Characteristics of hydrocarbon hydroxylase genes in a thermophilic aerobic biological system treating oily produced wastewater

Ruyin Liu, Yingxin Gao, Yifeng Ji, Yu Zhang and Min Yang

ABSTRACT

Alkane and aromatic hydroxylase genes in a full-scale aerobic system treating oily produced wastewater under thermophilic condition (45–50°C) in the Jidong oilfield, China, were investigated using clone library and quantitative polymerase chain reaction methods. Rather than the normally encountered integral-membrane non-haem iron monooxygenase (alkB) genes, only CYP153-type P450 hydroxylase genes were detected for the alkane activation, indicating that the terminal oxidation of alkanes might be mainly mediated by the CYP153-type alkane hydroxylases in the thermophilic aerobic process. Most of the obtained CYP153 gene clones showed distant homology with the reference sequences, which might represent novel alkane hydroxylases. For the aromatic activation, the polycyclic aromatic hydrocarbon-ring hydroxylating dioxygenase (PAH-RHD) gene was derived from Gram-negative PAH-degraders belonging to the Burkholderiales order, with a 0.72% relative abundance of PAH-RHD gene to 16S rRNA gene. This was consistent with the result of 16S rRNA gene analysis, indicating that Burkholderiales bacteria might play a key role in the full-scale process of thermophilic hydrocarbon degradation.

Key words | hydrocarbon hydroxylase, oily wastewater, thermophilic treatment

INTRODUCTION

Since hydrocarbon pollution can be observed universally due to the wide range of petroleum exploitation and utilization, biological degradation of hydrocarbons has been studied extensively. An oxidation pathway that involves a hydroxylase as the key enzyme has been found to be mainly responsible for the initial step of petroleum hydrocarbon metabolism. For the initial activation of alkanes, the integral-membrane non-haem iron monooxygenase (AIIkB) is particularly well characterized (van Beilen & Funhoff 2007). However, in some cases, it is difficult to associate the presence of alkB gene with the capacity of aliphatic hydrocarbon removal (Paise et al. 2011), and even to detect this gene in oil-contaminated sites (Palmroth et al. 2007). An alternative pathway initiated by the hydroxylase belonging to the CYP153 family of cytochrome P450 monooxygenase has been identified in some alkane-degraders lacking alkB, and relevant genes have been found in various environments (van Beilen et al. 2006; Wang et al. 2010). For the aerobic degradation of polycyclic aromatic hydrocarbons (PAHs), the initial step of PAH degradation is normally carried out by multicomponent aromatic-ring hydroxylating dioxygenases (RHD), various allele types of which have been reported in the Gram-negative (GN) and Gram-positive (GP) PAH-degraders (Habe & Omori 2003). Until now, most of the studies focus on the petroleum degradation in various oil-contaminated natural environments. In contrast, studies regarding hydrocarbon degradation in biological wastewater treatment processes are very limited, where continuous and efficient hydrocarbon-degrading microbial communities always have been established. The lack of knowledge of the full-scale microbial degradation process hampers the design of more robust biological approaches.

Wastewater generated during the production of oil in oilfields is commonly characterized by a high content of petroleum hydrocarbons, primarily composed of saturated aliphatic hydrocarbons (alkanes) and aromatic hydrocarbons (including PAHs), and elevated temperatures (Van Hamme et al. 2003). Biological processes are normally adopted for the treatment of the oil-production wastewater. In the Jidong oilfield of China, a full-scale biological system

doi: 10.2166/wst.2014.470
consisting of successive anaerobic and aerobic biofilm reactors is applied to treat produced water under thermophilic conditions (45–50 °C). In previous studies, we have revealed the microbial community associated with methanogenic hydrocarbon degradation in the anaerobic reactor (Liu et al. 2010), and identified some PAH-related catabolic genes predominant in the aerobic reactor (Wang et al. 2007). Nonetheless, we are still far from understanding the diversities and abundances of the dominant functional genes involved in the activation of alkanes and PAHs in this engineering system.

For biological wastewater treatment, continuous and steady removal of contaminants relies on the activity of microbial communities. The microbial degradation process is accelerated by enriching selected functional microorganisms. Therefore, a better understanding of microbial functional structure should be helpful to improve the treatment performance and operational stability. In this study, to understand the process of thermophilic aerobic microbial hydrocarbon degradation, culture-independent molecular approaches, including cloning and sequencing, and quantitative polymerase chain reaction (PCR) were used to characterize the hydrocarbon hydroxylase and 16S rRNA genes in the aerobic unit of a full-scale thermophilic biological system treating oil-production water in the Jidong oilfield.

MATERIALS AND METHODS

Sample collection

The biological treatment system was constructed and put into use in 2001 with a capacity of 5,000 tons of wastewater per day. The system consists of two identical series-wound tanks filled with fiber-type biocarriers, with the second tank being aerated. The hydraulic retention time (HRT) of each reactor is about 12 h, and the volumetric chemical oxygen demand (COD) load of the system is approximately 0.5 g L⁻¹ day⁻¹. The biofilm samples from the aerobic tank were taken for DNA extraction. The biomass on three biocarriers from the aerobic unit was shaved and then mixed as the biofilm sample.

Water quality analysis

Water samples of the influent and effluents of each tank were taken every 20 days over a period of 2 months. All of the samples were collected in sterile bottles and processed in the laboratory within 24 h after sampling (Reddy & Quinn 1999). The concentrations of PAHs in the extracts were determined by an Agilent 6890GC equipped with a 5973 mass selective detector using an HP-5 silica column under the selected ion monitoring mode (Chen et al. 2005). The recoveries for 16 priority PAHs were in the range of 70–96%. Conventional water quality parameters were measured according to Standard Methods (American Public Health Association et al. 1998).

DNA extraction

The biofilm samples were washed three times using phosphate-buffered saline (PBS, pH 7.0) and centrifuged at 4 °C, 10,400 g for 15 min. The genomic DNA was then extracted as described before (Liu et al. 2010). Genomic DNA was extracted from triplicate biofilm samples and then pooled.

PCR and cloning

The primers and PCR conditions are summarized in Table 1. PCR mixtures (50 μL) contained 1 μL of DNA template, 5 μL of PCR buffer, 200 μmol L⁻¹ of dNTP, 0.2 μmol L⁻¹ of each primer, 0.5 mg L⁻¹ bovine serum albumin (BSA), and 0.25U Taq polymerase (Takara, Dalian, China). Three separate reactions for each gene were run to minimize PCR bias in subsequent cloning steps, and all relevant PCR products from the same sample were further pooled together. Cloning was performed as described by Liu et al. (2010).

Phylogenetic analysis

Sequences were searched against the GenBank database by blastn for 16S rRNA genes and blastx for functional genes to determine the nearest matches. The presence of chimerae was determined by using the Bellerophon prior to phylogenetic analysis (Huber et al. 2004). The 16S rRNA gene sequences sharing 97% or greater similarity were considered to represent the same operational taxonomic unit (OTU) using MOTHUR software. Rarefaction curves of the functional gene clone libraries were produced using MOTHUR with the observed OTUs sharing 99% or greater similarity. Phylogenetic trees were constructed using MEGA by the neighbor-joining algorithm with bootstrap analyses.

Quantitative PCR assay

Real-time PCR analysis was performed using the primer sets Eub341F and Eub533R, PAH-RHD GN F, and PAH-RHD GN R, as well as P450fw1 and P450rv3. The 25 μL reactions typically contained 1× Sybr Green I, 1× dye (TaKaRa,
Dalian, China), 200 nmol L\(^{-1}\) each primer, 0.5 mg L\(^{-1}\) BSA, and 2 μL properly diluted DNA templates. Real-time PCR was run using an ABI7300 apparatus (ABI, USA) by the following program: 95°C for 30 s, 40 cycles consisting of (i) 95°C for 10 s, (ii) annealing temperature for 15 s, (iii) 72°C for 15 s, and (iv) 78°C for 26 s to collect the fluorescent signals, and finally the melting process automatically generated by the ABI7300 software. The standard plasmids carrying target genes were obtained by TA cloning and extracted using TIANpure Mini Plasmid kit (Tiangen, China). Triplicates of serially diluted standard plasmids were performed to obtain the standard curves. The amplification efficiencies of the above three primer sets were 98%, 97%, and 90%, respectively.

**Nucleotide sequence accession numbers**

The clone sequences that were determined have been deposited in the GenBank database under accession nos. KF548201–KF548304.

**RESULTS AND DISCUSSION**

**Performance of the aerobic treatment system**

As shown in Table 2, the process performance was excellent with 93% removal efficiency for total petroleum hydrocarbons (TPH) and 91% for total PAHs at 12-h HRT. However, no change in recalcitrant high-molecular-mass (HMM) PAHs occurred during the treatment. Although the anaerobic treatment removed petroleum hydrocarbons to some extent, the aerobic treatment contributed to 68% and 55% of the removals of TPH and total PAHs, respectively, showing that the aerobic process was a key wastewater treatment.

**Microbial community structure of the aerobic biofilm**

A 16S rRNA gene library (AB) was constructed to reveal the bacterial community structure of the biofilm in the aerobic

---

### Table 1 | Characteristics of PCR primer sets used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Target gene</th>
<th>Sequence 5′ → 3′</th>
<th>Annealing temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eub341F</td>
<td>16S rRNA</td>
<td>CCTACGGGAGGCAGGCAG</td>
<td>55</td>
</tr>
<tr>
<td>Eub518R</td>
<td></td>
<td>ATTACCGCGGCTGCTGG</td>
<td></td>
</tr>
<tr>
<td>27F</td>
<td>16S rRNA</td>
<td>AGAGTTTGATCCTGGCTCAG</td>
<td>55</td>
</tr>
<tr>
<td>1492R</td>
<td></td>
<td>GGTTACCCTGTACGACTT</td>
<td></td>
</tr>
<tr>
<td>PAH-RHD GP F</td>
<td>Gram-positive PAH-RHDα</td>
<td>CCGCGCGCAGAAATTTGNGNG</td>
<td>54</td>
</tr>
<tr>
<td>PAH-RHD GP R</td>
<td></td>
<td>GGGAAACCGTGCCCTGDAARAA</td>
<td></td>
</tr>
<tr>
<td>PAH-RHD GN F</td>
<td>Gram-negative PAH-RHDα</td>
<td>GAGATGCATACACGKGGTTGA</td>
<td>57</td>
</tr>
<tr>
<td>PAH-RHD GN R</td>
<td></td>
<td>AGCTGTTGTGCGGAGYAYWGTCMTT</td>
<td></td>
</tr>
<tr>
<td>alkB-1 F</td>
<td>alkB</td>
<td>AAYACNGCNAGYARCTNYYCAAYA</td>
<td>55</td>
</tr>
<tr>
<td>alkB-1 R</td>
<td></td>
<td>GCRNTGRTGRCAGTNGCGYTG</td>
<td></td>
</tr>
<tr>
<td>alkBFd</td>
<td>alkB</td>
<td>AACTACMTCGARCTYACGG</td>
<td>57</td>
</tr>
<tr>
<td>alkBRd</td>
<td></td>
<td>TGAMATGTGTYRCNTYTCCT</td>
<td></td>
</tr>
<tr>
<td>P450fw1</td>
<td>CYP153</td>
<td>GCAGCGCAGAAGCAGGAYRA</td>
<td>58</td>
</tr>
<tr>
<td>P450rv3</td>
<td></td>
<td>GCAGCGGCGGAGCGAGGCAG</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2 | Characteristics of the produced water during the treatment

<table>
<thead>
<tr>
<th>Water quality parameter</th>
<th>Influent</th>
<th>Anaero. eff.</th>
<th>Aero. eff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>7.2 ± 0.1</td>
<td>7.3 ± 0.1</td>
<td>7.1 ± 0.2</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>50 ± 1</td>
<td>50 ± 2</td>
<td>50 ± 2</td>
</tr>
<tr>
<td>COD (mg L(^{-1}))</td>
<td>270 ± 25</td>
<td>187 ± 20</td>
<td>50 ± 15</td>
</tr>
<tr>
<td>TPH (mg L(^{-1}))</td>
<td>28 ± 5</td>
<td>21 ± 3</td>
<td>2.0 ± 3</td>
</tr>
<tr>
<td>TDS(^{a}) (mg L(^{-1}))</td>
<td>1,400 ± 35</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SO(_{4})(^{2-}) (mg L(^{-1}))</td>
<td>25.5 ± 0.8</td>
<td>4.7 ± 1.3</td>
<td>10.3 ± 1.0</td>
</tr>
<tr>
<td>Total PAHs(^{d}) (μg L(^{-1}))</td>
<td>960 ± 19</td>
<td>620 ± 15</td>
<td>90 ± 10</td>
</tr>
<tr>
<td>HMM PAHs (ng L(^{-1}))</td>
<td>115 ± 20</td>
<td>123 ± 22</td>
<td>120 ± 18</td>
</tr>
</tbody>
</table>

\(^{a}\)Anaero. eff., effluent from anaerobic treatment.

\(^{b}\)Aero. eff., effluent from aerobic treatment.

\(^{c}\)TDS, total dissolved solid.

\(^{d}\)Total PAHs, the 16 PAHs included in the USEPA’s priority pollutant list and parts of alkyl-substituted PAHs.
unit. After removing the chimeric sequences, 44 clones were obtained and classified into 33 OTUs. Phylogenetic analysis showed that Betaproteobacteria was the most predominant group in the biofilm library, comprising 34.1% of clones (Figure 1). The majority of Betaproteobacteria clones belonged to the orders Burkholderiales and Rhodocyclales (25% and 9.1% of clones, respectively). Within the family Comamonadaceae of Burkholderiales, one OTU (AB100) was related to Brachymonas petroleovorans strain CHX, a bacterium isolated from the wastewater plant of a petroleum refinery, with the ability of degrading a range of light hydrocarbons (C5–C10) and aromatic compounds (Rouviere & Chen 2003), and two OTUs (AB22 and AB63) were affiliated with Diaphorobacter sp. KOTLB, which could utilize pyrene as sole sources of carbon and energy (Klankeo et al. 2009). Members of the Diaphorobacter genus were also found to predominate in a high-temperature (70°C) bioreactor for treating carbazole-containing wastewater (Tan & Ji 2010). Two OTUs (AB106 and AB14) that fell into Rhodocyclales were associated with anaerobic hydrocarbon-degraders. AB106 was affiliated with Azoarcus spp., members of which are known to anaerobically degrade toluene and ethylbenzene, as well as alkanes. AB14 was related to Georgfuchsia toluolica, which can degrade aromatic compounds with Fe(III), Mn(IV), or nitrate as an electron acceptor (Weelink et al. 2009). Aerobic and anaerobic hydrocarbon-degraders detected in the aerobic biofilm reactor were distributed into the orders Burkholderiales and Rhodocyclales, respectively, suggesting that the Betaproteobacteria play an important role during thermophilic petroleum degradation in the full-scale aerobic system.

### Hydrocarbon hydroxylase gene profiles

As the most intensively studied alkane hydroxylase, the integral-membrane non-haem iron monooxygenase gene (alkB) has been found in various petroleum-contaminated environments. Two sets of primers (alkB-1f/alkB-1r and alkBFd/alkBRd) successfully used to evaluate alkB gene diversity in environmental samples (Kloos et al. 2006; Powell et al. 2006) were used to amplify alkB genes in this study. Unexpectedly, neither of the primer sets gave a positive amplification signal. Nevertheless, the alkane hydroxylases belonging to the cytochrome P450 CYP153 family were successfully amplified using the P450fw1/P450rv3 degenerate primer pair (Table 1). The PCR products were cloned, and 21 positive colonies were randomly selected and sequenced. Of those, 19 clones were affiliated with P450 alkane hydroxylase gene sequences in GenBank. Despite the small number of CYP153 gene clones, the rarefaction curve was close to saturation, and the calculated homologous coverage was 89.5% (Figure S1, available online at http://www.iwaponline.com/wst/071/470.pdf).

As shown in the phylogenetic tree based on the alignment of deduced amino acid sequences (Figure 2), the diversity of P450 hydroxylase genes was rather low, and the obtained CYP153 gene relatives were concentrated into two clusters. One cluster comprising 13 similar sequences was separated from Parvibaculum sp. S13-5.
SCP39669; the other consisted of four sequences that showed homology with the CYP153 hydroxylase amino acid sequence of *Burkholderia xenovorans* LB400. The remaining two clones were affiliated with CYP153 sequences of *Sphingopyxis* and *Erythrobacter* species. Notably, most members of the two clusters showed a high similarity with each other but were distantly related to the reference sequences (Figure 2), suggesting that these clusters represent novel P450 genes present in the thermophilic aerobic reactor. Similarly, Wang *et al.* (2010) have found that hydroxylase genes (P450 and *alkB*) identified in the Sargasso Sea environmental DNA sequences also tend to cluster together, but are distinct from previous reported sequences. These results indicated that alkane-degrading bacteria possibly exhibit biogeographic distribution on the hydroxylase gene levels. However, possibly due to horizontal gene transfer, phylogenetic relationships of CYP153 genes are rather complex (Wang *et al.* 2010). For example, the CYP153 amino acid sequence of *Parvibaculum* sp. S13-5 was distantly related to that of *Parvibaculum* sp. S18-4 (Figure 2). Therefore, it is difficult to determine the phylogenetic positions of bacteria on the basis of the obtained P450 gene sequences.

CYP153 genes have been detected in many oil-degraders (Wang *et al.* 2010), and those retrieved from oil-polluted environments have been proved to encode alkane hydroxylases (Kubota *et al.* 2005). However, knowledge about their occurrence and genetic diversity is very scarce. In this study, interestingly, only CYP153-type P450 alkane hydroxylase genes, but no *alkB* gene, were detected in the aerobic unit of the full-scale thermophilic oily water treatment plant. It is possible that the copy number of *alkB* present in the biomass is below detection limit, or variations in *alkB* nucleotide sequences are responsible for the failure amplification (Jurelevicius *et al.* 2013). An alternative pathway initiated by CYP153 hydroxylases is common in alkane-degraders lacking AlkB (van Beilen *et al.* 2006). Also, the elevated temperature might have led to the dominant occurrence of the potential novel

![Figure 2](https://iwaponline.com/wst/article-pdf/71/1/75/470150/wst071010075.pdf)
hydroxylase genes, although further studies are required to evaluate the influence of temperature on alkane hydroxylases. Using the P450fw1/P450rv3 primer pair, the CYP153 genes were quantified. The ratio of CYP153 gene copy number to 16S rRNA gene copy number was 4.6% in the biofilm, remarkably higher than that (1.0%) in a laboratory system inoculated with the same aerobic biofilm biomass using glucose as the carbon source (Table S1, available online at http://www.iwaponline.com/wst/071/470.pdf). Thus, activation of alkanes in the aerobic unit may be mainly mediated by the CYP153 hydroxylases. However, expression of these CYP153 genes detected here is needed to verify their potential identity of alkane hydroxylases in future studies.

PAH-RHD genes were examined according to the method described by Cébron et al. (2008). RHD-GPf/RHD-GPr and RHD-GNf/RHD-GNr primer sets were used for qPCR and clone library analysis for GP and GN PAH-degraders, respectively. Only the GN primer set gave

---

**Figure 3** | Phylogenetic relationships among deduced amino acid sequences of PAH-RHD gene clones retrieved from the aerobic unit. Evolutionary dendrogram constructed using the neighbor-joining method. Bootstrap values (>50%) are indicated at branch points. The scale bar represents 0.5% estimated sequence divergence.
positive signals. This was in accordance with the results from the bacterial clone library that no GP bacterium related clone was found in the aerobic unit. The ratios of PAH-RHD gene copy number to 16S rRNA gene copy number were used as indicators of the PAH-degradation potential of the bacterial community present in the environmental samples (Cébron et al. 2008). In this study, the ratio was 0.72% in the biofilm sample retrieved from the aerobic unit treating oil-produced water, which was in the high end of the range of 0.001–0.922% observed in the highly PAH-contaminated soil and sediment samples (Cébron et al. 2008).

In the Jidong oilfield, low-molecular-mass (LMM) PAHs could be removed efficiently by the biological wastewater treatment process, whereas the operating conditions did not favor biodegradation of HMM PAHs. One major reason may be that the short HRT (12 h) was not sufficient for the biodegradation of HMM PAHs, which are characterized by low bioavailability. Leys et al. (2005) have hypothesized that, following contamination by PAHs, the initial degradation could be mainly performed by GN r-strategists, whereas K-strategist GP bacteria could outcompete them for the degradation of more persistent HMM PAHs. This hypothesis is supported by the evidence that the gene ratio of GN to GP PAH-RHDs dramatically decreased over time in many PAH-contaminated sites (Cébron et al. 2008; Yergeau et al. 2009). Hence, it can be reasonably expected that in the biological system the selection pressures exerted by the short HRT favor the proliferation of r-selected GN PAH-degraders.

A total of 41 partial PAH-RHD gene clones were further obtained and sequenced to characterize diversity of PAH-RHD gene. All sequences were related (95–98%) to the PAH-dioxygenase genes of Comamonas, Burkholderia, and Ralstonia species in the order Burkholderiales (Figure 3) retrieved from natural environments, suggesting that microorganisms found in nature and harboring these ring-degrading genes possibly have adjusted to the thermophilic habitats in this engineering system. Given abundant Burkholderiales bacteria in the aerobic unit based on the phylogenetic analysis of 16S rRNA genes, members of Burkholderiales should play a key role in the thermophilic aerobic degradation of LMM PAHs.

The major purpose of this study is to characterize the dominant hydrocarbon-degradation genes and microbial populations in the thermophilic oil-production wastewater treatment plant, and thus the clone library method was used in this study. Despite the relatively small number of functional gene clones, the rarefaction curves tend to saturate (Figure S1), suggesting low diversities of hydrocarbon-degradation genes in the aerobic system. However, future work is needed to get a comprehensive outlook for the molecular diversities in the system using a high-throughput sequencing method.

CONCLUSIONS

In this study, we demonstrated the microbial hydrocarbon-degradation process in a full-scale oily wastewater treatment plant by analyzing the degradation genes. The results indicate that activation of alkanes during the thermophilic process might be mainly mediated by the novel CYP153 hydroxylases, and members of Burkholderiales should be the key players responsible for the thermophilic aerobic degradation of LMM PAHs. The work may give some implications for oily wastewater treatment: (i) hydrocarbon-degrading bacteria in the order Burkholderiales, especially those with CYP153-type P450 hydroxylase genes, might be used to construct bioaugmented systems for high-temperature petroleum wastewater treatment and (ii) the short HRT (∼12 h) operation is effective to remove LMM PAHs, but insufficient for HMM PAHs.

ACKNOWLEDGEMENT

This work was supported by the National High Technology Research and Development (863) Program of China (No. 2012AA063401).

REFERENCES


First received 30 April 2014; accepted in revised form 10 November 2014. Available online 22 November 2014.