Gene 569 (2015) 203–217

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

A glycosyl hydrolase family 16 gene is responsible for the endogenous production of β-1,3-glucanases within decapod crustaceans

Stuart M. Linton [a,*], Melissa S. Cameron [a], Michael C. Gray [a], John A. Donald [a], Reinhard Saborowski [b], Martin von Bergen [c], Janina M. Tomm [c], Benjamin J. Allardyce [d]

[a] School of Life and Environmental Sciences, Deakin University, 75 Pigdons Road, Waurn Ponds, Victoria 3216, Australia
[b] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany
[c] Helmholtz Centre for Environmental Research - UFZ, Permoserstraße 15, 04318 Leipzig, Germany
[d] Institute for Frontier Materials, Deakin University, 75 Pigdons Road, Waurn Ponds, Victoria, 3216, Australia

ARTICLE INFO

Article history:
Received 27 November 2014
Received in revised form 18 April 2015
Accepted 22 May 2015
Available online 27 May 2015

Keywords:
β-1,3-glucanase
Laminarinase
Hemicellulase
Land crab
Gecarcoidea natalis
Cherax destructor

ABSTRACT

To identify the gene responsible for the production of a β-1,3-glucanase (laminarinase) within crustacea, a glycosyl hydrolase family 16 (GHF16) gene was sequenced from the midgut glands of the gecarcinid land crab, *Gecarcoidea natalis* and the freshwater crayfish, *Cherax destructor*. An open reading frame of 1098 bp for *G. natalis* and 1095 bp for *C. destructor* was sequenced from cDNA. For *G. natalis* and *C. destructor* respectively, this encoded putative proteins of 365 and 364 amino acids with molecular masses of 41.4 and 41.5 kDa. mRNA for an identical GHF16 protein was also expressed in the haemolymph of *C. destructor*. These putative proteins contained binding and catalytic domains that are characteristic of a β-1,3-glucanase from glycosyl hydrolase family 16. The amino acid sequences of two short 8–9 amino acid residue peptides from a previously purified β-1,3-glucanase from *G. natalis* matched exactly that of the putative protein sequence. This plus the molecular masses of the putative proteins matching that of the purified proteins strongly suggests that the sequences obtained encode for a catalytically active β-1,3-glucanase. A glycosyl hydrolase family 16 cDNA was also partially sequenced from the midgut glands of other amphibious (*Mictyris platycheles* and *Paragrapsus laevis*) and terrestrial decapod species (*Coenobita rugosus*, *Coenobita perlatus*, *Coenobita brevimanus* and *Birgus latro*) to confirm that the gene is widely expressed within this group. There are three possible hypothesised functions and thus evolutionary routes for the β-1,3-glucanase: 1) a digestive enzyme which hydrolyses β-1,3-glucans, 2) an enzyme which cleaves β-1,3-glycosidic bonds within cell walls to release cell contents or 3) an immune protein which can hydrolyse the cell walls of potentially pathogenic micro-organisms.

1. Introduction

Laminarinase or β-1,3-glucanase is a ubiquitous digestive hemicellulase enzyme that is present at high activities within the digestive fluid of aquatic and terrestrial crustaceans (Sova et al., 1970; Suzuki et al., 1987; Omondi and Stark, 1995; Figueiredo et al., 2001; Linton and Greenaway, 2004; Johnston and Freeman, 2005; Figueiredo and Anderson, 2009; Linton et al., 2009). It catalyses the hydrolysis of β-1,3-glycosidic bonds within β-1,3-glucans such as laminarin and callose. Laminarin is the major storage polysaccharide in brown algae (Bull and Chesters, 1966), diatoms (Pesentseva et al., 2008) and protozoans (Piavaux, 1977) while callose is present in the wound tissue of plants (Bacic et al., 1988; Ruiz-Herrera, 1992; Terra and Ferreira, 1994).

Within the decapods, β-1,3-glucanase has been purified and characterised from the midgut glands of the gecarcinid land crab, *Gecarcoidea natalis*, and the freshwater crayfish, *Cherax destructor* (Allardyce and Linton, 2008). *G. natalis* is an herbivorous species that consumes mainly leaf litter (Greenaway and Linton, 1995). It is able to digest substantial amounts of cellulose and hemicellulose (up to 50% of that consumed with a leaf litter diet) using endogenous cellulase and hemicellulase enzymes such as β-1,3-glucanase (Linton and Greenaway, 2007). The β-1,3-glucanase, which is active against laminarin, is thought to exist as a dimer of two 41 kDa subunits (Table 1). It is assumed that it is synthesised endogenously within the midgut gland since high β-1,3-glucanase activities are present in this tissue (Allardyce and Linton, 2008); however, the gene responsible for the endogenous production of this enzyme has yet to be identified in Crustacea. It is likely that the β-1,3-glucanase is the product of a glycosyl

Abbreviations: β-1,3-glucans, Glucose polymer in which the monosaccharides are joined by β-1,3-glycosidic bonds; cDNA, Complementary DNA; GAPDH, Glyceraldehyde 3 phosphate dehydrogenase; GHF16, Glycosyl hydrolase family 16; LC-MS/MS, Liquid chromatography- mass spectrometry; LGBP, Lipopolysaccharide and β-1,3-glucan binding protein; MS-222, Ethyl 3-aminobenzoate methanesulfonate; Nano-HPLC or UPLC, Ultra performance liquid chromatography; PCR, Polymerase chain reaction; RACE, Rapid amplification of cDNA ends.

* Corresponding author.
E-mail address: stuart.linton@deakin.edu.au (S.M. Linton).

http://dx.doi.org/10.1016/j.gene.2015.05.056
0378-1119/

Table 1
Molecular mass and glycosyl hydrolase family of the β-1,3-glucanase (laminarinase) characterised in various invertebrates.

Species	Glycosyl hydrolase family	Estimated molecular mass (kDa)	Accession number	Reference
Phylum Mollusca				
Class: Gastropoda				
Haliotis discus hannai (abalone)	16	33	AB488493	Kumagai and Ojima (2009)
Haliotis tuberculata (abalone)		60		Lépagnol-Descamps et al. (1998)
Class: Bivalvia				
Spisula sachalinensis (clam)	16	38	AY308829	Kozhemyako et al. (2004)
Chlamys albidus (scallop)	16	37	DQ093347	Kovalchuk et al. (2009)
Perna viridis (mussel)	16	50	FJ623758	Zakharenko et al. (2011)
Mizuhopecten yessoensis (scallop)	16	36	AY848857	Kovalchuk et al. (2006)
Phylum: *Arthropoda*, *Subphylum*: *Hexapoda*, *Class*: *Insecta*				
Tenebrio molitor (meal worm)	16	50	ACS36221	Genta et al. (2009)
Cryptopygus antarcticus (Antartic springtail)	16	29.9	EU559744	Song et al. (2010)
Spodoptera frugiperda (fall armyworm)	16	37.5	EF641300	Bragatto et al. (2010)
Phylum: *Arthropoda*, *Subphylum*: *Crustacea*, *Class*: *Malacostracta*, *Order*: *Decapoda*				
Gecarcoidea natalis		41, 71		Allardyce and Linton (2008)
Cherax destructor		41, 71		Allardyce and Linton (2008)
Phylum: *Echinodermata*, *Class*: *Holothuroidea*				
Stichopus japonicus	16	37.5		Zhu et al. (2008)

hydrolase family 16 (GHF16) gene given that this gene has been confirmed to produce a similar enzyme (of 30–50 kDa) in other invertebrates (Table 1). Interestingly, although a digestive β-1,3-glucanase gene has not yet been identified, crustaceans do possess a similar GHF16 gene that produces a lipopolysaccharide and β-1,3-glucan binding protein (LGBP) (Lee et al., 2000; Sritunyalucksana et al., 2002; Du et al., 2007; Lin et al., 2008; Amparyup et al., 2012). This protein is primarily expressed in the haemocytes and is believed to play an immune role. Specifically, LGBPs bind lipopolysaccharides and β-1,3-glucans and this, via the prophenol oxidase system, stimulates an immune response (Lee et al., 2000; Sritunyalucksana and Söderhäll, 2000; Amparyup et al., 2012).

Although the lipopolysaccharide and β-1,3-glucan binding proteins have quite a different role than the digestive β-1,3-glucanases, there are some clues to suggest that the two classes of proteins and thus their genes are similar. The crustacean GHF16 gene produces a protein with both a catalytic and a binding domain, as seen in β-1,3-glucanases (Lee et al., 2000; Amparyup et al., 2012) and thus is classed as a glycosyl hydrolase. It has been suggested that the catalytic domain is inactive and therefore the protein lacks activity; however, this conclusion is based on a different protein that was initially identified as LGBP. The initial protein described by Cerenius et al. (1994) consists of 1339 amino acids and has a putative molecular mass of 152 kDa. It possesses a β-glucan binding domain but no catalytic domain. In contrast, numerous subsequent studies on crustaceans have since identified LGBPs of 349–376 amino acids (approximately 36–41 kDa) with both catalytic and binding domains (Lee et al., 2000; Roux et al., 2002; Sritunyalucksana et al., 2002; Lin et al., 2008; Liu et al., 2009; Yeh et al., 2009; Zhao et al., 2009; Amparyup et al., 2012). The size discrepancy and presence of a catalytic domain in these proteins suggest that the more recently identified proteins are not, in fact, related to that described by Cerenius et al., 1994. In addition, a number of studies have detected expression of a GHF16 protein within the midgut gland of the crustaceans *Eriocheir sinensis*, *Fenneropenaeus chinensis*, *Litopenaeus vannamei* and *Penaeus stylirostris* (Gross et al., 2001; Liu et al., 2009; Zhao et al., 2009) (Roux et al., 2002). It is therefore suggested that, like the GHF16 enzymes expressed in the hepatopancreas of molluscs (Table 1), these proteins may actually possess β-1,3-glucanase activity.

The aim of this study was to establish if a GHF16 gene expressed in the midgut gland of decapod crustaceans was responsible for producing an active β-1,3-glucanase that is secreted into the digestive fluid. To confirm this, the GHF16 cDNA derived from the midgut gland of *G. natalis* and *C. destructor* was sequenced. This sequence was then matched to the characteristics of a β-1,3-glucanase that had been previously purified and characterised from these species (Allardyce and Linton, 2008). To establish the link between the gene and protein, short peptides of β-1,3-glucanase purified from *G. natalis* were sequenced and compared to the putative amino acid sequences. The GHF16 cDNA was also partially sequenced in a range of distantly related decapod crustaceans to establish that crustaceans generally possess and express such a gene. Taken together, this evidence would explain the β-1,3-glucanase activity identified in the digestive fluid and midgut glands of numerous crustaceans.

2. Materials and methods

2.1. Strategy for sequencing β-1,3-glucanase cDNA from *C. destructor* and *G. natalis*

A GHF16 β-1,3-glucanase (laminarinase) cDNA was sequenced from cDNA derived from the midgut gland of the Gecarcinid land crab, *G. natalis* and the freshwater crayfish, *C. destructor*. To do this, crabs were euthanised and midgut gland, muscle and gill tissue taken. The midgut gland is a digestive tissue responsible for the production of digest enzymes. In contrast, the muscle and gills are not involved in digestion, and thus were taken as control tissues. β-1,3-glucanase is likely to be expressed in the midgut gland but not in the muscle and gill. Total RNA was extracted from each tissue and used to synthesise cDNA. Degenerate primers, designed from the conserved regions of β-1,3-glucanase sequences, were used to partially amplify and sequence the β-1,3-glucanase (laminarinase) cDNA from both *G. natalis* and *C. destructor*. From these partial sequences, sequence specific primers were designed for 3′ and 5′ rapid amplification of cDNA ends (RACE). RACE PCR was then used to amplify and sequence the 3′ and 5′ ends of the β-1,3-glucanase cDNA. Finally, the entire open reading frame sequence was determined in three replicate animals. In addition to the nucleotide sequence, the protein sequence of a number of short fragments of the β-1,3-glucanase (laminarinase) purified previously from *G. natalis* was determined using Orbitrap mass spectrometry. These sequences were compared to the putative amino acid sequence to support the conclusion that the nucleotide sequence presented encoded for the enzyme that had been previously purified and characterised. To confirm that crustaceans more broadly possess and express a GHF16 β-1,3-glucanase, the cDNA was also partially sequenced in a range of amphibious and terrestrial decapods (*Mictyris platycheles*, *Paragrapsus laevis*, *Coenobita perlatus*, *Coenobitarugosus*, *Coenobita brevimanus* and *Birgus latro*).

2.2. Collection and maintenance of animals

Crayfish, *C. destructor* were purchased from a local aquarium supplier; *M. platycheles* and *P. laevis* were collected from the Barwon river estuary (Victoria, Australia); Christmas Island species (*G. natalis*, *C. perlatus*, *C. rugosus*and *C. brevimanus*) were collected from the rainforest on Christmas Island. All species were transported back to the laboratory (Deakin laboratory for *C. destructor*, *M. platycheles* and *P. laevis*; Parks Australia laboratory for Christmas Island species), euthanized and their tissues taken. Tissues taken from the Christmas Island species were stored in RNA later (Ambion #AM7020) and airfreighted back to the laboratory at Deakin. For *B. latro*, tissue was taken from animals which had been recently killed by vehicles driving on the Island's roads.

2.3. Extraction of tissue and total RNA

Animals were initially anaesthetised; to do this, animals were injected with the anaesthetic MS-222 (50 mg mL^{-1} at a dose of 250 mg kg^{-1}) into the infrabranchial sinus, and monitored until there was no righting reflex (ability of the animals to turn themselves over). The animals were then quickly killed by the destruction of the brain and subesophageal ganglion. Midgut gland, muscle and gill tissue were quickly removed with instruments which had been treated with RNAase Zap (Ambion #AM9780). For all tissues except for the midgut gland from *G. natalis* and *B. latro*, between 100 and 250 mg of tissue was homogenised in 1–2.5 mL of TriReagent (Sigma #93289-100ML) using zirconia/silica beads (Daintree scientific #110791110z) in a FastPrep-24 homogeniser (MP biochemical #6004500). RNA was then isolated as per the manufacturer's (Sigma) protocol, reconstituted in 20–200 μL of nuclease free water, and its concentration (absorbance at 260 nm) and purity (ratio of absorbance at 260:280 nm) were determined using a NanoDrop UV-visible spectrophotometer (Nanodrop 2000c Thermoscientific). From previous experience, RNA could not be extracted from the midgut glands of *G. natalis* and *B. latro* with enough purity for cDNA synthesis using TriReagent. For this reason, RNA from the midgut gland of these species was extracted using an ISOLATE II RNA mini kit (Bioline #BIO-52071). cDNA from all species was

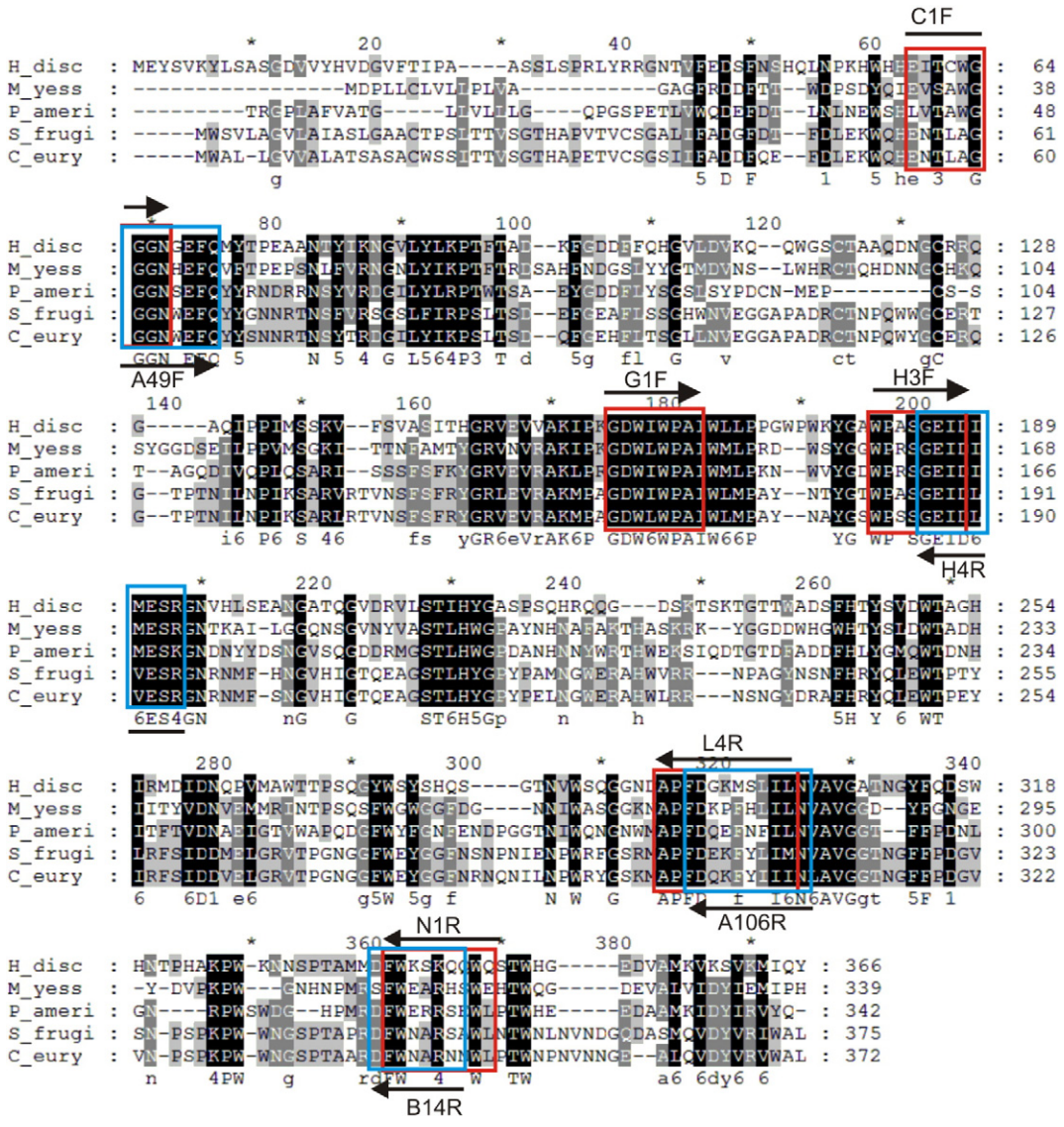


Fig. 1. Alignment of the putative amino acid sequences for β-1,3-glucanases from two molluscs (the abalone, *Haliotis discus* (H_disc) (accession number BAH84971) and the scallop, *Mizuhopecten yessoensis* (M_yess) (accession number AAW34372)) and three insects (the fall armyworm, *Spodoptera frugiperda* (S_frugi) (accession number ABR28478), the American cockroach, *Periplaneta americana* (P_ameri) (accession number ABR28480) and the orange sulphur butterfly, *Colias eurytheme* (C_eury) (accession number ACI32831)) that were used to construct degenerate primers. Total conservation is highlighted in black while partial conservation is highlighted in light grey. Sequences used to create the degenerate primers are indicated by blue and red boxes. Forward and reverse arrows and letter and number codes indicate the sequences from which the forward and reverse degenerate primers were constructed.

synthesised from 2 μg of RNA using a Biorad iScript reverse transcription supermix for RT-q PCR (Biorad #170-8840). It was then used in subsequent PCR reactions.

2.4. Amplification of an internal region of the β-1,3-glucanase gene

An internal region of the GHF16 gene from *G. natalis* and *C. destructor* was initially amplified using degenerate primers and PCR. To design the degenerate primers, the amino acid sequences of GHF16 β-1,3-glucanases from two molluscs (the scallop, *Mizuhopecten yessoensis*, accession number AAW34372 and the abalone *Haliotis discus hannai*, accession number BAH84971) and three insects (the fall armyworm *Spodoptera frugiperda*, accession number ABR28478, the American cockroach *Periplaneta americana*, accession number ABR28480 and the orange the sulphur butterfly *Colias eurytheme*, accession number ACI32831) were aligned using Clustal Omega (http://www.clustal.org/) (Fig. 1). Degenerate primers were then designed from highly conserved regions of the aligned sequences using the iCODEHOP website (http://dbmi-icode-01.dbmi.pitt.edu/i-codehop-context/) to reduce their degeneracy (Fig. 1) (Rose et al., 2003). Different combinations of forward and reverse primers were trialled in the PCR reactions (Tables 2, 3). All primers were synthesised by Geneworks Pty Ltd (www.geneworks.com.au).

2.5. PCR reactions

Using the degenerate primers and the polymerase chain reaction, PCR products of fragments of the GHF16 cDNA were amplified (Figs. S1-S3) (Table 3). PCR reactions were performed in a MJ-Research thermocycler using GoTaq Green master mix, primers and cDNA. Each 20 μL PCR reaction consisted of 1 μL of cDNA, 0.5 μL of each 100 μM forward and reverse primers (final concentration of primers = 2.5 μM), 10 μL of GoTaq Green two times master mix (Promega #M7122) and 8 μL of nuclease free water. The PCR reaction mixture was initially subjected to 92 °C for 5 min to melt the DNA and primers, before being cycled 40 times through the following; 92 °C for 1-min (melting), 45 °C for 2 min (annealing) and 72 °C for 2 min (extension). The reactions were held at 72 °C for 5 min for a final extension of products, and then maintained at 18 °C and stored at −20 °C.

Following PCR, a 10 μL aliquot of the reaction solution was electrophoresed at 100 V for 35 min in a 1% agarose gel containing TBE buffer (89 mM tris borate pH8.3, 2 mM EDTA). Products were visualised under UV light with SYBR safe gel stain (Life Technologies #S33102), which had been incorporated into the gel as per the manufacturer's instructions. The size of the products was determined from their migration distance against that of DNA standards (either E-Gel 1 Kb Plus DNA ladder (Life Technologies #10488-090) or E-gel low range quantitative DNA ladder (Life technologies #12373-031)). PCR products of interest were excised from the gel, purified, using a Nucleospin extract II kit (Macherey-Nagel #740 609.50) and directly sequenced using a BigDye Terminator v3.1 cycle sequencing kit (Life Technologies #4337454) and an Applied Biosystems ABI 3730 capillary sequencer as per the manufacturer's protocol.

Table 2
Sense and antisense degenerate primers that were used in various combinations to amplify the parts of the laminarinase cDNA by PCR. Primers were from the output of the iCODEHOP website following the alignment of β-1,3-glucanase (laminarinase) sequences (AAW34372, BAH84971, ABR28478, ABR28480, ACI32831) (Rose et al., 2003).

Primer name	DNA sequence of the primers
Sense primers	
Lam G1F	AGGGCGACTGGATCTGGCCNGCNAT
Lam H3F	CTGGCCCGGCTCNGGNGARATHG
Lam C1 F	GAGATCACCTGCTGGGGNGGNGGNAA
Lam A49F	GCGGCGGCAACTGGGARTTYCA
Antisense primers	
Lam L4R	TGATGAAGTAGAACTTCTGGTCRAANGGNGC
Lam N1R	GCCACTGGTGCCTGSNNTYCCARAA
Lam H4R	CCTGGACTCCATGATGTCDATYTCNCC
Lam A106R	TCAGCACGATGTAGAACTTYTGRTCRAA
Lam B14R	TTGTCGCGGCCGTTCCARAARTC

Table 3
Primer combinations which amplified GHF16 PCR products and the number of nucleotide base pairs successfully sequenced from these PCR products. PCR products were amplified using cDNA derived from the midgut gland of *G. natalis* (a) and *C. destructor* (b) and the haemolymph of *C. destructor* (c).

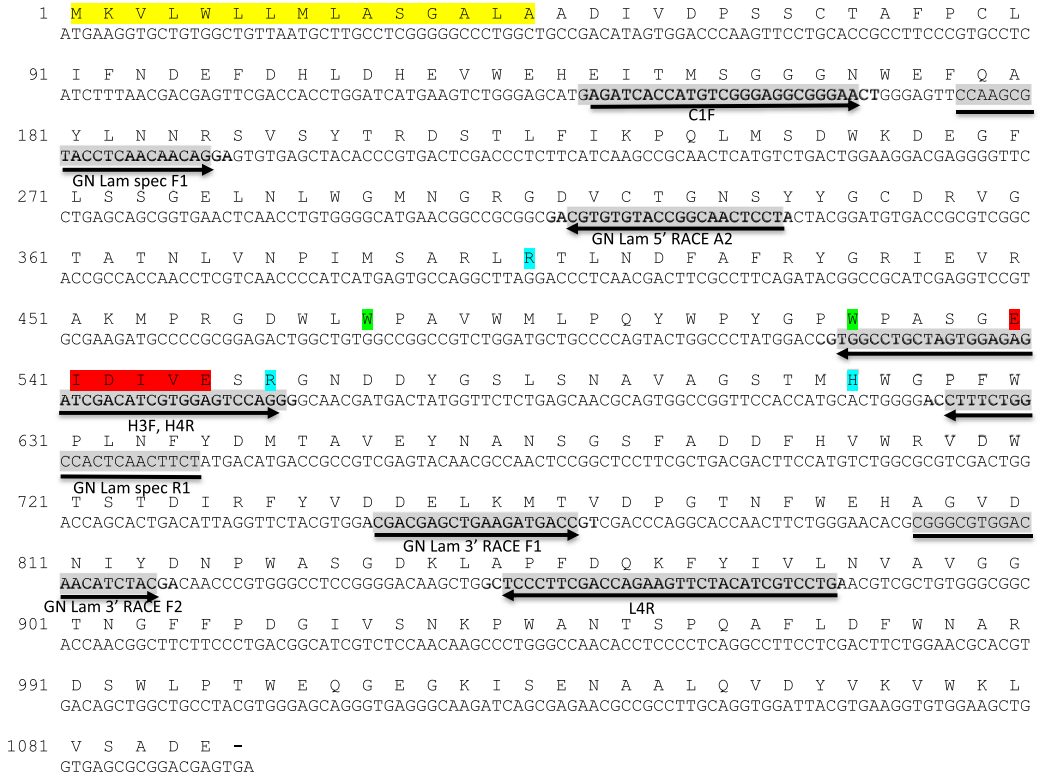
Primer pair	Nucleotide base pairs sequenced
(a) *Gecarcoidea natalis* (midgut gland)	
C1F, H4R	372 bp
H3F, L4R	340 bp
GN Lam spec F1, GN Lam spec R1	447 bp
3′ RACE	695 bp
5′ RACE	351 bp
(b) *Cherax destructor* (midgut gland)	
C1F, H4R	268 bp
G1F, L4R	358 bp
H3F, L4R	308 bp
H3F, N1R	434 bp
A49F, B14R	713 bp
3′ RACE	630 bp
5′ RACE	469 bp
(c) *Cherax destructor* (haemolymph)	
A49F, B14R	771 bp
C1F, L4R	718 bp
G1F, L4R	395 bp
H3F, L4R	335 bp
CD Lam 5′ S1, CD lam spec R2	493 bp
5′ RACE	354 bp

For each species, the overlapping nucleotides of the partial sequences were aligned using Clustal Omega (www.clustal.org) and assembled to elucidate the majority of the sequence. Finally 3′ and 5′ rapid amplification of cDNA ends (RACE) was used to sequence the respective ends of the cDNA (Fig. 2). As per the manufacturer's protocol (Takara), sequence specific primers were designed from the partial sequences obtained from the initial PCR reactions (Table 4). For 3′ RACE reactions, sense sequence specific primers between 19–25 nucleotides long, with a melting temperature between 50-60 °C, and a GC content between 40–60% were designed (Table 4). 5′ RACE reactions required an antisense sequence specific primer (10–14 nucleotides long, melting temperature between 30–40 °C and a GC content between 45–55%) with a 5′ phosphorylated primer end and two sets of nested sequence specific PCR primers with the same specifications as the 3′ RACE sequence specific primer (Table 4). A Takara 3′-Full RACE core set (Cat #6121) and a 5′-full RACE core set kits (Takara #6122) were used for these reactions. PCR and sequencing protocols used in the RACE reactions were as described above. The open reading frame of the sequence was determined by translating the sequence into the putative amino acid sequence using Gene Runner ver3.05. SMART domain analysis (http://smart.embl-heidelberg.de/) was used to classify the putative proteins into one of the glycosyl hydrolase families. Signal P3.0 programme (http://www.cbs.dtu.dk/services/SignalP/) was used to determine a potential hydrophobic signal sequence, and the ExPasy ProtParam tool (http://web.expasy.org/protparam/) used to estimate the molecular masses of the putative proteins. The tertiary structure of the protein was predicted using homology modelling of the putative amino acid sequences (Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/)).

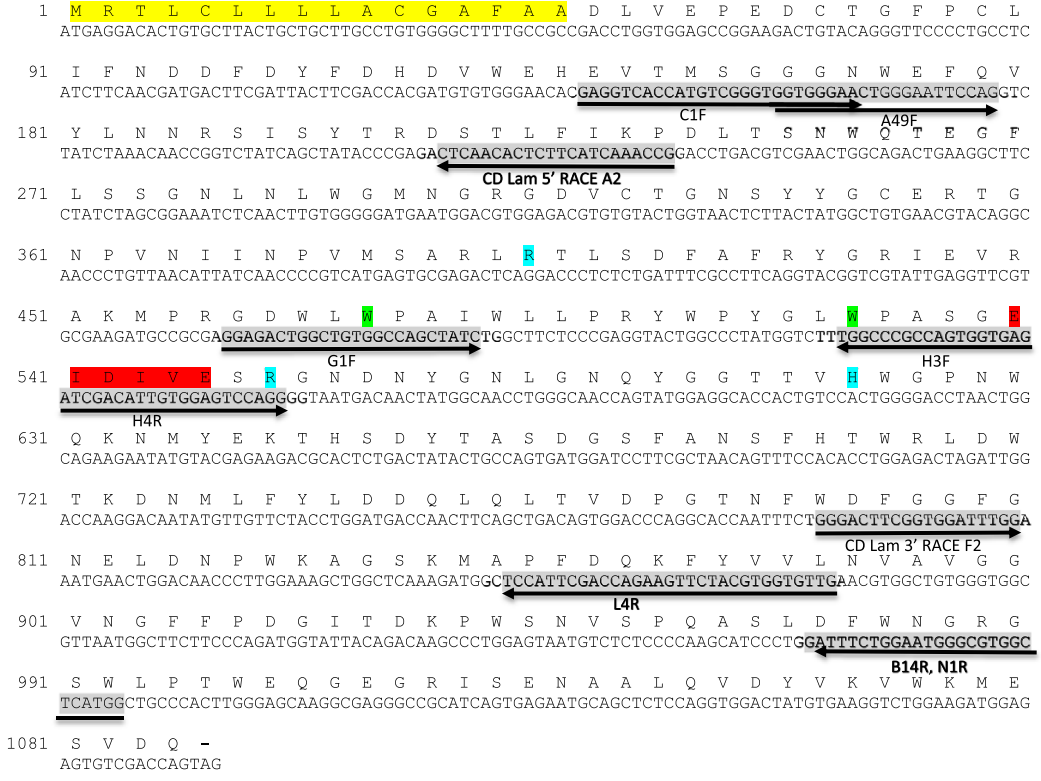
2.6. Partial sequencing of a GHF16 protein from the midgut glands of other decapod species

Partial sequences of GHF16 cDNA were also obtained for other amphibious (*M. platycheles* and *P. laevis*) and terrestrial decapods (*C. perlatus*, *C. rugosus*, *C. brevimanus* and *B. latro*). To do this, a fragment

a) *Gecarcoidea natalis nucleotide and putative amino acid sequence of a GHF16 β-1,3-glucanase derived from the midgut gland.*



b) *Cherax destructor nucleotide and putative amino acid sequence of a GHF16 β-1,3-glucanase derived from the midgut gland.*



c) *Cherax destructor nucleotide and putative amino acid sequence of a GHF16 β-1,3-glucanase protein derived from the haemolymph*

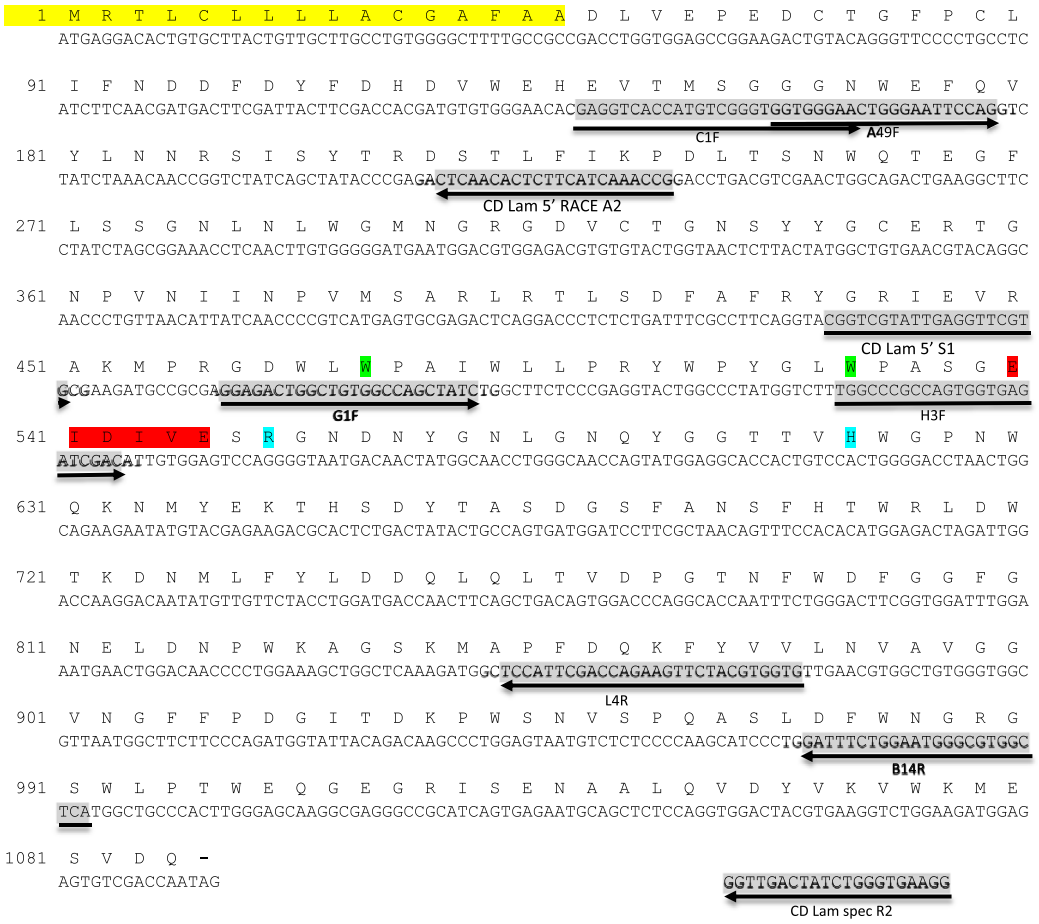


Fig. 2. Nucleotide and putative amino acid sequences of the β-1,3-glucanase cDNA derived from the midgut gland of the gecarcinid land crab, *Gecarcoidea natalis* (accession KJ995764) (a) and the midgut gland (b) (accession KJ995763) and haemolymph (c) (accession KJ995766) of the freshwater crayfish, *Cherax destructor*. The signal peptide is highlighted in yellow; amino acids that are critical to catalysis are highlighted in red; amino acids involved in modifying the pKa of the catalytic amino acids are highlighted in blue; while tryptophan amino acids involved in binding are highlighted in green. Bases highlighted in grey represent the sites of the PCR primers. Arrows and codes represent the forward and reverse primers that were used to generate the nucleotide sequences.

of the GHF16 cDNA derived from the midgut gland was amplified using degenerate primers and PCR (Fig. S3). The PCR product was excised from the gel, isolated and sequenced as described above.

2.7. Expression of GHF16 β-1,3-glucanase mRNA within various tissues

Expression of GHF16 mRNA within midgut gland, muscle and gill, was determined by amplifying a β-1,3-glucanase fragment using cDNA derived from these tissues, degenerate primer pairs (C1F H4R for *G. natalis* and H3F L4R *C. destructor*) and PCR. Fragments of either β-actin or GAPDH were also amplified from cDNA derived from each of the tissues. This served as a positive control to ensure that good quality cDNA was used in the reactions. RNA samples did not contain genomic DNA given that control PCR reactions using isolated total RNA in control reverse transcriptase reactions did not yield any PCR products for β-1,3-glucanase fragments.

2.8. Expression of a GHF16 protein within the haemocytes of C. destructor

To determine if a similar GHF16 gene is expressed in the haemocytes of *C. destructor*, the open reading frame of a GHF16 protein was sequenced using cDNA derived from the haemolymph. To do this, 50–100 μL haemolymph samples from 10 animals were collected and pooled into 1 mL of RNA later. This precipitated the protein and cells, which were then collected by centrifugation (12,000 *g* for 15 min). After centrifugation, the RNA later was carefully aspirated off and the RNA isolated from the pelleted material using ISOLATE II RNA mini kit (Bioline #BIO-52071). cDNA was synthesised using Biorad iScript reverse transcription supermix for RT-q PCR (Biorad #170-8840) and fragments of the GHF16 protein amplified using degenerate primer pairs, sequence specific primers, PCR and 5′ RACE as described above. PCR products were directly sequenced and the sequences assembled as described above (Fig. S2).

2.9. Can the putative protein be translated into a digestive enzyme?

Sequences of short peptides were determined from a β-1,3-glucanase purified previously from the midgut gland of *G. natalis* (Allardyce and Linton, 2008). These sequences were compared to the same sequences from the putative amino acid sequence to indicate if the putative amino sequence derived from the cDNA could potentially match that of a previously purified β-1,3-glucanase. The probability of

the two sequences possessing the same amino acids by chance was calculated from the chance of randomly selecting the correct amino acid out of 20 for each of amino acid residues in the peptide that matched the putative sequence. For *G. natalis* the calculated probability was

$$\text{Probability} = \left(\frac{1}{20}\right)^{17} = 7.63 \times 10^{-23}.$$

To do this, a β-1,3-glucanase sample, previously purified from the midgut gland of *G. natalis*, was run on a 12% polyacrylamide gel and stained with Coomassie blue as per the method of Allardyce and Linton(2008) and Linton and Shirley(2011). A protein band corresponding to the β-1,3-glucanase was excised from the gel and processed by in-gel tryptic digestion. Obtained peptides were eluted from the gel, lyophilised and resuspended in 0.1% formic acid. For LC-MS/MS analysis of peptides a nano-HPLC system (nanoAquity, Waters) online-coupled to an ion source into a LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) via a nanoelextrospray ion source (TriVersa NanoMate, Advion) was used. Peptides were injected on a trapping column (nanoAquity UPLC column, C18, 180 μm × 20 mm, 5 μm, Waters) and separated on a C18 UPLC column (nanoAquity UPLC column, C18, 75 μm × 100 mm, 1.7 μm, Waters) in a 45 min gradient. Full scan MS spectra were acquired in a positive ion mode in the LTQ-Orbitrap XL using a CID top 6 method as described in Rockstroh et al. (2011).

2.10. Phylogenetic analysis

Full length amino acid sequences for GHF16 proteins such as β-1,3-glucanase and lipopolysaccharide and β-glucan binding proteins from insects, molluscs and crustaceans were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank) and aligned using the Clustal W algorithm within the MEGA6 software package (http://www.megasoftware.net). A neighbourhood joining phylogenetic tree was then constructed using MEGA6 and assuming the Jones–Taylor–Thornton substitution model (Jones et al., 1992). The reliability of the phylogeny was tested with the bootstrap model with a 1000 replicates. Branches with boot strap values lower than 50 were collapsed to form polyphyletic groups.

3. Results

3.1. G. natalis and C. destructor GHF16 sequence

The GHF16 mRNA sequence for *G. natalis* had an open reading frame of 1098 bp while that for *C. destructor* had an open reading frame of 1095 bases. Translated, this would produce a putative protein of 365 amino acids with an estimated molecular mass of 41.4 kDa for *G. natalis*, and a 364 amino acid protein with an estimated molecular mass of 41.5 kDa for *C. destructor* (Fig. 2).

The putative amino acid sequences from both species contained a hydrophobic signal sequence that was 15 amino acids long (Fig. 2, marked in yellow) and the correct catalytic and binding residues to be a catalytically active β-1,3-glucanase (Juncosa et al., 1994; Henrissat and Davies, 1997; Genta et al., 2009; Kovalchuk et al., 2009; Song et al., 2010). Both sequences contained features that are characteristic of glycosyl hydrolase family 16 enzymes, in particular the amino acid sequence "Glu-Ile-Asp-Ile-Val-Glu" (Fig. 2, marked in red); the first glutamate residue acts as the nucleophile, the second glutamate acts as the general acid base catalyst, while the aspartate acts to modify the pKa of the catalytic residues (Juncosa et al., 1994; Genta et al., 2009). The sequences also contain tryptophan substrate binding residues (Fig. 2, marked in green) and 2 arginine residues and 1 histidine residue (Fig. 2, marked in blue) that help to modify the pKa of the catalytic glutamate residues (Genta et al., 2009; Kovalchuk et al., 2009; Song et al., 2010; Zakharenko et al., 2011). The predicted tertiary structure, produced from homology modelling of the putative amino acid sequences, had a β-jelly roll motif, which is also typical of family 16 glycosyl hydrolases (Genta et al., 2009) (Fig. 3). Like the β-1,3-glucanases from

Table 4
Sequence specific primers used for 3′ and 5′ RACE reactions, for checking the sequences and for positive control PCR reactions for *G. natalis* (a), *C. destructor* (b) and *B. latro* (c). For primers used in the nested PCR reactions of the 5′ RACE, the antisense primers pointed towards the 5′ end while the sense primers pointed towards the 3′ end, as per the 5′ RACE kit instructions. GAPDH primers were as described by Tsang et al. (2011). These primers were also used as positive controls for all amphibious and terrestrial decapod species. β-actin primers were used as positive controls for *Cherax destructor*.

(a) *Gecarcoidea natalis*

Primer name	DNA sequence of the primers
3′ RACE reactions	
Sense primers	
GN lam 3′ F1	CGACGAGCTGAAGATGACC
GN lam 3′ F2	CGGGCGTGGACAACATCTAC
5′ RACE reactions	
Phosphorylated sequence specific primer	
GN lam 5′ RT-P	GCTTGTTGGAGACG (Primer phosphorylated at the 5′ end)
Antisense primers	
GN Lam 5′ A1	TAGCAGGCCACGGTCCATAG
GN Lam 5′ A2	AGGAGTTGCCGGTACACACG
Sense primers	
GN Lam 5′ S1	CGACGAGCTGAAGATGACC
GN Lam 5′ S2	CGGGCGTGGACAACATCTAC
Sequence specific primers	
Sense primers	
GN Lam spec F1	CCAAGCGTACCTCAACAACAG
Antisense primers	
GN Lam spec R1	AGAAGTTGAGTGGCCAGAAAG
Positive control primers for GAPDH (Tsang et al., 2011)	
Sense primers	
GAPDH GA	ATGGTGTATATGTTCAAGTAYGAYTC
Antisense primers	
GAPDH GR	TCGCTAGATACAACATCATCYTCRGT

(b) *Cherax destructor*

Primer name	DNA sequence of the primers
3′ RACE reactions	
Sense primers	
CD Lam 3′ F2	GGGACTTCGGTGGATTTGG
5′ RACE reactions	
Phosphorylated sequence specific primer	
CD Lam 5′ RT-P	CAAATCCACCGAAG (Primer phosphorylated at the 5′ end)
Antisense primers	
CD Lam 5′ A1	GGCGAAATCAGAGAGGGTC
CD Lam 5′ A2	CGGTTTGATGAAGAGTGTTGAG
Sense primers	
CD Lam 5′ S1	CGGTCGTATTGAGGTTCGTG
CD Lam 5′ S2	TGACAACTATGGCAACCTGG
Sequence specific primers	
Antisense primers	
CD lam spec R2	CCTTCACCCAGATAGTCAACC
Positive control primers for β-actin	
Sense primers	
β-actin F	GGCTACTCCTTCACCACCAC
Antisense primers	
β-actin R	AGCTGTGGACGGTTTCATG

(c) *Birgus latro*

Primer name	DNA sequence of the primers
3′ RACE reactions	
Sense primers	
BL lam 3′ F1	TCGCCAACAGCTTCCACAC

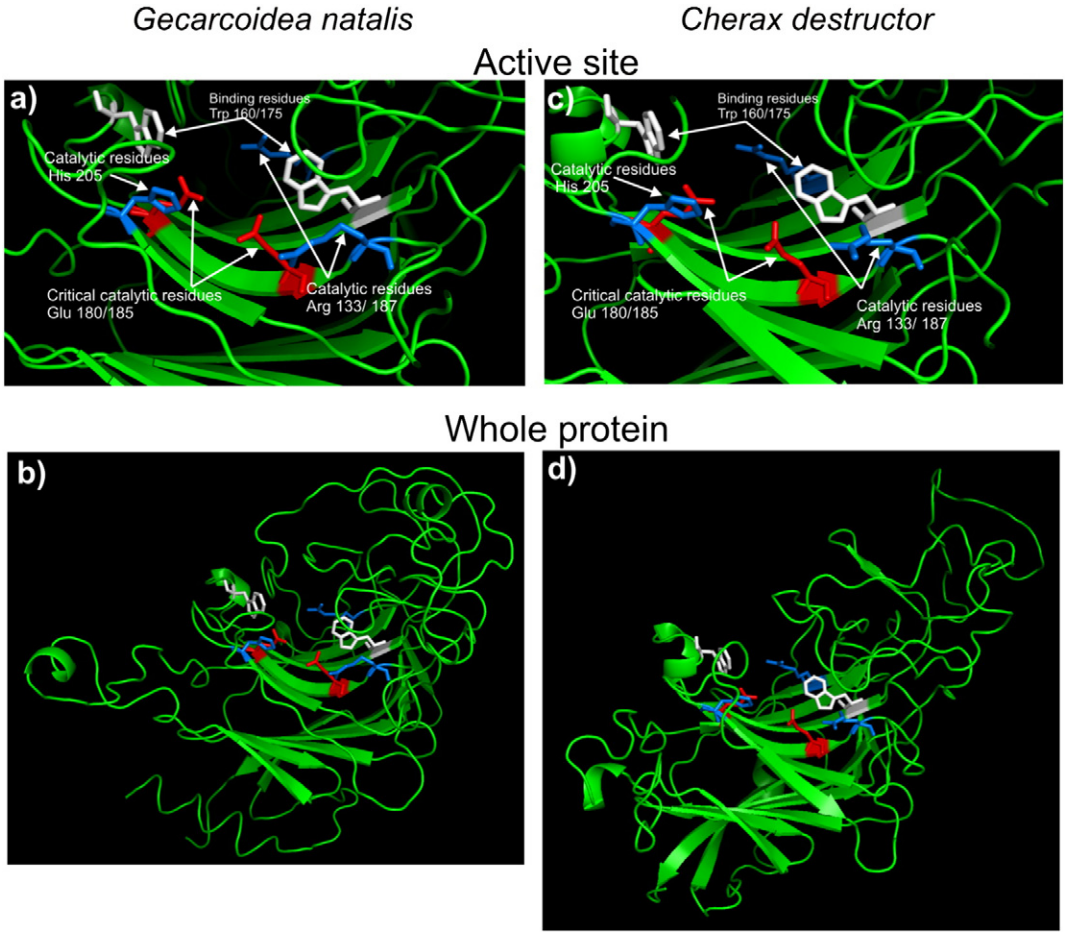


Fig. 3. Predicted tertiary structure of the β-1,3-glucanase proteins expressed in the midgut glands of *G. natalis* (a and b) and *C. destuctor* (c and d).Whole protein (b and d) and the active site (a and c) depicting the secondary structure of the protein plus the side chains of the amino acids involved in the catalysis and substrate binding. Critical catalytic residues (coloured red), the amino acid residues which aid these (coloured blue) and the tryptophan residues involved in substrate binding (coloured white) are labelled in the panels a and c. These residues have the same colour scheme in panels b and d. Structures were predicted from the putative amino acid sequence by homology modelling using the Phyre2 server (http://www.sbg.bio.ic.ac.uk/phyre2/). Figure prepared using Pymol (http://www.schrodinger.com/pymol/).

Tenebrio molitor and *Cryptopygus antarcticus*, the secondary and tertiary structures of the β-1,3-glucanase from both *G. natalis* and *C. destructor* were highly similar to that of an endo-β-1,3-glucanase from *Nocardiopsis* sp strain f96 (Protein data bank code c2hykA) (Genta et al., 2009; Song et al., 2010). Compared to the amino acid sequence from *Nocardiopsis*, the respective similarity for that of *G. natalis* and *C. destructor* was 42 and 40%.

C. destructor (A)	**FYLDDQLQ**
G. natalis (A)	**FYVDDELK**
G. natalis (P)	**FYVDDELK**

C. destructor (A)	**AALQVDYVK**
G. natalis (A)	**AALQVDYVK**
G. natalis (P)	**AALQVDYVK**

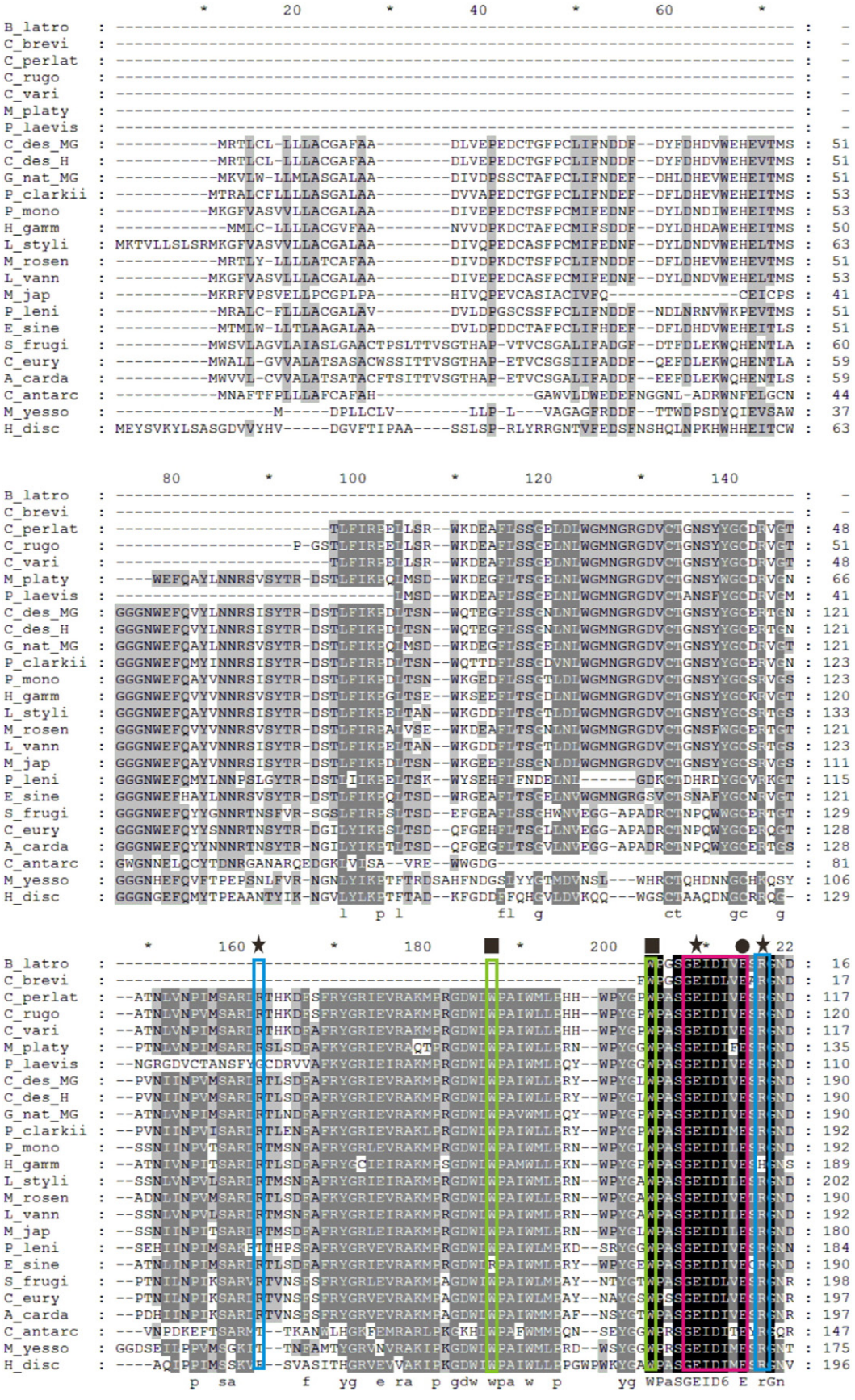
Fig. 4. Short peptide sequences (P) from a β-1,3-glucanase (laminarinase) purified from the midgut gland of the gecarcinid land crab, *Gecarcoidea natalis* aligned with the putative amino acid sequences (A) of a β-1,3-glucanase (laminarinase) from *Cherax destructor* and *Gecarcoidea natalis*. Totally conserved amino acid residues are coloured red while highly conserved amino acid residues are coloured green. The peptide sequences for *G. natalis* were determined by orbitrap mass spectrometry of a purified β-1,3-glucanase which had undergone a tryptic digest while the putative amino acid sequences for *C. destructor* and *Gecarcoidea natalis* were derived from the cDNA sequence.

3.2. Amino acid sequences of short peptides from a previously purified β-1,3-glucanase

In addition to the putative amino acid sequence determined from the cDNA sequence, two partial sequences of 8 and 9 amino acids long were obtained by Orbitrap mass spectrometry from a β-1,3-glucanase purified previously from the midgut gland of *G. natalis*. These fragments aligned perfectly with the putative sequence derived from the same tissue in *G. natalis* with the probability of the two proteins being the same by chance alone and thus being misidentified is 7.63×10^{-23} (Fig. 4). Similarly the amino acid residues were also highly conserved (first peptide) and totally conserved (second peptide) when compared with the putative amino acid sequence from *C. destructor* (Fig. 4). The probability that these two proteins share the same amino acids by chance is 6.10×10^{-19}. Furthermore, the estimated molecular masses of the β-1,3-glucanase proteins (41.4 kDa for *G. natalis* and 41.5 kDa for *C. destructor*) match that determined from β-1,3-glucanase purified previously (Allardyce and Linton, 2008) (Table 1). Thus, there is strong evidence that the GHF16 cDNA sequences presented here encode the β-1,3-glucanase enzymes which have been previously purified and characterised.

3.3. Alignment of sequences

A BLAST search of the putative amino acid sequences from *G. natalis* and *C. destructor* revealed that they were similar to that of β-glucan


B_latro
C_brevi
C_perlat
C_rugo
C_vari
M_platy
P_laevis
C_des_MG
C_des_H
G_nat_MG
P_clarkii
P_mono
H_gamm
L_styli
M_rosen
L_vann
M_jap
P_leni
E_sine
S_frugi
C_eury
A_carda
C_antarc
M_yesso
H_disc

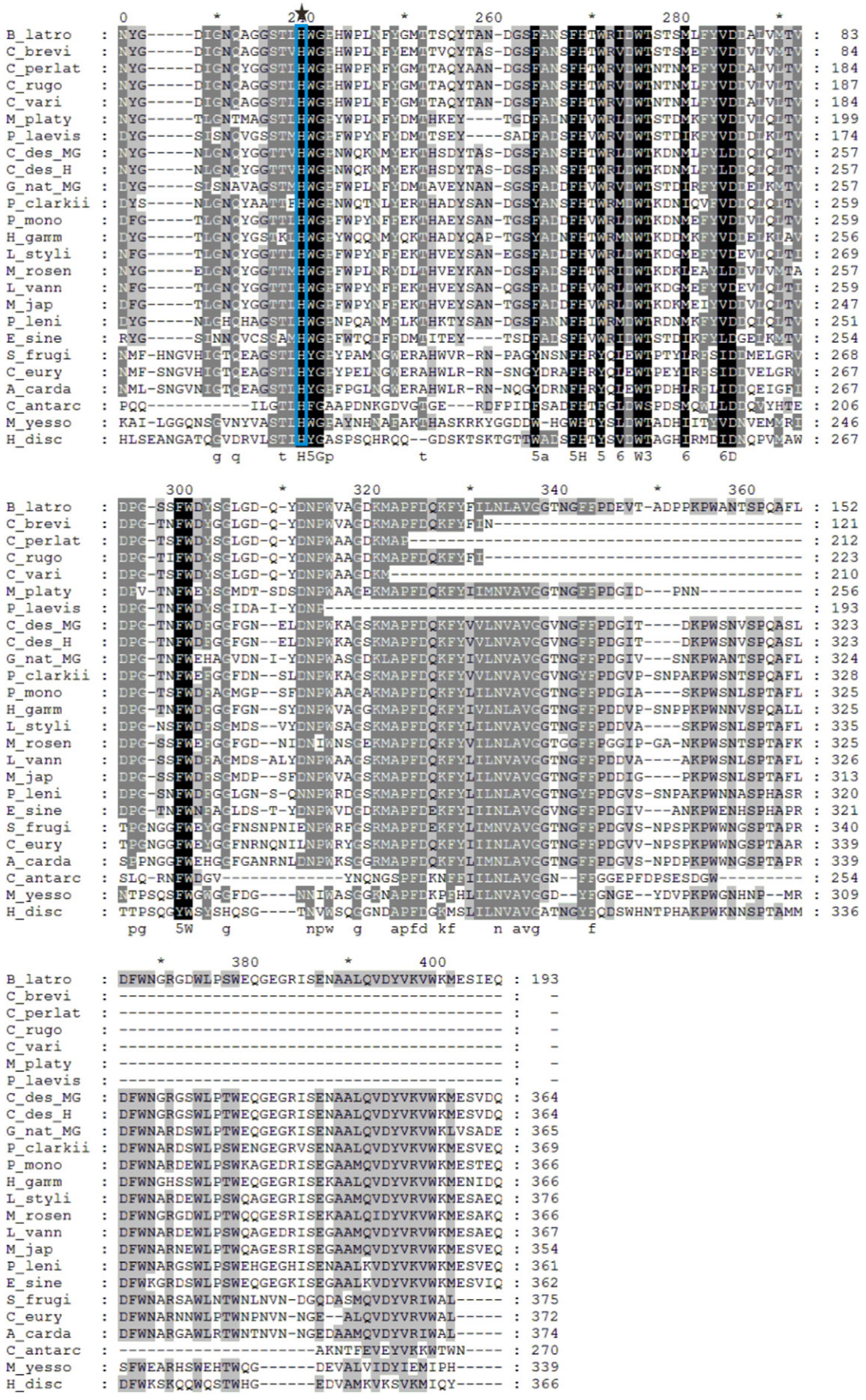


Fig. 5 (*continued*).

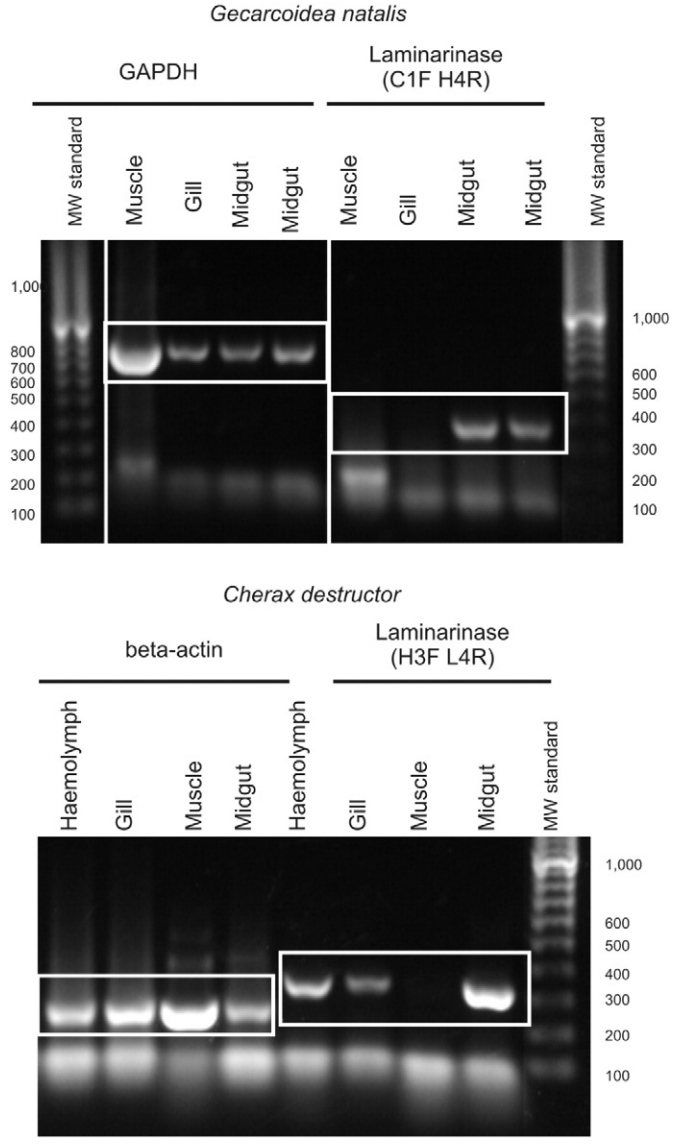


Fig. 6. Expression of β-1,3-glucanase (laminarinase) mRNA and a control mRNA, either beta actin or glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in different tissues indicated for *G. natalis* (a) and *C. destructor* (b). In *G. natalis*, the C1F H4R primer combination was used to amplify a fragment (≈350 bp) of β-1,3-glucanase (laminarinase) while in *C. destructor*, a fragment (≈300 bp) of the β-1,3-glucanase (laminarinase) cDNA was amplified using a H3F L4R primer combination. Control fragments for beta actin (≈200 bp) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (≈300 bp) were respectively amplified from cDNA derived from all tissues from *C. destructor* and *G. natalis*. Fragment size standards with the sizes indicated were run alongside the samples.

binding and lipopolysaccharide proteins from other crustaceans, and β-1,3-glucanase sequences from other invertebrates. Similarly a BLAST search using the nucleotide sequence from *G. natalis* revealed that it was respectively 75, 76 and 84% identical to that of β-1,3-glucan binding proteins from *E. sinensis*, *Marsupenaeus japonicus* and *Macrobrachium rosenbergii*. The nucleotide sequence from *C. destructor* had 97% identity with a 661 bp nucleotide sequence from the closely related crayfish, *Cherax quadricarinatus*; the function of the deduced protein has not been determined (Yudkovski et al., 2007).

3.4. Sequences from other crabs

Partial GHF16 cDNA sequences, ranging in size from 581 to 768 base pairs, were also obtained from cDNA prepared from the midgut gland of seven other crustaceans including amphibious and terrestrial brachyuran, and anomuran decapods. Like the sequences from *C. destructor* and *G. natalis*, the putative amino acid sequences contained the characteristic catalytic and binding domains of a glycosyl hydrolase family 16 protein (Juncosa et al., 1994; Genta et al., 2009) (Fig. 5). Thus, the cDNA may encode for a β-1,3-glucanase. Similarly a BLAST search of these putative amino acid sequences suggested that they were similar to that of β-glucan binding proteins from crustaceans and β-1,3-glucanases from other species.

3.5. Tissue expression

For *G. natalis*, β-1,3-glucanase mRNA was expressed in the midgut gland but not in gill and muscle (Fig. 6). In contrast, β-1,3-glucanase mRNA was expressed in the midgut gland, gill and haemolymph of *C. destructor*, but not in the muscle (Fig. 6). In all other species examined (*C. perlatus*, *C. rugosus* and *Coenobita variabilis*), β-1,3-glucanase mRNA was expressed in the midgut gland but not the gill or muscle (data not shown).

3.6. GHF16 β-glucan binding protein sequence from the haemolymph of *C. destructor*

Using the same degenerate and sequence specific GHF16 β-1,3-glucanase primers, a 1095 bp open reading frame was obtained from the cDNA prepared from the haemolymph of *C. destructor* (Fig. 2c). Translated this would produce a 364 amino acid residue protein with a molecular mass of 41.5 kDa (Fig. 2c). The nucleotide sequences of the open reading frames from both tissues, midgut gland and haemolymph, were 99% similar and the putative amino acid sequences were identical. The putative amino acid sequence derived from the haemolymph contained the same catalytic and binding amino acid residues as the putative amino acid sequence determined from the midgut gland (Figs. 2, 5).

3.7. Phylogenetic analysis of sequences

All of the sequences for the GHF16 proteins from insects, crustaceans and molluscs grouped together within a single branch of the phylogenetic tree. In contrast, sequences for true crustacean β-glucan binding proteins that are much larger (152 kDa compared to 30–50 kDa) and possess a binding but no catalytic domain formed an out-group away

Fig. 5. Alignment of the putative amino acid sequences for β-1,3-glucanases and lipopolysaccharide and β-glucan binding proteins. Black highlighting indicates total conservation while dark grey and light grey indicate partial conservation. Amino acid sequences that are characteristic of glycosyl hydrolase family 16 proteins and which are critical to catalysis are boxed by a red rectangle. The circle above the alignment indicates the glutamic acid residue which is critical to catalysing the reaction. Amino acids involved in catalysis that aid the critical glutamic acid residue are indicated by a black star above the alignment and blue rectangles around the amino acids. Amino acids involved in substrate binding are indicated by a black square above the alignment and green rectangles around the amino acids. Sequences for glycosyl hydrolase family 16 β-1,3-glucanase were determined in this study and were from *Birgus latro* (B_latro accession KJ995771), *Coenobita brevimanus* (C_brevi accession KJ995770), *Coenobita perlatus* (C_perlat accession KJ995767), *Coenobita rugosus* (C_rugo accession KJ995768), *Coenobita variabilis* (C_vari accession KJ995769), *Mictyris platycheles* (M_platy accession KJ995772), *Paragrapsus laevis* (P_laevis accession KJ995773), *Cherax destructor* (GHF16 sequence derived from the midgut gland (C_des_MG accession KJ995763) and haemolymph (C_des_H accession KJ995766)) and *Gecarcoidea natalis*(G_nat_MG accession KJ995764). These sequences were aligned with the β-1,3-glucan binding sequences from other crustaceans (*Procambarus clarkii* (P_clarkii), accession ACR20474; *Penaeus monodon* (P_mono), accession AEX08659; *Homarus gammarus* (H_gamm), accession CAE47485; *Litopenaeus stylirostris* (L_styli), accession AAM73871; *Macrobrachium rosenbergii* (M_rosen), accession ACT33045; *Litopenaeus vannamei* (L_vann) accession ABU92557; *Marsupenaeus japonicus*(M_jap), accession ABY89089; *Pacifastacusleniusculus* (P_leni) accession CAB65353; *Eriocheir sinensis* (E_sine) accession ACR56716) and β-1,3-glucanase sequences from representative insects (*Spodoptera frugiperda* (S_frugi), accession ABR28478; *Colias eurytheme* (C_eury), accession ACI32831; *Anthocharis cardamines* (A_carda), accession ACI32832; *Cryptopygus antarcticus* (C_antarc), accession FJ648734) and molluscs (*Mizuhopecten yessoensis* (M_yesso), accession AAW34372; and *Haliotis discus hannai* (H_disc), accession BAH84971; *Perna viridis* (P_viri) accession ACM68926).

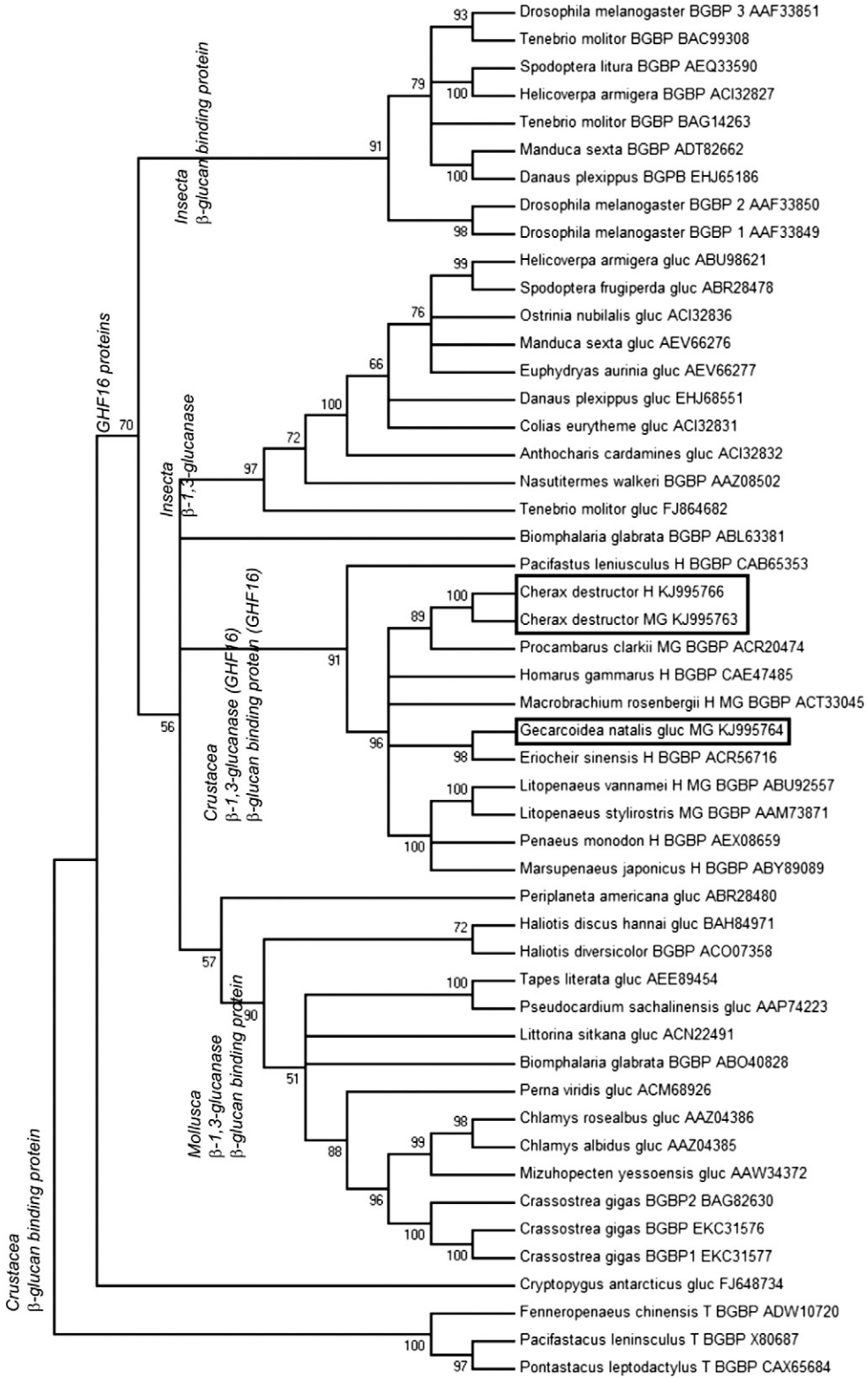


Fig. 7. Phylogeny of the beta-β-1,3-glucanase sequences from the midgut glands of various decapod species determined in this study with that of beta-glucan binding proteins and β-1,3-glucanases of glycosyl hydrolase family 16 (GHF16) from other crustaceans, molluscs and insects. Phylogeny was constructed by the maximum likelihood method assuming Jones–Taylor–Thornton model and the MEGA 6 software package. The credibility of the phylogeny was tested using the bootstrap method with a 1000 replicates and a cut off at 50. The values at each node represent the percentage bootstrap value. Accession numbers are given after the species name. BGBP indicates that the sequence was initially described as a lipopolysaccharide and β-glucan binding protein while gluc indicates that the sequence was initially described as a β-1,3-glucanase (laminarinase). Crustacean sequences determined in this study and were derived from either the haemolymph (H) or midgut gland (M) which are boxed. Branches containing either the β-1,3-glucanase or β-glucan binding protein sequences of the various invertebrate taxa are labelled.

from the GHF16 proteins (Fig. 7) (Cerenius et al., 1994). Within the phylogenetic tree, the crustacean sequences determined in this study were found to group with sequences of GHF16 β-glucan binding proteins from other crustaceans (Fig. 7). Also, there was no clear delineation between the sequences expressed in the haemolymph and midgut gland given that they did not form separate branches (Fig. 7). Indeed for *C. destructor* in particular, the amino sequences derived from the midgut and haemolymph were identical, and thus clustered on the same branch (Fig. 7). The amino acid sequences from *C. destructor* were most closely related to a putative β-glucan binding protein sequence from the crayfish *Procambarus clarkii*, which was specifically expressed in the midgut gland. Given the similarity of the sequences from the two tissues and thus the cross reactivity of the primers, it is likely that this sequence from *P. clarkii* is a β-1,3-glucanase and hence has been mis-identified.

Like the crustacean sequences, the molluscan β-1,3-glucanase sequences are grouped with β-glucan binding protein sequences (Fig. 7). Most notably and like *C. destructor*, the sequences for β-glucan binding protein and β-1,3-glucanase from an abalone species, *Haliotis* formed a monophyletic group (Fig. 7). In contrast for insects, the β-glucan binding protein sequences formed a separate branch to that for the β-1,3-glucanases (Fig. 7). The insect β-glucan binding protein lacks the catalytic amino acids present in the β-1,3-glucanases (Pauchet et al., 2009). The two genes are related but evolutionary distinct. Three β-1,3-glucanase sequences, one from the cockroach *P. americana*, the Antarctic spring tail, *C. antarcticus* and the gastropod *Biomphalaria glabrata* did not group with other species of their phyla (Fig. 7). This may indicate a separate evolutionary origin. Indeed it has been suggested that the β-1,3-glucanase gene for *C. antarcticus* may have been acquired by horizontal transfer from bacteria (Song et al., 2010).

4. Discussion

4.1. Sequence presented is a GHF16 β-1,3-glucanase

The cDNA sequenced from the midgut glands of *G. natalis* and *C. destructor* encoded a β-1,3-glucanase from glycosyl hydrolase family 16. The putative proteins from *G. natalis* and *C. destructor* respectively were 365 and 364 residues long, and had estimated molecule masses of 41.4 and 41.5 kDa (described previously in Allardyce and Linton, 2008). The proteins also contained the catalytic and binding domains of GHF16 proteins. Both the amino acid sequence of small peptides from *G. natalis* and the molecular masses of previously purified proteins, β-1,3-glucanases, from both species matched that determined from the putative amino acid sequences. The sequence and molecular mass similarities to active β-1,3-glucanases, taken together with the identification of putative catalytic domains, provide strong evidence that the cDNAs isolated here encode a protein with β-1,3-glucosidase activity rather than just a binding protein. This reasoning is similar to that proposed for the molluscs, where the β-1,3-glucanase protein within the crystalline style is most likely the product of the GHF16 gene expressed in the hepatopancreas, the tissue that synthesises the digestive enzymes (Kozhemyako et al., 2004; Kovalchuk et al., 2006, 2009; Kumagai and Ojima, 2009; Zakharenko et al., 2011). This is the first description of the sequence which is responsible for the endogenous production of a β-1,3-glucanase (laminarinase) in decapod Crustacea. This completes the characterisation of the β-13,-glucanase (laminarinase) which had been previously purified and characterised from *G. natalis* and *C. destructor* (Allardyce and Linton, 2008). The putative GHF16 amino acid sequence derived from the haemolymph of *C. destructor* was 100% identical to that from the midgut gland of the same species. Both sequences contained the catalytic and binding domains of GHF16 proteins. Thus, the protein expressed in the haemocytes may be an active β-1,3-glucanase.

4.2. β-1,3-glucanase synthesised by the midgut gland and secreted into the digestive fluid

The β-1,3-glucanase was expressed in the midgut gland, the organ responsible for the production of digestive enzymes, and is secreted into the digestive fluid (Allardyce and Linton, 2008). A GHF16 protein was expressed in the gills and haemolymph of *C. destructor*. It is likely that this represents expression of a GHF16 protein within the haemocytes that are present in both of these tissues. Perhaps the expression of the GHF16 protein within the gills of *C. destructor* was due to a bacterial infection in the animal sampled as expression of the GHF16 protein within the gills was not observed in other animals. A GHF16 β-1,3-glucanase was also expressed within the midgut gland of other amphibious and terrestrial decapods species, and this suggests that expression of this gene is widespread amongst the decapod crustacea. This would explain the high β-1,3-glucanase activities observed in the digestive fluid and midgut glands of numerous crustacean species (Sova et al., 1970)(Suzuki et al., 1987; Omondi and Stark, 1995; Figueiredo et al., 2001; Johnston and Freeman, 2005; Figueiredo and Anderson, 2009).

4.3. GHF16 genes expressed in the midgut gland of other species may also be β-1,3-glucanases that have been misidentified

Expression of a similar GHF16 protein within the midgut gland of the decapod crustaceans, *P. stylirostris*, *F. chinensis*, *M. rosenbergii* and *E. sinensis*, has also been described previously (Roux et al., 2002; Liu et al., 2009; Yeh et al., 2009). Based on sequence similarity, the putative protein from these crustaceans was identified as a β-glucan binding protein with no hypothesised β-1,3-glucanase activity. Given that these proteins, like the ones described in this study, have conserved binding and catalytic domains of a GHF16 enzyme, they may also be catalytically active. As demonstrated for the β-1,3-glucanase from *G. natalis*, these sequences may produce an active β-1,3-glucanase that is secreted into the digestive fluid and thus they may have been misidentified. In contrast, true β-glucan binding proteins lack a catalytic domain and hence lack activity (Cerenius et al., 1994).

4.4. Function of the GHF16 protein

Given that the β-1,3-glucanase contains a catalytic site, is expressed in the midgut gland, and is secreted into the digestive fluid, it is reasonable to assume that it is involved in the digestive process. However, its similarity to the immune binding proteins and the expression of the protein within the haemocytes of *C. destructor*raise some doubt as to its role within the midgut. If the enzyme performs a purely digestive role, it could be used to hydrolyse dietary β-1,3-glucans to release glucose in much the same way as the crustacean cellulases release glucose from dietary cellulose. Alternatively, it may assist indirectly in the digestive process by hydrolysing plant and fungal cell walls in order to release the highly digestible cell contents. Finally, it may act as an immune protein to potentially detect and inactivate ingested pathogens. It should also be noted that these options are not mutually exclusive and the enzyme may have a number of putative roles, which are discussed below.

4.4.1. Hydrolysis of β-1,3-glucans as a source of glucose

β-1,3-glucanase is an enzyme that can hydrolyse β-1,3-glucans such as laminarin and callose. Laminarin is the major storage polysaccharide within brown algae and callose and is produced in plants in response to wounding (Bull and Chesters, 1966; Bacic et al., 1988). β-1,3-glucans are also present in other organisms such as diatoms, protozoans and fungi (Bull and Chesters, 1966; Piavaux, 1977; Bacic et al., 1988; Ruiz-Herrera, 1992; Terra and Ferreira, 1994; Pesentseva et al., 2008). In species such as *Petrolisthes elongatus* and *Metapenaeus bennetae* that consume algae and possess high activities of β-1,3-glucanase, this

enzyme would hydrolyse and thus digest β-1,3-glucans to produce mainly glucose (Johnston and Freeman, 2005; Allardyce and Linton, 2008; Figueiredo and Anderson, 2009).

4.4.2. Physical disruption of fungal and plant cell walls

Alternatively, the β-1,3-glucanase may help to break open the cell walls of fungal and plant material. In particular, the enzyme may work with endo-β-1,4-glucanase to hydrolyse the β-1,3 and β-1,4-glycosidic bonds within the cellulose and hemicellulose that comprise the cell walls. This, in conjunction with the mechanical fragmentation by the gastric mill, would rupture the cell walls of plants and fungi to release the highly digestible cell contents that then could be easily digested and absorbed. This may partially explain the function of the β-1,3-glucanase and endo-β-1,4-glucanase within the anomuran land crab, *B. latro*. This species consumes seeds and fleshy fibrous fruits that are high in either protein, carbohydrate or lipid (Wilde et al., 2004). It also possesses both β-1,3-glucanase and endo-β-1,4-glucanase activities, but would derive little energy from the digestion of cellulose and hemicellulose, compared to the other nutrients (Wilde et al., 2004; Linton et al., 2014). *B. latro* also lacks a β-1,4-glucosidase and thus the ability to hydrolyse cellulose to glucose (Linton et al., 2014).

4.4.3. Immune protein

A potential third option is that the protein acts as an immune protein. The β-1,3-glucanase may work like lysozyme by hydrolysing the β-1,3-glucans within cell walls of fungi and bacteria to potentially render them inactive. Indeed, the β-1,3-glucanases secreted into the midgut by insects such *Tenebrio* and lepidopteran, *Helicoverpa armigera* can hydrolyse the cell walls of these organisms (Genta et al., 2009; Pauchet et al., 2009). As described for the GHF16 β-glucan binding proteins in the haemolymph, the GFH16 protein in the digestive fluid may also bind to the β-1,3-glucans in the cell walls of micro-organisms and stimulate the prophenol oxidase system to activate a humoral immune response (Lee et al., 2000; Sritunyalucksana and Söderhäll, 2000; Amparyup et al., 2012). However stimulation of the prophenol oxidase system within the digestive fluid has yet to be demonstrated and thus the primary function of the protein may be to hydrolyse cell walls. β-1,3-glucanase activities are present in the digestive fluid of the carnivorous decapods *Scylla serrata* and *Portunus pelagicus* (Figueiredo and Anderson, 2009). *S. serrata* and *P. pelagicus* consume animal material in the form of molluscs, crustaceans and polychaete worms (Figueiredo and Anderson, 2009). In these species, the β-1,3-glucanase would not function as a digestive enzyme as these species do not consume significant amounts of β-1,3-glucans. It is more likely that the enzyme would act as an immune protein as described above, particularly if they may eat rotting material that may contain significant amounts of bacteria and fungi. Terrestrial hermit crabs (*Coenobita* sp.) commonly feed on animal faeces and may utilise the β-1,3-glucanase present within the digestive fluid as an immune enzyme to inactivate potentially pathogenic micro-organisms (Greenaway, 2003).

4.5. Evolutionary relationship of the β-1,3-glucanases and β-glucan binding proteins

The evolution of the β-1,3-glucanase gene within both crustaceans and molluscs is unclear since the GHF16 gene for β-1,3-glucanase and β-glucan binding proteins are very closely related given they possess highly conserved catalytic and binding domains (Figs. 5, 7). Indeed, these two classes of protein are so similar that they are grouped together by phylogenetic analysis and thus may be the same protein. Most notably, the sequences for β-1,3-glucanase and β-glucan binding protein are clustered within one branch for both *Haliotis* and *C. destructor*. The putative amino acid sequences for the GHF16 protein from the haemolymph and midgut gland of *C. destructor* were also identical. Similarly the conservation of the catalytic and binding domains between sequences also suggests that the GHF16 protein from other crustaceans may have β-1,3-glucanase activity, and therefore deserves further investigation. The conservation of domains also explains the amplification of a GHF16 gene from the midgut gland and haemocytes of *F. chinensis* and *E. sinensis* (Liu et al., 2009; Zhao et al., 2009).

Given the proposed phylogeny, the original function of the protein is unclear; was the GHF16 protein initially an immune protein which then took on a digestive role? This would be similar to lysozymes in ruminant artiodactyl mammals. These animals express a chicken like lysozyme in the acid stomach. This enzyme evolved from an immune protein to be a digestive enzyme which digests bacteria derived from the rumen (Dobson et al., 1984; Callewaert and Michiels, 2010). Since the GHF16 protein cannot actively select its substrate, it is likely that the real function might be a mix of digestive and immune roles. As an extension of this, the protein may have evolved in two directions. In animals that consume increasing amounts of β-1,3-glucans (laminarin, callose etc.) the activity of β-1,3-glucanase may increase to digest these polysaccharides. Alternatively, it may be an immune protein that deals with potentially pathogenic micro-organisms, particularly in animals that consume rotting material or are long lived. Furthermore, the protein may have co-evolved with the prophenol oxidase system giving rise to the situation where the function of the protein is to amplify the signal of a potentially pathogenic micro-organism (Sritunyalucksana and Söderhäll, 2000). In this system, the binding of a β-1,3-glucan to a binding protein stimulates the activation of pro-phenol oxidase to phenol oxidase. Phenol oxidase catalyses the oxidation of monophenol compounds such as tyrosine that leads to the production of melanin, the products of which are cytotoxic and antimicrobial (Sritunyalucksana and Söderhäll, 2000; Amparyup et al., 2012). This may have occurred in insects, with the protein initially being a β-1,3-glucanase. The protein expressed within the haemolymph may have then lost the catalytic amino acids and evolved into a separate β-glucan binding protein whose function is to amplify the signal of a potentially pathogenic micro-organism (Pauchet et al., 2009). Given these hypotheses, the β-1,3-glucanase from species that do not consume significant amounts of β-1,3-glucans needs to be isolated and characterised in terms of its immune function; that is, its potential to hydrolyse the cell walls of micro-organisms and its ability to activate the prophenol oxidase system in both the haemolymph and digestive fluid. Furthermore, the potential of a GHF16 lipopolysaccharide and β-glucan binding protein expressed within the haemocytes and containing catalytic and binding domains requires re-examination for enzyme activity.

Research into cellulase and hemicellulase enzymes focusses on the characterisation of the enzymes for the digestion of cellulose and hemicellulose within plant cell walls. However what has yet to be addressed is the evolutionary origin of these enzymes. Perhaps like that suggested here, the enzymes original function may have been to hydrolyse the polysaccharides within the cell walls of micro-organisms. The same enzyme may then be utilised to digest plant material as the animal has adopted a leaf litter diet.

Acknowledgements

This work was funded by the Molecular and Medical Strategic Research Centre of Deakin University. The authors would also like to thank Peter Greenaway for the help with the collection of the tissue samples.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2015.05.056.

References

Allardyce, B.J., Linton, S.M., 2008. Purification and characterisation of endo-β-1,4-glucanase and laminarinase enzymes from the gecarcinid land crab *Gecarcoidea natalis* and the aquatic crayfish *Cherax destructor*. J. Exp. Biol. 211, 2275–2287.

Amparyup, P., Sutthangkul, J., Charoensapsri, W., Tassanakajon, A., 2012. Pattern recognition protein binds to lipopolysaccharide and β-1,3-glucan and activates shrimp prophenoloxidase system. J. Biol. Chem. 287, 10060–10069.

Bacic, A., Harris, P.J., Stone, B.A., 1988. Structure and function of plant cell wall. In: Stumpf, P.K., Conn, E.E. (Eds.), The Biochemistry of Plants. Academic Press, New York, pp. 297–371.

Bragatto, I., Genta, F.A., Ribeiro, A.F., Terra, W.R., Ferreira, C.I., 2010. Characterization of a β-1,3-glucanase active in the alkaline midgut of *Spodoptera frugiperda* larvae and its relation to β−glucan-binding proteins. Insect Biochem. Mol. Biol. 40, 861–872.

Bull, A.T., Chesters, C.G.C., 1966. The biochemistry of laminarin and the nature of laminarinase. Adv. Enzymol. Relat. Areas Mol. Biol. 28, 325–364.

Callewaert, L., Michiels, C., 2010. Lysozymes in the animal kingdom. J. Biosci. 35, 127–160.

Cerenius, L., Liang, Z., Duvic, B., Keyser, P., Hellman, U., Palva, E.T., Iwanaga, S., Soderhall, K., 1994. Structure and biological activity of a 1,3-β-D-glucan-binding protein in crustacean blood. J. Biol. Chem. 269, 29462–29467.

Dobson, D.E., Prager, E.M., Wilson, A.C., 1984. Stomach lysozymes of ruminants. I. Distribution and catalytic properties. J. Biol. Chem. 259, 11607–11616.

Du, X.-J., Zhao, X.-F., Wang, J.-X., 2007. Molecular cloning and characterization of a lipopolysaccharide and β-1,3-glucan binding protein from fleshy prawn (*Fenneropenaeus chinensis*). Mol. Immunol. 44, 1085–1094.

Figueiredo, M.S.R.B., Anderson, A.J., 2009. Digestive enzyme spectra in crustacean decapods (Paleomonidae, Portunidae and Penaeidae) feeding in the natural habitat. Aquac. Res. 40, 282–291.

Figueiredo, M.S.R.B., Kricker, J.A., Anderson, A.J., 2001. Digestive enzyme activities in the alimentary tract of redclaw crayfish, *Cherax quadricarinatus* (Decapoda: Parastacidae). J. Crust. Biol. 21, 334–344.

Genta, F.A., Bragatto, I., Terra, W.R., Ferreira, C., 2009. Purification, characterization and sequencing of the major β-1,3-glucanase from the midgut of *Tenebrio molitor* larvae. Insect Biochem. Mol. Biol. 39, 861–874.

Greenaway, P., 2003. Terrestrial adaptations in the Anomura (Crustacea: Decapoda). Mem. Mus. Vic. 60, 13–26.

Greenaway, P., Linton, S.M., 1995. Dietary assimilation and food retention time in the herbivorous terrestrial crab *Gecarcoidea natalis*. Physiol. Zool. 68, 1006–1028.

Gross, P.S., Bartlett, T.C., Browdy, C.L., Chapman, R.W., Warr, G.W., 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific White shrimp, *Litopenaeus vannamei*, and the Atlantic White shrimp, *L. setiferus*. Dev. Comp. Immunol. 25, 565–577.

Henrissat, B., Davies, G., 1997. Structural and sequence-based classification of glycoside hydrolases. Curr. Opin. Struct. Biol. 7, 637–644.

Johnston, D., Freeman, J., 2005. Dietary preference and digestive enzymes activities as indicators of trophic resourece utilization by six species of crab. Biol. Bull. 208, 36–46.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

Juncosa, M., Pons, J., Dot, T., Querol, E., Planas, A., 1994. Identification of active site carboxylic residues in *Bacillus licheniformis* 1,3-1,4-beta-D-glucan 4-glucanohydrolase by site-directed mutagenesis. J. Biol. Chem. 269, 14530–14535.

Kovalchuk, S.N., Sundukova, E.V., Kusaykin, M.I., Guzev, K.V., Anastiuk, S.D., Likhatskaya, G.N., Trifonov, E.V., Nurminski, E.A., Kozhemyako, V.B., Zvyagintseva, T.N., Rasskazov, V.A., 2006. Purification, cDNA cloning and homology modeling of endo-1,3-β-d-glucanase from scallop *Mizuhopecten yessoensis*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 143, 473–485.

Kovalchuk, S.N., Bakunina, I.Y., Burtseva, Y.V., Emelyanenko, V.I., Kim, N.Y., Guzev, K.V., Kozhemyako, V.B., Rasskazov, V.A., Zvyagintseva, T.N., 2009. An endo-(1-3)-β-d-glucanase from the scallop *Chlamys albidus*: catalytic properties, cDNA cloning and secondary-structure characterization. Carbohydr. Res. 344, 191–197.

Kozhemyako, V.B., Rebrikov, D.V., Lukyanov, S.A., Bogdanova, E., Marin, A., Mazur, A.K., Kovalchuk, S.N., Agafonova, E.V., Sova, V.V., Elyakova, L.A., Rasskazov, V.A., 2004. Molecular cloning and characterization of an endo-1,3-β-D-glucanase from the mollusk *Spisula sachalinensis*. Comp. Biochem. Physiol. B 137, 169–178.

Kumagai, Y., Ojima, T., 2009. Enzymatic properties and the primary structure of a [beta]-1,3-glucanase from the digestive fluid of the Pacific abalone *Haliotis discus hannai*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 154, 113–120.

Lee, S.Y., Wang, R., Soderhall, K., 2000. A lipopolysaccharide- and β-1,3-glucan-binding protein from hemocytes of the freshwater crayfish *Pacifastacus leniusculus*. J. Biol. Chem. 275, 1337–1343.

Lépagnol-Descamps, V., Richard, C., Lahaye, M., Potin, P., Yvin, J.-C., Kloareg, B., 1998. Purification and determination of the action pattern of *Haliotis tuberculata laminarinase*. Carbohydr. Res. 310, 283–289.

Lin, Y.-C., Vaseeharan, B., Chen, J.-C., 2008. Identification and phylogenetic analysis on lipopolysaccharide and β-1,3-glucan binding protein (LGBP) of kuruma shrimp *Marsupenaeus japonicus*. Dev. Comp. Immunol. 32, 1260–1269.

Linton, S.M., Greenaway, P., 2004. Presence and properties of cellulase and hemicellulase enzymes of the gecarcinid land crabs *Gecarcoidea natalis* and *Discoplax hirtipes*. J. Exp. Biol. 207, 4095–4104.

Linton, S.M., Greenaway, P., 2007. A review of feeding and nutrition of herbivorous land crabs: adaptations to low quality plant diets. J. Comp. Physiol. B. 177, 269–286.

Linton, S.M., Shirley, A.J., 2011. Isozymes from the herbivorous gecarcinid land crab, *Gecarcoidea natalis* that possess both lichenase and endo-β-1,4-glucanase activity. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 160, 44–53.

Linton, S.M., Allardyce, B.J., Hagen, W., Wencke, P., Saborowski, R., 2009. Food utilisation and digestive ability of aquatic and semi-terrestrial crayfishes, *Cherax destructor* and *Engaeus sericatus* (Astacidae, Parastacidae). J. Comp. Physiol. B. 179, 493–507.

Linton, S., Saborowski, R., Shirley, A., Penny, J., 2014. Digestive enzymes of two brachyuran and two anomuran land crabs from Christmas Island, Indian Ocean. J. Comp. Physiol. B. 1–20.

Liu, F., Li, F., Dong, B., Wang, X., Xiang, J., 2009. Molecular cloning and characterisation of a pattern recognition protein, lipopolysaccharide and β-1,3-glucan binding protein (LGBP) from Chinese shrimp *Fenneropenaeus chinensis*. Mol. Biol. Rep. 36, 471–477.

Omondi, J.G., Stark, J.R., 1995. Some digestive carbohydrases from the midgut gland of *Penaeus indicus* and *Penaeus vannamei* (Decapoda: Penaeidae). Aquaculture 134, 121–135.

Pauchet, Y., Freitak, D., Heidel-Fischer, H.M., Heckel, D.G., Vogel, H., 2009. Immunity or digestion. J. Biol. Chem. 284, 2214–2224.

Pesentseva, M.S., Kusaykin, M.I., Anastyuk, S.D., Sova, V.V., Zvyagintseva, T.N., 2008. Catalytic properties and mode of action of endo-(1 → 3)-β-d-glucanase and β-d-glucosidase from the marine mollusk *Littorina kurila*. Carbohydr. Res. 343, 2393–2400.

Piavaux, A., 1977. Distribution and localization of the digestive laminarinases in animals. Biochem. Syst. Ecol. 5, 231–239.

Rockstroh, M., Muller, S.A., Jende, C., Kerzhner, A., von Bergen, M., Tomm, J.M., 2011. Cell fractionation — an important tool for compartment proteomics. J. Integr. OMICS 1, 135–143.

Rose, T.M., Henikoff, J.G., Henikoff, S., 2003. CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. Nucleic Acids Res. 31, 3763–3766.

Roux, M.M., Pain, A., Klimpel, K.R., Dhar, A.K., 2002. The lipopolysaccharide and β-1,3-glucan binding protein gene is upregulated in white spot virus-infected shrimp (*Penaeus stylirostris*). J. Virol. 76, 7140–7149.

Ruiz-Herrera, J., 1992. Fungal Cell Wall: Structure, Synthesis and Assembly. CRC Press, London.

Song, J.M., Nam, K., Sun, Y.-U., Kang, M.H., Kim, C.-G., Kwon, S.-T., Lee, J., Lee, Y.-H., 2010. Molecular and biochemical characterizations of a novel arthropod endo-[beta]-1,3-glucanase from the Antarctic springtail, *Cryptopygus antarcticus*, horizontally acquired from bacteria. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155, 403–412.

Sova, V.V., Elyakova, L.A., Vaskovsky, V.E., 1970. The distribution of laminarinases in marine invertebrates. Comp. Biochem. Physiol. 32, 459–464.

Sritunyalucksana, K., Söderhäll, K., 2000. The proPO and clotting system in crustaceans. Aquaculture 191, 53–69.

Sritunyalucksana, K., Lee, S.Y., Söderhäll, K., 2002. A β-1,3-glucan binding protein from the black tiger shrimp, *Penaeus monodon*. Dev. Comp. Immunol. 26, 237–245.

Suzuki, M., Horii, T., Kikuchi, R., Ohnishi, T., 1987. Purification of laminarinase from Antarctic krill, *Euphausia superba*. Nippon Suisan Gakkaishi 2, 311–317.

Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. 109B, 1–62.

Tsang, L.M., Chan, T.-Y., Ahyong, S.T., Chu, K.H., 2011. Hermit to king, or hermit to all: multiple transitions to crab-like forms from hermit crab ancestors. Syst. Biol. 60, 616–629.

Wilde, J.E., Linton, S.M., Greenaway, P., 2004. Dietary assimilation and the digestive strategy of the omnivorous anomuran land crab *Birgus latro* (Coenobitidae). J. Comp. Physiol. B. 174, 299–308.

Yeh, M.-S., Chang, C.-C., Cheng, W., 2009. Molecular cloning and characterization of lipopolysaccharide- and β-1,3-glucan-binding protein from the giant freshwater prawn *Macrobrachium rosenbergii* and its transcription in relation to foreign material injection and the molt stage. Fish Shellfish Immunol. 27, 701–706.

Yudkovski, Y., Shechter, A., Chalifa-Caspi, V., Auslander, M., Ophir, R., Dauphin-Villemant, C., Waterman, M., Sagi, A., Tom, M., 2007. Hepatopancreatic multi-transcript expression patterns in the crayfish *Cherax quadricarinatus* during the moult cycle. Insect Mol. Biol. 16, 661–674.

Zakharenko, A.M., Kusaykin, M.I., Kovalchuk, S.N., Anastyuk, S.D., Ly, B.M., Sova, V.V., Rasskazov, V.A., Zvyagintseva, T.N., 2011. Enzymatic and molecular characterization of an endo-1,3-β-d-glucanase from the crystalline styles of the mussel *Perna viridis*. Carbohydr. Res. 346, 243–252.

Zhao, D., Chen, L., Qin, C., Zhang, H., Wu, P., Li, E., Chen, L., Qin, J., 2009. Molecular cloning and characterization of the lipopolysaccharide and β-1, 3-glucan binding protein in Chinese mitten crab (*Eriocheir sinensis*). Comp. Biochem. Physiol. B Biochem. Mol. Biol. 154, 17–24.

Zhu, B.-W., Zhao, J.-G., Yang, J.-F., Mikiro, T., Zhang, Z.-S., Zhou, D.-y., 2008. Purification and partial characterization of a novel [beta]-1,3-glucanase from the gut of sea cucumber *Stichopus japonicus*. Process Biochem. Metab. Eng. 43, 1102–1106.