Gene 651 (2018) 70–78

Contents lists available at ScienceDirect

Gene

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

Research paper

CpMCA, a novel metacaspase gene from the harmful dinoflagellate *Cochlodinium polykrikoides* and its expression during cell death

Hui Wang[a], Bum Soo Park[b], Weol-Ae Lim[c], Jang-Seu Ki[a,*]

[a] Department of Biotechnology, Sangmyung University, Seoul 03016, South Korea
[b] Institute of Natural Sciences, Sangmyung University, Seoul 03016, South Korea
[c] Ocean Climate and Ecology Research Division, National Institute of Fisheries Science, Busan 46083, South Korea

ARTICLE INFO

Keywords:
Cochlodinium polykrikoides
CpMCA
Algicide
HGT
Transcriptional response
PCD

ABSTRACT

Metacaspases (MCAs) are cysteine proteases that share sequence homology with caspases, and may play roles in programmed cell death (PCD). In the present study, we identified a novel MCA gene (*CpMCA*) from the red tide dinoflagellate *Cochlodinium polykrikoides*, and examined its molecular characteristics and gene expression in response to algicide-induced cell death. *CpMCA* cDNA is 1164 bp in length, containing a dinoflagellate spliced leader sequence (dinoSL), an 879-bp open reading frame (ORF), which codes for a 293-aa protein, and a poly (A) tail. Multi-sequence comparison indicated that CpMCA belongs to type I MCA, but it has a different structure at the N-terminal. Phylogenetic analysis showed that *C. polykrikoides* may have acquired the *MCA* gene from bacteria by means of horizontal gene transfer (HGT). In addition, expressions of *CpMCA* significantly increased following exposure to the common algicides copper sulfate and oxidizing chlorine, which trigger cell death in dinoflagellates, suggesting that *CpMCA* may be involved in cell death.

1. Introduction

MCAs belong to the cysteine-type C14 protease family (Carmona-Gutierrez et al., 2010; Watanabe and Lam, 2011). They share a degree of homology with caspases as they contain caspase-like domains. Until now, several MCAs have been characterized in bacteria, fungi, plant and algae (Piszczek et al., 2011; Ahmad et al., 2012; Wang et al., 2012; Choi and Berges, 2013). Based on their structural features, they are subdivided into type I and type II. Type I MCAs have an N-terminal prodomain that is not present in type II (Piszczek et al., 2012). Both type I and type II MCAs contain a linker that connects the two catalytic caspase-like domains of the protease: the p20 domain (COG 4249, KOG1546 in the NCBI Conserved Domain Database) and the p10 domain. The linker found in type II MCAs, varying in size from 90 to 150 amino acids, is much longer than that found in type I MCAs (~20 residues). The p10 domain is defined by a well conserved motif, whose consensus sequence is SGCXDXQTSADV, which shares 75% matching identity between species, and by other conserved short sequences (Rahman, 2010). Recently, Choi and Berges (2013) reported that a new group of MCAs (designated as Type III MCAs) from heterokontophyta and brown macroalga, which displaying evidence of distinct rearrangements of domain structures (the p10 domain is located in N-terminus). Interestingly, they suggested that the type II and III MCAs might derive from type I MCAs, whereas type I MCAs acquired from bacteria through HGT (Choi and Berges, 2013).

As caspases have a well-studied role in PCD (or apoptosis), a similar function has been hypothesized for MCAs (Tsiatsiani et al., 2011). Previous studies showed that the plant tomato *MCA1* and wheat *MCA4* were involved in PCD induced by the necrotrophic pathogen *Botrytis cinerea* and fungal pathogen *Puccinia striiformis*, respectively (Hoeberichts et al., 2003; Wang et al., 2012). Additional functions of MCAs have been discovered: for example, one MCA (*Yca1*) from yeast has been reported to be involved in cell cycle regulation and protein quality control (Lee et al., 2008; Tsiatsiani et al., 2011). Nevertheless, although considerable progress has been made in recent years to advance our knowledge and understanding of MCAs, detailed information is still lacking, particularly in protists, including microalgae, about their diversity and the mechanisms whereby they are able to function and accomplish their physiologic role.

As microeukaryotes, dinoflagellate algae are a large group of

Abbreviations: BLAST, Basic Local Alignment Search Tool; cDNA, complementary DNA; C_T, cycle threshold; DinoSL, dinoflagellates splice-leader; EST, expression sequence tag; HABs, harmful algal blooms; HSPs, heat shock proteins; KatG, catalase-peroxidase; HGT, horizontal gene transfer; MCA, metacaspase; NJ, neighbor-joining; ORF, open reading frame; PCD, programmed cell death; qRT-PCR, quantitative real time polymerase chain reaction; RACE, rapid amplification of cDNA ends; ROS, reactive oxygen species; SOD, superoxide dismutase; TUA, α-tubulin; UTR, untranslated region

* Corresponding author at: Department of Biotechnology, Sangmyung University, South Korea.
E-mail address: kijs@smu.ac.kr (J.-S. Ki).

https://doi.org/10.1016/j.gene.2018.02.002
Received 24 October 2017; Received in revised form 26 January 2018; Accepted 1 February 2018
Available online 03 February 2018
0378-1119/

unicellular protists existing as free-living, parasitic, or symbiotic forms that can be found in oceans and freshwater ecosystems worldwide (Hackett et al., 2004; Taylor et al., 2008). They possess unique chromosomal, genomic, and gene expression features, including large nuclear genome size, modified DNA bases, lack of nucleosomes, condensed chromosomes, high copy numbers of many genes, and special post-transcriptional regulation characteristics like the dinoSL *trans*-splicing (Hackett et al., 2004; Lin and Zhang, 2010; Brunelle and Van Dolah, 2011; Lin, 2011; Shoguchi et al., 2013; Ponmani et al., 2016). In addition, the dinoflagellate genome has been greatly influenced by lateral gene transfer (LGT) through endosymbiosis (Hackett et al., 2004; Wisecaver and Hackett, 2011). For these reasons, dinoflagellates have been widely used in studies related to gene and genome research, and to the molecular evolution of protists.

Furthermore, many dinoflagellates (e.g., *Alexandrium*, *Cochlodinium*, and *Prorocentrum*) are responsible for HABs, whose outbreaks can cause severe damages to aquaculture and marine ecosystems. High density blooms of *Cochlodinium polykrikoides* occur almost every year in East Asian countries, affecting aquaculture and damaging ecosystems, and the range of the species is expanding globally (Onitsuka et al., 2010; Lee et al., 2013). Therefore, many studies have been performed in *C. polykrikoides* during the last decades, and these works have included environmental surveys, investigating bloom forming mechanisms, and/or mitigation measures (Gobler et al., 2008; Gárate-Lizárraga, 2013; Guo et al., 2016a; Lim et al., 2017).

Dinoflagellate severe blooms can terminate because of a variety of reasons, such as nutrient depletion, viral/bacterial infections, use of biocides, zooplankton grazing and sedimentation. From a cellular perspective, bloom termination can be attributed to dinoflagellate PCD, whether it may be triggered by intrinsic stimuli or by extrinsic signals. Vardi et al. (1999) were the first to propose that the disappearance phase of a dinoflagellate bloom might be related to PCD when the local pH rose and the concentration of dissolved CO_2 became limiting. Since then, a dinoflagellate *MCA* gene was identified from *Prorocentrum donghaiense*, which was proposed to mediate PCD in cellular senescence (Zhang et al., 2006; Yang et al., 2008). The first full sequence of a dinoflagellate *MCA* was determined for the photosynthetic endosymbiont *Symbiodinium microadriaticum* (Aranda et al., 2016), but its function has not yet been studied sufficiently. More recently, Pokrzywinski et al. (2017) reported that a caspase-like protease activity was induced in dinoflagellates (e.g., *Prorocentrum minimum*, *Karlodinium veneficum* and *Gyrodinium instriatum*) when exposed to the bacterial algicide IRI-160AA, which causes cell death. These results showed that PCD may be triggered by a wide range of stressors in dinoflagellates; however, the molecular mechanism through which PCD occurs is still largely unknown, particularly with respect to the role that caspase-like proteins (e.g., MCAs) play in it.

In the present study, we determined the full-length cDNA of the *CpMCA* gene from *Cochlodinium polykrikoides*, and characterized the gene structural features (e.g., conserved motifs in the CpMCA protein and the genomic coding region) and phylogenetic relationship with other MCAs. In addition, we examined the transcriptional response of *CpMCA* in cells exposed to different algicides, non-oxidizing copper sulfate ($CuSO_4$) and oxidizing chlorine (NaOCl), because they are commonly used for controlling HABs and are able to induce cell death (Ebenezer et al., 2014). Finally, we discuss the potential role of *CpMCA* in the termination of dinoflagellate blooms, as it may be involved in the cell death pathway.

2. Materials and methods

2.1. Algae culture

A strain (Cp-01) of *C. polykrikoides* was obtained from the National Fisheries Research and Development Institute (NFRDI), Korea. Cells were cultured in f/2 medium at 20 °C in 12 h light/12 h dark cycles, with a photon flux density of ~65 μmol photons $m^{-2} s^{-1}$.

2.2. RNA extraction, cDNA synthesis, and DNA extraction

To extract RNA from *C. polykrikoides* cells, cultures were collected in centrifuge tubes and centrifuged at 10,000 rpm for 10 min, frozen immediately in liquid nitrogen and stored at −80 °C until RNA extraction. Preserved cells were physically broken by freeze-thawing in liquid nitrogen, and further homogenization was performed using zirconium beads (diameter 0.1 mm) with a Mini-bead beater (BioSpec Products Inc., Bartlesville, OK). Total RNA was isolated using Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions, and further purified on Mini Spin Columns of RNeasy Mini Kit (Qiagen, Valencia, CA). RNA quality and quantity were measured with an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). Reverse transcription (RT) was carried out using a TOPscript™ cDNA Synthesis Kit (Enzynomics). Total genomic DNA was extracted from *C. polykrikoides* following the cetyltrimethylammonium bromide (CTAB) protocol as described by Murray and Thompson (1980).

2.3. Sequence determination of the CpMCA gene

The full length sequence of *CpMCA* cDNA was determined by RACE PCR. Partial sequences of *CpMCA* cDNA were extracted from an EST database of *C. polykrikoides*, and were used for designing the primers to be used for full length amplification (Table 1). The 3′- and 5′-UTRs of the *CpMCA* transcript were determined using the 3′- and 5′-RACE, respectively. For the RACE, nest PCRs were employed, and the primers used in each PCR are listed in Table 1. Reaction conditions for the primary and secondary PCRs were as follows: pre-denaturation at 95 °C for 5 min; 35 cycles of 95 °C for 30 s, 55 °C/58 °C (primary/secondary, respectively) for 30 s, 72 °C for 2 min, and extension at 72 °C for 10 min. Positive core PCR products were purified, cloned into pMD20-T vector (Takara, Shiga, Japan), transformed into *E. coli* competent cells, and subjected to DNA sequencing. The full-length sequence of *CpMCA* cDNA was validated by PCR with specific primers (Table 1). The primers used in *CpMCA* genomic sequence determination were designed according to the cDNA sequence (Table 1).

2.4. CpMCA characterization and phylogenetic analysis

The 3′-end, partial sequences and 5′-end cDNA sequences of *CpMCA* were properly assembled by Sequencher v4.7. Protein motifs and conserved domains of the CpMCA protein were analyzed with online servers and public databases, including PROSITE (http://prosite.expasy.

Table 1
The primers used in the present study.

Gene	Primer	Nucleotide sequence (5′ → 3′)	Remark
CpMCA	5SL	DCCGTAGCCATTTTGGCTCAAG	5′-RACE/full length
CpMCA	CpMCA-5R1	CGGTGCCTGCTGATGTTTG	5′-RACE
CpMCA	CpMCA-5R2	CATCGGCCTCCTCATTGTC	5′-RACE
CpMCA	CpMCA-3F1	TTCGGGACGACGATTTCG	3′-RACE
CpMCA	CpMCA-3F2	CTCTGGTTGCACCGACAAG	3′-RACE/RT-PCR
CpMCA	CpMCA-R1	CTGTTGTACACTTCCCCAAC	RT-PCR
CpTUA	CpTUA-F	TTCTCGCGCATCGACCACAAG	RT-PCR
CpTUA	CpTUA-R	TCCATACCCTCGCCGACATAC	RT-PCR
	B26	GACTCTAGACGACATCGA$(T)_{18}$	3′-RACE
	B25	GACTCTAGACGACATCGA	3′-RACE
CpMCA	CpMCA-SF1	AGGCCACGATGGTCAAGTG	Full length/genomic DNA
CpMCA	CpMCA-SR1	AACACAAGTGTGCTCCGCT	Full length/genomic DNA
CpMCA	CpMCA-SR2	CTGGATCCCTCACTTGGTG	Full length/genomic DNA
CpMCA	CpMCA-IF1	CACGCAGAGCATCACGATC	Intergenic DNA
CpMCA	CpMCA-IR1	TCCAGTGGGTGCTTACCAG	Intergenic DNA