**A THESIS**

**FOR THE DEGREE OF DOCTOR OF PHYLOSOPHY**

# MOLECULAR CHARACTERIZATION AND EXPRESSION ANALYSIS OF POLYSACCHARIDE DEGRADING ENZYMES, CALCIUM REGULATORY AND PATTERN RECOGNITION PROTEIN FROM DISK ABALONE, *Haliotis discus discus*

**Chamilani Nikapitiya**

Department of Aquatic Life Medicine

GRADUATE SCHOOL

CHEJU NATIONAL UNIVERSITY

2008-08

# MOLECULAR CHARACTERIZATION AND EXPRESSION ANALYSIS OF POLYSACCHARIDE DEGRADING ENZYMES, CALCIUM REGULATORY AND PATTERN RECOGNITION PROTEIN FROM DISK ABALONE, *Haliotis discus discus*

**Chamilani Nikapitiya**

**(Supervised by Professor Jehee Lee)**

A thesis submitted in partial fulfillment of the requirement for the degree of

**DOCTOR OF PHILOSOPHY**

**2008-08**

The thesis has been examined and approved by

.....

**Thesis Director, Choon-Bok Song, Professor of Biotechnology**

.....

**Moon-Soo Heo, Professor of Biotechnology**

.....

**In-Kyu Yeo, Professor of Biotechnology**

.....

**Jeong Joon Bum, Professor of Biotechnology**

.....

**Jehee Lee, Professor of Biotechnology**

.....

**Date**

**Department of Aquatic Life Medicine**

**GRADUATE SCHOOL**

**CHEJU NATIONAL UNIVERSITY**

**REPUBLIC OF KOREA**

**CONTENTS**

## 요약문

전복은 전세계적으로 많이 양식되는 고부가가치의 패류로 알려져 있다. 본 연구에서는 전복의 생리학적, 생물학적 기능들을 이해하기 위해 다당류분해 (소화), 칼슘 조절 및 면역에 관련된 유전자들에 초점을 맞추었다.

본 연구의 첫번째로 까막전복의 cDNA library 로부터 6 가지 다당류 분해 효소들(PDEs)에 대해 EST 클론들을 선택하여 확인하고 기능적 특성을 분석하였다. 다당류 분해와 관련된 그 유전자들은 endo β-1,4-glucanase (HdEndg), α-amylase (HdAmy), alginate lyase (HdAlgl), β-mannanase (HdMann), laminarinase (HdLmn)와 arylsulfatase (HdArys)이다.

까막전복의 PDEs 염기서열들은 기존에 알려진 다당류 분해 효소의 서열과 비교 분석되어졌고, motif 들이 확인되어졌다. 그 염기서열들을 토대로 pMAL fusion protein purification system 을 이용하여 재조합 단백질을 생산하였고, 재조합된 PDEs 의 기능 분석 및 최적온도, 최적 pH, 열안정성과 같은 특성을 분석하였다.

HdEndg 을 코딩하는 염기서열은 1824 bp (608 aa)로 분석되었다. HdEndg 효소는 40 ºC 에서 최적 온도를 나타냈고, pH 4.5 에서 최적 pH 를 나타냈으며 이들 조건하에서 0.3 U/mg 의 특이활성을 나타냈다. HdAmy 는 1536 bp (511 aa)의 코딩 염기 서열을 가지고 있고 50 ºC 에서 최적 온도를 나타냈으며, 최적 pH 는 7.5 에서 나타났다. Alginate lyase 는 갈조류의 주된 구조 성분인 alginate 를 분해하는 효소로, HdAlgl 을 코딩하는 염기서열은 897 bp (298 aa)로 분석되었다. 그 정제된 재조합 단백질은 최적 온도인 40 ºC 에서 sodium alginate 에 대해 2 U/mg 의 활성을 나타냈다. HdMann 를 코딩하는 염기서열은 1125 bp (375 aa)로 분석되었고, 처음 18 개의 아미노산 잔기는 세포외로 분비시키는 신호서열로 확인되었다. 유전자 데이터베이스 (NCBI)로부터 아미노산 서열을 비교 분석한 결과 *Mitilus edulis* 와 *H. discus hannai* mananase 에 각각 49%, 48%의 유사성을 나타내었다. Sulfatase 는 glycosaminoglycans, proteoglycans, steroids, glycolipids 와 같은 다양한 기질로부터 sulfate ester bond 를 가수분해 한다. 그 서열 특성을 분석한 결과 모티브 및 기능적 아미노산 잔기들이 sulfatase family 의 주된 특성을 나타내는 것으로 확인 되었다. HdArys 를 코딩하는 염기서열은 1449 bp (481 aa)로 분석되었고, HdArys 의 아미노산 구성 중 24-433

aa 부분에 sulfatase domain 을 포함하고 있었다. 그리고 $^{72}$CSPSR$^{76}$ 에는 sulfatase signature motif 가 보여졌고, 기질과 결합, 활성, 금속 조절, 효소 안정화 부위 등이 포함되어 있었다. HdArys 는 *Helix pomatia* sulfatase I precursor 에 가장 높은 유사성(45%)를 나타냈고 *Rattus norvegicus, Bos Taurus, Homo sapiens* arylsulfatase B 와도 41%의 유사성을 보였다.

까막전복에서 PDEs의 in vivo 발현 분석을 위해, 초기 8주 동안 먹이를 공급하지 않았다가 그 후 28일 동안 먹이를 공급하였다. 아가미, 맨틀, 근육, 소화관, 간췌장 조직으로부터 mRNA 발현 분석을 위해 Semi quantitative RT-PCR를 수행 하였다. 그 결과 모든 PDEs들이 소화관에서보다 간췌장에서 강하게 발현되었음을 확인하였으며 이것은 소화효소들이 간췌장에 많이 존재하고 있기 때문으로 생각된다. 또한 발현 수준이 서로 다른 유전자들에서 다양하게 보여졌다. 먹이를 공급하지 않은 기간동안 체중이 최대 18.6%까지 소실되었다. 먹이를 공급한 8주 동안은 모든 다당류 분해 효소들이 대조구에 비해 의미있는 감소를 보였다. 먹이를 주고나서 2주 동안은 mRNA 전사 수준이 점차적으로 증가하는 것을 확인하였고 먹이를 준 8주 동안 전복들은 살아남을 수 있었다.

본 연구의 두 번째 파트에서는 칼슘 조절에 관여하는 regucalcin 유전자에 대한 분석과 mRNA 발현 수준을 분석하였다. 칼슘 이온은 많은 세포질과 생리적 기능들에서 근육 수축, 신경 세포 신호 전달 활성화, 세포 분열 및 세포 죽음 등에 2 차 전달자로서 관여를 한다. 전복의 regucalcin (HdReg)를 코딩하는 염기서열은 918 bp (305 aa)로 확인되었다. 그 HdReg 아미노산 서열은 기존의 $Ca^{2+}$ binding domain 처럼 EF-hand motif 를 포함하지 않았기에, 새로운 클래스의 $Ca^{2+}$ binding protein 으로 제안한다. 유전자 데이터베이스로부터 HdReg 의 아미노산 서열을 비교 분석한 결과, 닭과 제브라피쉬의 regucalcin 과 45%의 유사성을 나타냈으며, rat 과 mouse 의 regucalcin 과는 44%의 유사성을 보였다. In vivo 에서 아가미, 맨틀, 소화관, 근육으로부터 regucalcin 의 조직 특이적 발현 분석을 위해 semi quantitative RT-PCR 을 수행하였다. 0.5 mg $CaCl_2$/g 를 전복의 근육 안쪽에 주사하였고 HdReg 의 칼슘 조절 역할을 진단하였다. HdReg mRNA 는 아가미, 맨틀, 소화관, 근육 모두에서 발현 되어졌다.

Semi-quantitative RT-PCR 결과 $CaCl_2$ 를 근육 안쪽에 투입한 그룹에서 30 분 후 근육에서 regucalcin 의 mRNA 가 의미있게 유도되어졌고, 1 시간 후 발현 수준이 최고치에 도달했다. 2 시간 째에는 발현 레벨이 감소하는 경향을 보였다. 전체적으로 볼 때 칼슘을 주입 후 근육에서 regucalcin 의 발현 레벨이 특이적으로 증가함을 나타냈다.

본 연구의 마지막 파트는 면역에 관계된 pattern recognition 유전자에 초점을 맞추었고, 전복에서 다른 면역 조절인자들에 의한 면역 첼린져 후 면역관련 유전자들의 mRNA 발현 레벨을 확인하였다. pattern recognition protein은 미생물들의 표면상에 공통적인 epitope들의 인지에 의한 초기 면역 반응에서 중요한 역할을 하고 비자기 병원성 분자 패턴들이 붙었을 때 숙주 면역 반응을 시작한다.

까막전복의 whole body cDNA library 로부터 Pattern recognition protein (PRP) 유전자가 선택되어졌고 분석 결과 HdPRP 를 코딩하는 염기서열은 20 aa 의 아미노산 서열을 포함하는 1260 bp (420 aa)로 분석됐다. 그 아미노산 서열은 민물 달팽이인 *Biomphalaria glabrata* 의 glucan recognition protein (BGRP) 에 50%의 유사성을 나타내었다. Potential polysaccharide binding, putative cell adhesion 그리고 glucanase motifs 과 같은 motif 분석 결과 기존의 무척추동물 PRP motif 에 조금 변화된 형태의 motif 가 발견되었다.

In vivo에서 PRP의 mRNA 발현 수준을 확인하기 위해, 전복에 *Vibrio alginolyticus* (150 μl/animal, $OD_{600}$=1.0), lipopolysaccharide (4 μg/g animal) 와 β-1,3-glucan (6 μg/g animal)를 근육 안쪽으로 주입하였고, RT-PCR 결과 HdPRP가 아가미, 맨틀, 소화관, 간췌장, 혈구에서 발현됨을 확인하였다. *V. alginolyticus* 를 주입하였을 때는 12시간 후 아가미에서 발현이 증가하였고 48시간까지 증가함을 나타냈다. 이들 데이터는 전복의 PRP가 모든 선택된 조직들에서 지속적으로 발현되고 전복 방어 메커니즘에도 중요한 역할을 하는 것으로 판단된다.

결론적으로 본 연구에서 보여지듯이 까막전복에서 기능적으로 활성을 갖는 다양한 다당류 분해 효소들은 주로 간췌장에서 발현되었다. 또한 regucalcin이 체내에서 충분한 칼슘 조절을 위해 중요한 것으로 보여졌다. 그리고 전복에서 PRP의 발현은 다른 미생물들로부터 자신을 보호하기 위한 초기 면역 반응을 주로 도와줄 수 있는 것으로 보여진다.

## Summary

Abalone is an important and highly priced species in world wide aquaculture. Inorder to understand the physiological and biological functions of the abalone, we have forcused on polysaccharide degrading (digestive), calcium regulatory and immune response genes in this study.

In the first part of this study, we identified and functionally characterized the 6 EST clones of polysaccharide degrading enzymes denoted as PDEs from disk abalone *Haliotis discus discus* cDNA library. The genes that responsible for polysaccharide degrading are endo β−1,4-glucanase (HdEndg), α-amylase (HdAmy), alginate lyase (HdAlgl), β−mannanase (HdMann), laminarinase (HdLmn) and arylsulfatase (HdArys). Sequence characterization of the abalone PDEs showed that those genes were matched with other known poly saccharide degrading enzymes and respective motifs. The recombinant enzymes of selected PDEs were purified using pMAL fusion protein purification system. Recombinat PDEs were characterized for their specific function and other factors such as optimal temperature, pH and thermal stability. The HdEndg comprises 1824 bp ORF coding for 608 amino acids. The purified HdEndg enzyme showed the 0.3 U/mg specific activity at 40 $^{o}$C optimal temperature and pH 4.5. HdAmy has 1536 bp ORF coding for 511 amino acids. The optimum temperature and pH of HdAmy were 50 $^{o}$C, 7.5, respectively. Alginate lyase is responsible for the degradation of alginate, which is considered as the main structural component of brown seaweed. HdAlgl consists of 897 bp ORF codes for 298 amino acids. The purified enzyme showed 2 Unit/mg activity towards sodium alginate at 40 $^{o}$C optimum temperature. HdMann cDNA contains an ORF of 1125 bp that encodes 375 amino acids, of which the first 18 residues comprise the secretion signal peptide. According to searches for the NCBI, GenBank databases, the amino-acid sequence of HdMann was found to show 49% and 48% identity values to *Mitilus edulis and H. discus hannai* mananase, respectively.

Sulfatases are responsible for hydrolyzing the sulphate ester bonds from variety of substrates such as glycosaminoglycans, proteoglycans, steroids and glycolipids. The sequence characterization results revealed that HdArys encodes the main characteristic motifs and functional amino acid residues of the sulfatase family. HdArys cDNA contains 1449 bp ORF that encodes 481 amino acids. The HdArys amino acid sequence consists of a sulfatase domain (24-433 aa) and characteristic sulfatase signature motif was present at $^{72}$CSPSR$^{76}$. Also, functionally important residues involved in substrate binding, activation, metal coordination and enzyme stabilization were observed in the sequence. The HdArys showed the highest level amino acid sequence identity (45%) with *Helix pomatia* sulfatase I precursor while it shared 41% identity with *Rattus norvegicus, Bos taurus* and *Homo sapiens* arylsulfatase B.

For the *in vivo* expression analysis of PDEs in disk abalone, the animals were starved for 8 weeks and after that re-fed the animals continuously for 28 days. Semi quantitative RT-PCR was carried out to determine the mRNA expression analysis in the digestive tract and hepatopancreas tissues. All the PDEs were strongly expressed in hepatopancreas than digestive tract, suggesting that key organ which encountered digestive enzyme is hepatopancreas. Also, expression level was varies in different genes. Following the starvation period, maximum weight loss was 18.6%. During 8 weeks of starvation, all polysaccharide degrading enzymes were significantly ($p<0.05$) decreased compare to fed animals (control). During 2 weeks of re-feeding, gradual increase of mRNA transcription was observed, suggesting that abalones can survive for 8 week of starvation.

In the second part of study we have characterized the calcium regulatory regucalcin gene and its mRNA expression analysis. Calcium ion acts as a second messenger, in the regulation of many cellular and physiological functions such as muscle contraction, neuronal activation, cell differentiation and cell death in many organisms. Abalone regucalcin (HdReg) ORF consists of 918 nucleotides encoding 305 amino acids. The HdReg amino acid sequence

did not contain the EF-hand motif as a $Ca^{2+}$ binding domain, suggesting a novel class of $Ca^{2+}$ binding protein. According to searches for the NCBI, GenBank databases, it showed 45% identity to chicken and zebrafish, and 44% to rat and mouse regucalcin in deduced amino acid level. *In vivo* regucalcin tissue specific expression was investigated in gill, mantle, digestive tract, abductor muscle by semi quantitative RT-PCR. Abalones were injected with 0.5 mg $CaCl_2$/g of animal intramuscularly and determined the Ca regulatory role of HdReg. HdReg mRNA was expressed in gill, mantle, digestive tract, and abductor muscle. Semi-quantitative RT-PCR results showed that an intramuscular administration of $CaCl_2$ could significantly induce regucalcin mRNA in abductor muscle after 30 min of administration and reached maximum after 1 h. Subsequently, the expression level was decreased after 2 h. This indicates that the expression of regucalcin mRNA is constitutive, and specifically up regulated in abalone abductor muscle by $Ca^{2+}$ administration.

In the third part of this study was focused on immune related pattern recognition gene of disk abalone with respect to mRNA expression analysis after immune challenge by diffent immune modulators. Pattern recognition molecules play an important role in innate immunity by recognizing common epitopes on microorganisms surface and initiate the host immune response, when binding to non self pathogen associated molecular patterns (PAMPs). Pattern recognition protein (PRP) was isolated from a disk abalone, *H. discus discus,* whole body cDNA library. The HdPRP ORF consists of 1260 bp encodes 420 amino acids with 20 aa of a signal peptide sequence. The deduced aa sequence of HdPRP showed high identity to β−glucan recognition protein (BGRP) of the fresh water snail *Biomphalaria glabrata* (50%). Characteristic motifs such as potential polysaccharide binding, putative cell adhesion, and glucanase motifs were found in HdPRP with slight modifications to other invertebrate PRP motifs. To evaluate *in vivo* PRP mRNA expression, abalones were intra muscularly injected with *Vibrio alginolyticus* (150 μl/animal, $OD_{600}$=1.0), lipopolysaccharide (4 μg/g animal) and β-1,3-glucan (6 μg/g animal). RT-PCR results showed that HdPRP was constitutively

expressed in gill, mantle, digestive tract, hepatopancreas and hemocytes. Animals injected with *V. alginolyticus* showed that the expression was increased initially within 12 h post injection in gill and it was increased until 48 h. These data indicated that the abalone PRP was constitutively expressed in all the selected tissues and it acts as an acute inducible protein that could play an important role in abalone immune defense mechanism.

Finally, in this study demonstrated that disk abalone has functionally active different polysaccharide degrading enzymes, which are expressed mainly in hepatopancreas. Also, abalone has regucalcin that is important for efficient $Ca^{2+}$ balance in the body. Also, constitutive expression of PRP in abalone could support to maintain innate immune defense to protect from different microorganisms.

## LIST OF FIGURES

## LIST OF TABLES

# PART 1

# Molecular Characterization and Expression Analysis of Polysaccharide Degrading Enzymes from Disk Abalone *Haliotis discus discus*

## ABSTRACT

Disk abalone is belongs to marine herbivorous mollusks, feeds mainly on macro algae. Their digestive system is developed to ingest and digest those foods and many carbohydrate digestive enzymes are use to break down the structural polysaccharides of algae. To increase the understanding of polysaccharides break down in abalone, few of the disk abalone polysaccharide degrading enzymes related genes were studied in their molecular level. Also, attempted to know how tissue expression levels are affected by one of the exogenous factor, food and their co-ordination at the time of food scarcity and food availability.

In the present study, we identified EST clones of polysaccharide degrading enzymes from disk abalone *Haliotis discus discus* cDNA library. Total 6 cDNAs which responsible for polysaccharide degrading (Endo β−1,4-glucanase, HdEndg; α-amylase, HdAmy; alginate lyase, HdAlgl; β-mannanase, HdMann; laminarinase, HdLmn and arylsulfatase, HdArys were sequenced to determine full length sequence. The resulting full length sequences were compared with other known sequences available in National Center for Biotechnology Information (NCBI) data base. After having the full length coding sequence, sequence characterization, ClustalW pairwise and multiple analysis and phylogenetic analysis were performed in order to establish the relationship between disk abalone and known respective genes. All the genes showed their characteristic motifs, catalytic sites, substrate binding sites and conserved regions with other known species of respective genes.

Coding sequences of HdEndg, HdAmy, HdAlgl and HdArys were cloned into pMAL-c2X plasmid after amplified the coding sequences with designed primers. Successful clones were transformed into *E. coli* K12TB1 and the respective proteins were expressed by inducing with IPTG. The expected sizes of the proteins were as 107.5, 96.5, 75.5 and 94.5 kDa for HdEndg, HdAmy, HdAlgl and HdArys, respectively, together with the fusion protein where maltose binding protein (MBP) contributes to 42.5 kDa molecular weight. Therefore,

recombinant HdEndg, HdAmy, HdAlgl and HdArys molecular masses are 65, 54, 33 and 52 kDa, respectively, as predicted. Purified recombinant proteins were assayed for optimal temperature, pH and thermal stability using respective substrates for the protein. Different PDEs showed different temperature and pH optima and thermostabilities. Reverse transcription-polymerase chain reaction (RT-PCR) was conducted to analyze the mRNA expression of disk abalone polysaccharide digestible enzymes in digestive tract and hepatopancreas tissues. A semi-quantitative RT-PCR assay was conducted to estimate mRNA expression of digestive enzymes in hepatopancreas to find out the effect of food availability on the gene expression.

All the genes were strongly expressed in hepatopancreas than digestive tract, suggesting that key organ which encountered digestive enzyme is hepatopancreas. Also different expression levels were observed in different genes. Further, abalones were starved for 8 weeks and then re-fed for 2 weeks. Semi quantitative RT-PCR was carried out to evaluate starvation and re-fed effect on their respective gene transcription in hepatopancreas. Following the starvation period, maximum weight loss was 18.6%. During 8 weeks of starvation, significant decrease of all polysaccharide degrading enzymes gene expression was observed compare to fed animals (control). During 2 weeks of re-feeding, gradual increase of mRNA transcription was observed, suggesting that abalones can survive for more than 8 week of starvation.

# INTRODUCTION

Abalone is herbivorous marine gastropod, valued as a high commercial and nutritious palatable food. They are widely distributed and cultured in Australia, South Africa, United states, Mexico as well as Asian countries such as China, Japan, Korea, etc. (Gordon and Cook, 2001). They are typically feed on micro (7-8 mm newly settled juveniles) and macro algae (size reached >10 mm) as its main source of food. Polysaccharides are the major constituent in algae and sulfated polysaccharides are present as sulfate fucose (fucoidens) and as sulfate galactans (carrageenan and agar) in marine algae. Widely present polysaccharides are laminaran, fucoidan and alginate in brown algae and carragenan and agar in red algae (Melo et al., 2002). Apart from that, cellulose and hemicellulose are the main structural polysaccharides present in algae as well as other higher plants, where as starch is the main storage form of polysaccharides for later use as a food or energy source. Polysaccharides contain similar or different monosaccharide units bonded together by linear or branched glycosidic linkages to form long chains.

Abalones digestive system anatomically and biochemically adapted to feed on alage. Their digestive system uses enzymes to break down these structural polysaccharides of algae to oligo and monosacchride units for enhance absorption of nutrients to their body. Therefore, PDEs play an important role in abalones for efficient digestion of algae. PDEs such as alginate lyase, carboxymethylcellulase (CMCase), laminarinase, agarase, carragenase in the hepatopancreas of abalone, *Haliotis midae* have been reported previously (Erasmus et al., 1997).

**Cellulase**

Cellulose is the most abundant homo polysaccharide polymer on earth and wide variety of micro and macro-organisms used it as a food source. It is a beta linked linear polymer of 8000-12000 glucose units (Tomme et al., 1995). These polymers bound together to form highly ordered crystalline structures and less ordered (amorphous) regions. These

amorphous regions can easily be degraded by cellulolytic enzymes. Three major enzymes are involved in the degradation of cellulose to glucose are endo glucanase (endo-1,4-β-D-glucanase, EG, EC 3.2.1.4), cellobiohydrolase (exo-1,4-β-D-glucanase, CBH, EC 3.2.1.91), and β-glucosidase (1,4-β-D-glucosidase, BG, EC 3.2.1.21). Endoglucanase acts in random fashion, cleaving β−linked bonds within the cellulose molecule; CBH removes cellobiose units from the non reducing end of the cellulose chain and BG degrades cellobiose and cellooligosachcharides to glucose. In different organisms different cellulase systems exist and these enzymes work synergistically as a system (Nishida et al., 2007).

The cellulolytic enzymes have been isolated from plants, microorganisms (bacteria and fungi) (Tomme et al., 1995) and herbivorous invertebrates such as arthropods (Watanabe et al., 1997), nematodes (Smant et al., 1998) and mollusks (Xu et al., 2000; Suzuki et al., 2003). It was widely accepted that cellulose degrading ability was restricted to plants and microbes, and that animals, which fed on cellulose were able to digest it only because they harbored symbiotic cellulolytic bacteria, fungi or protozoa in their guts (Wang et al., 2003). However, purification of cellulases from essentially microbe free invertebrate gut regions and the subsequent cloning of genes encoding such proteins from several invertebrates have shown the cellulytic activity (Erasmus et al., 1997).

There are about approx. 100 families of glycosyl hydrolase families including cellulase on the basis of the primary structure (http://afmb.cnrs-mrs.fr/CAZY/). Among them around 12 families contain enzymes such as endo-β-1,4-glucanases or cellobiohydrolases that can hydrolyse cellulose. Most taxonomically diverse family 9 contains cellulases from plants, bacteria, a slime mold, crayfish, cockroach, termite, abalone and sea urchin but no any fungal species (Tokuda et al., 1999). Family 5 contains only bacteria, fungi and nematode cellulases. According to the primary structure analysis, the blue mussel cellulase is classified into the GHF45 subfamily 2, being distinct from the arthropods that are classified into GHF9 even though, the origin of mussel cellulases were similar to arthropods (Nishida et al., 2007).

Although, the primary structure data for metazoan cellulase genes are available, biochemical properties and physiological significance of metazoan GHF9-cellulases, especially those of mollusk enzymes, are still in poor.

**Amylase**

Starch is composed of the glucose polymers, amylose (linear) and amylopectin (branched). Starch-converting enzymes are basically divided into four groups: endo amylases, exo amylases, de branching enzymes and transferases. Endo amylases like α-amylases, which belongs to glycoside hydrolase family 13 are able to cleave α-1-4- glycosidic bonds that are present in the inner part of amylose or amylopectin chains (MacGregor et al., 2001).

They are found in phyla from Archaea (Pandey et al., 2000) to mammalia (Machius et al., 1995). Alpha amylase was isolated and purified from bacteria, protozoa such as *Entamoeba histolytica and Toxoplasma gondii*, and characterized some of the physicochemical properties of the enzyme. Alpha amylase was purified and characterized from the muscle and intestines of the parasitic helminth of pigs *Ascaris suum* (Zoltowska, 2001). Alpha amylase was cloned from *Drosophila melanogaster* and identified the characteristic conserved motifs typical to α-amylase in vertebrates (Boer and Hickey, 1986). Several α-amylase have also been cloned in fish species such as winter flounder, *Pseudopleuronectes americanus* (Douglas et al., 2000); barramundi, sea bass; *Lates calcarifer* (Ma et al., 2004); and spotted green puffer fish, *Tetraodon nigroviridis* (Bouneau et al., 2003). Aditionally, α-amylase sequences are available in the GenBank data base such as fish Japanese eel, *Anguilla japonica* (GenBank accession no. AB070721); zebrafish, *Danio rerio* (Accession no. BC06286); in mollusks *Pecten maximus* (P91778) and Pacific oyster, *Crassostrea gigas* (CAA69658). However, there are limited sequence information is currently available for invertebrates.

**Alginate lyase**

Alginate is a major structural hetero polysaccharide in brown algae that is composed of (1→4)-linked β-D-mannuronate (M) and C-5 epimer alpha-L-guluronate (G). Those are arranged as a homopolymeric (polyM block and polyG block) and or random heteropolymeric M or G blocks. Alginate is hydrolysed by alginate lyase, by a β-elimination mechanism forming 4-deoxy-L-erythro-hex-4-enopyranosyluronate at the new non reducing terminus. The enzyme has been detected from a wide variety of sources such as mollusks, seaweeds and bacteria and use for the analysis of the fine structure of alginates and production of protoplasts from seaweeds. Depending on the substrate specificity, the alginate lyase was primarily classified into poly (M)-lyase and poly (G)-lyase acting on the poly (M) and poly (G) blocks, respectively. Up to now, primary structure analyses of alginate lyases have been performed mainly on bacterial enzymes, and more than 20 genes of bacterial enzymes have been cloned. The amino acid sequence of a molluscan enzyme has been determined only in *Turbo cornutus* SP2 by the protein method. Recently, *H. discus hannai* alginate lyase cDNA was cloned and characterized (Schimizu et al., 2003).

**Endo-β-1,4-mannanase**

Endo-β-1,4-mannanase (β-mannanase, EC 3.2.1.78) is the enzyme that catalyzes the endolytic hydrolysis of β-1,4-mannosidic linkages in mannopolysaccharides such as β-1,4-mannan, glucomannan, and galactomannan. Beta-1,4-Mannan is a major hemicellulose of both higher plants and red algae, consists of β-1,4-linked D-mannopyranose residues. Aditionally, β-mannanase is an important enzyme in various industrial processes (Wong and Saddler, 1993; Christgau et al., 1994a) and manufacturing manno oligosaccharides (mannotriose and mannobiose), which are valuable for food additives and pharmaceutical ingredients (Kobayashi et al., 1987). Beta-mannanase is also useful for cell-engineering of red algae, since this enzyme can degrade β-1,4-mannan, a major structural polysaccharide of red algae (Percival and McDowell, 1967), and weaken the cell-wall structure (Gretz et al.,

1982). It has been reported that, protoplasts could be prepared from a red alga *Porphyra* sp. by using an abalone crude enzyme that contains β-mannanases (Polne-Fuller and Gibor, 1984; Gall et al., 1993; Dai et al., 2004; Dipakkore et al., 2005).

The cDNAs encoding β-mannanases have been cloned from bacteria (Gibbs et al., 1996; Cann et al., 1999; Sunna et al., 2000; Charrier and Rouland, 2001), fungi (Christgau et al., 1994b), and mollusks (Xu et al., 2002c; Ootsuka et al., 2006). Molluscan β-mannanases have been isolated from *Pomacea insularus* (Yamaura and Matsumoto, 1993), *Littorina brevicula* (Yamaura et al., 1996), and *Mytilus edulis* (Xu et al., 2002a), and characterized to some extent. The N-terminal sequences of *Pomacea* and *Littorina* enzymes were determined by protein sequencing methods (Yamaura and Matsumoto, 1993; Yamaura et al., 1996), whereas the complete primary structure of the *Mytilus* enzyme and Pacific abalone *Haliotis discus hannai* were deduced by the cDNA method (Ootsuka et al., 2006). However, there is no report on isolation of β-mannanase and its primary structure of disk abalone, *H. discus discus*.

**Laminarinase**

Laminarinase (endo-β-1,3-glucanase) is an enzyme which displays its main hydrolytic activity on the β-1,3-glucose polymer laminarin. On the basis of substrate specificity, this protein belongs to the laminarinase subfamily of the family 16 glycosyl hydrolases (Allouch et al., 2003), which are enzymes that are widely distributed among bacteria, fungi and higher plants. The enzyme displays its highest hydrolytic activity on the β-1,3-glucose polymer laminarin in brown seaweeds and has some hydrolytic activity on the β-1,3-1,4-glucose polymers lichenan and barley β-glucan (Gueguen et al., 1997). The β-1,3-glucanases can play various physiological roles. In plants, they have been implicated in cell differentiation and defense against fungal pathogens (Mackay et al., 1985). In fungi, β-1,3-glucanases are important in morphogenetic processes, β-glucan mobilization and fungal pathogen plant interactions (Bachman & McClay, 1996). In bacteria, β-1,3-glucanases

hydrolyze 1,3-β-glucosyl linkages, but they usually require a region of un substituted contiguous 1,3-β-linked glucosyl residues (Gueguen et al., 1997).

**Arylsulfatase**

Sulfatases are large protein family, which constitutes a group of enzymes capable of hydrolyzing the sulfate ester bond of a variety of biological compounds. Sulfatases hydrolyze sulfate esters from different sulfated substrates such as carbohydrates, proteoglycans, glycosaminoglycans, steroids and glycolipids (Parenti et al., 1997). Since, one of the main constituent of brown and red seaweeds is sulfated polysaccharides, it is important to study sulfatase family genes under PDEs, as they indirectly support for degradation of sulfated polysaccharides.

A sub set of sulfatases, referred to as arylsulfatase, are characterized by the ability to hydrolyze, besides their natural substrate, sulfate ester of chromogenic or fluorogenic aromatic compounds (aryl compounds), such as p-nitrophenyl sulfate, p-nitrocatechol sulfate and 4-methylumbelliferyl (4-MU) sulfate (Parenti et al., 1997). Many bacterial arylsulfatase have been known to play role in degradation of sulfated polysaccharides (Roberton, 1997; Barbeyron et al., 2004). Sulfated polysaccharides were degraded to sulfated disaccharides by lyase, followed by desulfatation of the resultant disaccharides and hydrolysis to the monosaccharide level by a glycosidase in bacteria. Arylsulfatase hydrolysed sulfate ester bonds in agar without any glycosidase activity, which is demonstrated by increasing of gelling strength of enzyme treated agar (Seo et al., 2001).

All sulfatases contained a Cα-formylglycine (FGly) residues at their catalytic site that is essential for enzymatic activity. FGly formation occurs post-translationally by the oxidation of a cysteine residue that is conserved in all eukaryotic sulfatases as well as in most prokaryotic sulfatases (Dierks et al., 1997). Sulfatase Modifying Factor I (SUM1) gene, which has been highly conserved during evolution, from bacteria to humans (Landgrebe et al., 2003) encodes the formylglycine generating enzyme, which is responsible for the cystein to

FGly conversion. The bacterial protein AtsB modifying factor which is not homologous to SUM1 is responsible for post-translational modification of Ser type sulfatases (Dierks et al., 1997).

Most of the studies on arylsulfatases have been conducted on mammals (human, rat, goat) and classification based on enzymes of lower vertebrate and invertebrate sulfatases were limited. So far human, rat, goat, avian sulfatases have been cloned and their coding sequences were characterized. It has been reported that several marine animals that feed on algae are known to secrete carbohydrate sulfatases as digestive enzymes, and cleave the sulfate ester bonds in dietary polysaccharides to improve digestion and absorption of marine polysaccharides (Milanesi et al., 1972; Hoshi, 1980). The hepatopancrase of various mollusks were found to be the rich source of carbohydrate sulfatase. Arylsulfatase gene has been identified and characterized in sea urchin (*Hemicentrotus pulcherimus*). Mollusks sulfatases such as *Helix pomatia* SULF1 and SULF2 have also been characterized however, the protein was not purified (Wittstock, 2000).

There is an extensive interest in polysaccharides as a major source for renewable energy and raw materials. Marine algae are the only sources for industrially important phycocolloids like agar, carrageenan and alginate and are proven to be rich source of structurally diverse bioactive compounds with valuable pharmaceutical and bio medical potential. Therefore, break down of these compounds is important. Presently, PDEs could convert polysaccharides to utilizable oligosaccharides or monosaccharides such as glucose efficiently, under mild conditions, have been the subject of much investigation. In the last decades, enzyme separation and characterization has become increasingly important because of the evolving application of enzymes in the brewing, food, textile, chemical, detergent and central role in agro-industry and biotechnology. New enzymes from microbes and plants are constantly investigated for food and pharmaceutical industries, and for analytical purposes. However, as those already discovered enzymes, which responsible for polysaccharide

degrading, have not completely met the requirements for the intended industrial applications. Therefore, continuation of the previous studies and search for new enzymes and study their nature of the biochemical properties would be beneficial for the improvement of their biochemical properties to suit for the industrial applications.

Knowledge of feeding strategies and adaptive digestive capacities of an organism under different food conditions and their effects on growth and survival are of major interest in aquaculture. It has been reported that control of digestive enzyme activities may, therefore be important in maximizing absorption efficiencies and food conversion efficiencies. Limitation of food and low food conversion efficiency would cause improper nutrients absorption to abalone and thereby cause slow growth rate of the animal. Therefore, better understanding of physiological and environmental factors that determine optimal food conversion efficiencies would be a better solution for the problem of food limitation. Although, abalone farming is primarily based on fresh algal foods, experimental evidence showed that juvenile abalone can eat pelleted feed and grow at a higher rate when compared to kelp food. Therefore, studies of PDEs are important, to understand the nature of PDEs in disk abalone, to maximize absorption and food conversion efficiencies, and to optimize the formulation of artificial diets of abalone. Also resulted oligosaccharides and monosaccharides after break down of polysaccharides could be used for many industrial, pharmaceutical and medical field applications.

Up to date, the basic knowledge of key physiological processes of digestion, and more broadly in growth and development, remains poorly understood in abalones. A better knowledge in regulation of expression of polysaccharide degrading genes is required, to understand the relationships between enzymatic properties and absorption processes of disk abalones. Therefore, researches are needed to determine how these genes are regulated and to elucidate the nature of environmental signals, their mediators, and the transduction pathways by which messages reach nuclei. In cockles, digestive enzyme activities were identified as a

physiological parameter affecting digestive capacity (Ibarrola et al., 1998). Until now nature of its regulation remains unknown for most PDEs.

Therefore, in this study, identified PDEs in disk abalone *H. discus discu*s were studied in molecular and biochemical level. The gene sequences were characterized and purified the recombinant proteins using pMAL fusion protein purification system. The recombinant proteins were biochemically characterized for their respective functions. I*n vivo* expression study was conducted to determine the mRNA expression of those genes in digestive tract and hepatopancreas tissues of disk abalone. Additionally, PDEs expression pattern was evaluated in hepatopancreas tissue after starvation and re-fed of alga. Moreover, CuZnSOD gene expression was analyzed to understand the effects of prolonged starvation and re-feeding on the antioxidant status and related oxidative stress of disk abalone.

## MATERIALS AND METHODS

### *Identification, cloning and sequence characterization of disk abalone PDEs*

#### Identification and cloning of PDE related genes from disk abalone cDNA library

Abalone cDNA library was constructed using total RNA isolated from the abalone digestive gland and cDNA library construction kit (Creator$^{TM}$ SMART$^{TM}$, Clontech, USA). The cDNA was normalized with Trimmer-Direct cDNA normalization kit according to the manufacture's protocol (Evrogen, Russia). After assemble the individual EST data into group of sequences, the resulting data were compared against the National Center for Biotechnology Information (NCBI) nonreduntant protein data base by using the programme BLAST-X. Sequences with a significant similarity to known PDEs were identified and compared with full length coding sequences available at NCBI GenBank for the genes corresponding putative function. Six cDNA clones, β-1,4-endoglucanase (HdEndg), α-amylase (HdAmy), β-mannanase (HdMann), Alginate lyase (HdAlg), laminarinase (HdLms) and Arylsulfatase (HdArys) were selected from the cDNA library. They were then transformed into XL1-Blue MRF' cells and plasmid DNA of the each putative PDEs were obtained by the Accuprep$^{TM}$ plasmid extraction kit (Bioneer Co., Korea). The full length of each clone was determined by sequencing with internal primers from the known sequence of 3' end using termination kit, Big Dye and an ABI 3700 sequencer (Macrogen Co., Korea). Primer list which was used for sequencing is given in the Table 1.

#### Sequence characterization of disk abalone PDEs

After determine the full length coding sequence, the sequence was compared with other sequences available in the NCBI database. Nucleotide and amino acid sequence of abalone PDEs were analysed using the DNAssit program (version 2.2). The protein sequences were aligned with the ClustalW multiple sequence alignment program (version 1.8) and phylogenetic tree was developed by MEGA 3.1 program using neighbor-joining

method. Signal peptide was predicted using a SignalP program (http://www.cbs.dtu.dk/). The catalytic regions were determined by ExPasy motif scan program (http://myhits.isb-sib.ch/cgi-bin/motif_scan). Protein conformation was molded by the Swiss-Model (http://swissmodel.expasy.org/) and viewed in the Swiss Pdb viewer (Schwede et al., 2003).

***In vivo PDEs mRNA expression analysis during starvation and re-feeding***

**Animals**

The disk abalones were purchased from Youngsoo abalone farm (Jeju, Republic of Korea). Individuals mean body mass were 45-50±1.8 g (with mean body length were 8x5 cm) and they were maintained in 40 L tanks with an aerated sand filtered sea water, having temperature of 19±1 $^{o}$C for a week to acclimatize to laboratory conditions. Maximum of 10 healthy disk abalones were kept per tank undisturbed and fed with fresh seaweed materials daily. All were individually tagged before being taken for the experiments and body mass were noted down.

**Starvation treatment and isolation of abalone tissues**

Starvation was continued until 8 weeks and after that re-fed the animals with seaweeds continuously up to 28 days. Separate tank was maintained for 8 weeks with continues feeding as a control. The digestive tract and hepatopancreas tissues were isolated from three healthy abalones, separately for tissue expression analysis. Hepatopancreas and digestive tract tissues were isolated from starved animals. All the tissues were immediately snap-frozen in liquid nitrogen and stored at -70 $^{o}$C.

**Table 1:** Primers used in this study

| Name | Object | Sequence (5' to 3' direction) |
|---|---|---|
| HdEndg F1 | Internal sequencing 1 | GCTGGACAAGAGGATCATATG |
| HdEndg F2 | Internal sequencing 2 | AGTCCTACAGCCAAGTGCTA |
| HdEndg F3 | Internal sequencing 3 | TGTCAACAACCCGGTGGTCTAC |
| HdEndg F4 | ORF amplification | $(GA)_3$TCTAGAGATGTTCCCGTCACCTTCTC (XbaI) |
| HdEndg R5 | ORF amplification | $(GA)_3$AAGCTTTCACATCAGTGGTTTTAACAGAAAGT (HindIII) |
| HdEndg F6 | RT-PCR | AGAGACCACTCAACACACTTGCCT |
| HdEndg F7 | RT-PCR | TGTACGCATCCGACCATTTCAGGA |
| HdAmy F7 | Internal sequencing 1 | CTTCCAGGAGGTAATCGACATG |
| HdAmy F8 | Internal sequencing 2 | CCGATAAAGATCCAATGGTGGC |
| HdAmy F9 | ORF amplification | $(GA)_3$GAATTCTCGATGTACAGTGATCCCCA (EcoRI) |
| HdAmy R10 | ORF amplification | $(GA)_3$ AAGCTTCTAACCAACGCGCTTTGG (HindIII) |
| HdAmy F11 | RT-PCR | TACCTGGTGTTTAGCGTTCCTGCT |
| HdAmy R12 | RT-PCR | ACCTCCAACACCGGTATGTGATT |
| HdAlgl F13 | Internal sequencing 1 | CATCTACTTCGACCCCAACTT |
| HdAlgl F14 | ORF amplification | $(GA)_3$GAATTCATGGACATTCCCATTACAAAACAC (EcoR I) |
| HdAlgl R15 | ORF amplification | $(GA)_3$CTGCAGCTAAAGAGTTACCGTAGTCGAAGTTG (Pst I) |
| HdAlgl F16 | RT-PCR | GCTGGCGACAAACTGTGTGTGAAA |
| HdAlgl R17 | RT-PCR | TGCCTCCCTTAACCCAGTCAAAGT |
| HdMann F18 | Internal sequencing | CTGGGCTGGGAGACTTTATA |
| HdMann F19 | RT-PCR | ACTATTAGTCTTGCCAGCCGTGGA |
| HdMann R20 | RT-PCR | TTTCACAGACTTGACCCACGGGAT |
| HdAryls F21 | Internal sequencing 1 | GGCACTTACTCAACGCTTC |
| HdAryls F22 | ORF amplification | $(GA)_3$GAATTCGCAGGACGTCCACGCCATA (EcoRI) |
| HdAryls R23 | ORF amplification | $(GA)_3$TCTAGATCAACACCAGCCAGGAGACCA (XbaI) |
| HdAryls F24 | RT-PCR | ATTCGTGGCTTCCTTGACTACCGT |
| HdAryls R25 | RT-PCR | TACATGAAGCCCACAGCTCTCGTT |
| HdLms F26 | RT-PCR | ATGGCGTTCTCTACCTCAAACCGA |
| HdLms R27 | RT-PCR | TTGGCGACGCTCCGTAATGAATTG |
| HdAct F28 | RT-PCR | GACTCTGGTGATGGTGTCACCCA |
| HdAct R29 | RT-PCR | ATCTCCTTCTGCATTCTGTCGGC |
| HdCuZnSOD F30 | RT-PCR | GGCAAACATGGCTTCCACGTTCAT |
| HdCuZnSOD R31 | RT-PCR | TCATCCACTCCAGCATGGACAACA |

**Determination of weight loss percentage (%) during starvation**

Disk abalone body weight was measured before being starvation started. Body weight was measured every 2 weeks interval up to 8 weeks of starvation period. Similarly experiment was continued for re-fed animals from day 1-28 and control animals seperatly, which did not undergone starvation. Experiment was conducted for three replicate animals. Weight loss was determined as percentage weight loss by taking the difference between initial (W1) and final weight (W2).

**RNA extraction and cDNA synthesis**

Total RNA was extracted from isolated tissues (100 mg each) using Tri Reagent™ (Sigma, USA) according to the manufacturer's protocol. The RNA concentration was determined by measurement of absorbance at 260 nm in a UV-spectrometer (Bio-Rad, USA). The respective cDNAs from selected tissues were synthesized using SuperScript III first strand cDNA synthesis kit (Invitrogen, USA). Briefly, RNA was incubated with 1 μL of 50 μM oligo$(dT)_{12–18}$ and 1 μL of 10 mM dNTP for 5 min at 65 °C. After incubation, 4 μL of 4× cDNA synthesis buffer, 1 μL of dithiothreitol (DTT, 0.1 M), 1 μL of RNaseOUT™ (40 U/μL) and 1 μL of cloned AMV reverse transcriptase (15 U/μL) were added and incubated for 1 h at 45 °C. The PCR reaction was terminated by adjusting the temperature to 85 °C for 5 min and the resulting cDNA was stored at −20 $^{o}$C.

**Tissue specific mRNA expression analysis by RT-PCR**

RT-PCR was performed to determine the mRNA expression in hepatopancreas and digestive tract tissues in control samples. Hepatopancreas tissue was analysed to findout mRNA expression pattern during starvation. Gene specific primers were used to amplify PDE cDNA fragments as listed in Table 1. Disk abalone actin expression was used as an internal PCR control by amplification of 492 bp fragment. All PCR reactions were carried out in a

25 μL reaction volume containing 4 μL of cDNA from each tissue, 2.5 μL of 10× TaKaRa Ex Taq™ buffer, 2.0 μL of 2.5 mM dNTP mix, 1.0 μL of each primer (20 pmol/μL), and 0.125 μL (5 U/μL) of TaKaRa Ex Taq™ DNA polymerase (TaKaRa, Japan). Preliminarily, the PCR cycle number was optimized in different tissues by performing different cycle numbers ($n$ =22–24, 26, and 28). After analyzing expression results from different cycles, 28 PCR cycles were selected to analysis of tissue specific and after starvation expression and 23 PCR cycles were used for actin expression analysis. The cycling protocol was: one cycle of 94 °C for 2 min, 28/23 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s, and one cycle of 72 °C for 5 min for the final extension. The PCR products were visualized on a 1.5% agarose gel stained with ethidium bromide. Electrophoretic images and the optical densities of amplified bands were analyzed using the Scion Image software (Release alpha 4.0.3.2).

**Data analysis**

For comparison of relative mRNA levels, statistical analysis was performed with one way ANOVA and mean comparisons were performed by Duncan multiple range tests using SPSS 11.5. Significant P values (P>0.05) were obtained by Duncan multiple range test. Results are shown as mean ± SE of three animals per group.

***Functional characterization of PDEs in disk abalone***

**Cloning of PDEs coding sequences into the pMAL-c2x expression vector**

The plasmid DNA of 4 cDNA clones, HdEndg, HdAmy, HdAlgl and HdArys were used as templates for ORF amplification. Primers were designed with corresponding restriction enzyme sites at the N and C-terminus of each four clones (Table 1). Breifly, 50 μL PCR reaction was conducted with 5 units of Herculase, 5 μL of 10x Herculase buffer, 1 μL of 10 mM dNTP, 50 ng of template and 20 pmol of each primer. After initial incubation at 94 °C for 2 min, 25 cycles were carried out with 30 s denaturation at 94 °C, 30 s of annealing at

50 °C, and 1 min 45 s of elongation at 72 °C, followed by a final extension at 72 °C for 5 min. The PCR product was analyzed using 1% agarose gel and ethidium bromide staining. The PCR product was purified using the Accuprep™ gel purification kit (Bioneer Co., Korea) and digested with the respective restriction enzymes. The expression vector, pMAL-c2X, was digested with the same restriction enzymes as the PCR product. The vector and the PCR product were purified by a 1% agarose gel using the Qiaex-II gel purification kit (QIAGEN Inc., USA). The ligation was carried out at 15 °C, overnight with 100 ng of pMAL-c2X vector, 70 ng of PCR product, 1 μL of 10× ligation buffer and 0.5 μL 1× T4DNA ligase (TaKaRa, Japan). The ligated products of each clone were transformed into DH10b cells. The correct recombinant (confirmed by restriction enzyme digestion and sequencing) was transformed into competent cells of *E. coli* K12 TB1.

**Over-expression and purification of recombinant proteins**

The recombinant PDEs, HdEndg, HdAmy, HdAlgl and HdArys were over-expressed in *E. coli* K12TB1 cells in the presence of isopropyl-ß-thiogalactopyranoside (IPTG). Briefly, a 5 mL of *E. coli* K12 TB1 starter culture was inoculated into 100 mL Luria broth with 100 μL ampicillin (100 mg/mL) and 10 mM glucose (2% final concentration). The culture was incubated at 30 °C with shaking at 200 rpm until the cell count reached 0.5 at 600 nm optical density. The culture was shifted to 20 °C for 15 min and induced by IPTG at 0.5 mM final concentration. After 6 h of induction, the cells were cooled on ice for 30 min, and harvested by centrifugation at 4000 × g for 20 min at 4 °C and removed the supernatant. Cells were re-suspended in 5 mL column buffer (Tris–HCl, pH 7.4, NaCl) and frozen at −20 °C for over night. After thawing in ice, the bacterial cells sonicated six times in short pulses of 10 s. The supernatant was diluted with a 1:5 column buffer after centrifuged at 9000× g for 30 min at 4 °C. The recombinant fusion protein which fused with maltose binding protein (MBP) was purified by pMAL™ protein fusion and purification system. In brief, resin was poured

into a 1 × 5 cm column and washed by column buffer (8 column volumes). The diluted crude extract was loaded at a flow rate of 1 mL/h. The column was then washed with 12 column volumes of column buffer, and the fusion protein was eluted with an elution buffer (column buffer + 2.5 mM maltose). The purified proteins were collected in 250 μL aliquots and respective samples were run on 10% SDS-PAGE with a protein marker (Bio-Rad, USA). Gels were stained using 0.05% Coomassie blue R-250, followed by a standard de-staining procedure. The concentrations of the purified proteins were determined by the method of Bradford using bovine serum albumin (BSA) as the standard (Bradford, 1976).

**Biochemical characterization of recombinant PDEs**

**Endoglucanase activity assay**

Modified method of Sugimura et al., 2003 was used to assay HdEndg activity (carboxymethyl cellulose;CMCase). Reaction mixture (0.5 mL) contained 1% (w/v) boiled CMC solution, 50 mM acetate buffer, pH 5.5 and appropriately diluted enzyme solution. After 30 min incubation at 40 $^{o}$C, the reducing sugar liberated in the reaction mixture was measured by the dinitrosalicylic acid method (Bernfeld, 1955. One unit (U) of HdEndg or CMCase activity is defined as the amount of enzyme which produces 1 μmole reducing sugar as glucose per min in the reaction mixture under the specified conditions. To determine the optimum temperature, HdEndg activity was determined in different temperature range from 10-80 $^{o}$C with 10 $^{o}$C intervals at pH 5.5. Thermal stability of the HdEndg was determined by measuring the remaining activity of the enzyme that had been incubated at 10-80 $^{o}$C with 10 $^{o}$C intervals for 30 min. Also thermal stability of the HdEndg was determined from 1-4 hrs at 30, 40, 50 and 60 $^{o}$C for every 1h interval. pH dependence of HdEndg activity was determined at 40 $^{o}$C in the reaction mixtures adjusted at pH 3-10 with different buffers. Buffers used were 50 mM citrate buffer (pH 3-6.5), 50 mM sodium phosphate buffer (pH 6-8) and 50 mM glycine-NaOH buffer (pH 8-10). To evaluvate the effect of metal ions and

chemical reagents on the HdEndg activity, $CaCl_2$, KCl, $MgSO_4$, $MnCl_2$, NaCl, EDTA, $ZnSO_4$, $CuSO_4$ and $FeSO_4$ were used in final concentration of 2 mM in the reaction.

**Alpha amylase activity assay**

The α-amylase activity in the recombinant purified protein was determined using modified method of Lim et al., 2003 by measuring the amount of reducing sugars released during incubation with starch. A 75 µL of enzyme was diluted with 10 mM Tris-HCL buffer (pH 7.0) to total volume of 0.5 mL and was added to 250 µL of 1% (w/v) starch dissolved in 10 mM Tris HCl buffer (pH 7.0), and the mixtures were incubated at 50 $^{o}$C for 30 min. One unit of the enzyme activity was defined as the amount of the enzyme that liberated 1 µmol of reducing sugar per min at 50 $^{o}$C. The amount of reducing sugar released was determined by the dinitrosalicylic acid method (Bernfeld, 1955). Temperature dependence on α-amylase activity was determined at different temperatures from 20-70 $^{o}$C for 30 min incubation period. Thermostability of α-amylase was obtained by pre incubating α-amylase samples in 10 mM HCl buffer (pH 7.0) at 40, 50 and 60 $^{o}$C for 1-4 hrs in every 1 h interval, and measuring residual activity under described above assay conditions. The effect of pH on the α-amylase activity was determined by using above amylase assay from pH 3.5-10.0. All the assays were performed at 50 $^{o}$C. Effects of metal ions and various reagents on the α-amylase activity were measured at pH 7.0 and 50 $^{o}$C in the presence of $CaCl_2$, KCl, $MgSO_4$, $MnCl_2$, NaCl, EDTA, $ZnSO_4$, $CuSO_4$ and $FeSO_4$.

**Aliginate lyase activity assay**

HdAlgl activity was determined using modified method of Shimizu et al., 2003. To determine the HdAlgl activity, 1% (w/w) sodium alginate was dissolved in distilled water and heated at 90 $^{o}$C for 1 h before use. The alginate lyase activity was assayed in a 1 mL reaction mixture containing 0.1% substrate, 50 mM sodium phosphate (pH 6.5) and an appropriate

amount of enzyme at 40 $^{o}$C. The reaction was initiated by the addition of 75 µL of enzyme and absorbance was measured at 235 nm. One unit of lyase activity is defined as the amount of enzyme that increases absorbance at 235 nm to 0.01 for 1 min. Temperature dependence of HdAlgl was measured at 20-60 $^{o}$C in 50 mM sodium phosphate buffer (pH 6.5). The effect of pH on the alginate lyase activity was determined by performing alginate lyase assay with varing the pH from pH 5.0-10.0.

**Sulfatase activity assay**

Arylsulfatase activity was assayed by measuring the amount of p-nitrophenol released from pNPS. The assay mixture containing 1 mL of enzyme solution and 250 µL of 25mM pNPS (pH 7.0) was incubated at 45 $^{o}$C for 60 min. The reaction was stopped by adding 1 mL of 0.5 M NaOH. The released p-nitrophenol was measured using absorbance at 410 nm. One unit of arylsulfatase activity was defined as the catalysis of producing 1.0 µmol of p-nitrophenol from pNPS per minute by the enzyme. Sulfatase activity was measured according to modified method of Dodgson and Price, 1962 using 4 sulfated marine polysaccharides (agar, dextran sulfate, chondroitin 6 sulfate sodium salt, chondroitin sulfate B sodium salt) to determine substrate specificity of the enzyme. The assay mixture containing 100 µL enzyme solution and 900 µL of 0.5% agar suspended in Tris HCl buffer (pH 7.0) was incubated at 45 $^{o}$C for 60 min. Then the sample was centrifuged at 3,000x g for 5 min. Supernatant (900 µL) was taken and mixed with 200 µL of concentrated HCl and 600 µL of 13% $BaCl_2.2H_2O$ and 3% Tween-80. After the incubation for 30 min at room temperature, the absorbance was measured at 420 nm.

# RESULTS

## Sequence characterization of PDEs

### Sequence characterization of disk abalone endoglucanse (HdEndg)

HdEndg full-length sequence was isolated from a disk abalone cDNA library and was deposited in the NCBI under the accession numbers of EF103351. The nucleotide and deduced amino acid sequences of HdEndg is shown in Figure 1. The start (ATG) and stop codons (ATG) were found in nucleotide position at 14 and 1838 of the full length cDNA. N-terminal contains a putative signal peptide of 17 amino acids in length based on SignalP V1.1, World Wide Web Prediction Sever, Center for Biological Sequence analysis (Nielsen et al., 1997). The complete nucleotide sequence shows an OFR of 1824 bp corresponding to putative protein of 608 amino acids. The predicted molecular mass of HdEndg was 65 kDa and estimated isoelectric point (p*I*) was 4.5. Also, two potential N- glycosylation sites were predicted at the positions of 80-83 and 148-150. A putative polyadenylation signal sequence AATAAA and a polyA tail were found in 3'-terminal region.

ClustalW pair-wise amino acid comparison revealed that HdEndg has the highest identity 50.3% to *H. discus hannai* (Table 2). Interestingly, 131 residues of the N-terminal mature peptide sequence showed 35% and 34.3% identity to 126 residues of *H. discus hannai* and 128 residues of *H. discus discus*, respectively. Similar to *H. discus hannai*, this N-terminal region was considered as the family II CBM, which was attached by a linker to catalytic domain (Figure 2). Additionally, C-terminal region of 460 residues in the HdEndg was identified as GHF9-type catalytic domain and it showed 56.3% and 56% identity with the corresponding regions of *H. discus hannai* (accession no: BAC67186) and *H. discus discus* (accession no: BAD 44734).

The amino acid residues, which are involved in catalysis and substrate binding in all GHF9 members are conserved in HdEndg. In HdEndg, the conserved active site histidine was replaced by Arg513 in first conserved region. The second region contained conserved

catalytically active residues at positions Asp557 and Glu566. In third conserved region, two catalytic active Asp residues were present at positions of Asp207 and Asp210. Therefore, HdEndg cDNA encodes full-length nucleotide sequences with the main characteristic features of the Endoglucanase GHF9 with CBM II family. ClustalW multiple sequence alignments of HdEndg with other known endoglucanases are shown in Figure 3. Results showed that characteristic features of GHF9 Endg possessing family II CBM of HdEndg were highly conserved in selected species with some modifications in the respective regions. Also, HdEndgs shares 50, 49, and 37% identity with the catalytic domains of *H. discus hannai*, *R. speratus*, and *Nasutitermes takasagoegensis*, respectively.

**Phylogenetic analysis**

To determine the relative position of the HdEndg in evolution, known Endoglucanase amino acid sequences from 21 representative species were used to construct a phylogenetic tree (Figure 4). Fungi *Mucor circinelloides* and bacteria *Cellulomonas fimi* Endg which belongs to GHF 5 and 6, respectively were used as out groups for the analysis. It showed that arthropods, mollusk, bacteria Endgs, which used here were made sub clusters in a separate main cluster, and plants were positioned in second main cluster. Even though, the HdEndg connected with mollusk sub cluster, previously identified *H. discus discus* HdEndg (accession no: BAD 44734) was not closely related with the *H. discus discus* in this study. HdEndg was more closely related to the *H. discus hannai.* However, *H. discus hannai* and *H. discus discus* were closely positioned in a same group. However, present sequence was diverted from those and separately positioned in a mollusk group. Therefore, phylogenetic analysis provides evidence that the HdEndg has been derived from a common ancestor and they diverged from one another for some extent with the evolution.

```
                                                                TATCAAGGACAGA     13
ATGCAGAGCAAGATCTTGGTGCTGGTGTCGTGGATGACCCTCGTGGCCGGCGATGTTCCCGTCACCTTC     82
M--Q--S--K--I--L--V--L--V--S--W--M--T--L--V--A--G--D--V--P--V--T--F--     23
TCCGTCGTATGGGCCAGTGCTGAAAAGGGATCCTTCTGTGTACCAGTTGCCCGGGATCTGCATGGATGG    151
S--V--V--W--A--S--A--E--K--G--S--F--C--V--P--V--A--R--D--L--H--G--W--     46
AGGATCCACCTAATGTCCAGTGACAGCATCACGCGACTAGATATAACTGAGGCTTCTGTTGTTACAACA    220
R--I--H--L--M--S--S--D--S--I--T--R--L--D--I--T--E--A--S--V--V--T--T--     69
TTAAACAATGGCAGGGAATTCATCCTCGAGAACAAATCCTGGAATGCTGATGAACACGTGGGTGATACG    289
L--N--N--G--R--E--F--I--L--E--N--K--S--W--N--A--D--E--H--V--G--D--T--     92
ATTTGTGTAGACTTCCTAGCCAACACTGACCACCGCCTTGGGGTGCTGACTGGGCATGCGTACCTCGAG    358
I--C--V--D--F--L--A--N--T--D--H--R--L--G--V--L--T--G--H--A--Y--L--E--    115
GGCTTCAAGAGAGACCACTCAACACACTTGCCTCTGACTGATACAGTTTCTTTGTCCAGTACGCCAGCC    427
G--F--K--R--D--H--S--T--H--L--P--L--T--D--T--V--S--L--S--S--T--P--A--    138
ACAGTGAGATCAACACACGGGGCTGCTAATACTAGCGGGACAGCAGGAAAATACGACTACAAAGAAGTT    496
T--V--R--S--T--H--G--A--A--N--T--S--G--T--A--G--K--Y--D--Y--K--E--V--    161
TTGGGTCTCTCAATATTGTTTTATGACGCGCAACGATCCGGCAAACTTCCGGCAAATAATCATGTCTCC    565
L--G--L--S--I--L--F--Y--D--A--Q--R--S--G--K--L--P--A--N--N--H--V--S--    184
TGGCGCGCAGATTCGTCCTTGACAGACAAAGGGGACAACGGGGAGGATCTGACCGGGGGATGGTATGAT    634
W--R--A--D--S--S--L--T--D--K--G--D--N--G--E--D--L--T--G--G--W--Y--D--    207
GCCGGAGATATGATCAAGTTTAACCTTCCTATGGCTTCGGCAACAACAATATTATCCTGGGGTTTCCTG    703
A--G--D--M--I--K--F--N--L--P--M--A--S--A--T--T--I--L--S--W--G--F--L--    230
AAATGGTCGGATGCGTACAAAACGGCTGGACAAGAGGATCATATGTATGATATGATTAAGTGGCCTTTG    772
K--W--S--D--A--Y--K--T--A--G--Q--E--D--H--M--Y--D--M--I--K--W--P--L--    253
GATTACTTTTTGAAATGTTGGAATCCTGTCAAGAAAGAATATTATGTCCAGGTAGGCGAAAGCAGCCTC    841
D--Y--F--L--K--C--W--N--P--V--K--K--E--Y--Y--V--Q--V--G--E--S--S--L--    276
GACAGCACATTTTGGGGTAGACCAGAAGACCTACCAGCCAATCTGAAAGTGTTCAAGGTAACGGCTTCA    910
D--S--T--F--W--G--R--P--E--D--L--P--A--N--L--K--V--F--K--V--T--A--S--    299
AAGCCCGGAAGTGACGTTGCTGGCATCACAGCTGCAGCTCTGGCAGCTGGGTCAGTGCTATACCAGACA    979
K--P--G--S--D--V--A--G--I--T--A--A--A--L--A--A--G--S--V--L--Y--Q--T--    322
AAAGATGCTTCGTATGGAGCACGTCTTCTGAGTGCAGCGGAGAGTCTTTATGCATTCGCCACGACATAC   1048
K--D--A--S--Y--G--A--R--L--L--S--A--A--E--S--L--Y--A--F--A--T--T--Y--    345
AGGGGTATATACAGTGTCAGTGTTCCAGCCGCGGCATCAGCATACGCGTCCACTGGTTATAACGACGAG   1117
R--G--I--Y--S--V--S--V--P--A--A--A--S--A--Y--A--S--T--G--Y--N--D--E--    368
CTGTGTGTGGCGGCAGTCTGGCTGTACAAGGCAACACGAGACAAGAAATATCTGGAAGATGCGAAAACT   1186
L--C--V--A--A--V--W--L--Y--K--A--T--R--D--K--K--Y--L--E--D--A--K--T--    391
TATCACAGTTCCGCCTCCCCCTGGGCCTACTCATGGGACGATACCACTGTTGGGTGTCAGCTGATGCTG   1255
Y--H--S--S--A--S--P--W--A--Y--S--W--D--D--T--T--V--G--C--Q--L--M--L--    414
TACGATATGACCCGTGACGTCACCTATAGGACTGAGGTCCAGCGATTCCTGACCAGTTGGAAGCCAGGT   1324
```

```
Y--D--M--T--R--D--V--T--Y--R--T--E--V--Q--R--F--L--T--S--W--K--P--G--    437
GGCTCTCTTCCCTATACACCTTGTGGAATGGCCTTTAGAAGCAAATGGGGGTCACTCAGATACGACGCA  1393
G--S--L--P--Y--T--P--C--G--M--A--F--R--S--K--W--G--S--L--R--Y--D--A--    460
AACGTCGCCTTCATTGCCTTAATGGCCGCTGAGGACGGAATAGATCAAGTGGAAAACCGACAGTGGGCG  1462
N--V--A--F--I--A--L--M--A--A--E--D--G--I--D--Q--V--E--N--R--Q--W--A--    483
CTGAGTCAGCTGAATTACATACTTGGGGACAACAAACAACATTTGAGCTTTGTTGTCGGATTTGGGTCC  1531
L--S--Q--L--N--Y--I--L--G--D--N--K--Q--H--L--S--F--V--V--G--F--G--S--    506
AAGTATCCGTTACAGCCTCGCCACGGAGCTAGCTCCTGTCCGGATCAACCCGCCACATGTGACTGGTCC  1600
K--Y--P--L--Q--P--R--H--G--A--S--S--C--P--D--Q--P--A--T--C--D--W--S--    529
AACTTTCACAGTCCCGGTCCAAACCCCCACGTCCTTAAAGGGGCTCTAGTTGGAGGCCCGGACGGAACT  1669
N--F--H--S--P--G--P--N--P--H--V--L--K--G--A--L--V--G--G--P--D--G--T--    552
GACACTTATTCTGACAAGAGGAGCGACTACGTCGGGAACGAGGTGGCTGTTGATTACAACGCCGGCTTC  1738
D--T--Y--S--D--K--R--S--D--Y--V--G--N--E--V--A--V--D--Y--N--A--G--F--    575
CAATCCACAATAGCAGGTTTGGTTCATTTGTCAACAACCCGGTGGTCTACCTGCCAGTCCTACAGCCAA  1807
Q--S--T--I--A--G--L--V--H--L--S--T--T--R--W--S--T--C--Q--S--Y--S--Q--    598
GTGCTAGACTTTCTGTTAAAACCACTGATGTGATCACACACAGCCGTCTCAACACTGTCAATAAAATTC  1876
V--L--D--F--L--L--K--P--L--M--*                                         608
TGTCTACAAAAAAAAAAAAAAAAAAAAAAA                                          1908
```

**Figure 1:** The nucleotide and deduced amino acid sequences of the HdEndg. The sequence has been deposited in NCBI under the accession number EF103351. The start (ATG) and stop (TAA) codons are bold and shaded. A putative signal peptide is indicated by dotted underline. A putative N-glycosylation sites are boxed. Catalytically important residues in GHF9 are bold. A putative polyadenylation signal AATAAA is bold and under lined.

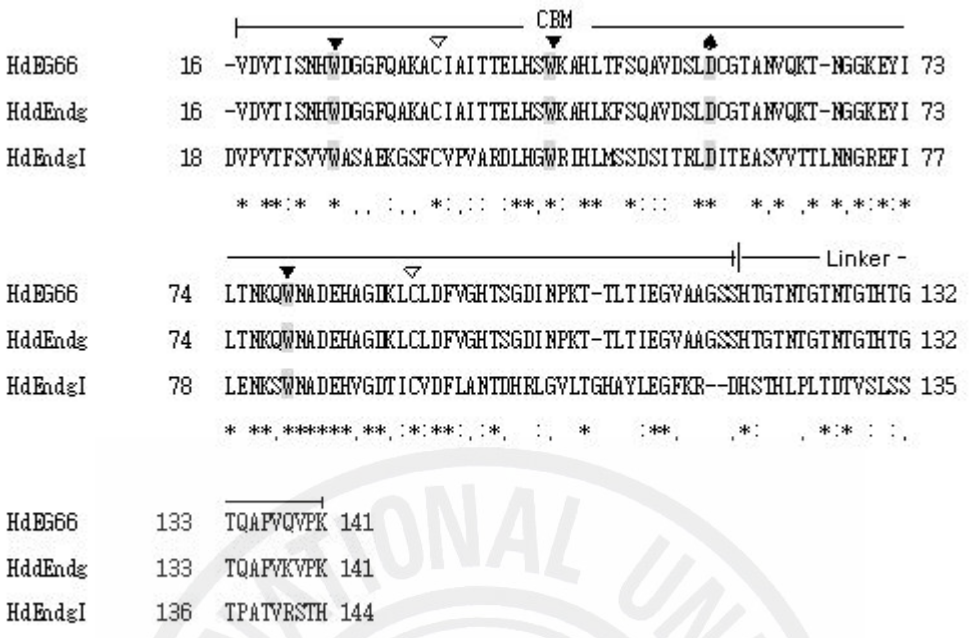
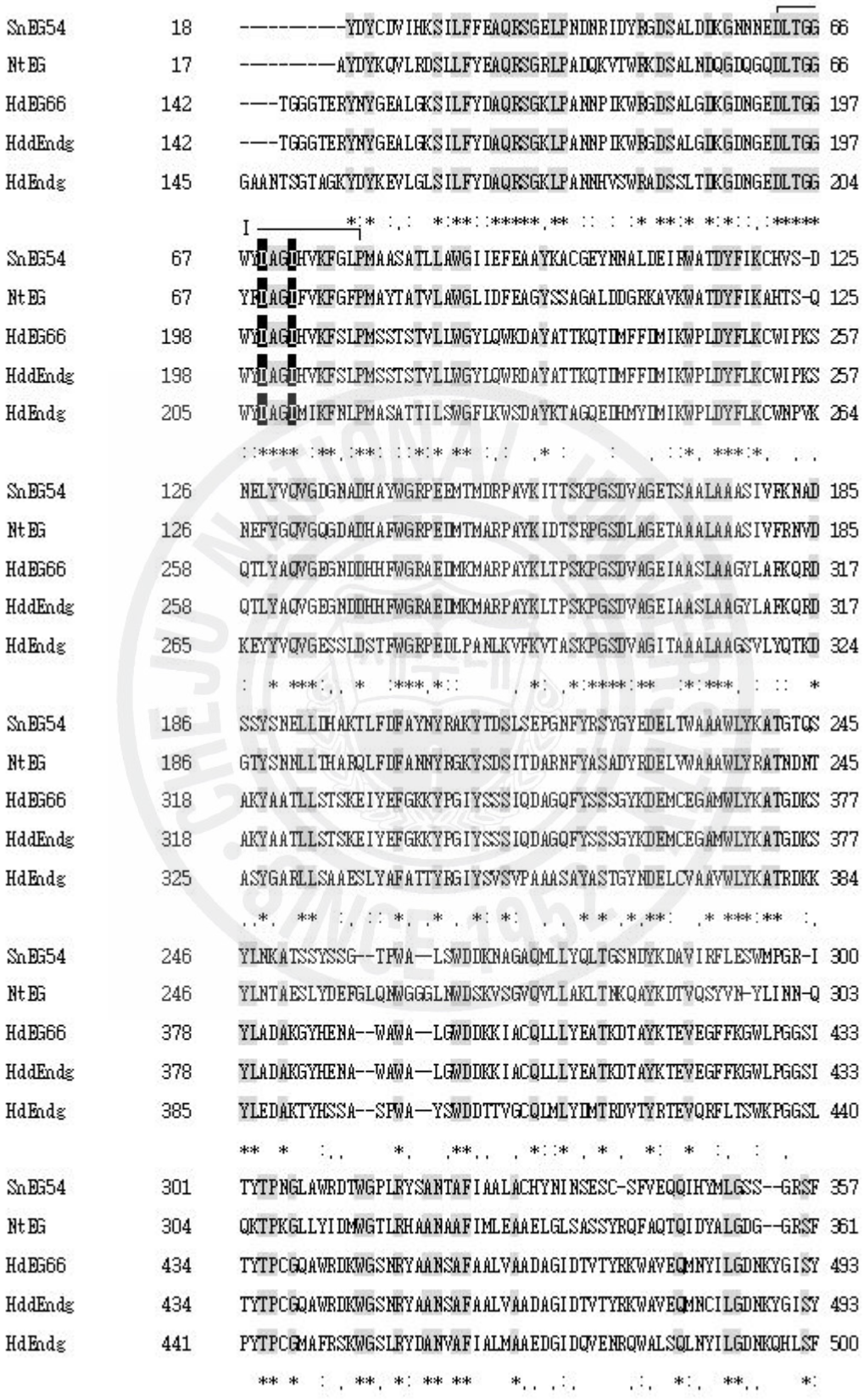



**Figure 2:** Comparison of HdEndg family II CBM amino acid sequence with known species family II CBM of endoglucanases. The sequence of N-terminal 127 residues of mature peptide was aligned with the CBM attached by a linker of *H. discus hannai* (HdEG66). Locations of three tryptophans strictly conserved in family II CBMs are shaded and indicated by ▼. The tryptophan residue substituted by aspartic acid is indicated by ♠. Conserved cystein residues are indicated in ▽. Identical residues among the sequence are indicated by asterisks. Accession numbers of the amino acid sequences are same as in figure 4.



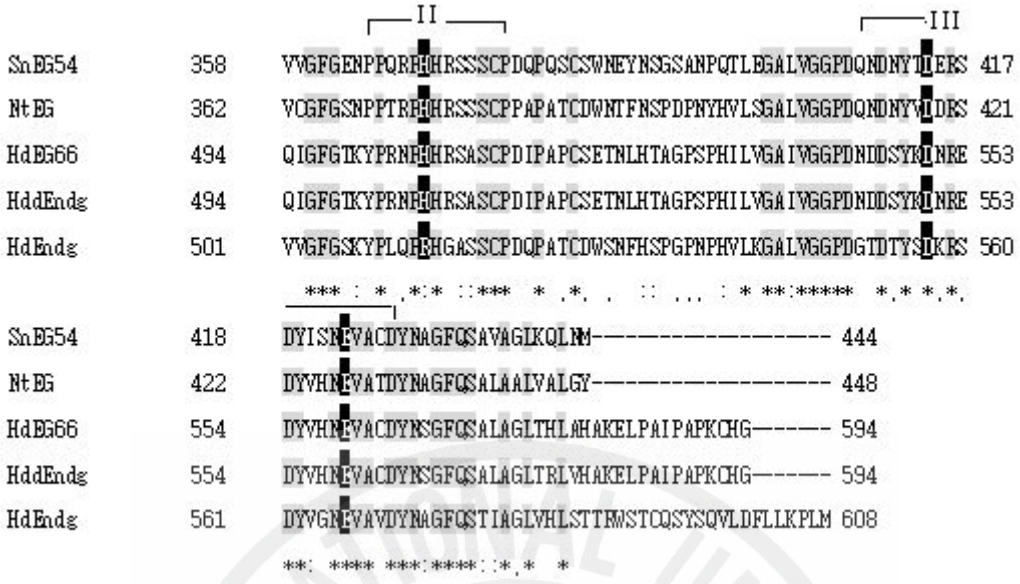



**Figure 3:** Multiple alignment of amino acid sequences for the *H. discus discus* HdEndg and other known species endoglucanases. The sequence of C-terminal 464 resides for the HdEndg was aligned with the sequences for catalytic domains of HdEG66, HdEndg, *Strongylocentrotus nudus* (SnEG54, accession BAF62178) and *Nasutitermes takasagoensis* (NtEG, accession AB013272). Residues participating in catalytic action as nucleophiles or proton donors in GHF9 enzymes are shaded in black. Consenses signatures for GHF9 are marked as I, II and III. Identical residues among the sequence are indicated by asterisks.

**Table 2:** Pairwise CLUSTALW analysis and comparison of HdEndg amino acid sequence with other known species endoglucanases.

| Species | Accession number | Identity (%) | Amino acids |
|---|---|---|---|
| Cellulase *Haliotis discus hannai* | BAC67186 | 50.3 | 594 |
| Cellulase *Haliotis discus discus* | BAD44734 | 49.6 | 594 |
| Cellulase *Corbicula japonica* | BAF38757 | 46.0 | 596 |
| Endoglucanase *Ampullaria crossean* | ABD24276 | 37.8 | 723 |
| Cellulase *Strongylocentrotus nudus* | BAF62178 | 37.0 | 444 |
| Endoglucanase *Eisenia andrei* | AAX92641 | 36.8 | 456 |
| Endoglucanase *Mastotermes darwiniensis* | CAD54729 | 36.1 | 448 |
| Endoglucanase *Nasutitermes takasagoensis* | AB013272 | 35.3 | 448 |
| Endoglucanase *Cherax quadricarinatus* | AF148497 | 35.0 | 448 |
| Endoglucanase *Nasutitermes walkeri* | AB013273 | 34.5 | 448 |

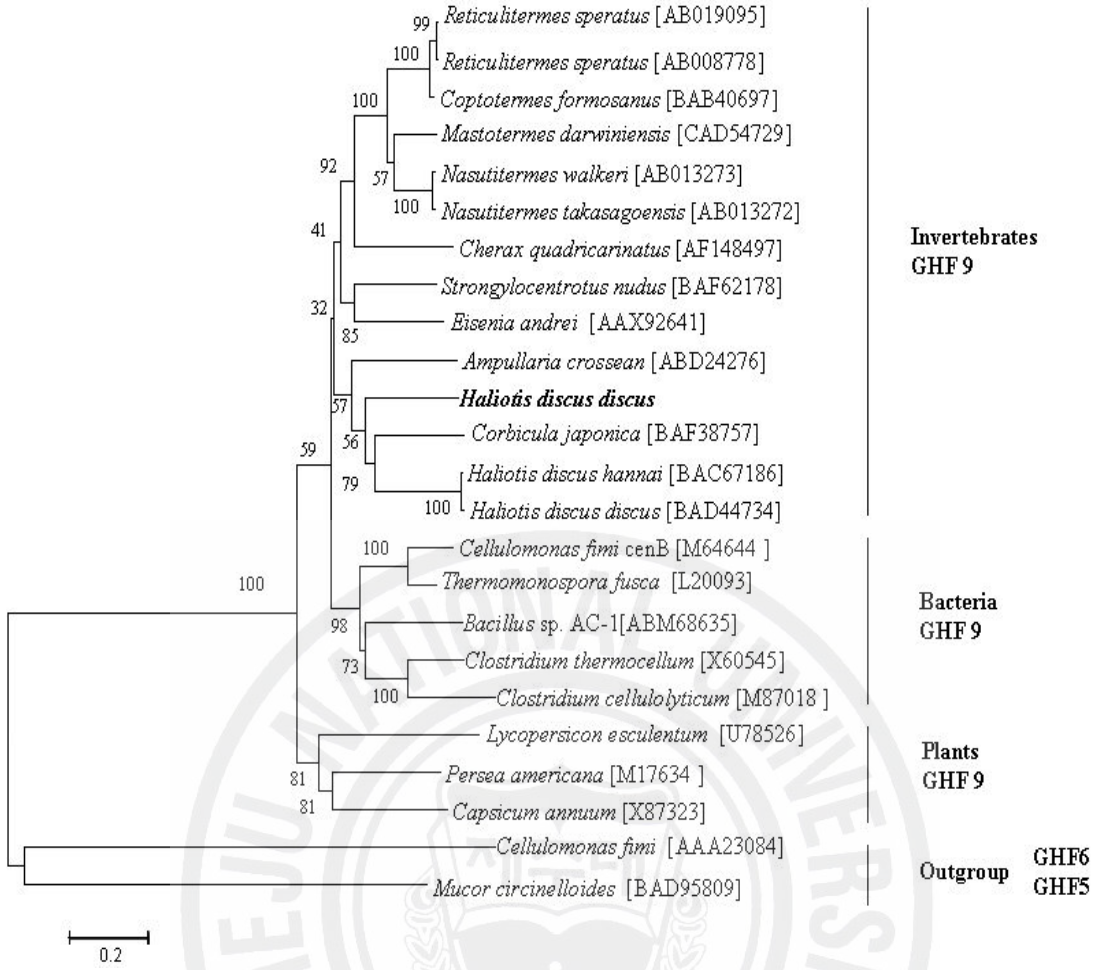



**Figure 4:** Phylogenetic analysis of HdEndg with other species GHF9 endoglucanases. The tree is based on an alignment corresponding to the full length amino acid sequences, using clustalW and MEGA (3.1). The numbers at the branches denotes the bootstrap majority consensus values on 1000 replicates. *Mucor circinelloides* and *Cellulomonas fimi* Endgs, which belongs to GHF 5 and 6, respectively were used as out groups for the analysis. The GenBank accession numbers are shown next to each species.

**Sequence characterization of disk abalone α-amaylase (HdAmy)**

The 1650 bp full length cDNA of HdAmy has 1536 bp ORF encoding 511 amino acids with putative 16 amino acids signal peptide (GenBank accession no. EF103352). Neucleotide and amino acid sequence of disk abalone is shown in Figure 5. The start (ATG) and stop codons (ATG) were found in nucleotide position at 18 and 96 of the full length cDNA represented in bold face. The predicted molecular mass mature protein of HdAmy was 54 kDa and estimated isoelectric point (p*I*) was 8.3. The active site residues (Asp214, Glu 250 and Asp315) as well as residues making up the calcium (Asn118, Arg175, Asp184, and His218) and chloride (Arg212, Asn313, Arg351) binding sites are fully conserved in HdAmy (Figure. 6). Also, the polyadenylation signal and poly A tail were identified in the 3' of the sequence. A pairwise alignment was generated by ClustalW and identity percentage of α-amylase amino acid to other known speceies is given in the Table 3. Disk abalone α-amylase showed the highest identity (66.5%) to *Pectan maximus* α-amylase. ClustalW multiple alignment results showed that the 8 cysteine residues involved in disulfide bond formation in mammalian amylases are well conserved in disk abalones as well as all species compared in this study.

**Phylogenetic analysis of HdAmy**

Phylogenetic analysis results showed that in general α-amylase genes were distinctly divided into two main clusters namely vertebrate and invertebrate (Figure 7). In invertebrate cluster, it was divided into arthropods and mollusks two sub clusters. The HdAmy was grouped with other mollusks sub cluster with *Pecten maximus* and *Crassostria gigas* suggesting that they all may have similar α-amylase function.

```
                                                          ATTGTAGCAACGTCTGCC     18
ATGATTCCTACCTGGTGTTTAGCGTTCCTGCTGCCAACGGCCCTAGCGTCGATGTACAGTGATCCCCAC     87
M--I--P--T--W--C--L--A--F--L--L--P--T--A--L--A--S--M--Y--S--D--P--H--     23
TGTTCAAGTGGTAGGAGTGCCATCACCCACCTGTTCGAGTGGACGTGGTCGGACATTGCGAAAGAATGT    156
C--S--S--G--R--S--A--I--T--H--L--F--E--W--T--W--S--D--I--A--K--E--C--     46
GAGAGGTTCCTTGGACCATATGGCTACTGTGGAGTTCAGATATCTCCTCCCAATGAGAACCGTGTAGTC    225
E--R--F--L--G--P--Y--G--Y--C--G--V--Q--I--S--P--P--N--E--N--R--V--V--     69
ACTAACCCAAGACGACCATGGTGGGAGAGGTACCAGCCTGTCAGCTACAAACTTCAGACGAGGAGCGGC    294
T--N--P--R--R--P--W--W--E--R--Y--Q--P--V--S--Y--K--L--Q--T--R--S--G--     92
TCAGAGCAGGACCTGAAGGACATGGTGTCCAGGTGCAACAAGGCCGGAGTCAGAATATATGCGGACACT    363
S--E--Q--D--L--K--D--M--V--S--R--C--N--K--A--G--V--R--I--Y--A--D--T--    115
GTGATTAATCACATGACCGGTGTTGGAGGTACGGGTCAAGGCACGGCTGGATCGTCGTACGACGCCAAC    432
V--I--N--H--M--T--G--V--G--G--T--G--Q--G--T--A--G--S--S--Y--D--A--N--    138
CAGCTGAAGTACCCAGGCGTTCCATACGGCCCAACCGACTTCAACGACAACTCCAACTGCCACACTGGC    501
Q--L--K--Y--P--G--V--P--Y--G--P--T--D--F--N--D--N--S--N--C--H--T--G--    161
GACATGCAAATTCATAACTACAATAACCCGGAAGAGGTGAGGAACTGCCGGTTGGTCGGGTTGGCGGAT    570
D--M--Q--I--H--N--Y--N--N--P--E--E--V--R--N--C--R--L--V--G--L--A--D--    184
CTGAAAGCCGGCAAGGATTACGTGAGGAGCGAGATTGAGGGTTACCTTAACCACCTGGTCGACATTGGA    639
L--K--A--G--K--D--Y--V--R--S--E--I--E--G--Y--L--N--H--L--V--D--I--G--    207
ATTGCTGGATTCAGAGTCGACGCTGCCAAACATATGTGGCCAGGGGACCTGACTGCTATATTCGGGAGT    708
I--A--G--F--R--V--D--A--A--K--H--M--W--P--G--D--L--T--A--I--F--G--S--    230
GTAAAGAATCTCCGTTCTGACGTGTTTGGGGGTGGCAAGAAGCCTTTCGTCTTCCAGGAGGTAATCGAC    777
V--K--N--L--R--S--D--V--F--G--G--G--K--K--P--F--V--F--Q--E--V--I--D--    253
ATGGGTGGTGAGCCGATCAAGGGAGACCAGTACCTTGGCAGTGGCAGGGTCACCAACTTCATTTTTGGT    846
M--G--G--E--P--I--K--G--D--Q--Y--L--G--S--G--R--V--T--N--F--I--F--G--    276
GCCAAGTTGGCTCAAGTCTTCAGGAAGCAAAACGCCATGAAGTATTTATCCAACTGGGGATCAGCTTGG    915
A--K--L--A--Q--V--F--R--K--Q--N--A--M--K--Y--L--S--N--W--G--S--A--W--    299
GGCATGTTGAAATCCGACGATTCTGTCGTCTTCATCGATAACCATGACAATCAGCGTGGTCATGGCGGC    984
G--M--L--K--S--D--D--S--V--V--F--I--D--N--H--D--N--Q--R--G--H--G--G--    322
GGCGGAGGTGTTCTCACCCACTTTGAGCCCCGTCCTTATAAGCTGGCTACAGCCTTCATGCTGGCCCAT   1053
G--G--G--V--L--T--H--F--E--P--R--P--Y--K--L--A--T--A--F--M--L--A--H--    345
CCATATGGTTTCCCCAGGGTGATGAGCTCATATAACTTCAACCAAGCCAACACTGACCAAGGACCCCCT   1122
P--Y--G--F--P--R--V--M--S--S--Y--N--F--N--Q--A--N--T--D--Q--G--P--P--    368
CAGAACGGCGACATGAGTACTAAGCCGGTAACCATCAATGGTATGGTCTGTGGTAACGGCTGGACGTGT   1191
Q--N--G--D--M--S--T--K--P--V--T--I--N--G--M--V--C--G--N--G--W--T--C--    391
GAGCACCGATGGCGCCAGATGTACAACATGGTTGCCTTCAGGAACATAGCGGGTTACTCCGGGTTATCA   1260
H--E--R--W--R--Q--M--Y--N--M--V--A--F--R--N--I--A--G--Y--S--G--L--S--    414
AATTGGTGGTCAGGCTCAGACTATCAGATAGCATTTTCCAGAGGAAACAAGGCCTTTATTGCCTTCAAC   1329
```

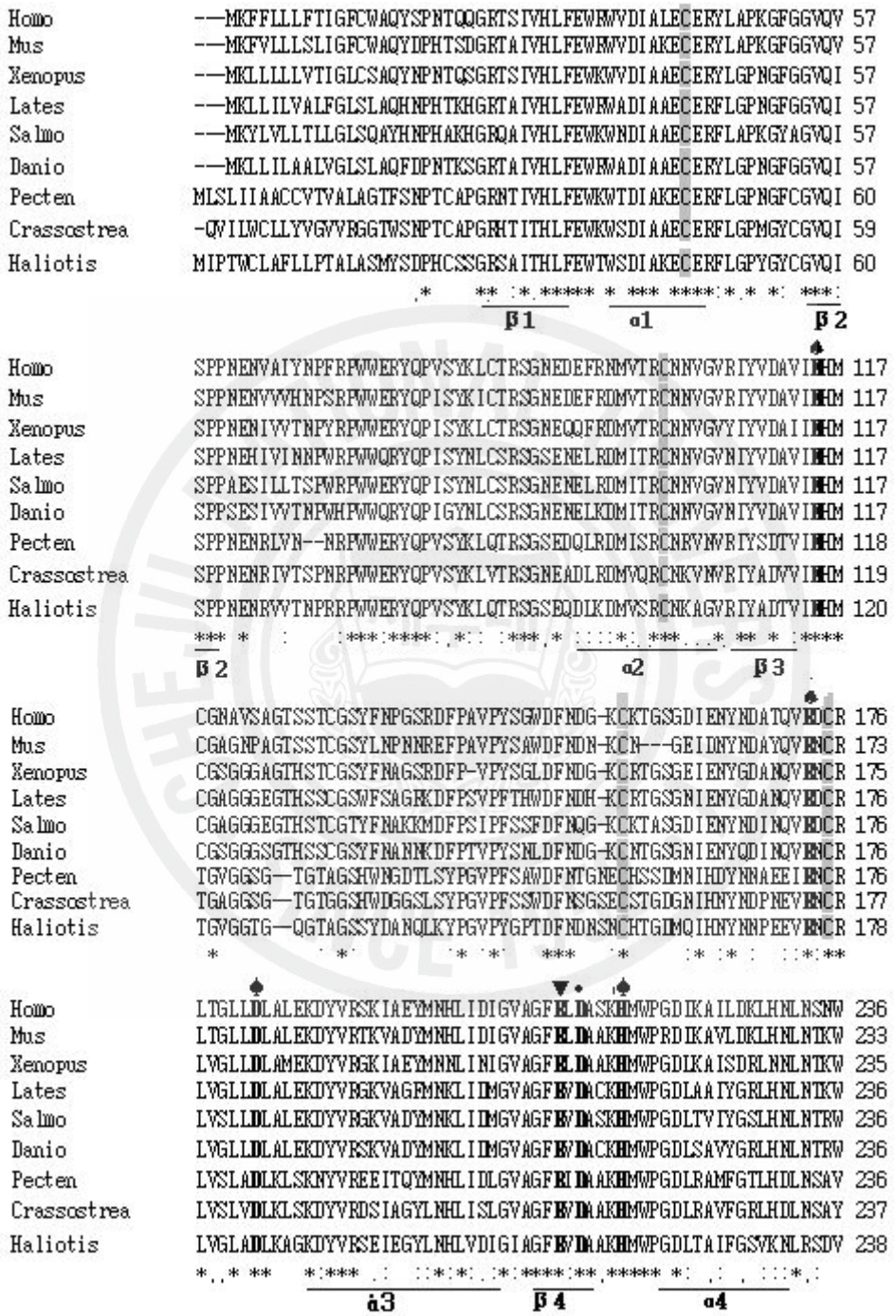
```
N--W--W--S--G--S--D--Y--Q--I--A--F--S--R--G--N--K--A--F--I--A--F--N--    437
CTCGAGGGATATGACCTTTCCAAATCCCTTAACACTGGACTTCCGTCTGGAAGCTATTGCGATGTCATT   1398
L--E--G--Y--D--L--S--K--S--L--N--T--G--L--P--S--G--S--Y--C--D--V--I--    460
TCTGGTAACCTCGAGAACGGCGGTTGTACTGGCAAGACAATTGACGTTGACGGAAGTGGCCATGCTTCG   1467
S--G--N--L--E--N--G--G--C--T--G--K--T--I--D--V--D--G--S--G--H--A--S--    483
ATTCATATCAGCAGCTCCGATAAAGATCCAATGGTGGCCATCCATGTGGGTGCGAAGAAGGGTTCGTCG   1536
I--H--I--S--S--S--D--K--D--P--M--V--A--I--H--V--G--A--K--K--G--S--S--    506
CCAAAGCGCGTTGGTTAGGAGCTGTAAATTGGTCAGGAGCTGTAATCTTACATGGAGACGTGCGAACGT   1605
P--K--R--V--G--*                                                        511
GAATAAACTGTTATTTCAAAAAAAAAAAAAAAAAAAAAAAAAAA                           1650
```

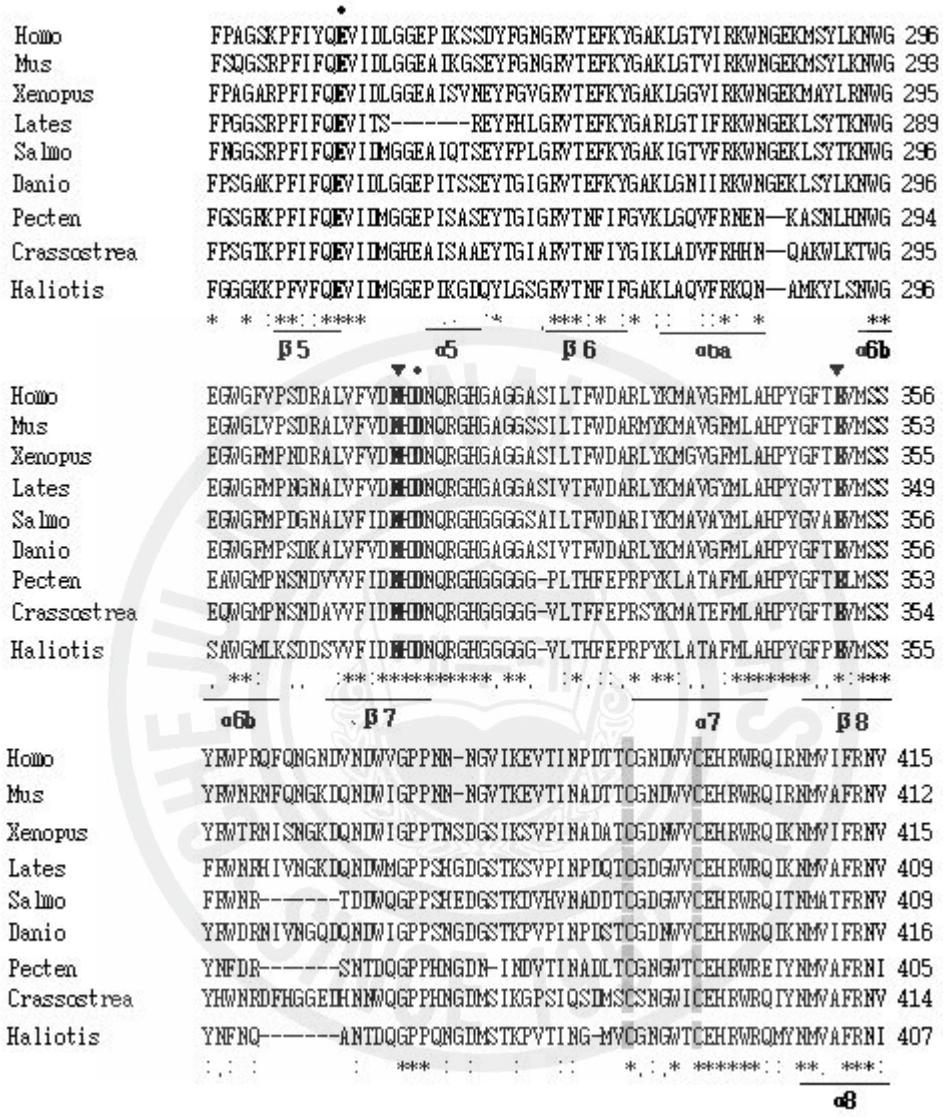
**Figure 5:** Full-length cDNA sequence and deduced amino acid sequence of HdAmy.

The start (ATG) and stop (TAA) codons are in bold. A putative signal peptide is indicated by dotted underline. Cysteine residues involved in disulfide bonds are shaded. Active site residues (Asp214, Glu250, Asp315) are shown in bold italics. The residues involved in calcium binding (Asn118, Arg175, Asp184, and His 218) are shown in bold, and the residues of the chloride binding site (Arg212, Asn313, Arg351) are shown in bold, double line boxed. Four conserved regions found in the α-amylase family are under lined. Putative polyadenylation signal is underlined and poly A tail is at the end.

**Table 3:** Pairwise CLUSTALW analysis and comparison of the deduced amino acid sequence of HdAmy with other known species α-amylases.

| Species | Accession number | Identity (%) | Amino acids |
|---|---|---|---|
| Alpha-amylase precursor *Pecten maximus* | P91778 | 66.5 | 508 |
| Alpha-amylase *Crassostrea gigas* | CAA69658 | 64.6 | 518 |
| Alpha amylase *Tribolium castaneum* | NP_001107848 | 57.1 | 490 |
| Alpha-amylase precursor *Xenopus laevis* | BC056841 | 56.5 | 511 |
| Alpha amylase pancreatic *Danio rerio* | BC062867 | 56 | 512 |
| Amylase 2, pancreatic *Mus musculus* | NM_009669 | 55.5 | 508 |
| Amylase, alpha 2B pancreatic *Sus scrofa* | NM_214195 | 54.4 | 511 |
| Amylase, alpha 2A pancreatic *Homo sapiens* | NM_000699 | 54.2 | 511 |
| Alpha amylase *Pseudopleuronectes americanus* | AF252633 | 54 | 512 |
| Alpha amylase *Salmo salar* | DQ331024 | 53.6 | 505 |
| Alpha-amylase *Drosophila melanogaster* | BAB32533 | 53.4 | 494 |
| Alpha-amylase *Lates calcarifer* | AF416651 | 53.2 | 505 |
| Amylase *Tetraodon nigroviridis* | AJ427289 | 52.6 | 513 |
| Amylase, alpha 2A; *Gallus gallus* | NP_001001473 | 50.0 | 512 |


Homo
Mus
Xenopus
Lates
Salmo
Danio
Pecten
Crassostrea
Haliotis
β1
α1
β2
β2
α2
β3
α3
β4
α4


Homo FPAGSKPFIYQEVIDLGGEPIKSSDYPGNGRVTEFKYGAKLGTVIRKWNGEKMSYLKNWG 296
Mus FSQGSRPFIFQEVIDLGGEAIKGSEYPGNGRVTEFKYGAKLGTVIRKWNGEKMSYLKNWG 298
Xenopus FPAGARPFIFQEVIDLGGEAISVNEYPGVGRVTEFKYGAKLGGVIRKWNGEKMAYLRNWG 295
Lates FPGGSRPFIFQEVITS------REYFHLGRVTEFKYGARLGTIFRKWNGEKLSYTKNWG 289
Salmo FNGGSRPFIFQEVIDMGGEAIQTSEYFPLGRVTEFKYGAKIGTVFRKWNGEKLSYTKNWG 296
Danio FPSGAKPFIFQEVIDLGGEPITSSEYTGIGRVTEFKYGAKLGNIIRKWNGEKLSYLKNWG 296
Pecten FGSGRKPFIFQEVIDMGGEPISASEYTGIGRVTNFIPGVKLGQVFRNEN—KASNLHNWG 294
Crassostrea FPSGTKPFIFQEVIDMGHEAISAAEYTGIARVTNFIYGIKLADVFRHHN—QAKWLKTWG 295
Haliotis FGGGKKPFVFQEVIDMGGEPIKGDQYLGSGRVTNFIFGAKLAQVFRKQN—AMKYLSNWG 296
β5 α5 β6 αba α6b
Homo EGWGFVPSDRALVFVDMHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGFTMVMSS 356
Mus EGWGLVPSDRALVFVDMHDNQRGHGAGGSSILTFWDARMYKMAVGFMLAHPYGFTMVMSS 353
Xenopus EGWGFMPNDRALVFVDMHDNQRGHGAGGASILTFWDARLYKMGVGFMLAHPYGFTMVMSS 355
Lates EGWGFMPNGNALVFVDMHDNQRGHGAGGASIVTFWDARLYKMAVGYMLAHPYGVTMVMSS 349
Salmo EGWGFMPDGNALVFIDMHDNQRGHGGGGSAILTFWDARIYKMAVAYMLAHPYGVAMVMSS 356
Danio EGWGFMPSDKALVFVDMHDNQRGHGAGGASIVTFWDARLYKMAVGFMLAHPYGFTMVMSS 356
Pecten EAWGMPNSNDVVVFIDMHDNQRGHGGGGG-PLTHFEPRPYKLATAFMLAHPYGFTMLMSS 353
Crassostrea EQWGMPNSNDAVVFIDMHDNQRGHGGGGG-VLTFFEPRSYKMATEFMLAHPYGFTMVMSS 354
Haliotis SAWGMLKSDDSVVFIDMHDNQRGHGGGGG-VLTHFEPRPYKLATAFMLAHPYGFPMVMSS 355
α6b β7 α7 β8
Homo YMWPRQFQNGNDVNDWVGPPNN-NGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNV 415
Mus YMWNRNFQNGKDQNDWIGPPNN-NGVTKEVTINADTTCGNDWVCEHRWRQIRNMVAFRNV 412
Xenopus YMWTRNISNGKDQNDWIGPPTNSDGSIKSVPINADATCGDNWVCEHRWRQIKNMVIFRNV 415
Lates FMWNRHIVNGKDQNDWMGPPSHGDGSTKSVPINPDQTCGDGWVCEHRWRQIKNMVAFRNV 409
Salmo FMWNR------TDDWQGPPSHEDGSTKDVHVNADDTCGDGWVCEHRWRQITNMATFRNV 409
Danio YMWDRNIVNGQDQNDWIGPPSNGDGSTKPVPINPDSTCGDNWVCEHRWRQIKNMVIFRNV 416
Pecten YNFDR------SNTDQGPPHNGDN-INDVTINADLTCGNGWTCEHRWREIYNMVAFRNI 405
Crassostrea YMWNRDFHGGEDHNNWQGPPHNGDMSIKGPSIQSDMSCSNGWICEHRWRQIYNMVAFRNV 414
Haliotis YNFNQ------ANTDQGPPQNGDMSTKPVTING-MVCGNGWTCEHRWRQMYNMVAFRNI 407
α8

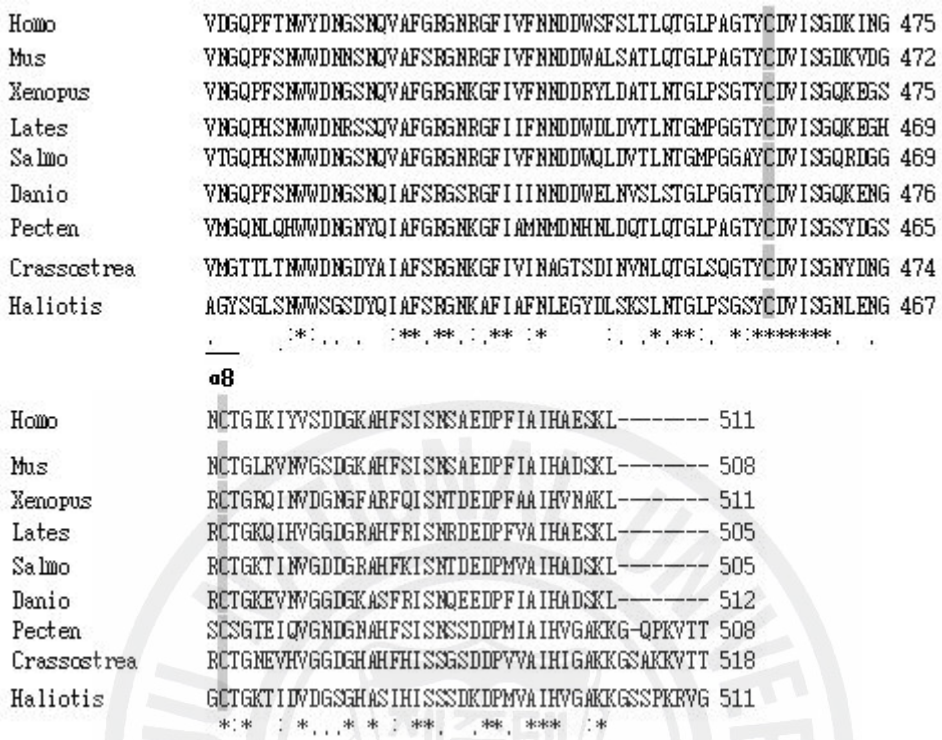



**Figure 6:** Multiple alignment of HdAmy amino acid sequence with known species α-amylases. Identical residues in all sequences are indicated by (*) under the column, conserved substitutions are indicated by (:), and semi-conserved substitutions are indicated by (.). Deletions are indicated by dashes. Cysteine residues are shaded. Active site residues (Asp214, Glu250, Asp315) are bold and labeled with (•) over the column. Calcium binding residues (Asn118, Arg175, Asp184, and His218) are shown in bold and labeled (♠), and chloride binding site residues (Arg212, Asn313, Arg351) are shown in bold and labeled (♦). Positions of α-helices and β-sheets are underlined. Acession numbers of the amino acid sequences are same as in figure 9.

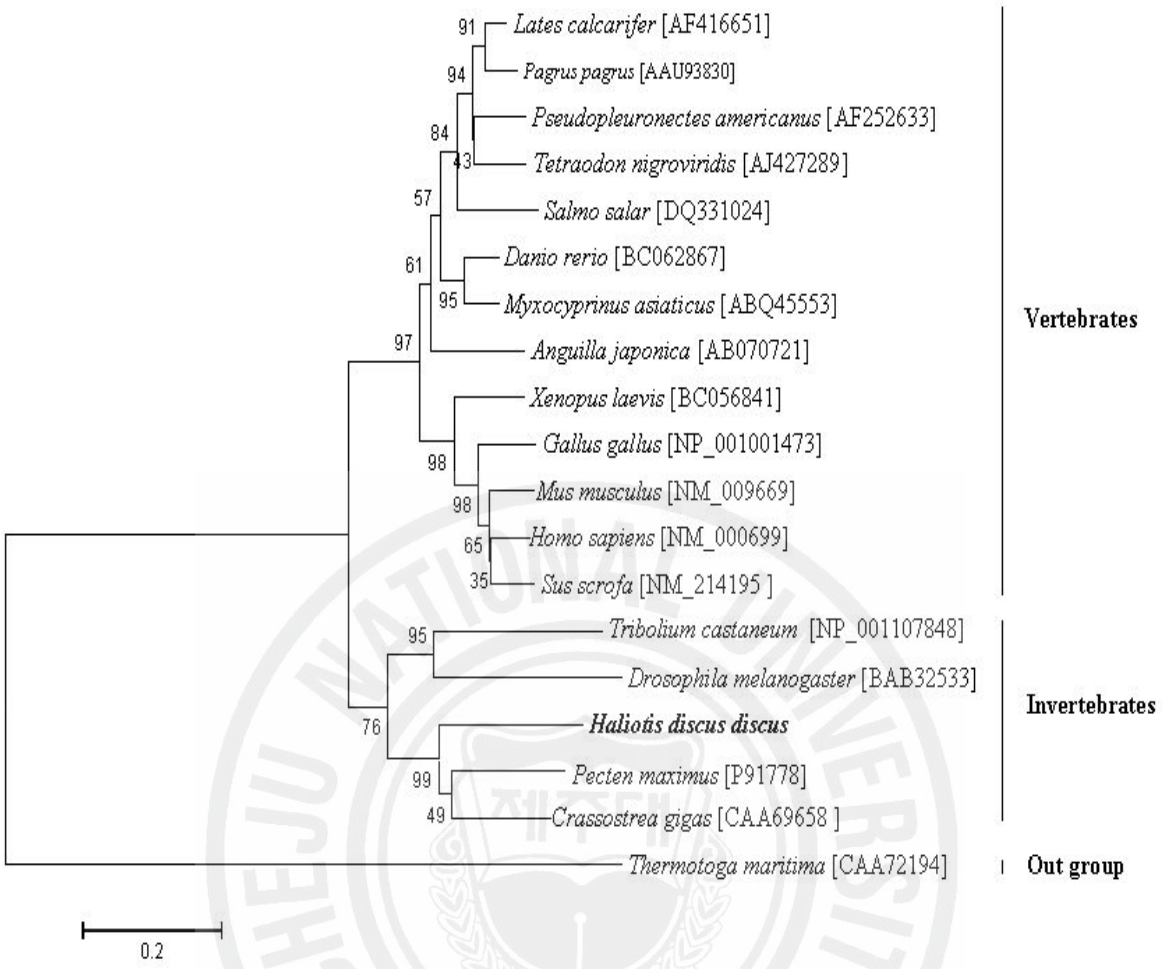



**Figure 7:** Phylogenetic analysis of HdAmy with other speceies α-amylases. The tree is based on an alignment corresponding to the full length amino acid sequences, using clustalW and MEGA (3.1). The numbers at the branches denotes the bootstrap majority consensus values on 1000 replicates. *Thermotoga maritime* α-amylase was used as out groups for the analysis. The GenBank accession numbers are shown next to each speceis.

**Sequence characterization of disk abalone aliginate lyase (HdAlgl)**

In this study, one EST clone was identified from disk abalone digestive gland cDNA library that showed similarity to alginate lyase. The nucleotide and deduced amino acid sequences are shown in Figure 8 and nucleotide sequence was deposited in the GenBank under accession number EF103359. It showed 897 bp ORF codes for 298 amino acids with 16 amino acid signal peptide. The molecular mass and predicted isoelectric point were 33 kDa and 6.7, respectively. The BLAST search using deduced amino acid sequence of alginate lyase revealed that our sequence was most similar to alginate lyase of *H. discus hannai* (accession no: BAC67186). Pairwise amino acid comparison results showed that HdAlgl has lower identity (12%) to *H. discus hannai.* Moreover, pairwise comparison results showed that identity of C-terminal HdAlgl (140-199 aa) to N-terminal *H. discus hannai* (15-74 aa) and *Turbo* SP2 (1-57 aa) sequences were 47 and 41%, respectively (Figure 9). Relatively high conservative regions, which present in other alginate lyases were also present in HdAlgl. Disulfide bond prediction results showed that there were 4 cystein residues responsible for forming 2 disulfide bonds, positions at $C^{34}$-$C^{92}$ and $C^{264}$-$C^{273}$. Further more, one N-linked glycosylation site ($^{257}$NVSQ$^{260}$) was identified in HdAlgl.

**Phylogenetic analysis of HdAlgl**

To determine the evolutional relationship of HdAlgl with other known alginate lyases, 7 representatives HdAlgl were analyzed to construct a phylogenetic tree (Figure 10). Phylogenetic tree showed two main clusters including mollusks and *Chlorella* virus alginate lyase in one cluster and bacterial alginate lyase in second cluster. HdAlgl was positioned in mollusk subcluster and *Chlorella* virus alginate lyase was sub clustered as a single species within the main cluster. Therefore, phylogenetic analysis provides evidence that the HdAlgl has been derived from the common ancestor with other alginate lyases, however, further identifications are needed to get clear identification for the position of mollusk alginate lyases in evolution.

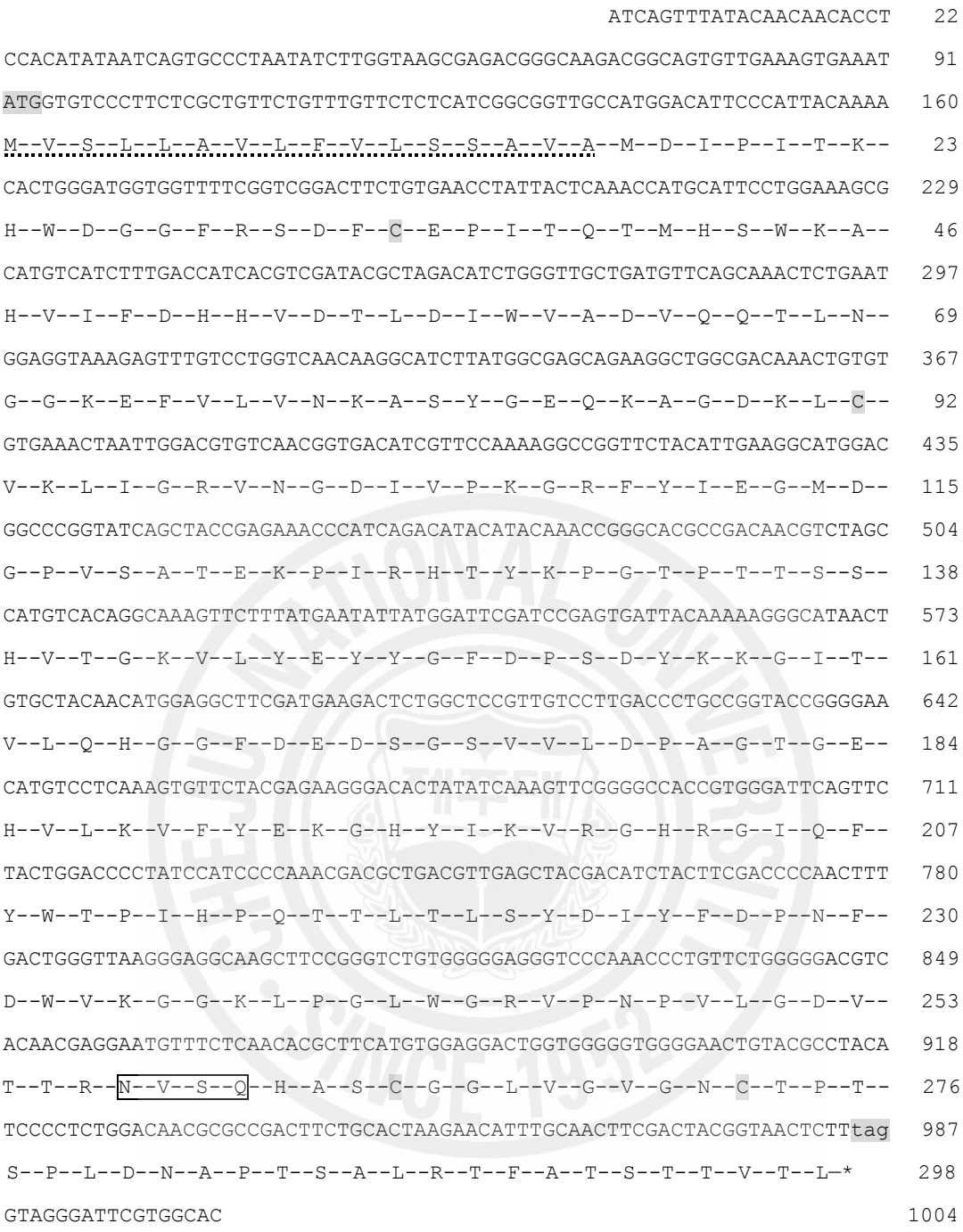



**Figure 8:** Nucleotide and deduced amino-acid sequences of HdAlgl. The start (ATG) and stop (TAA) codons are shaded. A putative signal peptide is indicated by a dotted underline. N-glycosylation site is boxed. Cysteine residues involved in disulfide bonds are shaded.

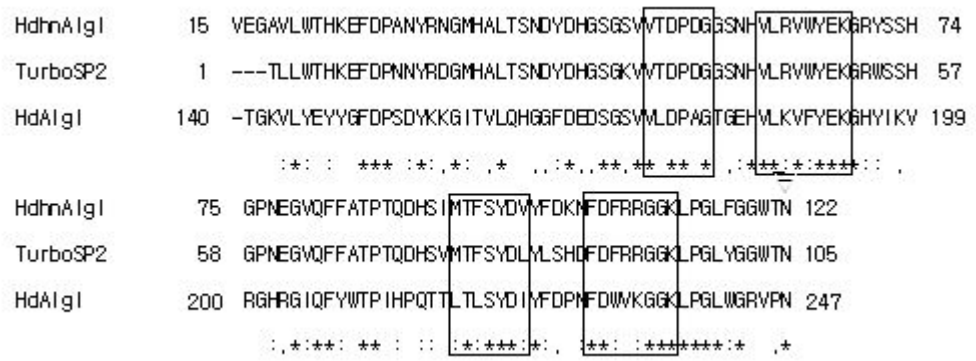



**Figure 9:** Comparison of amino acid sequences of alginate lyases. The amino acid sequence of HdAlgl (from 140-247 aa) was aligned with those *Turbo* SP2 (from 1-105 aa) and the *H. discus hannai* (HdhnAlgl) (from 15-122). Relatively high conservative regions among the sequences are boxed.

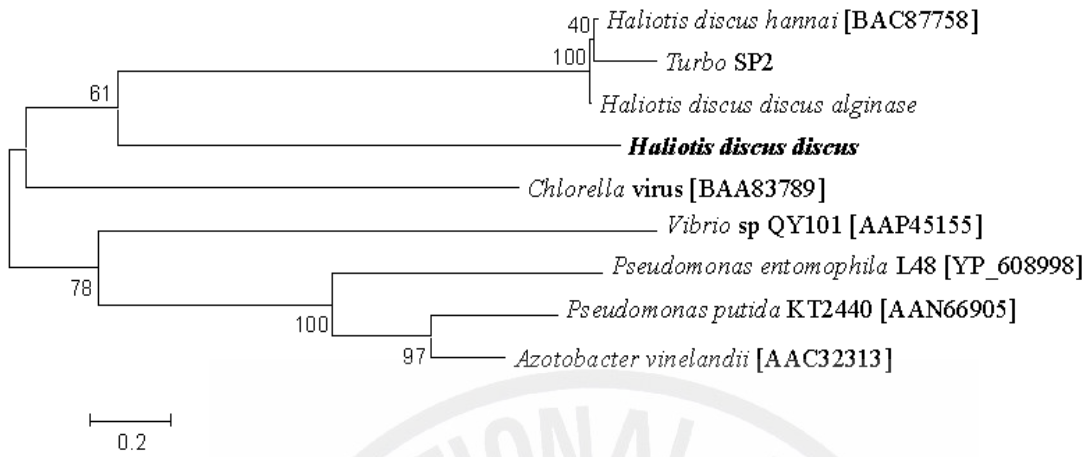



**Figure 10:** Phylogenetic analysis of HdAlgl with other known species alginate lyases. The tree is based on an alignment corresponding to the full length amino acid sequences, using clustalW and MEGA (3.1). The numbers at the branches denotes the bootstrap majority consensus values on 1000 replicates. The GenBank accession numbers are shown next to each species.

**Sequence characterization of disk abalone β-mannanase (HdMann)**

After successive sequencing reaction using forward primer, 1332 bp of full length nucleotide sequence of disk abalone β-mannanase was determined (Figure 11) and named as HdMann. The translational initiation codon, ATG, was found in nucleotide positions at 91–93, and the termination codon, TAA, in positions at 1217-1219. In 3' terminal region, a putative polyadenylation signal sequence, AATAAA, and a poly ($A^+$) tail were found. HdMann cDNA contains an ORF (1125bp) that encodes 375 amino acids, of which the first 18 residues comprise the secretion signal peptide. The calculated molecular mass of the 249-residue mature protein is 28 kDa and predicted isoelectric point was 5.6. Two potential N-glycolysation sites ($^{198}$NATH$^{201}$ and $^{362}$NNSD$^{365}$) were identified using PROSITE, program. Further, motif scan analysis showed that HdMann contained a domain belongs to glycoside hydrolase family 5 (GHF5). According to searches for the NCBI, GenBank databases, the amino acid sequence of HdMann was found to show 49% sequence identity with ManA from *M. edulis* (accession number Q8WPJ2), 48% with *H. discus hannai* (BAE78456) and *B. glabrata* β-mannanase (AAV91523). HdMann sequence was aligned with those of *M. edulis*, *B. glabrata* and *H. discus hannai* β-mannanases (Figure 12) and it was observed that residues participating in catalytic action as proton donor or nucleophile in GHF5 enzymes were conserved in HdMann at E183 (putative catalytic acid/base) and E310 (putative catalytic nucleophile).

**Phylogenetic analysis of HdMann**

Phylogenetic tree shows main two clusters, mollusks and bacteria and HdMann forms a monophyletic clade in the mollusk cluster (Figure 13). The fungi *Aspergillus clavatus* was used as an out group for this study.

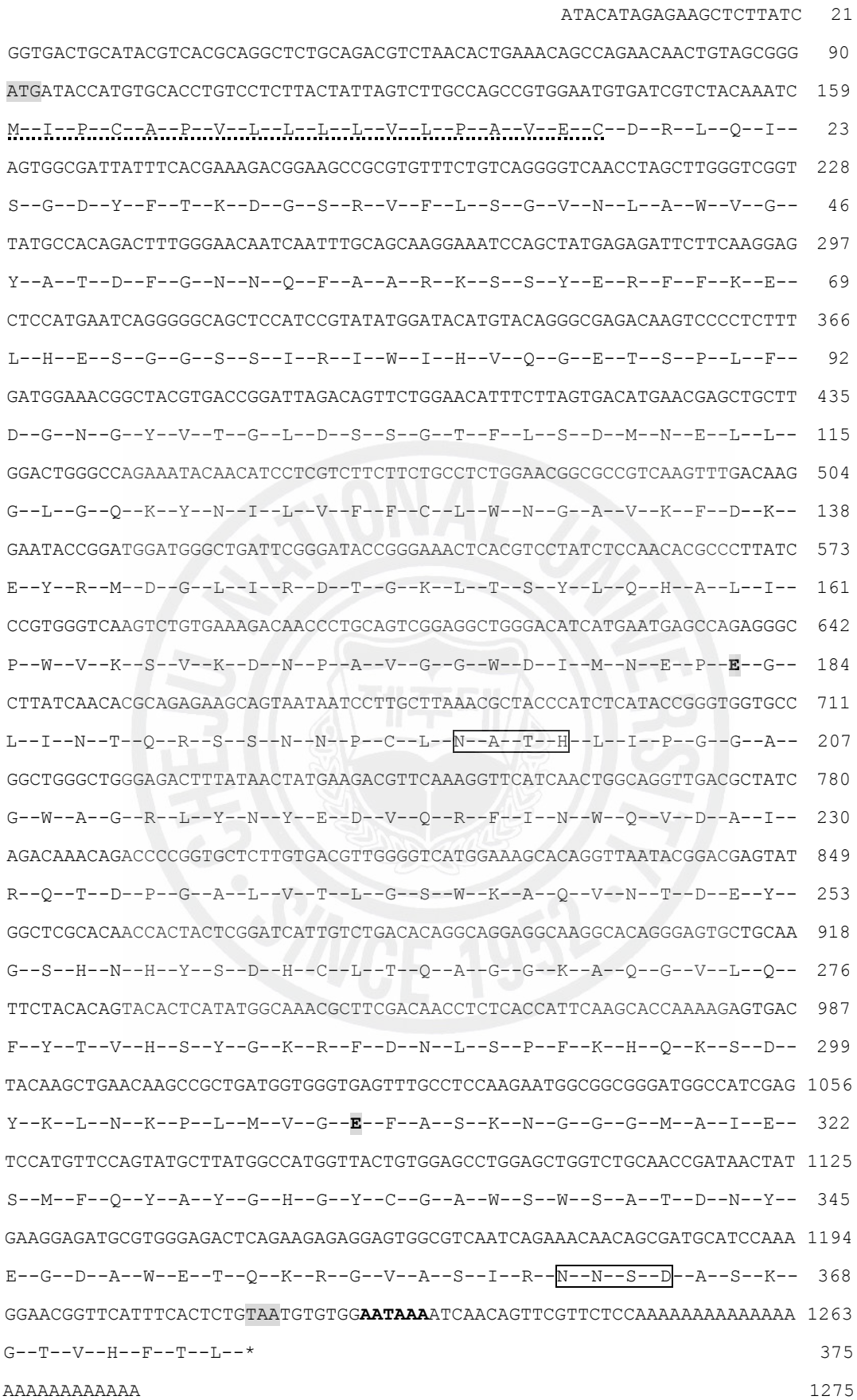

ATACATAGAGAAGCTCTTATC 21
GGTGACTGCATACGTCACGCAGGCTCTGCAGACGTCTAACACTGAAACAGCCAGAACAACTGTAGCGGG 90
ATGATACCATGTGCACCTGTCCTCTTACTATTAGTCTTGCCAGCCGTGGAATGTGATCGTCTACAAATC 159
M--I--P--C--A--P--V--L--L--L--L--V--L--P--A--V--E--C--D--R--L--Q--I-- 23
AGTGGCGATTATTTCACGAAAGACGGAAGCCGCGTGTTTCTGTCAGGGGTCAACCTAGCTTGGGTCGGT 228
S--G--D--Y--F--T--K--D--G--S--R--V--F--L--S--G--V--N--L--A--W--V--G-- 46
TATGCCACAGACTTTGGGAACAATCAATTTGCAGCAAGGAAATCCAGCTATGAGAGATTCTTCAAGGAG 297
Y--A--T--D--F--G--N--N--Q--F--A--A--R--K--S--S--Y--E--R--F--F--K--E-- 69
CTCCATGAATCAGGGGGCAGCTCCATCCGTATATGGATACATGTACAGGGCGAGACAAGTCCCCTCTTT 366
L--H--E--S--G--G--S--S--I--R--I--W--I--H--V--Q--G--E--T--S--P--L--F-- 92
GATGGAAACGGCTACGTGACCGGATTAGACAGTTCTGGAACATTTCTTAGTGACATGAACGAGCTGCTT 435
D--G--N--G--Y--V--T--G--L--D--S--S--G--T--F--L--S--D--M--N--E--L--L-- 115
GGACTGGGCCAGAAATACAACATCCTCGTCTTCTTCTGCCTCTGGAACGGCGCCGTCAAGTTTGACAAG 504
G--L--G--Q--K--Y--N--I--L--V--F--F--C--L--W--N--G--A--V--K--F--D--K-- 138
GAATACCGGATGGATGGGCTGATTCGGGATACCGGGAAACTCACGTCCTATCTCCAACACGCCCTTATC 573
E--Y--R--M--D--G--L--I--R--D--T--G--K--L--T--S--Y--L--Q--H--A--L--I-- 161
CCGTGGGTCAAGTCTGTGAAAGACAACCCTGCAGTCGGAGGCTGGGACATCATGAATGAGCCAGAGGGC 642
P--W--V--K--S--V--K--D--N--P--A--V--G--G--W--D--I--M--N--E--P--E--G-- 184
CTTATCAACACGCAGAGAAGCAGTAATAATCCTTGCTTAAACGCTACCCATCTCATACCGGGTGGTGCC 711
L--I--N--T--Q--R--S--S--N--N--P--C--L--N--A--T--H--L--I--P--G--G--A-- 207
GGCTGGGCTGGGAGACTTTATAACTATGAAGACGTTCAAAGGTTCATCAACTGGCAGGTTGACGCTATC 780
G--W--A--G--R--L--Y--N--Y--E--D--V--Q--R--F--I--N--W--Q--V--D--A--I-- 230
AGACAAACAGACCCCGGTGCTCTTGTGACGTTGGGGTCATGGAAAGCACAGGTTAATACGGACGAGTAT 849
R--Q--T--D--P--G--A--L--V--T--L--G--S--W--K--A--Q--V--N--T--D--E--Y-- 253
GGCTCGCACAACCACTACTCGGATCATTGTCTGACACAGGCAGGAGGCAAGGCACAGGGAGTGCTGCAA 918
G--S--H--N--H--Y--S--D--H--C--L--T--Q--A--G--G--K--A--Q--G--V--L--Q-- 276
TTCTACACAGTACACTCATATGGCAAACGCTTCGACAACCTCTCACCATTCAAGCACCAAAAGAGTGAC 987
F--Y--T--V--H--S--Y--G--K--R--F--D--N--L--S--P--F--K--H--Q--K--S--D-- 299
TACAAGCTGAACAAGCCGCTGATGGTGGGTGAGTTTGCCTCCAAGAATGGCGGCGGGATGGCCATCGAG 1056
Y--K--L--N--K--P--L--M--V--G--E--F--A--S--K--N--G--G--G--M--A--I--E-- 322
TCCATGTTCCAGTATGCTTATGGCCATGGTTACTGTGGAGCCTGGAGCTGGTCTGCAACCGATAACTAT 1125
S--M--F--Q--Y--A--Y--G--H--G--Y--C--G--A--W--S--W--S--A--T--D--N--Y-- 345
GAAGGAGATGCGTGGGAGACTCAGAAGAGAGGAGTGGCGTCAATCAGAAACAACAGCGATGCATCCAAA 1194
E--G--D--A--W--E--T--Q--K--R--G--V--A--S--I--R--N--N--S--D--A--S--K-- 368
GGAACGGTTCATTTCACTCTGTAATGTGTGGAATAAAATCAACAGTTCGTTCTCCAAAAAAAAAAAAAA 1263
G--T--V--H--F--T--L--* 375
AAAAAAAAAAAA 1275

**Figure 11:** Nucleotide and deduced amino-acid sequences of HdMann. The translational initiation codon ATG, termination codon TAA are shaded and a putative polyadenylation signal AATAA are bold. A putative signal peptide is indicated by a dotted underline. N-glycosylation sites are boxed. Residues participating in catalytic action as nucleophiles or proton donors in GHF5 are bold and shaded. Residue numbers for nucleotide and amino-acid are indicated in the right of each row.

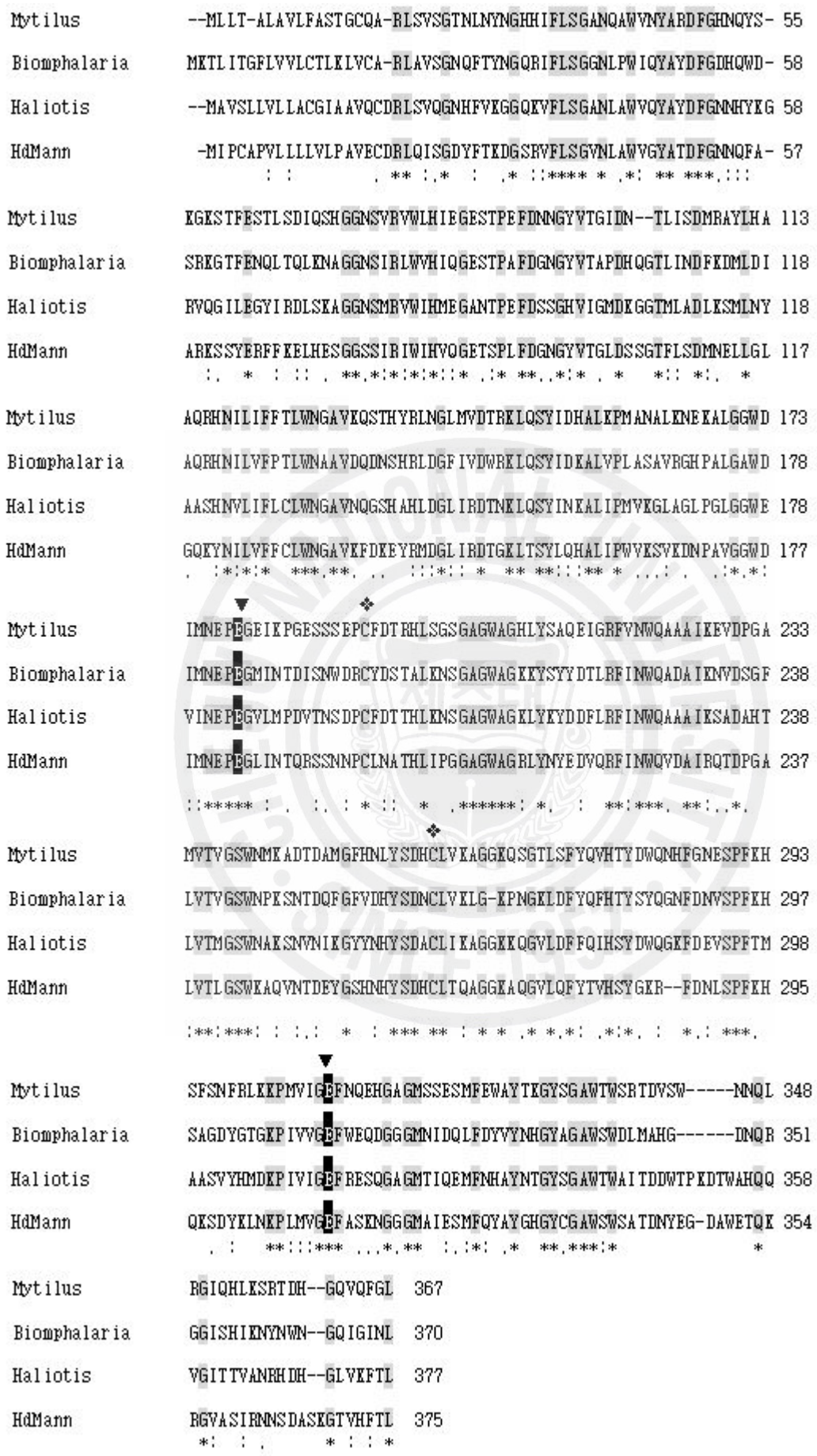

Mytilus --MLLT-ALAVLFASTGCQA-RLSVSGTNLNYNGHHIFLSGANQAWVNYARDFGHNQYS- 55
Biomphalaria MKTLITGFLVVLCTLKLVCA-RLAVSGNQFTYNGQRIFLSGGNLPWIQYAYDFGDHQWD- 58
Haliotis --MAVSLLVLLACGIAAVQCDRLSVQGNHFVKGGQKVFLSGANLAWVQYAYDFGNNHYKG 58
HdMann -MIPCAPVLLLLVLPAVECDRLQISGDYFTKDGSRVFLSGVNLAWVGYATDFGNNQFA- 57
Mytilus KGKSTFESTLSDIQSHGGNSVRVWLHIEGESTPEFDNNGYVTGIDN--TLISDMRAYLHA 113
Biomphalaria SREGTFENQLTQLKNAGGNSIRLWVHIQGESTPAFDGNGYVTAPDHQGTLINDFKDMLDI 118
Haliotis RVQGILEGYIRDLSKAGGNSMRVWIHMEGANTPEFDSSGHVIGMDKGGTMLADLKSMLNY 118
HdMann ARKSSYERFFKELHESGGSSIRIWIHVQGETSPLFDGNGYVTGLDSSGTFLSDMNELLGL 117
Mytilus AQRHNILIFFTLWNGAVKQSTHYRLNGLMVDTRKLQSYIDHALKPMANALKNEKALGGWD 173
Biomphalaria AQRHNILVFPTLWNAAVDQDNSHRLDGFIVDWRKLQSYIDKALVPLASAVRGHPALGAWD 178
Haliotis AASHNVLIFLCLWNGAVNQGSHAHLDGLIRDTNKLQSYINKALIPMVKGLAGLPGLGGWE 178
HdMann GQKYNILVFFCLWNGAVKFDKEYRMDGLIRDTGKLTSYLQHALIPWVKSVKDNPAVGGWD 177
Mytilus IMNEFEGGEIKPGESSSEPCFDTRHLSGSGAGWAGHLYSAQEIGRFVNWQAAAIKEVDPGA 233
Biomphalaria IMNEFEGGMINTDISNWDRCYDSTALKNSGAGWAGKKYSYYDTLRFINWQADAIKNVDSGF 238
Haliotis VINEFEGGVLMPDVTNSDPCFDTTHLKNSGAGWAGKLYKYDDFLRFINWQAAAIKSADAHT 238
HdMann IMNEFEGGLINTQRSSNNPCLNATHLIPGGAGWAGRLYNYEDVQRFINWQVDAIRQTDPGA 237
Mytilus MVTVGSWNMKADTDAMGFHNLYSDHCLVKAGGKQSGTLSFYQVHTYDWQNHFGNESPFKH 293
Biomphalaria LVTVGSWNPKSNTDQFGFVDHYSDNCLVKLG-KPNGKLDFYQFHTYSYQGNFDNVSPFKH 297
Haliotis LVTMGSWNAKSNVNIKGYYNHYSDACLIKAGGKKQGVLDFFQIHSYDWQGKFDEVSPFTM 298
HdMann LVTLGSWKAQVNTDEYGSHNHYSDHCLTQAGGKAQGVLQFYTVHSYGKR--FDNLSPFKH 295
Mytilus SFSNFRLKKPMVIGEFNQEHGAGMSSESMFEWAYTKGYSGAWTWSRTDVSW-----NNQL 348
Biomphalaria SAGDYGTGKPIVVGEFWEQDGGGMNIDQLFDYVYNHGYAGAWSWDLMAHG------DNQR 351
Haliotis AASVYHMDKPIVIGEFRESQGAGMTIQEMFNHAYNTGYSGAWTWAITDDWTPKDTWAHQQ 358
HdMann QKSDYKLNKPLMVGEFASKNGGGMAIESMFQYAYGHGYCGAWSWSATDNYEG-DAWETQK 354
Mytilus RGIQHLKSRTDH--GQVQFGL 367
Biomphalaria GGISHIKNYNWN--GQIGINL 370
Haliotis VGITTVANRHDH--GLVKFTL 377
HdMann RGVASIRNNSDASKGTVHFTL 375

**Figure 12:** Alignment of amino-acid sequences for HdMann and other known β -mannanases. The amino nacid sequence of HdMan is aligned with those of *M. edulis* (Q8WPJ2), *B. glabrata* (AAV91523), and *H. discus hannai* (BAE78456). The positions of identical residues, highly conserved substitutions, conservative substitutions, and gaps are indicated by asterisk (*), colon (:), dot (.), and dash (-), respectively. Residues participating in catalytic action as nucleophiles or proton donors in GHF5 enzymes, i.e., E183 and E310 are shaded. Cystein residues which responsible for di sulfide bond formation is marked in (❖).

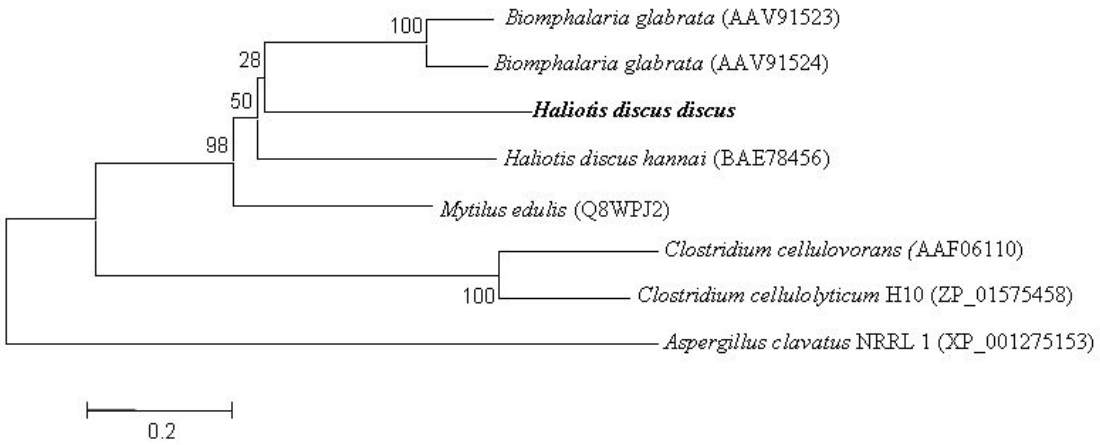



**Figure 13:** Phylogenetic analysis of HdMann with known species β-mannanase. The tree is based on an alignment corresponding to the full length amino acid sequences, using clustalW and MEGA (3.1). The numbers at the branches denotes the bootstrap majority consensus values on 1000 replicates. *Aspergillus clavatus* β-mannanase was used as out groups for the analysis. The GenBank accession numbers are shown next to each species.

**Sequence characterization of HdLms**

One EST clone was identified from disk abalone digestive gland cDNA library that showed similarity to laminarinase (β-1,3-endoglucanase). From the 903 bp partial sequence of abalone laminarinase, ORF having 378 bp that codes for 126 amino acids sequence was identified. The nucleotide and deduced amino acid sequences are shown in Figure 14. The calculated molecular mass and predicted isoelectric point were 14 kDa and 8.9, respectively. The BLAST search using deduced amino acid sequence of HdLms revealed that our sequence was similar to 1,3-endo β- glucanase of *Mizuhopecten yessoens* (AAW34372) and coelomic cytolytic factor 1 of *Eisenia fetida* (AF030028) with 15% identitiy. Disulfide bond prediction results showed that there are 2 cystein residues responsible for forming one disulfide bonds, positions at C46-C54.


**Figure 14:** The nucleotide and deduced amino acid sequences of the HdLms. The translational start codon ATG, termination codon TGA are bold and shaded. Cysteine residues involved in disulfide bonds are shaded. Residue numbers for both nucleotide and amino acid are indicated in the right of each row

**Sequence characterization of disk abalone arylsufatase (HdArys)**

The EST clone, which was identified from the disk abalone normalized cDNA library showed similarity to known sulfatase family proteins. The full length sequence was deposited in the NCBI under accession number of EF103354. The HdArys full length nucleotide and deduced amino acid sequence is shown in Figure 15. HdArys cDNA contains 1449bp an ORF that encodes 481 amino acids with first 21 residues comprise the secretion signal peptide. The calculated molecular mass and isoelectric point were 52 kDa and 5.8, respectively. The protein consists of a sulfatase domain (24-433 aa) which was analyzed by motif scan analysis. The characteristic sulfatase signature motif, was present at $^{72}$CSPSR$^{76}$. Other functionally important regions, as those containing residues involved in substrate binding and activation (Lys126, Tyr213, Lys289) and metal coordination (Asp34, Asp35, Asp271, Asn272) and enzyme stabilization (Arg76 and His128) were observed in the sequence. Five potential N-glycolysation sites ($^{106}$NITI$^{109}$, $^{169}$NNTT$^{172}$, $^{178}$NGTY$^{181}$, $^{198}$NQSQ$^{201}$, $^{250}$NVTK$^{253}$) were identified using PROSITE, program. Interestingly, leucine zipper pattern motif was observed at the amino acid position from 96-117. The HdArys full length consisted of 5' untranslated region (UTR) of 44 bp and 3'UTR of 70 bp. At 24 bp downstream of the stop codon, the polyadenylation cleavage motif was observed, followed by a poly A tail. In ClustalW pair wise amino acid sequence alignment, HdArys showed the highest level amino acid sequence identity (45%) with *Helix pomatia* sulfatase I precursor while it shared 41% identity with *Rattus norvegicus, Bos taurus* and *Homo sapiens* arylsulfatase B. The amino acid identity of HdArys with other known species is shown in the Table 4. Additionally, pairwise amino acid comparison was conducted for *Nematostella vectensis* predicted protein, predicted similar to arylsulfatase B of *Strongilocentrus perpu*, and similar to arylsulfatase I of *Danio rerio* and showed 42, 38 and 36% identity, respectively. ClustalW multiple alignment results with selected vertebrate and invertebrates shown in Figure16 and the characteristic conserved sulfatase signature motif and other important residues of the active site of sulfatases shared

conserved or similar residues at the same positions in different organisms. Furthermore, the catalytic domain of known members of the sulfatases family and the amino acid forming part of the active site of HdArys is conserved.

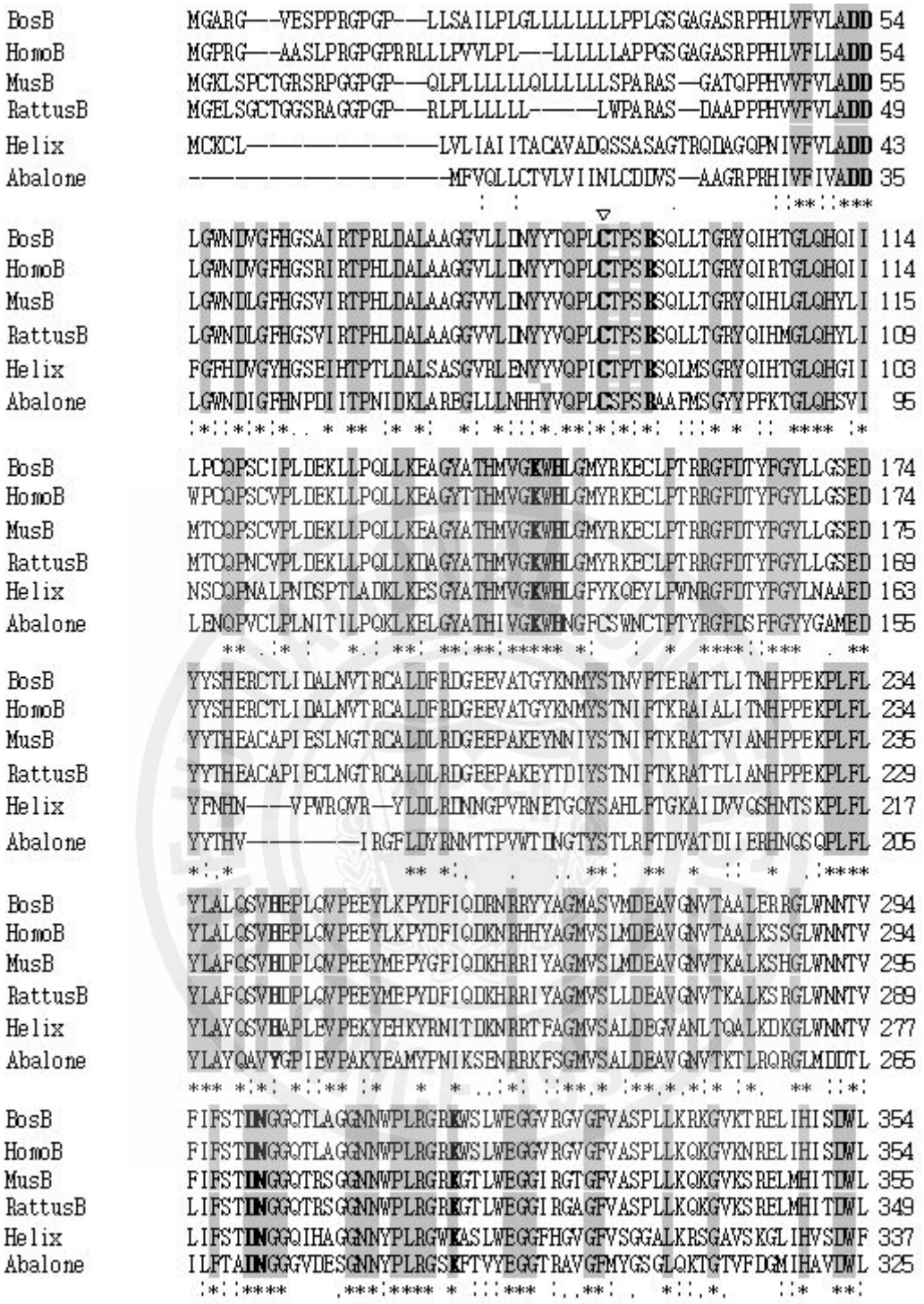
**Phylogenetic analysis of HdArys**

To determine the position of HdArys gene in evolution, 13 sequences which belongs to sulfatase family were analyzed to construct a phylogenetic tree using bacteria *Rhodopirellula baltica* SH1 arylsulfatase B as an out group. Based on the phylogenetic analysis, mammalian arylsulfatase was positioned in main cluster and was clearly sub divided into arylsulfatase B and J with in the main cluster. The Figure 17 shows that *H. pomatia* sulfatase 1 precursor diverged from vertebrate sulfatase family genes and HdArys gene was diverted from *Helix pomatia* sulfatase 1 precursor. The arylsulfatase B of arthropods were positioned in another main sub cluster and *R. baltica* SH1 was positioned as a out group in this study. The phylogenetic analysis provides the evidence that the HdArys has been derived from a common ancestor with other sulfates family genes.

```
                                    AGAAATAGGAAACCAACGTGTAATTCCTGTCCACGGAAGGAGAG   44
ATGTTTGTCCAGTTATTATGCACAGTTTTGGTCATCATCAACCTCTGTGATGACGTTTCTGCAGCAGGA  113
M--F--V--Q--L--L--C--T--V--L--V--I--I--N--L--C--D--D--V--S--A--A--G--   23
CGTCCACGCCATATTGTGTTCATCGTGGCGGATGATCTCGGATGGAACGACATTGGCTTTCACAACCCC  182
R--P--R--H--I--V--F--I--V--A--D--D--L--G--W--N--D--I--G--F--H--N--P--   46
GATATAATCACACCCAACATCGACAAGCTGGCAAGAGAAGGCTTGCTTCTGAATCATCACTATGTTCAA  251
D--I--I--T--P--N--I--D--K--L--A--R--E--G--L--L--L--N--H--H--Y--V--Q--   69
CCACTCTGCAGTCCATCGAGAGCTGCCTTTATGTCCGGCTACTACCCCTTCAAGACAGGTCTGCAGCAC  320
P--L--C--S--P--S--R--A--A--F--M--S--G--Y--Y--P--F--K--T--G--L--Q--H--   92
TCGGTCATTCTGGAGAACCAGCCCGTCTGTCTACCCCTGAATATCACAATCCTGCCACAGAAACTGAAG  389
S--V--I--L--E--N--Q--P--V--C--L--P--L--N--I--T--I--L--P--Q--K--L--K--  115
GAGCTTGGATATGCAACACACATTGTCGGCAAGTGGCACAATGGGTTCTGTAGTTGGAATTGCACCCCG  458
E--L--G--Y--A--T--H--I--V--G--K--W--H--N--G--F--C--S--W--N--C--T--P--  138
ACGTACCGTGGCTTTGACAGCTTCTTTGGCTACTACGGCGCCATGGAAGACTACTACACCCACGTCATT  527
T--Y--R--G--F--D--S--F--F--G--Y--Y--G--A--M--E--D--Y--Y--T--H--V--I--  161
CGTGGCTTCCTTGACTACCGTAACAACACCACCCCCGTTTGGACCGACAACGGCACTTACTCAACGCTT  596
R--G--F--L--D--Y--R--N--N--T--T--P--V--W--T--D--N--G--T--Y--S--T--L--  184
CGGTTTACTGACGTAGCCACTGACATCATCGAGCGTCACAACCAGAGTCAGCCATTGTTTCTGTACCTG  665
R--F--T--D--V--A--T--D--I--I--E--R--H--N--Q--S--Q--P--L--F--L--Y--L--  207
GCGTACCAAGCTGTCTACGGACCTATTGAGGTTCCCGCAAAGTATGAAGCAATGTATCCAAACATTAAA  734
A--Y--Q--A--V--Y--G--P--I--E--V--P--A--K--Y--E--A--M--Y--P--N--I--K--  230
TCAGAAAATCGTCGAAAGTTTTCGGGAATGGTCTCTGCTCTTGATGAAGCAGTTGGTAACGTAACTAAA  803
S--E--N--R--R--K--F--S--G--M--V--S--A--L--D--E--A--V--G--N--V--T--K--  253
ACGTTAAGACAAAGAGGGTTAATGGACGACACGCTGATTCTGTTCACTGCCGATAATGGCGGCGGGGTC  872
T--L--R--Q--R--G--L--M--D--D--T--L--I--L--F--T--A--D--N--G--G--G--V--  276
GACGAATCTGGGAACAACTACCCTCTGCGTGGAAGCAAGTTTACCGTGTACGAAGGCGGAACGAGAGCT  941
D--E--S--G--N--N--Y--P--L--R--G--S--K--F--T--V--Y--E--G--G--T--R--A--  299
GTGGGCTTCATGTATGGATCGGGTCTCCAAAAGACTGGAACTGTATTTGACGGGATGATCCACGCCGTG 1010
V--G--F--M--Y--G--S--G--L--Q--K--T--G--T--V--F--D--G--M--I--H--A--V--  322
GACTGGCTGCCCACCCTGACAGCAGCTGCCGGGGGGGACCCCAGTGTCCGACCGTGACGGCATCAATCTG 1079
D--W--L--P--T--L--T--A--A--G--G--T--P--V--S--D--R--D--G--I--N--L--  345
TGGCCTAGTCTCAGCACAGCCTCCCCGTCCCCCCGCACTGAGGTCGTCTACAACTACGACTCGCACCCC 1148
W--P--S--L--S--T--A--S--P--S--P--R--T--E--V--V--Y--N--Y--D--S--H--P--  368
CAGCCCGTTCAAGGACACGCTGCCATCAGAGTGGGTGACTACAAACTGATCGATGGCTACCCGGGACCC 1217
Q--P--V--Q--G--H--A--A--I--R--V--G--D--Y--K--L--I--D--G--Y--P--G--P--  391
TTCCCTGATTGGTACAAGCCTGAACAAGTCACATCTAGTTTGAACACCAGATTCAGCAGGGATTCGGCC 1286
F--P--D--W--Y--K--P--E--Q--V--T--S--S--L--N--T--R--F--S--R--D--S--A--  414
AATCAGTATCAGCTGTTCAATTTGAAAGATGACCCCAATGAGCGCAACGACCTCTCCAACTTTCGTCCG 1355
```

```
N--Q--Y--Q--L--F--N--L--K--D--D--P--N--E--R--N--D--L--S--N--F--R--P--  437
GACATGGTAAAGAAGCTTGCTGCCAGACTGGCCTGGTATAAGAAGCAGGCAGTACCACCCAACTTCCCT 1424
D--M--V--K--K--L--A--A--R--L--A--W--Y--K--K--Q--A--V--P--P--N--F--P--  460
GAGACCCCCGACGACCTGAGCAACCCTGCACTGTACGGCAATGTCTGGTCTCCTGGCTGGTGTTGAGAG 1493
E--T--P--D--D--L--S--N--P--A--L--Y--G--N--V--W--S--P--G--W--C--*       483
CTTCTTGTTGTACTGTCACTGAATAAAGTCGATATGTGAAAAAAAAAAAAAAAAAAAAAAAAA       1557
```

**Figure 15:** The nucleotide and deduced amino acid sequences of the HdArys. The start (ATG) and stop (TAA) codons are shaded. A putative polyadenylation signal AATAAA is bold and under lined. A putative signal peptide is indicated by dotted underline. A putative N-glycosylation sites are bold. Underlined is shown the identified sulfatase domain. Sulfatase signature is boxed. Lucine zipper motif is shaded. Residue numbers for both nucleotide and amino acid are indicated in the right of each row.


BosB MGARG--VESPPRGPGP--LLSAILPLGLLLLLLLPPLGSGAGASRPPHLVFVLADD 54
HomoB MGPRG--AASLPRGPGPRRLLLPVVLPL--LLLLLAPPGSGAGASRPPHLVFLLADD 54
MusB MGKLSPCTGRSRPGGPGP--QLPLLLLLQLLLLLLSPARAS--GATQPPHVVFVLADD 55
RattusB MGELSGCTGGSRAGGPGP--RLPLLLLL-----LWPARAS--DAAPPPHVVFVLADD 49
Helix MCKCL----------LVLIAIITACAVADQSSASAGTRQDAGQPNIVFVLADD 43
Abalone ----------MFVQLLCTVLVIINLCDDVS--AAGRPPHIVFIVADD 35
BosB LGWNDVGFHGSAIRTPRLDALAAGGVLLDNYYTQPLCTPSRSQLLTGRYQIHTGLQHQII 114
HomoB LGWNDVGFHGSRIRTPHLDALAAGGVLLDNYYTQPLCTPSRSQLLTGRYQIRTGLQHQII 114
MusB LGWNDLGFHGSVIRTPHLDALAAGGVVLDNYYVQPLCTPSRSQLLTGRYQIHLGLQHYLI 115
RattusB LGWNDLGFHGSVIRTPHLDALAAGGVVLDNYYVQPLCTPSRSQLLTGRYQIHMGLQHYLI 109
Helix FGFHDVGYHGSEIHTPTLDALSASGVRLENYYVQPICTPTRSQLMSGRYQIHTGLQHGII 103
Abalone LGWNDIGFHNPDIITPNIDKLAREGLLLNHHYVQPLCSPSRAAFMSGYYPFKTGLQHSVI 96
BosB LPCQPSCIPLDEKLLPQLLKEAGYATHMVGKWHLGMYRKECLPTRRGFDTYPGYLLGSED 174
HomoB WPCQPSCVPLDEKLLPQLLKEAGYTTHMVGKWHLGMYRKECLPTRRGFDTYPGYLLGSED 174
MusB MTCQPSCVPLDEKLLPQLLKEAGYATHMVGKWHLGMYRKECLPTRRGFDTYPGYLLGSED 175
RattusB MTCQPNCVPLDEKLLPQLLKDAGYATHMVGKWHLGMYRKECLPTRRGFDTYPGYLLGSED 169
Helix NSCQPNALPNDSPTLADKLKESGYATHMVGKWHLGFYKQEYLPWNRGFDTYPGYLNAAED 163
Abalone LENQPVCLPLNITILPQKLKELGYATHIVGKWHNGFCSWNCTPTYRGFDSFFGYYGAMED 155
BosB YYSHERCTLIDALNVTRCALDFRDGEEVATGYKNMYSTNVFTERATTLITNHPPEKPLFL 234
HomoB YYSHERCTLIDALNVTRCALDFRDGEEVATGYKNMYSTNIFTKRAIALITNHPPEKPLFL 234
MusB YYTHEACAPIESLNGTRCALDLRDGEEPAKEYNNIYSTNIFTKRATTVIANHPPEKPLFL 235
RattusB YYTHEACAPIECLNGTRCALDLRDGEEPAKEYTDIYSTNIFTKRATTLIANHPPEKPLFL 229
Helix YFNHN---VPWRQVR--YLDLRDNNGPVRNETGQYSAHLFTGKAIDVVQSHNTSKPLFL 217
Abalone YYTHV--------IRGFLDYRNNTTPVWTDNGTYSTLRFTDVATDIIERHNQSQPLFL 205
BosB YLALQSVHEPLQVPEEYLKPYDFIQDRNRRYYAGMASVMDEAVGNVTAALERRGLWNNTV 294
HomoB YLALQSVHEPLQVPEEYLKPYDFIQDKNRHHYAGMVSLMDEAVGNVTAALKSSGLWNNTV 294
MusB YLAFQSVHDPLQVPEEYMEPYGFIQDKHRRIYAGMVSLMDEAVGNVTKALKSHGLWNNTV 295
RattusB YLAFQSVHDPLQVPEEYMEPYDFIQDKHRRIYAGMVSLLDEAVGNVTKALKSRGLWNNTV 289
Helix YLAYQSVHAPLEVPEKYEHKYRNITDKNRRTFAGMVSALDEGVANLTQALKDKGLWNNTV 277
Abalone YLAYQAVYGPIEVPAKYEAMYPNIKSENRRKFSGMVSALDEAVGNVTKTLRQRGLMDDTL 265
BosB FIFSTDNGGQTLAGGNNWPLRGRKWSLWEGGVRGVGFVASPLLKRKGVKTRELIHISDWL 354
HomoB FIFSTDNGGQTLAGGNNWPLRGRKWSLWEGGVRGVGFVASPLLKQKGVKNRELIHISDWL 354
MusB FIFSTDNGGQTRSGGNNWPLRGRKGTLWEGGIRGTGFVASPLLKQKGVKSRELMHITDWL 355
RattusB LIFSTDNGGQTRSGGNNWPLRGRKGTLWEGGIRGAGFVASPLLKQKGVKSRELMHITDWL 349
Helix LIFSTDNGGQIHAGGNNYPLRGWKASLWEGGFHGVGFVSGGALKRSGAVSKGLIHVSDWF 337
Abalone ILFTADNGGGVIDESGNNYPLRGSKFTVYEGGTRAVGFMYGSGLQKTGTVFDGMIHAVDWL 325

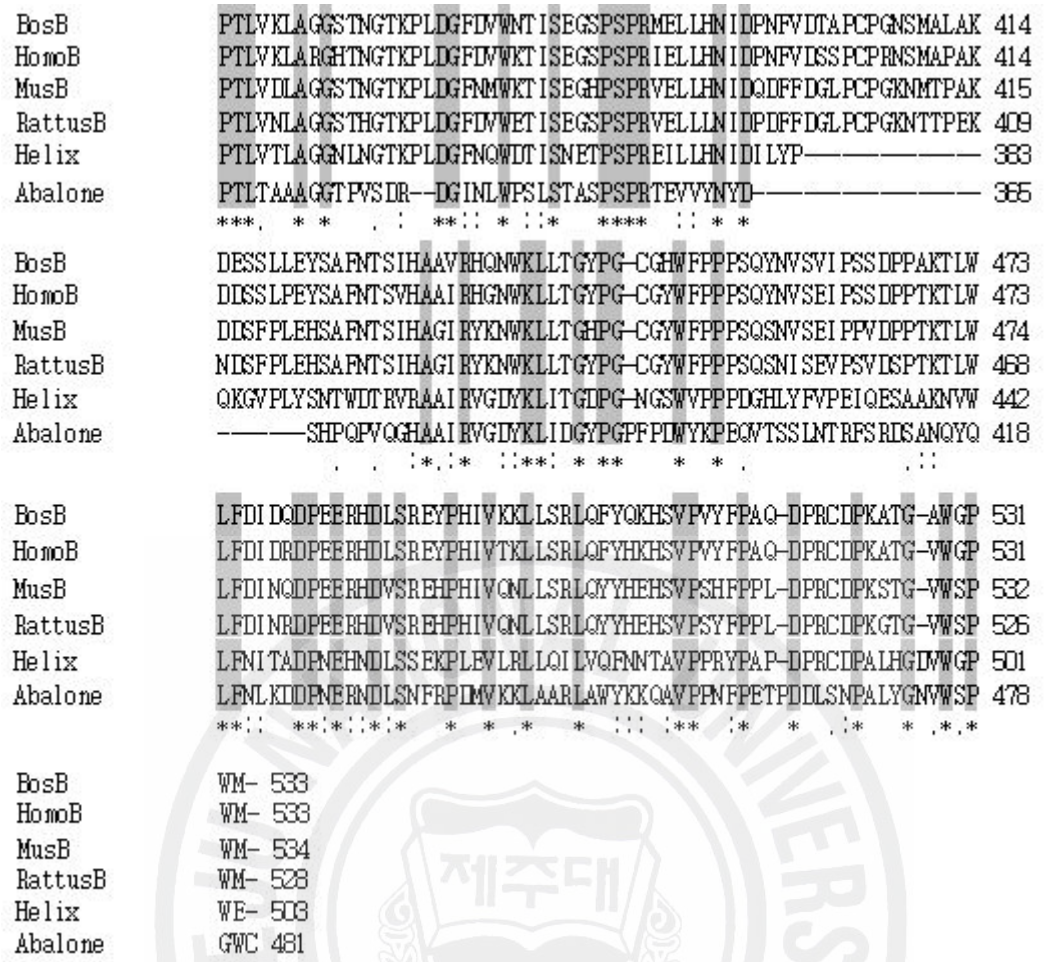



**Figure 16:** Multiple alignment of amino acid sequences for the HdArys and other species arylsulfatase. Amino acid sequences of aylsulfatase from *Helix pomatia* Sulfatase 1 precursor (accession no. AAF30402), *Bos taurus* Arylsulfatase B (accession no. NP_001094645), *Homo sapiens* arylsulfatase B (accession no. NM_000046), *Mus musculus* Arylsulfatase B *(*accession no. NP_033842), *Rattus norvegicus* arylsulfatase B (accession no. NP_254278) were aligned. Bold letters represents the amino acid homologous to essential amino acids found at active site of arylsulfatases. Conserved sulfatase signature is shaded. Cystein to be transformed into $C_\alpha$ formylglycine is labeled in (♣▽) over the column.

**Table 4:** Pairwise CLUSTALW analysis and comparison of the deduced amino acid sequence of HdArys with other known arylsulfatases.

| Species | Accession number | Identity (%) | Amino acids |
|---|---|---|---|
| **Sulfatase 1 precursor** *Helix pomatia* | AAF30402 | 45 | 503 |
| **Arylsulfatase B** *Mus musculus* | NP_033842 | 41 | 534 |
| **Arylsulfatase B** *Bos taurus* | NP_001094645 | 41 | 533 |
| **Arylsulfatase B** *Homo sapiens* | NM_000046 | 41 | 533 |
| **Arylsulfatase B** *Rattus norvegicus* | NP_254278 | 41 | 528 |
| **Arylsulfatase J** *Mus musculus* | EDL12284 | 38 | 572 |
| **Arylsulfatase I** *Homo sapiens* | AY875937 | 37 | 537 |
| **Arylsulfatase J** *Homo sapiens* | AM049401 | 35 | 596 |

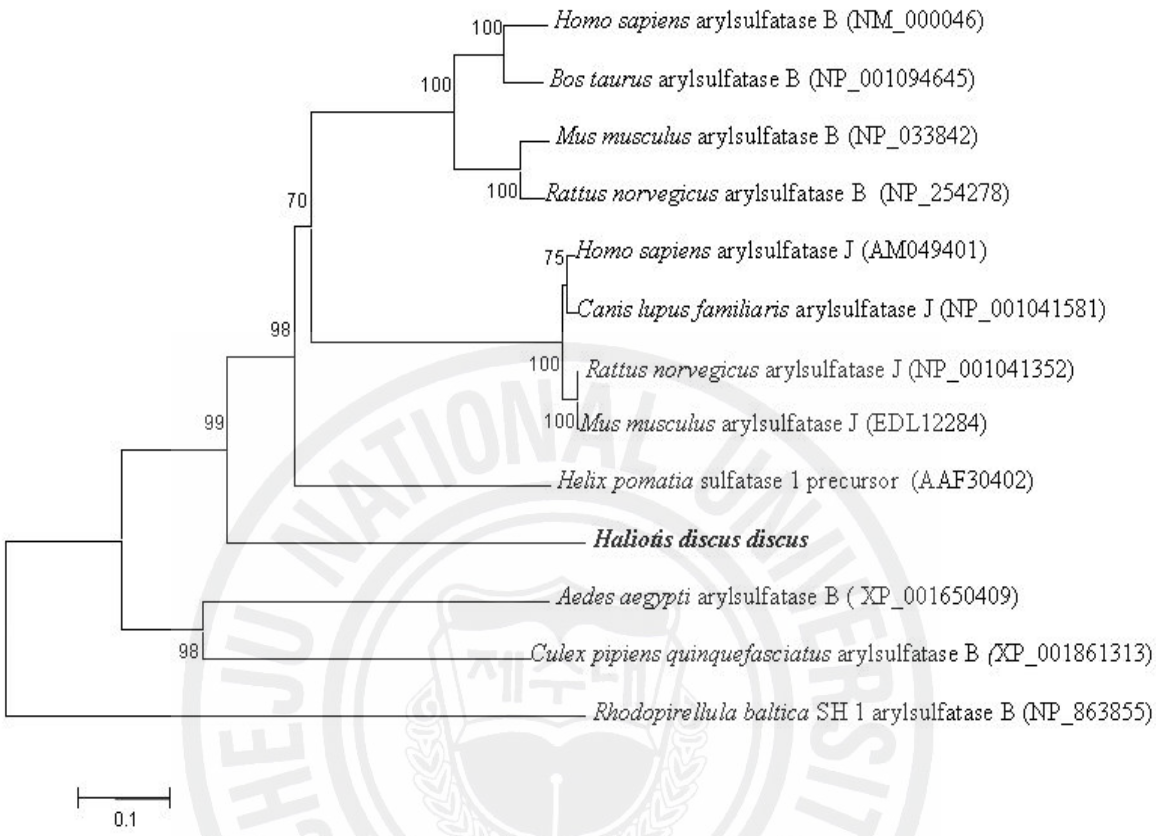



**Figure 17:** Phylogenetic analysis of HdArys with other species sulfatases and arylsulfatases. The tree is based on an alignment corresponding to the full length amino acid sequences, using clustalW and MEGA (3.1). The numbers at the branches denotes the bootstrap majority consensus values on 1000 replicates. *Rhodopirellulanbaltica* SH1 arylsulfatase B was used as out groups for the analysis. The GenBank accession numbers are shown next to each organism.

**3D-Structure analysis of HdEndg, HdAmy, HdMann and HdArys**

The 3D-structures of HdEndg, HdAmy, HdMann and HdArys were predicted by Swiss-Model and visualized in Swiss-Pdb viewer (Schwede et al., 2003). The 3D-structure of endoglucanase from termite, *N. takasagoensis*, at pH 2.5 was taken as a template and hypothetical confirmation of HdEndg was predicted (Starting from 156 aa residues). It showed catalytically module (folds [α/α]6] which is general to GHF9 (Figure 18) and catalaytically imporatnat residues. The sequence identity of the template with HdEndg was 49%. As the termite does not contained CBM module in their protein, HdEndg extended N-terminal sequence region has not taken to consideration for the structural analysis.

HdAmy 3D-structure was predicted (Starting from 22-502 aa residues) using pig pancreatic α-amylase as a template. The sequence identity of the template with HdAmy was 56%. The 3D-structure of HdAmy showed three main domains namely i) domain A as the catalytic domain, shaped as a $(\beta/\beta)_8$ barrel; ii) domain B is a long loop between $\beta_3$ and $\alpha_3$ and (iii) domain C comprised with a C-terminal beta sandwich (Greek key) (Figure 19).

X-Ray structure of β-mannanase from blue mussel *Mytilus edulis* was taken as a template and HdMann 3D structure was predicted (Starting from 20-375 aa residues). The sequence identity of HdMann with selected template was 50%. The predicted 3D-structure of HdMann showed $(\beta\alpha)_8$-barrel fold similar to GHF5 enzymes (Figure 20).

Figure 21 shows the predicted 3D-structure of the HdArys and its catalytically imporatnat residues, which compared with human arylsulfatase. The N-terminal domain consists of 10 stranded β sheets and 13 α helices. In 10 stranded β sheets all strands except β 10 are parallel to each other. The small C-terminal domain consists of 4 stranded β sheets which are anti parallel to the large β sheet domain.

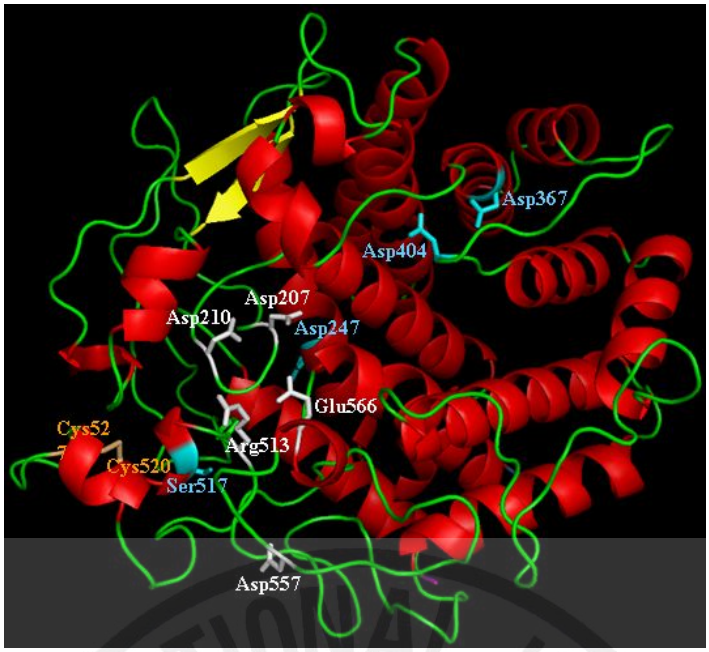



**Figure 18:** HdEndg 3D-structure predicted by Swiss-Model and visualized in Swiss-Pdb viewer. The cataltically important residues in GHF9 cellulases (His replaced by Arg513, Asp557, Glu566, Asp207 and Asp210) and Conserved $Ca^{2+}$ binding resdues at Asp247, Asp367, Asp404, and Ser517 are marked in white and blue, respectively.

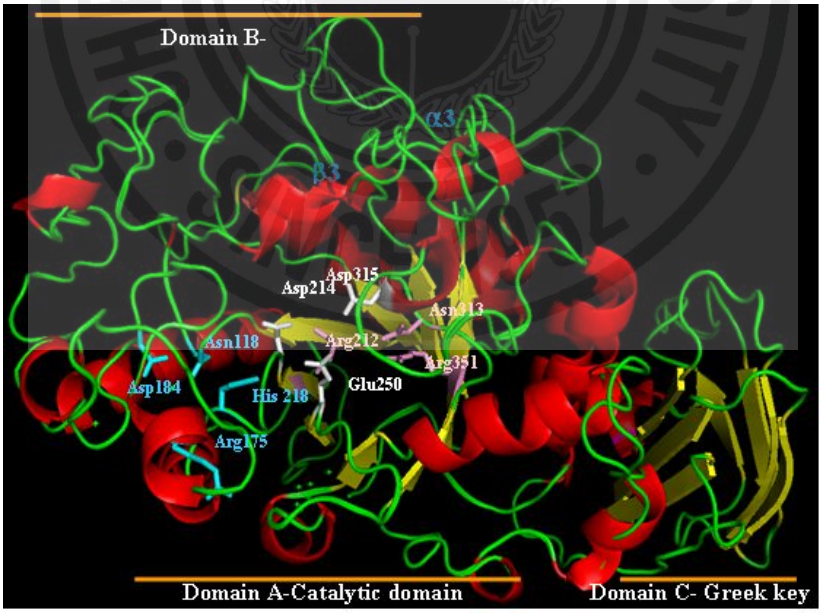



**Figure 19:** HdAmy 3D-structure predicted by Swiss-Model and visualized in Swiss-Pdb viewer. The catalytically active site residues (Asp214, Glu250 and Asp315) , residues responsible for the calcium (Asn118, Arg175, Asp184, and His218) and chloride (Arg212, Asn 313, Arg351) binding sites are marked in white, blue and pink, respectively.

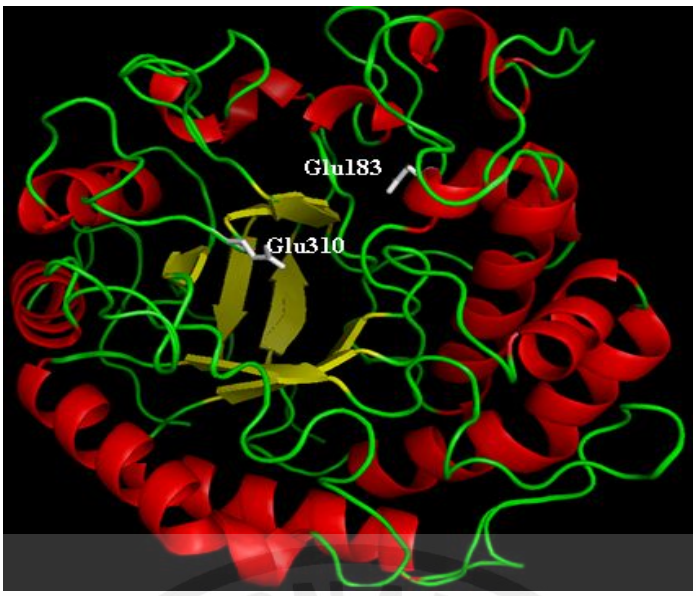



**Figure 20:** HdMann three dimensional structure predicted by Swiss-Model and visualized in Swiss-Pdb viewer. Residues participating in catalytic action as nucleophiles or proton donors in GHF5 enzymes (Glu183 and Glu310) are marked in white.

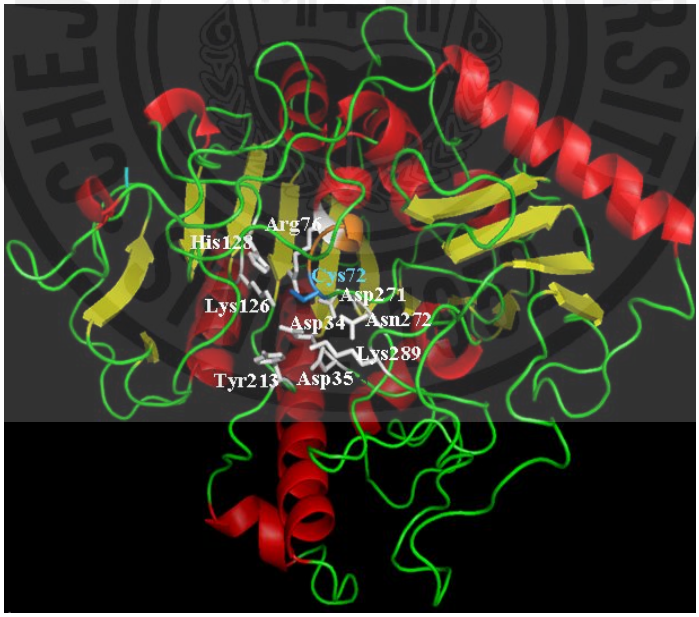



**Figure 21:** HdArys 3D-structure predicted by Swiss-Model and visualized in Swiss-Pdb viewer. The cystein (Cys72) residue which responsible for formation of formylglycine is marked in blue. Functionally important residues involved in substrate binding and activation (Lys126, Tyr213, Lys289), metal coordination (Asp34, Asp35, Asp271, Asn272) and enzyme stabilization (Arg76 and His128) marked.

## In vivo PDE mRNA expression analysis during starvation and re-feeding

### Determination of weight loss %

Weight loss percentage of the starved and re-fed abalones is shown in Figure 22. Maximum weight loss was observed at week 8 and it was about 18%. The weight loss was increased significantly at week 2 and increasing level was reduced from week 6-8. However, the weight loss was significantly decreased during the period of re-feeding. Interestingly, there was no significant difference in weight loss in control animals, which were fed continuously through out the experiment in a separate tank. During the subsequent re-feeding period, previously unfed abalones increased their body weight sufficiently to over come the weight loss caused during 8 week of starving condition.

### Tissue specific mRNA expression analysis

RT-PCR results showed that PDEs were expressed in different levels in hepatopancrease and digestive tract (Figure 23A). All the genes were expressed in higher level in hepatopancrease than the digestive tract tissue. In hepatopancreas, alginate lyase showed the highest relative mRNA expression. Mean time other PDEs namely β-1,4-endoglucanase, arylsulfatase, α-amylase, β-mannanase and laminarinase were expressed in descending order, respectively (Figure 23B). However, mannanase was only expressed hepatopancrease but not in the digestive tract.

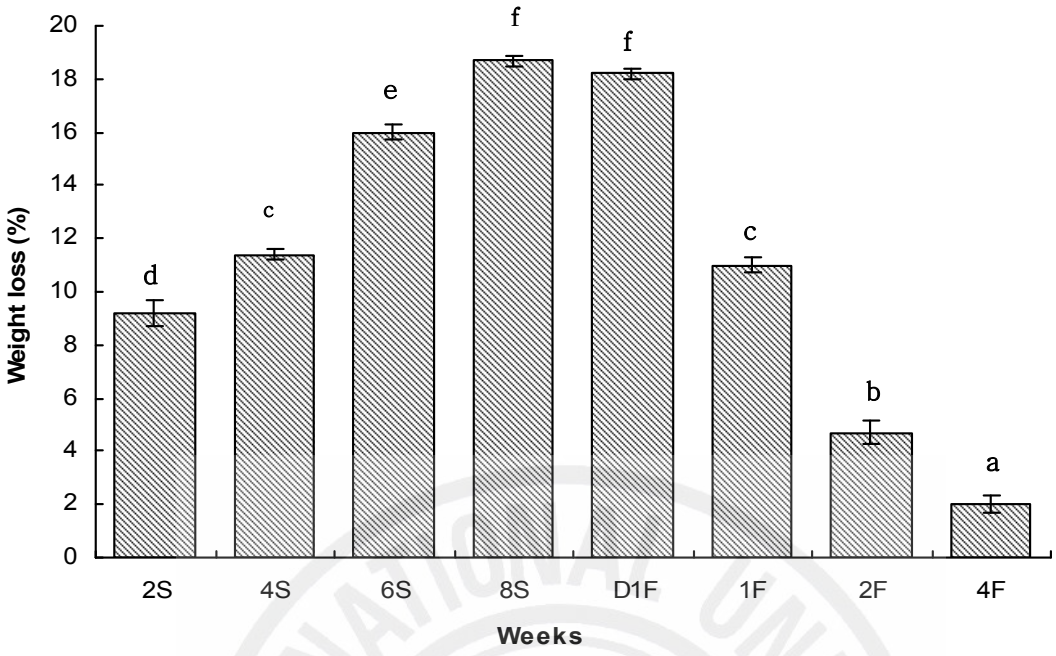



**Figure 22.** Weight loss of *H. discus discus* during starvation for 8 weeks and re-feeding for 4 weeks**.** 2 S-Week 2 starvation; 4S-Week 4 starvation; 6S-Week 6 starvation; 8S-Week 8 starvation D1F-Day 1 feeding; 1F-Week 1 feeding; 2F-Week 2 feeding;; 4F- Week 8 feeding**.** Bars represent the means ± SD. The expression levels correspond to the mean of three assays. Means with the same letters are not significant different at $p<0.05$.

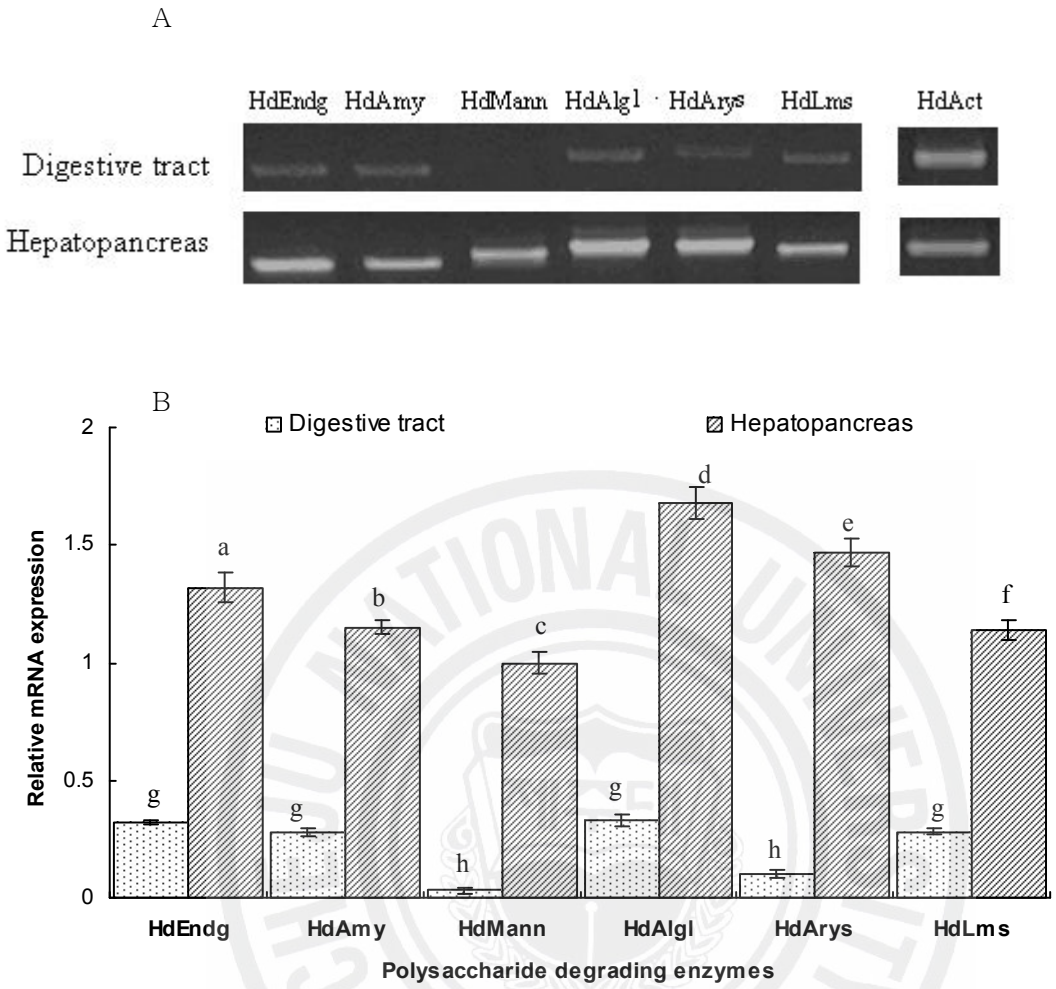



**Figure 23.** Tissue specific relative mRNA expression of polysaccharide degrading enzymes (HdEndg, HdAmy, HdMann, HdAlgl, HdArys and HdLms) in digestive tract and hepatopancreas of *H. discus discus*. **A.** Representative agarose gel showing amplified products of HdEndg, HdAmy, HdMann, HdAlgl, HdArys and HdLms and corresponding actin control. **B**. Semi quantitative RT-PCR analysis of PDEs relative to corresponding actin, as determined by densitometric analysis using Scion Image software. The expression levels correspond to the mean of three assays. Means with the same letters are not significant different at $p<0.05$ based on ANOVA. Bars represent the means $\pm$ SD.

**The mRNA expression analysis of PDEs in abalone hepatopancreas during starvation**

During 8 week of starvation, mRNA expression of all the PDE related genes were significantly decreased from the control level. Then, the mRNA expression was increased with the re-feeding period of the animal for day 1-14. Figure 24 shows the semi quantitative mRNA expression analysis results of PDEs in hepatopancreas during starvation and re-feeding. The HdEndg mRNA expression was decreased by 40% from 0-8 week of starvation and it was further decreased to 60% from control value (Figure 24A). The HdAmy mRNA level was decreased by 50% from control to day 1 of starvation (Figure 24B). Also, the decreasing fold of HdAlgl and HdLms was 47 and 70%, respectively from 0-8 week starvation (Figure 24C and Figure 24D). Moreover, HdMann and HdArys mRNA expression was decreased from 0-8 week of starvation by 40 and 30%, respectively (Figure 24E and Figure 24F). HdAlgl mRNA expression was further decreased by 10% from 8$^{th}$ week of starvation to day 1 of re-feeding and increased only upto 80% during 14 days of re-feeding. In contrast, HdLms mRNA expression was started to increase from day 1 of re-feeding. There was no HdMann and HdArys mRNA expression level difference was observed between 8$^{th}$ week of starvation and day 1 of re-feeding. Interestingly, after day 1 of re-feeding the mRNA expression levels of all the genes were increased and it reached to control level. As the expression level was reached to basal level after 14 days of re-feeding we did not analyze mRNA expression for further time points. The Figure 25 summarizes the effect of starvation and re-feeding on PDEs in disk abalone with relative to actin expression as internal control. It showed that HdLms, HdMann and HdAlgl mRNA expression was significantly decreased from 0-8 week of starvation compared to HdEndg, HdAmy and HdArys. Interestingly, all the genes were returned to its basal mRNA expression level, except HdAlgl at 14 days of re-feeding.

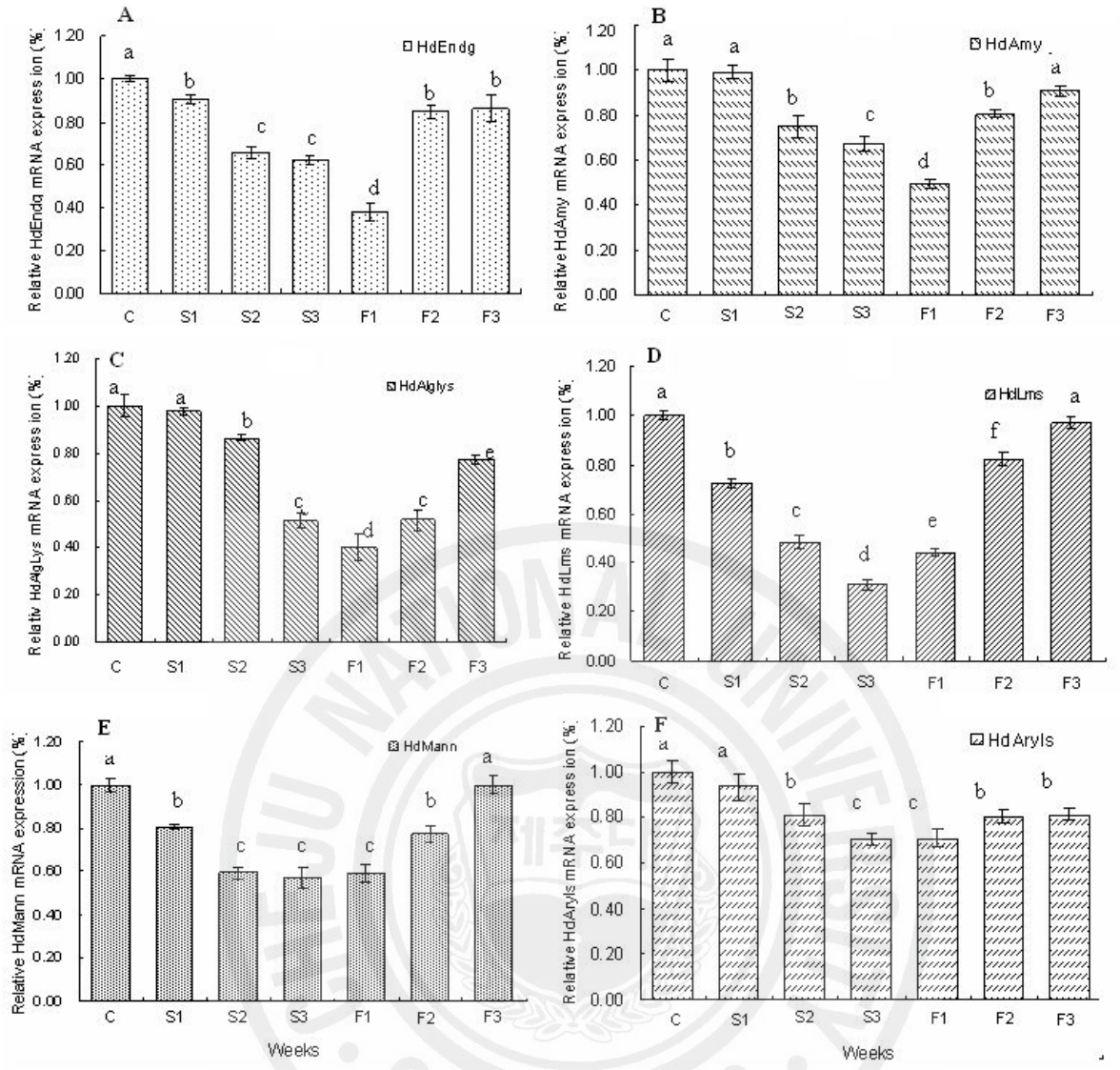



**Figure 24.** Semi quantitative mRNA expression analysis of PDEs in hepatopancreas at different time points of starvation and re-feeding A. HdEndg, B. HdAmy, C. HdAlgl, D. HdLms, E. HdMann and F. HdArys mRNA expression at different time points of starvation and re-feeding. Lane C: control, S1: Starvation 2nd week, S2: Starvation 4th week, S3: Starvation 8th week, D1F: 1st day after re-feeding, F2: 1st week after re-feeding, F3: 2nd week after re-feeding. Corresponding actin protein mRNA expression was used as an internal PCR control. The level of each gene expression is mean of three RT-PCR assays, which are calculated relative to that of the expression recorded for the corresponding control (shown as 1). Bars represent the means ± SD. The expression levels correspond to the mean of three assays. Means with the same letters are not significant different at $p<0.05$ based on ANOVA.

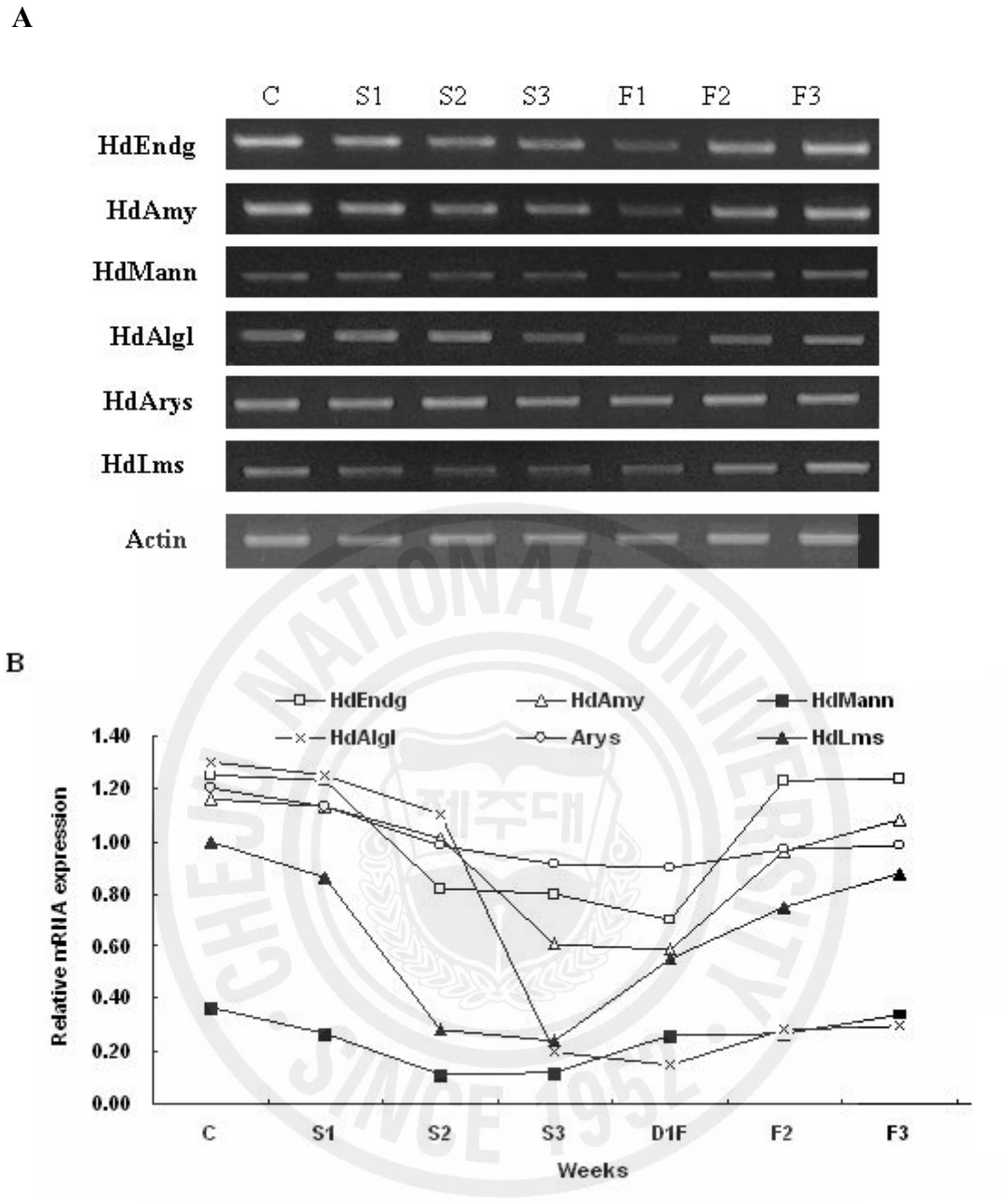



**Figure 25**. Comparison of relative mRNA expression of PDEs in hepatopancreas at different time points of starvation and re-feeding. **A.** Representative agarose gel showing amplified products of HdEndg, HdAmy, HdMann, HdAlgl, HdArys and HdLms and corresponding actin control. **B.** Semi quantitative RT-PCR analysis of PDEs relative to corresponding actin, as determined by dinsitometric analysis using Scion Image software. Lane C: control, S1: Starvation 2$^{nd}$ week, S2: Starvation 4th week, S3: Starvation 8th week, D1F: 1$^{st}$ day after re-feeding, F2: 1st week after re-feeding, F3: 2$^{nd}$ week after re-feeding. The expression levels correspond to the mean of three RT-PCR assays.

**RT-PCR analysis of antioxidant CuZnSOD expression during starvation**

The CuZnSOD mRNA expression during starvation and re-feeding is shown in Figure 26. CuZnSOD expression was increased by 30 and 40% after 2 and 8 weeks of starvation compared to control group. The CuZnSOD mRNA expression was reached to control level after re-feeding.

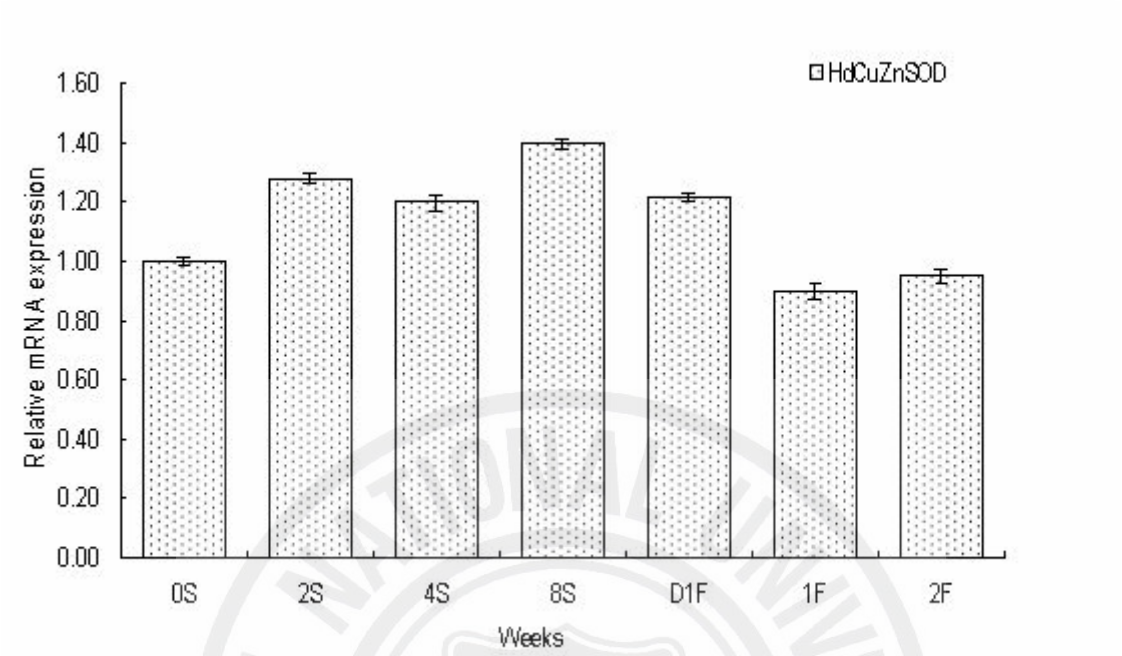



**Figure 26:** Relative mRNA expression levels of CuZnSOD (HdCuZnSOD) during starvation and re-feeding in abalone hepatopancreas. Lane 0S: control, 2S: starvation 2nd week, 4S: starvation 4th week, 8S: starvation 8th week, D1F: 1st day after re-feeding, 1F: 1st week after re-feeding, 2F: 2nd week after re-feeding. Corresponding actin expression was used as an internal PCR control. The level of HdCuZnSOD is mean of three RT-PCR assays, which are calculated relative to that of the expression recorded for the corresponding control (shown as 1). Bars represent the means $\pm$ SD. The expression levels correspond to the mean of three assays. Means with the same letters are not significant different at $p<0.05$ based on ANOVA.

## Functional characterization of abalone PDEs

### Purification of recombinant PDEs by pMAL protein purification system

The recombinant HdEndg, HdAmy, HdAlgl and HdArys enzymes were over-expressed in *E.coli* K12TB1. The SDS–PAGE analysis results showed that approximately 107.5, 96.5, 75.5 and 94.5 kDa strong protein band for HdEndg, HdAmy, HdAlgl and HdArys, respectively (Figure 27). Sizes of our purified protein were perfectly matched with predicted sizes of 65.0, 54, 33, 52 kDa, since maltose binding protein (MBP) has 42.5 kDa molecular weight.

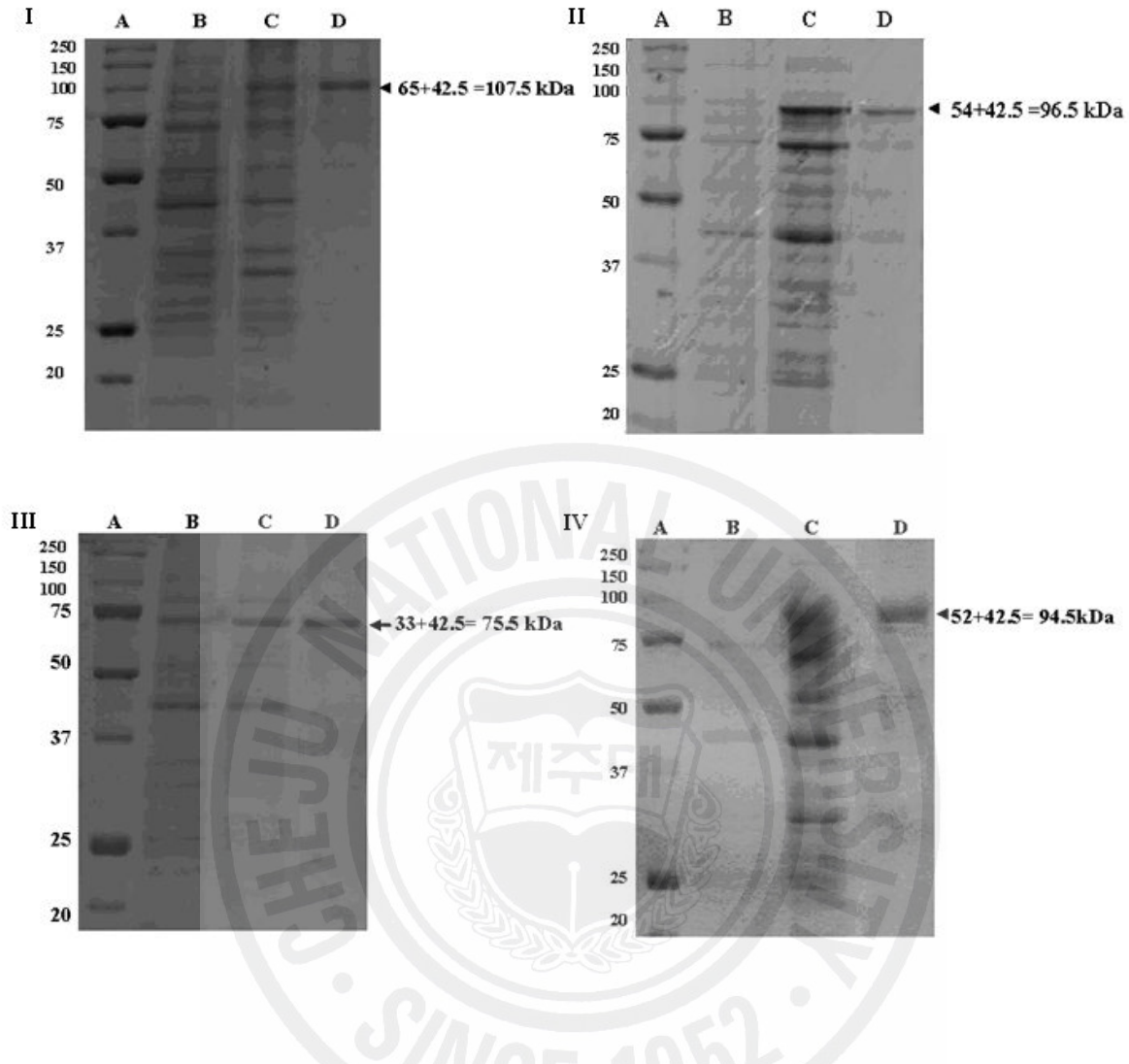



**Figure 27:** Purification of recombinant abalone PDEs in *E. coli* K12TB1 cells and polyacrylamide gel electrophoresis. I: purification of recombinant HdEndg; II: purification of recombinant HdAmy; III: purification of recombinant HdAlgl; IV: recombinant HdArys.

A: protein marker (Bio-Rad), B: before induction with IPTG, C: after IPTG induction at 25 $^{o}$C for 5 h, D: Purified recombinant protein using pMAL$^{TM}$ protein fusion and purification system.

**Enzymatic activity and biochemical properties of recombinant HdEndg**

The purified HdEndg showed 0.3 U/mg specific activity to 1% CMS as a substrate. HdEndg activity under different temperature and pH conditions were carried out to characterize the biochemical properties. The thermostability and thermoactivity of the purified HdEndg are shown in Figure 28A. The purified enzyme in 50 mM acetate buffer, pH 4.5 was quite stable upto 40 $^{o}$C for 30 min and reduced its activity by 35% at 50 $^{o}$C upon incubation for 30 min. It showed maximum activity at 40 $^{o}$C and the enzyme activity was drastically reduced thereafter. The relative enzyme activity at 60 $^{o}$C and 70 $^{o}$C was 38% and 0%, respectively under the assay conditions used. Further, thermal stability of the recombinant Endg at different temperatures with the respect of time is given in Figure 28B. The enzyme was fairly stable at 30 and 40 $^{o}$C for 4 hrs and relative activity at 30 and 40 $^{o}$C were 72% and 55%, respectively. Further, at 40 and 50 $^{o}$C, the enzyme was inactivated at 2h and 1h, respectively. The enzyme was stable at pH 4.0-8.0. It showed optimal activity at pH 4.5 and retained 65% activity at pH 3.5 and 72% activity at pH 8.5 (Figure 28C). The effect of metal ions and reagents on enzymatic activity was investigated as inhibitors or activators at optimum conditions (Figure 28D). The enzyme activity was enhanced by 20% (1.2 fold) by $Ca^{2+}$ (5mM) and decreased by 60% (0.6 fold) by $Mn^{2+}$. There was no significant effect by other reagents on enzymatic assay.

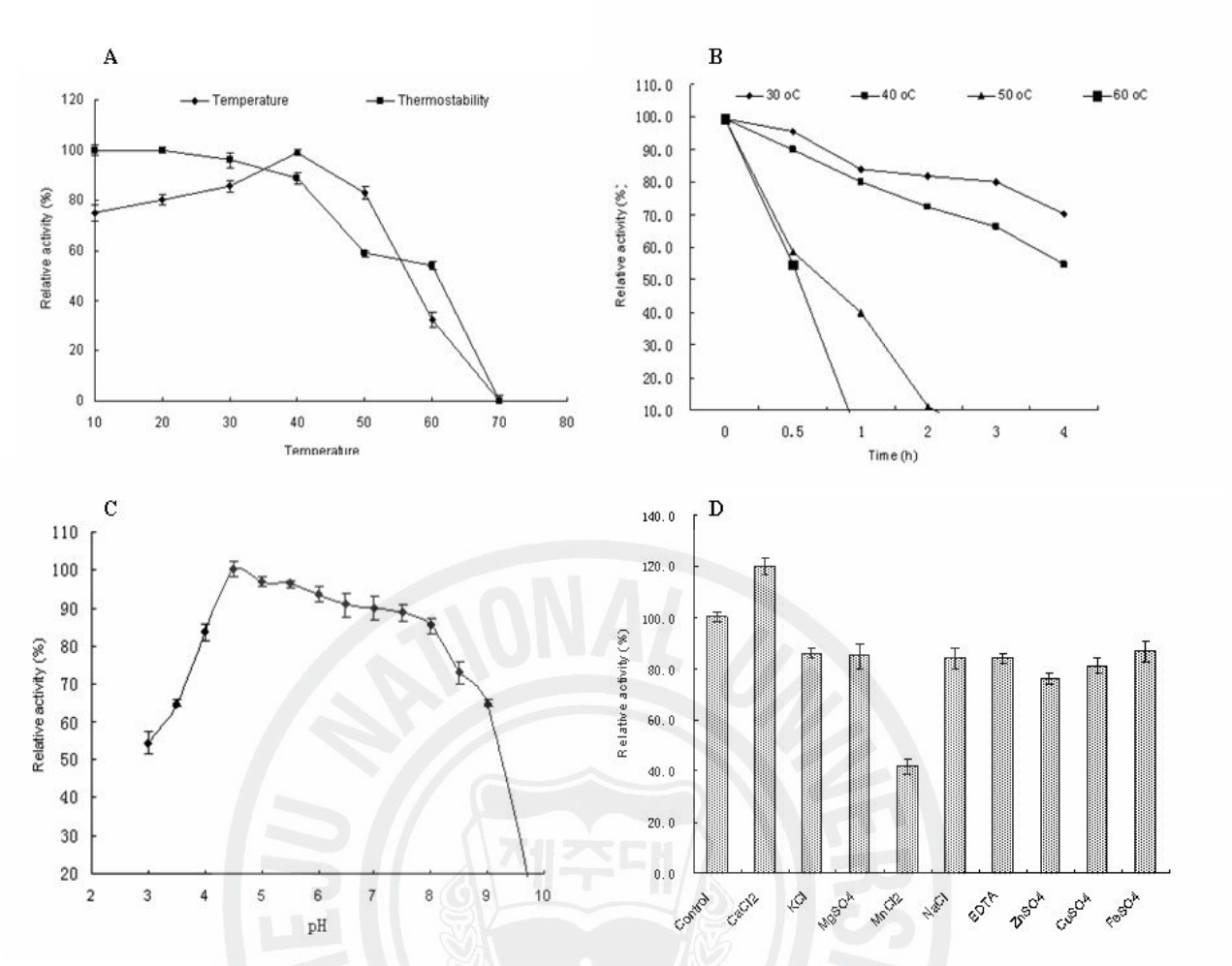



**Figure 28:** Characterization of biochemical properties of recombinant HdEndg.

**A:** Determination of optimum temperature and enzyme stability. Relative enzyme activity (% of maximum) of the purified HdEndg measured at different temperatures with carboxymethyl cellulose (CMC) as a substrate. Optimum temperature was determined by incubating the recombinant HdEndg in citrate buffer (50 mM, pH 4.5) containing 1% (w/v) CMC over a temperature ranges from 10-70 $^{o}$C in 10 $^{o}$C increments for 30 min and the reducing end groups exposed were determined using the dinitrosalicylic acid method. For enzyme stability, the enzyme solution in 50 mM acetate buffer, pH 4.5 was incubated for 30 min at various temperatures, and then the residual enzyme activities were assayed under standard conditions.

**B:** Determination of thermal stability at different temperatures for different time points. Thermostability of the recombinant HdEndg was determined by measurement of residual activity under standard conditions after incubation in acetate buffer (50 mM, pH 4.5) at

various temperatures (30, 40, 50, 60 $^{o}$C) for different times ranging from 0.5-4 h. **C:** Relative enzyme activity (% of maximum) of the purified endoglucanase at different pH values. Relative enzymatic activity was assayed by incubating the enzyme with the carboxymethyl cellulose substrate (CMC) at different pH values (50 mM buffer) for 30 min at 40 $^{o}$C prior to measuring the residual Endg activity under standard assay conditions. Activities at different pH are shown with solid diamond. Buffer used (50 mM): citrate, pH 3-6.5; phosphate, pH 6-8; glycine, pH 8-10. **D:** Effect of metal ions and reagents on enzymatic activity. The data presented are average of two independent experiments. Error bars represent $\pm$ SD.

**Enzymatic activity and biochemical properties of recombinant HdAmy**

The purified HdAmy showed 0.1 U/mg specific activity to 1% starch as a substrate. The recombinant HdAmy enzyme showed temperature optima at 50 $^{o}$C (Figure 29A). From 30-60 $^{o}$C, the activity was retained up to more than 70% and then began to decrease. The thermal stability of HdAmy is shown in Figure 29B. Also, thermal inactivation was occurred when the protein was at 60-70 $^{o}$C for 30 min. However, the protein can be stable for heat at 30, 40 and 50 $^{o}$C up to more than 60% and at 30 and 40 $^{o}$C the protein was stable even for1h. The effect of pH on the purified enzyme is shown in Figure 29C. Higher activity was obtained in the range of buffer from pH 5.5-8.5 and optimum pH was obtained at pH 7.5. The enzyme was totally inactivated at pH 10.5. The influence of different metal ions and reagents were tested as inhibitors or activators at optimum conditions (Figure 29D). The enzyme activity was enhanced by about 15% with 5 mM of $CaCl_2$. However, the relative activity was decreased by about 60% for EDTA, $Mn^{2+}$ $Fe^{2+}$, $Cu^{2+}$ and $Zn^{2+}$.

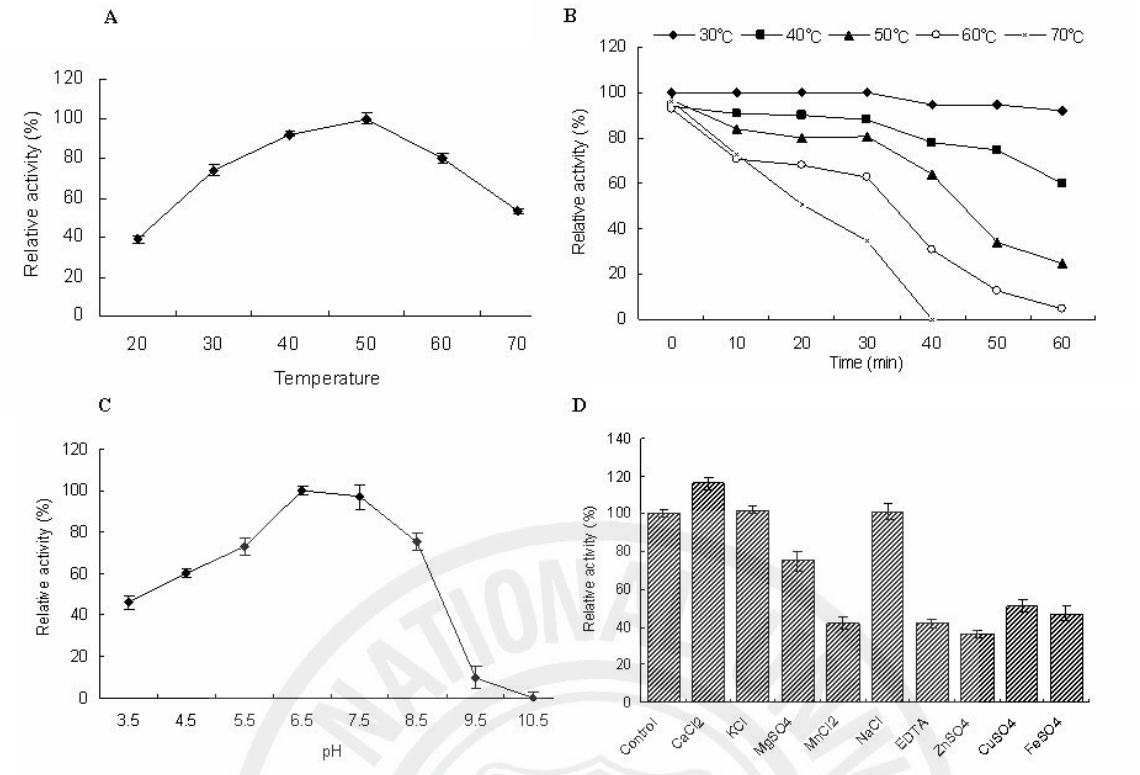



**Figure 29:** Characterization of biochemical properties of recombinant HdAmy.

**A:** Relative enzyme activity (% of maximum) of the purified HdAmy measured at different temperatures with 1% starch as a substrate. Optimum temperature was determined by incubating the recombinant HdAmy in phosphate buffer (50 mM, pH 6.5) containing 1% (w/v) starch over a temperature ranges from 12-70 $^{o}$C in 10 $^{o}$C increments for 30 min and the reducing end groups exposed were determined using the dinitrosalicylic acid method. **B:** Determination of thermal stability at different temperatures for different time points. Thermostability of the recombinant HdAmy was determined by measurement of residual activity under standard conditions after incubation in phosphate buffer (50 mM, pH 6.5) at various temperatures (30, 40, 50 and 60 $^{o}$C) for different times ranging from 0-60 min. **C:** Relative enzyme activity (% of maximum) of the purified HdAmy at different pH values. Relative enzymatic activity was assayed by incubating the enzyme with the 1% starch at different pH values (50 mM buffer) for 30 min at 50 $^{o}$C prior to measuring the residual amylase activity under standard assay conditions. Activities at different pH are shown with

solid diamond. Buffer used (50 mM): citrate, pH 3-6.5; phosphate, pH 6-8; glycine, pH 8-10. **D:** Determination of effect of metal ions and reagents on HdAmy activity. The data presented are average of two independent experiments. Error bars represent ±SD.

**Enzymatic activity and biochemical properties of recombinant HdAlgl**

The purified enzyme showed 2 Unit/mg activity towards sodium alginate as a substrate. The enzyme showed temperature optima at 40 $^{o}$C (Figure 30A). From 30-50 $^{o}$C, the activity was retained up to more than 80% and then began to decrease. The thermal stability of HdAlgl is shown in Figure 30B. It can be observed that when protein was heated at 50, 60 and 70 $^{o}$C, thermal inactivation was occurred after 30 min. However, the protein can be stable for heat at 30 and 40 $^{o}$C even for 1h without changing its activity. The effect of pH on the purified enzyme is shown in Figure 30C. Higher activity was obtained in the range of buffer from pH 6.0-8.0 and optimum pH was obtained at pH 7.0. The enzyme was totally inactivated at pH 10.0.

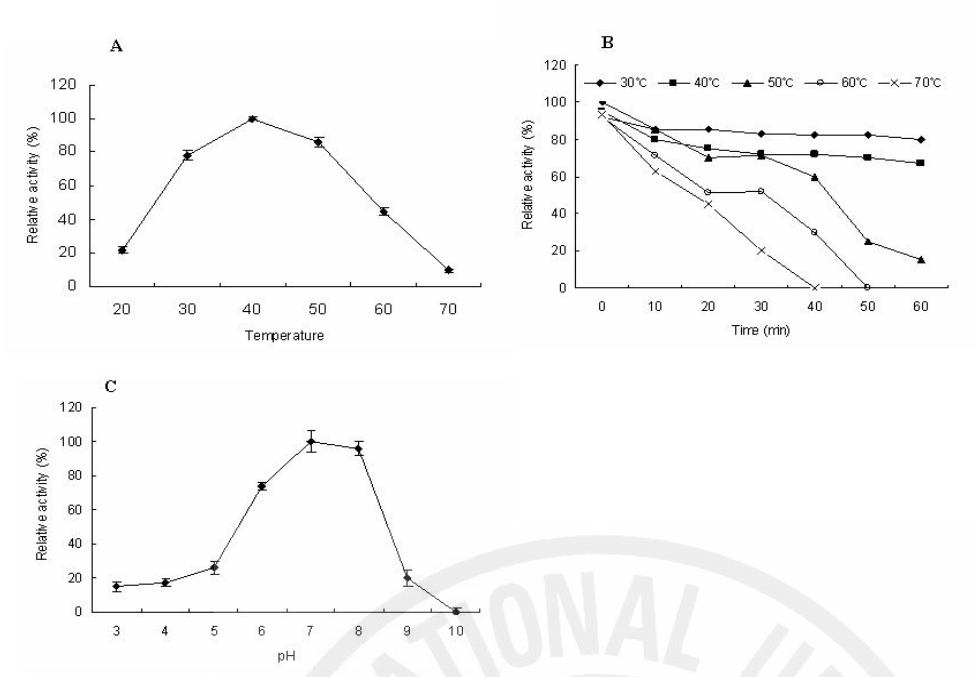



**Figure 30:** Characterization of biochemical properties of recombinant HdAlgl.

**A:** Relative enzyme activity (% of maximum) of the purified HdAlgl measured at different temperatures with .1% sodium alginate as a substrate. Optimum temperature was determined by incubating the recombinant HdAlgl in phosphate buffer (50 mM, pH 7.0) containing .1% (w/v) sodium alginate over a temperature ranges from 20-70 $^{o}$C in 10 $^{o}$C increment. The activity was monitored by measuring the absorbance at 235 nm. **B:** Determination of thermal stability at different temperatures for different time points. Thermostability of the recombinant HdAlgl was determined by measurement of residual activity under standard conditions after incubation in phosphate buffer (50 mM, pH 7.0) at various temperatures (30, 40, 50,60 and 70 $^{o}$C) for different times ranging from 0-60 min. **C:** Relative enzyme activity (% of maximum) of the purified HdAlgl at different pH values. Relative enzymatic activity was assayed by incubating the enzyme with the 1% sodium alginate at different pH values (50 mM buffer) at 40 $^{o}$C prior to measuring the residual alginate lyase activity under standard assay conditions. Buffer used (50 mM): citrate, pH 3-6.5; phosphate, pH 6-8; glycine, pH 8-10.

**Enzymatic activity and biochemical properties of recombinant HdArys**

The specific activity measured for the recombinant HdArys towards various substrates and relative activity % is given in Figure 31. It shows that the specific activity towards all the substrates was about 0.06 unit/mg at 45 $^{o}$C at pH 7.0. The relative enzymatic activity was the highest for the natural substrate, dermaten sulfate and it showed 97% activity for synthetic substrates, p-nitro phenyl sulfate and p-nitro chatachol sulfate. However, due to the low activity of recombinant HdArys, we did not further characterize the enzyme.

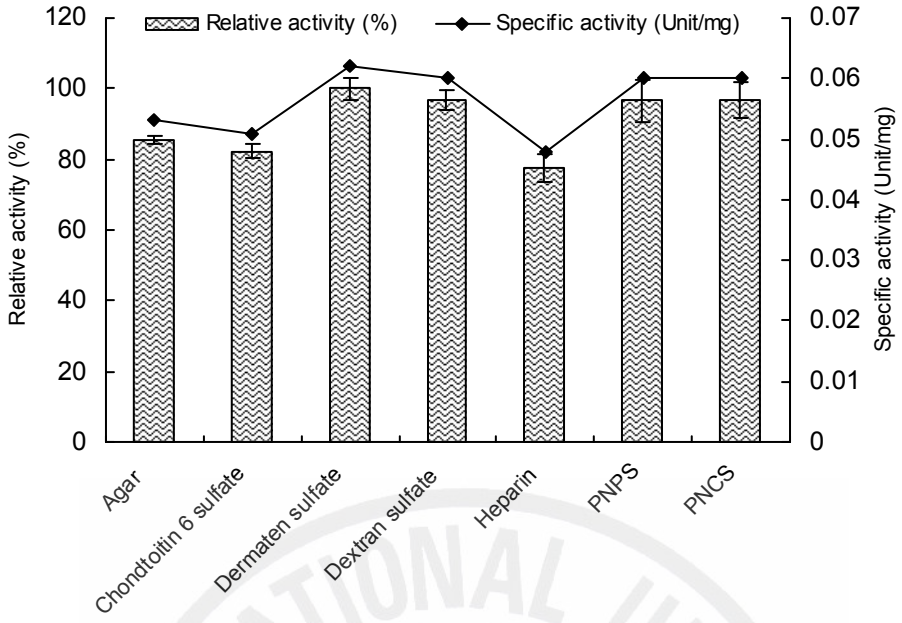



**Figure 31:** Substrate specificity and relative activity of recombinant HdArys.

# DISCUSSION

In this study, we described the molecular basis, *in vivo* expression, biochemical properties and functional role of several PDEs identified from disk abalone cDNA library.

The endoglucanases belong to the family 6, 7, 8, 9, 12 and 45 based on amino acid sequence. By sequence comparisons with other invertebrate and bacterial endoglucanases, HdEndg is considered to be GHF9 type catalytic domain based on active site regions and highly conserved characteristic active site residues. The cataltically important residues in GHF9 cellulases such as Asp557, Glu566, Asp207 and Asp210 in the HdEndg were all conserved. It showed His506, Asp550, Glu559, Asp200 and Asp203, in the *H. discus hannai* primary structure (Suzuki et al., 2003). In *Cherax quadricarinatus* showed catalytically active residues at His397, Asp546, Asp555, Asp441 and Glu450 (Byrne et al., 1999). Extended N-terminal region was found in abalone cDNA with 131 residues. Similar region was found in *H. discus hannai* with 121 residues and 125 residues with *Cellulomonas fimi* CenA (Wong et al., 1986). It showed 35% and 12% identity with CBM attached by a linker in *H discus hannai* (Suzuki et al., 2003) and *C.fimi* CenA (Wong et al., 1986)., respectively. This region was not found in other invertebrate cellulases. CBM family II is the largest family among five major CBM families (family I-IV and VI) and it pocess strictly conserved 4 tryptophans and highly conserved 2 cysteines that form a disulfide bridge. However, three out of four tryptophan residues were conserved at position 27, 46 and 83 and other tryptophan was replaced by aspartic acid at residue 60. Similar results were observed for *H. discus hannai* which 4 tryptophans were positioned at residues 24, 43 and 79 in their sequence and aspartic acid at residues 57. Even though, 2 cystein residues were present in the region at residues 36 and 94, they are not conserved with 2 cysteine, which was responsible for disulfide bridge formation. However, they were conserved with 2 cysteine residues at position 33 and 90 out of three in *H. discus hannai.* Further, it has been reported that linker region which connects the N-terminal extended region and the catalytic domain was rich in threonine (33%) and

glycine (16%) residues in *H. discus hannai* however, 21% threonine and only 4% glycine residues were present in the disk abalone. In contrast 17% serine residues were observed and it was absent in *H. discus hannai*. Above facts suggest that N-terminal extended region of the *H. discus discus* also belongs to a family II CBM followed by a linker. The invertebrates, bacteria and plant endoglucanases are grouped into distinct clades reflecting the evolution history of Endoglucanases. HdEndg has the common folding pattern of GHF9 and consist of α/α helix 6 barrel. Three dimensional structural analyses showed that *Nasutitermis takasagoenis* conserved Glu412 is the catalytic acid/base residue and the conserved Asp54 or Asp57 is the base (Khademi, 2002). These conserved residues were observed in HdEndg at positions 566, 207 and 210, respectively, when compare the sequence with extended N-terimnal domain. The two residues participating in acid base catalysis of *Clostridium thermocellum* endoglucanase CelD are Asp201 and Glu555 (Chauvaux et al., 1995). When pH changed from low to higher pH, there was no major structural confirmation changed. The confirmation of the side chains of the catalytic acid/base was changed by rotating from hydrophobic cavity to a relatively hydrophilic environment and this side chain displacement may decrease the enzyme activity at higher pH. Crystalographic analysis indicated that *C. thermocellum* endoglucanase CelD contained 3 $Ca^{2+}$ binding sites and one $Zn^{2+}$ binding site (Chauvaux et al., 1995). Comparison with those sites, HdEndg contained conserved $Ca^{2+}$ binding resdues at Asp247, Asp367, Asp404, and Ser517 and these residues may be responsible for the binding of $Ca^{2+}$ in HdEndg. It has been reported that unlike catalytic residues, none of the residues involved in $Ca^{2+}$ binding is strictly conserved among all catalytic domain of family E cellulase and at present it is difficult to predict from the sequence analysis which of the other members of family E may be stabilized in a similar manner by $Ca^{2+}$ (Chauvaux, et al., 1995).

There are several factors such as temperature, pH, thermostability which important for catalysis of the enzyme. It is important to determine an optimal pH, temperature for the

particular enzyme to know the optimum activity. The purified endoglucanase is active over a broad pH (4.5-8) and temperature up to 50 $^{o}$C and may suitable for industrial application after optimize its activity with different expression system. It has been reported that endoglucanases of some species were stable over a wide range of pH (3.5-7.5) and at temperature up to 65 $^{o}$C and therefore, makes the enzyme suitable for use in cellulose saccharification at moderate temperature and pH. Further, they have shown that due to the $Ca^{2+}$ binding to the enzyme, it stabilizes the enzyme against thermal denaturation and increasing its substrate binding affinity (Chauvaux, et al., 1992). However, these $Ca^{2+}$ sites do not follow the EF-hand pattern observed in many $Ca^{2+}$ binding proteins. It has been reported that *H. discus hannai* contained isoforms ranging from 75-100 kDa. Also there were evident that fungal strains possessed multiple forms of Endgs with different properties. Similarly there may be several isoforms of cellulases in *H. discus discus* in nature and HdEndg is one of the endoglucanase enzymes in the disk abalone. These results suggest that there may multiple endoglucanases, which responsible for cellulose digestion. As *C. quadricarinatus* (Loya-Javellana et al., 1993), this suggest that the endoglucanase partially breakdown plant cell wall polysaccharide and may used by disk abalone to obtain energy (glucose) directly from soluble or amorphous cellulose substrates, or potentially as a tool to gain access to other nutrients within plant cells, or to reduce the digestive viscosity of soluble polysaccharides leached from plant cell walls. It could be suggest that prior disruption of the insoluble or crystalline cellulose component present in plant detritus by microbial organisms may assist in subsequent processing (Crawford et al., 2004) by disk abalone.

Alpha amylase is belongs to glycosyl hydrolase family 13 and is responsible for hydrolysis of α-1,4-glycoside bonds in starch and related substrates. The molecule is highly conserved between organisms showing conserved active site residues, calcium binding and chloride binding residues. HdAmy shows 4 conserved regions (region 1, 111-115; region 2, 208-216; region 3, 248-251 and region 4, 309-316) present in α-amylases of other organisms

and catalytic residues of this family 13 are located in 2, 3 and 4 regions. Similar regions with different amino acid positions were observed in several bacteria (Lim et al., 2003), mammals, lower vertebrates such as Atlantic salmon (Froystad et al., 2006) and other invertebrates (Boer and Hickey., 1986). In contrast to HdAmy, the Atlantic salmon α-amylase sequence showed ten cystein residues which conserved in mammalian amylases that involved in disulfide bond formation. Similarly, conserved active site residues (Asp212, Glu248 and Asp315), calcium (Asn115, Arg173, Asp182 and His216) and chloride (Arg210, Asn 313, Arg352) binding sites residues in Atlantic salmoncies were conserved with different amino acid positions (Froystad et al., 2006). Alpha amylase and related amylolytic enzymes from different organisms shows similar 3D-structure, even though there primary structure is different. The central domain contains α/β barrel and a greek key motif as a separate domain C and with at least one additional domain. Domain B comprises several β strands of variable length, depending on the species. The substrate binding is localized to a cleft between α/β barrel and calcium is required to maintain the structural integrity of the enzyme. Moreover, Removal of calcium leads to decrease in thermostability and or decreased in enzymatic activity (Violet and Meunier, 1989). The HdAmy showed optimum activity at 50 $^{o}$C and pH 7.5. Hypothermophilic bacteria *Thermotoga maritima* MSB8 showed maximum activity at pH 7.0 and its optimum temperature for activity was 70 $^{o}$C (Lim et al., 2003). Purified recombinant protein of *Thermococcus hydrothermalis* showed activity on broad range of pH (5-5.5) and temperature 75-85 $^{o}$C. (Leveque et.al., 2000). In agreement with this, HdAmy also showed activity in a broad spectrum of pH (5.5-8.5) and temperature (30-60 $^{o}$C), even though, the enzyme was not stable after heating at 60 $^{o}$C for 30 min. Like α-amylases of many other organisms, recombinant HdAmy enhanced its activity in the presence of $Ca^{2+}$ and decreased by the presence of EDTA. Similar results were observed for *T. maritima* expressed in *E.coli* in presence of 10 mM $Ca^{2+}$ and 25 mM EDTA (Lim et al., 2003). Thermal inactivation at 60 and 70 $^{o}$C after 30 min incubation suggested that the enzyme was not

thermostable. Therefore, it could be suggested that, lower thermostabilty of HdAmy may be due to lower body temperature of the abalone.

At present primary structure analysis of alginate lyase have been performed mainly on bacterial enzymes. It has been reported that in *H. discus hannai* an Asn corresponding to the carbohydrate chain anchoring residue in *Turbo* SP2 (Asn105) and 5 Cys (Cys-106, 115, 145, 150, and 236) which has been indicated as important residues for catalytic activity and/or disulfide bonding are all conserved (Shimizu et al., 2003). However, in HdAlgl carbohydrate anchoring site was observed at Asn 247 and the 4 Cys residues. It was shown that all those important residues were absent or substituted by different amino acid residues in *Chlorella* virus CL2. This indicates that in the molluscan enzymes these residues are not directly participating in the alginate degrading activity because the *Chlotella* virus can depolymerize alginate. Also, it has been reported that conserved Asn residue may helps to stabilize molluscan alginate lyase enzyme (Shimizu et al., 2003). However, the recombinant protein may liable to heat denaturation, than native protein, since bacterially expressed protein not posses glycosylation. The highly conserved regions showed more than 57% amino acid identity with *Turbo* SP2 and *H. discus hannai* and those regions containing basic and aromatic amino acids have been suggested to act as catalytically and or substrate binding mechanism. To identify the residues involve in catalytic mechanism details studies are needed on the primary and gene structure of the molluscan alginate lyases. Low identity (10-15%) of HdAlgl with other mollusks and bacterial species may be due to varying length of the amino acid sequences and to get an idea about their evolution the phylogenetic relationship need to be studied. It has been reported that alginate lyase was produced in the hepatopancreas and secreted in to digestive fluid in *H. discus hannai* (Shimizu et al., 2003). Therefore, it could be suggest that the HdAlgl is produced in hepatopancreatic cells as a precursor and secreted the enzyme into digestive tract as a mature protein. The purified alginate lyase from combination of digestive fluid and hepatopancreas in *H. discus hannai*

showed 1324 U/mg specific activity (Shimizu et al., 2003). Though, we successfully expressed the HdAlgl in bacterial expression system and purified by pMAL-c2x protein purification system, the specific activity of the recombinant protein was low. The optimal pH and temperature of HdAlgl were 7.0 and 40 $^{o}$C and, respectively that similar to *H. discus hannai* alginate lyase, which was purified from the hepatopancreas and digestive fluid (optimal pH and temperature were pH 8.0 and 45 $^{o}$C, respectively).

Hemicelluloses are the second most abundant polysaccharide in nature and it constitute of xylans, xyloglucans and mannans. Endo β−1,4-mannanase is important for the enzymatic digestion of hemicelluloses (Ootsuka et al, 2006). The β mannanse belongs to GHF 5 and on the basis of the primary structure, *H. discus hannai* β−mannanase have been characterized into GHF5. It has been reported that 16 residues of signal peptide is responsible to release enzyme molecules into the luman of the stomach from digestive gland of *M. edulis* (Xu et al., 2002a) and the presence of signal peptide in HdMann may also responsible for this function. However, they are stored in granulae inside the cells of the digestive gland before release enzyme molecules into the lumen. Similar to HdMann, *H. discus hannai* and other organisms selected in this study showed conserved putative catalytic acid/base and putative catalytic nucleophile. Also as expected for GH 5 famiy enzymes, the predicted 3D-structure showed a (βα) 8-barrel fold. Thus, we concluded that HdMann is also an enzyme belonging to GHF5. Further, HdMann showed 6 cystein residues in the coding sequence suggesting that the presence of 3 disulfide bridges in this protein, however, there was only one disulfide bridge was observed in *M. edulis* protein sequence. Structural studies of *M. edulis* β-mannanase has been shown that conserved two cysteine residues present in the native protein, Cys192 and Cys259, which is responsible to form disulfide bond was not observed due to consequence of the X-ray radiation (Larsson et al., 2006). Besides the two catalytic residues, which in *M. edulis* are the nucleophile Glu308 and the acid/base Glu177 (Xu, et al., 2002b). Six other amino acid residues which highly conserved in the active site are Arg77, Trp79,

Asn176, (catalytic nucleophile), and His277, Tyr279 and Trp337 (Hilge et al., 1998)

It has been reported that mollusk, *Spisula sachalinensis* is a rich source of β-1,3-glucanases and database search of β-1,3-glucanase showed significant homology with the lipopolysaccharide and β-1,3-glucan binding proteins and coelomic cytolytic factors of arthropods and annelids, respectively (Kozhemyako et al., 2004). It is normally considered that the glucanase gene have evolved from ancient β−1,3-glucanase gene, however, the activity of the glucanase has lost during evolution and retained only the glucan binding activity. Therefore, HdLms may also have the glucan binding activity similar to other mollusks and further studies are needed. Low sequence identity may be due to the amino acid length difference between HdLms and known species. In contrast to HdLms, *S. sachalinensis* contained 316 aa, with predicted molecular mass of 38 kDa and Ip of 7.0 (Kozhemyako et al., 2004).

A core motif C/S-X-P-X-R (‘sulfatase signature’) is conserved across the signatures of the entire enzyme class including eukaryotic sulfatases and both bacterial Cys-type and Ser-type sulfatases (Dierks et al., 1997; Dierks et al., 1999). Active sites of all the sulfatases studied carry a Cα-formylglycine (FGly) that is post translationally generated by the oxidation of a conserved cystein in prokaryotic and eukaryotic sulfatases or some bacterial sulfatases the oxidation of serine residue. It has been reported that a characteristic sulfatase peptide segment, containing the cysteine to be modified and some highly conserved surrounding residues, is both necessary and sufficient to direct FGly formation (Dierks et al., 1999). The sequence identity between sulfatases ranging from 20-45% and residues around the active site being highly conserved suggesting that members of this family have a common origin and share a structural fold and the active site. To gain further insight into the mechanism of sulfate hydrolysis by sulfatases the active site of the HdAry was compared with human arylsulfatase B and it showed unique features. It has been shown that modified cystein residue 91 and a metal ion are located at the base of the substrate binding pocket and

that amino acid residue is able to covalently bind sulfate. Also, it has been reported that some of the conserved amino acid residues play a role in stabilizing the calcium ion and the sulfate ester in the active site. Therefore, it formed an intermediate enzyme-sulfate complex (Parenti et al., 1997). Due to the unique features of both human arylsulfatase and HdArys it could be suggests a same mechanism of substrate binding as well as for the reactions they catalyze. Even though, the leucine zipper pattern ($^{96}$LENQPVCLPLNITILPQKLKEL$^{117}$) was observed in HdArys sequence, we could not observe such pattern in other analyzed sequences suggesting that the protein could under go a dimerization. The sequence characterization results revealed that HdArys encodes a full length nucleotide sequence with the main characteristic motifs and functional amino acid residues of the sulfatase family. The high degree of predicted amino acid sequence similarity along the sequence reflects the functional correlation of the enzymes to variety of substrates. Abnormal accumulation of intermediate sulfate compounds results a several human genetic disorders due to the deficiency of single as well as multiple sulfatase activities (Dierks et al., 2003;Obaya, 2006).

HdArys was detected in both cellular pellets as well as in soluble fraction suggesting that the protein has much cellular localization. All members of the sulfatase family characterize the hydrolysis of sulfate ester bonds in a variety of substrates, ranging from complex molecules such as glycosaminoglycans and sulfolipids, to simple ones such as beta hydroxysteroid sulfates (Parenti et al., 1997). The sulfatase activity was measured using chromogenic or fluorogenic aromatic compounds sulfated compounds such as p-nitrophenol sulfate (pNPS), p-nitrocatechol sulfate or 4-methylumbelliferyl sulfate (4-SUM) without considering the physiological substrate and it has been reported that initial classification of human sulfatases as arylsulfatases and non arylsulfatases were determined according to this (Obaya, 2006). In this study, even though, HdArys showed low activity, as there was activity towards pNPS and nitrocataecholsulfate we could be suggested that this enzyme was belongs to arylsulfatase family.

In general purified recombinant PDEs in pMAL-c2X expression system showed low activity compared with reported PDEs in other species. This low activity may be due to errors occurred in protein folding or lack of post translational modifications in the *E. coli*. Therefore, efficient eukaryotic protein expression system where the post translational modification could take place might help to produce active form of proteins.

Abalone PDEs expression profile is constructed in the hepatopnacrease based on the food availability to advance understanding of digestive regulation. We observed a significant decrease of weight loss in disk abalone in 8 weeks starvation. This may be due to loss of carbohydrate and protein in the tissues. It has been reported that 13.44% wet weight loss of green abalone *H. fulgens* after 27 days of starvation was principally due to loss of protein and carbohydrate in the tissue suggesting that protein and carbohydrate are being used as energy sources (Viana et al., 2007). Similar results was observed by Durazo-Beltrán et al. 2004 and reported that 18.8% decrease in the weight of juvenile *H. fulgens* with no mortality after 60 days of starvation. Carefoot et al., 1993 reported that no mortality of the abalone (*H. kamtschatkana*) after 27 days of starvation and that a significant weight loss occurred under basal metabolism conditions. Upon resumption of feeding, returned to normal metabolism.
In this study, after 6$^{th}$ week of starvation the disk abalone gonad was totally disappeared and size of the hepatopancrase has been reduced to about 80% from the original. Further, colour has been changed from blackish green to brown due to the lack of food. However, after 2$^{nd}$ week of re-feeding, the appearance of the hepatopancrease of disk abalone recovered to normal suggesting that it reached to normal metabolism level. Similar physiolological response has been observed in mollusk bivalves, and depending on the food availability that increased the size of digestive tubules in cockles, resulting in a volume increase of the whole digestive organ (Ibarrola et al., 2000). The increase in size of the digestive structures was linked to synthesis of digestive enzymes, leading to an increase in their total activity (Ibarrola et al., 1996; Ibarrola et al., 1998). It has been reported that main organ, which responsible for

digestive enzyme production is hepatopancrase. All the PDE related genes were expressed at different levels in hepatopancrase and digestive tract tissue. A strong preferential expression of all the genes in hepatopancrase indicates that the both intra and extra cellular digestive function may takes place in hepatopancrase of the disk abalone. This result is in accordance with the digestive function of this organ in *C. gigas,* which carries out both extra and intracellular digestion (Morton, 1983; Henry et al., 1993). Higher expression of these genes indicated that all these genes appeared to be more significant for abalone digestive process and plays an important role in sea weed degradation as major cell wall constituents are polysaccharides.

In most bivalves, amylase is the most common digestive enzyme encountered in the digestive gland (Seiderer and Newell, 1979; Teo and Sabapathy, 1990; Yan et al., 1996; Le Moine et al., 1997) and, is also found in secretory granules and in gastric cells (Henry et al., 1993). Amylase is also incorporated into the crystalline stylet where it contributes to extracellular digestion of foods. Small amounts of α-amylase mRNA expression was observed in non digestive tissues with no digestive functions (gills, mantle, muscle and labial palps) in *C. gigas*, but the functional role of the amylase enzymes was unclear. It has been reported that digestive gland is the most important site for storages/secretion of enzymes in juvenile red abalone *H. rufescens* and the stomach-digestive gland region characterized by the presence of high amounts of complex carbohydrases such as cellulase (Esquivel and Felbeck, 2006). In previous studies also reported the presence of endogeneous CMCase, alginate lyase, laminarinase, agarase and carragenase in the digestive gland. It has been reported that *Mytilus edulis* also produces a β-mannanase, from the digestive gland of the mollusk (Xu et al., 2002a). The hypothesis of a physiological role of amylase, in addition to their strict digestive function, is supported by reports of amylase activities in mammals. When compare with all the genes the amylase expression was some what higher through out the starvation. This may be due to utilization of glycogen as an energy source during starvation.

It has been reported that, liver specific amylases displayed a strong affinity to glycogen in rats and were found to be capable of glycogen hydrolysis and glucose formation (Koyama et al., 2001). Furthermore, α-amylase has been reported to be capable of hydrolysis of oyster glycogen (Matsui et al., 1996). Higher expression of alginate lyase in hepatopancreas may be due to the type of abalone seaweed used and during this period abalone controls were fed continuously with brown sea weed where the main structural polysaccharide was alginate. The scarcity or total absence of food during shorter or longer periods is a characteristic for marine ecosystem and, that conditions the physiology of the animals that inhabit them. Mollusks, as part of their natural life cycle, are able to withstand long periods of time without food. They are well adapted to mobilize their metabolic reserves and body constituents to survive periods of food deprivation. Mollusk, especially the bivalves are subject to an annual cycle of accumulation and use of energy reserves associated with gametogenesis and this cycle is influenced by environmental parameters such as food availability and temperature (Gabbott, 1976). An understanding of the metabolic processes and the capacity of response in total absence of food allows us to be more informed as to the ecology of these species. Carbohydrates made the largest contributions to energy output, followed by proteins and lipids. The energy needed to maintain vital functions in organisms is obtained from catabolism of body components, with a reduction in dry mass of the organisms, and consequently their energy content. When the food is available the mRNA expression of digestive enzymes were high showing a regulation pattern for food availability. This pattern related to the increasing number of mRNA transcripts of the enzymes. Previous studies concerning the digestive enzymatic apparatus of bivalve molluscs have shown that the digestive gland is rich in α-amylase (Reid, 1966, Reid,1968; Mathers, 1973). This enzyme is scarce during the winter, however, its mRNA transcripts are abundant from March, when phytoplankton blooms begin, until September. In fact, its presence suggests a correlation with that of blooms in spring and summer (Paulet et al., 1997). Moreover, recent nutritional

experiments *in vivo* have shown that increased amylase activity is positively correlated with food inputs, as reported for another bivalve, *Cerastoderma edule* (Ibarolla et al., 1998). In this study, a relationship could be suggested between an external factor, i.e. food availability, and an internal factor, the increase in α-amylase and other PDE genes mRNA transcript levels. Therefore, these genes can be used as an applied molecular tool to monitor the food availability over the year and thereby to monitor the metabolic events occurring during this period in the hepatopancrease.

In the present study, decreasing pattern of mRNA expression over the starvation was observed. The usual response to nutritive stress is an increase or decrease in metabolic rate according to whether food is present or not, so that when faced with a prolonged absence of food the metabolic rate decreases to minimum values which are then maintained throughout the starvation period (standard metabolism). This strategy enables the animal to reduce its energy consumption and save energy at a time – starvation or nutritive stress – when energy gain is nonexistent or only very slight (Bayne and Newell, 1983). It has been reported that significant reduction in metabolic rate (oxygen consumption rate) over the first 14 days of starvation in *V. pullas*tra than in *R. decussates.*

We demonstrated the ability of the polysaccharide degrading genes to be regulated at the transcriptional level when food varied. Hence, food availability constitutes an external regulatory factor for those genes. Similar results have been reported for amylase A gene in the Pacific oyster *C. gigas*. Diet quality is likely to have an effect on amylase activity because different classes of algal species vary in carbohydrate and starch content (Moal et al., 1987). Regulation in amylase activity and active cytoplasmic mRNAs was reported at the cellular and molecular levels due to changes in the carbohydrate content of diets in rats (Wicker et al., 1984). Adaptation of digestive enzymes, such as amylase, was also observed in shrimp, depending on casein and protein sources in diets (Le Moullac et al., 1997). These examples emphasize the potential effects of food quality on regulation processes of polysaccharide

degrading genes. Arylsulfatase gene is also mainly expressed in the digestive gland at a higher level, however, its expression does not respond much to variation of food availability. This may be due to not direct involvement of arylsulfatase in abalone seaweed digestion.

ROS are produced in a series of biochemical reactions that will normally occur within the cellular compartments (i.e., mitochondrium and the endoplasmic reticulum are the most important ROS sources) (Halliwell and Gutteridge, 1999). Oxidative stress can be understood as a situation derived either from an enhanced rate of ROS generation and a reduced level of antioxidant defenses (Sies, 1985). It is well known that a number of different enzymes and non-enzymatic compounds participate in the antioxidant chain in biological systems. Among the enzymes, superoxide dismutase (SOD) converts superoxide anion ($O_2^-$) to hydrogen peroxide ($H_2O_2$), catalase (CAT) reduces $H_2O_2$ to water. SOD has been used as biomarkers of cellular stress and toxicity in different aquatic organisms such as mollusks, fish etc. starvation has been reported to have pro-oxidant effects in mammals and being considered responsible for most of the detrimental effects derived from food deprivation, as increased ROS generation is not adequately neutralized by antioxidant systems (Robinson et al., 1997). Studies regarding the influence of food deprivation on antioxidant defenses in invertebrates are limited. Therefore, to know whether the hepatopancrase oxidant-antioxident status may be altered by starvation and re-feeding we conducted semi quantitative RT-PCR analysis using gene specific F and R CuZnSOD primers. Since, SOD are the first line defense against toxic intra cellular radicals produced during normal cellular metabolism or under oxidative stress (Fridovich, 1995), we chose CuZnSOD gene for our experiment. It has been reported that enhanced SOD mRNA expression level in *Hydra vulgaris* which starved for 5 days as compared to control at both 18 and 30 $^{o}$C (Dash et al., 2007). Starvation enhanced reactive oxygen species generation in the liver of *Dentex* dentex and increased SOD activity during 5 weeks starvation was observed (Morales et al., 2004). Further studies reported that the antioxidant defenses were not able to scavenge total ROS. In sea bream (*S. aurata*)

deprived of food for 46 days, an increase of hepatic SOD activity has also been reported (Pascual et al., 2003). Similarly in this study it was observed that increased mRNA expression during starvation and decreasing level of expression during re-feeding. This suggests that starvation caused rise in the $O_2^{\cdot -}$ rate with the prooxidant situation in disk abalone and, SOD transcription level was increased to translate SOD to reduce the oxidative stress caused during starvation.

The disk abalone may able to adapt its digestion according to food availability in the environment. Such an adaptative response of digestive enzymes will probably affect digestion and absorption efficiencies for the corresponding substrate. Studying the variability of such a response at the level of an individual would be of great interest in the explanation of differences in growth rates among individuals. The present study shows the PDEs which are associated with abalone digestive processes, and its expression is mainly localized in the hepatopancrase, where it is regulated according to variation of food availability at the transcription level. Increase in food quantity also resulted in cellular growth of the digestive, and may leads to an increase of total enzyme activity with probably an increase of cell number. The disk abalone could survive for 8 weeks of starvation and alterations induced by starvation reversed after re-feeding such as weight and mRNA transcription levels related to PDEs and oxidative stress related genes.

The physiological significance of PDEs in the utilization of polysaccharide substrates as energy and carbon sources in abalone is an important task for future research. The longer term application of studies on evolution and function of PDEs in marine species nutrition may assist the development of improved artificial diets for cultured aquatic species through the incorporation of low-cost feed components sourced from plant materials.

# PART 2

# Molecular Characterization and Expression Analysis of Calcium Regulatory Protein Regucalcin from Disk Abalone *Haliotis discus discus*

## ABSTRACT

Regucalcin is a novel calcium ($Ca^{2+}$) binding protein and it has been demonstrated to play a multifunctional role in many organisms. Here, we report the molecular cloning of invertebrate regucalcin cDNA from disk abalone *Haliotis discus discus*. The full length cDNA showed 1321 bp of nucleotides with a polyadenylated sequence (AATAAA). Abalone regucalcin (HdReg) open reading frame (ORF) consists of 915 nucleotides encoding 305 amino acids (aa). Estimated molecular mass was 33 kDa and predicted isoelectric point (pI) was 4.9. The HdReg amino acid sequence did not contain the EF-hand motif as a $Ca^{2+}$ binding domain, suggesting a novel class of $Ca^{2+}$ binding protein. Moreover, it showed 45% identity to chicken and zebrafish, and 44% to rat and mouse regucalcin in deduced amino acid level. The tissue expression analysis of HdReg mRNA was investigated by RT-PCR and it was expressed in all the tissues tested such as gill, mantle, digestive tract, and abductor muscle. Semi-quantitative RT-PCR results showed that an intramuscular administration of calcium chloride ($CaCl_2$) (0.5mg $CaCl_2$/g of abalone) could significantly induce regucalcin mRNA in abductor muscle after 30 min of administration and reached maximum after 1h. Subsequently, the expression level was decreased after 2h. This indicates that the expression of regucalcin mRNA is constitutive, and specifically up regulated in abalone abductor muscle by $Ca^{2+}$ administration.

## INTRODUCTION

Calcium ($Ca^{2+}$) is an important ion for many organisms. It acts as a second messenger in the regulation of many cellular and physiological functions such as muscle contraction, neuronal activation, cell differentiation and cell death. There are many energy dependent $Ca^{2+}$ transporters and $Ca^{2+}$ channels are present in the plasma membrane and membranes of intracellular organs. Intra-cellular $Ca^{2+}$ levels and $Ca^{2+}$ signals are mediated by those energy dependent $Ca^{2+}$ transporters and $Ca^{2+}$ channels (Osterloh et al., 1998). Numbers of $Ca^{2+}$ binding proteins are involved in regulation of $Ca^{2+}$ levels within the cytoplasm. Currently, the largest groups of $Ca^{2+}$ binding proteins are the EF-hand superfamily genes containing EF-hand motifs (Nakayama and Kretsinger, 1994). It has been reported that most of these proteins such as calmodulin, troponin C, and myosin light chain maintain intracellular $Ca^{2+}$ homeostasis and transmit $Ca^{2+}$ signals to specific target proteins by the activation of their $Ca^{2+}$ bound conformation (Osterloh et al., 1998).

Regucalcin is a regulatory protein of $Ca^{2+}$ and also known as senescence marker protein-30 (SMP30). It was discovered in 1978 as a novel $Ca^{2+}$ binding protein in rat liver cells (Yamaguchi and Yamamoto, 1978; Yamaguchi and Sakurai, 1992). EF-hand motif as a $Ca^{2+}$ binding domain was absent in regucalcin and is different from calmodulin and other $Ca^{2+}$ related proteins. Multifunctional roles of regucalcin, related to regulation of cellular functions in rat liver, kidney, and brain neuronal cells have been reported. It regulates intracellular $Ca^{2+}$ homeostasis by enhancing $Ca^{2+}$ pumping activity in the plasma membrane (Kurota and Yamaguchi, 1997; Takahashi and Yamaguchi, 1997) and mitochondria of liver and renal cortex cells through activation of $Ca^{2+}$ dependent ATPases (Yamaguchi, 2005). Regucalcin has a reversible effect on the $Ca^{2+}$ induced activation and inhibition of many liver and renal cortex cell enzymes (Yamaguchi and Sakurai, 1992). It also inhibits various protein kinases (including $Ca^{2+}$ /calmodulin- dependent protein kinase, protein kinase C and tyrosine kinase) and protein phosphatases due to binding of $Ca^{2+}$, indicating a regulatory role in signal

transduction within the cell. Moreover, it has been reported that regucalcin can inhibit synthesis of RNA in the nuclei of normal and regenerating rat liver in vitro (Yamaguchi and Ueoka, 1997).

Regucalcin gene was identified in many vertebrates and its cDNA was cloned for human (*Homo sapiens*, **NP_004674**); rat (*Rattus norvegicus*, **BAA07490**); rabbit (*Oryctolagus cuniculus*, **BAA88079**); mouse (*Mus musculus*, **NP_033086**); bovine (*Bos Taurus*, **BAA88080**), and toad (*Xenopus laevis*, **BAA90694**). The deduced amino acid sequence of anterior fat body protein (8-277aa region) from insect flesh fly (*Sarcophaga peregrina*) exhibited similarity to that of mammalian regucalcin (SMP30) (Nakajima and Natori, 2000). Moreover, on the search of NCBI GenBank, we observed other organisms containing the regucalcin family or a regucalcin homologue, such as bacteria *Bacillus cereus* ATCC 10987 regucalcin family protein, (**NP_978918**); fungi *Aspergillus fumigatus* Af293 regucalcin putative, (**XP_751966**); and, invertebrates like *Drosophila melanogaster* regucalcin homologue, (**BAA99283**). However, their genetic and functional characteristics are not yet clearly understood.

Shimokawa and Yamaguchi (1992) have reported that the expression of regucalcin mRNA is specific in the liver, and that it is increased by the administration of $CaCl_2$ to rats. $Ca^{2+}$ is a regulatory agent involved in many physiological processes of invertebrates such as mollusks and it is also the primary cation in the formation of shell structures. It is a product of $Ca^{2+}$ metabolism, which contains more than 90% of $CaCO_3$ crystals (Addadi and Weiner, 1997). In marine bivalves, calcium is taken up by gill from the external medium and transported to the mantle epithelium. L-type $Ca^{2+}$ channels, which are regulated by calmodulin, have been suggested to be involved in calcium transport process for calcification in some marine invertebrates (Zoccola et al., 1999). However, the mechanism of $Ca^{2+}$ uptake, accumulation, transport, incorporation, and the particular regulators involved in $Ca^{2+}$ metabolism remains an interesting field for investigation. Hence, the structural analysis of

novel class of $Ca^{2+}$ binding proteins would be helpful to investigate the novel architecture of $Ca^{2+}$ recognition and in understanding the mechanism of intracellular $Ca^{2+}$ signal transduction.

The relationship between vertebrate and invertebrate regucalcin is curious, especially since no reported data on invertebrate regucalcin gene and its functional characterization exist according to our knowledge. To gain insight of the regucalcin gene in mollusks, we established the primary molecular structure of disk abalone, *H. discus discus*, and compared its sequence with known regucalcin molecules from other animals. Further, the expression of its mRNA was analyzed with or without intramuscular injection of $CaCl_2$.

## MATERIAL AND METHODS

### Cloning and sequencing of abalone regucalcin cDNA

*H. discus discus* regucalcin cDNA was obtained from the normalized cDNA library, which was synthesized using a Creator SMART cDNA library construction kit (Clontech, USA). The cDNA was normalized with Trimmer-Direct cDNA normalization kit according to the manufacture's protocol (Evrogen, Russia). The plasmid DNA of the putative regucalcin was obtained by Accuprep$^{TM}$ plasmid extraction kit (Bioneer Co., Korea). The full length sequence was obtained by designing the internal primer HdReg-1I (Table 1) from the known sequence of 3' end and sequenced using termination kit, Big Dye and an ABI 3700 sequencer (Macrogen Co., Korea).

### Abalones

The disk abalones were obtained from Fisheries Resources Research Institute (Jeju, Republic of Korea). Individuals were averaging 80.0±5.0 mm and 60.0±5.0 g in length and body mass, respectively. They were maintained in 40 L tanks with an aerated seawater having temperature of 20±2 $^{o}$C and fed with seaweeds for 7 days to acclimatize to laboratory conditions before $CaCl_2$ injection.

### $CaCl_2$ administration and tissue collection

$CaCl_2$ was dissolved in sterile distilled water and abalones were injected with 0.5 mg $CaCl_2$/g of abalone intramuscularly. Untreated abalones were kept as controls for the experiment. Gill, mantle, digestive tract, and abductor muscle tissues were collected separately from uninjected animals. Same tissues were collected from $CaCl_2$ injected abalones at 30 min, 1h, and 2h. Three (n=3) abalones were used at each time point of $CaCl_2$ injection and the control. Tissues were immediately frozen in liquid nitrogen and stored at -70 $^{o}$C until RNA isolation. Three independent experiments were conducted for the study.

**RNA extraction and cDNA construction**

Total RNAs of the gill, mantle, digestive tract and abductor muscle tissues (100 mg each) were isolated using Tri Reagent$^{TM}$ (Sigma, USA) according to the manufacturer's protocol. The first strand cDNA synthesis was carried out based on SuperScript III First Strand cDNA synthesis kit (Invitrogen, USA). In brief, 1 μL of 50 μM oligo$(dT)_{20}$ and 1 μL of 10 mM dNTP were incubated with isolated 2.5 μg of RNA for 5 min at 65 $^{o}$C. After incubation, 2 μL of 10 x cDNA synthesis buffer, 4 μL of 25 mM $MgCl_2$, 2 μL of dithiothreitol (DTT, 0.1 M), 1 μL of RNAsaseOUT$^{TM}$ (40 U/μL), and 1 μL of SuperScript III reverse transcriptase (15 U/μL) were added and incubated for 1 h at 50 $^{o}$C. The PCR reaction was terminated by adjusting the temperature to 85 $^{o}$C for 5 min and the resulting cDNA was stored at -20 $^{o}$C for further experiments.

**Semi quantitative RT-PCR analysis**

RT-PCR was performed to study the regucalcin mRNA expression in disk abalone tissues including gill, mantle, abductor muscle, and digestive tract. Furthermore, tissue specific relative mRNA expression among different tissues and relative mRNA expression levels of $CaCl_2$ injected tissues were analyzed using semi-quantitative RT-PCR. For comparison of relative regucalcin mRNA levels, statistical analysis was performed with one way ANOVA and mean comparisons were performed by Duncan multiple range tests using SPSS 11.5. Significant P values were obtained by Duncan multiple range test. A pair of regucalcin gene specific primers HdReg-1F and HdReg-1R was used to amplify a 444 bp fragment of HdReg. As an internal control, 420 bp disk abalone ribosomal protein cDNA (GenBank: EF103427) was amplified using ribosomal 2F and ribosomal 2R primers (Table 5).

RT-PCR was carried out in a TaKaRa PCR thermal cycler in 25 μL reaction volume containing 2 μL of cDNA from each tissue, 2.5 μL of 10x NEB buffer, 2.0 μL of 2.5 mM dNTP mix, 1.0 μL of each primer (20 pmol/μL, including regucalcin and ribosomal protein

forward and reverse primer), and 0.13 μL (5 U/μL) of NEB *Taq*™ DNA polymerase (New England BioLab). The PCR temperature conditions for regucalcin were 94 °C for 2 min followed by 30 cycles of 94 °C for 30 sec, 57 °C for 30 sec, and 72 °C for 30 sec. For ribosomal protein, it was 94 °C for 2 min followed by 23 cycles of 94 °C for 30 sec, 55 °C for 30 sec, and 72 °C for 30 sec. After the final cycle, samples were incubated for a further 5 min for 72 °C, and then held at 4 °C prior to analysis. The PCR products were analyzed by 1.5% agarose gel electrophoresis containing ethidium bromide and 100 bp molecular marker (TaKaRa, Japan). Electrophoretic images and the optical densities of amplified bands were analyzed using the Scion Image software (Release alpha 4.0.3.2).

**Sequence analysis**

Sequence similarity analysis was performed with BLAST algorithm at the NCBI BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). The HdReg deduced amino acid sequence was analyzed for signal sequence and motif prediction using ExPASy (http://www.au.expasy.org). Multiple alignment of the HdReg was performed using ClustalW multiple alignment program (http://www.ebi.ac.uk/clustalw/). MEGA 3.1 (Kumar et al. 2004) was used to produce the phylogenetic tree using the neighbor joining (NJ) method.

**Table 5:** Primers used for HdReg RT-PCR expression analysis

| Name | Object | Sequence (5' to 3' direction) |
|---|---|---|
| HdReg-1I | Internal sequencing | GATATGGGTTACCCGGATG |
| HdReg-1F | RT-PCR amplification | CGCCAATATGTTCAACGACGGCAA |
| HdReg-1R | | TTGACAGTACGGATCACCTTGCCA |
| Ribosomal-2F | RT-PCR positive control | GGGAAGTGTGGCGTGTCAAATACA |
| Ribosomal-2R | | TCCCTTCTTGGCGTTCTTCCTCTT |

## RESULTS

### Sequence characterization of regucalcin

We characterized the primary molecular structure of *H. discus discus* regucalcin gene. The regucalcin cDNA was obtained from the disk abalone normalized cDNA library and it showed similarity to known sequences of regucalcin by BLAST analysis. The full length sequence was obtained using terminator reaction kit, Big Dye, and ABI 3700 sequencer (Macrogen Co., Korea). The full length HdReg cDNA consisted of 1321 bp with a 5' untranslated region (UTR) of 75 bases followed by an ORF of 915 bp and a 3' UTR of 329 bp (Figure. 32). It contained a polyadenylation signal with AATAAA sequence and a poly (A) tail. The ORF can encode a polypeptide of 305 amino acids. The estimated molecular mass and predicted isoelectric point were 33 kDa and 4.9, respectively. A search for potential N-glycosylation sites using the PROSITE program (http://kr.expasy.org/prosite/) showed N-glycosylation site at the amino acid positions 151-154 (NITI). The amino acid composition of regucalcin shows a relatively high content of valine (10%), aspartic acid (9%), glycine (9%), and serine (8%). The pairwise CLUSTALW analysis of deduced amino acid sequences was performed to show identity percentage (%) of HdReg with different species of regucalcin (Table 6). It showed 45% identity with chicken and zebrafish and 44% to rat and mouse regucalcin. Moreover, it showed 42% identity to human regucalcin. *S. peregrine* anterior fat body protein and *D. melanogaster* regucalcin homologue were 33 and 32% identical to HdReg, respectively. ClustalW multiple sequence alignment of HdReg with other regucalcin species in database is shown in Figure 33. The overall identity of HdReg with those selected vertebrate species was 34%. Of 29 aspartic acid residues in HdReg, 11 residues (38%) were conserved with vertebrate species. However, only 3 of 12 (25%) glutamic acid residues were conserved with all other vertebrate regucalcin aligned in this study. Moreover, using the Kyte and Doolittle method (1982), we analyzed the hydrophilicity profile of regucalcin and hydrophilic as well as hydrophobic residues in both N-terminal and C-terminal regions of the

regucalcin molecule were detected. It showed 66% of hydrophilic residues at positions between 100 and 200 of the amino acid sequence and as a molecule, regucalcin could be considered a hydrophilic molecule.

Phylogenetic analysis was carried out to find out the evolutionary position of HdReg. The analysis was carried out using MEGA 3.1 software with Neighbor-joining (NJ) method and boot strap values were taken from 1000 replicates. The tree was constructed using 15 organisms. Analysis results showed that the deduced amino acid sequence of HdReg was clearly placed in a single cluster and all the other vertebrate and invertebrate regucalcin were placed in distinct separate clusters (Figure 34). Two insects, *S. peregrine* anterior fat body protein and *D. melanogaster* regucalcin homologues, were placed in one sub cluster. The two bacterial regucalcins *Xanthomonas axonopodis* pv. citri str. 306 (**AAM36628**) regucalcin and *X. oryzae* pv. oryzae MAFF 311018 (**BAE69525**) regucalcin were used as an out-group for the analysis.

**Expression of regucalcin in different tissues**

The tissue specific HdReg expression was analyzed *in vivo* by amplifying 444 bp regucalcin cDNA fragment. A constitutive expression gene, 420 bp abalone ribosomal protein was used as a housekeeping gene. RT-PCR results showed that regucalcin mRNA was constitutively expressed in all the tissues tested such as gill, mantle, digestive tract, and abductor muscle. The tissue distribution of regucalcin mRNA in abalone is shown in Figure 35A and 35B and tissue specific variation of mRNA expression was observed. Significantly highest and lowest ($P<0.05$) mRNA expression levels were detected in abductor muscle and mantle, respectively.

```
                                                                AAGATA     6
GAAGGAGTTGGTACCCAGCGACAGAGAGAACATATTCGCGAGTATTAAGCGACTCAGCATATTACGATT   75
ATGTCTGAGAAGATCGAGGTCGTTGTGAAGAACTGTGTTAAGCATGGGGAGGGTCCCCACTGGGACGAC  144
M--S--E--K--I--E--V--V--V--K--N--C--V--K--H--G--E--G--P--H--W--D--D--   23
ACTACGCAGTGCCTGTACTACGTGGACGTGTTCGTTAACGACGTCCACAGATACAACGCCGTCACGGGA  213
T--T--Q--C--L--Y--Y--V--D--V--F--V--N--D--V--H--R--Y--N--A--V--T--G--   46
CAGGACAGCCGACTAAACCTTGGAGACTCTGTGGTGAGCGTCATCATCCCCCGCCAGAAGGGAGGCTTT  282
Q--D--S--R--L--N--L--G--D--S--V--V--S--V--I--I--P--R--Q--K--G--G--F--   69
GTGGTGAGTCAGAAGACAGACCTCGGCCTCGTGGACTCATGGGAGAGCGGCAAGGTCACCAGCATTGCG  351
V--V--S--Q--K--T--D--L--G--L--V--D--S--W--E--S--G--K--V--T--S--I--A--   92
AAGGTCGAGGCCGACGCCAATATGTTCAACGACGGCAAAGTGGATGCTTCCGGTAGACTCTGGATTGGG  420
K--V--E--A--D--A--N--M--F--N--D--G--K--V--D--A--S--G--R--L--W--I--G--  115
TCTTGTTTCTGTGGCAACGGTATCGCTGACCCCAACAAGTGGCTGAAGCTTCAGGGCTCATTGTACAGC  489
S--C--F--C--G--N--G--I--A--D--P--N--K--W--L--K--L--Q--G--S--L--Y--S--  138
CTGGATAGCGATGGCAGCGTGAAAAAACACATCAGTAACATCACCATCAGCAACGGCATGGACTGGACT  558
L--D--S--D--G--S--V--K--K--H--I--S--N--I--T--I--S--N--G--M--D--W--T--  161
GACGACAACAGCGTCATGTACTACATCGACTCCGTCCCCAGGCAGGTCTGGGCCTTCGACTTTGACATC  627
D--D--N--S--V--M--Y--Y--I--D--S--V--P--R--Q--V--W--A--F--D--F--D--I--  184
ACAGCTGGGGCCATCAGCAACAAGAGAACTGTAGCAGATTTCAGCTCAACTGAAATCCCAGATATGGGT  696
T--A--G--A--I--S--N--K--R--T--V--A--D--F--S--S--T--E--I--P--D--M--G--  207
TACCCGGATGGAATGGCCATTGATACGGAGGGGAAATTGTGGGTGGCCTGTTTTGACGCTGGGAAAGTT  765
Y--P--D--G--M--A--I--D--T--E--G--K--L--W--V--A--C--F--D--A--G--K--V--  230
GTCCGACTCGACCCCGAGACTGGCAAGGTGATCCGTACTGTCAAGTTCCCATGTGAAAAGATAACTTCC  834
V--R--L--D--P--E--T--G--K--V--I--R--T--V--K--F--P--C--E--K--I--T--S--  253
TGTTGTTTCGGCGGGGACGACCTCACCGACCTCTACGTCACGTCATCGCCGTACATGATGACCGAAGCA  903
C--C--F--G--G--D--D--L--T--D--L--Y--V--T--S--S--P--Y--M--M--T--E--A--  276
GAATTGGAGAAAACACCACTGGCTGGCTCAATCTTCAAGGTCACCGGGTTAGGAGTTCGAGGCCGCAAA  972
E--L--E--K--T--P--L--A--G--S--I--F--K--V--T--G--L--G--V--R--G--R--K--  299
GCTAACATGTTCTCTGGAtaaATGTTTTCCGTTCAATGCATTTTCAAAGACAAATATTTTCAACCTTAC 1041
A--N--M--F--S--G--***                                                  305
GCTTTGTGGACCGGGGCCTTGGTCCTGAATGCCCCAACACTTTCCAACATGAAACTAACATCTTGAAGC 1110
AGTCCAACTTTGTGAGATACCCCTAGCTACATTCTGGTCTACCCGCAGCCTTCGAAATTCACGCTAGTC 1179
CGCTAGCCCGAGACTACTTAGTTGTACTAAGGGCAAGTAAAAATGATGACTACTACTATTTGAGTGTTA 1248
TAACAGTTTACTTATACTACAACTTGGAATTGAATAAAACTGATAAAATATGAAAAAAAAAAAAAAAAA 1317
AAAA                                                                   1321
```

**Figure 32:** The complete nucleotide and deduced amino acid sequences of HdReg. The predicted amino acid sequence was presented below the nucleotide sequence. Position one of

the amino acid sequence represents methionine (Met) residue. The termination codon (TAA) was indicated by ***. The putative polyadenylation signal AATAAA is boxed with dash lines and N-glycosylation site is boxed. The poly A tail is at the end marked in bold letters. The aspartic acid residues were marked in bold. This disk abalone regucalcin cDNA sequence has been submitted to GenBank under the accession number EF103358.

**Table 6:** Pairwise CLUSTALW analysis and comparison of the deduced amino acid sequence of HdReg protein with other vertebrate regucalcin.

| Species | NCBI accession no. | Identity (%) | Amino acids |
|---|---|---|---|
| *Gallus gallus* | BAA90693 | 45 | 299 |
| *Danio rerio* | NP_991309 | 45 | 295 |
| *Rattus norvegicus* | BAA07490 | 44 | 299 |
| *Mus musculus* | NP_033086 | 44 | 299 |
| *Sus scrofa* | NP_001070688 | 43 | 299 |
| *Xenopus laevis* | BAA90694 | 43 | 299 |
| *Oryctolagus cuniculus* | BAA88079 | 42 | 299 |
| *Bos taurus* | BAA88080 | 42 | 299 |
| *Homo sapiens* | NP_004674 | 42 | 299 |

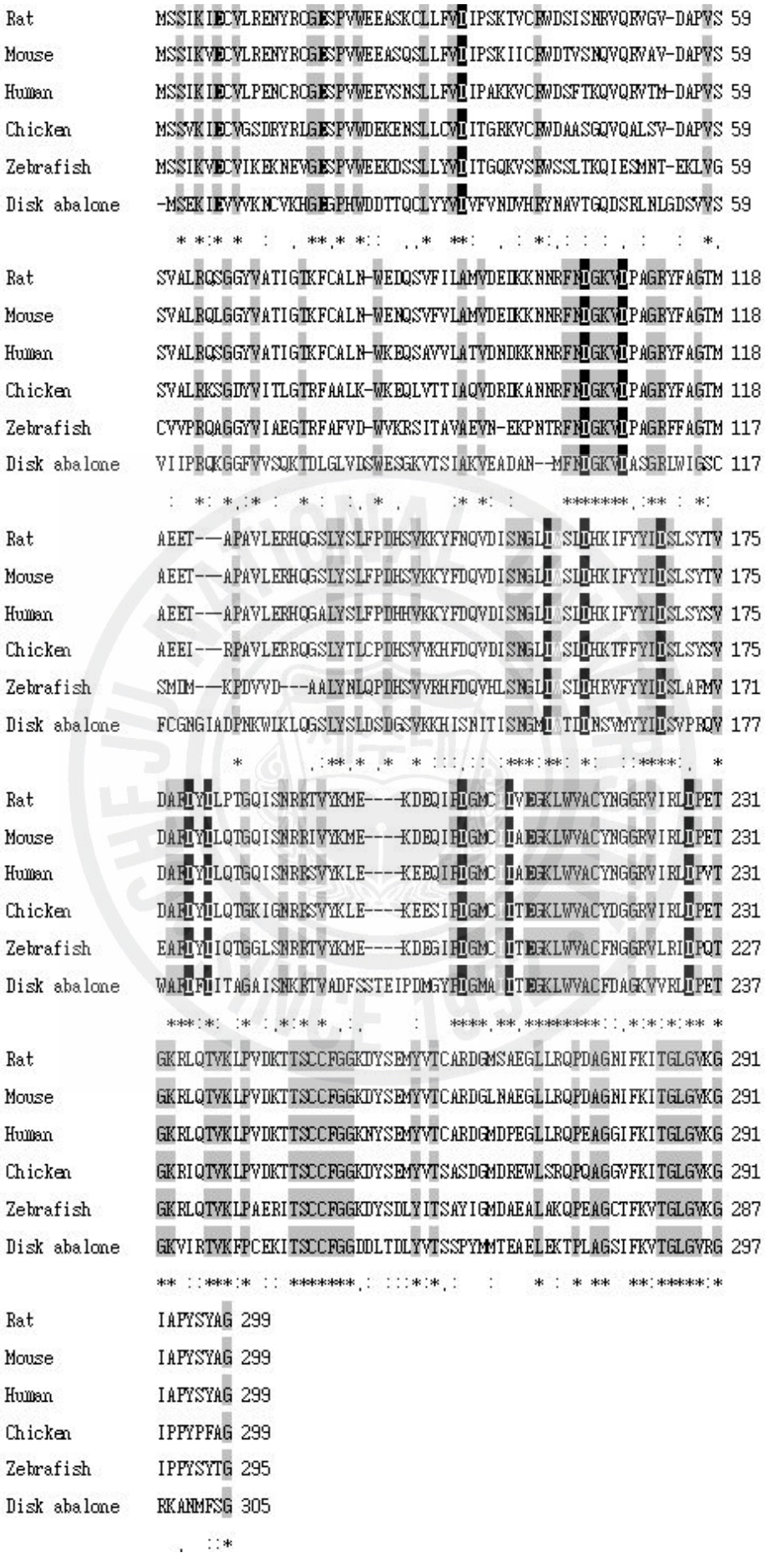

Rat
Mouse
Human
Chicken
Zebrafish
Disk abalone

**Figure 33:** ClustalW multiple sequence alignment of HdReg with known regucalcin. The GenBank accession numbers for the vertebrate regucalcin sequences are same as the numbers used in Figure 34. Asterisk marks indicate identical amino acids and the numbers to the right indicate the amino acid position of regucalcin in the corresponding species. Conserved aspartic acid residues are highlighted in black and conserved glutamic acid residues are in bold type among other vertebrate regucalcin aligned.

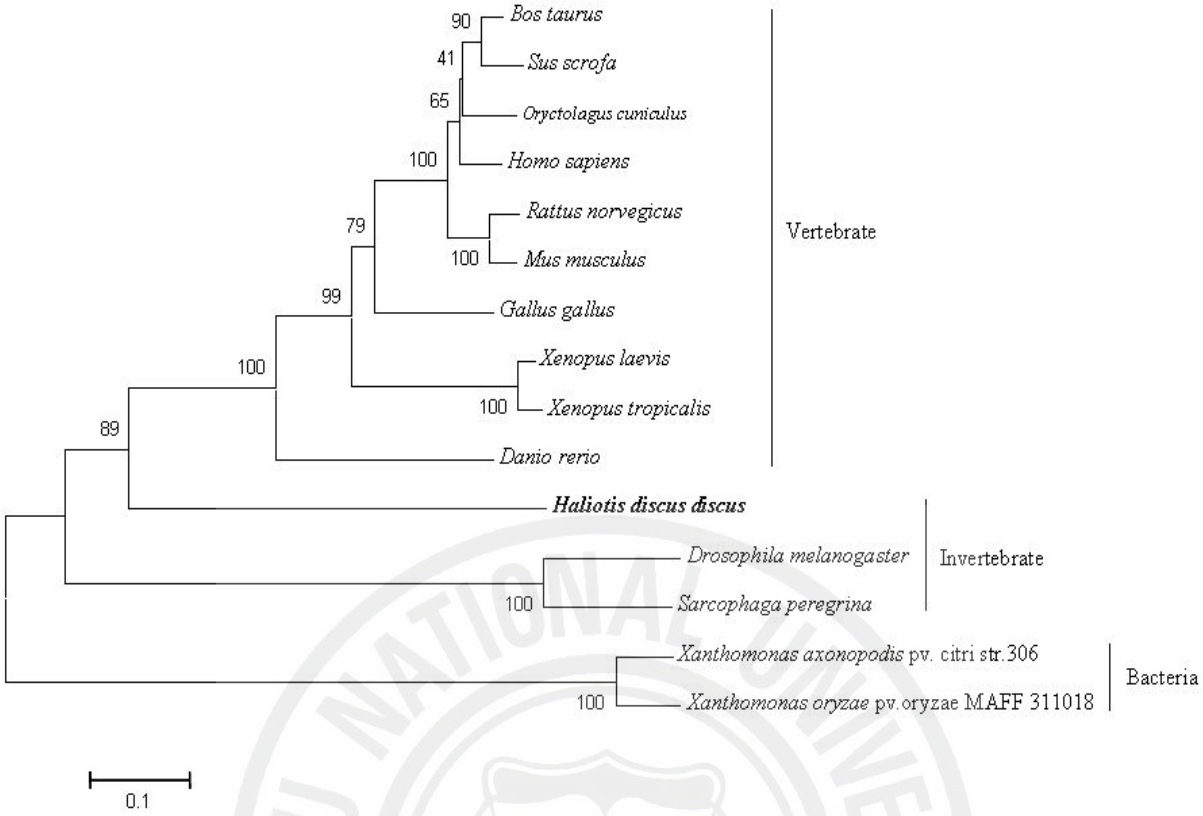
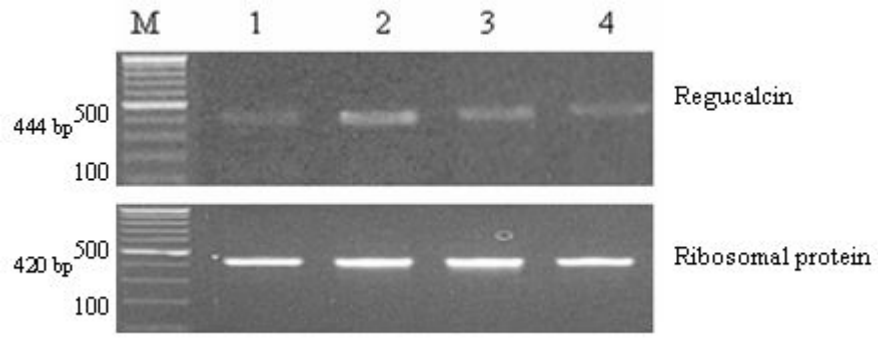



**Figure 34:** Phylogenetic tree of 15 regucalcins, including regucalcin from disk abalone (*H. discus discus*), GenBank accession nos. **EF103356**; chicken (*Gallus gallus*), **BAA90693**; bovine (*Bos taurus*), **BAA88080**; zebrafish (*Danio rerio*), **NP_991309**; rat (*Rattus norvegicus*), **BAA07490**; mouse (*Mus musculus*), **NP_033086**; human (*Homo sapience*), **NP_004674**; pig (*Sus scrofa*), **NP_001070688**; rabbit (*Oryctolagus cuniculus),* **BAA88079**; African clawed frog (*Xenopus laevis*), **BAA90694**; frog (*Sirulana tropicalis*), **NP_001006822**; *Xanthomonas axonopodis* pv.citri str. 306, **AAM36628**; *Xanthomonas oryzae* pv. oryzae MAFF 311018, **BAE69525**; anterior fat body protein of flesh fly (*Sarcophaga peregrina*), **BAA99282** and *Drosophila melanogaster* regucalcin homologue, **BAA99283**. A phylogenetic tree was constructed by the Neighbor-Joining method with MEGA 3.1. Numbers at the nodes are bootstrap values representing their robustness (1000 replications).

A




B

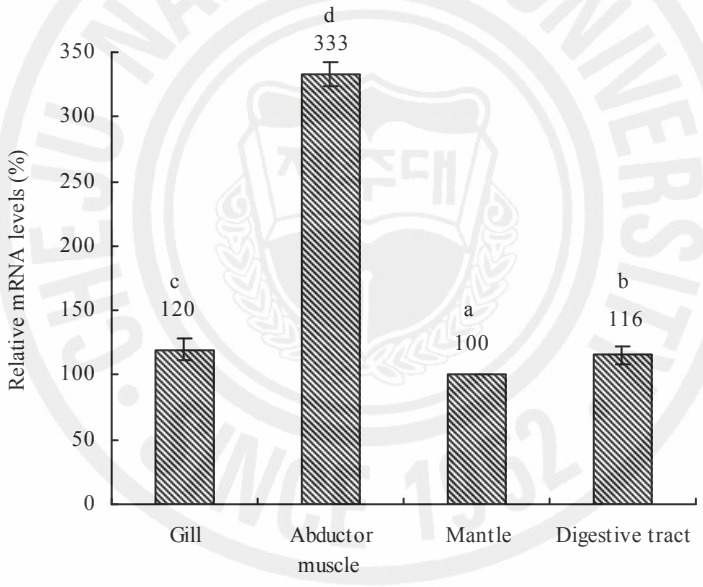
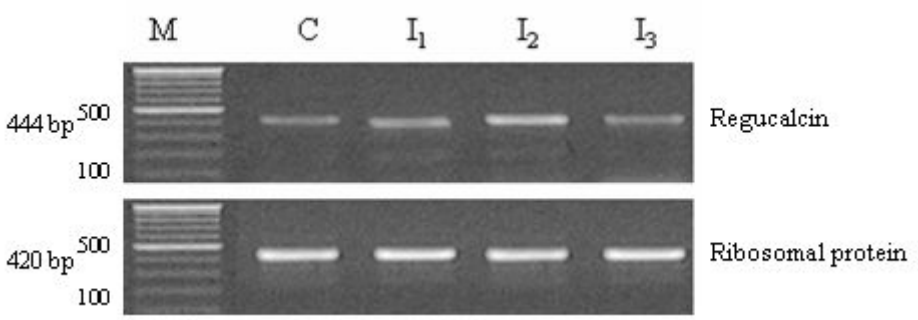



**Figure 35.** Tissue specific expression of HdReg in different tissues. A: RT-PCR analysis of HdReg in different tissues. Lane 1- Gill, 2- Abductor muscle, 3- Mantle, 4- Digestive tract and M- 100 bp marker. Ribosomal protein fragment was used as a housekeeping gene. B: Relative mRNA levels of Regucalcin in *H. discus discus* tissues. The levels of HdReg mRNA are the means of three assays (n=3), which are calculated relative to that of the expression recorded for the mantle (shown as 100%). Bars represent the means $\pm$ SD. Significance of the differences ($p<0.05$) are illustrated as a,b,c, and d.

**Response of intramuscular $CaCl_2$ injection**

The effect of calcium administration on HdReg mRNA expression was detected by RT-PCR analysis. A solution of $CaCl_2$ (0.5 mg/1g of abalone) was administered to abalones intramuscularly, and the tissues were collected at 30 min, 1 h, and 2 h after the administration. Even though constitutive expression was observed in HdReg in gill, mantle and digestive tract, clear induction was observed only in abductor muscle after $CaCl_2$ administration. Therefore, only the abductor muscle RT-PCR expression and respective semi quantitative analysis results are shown in Figure 36A and 36B, respectively. The results showed that abductor muscle regucalcin mRNA expression was induced significantly ($P<0.05$) at 30 min, 1 h, and 2 h after $CaCl_2$ administration when compared to uninjected abalones (Figure 36B). Induction was significantly higher (3.88-fold) at 1h of $CaCl_2$ administrated abalones than the uninjected abalones. Moreover, the relative mRNA expression level was significantly lower in 2h of $CaCl_2$ injected animals than the 30 min and 1h.

A




B

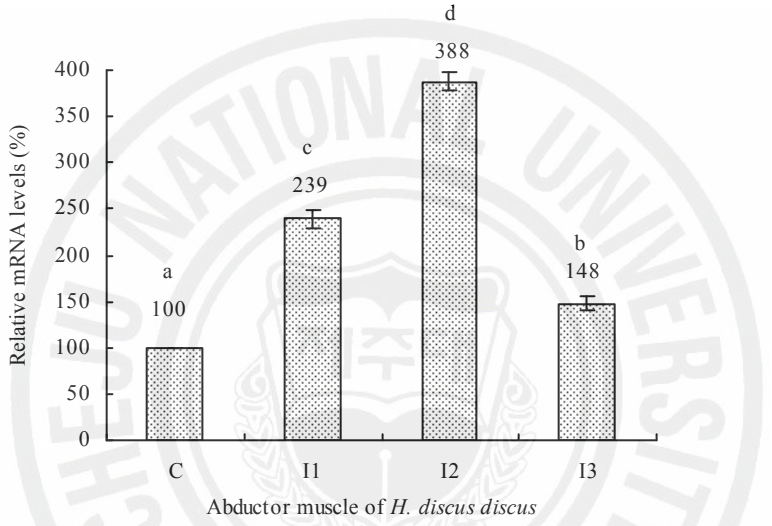



**Figure 36.** Expression of HdReg mRNA after $CaCl_2$ administration A: RT-PCR analysis of *H. discus discus* abductor muscle regucalcin gene expression at different time points after $CaCl_2$ administration. Lane C: control, $I_1$: 30 min, $I_2$: 1h, $I_3$: 2 h after administration. M: 100 bp marker. Abalone ribosomal protein fragment was used as a house keeping gene. B: Relative mRNA levels of HdReg after $CaCl_2$ administration. C-control, I1-30 min, I2-1 h and I3-2 h after administration. The levels of HdReg mRNA are means of three assays (n=3), which are calculated relative to that of the expression recorded for the control (shown as 100%). Bars represent the means $\pm$ SD. Significance of the differences ($p<0.05$) are shown as a, b, c, and d.

## DISCUSSION

In this study, we isolated and studied the primary structure of regucalcin gene from *H. discus discus* (Phylum mollusca, class Gastropoda). Full length cDNA of *H. discus discus* regucalcin was sequenced and characterized. Tissue specific mRNA expression was also investigated. This is the first report describing gene characterization of mollusk regucalcin.

Reucalcin is novel calcium binding protein that differ from calmodulin and other calcium binding proteins. It is mainly distributed in hepatic cytosol of rats and plays a physiological role in the cell, different from that of calmodulin. (Yamaguchi and Sakurai, 1992). Among the various regucalcins that have been cloned and described from vertebrates, the closest amino acid sequence identity for abalone was chicken (45%) and zebrafish (45%). It was 44% identical to rat and mouse and 42% identical to human regucalcin. The maximum identity of deduced amino acid sequence of anterior fat body protein (8-277 aa region) from flesh fly and rat SMP30 was 33% (Nakajima and Natori, 2000). Although the insects belong to invertebrates, a low identity was (33%) observed between HdReg and flesh fly anterior fat body protein. The HdReg contained 305 aa residues and molecular mass was calculated to be 33 kDa. It has been reported that estimated molecular mass of rat regucalcin was 33 kDa and was composed of 299 aa residues. Most of the coding sequences reported in vertebrate regucalcin were 299 aa such as chicken, rat, mouse, and human. Some of the regucalcin family proteins deposited in NCBI data base showed varying length of amino acids such as *B. cereus* ATCC 10987 (300 aa), *A. Fumigatus* (281 aa), and *D. melanogaster* (303 aa). Furthermore, the amino acid sequences of regucalcin did not show significant homology, when compared with the GenBank databases containing several other $Ca^{2+}$ binding proteins such as calmodulin, S-100beta, and calbindin-D28k. HdReg does not contain the typical structure of EF-hand motif (helix-loop-helix domain) as a $Ca^{2+}$ binding domain. The EF-hand loop consists of 12 aa and 5 of them have a carboxyl group (or hydroxyl group) in their side chain, precisely spaced so as to coordinate the $Ca^{2+}$ (Yamaguchi, 2000). Therefore, the

regucalcin was considered as a novel class of calcium-binding protein, which differs from calmodulin and other calcium binding proteins (Yamaguchi and Sakurai, 1992).

The amino acid composition of HdReg shows relatively high content of valine (10%), aspartic acid (9%), glycine (9%), and serine (8%). Aspartic acid (8%) and glutamic acid (5%) present in rat regucalcin could be related to $Ca^{2+}$ binding (Yamaguchi, 2000). Similar to rat regucalcin, 9% aspartic acid and 3% glutamic acid were observed in HdReg, and these residues could be important in HdReg. Moreover, in previous studies the hydrophilic region could be a functional domain, which is related to the binding of $Ca^{2+}$ (Shimokawa and Yamaguchi, 1993). HdReg as a molecule, showed hydrophilic character and at positions between 100 and 200 showed comparatively higher hydrophilic region (66%). Interestingly, a similar result was observed for rat liver regucalcin (Shimokawa and Yamaguchi, 1993).

Even though a potential N-glycosylation site at the position of 151-154 aa (NITI) was observed in HdReg, whether it is necessary for the regucalcin activity is unknown. Therefore, further investigations of its functional characteristics are required. In mouse and human regucalcin, N-glycosylation sites were found at positions of $^{164}$NQS$^{166}$ and $^{254}$NYS$^{256}$ in their sequences. However, no potential N-linked glycosylation site (Asn-X-Ser (or Thr) in rat regucalcin sequence (Shimokawa and Yamaguchi, 1993) was present. Moreover, a signal peptide was absent in all the known regucalcin sequences and HdReg. SOSUI prediction program of trans-membrane region analysis (Nagoya University, Japan (http://bp.nuap.nagoya-u.ac.jp/sosui/) showed that HdReg and other known regucalcin sequences were cytoplasmic proteins and not trans-membrane proteins.

Comparison analysis of regucalcin nucleotide sequences from seven vertebrate species (human, rat, mouse, rabbit, bovine, chicken, and toad livers) showed that they were highly conserved in their coding region. Conservation of regucalcin genes throughout the evolution is 69.9-91.3% identity (Misawa and Yamaguchi, 2000; Yamaguchi, 2000). In this study, the overall identity of HdReg with those selected vertebrate species was 34%. The

structural features suggest that the HdReg is partially similar to vertebrate regucalcin, but its functional characteristics may not completely similar to vertebrates as suggested by Nakajima and Natori (2000). Further, a constructed phylogenetic tree shows that the abalone is placed in a single cluster separated from a vertebrate main cluster. Moreover, it was separated from the invertebrate sub-cluster. This shows that there is a phylogenetic distance between vertebrate and invertebrates as well as among vertebrate and invertebrates even though they derived from a common ancestor.

Abalone gill, mantle, digestive tract, and abductor muscle were used for the expression analysis and HdReg was expressed in all the tissues tested. This suggests that HdReg is not confined to the specific tissue and may function in all tissues to some degree. However, expression level was higher in abductor muscle compared to other tissues tested and data from three independent experiments were similar. Shimokawa and Yamaguchi, 1992 have suggested that specific expression of rat regucalcin is mainly in liver and very limited in kidney. Further investigation showed that the regucalcin mRNA expression was observed in rat heart muscle and it was present in the cytoplasm of heart muscle cells (Yamaguchi and Nakajima 2002). $Ca^{2+}$ is an important element for many cellular processes in all the organisms including mollusks. The abalone mRNA was highly expressed in abductor muscle, showing that regucalcin may be involved in muscle contraction and relaxation for locomotion, attaching to substrates, and shell opening and closing. Although we have analyzed the tissue specific HdReg mRNA expression levels in different tissues, the detailed mechanism or functional role of regucalcin in $Ca^{2+}$ uptake, accumulation and release remains unelucidated. HdReg mRNA expression was significantly up regulated after 30 min of calcium administration and expression was higher at 1h than 30 min and 2h. However, other tissues tested did not show any significant induction. These results suggest that the expression of regucalcin mRNA rapidly responds in abductor muscle after $Ca^{2+}$ administration and regulatory system may suppress over-expressed regucalcin in abductor muscle or the injected

$Ca^{2+}$ may diffuse away to reach normal concentration. However, whether the dose of $Ca^{2+}$ administration is sufficient to increase the $Ca^{2+}$ content in other tested tissues is to be explored. Regucalcin mRNA expression (Shimokawa and Yamaguchi, 1992) and its protein content analysis (Yamaguchi and Isogai, 1993) were studied in liver and kidney cortex of rats and their *in vivo* regucalcin expression has been shown to be stimulated by the administration of $CaCl_2$ (Shimokawa and Yamaguchi, 1992; Shimokawa and Yamaguchi, 1993; Yamaguchi and Kurota, 1995). Moreover, Shimokawa and Yamaguchi, 1992 reported that several steps may be related to the regulation of regucalcin mRNA levels during the process of transcription, mRNA stability, translation, and posttranslational events.

In conclusion, we have cloned a novel regucalcin gene from *H. discus discus* as a first reported regucalcin gene in invertebrate mollusks. Then, we characterized the primary structure of regucalcin and observed common characteristics related to regucalcin. Regucalcin mRNA expression was constitutive, and specifically up regulated in abalone abductor muscle by $CaCl_2$ administration. However, the exact function of the HdReg and mechanism of its regulation needs to be investigated.

# PART 3

## Molecular Characterization and Expression Analysis of Pattern Recognition Protein from Disk Abalone *Haliotis discus discus*

## ABSTRACT

Pattern recognition molecules play an important role in innate immunity by recognizing common epitopes on the surface of invading microorganism. Pattern recognition protein (PRP) was isolated from a disk abalone, *Haliotis discus discus,* normalized cDNA library. It encodes 420 amino acids (aa) including 20 aa of a signal peptide sequence. The mature protein has an estimated molecular mass of 45 kDa and predicted pI of 5.0. The deduced aa sequence of HdPRP showed the highest identity (50%) to β-glucan recognition protein (BGRP) of the fresh water snail *Biomphalaria glabrata*. Characteristic potential polysaccharide binding, cell adhesion, and glucanase motifs were found in HdPRP, similar to invertebrate PRP motifs. Reverse transcription polymerase chain reaction (RT-PCR) results showed that HdPRP was constitutively expressed in gill, mantle, digestive tract, hepatopancreas and hemocytes, suggesting an innate immune role in these tissues. Animals injected with *Vibrio alginolyticus* bacteria showed that the mRNA expression was increased at 12 h post injection in gill and continued until 48 h. Abalone HdCuZnSOD gene expression was analyzed to show, whether the oxidative stress has been occurred after exposed to different pathogen associated molecular patterns (PAMPs). It showed that the mRNA expression was significantly increased in *Vibrio*, lipopolysaccharides (LPS) and β-1,3-glucan injected animals compared to control animals. Abalone PRP can recognize different PAMPs and may activate different immune genes to defense against these pathogens. It acts as an acute inducible protein that could play an important role in abalone immune defense mechanism.

## INTRODUCTION

Invertebrate animals lack antibodies and adaptive immune responses. However, they have efficient innate immune systems to defend themselves against invading foreign materials (Vargas-Albores and Yepiz-Plascencia, 2000). Innate immune reactions are the first line of defense for vertebrates and invertebrates and are initiating by several kind of non-self pathogen-associated molecular patterns (PAMPs) particularly, LPS and flagellin from gram-negative bacteria, lipoteichoic acid (LTA) and peptidoglycan from gram positive bacteria and β-1,3-glucans, and mannans from fungal cell wall (Fabrick, 2003). PAMPs are molecules present on the surface of microorganisms, generally as structural components which are absent in host animals.

Function of the pattern recognition protein (PRP) is to initiate the host immune response, when binding to non-self pathogen-associated molecular patterns (PAMPs). Invertebrate immune defense system consists of both cellular and humoral immune responses. The cellular immune responses contain encapsulation, phagocytosis, and nodule formation. Humoral immune responses contain clotting system of arthropods, the synthesis of a broad spectrum of potent antimicrobial proteins, and the prophenoloxidase activating system (proPO system) (Cerenius et al., 1994; Lee et al., 2000; Kim et al., 2000; Yu and Kanost 2003). It has been reported that pattern recognition molecules act as biosensors in the activation of innate immune responses in both vertebrates and invertebrates (Fabrick, 2003). Recently, several PRPs have been isolated and characterized in variety of invertebrates species, such as LPS-and/or β−1,3-glucan binding proteins, peptidoglycan recognition proteins, gram-negative bacteria binding proteins (GNBP), C-type lectins, galactoside binding lectins (galectins), thioester containing proteins, fibrinogen like domain immunolectins, scavenger receptors and hemolin (Christophides et al., 2004; Dziarski, 2004). It has been reported that common motifs such as bacterial glucanase like (Cerenius et al., 1994; Kim et al., 2000; Ochiai and Ashida 2000), bacteriophage lysozyme like (Yoshida et al.,

1996) and immunoglobulin like (Fearon and Locksley, 1996) motifs exists in some of their primary structure. Moreover, their function in the immune response has also been studied.

Lipopolysaccharide and β-1,3-glucan binding protein (LGBP) play an important role in the innate immune response of invertebrates as a PRP. These molecules lack the binding specificity of antibodies even though they recognize and bind PAMPs. LPS is the principal component of the cell wall of gram negative bacteria and studies on *Drosophila* demonstrated that the GNBP (also known as LGBP), functions as a recognition receptor for LPS and β-1,3-glucan (Kim et al., 2000). The prophenoloxidase (proPO) activating system is an important non-self recognition system in invertebrates which can be activated by LPS or peptidoglycan from bacteria and β-1,3-glucans from fungi (Ashida et al., 1983; Smith et al., 1984; Hoffmann et al., 1996). Non-self-molecules are recognized by endogenous PRPs and their receptors, and they cause activation of the proPO system (Sritunyalucksana and Soderhall, 2000). Active form of proPO, phenoloxidase (PO), plays an important role, as it can melanize pathogens, sclerotize the cuticle, and heal wounds in invertebrates (Wang et al., 2007).

The opsonic effect of the LPS binding protein in the cockroach *Periplaneta americana* (Jomori and Natori, 1992), opsonic effect and degranulation of blood cells by the β-1,3-glucan binding protein (BGBP) in the crayfish *Pacifastacus leniusculus*, (Lee et al., 2000) and the hemocyte nodule formation by the LPS binding protein in the silkworm *Bombyx mori* (Koizumi et al., 1999) have been reported as special biological properties of PRPs. Further, BGRPs were able to bind to fungal cell wall component glucans and have been characterized in many species. For example, a BGBP gene cloned from silkworm *B. mori* had strong specific affinity for β-1,3-glucan (Ochiai and Ashida, 2000). Moreover, a 52 kDa BGBP isolated form *Manduca sexta* was shown to aggregate bacteria and fungi and to stimulate proPO activation (Jiang et al., 2004).

Several PRPs have been cloned and characterized in crustaceans such as crayfish *P.*

*leniusculus* (Lee et al., 2000), tiger shrimp *P. monodon* (Sritunyalucksana and Soderhall, 2000) and blue shrimp *L. stylirostris* (Roux et al., 2002). Cloning of LGBP cDNA from freshwater crayfish *P. leniusculus* suggested that it plays a role in proPO activation (Lee et al., 2000). Recently a LGBP cDNA was cloned from the hemocyte and hepatopancreas of white shrimp *Litopenaeus vannamei* (Cheng et al., 2005). BGBP that can specifically bind to β-glucan has been isolated from black tiger shrimp *P. monodon* (Sritunyalucksana and Soderhall, 2000). However, there is little information reported about mollusk PRPs and according to our knowledge no information about gene characterization or gene expression has been studied about abalone PRPs. Abalones are large algivorous marine mollusks of the genus, *Haliotis* (Phylum mollusca, Class Gastropoda). They are one of the most commercially important gastropods in aquaculture.

In present study, a PRP cDNA from disk abalone *Haliotis discus dscus* was cloned, sequenced and characterized. RT-PCR was carried out to evaluate tissue specific mRNA expression in different abalone tissues. To evaluate its mRNA expression, abalones were injected using different immuno modulators such as *V. alginolyticus*, LPS from *Escherichia coli* and β-1,3-glucan from *Laminaria digitata* which are the major polysaccharides in cell walls of gram negative bacteria and fungi, respectively. Further, oxidative stress after different immuno modulators injection was evaluated by abalone CuZnSOD mRNA expression, using above injected abalone tissues.

## MATERIAL AND METHODS

### Cloning and sequencing of abalone PRP cDNA

Disk abalone cDNA library was constructed from the digestive tract tissues of the abalone using a cDNA library construction kit (Creator$^{TM}$ SMART$^{TM}$, Clontech, USA) and cDNA was normalized with Trimmer-Direct cDNA normalization kit according to the manufacture's protocol (Evrogen, Russia). By the results of expressed sequence tag (EST) blast, putative PRP function having cDNA clone was identified. The plasmid DNA of the putative PRP was obtained by the Accuprep$^{TM}$ plasmid extraction kit (Bioneer Co., Korea). The full length sequence was determined by sequencing reaction using termination kit, Big Dye and an ABI 3700 sequencer (Macrogen Co., Korea). After determine the full length, the sequence was compared with other sequences available in the National Center for Biotechnology Information (NCBI) database and sequence similarities were analyzed with BLAST algorithm at the NCBI BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). Signal peptide of HdPRP deduced amino acid sequence was predicted through a SignalP program (http://www.cbs.dtu.dk/) and motif prediction was done using ExPASy (http://www.au.expasy.org). Pair wise and multiple alignment of the HdPRP were performed using ClustalW multiple alignment 1.8 programe (Thompson et al., 2004). MEGA 3.1 (Kumar et al., 2004) was used to produce the phylogenetic tree using the neighbor joining (NJ) method.

### Animals

The disk abalones were purchased from Youngsoo abalone farm (Jeju, Republic of Korea). Individuals body mass were 45-50 g and they were maintained in 40 L tanks with an aerated sea water having temperature of 19±2 $^{o}$C for a week to acclimatize to laboratory conditions before being injected by immuno modulators which contains different PAMPs such as *V. alginolyticus*, LPS and β-1,3-glucan.

**Bacterial challenge, LPS, β-1,3-glucan injection and tissue collection**

For bacterial challenge, abalones were injected with *V. alginolyticus* suspended in PBS (150 μl/animal, $OD_{600}$=1.0) intramuscularly. For LPS induced mRNA expression analysis, abalones were injected with LPS (4 μg/g animal, 055:B5 from *E. coli*, Sigma-Aldrich, USA). For β-1,3-glucan induction, the abalones were injected with 6 μg/g animal of laminarin (beta-1,3-glucan) from *L. digitata* (Sigma-Aldrich, USA). All treated animals were kept 12, 24 and 48 h post injection (p.i.) and gill was taken from bacterial challenged, LPS injected and β-1,3-glucan injected abalones to find out the transcriptional induction of HdPRP. Untreated and PBS injected animals were kept separately as control groups. Gill, digestive tract, mantle, hepatopancreas and hemolymph tissue samples were collected from untreated control group and gill was taken at 12 h p.i. of PBS. Using sterilized syringe, the hemolymph (2-3 ml/abalone) was collected from the abalone pericardial cavity. Then, it was immediately centrifuged at 3000 x g for 10 min at 4 $^{o}$C and hemocytes were collected for RNA extraction. Tissues were immediately snap-frozen in liquid nitrogen and stored at -70 $^{o}$C until RNA isolation. Three independent experiments were conducted and three independent tissue preparations were done for each experiment (n=3).

**RNA isolation, cDNA construction and semi quantitative RT-PCR analysis**

Total RNA of the gill, mantle, digestive tract, hepatopancreas and hemocyte was isolated using Tri Reagent$^{TM}$ (Sigma, USA) according to the manufacture's protocol. Using 2 μg of RNA, the first strand cDNA synthesis was carried out by SuperScript III first strand cDNA synthesis kit (Invitrogen, USA). The resulting cDNA was stored at -20 $^{o}$C for further experiments.

Semi-quantitative RT-PCR was carried out to study the tissue specific distribution of HdPRP mRNA expression in disk abalone and the transcriptional induction. Electrophoretic images and the optical densities of amplified bands were analyzed using the Scion Image

analysis software (Release alpha 4.0.3.2), Scion Corporation, Maryland, USA. The relative mRNA expression levels were calculated as a ratio of the target gene mRNA and ribosomal protein mRNA expression. A pair of HdPRP gene specific primers HdPRP–1F and HdPRP–1R was used to amplify a 408 bp fragment of HdPRP. A pair of gene-specific primers of HdSOD-2F and HdSOD-2R was designed to amplify a 254 bp fragment of abalone CuZnSOD (Table 1). As a house keeping gene, disk abalone ribosomal protein expression (GenBank accession no. **EF103427**) was expressed by amplification of 420 bp cDNA fragment using ribosomal 3F and ribosomal 3R primers (Table 7).

RT-PCR was carried out in a TaKaRa PCR thermal cycler in 25 µl reaction volume containing 2 µl of cDNA from each tissue, 2.5 µl of 10x *TaKaRa Ex Taq*$^{TM}$ buffer, 2.0 µl of 2.5 mM dNTP mix, 1.0 µl of each primer (20 pmol/µl ), and 0.13 µl (5 U/µl) of *TaKaRa Ex Taq*$^{TM}$ DNA polymerase (TaKaRa, Japan). The PCR temperature profile was 94 $^{o}$C for 2 min followed by 30 cycles of 94 $^{o}$C for 30 sec, 57 $^{o}$C for 30 sec and 72 $^{o}$C for 30 sec. Same PCR temperature profile was used for HdCuZnSOD amplification with 24 cycles. For ribosomal protein, it was 94 $^{o}$C for 2 min followed by 23 cycles of 94 $^{o}$C for 30 sec, 55 $^{o}$C for 30 sec and 72 $^{o}$C for 30 sec. After the final cycle, samples were incubated for a further 5 min at 72 $^{o}$C then held at 4 $^{o}$C prior to analysis. The PCR products were analyzed by electrophoresis on 1.5% agarose gel containing ethidium bromide and 100 bp molecular marker (TaKaRa, Japan).

**Statistical analysis**

For comparison of relative HdPRP and HdCuZnSOD mRNA levels, statistical analysis was performed with one way ANOVA and mean comparisons were performed by Duncun's multiple range test using SPSS 11.5 program. Significant P values ($P>0.05$) were obtained by Duncan multiple range test and results were shown as mean $\pm$ SE of three animals per group.

**Table 7:** Primers used for HdPRP RT-PCR expression analysis

| Name | Object | Sequence (5' to 3' direction) |
|---|---|---|
| HdPRP – 1F | RT-PCR amplification | GCAACACCACCAACAACGACATCA |
| HdPRP – 1R | | AGCCGATATTTCCACCTTGCCGTA |
| HdSOD-2F | RT-PCR amplification | GGCAAACATGGCTTCCACGTTCAT |
| HdSOD-2R | | TCATCCACTCCAGCATGGACAACA |
| Ribosomal-3F | RT-PCR positive control | GGGAAGTGTGGCGTGTCAAATACA |
| Ribosomal-3R | | TCCCTTCTTGGCGTTCTTCCTCTT |

## RESULTS

### Sequence characterization of HdPRP full length cDNA

The HdPRP consisted 1454 bp full length with an ORF of 1263 bp. It showed a polyadenylation signal with sequence of AATAAA and a poly(A) tail (Figure 37). The ORF is capable of encoding a polypeptide of 420 aa with an estimated molecular mass of 47 kDa and predicted isoelectric point (*pI*) of 5.0. The deduced aa sequence of HdPRP was compared with several GNBP, BGRP, BGBP, LGBP, β-1,3-glucanase (BGNS) and coelomic cytolytic factor (CCF) and CCF like proteins of other species that are available in GenBank. Their GenBank accession numbers and aa identity percentages (%) were shown in the Table 8. The pairwise CLUSTALW analysis of HdPRP showed 50% identity with the BGRP of the *B. glabrata* (GenBank accession no. **ABL63381**). The HdPRP cDNA sequence was submitted to the NCBI GenBank under accession number **EF103355**.

The HdPRP sequence contains a 20 aa putative signal peptide, a potential recognition motif for β-1,3-linkage of polysaccharides, putative cell adhesion site, a glucanase motif, and a protein kinase C phosphorylation site (Figure 37). Moreover, LPS binding (LPB) motif was found in HdPRP sequence at the position from 244 to 261 and it was 50% similar to bivalvia, *Chlamys farreri* LPB motif. The motif scan analysis showed that amino acid region from 115 to 376 belongs to the glycoside hydrolase family 16. Further, HdPRP sequence contained β-1,3-glucanase site with active residues of W (Trip), E (Glu), I (Ile) and D (Asp) at positions of 229, 234, 235 and 236, respectively. Moreover, there was no characteristic putative cell adhesive site, RGD (Arg-Gly-Asp), however, the modified MGD was found in the sequence at position of 209 when compared with shrimp PRP sequences. Moreover, recognition motif for beta-1,3-linkages of polysaccharide was observed at the position of 292 to 310. Three potential N-glycosylation sites were identified at positions of 48-51, 119-122 and 330-333 by using ExPASy Proteomics tools.

Further, a multiple sequence alignment was created by CLUSTALW program using 8 invertebrate species including *H. discus discus* (Figure 38). The HdPRP sequence was slightly different from analyzed other PRPs, in that upstream part of the abalone PRP was little longer than the other PRPs. Even though, the conserved motifs were observed among the sequences, the length of the amino acids was varied from species to species.

**Phylogenetic analysis of HdPRP**

A phylogenetic tree, based on deduced amino acid sequences of different PRPs including BGBP, LGBP, GNBP, BGRP and CCF like proteins of 21 known invertebrate species was created using neighbor joining (NJ) method (Figure 39) to determine the position of HdPRP in evolution. Phylogenetic analysis showed that insect GNBP and BGNS were grouped into one cluster and crustacean BGBP and LGBP were grouped into another cluster. However, HdPRP was grouped with PRPs identified from other mollusk species, *B. glabrata* GRP and *C. farreri* LGBP. HdPRP and *C. farreri* LGBP was sub-grouped with bootstrap value of 56, despite of their low identity (40% aa identity). Interestingly, GNBP of mollusk *B. glabrata* (GenBank accession no. **ABO40828**) did not position with same mollusk cluster and was grouped with annelids CCF-like proteins.

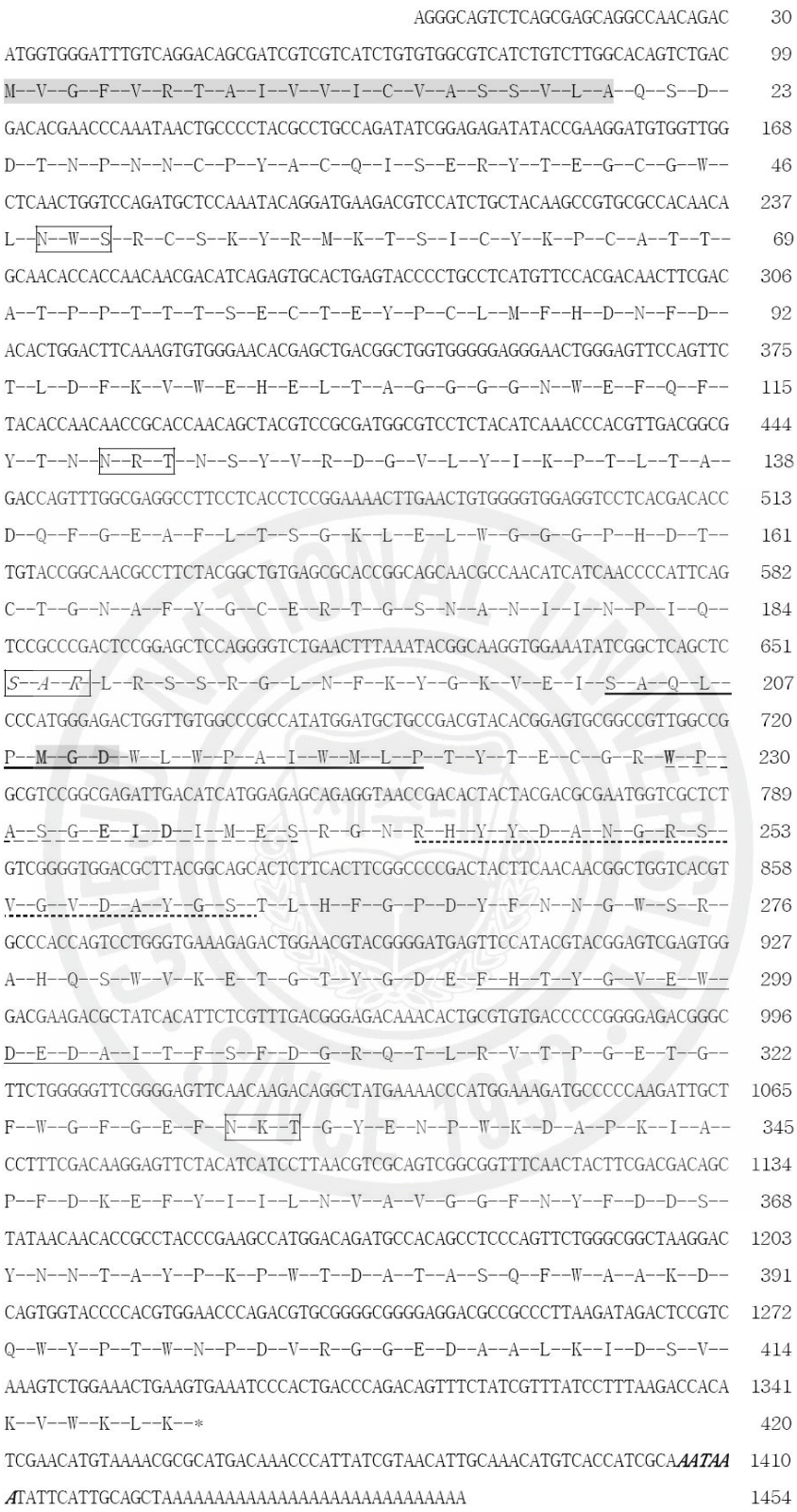



**Figure 37**. Nucleotide and deduced amino acid sequences of HdPRP. The nucleotide sequence is numbered from 5'end, and the single letter aa code is presented below the

corresponding codon. The putative signal sequence (aa position 1-20) is shaded. Potential N-glycosylation sites are boxed. The underlined aa sequence represents a potential recognition motif for beta 1,3-linkage of polysaccharides. Putative cell adhesive site is shaded and bold. Polysaccharide binding (PsB) motif and glucanase motif (GM) is bold under lined and dash under lined respectively. LPS binding motif is dotted under line. Conserved potential kinase C phosphorylation site is Italic and boxed. Amino acid residues that compose functional domain of glucanase are marked in bold. A putative polyadenylation sequence, AATAAA, is bold underlined, and the termination code is marked with an asterisk.

**Table 8:** Pairwise CLUSTALW analysis and comparison of the deduced amino acid sequence of HdPRP protein with invertebrate other PRPs.

| Species | Name of the PRP | NCBI accession no. | Identity (%) | Amino acids |
|---|---|---|---|---|
| *Biomphalaria glabrata* | BGRP | ABL63381 | 50 | 393 |
| *Aedes aegypti* | GNBP | XP_001659797 | 45 | 394 |
| *Diatraea saccharalis* | BGNS | ABR28479 | 45 | 375 |
| *Nasutitermes dixoni* | GNBP | AAZ08481 | 43 | 379 |
| *Nasutitermes magnus* | GNBP | AAZ08486 | 43 | 379 |
| *Nasutitermes fumigatus* | GNBP | AAZ08483 | 43 | 379 |
| *Litopenaeus venname* | BGBP | AAW51361 | 42 | 367 |
| *Fenneropenaeus chinensis* | LGBP | AAX63902 | 42 | 366 |
| *Penaeus monodon* | BGBP | AAM21213 | 42 | 366 |
| *Litopenaeus stylirostris* | LGBP | AAM73871 | 42 | 376 |
| *Marsupenaeus japonicus* | BGBP | BAD36807 | 41 | 366 |
| *Chlamys farreri* | LGBP | AAP82240 | 40 | 440 |
| *Pacifastacus leniusculus* | LGBP | AJ250128 | 38 | 361 |
| *Eisenia foetida* | coelomic factor 1 | AF030028 | 35 | 384 |
| *Biomphalaria glabrata* | GNBP | ABO40828 | 34 | 435 |

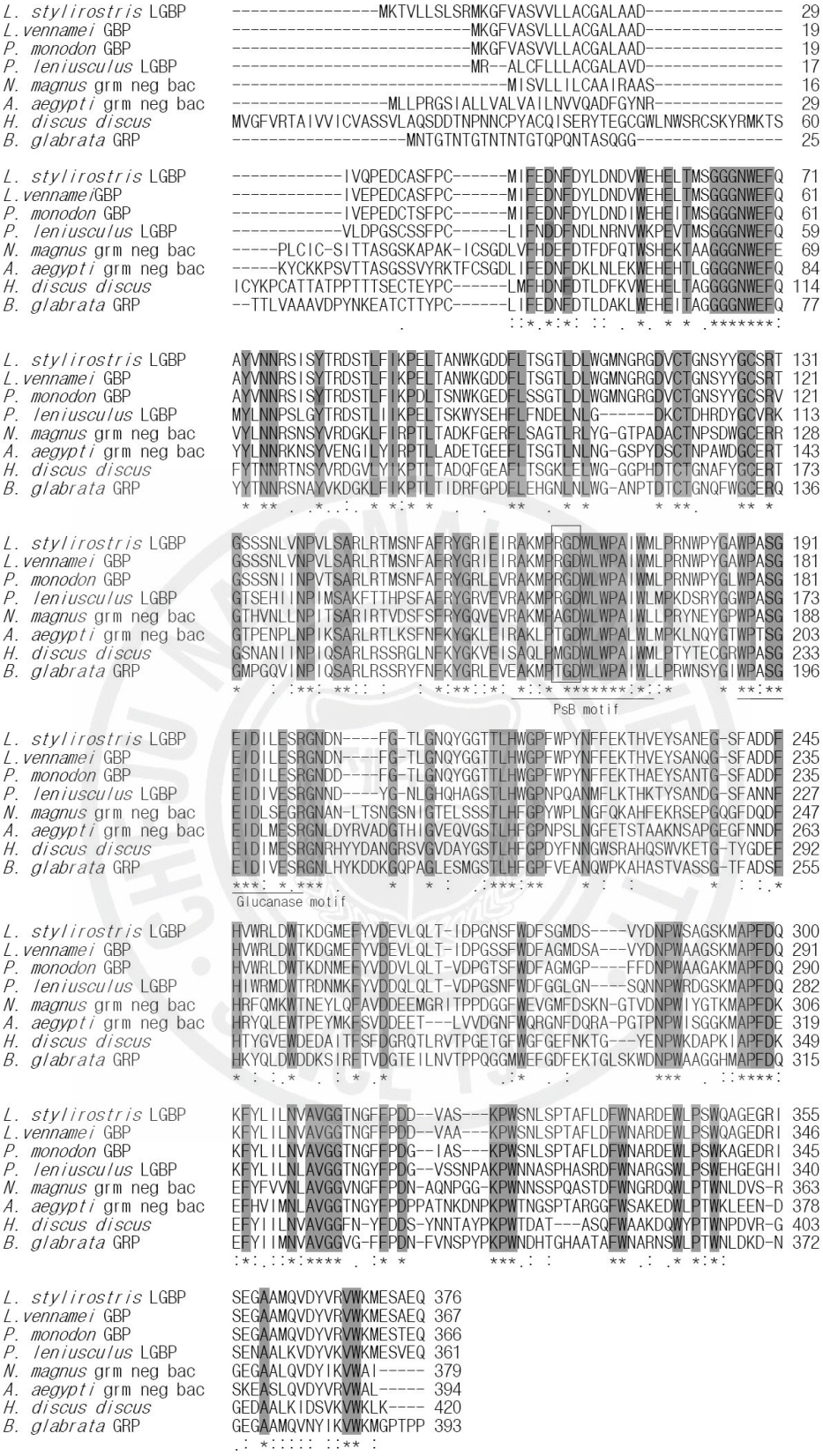



**Figure 38.** Multiple sequence alignments of HdPRP with 7 other homologous pattern recognition amino acid sequences. BGRP of the *Biomphalaria glabrata* (**ABL63381**), GNBP

of *Aedes aegypti* (**XP_001659797**), GNBP of *Nasutitermes magnus* (**AAZ08486**), BGBP of *Litopenaeus vennname* (**AAW51361**), BGBP of *Penaeus monodon* (**AAM21213**), LGBP of *Litopenaeus stylirostris* (**AAM73871**) and LGBP of *Pacifastacus leniusculus* (**AJ250128**). Asterisk marks with gray colour shading indicates identical amino acids and numbers to the right indicate the aa position of PRP that in the corresponding species. Potential glucanase and polysaccharide binding (PsB) motifs are under lined. Gaps were introduced for optimal alignment.

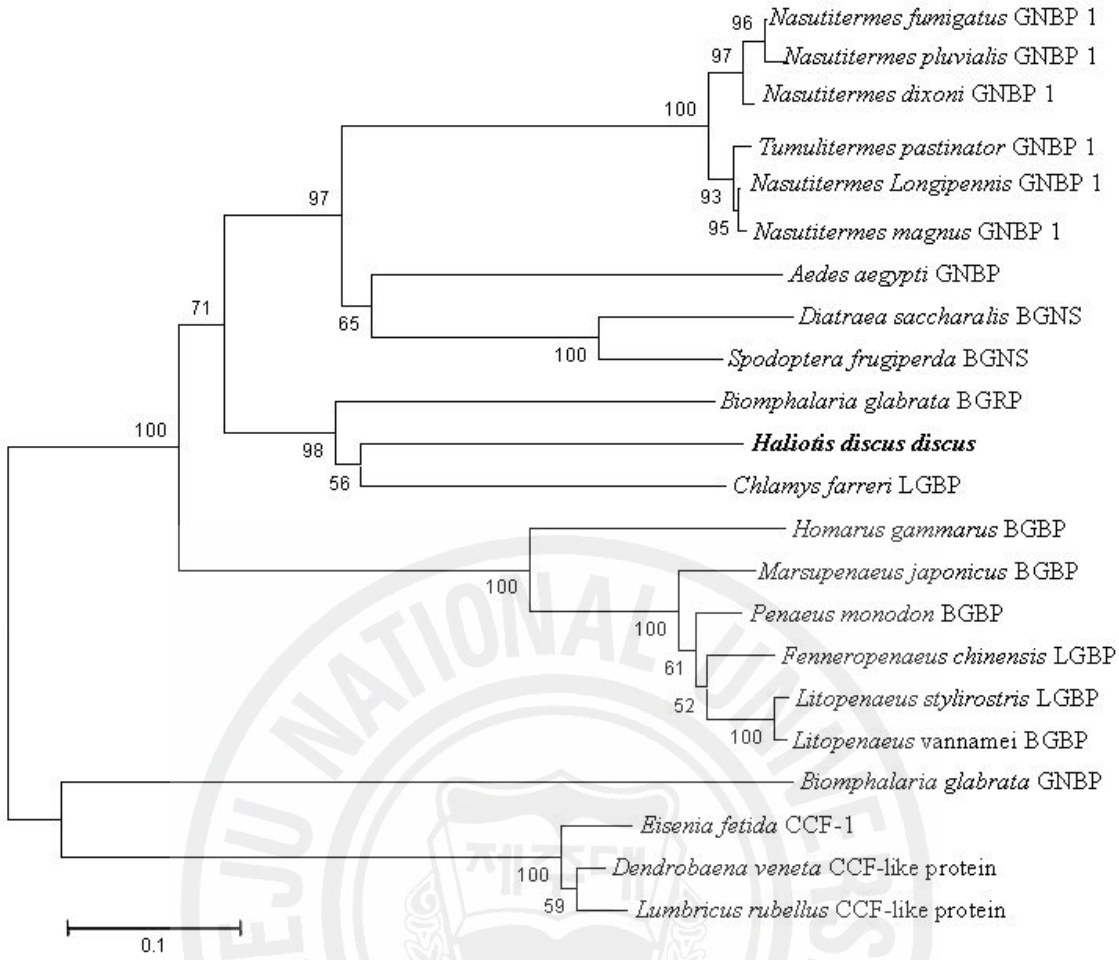



**Figure 39.** A Phylogenetic tree of HdPRP with other 21 invertebrate species PRP reconstructed by the Neighbor-Joining method (MEGA 3.1, [18]). The tree is based on an alignment corresponding to the full length aa sequences, using ClustalW and MEGA (3.1). The numbers are shown at the branches denotes the bootstrap majority consensus values on 1000 replicates. The GenBank accession numbers for the sequence designations as follows. *Biomphalaria glabrata* BGRP (**ABL63381**), *Diatraea saccharalis* beta 1,3-glucanase (BGNS) (**ABR28479**), *Spodoptera frugiperda* BGNS (**ABR28478**), *Aedes aegypti* GNBP (**XP 001659797**), *Nasutitermes dixoni* GNBP 1 (**AAZ08481**), *N. magnus* GNBP 1 (**AAZ08486**), *Litopenaeus venname* BGBP (**AAW51361**), *Penaeus monodon* BGBP (**AAM21213**), *L. stylirostris* LGBP (**AAM73871**), *Chlamys farreri* LGBP (**AAP82240**), *Fenneropenaeus chinensis* LGBP (**AAX63902**), *Eisenia foetida* coelomic cytolytic factor 1 (CCF1) (**AF030028**), *N. fumigatus* GNBP 1 (**AAZ08483**), *N. pluvialis* GNBP 1 (**AAZ08487**), *Tumulitermes pastinator* GNBP 1 (**AAZ08490**), *N. Longipennis* GNBP 1 (**AAZ08485**), *Homarus gammarus* BGBP (**CAE47485**), *Marsupenaeus japonicus* BGBP (**BAD36807**), *B. glabrata* GNBP (**ABO40828**), *Dendrobaena veneta* CCF-like protein (**AAY85745**), *Lumbricus rubellus CCF-like protein* (**AAY85746**).

**Tissue expression analysis of HdPRP mRNA**

Semi-quantitative RT-PCR analysis was carried out to assess the tissue specific mRNA level of HdPRP in untreated (control) disk abalone. HdPRP mRNA was expressed in gill, mantle and digestive tract at 30 cycles of PCR (Figure 40A) where as in hepatopancreas and hemocyte, the HdPRP mRNA expression was obtained at 35 cycles of PCR (data not shown). Semi quantitative expression analysis showed that the relative expression level was significantly higher ($p<0.05$) in gill compare to digestive tract and mantle (Figure 40B).

**HdPRP mRNA expression analysis after *V. alginolyticus*, LPS and β-1,3-glucan induction**

HdPRP mRNA expression at different time points after *V. alginolyticus*, LPS and β -1,3-glucan are shown in Figure 41A, 42A and 43A, respectively. When *V. alginolyticus* is injected into abalone, HdPRP mRNA expression was significantly ($p<0.05$) increased at 12 h p.i. compared to control group. Increasing trend was continued up to 48 h p.i. The highest mRNA expression was observed at 24 h p.i. and was 4.32 fold higher than the control (Figure 41B). We tried out LPS induction of abalone to find out whether HdPRP can recognize gram negative bacterial cell wall compounds. It also showed significantly higher ($p<0.05$) mRNA expression in HdPRP at 12 h p.i. of LPS than the control (Figure 42B). The highest expression was observed at 48 h p.i. and it was 4.84 fold higher than the control. The relative expression pattern of HdPRP in gill after glucan injection is shown in Figure 43B. The HdPRP transcript was increased at 12 h p.i and reached maximum at 24 h p.i. and then subsequently decreased at 48 h p.i. Significant difference ($p<0.05$) in HdPRP expression was observed between the control and injected groups of 12, 24 and 48 h p.i.

**HdCuZnSOD mRNA expression analysis after *V. alginolyticus*, LPS and β-1,3-glucan induction**

HdCuZnSOD expression was analyzed in *vivo* for bacteria, LPS and β-1,3-glucan injected abalones at 12, 24, and 48 h p.i. to find out whether the oxidative stress has been occurred in abalone after exposed to different PAMPs. HdCuZnSOD mRNA expression at different time points after *V. alginolyticus*, LPS and glucan are shown in Figure 41C, 42C and 43C, respectively. In all the treatments HdCuZnSOD were significantly increased after 12 h p.i. compared to control. In *Vibrio* injected abalone gill, HdCuZnSOD mRNA expression was significantly ($p<0.05$) increased at 12 h post injection compared to control and thereafter, it was significantly decreased (Figure 41D). Interestingly, it was significantly decreased ($p<0.05$) into 0.32 fold at 48 h p.i. compare to control. Also significantly ($p<0.05$) higher relative HdCuZnSOD mRNA expression was observed in LPS injected abalones after 12 h p.i. than the control and thereafter, a decreasing trend was observed (Figure 42D). However, relative HdCuZnSOD mRNA expression was significantly increased with the time in glucan injected abalone gill and relative HdCuZnSOD transcription was 2.27 fold higher at 48 h p.i. than the control (Figure 43D).

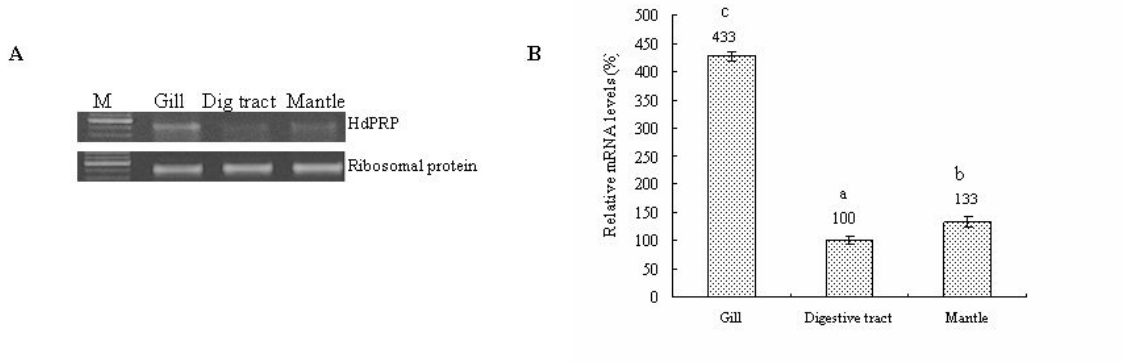



**Figure 40**: Tissue specific expression analysis of HdPRP in disk abalone by RT-PCR. A: HdPRP mRNA levels in Lane 1-gill; 2-digestive tract and 3-mantle. M-100 bp molecular marker. Corresponding ribosomal protein mRNA expression was used as an internal PCR control. B: Relative mRNA levels of HdPRP tissues. The relative mRNA expression levels of abalones were calculated using intensity value ratio of the target gene mRNA expression and ribosomal protein mRNA expression and then calculated relative to that of the expression recorded for the digestive tract (shown as 100%). The levels of HdPRP mRNA are means of three assays. Means with the different letters are significantly different at $p<0.05$ level. Bars represent the means $\pm$ SD.

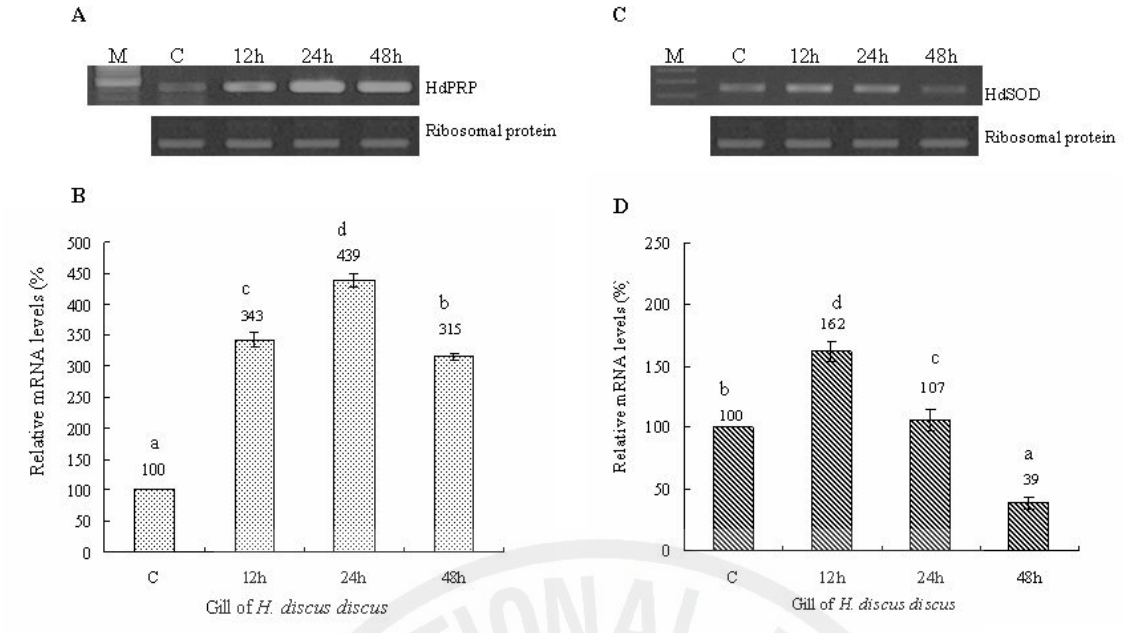



**Figure 41:** Expression of HdPRP and HdSOD mRNA induced by *V. alginolyticus*. (A) RT-PCR analysis of *H. discus discus* Gill PRP gene expression at different time points after *V. alginolyticus* injection. (B) Relative mRNA levels of HdPRP after *V. alginolyticus* injection. (C) RT PCR analysis of *H. discus discus* Gill CuZnSOD gene expression at different time points after *V. alginolyticus* injection. (D) Relative mRNA levels of HdSOD after *V. alginolyticus* injection. Lane C: control, I1: 12 h, I2: 24 h, I3: 48 h after administration. M: 100 bp marker. Corresponding ribosomal protein mRNA expression was used as an internal PCR control. The levels of HdPRP and HdSOD mRNA are means of three assays. The relative mRNA expression levels of abalones were calculated using intensity value ratio of the target gene mRNA expression and ribosomal protein mRNA expression and then calculated relative to that of the expression recorded for the corresponding control (shown as 100%). Means with the different letters are significantly different at $p<0.05$ level. Bars represent the means $\pm$ SD.

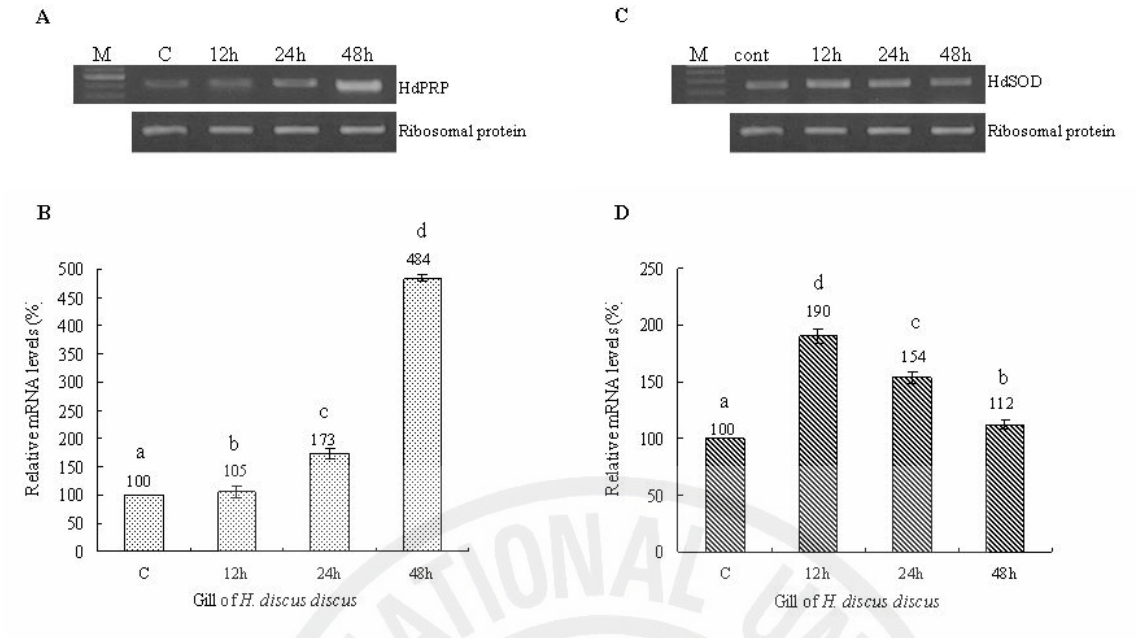



**Figure 42:** Expression of HdPRP and HdSOD mRNA induced by LPS. (A) RT- PCR analysis of *H. discus discus* Gill PRP gene expression at different time points after LPS injection. (B) Relative mRNA levels of HdPRP after LPS injection. (C) RT PCR analysis of *H. discus discus* Gill CuZnSOD gene expression at different time points after LPS injection. (D) Relative mRNA levels of HdSOD after LPS injection. Lane C: control, I1: 12 h, I2: 24 h, I3: 48 h after administration. M: 100 bp marker. Corresponding ribosomal protein mRNA expression was used as an internal PCR control. The levels of HdPRP and HdSOD mRNA are means of three assays. The relative mRNA expression levels of abalones were calculated using intensity value ratio of the target gene mRNA expression and ribosomal protein mRNA expression and then calculated relative to that of the expression recorded for the corresponding control (shown as 100%). Means with the different letters are significantly different at $p<0.05$ level. Bars represent the means $\pm$ SD.

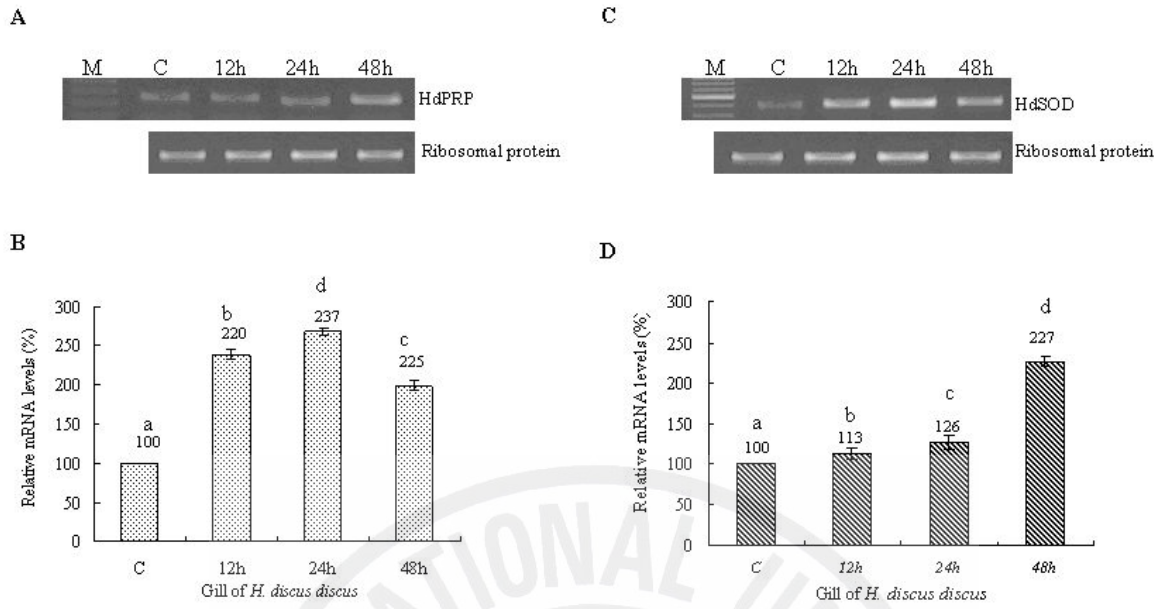



**Figure 43:** Expression of HdPRP and HdSOD mRNA induced by β-1,3-glucan. (A) RT-PCR analysis of *H. discus discus* Gill PRP gene expression at different time points after β-1,3-glucan injection. (B) Relative mRNA levels of HdPRP after β-1,3-glucan injection. (C) RT PCR analysis of *H. discus discus* Gill CuZnSOD gene expression at different time points after beta-1,3-glucan injection. (D) Relative mRNA levels of HdSOD after β-1,3-glucan injection. Lane C: control, I1: 12 h, I2: 24 h, I3: 48 h after administration. M: 100 bp marker. Corresponding ribosomal protein mRNA expression was used as an internal PCR control. The levels of HdPRP and HdSOD mRNA are means of three assays. The relative mRNA expression levels of abalones were calculated using intensity value ratio of the target gene mRNA expression and ribosomal protein mRNA expression and then calculated relative to that of the expression recorded for the corresponding control (shown as 100%). Means with the different letters are significantly different at $p<0.05$ level. Bars represent the means $\pm$ SD.

## DISCUSSION

In present study, we cloned and characterized the full length cDNA encoding PRP from abalone (*H. discus discus*). Also mRNA expression was analyzed after challenged by different immune modulators namely, *V. alginolyticus*, LPS and β-1,3-glucan.

BLAST analysis results showed that HdPRP had moderate similarity with PRP of invertebrate mollusks and arthropods including crustaceans and other members of GNBP family. Multiple alignment of the HdPRP gene with homologous genes of insects and crustacean species revealed that these sequence motifs remained conserved across species boundaries with slight modifications, indicating the functional conservation of these sequence motifs. LPS binding motif was found in HdPRP at the position from 244 to 261 which was at position 353-370 in bivalvia, *C. farreri* (Su et al., 2004). However, the glucanase motif of abalone PRP was not matched with *C. farreri*. Conserved potential recognition motif for β-1,3-linkage of polysaccharide was observed in abalone at the position of 292-310 which was slightly modified in amino acid sequences of blue shrimp *L. stylirostris* (Roux et al., 2002), white shrimp *L. vennamei* (Cheng et al., 2005) and cray fish *P. leniusculus* (Lee et al., 2000). Glucanase motif of earth worm *E. foetida* was observed in HdPRP with slight changes in amino acid residues. It has been reported that the domain comprising amino acids 149-227 which contained both polysaccharide binding and glucanase motif, in *E. foetida*, binds LPS and β-1,3-glucans from gram negative bacteria and yeast, respectively (Bilej et al., 2001). All four amino acid residues (W, E, I, D) which are considered to be necessary for the catalytic mechanism of bacterial glucanases are conserved in HdPRP (in this study), cray fish LGBP, earthworm CCF-1 (Bilej et al., 2001) and GNBPs. Even though, they have their glucanase motif, none of these PRPs have any glucanase activity. It has been reported that PRPs containing glucanase-like motifs developed from a primitive glucanase in invertebrates and then evolved into proteins without having glucanase activity. However, they can bind glucans and after binding may operate as elicitors of defense response (Cerenius et al., 1994).

Alignment of the HdPRP sequence to other invertebrate proteins containing glucanase-like domains reveals a high homology at the N-terminal region of all sequences. It suggests that this region is involved in binding to the cell wall component of microorganisms, which has been shown for the BGBP of *B. mori* (Ochiai and Ashida, 2000).

The two distinct carbohydrate recognition domains were present in HdPRP and may extend the protection of abalone against different micro organisms. It has been reported that carbohydrate recognition domains in scallop may facilitate the interaction of immunocytes with the pathogen and subsequent induction of cellular defense (Su et al., 2004). It has been reported that BGBP reacts with BG and the BG-BGBP complex induces degranulation of hemocytes, subsequently enhances the activation of proPO system which brings about synthesis of melanin through oxidation of phenols (Vargas-Albores and Yepiz-Plascencia, 2000). It has been shown that PRPs such as *M. sexta* BGRP, earth worm CCF-1, crayfish LGBP and silkworm PGRP mediate activation of the prophenoloxidase activating system (the proPO system) which is considered being a recognition and defense system in many invertebrates. Similar to HdPRP, it has been reported that there was no putative cell adhesive site, RGD in scallop LGBP (Su et al., 2004; Johansson, 1999). In cray fish, the RGD cell adhesion motif was found to serve as a ligand for cell surface intergrins and to mediate blood cell adhesion and cellular immunity (Johansson, 1999). The polysaccharide binding (PsB) motif $^{150}$HAKMPVGDWLWPAIWM$^{165}$ was observed in *Eisenia foetida* from 150-165 amino acid position (Bilej et al., 2001). In abalone, similar motif $^{204}$SAQLPMGDWLWPAIWM$^{219}$ was observed at the position from 204 to 219 with slight modifications at the position 150, 152, 153 155. However, this motif is conserved in other species with slight modifications within the motif in their sequences.

Phylogenetic analysis of the different PRPs of 22 species resulted in a tree which forms monophyletic clade with *B. glabrata* BGRP and *C. farreri* LGBP. HdPRP was grouped with *B. glabrata* BGRP and *C. farreri* LGBP with bootstrap value of 98 and it was closer to

*C. farreri* LGBP with bootstrap value 56, despite of their low identity (40% aa identity). Interestingly, two homologous of *B. glabrata* BGRP and GNBP were far away from one another and *B. glabrata* GNBP was positioned with phylum annelid CCF-like proteins. This may be due to functional difference between those genes. Similar analysis was observed by Zhang et al., 2007 and his computational analysis suggested that the *B. glabrata* GNBP was secreted where as other two genes (*B. glabrata,* GenBank accession no. **EF121824** and scallop, GenBank accession no. **AAP82240**) were intracellular proteins.

Because the recognition of microbial cell wall components is an essential initial step for intra-cellular immune signaling, we were interested to study the gene expression of PRP by injection of *V. aliginolyticus,* LPS from *E. coli* and β-1,3- glucan from *L. digitata* as PAMPs. RT-PCR analysis showed that HdPRP mRNA are constitutively expressed in all the tissues tested, although in hemocyte and hepatopancreas HdPRP mRNA expression could be observed at 35 cycles of RT-PCR. Semi quantitative RT-PCR analysis showed that fleshy prawn *F. chinensis* LGBP was mainly expressed in hemocyte followed by hepatopancreas and gill at 22 cycles of PCR (Du et al., 2007). Also several PRP genes have been described in hemocytes and hepatopancreas in many invertebrate species (Lee et al., 2000; Sritunyalucksana et al., 2002; Cheng et al., 2005). In contrast, disk abalone gill is the major tissue which observed higher mRNA expression than the other tissues tested. This suggests that gill may be the first line defense in disk abalone due to the frequent exposure to the environment and to defend from the invading foreign materials.

It has been reported that constitutive occurrence of some PRPs in insects keep the animal in a constant state of readiness by providing immediate detection of an invading infectious threat. The induction of PRPs provides additional recognition capability until the infection has been cleared (Fabrick et al., 2003). Initially high level (12 h p.i.) of HdPRP mRNA expression, after *V. alginolyticus,* LPS and β-1,3-glucan injection suggests that HdPRP mRNA could be markedly induced by those PAMPs stimulation and perhaps

involved in the recognition of different immune modulators. Similar gene expression pattern results were observed in scallop *C. farreri* hemocytes after challenged using *V. anguilarum.* LGBP was up-regulated initially after stimulation, and subsequently came back to the normal level as the LPS was cleared (Su et al., 2004). It has been reported that BGBP gene cloned from silk worm *B. mori* had strong specific affinity for β-1,3- glucan (Ochiai and Ashida, 2000). Also in *Drosophila*, GNBP also known as LGBP also function as a recognition receptor for LPS and β-1,3-glucan (Kim et al., 2000). Similarly, it has been reported that as the white spot virus infection progressed in the virus injected *P. stylirostris*, the LGBP gene showed an increasingly higher level of expression than the healthy animals (Cheng et al., 2005).

It is interesting to note that HdPRP isolated from abalone recognized gram negative bacteria *Vibrio*, LPS and β-1,3-glucan from *L. digitata* confirming the presence of LPS and glucan recognition and binding sites in HdPRP. A BGBP isolated and purified from the hemocytes of tiger shrimp *P. monodon* can bind only to β-1,3-glucans such as zymosan and curdlan, however, not to LPS (Sritunyalucksana et al., 2002) indicating that its binding is specific for β-1,3-glucans. A LGBP isolated from the hemocytes of crayfish *P. leniucsulus* has binding activity to LPS as well as β-1,3-glucans such as curdlan and laminarin (Lee et al., 2000).

In general, the recognition proteins are present in the plasma or cell surface before the invasion of foreign objects (Vargas-Albores and Yepiz-Plascencia, 2000) and they cannot destroy foreign matters. They trigger other defense system responses such as phagocytosis, encapsulation, nodule formation, activation of proPO cascade etc. Pattern recognition molecules (LBP or BGBP) of crustaceans, recognize LPS or BG, and directly activate the hemocytes, degranulate them, and subsequently induce the activation of proPO system (Soderhall et al., 1990; Soderhall and Cerenius, 1991). It has been reported that, even though, microbial components directly can activate crustaceans defensive cellular functions,

recognition proteins amplify these stimuli, resembling to secondary activities of vertebrate antibodies (Su et al., 2004). Overall data in present study revealed that when foreign object is injected/entered in to the body, the HdPRP gene shows an increasingly higher level of transcription in the injected animal tissues than in the control. This suggests that HdPRP functions as recognizing different PAMP and may activate different immune genes to defense against these pathogens.

Phagocytosis is one of the major celluar immune defense mechanism involved in invertebrates after recognition of foreign materials by PRPs (Sritunyalucksana and Soderhall, 2000). We analyzed CuZnSOD expression in *vivo* for *Vibrio*, LPS and β-1,3-glucan induced abalone to find out whether the foreign intruders has been caused oxidative stress to the abalone. It has been reported that many reactive oxygen species such as superoxide anion ($O_2^-$), hydrogen peroxide ($H_2O_2$) hydroxyl radicals ($OH^\cdot$) and singlet oxygen ($^1O_2$) are produced during phagocytosis (Chiu et al., 2007; Munoz et al., 2000) and $O_2^-$ plays an important role in microbiocidal activity (Bell and Smith, 1993). Though, they play an important role in host defense mechanism, it can damage host cells due to oxidative stress. Therefore, to maintain low levels of ROS, the host cells make a protective mechanism by producing antioxidants. In this study, CuZnSOD mRNA expression was significantly increased after 12 h p.i. compared to control in all the treatments. This may be due to the effort of recognizing the bacterial pathogen or other foreign material and clearance of infection by activating cellular and/or humoral immune responses. Significantly higher CuZnSOD mRNA expression in glucan injected animals at 12, 24 and 48 h p.i. could be suggested that there is a protective role against oxidative stress caused by superoxide produced after exposure to foreign intruders (Bachere et al., 1995; Holmblad and Soderhall, 1999). The phagocytosis activity in abalone injected with bacteria and heterogeneous foreign substances would generate a mass amount of ROS which need to be eliminated by extra SOD translated from more SOD transcripts.

In conclusion, 1263 bp PRP cDNA cloned from *H. discus discus* encodes 420 aa (a mature protein of 400 aa). HdPRP was constitutively expressed and contribute as an inducible acute-phase protein that could play a critical role in abalone-pathogen interaction. According to our knowledge this is the first report on abalone PRP cloning, gene characterization and expression analysis. Hence, information of this HdPRP may be useful in studies of PRPs in other marine invertebrates. This study could be further expanded to find out whether HdPRP responds to viral pathogens as studied by Roux et al, 2002, which may open up new insights into the role of PRPs in viral pathogenesis in invertebrates.

## REFERENCES


Addadi L, Weiner S. 1997. A pavement of pearl. Nature **389**: 912-914.

Allouch J, Helbert W, Henrissat B, Czjzek M. 2004. Parallel substrate binding sites in a β-agarase suggest a novel mode of action on double-helical agarose. Structure **12**: 623–632.

Ashida M, Ishizaki Y, Iwahana H. 1983. Activation of pro-phenoloxidase by bacterial cell walls or beta-1,3-glucans in plasma of the silkworm, *Bombyx mori*. Biochemical and Biophysical Research Communication **113 (2)**: 562– 68.

Bachere E, Miahle E, Rodriguez J. 1995. Identification of defense effectors in the haemolymph of crustaceans with particular reference to the shrimp *Penaeus japonicus* (Bate): prospects and application. Fish and Shellfish Immunology **5**: 597-612.

Bachman ES, McClay DR. 1996. Molecular cloning of the first metazoan β-1,3-glucanase from eggs of the sea urchin *Strongylocentrotus purpuratus*. Proceedings of the National Academy of Sciences of the United States of America **93**: 6808–6813.

Baybeyron T, Potin P, Richard C, Collin O, Kloareg B. 1995. Arylsulfatase from Alteromonas carrageenovora Microbiology **141**: 2897-2904

Bayne BL, Newell RC. 1983. Physiological energetics of marine molluscs. In: Wilbur KM (ed) The Mollusca. Academic Press, New York. **4** : 407–515.

Bell KL, Smith VJ. 1993. *In vitro* superoxide production by hyaline cells of the shore crab *Carcinus maenas* (L.). Developmental and Comparative Immunology.**17**: 211-219.

Bernfeld P. 1955. Amylases α and β. Methods Enzymology 1. 149-155.

Bilej M, Baetselier PDe, Dijck, EV, Stijlemans B, Colige A. 2001 Distinct carbohydrate recognition domains of an invertebrate defense molecule recognize gram-negative and gram-positive bacteria. Journal of Biological Chemistry **276 (49)**: 45840-45847.

Boer PH, Hickey DA. 1986. The alpha amylase gene in *Drosophila melanogaster*, nucleotide sequence, gene structure and expression motifs. Nucleic acids research **14(21):** 8399-8411.

Bouneau L, Lardier G., Fischer C, Ronsin M, Weissenbach J, Bernot A. 2003. Analysis of 148 kb of genomic DNA of *Tetraodon nigroviridis* covering an amylase gene family. Mitochondrial DNA **14(1)**, 1–13.

Bradford MM. 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry **72**, 248-254.

Byrne KA, Lehnert SA, Johnson SE, Moore SS. 1999. Isolation of cDNA encoding a putative cellulase in the red claw crayfish *Cherex quadricarinatus*. Gene **239**:317-324.

Cann IKO, Kocherginskaya S, King MR, White BA, Mackie RI. 1999. Molecular cloning, sequencing, and expression of a novel multidomain mannanase gene from

*Thermoanaerobacterium polysaccharolyticum*, Journal of Bacteriology. **181**: 1643–1651.

Carefoot TH, Quian PY, Taylor BE, West T, Osborne J. 1993. Effect of starvation on energy reserves and metabolism in the northern abalone, *Haliotis kamtschatkana*. Aquaculture **11**: 315–325.

Cerenius L, Liang Z, Duvic B, Keyser P, Hellman U, El Palva T, Iwanaga S. Soderhall K. 1994. Structure and biological activity of a 1,3-beta-D-glucan-binding protein in crustacean blood. Journal of Biological Chemistry.**269 (47)**: 29462–7.

Charrier M, Rouland C. 2001. Mannan-degrading enzymes purified from the crop of the brown garded snail *Helix aspersa* Müller (Gastropoda Pulmonata), Journal of Experimental Zoology **290:** 125–135.

Chauvaux S, Beguin P, Aubert JP. 1992. Site-directed mutagenesis of essential carboxylic residues in *Clostridium thermocellum* endoglucanase celd. Journal of Biological Chemistry **267**:4472-4478.

Chauvaux S, Souchons H, Alzari PM, Chariot P, Beguin P. 1995. Structural and functional analysis of the metal binding sites of *Clostridium thermocellum* Endoglucanase CelD. The Journal of Biological Chemistry **270(17)**: 9757-9762.

Cheng W, Liu CH, Tsai C, Chen J. 2005. Molecular cloning and characterization of a pattern recognition molecule, lipopolysaccharide-and beta 1,3- glucan binding protein (LGBP) from the white shrimp *Litopenaeus vannamei*. Fish and Shellfish Immunology **18**: 297-310.

Chiu CH, Guu YK, Liu CH, Pan TM, Cheng W. 2007. Immune responses and gene expression in white shrimp, *Litopenaeus vannamei*, induced by *Lactobacillus plantarum.* Fish and Shellfish Immunology **23**: 364-377.

Christgau S, Andersen LN, Kauppinen S, Heldt-Hansen HP, Dalboege H. 1994a. Purified enzyme exhibiting mannanase activity; application in oil, paper, pulp, fruit and vegetable juice industry and in carrageenan extraction, Patent Novo-Nordisk, 9425576, 10 November 1994.

Christgau S, Kauppinen S, Vind J, Kofod LV, Dalboge H. 1994b. Expression cloning, purification and characterization of a β-1,4-mannanase from *Aspergillus aculeatus*, Biochemical Molecular Biology International **33:** 917–925.

Christophides GK, Vlachou D, Kafatos FC. 2004. Comparative and functional genomics of the innate immune system in the malaria vector *Anopheles gambiae*. Immunological Review **198**: 127-148.

Crawford AC, Kricker JA, Anderson AJ, Neil RH, Mather PB. 2004. Structure and functiona of a cellulase gene in redclaw crayfish*, Cherax quadricarinatus* **340(2)**: 267-274.

Dai J, Yang Z, LiuW, Bao Z, Han B, Shen S, Zhou L. 2004. Seedling production using enzymatically isolated thallus cells and its application in Porphyra cultivation.

Hydrobiologia **512**: 127–131.

Dierks T. 2003. Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. Cell **113**:435-444.

Dierks T, Miech C, Hummerjohann J, Schmidt B, Kertesz MA,Von Figura K. 1998. Posttranslational formation of formylglycine in prokaryotic sulfatases by modification of either cysteine or serine. Journal of Biological Chemistry. **273**: 25560–25564.

Dierks T, Lecca MR, Schlotterhose P, Schmidt B, von Figura K. 1999. Sequence determinant directing conversion of cystein to formylglycine in eukaryotic sulfatases. EMBO Journal **18**: 2084-2091.

Dierks T, Schmidt B, von Figura K. 1997. Conversion of cysteine to formylglycine : a protein modification in the endoplasmic reticulum. Proceedings of the National Academy of Sciences of the United States of America. **94**: 11963-11968.

Dipakkore S, Reddy CRK, Jha B. 2005. Production and seeding of protoplasts of *Porphyra okhaensis* (Bangiales, Rhodophyta) in laboratory culture, Journal of Applied Phycology. **17:** 331–337.

Dodgson KS, Price RG. 1962. A note on the determination of the ester sulfate content of sulfated polysaccharides. Biochemical Journal **84**: 350-356.

Douglas SE, Mandla S, Gallant JW. 2000. Molecular analysis of the amylase gene and its expression during development in the winter flounder, *Pleuronectes americanus*, Aquaculture **190**: 247–260.

Du XJ, Zhao XF, Wang JX. 2007. Molecular cloning and characterization of a lipopolysaccharide and β-1,3-glucan binding protein from fleshy prawn (*Fenneropenaeus chinesis*). Molecular Immunology **44**: 1085-1094.

Durazo-Beltrán E., D'Abramo LR, Toro-Vazquez JF, Vasquez-Peláez C, Viana MT. 2003. Effect of quantity and quality of triacylglycerols in formulated diets on growth and fatty acid composition in tissue of green abalone (*Haliotis fulgens*). Aquaculture **224**: 257–270.

Dziarski R. Peptidoglycan recognition proteins (PGRPs). 2004. Molecular Immunology **40**: 877-886.

Erasmus JH, Peter AC, Vernon EC. 1997. The role of bacteria in the digestion of seaweed by the abalone *Haliotis midase* Aquaculture **155:** 377-386.

Esquivel ZG, Felbeck H. 2006. Activity of digestive enzymes along the gut of juvenile red abalone, Haliotis rufescens, fed natural and balanced diets. Aquaculture **261**: 615-625.

Fabrick JA, Baker JE, Kanost MR. 2003. cDNA cloning, purification, properties, and function of a beta-1,3-glucan recognition protein from a pyralid moth, *Plodia interpunctella*. Insect Biochemistry and Molecular Biology. **33**: 579-94.

Fearon DT, Locksley RM. 1996.The instructive role of innate immunity in the acquired immune response. Science **272**: 50-54.

Fridovich I. 1995. Superoxide radical and superoxide dismutases. Annual Review of Biochemistry **64**: 97–112.

Froystad MK, Lilleeng E, Sundby A, Krogdahl A. Cloning and characterization of α-amylase from Atlantic salmon (*Salmo salar* L.). Comparative Biochemistry and Physiology, Part A **145**: 479-492.

Gabbott PA. 1976. Energy metabolism. In: Bayne BL (ed) Marine Mussels: their Ecology and Physiology. Cambridge Unversity Press. 293-357.

Gall EA, Chiang YM, Kloareg B. 1993. Isolation and regeneration of protoplasts from *Porphyra dentata* and *Porphyra crispate*. European Journal of. Phycology. **28**: 277–283.

Gibbs MD, Elinder AU, Reeves RA, Bergquist PL. 1996. Sequencing, cloning and expression of a β-1,4-mannanase gene, manA, from the extremely thermophilic anaerobic bacterium, *Caldicellulosiruptor* Rt8B. 4, FEMS Microbiology Letters. **141**, 37–43.

Gordon HR, Cook PA. 2001. World abalone supply, market and pricing: historical, current and future. Journal of Shellfish Research **20:** 567-570.

Gretz MR, Sommerfeld MR, Aronson JM. 1982. Cellwall composition of the generic phase of *Bangia atropurpurea*. Botanical Mar. **25**: 529–535.

Gueguen Y, Voorhorst WG., Van der Oost J, De Vos WM. 1997. Molecular and biochemical characterization of an endo-β-1,3-glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. Journal of Biological Chemistry **272**: 31258–31264.

Halliwell B, Gutteridge JMC. 1999. Free Radicals in Biology and Medicine. 3rd edition. Oxford University Press, Oxford.

Henry M, Benlimame N, Boucaud-Camou E, Mathieu M, Donval A. 1993. The amylase-secreting cells of the stomach of the scallop, *Pecten maximus*: ultrastructural, immunohistochemical and immunocytochemical characterizations. Tissue Cell **25**: 537–548.

Hilge M, Gloor SM, Rypniewski W, Sauer O, Heightman TD, Zimmermann W, Kaspar W, Piontek K. 1998. High-resolution native and complex structure of thermostable β-mannanase from *Thermomonospora fusca* substrate specificity in glycosyl hydrolase family 5, *Structure* **6(11)**: 1433–1444.

Hoffmann JA, Reichart JM, Hetru C. 1996. Innate immunity in higher insects. Current Opinion in Immunology **8(1)**: 8-13.

Holmblad T, Soderhall K. 1999. Cell adhesion molecules and antioxidative enzymes in a crustacean, possible role in immunity. Aquaculture **172**:111-123.

Hoshi M, Moriya T. 1980. Arylsulfatase of sea-urchin sperm. 2. Arylsulfatase as a lysin of sea-urchins, Developmental Biology **74**: 343–350

Ibarrola I, Iglesias JIP, Navarro E. 1996. Differential absorption of biochemical components in the diet of the cockle *Cerastoderma edule*: enzymatic responses to variations in seston composition. Canadian Journal of Zoology **74:** 1887–1897.

Ibarrola I, Larretxea X, Iglesias JIP, Urrutia MB, Navarro E. 1998. Seasonal variation of digestive enzyme activities in the digestive gland and the crystallin style of the common cockle *Cerastoderma edule*. Comparative Biochemistry and Physiology-Part A: Molecular and Intergrative Physiology **121**: 25–34.

Ibarrola I, Etxeberria M, Iglesias JIP, Urrutia MB, Angulo A. 2000. Acute and acclimated responses of the cockle *Cerastoderma edule* (L.) to changes in food quality and quantity. II. Enzymatic, cellular and tissular responses of the digestive gland. Journal of Experimental Marine Biology and Ecology **252**: 199–219.

Jiang H, Mab C, Lua ZQ, Kanost MR. 2004. Beta-1,3-Glucan recognition protein-2 (betaGRP-2) from *Manduca sexta*: an acute-phase protein that binds beta-1,3-glucan and lipoteichoic acid to aggregate fungi and bacteria and stimulate prophenoloxidase activation. Insect Biochemistry and Molecular Biology **34**: 89–100.

Johansson MW. Cell adhesion molecules in invertebrate immunity. 1999. Developmental and Comparative Immunology **23 (4–5)**: 303–315.

Jomori T, Natori S. 1992. Function of the lipopolysaccharide-binding protein of *Periplaneta americana* as an opsonin. FEBS Letters **296 (3)**: 283–286.

Khademi S, Guarino LA, Watanabe H, Tokuda G, Meyer EF. 2002. Structure of an endoglucanase from termite, *Nasutitermes takasagoensis*. Acta Crystallographica Section D: **D58**, 653–659.

Kim YS, Ryu JH, Han SJ, Choi KH, Nam KB, Jang IH, Lemaitre B, Brey PT, Lee WJ. 2000. Gram-negative bacteria binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. Journal of Biological Chemistry. **275**: 32721–32727.

Kobayashi Y, Echigen R, Mada M, Mutai M. 1987. Effects of hydrolyzates of konjac mannan and soybean oligosaccharides on intestinal flora in man and rats. In: Mitsuoka T. (ed.), Intestinal Flora and Food Factors. Gakkai Shuppan Centre, Tokyo, 79–97.

Koyama I, Komine S, Hokari S, Yakushijin M, Matsunaga T, Komoda T. 2001. Expression of α-amylase gene in rat liver: liver-specific amylase has a high affinity to glycogen. Electrophoresis **22**: 12–17

Koizumi N, Imamura M, Kadotani T, Yaoi K, Iwashana H, Sato R. 1999. The lipopolysaccharide-binding protein participating in hemocyte nodule formation in the silkworm *Bombyx mori* is a novel member of the C-type lectin super family with two different tandem carbohydrate-recognition domains. FEBS Letters; **443(2)**:

139–143.

Kozhemyako VB, Rebrikov DV, Lukyanov SA, Bogdanova EA, Marin A, Mazur AK, Kovalchuk SN, Agafonova EV, Sova VV, Elyakova LA, Rasskazov VA. 2004. Molecular cloning and characterization of an endo-1,3-β-D-glucanase from the mollusk *Spisula sachalinensis*. Comparative Biochemistry and Physiology-Part B: Biochemistry and Molecular Biology **137**: 169–178.

Kumar S, Tamura K, Nei M. 2004 MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Bioinformatics **5(2)**:150-163.

Kurota H,Yamaguchi M. 1997. Activatory effect of calcium-binding protein regucalcin on ATP-dependent calcium transport in the basolateral membranes of rat kidney cortex. Molecular and Cellular Biochemistry **169**: 149–156.

Kyte J, Doolittle R F, 1982. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology 157:105.

Landgrebe J, Dierks T, Schmidt B, von Figura K. 2003. The human SUMF1 gene, required for posttranslational sulfatase modification, defines a new gene family which is conserved from pro-to eukaryotes. Gene **316**:47-56.

Larsson AM, Anderson L, Xu B, Munoz IG, Isabel U, Janson J, Stalbrand H, Stahlberg J. 2006. Three-dimensional crystal structure and enzymic characterization of β-mannanase Man5A from blue mussel *Mytilus edulis* . Journal of Molecular Biology **357**: 1500–1510.

Lee SY, Wang R, Soderhall K. 2000. A lipopolysaccharide-and beta-1, 3-glucan-binding protein from hemocytes of the fresh water crayfish *Pacifastacus leniusculus*. Journal of Biological Chemistry; **275 (2):** 1337–43.

Le Moine S, Sellos D, Moal J. 1997. Amylase in *Pecten maximus* (Mollusca, bivalves): protein and cDNA characterization expression in the digestive gland. Molecular Marine Biology and Biotechnology **6**: 228–237.

Le Moullac G, Klein B, Sellos D, Van Wormhoudt A. 1997. Adaptation of trypsin, chymotrypsin and alpha amylase to casein level and protein source in *Penaus vannamei* (Crustacea Decapoda). Journal of Experimental Marine Biology and Ecology 208:107–125.

Leveque E, Haye B, Belarbi A. 2000. Cloning and expression of an α-amylase encoding gene from the hyperthermophilic archaebacterium *Thermococcus hydrothermalis* and biochemical characterisation of the recombinant enzyme. FEMS Microbiology Letters **186**: 67-71.

Lim WJ, Park SR, An CL, Lee JY, Hong SY, Shin EC, Kim EJ, Kim JO, Kim H, Yun HD. 2003. Cloning and characterization of a thermostable intracellular α-amylase gene from the hypothermophilic bacterium *Thermotoga maritime* MSB8. Research in Microbiology **154**:681-687.

Loya-Javellana GN, Fielder DR, Thorne MJ. 1993, Food choice by free-living stages of the tropical freshwater crayfish. *Cherax quadricarinatus* (Parastacidae: Decapoda), *Aquaculture* **118**: 299–308.

Ma P, Liu Y, Reddy KP, Chan WK, Lam TJ. 2004. Characterization of the seabass pancreatic alpha-amylase gene and promoter. General and Comparative Endocrinology. **137**:78–88.

MacGregor EA, Janecek S, Svensson, B. 2001. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. **1546**: 1-20.

Machius M, Wiegand G, Huber R. 1995. Crystal structure of calcium-depleted *Bacillus licheniformis* alpha-amylase at 2.2 A resolution. Journal of Molecular Biology **246(4):** 545–559.

Mackay RM, Baird S, Dove MJ, Erratt JA, Gines M, Moranelli F. 1985. Glucanase gene diversity in prokaryotics and eukariyotic organisms. Biosystems. 18: 279-292.

Mathers NF. 1973. A comparative histochemical survey of enzymes associated with the processes of digestion in *Ostrea edulis* and *Crassostrea angulata* (Mollusca-Bivalvia). Journal of Zoology (London) **169**: 169–179.

Matsui M, Kakut M, Misaki A. 1996. Fine structural features of oyster glycogen: mode of multiple branching. Carbohydrate Polymer **31(4)**: 227-235.

Melo MRS, Feitosa JPA, Freitas ALP, Paula de RCM. 2002. Isolation and characterization of soluble sulfated polysaccharide from the red seaweed *Gracilaria cornea*. Carbohydrate Polymer **49**: 491-498.

Milanesi AA, Bind JWC. 1972. Lysosomal enzymes in aquatic species II. Distribution and particle properties of thermally acclimated muscle lysosomes of rainbow trout *Salmo gairdeneri*, Comparative Biochemistry and Physiology **41**: 473–491.

Misawa H, Yamaguchi M. 2000. The gene of $Ca^{2+}$ binding protein regucalcin is highly conserved in vertebrate species. International Journal of Molecular Medicine **6**: 191-196.

Moal J, Martin-Jezequel V, Harris RP, Samain JF, Poulet SA. 1987. Interspecific and intraspecific variability of the chemical composition of marine phytoplancton. Oceanol. Acta **10**: 339– 346.

Morales AE, Pérez-Jiménez A, Hidalgo MC, Abellán E, Cardenete G.. 2004. Oxidative stress and antioxidant defenses after prolonged starvation in *Dentex dentex* liver. Comparative Biochemistry and Physiology, Part C **139**:153–161.

Morton B. 1983. Feeding and Digestion in Bivalvia. The Mollusca, Physiology, Part 2, 5: 65–147.

Munoz M, Cedeno R, Rodriguez J, Van der Knaap WPW, Mialhe E, Bachere E. 2000. Measurement of reactive oxygen intermediate production in haemocyte of the penaeid shrimp, *Penaeus vannamei*. Aquaculture **191**: 89-107.

Nakajima Y, Natori S. 2000. Identification and characterization of an anterior fat body protein in an insect. Journal of Biochemistry **127**: 901-908.

Nakayama S, Kretsinger RH. 1994. Evolution of the EF-hand family of proteins. Annual Review of Biophysics and Biomolecular Structure **23**: 473-507.

Nishida Y, Suzuki K, Kumagai Y, Tanaka H, Inoue A, Ojima T. 2007. Isolation and primary structure of a cellulase from the Japanese sea urchin *Strongylocentrotus nudus*. Biochimie **89**: 1002-1011

Obaya AJ. 2006. Molecular cloning and initial characterization of three novel human sulfatases. Gene **372**: 110–117.

Ochiai M, Ashida M. 2000. The binding domain and the cDNA cloning of β-1,3-glucan recognition protein from the silkworm, *Bomryx mori*. Journal of Biological Chemistry **275 (7)**: 4995–5002.

Ootsuka S, Saga N, Suzuki K, Inoue A, Ojima T. 2006. Isolation and cloning of an endo-β-1,4-mannanase from Pacific abalone Haliotis discus discus hannai. Journal of Biotechnology 125: 269-280.

Osterloh D, Ivanenkov VV, Gerke V. 1998. Hydrophobic residues in the C-terminal region of S100A1 are essential for target protein binding but not for dimerization. Cell Calcium **24 (2)**: 137-151.

Pandey A, Nigam P, Soccol CR, Soccol VT, Singh D, Mohan R. 2000. Advances in microbial amylases. Applied Biochemistry and Biotechnology **31** (Pt 2): 135-152.

Parenti G, Meroni G, Ballabio A. 1997. The sulfatase gene family. Current Opinion in Genetics & Development **7**: 386–391.

Pascual P, Pedrajas JR, Toribio F, López-Barea J, Peinado J. 2003. Effect of food deprivation on oxidative stress biomarkers in fish (*Sparus aurata*). Chemico-Biological International **145**: 191–199.

Paulet YM, Bekhadra F, Devauchelle N, Donval A, Dorange G.. 1997. Seasonal cycles, reproduction and oocyte quality in Pecten maximus from the Bay of Brest Cycles. Annales de l'Institute Oceanographique **73**: 101-112.

Percival E, McDowell RH. 1967. Other neutral polysaccharides, food reserve and structural food reserve. In: Percival E, McDowell RH. (eds.), Chemistry and Enzymology of Marine Algal Polysaccharides. Academic Press, London, 73–96.

Polne-Fuller M, Gibor A. 1984. Developmental studies in *Porphyra*. I. Blade differentiation in *Porphyra perforata* as expressed by morphology, enzymatic digestion, and protoplast regeneration, Journal of Phycology. **20:** 609–616.

Reid RGB. 1966. Digestive tract enzymes in the bivalvia *Lima hians* (Gmelin) and *Mya arenaria* L. Comparative Biochemistry and Physiology **17**: 417–433.

Reid RGB. 1968. The distribution of digestive tract enzymes in lamellibranchiate bivalves. Comparative Biochemistry and Physiology **24**: 727–744.

Roberton AM, Wright DP. 1997. Bacterial glycosulfatases and sulphomucin degradation. Canadian Journal of Gastroenterology **11**: 361-366.

Robinson MK, Rustum RR, Chambers EA, Rounds JD, Wilmore DW, Jacobs DO. 1997. Starvation enhances hepatic free radical release following endotoxemia. Journal of Surgery Research **69**: 325–330.

Roux MM, Pain A, Klimpel KR, Dhar AK. 2002. The lipopolysaccharide and beta 1,3-glucan binding protein gene is up regulated in white spot virus infected shrimp, Penaeus *stylirostris*. Journal of Virology 7140-7149.

Schwede T, Kopp J, Guex N, Peitsch MC. 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Research **31**: 3381-3385.

Seiderer LJ, Newell RC. 1979. Adjustment of the activity of α-amylase extracted from the stylet of the black mussel *Choromytilus meridionalis* (Krauss) in response to thermal acclimation. Journal of Experimental Marine Biology and Ecology. **39**: 79–86.

Seo HJ, Kim JH, Byun DS, Kim HR. 2001. Purification and characterization of arylsulfatase produced from Sphingomonas sp. Nov. 75$^{th}$ meeting of Japaneese Society of Fisheries Science, Yokohama, Japan 1575-1578.

Shimizu E, Ojima T, Nishita K. 2003. cDNA cloning of an alginate lyase from abalone, *Haliotis discus discus hannai* **338**: 2841-2852.

Shimokawa N, Yamaguchi M. 1992. Calcium administration stimulates the expression of calcium-binding protein regucalcin mRNA in rat liver. FEBS Letters **305**: 151-154.

Shimokawa N, Yamaguchi M. 1993. Expression of hepatic calcium binding protein regucalcin mRNA is mediated through $Ca^{2+}$/calmodulin in rat liver. FEBS Letters **316**: 79-84.

Sies H. 1985. Oxidative stress: introductory remarks. In: Sies, H. (ed), Oxidative Stress. Academic press, Incorporated, London.

Smant G., Stokkermans JPWG., Yan Y, De Boer JM, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, Bakker J. 1998. Endogenous cellulases in animals: isolation of β-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proceedings of the National Academy of Science of the United States of America **95**: 4906-4911

Smith VJ, Soderhall K, Hamilton M. 1984. Beta 1,3-glucan induced cellular defense reaction in the shore crab, *Carcinus maenas*. Comparative Biochemistry and Physiology, Part A **77**:636-9.

Soderhall K, Aspan A, Duvic B. 1990. Is the proPO-system involved in defense and recognition in arthropods? Research in Immunology **141**: 946-948.

Soderhall K, Cerenius L. 1991. Crustacean immunity. Annual Review of Fish Disease. **2**:3-23.

Sritunyalucksana K, Lee SY, Soderhall K. 2002 A beta-1,3-glucan binding protein from the

black tiger shrimp *Penaeus mondon*. Developmental and Comparative Immunology **26**: 237–45.

Sritunyalucksana K, Soderhall K. 2000. The proPO and clotting system in crustaceans. Aquaculture **191**:53-69.

Su J, Song L, Xu W, Wu W, Wu L, Li H, Xiang J. 2004. cDNA cloning and mRNA expression of the lipopolysaccharide-and beta-1,3-glucan binding protein gene from scallop *Chlamys farreri*. Aquaculture 239: 69-80.

Sugimura M, Watanabe H, Lo N, Saito H. 2003. Purification, characterixation, cDNA cloning and nucleotide sequencing of a cellulase from the yellow spotted longicorn beetle, *Psacothea hilaris*. Europian Journal of Biochemistry **270**: 3455-3460.

Sunna A, Gibbs MD, Chin CWJ, Nelson PJ, Bergquist PL. 2000. A gene encoding a novel multidomain β-1,4-mannanase from *Caldibacillus cellulovorans* and action of the recombinant enzyme on kraft pulp, Applied and Environmental Microbiology. **66:** 664–670.

Suzuki K, Ojima T, Nishita K. 2003. Purification and cDNA cloning of a cellulase from abalone *Haliotis discus hannai.* European Journal of Biochemistry. **270**: 771-778.

Takahashi H, Yamaguchi M. 1997. Stimulatory effect of regucalcin on ATP-dependent calcium transport in rat liver plasma membrane. Molecular and Cellular Biochemistry **168:** 149-153.

Teo LH, Sabapathy V. 1990. Preliminary report on the digestive enzymes present in the digestive gland of *Perna viridis*. Marine Biology 106: 403–407.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Research **22**: 4673-4680.

Tokuda G, Lo N, Watanabe H, Slaytor M, Matsumoto T, Noda H. 1999. Metazoan cellulase genes from termites: intron/exon structure and sites of expression Biochimica et Biophysica Acta **1447**: 146-159.

Tomme P, Warren RAJ, Gilkes NR. 1995. Cellulose hydrolysis by bacteria and fungi, Advanced Microbial Physiology **37**: 1-81.

Vargas-Albores F, Yepiz-Plascencia G.M. 2000. Beta glucan binding protein and its role in shrimp immune response. Aquaculture **191**: 13-21.

Viana MT, Abramo LRD, Gonzalez MA, Garcia-Suarez JV, Shimada A, Vasquez-Pelaez. C. 2007. Energy and nutrient utilization of juvenile green abalone (*Haliotis fulgens*) during starvation. Aquaculture **264**: 323-329.

Violet M, Meunier JC. 1989. Kinetic studies of the irreversible thermal inactivation of *Bacillus licheniformis* α-amylase. Biochemical Journal **263**:665-670.

Watanabe H, Nakamura M, Tokuda G, Yamaoka I, Scrivener AM, Noda H. 1997. Site of

secretion and properties of endogenous endo-β-1,4-glucanase components from *Reticulitermes speratus* (Kolbe), a Japanese subterranean termite. Insect Biochemistry and Molecular Biology **27 (97):** 305-313.

Wang YC, Chang PS, Chen HY. 2007. Tissue expressions of nine genes important to immune defence of the Pacific white shrimp *Litopenaeus vannamei*. Fish and Shellfish Immunology **23**:1161-77.

Wang, J., Ding, M., Chen, Y. Li, Q., Xu, G., Zhao, F. 2003. Isolation of a multifunctional enzogeneous cellulase gene from mollusc, *Ampullaria crossean*. Acta Biochim Biophys Sin **35**: 941-946.

Wicker C, Puigserver A, Scheele G. 1984. Dietary regulation of levels of actine mRNA coding for amylase and serine protease zymogens in the rat pancreas. European Journal of Biochemistry **139**: 381–387.

Wittstock U, Fischer M, Swendsen I, Halkier A. 2000. Cloning and characterization of two cDNAs encoding sulfatase in the Roman snail, *Helix pomatia*. IUBMB Life. **49 (1):** 71-76.

Wong WKR, Gerhard B, Guo ZM, Kilburn DG, Warren AJ, Miller RCJr. 1986. Characterization and structure of an endoglucanase gene cenA of *Cellulomonas fimi*. Gene **44**: 315-324.

Wong KKY, Saddler JN. 1993. Applications of hemicellulases in the food, feed, and pulp and paper industries. In: Coughlan MP, Hazlewood GP. (eds), Hemicellulose and Hemicellulases. Portland Press Ltd., London, United Kingdom. 127–143.

Yamaguchi M. 2000. Break throughs and views. The role of regucalcin in nuclear regulation of regenerating liver. Biochemical and Biophysical Research Communications **276**: 1-6.

Yamaguchi M. 2005. Role of regucalcin in maintaining cell homeostasis and function (review). International Journal of Molecular Medicine **15(3**): 371-389.

Yamaguchi M, Isogai M. 1993. Tissue concentration of calcium binding protein regucalcin in rats by enzyme-linked immunoadsorbent assay. Molecular and Cellular Biochemistry **122**: 65-68.

Yamaguchi M, Kurota H. 1995. Expression of calcium-binding protein regucalcin mRNA in the kidney cortex of rats: the stimulation by calcium administration. Molecular and Cellular Biochemistry **146**: 71-77.

Yamaguchi M, Nakajima R. 2002. Role of regucalcin as an activator of sarcoplasmic reticulum $Ca^{2+}$-ATPase activity in rat heart muscle. Journal of Cell Biochemistry **86**: 184-193.

Yamaguchi M' Sakurai T. 1992. Reversible effect of calcium-binding protein regucalcin on the $Ca^{2+}$ induced inhibition of deoxyuridine 5'-triphosphatase activity in rat liver cytosol. Molecular and Cellular Biochemistry **110**: 25-29.

Yamaguchi M, Ueoka S.1997. Inhibitory effect of calcium-binding protein regucalcin on ribonucleic acid synthesis in isolated rat liver nuclei. Molecular and Cellular Biochemistry **173**: 169-175.

Yamaguchi M, Yamamoto T. 1978. Purification of calcium binding substances from soluble fraction of normal rat liver. Chemical and Pharmaceutical Bulletin (Tokyo) **26**: 1915-1918.

Yamaura I, Matsumoto T. 1993. Purification and some properties of *endo*-1,4-β-d-mannanase from a mud snail, *Pomacea insularus* (de Ordigny), *Biosci. Biotech. Biochem.* **57:** 1316–1319.

Yamaura I, Nozaki Y, Matsumoto T, Kato T. 1996. Purification and some properties of an endo-1,4-beta-d-mannanase from a marine mollusc *Littorina brevicula*. Bioscience Biotechnology and Biochemistry **60**: 674–676.

Yan T, Theo LH, Sin YM. 1996. Effects of metals on alpha-amylase activity in the digestive gland of the green mussel, *Perna viridis* L. Bulletin of Environmental Contamination and Toxicology **56**: 677–682.

Yoshida H, Kinoshita K, Ashida M. 1996. Purification of a peptydoglycan recognition protein from the hemolymph of the silkworm, *Bombyx mory*. Journal of Biological Chemistry. **271**:13854-13860.

Yu XQ, Kanost MR. 2003. *Manduca sexta* lipopolysaccharide-specific immulectin-2 protects larvae from bacterial infection. Development & Comparative Immunology **27**:189–96.

Xu B, Hagglund P, Stalbrand H, Janson JC. 2002a. Endo-beta-1,4-mannanases from blue mussel, *Mytilus edulis*: purification, characterization, and mode of action. Journal of Biotechnology **92**: 267–277.

Xu B, Mun oz IG, Janson JC. Stahlberg J. 2002b. Crystallization and X-ray analysis of native and selenomethionyl β-mannanase Man5A from blue mussel, *Mytilus edulis*, expressed in *Pichia pastoris*. Acta Crystallography. section D **58**: 542–545.

Xu B, Sellos D, Janson J. 2002c. Cloning and expression in *Pichia pastoris* of a blue mussel (*Mytilus edulis*) β-mannanase. European Journal of Biochemistry **269**: 1753-1760.

Zhang SM, Zeng, Y, Loker EC. 2007. Characterization of immune genes from the schistosome host snail *Biomphalaria glabrata* that encode peptidoglycan recognition proteins and gram-negative bacteria binding protein. Immunogenetics **59**: 883-898.

Zoccola D, Tambutte E, Senegas-Balas F, Michiels JF, Failla JP, Jaubert J, Allemand D. 1999. Cloning of a calcium channel alpha 1 subunit from the reef-building coral, *Stylophora pistillata*. Gene **227**: 157-167.

Zoltowska K. 2001. Purification and characterization of alpha amylases from the intestine and muscle of *Ascaris suum* (Nematoda). Acta Biochimica Polonica **48 (3)**: 763-774.

**ACKNOWLEDGMENTS**

I would like to express my sincere appreciation and gratitude to my supervisor, Professor Jehee Lee for giving me the opportunity to study in Molecular Genetic Laboratory under his supervision, and I am greatly indebted to him for his guidance, advice, encouragement and untiring support extended to me in all matters throughout my study period at the Cheju National University, in the Republic of Korea. I wish to express my gratitude to Prof. Choon-Bok Song, Prof. Moon-Soo Heo, Prof. In-Kyu Yeo, Prof. You-Jin Jeon, Prof. Joon Bum Jeong of the Department of Biotechnology for teaching, sharing their research experiences and allowing me to use their research laboratories. I am thankful to Prof Ki-Wan Lee, Prof Kwang-Sik Choi, Prof Young-Don Lee, Prof Kyeong-Jun Lee and Prof Gi-Young Kim for the helpful discussions and allowing me to use their laboratory facilities.

I am grateful to Prof. Upali Samarajeewa and Prof. Chamara Illeperuma, Dept. of Food Science and Technology, University of Peradeiniya, Sri Lanka and Dr. K.H. Sarananda of Food Research Unit, Dept. of Agriculture, Sri Lanka for leading me to engage in research activities and higher studies. I am especially thankful to Dr. Nalin Siriwardana and Dr. Prashani Ekanayaka for introducing me to Prof Jehee Lee, of the Republic of Korea, where I could develop my carrier as a successful researcher.

It is a great experience and pleasure to spend my study life with past and present Korean colleagues of my laboratory, and I am thankful to Dr. Hyun-Sil Kang, Hojin Park and Kyoung-Im Kang for all the assistance given in all my difficulties. My special thanks are due to Chul-hong Oh, for his untiring support extended to me in many ways, during my academic life in Korea. I appreciate the cooperation extended by my lab colleagues Young-Deuk Lee, Su-Kyoung Lee, Young-Hwa Kim, Yu-Chul Kim, Hyunjae Kim. I also appreciate Mrs. Ilson Whang for her kindness, helpful suggestions and support towards the success of my studies in Korea. I appreciate the encouragement given by all former and present international students,

and especially thankful to Dr. Wang Ning and Wan Qiang. I wish to thank colleagues of other laboratories, especially Maeng-Jin Kim, Soo-Jin Heo, Seon-Heui Cha, Gin-Nae Ahn, Song-hun Han, Hye Young Yang and Jan-Tae Won for their numerous supports in many ways.

I appreciate the encouragement and support given by past and present Sri Lankan friends specially Dr. Rohan Karawita, Dr. Yasantha Athukorala, Mahinda Senevirathna, Mahanama De Zoysa and Nadun Karunathilaka, to make my life in Jeju a memorable. I am thankful and appreciate the support extended by past and present officers in the Department of Biotechnology especially Messers Oh-Sang Gyu, Min-Ju Kim and Kyoung-Ju Kim, the NURI program for the scholarship offered for my studies and the financial support extended to me to participate in various International and local conferences. I am thankful to Prof Song Sung-Dae, the Director, and the staff of the Graduate School of CNU for their generous co-operation extended during my studies study in South Korea.

I also appreciate the help extended by Ms Priyantha Samaraweera, and especial thanks are due to Ms Thushari Hettiarachchi, for her love, care and support extended to me in many ways to make my stay in Korea a comfortable and happy. Finally, last, but not least, a debt of gratitude is owed to my loving parents and family members for being a source of strength and inspiration for all my work.