	Small nuclear ribonucleoprotein D2-like protein OS=*Chlamys farreri*	0.016696	0	0.21	-0.06	-0.07

Rows : - Objective function : R=0.656

- Sum of all pairwise distances of neighboring rows (path length): S=11.776
- Linkage rule: Average linkage
- Tree Seriation rule: Multiple-fragment heuristic (MF)

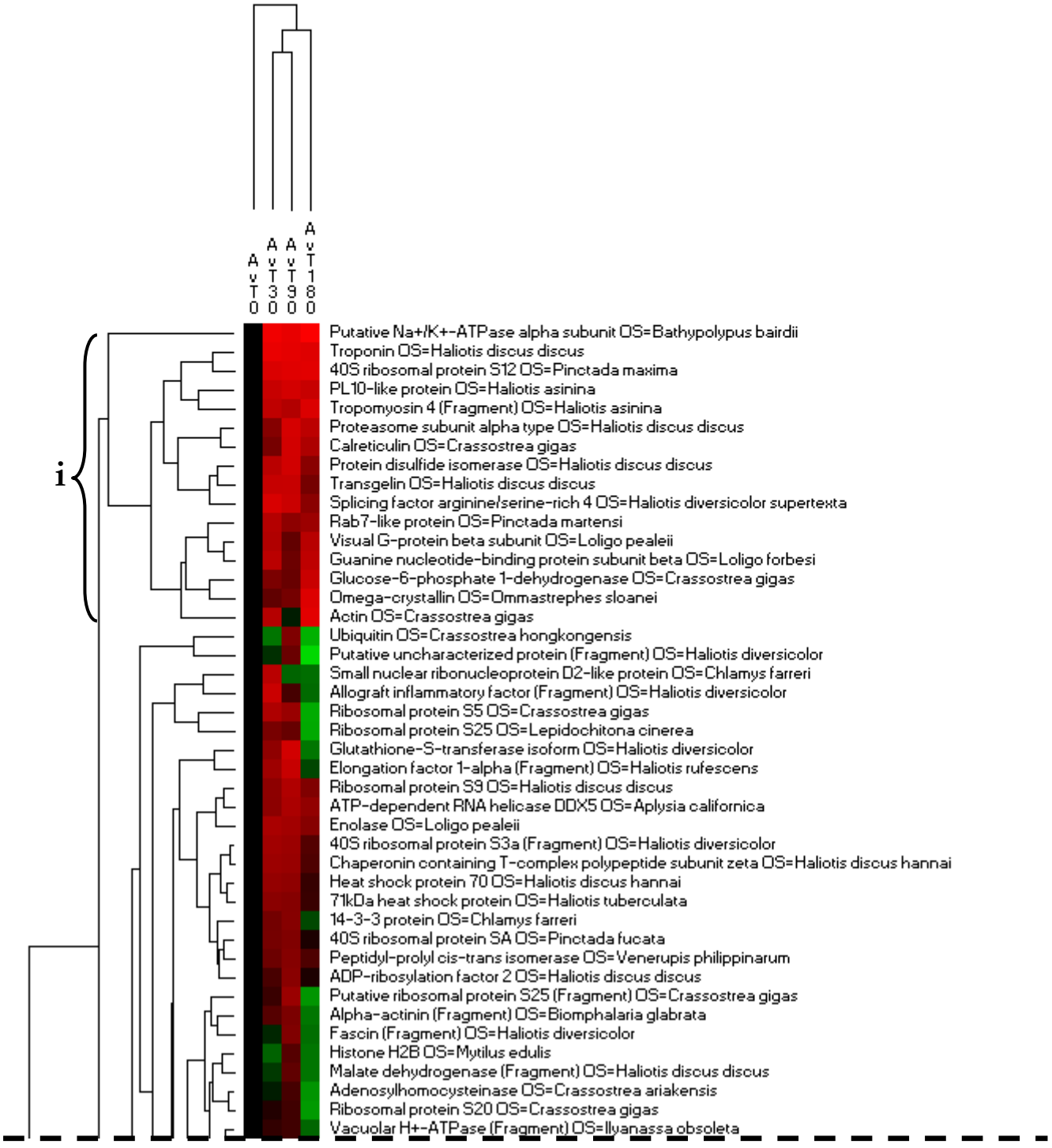
Columns : - Objective function : R=0.776

- Sum of all pairwise distances of neighboring columns (path length): S=4.906
- Linkage rule: Average linkage

Dissimilarity : - Euclidean distance

The colors scale:

Min = -0.57 0.00 Max = 0.57



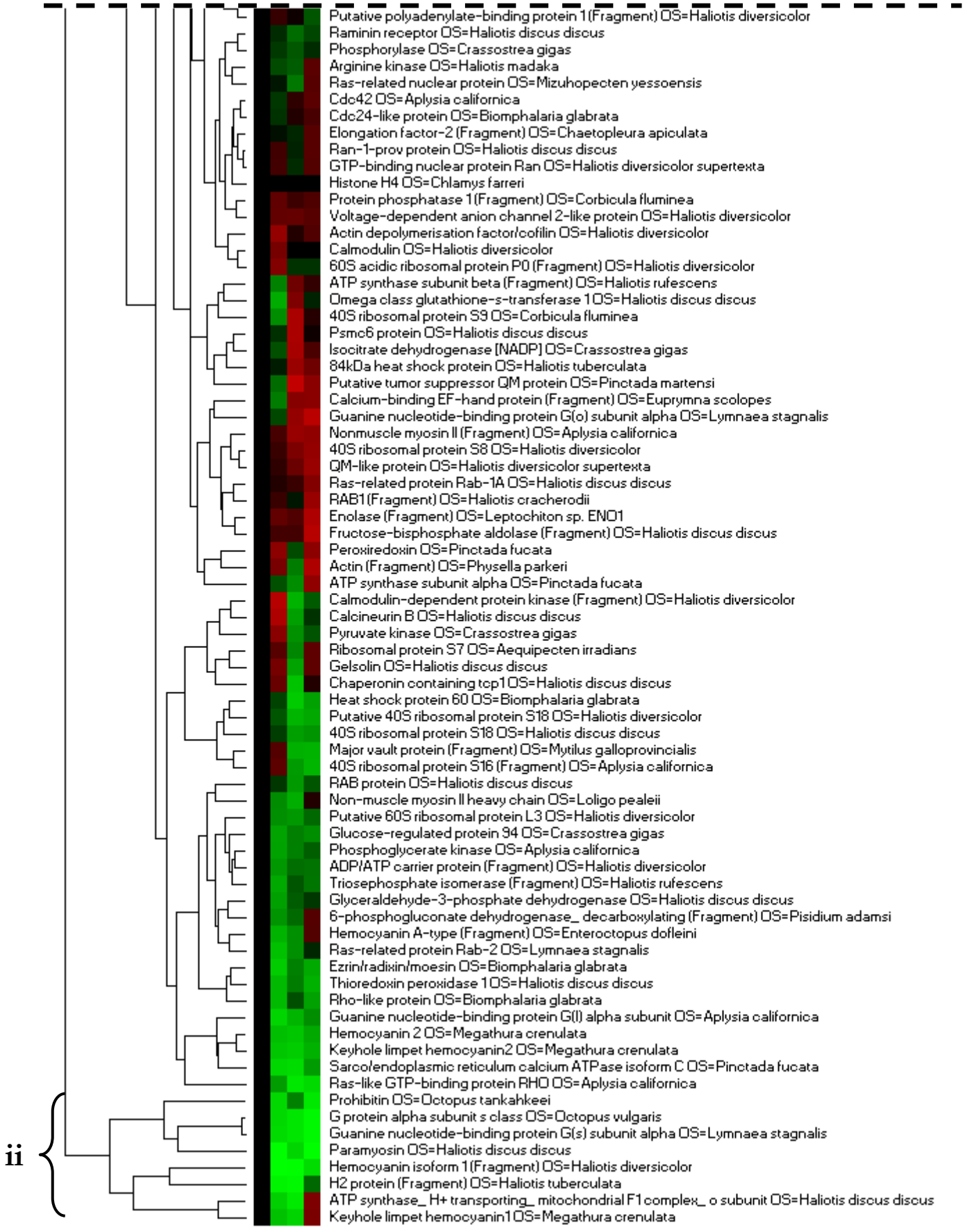

Putative polyadenylate-binding protein 1 (Fragment) OS=Haliotis diversicolor
Raminin receptor OS=Haliotis discus discus
Phosphorylase OS=Crassostrea gigas
Arginine kinase OS=Haliotis madaka
Ras-related nuclear protein OS=Mizuhopecten yessoensis
Cdc42 OS=Aplysia californica
Cdc24-like protein OS=Biomphalaria glabrata
Elongation factor-2 (Fragment) OS=Chaetopleura apiculata
Ran-1-prov protein OS=Haliotis discus discus
GTP-binding nuclear protein Ran OS=Haliotis diversicolor supertexta
Histone H4 OS=Chlamys farreri
Protein phosphatase 1 (Fragment) OS=Corbicula fluminea
Voltage-dependent anion channel 2-like protein OS=Haliotis diversicolor
Actin depolymerisation factor/cofilin OS=Haliotis diversicolor
Calmodulin OS=Haliotis diversicolor
60S acidic ribosomal protein P0 (Fragment) OS=Haliotis diversicolor
ATP synthase subunit beta (Fragment) OS=Haliotis rufescens
Omega class glutathione-s-transferase 1 OS=Haliotis discus discus
40S ribosomal protein S9 OS=Corbicula fluminea
Psmc6 protein OS=Haliotis discus discus
Isocitrate dehydrogenase [NADP] OS=Crassostrea gigas
84kDa heat shock protein OS=Haliotis tuberculata
Putative tumor suppressor QM protein OS=Pinctada martensi
Calcium-binding EF-hand protein (Fragment) OS=Euprymna scolopes
Guanine nucleotide-binding protein G(o) subunit alpha OS=Lymnaea stagnalis
Nonmuscle myosin II (Fragment) OS=Aplysia californica
40S ribosomal protein S8 OS=Haliotis diversicolor
QM-like protein OS=Haliotis diversicolor supertexta
Ras-related protein Rab-1A OS=Haliotis discus discus
RAB1 (Fragment) OS=Haliotis cracherodii
Enolase (Fragment) OS=Leptochiton sp. ENO1
Fructose-bisphosphate aldolase (Fragment) OS=Haliotis discus discus
Peroxiredoxin OS=Pinctada fucata
Actin (Fragment) OS=Physella parkeri
ATP synthase subunit alpha OS=Pinctada fucata
Calmodulin-dependent protein kinase (Fragment) OS=Haliotis diversicolor
Calcineurin B OS=Haliotis discus discus
Pyruvate kinase OS=Crassostrea gigas
Ribosomal protein S7 OS=Aequipecten irradians
Gelsolin OS=Haliotis discus discus
Chaperonin containing tcp1 OS=Haliotis discus discus
Heat shock protein 60 OS=Biomphalaria glabrata
Putative 40S ribosomal protein S18 OS=Haliotis diversicolor
40S ribosomal protein S18 OS=Haliotis discus discus
Major vault protein (Fragment) OS=Mytilus galloprovincialis
40S ribosomal protein S16 (Fragment) OS=Aplysia californica
RAB protein OS=Haliotis discus discus
Non-muscle myosin II heavy chain OS=Loligo pealeii
Putative 60S ribosomal protein L3 OS=Haliotis diversicolor
Glucose-regulated protein 94 OS=Crassostrea gigas
Phosphoglycerate kinase OS=Aplysia californica
ADP/ATP carrier protein (Fragment) OS=Haliotis diversicolor
Triosephosphate isomerase (Fragment) OS=Haliotis rufescens
Glyceraldehyde-3-phosphate dehydrogenase OS=Haliotis discus discus
6-phosphogluconate dehydrogenase_ decarboxylating (Fragment) OS=Pisidium adamsi
Hemocyanin A-type (Fragment) OS=Enteroctopus dofleini
Ras-related protein Rab-2 OS=Lymnaea stagnalis
Ezrin/radixin/moesin OS=Biomphalaria glabrata
Thioredoxin peroxidase 1 OS=Haliotis discus discus
Rho-like protein OS=Biomphalaria glabrata
Guanine nucleotide-binding protein G(I) alpha subunit OS=Aplysia californica
Hemocyanin 2 OS=Megathura crenulata
Keyhole limpet hemocyanin2 OS=Megathura crenulata
Sarco/endoplasmic reticulum calcium ATPase isoform C OS=Pinctada fucata
Ras-like GTP-binding protein RHO OS=Aplysia californica
Prohibitin OS=Octopus tankahkeei
G protein alpha subunit s class OS=Octopus vulgaris
Guanine nucleotide-binding protein G(s) subunit alpha OS=Lymnaea stagnalis
Paramyosin OS=Haliotis discus discus
Hemocyanin isoform 1 (Fragment) OS=Haliotis diversicolor
H2 protein (Fragment) OS=Haliotis tuberculata
ATP synthase_ H+ transporting_ mitochondrial F1 complex_ o subunit OS=Haliotis discus discus
Keyhole limpet hemocyanin1 OS=Megathura crenulata
ii

Figure 3-5: A heat map showing statistical seriation of the average expression profiles for proteins identified by iTRAQ in *H. midae* haemocytes during exposure to thermal stress. Expression values for each protein were log2 transformed, quantile normalised and normalised to the control prior to input into PermutMatrix software for seriation. Protein descriptions were obtained from the UniProt database and are associated with accession numbers indicated in Appendix B. The image generated using PermutMatrix was too large to fit onto a single page, and was therefore split across two pages at the dotted line. Brackets indicate groups of proteins that overall i) increase and ii) decrease relative to the control at all sampled time points during heat shock, which were used in downstream gene ontology analysis.

Table 3-3: Proteins with differential expression profiles at time points (min) sampled during exposure to heat shock. Proteins identified by iTRAQ were statistically seriated in order to identify groups of proteins with similar heat shock induced expression profiles. The proteins that clustered together at the top and bottom branches of the dendrogram are presented. These are proteins that responded to heat stress by increasing or decreasing expression at every sampled time point during the experiment. Expression values are average, log2 transformed, quantile normalised iTRAQ label intensity values, relative to the control (0).

			Relative Expression			
	Accession	**Description**	**0**	**30**	**90**	**180**
Up-regulated						
	C6JUN0_PINMA	40S ribosomal protein S12 OS=*Pinctada maxima*	0	0.31	0.33	0.34
	C4NY64_CRAGI	Actin OS=*Crassostrea gigas*	0	0.19	-0.01	0.35
	A5LGG9_CRAGI	Calreticulin OS=*Crassostrea gigas*	0	0.08	0.28	0.18
	Q0KHB8_CRAGI	Glucose-6-phosphate 1-dehydrogenase OS=*Crassostrea gigas*	0	0.09	0.06	0.25
	GBB_LOLFO	Guanine nucleotide-binding protein subunit beta OS=*Loligo forbesi*	0	0.21	0.07	0.22
	CROM_OMMSL	Omega-crystallin OS=*Ommastrephes sloanei*	0	0.05	0.08	0.32
	C7EAA2_HALAI	PL10-like protein OS=*Haliotis asinina*	0	0.24	0.28	0.24
	B6RB41_HALDI	Proteasome subunit alpha type OS=*Haliotis discus discus*	0	0.10	0.29	0.21
	B6RB63_HALDI	Protein disulfide isomerase OS=*Haliotis discus discus*	0	0.21	0.28	0.11
	F8V2T8_9MOLL	Putative Na+/K+-ATPase alpha subunit OS=*Bathypolypus bairdii*	0	0.41	0.37	0.51

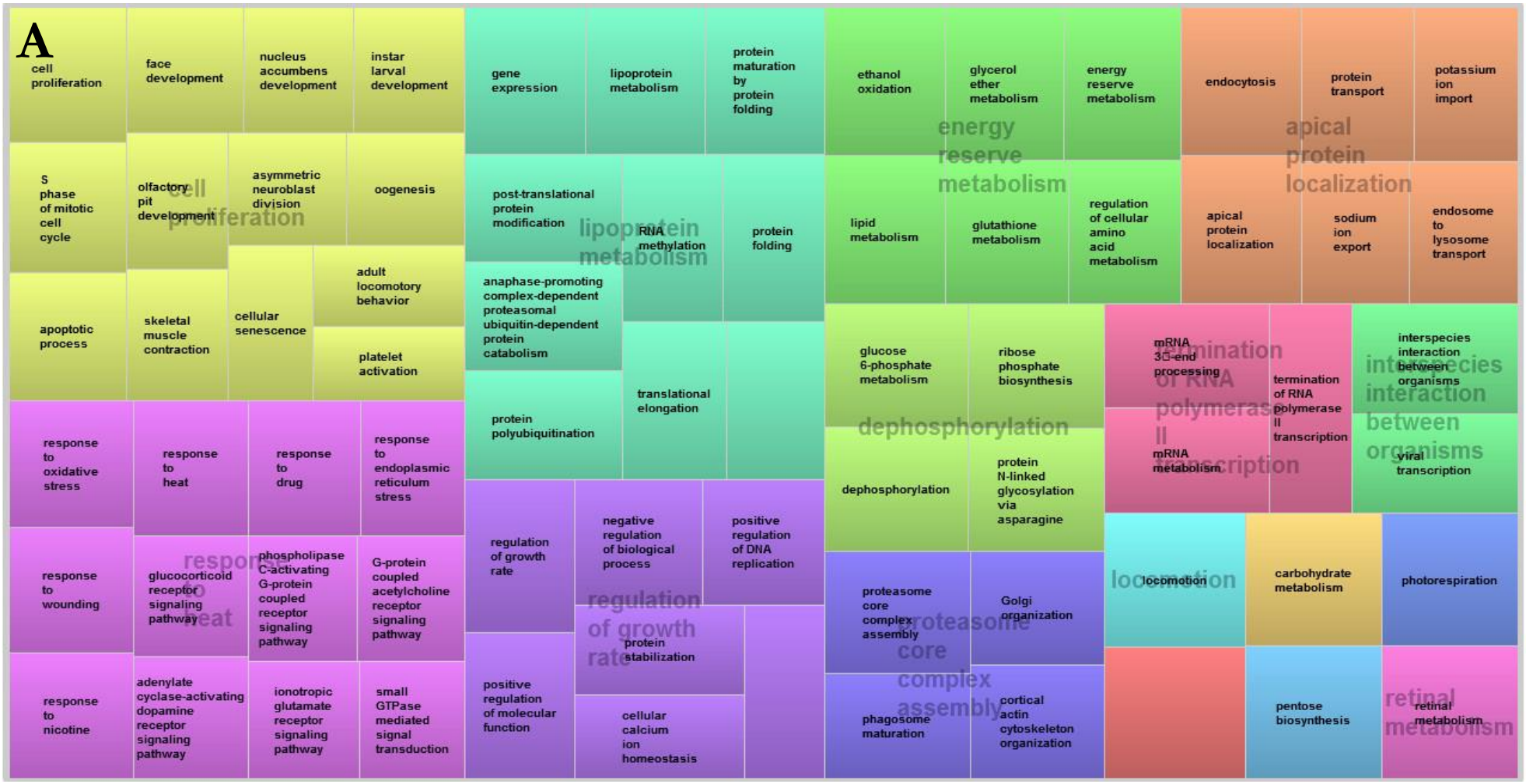
	C1KBI9_9BIVA	Rab7-like protein OS=*Pinctada martensi*	0	0.19	0.12	0.14
	E6Y2Z7_HALDV	Splicing factor arginine/serine-rich 4 OS=*Haliotis diversicolor supertexta*	0	0.30	0.25	0.11
	B6RB35_HALDI	Transgelin OS=*Haliotis discus discus*	0	0.24	0.24	0.08
	Q7YZR1_HALAI	Tropomyosin 4 (Fragment) OS=*Haliotis asinina*	0	0.22	0.19	0.31
	B6RB46_HALDI	Troponin OS=*Haliotis discus discus*	0	0.38	0.35	0.33
	Q6TQF6_DORPE	Visual G-protein beta subunit OS=*Loligo pealeii*	0	0.18	0.06	0.20
Down-regulated						
	B6RB43_HALDI	ATP synthase_ H+ transporting_ mitochondrial F1 complex_ o subunit OS=*Haliotis discus discus*	0	-0.27	-0.38	0.09
	Q9NL91_OCTVU	G protein alpha subunit s class OS=*Octopus vulgaris*	0	-0.30	-0.36	-0.47
	GNAS_LYMST	Guanine nucleotide-binding protein G(s) subunit alpha OS=*Lymnaea stagnalis*	0	-0.31	-0.36	-0.47
	Q8I0U4_HALTU	H2 protein (Fragment) OS=*Haliotis tuberculata*	0	-0.44	-0.48	-0.06
	C7FEG7_HALDV	Hemocyanin isoform 1 (Fragment) OS=*Haliotis diversicolor*	0	-0.57	-0.54	-0.31
	Q6KC56_MEGCR	Keyhole limpet hemocyanin1 OS=*Megathura crenulata*	0	-0.23	-0.29	0.09
	E5RSV6_HALDI	Paramyosin OS=*Haliotis discus discus*	0	-0.54	-0.29	-0.54
	H8XWJ5_9MOLL	Prohibitin OS=*Octopus tankahkeei*	0	-0.30	-0.10	-0.36

3.3.2.1. Gene ontology analysis

To provide functional classification of the identified proteins, Blast2GO analysis was carried out to assign GO terms to the identified proteins. However, these GO terms were very general and therefore in order to better interpret this data, over-simplified or redundant GO terms were removed using the tool REViGO (Supek et al., 2011). However, REViGO is only effective on small sets of GO terms, and so proteins indicated by brackets on the dendrogram in Figure 3-5 were analysed. The REViGO outputs were plotted as treemaps (Figure 3-6).

Proteins which were up-gregulated at all sampled time points were associated with GO terms for biolgical processes related to the stress response (for example 'response to endoplasmic reticulum stress', 'response to heat' and 'response to oxidative stress'), as well as G protein signalling related to the stress response ('small GTPase mediated signal transduction) (Figure 3-6-A). In addition, metabolic processes appear to have been affected ('lipoprotein metabolism', 'energy reserve metabolism'). The presence of the supergroup 'proteasome core complex assembly' suggests changes in protein degradation rates by the proteasome. DNA, RNA and protein synthesis may also be affected by heat stress as several superclusters refer to the regulation of transcription or translation. These biological processes are facilitated by the molecular functions indicated in Figure 3-6-B, which identify a number of GO term clusters associated with binding (for example 'RNA binding, 'nucleotide binding' and 'protein binding involved in protein folding').

Proteins which were down-regulated at all sampled time points were associated with GO terms for biological processes related to similar stress response superclusters (for example 'response to stress') (Figure 3-6-C). These GO term clusters also indicate that DNA biosynthesis may be negatively impacted by heat shock, as suggested by the supercluster related to 'DNA biosynthesis'. The supercluster referring to 'female pregnancy' contains GO clusters pertaining to development, morphogenesis and sensory perception. Once again, the GO superclusters for the molecular function GO terms associated with these proteins primarily refer to binding, such as 'protein complex binding' and 'G protein beta subunit binding'. A supercluster containing GO clusters associated with the actin cytoskeleton and GTP/ATPase activity is also present.


A
cell proliferation
face development
nucleus accumbens development
instar larval development
gene expression
lipoprotein metabolism
protein maturation by protein folding
ethanol oxidation
glycerol ether metabolism
energy reserve metabolism
endocytosis
protein transport
potassium ion import
S phase of mitotic cell cycle
olfactory pit development
asymmetric neuroblast division
oogenesis
post-translational protein modification
RNA methylation
protein folding
lipid metabolism
glutathione metabolism
regulation of cellular amino acid metabolism
apical protein localization
sodium ion export
endosome to lysosome transport
cell proliferation
lipoprotein metabolism
energy reserve metabolism
apical protein localization
apoptotic process
skeletal muscle contraction
cellular senescence
adult locomotory behavior
platelet activation
anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolism
protein polyubiquitination
translational elongation
glucose 6-phosphate metabolism
ribose phosphate biosynthesis
mRNA 3□-end processing
termination of RNA polymerase II transcription
interspecies interaction between organisms
response to oxidative stress
response to heat
response to drug
response to endoplasmic reticulum stress
dephosphorylation
mRNA metabolism
viral transcription
response to wounding
glucocorticoid receptor signaling pathway
phospholipase C-activating G-protein coupled receptor signaling pathway
G-protein coupled acetylcholine receptor signaling pathway
regulation of growth rate
negative regulation of biological process
positive regulation of DNA replication
dephosphorylation
protein N-linked glycosylation via asparagine
locomotion
carbohydrate metabolism
photorespiration
response to heat
regulation of growth rate
termination of RNA polymerase II transcription
interspecies interaction between organisms
proteasome core complex assembly
response to nicotine
adenylate cyclase-activating dopamine receptor signaling pathway
ionotropic glutamate receptor signaling pathway
small GTPase mediated signal transduction
positive regulation of molecular function
protein stabilization
cellular calcium ion homeostasis
proteasome core complex assembly
Golgi organization
phagosome maturation
cortical actin cytoskeleton organization
pentose biosynthesis
retinal metabolism
retinal metabolism


B
complement component C1q binding
protein binding involved in protein folding
troponin T binding
troponin C binding
threonine-type endopeptidase activity
4-nitrophenylphosphatase activity
mRNA binding
RNA binding
sodium ion binding
calcium ion binding
RS domain binding
calcium-dependent protein binding
calmodulin binding
protein domain specific binding
4-nitrophenylphosphatase activity
GTPase activity
translation initiation factor activity
DNA binding
iron ion binding
type 1 angiotensin receptor binding
androgen receptor binding
ATP-dependent RNA helicase activity
structural constituent of cytoskeleton
protein disulfide isomerase activity
calcium channel regulator activity
thyroid hormone binding
phosphatidylinositol 3-kinase binding
integrin binding
chaperone binding
GTPase binding
retinal binding
GDP binding
structural constituent of eye lens
glucose-6-phosphate dehydrogenase activity
protein disulfide oxidoreductase activity
retinal dehydrogenase activity
ADP binding
signal transducer activity
hormone binding
peptide binding
carbohydrate binding
procollagen-proline 4-dioxygenase activity
3-chloroallyl aldehyde dehydrogenase activity
aldehyde dehydrogenase [NAD(P)+] activity
NAD+ binding
nucleotide binding
electron carrier activity
glucose binding
protein binding


C
response to salt stress
response to drug
response to caffeine
response to auxin stimulus
response to amphetamine
response to trehalose stimulus
response to nitric oxide
response to endoplasmic reticulum stress
cellular response to prostaglandin E stimulus
progesterone receptor signaling pathway
response to misfolded protein
response to ethanol
adenylate cyclase-activating dopamine receptor signaling pathway
cellular response to interleukin-6
cellular response to catecholamine stimulus
female pregnancy
post-embryonic body morphogenesis
embryonic hindlimb morphogenesis
organ regeneration
mammary gland alveolus development
negative regulation of blood pressure
sensory perception of sweet taste
chaeta development
lateral root development
skin development
visual behavior
sensory perception of smell
ovarian follicle atresia
embryo sac egg cell differentiation
adult heart development
endochondral ossification
blood coagulation
positive regulation of osteoclast differentiation
DNA biosynthesis
ubiquitin-dependent protein catabolism
DNA replication
RNA methylation
nucleotide biosynthesis
regulation of cellular process
negative regulation of transcription by competitive promoter binding
negative regulation of apoptotic process
GTP catabolism
regulation of transcription from RNA polymerase II promoter
negative regulation of cell growth
transport
oxygen transport
cell division
water transport
mitochondrial ATP synthesis coupled proton transport
intracellular transport
genetic imprinting
activation of adenylate cyclase activity
regulation of insulin secretion
tissue homeostasis
calcium ion homeostasis
positive regulation of sodium ion transport
negative regulation of renal sodium excretion
proteasome core complex assembly
histone deacetylation
muscle filament sliding
karyogamy
myofibril assembly
atrial cardiac muscle tissue morphogenesis
energy reserve metabolism
metabolism
response to stimulus
respiratory electron transport chain
photorespiration


D
G-protein beta-subunit binding
G-protein alpha-subunit binding
alpha-tubulin binding
actin-dependent ATPase activity
microfilament motor activity
metal ion binding
GTP binding
GTP binding
oxidoreductase activity
oxidoreductase activity
ionotropic glutamate receptor binding
G-protein beta/gamma-subunit complex binding
toxin receptor binding
actin-dependent ATPase activity
sequence-specific DNA binding RNA polymerase II transcription factor activity
adenylate cyclase activity
adenylate cyclase activity
steroid binding
steroid binding
histone deacetylase binding
G-protein beta-subunit binding
mu-type opioid receptor binding
corticotropin-releasing hormone receptor 1 binding
GTPase activity
ATPase activity
structural constituent of muscle
structural constituent of muscle
insulin-like growth factor receptor binding
beta-2 adrenergic receptor binding
protein domain specific binding
protein complex binding
transmembrane transporter activity
oxygen transporter activity
oxygen transporter activity
signal transducer activity
signal transducer activity
drug binding
drug binding
transcription regulatory region DNA binding
transcription regulatory region DNA binding

Figure 3-6: Treemaps generated by REViGO using GO terms associated with proteins which responded to heat stress. These proteins were identified by statistical seriation of the protein expression profiles determined by iTRAQ analysis. GO terms for proteins which were up-regulated by heat stress are indicated in treemaps for A) biological process and B) molecular function. GO terms for proteins which were down-regulated by heat stress are indicated in treemaps for C) biological process and D) molecular function. Each smaller rectangle is a representative of a cluster of GO terms associated with these proteins. These are joined into 'superclusters', defined by different colours, which are loosely related under the representative term in large text. The size of the rectangles is indicative of the uniqueness of the GO term in the underlying database, in this case the UniProt repository. Not all of the terms are reported in each treemap due to constraints on space.

3.3.3. iTRAQ validation by western blot hybridisation

In order to validate the iTRAQ data, the expression profile of selected proteins during exposure to heat shock was determined using a complementary molecular technique. Only 3 proteins out of the 11 significantly regulated proteins (Table 3-2) had more than 60% homology between the molluscan protein sequence and the proteins used to produce commercially available polyclonal antibodies. These antibodies were therefore suitable for use in expression analysis by western blot hybridisation. These proteins were heat shock protein 70 (Hsp70), glucose-6-phosphate 1-dehydrogenase (G6P1D) and glutathione-S-transferase (GSTO1). During optimisation of western blotting conditions for each antibody, the molecular weight of the detected bands was determined (Table 3-4). For each protein, the expected protein size (based on amino acid sequence) was obtained from the UniProt database and compared to the measured band size determined by western blot. The antibody for Hsp70 gave a single, clear band with the correct measured band size (approximately 70 kDa). The western blot probed with G6P1D antibody had multiple bands, but the most reproducibly detected band was visible at 41.5 kDa, which is approximately 10 kDa smaller than expected. GSTO1 had a single visible band; however it was measured at 88.4 kDa whereas the predicted protein size for GSTO1 was 27.4 kDa.

Table 3-4: The observed (measured) band size obtained for each of the proteins selected for validation of iTRAQ data by western blot analysis, compared to the expected size based on the predicted molecular weight of the protein according to the UniProt entry for that protein.

Protein	Expected band size (Da)	Measured band size (Da)
Hsp70	71646	72400
G6P1D	52919	41500
GSTO1	27366	88400

The Hsp70 western blot (Figure 3-7) indicated no statistically significant difference in expression of this protein between the control and heat shocked samples following one way ANOVA, however the amount of Hsp70 in the haemocytes sampled at 30 minutes appeared to have increased compared to the control, followed by elevated expression for the duration of the heat shock.

Western blot analysis of G6P1D detected a statistically significant ($p=0.0265$) increase in expression of the protein after 30 minutes of heat shock exposure (Figure 3-8), followed by a decrease back to approximately basal levels at 90 minutes. A doublet band was detectable in some of the samples and biological replicates. The doublet was visible in 3 out of 4 biological replicates in the 90 minute sample (panel A, B and C). It was also faintly visible in one of the 30 minute biological repeats (panel B), and in at least one control sample (panel B and faintly in panel C). The 180 minute sample in panel C also has a faint doublet band. The upper band of the doublet appears to be at a marginally higher molecular weight than the band detected in the control sample.

Western blot analysis of GSTO1 expression did not detect a statistically significant difference in expression of this protein over the course of the heat shock experiment, although a rapid increase in the amount of GSTO1 was detected at 30 minutes (Figure 3-9). By 180 minutes, the amount of GSTO1 had returned to approximately basal levels. There was a large discrepancy between the predicted and observed molecular weights of this protein, as the measured band was nearly 4 times larger than the UniProt predicted size of GSTO1.

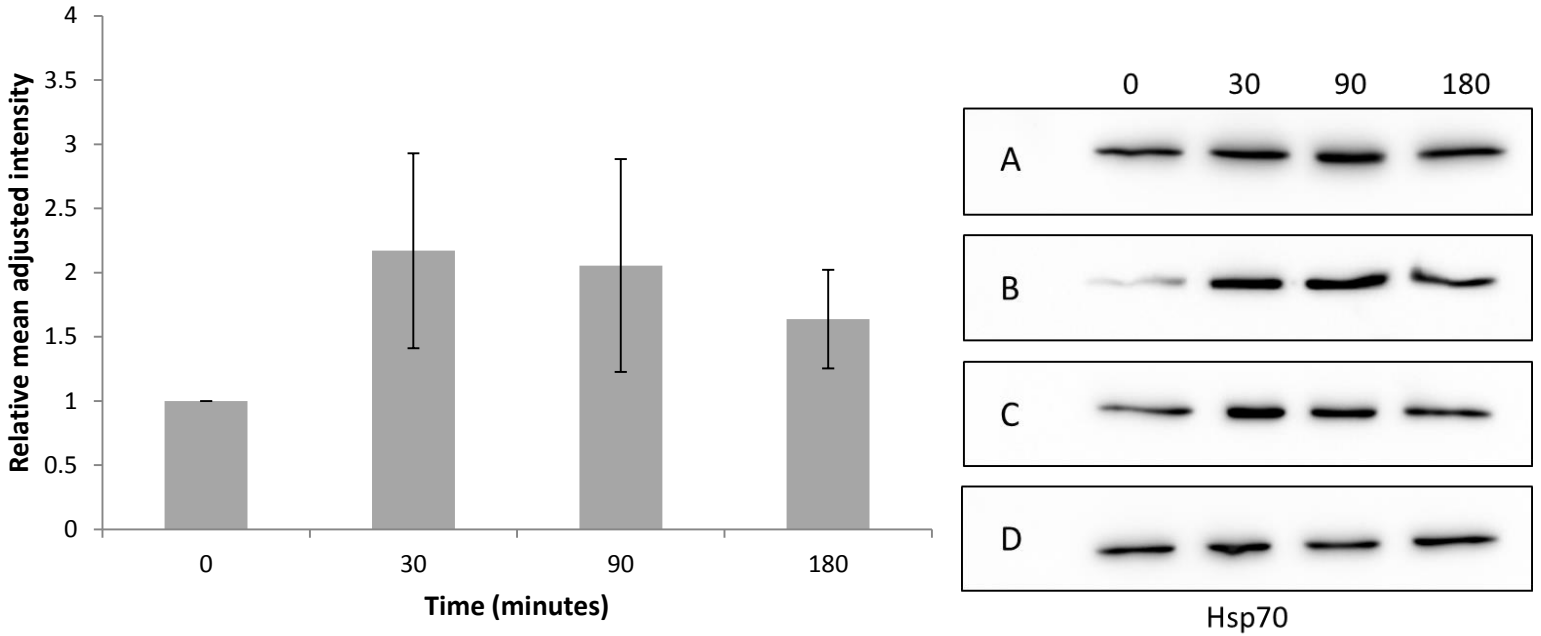


Figure 3-7: Western blot analysis of Hsp 70 expression in *H. midae* haemocytes sampled during a timed heat shock experiment. The bar graph shows the average volume intensity data obtained using ImageLab Software (Bio-Rad) relative to the control (0) sample, and the inlaid images are the chemiluminescent western blot signals detected with a ChemiDoc XRS (Bio-Rad). A, B, C and D are images obtained from four biological replicates (n=4). Error bars represent standard error of the mean.

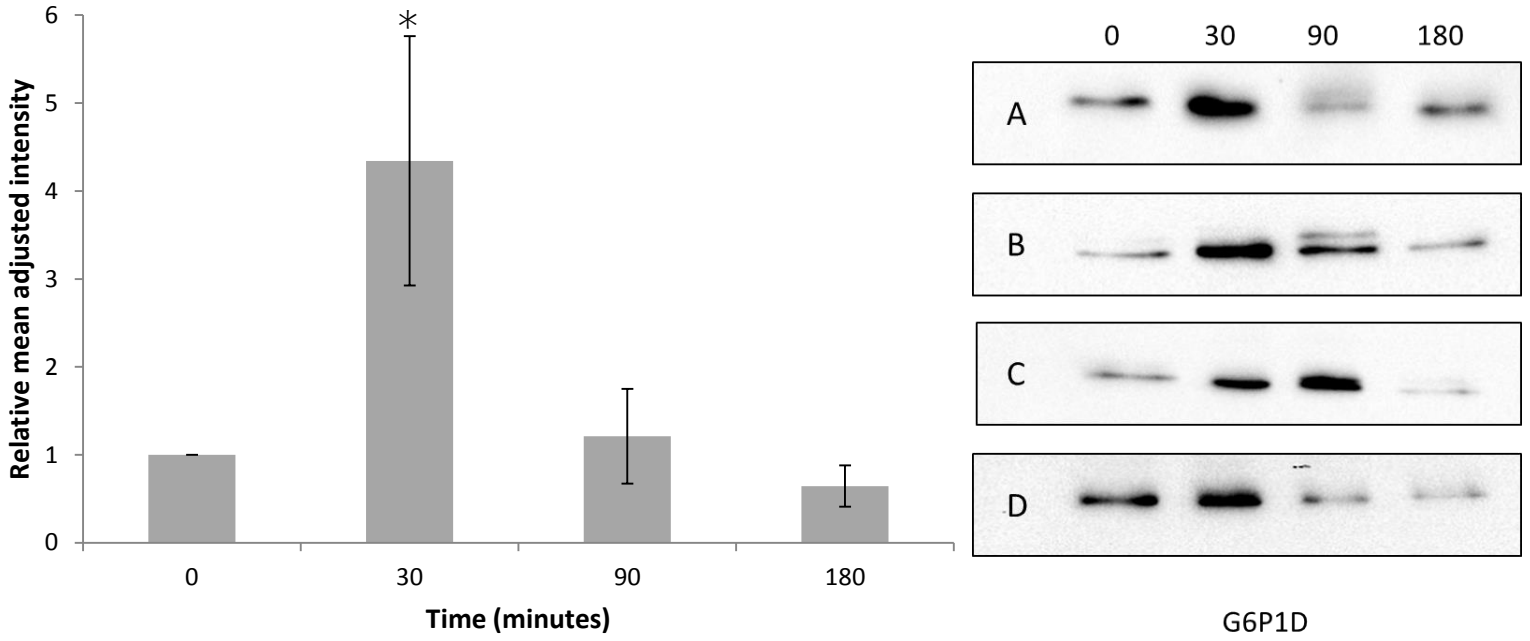


Figure 3-8: Western blot analysis of G6P1D expression in *H. midae* haemocytes sampled during a timed heat shock experiment. The bar graph shows the average volume intensity data obtained using ImageLab Software (Bio-Rad) relative to the control (0) sample, and the inlaid images are the chemiluminescent western blot signals detected with a ChemiDoc XRS (Bio-Rad). A, B, C and D are images obtained from four biological replicates (n=4). Error bars represent standard error of the mean. Volume analysis in lanes showing doublet bands was performed on the lower band, with the same molecular weight as the control (0). Statistical significance was determined by one way ANOVA and a p=0.0265 is indicated on the graph by an asterisk (*).

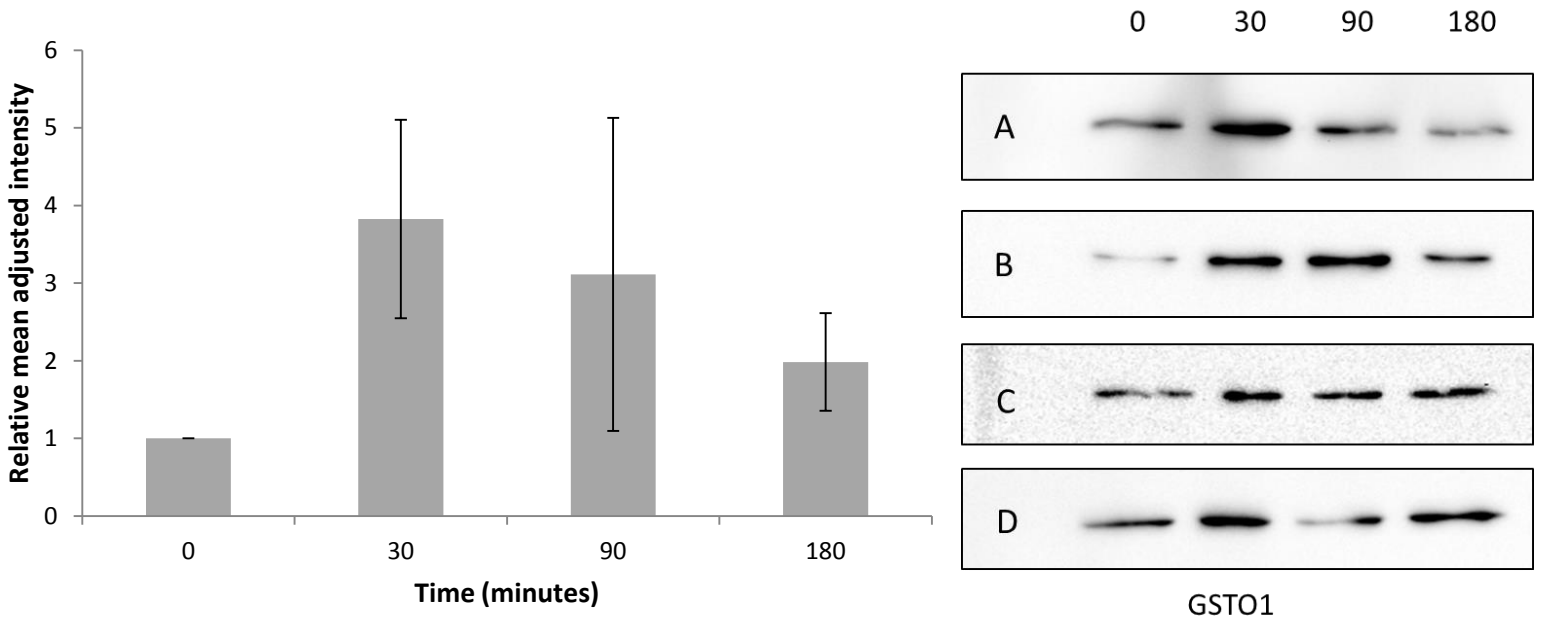


Figure 3-9: Western blot analysis of GSTO1 expression in *H. midae* haemocytes sampled during a timed heat shock experiment. The bar graph shows the average volume intensity data obtained using ImageLab Software (Bio-Rad) relative to the control (0) sample, and the inlaid images are the chemiluminescent western blot signals detected with a ChemiDoc XRS (Bio-Rad). A, B, C and D are images obtained from four biological replicates (n=4). Error bars represent standard error of the mean.

3.4. Discussion

Developing our understanding of the abalone immune system is fundamental to the improvement of abalone farming practices. To date there have been few molecular studies investigating the abalone stress response caused by rapidly increased water temperature, despite reported fluctuations in abalone tank water temperature (Vosloo and Vosloo, 2010). Proteomic technology is particularly valuable in assessing the effects of environmental changes because it provides a 'snapshot' of the cellular response to stress (Tomanek, 2011). In the present study, the high throughput proteomic technique iTRAQ was used to identify differentially expressed proteins in the haemocytes of the South African abalone, *H. midae*, during exposure to heat shock. As with any proteomic study, this methodology required the use of a database of amino acid sequences in order to facilitate peptide spectrum matching to identify proteins.

In the present study, the database chosen comprised of all molluscan protein sequences in the UniProt database, including uncurated trEMBL sequences. This database choice ensured the maximum likelihood that proteins from *H. midae* will match to at least one protein sequence in the database. In order to corroborate the database search results, *de novo* peptide sequencing was performed and the *de novo* peptides were aligned with the PSMs from the database search. Once the protein identities were determined, the data was substantiated by combining two independent iTRAQ experiments, ensuring that only proteins with good quality peptides detected in both independent experiments were considered further. In this manner, the quality of the PSMs and the number of protein identities were maximised. The combined list comprised 116 protein entries with associated quantitative data. The expression profiles of these proteins during exposure to heat shock were then further investigated to identify putative stress response proteins in *H. midae*.

3.4.1. Proteins with statistically significant differences in expression during heat shock

Of the 116 identified proteins, a total of 11 had a statistically significant response to heat stress – eight responded in all four biological replicates and three responded in three of the

four biological replicates. Of these, several proteins had a putative stress response function of particular interest.

The protein P30842|CROM_OMMSL is an omega crystallin found in the squid eye, where it is responsible for optical properties of the lens (Horwitz et al., 2006). In molluscs such as the scallop, omega crystallin is homologous to aldehyde dehydrogenase in the cytoplasm (ALDH1A1) and mitochondria (ALDH2) of humans (Zinovieva et al., 1993). Aldehyde dehydrogenases (ALDH) are enzymes that catalyse the oxidation of aldehydes, which may be produced as a result of oxidative stress to the lipid membranes of the cells (Ellis, 2007). A sequence alignment revealed that there is a 53.4% identity between the squid omega-crystallin protein and the only molluscan non-cephalopod aldehyde dehydrogenase in the database (Uniprot identifier K1QNT7_CRAGI, *Crassostrea gigas*). This homology could explain why a squid eye lens protein was identified in the haemocytes of *H. midae,* especially as the identity in this case is based on three unique peptides representing only 8% coverage of the protein in question. For the purposes of this study, it was henceforth assumed that this protein therefore is an ALDH, which makes better sense in context to its origin and putative function.

The ALDH protein has yet to be properly characterised in the abalone. Although lacking full characterisation, differential expression of ALDH has been reported in the mollusca. Proteomic investigations have described the response of ALDH to the toxicity of both bisphenol A and diallyl phthalate in the abalone *Haliotis diversicolor supertexta* (Zhou et al., 2010), as well as cold acclimation in the warm adapted mussel *Mytilus galloprovincialis* and cold adapted *Mytilus trossulus* (Fields et al., 2012b). Given that so few proteomic investigations have been reported in the mollusca, and this protein was identified in several of them, ALDH clearly plays a role in the molluscan stress response where it may respond to oxidative stress. The large increase in expression of this protein detected by iTRAQ in this study may therefore confirm that exposure to acute heat stress leads to oxidative stress in *H. midae*. Indeed, another important oxidative stress response protein, glutathione S-transferase (GST), also responded to heat shock with statistical significance.

The free radical scavenging activity of GST has been described in clams (Hoarau et al., 2004) and oysters (Boutet et al., 2004), while two omega class GSTs have been described and partially characterised in disc abalone (*Haliotis discus discus*) (Wan et al., 2009). Abalone

omega GSTs appear to respond to heat stress, heavy metal toxicity and endocrine disrupting chemical exposure. Wan et al. (2009) determined that each of these omega class GSTs behaved differently when exposed to acute thermal stress, and consequently, it is currently unclear how omega GSTs are regulated in abalone. The dramatic elevation of omega GST transcription detected in abalone subjected to heat shock by Wan et al. (2009) is similar to the increase in GST isoform O expression detected by iTRAQ in *H. midae* haemocytes after 30 minutes of heat exposure. The cellular oxidative stress response is a well-documented and vital stress response mechanism, which has yet to be fully characterised in the abalone.

Glucose 6 phosphate 1 dehydrogenase (G6P1D) is the rate-determining enzyme in the pentose phosphate pathway that provides energy in the form of NADPH to cells (Lama et al., 2013). NADPH is involved in important processes such as amino acid and nucleic acid synthesis, and the removal of oxidised glutathione (GSSG) by reducing it to GSH (da Silva *et. al.*, 2005). The pentose phosphate pathway (PPP) appears to be inhibited by oxidative stress in the marine snail *Littoria* (Lama et al., 2013) and clams (Wu et al., 2013). In the present study, the amount of G6P1D in *H. midae* haemocytes increased after 180 minutes of exposure to heat stress. Lama et al. (2013) reported that G6P1D was phosphorylated under normal conditions, but under anoxic stress conditions underwent dephosphorylation and became more active as a result, potentially to facilitate the flow of NADPH into the oxidative stress response. Although it is possible to detect post-translational modifications (PTMs) using iTRAQ and mass spectrometry, enrichment for specific modifications is usually required to prevent their loss during analysis (Mann and Jensen, 2003). As phosphorylation was not enriched for in this study, the phosphorylation state of G6P1D could not be determined by iTRAQ.

The heat shock protein (Hsp) family is a well characterised family of stress response proteins that act as chaperones in order to restore homeostasis to the cell during exposure to a diverse range of stressors (Feder and Hofmann, 1999). It consists of several members with a molecular weight around 70kDa, some of which are inducible by stress (hsp70) and some of which are constitutively expressed (Wang et al., 2013), such as heat shock cognate (hsc) 70. Based on the limited molluscan sequences available for both Hsp70 and Hsc70 (Boutet et al., 2003a, 2003b), there is high homology between the nucleic acid sequence of Hsp70 and Hsc70 (97% identity in *Crassostrea gigas*, and 98% in *Ostrea edulis*). Both possess

the three characteristic amino acid sequence motifs of the Hsp family (Boutet et al., 2003b), making it potentially difficult to differentiate between the constitutive and inducible forms of this protein. The main difference between these proteins is that the promoter region of hsp70 has heat shock elements (HSEs) which are stress inducible, and therefore this protein will be up-regulated in response to stress. Hsp70 has been characterised in the mollusca, including several members of the *Haliotis* genus (Cheng et al., 2007; Farcy et al., 2007). The expression of Hsp70 mRNA increases during bacterial challenge with the pathogen *Vibrio anguillarum* in *Haliotis discus hannai* (Cheng et al., 2007), and after 30 minutes of exposure to acute heat stress in *Haliotis tuberculata* (Farcy et al., 2007). Under the heat shock conditions of this study, the identified protein showed an early increase in expression at 30 minutes, followed by an apparent decrease in expression later in the heat shock time course. The stress inducible nature of this protein leads to the conclusion that it is an Hsp70, rather than the constitutively expressed homologue. Brun et al. (2008) observed a more long-term Hsp70 protein response in two species of scallop, *Argopecten irradians* and *Placopecten magellanicus*, where it was significantly regulated 21 days after a 3 hour heat shock. However, the induction temperature and duration of Hsp70 expression is entirely dependent on the thermal range of the organism in question (Feder and Hofmann, 1999), possibly explaining the rapid, short-lived response of the *H. midae* Hsp70 detected in this study.

Ribosomal proteins are part of the essential DNA translation machinery of the cell, although there is recent evidence that suggests that ribosomes perform a variety of 'extra-ribosomal' activities by altering the trans-activation of regulatory proteins (Lindström, 2009). These include proteins involved in apoptosis and DNA repair regulation. The ribosomal 40S subunit S8 (B3TK66_HALDV) was significantly regulated in this study, although its significance in the abalone stress response is unknown. Ribosomal proteins made up a large proportion of the proteins responding to heavy metal exposure detected by Muralidharan et al. (2012) in the Sydney rock oyster, although they did not determine a function. Silvestre et al. (2010) identified temperature dependent differential expression of ribosomal subunit 40S in sturgeon larva using 2D PAGE, and suggested that temperature affects protein synthesis. In the present study, iTRAQ also detected differential expression of proteins involved in RNA metabolism - PL-10 like protein (C7EAA2_HALAI), which is a member of the DEAD-box family of RNA helicases (Rosner and Rinkevich, 2007), and small nuclear ribonucleoprotein D2-like protein (Q8MUE3_9BIVA). Their

significance in the abalone stress response is unknown and to date the effect of stress on nucleic acid or protein synthesis in abalone has not been investigated.

3.4.2. Determination of heat-responding biological processes and molecular functions by gene ontology analysis

Once the proteins with statistically significant responses to heat stress were identified, the data was then subjected to statistical seriation in order to identify groups of proteins with similar responses to heat stress. Proteins were ordered in a dendrogram according to their expression patterns. The protein groups that responded with the highest overall up- or down-regulation during heat stress according to the dendrogram were selected for gene ontology (GO) analysis in order to determine their functional significance. This information assisted in identifying biochemical pathways in abalone haemocytes that are affected by exposure to thermal stress. Thus, three interconnected general cellular responses that appear to be affected by heat stress in *H. midae* have been identified using this information: signalling, response to stress and metabolic regulation. The proteins identified by iTRAQ that contribute to each of these functions will be discussed below. As many proteins have multifaceted functions, the most commonly reported function for each protein will be discussed.

3.4.2.1. Proteins involved in cell signalling and regulation

It is critically important for cells to transduce information regarding their environment, as the ability to respond to environmental changes determines a cell's ability to survive. This is arguably the most important feature of the cells comprising the immune system, as many immune functions are predicated on the ability of the self to detect and respond to non-self or potentially harmful environmental conditions. This communication usually depends on the presence of cell surface receptors which are capable of receiving and decoding chemical signals from the extracellular space. The molluscan innate immune system appears to be effected via signalling receptors on the surface of haemocytes (Humphries and Yoshino, 2003). Despite this, relatively little is understood about the mechanisms by which molluscan haemocytes determine or initiate the need for a stress response (Lacoste et al., 2001).

Several GO clusters related to calcium dependent signalling were identified in this analysis, associated with both up- and down-regulated proteins. Calcium signalling is both a trigger and a controller of vital cellular processes, and effective Ca^{2+} signalling therefore requires that calcium ions be available in a regulated manner from sources maintained both inside and outside the cell (Berridge et al., 2000). The system of calcium release and control occurs via voltage, receptor, or store operated channels (Berridge et al., 2000); however, different cell types can have distinct, unique calcium triggering mechanisms - some of which are not well understood. In order to maintain control, most of the released calcium is immediately bound to buffering molecules such as calbindin-D or calreticulin, where it is held until it is required for downstream signalling.

Differential expression of several G protein subunits implicates another important signal transduction pathway in the abalone heat stress response. GO terms relating to G protein-regulated cell signalling are associated with proteins in both the up- and down-regulated groups. Both visual G protein B subunit (Q6TQF6_LOLPE) and guanine nucleotide-binding protein subunit B (GBB_LOLFO) are signal transduction proteins in the guanine nucleotide binding protein (G protein) family, and function as molecular 'switches' capable of eliciting a response to extracellular stimuli (Ridley, 2001). G proteins exist in mammalian cells in alpha, beta and gamma subunits (Ridley, 2001). Each subunit has several isoforms and isoform splice variants which combine into hundreds of different stimulus-specific combinations (Hepler and Gilman, 1992; Vetter and Wittinghofer, 2001). A differentially regulated G protein B subunit which responded to heat shock was identified in Chapter 2 using 2D PAGE. In the heat map, a G protein alpha subunit s-class (Q9NL91_OCTVU) grouped with the down-regulated proteins, confirming the expression pattern determined by 2D PAGE. The rho proteins are yet another member of the GTP-ase family, falling under the superfamily category of the ras-like proteins. These small signalling proteins are responsible for, in particular, movement of the actin cytoskeleton through regulation of stress fibres and focal adhesions to the extracellular matrix (Ridley, 2001).

Different G proteins are capable of regulating very specific cellular mechanisms, ranging from translation to exocytosis (Gasper et al., 2009). G proteins also operate in an ATP-dependent manner and would require support in the form of increased ATP production in order to facilitate cellular signalling, which is indicated by the differential regulation of ATP

producing proteins. Although cell signalling is evidently important in the abalone stress response, there is currently a lack of information describing the mechanisms by which signalling contributes to the stress response of abalone.

3.4.2.1. Proteins involved in the stress response

The response to stress is a diverse and complex system in any cell. Different stressors activate or inhibit a wide variety of cellular stress response mechanisms. The following discussion will focus on stress response proteins for two distinct stress response biochemical pathways identified by GO analysis – the oxidative stress response and the unfolded protein response.

The oxidative stress response is extremely important in returning the cell to homeostasis following excessive production of harmful reactive oxygen species. Although exposure to increased temperature induces oxidative stress in a number of aquatic animals, the mechanisms underlying this connection remain unclear (Lushchak, 2011). There is evidence that oxidative stress induces the expression of some Hsps (Larkindale and Knight, 2002). Indeed, the link between heat stress and the oxidative stress response has been more extensively studied in organisms that are unable to move away from stressful environmental conditions, where the ability to acquire thermotolerance is vital to survival, such as in plants (Larkindale and Knight, 2002; Mohanty et al., 2012). Although abalone are motile, a study by Tarr (1995) determined that 81.5% of tagged abalone on the South African coast remained in the exact same position after three years of observation. Under aquaculture conditions abalone are also unable to escape potentially stressful environmental changes. Differential antioxidant activity of catalase, superoxide dismutase and glutathione peroxidase has been recorded in *H. midae* at various oxygen concentrations and temperatures (Vosloo et al., 2013a), suggesting that these stress responses are also linked in abalone.

In this study, the detection of several antioxidant proteins which appear to be regulated during acute heat exposure further supports this link. The up-regulated group of proteins contains previously discussed anti-oxidant proteins with statistically significant regulation G6P1D (Q0KHB8_CRAGI) and ALDH (CROM_OMMSL). Another statistically significant protein in this group is haemocyanin, an oxygen carrier molecule that is found

in both the mollusca and arthropods, where it performs a role analogous to haemoglobin in vertebrates and most insects (Markl, 2013). Haemocyanin uses Cu^{2+} to bind free O^{2-} in the haemolymph, which is what gives abalone blood its characteristic blue colour (Swerdlow et al., 1996). Due to its association with oxygen, molluscan haemocyanin also performs regulatory and possibly stress response functions, and is known to have phenoloxidase and catalase activity (Siddiqui et al., 2006; Vosloo et al., 2013a).

The connection between the cytoskeleton and oxidative stress was discussed in Chapter 2. Actin (C4NY64_CRAGI) exhibited increased relative expression at each of the time points sampled, and given the prevalence of cytoskeletal proteins identified by 2D PAGE (Chapter 2), it was of particular interest that several cytoskeletal proteins were also identified by iTRAQ. Transgelin (B6RB35_HALDI) has a similar expression pattern to actin, and although it has not been completely functionally characterised, it is thought to be sensitive to changes in cell shape, capable of cross-linking actin fibres and acting as an actin stress fibre-associated protein (Assinder et al., 2009). Although there is a current lack of information describing the effects of stress on the abalone cytoskeleton, a number of proteins identified in this study may represent a good starting point for further characterisation of the abalone haemocyte structural protein network.

According to this analysis, tropomyosin and troponin have similar heat-induced expression profiles to that of actin and transgelin. In vertebrates, troponin is found in the skeletal and cardiac musculature (Hooper and Thuma, 2005). In molluscs, troponin is an important cytoskeletal regulatory protein that responds to bacterial and viral stimuli in clams (Chen et al., 2011) and *H. tuberculata* (Travers et al., 2010). Tropomyosin is associated with troponin in the groove between actin filaments, where it prevents the activity of myosin and consequently, cytoskeletal contraction (Gunning et al., 2005). Myosin is an ATP-dependent motor protein that associates with the actin cytoskeleton to regulate cellular movement (Hooper and Thuma, 2005). Troponin is under the control of intracellular calcium ions which bind to troponin in a concentration dependent manner to cause structural changes in troponin, releasing tropomyosin and allowing myosin to act on the actin fibres (Hitchcock, 1977). Together with the results of the 2D PAGE analysis, the presence of these proteins and their apparent regulation during exposure to heat stress corroborates the hypothesis that cytoskeletal remodelling is an important consequence of hyperthermia in abalone. However, it remains unclear whether differential expression of

cytoskeletal elements is a direct or regulatory response to heat stress, or whether the cytoskeleton is being destabilised or degraded by the stress.

Molecular chaperones are a vital component of the cellular response to stress. Chaperones, including Hsp70, are responsible for ensuring the correct folding of proteins in the cell. During exposure to harmful environmental conditions, proteins can become denatured and aggregate in subcellular compartments such as the endoplasmic reticulum (ER), where they may become toxic (Hetz, 2012). Interestingly, the ER is also the largest store of intracellular Ca^{2+} in eukaryotic cells, and plays a central role in the regulation of Ca^{2+} signalling (Lam and Galione, 2013). One of the up-regulated proteins identified in this study is the ER resident molecular chaperone calreticulin (A5LGG9_CRAGI), which is also known to play a significant role in maintaining calcium homeostasis (Michalak et al., 1999) and may therefore play an important role in the signalling events that occur as the cells attempt to return to homeostasis. Calreticulin ensures the correct folding of proteins in the ER subcellular component, which can become overwhelmed by misfolded proteins under stress conditions, and can trigger the stress response pathway known as the unfolded protein response (UPR) (Schröder, 2008). Another up-regulated protein, protein disulphide isomerase (B6RB63_HALDI), is integrally linked with protein folding in the ER, where it catalyses the formation of disulphide bonds between amino acids during protein folding, and according to this analysis, is also up-regulated during exposure to acute heat stress in *H. midae*. Protein disulphide isomerase (PDI) also functions as an antioxidant, and has been implicated in the stress response mechanism of mussels (McDonagh and Sheehan, 2007) during exposure to oxidative stress. Neither calreticulin nor PDI have been characterised in abalone, despite their importance in the protein folding cellular machinery.

One of the outcomes of persistent ER stress is the activation of the proteasome – the major cellular structure responsible for degradation of terminally misfolded proteins (Ahner et al., 2004). Up-regulation of the alpha subunit of the proteasome (B6RB41_HALDI) suggests that acute heat stress leads to an increased demand for proteasome-assisted degradation of misfolded proteins. Differences in proteasome subunit expression during stress exposure were detected in mussels by Tomanek and Zuzow (2010), who associated variable levels of proteasome production during stress with the overall thermal tolerance of a species. This parallels the findings of Gracey *et. al.*, (2008)

who reported changes in transcriptional regulation of genes involved in DNA synthesis and cell division in mussels during exposure to abiotic stresses associated with tidal changes.

3.4.2.2. Proteins involved in energy production and metabolic regulation

Thermal stress causes the perturbation of cellular processes, and consequently, it is important that the stressed cell can meet its energy needs in order to return to homeostasis. Most proteins that have already been discussed also have a putative metabolic function according to their associated GO terms. In addition to these, a putative Na^+/K^+ -ATPase alpha subunit (UniProt identifier F8V2T8_9MOLL), which is a metabolic protein which uses ATP to move Na+ and K+ across the cellular membrane via a proton gradient, is up-regulated by heat stress in *H. midae*. As ion concentrations contribute to and regulate homeostasis in the cell (Kong et al., 2007), it is extremely important for a cell under stress to maintain a proper balance of abundant ions such as Na^+ and K^+. This protein has been previously characterised in the mud crab *Scylla serrate* exposed to cold stress, where its activity initially decreased and was then followed by an increase after 6 hours of exposure (Kong et al., 2007). This protein currently lacks proper characterisation in the mollusca, especially given the role of ion signalling in stress response mechanisms (Berridge et al., 2003). Expression of another protein involved in ATP synthesis, the ATP synthase H^+ transporting mitochondrial complex O subunit (ATP5O) (B6RB43_HALDI), was statistically significant and found to group with the down-regulated proteins. This protein appears to be a mitochondrion resident ATP synthase, which makes use of proton energy gradients across mitochondrial membranes to synthesise ATP from ADP. Although we currently lack sufficient experimental data which quantify the metabolic costs of short- and long-term exposure to heat shock, especially in marine organisms which are likely to experience increasing temperature as the effects of global climate change become more severe, the heat shock response is known to be energetically costly (Tomanek, 2010). Since *H. midae* has been shown to favour ATP production at high temperatures to ensure survival at the cost of growth (Vosloo et al., 2013b), it seems that elevated temperature may indeed affect the productivity of abalone farms.

3.4.3. Validation of iTRAQ using western blot hybridisation

Proteomics data, like any scientific data, has a certain amount of inherent error due to human intervention or experimental limitations. Despite including proper controls and using suitably stringent data analysis techniques, it is possible that bias in the methodology can lead to errors in the results. A complementary molecular technique should therefore be employed to confirm the outcomes of a complex experiment employing a quantitative proteomics approach.

Western blot hybridisation was used to attempt to validate the iTRAQ data obtained in this study, because it measures the same feature (the amount of protein in the cell) and is reasonably robust. Again, working with a non-model organism presents obstacles because commercially available antibodies produced using proteins from other species may not bind to the proteins in *H. midae*. For this reason, antibodies produced to proteins that were highly homologous to the protein sequence identified by iTRAQ were selected. Only 3 of the statistically significant molluscan proteins had sufficient homology to human proteins used to produce primary antibodies commercially. The proteins with potentially suitable primary antibodies available for validation of the iTRAQ results were Hsp70, G6P1D and GSTO1.

The Hsp70 primary antibody had been previously used in this laboratory with good success. In this case, the western blot assay showed a single band at the correct molecular weight. Western blotting of heat shocked abalone haemocyte protein with this antibody revealed no statistically significant change in the amount of Hsp70 detected, however, the pattern of expression appears to match that determined by iTRAQ. After 30 minutes of exposure the amount of Hsp70 detected in the haemocytes appeared approximately 2-fold greater than the control, and elevated levels of Hsp70 appeared to persist for the duration of the heat shock.

The G6P1D antibody consistently detected a band which was 10 kDa smaller than the predicted size of this protein, and featured a prominent doublet in some of the samples. The predicted size of G6P1D (Q0KHB8_CRAGI) is approximately 53 kDa in size; however, the bands detected by western blot were approximately 41.5 kDa. An additional, very faint 84 kDa band was visible in two of the four biological repeats. The 41.5 kDa

band was reliably detected and found to increase significantly at 30 minutes of heat exposure. A doublet band was visible in the 90 minute samples, suggesting post-translational modification of the protein. G6P1D is a known phosphoprotein (Lama et al., 2013), which has variable activity under different phosphorylation states. Since phosphorylation was not specifically enriched in the iTRAQ methodology employed in this study, peptides that retained phosphorylated amino acids may have fallen outside the mass error range in the Peaks search, and therefore may not have been included in the analysis. iTRAQ analysis determined that G6P1D expression increased during heat shock, with the highest levels occurring at 180 minutes exposure. Although the western blot also shows an increase in expression, maximum G6P1D levels were detected much earlier, with the largest expression levels relative to the control occurring at 30 minutes. This discrepancy may be due to a number of factors, including the time of year at which the animals were sampled and possible differences in the stress status of the animals prior to experimentation.

The antibody used in the western blot analysis to validate the GSTO1 expression levels obtained by iTRAQ failed to detect a band of the expected size. Although the predicted molecular weight of GSTO1 is approximately 27 kDa, an 88.4 kDa band was constantly detected. According to the expression profile determined by western blot analysis, GSTO1 was induced by heat shock and remained up-regulated for the duration of the heat shock experiment. This antibody therefore failed to confirm the iTRAQ data, which detected an initial increase in GSTO1, followed by a decrease after 180 minutes of heat stress. One of the major issues with studying this protein is that the glutathione S-transferases are sub-categorised into up to 10 classes (Wan et al., 2009), with some homology between the classes and superfamilies. Since the protein identity was made with peptides comprising 27% and 16% coverage of the total protein (iTRAQ #1 and #2 respectively, Appendix B), and the antibody used in the western blot analysis was raised to human GSTO1, it is possible that expression of the incorrect GST protein was assayed in the western blot.

3.4.4. Concluding remarks and future work

The lack of current literature characterising many of the proteins identified in this study represents an interesting challenge for future research. Several classes or families of proteins were identified which provide a potential starting point for subsequent abalone stress studies. Despite the limitations, the proteins that have been identified in this study encompass a range of diverse cellular functions and could potentially shed light on the mechanisms by which *H. midae* responds to stress. Signalling molecules play a role in conducting the cellular mechanisms of the stress response. Cytoskeletal reorganisation appears to be a major outcome of exposure to acute heat stress, accompanied by a range of signalling molecules. Molecular chaperones which ensure correct folding of nascent proteins are also regulated as the haemocytes attempt to return to homeostasis. Finally, the metabolic function of the haemocytes appears to be perturbed by acute heat stress.

An important step in a large-scale proteomics study is validation of the results using a similar molecular technique in order to attempt to confirm the expression patterns of the identified proteins. For this study, western blot hybridisation was utilised with limited success to validate the iTRAQ results. Although western blotting is a valuable tool for validating proteomic studies, it is important to note that primary antibodies raised to a protein from a different organism, even if high homology may exist with the cognate protein from the target organism, may not be suitable.

Proteomic profiling has great potential in furthering our understanding of the molecular systems at work in a cell or tissue. Since proteomic analysis of non-model organisms currently necessitates alternative database strategies, additional stringency is required in order to prevent errors in the identification of proteins. Therefore, the proteins that are detected and reported herein represent those with the highest statistical probability of correctly matching to *H. midae* proteins. Combining two iTRAQ experiments has the effect of removing low-quality matches with few unique peptides which may be incorrect; however it may also result in the loss of many identities that may have biological importance. As this study represents a pilot study for this type of research, an overly cautious approach was followed in order to obtain the highest quality results.

Chapter 4
CHARACTERISATION OF THE RESPONSE OF A MOLECULAR CHAPERONE TO ACUTE THERMAL STRESS IN *H. MIDAE* HAEMOCYTES

4.1. Introduction

In Chapter 3, the quantitative proteomic technique iTRAQ was used to identify proteins that are differentially expressed during exposure to acute heat stress in the abalone *H. midae*. Abalone are currently considered non-model organisms as the full sequence of the abalone genome is unavailable. Therefore, MS analysis of *H. midae* necessitated the use of a database comprised of all known molluscan protein sequences. Each protein identity determined using this database referred to a protein sequence that was homologous, but not necessarily identical, to its counterpart in *H. midae*. MS analysis identifies proteins based on peptide spectrum matching, which aligns theoretical peptides from a database of known protein sequences to spectra measured from peptides detected in a sample. Computationally this entails that, within a margin of error, the identity of an *H. midae* protein is likely to correspond to a particular protein in the database. Many of the proteins identified in Chapter 3 had no abalone-specific protein sequence available in the database. This therefore represents a gap in the current knowledge of the abalone molecular stress response which must be addressed for the sake of future abalone studies.

In order to further characterise the heat stress response in *H. midae*, a protein was selected from the list of differentially regulated proteins identified by iTRAQ (Chapter 3) for further expression analysis. By selecting a protein with a previously determined heat-induced expression profile, this study will also provide an additional opportunity to validate the iTRAQ expression data using complementary molecular techniques. Furthermore, a protein with no currently available abalone-specific amino acid sequence was selected in order to improve our current understanding of the abalone molecular heat stress response.

All animals are exposed to stresses in their environment, and the ability to respond to stress at a molecular level is an evolutionarily ancient mechanism (Prohászka and Füst, 2004; Kültz, 2003). Proteins are the bioactive components of the cell that respond to external stimuli, and their function is determined by their amino acid composition and tertiary structure. These physical properties of a protein specifically determine the native tertiary structure of the protein under normal cellular conditions. These properties also determine the temperature at which a protein will denature and subsequently unfold, misfold and aggregate (Sørensen et al., 2003). Thermal stress has the undesirable effect of destabilising the physical structure of all proteins, thus altering the rates of physiological

reactions in the cell (Serafini et al., 2011). The accumulation of inappropriately folded proteins in the cell is toxic, and can lead to the triggering of apoptosis (Buck et al., 2007). Proteins known as molecular chaperones assist non-native proteins in attaining their native state under normal cellular conditions (Zhang and Kaufman, 2006). The heat shock protein family are molecular chaperones which are specifically induced by detrimental conditions such as heat shock (Sørensen et al., 2003). For example, one of the major heat shock proteins, Hsp70, has been extensively studied in a number of species, including the abalone *Haliotis discus hannai* (Cheng et al., 2007) and *H. tuberculata* (Farcy et al., 2007), and is differentially expressed during acute heat stress in *H. midae* (Chapter 3). There are many additional important abalone molecular chaperones that as yet lack gene sequence information or functional characterisation.

Of the putatively identified proteins in Chapter 3, the calcium binding molecular chaperone calreticulin (CRT) was amongst those that were observed to be differentially regulated during exposure to acute heat stress. CRT is known to be induced by conditions that cause accumulation of proteins in the endoplasmic reticulum (ER), triggering the unfolded protein response (UPR) (Kawabe and Yokoyama, 2010). CRT is located in the ER lumen, where it binds Ca^{2+} with a high affinity (Michalak et al., 1999). An increase in CRT expression causes an increase in the amount of Ca^{2+} in the ER (Bastianutto et al., 1995; Qiu and Michalak, 2009). Ca^{2+} is involved in signal transduction for a vast array of cellular functions (Berridge et al., 2003), and the mechanics of its storage and release are extremely complex (Berridge et al., 2000). The lumen of the ER is an internal calcium storage location for the cell, and the release of Ca^{2+} from internal stores regulates many diverse cellular processes (Lam and Galione, 2013).

Calreticulin has had limited attention in the field of molluscan molecular research, and to date the only available molluscan calreticulin amino acid sequences in the UniProt database are for the sea hare *Aplysia californica* (Kennedy et al., 1992), the oysters *Pinctada fucata* (Fan et al., 2008) and *Crassostrea gigas* (Kawabe and Yokoyama, 2010) and recently the mussel *Hyriopsis cumingii* (accession T1PZZ4_9BIVA). These studies indicate that the molluscan calreticulin gene is differentially regulated during abiotic stress and therefore identify calreticulin as a candidate stress response protein in the mollusca. The immune capacity of the mussels *M. edulis* and *M. galloprovincialis* has been associated with the level of calreticulin protein present in the animal (López et al., 2002).

4.1.1. Aim of this study

Future proteomic studies of non-model organisms such as *H. midae* depend on the availability of suitable protein databases for MS analysis. Although iTRAQ determined that CRT was differentially regulated in abalone haemocytes during exposure to acute heat stress (Chapter 3), the abalone CRT amino acid sequence is currently not present in the UniProt database. This important protein is poorly understood in abalone. A 387 bp portion of the *H. midae* calreticulin mRNA sequence was identified previously in our laboratory (Jarrod Lyons, 2007 – unpublished data). This information facilitated further characterisation of CRT in *H. midae* haemocytes.

The aim of this study is to characterize the expression of the molecular chaperone calreticulin in the South African abalone *H. midae* during exposure to acute heat shock. In so doing, the heat shock induced expression profile of calreticulin observed by iTRAQ analysis will be validated using different molecular techniques. This study will also assess the suitability of calreticulin as a molecular indicator of heat stress in *H. midae*. In order to accomplish these aims, the following objectives were addressed:

i) A portion of *H. midae* CRT cDNA was amplified by polymerase chain reaction (PCR), and cloned into a cloning vector for sequencing;
ii) The sequence information was used to characterise the features of *H. midae* calreticulin using available bioinformatics tools;
iii) Expression profiling of the CRT response to acute heat stress was carried out using both mRNA based techniques and western blot hybridisation with a commercially available anti-CRT antibody.

4.2. Materials and Methods

Solutions used in this chapter are listed in Appendix A.

For the purposes of this study, a biological replicate is defined as samples obtained from distinct groups of abalone under the same environmental conditions. A technical replicate is the replication of the same experimental technique on the same sample.

4.2.1. Abalone stocks and maintenance

Abalone were maintained at the DAFF Research aquarium as described in 2.2.1.

4.2.2. Sequencing of an *H. midae* CRT mRNA

Rapid amplification of cDNA ends (RACE) is a method that uses gene specific primers with commercially produced 'anchoring' primers to extend the PCR amplified regions of a target gene in the 3' and 5' direction. The PCR products of 3' and 5' RACE can then be cloned into an appropriate vector for sequencing. This method was used in *H. midae* to obtain additional cDNA sequence information for the CRT gene.

4.2.2.1. RACE primer design

The partial (397bp) sequence of *H. midae* calreticulin was previously determined in our laboratory (Lyons, 2007). This sequence was used to design appropriate gene specific primers (GSPs) for 3' and 5' RACE PCR using the program DNAMAN, version 5.2.10 (Lynnon BioSoft©, Quebec, Canada). The sense primer (5' – GGA CAA GCC AGA ACA CAT CCC CGA CCC TGA TGC C – 3') was designed to the 3' end of the known sequence, and had a Tm of 70.7 °C. The high Tm facilitated 'touchdown' RACE PCR, which allowed gene specific product to accumulate in the early phases of the PCR, increasing the likelihood of producing the intended product. The antisense primer (5' – TCC GCT CTC GAC CTT CTC GTT ATC G – 3') was designed to the 5' end of the known sequence, and had a Tm of 62.7.

4.2.2.2. RNA isolation and RACE ready cDNA conversion

RNA was isolated from *H. midae* haemocytes according to Chomczynski and Sacchi (1987) (see 4.2.3.2 for the full method). The RNA was quantified using Nanodrop® ND-1000 UV-Vis spectrophotometer (Nanodrop technologies) and the quality was determined by electrophoresing 500 ng of total RNA on a 1.2% agarose/formaldehyde denaturing gel to ensure there was no degradation. The SMART™ RACE cDNA Amplification Kit (Clontech Laboratories Inc) was then used for RACE ready cDNA conversion according to the manufacturer's instructions. Described briefly, the 3' (5'–AAG CAG TGG TAT CAA CGC AGA GTA C(T) 30V N–3' (N = A, C, G, or T; V = A, G, or C)) and 5' (5'–(T)25V N–3' (N = A, C, G, or T; V = A, G, or C)) CDS primer A were each annealed to 1 μg of total RNA in clean 1.5 ml microcentrifuge tubes at 72 °C for 3 minutes, then cooled at 42 °C for 2 minutes. The SMART IIA oligo adapter (1 μl) was added to the 5' RACE cDNA conversion only. Superscript™ II RNase H- Reverse Transcriptase (Invitrogen) (100 U) was then added to each tube and the reaction was incubated at 42 °C for 1.5 hours before inactivation by incubation at 70 °C for 10 minutes. The resulting cDNA was diluted with tricine EDTA (TE) buffer (Appendix A), and used directly for RACE PCR.

4.2.2.3. RACE PCR

RACE PCR was performed using the RACE ready cDNA described above as a template. The primers used for 3' and 5' RACE PCR were the 3' and 5' gene specific primers respectively (as described 4.2.2.1), and the Universal Primer Mix (UPM) provided in the SMART™ RACE cDNA Amplification Kit (Clontech) (5'-CTA ATA CGA CTC ACT ATA GGG CAA GCA GTG GTA TCA ACG CAG AGT-3'). PCR reactions were prepared using a master mix of the Advantage 2 Polymerase mix and PCR buffer provided in the kit at a 1X total concentration. Touchdown PCR parameters were as follows: 94 °C for 5 min, followed by 5 cycles of 94 °C for 30 seconds, 72 °C for 3 minutes, followed by 5 cycles of 94 °C for 30 seconds, 70 °C for 30 seconds, 72 °C for 3 minutes, followed by 30 cycles of 94°C for 30 seconds, 68 °C for 30 seconds, 72 °C for 3 minutes and a final elongation step of 72°C for 7 minutes. PCR products were visualised under UV light after electrophoresis on a 1.2% agarose TAE gel with ethidium bromide added.

4.2.2.4. Cloning of RACE PCR product into PTZ57R/T plasmid

The visualised PCR product corresponding to the expected size was extracted from the gel and cleaned using the Biospin gel extraction kit (Biospin). The cleaned PCR product was then ligated into the A/T cloning vector PTZ57R/T, supplied with the InsTAclone™ PCR cloning kit (Fermentas) according to the manufacturer's instructions. The ligation mix was transformed into competent *Escherichia coli* JM109 cells which were inoculated onto 0.002% w/v XGal, 0.5 mM IPTG, 100 μg/ml ampicillin Luria agar plates. The plates were incubated at 37 °C overnight, and from the white colonies positive transformants were picked and re-streaked onto fresh plates, which were also incubated at 37 °C overnight in order to confirm that single white colonies had been isolated. Individual colonies from each positive transformant were then inoculated into 5 ml Luria broth + 100 μg/ml ampicillin and grown to exponential phase by incubation at 37 °C overnight.

The plasmids were purified out of the transformed *E. coli* JM109 by alkaline lysis (Ausubel *et al.,* 1989). The insert was amplified out of the purified plasmids by PCR using universal M13 primers (M13 forward 5'-CGC CAG GGT TTT CCC AGT CAC GAC-3', reverse 5'-GAG CGG ATA ACA ATT TCA CAC AGG-3') which flank the multiple cloning site on PTZ57R/T. The PCR profile was as follows: 94 °C for 5min, followed by 25 cycles of 94 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 1 minute, followed by a final elongation step of 72 °C for 7 minutes. Clones were chosen based on the presence of a band in the range corresponding to the size of the original PCR product. The chosen clones were once again grown up in Luria broth + Amp (100 μg/ml) and the plasmid was isolated with the peqGOLD plasmid mini prep kit I (peqLab Ltd.) according to manufacturer's instructions. The concentration of the final eluent was determined by spectrophotometry using the NanoDrop® system (Thermo Scientific). PCR with universal M13 primers (as above) was used once again to confirm the presence of the insert.

4.2.2.5. Sequencing

20 µl of a 100 ng/µl sample of the purified clones were sent to Macrogen Inc. (1001 World Meridian Center 60-24, Gasan-dong Geumchun-gu Seoul Korea 153-021) for sequencing. The resultant sequence files were edited using Chromas version 2.01 (©Technelysium Ltd, Queensland, Australia) to remove the flanking regions with poor signal.

Initially an additional ≈400 bp of sequence was obtained from the 3' RACE PCR product, and so a gene specific primer (5'- TCA GGA TGA GGA GGA GAG G-3') was designed towards the 3' end of this new sequence using DNAMAN© software (version 5.2.10 Lynnon Biosoft) in order to extend the read length. The sequence files returned by Macrogen Inc. were assembled into a consensus sequence using DNAMAN.

4.2.2.6. Sequence identity confirmation and phylogenetic analysis

In order to confirm the identity of the consensus sequence obtained by RACE, the BLAST tool was used to find regions of similarity between it and other known sequences (www.ncbi.nlm.nih.gov). A deduced amino acid sequence was determined by predicting the open reading frame for the consensus sequence using DNAMAN software, which was confirmed using the open reading frame finder tool on the NCBI website. The open reading frame was translated using the standard genetic code table, and the resultant deduced amino acid sequence was then compared to known protein sequences using the NCBI's BLASTx tool to identify the protein. Protein family motifs in the deduced amino acid sequence were found using a Prosite scan (http://www.expasy.ch/tools/scanprosite/) which identifies family or superfamily motifs in the amino acid sequence by comparing it to sequences stored in the Prosite database.

4.2.3. Transcriptional regulation of CRT during exposure to acute heat shock

In order to better understand the role of CRT in the heat stress response of *H. midae*, an experiment was carried out to assay the transcriptional regulation of CRT during exposure to thermal stress. Haemocytes were sampled at 15, 30, 45, 60 and 120 minutes of exposure and compared to the untreated control group (T0) using real time qPCR.

4.2.3.1. Heat shock and isolation of haemocytes for real time qPCR

Healthy *H. midae* kept at the DAFF research aquarium were placed in 50 L glass tanks at ambient circulating water temperature to acclimatise for 24 hours overnight. Haemolymph was removed from the T0 sample group using a 21 G x 1½" needle via the pedal sinus, before the remainder of the animals were moved to an identical glass tank in which the water had been heated to 23 °C. Samples were taken at 15, 30, 45, 60 and 120 minutes after placement in the heated tank. At least 3 ml of haemolymph was obtained at each time point from no less than 5 animals. The remaining haemolymph was divided into 2 x 1 ml aliquots in 1.5 ml eppendorf tubes, and centrifuged at 14 000 rpm for 2 minutes. The supernatant was removed and the remaining haemocyte pellets frozen in liquid nitrogen, where they remained until they could be stored at -80 °C.

4.2.3.2. RNA isolation from haemocytes

RNA isolation proceeded according to Chomczynski & Sacchi (1987) with some modifications. All glassware used in RNA isolation was cleaned first with 70% EtOH, before it was autoclaved twice in order to remove RNAses, and all solutions were made up using high grade RNAse free reagents, and DEPC treated water (Appendix A).

The haemocyte pellets were removed from -80 °C and thawed on ice. The pellets were resuspended in Solution D (Appendix A) by vortexing for 1 minute, and further homogenized by passing the pellet through a 1 ml pipette tip until fully resuspended. The RNA was then removed by phenol:chloroform:isoamyl alcohol extraction and the aqueous phase was gently mixed with 100% isopropanol and incubated overnight at -20 °C to precipitate the RNA. The precipitate was pelleted in a centrifuge at 14 000 rpm for 20

minutes in a centrifuge at 4 °C. The pellet was once again dissolved in Solution D, precipitated again in 100% isopropanol at -20 °C, then pelleted by centrifugation at 14 000 rpm at 4 °C. The pellet was then washed with ice cold ethanol, once at 100% then once at 75% concentration, before being dried in a fume hood and dissolved in 25 µl DEPC-dH_2O. The resuspended total RNA was then DNase I (Thermo Scientific) treated for 3 hours at 37 °C, before recovery by phenol: chloroform: isoamyl alcohol extraction. The aqueous phase was removed to a new 1.5 ml microcentrifuge tube, and RNA was precipitated by incubation with 3 M sodium acetate, then washed with ice cold absolute ethanol and pelleted by centrifugation at 14 000 rpm for 10 minutes at 4 °C. The pellet was air dried in a fume hood and resuspended in 25 µl DEPC-dH_2O. The quantity of RNA was determined by Nanodrop® (Thermo Scientific), and the quality of each sample was checked on a 1.2% w/v denaturing agarose/formaldehyde mini gel. Photographs of the RNA gels taken with the Bio-Rad GelDoc EQ-system™ (Bio-Rad) are included in Appendix B. The RNA was then used directly for cDNA conversion.

4.2.3.3. cDNA conversion

Total RNA was converted to cDNA using the ImProm-II™ Reverse Transcription system (Promega) according to the manufacturer's instructions. The primers (1 µl of oligo dT and 1 µl of random hexamer primers) were annealed to 2 µg of RNA from each sample by incubation at 70 °C for 5 minutes. The reaction components (5x reaction buffer, $MgCl_2$, dNTP's, RNAse inhibitor and reverse transcriptase) were added to each reaction mix to the specified concentrations. The reactions were incubated at 25 °C for 5 minutes, then in a water bath at 42 °C overnight. The reaction was inactivated by incubation at 70 °C for 15 minutes in a heating block. The cDNA was then used directly in real time qPCR reactions.

4.2.3.4. Real time qPCR

4.2.3.4.1. Primer design for real time qPCR

Real time primers were designed to the *H. midae* calreticulin sequence using Beacon designer (Premier Biosoft). The primers were designed towards the 3' end of the known sequence with a Tm=62.6 and 62.7 °C, and an amplicon size of 110 bp (Forward 5'- CCG

ATG ACT GGG ATG AGC GTG AG-3', Reverse 5'- TCA TCA GGG CTT CTT GGC ATC AGG-3'). The design software algorithm automatically selects the best pair of primers according to the optimal parameters for a SYBR® green reaction. The primers were tested for specificity using the NCBI database BLAST algorithm, and found to correspond specifically to calreticulin sequences in the database.

4.2.3.4.2. Reference gene selection

The suitability of potential reference genes for real-time qPCR analysis was determined by assessing their expression profiles under experimental conditions. PCR primers designed to the *H. midae* 18s RNA subunit sequence by Bronwyn Arendse-Bailey (based on microarray work performed during the course of her PhD research (personal communication, unpublished data)) had been previously used in our laboratory as a reference gene. Another gene identified by Brownyn Arendse- Bailey (personal communication, unpublished data) in a microarray as 'ribosomal factor L28' was constitutively expressed during her experimental conditions and was therefore selected as a candidate reference gene. The expression of these genes during exposure to acute heat shock was determined by real-time qPCR, and the gene showing the most stable expression was selected as the reference gene.

4.2.3.4.3. Polymerase chain reaction (PCR)

Real time qPCR was carried out in a Rotor-Gene™ 6000 (Corbett Life Sciences) machine. The PCR profile was 40 cycles of 95 °C for 10 sec, 60 °C for 15 sec and 72 °C for 20 sec. Melt curve analysis of the PCR product was performed by increasing the temperature in 1 °C increments from 72 °C to 95 °C.

Real time qPCR reactions proceeded in thin walled real time qPCR tubes (Corbett Life Sciences) in a 72 well rotor. Each 12.5 µl PCR reaction contained 1x Sensimix (containing 3 mM $MgCl_2$), 1 x SYBR® Green, 0.2 µM forward and reverse primers, 1 µl of cDNA template and sterile Millipore™ milliQ H_2O. A no template control (NTC) and no reverse transcriptase control (no RT) were included in all real time qPCR experiments. In no template control reactions, 1 µl of milliQ H_2O was substituted for the cDNA template. In no RT controls, 1 µl of pooled RNA from all time points was substituted for the cDNA

template. In reactions pertaining to the gene of interest, a purified plasmid containing a fragment of *H. midae* calreticulin was included as template in a positive control. All time points, standard curve points and controls were performed in triplicate, and four biological replicates of the heat shock experiment were performed.

4.2.3.4.4. Statistical data analysis

The standard curve equation, threshold, and resultant Ct values were calculated using the Rotor-Gene™ 6000 series software package (version 1.7, Corbett Life Sciences). The data was exported to Microsoft® Office Excel (2010), and mean Ct values were then used in the Pfaffl (Pfaffl, 2001) equation to calculate relative expression. Calreticulin Ct values were normalised to reference gene Ct values, and this was then normalised to time 0 for each time point according to the Pfaffl equation:

$$\text{ratio} = \frac{(E_{\text{target}})^{\Delta \text{CP}_{\text{target}}(control-sample)}}{(E_{\text{ref}})^{\Delta \text{CP}_{\text{ref}}(control-sample)}}$$

Where 'E' is the efficiency of the real time reaction, 'target' denotes the gene of interest and 'ref' denotes the reference gene. In this case the ΔCP (or cycle threshold, otherwise denoted as Ct) is the difference in Ct values for the 'control' T0, the untreated sample, and the time point in question or 'sample'.

Statistical analysis for comparisons between the control and experimental samples was performed using a t-test with $p<0.05$ accepted as a significant difference between values (SigmaStat).

4.2.4. Examination of CRT protein regulation during exposure to acute heat shock

The aim of this experiment was to determine whether CRT protein is differentially regulated during exposure to heat stress. This was achieved by western blot analysis using commercially available polyclonal antibodies raised to a human CRT synthetic peptide in rabbits. The selection of the primary antibodies was informed by the deduced *H. midae* CRT amino acid sequence determined in 4.2.2. Commercially available CRT antibodies were screened by aligning the antigen sequence provided by the manufacturer with the

corresponding region of the known *H. midae* CRT amino acid sequence. A polyclonal rabbit anti-CRT primary antibody (Abcam cat# ab2907), generated to recombinant human calreticulin produced in the Baculovirus insect cell system, was selected based on 79% homology to the *H. midae* CRT amino acid sequence.

4.2.4.1. Heat shock and extraction of haemocytes

As described in 4.2.3.1, haemolymph was sampled from *H. midae* after 0, 30, 60, 120, 180 minutes of exposure to heat shock and stored at -80 °C. A recovery time point (24 h) was sampled from animals that had been exposed to 180 minutes of heat shock, and then were allowed to recover for 24 hours at the ambient water temperature (13 °C). The heat shock experiment was repeated 3 times (n=3) on naïve animals.

4.2.4.2. Detection and expression profiling of CRT in H. midae by western blot hybridisation

Haemocyte processing, total protein isolation and western blot analysis was carried out according to the method described in 3.2.6.4. Protein samples were loaded on to a 10% separating polyacrylamide gel and electrophoresed at 100 V. Once sufficiently separated, the protein was transferred electrophoretically onto a nitrocellulose membrane in a vertical wet transfer tank (Bio-Rad Mini Trans-Blot® Cell) in Towbin buffer (Appendix A). Equal loading of samples on the nitrocellulose membrane was confirmed by reversible total protein staining with Ponceau S. The CRT primary antibody was then applied to the nitrocellulose membrane at a concentration of 1:2000 in 5% blocking buffer (Appendix A) and incubated overnight at 4 °C. The secondary antibody was applied at a concentration of 1:20000 in 5% blocking buffer (Appendix A) and incubated for 2 hours at room temperature with shaking. WesternBright ECL HRP substrate was applied according to the manufacturer's instructions (Advansta) for chemiluminescent detection of CRT. The Bio-Rad Chemidoc XR captured the chemiluminescent signal, which was processed in ImageLab software (Bio-Rad). The intensity of the bands was measured using the rectangular volume analysis tool in ImageLab, with background subtraction set to 'global'. Expression analysis of *H. midae* CRT during exposure to heat shock was calculated by determining the relative change in intensity of each experimental sample to the control

sample (T0). A t-test was used to statistically verify any observed differences in expression between control and experimental groups.

4.3. Results

4.3.1. Sequence identity confirmation and phylogenetic analysis of *H. midae* CRT

Rapid amplification of cDNA ends (RACE) was performed in order to obtain the full length cDNA sequence of *H. midae* CRT. Although 5' RACE PCR was attempted on several occasions, the reaction failed to produce a PCR product. However, 3' RACE PCR yielded an ~1100 bp PCR product which was sub-cloned into the pTZ57R/T vector and transformed into *E. coli* JM109 (Figure 4-1).

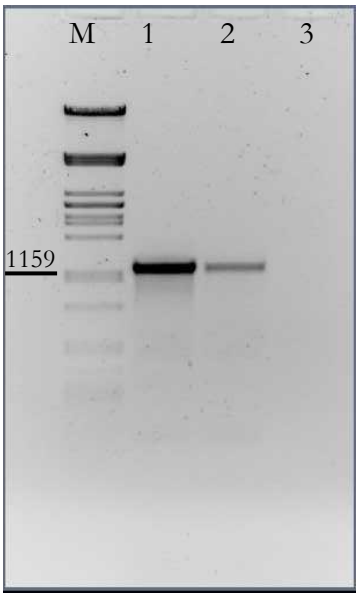


Figure 4-1: PCR amplification of a ~1100bp fragment of *H. midae* CRT on a 1.2% TAE agarose gel. The insert was amplified from the plasmid using universal M13 F and R primers. Lane M - Lambda DNA digested with *Pst* 1, lane 1 - 10 ng template, lane 2 - 1 ng template, lane 3 - no template control (NTC). The 1159 bp molecular weight marker is shown to facilitate size determination of the PCR product.

The nucleotide consensus sequence was assembled in DNAMAN, with a total size of 1366 bp (denoted as HM-CRT). A putative polyadenylation signal and poly-A tail are present in the nucleotide sequence (Figure 4-2), confirming that the 3' end of the sequence is complete. However, there is a fragment missing at the 5' end as 5' RACE was unsuccessful and therefore, the sequence reported represents a portion of the total *H. midae* CRT sequence.

The amino acid sequence of HM-CRT was deduced from the nucleotide sequence using DNAMAN software, and confirmed with the open reading frame (ORF) finder tool on the NCBI website. The algorithm for ORF detection assigned a methionine (ATG) as the putative start position by default. As the 5' end of the HM-CRT sequence is missing, this methionine does not actually represent the start of the *H. midae* CRT sequence. The longest ORF was identified in the +3 reading (Figure 4-3). The deduced 281 amino acid sequence of the identified ORF terminates in an endoplasmic reticulum retrieval signal (HDEL) at position 821 on the nucleic acid sequence, followed by a putative stop codon (TAG) (Figure 4-2). A Prosite scan revealed the calreticulin family repeated motif signature (PS00805) in the deduced amino acid sequence, which is highly conserved in the calreticulin sequences of other organisms.

```
3     ATG TTC GGC CCG GAT ATT TGC GGT CCA GGA ACA AAG AAG GTT CAT   47
0      M   F   G   P   D   I   C   G   P   G   T   K   K   V   H    14

48    GTC CTT TTC AAC TAC AAG GGA AAG AAT CTT TTG ACA AAG AAA GAT   92
15     V   L   F   N   Y   K   G   K   N   L   L   T   K   K   D    29

93    ATC CGA TGC AAG GAT GAT GTT TCC ACT CAC CTG TAC ACC CTG ATT   137
30     I   R   C   K   D   D   V   S   T   H   L   Y   T   L   I    44

138   GTG AAT GCT GAC AAC ACA TAC GAG GTG AAG ATC GAT AAC GAG AAG   182
45     V   N   A   D   N   T   Y   E   V   K   I   D   N   E   K    59

183   GTC GAG AGC GGA GAG CTT GAG GCA GAT TGG GAT TTC CTT CCA GCC   227
60     V   E   S   G   E   L   E   A   D   W   D   F   L   P   A    74

228   AAG AAA ATC AAG GAC CCC GAC GCA AAG AAG CCC GAT GAC TGG GAT   272
75     K   K   I   K   D   P   D   A   K   K   P   D   D   W   D    89

273   GAG CGT GAG AAG ATC GAT GAT CCT GAT GAC ACC AAG CCC GAG GAT   317
90     E   R   E   K   I   D   D   P   D   D   T   K   P   E   D   104

318   TGG GAC AAG CCA GAA CAC ATC CCC GAC CCT GAT GCC AAG AAG CCT   362
105    W   D   K   P   E   H   I   P   D   P   D   A   K   K   P   119

363   GAT GAC TGG GAT GAT GAG ATG GAT GGT GAA TGG GAG CCC CCA ATG   407
120    D   D   W   D   D   E   M   D   G   E   W   E   P   P   M   134

408   GTA GAT AAC CCT GAT TAC AAG GGA GAA TGG AAA CCC AAG CAG ATT   452
135    V   D   N   P   D   Y   K   G   E   W   K   P   K   Q   I   149

453   GAC AAC CCA GAC TAC AAA GGC AAA TGG GTC CAC CCT GAA ATT GAC   497
150    D   N   P   D   Y   K   G   K   W   V   H   P   E   I   D   164

498   AAC CCT GAC TAC TCT CCC GAT GAA AGC CTC TAC AGA TAC ACA GAT   542
165    N   P   D   Y   S   P   D   E   S   L   Y   R   Y   T   D   179

543   ATT GGT GCT ATC GGC TTT GAT TTG TGG CAG GTT AAA TCT GGA ACC   587
180    I   G   A   I   G   F   D   L   W   Q   V   K   S   G   T   194

588   ATC TTT GAT AAT GTC CTC ATT ACT GAT GAT GTA GAC TTC GCA GAA   632
195    I   F   D   N   V   L   I   T   D   D   V   D   F   A   E   209
```

```
633   GAA TTT GGA AAA TCT ACA TGG GGC AAA ACA AAG GAC CCT GAA AAG   677
210    E   F   G   K   S   T   W   G   K   T   K   D   P   E   K    224

678   AAG ATG AAA GAT GCT CAG GAT GAG GAG GAG AGG AAA CAG AGG GAG   722
225    K   M   K   D   A   Q   D   E   E   E   R   K   Q   R   E    239

723   GAG GAA GAG AAG AAG AGG AAA GAG GAG GAA GAT GCC AAG AAG GAA   767
240    E   E   E   K   K   R   K   E   E   E   D   A   K   K   E    254

768   GAT GGT GAT GAG GCA GAG GAG GAA GAA GAA GAG GAT GAC TCA GAT   812
255    D   G   D   E   A   E   E   E   E   E   E   D   D   S   D    269

813   AAA AAA GAA GAA CCT GAT CAT GAC GAG TTG TAG AGA TCT TCC AGA   857
270    K   K   E   E   P   D   H   D   E   L   *

858   GAT TAA CCA ATC ATC TTC ATC CGT CAT TTG TCA CCA TTC ATT TCA   902
903   TGC CAT TCA CAT GAC GTA TAA TAT GGG AAG TCT TTT AGT CGG GGT   947
948   CAA AAG CTT TCA TCA AAG CTC TTA AGG AAA TTT GTT AAG TTT TTG   992
993   TTT TTT ATT TTG TAT TCA TCG GCC ATT GTG TGT GAG ATG TTT TAA  1037
1038  AGA AAA GTC AAC AGC GTT TCA CCC TAA TAT GCC ACC ATG AAC TCG  1082
1083  CTA AAA ACA AAA TGT GTG TGT GAA GAT GAT ATG CAT GGG ACA GTC  1127
1128  CTA TGT TCT TGG TAA TGG CAT TGC TGT CTA CAT TGT TAG GTT GTC  1172
1173  TGT CAT ATT AGA TGA TTT CAC TCT TTG TTT TGT ATT TAA ATA TTT  1217
1218  AAT GTA CAA TTT AAG TGA TGT ATA AAG AAC TAT TGT AAA TTG AAT  1262
1263  GCT ATT CAT GTA AAT TGC TCT GAA ACA TAA CTA ATG TAG GGG TGT  1307
1308  TTA TAC CAA AGT GAG GTA ATA AAT GCT TCT TTT TAA AAA AAA AAA  1352
1353  AAA AAA AAA AAA
```

Figure 4-2: (Continued from previous page) Partial nucleotide and derived amino acid sequence of *H. midae* CRT. Three CRT family repeat sequences are underlined. An ER retrieval sequence HDEL is shown in italics, while the asterisk (*) denotes the stop codon. The putative polyadenylation signal is shown in bold and underlined. The putative amino acid sequence was deduced using DNAMAN software and the ORF finder tool on the NCBI website. Although the open reading frame algorithm begins the derived amino acid sequence with a methionine (ATG), it should be noted that this does not represent the start of the full coding sequence of HM-CRT.

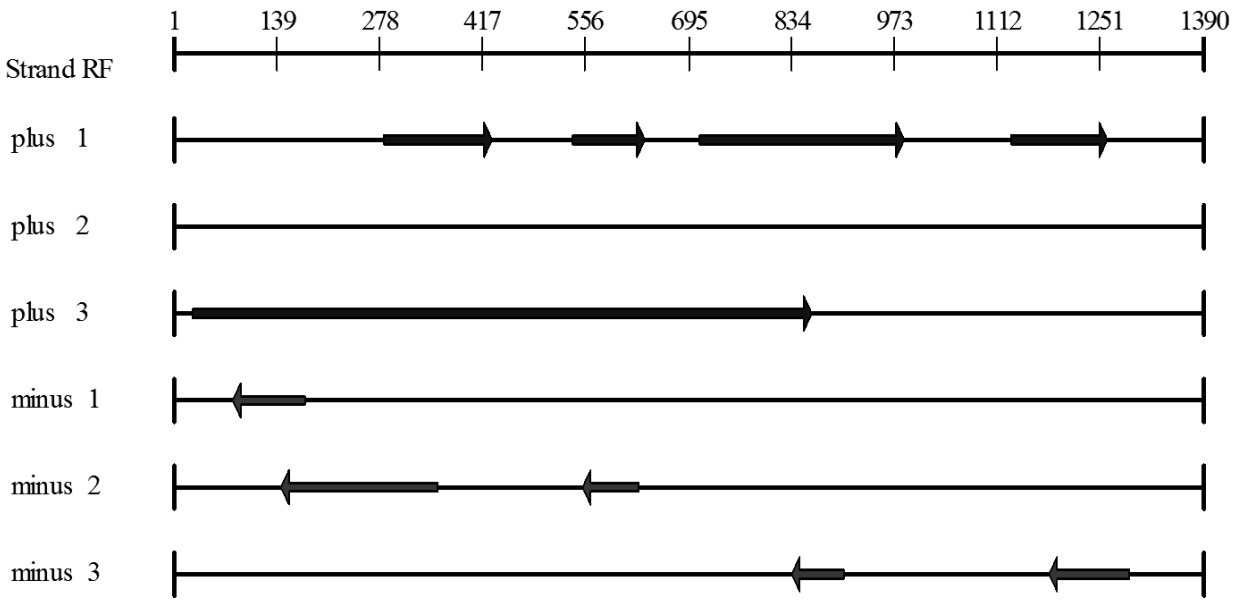


Figure 4-3: Open reading frames identified in the assembled partial nucleotide sequence of HM-CRT using the universal genetic code in DNAMAN, and confirmed using the ORF finder tool on the NCBI website. Frame plus 3 represents the open reading frame used to derive the partial deduced amino acid sequence of calreticulin.

A BLASTx search of the NCBI database (Table 4-1) using the translated amino acid sequence of HM-CRT indicated that the CRT from *Aplysia californica* has the greatest sequence similarity (87%), followed by *Crassostrea gigas* with 86% similarity. *Pinctada fucata*, *Eisenia fetida* and *Gallus gallus* CRT amino acid sequences all share over 80% similarity with HM-CRT. More distantly related animals such as *Cricetulus griseus*, *Homo sapiens* and *Mus musculus* share 78% sequence homology with the deduced sequence.

Table 4-1: Similarity of CRT from selected organisms to the deduced *Haliotis midae* CRT amino acid sequence as determined by BLASTx analysis.

Species	Accession number	Identity (%)	Total Score	E-value
Aplysia californica	AAB24569.1	87%	280	4e-86
Crassostrea gigas	BAF63639.1	86%	280	1e-85
Pinctada fucata	ABR68546.1	83%	259	9e-78
Eisenia fetida	ABI74618.1	81%	260	5e-78
Gallus gallus	AAS49610.1	81%	259	5e-78
Cricetulus griseus	AAM48568.1	78%	251	1e-74
Homo sapiens	EAW84330.1	78%	254	8e-76
Mus musculus	NP_031617.1	78%	250	2e-74
Ixodes jellisoni	AAR29949.1	77%	253	2e-75
Oryctolagus cuniculus	NP_001075704.1	77%	251	1e-74
Xenopus laevis	NP_001080765.1	76%	251	1e-74

4.3.2. Transcriptional regulation of *H. midae* CRT during exposure to heat shock

The relative expression of CRT in *H. midae* haemocytes during exposure to thermal stress was determined by real-time qPCR. The Ct values for CRT were normalised against Ct values for the reference gene, ribosomal factor L28, over the same time course. The selection of the reference gene was based on its constitutive expression under experimental conditions (data shown in Appendix B).

Figure 4-4 shows the relative fold change in CRT mRNA expression during exposure to heat shock, as determined by real-time qPCR. The raw data were normalised and calibrated to the control (0) by the Pfaffl method. The overall fold change in CRT expression in response to heat shock is within the range of 0.2 to 1.6 fold. CRT expression decreased after 15 minutes of heat shock, followed by an increase at subsequent sample points. The average fold change differed significantly between the control (0) and 15 minute sample (Mann-Whitney U test, $p < 0.05$). The no template (NTC) and no reverse transcriptase control (no RT) either had Ct

values that were more than 10 cycles later than the samples containing cDNA template, or they did not cross the cycle threshold during the reaction. NTC reactions did not generate a peak during the melt curve analysis.

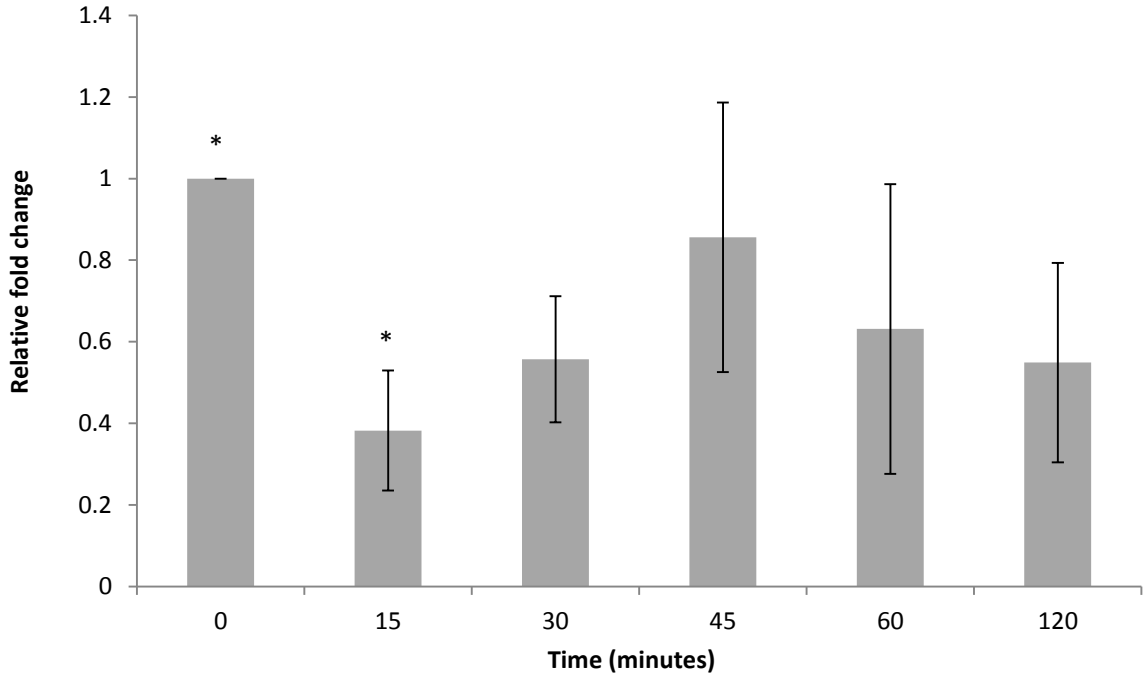


Figure 4-4: Expression of the calreticulin gene in *H. midae* during a heat shock time trial as determined by real-time qPCR, normalised to the reference gene ribosomal factor L28 and normalised to T0 using the Pfaffl method. Data are represented as mean relative fold change ± standard error for triplicate technical replicates of the PCR at each time point and across four biological replicates (n=4). Statistical analysis between T0 and each other time point was performed using a Mann-Whitney U test in SigmaStat 3.10.0. Asterisks represent statistical significance vs T0 at $p<0.05$.

4.3.3. Translational regulation of *H. midae* CRT during exposure to heat shock

The translational regulation of *H. midae* CRT during exposure to acute elevated temperature was measured using western blot hybridisation. Equal loading was confirmed by reversible Ponceaus S stain of the protein on the nitrocellulose membrane. This confirmed that each protein sample was loaded and transferred equally to the nitrocellulose membrane (Appendix B).

Following chemiluminescent detection, the density of the CRT band was determined and plotted (Figure 4-5). There was an increase in the amount of CRT after 30 minutes of heat shock, and after 2 hours the amount of CRT protein in the haemocytes had approximately doubled, which was statistically significant according to the t-test ($p < 0.05$).

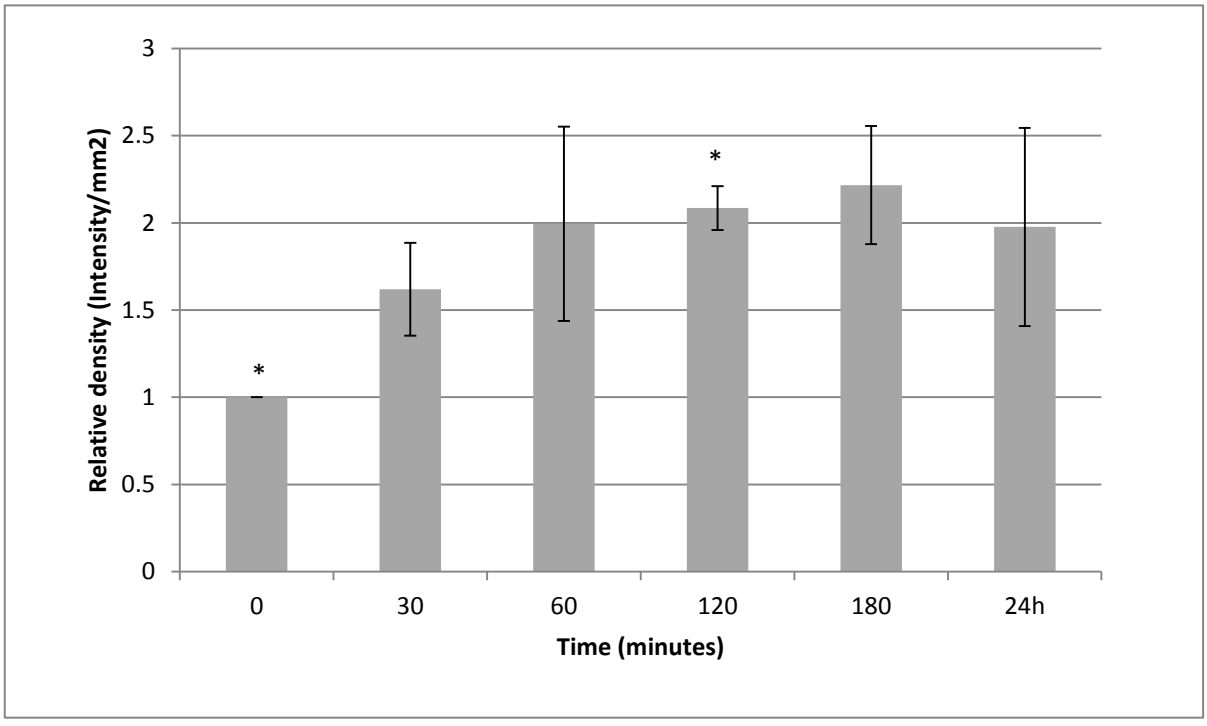


Figure 4-5: Expression of calreticulin protein in *H. midae* during a heat shock time course determined by western blot analysis. Data are represented as average intensity of a 60kDa band ± standard error at each time point and across three biological repeats. Statistical analysis was performed using a Students t-test in SigmaStat 3.10.0. Asterisks (*) represent a statistically significant difference with $p<0.05$.

4.4. Discussion

Proteomics is a rapidly expanding field of molecular enquiry, which is capable of simultaneously identifying hundreds of proteins that are differentially regulated during exposure to detrimental conditions. MS based proteomic studies still rely on a genomic database for the identification of proteins, which has somewhat slowed the progress of proteomic investigations of non-model organisms such as *H. midae* (Rodrigues et al., 2012). Fortunately, new sequences are constantly being added to proteomic databases and therefore there is great potential for future research (Zhou et al., 2012). In the interim, MS analysis of non-model organisms can be performed by searching entire databases, or subsets thereof, in order to assign protein identities based on sequence homology between species. This approach was adopted in Chapter 3 in order to identify proteins in *H. midae* haemocytes that are differentially expressed during exposure to acute heat shock.

The molecular chaperone calreticulin was identified by iTRAQ and appeared to be differentially expressed in the haemocytes of *H. midae* experiencing acute heat stress. Molecular chaperones such as calreticulin play a vital role in the endoplasmic reticulum (ER) response to stress, and are responsible for the proper folding of misfolded/unfolded proteins. CRT also has the capacity to bind Ca^{2+}, which is a signalling molecule implicated in many cellular processes. The differential regulation of CRT between two species of mussel has been associated with the animals' ability to cope with stress (López et al., 2002), which indicates that CRT in abalone may be involved in the molecular stress response. However, this protein had not been previously characterised in abalone. The aim of this study, therefore, was to characterise the calreticulin (CRT) gene in *H. midae* by analysis of CRT regulation during heat shock in *H. midae* haemocytes, in order to gain insight into the abalone heat stress response and to facilitate further validation of the iTRAQ results obtained in Chapter 3.

There is currently no sequence information available for abalone CRT, and so in order to assess the effect of heat shock on the expression of *H. midae* CRT using alternative molecular methods it was first necessary to obtain sequence information for this protein. The HM-CRT cDNA sequence was obtained using RACE to extend the known sequence to the 3' and 5' ends. Despite several attempts, 5' RACE proved unsuccessful. There are a number of possible reasons for this but the most likely is either degradation of the 5' end of the mRNA, or termination of the cDNA conversion reaction before completion. The difficulty of

successfully performing 5' RACE is well established, and has prompted the development of alternative RACE methods for improved 5' RACE success (Bower and Johnston, 2010). Despite this, a successful 3' RACE procedure allowed for the eventual sequencing of 1366 bp of the CRT cDNA from *H. midae* haemocytes. Open reading frame prediction was carried out on this sequence to determine the putative amino acid sequence of *H. midae* CRT.

The putative HM-CRT fragment contained an open reading frame coding for 281 amino acids, which had high similarity to other molluscan CRT sequences. The amino acid sequences of CRT from *A. californica* (NCBI accession S51239.1) and *P. fucata* (NCBI accession EF551334.1) are 405 and 414 amino acids in length respectively. Approximately 130 amino acids are therefore missing from the 5' end of the CRT amino acid sequence of *H. midae*. Multiple sequence alignments of calreticulin amino acid sequences from other animals indicate that this region is particularly poorly conserved and therefore further attempts to obtain the 5' end sequence using degenerate primers would yield poor results. At this time there are no complete calreticulin mRNA or deduced amino acid sequences for any members of the *Haliotis* genus available on either the NCBI database or UniProt. The unique sequence reported here was instrumental in subsequent molecular investigations conducted to characterise the transcriptional and translational regulation of this *H. midae* molecular chaperone during exposure to acute heat shock.

The cDNA sequence of *H. midae* CRT was used to design PCR primers for real time qPCR in order to quantify transcription of the CRT gene under heat shock conditions. Real time qPCR has rapidly become one of the most widely used molecular diagnostic techniques since its inception. For the purposes of this study the guidelines of Bustin et al. (2009) and the normalisation method of Pfaffl (2001) were followed. While quantitative real time PCR (or qPCR) is extensively used in model organisms, there is less information available for non-model organisms, particularly when the sample is acquired from live animals and not from a cell culture. The response to thermal stress of both the 18S rRNA gene and ribosomal factor L28 was measured using real-time qPCR. Although the 18S rRNA gene was differentially expressed during exposure to experimental conditions, ribosomal factor L28 was stably expressed and was therefore selected as a reference gene.

RT-qPCR examination of CRT expression in heat-shocked *H. midae* revealed a significant change in CRT mRNA expression soon after the animals were initially exposed to heated

water. One way ANOVA failed to detect a significant difference in CRT expression between time points, which may be attributed to the low number of replicates performed (n=4). The power of the performed test (0.050) was below the desired power (0.800), which indicates that the test is likely to not detect a difference when one actually exists (SigmaStat 3.10). Therefore a t-test was performed, comparing the control sample (0) to each experimental time point. A Mann-Whitney Rank-sum test showed that the difference between the CRT mRNA in the control group and the group sampled after 15 minutes of heat shock was significant ($p = 0.029$). This indicates a rapid, transient decrease in the transcriptional expression of CRT in *H. midae* haemocytes following exposure to acute heat shock. While the transcriptional expression of CRT is still lower compared to the control after 30 minutes of heat shock, it appears that following this early decrease the amount of *H. midae* CRT returned to the basal level.

When the endoplasmic reticulum is placed under stress and misfolded/unfolded proteins begin to accumulate, the unfolded protein response or UPR is triggered. The primary functions of the UPR are to activate the transcription of chaperones, and repress translation of proteins in the ER (Groenendyk et al., 2005). According to the literature, therefore, CRT should be transcriptionally up-regulated in response to stress (Qiu and Michalak, 2009) under the rigid control of a suite of transcription factors. Indeed, an ER stress response element (ERSE) lies adjacent to the TATA box of the calreticulin promoter in many species (Waser et al., 1997). The data obtained in this study seem to indicate that the CRT in *H. midae* is initially transcriptionally repressed during acute short-term heat stress.

It is possible that insufficient time intervals were sampled in this study to detect more significant differential expression of the CRT gene. In order to respond to environmental stress effectively, chaperones like CRT could be transcriptionally up-regulated extremely rapidly following exposure to stress. Hooper et al. (2007a) caution that the choice of time points when measuring the expression of a gene or protein experimentally may affect the appearance of the results. A study of CRT expression in a marine invertebrate reported that a 4 hour exposure to elevated temperature elicited the highest fold change of CRT (Luana et al., 2007). This may indicate that the stress response which regulates CRT transcription requires a longer exposure period to a stress than what was investigated in this study.

Having obtained a putative amino acid sequence for most of the *H. midae* CRT protein, a suitable commercially available anti-CRT antibody was purchased for use in western blot assays to quantify the expression of CRT protein in *H. midae* haemocytes during exposure to acute heat shock. A 60 kDa CRT band detected on the western blot was the appropriate size in accordance to the published literature (Coppolino and Dedhar, 1998). Although CRT is a 46.6 kDa protein, it migrates anomalously during PAGE, and consequently, appears to have an apparent molecular weight of 60 kDa, possibly due to the acidic nature of the protein (Coppolino and Dedhar, 1998).

Previously, the effect of acute heat shock on the expression of calreticulin in the haemocytes of *H. midae* was measured using iTRAQ and mass spectrometry (Chapter 3). iTRAQ analysis of four biological replicates of haemocytes sampled at 0, 30, 90 and 180 minutes from animals exposed to heated water showed that calreticulin underwent the greatest fold change in relation to the control after 90 minutes (an average $\log_2$ fold change of 0.277 vs. the control). In the present study, an increasing amount of CRT protein was detected in the heat shocked samples by western blot analysis, and after 120 minutes there was a statistically significant difference between the amount of CRT in the control and heat shocked samples (an average $\log_2$ fold change of 1.4). Although the sampling times differ slightly between the two experiments, the expression pattern of CRT detected in heat shocked abalone is the same although the magnitude of expression detected by western blot is much higher. The quantitative information obtained by iTRAQ for calreticulin has therefore been confirmed using a complementary molecular technique. In addition, the data confirms that CRT is regulated in response to stress in *H. midae*, and that exposure to acute heat stress for as little as 120 minutes may cause ER stress which leads to the triggering of the UPR.

Chaperones such as CRT are up-regulated during ER stress in an attempt to manage the stress and return the cell to homeostasis (Qiu and Michalak, 2009). During ER stress, chaperones are needed in greater quantities because stress causes an imbalance between cellular needs and the capability of the cell to produce properly folded proteins (Ni and Lee, 2007). When ER protein folding capability is compromised the UPR is activated, which enhances the cell's ability to produce properly folded proteins during exposure to stressful conditions (Bateman, 2006). Increased expression of CRT was detected in *C. elegans* following exposure to the UPR-inducing chemical tunicamycin (Lee et al., 2007), indicating that the role of CRT is linked to the UPR and therefore to the cellular stress management machinery. If the UPR is unable to

return the cell to homeostasis, activation of the apoptotic pathway is signalled and the cell is targeted for degradation. The data presented here indicates that exposing abalone to 23°C water for approximately 2 hours is sufficient to induce the expression of CRT in haemocytes, although it is not clear whether its expression is enhanced via the UPR or another stress response pathway. CRT performs several other cellular functions besides its role in folding mis/unfolded glycoproteins (Gelebart et al., 2005; Qiu and Michalak, 2009). Primarily, the ability of CRT to bind Ca^{2+} with a high capacity is vital in a cell undergoing stress as the Ca^{2+} ion is implicated in the regulation of many cellular processes (Groenendyk et al., 2005).

An increase in the amount of CRT protein leads to an increase in the amount of Ca^{2+} in the cell (Qiu and Michalak, 2009). The influx of Ca^{2+} is regulated with the assistance of intermediary proteins which bind the free Ca^{2+} to calcium binding proteins (Nikapitiya et al., 2010). In gastropod phagocytic cells, such as haemocytes, regulation of Ca^{2+} has been associated with phagocytic activity through the activity of calcium binding proteins such as calmodulin (Nikapitiya et al., 2010). The role of CRT in the regulation of the Ca^{2+} signalling cascades in abalone haemocytes is unknown. Although molecular regulation of calcium signalling is important to the stress response of the mollusca, this field remains underrepresented in the literature.

4.4.1. Conclusion and future work

Heat induced transcriptional and translational regulation of an abalone calreticulin is reported here. iTRAQ analysis had determined that CRT was differentially expressed during *H. midae* exposure to acute heat shock. However, this data could not be verified because there was no amino acid sequence available for abalone CRT, limiting further characterisation of this protein. Sequence information for the *H. midae* CRT gene was obtained, and ORF prediction software was used to determine the putative CRT amino acid sequence. Gene specific primers were designed for quantitative real-time PCR analysis of calreticulin gene expression during exposure to heat shock. Following this, a commercially available polyclonal CRT antibody was used to quantify CRT protein expression in haemocytes from heat stressed abalone. Interestingly, CRT mRNA expression decreased transiently following brief exposure to heat shock, and subsequently returned to basal levels; whereas CRT protein expression increased after 30 minutes and remained elevated even after a 24 hour recovery period. This indicates that the transcriptional and translational regulatory mechanisms for CRT in *H. midae* may be regulated temporally by different systems.

In addition to providing insight into the expression of an important chaperone, and calcium binding protein, this data also validates the iTRAQ data obtained in Chapter 3. Expression analysis of CRT protein by iTRAQ and western blot both showed similar regulation over the same time scale. Validation is important for proteomics based studies, especially where there is no species-specific database available. This study represents one of the first to describe calreticulin expression in abalone. Together, this information may lead to a better understanding of the abalone molecular stress response. The unfolded protein response (UPR) is implicated as a stress response mechanism in abalone exposed to acute heat stress, as CRT is known to be induced by conditions that trigger the UPR. This may also indicate that acute heat stress triggers the UPR, which can lead to apoptosis, suggesting that cellular damage caused by fluctuating temperature on abalone farms should be of concern to farmers.

Further research on the signalling mechanisms responsible for co-ordinating the abalone's cellular defence against stresses such as heat shock will be valuable. Furthermore, the capability of abalone to cope with stress conditions can be linked to the expression of chaperones like calreticulin. Further characterisation of chaperones involved in the maintenance of cellular homeostasis, and their response to the abiotic stresses common to

abalone farms, would provide further insight into the molecular processes that link the abalone stress response and immune system.

Chapter 5
GENERAL DISCUSSION

Mass mortality of farmed abalone following the outbreak of viral or bacterial pathogens threatens the sustainability of this economically important industry. One of the factors contributing to the severity of disease outbreaks is the poor understanding of the molecular basis of the abalone stress response. Investigations into the stress response mechanisms of abalone have postulated a link between stress and increased susceptibility to disease, with elevated temperature implicated as a potential immunosuppressive factor.

The tank water temperature on abalone farms fluctuates, and it is these changes in water temperature that are associated with the increased death rate in farmed abalone known as 'summer mortality'. Therefore it would be beneficial to investigate the effect that acute elevated temperature has on the abalone molecular stress response. This information could also assist in the optimisation of abalone production conditions to minimise mortality caused by opportunistic infection.

Modern proteomics technology has the capacity to identify differentially regulated proteins in a complex sample, but to date has been underutilised in non-model organisms like the abalone. The advantage of proteomics is its capacity to identify and quantify hundreds of proteins simultaneously, imparting a holistic impression of the molecular mechanisms at work in an organism. One of the primary reasons that aquaculture research has been slow to adopt proteomics technology is that protein identification depends on the availability of a genomic database. Many of the most commonly farmed species lack sequence information for most of their genome, and therefore protein identification by mass spectrometry can be challenging. However, with improved mass spectrometer sensitivity and computational strategies, proteomic analysis is becoming an appealing option for the assessment of environmental perturbation in non-model organisms.

The aim of this study was to achieve a better understanding of the abalone heat induced stress response by identifying proteins that are differentially regulated in *H. midae* during exposure to acute heat stress using proteomics technology.

Initially, an appropriate temperature for heat shock was determined for *H. midae* using total haemocyte counts (Chapter 2). A heat shock water temperature of 23 °C, or 10°C above the ambient water temperature, was chosen. The duration of the detectable molecular response to heat shock was determined using 2D PAGE analysis. Exposure to 23°C water for 3 hours was found to illicit a detectable response at the protein level, without permanent injury to the abalone following exposure. Analysis of protein expression in a sample taken from abalone that had been allowed to recover from exposure to heat shock for 24 hours indicated that the cellular response to heat stress persisted in this group. This information confirmed that exposure to 23°C water elicited a response in haemocytes that was detectable using 2D PAGE.

Having determined appropriate conditions to instigate a stress response in *H. midae*, the proteomic techniques 2D PAGE and iTRAQ were employed to distinguish proteins that were differentially expressed during exposure to heat shock. In so doing, several putative stress response biochemical pathways have been identified that would be of interest for future investigations of abalone. A generalised schematic for the hypothesised mechanism of the *H. midae* heat stress response based on this data is presented in Figure 5-1. This hypothesised stress response mechanism is based on the proteins which responded to the experimental conditions used in this study (Chapters 2 and 3). Although some of the identified proteins have been characterised in other molluscs, there is still a lack of information available for many of these proteins in abalone. A discussion of these stress response pathways as presented in the diagram will follow, with reference to existing molluscan studies relating to those pathways.

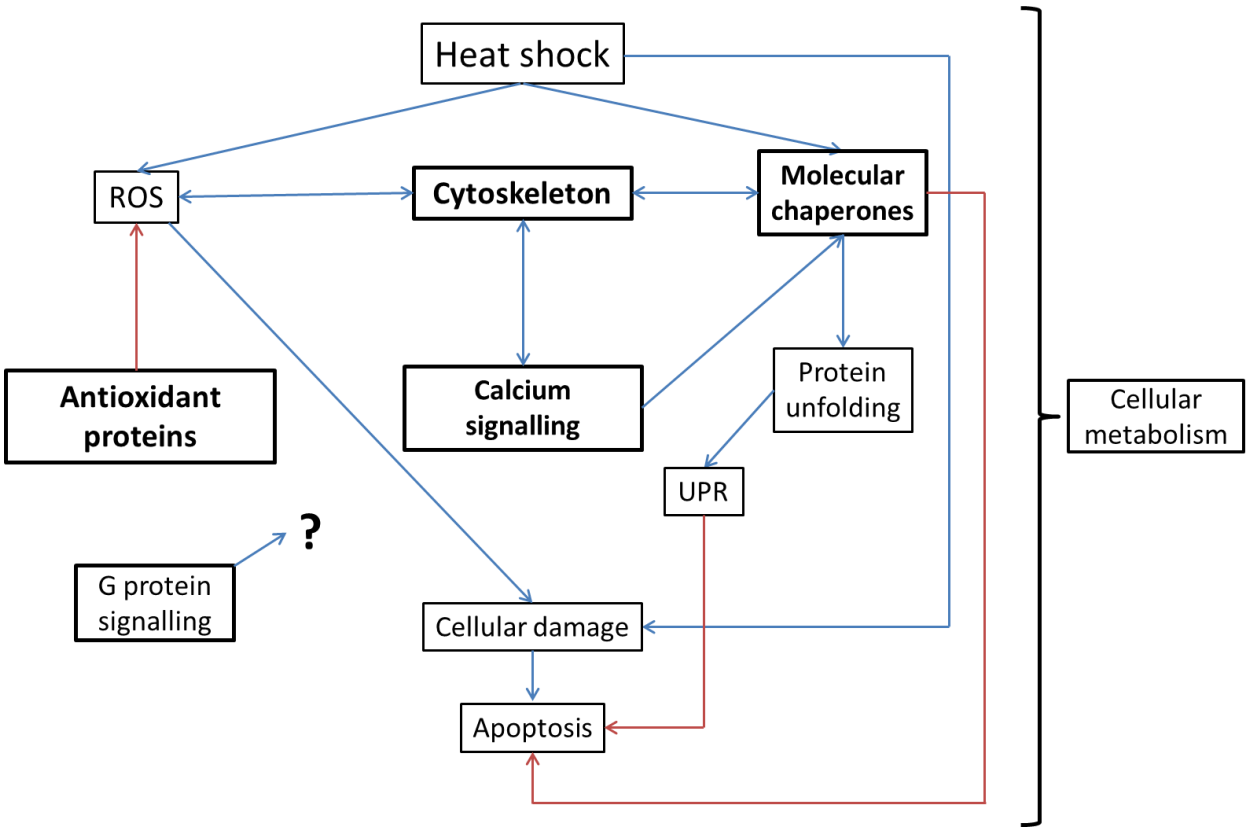


Figure 5-1: Proposed schematic of the hypothesised *H. midae* molecular heat stress response, based on proteomic data obtained using 2D PAGE and iTRAQ. The blue and red arrows represent induction and repression of molecular pathways during exposure to acute heat shock, respectively. Literature contributing to this schematic has been presented in the body of this chapter and in relevant chapters of this thesis. Although G-protein signalling is hypothesised to contribute to the stress response based on experimental data from this study, the mechanism of this interaction remains unknown and is therefore represented by a question mark. The regulation of cellular metabolism affects most cellular activities, and is therefore represented by a bracket on this schematic diagram.

2D PAGE analysis of proteins extracted from heat shocked *H. midae* haemocytes identified 5 differentially regulated proteins with cytoskeletal function (Chapter 2). This suggests that cytoskeletal remodelling may be an effect of acute heat shock. Cytoskeletal proteins are commonly detected in proteomic stress response studies, as they are both abundant in the cell and well represented in the protein sequence database. Investigations into the effects of thermal stress in mussels have led to the hypothesis that cytoskeletal elements may be especially sensitive to changes in temperature (Tomanek and Zuzow, 2010; Fields et al., 2012a), possibly as a result of damage caused by oxidative stress (Sheehan and McDonagh, 2008). However there is also evidence to suggest that the cytoskeleton does not merely respond passively to oxidative stress, but rather plays an active role in the regulation of the stress response (Dalle-Donne et al., 2001). Immunosuppression in stressed abalone has been

reported in abalone following exposure to a range of abiotic stresses (Cheng et al., 2004a, 2004b; Hooper et al., 2007a). Immunosuppression in Australian abalone following exposure to severe heat stress (10 °C above ambient for one week) has recently been documented by Hooper et al. (2014). A decrease in phagocytic ability has been detected in the oyster *Crassostrea virginica* following exposure to sudden temperature elevation (Hégaret et al., 2003a). The apparent disruption of the abalone haemocyte cytoskeleton by heat stress detected in this study suggests a physical mechanism by which stress could be linked to immunosuppression, as the molluscan immune response depends in part on the cytoskeleton-dependent phagocytic activity of haemocytes. The role of some heat induced molecular chaperones includes stabilisation of the cytoskeleton, although the molecular mechanisms by which these proteins confer thermotolerance are not fully understood (Wettstein et al., 2012).

The heat shock response of *H. midae* was also investigated using iTRAQ labelling and MS analysis which identified and quantified the expression of 116 abalone haemocyte proteins (Chapter 3). Of these, 11 proteins exhibited statistically significant changes in expression in heat shocked abalone. Differentially regulated proteins were also identified using cluster analysis of the average heat induced expression profiles of these proteins. GO analysis of the differentially regulated proteins suggested that biochemical pathways relating to stress response, cell signalling and cellular metabolism were particularly affected by exposure to heat shock. The complexity and interconnectedness of the affected biochemical pathways in this case was further complicated by the relative lack of functional studies in molluscs for many of the differentially regulated proteins.

The molluscan oxidative stress response has previously been studied in a number of species. Although monitoring of oxidative stress biomarkers would be beneficial in most aquatic systems, the antioxidant network in aquatic organisms is extremely complex (Regoli and Giuliani, 2013) and the mechanisms of ROS production by abalone haemocytes has not yet been sufficiently researched (Hooper et al., 2007a). De Zoysa et al. (2009) characterised the oxidative stress response of *H. discus discus* by measuring transcriptional regulation of various antioxidant genes. The antioxidant capacity of *H. midae* under various oxygen levels and water temperatures was measured by determining the activity of superoxide dismutase, glutathione peroxidase and catalase using commercially available kits (Vosloo et al., 2013a). Although some antioxidant genes have been characterised in abalone at the transcriptional level (De Zoysa et al., 2008, 2009; Ekanayake et al., 2008; Pushpamali et al., 2008), proper

characterisation of the abalone antioxidant response at the protein level is currently lacking. In this study, we determined that several proteins involved in the oxidative stress response were differentially regulated during acute heat stress, again indicating that heat shock causes oxidative stress in the haemocytes. As oxidative stress is thought to be a major contributing factor to cytoskeletal remodelling in molluscs (Sheehan and McDonagh, 2008), this corroborates the 2D PAGE data. Although many proteins involved in the molluscan oxidative stress response have been identified, in many cases their exact molecular function awaits clarification (Loker, 2010). Omega crystallin, for example, was detected by iTRAQ but the only available molluscan sequence for identification of this protein in the database was a squid eye lens protein. However, molluscan aldehyde dehydrogenase is homologous to omega crystallin in scallops (Horwitz et al., 2006). Aldehyde dehydrogenase has previously been detected in mussels exposed to environmental pollutants (Zhou et al., 2010), and is therefore a novel and interesting potential target for future research.

As well as these oxidative stress proteins, molecular chaperones including heat shock protein 70 (Hsp70) were differentially expressed during exposure to heat shock. Hsp70 is a well-known stress inducible protein which has been characterised in other species of abalone (Cheng et al., 2007; Farcy et al., 2007) and previously used as a molecular stress indicator in *H. midae* (Vosloo and Vosloo, 2010). While stressed, the cell can become overwhelmed by the accumulation of misfolded proteins. Molecular chaperones, like heat shock proteins, are responsible for arranging nascent proteins in their native state, and for re-folding incorrectly folded proteins that may otherwise form toxic aggregates (Bateman, 2006). The cellular machinery responsible for the removal of misfolded proteins during ER stress is known as the unfolded protein response (UPR). During exposure to heat stress, increased heat shock protein expression confers thermotolerance and allows the stressed cell to attempt to return to homeostasis (Kiang and Tsokos, 1998). These and other stress response pathways are under the control of receptors that are capable of detecting stress and initiating the appropriate response (Lacoste et al., 2001). In order to co-ordinate the complex cellular processes required for restoring homeostasis, cells exposed to stress make use of an interconnected network of molecular signalling pathways.

Cell signalling activity in heat stressed *H. midae* appears to be mediated in part through calcium signalling and small G proteins. The concentration of calcium in intracellular spaces such as the endoplasmic reticulum affects a number of cellular functions. To further our

understanding of the abalone stress response, and facilitate future molecular studies, a known stress response protein was selected for further characterisation and expression analysis in *H. midae* (Chapter 4). Calreticulin appears to be differentially regulated during heat shock according to the iTRAQ analysis reported in Chapter 3, and consequently, the heat stress induced expression profile of CRT was investigated using alternate molecular methods in order to validate the iTRAQ data. CRT is a molecular chaperone and calcium carrier molecule that has been implicated in the stress response in a number of animals. CRT currently lacks an abalone DNA or protein sequence, and it was therefore first necessary to determine the cDNA sequence for *H. midae* CRT. A 1366 bp portion of *H. midae* CRT was successfully sequenced, which allowed for a partial amino acid sequence to be deduced for *H. midae* CRT. This sequence had CRT family signature repeat motifs, and significant homology to CRT from other molluscs in the NCBI database. Using the cDNA sequence, gene specific primers were synthesised so that real-time qPCR could be used to determine the effects of heat stress on the transcription of the CRT gene in *H. midae* haemocytes. Interestingly, the amount of CRT mRNA appeared to decrease after 15 minutes of exposure to heat stress, and thereafter returned to basal levels. In contrast, the amount of *H. midae* CRT protein increased while the abalone were exposed to heat stress. CRT is part of the cellular stress response mechanism that makes up the unfolded protein response in the endoplasmic reticulum, and therefore its differential regulation may indicate that the ER is under stress due to the accumulation of misfolded proteins (Schröder, 2008). An increase in CRT protein directly influences the cell's Ca^{2+} buffering capacity (Qiu and Michalak, 2009), which supports the REViGO analysis data and suggests that calcium signalling plays a role in the abalone heat shock response. CRT appears to be reproducibly regulated in response to acute heat shock in *H. midae*, and may prove to be a useful biomarker for monitoring the stress status of abalone.

The effects of stress on the metabolism of abalone and other molluscs should be of particular interest to farmers, as it has a direct influence on the growth rate and therefore profitability of farmed abalone. The disruption of ATP production by abiotic stresses such as heat shock would negatively impact the ability of abalone haemocytes to perform their immune function (Coyne, 2011). The detection of several differentially regulated proteins involved in the production of cellular energy in the form of ATP, or proteins that depend on the availability of cellular ATP, indicate that abalone metabolism is affected by exposure to acute thermal stress. Vosloo et al. (2013b) have studied the metabolic effect of long-term (one month) exposure to 19 °C water in *H. midae* and concluded that while the animals can compensate for

the increased energy demands during stress, their growth is likely to be affected. This parallels the findings of Fields et al. (2012b), who concluded that regulating oxidative stress caused by exposure to acute heat shock forces a trade-off in cellular metabolism at the cost of ATP production in the blue mussel. In order to gain a better understanding of the metabolic costs of temperature stress, we now need to measure metabolic changes alongside protein expression in abalone. This type of investigation will have broader application in the future, as anthropogenically induced climate change may significantly alter the water temperature of coastal marine systems (Harley et al., 2006).

Due to the exploratory nature of this research, some validation of the iTRAQ results was deemed to be necessary in order to confirm that the experimental design was sufficiently stringent. The iTRAQ expression data would be deemed to have been validated if the heat induced expression profiles of proteins could be confirmed using a complementary molecular technique. Unfortunately, although the need for 'minimum information about a proteomics experiment' (MIAPE) guidelines were drafted in 2007 (Taylor et al., 2007), there is still no defining documentation that provides guidelines for the validation of an iTRAQ proteomics experiment. Western blot analysis was carried out using commercially available polyclonal antibodies to proteins identified by iTRAQ. Since there are no available antibodies to *Haliotis* proteins specifically, antibodies were selected where the antigen showed high (>60%) homology to known molluscan protein sequences for that protein. Since there were no suitable antibodies available for many of the significant proteins identified by iTRAQ, validation was only attempted for three proteins: Hsp70, G6P1D and GSTO1. Analysis of Hsp70 expression in heat shocked haemocytes by western blot revealed the same expression pattern as the iTRAQ analysis. Western blot analysis of the G6P1D response to heat shock determined that the protein level increased by a similar order of magnitude as determined by iTRAQ, although the western blot detected an increase in G6P1D expression much earlier than the iTRAQ. Western blot analysis of GSTO1 expression was problematic, as the molecular weight of the detected protein was much larger than expected, and the expression profile differed from the iTRAQ expression profile for this protein. Based on the mixed success of this validation experiment, we determined that validation of iTRAQ data by western blot depends largely on the availability of suitable primary antibodies for the target protein. If the target protein sequence is unknown in abalone, western blotting can be prone to error. Future investigations should therefore take care when selecting appropriate antibodies for validation of proteomic data.

H. midae has an innate immune system which is considered evolutionarily ancient compared to the acquired immune system of more complex animals. It is therefore tempting to consider their immune system as simple and primitive. However, the persistence of the innate immune system throughout evolutionary history is a testament to its efficacy and reliability. Although abalone are susceptible to diseases under farming conditions, their immune system is nonetheless a sophisticated and highly effective molecular mechanism that is capable of responding to stresses rapidly and efficiently. This study has identified several proteins that respond to acute heat shock in *H. midae* haemocytes. Based on this data, the overwhelming characteristic of the abalone cellular response to heat shock is its complexity and interconnectedness. Although acute heat shock may cause damage to the haemocytes, and potentially lead to immunosuppression, abalone appear to be capable of withstanding exposure to the stress for a limited period of time. It may be of concern to abalone farmers that changes in temperature may leave abalone vulnerable to opportunistic pathogenic attack until they are able to repair cellular damage. Furthermore, it appears that the metabolic cost of surviving stressful conditions could hinder the growth rate of *H. midae*. Ideally a farming process that controls water temperature could reduce the incidence of mortality, thereby improving the output of abalone farms.

To our knowledge, this study is one of the first to report the use of proteomic technology to characterise the effect of environmental stress on a mollusc. iTRAQ is currently a novel tool for identifying and quantifying proteins involved in the stress response of abalone since the method has thus far been limited to investigation of model organisms. As more genomic sequences become available for cultivated seafood species, proteomics will have the capability to realise its enormous potential in aquaculture research. In the iTRAQ workflow presented here, many potentially biologically relevant protein identities were discarded because they were not detected in both iTRAQ experiments. This is a consequence of the increased stringency employed in this study which ensured that fewer erroneous protein identities were made. This strategy was adopted as guidelines for a study such as this on a non-model organism have not been established.

As aquaculture is a rapidly growing means of food production, it is increasingly important that we understand the molecular mechanisms by which these animals respond to environmental stress. The tools reported in this study will be useful for future molecular investigations as

they are applicable to a number of important aquacultured species. A proteomic investigation into the effect of anthropogenic global warming using 2D PAGE has already been carried out on mussels (Fields et al., 2012b). Investigations into the effect of thermal stress will have growing significance in the future, as global climate change begins to affect the temperature of the oceans.

In conclusion, this study fulfilled the research objectives outlined in Chapter 1 by i) identifying proteins that are differentially regulated during exposure to acute heat stress, which ii) identified biochemical pathways which may be involved in the abalone heat stress response. This information provides as an initial insight into the effect of stress on abalone, which contributes to our understanding of the molecular basis of the abalone stress response. A better understanding of the abalone stress response may inform abalone farming practices, so that the productivity of abalone farms may increase. The use of proteomics in this investigation provided a platform for future studies of important cultivated non-model organisms. This work illuminates only a portion of the complex network of proteins that make up the abalone stress response, which we hope to one day understand fully.

Appendix A

Solutions

Anti-coagulation buffer

0.45 M NaCl

0.1 M Glucose

30 mM Sodium citrate

26 mM Citric acid

10 mM EDTA

Adjust pH to 4.6

Blocking buffer

5% w/v skim milk powder in 1x TBS

Coomassie Brilliant blue stain

0.5 g Coomassie R250

225 ml dH2O

225 ml methanol

50 ml acetic acid

DEPC-dH20

1000 ml dH_2O

1 ml DEPC

Stand overnight, then autoclave.

Destain solution

250 ml methanol

100 ml acetic acid

650 ml d H2O

10% polyacrylamide gel (separating gel)

2.5 ml Tris HCl SDS pH 8.8

2.49 ml 37% Acrylamide

50 ul Ammonium Persulphate

10 ul TEMED

4.95 ml dH_2O

4% polyacrylamide gel (stacking gel)

0.5 ml Tris HCl SDS pH 6.8

0.4 ml 37% Acrylamide

25 ul Ammonium Persulphate

2 ul TEMED

3.073 ml dH_2O

Ponceaus stain

0.1% w/v Ponceaus powder

5% v/v Acetic acid

Make up to 500 ml with dH_2O

Protein Lysis Buffer

200 mM NaCl

1 mm EDTA

0.1% Triton X100

1 mM PMSF

1 mM DTT

20 mM Tris HCl pH 8.0

Rehydration Buffer

5.6 ul ASB14 (10% stock -20°C)

7.0 ul Biorad Amphylytes (at 20% stock)

9.3 ul DTT (30% stock -20°C)

(250ug) Isolated total protein

ULB (up to 140ul)

5 X Sample application buffer

1 M Tris-HCl pH 6.8

50% glycerol

25% SDS

5 mg bromophenol blue

21% w/v dithiothreatol (DTT)

Solution D

4 M Guanidium thiocyanate

25 mM Sodium Citrate

0.5% Sarkosyl

Filter sterilize before adding:

70 ul β-Mercaptoethanol

Make up to 10 ml with DEPC-dH2O

5 X SDS electrophoresis buffer

15.1 g Tris

72.0 g Glycine

5.0 g SDS

Make up to 1000 ml with dH_2O

Towbin buffer

6.06 g Tris

28.84 g Glycine

400 ml Methanol

Make up to 2 litres with dH_2O

Tris buffered saline solution (TBS) (10x stock)

50 mM Tris

150 mM NaCl

300 ml d H_2O

Adjust pH to 7.4 with HCl

Make up to 500 ml with dH_2O

Autoclave, dilute 1:10 in dH_2O to make 1X TBS

TBST

1x TRIS containing 0.1% v/v TWEEN 20

Tricene EDTA Buffer

10mM Tris-Cl, pH 8.0

1mM EDTA

Urea Lysis Buffer

24 g Urea

7.612 g Thiourea

1.0 g CHAPS

15 ml dH2O

Heat at 35°C to aid dissolving

Once dissolved, make up volume to 50 mls with dH2O.

Aliquot out into 2 ml/1.5 ml eppendorf tubes and store at -20°C

Appendix B

Supplementary data

Contents:

B1 2D PAGE scanned gel images

The following images were used in the Melanie analysis for 2D PAGE in Chapter 2 in order to identify heat stress regulated proteins in *H. midae* haemocytes.

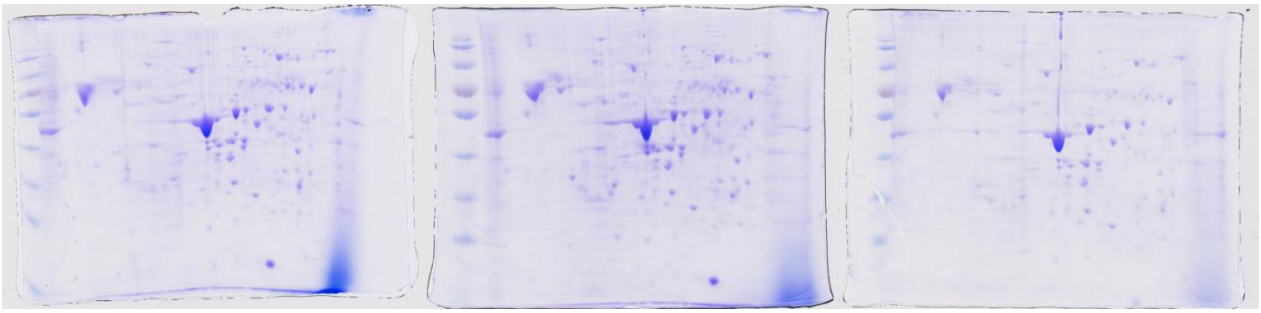


Figure B-1: 2D PAGE gel images for biological replicate #1 for protein isolated from abalone haemocytes sampled at time 0 (control), images from left to right are technical replicate 1, 2 and 3.

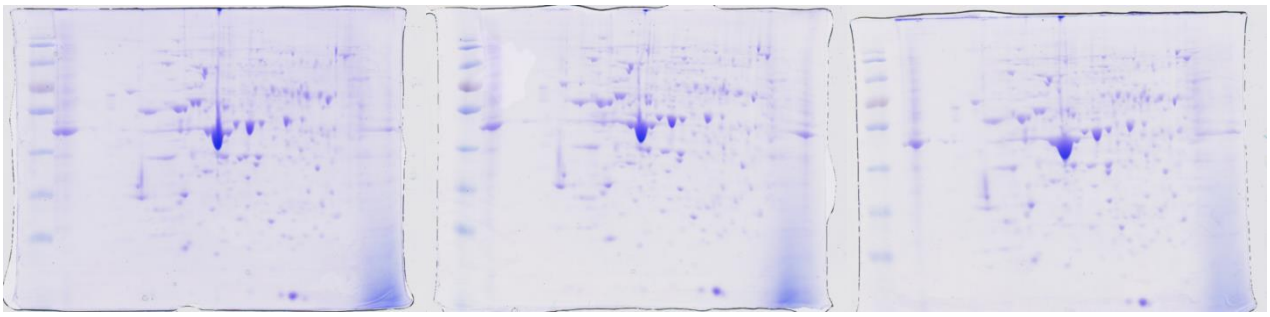


Figure B-2: 2D PAGE gel images for biological replicate #1 for protein isolated from abalone haemocytes sampled at 30 minutes, images from left to right are technical replicates 1, 2 and 3.

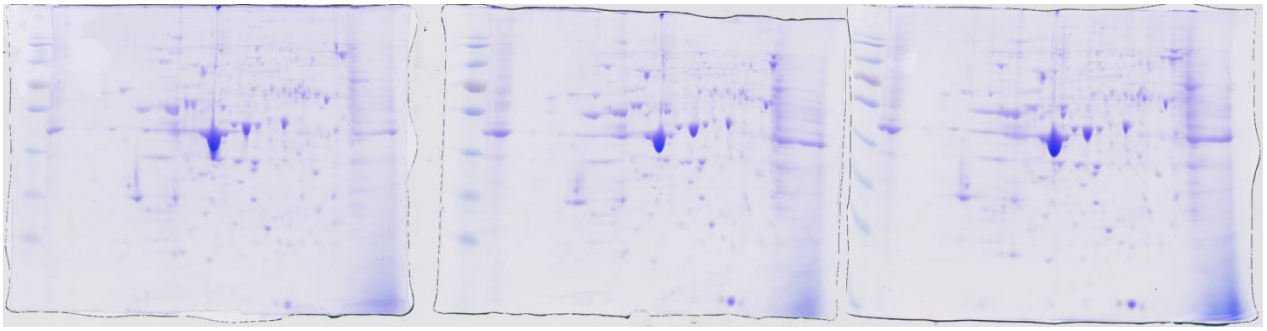


Figure B-3: 2D PAGE gel images for biological replicate #1 isolated from abalone haemocytes sampled at 180 minutes, images from left to right are technical replicates 1, 2 and 3.

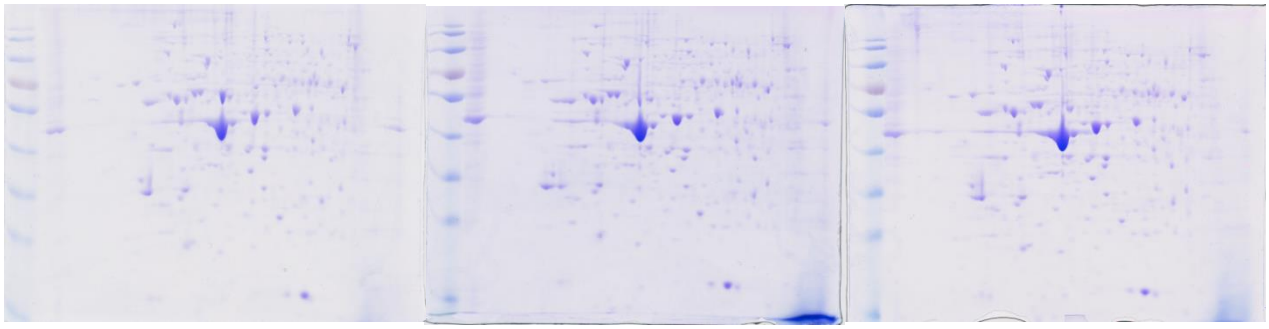


Figure B-4: 2D PAGE gel images for biological replicate #2 isolated from abalone haemocytes sampled at time 0 (control), images from left to right are technical replicates 1, 2 and 3.

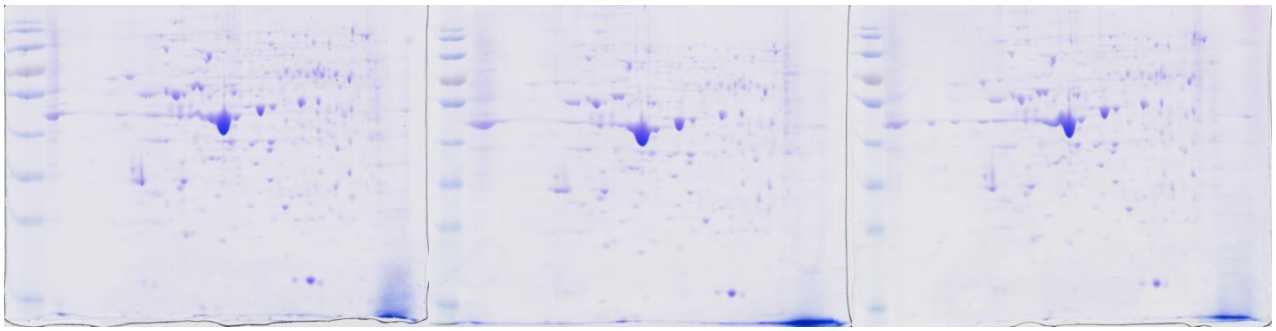


Figure B-5: 2D PAGE gel images for biological replicate #2 isolated from abalone haemocytes sampled at 30 minutes, images from left to right are technical replicates 1, 2 and 3.

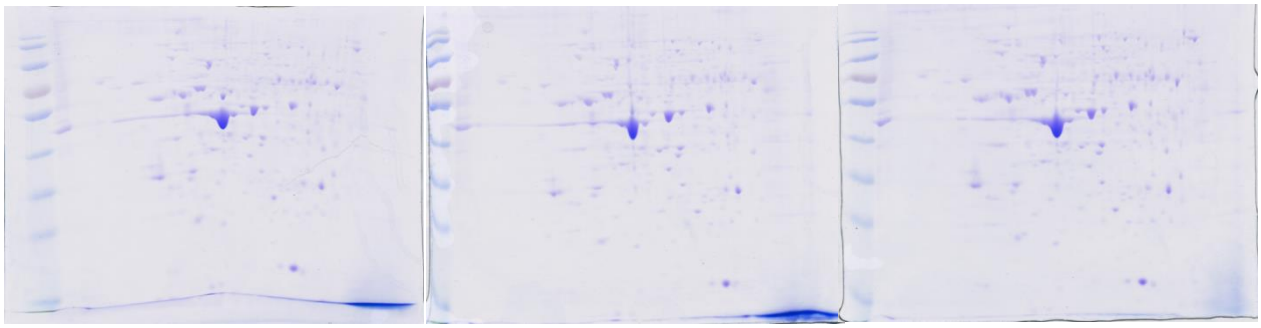


Figure B-6: 2D PAGE gel images for biological replicate #2 isolated from abalone haemocytes sampled at 180 minutes, images from left to right are technical replicates 1, 2 and 3.

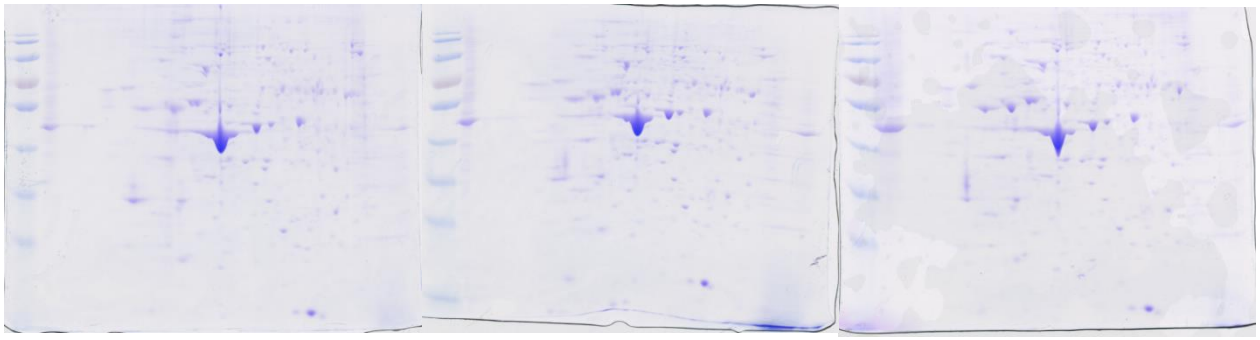


Figure B-7: 2D PAGE gel images for biological replicate #3 isolated from abalone haemocytes sampled at time 0 (control), images from left to right are technical replicates 1, 2 and 3.

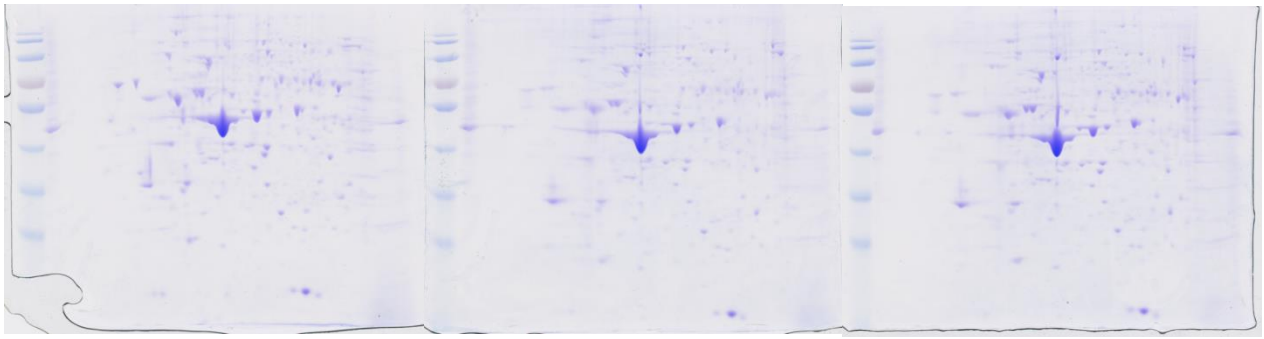


Figure B-8: 2D PAGE gel images for biological replicate #3 isolated from abalone haemocytes sampled at 30 minutes, images from left to right are technical replicates 1, 2 and 3.

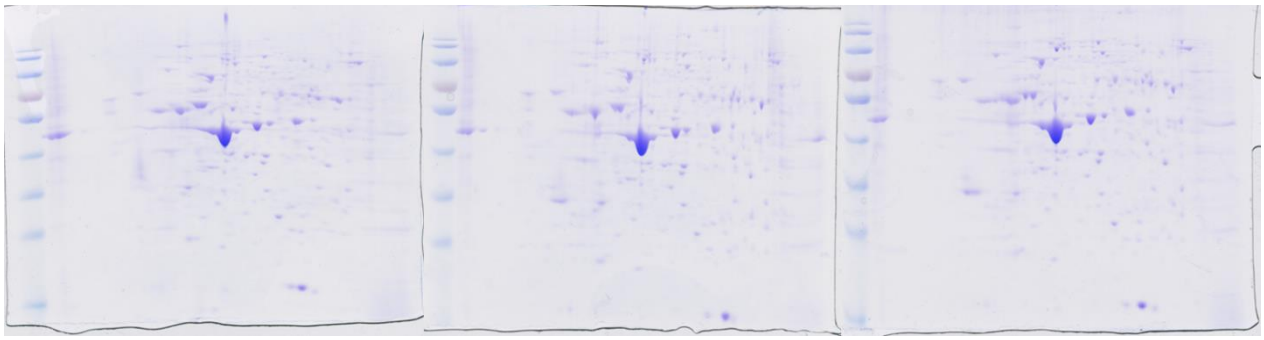


Figure B-9: 2D PAGE gel images for biological replicate #3 isolated from abalone haemocytes sampled at 180 minutes, images from left to right are technical replicates 1, 2 and 3.

B2 - iTRAQ supplementary protein identification information

Table B-1: PEAKS quantitative data output for proteins that were in common between two independent iTRAQ experiments. The information alternates for each protein identity in iTRAQ #1 and iTRAQ #2. The accession number and description of each protein identity refers to the database of molluscan sequences compiled using UniProt, which was used for peptide spectrum matching.

Description	Accession number	-10logp iTRAQ1	-10logp iTRAQ 2	Coverage iTRAQ1 (%)	Coverage iTRAQ 2 (%)	# peptides iTRAQ 1	# peptides iTRAQ 2	#Unique peptides iTRAQ 1	#Unique peptides iTRAQ 2	Mass iTRAQ 1	Mass iTRAQ 2
14-3-3 protein OS=Chlamys farreri PE=2 SV=1	G8A4W9|G8A4W9_9BIVA	109.88	63.84	14	16	7	6	7	6	30598	30598
40S ribosomal protein S12 OS=Pinctada maxima PE=2 SV=1	C6JUN0|C6JUN0_PINMA	68.7	49.1	18	11	3	2	2	2	14992	14992
40S ribosomal protein S16 (Fragment) OS=Aplysia californica PE=2 SV=1	Q8T6A2|Q8T6A2_APLCA	56.7	48.76	33	24	4	3	4	3	8704	8704

40S ribosomal protein S18 OS=Haliotis discus discus PE=2 SV=1	B6RB81|B6RB81_HALDI	97.29	68.5	22	27	5	5	3	5	17925	17925
40S ribosomal protein S3a (Fragment) OS=Haliotis diversicolor PE=2 SV=1	B3TK56|B3TK56_HALDV	111.13	50.87	37	19	8	4	5	4	19514	19514
40S ribosomal protein S8 OS=Haliotis diversicolor PE=2 SV=1	B3TK66|B3TK66_HALDV	175.81	71.86	28	29	5	6	5	6	23987	23987
40S ribosomal protein S9 OS=Corbicula fluminea GN=rsp9 PE=2 SV=1	Q32ZL3|Q32ZL3_9BIVA	66.13	37.71	20	19	6	5	2	5	22318	22318
40S ribosomal protein SA OS=Pinctada fucata PE=2 SV=1	A3RLT6|RSSA_PINFU	175.38	128.85	32	34	11	12	11	9	33520	33520
60S acidic ribosomal protein P0 (Fragment) OS=Haliotis diversicolor PE=2 SV=1	B3TK58|B3TK58_HALDV	175.53	115.12	48	28	13	8	11	8	28534	28534

6-phosphogluconate dehydrogenase_ decarboxylating (Fragment) OS=Pisidium adamsi GN=PGD PE=2 SV=1	Q8T7M6|Q8T7M6_9BIVA	44.71	49.93	39	23	5	7	5	4	19889	19889
71kDa heat shock protein OS=Haliotis tuberculata GN=hsp71 PE=2 SV=1	Q17UC1|Q17UC1_HALTU	312.14	218.71	84	47	103	38	62	11	71357	71357
84kDa heat shock protein OS=Haliotis tuberculata GN=hsp84 PE=2 SV=1	Q17UC2|Q17UC2_HALTU	257.14	139.33	44	30	39	25	3	8	84015	84015
Actin (Fragment) OS=Physella parkeri PE=3 SV=1	Q6RYC7|Q6RYC7_9GAST	263.28	201.29	72	75	40	29	3	4	22886	22886
Actin depolymerisation factor/cofilin OS=Haliotis diversicolor GN=ADF PE=2 SV=1	B3SND4|B3SND4_HALDV	98.36	63.85	34	25	10	6	9	5	18147	18147

Actin OS=Crassostrea gigas GN=act PE=2 SV=1	C4NY64|C4NY64_CRAGI	337.47	235.29	90	66	109	45	9	2	41806	41806
Adenosylhomocysteinase OS=Crassostrea ariakensis PE=2 SV=1	G8XUN4|G8XUN4_CRAAR	37.29	32.17	6	8	3	4	3	4	47529	47529
ADP/ATP carrier protein (Fragment) OS=Haliotis diversicolor PE=2 SV=1	B3TK64|B3TK64_HALDV	105.92	88	18	33	5	8	4	5	16707	16707
ADP-ribosylation factor 2 OS=Haliotis discus discus PE=2 SV=1	B6RB72|B6RB72_HALDI	184.98	113.38	58	40	12	7	12	6	20940	20940
Allograft inflammatory factor (Fragment) OS=Haliotis diversicolor GN=AlInFa PE=2 SV=1	B3SN77|B3SN77_HALDV	85.46	63.4	24	21	4	4	3	2	15608	15608
Alpha-actinin (Fragment) OS=Biomphalaria glabrata PE=2 SV=1	Q8ITH0|Q8ITH0_BIOGL	257.74	208.01	42	43	38	42	33	38	89505	89505

Arginine kinase OS=Haliotis madaka PE=2 SV=1	P51544|KARG_HALMK	145.45	98.64	37	23	18	14	10	6	39814	39814
ATP synthase subunit alpha OS=Pinctada fucata PE=2 SV=1	Q000T7|Q000T7_PINFU	144.31	89.91	40	21	22	14	20	11	59680	59680
ATP synthase subunit beta (Fragment) OS=Haliotis rufescens PE=2 SV=1	Q45Y90|Q45Y90_HALRU	281.37	220.96	85	69	58	23	38	4	46210	46210
ATP synthase_ H+ transporting_ mitochondrial F1 complex_ o subunit OS=Haliotis discus discus PE=2 SV=1	B6RB43|B6RB43_HALDI	88.68	67.59	30	43	7	10	7	9	19478	19478
ATP-dependent RNA helicase DDX5 OS=Aplysia californica PE=2 SV=1	E5F064|E5F064_APLCA	131.04	36.23	9	7	8	4	6	4	66554	66554

Calcineurin B OS=Haliotis discus discus PE=2 SV=1	B6RB21|B6RB21_HALDI	103.63	81.77	26	27	4	4	4	3	19329	19329
Calcium-binding EF-hand protein (Fragment) OS=Euprymna scolopes PE=2 SV=1	Q68LN2|Q68LN2_9MOLL	98.47	54.09	86	54	5	4	2	2	6773	6773
Calmodulin OS=Haliotis diversicolor GN=CaM PE=2 SV=1	B3SND3|B3SND3_HALDV	153.36	131.31	50	46	9	10	6	8	16838	16838
Calmodulin-dependent protein kinase (Fragment) OS=Haliotis diversicolor PE=2 SV=1	B3TK60|B3TK60_HALDV	131.89	112.87	44	23	9	7	9	6	23608	23608
Calreticulin OS=Crassostrea gigas GN=CRT PE=2 SV=1	A5LGG9|A5LGG9_CRAGI	115.9	69.15	23	13	11	7	8	3	48187	48187
Cdc24-like protein OS=Biomphalaria	E6Y9P6|E6Y9P6_BIOGL	168.3	96.49	24	28	6	5	5	4	21234	21234

glabrata PE=2 SV=1											
Cdc42 OS=Aplysia californica PE=2 SV=1	Q86DH9|Q86DH9_APLCA	168.3	96.49	36	32	8	6	7	5	21407	21407
Chaperonin containing T-complex polypeptide subunit zeta OS=Haliotis discus hannai PE=2 SV=1	H9AWU2|H9AWU2_HALDH	164.71	113.55	28	18	12	9	11	6	58458	58458
Chaperonin containing tcp1 OS=Haliotis discus discus PE=2 SV=1	B6RB18|B6RB18_HALDI	191.15	101.85	33	20	19	10	16	7	59257	59257
Elongation factor 1-alpha (Fragment) OS=Haliotis rufescens PE=2 SV=1	Q45Y88|Q45Y88_HALRU	222.51	183.41	43	49	22	22	4	4	45146	45146
Elongation factor-2 (Fragment) OS=Chaetopleura apiculata PE=2 SV=1	Q9BNW3|Q9BNW3_CHAAP	94.03	42.63	5	6	6	6	4	4	81473	81473

Enolase (Fragment) OS=Leptochiton sp. ENO1 PE=3 SV=1	A8DU73|A8DU73_9MOLL	143.94	116.98	24	25	9	14	4	9	39602	39602
Enolase OS=Loligo pealeii PE=2 SV=1	O02654|ENO_LOLPE	204.03	153.86	41	38	20	22	13	17	47426	47426
Ezrin/radixin/moesin OS=Biomphalaria glabrata PE=2 SV=1	Q963F9|Q963F9_BIOGL	123.45	58.79	31	23	18	15	16	11	69472	69472
Fascin (Fragment) OS=Haliotis diversicolor PE=2 SV=1	B3TK44|B3TK44_HALDV	119.76	113.55	21	25	8	11	8	10	36334	36334
Fructose-bisphosphate aldolase (Fragment) OS=Haliotis discus discus PE=2 SV=1	B6RB70|B6RB70_HALDI	189.18	162.39	46	29	12	11	4	3	26118	26118
G protein alpha subunit s class OS=Octopus vulgaris GN=OvGas PE=2 SV=1	Q9NL91|Q9NL91_OCTVU	132.54	102.76	20	17	10	12	6	6	44563	44563

Gelsolin OS=Haliotis discus discus PE=2 SV=1	B6RB97|B6RB97_HALDI	228.85	113.46	67	32	18	10	16	10	23370	23370
Glucose-6-phosphate 1-dehydrogenase OS=Crassostrea gigas GN=g6PD PE=2 SV=1	Q0KHB8|Q0KHB8_CRAGI	132.99	103.23	31	19	14	10	12	10	52919	52919
Glucose-regulated protein 94 OS=Crassostrea gigas GN=GRP94 PE=2 SV=1	A5LGG7|A5LGG7_CRAGI	190.92	81.95	24	13	22	13	14	8	91624	91624
Glutathione-S-transferase isoform OS=Haliotis diversicolor PE=2 SV=1	B3TK24|B3TK24_HALDV	101.49	73.22	27	16	7	5	7	4	24726	24726
Glyceraldehyde-3-phosphate dehydrogenase OS=Haliotis discus discus PE=2 SV=1	B6RB30|B6RB30_HALDI	302.82	231.66	82	42	56	17	51	15	31477	31477
GTP-binding nuclear protein Ran OS=Haliotis diversicolor supertexta	E6Y2Z1|E6Y2Z1_HALDV	125.53	94.97	22	37	5	9	5	8	25245	25245

PE=2 SV=1											
Guanine nucleotide-binding protein G(I) alpha subunit OS=Aplysia californica PE=2 SV=1	Q0ZA02|Q0ZA02_APLCA	129.04	89.51	21	26	10	10	4	4	40534	40534
Guanine nucleotide-binding protein G(o) subunit alpha OS=Lymnaea stagnalis PE=2 SV=3	P30683|GNAO_LYMST	119.12	98.81	27	23	11	11	5	8	40272	40272
Guanine nucleotide-binding protein G(s) subunit alpha OS=Lymnaea stagnalis PE=2 SV=1	P30684|GNAS_LYMST	132.92	102.76	23	29	12	16	8	9	44099	44099
Guanine nucleotide-binding protein subunit beta OS=Loligo forbesi PE=2 SV=1	P23232|GBB_LOLFO	138.6	82.43	20	25	7	8	6	6	37323	37323

H2 protein (Fragment) OS=Haliotis tuberculata GN=H2 PE=4 SV=1	Q8I0U4|Q8I0U4_HALTU	358.49	256.94	50	20	252	76	79	22	380543	380543
Heat shock protein 60 OS=Biomphalaria glabrata GN=HSP60 PE=2 SV=1	B8Y4I8|B8Y4I8_BIOGL	58.56	84.45	13	16	8	14	8	7	60886	60886
Heat shock protein 70 OS=Haliotis discus hannai PE=2 SV=1	Q2MJK5|Q2MJK5_HALDH	309.75	218.71	79	46	99	37	58	12	71646	71646
Hemocyanin 2 OS=Megathura crenulata GN=KLH2 PE=4 SV=1	Q1MVA1|Q1MVA1_MEGCR	203.55	180.93	15	17	58	66	38	40	391524	391524
Hemocyanin A-type (Fragment) OS=Enteroctopus dofleini GN=HA PE=4 SV=1	Q5XLV2|Q5XLV2_ENTDO	111.65	113.77	12	18	34	62	25	42	332011	332011
Hemocyanin isoform 1	C7FEG7|C7FEG7_HALDV	320.93	283.94	57	23	317	102	211	49	381844	381844

(Fragment) OS=Haliotis diversicolor GN=H1 PE=4 SV=1											
Histone H2B OS=Mytilus edulis GN=H2B PE=3 SV=1	Q6WV83|Q6WV83_MYTED	176.25	159.9	35	56	13	10	2	9	13788	13788
Histone H4 OS=Chlamys farreri GN=H4 PE=3 SV=1	A2CI30|A2CI30_9BIVA	190.04	169.9	66	66	12	11	2	2	11367	11367
Isocitrate dehydrogenase [NADP] OS=Crassostrea gigas PE=2 SV=1	Q5QGY7|Q5QGY7_CRAGI	112.96	49.33	14	4	8	2	5	2	50892	50892
Keyhole limpet hemocyanin1 OS=Megathura crenulata GN=klh1 PE=2 SV=1	Q6KC56|Q6KC56_MEGCR	144.4	118.33	14	14	52	56	36	34	358913	358913
Keyhole limpet hemocyanin2 OS=Megathura crenulata	Q6KC55|Q6KC55_MEGCR	203.55	180.93	15	17	59	66	39	40	391542	391542

GN=klh2 PE=2 SV=1											
Major vault protein (Fragment) OS=Mytilus galloprovincialis PE=2 SV=1	A0SXG1|A0SXG1_MYTGA	84.26	55.69	18	15	5	5	5	5	31712	31712
Malate dehydrogenase (Fragment) OS=Haliotis discus discus PE=2 SV=1	B6RB90|B6RB90_HALDI	208.78	131.47	74	53	15	11	11	6	25804	25804
Nonmuscle myosin II (Fragment) OS=Aplysia californica PE=2 SV=1	Q45R40|Q45R40_APLCA	370.84	252.1	82	36	306	105	238	66	220195	220195
Non-muscle myosin II heavy chain OS=Loligo pealeii PE=2 SV=1	Q8SWQ7|Q8SWQ7_LOLPE	336.93	185.12	75	35	295	93	235	59	227150	227150
Omega class glutathione-s-transferase 1 OS=Haliotis discus discus PE=2 SV=1	B6RAZ8|B6RAZ8_HALDI	122.41	77	35	17	9	6	9	5	27366	27366

Omega-crystallin OS=Ommastrephes sloanei PE=2 SV=3	P30842|CROM_OMMSL	55.43	33.98	5	8	2	4	2	3	56091	56091
Paramyosin OS=Haliotis discus discus GN=hald2 PE=2 SV=1	E5RSV6|E5RSV6_HALDI	84.19	66.68	10	8	11	10	6	5	99518	99518
Peptidyl-prolyl cis-trans isomerase OS=Venerupis philippinarum PE=2 SV=1	C8CBN2|C8CBN2_VENPH	57.86	53.93	19	24	4	5	2	3	20151	20151
Peroxiredoxin OS=Pinctada fucata PE=2 SV=1	D3K380|D3K380_PINFU	103.47	77.81	32	16	10	6	2	4	22317	22317
Phosphoglycerate kinase OS=Aplysia californica GN=PGK PE=2 SV=1	O61471|PGK_APLCA	88.19	74.8	12	16	6	7	5	5	43653	43653
Phosphorylase OS=Crassostrea gigas PE=2 SV=1	Q596I0|Q596I0_CRAGI	111.16	67.11	32	30	28	36	27	30	98636	98636

PL10-like protein OS=Haliotis asinina PE=2 SV=1	C7EAA2|C7EAA2_HALAI	261.3	145.73	33	24	30	22	19	13	86007	86007
Prohibitin OS=Octopus tankahkeei PE=2 SV=1	H8XWJ5|H8XWJ5_9MOLL	84.99	50.42	19	19	5	8	5	7	30207	30207
Proteasome subunit alpha type OS=Haliotis discus discus PE=2 SV=1	B6RB41|B6RB41_HALDI	125.65	101.97	26	37	6	10	5	8	26152	26152
Protein disulfide isomerase OS=Haliotis discus discus PE=2 SV=1	B6RB63|B6RB63_HALDI	274.88	206.68	52	34	36	28	28	22	55235	55235
Protein phosphatase 1 (Fragment) OS=Corbicula fluminea GN=PPP1 PE=4 SV=1	G9HQ39|G9HQ39_9BIVA	180.02	131.77	43	36	5	6	5	5	18086	18086
Psmc6 protein OS=Haliotis discus discus PE=2 SV=1	B6RB65|B6RB65_HALDI	103.97	89.13	15	12	3	3	3	2	30677	30677

Putative 40S ribosomal protein S18 OS=Haliotis diversicolor GN=RPS18 PE=2 SV=1	G9K381|G9K381_HALDV	97.29	68.5	22	27	5	5	3	5	17925	17925
Putative 60S ribosomal protein L3 OS=Haliotis diversicolor GN=RPL3 PE=2 SV=1	G9K380|G9K380_HALDV	201.17	139.66	39	38	14	21	13	16	46111	46111
Putative Na+/K+-ATPase alpha subunit OS=Bathypolypus bairdii PE=2 SV=1	F8V2T8|F8V2T8_9MOLL	128.44	61.33	15	4	14	6	6	2	113792	113792
Putative polyadenylate-binding protein 1 (Fragment) OS=Haliotis diversicolor GN=PABP1 PE=2 SV=1	G9K385|G9K385_HALDV	123.35	48.64	11	6	6	3	5	3	62318	62318
Putative ribosomal protein S25 (Fragment) OS=Crassostrea gigas	Q70MT3|Q70MT3_CRAGI	53.32	56.89	17	26	2	4	2	3	13679	13679

PE=2 SV=1											
Putative tumor suppressor QM protein OS=Pinctada martensi PE=2 SV=1	A5HIE9|A5HIE9_9BIVA	35.25	23.26	9	13	3	4	3	4	25060	25060
Putative uncharacterized protein (Fragment) OS=Haliotis diversicolor PE=2 SV=1	B3TK33|B3TK33_HALDV	159.14	109.86	44	26	14	7	14	6	26128	26128
Pyruvate kinase OS=Crassostrea gigas GN=pk PE=2 SV=1	Q0KHB6|Q0KHB6_CRAGI	76.02	40.5	12	12	6	9	6	6	61346	61346
QM-like protein OS=Haliotis diversicolor supertexta PE=2 SV=1	B8XW76|B8XW76_HALDV	66.5	79.35	22	21	7	7	7	7	25097	25097
RAB protein OS=Haliotis discus discus PE=2 SV=1	B6RB76|B6RB76_HALDI	167.94	148.46	46	46	9	10	6	8	22495	22495
RAB1 (Fragment) OS=Haliotis cracherodii	A7L3I9|A7L3I9_HALCR	168.9	126.66	82	71	14	14	4	4	16030	16030

PE=2 SV=1											
Rab7-like protein OS=Pinctada martensi PE=2 SV=1	C1KBI9|C1KBI9_9BIVA	163.04	92	43	27	8	5	8	4	23118	23118
Raminin receptor OS=Haliotis discus discus PE=2 SV=1	B6RB17|B6RB17_HALDI	179.74	142.67	27	42	10	13	9	9	35225	35225
Ran-1-prov protein OS=Haliotis discus discus PE=2 SV=1	B6RB29|B6RB29_HALDI	125.53	94.97	22	37	5	9	5	8	25187	25187
Ras-like GTP-binding protein RHO OS=Aplysia californica GN=RHO PE=2 SV=1	P01122|RHO_APLCA	100.47	62.13	18	24	4	4	3	4	21661	21661
Ras-related nuclear protein OS=Mizuhopecten yessoensis GN=Ran PE=2 SV=1	H1ADR7|H1ADR7_MIZYE	99.41	86.61	16	18	4	5	4	5	25059	25059

Ras-related protein Rab-1A OS=Haliotis discus discus PE=2 SV=1	B6RB23|B6RB23_HALDI	214.43	153.59	84	74	19	18	6	6	22803	22803
Ras-related protein Rab-2 OS=Lymnaea stagnalis GN=RAB2 PE=2 SV=1	Q05975|RAB2_LYMST	199.23	84.97	52	18	11	4	10	3	23535	23535
Rho-like protein OS=Biomphalaria glabrata PE=2 SV=1	E6Y9P8|E6Y9P8_BIOGL	120.3	80.6	24	25	6	5	5	5	21717	21717
Ribosomal protein S20 OS=Crassostrea gigas GN=rps20 PE=4 SV=1	Q70MN4|Q70MN4_CRAGI	93.71	64.81	25	16	4	3	3	2	13284	13284
Ribosomal protein S25 OS=Lepidochitona cinerea PE=4 SV=1	E3P7H5|E3P7H5_9MOLL	61.76	62.08	23	31	3	4	3	4	13081	13081
Ribosomal protein S5 OS=Crassostrea gigas PE=2 SV=1	Q4H451|Q4H451_CRAGI	103.57	73.8	16	12	3	2	3	2	22965	22965

Ribosomal protein S7 OS=Aequipecten irradians PE=2 SV=1	Q8ITA9|Q8ITA9_AEQIR	69.01	57.59	38	18	10	4	10	2	22359	22359
Ribosomal protein S9 OS=Haliotis discus discus PE=2 SV=1	A9LMJ6|A9LMJ6_HALDI	120.99	83.42	31	28	9	8	5	7	22348	22348
Sarco/endoplasmic reticulum calcium ATPase isoform C OS=Pinctada fucata GN=SERCA PE=2 SV=1	B2KKR2|B2KKR2_PINFU	37.55	68.91	5	4	5	4	3	2	110388	110388
Small nuclear ribonucleoprotein D2-like protein OS=Chlamys farreri PE=4 SV=1	Q8MUE3|Q8MUE3_9BIVA	57.6	46.82	14	22	2	2	2	2	13677	13677
Splicing factor arginine/serine-rich 4 OS=Haliotis diversicolor supertexta PE=2 SV=1	E6Y2Z7|E6Y2Z7_HALDV	138.2	50.39	35	12	8	2	7	2	18398	18398

Thioredoxin peroxidase 1 OS=Haliotis discus discus PE=2 SV=1	B1N693|B1N693_HALDI	80.03	121.81	10	28	2	10	2	7	27961	27961
Transgelin OS=Haliotis discus discus PE=2 SV=1	B6RB35|B6RB35_HALDI	169.35	107.69	47	20	12	6	11	5	20766	20766
Triosephosphate isomerase (Fragment) OS=Haliotis rufescens PE=2 SV=1	Q45Y86|Q45Y86_HALRU	175.51	138.04	36	33	7	8	5	3	22629	22629
Tropomyosin 4 (Fragment) OS=Haliotis asinina PE=2 SV=1	Q7YZR1|Q7YZR1_HALAI	156.4	112.62	56	64	11	9	7	2	12349	12349
Troponin OS=Haliotis discus discus PE=2 SV=1	B6RB46|B6RB46_HALDI	43.03	31.71	18	15	3	4	2	3	19455	19455
Ubiquitin OS=Crassostrea hongkongensis GN=UbL40 PE=2 SV=1	D2XEB0|D2XEB0_CRAHO	153.91	117.23	48	59	7	11	7	6	14744	14744

Vacuolar H+-ATPase (Fragment) OS=Ilyanassa obsoleta PE=2 SV=1	O96703|O96703_9CAEN	141.01	81.45	52	44	6	9	2	8	14015	14015
Visual G-protein beta subunit OS=Loligo pealeii PE=2 SV=1	Q6TQF6|Q6TQF6_LOLPE	138.6	82.43	20	25	7	8	6	6	37309	37309
Voltage-dependent anion channel 2-like protein OS=Haliotis diversicolor PE=2 SV=1	D7RP02|D7RP02_HALDV	186.74	105.45	38	30	13	11	11	8	30632	30632

B3 - Peaks search supplementary information

Supplementary information for the Peaks searches, *de novo* searches and Peaks quantitative workflow for both iTRAQ experiments have been uploaded to online cloud storage services, and are available via either of the following links:

Via Dropbox:

http://tinyurl.com/l8q8e7m

Via Microsoft OneDrive:

http://1drv.ms/1jtI3B5

B4 - Supplementary CLC workbench figures

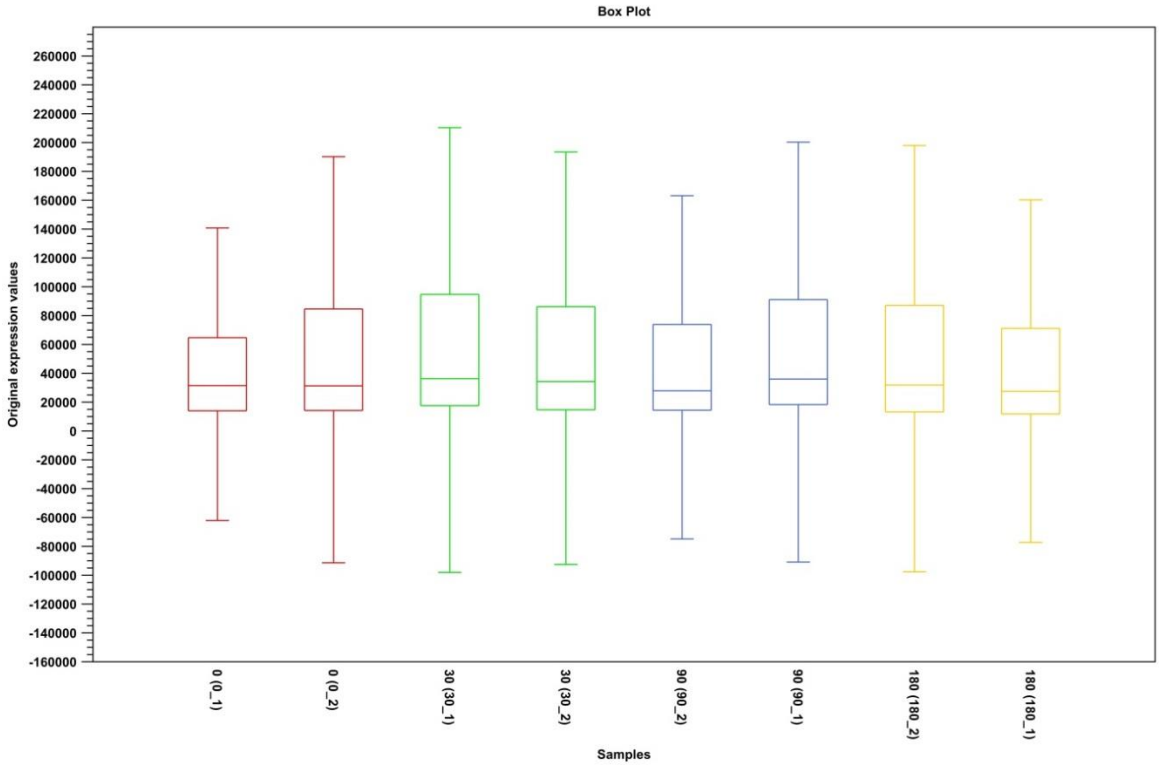


Figure B-10: Box plot showing mean and dispersion of original expression values for iTRAQ experiment 1 before quantile normalisation.

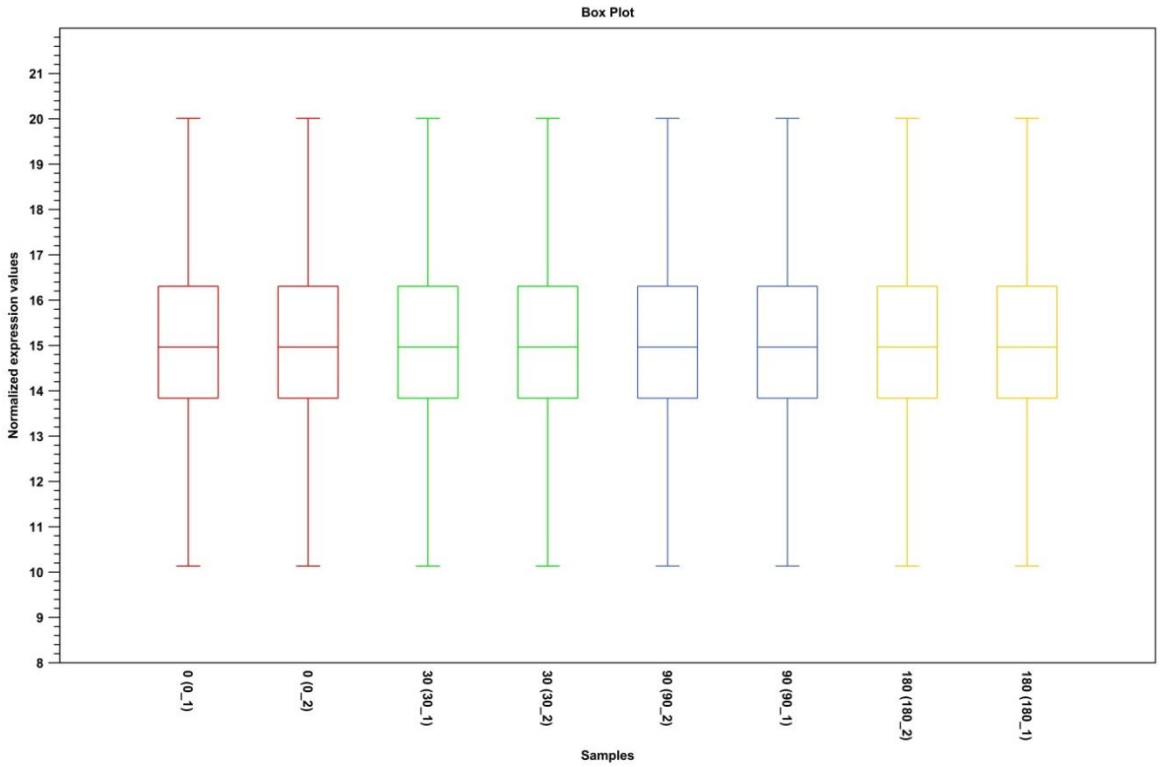


Figure B-11: Box plot showing mean and dispersion of quantile normalised expression values for iTRAQ experiment 1

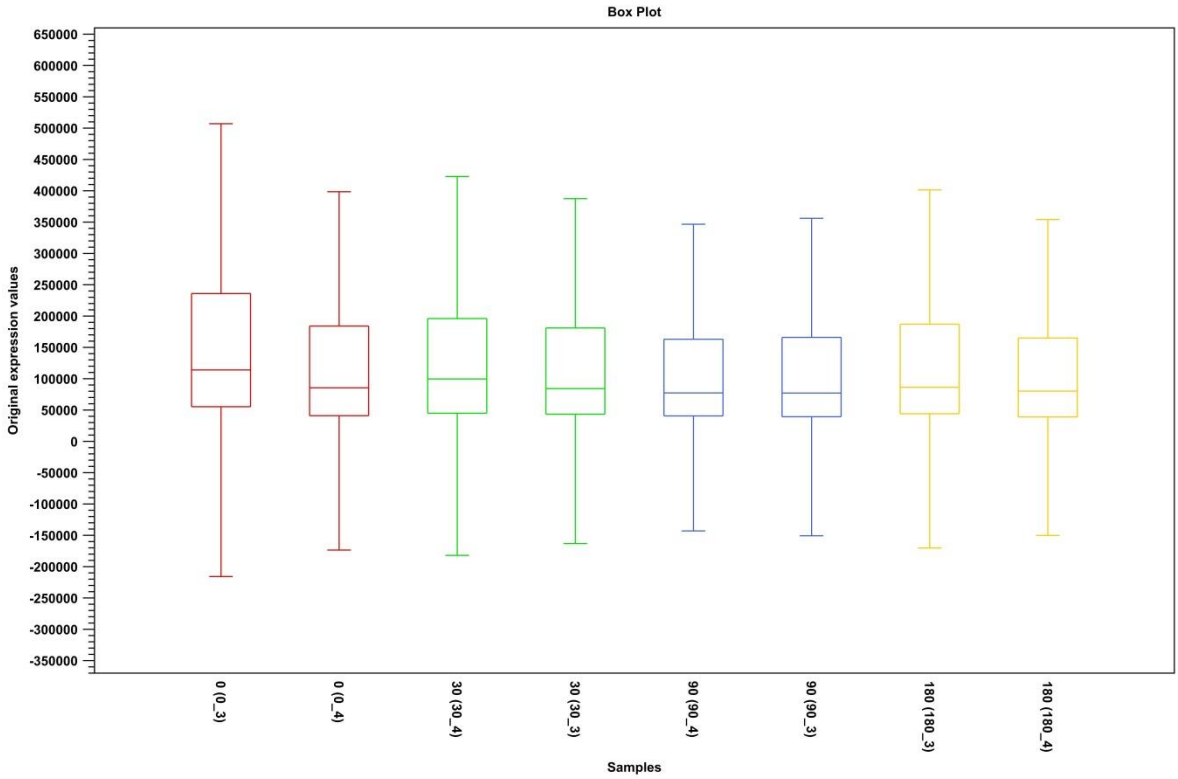


Figure B-12: Box plot showing mean and dispersion of original expression values for iTRAQ experiment 2 before quantile normalisation.

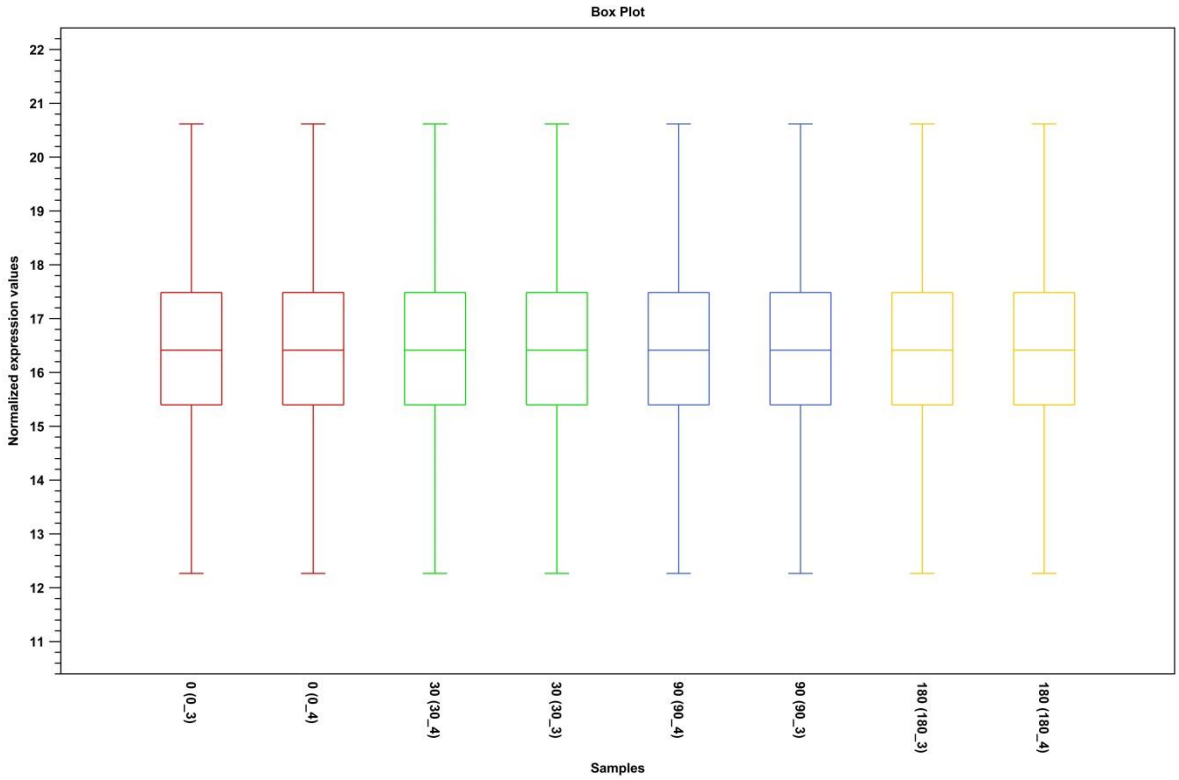


Figure B-13: Box plot showing mean and dispersion of quantile normalised expression values for iTRAQ experiment 2

B5 - Hsp70d, G6P1D and GSTO1 western blot loading controls

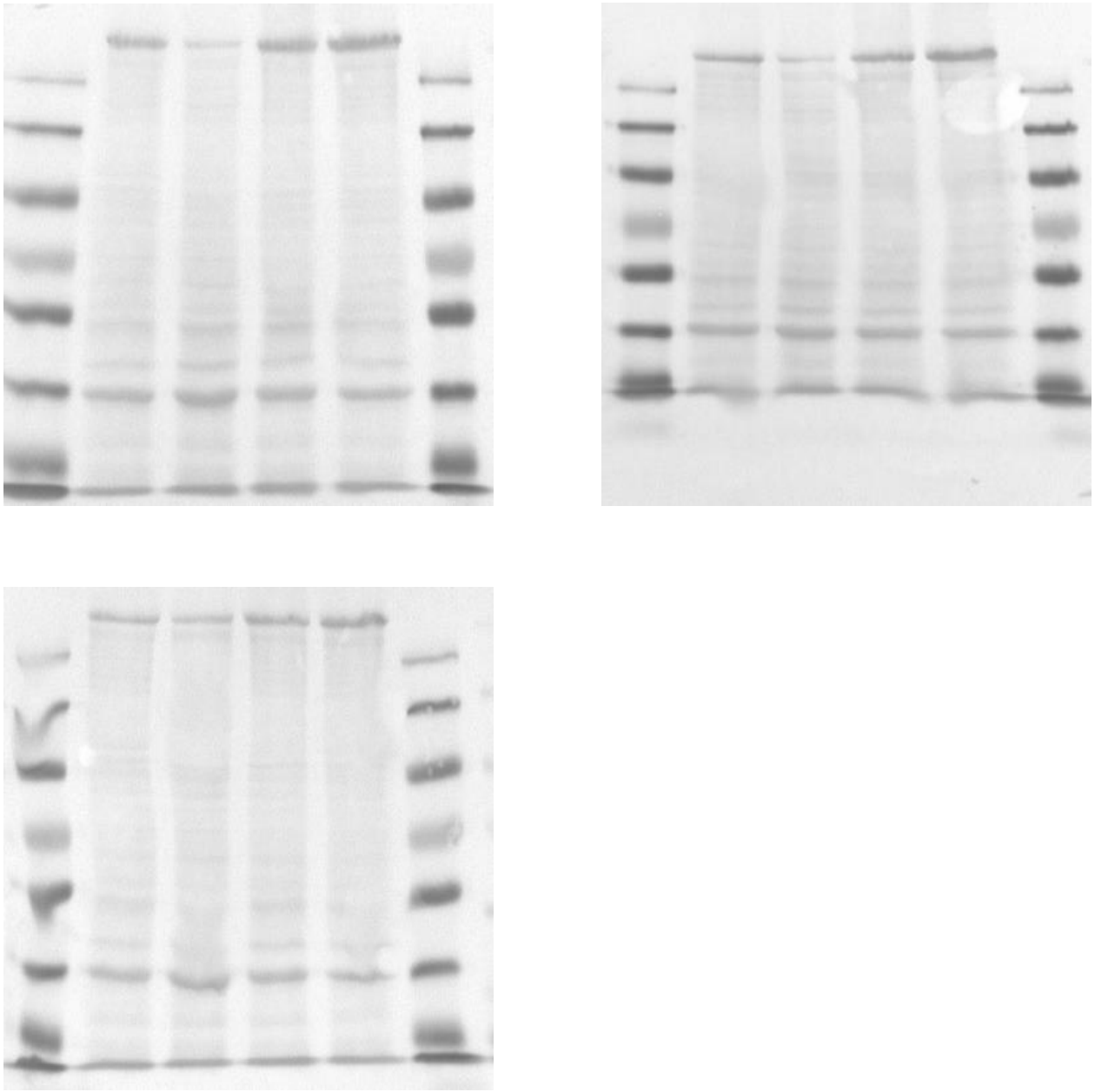
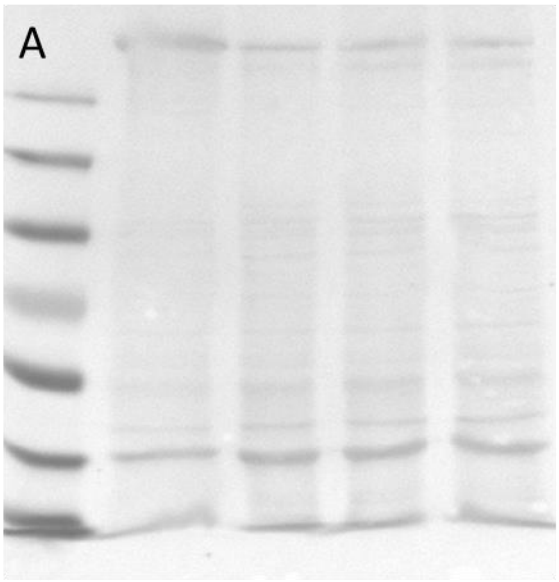
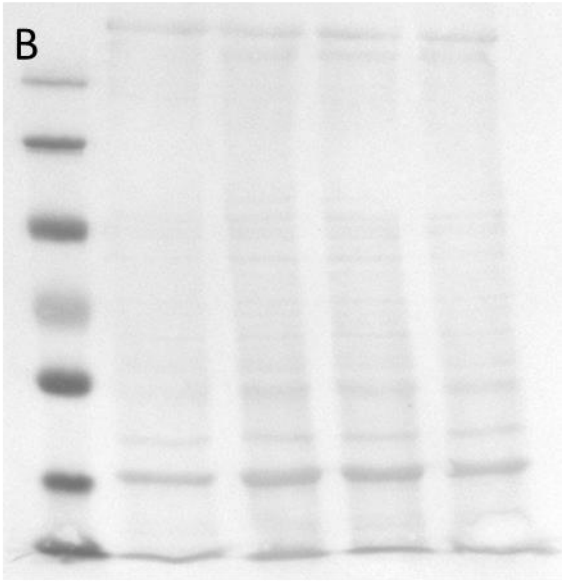


Figure B-14: Ponceau S total protein stains of protein transferred electrophoretically onto nitrocellulose membrane for the purposes of western blot hybridisation in order to detect and quantify the expression of A) Hsp70 B) G6P1D and C) GSTO1 in abalone haemocytes during exposure to acute heat stress. These images were used as loading controls for the first biological replicate performed. From left to right the lanes are: molecular weight marker, 0, 30, 90, 180, molecular weight marker.





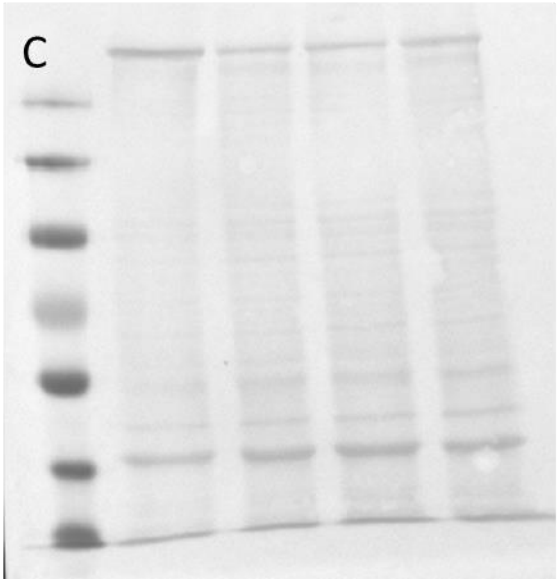


Figure B-15: Ponceau S total protein stains of protein transferred electrophoretically onto nitrocellulose membrane for the purposes of western blot hybridisation in order to detect and quantify the expression of A) Hsp70 B) G6P1D and C) GSTO1 in abalone haemocytes during exposure to acute heat stress. These images were used as loading controls for the second biological replicate performed. From left to right the lanes are: molecular weight marker, 0, 30, 90 and 180.

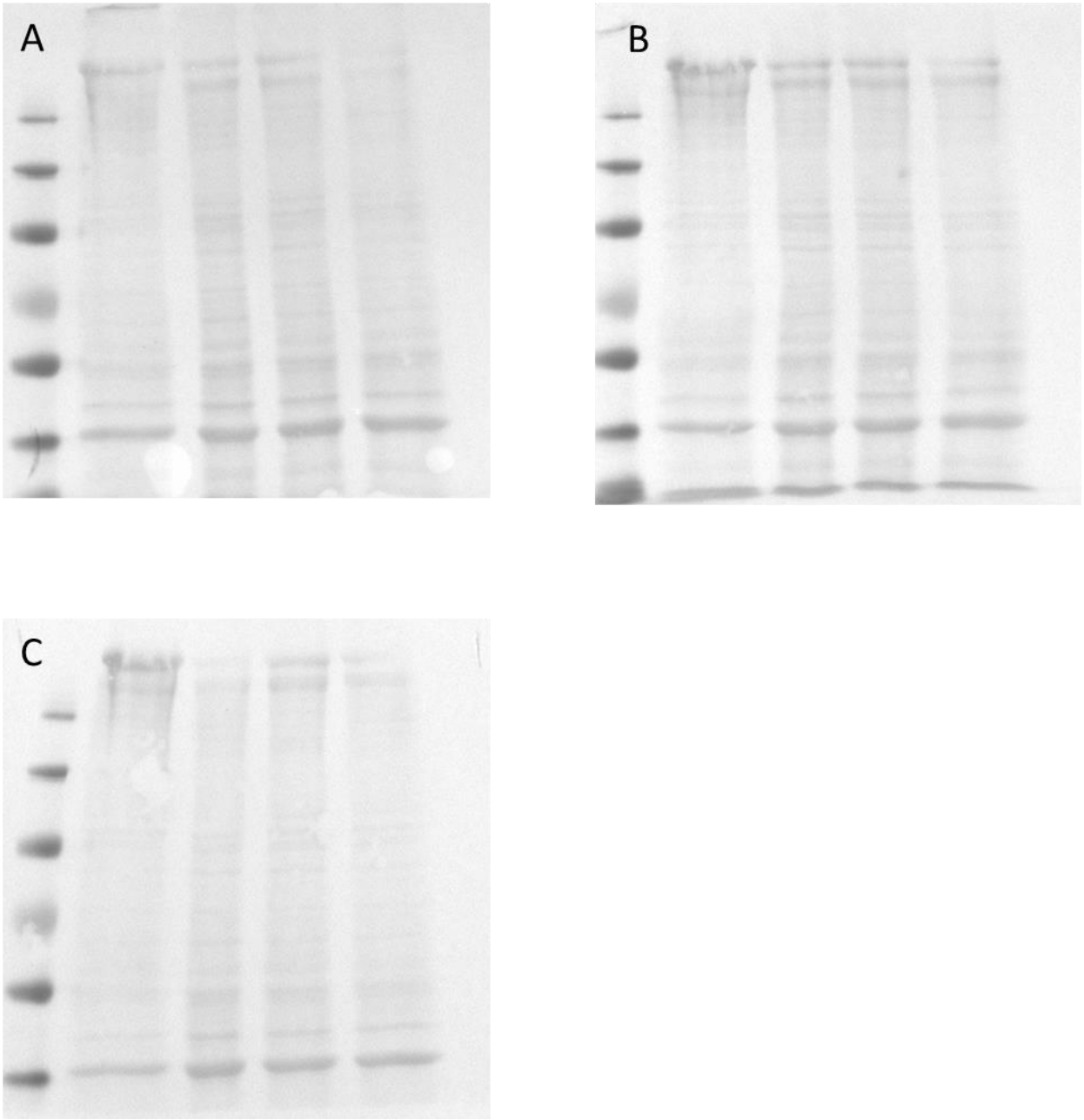


Figure B-16: Ponceau S total protein stains of protein transferred electrophoretically onto nitrocellulose membrane for the purposes of western blot hybridisation in order to detect and quantify the expression of A) Hsp70 B) G6P1D and C) GSTO1 in abalone haemocytes during exposure to acute heat stress. These images were used as loading controls for the third biological replicate performed. From left to right the lanes are: molecular weight marker, 0, 30, 90 and 180.

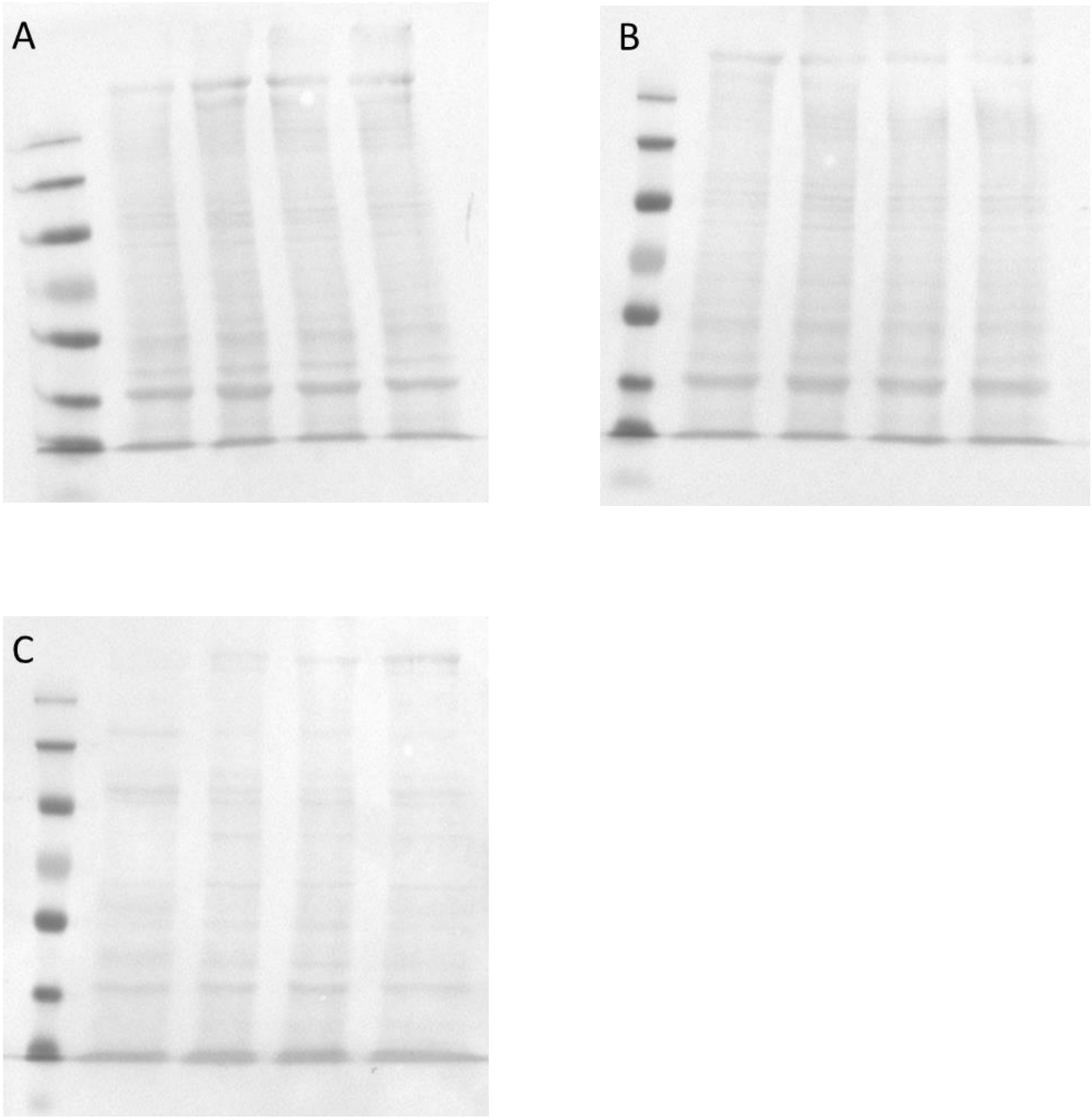


Figure B-17: Ponceau S total protein stains of protein transferred electrophoretically onto nitrocellulose membrane for the purposes of western blot hybridisation in order to detect and quantify the expression of A) Hsp70 B) G6P1D and C) GSTO1 in abalone haemocytes during exposure to acute heat stress. These images were used as loading controls for the fourth biological replicate performed. From left to right the lanes are: molecular weight marker, 0, 30, 90 and 180.

B6 - Real Time PCR supplementary information

RNA quality control:

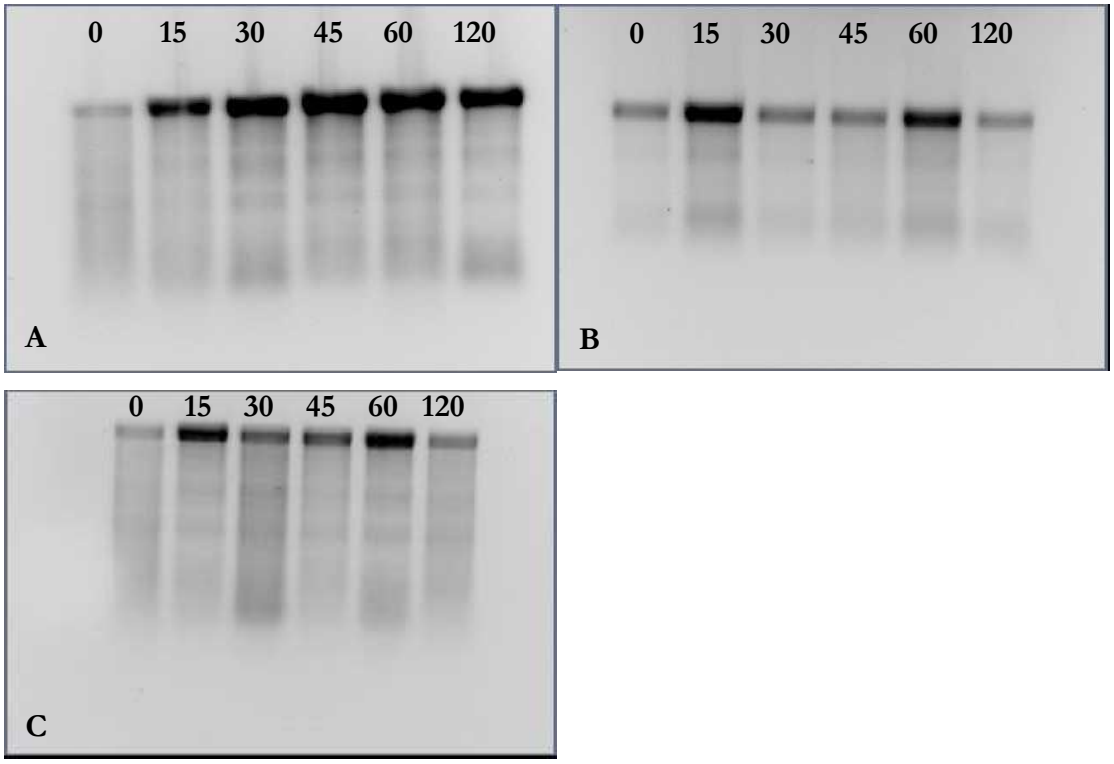


Figure B-18: 1.2% agarose/formaldehyde RNA gels showing quality of RNA used for cDNA conversion and qPCR reactions performed in Chapter 4. Lanes from left to right are RNA isolated from haemocytes sampled from abalone at time points 0, 15, 30, 45, 60 and 120 minutes during exposure to acute thermal stress. Gel A is RNA from biological replicate 1, Gel B from biological replicate 2 and Gel C from biological replicate 3.

Real Time PCR standard curves:

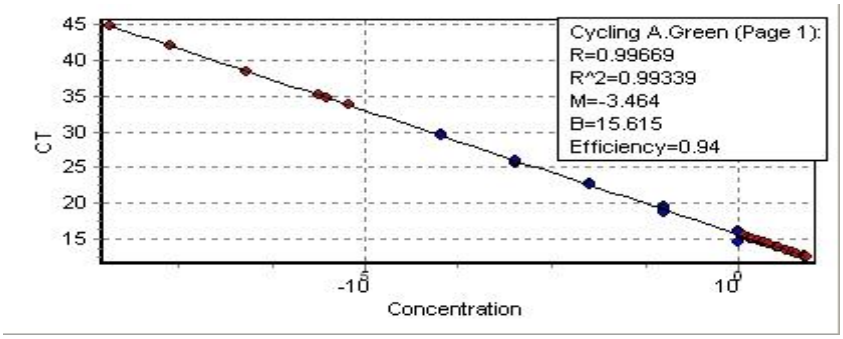


Figure B-19: A representative standard curve generated for calreticulin gene specific primers and using serially diluted pooled cDNA as a template. Standard curves were generated using the Rotor-Gene™ 6000 (Corbett Life Sciences) by qPCR as described in Chapter 4.

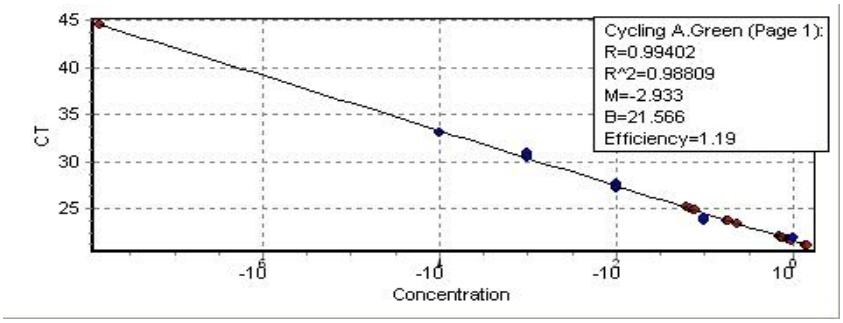


Figure B-20: A representative standard curve generated for ribosomal protein L28 (reference gene) gene specific primers using serially diluted pooled cDNA as a template. Standard curves were generated using the Rotor-Gene™ 6000 (Corbett Life Sciences) by qPCR as described in Chapter 4.

Real Time PCR melt curves:

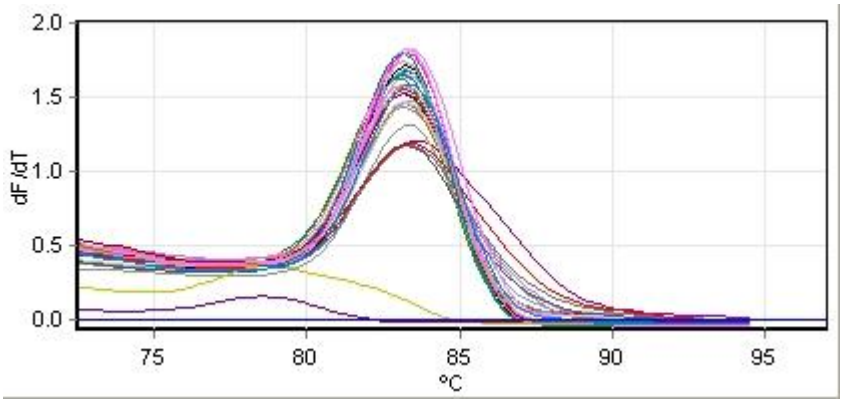


Figure B-21: A representative melt curve for the qPCR product generated using gene specific primers for calreticulin as described in Chapter 4. Melt curve was generated using the Rotor-Gene™ 6000 (Corbett Life Sciences).

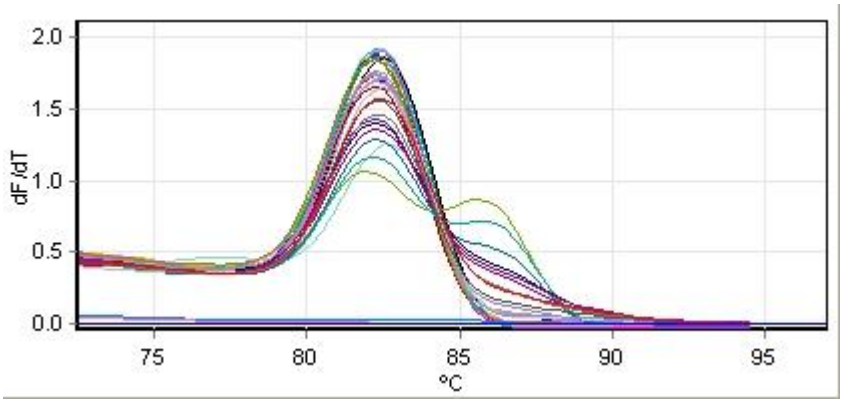


Figure B-22: A representative melt curve for the qPCR product generated using gene specific primers for ribosomal factor L28 (reference gene) as described in Chapter 4. Melt curve was generated using the Rotor-Gene™ 6000 (Corbett Life Sciences).

Reference gene selection:

The reference ribosomal factor L28 was stably expressed for the duration of the exposure to heat shock, according to MIQUE guidelines (Bustin et al., 2009). The real time qPCR reaction efficiencies for both the reference gene and gene of interest are shown along with M and R^2 values for the standard curves and melt analysis of the PCR product. Figure B-23 shows the average Ct values of ribosomal factor L28 obtained from three technical replicates assayed at each time point for each biological repeat (n=3). The average Ct value within each biological replicate did not differ by more than 5 Ct points across all time points. Furthermore, the reference gene did not cross the threshold point later than 30 cycles for any time point in any biological replicate.

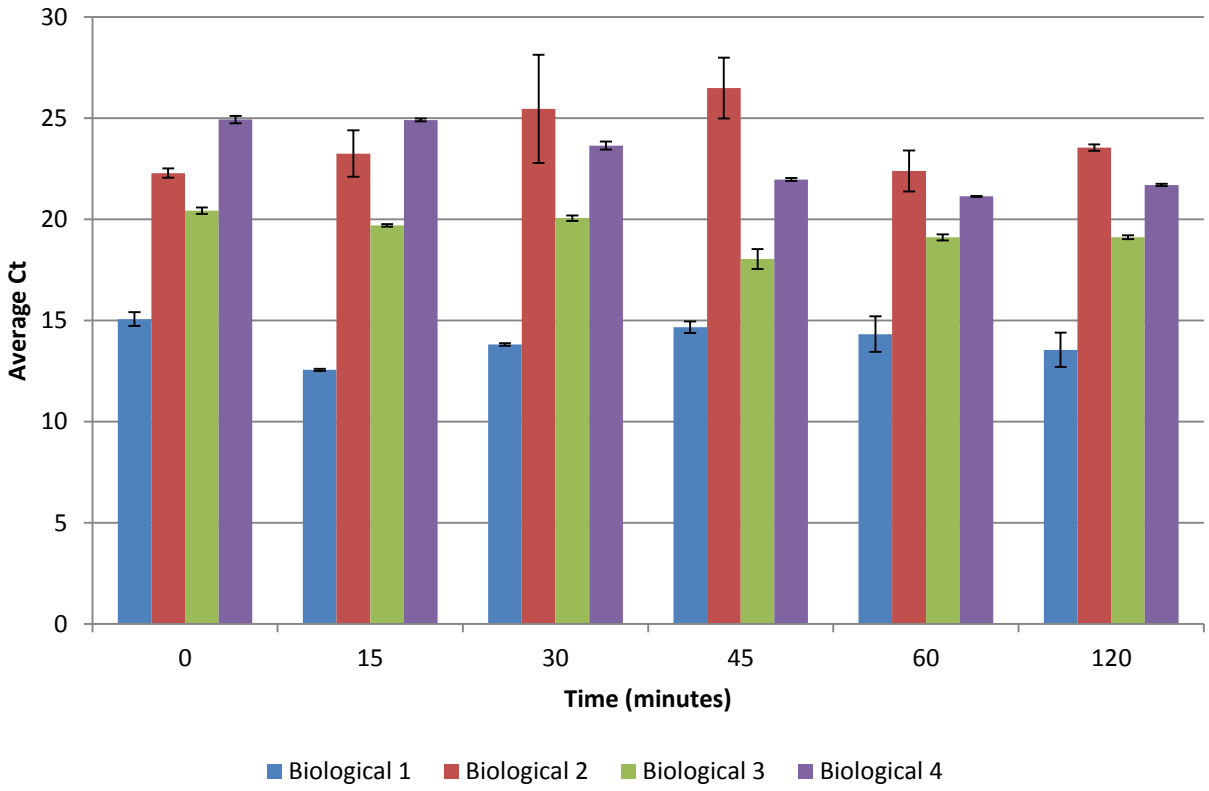


Figure B-23: Average Ct values for ribosomal factor L28 during exposure to acute heat shock as determined by real time qPCR. Based on its stable expression under experimental conditions, ribosomal factor L28 was selected as the reference gene for this study.

B7 - Calreticulin western blot loading controls

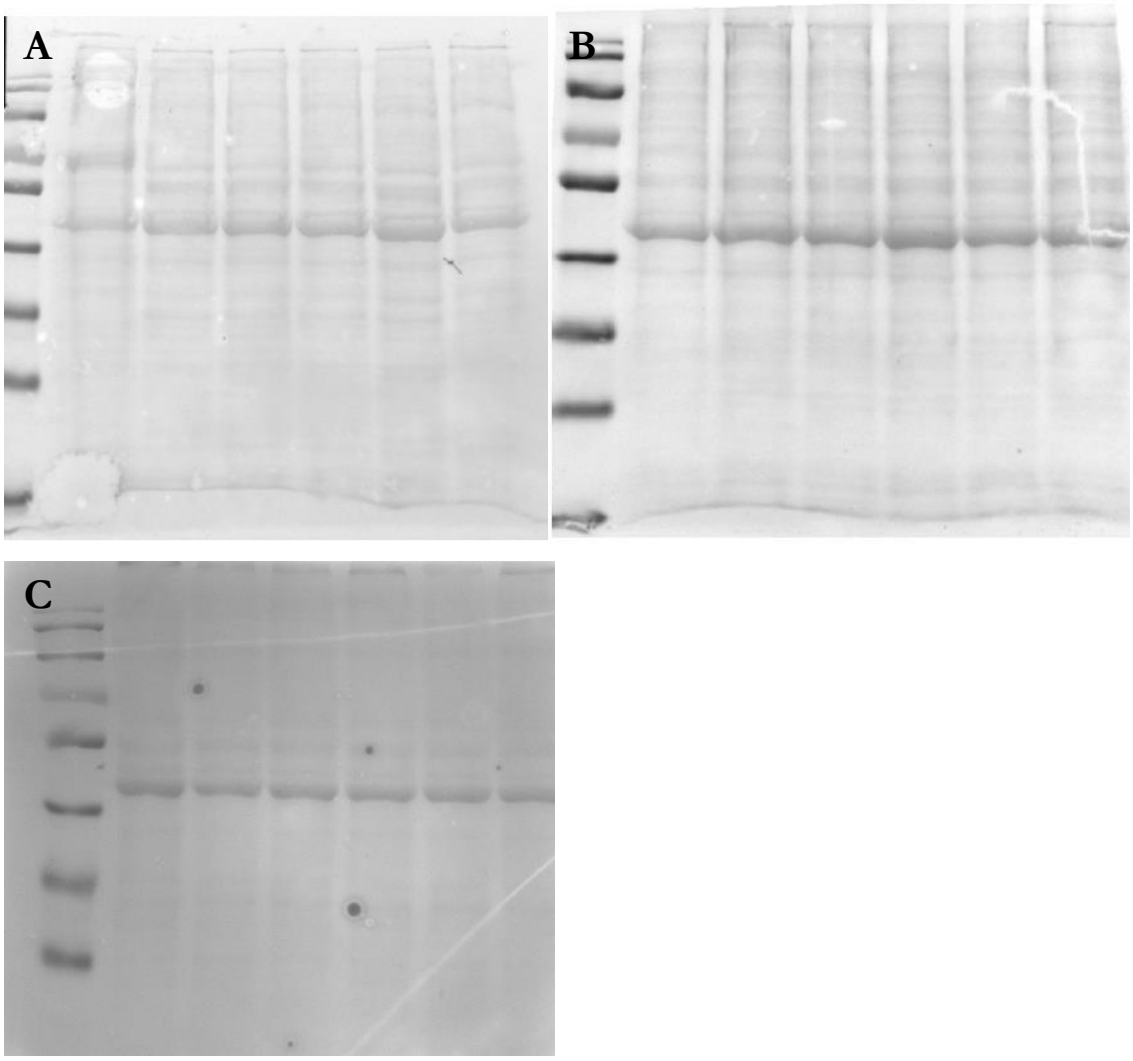


Figure B-24: Ponceau S total protein stains of protein transferred electrophoretically onto nitrocellulose membrane for the purposes of western blot hybridisation in order to detect and quantify the expression of calreticulin in abalone haemocytes during exposure to acute heat stress. From left to right, the lanes are: molecular weight marker, 0, 30, 60, 120, 180 and 24h samples. Images represent loading controls used in A) biological replicate 1, B) biological replicate 2 and C) biological replicate 3.

Appendix C

Literature cited

Abele, D., Tesch, C., Wencke, P., and Pörtner, H.-O. (2001). How does oxidative stress relate to thermal tolerance in the Antarctic bivalve *Yoldia eightsi*? Antarctic Science 13, 111–118.

Adamo, S. A. (2012). The effects of the stress response on immune function in invertebrates: An evolutionary perspective on an ancient connection. Hormones and Behavior *62*, 324–330.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Aggarwal, K., Choe, L. H., and Lee, K. H. (2006). Shotgun proteomics using the iTRAQ isobaric tags. Briefings in Functional Genomics & Proteomics *5*, 112–120.

Ahner, A., Brodsky, J. L., and others (2004). Checkpoints in ER-associated degradation: excuse me, which way to the proteasome? Trends in Cell Biology *14*, 474–478.

Alexandratos, N., and Bruinsma, J. (2012). World agriculture towards 2030/2050: the 2012 revision.

Assinder, S. J., Stanton, J.-A. L., and Prasad, P. D. (2009). Transgelin: an actin-binding protein and tumour suppressor. The International Journal of Biochemistry & Cell Biology *41*, 482–486.

Bachère, E. (2003). Anti-infectious immune effectors in marine invertebrates: potential tools for disease control in larviculture. Aquaculture *227*, 427–438.

Bastianutto, C., Clementi, E., Codazzi, F., Podini, P., De Giorgi, F., Rizzuto, R., Meldolesi, J., and Pozzan, T. (1995). Overexpression of calreticulin increases the Ca^{2+} capacity of rapidly exchanging Ca^{2+} stores and reveals aspects of their lumenal microenvironment and function. The Journal of Cell Biology *130*, 847–855.

Bateman, D. A. (2006). Unfolded Protein Response (UPR): Cellular control for our errors in life. Hypothesis *4*, 33.

Berridge, M. J., Bootman, M. D., and Roderick, H. L. (2003). Calcium signalling: dynamics, homeostasis and remodelling. Nature Reviews Molecular Cell Biology *4*, 517–529.

Berridge, M. J., Lipp, P., and Bootman, M. D. (2000). The versatility and universality of calcium signalling. Nature Reviews Molecular Cell Biology *1*, 11–21.

Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis *21*, 1037–1053.

Boutet, I., Tanguy, A., and Moraga, D. (2003a). Organization and nucleotide sequence of the European flat oyster *Ostrea edulis* heat shock cognate 70 (*hsc70*) and heat shock protein 70 (*hsp70*) genes. Aquatic Toxicology *65*, 221–225.

Boutet, I., Tanguy, A., and Moraga, D. (2004). Characterisation and expression of four mRNA sequences encoding glutathione S-transferases pi, mu, omega and sigma classes in the Pacific oyster *Crassostrea gigas* exposed to hydrocarbons and pesticides. Marine Biology *146*, 53–64.

Boutet, I., Tanguy, A., Rousseau, S., Auffret, M., and Moraga, D. (2003b). Molecular identification and expression of heat shock cognate 70 (hsc70) and heat shock protein 70 (hsp70) genes in the Pacific oyster *Crassostrea gigas*. Cell Stress & Chaperones *8*, 76.

Bower, N. I., and Johnston, I. A. (2010). Targeted rapid amplification of cDNA ends (T-RACE)—an improved RACE reaction through degradation of non-target sequences. Nucleic Acids Research *38*, e194–e194.

Branch, G. M., and Clark, B. M. (2006). Fish stocks and their management: the changing face of fisheries in South Africa. Marine Policy *30*, 3–17.

Britz, P. J., Hecht, T., and Mangold, S. (1997). Effect of temperature on growth, feed consumption and nutritional indices of *Haliotis midae* fed a formulated diet. Aquaculture *152*, 191–203.

Brun, N. T., Bricelj, V. M., MacRae, T. H., and Ross, N. W. (2008). Heat shock protein responses in thermally stressed bay scallops, *Argopecten irradians*, and sea scallops, *Placopecten magellanicus*. Journal of Experimental Marine Biology and Ecology *358*, 151–162.

Buck, T. M., Wright, C. M., and Brodsky, J. L. (2007). The activities and function of molecular chaperones in the endoplasmic reticulum. In: Seminars in Cell & Developmental Biology, 751–761.

Burkhart, J. M., Vaudel, M., Zahedi, R. P., Martens, L., and Sickmann, A. (2011). iTRAQ protein quantification: A quality-controlled workflow. Proteomics *11*, 1125–1134.

Bustin, S. A. *et al.* (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical Chemistry *55*, 611–622.

Byrne, M., Ho, M., Wong, E., Soars, N. A., Selvakumaraswamy, P., Shepard-Brennand, H., Dworjanyn, S. A., and Davis, A. R. (2011). Unshelled abalone and corrupted urchins: development of marine calcifiers in a changing ocean. Proceedings of the Royal Society B: Biological Sciences *278*, 2376–2383.

Campos, A., Tedesco, S., Vasconcelos, V., and Cristobal, S. (2012). Proteomic research in bivalves: Towards the identification of molecular markers of aquatic pollution. Journal of Proteomics.

Canesi, L., Ciacci, C., Lorusso, L. C., Betti, M., Guarnieri, T., Tavolari, S., and Gallo, G. (2006). Immunomodulation by 17 B-estradiol in bivalve hemocytes. American Journal of

Physiology-Regulatory, Integrative and Comparative Physiology *291*, R664–R673.

Cao, A., Fuentes, J., Comesaña, P., Casas, S. M., and Villalba, A. (2009). A proteomic approach envisaged to analyse the bases of oyster tolerance/resistance to bonamiosis. Aquaculture *295*, 149–156.

Caraux, G., and Pinloche, S. (2005). PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. Bioinformatics *21*, 1280–1281.

Chen, G., Zhang, C., Li, C., Wang, C., Xu, Z., and Yan, P. (2011). Haemocyte protein expression profiling of scallop *Chlamys farreri* response to acute viral necrosis virus (AVNV) infection. Developmental & Comparative Immunology *35*, 1135–1145.

Cheng, P., Liu, X., Zhang, G., and He, J. (2007). Cloning and expression analysis of a HSP70 gene from Pacific abalone (*Haliotis discus hannai*). Fish Shellfish Immunol. *22*, 77–87.

Cheng, W., Hsiao, I., Hsu, C.-H., Chen, J.-C., and others (2004a). Change in water temperature on the immune response of Taiwan abalone *Haliotis diversicolor supertexta* and its susceptibility to *Vibrio parahaemolyticus*. Fish & Shellfish Immunology *17*, 235–243.

Cheng, W., Li, C. H., and Chen, J. C. (2004b). Effect of dissolved oxygen on the immune response of *Haliotis diversicolor supertexta* and its susceptibility to *Vibrio parahaemolyticus*. Aquaculture *232*, 103–115.

Chomczynski, P., and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Annal. Biochem. 162(1): 156-9.

Clark, M. S., and Peck, L. S. (2009). HSP70 Heat shock proteins and environmental stress in Antarctic marine organisms: a mini-review. Marine Genomics *2*, 11–18.

Cook, P. A., and Gordon, R. H. (2010). World abalone supply, markets, and pricing. Journal of Shellfish Research *29*, 569–571.

Coppolino, M. G., and Dedhar, S. (1998). Calreticulin. The International Journal of Biochemistry & Cell Biology *30*, 553–558.

Coyne, V. (2011). The importance of ATP in the immune system of molluscs. Invertebrate Survival Journal *8*, 48–55.

Cunningham, L. (2005). Assessing the contribution of aquaculture to food security: a survey of methodologies, FAO.

Dahlhoff, E., and Somero, G. N. (1993). Effects of temperature on mitochondria from abalone (genus Haliotis): adaptive plasticity and its limits. Journal of Experimental Biology *185*, 151–168.

Dalle-Donne, I., Rossi, R., Milzani, A., Di Simplicio, P., and Colombo, R. (2001). The actin cytoskeleton response to oxidants: from small heat shock protein phosphorylation to changes in the redox state of actin itself. Free Radical Biology and Medicine *31*, 1624–

1632.

Dang, V. T., Speck, P., and Benkendorff, K. (2012). Influence of elevated temperatures on the immune response of abalone, *Haliotis rubra*. Fish & Shellfish Immunology *32*, 732–740.

Diaz, F., Re, A. D., Medina, Z., Re, G., Valdez, G., and Valenzuela, F. (2006). Thermal preference and tolerance of green abalone *Haliotis fulgens* (Philippi, 1845) and pink abalone *Haliotis corrugata* (Gray, 1828). Aquaculture Research *37*, 877–884.

Dimitriadis, V. K., Gougoula, C., Anestis, A., Pörtner, H. O., and Michaelidis, B. (2011). Monitoring the biochemical and cellular responses of marine bivalves during thermal stress by using biomarkers. Marine Environmental Research.

Donaghy, L., and Volety, A. K. (2011). Functional and metabolic characterization of hemocytes of the green mussel, *Perna viridis*: *in vitro* impacts of temperature. Fish & Shellfish Immunology *31*, 808–814.

Dowling, V., Hoarau, P. C., Romeo, M., O'Halloran, J., Van Pelt, F., O'Brien, N., and Sheehan, D. (2006). Protein carbonylation and heat shock response in *Ruditapes decussatus* following *p*, *p*'-dichlorodiphenyldichloroethylene (DDE) exposure: A proteomic approach reveals that DDE causes oxidative stress. Aquatic Toxicology *77*, 11–18.

Ekanayake, P. M., De Zoysa, M., Kang, H.-S., Wan, Q., Jee, Y., Lee, Y.-H., Kim, S.-J., and Lee, J. (2008). Cloning, characterization and tissue expression of disk abalone (*Haliotis discus discus*) catalase. Fish & Shellfish Immunology *24*, 267–278.

Ellis, E. M. (2007). Reactive carbonyls and oxidative stress: potential for therapeutic intervention. Pharmacology & Therapeutics *115*, 13–24.

Eng, J. K., Searle, B. C., Clauser, K. R., and Tabb, D. L. (2011). A face in the crowd: recognizing peptides through database search. Molecular & Cellular Proteomics *10*.

Erasmus, J. H., Cook, P. A., and Coyne, V. E. (1997). The role of bacteria in the digestion of seaweed by the abalone *Haliotis midae*. Aquaculture *155*, 377–386.

Ernoult, E., and Guette, C. (2011). OFFGEL-Isoelectric Focusing Fractionation for the Analysis of Complex Proteome.

Fan, W., Hu, Y., Li, C., Xie, L., and Zhang, R. (2008). Cloning, Characterization, and Expression Analysis of Calreticulin from Pearl Oyster *Pinctada fucata*. Tsinghua Science & Technology *13*, 466–473.

Farcy, E., Serpentini, A., Fiévet, B., and Lebel, J. M. (2007). Identification of cDNAs encoding HSP70 and HSP90 in the abalone *Haliotis tuberculata*: Transcriptional induction in response to thermal stress in hemocyte primary culture. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *146*, 540–550.

Feder, M. E., and Hofmann, G. E. (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annual Review of Physiology

61, 243–282.

Fields, P. A., Cox, K. M., and Karch, K. R. (2012a). Latitudinal variation in protein expression after heat stress in the salt marsh mussel *Geukensia demissa*. Integrative and Comparative Biology.

Fields, P. A., Zuzow, M. J., and Tomanek, L. (2012b). Proteomic responses of blue mussel (Mytilus) congeners to temperature acclimation. The Journal of Experimental Biology *215*, 1106–1116.

Finehout, E. J., Cantor, J. R., and Lee, K. H. (2005). Kinetic characterization of sequencing grade modified trypsin. Proteomics *5*, 2319–2321.

Fleury, E. *et al.* (2009). Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. BMC Genomics *10*, 341.

Forné, I., Abián, J., and Cerdà, J. (2010). Fish proteome analysis: Model organisms and non-sequenced species. Proteomics *10*, 858–872.

Gasper, R., Meyer, S., Gotthardt, K., Sirajuddin, M., and Wittinghofer, A. (2009). It takes two to tango: regulation of G proteins by dimerization. Nature Reviews Molecular Cell Biology *10*, 423–429.

Gelebart, P., Opas, M., and Michalak, M. (2005). Calreticulin, a Ca $^{2+}$-binding chaperone of the endoplasmic reticulum. The International Journal of Biochemistry \& Cell Biology *37*, 260–266.

Gevaert, K., and Vandekerckhove, J. (2011). Gel-Free Proteomics Methods and Protocols, Springer.

Girón-Pérez, M. (2010). Relationships between innate immunity in bivalve molluscs and environmental pollution. ISJ *7*, 149–156.

Gómez-Mendikute, A., Etxeberria, A., Olabarrieta, I., and Cajaraville, M. P. (2002). Oxygen radicals production and actin filament disruption in bivalve haemocytes treated with benzo (*a*) pyrene. Marine Environmental Research *54*, 431–436.

Gordon, H. R., and Cook, P. A. (2004). World abalone fisheries and aquaculture update: supply and market dynamics. Journal of Shellfish Research *23*, 935–940.

Groenendyk, J., Michalak, M., and others (2005). Endoplasmic reticulum quality control and apoptosis. ACTA BIOCHIMICA POLONICA-ENGLISH EDITION- *52*, 381.

Gunning, P. W., Schevzov, G., Kee, A. J., and Hardeman, E. C. (2005). Tropomyosin isoforms: divining rods for actin cytoskeleton function. Trends in Cell Biology *15*, 333–341.

Görg, A., Postel, W., and Günther, S. (1988). Two-dimensional electrophoresis. The current state of two-dimensional electrophoresis with immobilized pH gradients.

Electrophoresis *9*, 531–546.

Görg, A., Weiss, W., and Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. Proteomics *4*, 3665–3685.

Gracey, A. Y., Chaney, M. L, Boomhower, J. P., Tyburczy, W. R., Connor, K, Somero, G. N. 2008. Rhythms of gene expression in a fluctuating intertidal environment. *Current biology*. *18,* 1501-1507.

Habermann, B., Oegema, J., Sunyaev, S., and Shevchenko, A. (2004). The power and the limitations of cross-species protein identification by mass spectrometry-driven sequence similarity searches. Molecular & Cellular Proteomics *3*, 238–249.

Han, Z., Sun, J., Zhang, Y., He, F., Xu, Y., Matsumura, K., He, L.-S., Qiu, J.-W., Qi, S.-H., and Qian, P.-Y. (2013). iTRAQ-based proteomic profiling of the barnacle *Balanus amphitrite* in response to the antifouling compound meleagrin. Journal of Proteome Research *12*, 2090–2100.

Harley, C. D., Randall Hughes, A., Hultgren, K. M., Miner, B. G., Sorte, C. J., Thornber, C. S., Rodriguez, L. F., Tomanek, L., and Williams, S. L. (2006). The impacts of climate change in coastal marine systems. Ecology Letters *9*, 228–241.

Harry, J. L., Wilkins, M. R., Herbert, B. R., Packer, N. H., Gooley, A. A., and Williams, K. L. (2000). Proteomics: capacity versus utility. Electrophoresis *21*, 1071–1081.

Hauck, M., and Sweijd, N. (1999). A case study of abalone poaching in South Africa and its impact on fisheries management. ICES Journal of Marine Science: Journal Du Conseil *56*, 1024–1032.

Hawkridge, A. M., and Muddiman, D. C. (2009). Mass spectrometry-based biomarker discovery: toward a global proteome index of individuality. Annual Review of Analytical Chemistry (Palo Alto, Calif.) *2*, 265.

Hecht, T. (1994). Behavioural thermoregulation of the abalone, *Haliotis midae*, and the implications for intensive culture. Aquaculture *126*, 171–181.

Hégaret, H., Wikfors, G. H., and Soudant, P. (2003a). Flow cytometric analysis of haemocytes from eastern oysters, *Crassostrea virginica*, subjected to a sudden temperature elevation: II. Haemocyte functions: aggregation, viability, phagocytosis, and respiratory burst. Journal of Experimental Marine Biology and Ecology *293*, 249–265.

Hégaret, H., Wikfors, G. H., and Soudant, P. (2003b). Flow-cytometric analysis of haemocytes from eastern oysters, *Crassostrea virginica*, subjected to a sudden temperature elevation I. Haemocyte types and morphology. Journal of Experimental Marine Biology and Ecology *293*, 249–265.

Hepler, J. R., and Gilman, A. G. (1992). G proteins. Trends in Biochemical Sciences *17*, 383–387.

Hetz, C. (2012). The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nature Reviews Molecular Cell Biology *13*, 89–102.

Hine, P. (1999). The inter-relationships of bivalve haemocytes. Fish & Shellfish Immunology *9*, 367–385.

Hitchcock, S. E. (1977). Regulation of motility in nonmuscle cells. J. Cell Biol *74*, 15.

Hoarau, P., Garello, G., Gnassia-Barelli, M., Roméo, M., and Girard, J.-P. (2004). Effect of three xenobiotic compounds on Glutathione*S*-Transferase in the clam *Ruditapes decussatus*. Aquatic Toxicology *68*, 87–94.

Hofmann, G., and Somero, G. (1995). Evidence for protein damage at environmental temperatures: seasonal changes in levels of ubiquitin conjugates and hsp70 in the intertidal mussel *Mytilus trossulus*. Journal of Experimental Biology *198*, 1509–1518.

Hooper, C., Day, R., Slocombe, R., Benkendorff, K., and Handlinger, J. (2011). Effect of movement stress on immune function in farmed Australian abalone (hybrid *Haliotis laevigata* and *Haliotis rubra*). Aquaculture *315*, 348–354.

Hooper, C., Day, R., Slocombe, R., Benkendorff, K., Handlinger, J., and Goulias, J. (2014). Effects of severe heat stress on immune function, biochemistry and histopathology in farmed Australian abalone (hybrid *Haliotis laevigata* x *Haliotis rubra*). Aquaculture.

Hooper, C., Day, R., Slocombe, R., Handlinger, J., and Benkendorff, K. (2007a). Stress and immune responses in abalone: limitations in current knowledge and investigative methods based on other models. Fish & Shellfish Immunology *22*, 363–379.

Hooper, C., Hardy-Smith, P., and Handlinger, J. (2007b). Ganglioneuritis causing high mortalities in farmed Australian abalone (*Haliotis laevigata* and *Haliotis rubra*). Australian Veterinary Journal *85*, 188–193.

Hooper, S. L., and Thuma, J. B. (2005). Invertebrate muscles: muscle specific genes and proteins. Physiological Reviews *85*, 1001–1060.

Horwitz, J., Ding, L., Vasiliou, V., Cantore, M., and Piatigorsky, J. (2006). Scallop lens Omega-crystallin (ALDH1A9): A novel tetrameric aldehyde dehydrogenase. Biochemical and Biophysical Research Communications *348*, 1302–1309.

Hubner, N. C., Ren, S., and Mann, M. (2008). Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. Proteomics *8*, 4862–4872.

Humphries, J. E., and Yoshino, T. P. (2003). Cellular receptors and signal transduction in molluscan hemocytes: connections with the innate immune system of vertebrates. Integrative and Comparative Biology *43*, 305–312.

Huot, J., Houle, F., Marceau, F., and Landry, J. (1997). Oxidative stress-induced actin reorganization mediated by the p38 mitogen-activated protein kinase/heat shock protein

27 pathway in vascular endothelial cells. Circulation Research *80*, 383–392.

Jones, A. R., and Hubbard, S. J. (2010). An introduction to proteome bioinformatics. In: Proteome Bioinformatics, Springer, 1–5.

Kawabe, S., and Yokoyama, Y. (2010). Molecular cloning of *calnexin* and *calreticulin* in the Pacific oyster *Crassostrea gigas* and its expression in response to air exposure. Marine Genomics *3*, 19–27.

Kennedy, S. (2002). The role of proteomics in toxicology: identification of biomarkers of toxicity by protein expression analysis. Biomarkers *7*, 269–290.

Kennedy, T. E., Kuhl, D., Barzilai, A., Sweatt, J. D., and Kandel, E. R. (1992). Long-term sensitization training in *Aplysia* leads to an increase in calreticulin, a major presynaptic calcium-binding protein. Neuron *9*, 1013–1024.

Kiang, J. G., and Tsokos, G. C. (1998). Heat shock protein 70 kDa: molecular biology, biochemistry, and physiology. Pharmacology & Therapeutics *80*, 183–201.

Kong, X., Wang, G., and Li, S. (2007). Antioxidation and ATPase activity in the gill of mud crab *Scylla serrata* under cold stress. Chinese Journal of Oceanology and Limnology *25*, 221–226.

Kültz, D. (2003) Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. Journal of Experimental Biology *206*, 3119-3124.

Kültz, D. (2005) Molecular and evolutionary basis of the cellular stress response. Annu. Rev. Physiol. *67*: 225-257.

Kregel, K. C. (2002). Invited review: heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. Journal of Applied Physiology *92*, 2177–2186.

Kubota, K., Kosaka, T., and Ichikawa, K. (2009). Shotgun Protein Analysis by Liquid Chromatography-Tandem Mass Spectrometry. Methods in Molecular Biology (Clifton, NJ) *519*, 483.

Lacoste, A., De Cian, M.-C., Cueff, A., and Poulet, S. A. (2001). Noradrenaline and alpha-adrenergic signaling induce the hsp70 gene promoter in mollusc immune cells. Journal of Cell Science *114*, 3557–3564.

Lacoste, A., Malham, S. K., Gélébart, F., Cueff, A., and Poulet, S. A. (2002). Stress-induced immune changes in the oyster *Crassostrea gigas*. Developmental & Comparative Immunology *26*, 1–9.

Lam, A., and Galione, A. (2013). The endoplasmic reticulum and junctional membrane communication during calcium signalling. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research.

Lama, J. L., Bell, R. A., and Storey, K. B. (2013). Glucose-6-phosphate dehydrogenase regulation in the hepatopancreas of the anoxia-tolerant marine mollusc, *Littorina littorea*. PeerJ *1*, e21.

Larkindale, J., and Knight, M. R. (2002). Protection against heat stress-induced oxidative damage in *Arabidopsis* involves calcium, abscisic acid, ethylene, and salicylic acid. Plant Physiology *128*, 682–695.

Lee, D., Singaravelu, G., Park, B. J., and Ahnn, J. (2007). Differential Requirement of Unfolded Protein Response Pathway for Calreticulin Expression in *Caenorhabditis elegans*. Journal of Molecular Biology *372*, 331–340.

Lee, K.-K., Liu, P.-C., Chen, Y.-C., and Huang, C.-Y. (2001). The implication of ambient temperature with the outbreak of vibriosis in cultured small abalone *Haliotis diversicolor supertexta* Lischke. Journal of Thermal Biology *26*, 585–587.

Lesser, M. P. (2006). Oxidative stress in marine environments: biochemistry and physiological ecology. Annu. Rev. Physiol. *68*, 253–278.

Li, J., He, Q., Sun, H., and Liu, X. (2012). Acclimation-dependent expression of heat shock protein 70 in Pacific abalone (*Haliotis discus hannai* Ino) and its acute response to thermal exposure. Chinese Journal of Oceanology and Limnology *30*, 146–151.

Liebler, D. C. (2002). Introduction to proteomics: tools for the new biology, Springer.

Lindström, M. S. (2009). Emerging functions of ribosomal proteins in gene-specific transcription and translation. Biochemical and Biophysical Research Communications *379*, 167–170.

Linscheid, M. W., Ahrends, R., Pieper, S., and Kühn, A. (2009). Liquid Chromatography-Mass Spectrometry-Based Quantitative Proteomics. In: Proteomics, Springer, 189–205.

Liska, A. J., Popov, A. V., Sunyaev, S., Coughlin, P., Habermann, B., Shevchenko, A., Bork, P., Karsenti, E., and Shevchenko, A. (2004). Homology-based functional proteomics by mass spectrometry: Application to the Xenopus microtubule-associated proteome. Proteomics *4*, 2707–2721.

Lodish, H., and Zipursky, S. L. (2001). Molecular cell biology, Wiley Online Library.

Loker, E. S. (2010). Gastropod immunobiology. In: Invertebrate Immunity, Springer, 17–43.

López, J., Marina, A., Vazquez, J., and Alvarez, G. (2002). A proteomic approach to the study of the marine mussels *Mytilus edulis* and *M. galloprovincialis*. Marine Biology *141*, 217–223.

Lottspeich, F. (2009). Introduction to proteomics. In: Proteomics, Springer, 3–10.

Luana, W., Li, F., Wang, B., Zhang, X., Liu, Y., and Xiang, J. (2007). Molecular characteristics and expression analysis of calreticulin in Chinese shrimp *Fenneropenaeus*

chinensis. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *147*, 482–491.

Lushchak, V. I. (2011). Environmentally induced oxidative stress in aquatic animals. Aquatic Toxicology *101*, 13–30.

Macey, B., Christison, K., and Mouton, A. (2011). *Halioticida noduliformans* isolated from cultured abalone (*Haliotis midae*) in South Africa. Aquaculture *315*, 187–195.

Malagoli, D., Casarini, L., Sacchi, S., and Ottaviani, E. (2007). Stress and immune response in the mussel *Mytilus galloprovincialis*. Fish & Shellfish Immunology *23*, 171–177.

Malham, S. K., Lacoste, A., Gelebart, F., Cueff, A., and Poulet, S. A. (2003). Evidence for a direct link between stress and immunity in the mollusc *Haliotis tuberculata*. Journal of Experimental Zoology Part A: Comparative Experimental Biology *295*, 136–144.

Mann, M., Hendrickson, R. C., and Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. Annual Review of Biochemistry *70*, 437–473.

Mann, M., and Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. Nature Biotechnology *21*, 255–261.

Markl, J. (2013). Evolution of Molluscan Hemocyanin Structures. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics.

Marmaras, V. J., and Lampropoulou, M. (2009). Regulators and signalling in insect haemocyte immunity. Cellular Signalling *21*, 186–195.

Martello, L. B., Friedman, C. S., and Tjeerdema, R. S. (2000). Combined effects of pentachlorophenol and salinity stress on phagocytic and chemotactic function in two species of abalone. Aquatic Toxicology *49*, 213–225.

Martin, G. G., Oakes, C. T., Tousignant, H. R., Crabtree, H., and Yamakawa, R. (2007). Structure and function of haemocytes in two marine gastropods, *Megathura crenulata* and *Aplysia californica*. Journal of Molluscan Studies *73*, 355–365.

May, R. C., and Machesky, L. M. (2001). Phagocytosis and the actin cytoskeleton. Journal of Cell Science *114*, 1061–1077.

McDonagh, B., and Sheehan, D. (2007). Effect of oxidative stress on protein thiols in the blue mussel Mytilus edulis: proteomic identification of target proteins. Proteomics *7*, 3395–3403.

McDonald, W. H., and Yates, J. R. (2002). Shotgun proteomics and biomarker discovery. Disease Markers *18*, 99–105.

Meistertzheim, A.-L., Tanguy, A., Moraga, D., and Thébault, M.-T. (2007). Identification of differentially expressed genes of the Pacific oyster *Crassostrea gigas* exposed to prolonged thermal stress. FEBS Journal *274*, 6392–6402.

Merino, G. *et al.* (2012). Can marine fisheries and aquaculture meet fish demand from a growing human population in a changing climate? Global Environmental Change *22*, 795–806.

Michalak, M., Corbett, E. F., Mesaeli, N., Nakamura, K., and Opas, M. (1999). Calreticulin: one protein, one gene, many functions. Biochemical Journal *344*, 281.

Moehl, J., and Machena, C. (2000). African Aquaculture: A Regional Summary with Emphasis on Sub-Saharan Africa.

Mohanty, P., Kreslavski, V. D., Klimov, V. V., Los, D. A., Mimuro, M., Carpentier, R., and Allakhverdiev, S. I. (2012). Heat Stress: Susceptibility, Recovery and Regulation. Photosynthesis, 251–274.

Morimoto, R. I. (1998). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. Genes & Development *12*, 3788–3796.

Morimoto, R. I., Kline, M. P., Bimston, D. N., and Cotto, J. J. (1996). The heat-shock response: regulation and function of heat-shock proteins and molecular chaperones. Essays in Biochemistry *32*, 17–29.

Muralidharan, S., Thompson, E., Raftos, D., Birch, G., and Haynes, P. A. (2012). Quantitative proteomics of heavy metal stress responses in Sydney rock oysters. Proteomics *12*, 906–921.

Nesatyy, V. J., and Suter, M. J.-F. (2007). Proteomics for the analysis of environmental stress responses in organisms. Environmental Science & Technology *41*, 6891–6900.

Nesvizhskii, A. I. (2007). Protein identification by tandem mass spectrometry and sequence database searching. In: Methods in Molecular Biology, Springer, 87.

Ni, M., and Lee, A. S. (2007). ER chaperones in mammalian development and human diseases. FEBS Letters *581*, 3641–3651.

Nikapitiya, C., De Zoysa, M., Whang, I., Kim, S.-J., Choi, C. Y., Lee, J.-S., and Lee, J. (2010). Characterization and expression analysis of EF hand domain-containing calcium-regulatory gene from disk abalone: Calcium homeostasis and its role in immunity. Fish & Shellfish Immunology *29*, 334–342.

Osovitz, C. J., and Hofmann, G. E. (2005). Thermal history-dependent expression of the *hsp70* gene in purple sea urchins: Biogeographic patterns and the effect of temperature acclimation. Journal of Experimental Marine Biology and Ecology *327*, 134–143.

Ottaviani, E., Franceschi, C. 1997. The invertebrate phagocytic immunocyte: clues to a common evolution of immune and neuroendochrine systems. Immunology today. *18(4)*, 169-174.

Pallares, S., Arribas, P., Cespedes, V., Millan, A., and Velasco, J. (2012). Lethal and sublethal behavioural responses of saline water beetles to acute heat and osmotic stress.

Ecological Entomology *37*, 508–520.

Peracino, B., Borleis, J., Jin, T., Westphal, M., Schwartz, J. M., Wu, L., Bracco, E., Gerisch, G., Devreotes, P., and Bozzaro, S. (1998). G protein beta subunit-null mutants are impaired in phagocytosis and chemotaxis due to inappropriate regulation of the actin cytoskeleton. The Journal of Cell Biology *141*, 1529–1537.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research *29*, e45–e45.

Pietrogrande, M. C., Marchetti, N., Dondi, F., and Righetti, P. G. (2003). Spot overlapping in two-dimensional polyacrylamide gel electrophoresis maps: Relevance to proteomics. Electrophoresis *24*, 217–224.

Prohászka, Z., and Füst, G. (2004). Immunological aspects of heat-shock proteins—the optimum stress of life. Molecular Immunology *41*, 29–44.

Pushpamali, W. A., De Zoysa, M., Kang, H.-S., Oh, C. H., Whang, I., Kim, S. J., and Lee, J. (2008). Comparative study of two thioredoxin peroxidases from disk abalone (*Haliotis discus discus*): Cloning, recombinant protein purification, characterization of antioxidant activities and expression analysis. Fish & Shellfish Immunology *24*, 294–307.

Qiu, Y., and Michalak, M. (2009). Transcriptional control of the calreticulin gene in health and disease. The International Journal of Biochemistry & Cell Biology *41*, 531–538.

Raemaekers, S., Hauck, M., Bürgener, M., Mackenzie, A., Maharaj, G., Plagányi, É. E., and Britz, P. J. (2011). Review of the causes of the rise of the illegal South African abalone fishery and consequent closure of the rights-based fishery. Ocean & Coastal Management *54*, 433–445.

Regoli, F., and Giuliani, M. E. (2013). Oxidative pathways of chemical toxicity and oxidative stress biomarkers in marine organisms. Marine Environmental Research.

Ridley, A. J. (2001). Rho family proteins: coordinating cell responses. Trends in Cell Biology *11*, 471–477.

Robinson, N., Smith, B., Cooke, I., and Strugnell, J. (2012). A snail's pace: A preliminary analysis of the effects of stress and genetics on movement of *Haliotis*. Aquaculture.

Rodrigues, P. M., Silva, T. S., Dias, J., and Jessen, F. (2012). Proteomics in aquaculture: applications and trends. Journal of Proteomics *75*, 4325–4345.

Rodriguez-Ortega, M. J., Grøsvik, B. E., Rodriguez-Ariza, A., Goksøyr, A., and López-Barea, J. (2003). Changes in protein expression profiles in bivalve molluscs (*Chamaelea gallina*) exposed to four model environmental pollutants. Proteomics *3*, 1535–1543.

Rosner, A., and Rinkevich, B. (2007). The DDX3 subfamily of the DEAD box helicases: divergent roles as unveiled by studying different organisms and in vitro assays. Current Medicinal Chemistry *14*, 2517–2525.

Sales, J., and Britz, P. (2001). Research on abalone (*Haliotis midae* L.) cultivation in South Africa. Aquaculture Research *32*, 863–874.

Samali, A., and Orrenius, S. (1998). Heat shock proteins: regulators of stress response and apoptosis. Cell Stress & Chaperones *3*, 228.

Sanchez, B. C., Ralston-Hooper, K., and Sepúlveda, M. S. (2011). Review of recent proteomic applications in aquatic toxicology. Environmental Toxicology and Chemistry *30*, 274–282.

Santoro, M. G. (2000). Heat shock factors and the control of the stress response. Biochemical Pharmacology *59*, 55–63.

Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., Pujos-Guillot, E., Verheij, E., Wishart, D., and Wopereis, S. (2009). Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. Metabolomics *5*, 435–458.

Schröder, M. (2008). Endoplasmic reticulum stress responses. Cellular and Molecular Life Sciences *65*, 862–894.

Schulze, W. X., and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. Annual Review of Plant Biology *61*, 491–516.

Seidler, J., Zinn, N., Boehm, M. E., and Lehmann, W. D. (2010). De novo sequencing of peptides by MS/MS. Proteomics *10*, 634–649.

Serafini, L., Hann, J. B., Kültz, D., and Tomanek, L. (2011). The proteomic response of sea squirts (genus *Ciona*) to acute heat stress: A global perspective on the thermal stability of proteins. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics *6*, 322–334.

Sheehan, D., and McDonagh, B. (2008). Oxidative stress and bivalves: a proteomic approach. Invertebrate Survival Journal *5*, 110–123.

Shevchenko, A., Valcu, C.-M., and Junqueira, M. (2009). Tools for exploring the proteomosphere. Journal of Proteomics *72*, 137–144.

Siddiqui, N. I., Akosung, R. F., and Gielens, C. (2006). Location of intrinsic and inducible phenoloxidase activity in molluscan hemocyanin. Biochemical and Biophysical Research Communications *348*, 1138–1144.

De Silva, S. S. (2012). Aquaculture: a newly emergent food production sector—and perspectives of its impacts on biodiversity and conservation. Biodiversity and Conservation *21*, 3187–3220.

Silvestre, F., Linares-Casenave, J., Doroshov, S. I., and Kültz, D. (2010). A proteomic analysis of green and white sturgeon larvae exposed to heat stress and selenium. Science of the Total Environment *408*, 3176–3188.

Slattery, M., Ankisetty, S., Corrales, J., Marsh-Hunkin, K. E., Gochfeld, D. J., Willett, K. L., and Rimoldi, J. M. (2012). Marine Proteomics: A Critical Assessment of an Emerging Technology. Journal of Natural Products.

Small, J.-V., Rottner, K., Hahne, P., Anderson, K. I., and others (1999). Visualising the actin cytoskeleton. Microscopy Research and Technique *47*, 3–17.

Song, X., Bandow, J., Sherman, J., Baker, J. D., Brown, P. W., McDowell, M. T., and Molloy, M. P. (2008). iTRAQ experimental design for plasma biomarker discovery. Journal of Proteome Research *7*, 2952–2958.

Sørensen, J. G., Kristensen, T. N., and Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. Ecology Letters *6*, 1025–1037.

Sundström, A. (2012). Corruption and regulatory compliance: Experimental findings from South African small-scale fisheries. Marine Policy.

Supek, F., Bovsnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800.

Suzuki, N., and Mittler, R. (2006). Reactive oxygen species and temperature stresses: a delicate balance between signaling and destruction. Physiologia Plantarum *126*, 45–51.

Swerdlow, R. D., Ebert, R. F., Lee, P., Bonaventura, C., and Miller, K. I. (1996). Keyhole limpet hemocyanin: structural and functional characterization of two different subunits and multimers. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *113*, 537–548.

Tabb, D. L. (2007). What's driving false discovery rates? Journal of Proteome Research *7*, 45–46.

Tarr, R. (1995). Growth and movement of the South African abalone Haliotis midae: a reassessment. Marine and Freshwater Research *46*, 583–590.

Taylor, C. F. *et al.* (2007). The minimum information about a proteomics experiment (MIAPE). Nature Biotechnology *25*, 887–893.

Ten Doeschate, K., and Coyne, V. (2008). Improved growth rate in farmed *Haliotis midae* through probiotic treatment. Aquaculture *284*, 174–179.

Thompson, E. L., Taylor, D. A., Nair, S. V., Birch, G., Haynes, P. A., and Raftos, D. A. (2011). A proteomic analysis of the effects of metal contamination on Sydney Rock Oyster (*Saccostrea glomerata*) haemolymph. Aquatic Toxicology *103*, 241–249.

Timms, J. F., and Cutillas, P. R. (2010). Overview of quantitative LC-MS techniques for proteomics and activitomics. Methods Mol. Biol *658*, 19–45.

Tirard, C., Grossfeld, R., Levine, J., and Kennedy-Stoskopf, S. (1995). Effect of hyperthermia in vitro on stress protein synthesisand accumulation in oyster haemocytes.

Fish & Shellfish Immunology *5*, 9–25.

Tkácová, J., and Angelovicová, M. (2012). Heat Shock Proteins (HSPs): a Review. Cell *17*, 18.

Tomanek, L. (2010). Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. Journal of Experimental Biology *213*, 971–979.

Tomanek, L. (2011). Environmental proteomics: changes in the proteome of marine organisms in response to environmental stress, pollutants, infection, symbiosis, and development. Annual Review of Marine Science *3*, 373–399.

Tomanek, L., and Zuzow, M. J. (2010). The proteomic response of the mussel congeners Mytilus galloprovincialis and M. trossulus to acute heat stress: implications for thermal tolerance limits and metabolic costs of thermal stress. The Journal of Experimental Biology *213*, 3559–3574.

Travers, M. A., Mirella da Silva, P., Le Goic, N., Marie, D., Donval, A., Huchette, S., Koken, M., and Paillard, C. (2008). Morphologic, cytometric and functional characterisation of abalone (*Haliotis tuberculata*) haemocytes. Fish & Shellfish Immunology *24*, 400–411.

Travers, M., Shin, Y.-J., Jennings, S., and Cury, P. (2007). Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. Progress in Oceanography *75*, 751–770.

Travers, M.-A., Meistertzheim, A.-L., Cardinaud, M., Friedman, C. S., Huchette, S., Moraga, D., and Paillard, C. (2010). Gene expression patterns of abalone, *Haliotis tuberculata*, during successive infections by the pathogen *Vibrio harveyi*. Journal of Invertebrate Pathology *105*, 289–297.

Troell, M., Robertson-Andersson, D., Anderson, R., Bolton, J., Maneveldt, G., Halling, C., and Probyn, T. (2006). Abalone farming in South Africa: an overview with perspectives on kelp resources, abalone feed, potential for on-farm seaweed production and socio-economic importance. Aquaculture *257*, 266–281.

United-Nations (2004). World Population to 2300, New York: United Nations, Department of Economic and Social Affairs.

Unwin, R. D. (2010). Quantification of proteins by iTRAQ. Methods Mol. Biol *658*, 205–215.

Unwin, R. D., Carrick, E., and Whetton, A. D. (2012). Relative Quantification Mass Spectrometry Using iTRAQ Isobaric Tags. New Frontiers of Network Analysis in Systems Biology, 77–95.

Unwin, R. D., Griffiths, J. R., and Whetton, A. D. (2010). Simultaneous analysis of relative protein expression levels across multiple samples using iTRAQ isobaric tags with 2D nano

LC-MS/MS. Nature Protocols *5*, 1574–1582.

Vetter, I. R., and Wittinghofer, A. (2001). The guanine nucleotide-binding switch in three dimensions. Science *294*, 1299–1304.

Vosloo, D., van Rensburg, L., and Vosloo, A. (2013a). Oxidative stress in abalone: The role of temperature, oxygen and l-proline supplementation. Aquaculture *416*, 265–271.

Vosloo, D., and Vosloo, A. (2010). Response of cold-acclimated, farmed South African abalone (*Haliotis midae*) to short-term and long-term changes in temperature. Journal of Thermal Biology *35*, 317–323.

Vosloo, D., Vosloo, A., Morillion, E. J., Samuels, J. N., and Sommer, P. (2013b). Metabolic readjustment in juvenile South African abalone (*Haliotis midae*) acclimated to combinations of temperature and dissolved oxygen levels. Journal of Thermal Biology *38*, 458–466.

Waldbusser, G. G., Brunner, E. L., Haley, B. A., Hales, B., Langdon, C. J., and Prahl, F. G. (2013). A developmental and energetic basis linking larval oyster shell formation to acidification sensitivity. Geophysical Research Letters, 1–6.

Wan, Q., Whang, I., and Lee, J. (2012). Molecular and functional characterization of HdHSP20: A biomarker of environmental stresses in disk abalone *Haliotis discus discus*. Fish & Shellfish Immunology *33*, 48–59.

Wan, Q., Whang, I., Lee, J.-S., and Lee, J. (2009). Novel omega glutathione S-transferases in disk abalone: characterization and protective roles against environmental stress. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology *150*, 558–568.

Wang, K.-J., Ren, H.-L., Xu, D.-D., Cai, L., and Yang, M. (2008). Identification of the up-regulated expression genes in hemocytes of variously colored abalone (*Haliotis diversicolor* Reeve, 1846) challenged with bacteria. Developmental & Comparative Immunology *32*, 1326–1347.

Wang, N., and Li, L. (2008). Exploring the precursor ion exclusion feature of liquid chromatography- electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. Analytical Chemistry *80*, 4696–4710.

Wang, P., Bouwman, F. G., and Mariman, E. (2009). Generally detected proteins in comparative proteomics-A matter of cellular stress response? Proteomics *9*, 2955–2966.

Wang, X., Zhou, Z., Yang, A., Dong, Y., Chen, Z., Guan, X., Jiang, B., and Wang, B. (2013). Molecular Characterization and Expression Analysis of Heat Shock Cognate 70 After Heat Stress and Lipopolysaccharide Challenge in Sea Cucumber (*Apostichopus japonicus*). Biochemical Genetics, 1–15.

Wang, Y., Hu, M., Cheung, S., Shin, P., Lu, W., and Li, J. (2012). Immune parameter changes of hemocytes in green-lipped mussel *Perna viridis* exposure to hypoxia and

hyposalinity. Aquaculture *356*, 22–29.

Waser, M., Mesaeli, N., Spencer, C., and Michalak, M. (1997). Regulation of calreticulin gene expression by calcium. The Journal of Cell Biology *138*, 547–557.

Wettstein, G., Bellaye, P., Micheau, O., and Bonniaud, P. (2012). Small heat shock proteins and the cytoskeleton: An essential interplay for cell integrity? The International Journal of Biochemistry & Cell Biology *44*, 1680–1686.

Winston, G. W., Moore, M. N., Kirchin, M. A., and Soverchia, C. (1996). Production of reactive oxygen species by Hemocytes from the marine mussel, *Mytilus edulis*: Lysosomal localization and effect of xenobiotics. Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology *113*, 221–229.

Wiśnewski, J. R., Zougman, A., Nagaraj, N., Mann, M. 2009. Universal sample preparation method for proteome analysis. Nature Methods. *6,* 359-362.

Wu, H., Liu, X., Zhang, X., Ji, C., Zhao, J., and Yu, J. (2013). Proteomic and metabolomic responses of clam *Ruditapes philippinarum* to arsenic exposure under different salinities. Aquatic Toxicology.

Xu, C., and Ma, B. (2006). Software for computational peptide identification from MS-MS data. Drug Discovery Today *11*, 595–600.

Yao, C.-L., and Somero, G. N. (2012). The impact of acute temperature stress on hemocytes of invasive and native mussels (*Mytilus galloprovincialis* and *Mytilus californianus*): DNA damage, membrane integrity, apoptosis and signaling pathways. The Journal of Experimental Biology *215*, 4267–4277.

Yates, J. R. (1998). Mass spectrometry and the age of the proteome. Journal of Mass Spectrometry *33*, 1–19.

Zhang, G. *et al.* (2012a). The oyster genome reveals stress adaptation and complexity of shell formation. Nature *490*, 49–54.

Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G. A., and Ma, B. (2012b). PEAKS DB: De Novo sequencing assisted database search for sensitive and accurate peptide identification. Molecular & Cellular Proteomics *11*.

Zhang, K., and Kaufman, R. J. (2006). The unfolded protein response A stress signaling pathway critical for health and disease. Neurology *66*, S102–S109.

Zhou, J., Cai, Z. H., Li, L., Gao, Y. F., and Hutchinson, T. H. (2010). A proteomics based approach to assessing the toxicity of bisphenol A and diallyl phthalate to the abalone (*Haliotis diversicolor supertexta*). Chemosphere *79*, 595–604.

Zhou, X., Ding, Y., and Wang, Y. (2012). Proteomics: present and future in fish, shellfish and seafood. Reviews in Aquaculture *4*, 11–20.

Zieske, L. R. (2006). A perspective on the use of iTRAQ™ reagent technology for protein complex and profiling studies. Journal of Experimental Botany *57*, 1501–1508.

Zinovieva, R., Tomarev, S., and Piatigorsky, J. (1993). Aldehyde dehydrogenase-derived omega-crystallins of squid and octopus. Specialization for lens expression. Journal of Biological Chemistry *268*, 11449–11455.

De Zoysa, M., Pushpamali, W. A., Whang, I., Kim, S. J., and Lee, J. (2008). Mitochondrial thioredoxin-2 from disk abalone (*Haliotis discus discus*): Molecular characterization, tissue expression and DNA protection activity of its recombinant protein. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *149*, 630–639.

De Zoysa, M., Whang, I., Lee, Y., Lee, S., Lee, J. S., and Lee, J. (2009). Transcriptional analysis of antioxidant and immune defense genes in disk abalone (*Haliotis discus discus*) during thermal, low-salinity and hypoxic stress. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *154*, 387–395.