Molecular Cloning and Phylogenetic Analysis of a Chitin Deacetylase Isolated from the Epidermis of the Red Snow Crab *Chionoecetes japonicus*

Kakeru Fujimori, Hideto Fukushima, Masahiro Matsumiya

Department of Marine Science and Resources, College of Bioresource Sciences, Nihon University, Fujisawa, Japan
Email: matsumiya@brs.nihon-u.ac.jp

**Abstract**

Chitin deacetylase (CDA; EC 3. 5. 1. 41) catalyzes the deacetylation of chitin. In this study, we successfully cloned and sequenced a chitin deacetylase gene from the red snow crab *Chionoecetes japonicus*. By using reverse transcription-polymerase chain reaction (RT-PCR) and 5’ and 3’ rapid amplification of cDNA ends, we obtained a 2141-bp amplicon containing a chitin deacetylase gene (*CjCDA*) from the epidermis of *C. japonicus*. The amplicon contains a 1575-bp open reading frame that is predicted to encode a 525-amino acid protein. The structure predicted from the deduced amino acid sequence included an N-terminal signal peptide, chitin-binding domain (CBD), low-density lipoprotein receptor class A domain (LDL-A), and catalytic domain. Comparative analysis of the deduced amino acid sequence of *CjCDA* revealed the highest homology (74%) to gastrolith protein 59 of *Cherax quadricarinatus*. We used RT-PCR to evaluate the expression of *CjCDA* in various tissues of *C. japonicus*, and we observed that *CjCDA* was expressed only in the epidermis. A phylogenetic analysis, using the amino acid sequences of *CjCDA* and other known chitin deacetylases, showed that *CjCDA* belonged to a group of crustacean chitin deacetylases. To our knowledge, this is the first study reporting the cDNA cloning of a chitin deacetylase from a crab.

**Keywords**

Chitin Deacetylase, Molecular Cloning, *Chionoecetes japonicus*, Phylogenetic Analysis, Expression Analysis

**1. Introduction**

Chitin is a β-1,4-linked linear polysaccharide of N-acetyl-D-glucosamine
(GlcNAc). Chitin is widespread in nature, and is present in the cell walls of fungi, the exoskeletons of insects and crustaceans, and the cuttlebones of mollusks, and it is the second most abundant component of biomass after cellulose [1]. Chitosan, which is deacetylated chitin [2], is a β-1,4-linked polymer of glucosamine (GlcN) that is soluble in dilute acids and viscous [3] and is widely used in textile goods, cosmetics, as food additives, and in medical applications [4] [5] [6] because of its antibacterial effects, biodegradability, and biocompatibility [7] [8]. Chitosan is produced industrially via deacetylation of chitin obtained from the shells of crustaceans, such as shrimp and crab, by heat-treating chitin in a concentrated alkaline solution [9]. This alkali treatment does not lead to complete deacetylation but yields chitosan containing about 30% of GlcNAc at random [3]. In addition, chitosan production from 1 kg of chitin requires not only 6.3 kg of HCl and 1-8 kg of NaOH but also nitrogen, 0.5 t of process water, and 0.9 t of cooling water [10]. In contrast, enzymatic deacetylation proceeds under mild conditions and can either completely deacetylate or partially deacetylate chitin at specific sites [11] [12]. Therefore, chitosan production using enzymes is expected to replace alkaline production.

Chitin deacetylase (CDA) is an enzyme that catalyzes the hydrolysis of the acetamido groups of GlcNAc in chitin, producing GlcN and acetic acid [13]. CDA is a member of carbohydrate esterase family 4 (CE4). Members of CE4 share a conserved region called the NodB homology domain or polysaccharide deacetylase domain [14]. CDA was first discovered in the fungus Mucor rouxii [15] [16], and was subsequently found in various fungi and characterized. This enzyme was shown to play important roles in various physiological processes, including cell wall formation, spore formation, and fruiting body growth [17] [18] [19]. In insects, a CDA was discovered in Mamestra configurata [20], and was subsequently identified in the peritrophic membranes of numerous insects. In addition, in beetles, chitin is converted to chitosan when the trachea is extended, which suggests the involvement of CDA in this process [13] [21].

There have been a limited number of reports on CDA genes in crustaceans; specifically, in two varieties of shrimp, red claw shrimp (Cherax quadricarinatus) [22] and black tiger shrimp (Penaeus monodon) [23]. The CDA in P. monodon is reported to be expressed mainly in the gills and play a role in pathogen defense [23]. We attempted, for the first time, to amplify full-length CDA genes from crabs, which has not been often reported, and to determine the structure and phylogenetic relationships among CDAs.

2. Materials and Methods

2.1. Materials

Red snow crab Chionoecetes japonicas (male, carapace width; 12 cm, weight; 725 g, collected at Ishikawa prefecture Japan on November in 2015) was purchased from Nakagawa Co., Ltd. (Japan). Sample was transported living and then stored at −80°C.
2.2. Cloning of the Chitin Deacetylase cDNA from *C. japonicus*

The sequences of all primers used are presented in Table 1. Total RNA was extracted from the epidermis of the leg muscle of *C. japonicus* using ISOGEN II reagent (Nippon Gene, Tokyo, Japan) according to the manufacturer’s instructions. First-strand cDNA was synthesized using 500 ng of total RNA and oligo dT primers with PrimeScript II Reverse Transcriptase (RNase H-free) (Takara Bio, Shiga, Japan) according to the manufacturer’s instructions. Degenerate primers were designed for the reverse transcriptase-polymerase chain reaction (RT-PCR) from conserved sequences in insect and crustacean chitin deacetylases. In the first PCR, *C. japonicus* cDNA was used as a template and CDA F-1 and CDA R-1 were used as primers. The PCR conditions were as follows: 95°C for 2 min, followed by 30 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 50 s. The primers used are listed in Table 1, and the primer combinations are shown in Figure 1.

For the 3’ rapid amplification of cDNA ends (RACE), we designed primers specific to *CjCDA* (*CjCDA 3’-1 and CjCDA 3’-2; Table 1) based on the detected sequences. We amplified cDNA fragments encoding the 3’ region of *CjCDA* using

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
<th>Number of bases (bp)</th>
<th>Annealing temperatures (°C)</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDA F-1</td>
<td>TGYMNGAYGTNATHCARTGYAC</td>
<td>23</td>
<td>58.4</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA F-2</td>
<td>GGNYTNCARGCNHTNMGTGYCC</td>
<td>23</td>
<td>63.7</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA F-3</td>
<td>GAYTYWNSGAYGGNWSNGAYGA</td>
<td>23</td>
<td>61.3</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA F-4</td>
<td>ACNTTYGAYGAYGCNATHAA</td>
<td>20</td>
<td>51.3</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA R-1</td>
<td>CCAYTGDATNACCTGNGTCATNGT</td>
<td>24</td>
<td>58.5</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA R-2</td>
<td>GCNGTDATNGTNSWRTCTANAARAA</td>
<td>26</td>
<td>57.1</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA R-3</td>
<td>AAYTGNBWRTRNCNCCNACNGKNA</td>
<td>25</td>
<td>60.8</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CDA R-4</td>
<td>TAYTTRTNNSWNACRARRAANAGT</td>
<td>23</td>
<td>52.4</td>
<td>Primary PCR</td>
</tr>
<tr>
<td>CjCDA 5’-1</td>
<td>GCAGCGGCTTTACCTTCTTTCG</td>
<td>22</td>
<td>58.6</td>
<td>5’RACE</td>
</tr>
<tr>
<td>CjCDA 5’-2</td>
<td>TCTGAGCCTCCAGTGCAAG</td>
<td>20</td>
<td>58.4</td>
<td>5’RACE</td>
</tr>
<tr>
<td>AAP</td>
<td>GGCACCGGCTCGACTAGTACGGGIGGIGGGIGGGIGGIGG</td>
<td>36</td>
<td>76.5</td>
<td>5’RACE</td>
</tr>
<tr>
<td>AUAP</td>
<td>GGCACCGGCTCGACTAGTAC</td>
<td>20</td>
<td>62.5</td>
<td>5’RACE</td>
</tr>
<tr>
<td>CjCDA 3’-1</td>
<td>ACAACCCCCACACGTGCTC</td>
<td>19</td>
<td>60.4</td>
<td>3’RACE</td>
</tr>
<tr>
<td>CjCDA 3’-2</td>
<td>GCCTGTCCATCAAGTGCC</td>
<td>20</td>
<td>60.4</td>
<td>3’RACE</td>
</tr>
<tr>
<td>3’RACE</td>
<td>CTGTAATGCTGCGACTGAG</td>
<td>22</td>
<td>58.6</td>
<td>3’RACE</td>
</tr>
<tr>
<td>CjCDA Full F</td>
<td>GATCAGACGGGGAACAC</td>
<td>20</td>
<td>58.4</td>
<td>Full length PCR</td>
</tr>
<tr>
<td>CjCDA Full R</td>
<td>CTGGACACACATCGAT</td>
<td>20</td>
<td>52.2</td>
<td>Full length PCR</td>
</tr>
<tr>
<td>β-actin F</td>
<td>ATGTACGTGGCCATCCG</td>
<td>19</td>
<td>58.2</td>
<td>Tissue expression</td>
</tr>
<tr>
<td>β-actin R</td>
<td>CTGGTGCCGAGGTGAT</td>
<td>19</td>
<td>58.2</td>
<td>Tissue expression</td>
</tr>
<tr>
<td>CjCDA F</td>
<td>GCCCAACATTACCCGCTTTCG</td>
<td>22</td>
<td>60.4</td>
<td>Tissue expression</td>
</tr>
<tr>
<td>CjCDA R</td>
<td>ATACCTCTCAGTCAGTGCTG</td>
<td>21</td>
<td>60.4</td>
<td>Tissue expression</td>
</tr>
</tbody>
</table>
K. Fujimori et al.

Figure 1. Schematic representation of the cDNA structure of CjCDA and location of the primers. Arrowheads indicate the primers, and lines between the arrowheads indicate the amplified cDNA fragments.

*C. japonicus* cDNA as the template and the primer pairs *CjCDA* 3′-1, 3′ RACE and *CjCDA* 3′-2, 3′ RACE. The PCR conditions were as follows: 95°C for 2 min, followed by 30 cycles of 95°C for 30 s, 58°C for 30 s, and 72°C for 80 s. For the 5′ RACE, specific primers (*CjCDA* 5′-1 and *CjCDA* 5′-2; Table 1) were designed based on the nucleotide sequences obtained from the RT-PCR. Then, the cDNA fragments encoding the 5′ regions of *CjCDA* were amplified by PCR. In the first PCR, the newly synthesized first-strand cDNA was used as a template, and AAP and *CjCDA* 5′-1 were used as the primers. In the nested PCR, the first PCR products were used as templates and AUAP and *CjCDA* 5′-2 were used as primers. The PCR conditions were as follows: 95°C for 2 min, followed by 30 cycles of 95°C for 30 s, 56°C for 30 s, and 72°C for 20 s. The nucleotide sequences of cDNA fragments containing a full-length open reading frame (ORF) were confirmed by PCR using the specific primers *CjCDA* Full F and *CjCDA* Full R (Table 1) and Platinum Pfx DNA Polymerase (Invitrogen, Carlsbad, CA).

### 2.3. Nucleotide Sequence Analysis

The RT-PCR and 3′ and 5′ RACE amplification products as well as the full-length amplification products were subcloned into the pGEM-T Easy Vector (Promega, Madison, WI), according to the manufacturer’s instructions. The sequences were determined on an ABI PRISM 3130 genetic analyzer (Applied Biosystems, Foster City, CA) using the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems).

### 2.4. Expression Analysis

Total RNA was prepared from the muscle, epidermis, hepatopancreas, gills, shell, tendon, and intersegmental membrane as described in Section 2.2. Then, first-strand cDNA was amplified from the RNA isolated from each tissue as described above. For tissue-specific expression, we designed primers specific to *CjCDA* (*CjCDA* F and *CjCDA* R; Table 1) based on the detected sequences. Then, *CjCDA* was amplified using the first-strand cDNA as template and primers *CjCDA* F and *CjCDA* R (Table 1). The PCR conditions were as follows: 95°C for 2 min, followed by 30 cycles of 95°C for 30 s, 50°C for 30 s, and 72°C
for 30 s. To determine the amount of total RNA in each tissue, we amplified β-actin mRNA using a specific primer pair as a control for normalization (Table 1).

2.5. Phylogenetic Analysis

To determine the relationship between the CDA isolated from the epidermis of C. japonicus and other CDAs from insects, crustaceans, and mollusks, we constructed a phylogenetic tree based on the sequences of enzyme precursors by the neighbor-joining method using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). A fungal CDA (GenBank: XM_007272808.1) was used as an outgroup.

3. Results and Discussion

3.1. cDNA Cloning

From the epidermis of the leg muscle of red snow crab (Chionoecetes japonicas), we obtained a 2141-bp chitin deacetylase (CDA) gene containing a 1575-bp open reading frame (ORF) that encodes 525 amino acids. The sequence determined for the cDNA encoding the chitin deacetylase CjCDA was registered with the DNA Data Bank of Japan (DDBJ) database (accession no. LC342072). A poly(A) tail specific to eukaryotes was found at the 3’-terminus of CjCDA. Comparison of the deduced amino acid sequence of CjCDA (CjCDA) to CDAs in other organisms by BLAST showed the highest similarity (74%) with gastrolith protein 59 of C. quadricarinatus (CqCDA) [22] and the second highest similarity (72%) with chitin deacetylase1 of P. monodon (PmCDA) [23]. In addition, CjCDA showed ~60% similarity was known CDAs in insects. The structural prediction for CjCDA showed an N-terminal signal peptide (amino acids 1 - 18), chitin-binding domain (CBD; amino acids 31 - 90), low-density lipoprotein receptor class A domain (LDL-A; amino acids 103 - 142), and glycoside hydrolase/deacetylase, beta/alpha-barrel (Glyco_hydro/deAcase_b/a-brl; amino acids 145 - 461) (Figure 2). Because amino acids 183 - 293 in the Glyco_hydro/deAcase_b/a-brl are presumed to be the NodB homology domain (NODB_dom) and several characteristic motifs, including TFDD, H[S/T]xxHP, RxP[Y/F], DxxDW, and GxxxFxx [13], are present, CjCDA is considered to be a member of carbohydrate esterases family 4 (CE4). Comparison of CjCDA to the deduced amino acid sequences of CqCDA, PmCDA, and known CDAs in insects showed that CjCDA shares a very similar domain structure with these proteins (Figure 3). The CDAs in insects have been classified into five groups based on their structure (Figure 4) [24]. Groups I and II contain a signal peptide, CBD, LDL-A, and catalytic domain. Groups III and IV contain a signal peptide, CBD, and catalytic domain but no LDL-A. Group V contains only a signal peptide and catalytic domain, and these CDAs are specific to the midgut [24]. Because CjCDA contains a CBD, LDL-A, and catalytic domain, it can be classified into Group I. CqCDA and PmCDA share similar domain structure, and all crustacean
Figure 2. Full-length sequence and deduced amino acid sequence of CjCDA. Signal peptides are indicated by a dotted line. The chitin-binding domain is highlighted in pale gray, and the low-density lipoprotein receptor class A domain is highlighted in dark gray. The glycoside hydrolase/deacetylase, beta/alpha-barrel including the NodB homology domain is underlined with a solid line.

3.2. Tissue Expression

Tissue expression analysis of CjCDA in C. japonicus by RT-PCR using the β-actin gene as a control revealed that CjCDA is expressed only in the epidermis of muscle (Figure 5), and no expression was observed in gills, muscle, or other examined tissues. Previous studies revealed that PmCDA in P. monodon is mainly expressed in the gills and functions in the immune response to the pathogen that causes white spot disease [23] and the CDA in the fruit fly Drosophila...
Comparison of the deduced amino acid sequence of CjCDA to the deduced amino acid sequences of CDAs from crustaceans and insects (CjCDA: Cherax quadricarinatus, ALC79575.1; PmCDA: Penaeus monodon, ALO20448.1; TcCDA: Tribolium castaneum, ABU25223.1; BmCDA: Bombyx mori, ADO24153.1). CE4 domain motifs are enclosed in boxes.

**Figure 3.**

**melanogaster** is involved in the control of trachea extension [21]. In addition, the CDAs of *D. melanogaster* and a species of locust, *Locusta migratoria*, are expressed in the cuticle [21] [25]. Because inhibition of CDA expression in a species of beetle (*Tribolium castaneum*) prevented ecdysis, CDA is thought to play an important role in ecdysis [24]. When arthropods molt, the epidermis separates from the cuticle and a new cuticle is secreted on the surface of the epidermis. Thus, we hypothesize that CjCDA has some function during *C. japonicus* molting.

### 3.3 Phylogenetic Analysis

As a result of phylogenetic analysis based on the deduced amino acid sequences of CDAs from various arthropods, including insects and crustaceans, as well as...
mollusks and the fungus *Colletotrichum gloeosporioides*, which was used as an outgroup, CDAs in arthropods were shown to belong in Groups I-V, as reported previously [24]. However, the CDAs from mollusks were not in Groups I-V. All CDAs from crustaceans, including CjCDA, belonged to Group I (Figure 4).

**Figure 4.** Domain structure of CDAs. Black boxes indicate signal peptides, white boxes indicate chitin-binding domains, dark gray boxes indicate low-density lipoprotein receptor class A domains, and pale gray boxes indicate the catalytic domains of chitin deacetylase-like proteins.

Figure 5. Agarose gel electrophoresis showing the tissue expression pattern of CjCDA.

**Figure 5.** Agarose gel electrophoresis showing the tissue expression pattern of CjCDA.
There are no reports of CDAs from crustaceans belonging to any Group other than Group I. However, because some species have multiple CDA isozymes, like *T. castaneum*, there is a possibility that *C. japonicus* may contain multiple isozymes with different structures, like *T. castaneum*.

### 4. Conclusion

From the epidermis of the red snow crab, we obtained a full-length CDA gene, *CjCDA*, which contained a 1575-bp ORF encoding a 525-amino acid protein. *CjCDA* had a Group I domain structure, as it contained a signal peptide, CBD, LDL-A, and catalytic domain. Because *CjCDA* was expressed only in the epidermis, we hypothesize that it is involved in ecdysis. This study is the first report of the cloning of a full-length CDA gene from a crab.

### References


