Proteomic identification, cDNA cloning and enzymatic activity of glutathione *S*-transferases from the generalist marine gastropod, *Cyphoma gibbosum*

Kristen E. Whalen[*], Dexter Morin[†], Ching Yu Lin[‡#], Ronald S. Tjeerdema[‡], Jared V. Goldstone[*], Mark E. Hahn[*]

[*]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

[†]Department of Molecular Biosciences, University of California, Davis, CA 95616

[‡]Department of Environmental Toxicology, University of California, Davis, CA 95616

Current address: Institute of Environmental Health, College of Public Health, National Taiwan University, Taipei, Taiwan

Corresponding authors contact information:

Kristen Whalen

kwhalen@whoi.edu

MS 32, 45 Water Street, Woods Hole, MA 02543

Fax (508) 457-2134

Phone (508) 289-3479

Mark Hahn

mhahn@whoi.edu

MS 32, 45 Water Street, Woods Hole, MA 02543

Fax (508) 457-2134

Phone (508) 289-3242

Short title: Characterization of glutathione *S*-transferases from *Cyphoma gibbosum*

Footnote: The nucleotide and translated amino acid sequences for *C. gibbosum* GSTs have been deposited in GenBank with the following accession nos. EU008563 (CgGSTM1) and EU008562 (CgGSTM2).

Abbreviations footnote: CDNB, 1-chloro-2,4-dinitrobenzene; GST, glutathione S-transferase; nrDB, non-redundant database; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase polymerase chain reaction

ABSTRACT

Glutathione *S*-transferases (GST) were characterized from the digestive gland of *Cyphoma gibbosum* (Mollusca; Gastropoda), to investigate the possible role of these detoxification enzymes in conferring resistance to allelochemicals present in its gorgonian coral diet. We identified the collection of expressed cytosolic *Cyphoma* GST classes using a proteomic approach involving affinity chromatography, HPLC and nanospray liquid chromatography-tandem mass spectrometry (LC-MS/MS). Two major GST subunits were identified as putative mu-class GSTs; while one minor GST subunit was identified as a putative theta-class GST, apparently the first theta-class GST identified from a mollusc. Two *Cyphoma* GST cDNAs (CgGSTM1 and CgGSTM2) were isolated by RT-PCR using primers derived from peptide sequences. Phylogenetic analyses established both cDNAs as mu-class GSTs and revealed a mollusc-specific subclass of the GST-mu clade. These results provide new insights into metazoan GST diversity and the biochemical mechanisms used by marine organisms to cope with their chemically defended prey.

Keywords: *Cyphoma gibbosum*, glutathione S-transferase, gorgonian, natural product, chemical ecology, allelochemical, proteomic

INTRODUCTION

Glutathione S-transferases (GST, EC 2.5.1.18) form a large superfamily of multifunctional enzymes capable of conjugating a broad range of toxic electrophilic xenobiotics with glutathione [1]. GSTs function primarily as detoxification enzymes, generally rendering the resultant products more water soluble (nonreactive conjugate), thereby facilitating excretion. The catalytic versatility and diversity of GSTs can be attributed to the nonspecific nature of the hydrophobic substrate binding site (H-site), and the extensive gene duplication and divergence that has occurred in this superfamily [2, 3]. The soluble GSTs in metazoans are divided into eight classes (alpha, kappa, mu, pi, sigma, theta, omega and zeta) based on sequence identity, immunological and kinetic properties [1, 4, 5]. While the majority of GST enzymatic characterization has focused on mammalian forms, non-vertebrate models offer an exciting opportunity to examine the evolution of GSTs and their adaptive responses to environmental chemicals, including natural products.

Biochemical adaptations involving xenobiotic response genes, such as GSTs, may help explain consumer resistance to dietary chemical threats and provide further insight into the enzymatic mechanisms underlying foraging decisions. For example, the ability of terrestrial invertebrates (insects) to tolerate naturally occurring dietary toxins (i.e., allelochemicals in their host plants) has been linked in part to high constitutive activity of their GSTs (reviewed in [6]). The induction of GSTs in response to dietary allelochemicals may serve as an additional adaptive mechanism to protect against toxicity. While the role of GSTs in plant-herbivore interactions has clearly influenced the ecology of terrestrial consumers, far less is known about the

detoxification enzymes that allow marine invertebrates to exploit allelochemically defended prey.

Only a few studies have addressed the induction of GSTs upon exposure to allelochemicals in marine invertebrates [7-9]. One such study [9] examined GST activity in the digestive gland of a generalist gastropod, *Cyphoma gibbosum*, which exclusively feeds on several families of chemically defended gorgonian corals. The authors reported significantly higher GST activity in field-collected *C. gibbosum* feeding on gorgonians *Gorgonia ventalina* and *Briareum asbestinum,* suggesting that GST expression varies in response to different suites of gorgonian allelochemicals. Additionally, GST activity from *C. gibbosum* cytosolic preparations was among the highest ever reported from a molluscan digestive gland and was similar to values described from *Papilio polyxenes*, a specialist insect that feeds solely on chemically defended plants [10, 11]. In a subsequent study [12], thin-layer chromatographic profiles of nonpolar tissue extracts from *C. gibbosum* feeding on *G. ventalina* did not mirror those of its octocoral prey, lending further support to the idea that this gastropod predator has the capacity to biotransform dietary compounds to readily excretable metabolites. However, the GST isoforms responsible for the detoxification of gorgonian allelochemicals are not known, and in general GST diversity in molluscs, including marine gastropods, is poorly understood.

GST cDNA sequences representing alpha, mu, pi, omega, and sigma GST classes have been identified in molluscs. The majority of sequences are grouped within the pi class, with representatives from cephalopods (squid, octopus) and seven bivalve species ([13-20]; GenBank accession nos. EF194203, EF520700, DQ530213, DQ530212). The endogenous function of pi-

class GSTs in molluscs has yet to be defined; however, studies in mammalian systems indicate that pi GSTs are more involved than other GST classes in the detoxification of prostaglandins and other electrophilic α, β-unsaturated carbonyl compounds [21, 22]. The highest concentrations of prostaglandins in nature have been found in Caribbean gorgonians [23], where the acetoxy acids, hydroxyl methyl esters and hydroxyl acids of 15(R)-prostaglandin A_2 function as feeding deterrents against generalist reef fish [24, 25]. Prostaglandins in the A series can significantly induce GST activity in mammalian cells [26]. We hypothesized that *C. gibbosum* GSTs may conjugate gorgonian allelochemicals, like prostaglandins, potentially alleviating their toxicity.

The objective of the present study was to isolate and characterize the GSTs in *C. gibbosum* digestive gland that may protect this marine consumer from allelochemicals found in its prey. Initially, we targeted pi-class GSTs because of their role in prostaglandin metabolism. However, our attempts to obtain pi-class GSTs from *C. gibbosum* using RT-PCR with degenerate primers designed from an alignment of molluscan GST pi cDNA sequences [15] failed to yield any GST sequences. Therefore, we initiated a proteomic approach involving affinity chromatography coupled with HPLC and mass spectrometry to provide an unbiased assessment of GST protein diversity in the digestive gland of *C. gibbosum*. Here we report the identification of several mu-class GST proteins in *C. gibbosum* digestive gland, the cloning and phylogenetic characterization of two complete mu-class GST cDNAs from *C. gibbosum,* and partial peptide sequences of a theta-class GST, apparently the first member of this class identified from a mollusc. These results contribute to a better understanding of GST diversity in molluscs and of the biochemical

resistance mechanisms used by marine consumers to cope with their allelochemically defended prey.

MATERIALS AND METHODS

Reagents

CDNB, dithiothreitol (DTT), potassium phosphate, potassium chloride, EDTA, protease inhibitor cocktail (4-(2-aminoethyl)benzenesulfonyl fluoride, aprotinin, bestatin hydrochloride, E-64, leupeptin, pepstatin A), SDS, boric acid, NaCl, sodium acetate, GSH, GSH-agarose (G4510) was purchased from Sigma (St. Louis, MO). Bradford reagents and molecular weight standards for protein gels were purchased from BioRad (Hercules, CA). Novex® Tris-glycine gels and Sample Treatment Buffer were purchased from Invitrogen (Carlsbad, CA). Silver SNAP® Stain for Mass Spectrometry was purchased from Pierce (Rockford, IL).

Animals

A total of 42 adult *Cyphoma gibbosum* (ca 2-3 cm length) were collected from five shallow reefs (< 20m) near the Perry Institute of Marine Science (PIMS), Lee Stocking Island, Exuma Cays, Bahamas in January 2006 and transported to wet laboratory facilities provided by PIMS. Individuals were allowed to feed on a control diet (e.g., alginic acid and freeze-dried squid paste prepared as described in [27], or one of six gorgonian diets - *Briareum asbestinum, Eunicea mammosa, Gorgonia ventalina, Pseudopterogorgia acerosa, Pseudopterogorgia americana, Plexaura homomalla*) for four days. Upon completion of the feeding assay the digestive glands were immediately dissected and either preserved in RNALater® (Ambion, Austin, TX) or frozen

in liquid nitrogen and transported back to the Woods Hole Oceanographic Institution. Tissues were maintained at -80°C until further analysis.

Enzyme purification

Individual digestive glands (n = 42) were homogenized 1:4 (w/v) in ice-cold homogenization buffer (0.1 M potassium phosphate, 1 mM EDTA, 1 mM DTT, 1.15% potassium chloride, protease inhibitor cocktail (1X); pH 7.5) with an IKA Ultra Turrax T8 homogenizer (Wilmington, NC) for 30 sec on ice. All subsequent steps were carried out at 4°C. Cytosol was isolated by centrifugation of the crude homogenate at 750 x *g* for 10 min then, without stopping, at 12,000 x *g* for 10 min using a Beckman J2-21 centrifuge (Fullerton, CA). The supernatant was carefully removed, avoiding both the fatty layer and the pellet and transferred to an ultracentrifuge tube and centrifuged at 100,000 x *g* for 70 min using a Beckman L8-60M ultracentrifuge (Fullerton, CA). An aliquot of the supernatant was taken for cytosolic protein determination using the BCA protein assay method [28] with BSA as the standard and the remaining cytosol was flash frozen in liquid nitrogen until further analysis. Samples were stored for several months in this manner suffered no apparent loss of enzymatic activity.

GSTs were purified using size exclusion chromatography, affinity chromatography and HPLC following a modified method from [29] . Cytosolic samples from individual snail digestive glands were purified separately. Initially, a subset of cytosol samples (n = 8) were used to determine the optimal buffer conditions and volumes for GST purification. Eluted 1-mL fractions from both size-exclusion and affinity columns were sampled for activity towards

CDNB and assayed for protein using the Bradford assay [30]. These results helped streamline the GST purification process for the remaining cytosol samples.

A PD-10 desalting column (bed volume 8.3 mL, bed height 5 cm, 5K NMWL; GE Healthcare, Piscataway, NJ) containing Sephadex G-25 matrix was equilibrated in Buffer A (50 mM Tris buffer, 1 mM EDTA, 1 mM DTT; pH 6.0) and 2.5 mL of crude cytosol were applied to the column. Cytosolic samples less than 2.5 mL were brought up to this volume with Buffer A and then applied to the column. GST proteins were eluted by gravity with Buffer A and fractions containing GST activity were combined (approx. 6 mL of elute) and then applied to a GSH-agarose affinity column (bed volume 0.5 mL, 0.8 x 4 cm i.d.) equilibrated in Buffer A. The affinity column was washed with 7 mL of Buffer B (Buffer A + 0.5 M NaCl) to rinse away non-specific proteins. Retained GSTs were then eluted with 5 mL of Buffer C (50 mM Tris-base, 1 mM EDTA, 1 mM DTT, 0.5 M NaCl, 50 mM glutathione; pH 9.5) and fractions containing GST activity were then combined, buffer exchanged to low salt concentration, and concentrated with Amicon Ultra-4 centrifugational filters (5K NMWL membrane; Millipore, Billerica, MA). Protein concentrations of Amicon concentrates were determined with the NanoOrange protein quantitation kit (Molecular Probes, Eugene, OR).

Affinity-purified GSTs from individuals feeding on the same diet were pooled (100 μL injection volume) and injected onto a reverse phase Vydac protein/peptide column (model #218 TP 52; C18 μm 250 mm x 2.1 mm) and separated on a Waters 600 MultiSolvent Delivery System, with a flow rate of 0.5 mL/min. Peaks were detected using a Waters 2487 Dual Wavelength Absorbance Detector (λ = 214 nm). Mobile phase A consisted of 38% acetonitrile, 62% water

and 0.1% trifluoroacetic acid (TFA). Mobile phase B consisted of 80% acetonitrile, 20% water and 0.1% TFA. The initial mobile phase consisted of 100% A. GST subunits were separated using a linear gradient from 0 to 40% B in 22 min, and 40 to 100% B in 37 min. Peaks were hand-collected, centrifugally evaporated to dryness, resuspended in 1X sample treatment buffer and separated by 12% Novex® Tris-glycine SDS-PAGE gel electrophoresis according to manufacturer's instructions (Invitrogen, Carlsbad, CA). Bands were visualized using the SilverSNAP® staining kit for mass spectrometry, excised from the gel and delivered to the UCDavis Proteomics Facility, Davis, CA for in-gel digestion and LC-MS/MS analysis.

Nanospray LC-MS/MS and database analysis

Protein bands were excised from the gel and digested with trypsin as described in [31]. The trypsin-digested samples were analyzed using a Eksigent Nano LC 2-D system (Eksigent, Dublin, CA) coupled to an LTQ ion trap mass spectrometer (Thermo-Fisher, San Jose, CA) interfaced with a New Objective Picoview Nano-spray ionization source (Woburn, MA) to identify peptide fragments. Digested peptides were loaded on a reverse-phase Agilent Nano-trap (Zorbax 300SB-C18, 300Å, Agilent Technologies, Santa Clara, CA) at a loading flow rate of 5 μL/min for 10 min. The buffers used for the reverse-phase chromatography were 0.1% formic acid in water (solvent A) and 95% acetonitrile in 0.1% formic acid in water (solvent B). Peptides were eluted from the trap and chromatographically separated on a reverse-phase capillary column (Pico Frit, 75 μm x 15 cm, tip 5 μm: New Objective) packed in-house with Magic C18 AQ (3 μm, 100Å:Michrom BioResources, Auburn, CA) with a 40-min linear gradient of 2-80% solvent B (as described below) at a flow rate of 300nL min^{-1}. The LTQ parameters were as follows: electrospray potential, 1.8 kV; source temperature, 180º; collision

energy, 35%; dynamic exclusion duration, 1min. The MS survey scan followed by ten MS/MS scans were consecutively acquired over the LC gradient. Upon completion of an LC-MS/MS run, the MS/MS spectra from each survey were charge state deconvoluted and searched against the non-redundant NCBI protein database using BioWorks version 3.3 (Thermo Scientific, Waltham, MA) and MASCOT (Matrix Science, London, UK) and against the MSDB database (Imperial College, London, UK) using GPM software (http://www.thegpm.org). Protein fragments with significant hits to eukaryotic GSTs were then manually validated.

GST activity assay

Enzyme activity was measured using CDNB as a substrate by the method of [32] optimized for *C. gibbosum* [9] in a microplate format. The reaction mixture (in a final volume of 200 μL) contained 0.1 M potassium phosphate buffer, 1.0 mM EDTA, pH 7.5, 1 mM CDNB, 1 mM reduced glutathione (GSH) and 2μg of protein. CDNB was solubilized in ethanol and constituted 1% of the final reaction mixture volume. The reaction incubated at 25°C was initiated by the addition of CDNB and performed in triplicate. The conjugation of CDNB with GSH was measured as the increase in absorbance at 340 nm ($\Delta\varepsilon_{340}$ 0.00503 μM^{-1} cm^{-1}) using a tunable microplate reader (Versamax, Molecular Devices, Sunnyvale, CA). Activity was calculated using protein concentrations determined via the Bradford assay with BSA as a standard.

RNA isolation and cDNA synthesis

Total RNA was purified from *C. gibbosum* digestive glands using the RNeasy Maxi Kit and DNAse treated using a RNase-free DNAse Kit (Qiagen, Valencia, CA) following the

manufacturer's instructions. The integrity of each RNA sample was checked by electrophoresis on 1% agarose-formaldehyde gel in MOPS buffer and visualized with ethidium bromide under UV light. Total RNA concentration was determined using a Nanodrop spectrophotometer (Wilmington, DE). Poly(A)+ RNA was isolated using the MicroPoly(A)Purist mRNA purification kit (Ambion, Austin, TX) according to the manufacturer's instructions. Poly(A)+ RNA was pooled from seven individuals feeding on either a control diet or one of six gorgonian species (0.14 μg poly(A) RNA/individual) to ensure representation of all GSTs expressed under various dietary conditions. One microgram of pooled RNA was primed with modified oligo (dT) primers and used to create an adaptor-ligated ds cDNA library synthesized using the Marathon cDNA Amplification Kit (BD Biosciences, Palo Alto, CA) according to the manufacturer's instructions.

Primer design, RACE-PCR, cloning, sequencing

Degenerate primers were designed to *C. gibbosum* GST peptide sequences identified from LC-MS/MS with homology to mu-class GSTs from abalone (*Haliotis discus discus*) and oyster (*Crassostrea gigas*) (GSTm_R10) (Table 1), and theta-class GSTs from sea bream (*Pagrus major, Sparus aurata*), plaice (*Pleuronectes platessa*) and bass (*Micropterus salmoides*) (primers not shown). Primers were obtained from Sigma Genosys (St. Louis, MO). PCR products were generated by rapid amplification of cDNA ends (RACE) using GST degenerate primers in combination with specific oligonucleotides designed to adaptor sequences located on the 3' and 5' ends of the cDNA. Amplification of PCR products was carried out according to the Advantage 2 PCR Enzyme Kit (Clontech, Mountain View, CA) and cycling parameters were as follows: 94°C for 30 sec; 30 cycles of 94°C for 5 sec, annealing at 55°C for 30 sec, extension at

68°C for 2 min. Primers were used at the following concentrations: 10 μM of adaptor primer and 100 μM of degenerate primer in a 50 μL reaction. Sequences obtained by RACE were then used to design specific primers (GSTm_F13 and GSTm_F15) to clone full-length *C. gibbosum* GST cDNAs. PCR products were cloned into pGEM-T (Promega, Madison, WI) and sequenced in both directions using an ABI 3730XL capillary sequencer at the Keck facility of the Josephine Bay Paul Center for Comparative Molecular Biology & Evolution at the Marine Biological Laboratory (Woods Hole, MA). A minimum of twelve clones were sequenced for each PCR fragment.

Sequence Analysis

Clones were grouped based on similar nucleotide sequence using Sequencher (Gene Codes Co., Ann Arbor, MI) and used to generate a consensus sequence that was compared to known GST coding sequences using BLASTX. Multiple alignments of deduced amino acid sequences were performed using ClustalW [33]. Ambiguous alignment positions were excluded from further analyses. Phylogenetic relationships were investigated using Bayesian techniques as implemented in the computer program MrBayes (v 3.1.2; [34]). MrBayes estimates posterior probabilities using Metropolis-Hastings coupled Monte Carlo Markov chains (MC^3). We performed MC^3 estimates with uninformative prior probabilities using the WAG model of amino acid substitution [35] and prior uniform gamma distributions approximated with four categories (WAG+I+Γ). Four incrementally heated, randomly seeded Markov chains were run for $3x10^6$ generations, and topologies were sampled every 100^{th} generation. The MC^3 burnin values were conservatively set at $1x10^6$ generations. Posterior probabilities of topologies and clades were estimated from the sampled topologies after removal of the initial MC^3 burnin.

RESULTS

GST purification and activity

Digestive glands of individual *C. gibbosum* feeding for four days on one of six gorgonian diets (n = 4-7 snails per gorgonian diet) or a control diet (n = 13 snails) averaged 0.22 ± 0.11 g snail^{-1} (n = 42; mean ± SD). The yield of cytosolic protein was 125.6 ± 52.8 mg protein per g digestive gland wet weight. Glutathione transferase activity of crude cytosol ranged from 0.57 to 5.82 U mg^{-1} protein, with an average of 2.67 ± 1.27 U (mg protein)$^{-1}$ (n = 42), using CDNB as a substrate. Following size exclusion chromatography and affinity purification of the GSTs from *C. gibbosum* digestive gland, an average of 25 ± 14 μg of GST protein was collected from each digestive gland (n = 36), representing approximately 0.2% of the total cytosolic protein. A 220-fold purification of GST protein was obtained using size-exclusion and affinity chromatography (Table 2). Representative GST elution profiles from the Sephadex G-25 and GSH-agarose columns are presented in Figure 1. The majority of the GST activity eluted from GSH-agarose in fractions 18-20. A small amount of GST activity also was noted in the flow-through fractions from the affinity column (Figure 1); however, the resultant activity measurements represent a minor contribution to the entire GST activity of the sample and therefore these fractions were not collected for subsequent analysis.

Gel electrophoresis and GST identification

To provide an initial characterization of affinity-purified proteins, bound protein fractions obtained from GSH-agarose affinity columns from two individuals (feeding either on *P. americana* or the control diet) were concentrated, combined, and 15 μL of the pooled sample was analyzed on a 12% Novex® Tris-glycine SDS-PAGE gel (Figure 2). Three distinct bands

ranging in size from 21 to 24 kDa were detected by silver staining and excised from the gel for identification via LC-MS/MS. All three bands contained peptide sequences that closely matched those of mu-class GSTs in the database (Table 3). This pooled GST affinity-purified fraction was used as standard for subsequent SDS-PAGE gels.

HPLC and LC-MS/MS based GST identification

Affinity-purified fractions from individual digestive glands were concentrated and then equal volumes of affinity-purified GSTs were pooled within each snail diet for further HPLC separation of GST subunits. HPLC analysis of affinity-purified GSTs identified a total of fourteen unique (2 major and 12 minor) peaks. The relationship between snail diet and GST subunit expression will be described elsewhere (K. Whalen *et al*., manuscript in preparation). Here, we present the proteomic characterization of HPLC separated GST subunits. Figure 3 shows a representative HPLC separation of digestive gland GST subunits from four *C. gibbosum* feeding on *B. asbestinum*. For peak collection purposes, the fourteen peaks were grouped into seven fractions (B-H; A= injection peak) (Figure 3).

The HPLC fractions were separated on 12% Tris-glycine SDS-PAGE gels (Figure 4) and bands were excised for proteomic analysis. Peptide sequences were determined for fractions B-H using nanospray LC-MS/MS and identified by comparison to proteins in the Genbank database. Fractions D and F, with elution times of 18 and 22 minutes, respectively, represent the two major GST subunits. Both of these fractions yielded peptides that closely matched mu-class GSTs (Table 3; Figure 5). Peptides in the remaining fractions also matched GST mu-class proteins, with the exception of fraction B, which yielded eight peptides that matched theta-class GSTs

from fish (Table 3; Figure 6). Manual validation of LC-MS/MS spectra verified five of the *C. gibbosum* peptide sequences as identical to the fish GST-theta peptides (Table 3).

C. gibbosum cDNA cloning and sequence analysis

To obtain full-length cDNA sequences for *C. gibbosum* GSTs, polypeptides identified by mass spectrometry were used to design degenerate primers. The mu-class GST degenerate primer (GSTm_R10), designed to the peptide sequence AYMASDK, yielded two 600-bp fragments that were confirmed by BLAST searches to be partial cDNAs encoding mu-class GSTs. Complete cDNA sequences were generated by 3' RACE, revealing open reading frames encoding predicted proteins of 215 amino acids each. The predicted proteins displayed 61% and 62% amino acid identity to Pacific oyster GST mu for transcript 1 (CgGSTM1; GenBank accession no. EU008563) and transcript 2 (CgGSTM2; GenBank accession no. EU008562), respectively. The two *C. gibbosum* GST mu cDNAs are 78% identical and 88% similar (BLOSUM62) to each other at the protein level. Translated nucleotide sequences predict a molecular mass of 25.0 kDa for CgGSTM1 and 25.2 kDa for CgGSTM2. Figure 5 shows an alignment of these two sequences with several invertebrate and vertebrate mu-class GST sequences. Amino acid residues involved in glutathione binding as determined using the Pfam database [36] are entirely conserved among *C. gibbosum*, vertebrate and invertebrate sequences included in the alignment, while substrate binding sites are poorly conserved.

Two *C. gibbosum* tryptic peptides identified in the LC-MS/MS analysis (ITQSNAILR and AYMASDK), were found to be identical to corresponding *C. gibbosum* GST translated

nucleotide sequences (nucleotides 583-603 in CgGSTM1 and CgGSTM2; and 205-231 in CgGSTM1, respectively) (Table 3).

To infer relationships between *C. gibbosum* predicted protein sequences and other GST sequences, a phylogenetic tree was constructed using Bayesian techniques. *Cyphoma* GSTM1 and GSTM2 clearly group within the mu-class GST clade and appear most closely related to other molluscan mu GSTs, which together form a strongly supported subgroup within this clade (Figure 7).

DISCUSSION

Dietary toxins present physiological challenges to marine consumers, such as *Cyphoma gibbosum*, that feed solely on chemically defended gorgonians. The high concentration of allelochemicals in gorgonian corals [27, 37-40], coupled with the findings that digestive glands of *C. gibbosum* contain high levels of GST activity [9], prompted an investigation to identify and biochemically characterize those GST enzymes possibly responsible for detoxifying gorgonian allelochemicals.

The cytosolic GST activities in *C. gibbosum* digestive gland (Table 2 and [9]) are among the highest recorded for any molluscan tissue (reviewed in [9, 41]). The majority of studies reporting GST activity from molluscs have used GST activity measurements as an indicator of pollutant exposure [13, 42-45]. Although the level of GST activity in *C. gibbosum* is high even in comparison to that of molluscs exposed to pollutants known to induce GST expression [46-48], this high activity is unlikely to be a response to anthropogenic contamination, given the remote

site of collection. Rather, the high GST activity may reflect an adaptation that facilitates consumption of allelochemical-rich prey. Additional findings indicate that *C. gibbosum* GSTs are expressed constitutively at high levels regardless of the gorgonian diet, providing further support for this hypothesis [49]. Whether high GST activity is common among marine species that feed exclusively on chemically defended food is unknown, and may depend upon the specific suite of allelochemicals present in the diet. Regardless of its origin, the high GST activity in *C. gibbosum* suggested that this species would be a rich source of GST enzyme(s), enabling proteomic studies to further characterize GSTs potentially involved in allelochemical metabolism.

The results of the proteomic and molecular analyses in the present study clearly show that the predominant cytosolic GSTs expressed in *C. gibbosum* digestive gland are mu-class GSTs. These include the major HPLC peaks (fractions D and F) as well as several minor peaks. According to the current system of GST classification, GSTs sharing greater than 60% identity fall within the same class, while those with less than 30% identity are assigned to separate classes [1]. The classification of *C. gibbosum* GSTs as members of the mu class is supported by both the homology searches of *C. gibbosum* tryptic peptides (Table 3) and the phylogenetic analysis of translated *C. gibbosum* GST cDNAs (Figure 7). The presence of mu-class GSTs in molluscs had been predicted based on immunoblot analysis [43, 50, 51], but only recently have molluscan GST mu nucleotide sequences been determined ([16, 18], Genbank accession no. ABF67506). Our identification of two GST cDNAs in *C. gibbosum* provides the first published description of GST-mu forms from gastropods. Additionally, phylogenetic analysis revealed a mollusc-specific subclass within the GST-mu clade.

Proteomic results indicate the possibility of additional GST mu subunits beyond the two cDNAs identified here. Only two of the twelve *Cyphoma* tryptic peptides identified as matching abalone and oyster mu-class GSTs were found to be encoded by the two cDNA clones isolated from *C. gibbosum*. Not surprisingly these peptides corresponded to fairly conserved regions of the GST protein. However, three peptides (KAAYFEALPAK; SFLGDQQFFAGSK; and IMQPGSLDAFPTLLAFMGRIEALPAIK) identified as matching abalone mu-class GST isoform-1 are quite divergent from the translated *C. gibbosum* sequences, indicating that at least one, and possibly as many as three, additional GST mu subunits are present in *Cyphoma*. HPLC analysis of affinity-purified extracts suggests that there may be additional mu-class subunits, perhaps as many as thirteen (represented by the 13 peaks in fractions C-H, Figure 3), the majority of which are minor components. In addition, visualization of hand-collected HPLC fractions by SDS-PAGE and silver staining (Figure 4) revealed the major peak in fraction F to be a composite of two separate GST subunits that were unresolvable by HPLC. Peak F from affinity-purified extracts from each of the seven diets gave this similar double banding pattern (data not shown), indicating that both subunits are universally expressed regardless of snail diet. Both bands were later confirmed to be mu-class GSTs by LC-MS/MS, thus increasing the count of potential mu-class subunits identified by HPLC analysis to fourteen. Whether the HPLC peaks represent truly unique subunits or HPLC-resolved post-translationally modified variants of the same subunit remains unclear. Multiple GST mu subunits in a single species are not uncommon; five distinct mu-class subunits have been identified from humans [1]. Overall, our results provide conclusive evidence for at least three distinct GST-mu isoforms, and perhaps more, in *C. gibbosum*.

In this study, we identified a theta-like GST from *C. gibbosum* digestive gland and by manual validation confirmed the sequence of five peptides to be identical to theta and theta-class-related GSTs from fish. The HPLC fraction B shows a single peak, suggesting that only one subunit is expressed. Several attempts to amplify theta-like GSTs from *C. gibbosum* digestive gland mRNA with degenerate primers designed to these peptide sequences were not successful.

The identification of a putative theta-class GST in *C. gibbosum* was unexpected. To the best of our knowledge, no theta-like GSTs have been identified previously in a molluscan species, although theta-class GSTs exist in vertebrates, arthropods, polychaetes, algae and bacteria [52, 53]. Theta-class GSTs in general have been notoriously hard to identify because they normally do not bind to affinity matrices such as GSH-agarose [1]. In addition, most lack detectable activity toward CDNB, and thus can be missed if GST activity is the only means of GST isolation [54, 55]. The unique activity of theta-class GSTs can be traced to the presence of the essential Ser-11, responsible for glutathione deprotonation and activation, in place of the tyrosine found in the alpha, mu, and pi class GSTs [56]. Theta-class GSTs can be further distinguished from alpha/mu/pi class isoforms because of their high affinity for glutathione (high K_m), but low affinity for glutathione-conjugates. The diminished product retention in the active site of theta GSTs favors increased substrate turnover in comparison to alpha/mu/pi forms, which have a greater capacity to sequester conjugated products [56]. Generally, it is believed that the theta-class GSTs gave rise to the alpha/mu/pi classes via gene duplication events [52]; however, this has been called into question more recently [57].

It is interesting to note that *Cyphoma* theta-like peptides matched only theta-class GSTs from fish rather than those from terrestrial invertebrates. Using phylogenetic relationships inferred from Bayesian analyses, Lee et al. [58] found that theta class GSTs formed two distinct, well-supported clades, one ("theta A") containing only fish representatives and the other ("theta B") including theta-class GSTs from mammals, fish, birds, and invertebrates. Recent evidence has confirmed the deeply rooted nature of the "theta A' clade by the identification of a rho-GST—once thought to be a fish-specific class [58]—in a primitive cephalochordate [59]. The *Cyphoma* GST theta-like peptides matched both rho- and theta-class fish sequences in the "theta A" clade, providing the first evidence for invertebrate members of this group. Together with the findings of Lee et al. [58], these results suggest that the two theta clades, both of which include vertebrate and invertebrate representatives, resulted from an ancient divergence.

Several natural compounds, including alpha-tocopherol, coumarin, and indole-3-carbinol, have been identified as potent inducers of theta GSTs in some mammals [60]. However, it is the dehalogenase activity of theta-class GSTs that makes this class of enzymes so unique [56]. For example, both bacterial and mammalian theta-class GSTs are capable of metabolizing dichloromethane (DCM) to formaldehyde [61]. The oceans are the largest source of organohalogens [62], with representatives from cnidarians including briarane diterpenes [63], bromo-, chloro- and indo-vulones, clavulones and punaglandins (related to mammalian prostaglandins) [64-66], and chlorinated sterols [67]. It is possible that *C. gibbosum* theta-like GSTs may have evolved to protect this consumer against halogenated compounds from its gorgonian host.

Identifying allelochemical substrates for GSTs is challenging [6]. Many plant allelochemicals and marine natural products contain the appropriate functional groups that can be directly conjugated with GSH [68, 69]. *Cyphoma gibbosum* regularly feeds on the gorgonian *Plexaura homomalla,* which contains high concentrations (up to 8% of the dry weight) of prostaglandins (15(R)-PGA_2) [23, 70]. Prostaglandins are known to suppress cell proliferation; however, overexpression of mu-class GSTs resulting in increased conjugation of prostaglandins may inhibit the antiproliferative effects of these compounds [71]. Naturally occurring prostaglandins (PGA_2 and PGJ_2) undergo enzymatic conjugation by purified human GSTM1a-1a [22]. Thus, high expression of GST mu isoforms may allow *C. gibbosum* to tolerate the chemical defenses of its host and subsequently feed longer than would otherwise be possible. Preliminary studies have demonstrated that prostaglandins found in gorgonians significantly inhibit *Cyphoma* GST activity *in vitro* [49], consistent with the idea that gorgonian prostaglandins may be substrates for *Cyphoma* GSTs.

In summary, a proteomic approach was successful at identifying peptides representing the collection of expressed GST subunits in *C. gibbosum* digestive gland. This technique is a valuable alternative to designing degenerate primers to all of the GST classes suspected of being expressed in the sample. Using proteomics, we were able to identify several mu-class GSTs and one theta-class GST subunit(s), the latter class identified for the first time in a mollusc. The manually validated GST peptides were used to design primers to amplify two full-length *C. gibbosum* mu-class GST cDNAs. Future studies will focus on identifying gorgonian allelochemicals that are substrates for *C. gibbosum* GSTs using a bioassay-guided fractionation approach; expression of recombinant proteins will allow further characterization of individual

GST isoforms. The high GST activity and diversity of GST isoforms in *C. gibbosum* may protect this generalist predator against dietary chemicals while conferring a selective advantage over other consumers that avoid chemically defended prey.

ACKNOWLEDGEMENTS

We thank C. Gaebe, T. Rioux, L. Robinson, A. Tarrant and the staff at the Caribbean Marine Research Center for their assistance in animal collection. We also thank Y-J. Lee and M. Salemi at the UC Davis Proteomics Facility for help in interpreting the results of LC-MS/MS analyses. Support was provided by the WHOI-Cole Ocean Ventures Fund (KEW), the WHOI Ocean Life Institute (KEW and MEH), a grant from Walter A. and Hope Noyes Smith (MEH), the National Science Foundation Graduate Research Fellowship (KEW), and by the National Institutes of Health (P42-ES007381 and R01-ES015912 to JVG).

References Cited

[1] D. Sheehan, G. Meade, V.M. Foley, C.A. Dowd, Biochem. J. 360 (2001) 1-16.

[2] Y. Ivarsson, A.J. Mackey, M. Edalat, W.R. Pearson, B. Mannervik, J. Biol. Chem. 278 (2003) 8733-8738.

[3] W.R. Pearson, Methods Enzymol. 401 (2005) 186-204.

[4] J.D. Hayes, J.U. Flanagan, I.R. Jowsey, Annu. Rev. Pharmacol. Toxicol. 45 (2005) 51-88.

[5] D.W. Nebert, V. Vasiliou, Hum. Genomics 1 (2004) 460-464.

[6] X. Li, M.A. Schuler, M. Berenbaum, Annu. Rev. Entomol. 52 (2007) 231-253.

[7] B.C. DeBusk, S.S. Chimote, J.M. Rimoldi, D. Schenk, Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol. 127 (2000) 133-142.

[8] J.M. Kuhajek, D. Schlenk, Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol. 134 (2003) 473-479.

[9] N.H. Vrolijk, N.M. Targett, Marine Ecology Progress Series 88 (1992) 237-246.

[10] K. Lee, Insect Biochem. 21 (1991) 353-361.

[11] K. Lee, M. Berenbaum, Biochem. Syst. Ecol. 20 (1992) 197-207.

[12] G. Cronin, M.E. Hay, W. Fenical, N. Lindquist, Marine Ecology Progress Series 119 (1995) 177-189.

[13] P. Hoarau, G. Damiens, M. Romeo, M. Gnassia-Barelli, M.J. Bebianno, Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol. 143 (2006) 196-203.

[14] H.-L. Yang, Q.-Y. Zeng, E.-Q. Li, S.-G. Zhu, X.-W. Zhou, Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 139 (2004) 175-182.

[15] P. Doyen, P. Vasser, F. Rodius, Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol. 140 (2005) 300-308.

[16] B. Myrnes, I.W. Nilsen, Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol. 144 (2007) 403-407.

[17] H. Munasinghe, H.S. Kang, J. Lee, Journal of the World Aquaculture Society 37 (2006) 96-106.

[18] I. Boutet, A. Tanguy, D. Moraga, Marine Biology 146 (2004) 53-64.

[19] S.I. Tomarev, R.D. Zinovieva, K. Guo, J. Piatigorsky, J. Biol. Chem. 268 (1993) 4534-4542.

[20] S.I. Tomarev, R.D. Zinovieva, J. Piatigorsky, J. Biol. Chem. 266 (1991) 24226-24231.

[21] K. Berhane, M. Widersten, A. Engstrom, J.W. Kozarich, B. Mannervik, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 1480-1484.

[22] J.J. Bogaards, J.C. Venekamp, P.J. van Bladeren, Chem. Res. Toxicol. 10 (1997) 310-317.

[23] A.J. Weinheimer, R.L. Spraggins, Tetrahedron Lett. 59 (1969) 5185-5189.

[24] J.R. Pawlik, W. Fenical, Marine Ecology Progress Series 52 (1989) 95-98.

[25] D.J. Gerhart, Marine Ecology Progress Series 19 (1984) 181-187.

[26] K. Uchida, Mechanisms of Ageing and Development 116 (2000) 135-140.

[27] W. O'Neal, J.R. Pawlik, Marine Ecology Progress Series 240 (2002) 117-126.

[28] P.K. Smith, R.I. Krohn, G.T. Hermanson, A.K. Mallia, F.H. Gartner, M.D.e.a. Provenzano, Anal. Biochem. 150 (1985) 76-85.

[29] R.T. Donham, D. Morin, W.T. Jewell, S.A. Burns, A.E. Mitchell, M.W. Lame, H.J. Segall, R.S. Tjeerdema, Aqua. Toxicol. 71 (2005) 203-214.

[30] M.M. Bradford, Anal. Biochem. 72 (1976) 248-254.

[31] A. Shevchenko, M. Wilm, M. Mann, Anal. Chem. 68 (1996) 850-858.

[32] W.H. Habig, M.J. Pabst, W.B. Jakoby, J. Biol. Chem. 249 (1974) 7130-7139.

[33] J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res. 22 (1994) 4673-4680.

[34] F. Ronquist, J.P. Huelsenbeck, Bioinformatics 19 (2003) 1572-1574.

[35] S. Whelan, N. Goldman, Mol. Biol. Evol. 18 (2001) 691-699.

[36] A.J. Marchler-Bauer A, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH., Nucleic Acids Res. 35 (2007) 237-240.

[37] A.D. Rodriguez, Tetrahedron 51 (1995) 4571-4618.

[38] P.W. Sammarco, J.C. Coll, Marine Ecology Progress Series 88 (1992) 93-104.

[39] J.C. Coll, Chemical Reviews 92 (1992) 613-631.

[40] V.J. Paul, M.P. Puglisi, Nat. Prod. Rep. 21 (2004) 189-209.

[41] G. Le Pennec, M. Le Pennec, Aquat. Toxicol. 2003 (2003) 131-142.

[42] X.N. Verlecar, N. Pereira, S.R. Desae, K.B. Jena, Singdha, Current Science 91 (2006) 1153-1157.

[43] D. Sheehan, J. McIntosh, A. Power, P.J. Fitzpatrick, Biochem. Soc. Trans. 23 (1995) 419-422.

[44] P.J. Fitzpatrick, J. O'Halloran, D. Sheehan, D.R. Walsh, Biomarkers 2 (1997) 51-56.

[45] A. Kaaya, S. Najimi, D. Ribera, J.F. Narbonne, A. Moukrim, Bull. Environ. Contam. Toxicol. 62 (1999) 623-629.

[46] M. Boryslawskyj, A.C. Garrood, J.T. Pearson, Mar. Environ. Res. 24 (1988) 101-104.

[47] P.J. Fitzpatrick, D. Sheehan, D.R. Livingstone, Mar. Environ. Res. 39 (1995) 241-244.

[48] R.F. Lee, Marine Ecology Progress Series 46 (1988) 33-36.

[49] K.E. Whalen, Functional characterization and expression of molluscan detoxification enzymes and transporters involved in dietary allelochemical resistance, Biology Department, Joint Program in Oceanography at the Massachusetts Institute of Technology and the Woods Hole Oceanographic Institution, 2008, p. 362.

[50] P. Hoarau, G. Garello, M. Bnassia-Barelli, M. Romeo, J.-P. Girard, Aquat. Toxicol. 68 (2004) 87-94.

[51] M.-L. Vidal, P. Rouimi, L. Debrauwer, J.-F. Narbonne, Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol. 131 (2002) 477-489.

[52] S.E. Pemble, J.B. Taylor, Biochem. J. 287 (1992) 957-963.

[53] J.-S. Rhee, Y.-M. Lee, D.-S. Hwang, E.-J. Won, S. Raisuddin, K.-H. Shin, J.-S. Lee, Aquat. Toxicol. 83 (2007) 104-115.

[54] K.E. Comstock, M. Widersten, X.-Y. Hao, D. Henner, B. Mannervik, Arch. of Biochem. and Biophys. 311 (1994) 487-495.

[55] M.F. Lopez, W.F. Patton, W.B. Sawlivich, H. Erdjument-Bromage, P. Barry, K. Gmyrek, T. Hines, P. Tempst, W.M. Skea, Biochem. Biophys. Acta 1205 (1994) 29-38.

[56] S. Landi, Mutation Research 463 (2000) 247-283.

[57] B. Blanchette, X. Feng, B.R. Singh, Mar. Biotechnol. 9 (2007) 513-542.

[58] Y.-M. Lee, J. Soo Seo, S.-O. Jung, I.-C. Kim, J.-S. Lee, Biochem. Biophys. Res. Commun. 346 (2006) 1053-1061.

[59] C. Fan, S. Zhang, Z. Liu, L. Li, J. Luan, G. Saren, Int. J. Biochem. Cell. Biol. 39 (2007) 450-461.

[60] P.J. Sherratt, M.M. Manson, A.M. Thomson, E.A.M. Hissink, G.E. Neal, P.J. Van Bladeren, T. Green, J.D. Hayes, Biochem. J. 335 (1998) 619-630.

[61] D. Gisi, T. Leisinger, S. Vuilleumier, Arch. Toxicol. 73 (1999) 71-79.

[62] G.W. Gribble, Chemosphere 52 (2003) 289-297.

[63] N.K. Kubota, Y. Kobayashi, H. Iwamoto, Y. Fukazawa, Y. Uchio, Bull. Chem. Soc. Jpn. 79 (2006) 634-636.

[64] T. Rezanka, V.M. Dembitsky, Eur. J. Org. Chem. (2003) 309-316.

[65] B.J. Baker, P.J. Scheuer, J. of Nat. Prod. 57 (1994) 1346-1353.

[66] A.F. Rowley, C.L. Vogan, G.W. Taylor, A.S. Clare, J. Exp. Bio. 208 (2005) 3-14.

[67] M. Iwashima, K. Nara, Y. Nakamichi, K. Iguchi, Steroids 66 (2001) 25-32.

[68] L.B. Brattsten, in: G.A. Rosenthal, M.R. Berenbaum (Eds.), Herbivores: their interactions with secondary plant metabolites, Academic Press, New York, 1992, pp. 175-242.

[69] D. Schlenk, D.R. Buhler, Aquat. Toxicol. 13 (1988) 167-182.

[70] J.N. Dominguez, D.R. Adams, J. Famerich, Rev. Latinoam. Quim. 11 (1980) 56-58.

[71] S. Tsuchilda, K. Sato, CRC Crit. Rev. Biochem. Mol. Biol. 27 (1992) 337-384.

FIGURE LEGENDS

Figure 1. Purification of cytosolic GSTs from *C. gibbosum* digestive gland.

(A) The crude cytosol (2.5 mL) from one individual was applied to a size exclusion PD-10 desalting column (bed volume 8.3 mL, 5 cm height) containing Sephadex G-25 matrix (85 – 260 μm) equilibrated in 50 mM Tris buffer, 1mM EDTA, 1mM DTT at pH 6.0 (Buffer A). Fractions were collected in 1.2 mL aliquots by gravity at a flow rate of 1 mL min^{-1}. Fractions 3 – 7 were pooled for further purification. **(B)** The pooled fractions (approx. 6 mL) from the size exclusion column were applied to the GSH-agarose (sulfur to epoxide-activiated 4% cross-lined beaded agarose) affinity column (bed volume 0.5 mL, 0.8 x 4 cm i.d.) equilibrated in Buffer A. The affinity column was first rinsed with 6 mL of Buffer A and then 7 mL of Buffer B (Buffer A + 0.5 M NaCl) to rinse away non-specific proteins. GSTs were eluted with Buffer C (Buffer B + 50 mM GSH, pH 9.5) in 1 mL aliquots at a flow rate of 0.5 mL min^{-1}. Fractions 18 -20 were collected and pooled for further separation by HPLC. GST activity (bars) was measured as the increase in absorbance at 340 nm and protein concentration (●) was estimated by absorbance at 595 nm. Fractions from **(A)** and **(B)** were also examined for ethacrynic acid activity; however, no activity was observed (data not shown).

Figure 2. SDS-PAGE of affinity purified GSTs from *C. gibbosum* digestive gland

SDS-PAGE of affinity-purified extract from *C. gibbosum* digestive gland obtained from size exclusion and affinity chromatography. Lane M, Bio-Rad Kaleidoscope protein standards; lane Aff., affinity-purified extract. Bands visualized by silver staining.

Figure 3. HPLC spectrum of GST subunits

Representative HPLC spectrum of pooled affinity-purified extracts from *C. gibbosum* feeding on *B. asbestinum*. GST subunits were separated on a reverse phase VYDAC protein/peptide column (C18 μm 250 mm x 2.1 mm) with a flow rate of 0.5 mL min^{-1}. Mobile phase A consisted of acetonitrile/water/TFA (38:62:0.1, v/v) and mobile phase B consisted of acetonitrile/water/TFA (80:20:0.1, v/v). GST subunits were separated using a linear gradient from 0 to 40% B in 22 min, and 40 to 100% B in 37 min and visualized at 214 nm. Fourteen unique peaks were identified (not all visible in representative HPLC spectrum) and grouped into fractions B-H. Fraction A consisted of the injection peak (not shown); fraction B was identified as a theta-like GST; fractions C – H were identified as mu-class GSTs.

Figure 4. SDS-PAGE of HPLC separated GST subunits

Seven affinity extracts, pooled by diet, were separated by HPLC and fractions were hand-collected and subject to SDS-PAGE. For simplicity, only one diet's HPLC fractions are shown. Note the appearance of two bands for HPLC fraction F. Bands were visualized by silver staining. The figure is composed of five gels. The same molecular weight ladder and affinity-purified GST fraction were used as standards on each gel. MW: molecular weight standard; Aff: affinity-purified extract (see Figure 2); B-H: hand-collected HPLC fractions.

Figure 5. Multiple alignment of deduced amino acid sequences of CgGSTM1 and CgGSTM2 with homologous peptide fragment identification by LC-MS/MS

Comparison of translated mu-class GST cDNA sequences from *Cyphoma gibbosum* with selected vertebrate and invertebrate sequences available in GenBank. Amino acid residues identical to CgGSTM1 are indicated by a dot. Squares (■) and triangles (▲) indicate residues defining the G-site (GSH binding) and the H-site (substrate binding) respectively, as reported by NCBI Conserved Domain Database [36]. Shaded boxes represent peptide sequences with homology to tryptically digested peptides from HPLC fractions and affinity-purified bands identified by searching the NCBI protein nrDB (refer to Table 2). The alignment was constructed with the Clustal W algorithm and contains the following Genbank accession nos: *Cyphoma gibbosum* (EU008562, EU008563), *Haliotis discus discus* (ABF67506), *Crassostrea gigas* (CAD90167), *Homo sapiens* (1XW5B, 1XWKA), *Rattus norvegicus* (1B4PA), *Gallus gallus* (1GSUB), *Fasciola hepatica* (1FHE), *Schistosoma japonicum* (1UA5A).

Figure 6. Multiple alignment of theta-class GSTs from fish species with homologous peptide fragment identification by LC-MS/MS

Homology of trypically digested peptides from HPLC fraction B compared with published fish GSTs. Amino acid residues identical in all four fish species are indicated by a dot. Shaded boxes represent peptide sequences with homology to tryptically digested peptides from HPLC fraction B identified by searching the NCBI protein nrDB. The alignment was constructed using Clustal W and contains the following Genbank

accession nos: *Pagrus major* (BAD98442), *Micropterus salmoides* (AAQ91198), *Sparus aurata* (AAQ56182), and *Pleuronectes platessa* (CAA64495).

Figure 7. Phylogenetic tree depicting the relationship of *Cyphoma gibbosum* mu-class GSTs with other invertebrate and vertebrate GSTs.

Tree was constructed using MrBayes. Fruitfly, Housefly and Honeybee sigma-GSTs were defined as an outgroup. Values at branch points are posterior probabilities derived from 3e6 generations (see methods). Note the clustering of *Cyphoma* GSTMs (bold) with other molluscan GSTMs. The tree contains the following JGI and Genbank accession nos: *Apis mellifera* (XP_624682), *Boophilus microplus* (AAD15991), *Caenorhabditis elegans* (NP_499006, NP_503889), *Corbicula fluminea* (AAX20374), *Crassostrea gigas* (CAD90167), *Cyphoma gibbosum* (EU008563 and EU008562), *Danio rerio* (XP_690427), *Drosophila melanogaster* (NP_725653), *Fasciola hepatica* (P56598), *Haliotis discus discus* (ABF67506, ABF67507), *Haemaphysalis longicornis* (AAQ74441), *Homo sapiens* (NP_665683, AAV38750, NP_000840, NP_000841, NP_000843), *Lottia gigantea* (JGI scaffold_43000011), *Monodelphis domestica* (P46437), *Mytilus edulis* (AAS60226), *Mus musculus* (P13745, NP_038569, NP_034488), *Strongylocentrotus purpuratus* (XP_795664, XP_785553), *Sarcoptes scabiei* (AAO15607), *Tetraodon nigroviridis* (CAF91521), *Unio tumidus* (AAX20373), *Xenopus tropicalis* (NP_001004964).

Table 1. Primers used in present study

Gene	Primers	Direction	Sequences (5' to 3')	Purpose
GST Mu primer	GSTm_R10	Forward	AAY TTR TCR CTI GCC ATR TAI GC	Degenerate, 5'RACE
	RACE_F1	Reverse	AAT ACG ACT CAC TAT AGG	Adaptor RACE primer
CgGSTM1	GSTm_F13	Forward	CGC TCA AGC AGA CGT GAT TTG TTG ATA C	Full-length cloning
	AP1	Reverse	CCA TCC TAA TAC GAC TCA CTA TAG GGC	
CgGSTM2	GSTm_F15	Forward	GCT GCG TTT TCT TCT GAT AAC CAA GTC ACT C	Full-length cloning
	AP1	Reverse	CCA TCC TAA TAC GAC TCA CTA TAG GGC	

Table 2

Table 2. Activities and protein concentration from a representative *C. gibbosum* GST purification

Fraction	Volume (mL)	Protein (mg/mL)	GST Activity (U/mL)	Total Activity (U)	Specific Activity (U/mg)	Yield	Fold
Cytosol	1.8	7.65	4.36	7.85	0.57	100%	1
G-25 Column						41%	32
Fraction 3	1.0	0.03	0.15	0.15	5.06		
4	1.0	0.70	1.15	1.15	1.64		
5	1.0	0.58	1.00	1.00	1.72		
6	1.0	0.16	0.61	0.61	3.80		
7	1.0	0.05	0.30	0.30	6.00		
GSH-agarose						44%	220
Fraction 18	1.1	0.010	0.28	0.31	28.2		
19	1.1	0.036	2.74	2.97	75.0		
20	0.9	0.010	0.23	0.21	23.3		

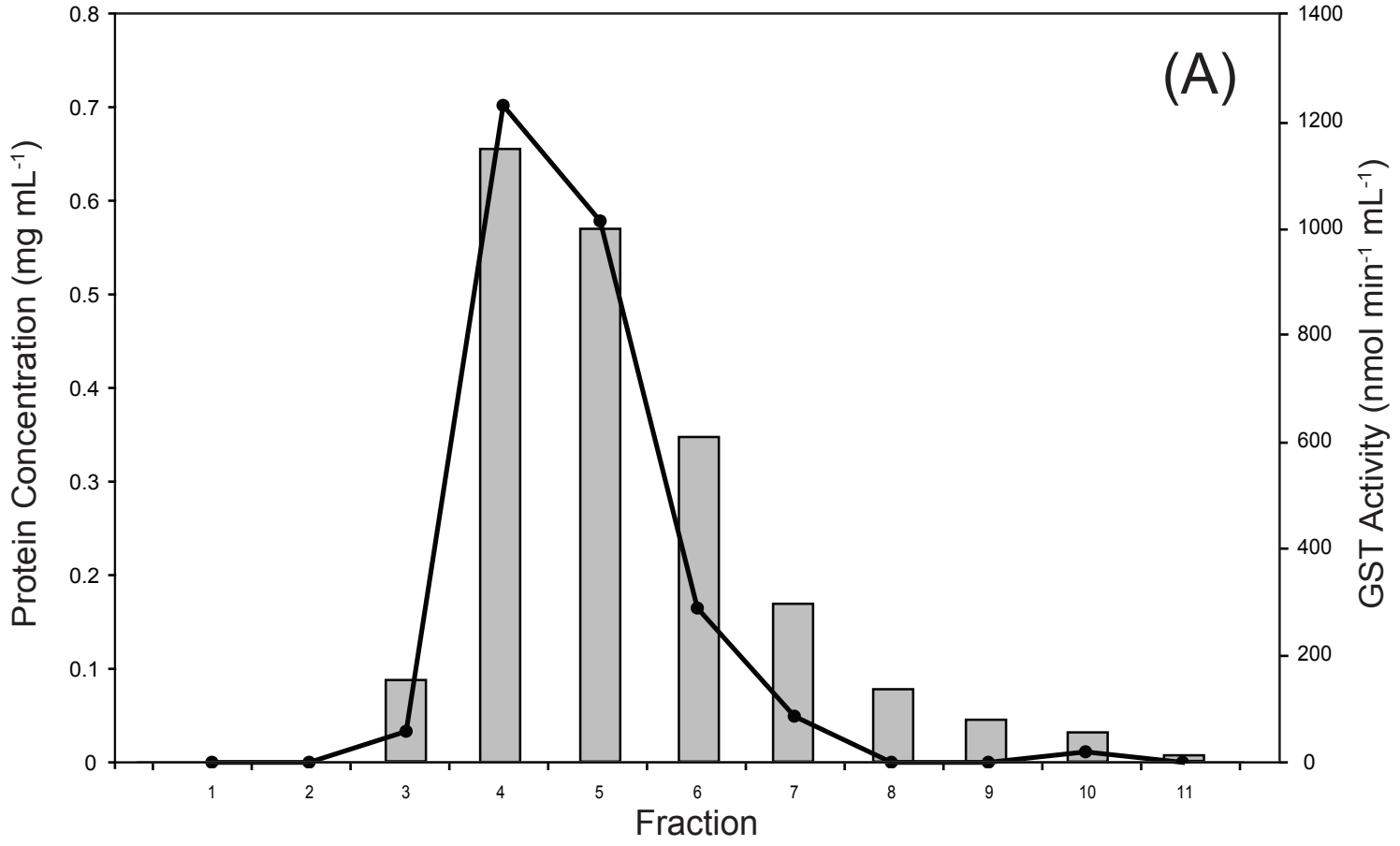
Fractions shown here correspond to the same fractions in Figure 1. Units are expressed as μmol/min.

Table 3. Homologous sequences identified from the partial tryptic digests of affinity-purified GSTs separated by HPLC

HPLC fraction or Affinity-protein band	Matched sequence	Class	Reference Species	NCBI Accession #
D, E, F_a, F_b, H, Affinity-protein band 1, 2	GLAQPIR	Mu	Abalone	ABF67506
E, H	GLAQPIRLLLK	Mu	Abalone	ABF67506
C, D, E, F_a, F_b, G, H, Affinity-protein band 1, 2, 3	ITQSNAILR†	Mu	Abalone	ABF67506
G	ITQSNAILRYIAR	Mu	Abalone	ABF67506
Affinity-protein band 3	KAAYFEALPAK	Mu	Abalone	ABF67506
G	SFLGDQQFFAGSK	Mu	Abalone	ABF67506
C, E	IMQPGSLDAFPTLLAFMGRIEALPAIK	Mu	Abalone	ABF67506
Affinity-protein band 1	IEALPAIKTFMSSAK	Mu	Abalone	ABF67506
F_a, Affinity-protein band 2	FIRRPINNK	Mu	Abalone	ABF67506
D	RPINNKSALFK	Mu	Abalone	ABF67506
F	GLGQPIR	Mu	Pacific oyster	CAD90167
F	AYMASDK†	Mu	Pacific oyster	CAD90167
B	SQEVMDMNPR†	Theta	Bass	AAQ91198
B	MFEGLTLNQK†	Theta	Bass	AAQ91198
B	VMIALLEEK†	Theta	Gilt Seabream	AAQ56182
B	VLNESYAACMYLESQFK†	Theta	Plaice	CAA64495
B	LWEGYLQK	Theta	Red Seabream, Bass	BAD98443, AAQ91198
B	GQLPAFK	Theta	Gilt Seabream, Bass	AAQ56182, BAD98443
			Red Seabream	AAQ91198
B	MADVIYYNWK†	Theta	Gilt Seabream, Bass	AAQ56182, BAD98443
			Red Seabream	AAQ91198
B	FGLCEER	Theta	Gilt Seabream, Bass	AAQ56182, BAD98443
			Red Seabream	AAQ91198

†Indicates sequences identical to those of *Cyphoma gibbosum* peptides determined by manual validation.
Affinity-purified protein bands are shown in Figure 2.
For the HPLC fraction F, F_a corresponds to the lighter band and F_b the darker band in Figure 4.

Figure 1 (revised)



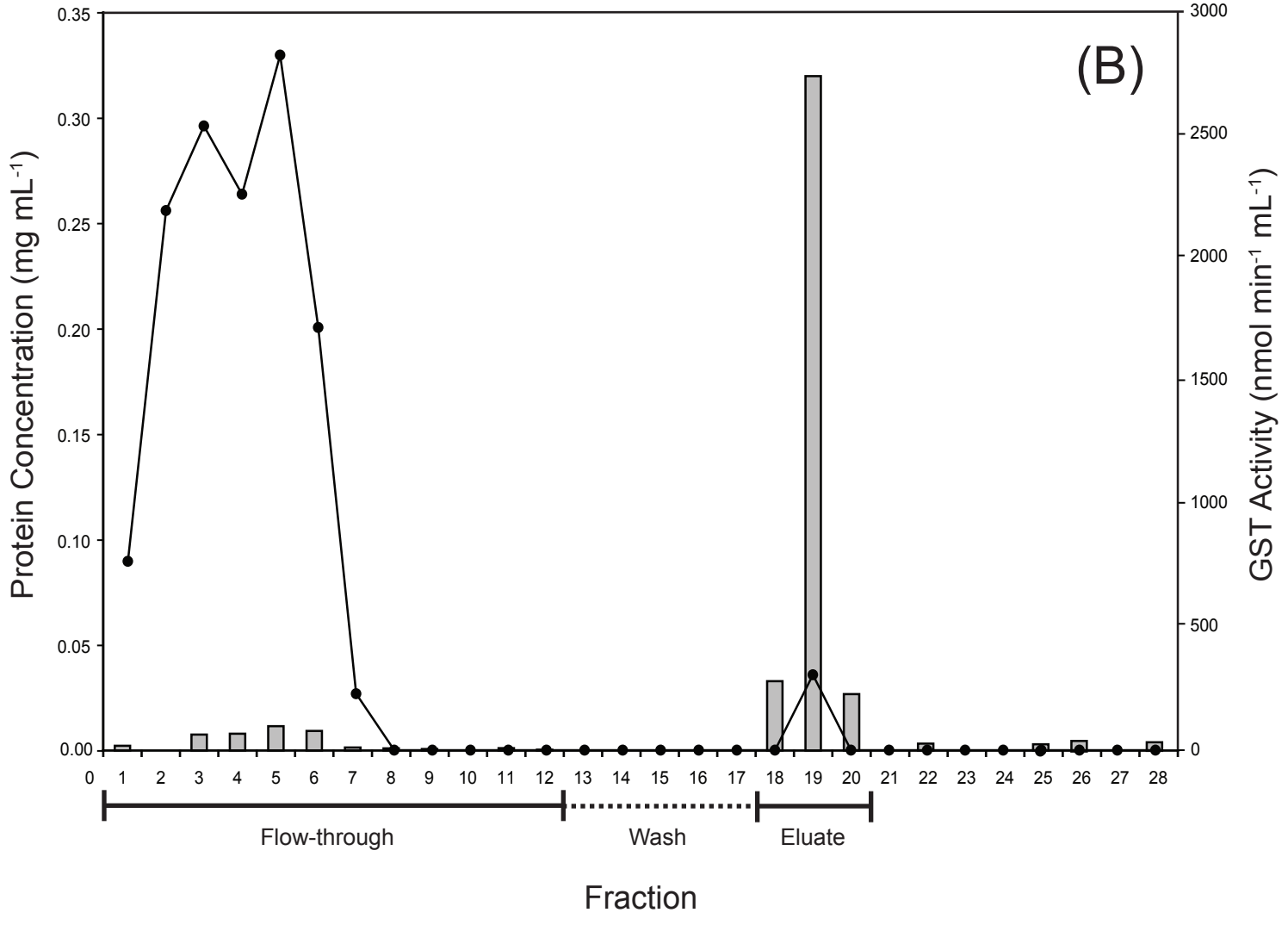


Figure 2

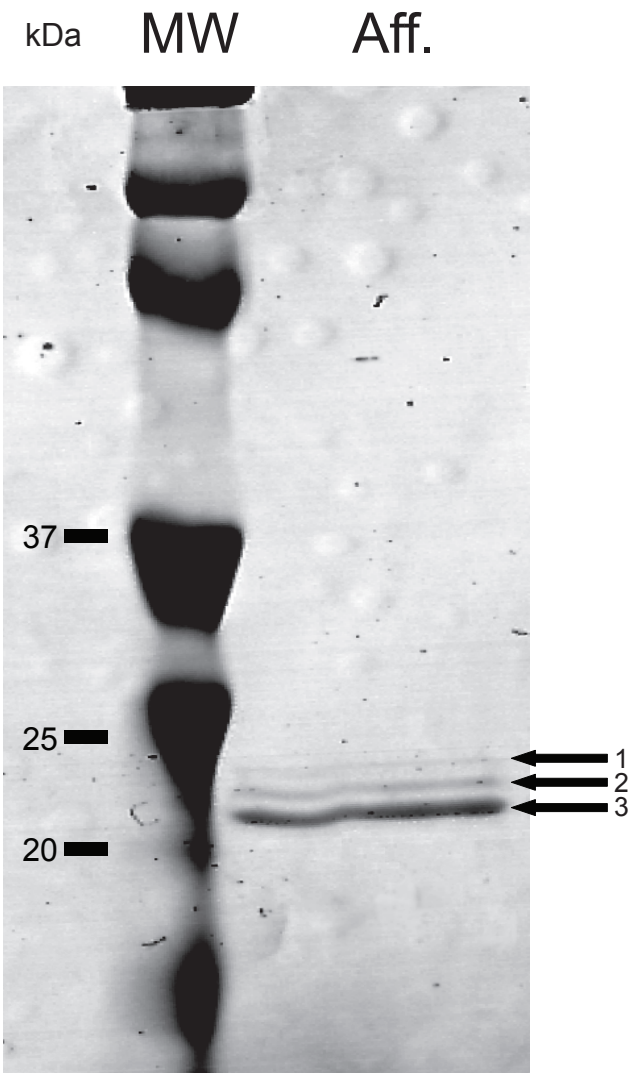


Figure 3

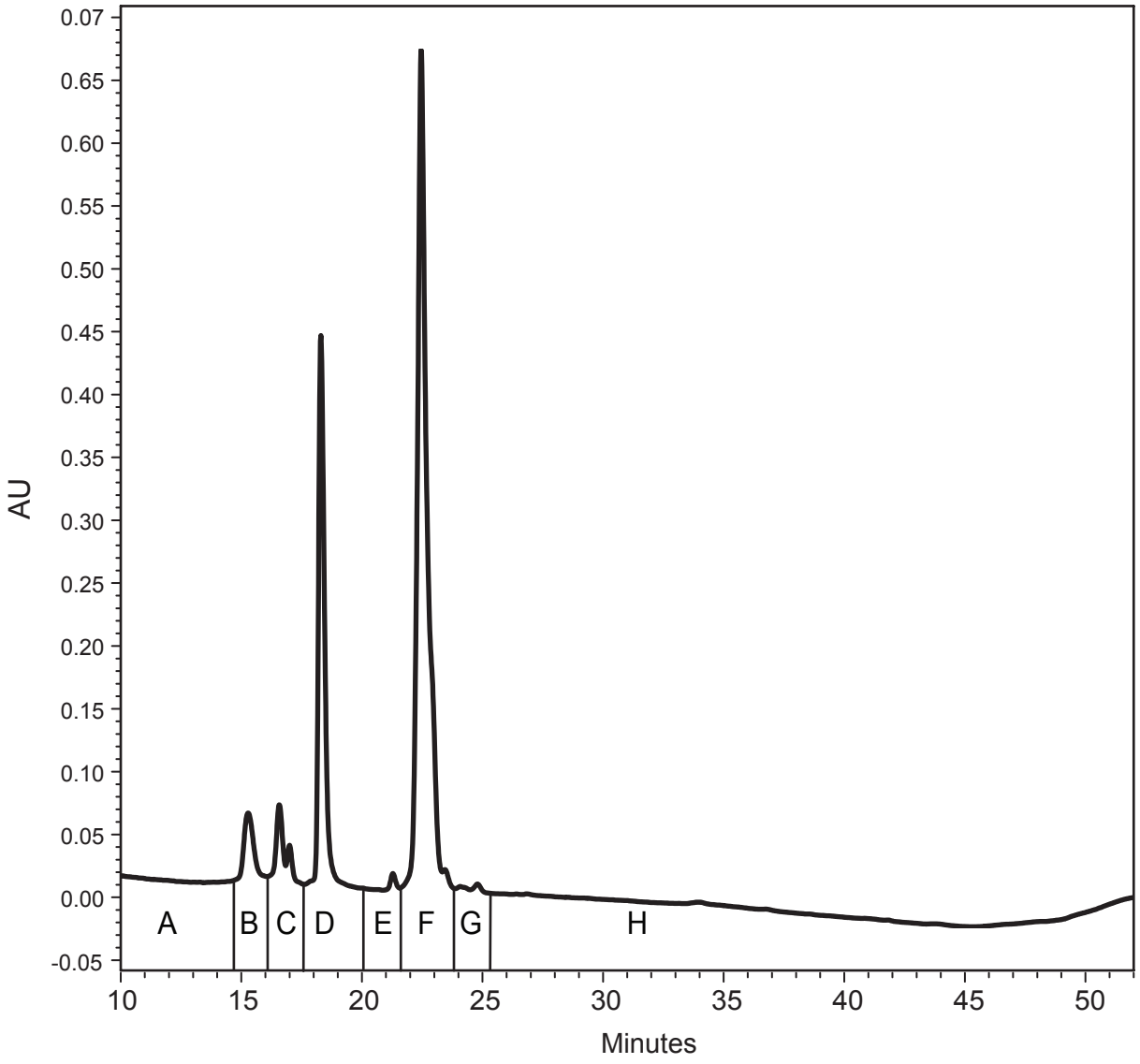

0.07
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
AU
A
B
C
D
E
F
G
H
10
15
20
25
30
35
40
45
50
Minutes

Figure 4

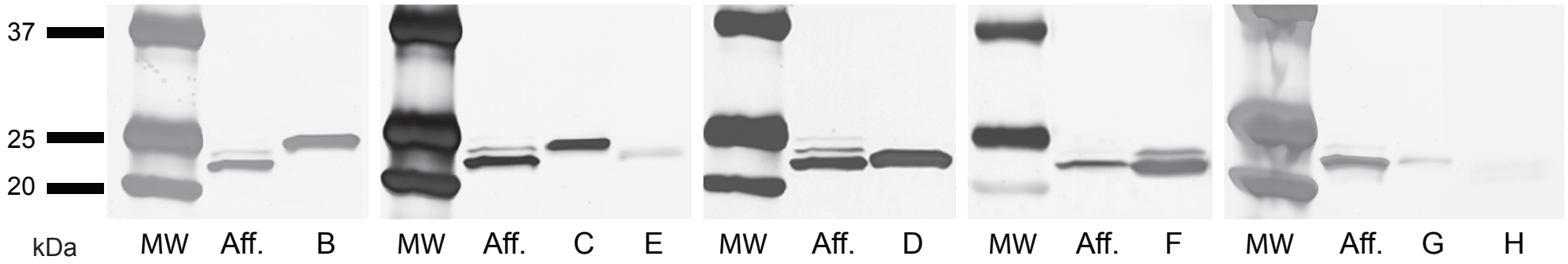


Figure 5

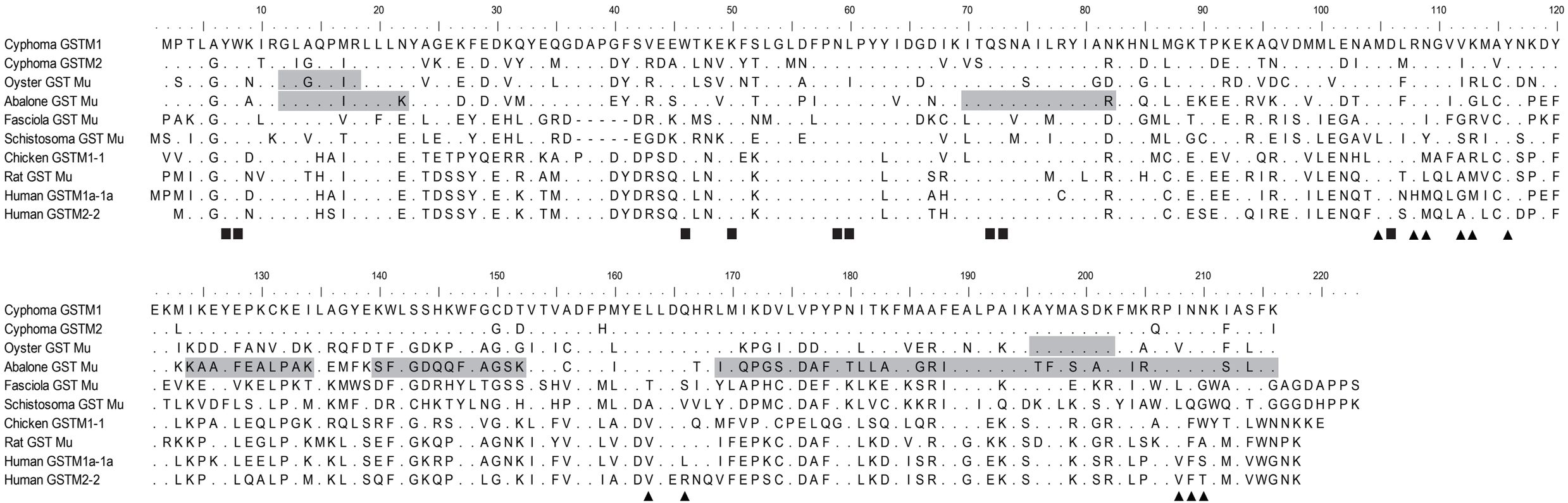

10 20 30 40 50 60 70 80 90 100 110 120
Cyphoma GSTM1 MPTLAYWKIRGLAQPMRLLLNYAGEKFEDKQYEQGDAPGFSVEEWTKEKFSLGLDFPNLPYYIDGDIKITQSNAILRYIANKHNLMGKTPKEKAQVDMMLENAMDLRNGVVKMAYNKDY
Cyphoma GSTM2 ....G...T..IG..I.....VK.E.D.VY..M....DY.RDA.LNV.YT..MN..........V.VS.........R..D.L...DE..TN.....DI...M....I..V.....
Oyster GST Mu .S..G..N....G..I.....V..E.D.V...L....DY.R...LSV.NT...A...I.....D.......S.....GD..G.L....RD.VDC...V.....F....IRLC.DN..
Abalone GST Mu ....G..A.....I....K....D.D.VM.......EY.R.S...V..T...PI.....V..N..............R..Q.L.EKEE.RVK..V..DT...F...I.GLC..PEF
Fasciola GST Mu PAK.G...L.....V..F.E.L..EY.EHL.GRD-----DR.K.MS...NM...L.........DKC.L...V..M....D..GML.T..E.R.RIS.IEGA.....I.FGRVC..PKF
Schistosoma GST Mu MS.I.G....K..V..T....E.LE..Y.EHL..RD-----EGDK.RNK..E...E.........V.L...M..I....D...ML.GC...R.EIS.LEGAVL.I.Y..SRI..S..F
Chicken GSTM1-1 VV..G..D.....HAI....E.TETPYQERR.KA.P..D.DPSD..N..EK.........L....V.L..........R...MC.E.EV..QR..VLENHL....MAFARLC.SP.F
Rat GST Mu PMI.G..NV...TH.I....E.TDSSY.E.R.AM....DYDRSQ.LN...K.........L...SR.......M..L.R..H.C.E.EE.RIR..VLENQ...T.LQLAMVC.SP.F
Human GSTM1a-1a MPMI.G..D.....HAI....E.TDSSY.E.K.TM....DYDRSQ.LN...K.........L...AH.........C...R....C.E.EE..IR..ILENQT..NHMQLGMIC..PEF
Human GSTM2-2 M..G..N.....HSI....E.TDSSY.E.K.TM....DYDRSQ.LN...K.........L...TH..........R....C.ESE..QIRE.ILENQF..S.MQLA.LC.DP.F
130 140 150 160 170 180 190 200 210 220
Cyphoma GSTM1 EKMIKEYEPKCKEILAGYEKWLSSHKWFGCDTVTVADFPMYELLDQHRLMIKDVLVPYPNITKFMAAFEALPAIKAYMASDKFMKRPINNKIASFK
Cyphoma GSTM2 ..L.....G.D.....H.....Q.....F...I
Oyster GST Mu ..IKDD.FANV.DK.RQFDTF.GDKP..AG.GI.IC...L.....KPGI.DD...L...VER..N..K.........A..V...F.L..
Abalone GST Mu ..KKAA.FEALPAK.EMFKSF.GDQQF.AGSK...C...I.....T.I.QPGS.DAF.TLLA..GRI.....TF.S.A..IR.....S.L..
Fasciola GST Mu .EVKE..VKELPKT.KMWSDF.GDRHYLTGSS.SHV..ML..T..SI.YLAPHC.DEF.KLKE.KSRI....K.....E.KR.I.W.L.GWA...GAGDAPPS
Schistosoma GST Mu .TLKVDFLS.LP.M.KMF.DR.CHKTYLNG.H..HP..ML.DA..VVLY.DPMC.DAF.KLVC.KKRI..I.Q.DK.LK.S.YIAW.LQGWQ.T.GGGDHPPK
Chicken GSTM1-1 ..LKPA.LEQLPGK.RQLSRF.G.RS..VG.KL.FV..LA.DV...Q.MFVP.CPELQG.LSQ.LQR....EK.S...R.GR...A..FWYT.LWNNKKE
Rat GST Mu .RKKP..LEGLP.KMKL.SEF.GKQP..AGNKI.YV..LV.DV.....IFEPKC.DAF..LKD.V.R..G.KK.SD..K.GR.LSK..FA.M.FWNPK
Human GSTM1a-1a ..LKPK.LEELP.K.KL.SEF.GKRP..AGNKI.FV..LV.DV..L..IFEPKC.DAF..LKD.ISR..G.EK.S...K.SR.LP..VFS.M.VWGNK
Human GSTM2-2 ..LKP..LQALP.M.KL.SQF.GKQP..LG.KI.FV..IA.DV.ERNQVFEPSC.DAF..LKD.ISR..G.EK.S...K.SR.LP..VFT.M.VWGNK

Figure 6

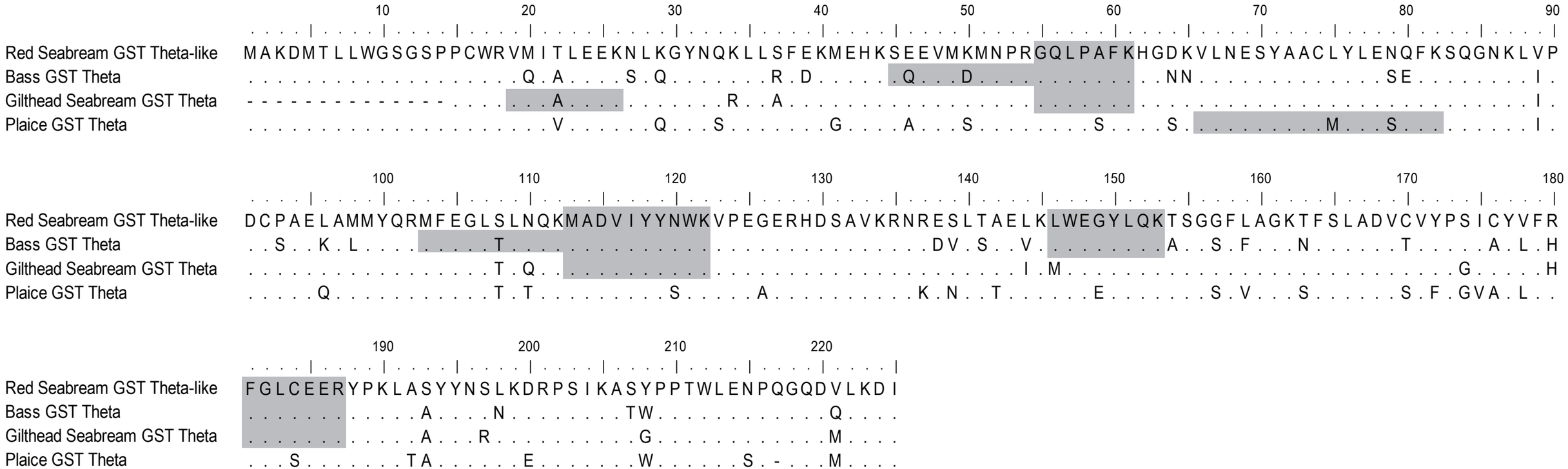
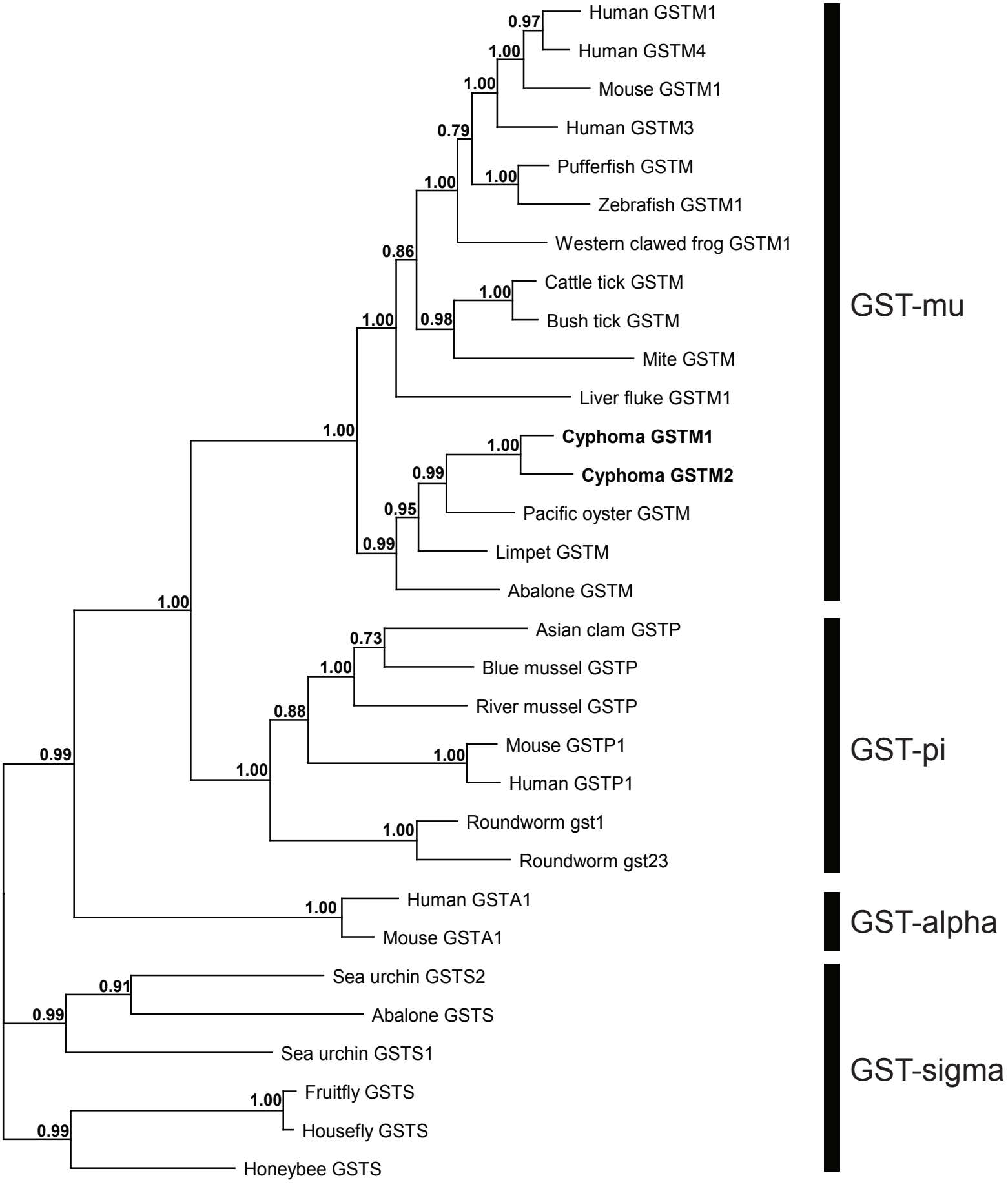

10 20 30 40 50 60 70 80 90
Red Seabream GST Theta-like MAKDMTLLWGSGSPPCWRVMITLEEKNLKGYNQKLLSFEKMEHKSEEVMKMNPRGQLPAFKHGDKVLNESYAACLYLENQFKSQGNKLVP
Bass GST Theta ..........Q.A....S.Q.....R.D.....Q...D..........NN..........SE.......I.
Gilthead Seabream GST Theta ----------....A.....R..A..................I.
Plaice GST Theta ..........V.....Q...S.....G....A...S.....S....S.........M...S.....I.
100 110 120 130 140 150 160 170 180
Red Seabream GST Theta-like DCPAELAMMYQRMFEGLSLNQKMADVIYYNWKVPEGERHDSAVKRNRESLTAELKLWEGYLQKTSGGFLAGKTFSLADVCVYPSICYVFR
Bass GST Theta ..S..K.L.....T..............DV.S..V.........A..S.F...N.....T.....A.L.H
Gilthead Seabream GST Theta ..........T.Q..........I.M..........G.....H
Plaice GST Theta .....Q..........T.T.........S.....A..........K.N..T.....E.....S.V...S.....S.F.GVA.L..
190 200 210 220
Red Seabream GST Theta-like FGLCEERYPKLASYYNSLKDRPSIKASYPPTWLENPQGQDVLKDI
Bass GST Theta .........A....N.........TW..........Q....
Gilthead Seabream GST Theta .........A...R..........G..........M....
Plaice GST Theta ...S.......TA......E.......W......S.-...M....

Figure 7


Human GSTM1
Human GSTM4
Mouse GSTM1
Human GSTM3
Pufferfish GSTM
Zebrafish GSTM1
Western clawed frog GSTM1
Cattle tick GSTM
Bush tick GSTM
Mite GSTM
Liver fluke GSTM1
Cyphoma GSTM1
Cyphoma GSTM2
Pacific oyster GSTM
Limpet GSTM
Abalone GSTM
GST-mu
Asian clam GSTP
Blue mussel GSTP
River mussel GSTP
Mouse GSTP1
Human GSTP1
Roundworm gst1
Roundworm gst23
GST-pi
Human GSTA1
Mouse GSTA1
GST-alpha
Sea urchin GSTS2
Abalone GSTS
Sea urchin GSTS1
Fruitfly GSTS
Housefly GSTS
Honeybee GSTS
GST-sigma
0.1