Gene 576 (2016) 618–625

Contents lists available at ScienceDirect

Gene

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

Characterization of a novel gene involved in cadmium accumulation screened from sponge-associated bacterial metagenome

Tetsushi Mori [a,1], Koji Iwamoto [b,1,2], Satoshi Wakaoji [a], Hiroya Araie [b], Yotaro Kohara [a], Yoshiko Okamura [a,3], Yoshihiro Shiraiwa [b], Haruko Takeyama [a,*]

[a] Faculty of Science and Engineering, Waseda University, Tokyo, Japan
[b] Graduate School of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan

ARTICLE INFO

Available online 17 October 2015

Keywords:
Cd accumulation
Sponge-associated bacterial metagenome
Pentapeptide motif

ABSTRACT

Metagenome research has brought much attention for the identification of important and novel genes of industrial and pharmaceutical value. Here, using a metagenome library constructed from bacteria associated with the marine sponge, *Styllisa massa*, a high-throughput screening technique using radioisotope was implemented to screen for cadmium (Cd) binding or accumulation genes. From a total of 3301 randomly selected clones, a clone 247-11C was identified as harboring an open reading frame (ORF) showing Cd accumulation characteristics. The ORF, termed as ORF5, was further analyzed by protein functional studies to reveal the presence of a protein, Cdae-1. Cdae-1, composed of a signal peptide and domain harboring an E(G/A)KCG pentapeptide motif, enhanced Cd accumulation when expressed in *Escherichia coli*. Although showing no direct binding to Cd in vitro, the presence of important amino acid residues related to Cd detoxification suggests that Cdae-1 may possess a different mechanism from known Cd binding proteins such as metallothioneins (MTs) and phytochelatins (PCs). In summary, using the advantage of bacterial metagenomes, our findings in this work suggest the first report on the identification of a unique protein involved in Cd accumulation from bacteria associated with a marine sponge.

1. Introduction

Cadmium (Cd) has been regarded as an important trace element due to its industrial applicability in nickel–cadmium batteries, Cd pigments, Cd coatings and as stabilizers in plastics and alloys (Morrow, 2010). However, long-term exposure to Cd or uptake at high levels has resulted in serious health and ecological problems (Jarup and Akesson, 2009; Boyd, 2010). Currently, the removal of Cd from the environment is conducted using chemical, membrane, ion exchange, solvent extraction and adsorption techniques (Rao et al., 2010). Alternatively, many organisms have adopted resistance mechanisms such as exclusion (Zhu et al., 2011), compartmentalization (Dehn et al., 2004), the formation of complexes (Inouhe et al., 1996) and the synthesis of metal binding proteins (Mejare and Bulow, 2001) to overcome Cd toxicity and heavy metal stress. Such mechanisms have attracted much attention as these systems provide an alternative to conventional Cd removal techniques and can be utilized to further overcome current bioremediation challenges.

Among these resistance mechanisms, the introduction or overexpression of metal-binding proteins have been widely exploited to increase Cd binding capacity, tolerance or accumulation. Two of the most well characterized binding proteins are metallothioneins (MTs) and phytochelatins (PCs). MTs, characterized as low-molecular cytosolic gene-encoded polypeptides, bind to a range of heavy metals including Cd^{2+}, Pb^{2+}, Bi^{3+}, Ag^{2+}, Cu^{2+}, Zn^{2+} and Ni^{2+} while PCs, glutathione polymers enzymatically synthesized by phytochelatin synthases (PCS), are chelators important for heavy metal detoxification (Henkel and Krebs, 2004; Dar et al., 2013; Rigouin et al., 2013). Both MTs and PCs have also been widely reported in various organisms including plants, yeasts, algae and fungi. In bacteria, the identification of Cd-binding proteins including MT and PC homologs has also been reported (Blindauer, 2011; Capasso et al., 1996; Harada et al., 2004; Tsuji et al. 2004). Bacteria also serve as an important expression system for the overexpression of Cd-binding proteins obtained from plants (Kim et al., 2009), yeast (Preveral et al., 2009) or synthetically synthesized peptides (Bae et al., 2002), further suggesting the importance of bacteria in Cd bioremediation. However, since most of the work related to Cd binding proteins has focused on currently identified proteins such as MTs or PCs, there is still

Abbreviations: Cd, cadmium; MTs, metallothioneins; PCs, phytochelatins; Cdae, Cadmium accumulation element.

* Corresponding author at: Faculty of Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku, Tokyo 162-8480, Japan.
E-mail address: haruko-takeyama@waseda.jp (H. Takeyama).

[1] These authors contributed equally to this work.

[2] Present address: Malaysia–Japan International Institute of Technology, University Teknologi Malaysia, Jalan Sultan Yahya Petra, Kuala Lumpur 54100, Malaysia.

[3] Present address: Graduate School of Advanced Sciences of Matter, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8530, Japan.

http://dx.doi.org/10.1016/j.gene.2015.10.018
0378-1119/

a need for the discovery of novel Cd binding proteins or proteins enhancing Cd accumulation.

Thus, to obtain such proteins, we focused our search on microbial metagenomes. In recent years, metagenomic research has supported the identification of novel and important genes from bacterial communities of both terrestrial and marine environments. Marine microbial metagenomes in particular, known for its unique and large genetic diversity, has served as a resource for genes such as lipases (Selvin et al., 2012), esterases (Okamura et al., 2010), fumarases (Jiang et al., 2010), beta-glucosidases (Fang et al., 2010), applicable to pharmaceutics, research and industry (Kennedy et al., 2010; Hentschel et al., 2012). Furthermore, metagenomic based research also provides an advantage in the identification and discovery of genes that may harbor non-elucidated characteristics or undetermined phenotypic properties since screening and activity assays are frequently conducted in bacterial hosts such as *E. coli* and *Bacillus*. However, although metagenomic researches are currently widely conducted, proteins related to Cd binding or accumulation have not been reported thus far.

Therefore, in this research, to conduct a comprehensive search for genes related to Cd binding or accumulation, we conducted the screening of such genes from the metagenome library of bacteria associated with the marine sponge *Styllisa massa*. We focused on bacteria associated with marine sponges, since marine sponges are known to be one of the largest producers of secondary metabolites (Thomas et al., 2010) and holds high potential of harboring unique functional genes including those related to heavy metal accumulation (Selvin et al., 2009; Nelson and Slinger-Cohen, 2014). Subsequently, functional analysis and preliminary sequence comparison studies to determine the novelty of the discovered protein were conducted. Here, we report one of the first reports on the identification of a unique protein involved in the accumulation of Cd from a sponge-associated bacterial metagenome.

2. Materials and methods

2.1. Sponge, bacterial strains and plasmids

The marine sponge, *S. massa*, was collected from the offshore of Ishigaki island, Okinawa, Japan. The *E. coli* strains, EPI300™ (Epicentre Biotechnologies) were used in metagenome library construction, DH5α (TOYOBO) and EC100 ™ (Epicentre Biotechnologies) in cloning and BL21 (DE3) (Novagen) in recombinant protein expression, respectively. The plasmid pCC1FOS (Epicentre Biotechnologies) was used for metagenome library construction and in standard cloning procedures. The Zero Blunt TOPO PCR Cloning Kit for Sequencing (Life Technologies) was used for the cloning of PCR amplicons, and the pET25b (Novagen) vector was used for protein expression.

2.2. Library construction

The preparation of marine sponge bacterial fraction and DNA extraction were performed as described by Okamura et al. (2010). Fosmid library construction was conducted using the CopyControl Fosmid Library Production Kit (Epicentre Biotechnologies) based on the manufacturer's protocol. Briefly, blunt-ended and 5′-phosphorylated DNA was separated by pulsed-field gel electrophoresis (1% LMP agarose/1× TBE gel, 0.5× TBE buffer, 0.5 s pulse, 9 V/cm, 14 °C, 120 °, 3 h) and approximately 40 kbp of DNA was recovered with GELase. The obtained DNA fragments were ligated with the pCC1FOS vector, packaged, titered and were infected into *E. coli* EPI300™ cells. Upon plating on LB agar plates containing 12.5 μg/mL chloramphenicol, colonies were selected manually or by using the BioPick automated colony picking system (Genomic Solutions). The metagenome library was stored at −80 °C in a 96-well plate format. Fosmid DNA were extracted from randomly selected clones following standard alkaline lysis procedure and digested by *Bam*HI and *Eco*RI/*Hin*dIII to estimate the average size of the DNA inserts.

2.3. Screening for Cd accumulation clones from metagenome library

Screening of Cd accumulation from the metagenome library clones was conducted using the microplate-BAS method. Prior to the screening of Cd accumulation genes, 40 randomly selected 96-well plates; comprised of 3301 metagenome library clones were cultured in LB medium containing 12.5 μg/mL of chloramphenicol overnight at 37 °C with agitation. The overnight cultures were diluted with 4 folds of culture medium described above with addition of 2.5 μM Cd including 37 kBq/mL of radioactive Cd (^{109}Cd, PerkinElmer Life and Analytical Science) and induction solution (Epicentre Biotechnologies) and were cultured at 37 °C for a subsequent 5 h. After 5 h, 50 μL of the bacterial culture was transferred to Multiscreen-GV filter plates (0.22 μm) and washed 3 times with 200 μL of 3% NaCl using a MultiScreen HTS Vacuum Manifold system (EMD Millipore). Upon drying, the plates were placed on imaging plates, overnight in the dark and the accumulation of Cd within the clones was detected using the Bio-imaging Analyzer BAS-1800 II (FUJIFILM). Positive clones showing high Cd accumulation were selected upon 2 rounds of screening. The selection criteria for Cd accumulation clones were determined based on the accumulation of Cd above the total average detection signal value. As such, the cut off point for positive clones was set to 2 folds and 3 folds for the 1st and 2nd round screening, respectively.

The clones obtained from the 2nd screen were reanalyzed to determine Cd accumulation using the silicone oil centrifugation method. 0.4 mL sampling tubes were prepared containing a dense bottom layer, composed of 50 μL silicone oil (Toray Dow Corning Silicone; SH550:SH556 = 2:1). 200 μL of cells incubated with 2.5 μM Cd including ^{109}Cd radioisotope were pipetted over the silicone oil layer, centrifuged at 10,000 *g* for 1 min and the sample tubes were frozen in liquid nitrogen. The bottom layer with the cells was clipped into measurement tubes and the radioactivity was measured using the COBRAII γ-counter (Packard Instrument). For the plasmid reintroduction assay, 200 mL of positive cultures were grown in 2-YT medium containing 12.5 μg/mL chloramphenicol and induction solution, autoinduced and cultured overnight. Plasmids were extracted using the Qiagen Plasmid Buffer Set and the Qiagen-tip 100 (Qiagen) based on the manufacturer's protocol. The extracted plasmid was subsequently transformed by electroporation using the Gene Pulser II (Biorad) into *E. coli* EC100™ electrocompetent cells. The clone showing the highest Cd accumulation was cloned and sequenced to determine the regions harboring the Cd accumulation gene.

2.4. Identification of the Cd accumulation gene

Based on the sequencing results, plasmid from the Cd accumulating clone was enzyme digested with the restriction enzyme *Xho*I to determine the region in which Cd accumulation occurs. The enzyme digested

Table 1
Primer sequences used in the amplification of the target fragments for the functional analysis of ORF5.

Target region		Primer sequences
ORF5	Forward	ATGCAGGTTCGTGCGCGGACACG
	Reverse	TTAGTGCCCGCACTTGCCCTC
ORF5 (T1)	Forward	ATGCAGGTTCGTGCGCGGACA
	Reverse	CGCCCTTATGGGCATCGGCGA
ORF5 (T2)	Forward	ATGCAGGTTCGTGCGCGGACA
	Reverse	GGCGGTCACGGAAACGGCGAA
Cdae-1	Forward	AAAAA**CATATG**AGCGAAGAGAAGAAATCC
	Reverse	AAAAA**CTCGAG**GCTGCCGCGCGGCACCAGG TGCCCGCACTTGCCCTCGC
Cdae-1R3	Forward	AGC**CATATG**GTCGAGTATGGAGGCGGC
	Reverse	**GGATCC**TTAGTGCCCGCACTTGCCCTC

Bold sequences show the restriction enzyme sites. Underlined sequence shows the thrombin cleavage site.