Functional and transcriptional analyses of the initial oxygenase genes for acenaphthene degradation from *Sphingomonas* sp. strain A4

Atsushi Kouzuma,¹ Onruthai Pinyakong,² Hideaki Nojiri,¹ Toshio Omori,³ Hisakazu Yamane¹ and Hiroshi Habe^{1†}

Correspondence

Hiroshi Habe

hiroshi.habe@aist.go.jp

¹Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

²Department of Microbiology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

³Department of Industrial Chemistry, Faculty of Engineering, Shibaura Institute of Technology, 3-9-14 Shibaura, Minato-ku, Tokyo 108-8548, Japan

Sphingomonas sp. strain A4 is capable of utilizing acenaphthene as its sole carbon and energy source. To isolate the genes responsible for acenaphthene degradation, transposon mutagenesis was performed on strain A4 and four mini-Tn5-inserted mutants lacking the ability to utilize acenaphthene were isolated. In three of the four mini-Tn5 inserted mutants, the mini-Tn5s were inserted into the same locus (within about 16 kb) as the *arhA1A2* genes, which had previously been identified as the genes encoding the terminal oxygenase components for the initial oxygenation of acenaphthene. The nucleotide sequence analysis of the corresponding 16.4 kb DNA fragment revealed the existence of 16 ORFs and a partial ORF. From these ORFs, the genes encoding the ferredoxin (ArhA3) and ferredoxin reductase (ArhA4) complementary to ArhA1A2 were identified. RT-PCR analysis suggested that a 13.5 kb gene cluster, consisting of 13 ORFs and including all the *arhA* genes, forms an operon, although it includes several ORFs that are apparently unnecessary for acenaphthene degradation. Furthermore, using gene disruption and quantitative RT-PCR analyses, the LysR-type activator, ArhR, required for expression of the 13.5 kb gene cluster was also identified. Transcription of the gene cluster by ArhR was induced in the presence of acenaphthene (or its metabolite), and a putative binding site (T-N₁₁-A motif) for ArhR was found upstream from the transcription start point of *arhA3*.

Received 6 January 2006

Revised 4 April 2006

Accepted 13 April 2006

INTRODUCTION

The genus *Sphingomonas* consists of strictly aerobic, chemoheterotrophic, yellow-pigmented, Gram-negative, rod-shaped bacteria that contain glycosphingolipids as cell envelope components and belong to the α -4 subgroup of the Proteobacteria (Takeuchi *et al.*, 2001). Sphingomonads are ubiquitous in the environment, occurring in soil, water and sediments (White *et al.*, 1996). They are attracting interest as

a result of their high potential for bioremediation. In recent years, many *Sphingomonas* strains, which have been isolated from a variety of contaminated environments, have been described with respect to their ability to degrade xenobiotics and aromatic compounds, such as biphenyl (Fredrickson *et al.*, 1995), dibenzo-*p*-dioxin (Wittich *et al.*, 1992), carbazole (Habe *et al.*, 2002), tetralin (Hernández *et al.*, 1999), pentachlorophenol (Saber & Crawford, 1985), 4-nonylphenol (Gabriel *et al.*, 2005), γ -hexachlorocyclohexane (Imai *et al.*, 1989) and herbicides (Adkins, 1999; Kohler, 1999). In addition, *Sphingomonas* strains are often isolated from contaminated soils because of their ability to degrade polycyclic aromatic hydrocarbons (PAHs), which are ubiquitous environmental pollutants with toxic, mutagenic and carcinogenic properties (Bastiaens *et al.*, 2000; Khan *et al.*, 1996; Mueller *et al.*, 1990; Pinyakong *et al.*, 2000).

Recent genetic analyses of the aromatic degradation pathways in PAH-degrading sphingomonads have increasingly

[†]Present address: National Institute of Advanced Industrial Science and Technology (AIST), Central 5-2, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan.

Abbreviations: 1,8-NA, 1,8-naphthalic anhydride; 1,8-NDCA, 1,8-naphthalenedicarboxylic acid; CFMM, carbon-free mineral medium; LTTR, LysR-type transcriptional regulator; MSTFA, *N*-methyl-*N*-trimethylsilyltrifluoroacetamide; PAH, polycyclic aromatic hydrocarbon.

The GenBank/EMBL/DDBJ accession number for the sequence of the *arh* genes and their flanking regions reported in this paper is AB240454.

revealed that most of the genes necessary for degrading an aromatic compound are scattered in several clusters and not organized in coordinately regulated operons (recently reviewed by Pinyakong *et al.*, 2003a). This raises the question of how these complexly arranged genes are efficiently regulated and transcribed. The system regulating PAH-degradative genes in sphingomonads has not been investigated fully, probably owing to their complexity. However, it was recently reported that the LysR-type transcriptional regulators (LTTRs) ThnR, PcpR and LinR, which activate the expression of target degradative genes, were isolated from *Sphingomonas* strains that degraded tetralin, pentachlorophenol and γ -hexachlorocyclohexane, respectively (Martínez-Pérez *et al.*, 2004; Cai & Xun, 2002; Miyauchi *et al.*, 2002).

To investigate the diverse and complex aromatic-degradation systems of sphingomonads, including their regulatory mechanisms, it is necessary first to isolate the gene(s) responsible for each degradation step and then to examine the regulatory mechanisms related to their functions. For this purpose, we investigated acenaphthene degradation by *Sphingomonas* sp. strain A4 at the molecular level (Pinyakong *et al.*, 2004). Strain A4 was isolated based on its ability to utilize acenaphthene as a sole source of carbon and energy (Komatsu *et al.*, 1993). Acenaphthene, a PAH that possesses a single alicyclic five-membered ring sharing three carbon atoms with two aromatic rings, is an abundant constituent of coal tar and creosote (Wise *et al.*, 1988). Commercial creosotes contain 0.95–6.1% (w/w) acenaphthene (Kohler *et al.*, 2000). Nevertheless, the biodegradation of the compound has been poorly studied, especially in terms of bacterial catabolic genes. To our knowledge, there are only a few acenaphthene-utilizing bacterial strains, almost all of which are *Sphingomonas* spp. isolated by our group (Komatsu *et al.*, 1993) or by Shi *et al.* (2001), except for *Alcaligenes* spp. reported by Selifonov *et al.* (1993). Previously, we demonstrated that the degradation of acenaphthene by strain A4 proceeds via 1-acenaphthenol and 1-acenaphthenone (Komatsu *et al.*, 1993). Recently, we isolated the terminal oxygenase genes, *arhA1* and *arhA2*, required for the initial oxygenation of acenaphthene into 1-acenaphthenol, using the shotgun cloning method (Pinyakong *et al.*, 2004). That was the first report to identify the genes responsible for PAH oxygenation in sphingomonads, along with a report on *phnA* genes published almost simultaneously by Demanèche *et al.* (2004). We also demonstrated that several PAHs other than acenaphthene, i.e. acenaphthylene, naphthalene, phenanthrene, anthracene and fluoranthene, could be dioxygenated to their corresponding *cis*-dihydrodiol products in resting *Escherichia coli* cells coexpressing *arhA1A2* with the genes encoding ferredoxin (*ahdA3*) and ferredoxin reductase (*ahdA4*) derived from another sphingomonad (Pinyakong *et al.*, 2004). However, the genes for the intrinsic electron-transport protein for ArhA1A2, the genes for the degradation pathway of their metabolic products, and their regulatory genes, have not been isolated.

In this study, we performed transposon mutagenesis on strain A4 to isolate the genes, other than *arhA1A2*, responsible for acenaphthene degradation. We also investigated the regulatory mechanism of a newly isolated gene cluster, which includes the *arhA* genes, as the first step for elucidating the entire acenaphthene degradation system in strain A4.

METHODS

Chemicals. The chemicals used in this study were of the highest purity available commercially (Kanto Chemical; Wako Pure Chemical; Tokyo Kasei Kogyo; Nacalai Tesque; Sigma-Aldrich).

Bacterial strains, plasmids and culture conditions. The bacterial strains and plasmids used in this study are listed in Table 1. *Sphingomonas* sp. strain A4 was cultivated at 30 °C in Luria–Bertani (LB) medium (Sambrook *et al.*, 1989) or carbon-free mineral medium (CFMM) as described previously (Pinyakong *et al.*, 2004) supplemented with 0.1% (w/v) acenaphthene, 0.1% (w/v) acenaphthylene, 0.1% (w/v) 1,8-naphthalic anhydride (1,8-NA), or 0.1% or 0.2% (w/v) fructose. Acenaphthene, acenaphthylene and 1,8-NA solutions were prepared by dissolving the appropriate compounds in DMSO (acenaphthene and acenaphthylene, 100 mg ml⁻¹; 1,8-NA, 25 mg ml⁻¹). The *E. coli* strains were grown at 37 °C in LB medium. When necessary, ampicillin (Ap), chloramphenicol (Cm), kanamycin (Km) and tetracycline (Tc) were used at 100, 34, 50 and 10 µg ml⁻¹, respectively. Gentamicin (Gm) was used at 10 µg ml⁻¹ for *E. coli* strains and at 30 µg ml⁻¹ for mutants of strain A4. IPTG was used at a final concentration of 100 µM.

DNA manipulations. Plasmid isolation, restriction endonuclease digestion, cloning, the transformation of *E. coli* and Southern hybridization were carried out using standard protocols (Sambrook *et al.*, 1989). For Southern hybridization, DNA fragments within *arhR* and ORF16 were amplified using PCR. The PCR products were labelled using a DIG DNA labelling kit (Roche Diagnostics) and used as probes. The primers used in this study are listed in Table 2.

Transposon mutagenesis. Transposon mutagenesis was carried out using filter-mating between strain A4 and *E. coli* S17-1 λ pir harbouring suicide plasmid pBSL202 (Alexeyev *et al.*, 1995). Both strains were grown on LB medium until the stationary phase and then suspended in 10 mM MgSO₄ after being washed twice with the same solution. Equal numbers of cells from these suspensions were mixed and incubated on a filter membrane (0.45 µm pore size) on an LB agar plate at 30 °C for 6 h. The cells were resuspended in 10 mM MgSO₄ and then spread on CFMM plates supplemented with 0.2% (w/v) fructose and Gm. After incubation at 30 °C for 5 days, 3% (w/v) acenaphthene solution dissolved in acetone was sprayed on the cells that had grown on the selective medium, and the cells were incubated overnight at 30 °C. The mutants that failed to produce a clear zone in the acenaphthene layer were isolated. For the indigo formation assay, the mutants were pre-incubated on an LB plate containing Gm at 30 °C for 4 days, and then a small amount of indole was put on the inside of the lid of the plate. After further incubation at 30 °C for 24 h, we observed whether the colour of the cells had changed to indigo blue.

The DNA regions flanking the mini-Tn5 insertion sites in mutants of strain A4 were cloned by digesting total DNA from the mutants with *EcoRI*, ligating them to *EcoRI*-digested pUC19, and transforming *E. coli* DH5 α . The clone libraries were selected on LB agar plates containing both Ap and Gm. Positive clones were subjected to nucleotide

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics*	Source or reference
<i>Sphingomonas</i> sp.		
A4	Ace ⁺ , 1,8-NA ⁺	Komatsu <i>et al.</i> (1993)
A4-PCM1	Ace ⁻	Pinyakong <i>et al.</i> (2004)
AG2-45	Mini-Tn5 insertion mutant of A4, Ace ⁻ , 1,8-NA ⁺ , Gm ^r	This study
AG2-48	Mini-Tn5 insertion mutant of A4, Ace ⁻ , 1,8-NA ⁺ , Gm ^r	This study
AG3-15	Mini-Tn5 insertion mutant of A4, Ace ⁻ , 1,8-NA ⁺ , Gm ^r	This study
AG3-69	Mini-Tn5 insertion mutant of A4, Ace ⁻ , 1,8-NA ⁻ , Gm ^r	This study
A4D15	ORF15::Gm ^r , Ace ⁺	This study
A4DR	<i>arhR</i> ::Gm ^r , Ace ⁻	This study
A4DRC	A4DR carrying pBBadR, Ace ⁺ , Gm ^r , Tc ^r	This study
<i>E. coli</i>		
S17- λ pir	<i>hsdR thi pro recA</i> ; RP4 integrated into the chromosome (<i>kan</i> ::Tn7 <i>tet</i> ::Mu) λ pir	de Lorenzo & Timmis (1994)
DH5 α	<i>supE44 ΔlacU169 (ϕ80 <i>lacZ</i>ΔM15) <i>hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i></i>	Laboratory collection
JM109	<i>e14⁻ (mcrA) recA1 endA1 gyrA96 thi-1 <i>hsdR17</i> (r_k⁻ m_k⁻) <i>supE44 relA1</i></i>	Laboratory collection
Plasmids		
pUC18/19	Cloning vector, Ap ^r	Yanisch-Perron <i>et al.</i> (1985)
pSTV28	Cloning vector, Cm ^r	Takara
pTKm	pT7Blue containing the Km ^r cassette	Yoshida <i>et al.</i> (2003)
pUC19Km ^r	Cloning vector, Km ^r	This study
pBSL202	Suicide vector, Ap ^r , mini-Tn5 Gm ^r	Alexeyev <i>et al.</i> (1995)
pEG1	pUC19 with 2.7 kb <i>EcoRI</i> fragment containing the flanking region of mini-Tn5 insertion site of AG3-15	This study
pEG2	pUC19 with 8.5 kb <i>EcoRI</i> fragment containing the flanking region of mini-Tn5 insertion site of AG2-45	This study
pEG3	pUC19 with 1.5 kb <i>EcoRI</i> fragment containing the flanking region of mini-Tn5 insertion site of AG2-48	This study
pEG4	pUC19 with 4.2 kb <i>EcoRI</i> fragment containing the flanking region of mini-Tn5 insertion site of AG3-69	This study
pK19 <i>mobsacB</i>	Vector for gene disruption containing <i>sacB</i> , Km ^r	Schäfer <i>et al.</i> (1994)
pBBad22T	Broad-host-range vector containing P _{<i>araBAD</i>} promoter and <i>araC</i> , Tc ^r	Sukchawalit <i>et al.</i> (1999)
SuperCos1	Cosmid vector, Ap ^r , Km ^r	Stratagene
pUK17	pUC19Km ^r with 16.4 kb <i>KpnI</i> fragment from pSC288	This study
pSC288	SuperCos1 cosmid library clone containing <i>arhA1A2</i> and its flanking region from strain A4	Pinyakong (2003)
pU288E1	pUC18 with 5.0 kb <i>EcoRI</i> fragment containing <i>arhA1A2</i>	Pinyakong <i>et al.</i> (2004)
pBSCN1	pBluescript II SK(-) with 5 kb <i>Clal</i> - <i>NotI</i> fragment from pSC288	This study
pBSCN2	pBluescript II SK(-) with 4.5 kb <i>Clal</i> - <i>NotI</i> fragment from pSC288	This study
pU351E1	pUC18 with 3.4 kb <i>EcoRI</i> fragment from pSC288	This study
pKGR16	pK19 <i>mobsacB</i> with 1.4 kb <i>HindIII</i> - <i>Sall</i> fragment from pBSCN2, <i>arhR</i> ::Gm ^r Km ^r	This study
pKG15	pK19 <i>mobsacB</i> with 1.0 kb <i>SphI</i> fragment from pU351E1, ORF15::Gm ^r Km ^r	This study
pBBadR	pBBad22T with 2.0 kb <i>SmaI</i> - <i>HindIII</i> fragment carrying <i>arhR</i> from pBSCN2, Tc ^r	This study
pBBadA12	pBBad22T with 2.9 kb <i>NcoI</i> fragment carrying <i>arhA1</i> and <i>arhA2</i> from pU288E1, Tc ^r	This study
pBBadA13	pBBadA12 with 0.7 kb <i>SmaI</i> - <i>NheI</i> fragment carrying <i>arhA3</i> from pBSCN2, Tc ^r	This study
pBBadA14	pBBadA13 with 1.5 kb <i>Sall</i> fragment carrying <i>arhA4</i> from pBSCN1, Tc ^r	This study
pUARhA1A2	pUC18 with 2.6 kb <i>Sall</i> - <i>PstI</i> fragment carrying <i>arhA1A2</i> from pU288E1, Ap ^r	Pinyakong <i>et al.</i> (2004)
pSArhA3	pSTV28 with 0.7 kb <i>KpnI</i> - <i>HindIII</i> fragment carrying <i>arhA3</i> from pBBadA13, Ap ^r	This study
pSArhA3A4	pSTV28 with 2.3 kb <i>KpnI</i> - <i>HindIII</i> fragment carrying <i>arhA3A4</i> from pBBadA14, Cm ^r	This study

*Ace⁺/Ace⁻ or 1,8-NA⁺/1,8-NA⁻ represents the ability/inability to grow on acenaphthene or 1,8-naphthalic anhydride as the sole source of carbon and energy. Ap^r, Gm^r, Km^r, Cm^r and Tc^r represent resistance to ampicillin, gentamicin, kanamycin, chloramphenicol and tetracycline, respectively.

Table 2. Primers

Primer	Sequence (5' to 3')	Annealed region (nucleotide positions*)
For Southern hybridization		
arhR for	TTGGTGTGGTGAAGATCCG	<i>arhR</i> (432–451)
arhR rev	GGCCTCCTTGAAAAGTATGC	<i>arhR</i> (1324–1342)
ORF15 for	ATGTCTGGTCTAATTATTGAGC	ORF15 (6245–6266)
ORF15 rev	TTAATAGGTCCTGAGCCTAAGG	ORF15 (7132–7153)
For sequence analyses		
GMR	AGTTGGGCATACGGGAAGA	Gm ^r cassette
For RT-PCR		
arhRT1F	AGGCAGTGGCAAAAGTTGAG	<i>arhA3</i> (2366–2385)
arhRT1R	GAGATGATGACGCTCTTGGT	ORF11 (3322–3341)
arhRT2F	CGTCTGGATCAACCGCTATT	ORF11 (3210–3229)
arhRT2R	CCGCTCACAGAAAGCTATCA	<i>arhA4</i> (4421–4440)
arhRT3F	CATTGGATTTCCGGCGGCTTT	<i>arhA4</i> (4309–4328)
arhRT3R	GCGCGCATCAGTTTAAAGATAG	ORF13 (5677–5697)
arhRT4F	TCAATGGGCTTACGATGCGA	ORF13 (4780–4799)
arhRT4R	CCCTCGTTCCGGATAAATTC	ORF14 (6097–6116)
arhRT5F	ATGCAAGATCCACGAGGCAA	ORF14 (5982–6001)
arhRT5R	AAGGTGACCTCGACTCCATT	ORF15 (6500–6519)
arhRT6F	ATGTCTGGTCTAATTATTGAGC	ORF15 (6245–6266)
arhRT6R	GGCCGAACATGATCTCGAT	ORF16 (7842–7860)
arhRT7F	ATGAAGCAGAAGGCTATGGG	ORF15 (7107–7126)
arhRT7R	GACCCGGTCAATTTCTGAGA	ORF1 (8545–8564)
arhRT8F	AACAGATCCAACGAACTGGG	ORF1 (8478–8497)
arhRT8R	CCCTGAATGGTAAGGAAGTC	ORF1 (10546–10565)
arhRT9F	ACTTCCTTACCATTTCAGGGT	ORF1 (10547–10566)
arhRT9R	ATTGCAAACACGCTTGGTG	<i>arhA1</i> (11723–11742)
arhRT10F	GGATGACCTGATCGACACCAA	<i>arhA1</i> (11708–11728)
arhRT10R	TCATATTACGGCCCTGCACA	<i>arhA2</i> (13620–13639)
arhRT11F	TGTGCAGGGCCGTAATATGA	<i>arhA2</i> (13620–13639)
arhRT11R	GTTTCGAGCTGGATGAAATCG	ORF6 (15659–15678)
For quantitative RT-PCR		
arhA1qF	TCTACGGTTCGCATCTCAAC	<i>arhA1</i> (12565–12584)
arhA1qR	CTGCCAGATCTTCAACGTGT	<i>arhA1</i> (12625–12644)
arhA3qF	AGGCAGTGGCAAAAGTTGAG	<i>arhA3</i> (2366–2385)
arhA3qR	CTGGTCGCGAATAATTCTCC	<i>arhA3</i> (2458–2477)
arhRqF	ATGTGCACCTGAACCTTGCTG	<i>arhR</i> (1208–1227)
arhRqR	TTTTGAGCGCGTTGCTGACA	<i>arhR</i> (1302–1321)
16SqF	GTGAGTGATGAAGGCCTTAG	<i>rrn</i> [†]
16SqR	CTTTACGCCCAGTAATTCCG	<i>rrn</i> [†]
For primer extension		
arhA3-PE1	CTGGTCGCGAATAATTCTCC	<i>arhA3</i> (2458–2477)

*The positions correspond to the nucleotide positions in AB240454.

†The primers for *rrn* were designed from the sequence of the 16S rRNA gene in AB104886.

sequencing using primer GMR (Table 2), which is specific for the Gm^r cassette in the mini-Tn5, and an ABI PRISM 310 genetic analyser (Applied Biosystems) according to the manufacturer's instructions.

Nucleotide sequence analysis and annotation. Previously, we obtained cosmid pSC288 containing *arhA1A2* and its flanking region in a 40 kb insert from a genomic library of strain A4

(Pinyakong, 2003). In this study, the 16·4 kb *KpnI* fragment of pSC288 containing the *arhA1A2* genes was subcloned into pUC19Km^r to yield pUK17 (Table 1), and the nucleotide sequence of this fragment was determined by the Dragon Genomics Center of Takara Bio. To construct pUC19Km^r, the Km^r cassette from pTKm (Yoshida *et al.*, 2003) was excised by *EcoRV* digestion, and was ligated to the *ScaI* site of pUC19. The nucleotide sequences were

analysed using DNASIS-Mac software (version 3.7; Hitachi Software). We searched for homology using the BLAST programs available at the web site of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast/blast_references.shtml). The deduced amino acid sequences of the observed ORFs were aligned using ClustalW (Thompson *et al.*, 1994), available at the NPSA web site (<http://npsa-pbil.ibcp.fr/NPSA/>).

Gene disruption and complementation. To disrupt *arhR* and ORF15, we first constructed plasmids pKGR16 and pKG15, respectively (Table 1, Fig. 1). The Gm^r cassette used for gene disruption was prepared as described previously (Pinyakong *et al.*, 2004). To construct pKGR16, the Gm^r cassette was ligated to the blunt-ended *Bam*HI site of pBSCN2 (Table 1, Fig. 1), as the *arhR* gene and the Gm^r cassette are transcribed in the same direction. The 2.0 kb *Hind*III–*Sal*I fragment of the resultant plasmid containing *arhR*::Gm^r was ligated to *Hind*III/*Sal*I-digested pK19*mobsacB* to yield pKGR16. To construct pKG15, the 1.0 kb *Sph*I fragment of pU351E1 (Table 1, Fig. 1) containing part of ORF15 was ligated to *Sph*I-digested pK19*mobsacB*, in which the *Sma*I site had been eliminated by *Sal*I/*Eco*RI digestion, blunting and self-ligation. The resultant plasmid was digested with *Sma*I and ligated to the Gm^r cassette, as ORF15 and the Gm^r cassette are transcribed in the same direction, to yield pKG15. Plasmids pKGR16 and pKG15 were introduced into strain A4 by filter-mating, and then the double-crossover recombinants were screened using a method described previously (Pinyakong *et al.*, 2004). The resultant *arhR* and ORF15 disruptants were designated strains A4DR and A4D15, respectively. The insertion of the Gm^r cassette into *arhR* and ORF15 was confirmed by Southern hybridization. To complement strain A4DR with *arhR* *in trans*, the 2.0 kb *Sma*I–*Hind*III fragment of pBSCN2 (Table 1, Fig. 1) containing *arhR* was ligated into *Sma*I/*Hind*III-digested pBBad22T (Table 1), as the *arhR* gene was located downstream from the arabinose-inducible *P*_{araBAD} promoter. The resultant plasmid, pBBadR (Table 1, Fig. 1), was introduced into strain A4DR by filter-mating to generate strain A4DRC.

Expression of the *arhA* genes in strain A4-PCM1 and assay for indigo formation activity. To express the *arhA* genes, *arhA1A2*, *arhA1A2A3* and *arhA1A2A3A4* were ligated downstream from the *P*_{araBAD} promoter in pBBad22T (Table 1) to yield pBBadA12, pBBadA13 and pBBadA14, respectively (Table 1, Fig. 1). To obtain pBBadA12, the 2.9 kb *Nco*I fragment of pU288E1 (Table 1, Fig. 1) containing *arhA1A2* was ligated into *Nco*I-digested pBBad22T. To obtain pBBadA13, the 0.7 kb *Sma*I–*Nhe*I fragment of pBSCN2 (Table 1, Fig. 1) containing *arhA3* was ligated into *Sma*I/*Xba*I-digested pBBadA12. To obtain pBBadA14, the 1.5 kb *Sal*I fragment of pBSCN1 (Table 1, Fig. 1) containing *arhA4* was ligated into *Sal*I-digested pBBadA13. Plasmids pBBadA12, pBBadA13, pBBadA14 and pBBad22T were introduced into strain A4-PCM1 (Pinyakong *et al.*, 2004), which is a spontaneous mutant defective in the ability to utilize acenaphthene. The strains harbouring the different plasmids were pre-incubated at 30 °C for 4 days on an LB agar plate containing Tc and 0.2 % arabinose. After a small amount of indole was added on the lid of the plate, the cells were incubated at 30 °C for 5 more days and then checked for the ability to produce indigo.

Construction of *E. coli* clones expressing the *arhA* genes and an assay for acenaphthene oxygenase activity. For the purpose of expression of the *arhA3* or *arhA3A4* genes in *E. coli*, the 0.7 or 2.3 kb *Kpn*I–*Hind*III fragment containing the respective genes from pBBadA13 or pBBadA14 was cloned into the corresponding site of pSTV28, resulting in pSArhA3 or pSArhA3A4 (Table 1). *E. coli* JM109 was then cotransformed with plasmids pSArhA3 or pSArhA3A4 and pUArhA1A2 (Pinyakong *et al.*, 2004; Table 1) containing *arhA1A2*. The resultant *E. coli* clones were grown at 37 °C in 100 ml LB medium containing Ap and Cm to an OD₆₀₀ of 0.7–0.8. IPTG was added at a final concentration of 100 µM followed by further cultivation at 30 °C for 4 h. The cells obtained were then washed twice with CFMM and resuspended in the same medium to yield a cell suspension of OD₆₀₀ 9–8. A 5 ml sample of each cell suspension was placed in the reaction tube and supplemented with 0.01 % (w/v) acenaphthene. After incubation at

Fig. 1. Physical map of the 16.4 kb *Kpn*I fragment of strain A4 and the location of the DNA fragments amplified by RT-PCR. Boxed arrows indicate the length and direction of the ORFs. The positions of the mini-Tn5 insert in the mutant strains (AG2-45, AG2-48 and AG3-15) are shown as solid triangles. Line segments under the physical map show inserts in the plasmids used in this study. To obtain pBBadA13, the 0.7 kb *Sma*I–*Nhe*I fragment containing *arhA3* was cloned into pBBadA12. For pBBadA14, the 1.5 kb *Sal*I fragment containing *arhA4* was cloned into pBBadA13. The locations of the Gm^r cassettes in plasmids pKGR16 and pKG15 are indicated. The bidirectional arrows indicate the locations of the DNA fragments amplified by RT-PCR with total RNA from acenaphthene-grown strain A4 cells and the primer sets shown in Table 2.

30 °C for 2 or 16 h, 0.01 % (w/v) fluoranthene, which served as an internal standard, was added to the reaction tube and the reaction mixture was extracted with an equal volume of ethyl acetate. The extracts were dried over anhydrous Na₂SO₄ and the solvent was removed under reduced pressure. The metabolite was derivatized with *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (MSTFA) at 70 °C for 20 min, and analysed by GC-MS as described previously (Pinyakong *et al.*, 2004). Control experiments using *E. coli* JM109 carrying pUArhA1A2 and pSTV28 (Table 1), or pUC18 (Table 1) and pSTV28, were performed in parallel. To detect acenaphthene-*cis*-1,2-diol, the metabolite was derivatized with methanboronic acid dissolved in dehydrated pyridine (2 mg ml⁻¹) at 70 °C for 30 min.

RNA preparation. Total RNA from strain A4 and its derivatives was extracted using a NucleoSpin RNA II (Macherey-Nagel), according to the manufacturer's instructions. The culture conditions used for total RNA extraction from the cells were as follows. For RT-PCR and primer extension, strain A4 was inoculated into CFMM supplemented with 0.1 % (w/v) acenaphthene, and the cells were harvested at approximately exponential phase. For quantitative RT-PCR, strains A4, A4DR and A4DRC were cultured on CFMM supplemented with 0.1 % (w/v) fructose to an OD₆₀₀ of 0.5 and were then supplemented with 2 % (w/v) acenaphthene solution dissolved in DMSO at a final concentration of 0.02 % (w/v) or the same volume of DMSO. After further cultivation for 2 h, the cells were harvested. The RNA was treated with RQ1 RNase-free DNase (Promega) before being used in further experiments.

RT-PCR. RT-PCR was performed with a One Step RNA PCR kit (AMV) (Takara). The DNA regions between *arhA3* and ORF6 (Fig. 1) were amplified using the primer sets shown in Table 2. The reaction mixture (25 µl) contained 2.5 µl 10-fold One Step RNA PCR buffer, 5 mM MgCl₂, 1 mM dNTP, 20 U RNase inhibitor, 2.5 U AMX reverse transcriptase XL, 2.5 U AMV-optimized *Taq*, 0.4 µM each primer, and 0.1 µg total RNA, prepared as described above. After the RT reaction at 50 °C for 30 min, PCR was performed using the following conditions: 94 °C for 2 min, and 30 cycles of 94 °C for 30 s, 65 °C for 30 s and 72 °C for 2.5 min. Negative control reactions were performed similarly, except the reverse transcriptase was omitted from the reaction mixture.

Quantitative RT-PCR. First-strand cDNA was synthesized in four separate RT reactions using the reverse primers for *arhA1*, *arhA3*, *arhR* and the 16S rRNA gene (*rrn*) (*arhA1*qR, *arhA3*qR, *arhR*qR and 16SqR, respectively; Table 2). The RT reaction was performed using SuperScript III reverse transcriptase (Invitrogen). The RT reaction mixture (20 µl) contained 0.3 µg total RNA from each cell culture, 0.1 µM each primer (Table 2), 4 µl fivefold First Strand buffer, 1 µl 0.1 M DTT, 1 µl RNaseOUT recombinant RNase inhibitor and 200 U SuperScript III reverse transcriptase. The reaction mixture was heated to 65 °C for 5 min and then chilled on ice. The RT enzymes were added, and the reaction mixture was incubated at 50 °C for 1 h. The reaction was terminated by incubation at 70 °C for 15 min. The cDNA solution was diluted 10-fold in deionized water before PCR amplification.

Real-time PCR was carried out with the ABI PRISM 7700 Sequence Detection System (Applied Biosystems) according to the manufacturer's instructions. The PCR mixture (22 µl) contained 2 µl of the cDNA solution, 0.4 µM each primer set (Table 2) and 11 µl SYBR Green PCR Master Mix (Applied Biosystems). Real-time PCR was performed using the following conditions: 50 °C for 2 min, 94 °C for 10 min, and 40 cycles of 94 °C for 5 s, 65 °C for 5 s and 72 °C for 30 s. For the standard curves, the DNA fragments of *arhA1*, *arhA3*, *arhR* and *rrn* were amplified by PCR with total DNA from strain A4 and the primer sets for the respective genes (Table 2), and dilution series were subjected to real-time PCR analyses as templates. The specificity

of the real-time PCR was verified from the dissociation curve analysis and agarose gel electrophoresis of the PCR products. Negative control reactions were performed similarly, except the reverse transcriptase was omitted from the RT reaction mixture. The mRNA levels of the target genes (*arhA1*, *arhA3* and *arhR*) were normalized to that of the reference gene (*rrn*) to correct for sample-to-sample variation in the amount of total RNA.

Primer extension analysis. The total RNA from strain A4 cells grown on acenaphthene was subjected to an RT reaction with SuperScript III reverse transcriptase (Invitrogen) and an IRD800-labelled primer (Aloka), *arhA3*-PE1 (Table 2), using the method described above. The primer extension products were purified using phenol/chloroform extraction and ethanol precipitation. The products were subjected to electrophoresis together with a sequence reaction using the same primer and a Li-Cor model 4200I-2 Auto-DNA sequencer running Base ImagIR data collection software 4.0 (LI-COR), according to the manufacturer's instructions.

RESULTS

Isolation and characterization of the mini-Tn5-inserted mutants

To identify the genes involved in acenaphthene degradation, we carried out Tn5 mutagenesis on strain A4. Of approximately 10 000 strains with mini-Tn5 inserts, four mutants (strains AG2-45, AG2-48, AG3-15 and AG3-69) that no longer produced a clear zone with acenaphthene were obtained. These four mutants also showed deficient growth on acenaphthene in liquid culture (data not shown). Strains AG2-45, AG2-48 and AG3-15 maintained the ability to grow on 1,8-NA, which is a metabolic intermediate of acenaphthene (Selifonov *et al.*, 1996), whereas strain AG3-69 had also lost the ability to grow on 1,8-NA. As strain A4 possesses the ability to produce indigo from indole, which involves the ring-hydroxylating oxygenase *ArhA1A2* (Pinyakong *et al.*, 2004), we examined the ability of these four mutants to produce indigo in order to estimate their ring-hydroxylating oxygenase activities. Strains AG2-45, AG2-48 and AG3-15 had also lost the ability to produce indigo (data not shown). These observations suggested that the genes involved in the initial oxygenation of acenaphthene did not function in strains AG2-45, AG2-48 and AG3-15.

To determine the location of the mini-Tn5 insertion in these four mutants, the region flanking the mini-Tn5 was cloned and partially sequenced. In strain AG2-45, the mini-Tn5 was inserted into the DNA region homologous (40 % amino acid identity) to part of the gene encoding an LTTR, DntR, from *Burkholderia* sp. strain DNT (a 2,4-dinitrotoluene degrader) (Lessner *et al.*, 2003). In strains AG2-48 and AG3-15, the mini-Tn5 was inserted into ORF1 and *arhA1*, respectively, both of which we have sequenced previously (Pinyakong *et al.*, 2004). In strain AG3-69, the mini-Tn5 was inserted into the DNA region homologous (59 % amino acid identity) to part of the gene encoding a ferredoxin reductase, RedA2, from *Sphingomonas wittichii* strain RW1 (a dibenzo-*p*-dioxin degrader) (Armengaud & Timmis, 1998) (data not shown).

Nucleotide sequence analysis and annotation of the 16.4 kb DNA region

Previously, we partially sequenced the 40 kb insert in cosmid pSC288 (Pinyakong, 2003), which contains a 5.0 kb *EcoRI* fragment carrying *arhA1A2* and its flanking region. Consistent with part of the preliminary sequence of pSC288, the DNA region flanking the mini-Tn5 in AG2-45 appeared to be located in the same locus as the 5.0 kb *EcoRI* fragment, within approximately 16 kb (Fig. 1). To determine the complete nucleotide sequence of this locus, the corresponding 16.4 kb *KpnI* fragment of pSC288 was subcloned and sequenced, revealing 16 ORFs and a partial ORF at this locus (Fig. 1, Table 3).

The complete deduced amino acid sequence of ORF7 (designated *arhR*), into which the mini-Tn5 was inserted in strain AG2-45, also shared the highest identity with an LTTR (DntR) from *Burkholderia* sp. strain DNT (Lessner *et al.*, 2003) (Table 3). The double-crossover disruptant of *arhR* (strain A4DR, Table 1) showed deficient growth on acenaphthene, and this deficiency was complemented in strain A4DR transformed with plasmid pBBadR (Table 1, Fig. 1) expressing the *arhR* gene (strain A4DRC, Table 1). These observations revealed that ArhR is clearly involved in acenaphthene metabolism in strain A4.

The deduced amino acid sequences of ORF10 and ORF12 shared the highest homology (50 % and 46 % amino acid identity) with a ferredoxin and a ferredoxin reductase (ThnA3 and ThnA4), respectively, from *Sphingopyxis macrogoltabida* strain TFA (a tetralin degrader) (Moreno-Ruiz *et al.*, 2003). We postulated that ORF10 and ORF12 encoded an intrinsic ferredoxin and ferredoxin reductase for ArhA1A2 and tentatively designated them *arhA3* and *arhA4*, respectively.

The deduced amino acid sequence of ORF11 shared the highest identity with a putative 1-hydroxy-2-naphthoaldehyde dehydrogenase (PhnF) from *Alcaligenes faecalis* strain AFK2 (aphenanthrene degrader) (Kiyohara *et al.*, 1982). However, the product of ORF11 was approximately 280 aa shorter than other aldehyde dehydrogenases, including PhnF, and lacked both the substrate- and NADH-binding domains (Gruez *et al.*, 2004), which are considered necessary for its enzymic activity. Therefore, ORF11 seems to be a pseudogene generated by deletion of the corresponding DNA region.

The ORF13 product showed homology with the large subunit of the terminal oxygenase component of salicylate 1-hydroxylase (AhdA1d) from *Sphingobium* sp. P2 (Pinyakong *et al.*, 2003b). However, ORF13 also seems to be a pseudogene because its products lacked part of the Rieske [2Fe-2S] cluster domain which is required for oxygenase activity (Ferraro *et al.*, 2005).

The ORF15 product displayed high homology with known 2,3-dihydroxybiphenyl 1,2-dioxygenases. However, the double-crossover disruptant of ORF15 (strain A4D15)

did not lose its ability to grow on acenaphthene, revealing that ORF15 is not essential for acenaphthene utilization.

The ORF6 product showed the highest homology with ferredoxin reductase (NahAa) from *Pseudomonas putida* NCIB 9816-4 (a naphthalene degrader) (Dennis & Zylstra, 2004), but this homology was extremely low (18.1 % amino acid identity). Especially, the C-terminal amino acid sequence of ORF6 showed no significant homology with that of NahAa, which contains the NAD-binding domain. In a BLAST analysis of the C-terminal region of ORF6, where the NAD-binding domain should be, we failed to detect any significant homologous proteins or peptides. It is still unclear whether the ORF6 product acts as a ferredoxin reductase. This ORF may have a function quite different from that of a ferredoxin reductase gene.

Assay of ArhA oxygenase activity

To examine whether the tentatively designated ArhA3 (ORF10) and ArhA4 (ORF12) really act as ferredoxin and ferredoxin reductase for ArhA1A2, respectively, we constructed three plasmids, pBBadA12, pBBadA13 and pBBadA14 (Table 1, Fig. 1), which coexpressed the *arhA1A2*, *arhA1A2A3* and *arhA1A2A3A4* genes, respectively, from the arabinose-inducible *P_{araBAD}* promoter. Plasmids pBBadA12 and pBBad22T (Table 1) were used as negative controls. These four plasmids were introduced into strain A4-PCM1, which is a spontaneous mutant defective in the ability to utilize acenaphthene (Pinyakong *et al.*, 2004). In strain A4-PCM1, the loss of the DNA region from *arhR* to *arhA1* (Fig. 1) was confirmed by Southern hybridization (data not shown). After pre-incubation on an LB agar plate containing Tc and 0.2 % arabinose, the ability of strain A4-PCM1 cells harbouring these four plasmids to produce indigo from indole was investigated by adding indole on the lid of the plate. After further incubation for 5 days, slight indigo formation was observed for strain A4-PCM1(pBBadA13), compared with the negative controls (data not shown). By contrast, significant indigo formation was observed only for strain A4-PCM1(pBBadA14) (data not shown). These observations showed that ArhA3 and ArhA4 function as the electron-transport system for ArhA1A2. To investigate acenaphthene oxygenase activity in the cells with or without ArhA3 and ArhA4 more quantitatively, we constructed *E. coli* strains harbouring plasmids pUArhA1A2 (Pinyakong *et al.*, 2004) and pSArhA3 or pSArhA3A4 (Table 1). *E. coli* strains harbouring pUC18 and pSTV28 or pUArhA1A2 and pSTV28 were used in control experiments. After resting-cell reactions with these *E. coli* strains and acenaphthene as a substrate, the metabolites were derivatized with MSTFA and subjected to GC-MS analysis. Two major products, 1-acenaphthenol and acenaphthene-1,2-diol, were detected and their relative quantities were determined (Table 4). The amounts of 1-acenaphthenol and acenaphthene-1,2-diol after 16 h of incubation were increased 2.5-fold and 10-fold, respectively, in the cells expressing *arhA1A2A3* as compared with the cells expressing only *arhA1A2*, and further increased in the

Table 3. Putative ORFs flanking the *arh* genes and their predicted products

ORF no. (gene name)	Probable function	Source		No. of amino acids/relatives	Amino acid identity (%)	Accession no.
		Protein	Micro-organism			
ORF7 (<i>arhR</i>)	LTTR	DntR	<i>Burkholderia</i> sp. DNT	303/301	40 (121/302)	AAP70493
ORF8	Transposase	ELI2880	<i>Erythrobacter litoralis</i> HTCC2594	65/284	19 (54/284)	ZP_00377639
ORF9	Transposase	ELI2879	<i>Erythrobacter litoralis</i> HTCC2594	88/88	90 (79/88)	ZP_00377638
ORF10 (<i>arhA3</i>)	Rieske-type ferredoxin	ThnA3	<i>Sphingopyxis macrogoltabida</i> TFA	109/104	50 (56/111)	AAD52963
ORF11*	1-Hydroxy-2-naphthoaldehyde dehydrogenase	PhnF	<i>Alcaligenes faecalis</i> AFK2	220/511	24 (124/511)	BAA76331
ORF12 (<i>arhA4</i>)	Ferredoxin reductase	ThnA4	<i>Sphingopyxis macrogoltabida</i> TFA	353/339	46 (164/353)	AAN26446
ORF13*	Large subunit of oxygenase	AhdA1d	<i>Sphingobium</i> sp. P2	351/424	44 (188/425)	BAC65433
ORF14	Small subunit of oxygenase	BphA2d	<i>Novosphingobium aromaticivorans</i> F199	161/162	31 (52/163)	AAD04001
ORF15	2,3-Dihydroxybiphenyl 1,2-dioxygenase	BphC	<i>Novosphingobium aromaticivorans</i> F199	302/299	67 (203/302)	AAD04006
ORF16	Transposase	Saro02000677	<i>Novosphingobium aromaticivorans</i> F199	130/129	86 (111/129)	ZP_00304890
ORF1	Hypothetical protein	Saro02002648	<i>Novosphingobium aromaticivorans</i> F199	824/827	55 (466/850)	ZP_00302784
ORF2†	Hypothetical protein					
ORF3 (<i>arhA1</i>)	Large subunit of dioxygenase	PhnAc	<i>Burkholderia</i> sp. RP007	443/450	56 (250/450)	AAD09872
ORF4 (<i>arhA2</i>)	Small subunit of dioxygenase	PhnAd	<i>Burkholderia</i> sp. RP007	191/195	44 (86/195)	AAD09873
ORF5	Hypothetical protein	XlnF	<i>Pseudomonas alcaligenes</i> NCIMB 9867	306/287	47 (143/306)	AAG39451
ORF6	Ferredoxin reductase	NahAa	<i>Pseudomonas putida</i> NCIB 9816-4	367/328	18 (67/370)	NP_863070
ORF17‡	Transposase	ORF11	<i>Rhizobium</i> sp. NGR234			CAA92408

*These ORFs are considered pseudogenes based on their deduced amino acid sequences (see text).

†The N-terminal sequence shows homology [65 % amino acid identity (91/140)] with hydratase/aldolase (PhnE) from *Burkholderia* sp. RP007 (accession no. Q9ZHH6), but the C-terminal sequence shows homology [58 % amino acid identity (35/60)] with a putative isomerase from *Burkholderia* sp. DBT1 (accession no. AAK96187).

‡Part of an ORF.

cells expressing *arhA1A2A3A4*. These results clearly showed that *ArhA3* and *ArhA4* were necessary for maximal acenaphthene oxygenase activity.

Analysis of the operon structure of the *arh* genes and their flanking region by RT-PCR

To investigate how the DNA region between *arhA3* and ORF6 is transcribed, RT-PCR was performed with total RNA from acenaphthene-grown strain A4 cells and the primer set designed to amplify this region (Table 2, Fig. 1). DNA fragments of the expected size were amplified with all the primer sets used. No amplification was observed when the reverse transcriptase was omitted from the reaction mixture (date not shown). These results suggested that the DNA region from *arhA3* to ORF6 is transcribed within the same transcription unit, although some promoters might be included within this region.

Involvement of *ArhR* in *arh* gene expression

To investigate how the disruption of *arhR* affects the expression of the *arh* genes, quantitative RT-PCR analyses of the *arhR*, *arhA3* and *arhA1* genes were performed. The differences in their mRNA levels in strains A4, A4DR (the *arhR* disruptant) and A4DRC (A4DR carrying *arhR* in trans) were compared under two conditions, i.e. with and without induction by acenaphthene (Fig. 2). In the wild-type cells, acenaphthene induced transcription of *arhA3* and *arhA1* (Fig. 2a, b). By contrast, in strain A4DR cells, the mRNA levels of *arhA3* and *arhA1* were much lower than those of the wild-type cells, even when the cells were not induced by acenaphthene. In strain A4DRC, the mRNA levels of *arhA3* and *arhA1* were restored to almost the same as those in the wild-type cells. These results strongly suggest that *ArhR* acts as an activator for the transcription of *arhA* genes and that transcription was induced more in the presence of acenaphthene (or its metabolites). By contrast, the transcription of *arhR* was not induced by acenaphthene under the experimental conditions, and it was not affected by the disruption

of *arhR* (Fig. 2c). Therefore, the *arhR* gene appears to be regulated in a different manner from the *arhA* genes, perhaps constitutively.

Determining the transcription start point of the *arhA3* gene

Based on the RT-PCR and quantitative RT-PCR (Fig. 2) results, we predicted the existence of at least one acenaphthene-inducible promoter, which is activated by *ArhR*, in the region upstream from the *arhA3* gene. Therefore, we performed a primer extension analysis to determine the transcription start point of the *arhA3* gene. One transcription start point was found 44 bp upstream from the translation start point of *arhA3* (Fig. 3a, b). The possible promoter sequence was assigned to TTGACG (−35 region) and TATCCG (−10 region) (Fig. 3b). This promoter sequence is similar to the consensus sequence of the *E. coli* σ^{70} -dependent promoter (Record *et al.*, 1996). Upstream from this promoter, between bases −56 and −70, there is a consensus sequence for LTTR binding (ATCACTGTCCGTGAT), which is a palindromic T-N₁₁-A motif (beginning and end underlined) (Schell, 1993) (Fig. 3b). This suggests that *ArhR* binds to this motif and regulates transcription from this promoter region. We did not find any other sequence similar to this motif in the sequenced 16.4 kb DNA region.

DISCUSSION

In this study, we isolated four mutants, each with a mini-Tn5 insert (strains AG2-45, AG2-48, AG3-15 and AG3-69) and defective in the ability to utilize acenaphthene. The mini-Tn5 was inserted within about 16 kb of the same locus in three of the four mutants (strains AG2-45, AG2-48 and AG3-15). The nucleotide sequence analysis of the corresponding 16.4 kb *KpnI* fragment revealed the existence of 16 ORFs and a partial ORF, most of which showed homology to known aromatic-degradative genes or transposase genes

Table 4. Acenaphthene oxygenase activity in recombinant *E. coli* cells expressing the *arhA* genes

<i>E. coli</i> strain	1-Acenaphthenol production (%) [*]		Acenaphthene-1,2-diol [†] production (%) [*]	
	2 h	16 h	2 h	16 h
JM109(pUC18)(pSTV28)	18.4 ± 1.0	16.8 ± 1.1	ND	ND
JM109(pUArhA1A2)(pSTV28)	26.3 ± 0.4	34.5 ± 1.1	0.32 ± 0.02	3.86 ± 0.13
JM109(pUArhA1A2)(pSArhA3)	46.9 ± 1.7	87.5 ± 3.5	8.78 ± 0.20	38.9 ± 0.7
JM109(pUArhA1A2)(pSArhA3A4)	45.4 ± 1.4	100 ± 2.3	9.29 ± 0.35	47.4 ± 1.5

^{*}Calculated from the GC-MS ion peak areas of the MSTFA derivatives of the product formed after 2 h or 16 h of incubation and expressed as percentages of the peak area of 1-acenaphthenol formed by JM109(pUArhA1A2)(pSArhA3A4) after 16 h of incubation. The values are means ± SD calculated from triplicate assays. ND, Not detected.

[†]The MSTFA derivative of this product had a GC retention time of 9.6 min and showed mass spectral peaks at *m/z* 330 (relative abundance 38 %), 315 (23 %), 241 (29 %), 168 (25 %), 147 (100 %). These fragment components were matched with those of the derivatives of acenaphthene-*cis*-1,2-diol or acenaphthene-*trans*-1,2-diol reported by Brooks *et al.* (1982), although their relative abundances were different. We identified acenaphthene-*cis*-1,2-diol as its methanboronate by comparing its mass spectral peaks with those reported by Brooks *et al.* (1982).

Fig. 2. Quantitative RT-PCR analyses of the *arh* genes in strain A4 and its derivatives. The mRNA levels of (a) *arhA3*, (b) *arhA1* and (c) *arhR* in the cells induced (grey bars) or not induced (white bars) with acenaphthene are shown. Cells were induced with 0.02 % (w/v) acenaphthene for 2 h after reaching an OD₆₀₀ of 0.5. Non-induced cells were supplemented with a volume of DMSO equal to that of the acenaphthene solution. The values of the mRNA levels in strain A4, A4DR and A4DRC cells are relative to that of strain A4 cells not induced by acenaphthene (taken as 1.0). The error bars represent the standard deviation calculated from triplicate assays.

(Table 3). The gene organization in this locus was interesting: the ORFs located in the region from *arhA3* to ORF6, all of which are oriented in the same transcription direction (Fig. 1), include a gene not essential for acenaphthene degradation (ORF15) and ORFs apparently not involved in aromatic degradation, such as two putative

pseudogenes (ORF11 and ORF13) and a putative transposase gene or its remnant (ORF16). RT-PCR analysis revealed that the entire region between *arhA3* and ORF6 is actually transcribed in the same direction when strain A4 is grown on acenaphthene (Fig. 1), suggesting that the region from *arhA3* to ORF6 forms an operon. This loose operon structure, including the long regions unnecessary for acenaphthene degradation, seems inefficient in terms of the energy required for transcription, in contrast to the well-studied operons for the aromatic degradation of pseudomonads (e.g. the *nah* and *sal* operons for naphthalene degradation; Dennis & Zylstra, 2004). Apparently, such loose genetic structures have been found in many PAH-degrading sphingomonads (Pinyakong *et al.*, 2003a). However, their actual transcriptional units have not been extensively investigated. Therefore, further study is necessary to prove that such an operon structure is typical for other PAH-degrading sphingomonads.

In this study, ArhR, which is an LTTR that activates the gene cluster from *arhA3*, was also isolated. LTTRs are one of the most common transcriptional regulators in prokaryotes. To

Fig. 3. Determination of the transcription start point of *arhA3* by primer extension. (a) The arrow indicates the primer extension product generated with primer *arhA3*-PE1 (Table 2) annealed to a specific site inside the *arhA3* gene. Lanes G, A, T and C correspond to the sequence ladder generated with the same primer, and the sequence pattern is shown on the right. (b) The positions of the transcription start point and the start codon of the *arhA3* gene are shown. The putative promoter sequences of the -35 and -10 regions are underlined. The putative ArhR-binding site is boxed. The arrows in the box indicate the interrupted palindromic sequences.

date, numerous LTTRs have been reported, and their transcriptional regulation mechanisms have been investigated in detail (Schell, 1993; Tropel & van der Meer, 2004). Usually, LTTRs act as transcriptional activators and increase transcription from target promoters between 6- and 200-fold, only in the presence of an inducer (Schell, 1993). We demonstrated that ArhR also activates transcription of the *arhA* genes in the presence of acenaphthene (or its metabolites) (Fig. 2a, b). Nevertheless, the mRNA levels of *arhA3* and *arhA1* in strain A4 increased only 2.4- and 6.2-fold, respectively, in the presence of acenaphthene. Moreover, in strain A4DR (the *arhR* disruptant), the mRNA levels of *arhA3* and *arhA1* were much lower (less than 10%) than those in strain A4 in the absence of acenaphthene. These facts suggest that ArhR increases some transcription of its target genes even in the absence of an inducer. Although this is unusual for LTTRs, a similar tendency has been observed for the activation of *clcABD* (3-chlorocatechol degradative operon) by CatR, which is the LTTR that activates both *catBCA* (catechol degradative operon) and *clcABD* (McFall *et al.*, 1998). In that study, an *in vitro* transcription assay was used to demonstrate that CatR could activate the *clcABD* promoter in a concentration-dependent manner in the absence of an inducer and that the activation was increased two- to threefold in the presence of inducer (McFall *et al.*, 1998, 1997). This was thought to be due to the binding of CatR to the *clcABD* promoter in its active conformation in the absence of an inducer (McFall *et al.*, 1998). In strain A4, ArhR is thought to activate its target promoter in a manner similar to that of CatR activation of the *clc* promoter. The difference in the fold-increase of the mRNA level in the presence of acenaphthene between *arhA3* (2.4-fold) and *arhA1* (6.2-fold) remains to be explained, but it is possible that another, more inducible promoter(s) exists between these genes. Also, there may be a difference in stability between the mRNA fragments transcribed from *arhA3* and *arhA1*.

In general, the genes for LTTRs are located just upstream from their target metabolic operons and are transcribed in the opposite direction; LTTRs repress their own expression, probably because the binding sites responsible for activating the target operon overlap the promoter of the genes for LTTRs (Schell, 1993). In strain A4, *arhR* is located upstream from *arhA3* in the gene cluster; however, the intergenic region between *arhR* and *arhA3* is about 1 kb in length and includes putative transposase genes or their remnants (ORF8 and ORF9). Therefore, autorepression as with typical LTTRs does not seem likely to occur with the expression of *arhR*. In fact, the transcription of *arhR* was not affected by the disruption of *arhR* under the conditions examined in this study (Fig. 2c). Therefore, the existence of ORF8 and ORF9 may affect the expression of the *arhR* gene.

For acenaphthene degradation in strain A4, ArhR first activates the expression of the *arhA* genes, and then the ArhA oxygenase converts acenaphthene into 1-acenaphthenol and acenaphthene-1,2-diol by two sequential monooxygenations.

In our previous work, it was shown that 1-acenaphthenol can be dehydrogenated into 1-acenaphthenone (Komatsu *et al.*, 1993; Pinyakong *et al.*, 2004). Although the degradation pathway downstream of 1-acenaphthenone and acenaphthene-1,2-diol was not clearly identified in strain A4, Selifonov *et al.* (1996) reported that 1,2-acenaphthoquinone was produced by two sequential monooxygenations and subsequent nonspecific dehydrogenations in resting cells of *Pseudomonas aeruginosa* strain PAO1 that expressed naphthalene dioxygenase genes (*nahA*). Furthermore, they reported that 1,2-acenaphthoquinone was oxidized spontaneously to 1,8-naphthalenedicarboxylic acid (1,8-NDCA), which was identified as its anhydride (1,8-NA). This degradation pathway may also be present in strain A4, because one mini-Tn5-inserted mutant, strain AG3-69, lost the ability to utilize 1,8-NA in addition to acenaphthene, suggesting that acenaphthene is degraded via 1,8-NDCA (or 1,8-NA) in strain A4. There have been no reports on the degradation pathway of 1,8-NDCA, although strain AG3-69 should provide an important clue for clarifying this. In strain AG3-69, the mini-Tn5 was inserted in the DNA region homologous to the *redA2* gene, which encodes the ferredoxin reductase component of the dioxin dioxygenase in *Sphingomonas wittichii* strain RW1 (Armengaud *et al.*, 1998). Therefore, this *redA2* homologue is likely involved in the oxygenation of 1,8-NDCA or its metabolites in combination with other oxygenase components, which have not been identified. We performed a preliminary nucleotide sequence analysis of the flanking region of this *redA2* homologue, but no other genes homologous to known aromatic degradative genes were found (data not shown). This suggests that the genes responsible for the further degradation of acenaphthene metabolites are also dispersed throughout the strain A4 genome, as are the dioxin dioxygenase genes of strain RW1 (Armengaud *et al.*, 1998). Apparently, ArhR is not involved in regulating these genes, because the *arhR* disruptant (strain A4DR) did not lose the ability to utilize 1,8-NA (data not shown). Although some difficulty will arise because of the scattered gene organization that is common among sphingomonads, further isolation of the degradative genes and investigation of their regulatory mechanisms are needed to clarify the entire acenaphthene degradation system in strain A4.

ACKNOWLEDGEMENTS

This study was carried out as a part of 'The Project for Development of Technologies for Analysing and Controlling the Mechanism of Biodegrading and Processing' which was entrusted by the New Energy and Industrial Technology Development Organization (NEDO).

REFERENCES

- Adkins, A. (1999). Degradation of the phenoxy acid herbicide diclofop-methyl by *Sphingomonas paucimobilis* isolated from a Canadian prairie soil. *J Ind Microbiol Biotechnol* **23**, 332–335.

- Alexeyev, M. F., Shokolenko, I. N. & Croughan, T. P. (1995). New mini-Tn5 derivatives for insertion mutagenesis and genetic engineering in gram-negative bacteria. *Can J Microbiol* **41**, 1053–1055.
- Armengaud, J. & Timmis, K. N. (1998). The reductase RedA2 of the multi-component dioxin dioxygenase system of *Sphingomonas* sp. RW1 is related to class-I cytochrome P₄₅₀-type reductases. *Eur J Biochem* **253**, 437–444.
- Armengaud, J., Happe, B. & Timmis, K. N. (1998). Genetic analysis of dioxin dioxygenase of *Sphingomonas* sp. strain RW1: catabolic genes dispersed on the genome. *J Bacteriol* **180**, 3954–3966.
- Bastiaens, L., Springael, D., Wattiau, P., Harms, H., deWachter, R., Verachtert, H. & Diels, L. (2000). Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria using PAH-sorbing carriers. *Appl Environ Microbiol* **66**, 1834–1843.
- Brooks, C. J. W., Cole, W. J., Borthwick, J. H. & Brown, G. M. (1982). Characterization of dihydroarene diols and related compounds by gas chromatography-mass spectrometry: comparison of derivatives. *J Chromatogr* **239**, 191–216.
- Cai, M. & Xun, L. (2002). Organization and regulation of pentachlorophenol-degrading genes in *Sphingobium chlorophenolicum* ATCC 39723. *J Bacteriol* **184**, 4672–4680.
- de Lorenzo, V. & Timmis, K. N. (1994). Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol* **235**, 386–405.
- Demanèche, S., Meyer, C., Micoud, J., Louwagie, M., Willison, J. C. & Jouanneau, Y. (2004). Identification and functional analysis of two aromatic-ring-hydroxylating dioxygenases from a *Sphingomonas* strain that degrades various polycyclic aromatic hydrocarbons. *Appl Environ Microbiol* **70**, 6714–6725.
- Dennis, J. J. & Zylstra, G. J. (2004). Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from *Pseudomonas putida* strain NCIB 9816-4. *J Mol Biol* **341**, 753–768.
- Ferraro, D. J., Gakhar, L. & Ramaswamy, S. (2005). Rieske business: structure-function of Rieske non-heme oxygenases. *Biochem Biophys Res Commun* **338**, 175–190.
- Fredrickson, J. K., Balkwill, D. L., Drake, G. R., Romine, M. F., Ringelberg, D. B. & White, D. C. (1995). Aromatic-degrading *Sphingomonas* isolates from the deep subsurface. *Appl Environ Microbiol* **61**, 1917–1922.
- Gabriel, F. L., Giger, W., Guenther, K. & Kohler, H. P. (2005). Differential degradation of nonylphenol isomers by *Sphingomonas xenophaga* Bayram. *Appl Environ Microbiol* **71**, 1123–1129.
- Gruez, A., Roig-Zamboni, V., Grisel, S., Salomoni, A., Valencia, C., Campanacci, V., Tegoni, M. & Cambillau, C. (2004). Crystal structure and kinetics identify *Escherichia coli* YdcW gene product as a medium-chain aldehyde dehydrogenase. *J Mol Biol* **343**, 29–41.
- Habe, H., Ashikawa, Y., Saiki, Y., Yoshida, T., Nojiri, H. & Omori, T. (2002). *Sphingomonas* sp. strain KA1, carrying a carbazole dioxygenase gene homologue, degrades chlorinated dibenzo-*p*-dioxins in soil. *FEMS Microbiol Lett* **211**, 43–49.
- Hernández, M. J., Reineke, W. & Santero, E. (1999). Genetic analysis of biodegradation of tetralin by a *Sphingomonas* strain. *Appl Environ Microbiol* **65**, 1806–1810.
- Imai, R., Nagata, Y., Senoo, K., Wada, H., Fukuda, M., Takagi, M. & Yano, K. (1989). Dehydrochlorination of γ -hexachlorocyclohexane (γ -BHC) by γ -BHC-assimilating *Pseudomonas paucimobilis*. *Agric Biol Chem* **53**, 2015–2017.
- Khan, A. A., Wang, R. F., Cao, W. W., Franklin, W. & Cerniglia, C. E. (1996). Reclassification of a polycyclic aromatic hydrocarbon-metabolizing bacterium, *Beijerinckia* sp. strain B1, as *Sphingomonas yanoikuyae* by fatty acid analysis, protein pattern analysis, DNA-DNA hybridization, and 16S ribosomal DNA sequencing. *Int J Syst Bacteriol* **46**, 466–469.
- Kiyohara, H., Nagao, K., Kouno, K. & Yano, K. (1982). Phenanthrene-degrading phenotype of *Alcaligenes faecalis* AFK2. *Appl Environ Microbiol* **43**, 458–461.
- Kohler, H. P. (1999). *Sphingomonas herbicidovorans* MH: a versatile phenoxyalkanoic acid herbicide degrader. *J Ind Microbiol Biotechnol* **23**, 336–340.
- Kohler, M., Künnifer, T., Schmid, P., Gujer, E., Crockett, R. & Wolfensberger, M. (2000). Inventory and emission factor of creosote, polycyclic aromatic hydrocarbons (PAH), and phenols from railroad ties treated with creosote. *Environ Sci Technol* **34**, 4766–4772.
- Komatsu, T., Omori, T. & Kodama, T. (1993). Microbial degradation of the polycyclic aromatic hydrocarbons acenaphthene and acenaphthylene by a pure bacterial culture. *Biosci Biotechnol Biochem* **57**, 864–865.
- Lessner, D. J., Parales, R. E., Narayan, S. & Gibson, D. T. (2003). Expression of the nitroarene dioxygenase genes in *Comamonas* sp. strain JS765 and *Acidovorax* sp. strain JS42 is induced by multiple aromatic compounds. *J Bacteriol* **185**, 3895–3904.
- Martínez-Pérez, O., Moreno-Ruiz, E., Floriano, B. & Santero, E. (2004). Regulation of tetralin biodegradation and identification of genes essential for expression of *thn* operons. *J Bacteriol* **186**, 6101–6109.
- McFall, S. M., Klem, T. J., Fujita, N., Ishihama, A. & Chakrabarty, A. M. (1997). DNase I footprinting, DNA bending and *in vitro* transcription analyses of ClcR and CatR interactions with the *clcABD* promoter: evidence of a conserved transcriptional activation mechanism. *Mol Microbiol* **24**, 965–976.
- McFall, S. M., Chugani, S. A. & Chakrabarty, A. M. (1998). Transcriptional activation of the catechol and chlorocatechol operons: variations on a theme. *Gene* **223**, 257–267.
- Miyauchi, K., Lee, H. S., Fukuda, M., Takagi, M. & Nagata, Y. (2002). Cloning and characterization of *linR*, involved in regulation of the downstream pathway for γ -hexachlorocyclohexane degradation in *Sphingomonas paucimobilis* UT26. *Appl Environ Microbiol* **68**, 1803–1807.
- Moreno-Ruiz, E., Hernández, M. J., Martínez-Pérez, O. & Santero, E. (2003). Identification and functional characterization of *Sphingomonas macrogolita* strain TFA genes involved in the first two steps of the tetralin catabolic pathway. *J Bacteriol* **185**, 2026–2030.
- Mueller, J. G., Chapman, P. J., Blattmann, B. O. & Pritchard, P. H. (1990). Isolation and characterization of a fluoranthene-utilizing strain of *Pseudomonas paucimobilis*. *Appl Environ Microbiol* **56**, 1079–1086.
- Pinyakong, O. (2003). *Biochemical and genetic analyses of polycyclic aromatic hydrocarbons degradation in sphingomonads*. PhD thesis, University of Tokyo, Tokyo.
- Pinyakong, O., Habe, H., Supaka, N. & 7 other authors (2000). Identification of novel metabolites in the degradation of phenanthrene by *Sphingomonas* sp. strain P2. *FEMS Microbiol Lett* **191**, 115–121.
- Pinyakong, O., Habe, H. & Omori, T. (2003a). The unique aromatic catabolic genes in sphingomonads degrading polycyclic aromatic hydrocarbons (PAHs). *J Gen Appl Microbiol* **49**, 1–19.
- Pinyakong, O., Habe, H., Yoshida, T., Nojiri, H. & Omori, T. (2003b). Identification of three novel salicylate 1-hydroxylases involved in the phenanthrene degradation of *Sphingobium* sp. strain P2. *Biochem Biophys Res Commun* **301**, 350–357.
- Pinyakong, O., Habe, H., Kouzuma, A., Nojiri, H., Yamane, H. & Omori, T. (2004). Isolation and characterization of genes encoding

polycyclic aromatic hydrocarbon dioxygenase from acenaphthene and acenaphthylene degrading *Sphingomonas* sp. strain A4. *FEMS Microbiol Lett* **238**, 297–305.

Record, M. T., Jr, Reznikoff, W. S., Craig, M. L., McQuade, K. L. & Schlax, P. J. (1996). *Escherichia coli* RNA polymerase ($E\sigma^{70}$), promoters, and the kinetics of the steps of transcription initiation. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, pp. 792–820. Edited by F. C. Neidhardt & others. Washington, DC: American Society for Microbiology.

Saber, D. L. & Crawford, R. L. (1985). Isolation and characterization of *Flavobacterium* strains that degrade pentachlorophenol. *Appl Environ Microbiol* **50**, 1512–1518.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G. & Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69–73.

Schell, M. A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* **47**, 597–626.

Selifonov, S. A., Slepchenkin, A. V., Adanin, V. M., Grechkina, G. M. & Starovoitov, I. I. (1993). Acenaphthene catabolism by strains of *Alcaligenes eutrophus* and *Alcaligenes paradoxus*. *Microbiology* [English translation of *Mikrobiologiya*] **62**, 85–92.

Selifonov, S. A., Grifoll, M., Eaton, R. W. & Chapman, P. J. (1996). Oxidation of naphthenoaromatic and methyl-substituted aromatic compounds by naphthalene 1,2-dioxygenase. *Appl Environ Microbiol* **62**, 507–514.

Shi, T., Fredrickson, J. K. & Balkwill, D. L. (2001). Biodegradation of polycyclic aromatic hydrocarbons by *Sphingomonas* strains isolated from the terrestrial subsurface. *J Ind Microbiol Biotechnol* **26**, 283–289.

Sukchawalit, R., Vattanaviboon, P., Sallabhan, R. & Mongkolsuk, S. (1999). Construction and characterization of regulated L-arabinose-inducible broad host range expression vectors in *Xanthomonas*. *FEMS Microbiol Lett* **181**, 217–223.

Takeuchi, M., Hamana, K. & Hiraishi, A. (2001). Proposal of the genus *Sphingomonas* sensu stricto and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analyses. *Int J Syst Evol Microbiol* **51**, 1405–1417.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Tropel, D. & van der Meer, J. R. (2004). Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol Mol Biol Rev* **68**, 474–500.

White, D. C., Sutton, S. D. & Ringelberg, D. B. (1996). The genus *Sphingomonas*: physiology and ecology. *Curr Opin Biotechnol* **7**, 301–306.

Wise, S. A., Benner, B. A., Burd, G. D., Chester, S. N., Rebbert, R. E. & Schantz, M. M. (1988). Determination of polycyclic aromatic hydrocarbons in a coal tar standard reference material. *Anal Chem* **60**, 887–894.

Wittich, R. M., Wilkes, H., Sinnwell, V., Francke, W. & Fortnagel, P. (1992). Metabolism of dibenzo-*p*-dioxin by *Sphingomonas* sp. strain RW1. *Appl Environ Microbiol* **58**, 1005–1010.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Yoshida, T., Ayabe, Y., Yasunaga, M., Usami, Y., Habe, H., Nojiri, H. & Omori, T. (2003). Genes involved in the synthesis of the exopolysaccharide methanolan by the obligate methylotroph *Methylobacillus* sp. strain 12S. *Microbiology* **149**, 431–444.