Recombinant production of docosahexaenoic acid in a polyketide biosynthesis mode in *Escherichia coli*

Yoshitake Orikasa[1], Takanori Nishida[1], Akiko Yamada[2*], Reiko Yu[2*], Kazuo Watanabe[2], Akira Hase[3], Naoki Morita[4] & Hidetoshi Okuyama[1]**

[1] Graduate School of Environmental Earth Science, Hokkaido University, Sapporo 060-0810, Japan

[2] Sagami Chemical Research Center, Ayase 252-1193, Japan

[3] Biological Laboratory, Hakodate Campus, Hokkaido University of Education, Hakodate 040-8567, Japan

[4] Research Institute of Genome-based Biofactory, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo 062-8517, Japan

*Present addresses: Bio Science Division, Instruments Company, Nicon Corporation, Yokohama 244-8533, Japan (A. Y.) and Mitsubishi Chemical Group Science and Technology Research Center Inc., Yokohama 227-8502, Japan (R. Y.)

**Author for correspondence (Fax: +81 11 707 2347; E-mail: hoku@ees.hokudai.ac.jp)

Key Words: Docosahexaenoic acid, DHA biosynthesis genes, Eicosapentaenoic acid, *Moritella marina* strain MP-1, Polyunsaturated fatty acid

Abstract The docosahexaenoic acid (DHA) biosynthesis gene cluster (pDHA3) from the DHA-producing *Moritella marina* strain MP-1 includes the genes *pfaA*, *pfaB*, *pfaC*, and *pfaD*, which are similar to the genes of polyketide biosynthesis. When this cluster was coexpressed in *Escherichia coli* with *M. marina* MP-1 *pfaE*, which encodes phosphopantetheinyl transferase, DHA was biosynthesized. The maximum production of DHA (5% of total fatty acids) was observed at 15 °C. This is the first report of the recombinant production of DHA in a polyketide biosynthesis mode.

Introduction

Polyunsaturated fatty acids (PUFAs) such as eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) are formed in organisms by a combination of the elongation and desaturation of either C_{16} or C_{18} fatty acids or by enzyme systems consisting of polyketide synthases (PKSs), which form PUFAs from C_2 compounds (possibly acetyl-CoA). Therefore, the recombinant production of PUFAs that are of pharmaceutical and nutritional importance (Bezard et al. 1994, Gill and Valivety 1997) has been attempted using these two systems: reconstruction of the stepwise reaction processes of elongation and desaturation of predominantly C_{18} fatty acids that occur in transgenic plants (Wu et al. 2005, Robert 2006) or the introduction of a single set of PKS genes into host bacteria (Metz et al. 2001, Orikasa et al. 2004).

Recently, significant levels of arachidonic acid, EPA, and DHA were produced in *Brassica juncea* (Wu et al. 2004) and *Arabidopsis thaliana* (Robert 2006), and EPA and DHA in *Glycine max* (Kinney et al. 2004) when the plants were metabolically engineered in a stepwise manner. Using the PKS enzyme system, only the transgenic production of EPA was successfully achieved in *Escherichia coli* (Orikasa et al. 2004, Nishida et al. 2006) and cyanobacteria (Takeyama et al. 1997, Yu et al. 2000). Orikasa et al. (2004) developed various recombinant systems involving clustered polyketide-like synthesis genes (the EPA gene cluster) that had been obtained from EPA-producing *Shewanella pneumatophori* SCRC-2738 (formerly *Shewanella* sp. SCRC-2738 [Hirota et al. 2005]; hereafter called SCRC-2738). Significantly high levels of EPA, comprising more than 20% of the total fatty acids, were thus formed in *E. coli* (Orikasa et al. 2004).

Tanaka et al. (1999) cloned clustered genes (the DHA gene cluster) as a cosmid genomic clone of p3D5, with a 40-kbp insert from *Moritella marina* strain MP-1 (hereafter called MP-1). This DHA gene cluster consisted of four genes, *pfaA*, *pfaB*, *pfaC*, and *pfaD*, which are highly homologous to those included in the EPA gene cluster of SCRC-2738. They are also highly homologous to the corresponding genes of other EPA-producing bacteria (Allen and Bartlett 2002, Gentile et al. 2003) and to those of bacteria expected (but not yet confirmed) to produce DHA (Methe et al. 2005).

However, the DHA gene cluster from MP-1 lacks a gene corresponding to *pfaE*. The domain structure of each *pfa* gene of the DHA gene cluster of MP-1 is as follows: *pfaA* encodes a multifunction protein including domains for 3-ketoacyl synthase (KS), malonyl-CoA:acyl carrier protein (ACP) acyltransferase (AT), five ACP repeats, and 3-ketoacyl–ACP reductase (KR). *pfaC* encodes a protein with domains of two KS repeats, chain-length factor, and two 3-hydroxydecanoyl–ACP dehydratases (HD). A very recent BLAST search indicated that *pfaC* of the DHA gene cluster and also of the EPA gene cluster of SCRC-2738 includes one additional HD region (see Figure 1). Genes *pfaB* and *pfaD* encode proteins with KR and AT domains and an enoyl reductase (ER) domain, respectively (Tanaka et al. 1999).

Although we constructed a high-copy-number plasmid vector (pDHA2) that carried *pfaA–D* by modifying p3D5 of MP-1 and p3D5 and pDHA2 were transformed into *Escherichia coli* already carrying a *pfaE* gene to produce DHA, no DHA was produced. We then identified a single-nucleotide replacement in the *pfaA* gene of the DHA gene cluster in p3D5 and pDHA2, which inactivated the gene. However, a plasmid vector (pDHA3) with the native nucleotide sequence was subsequently constructed. Thus, the recombinant production of DHA was achieved by combining pDHA3 and *pfaE* from MP-1.

Materials and methods

Bacterial strains and culture conditions

Table 1 summarizes the bacterial strains and vectors used in this study. *Escherichia coli* DH5α recombinant cells were cultured with shaking in Luria–Bertani (LB) medium supplemented with the indicated antibiotics at 37 °C for 16 h. An aliquot of the 37 °C-grown precultured *E. coli* DH5α cells carrying one or two vectors was transferred to 3 mL of fresh LB medium in a test tube and then cultured with shaking at 180 rpm at 15 °C for 96 h to allow DHA production. Cells carrying pSTV29 and those carrying pET21a or the pCR2.1®-TOPO® vector (pTOPO, Invitrogen) were cultured in LB

medium containing chloramphenicol at 30 μg/mL and ampicillin at 50 μg/mL, respectively. When the temperature-dependent production of DHA was investigated, the recombinant *E. coli* DH5α cells were cultured in 10 mL of LB medium in a 50 mL flask at 180 rpm at 10 °C, 15 °C, 20 °C, or 25 °C. The cells were harvested at mid to late exponential phase and subjected to fatty-acid analysis. *Moritella marina* strain MP-1 (ATCC 15381) was cultured in LB medium containing 3.0% NaCl at 15 °C for 96 h.

PCR, plasmid construction, and nucleotide determination

A cosmid genomic clone of p3D5 with 40 kbp from MP-1, consisting of at least 22 open reading frames (ORFs), included the four genes *pfaA–D* (see Figure 1 and Tanaka et al. 1999), which are considered necessary for the biosynthesis of DHA in this bacterium. To remove ORFs other than those of *pfaA–D*, p3D5 was treated with *Bam*HI, producing a DNA fragment of approximately 11.6 kbp that included *pfaB–D* and partial sequences of *pfaA* and ORF 12. This fragment was cloned into the *Bam*HI site of pBluescript IISK+ (clone pBSK/Bam-Bam). The pBSK/Bam-Bam clone was treated with *Sac*I and *Kpn*I, producing a DNA fragment of approximately 11.7 kbp that encoded *pfaB–D* and partial sequences of *pfaA* and ORF 12. This fragment was subcloned into the *Sac*I–*Kpn*I sites of pSTV29 (clone pSTV29/Sac-Kpn). p3D5 was then treated with *Xho*I. The resulting DNA fragment of approximately 11.9 kbp, which included *pfaB* and partial sequences of *pfaC* and *pfaA*, was subcloned into the *Xho*I site of pSTV29/Sac-Kpn (clone pSTV29/Sac-Xho). p3D5 was treated with *Kpn*I, producing a fragment of approximately 4 kbp that included ORFs 5–7 and partial sequences of ORF 4 and *pfaA*. This fragment was subcloned into the *Kpn*I site of pSTV29/Sac-Xho to produce a pSTV29 vector that included an insert of approximately 20 kbp encoding *pfaA–D*, three unidentified ORFs (ORFs 5–7; see Fig. 1), and two truncated ORFs (ORFs 4 and 12). This clone was designated pDHA2.

A comparison of our nucleotide sequence (AB025342) for the whole DHA gene cluster from MP-1 with the corresponding sequence reported by Facciotti et al. (1998) indicated that there was a disagreement in one nucleotide of the *pfaA* gene region in the p3D5 and pDHA2 sequences.

Resequencing the 250-bp region of the genomic DNA of MP-1, including the contentious sequence, revealed the base replacement of "C" with "T" (see below) at nucleotide 1,190 of the *pfaA* gene in p3D5 and in its derivative, pDHA2. The vector carrying the corrected sequence of the *pfaA* gene was prepared as follows. The region including the replacement nucleotide, between restriction sites *Pme*I and *Mlu*I within the *pfaA* gene, was amplified by PCR using the forward and reverse primers 5′-GGTGAAGGTATTGGCATGAT-3′ and 5′-ATCTTCACGTGGCATCCAAG-3′, and the genomic DNA of MP-1 as the template. The resulting DNA fragment (pfaApm) of 484 bp was cloned into the pTOPO vector (pTOPO::pfaApm). pTOPO::pfaApm was treated with *Pme*I and then with *Mlu*I. The digests were purified by electrophoresis on agarose gels. The resulting DNA fragment of 450 bp was ligated to *Pme*I- and *Mlu*I-digested pDHA2, producing pDHA3. pDHA3 is a pSTV29 vector carrying the complete sequence of *pfaA–D*.

The phosphopantetheinyl transferase (PPTase) gene (*pfaE*) of MP-1, which had been integrated into pET21a (pET21a::*pfaE*), was prepared as described previously (Orikasa et al. 2006). The chromosomal DNA of MP-1 was isolated according to the procedure described previously (Tanaka et al. 1999). Determination of the nucleotide sequences and their analysis were carried out as described previously (Orikasa et al. 2006).

GLC and GC-MS analyses of fatty acids

Fatty acids were analyzed as their methyl esters (FAMEs), as described previously (Orikasa et al. 2006). To identify DHA, total FAMEs were converted to their pyrrolidide adducts and then subjected to gas chromatography/mass spectrometry (GC/MS), as described previously (Orikasa et al. 2006).

Results and discussion

Recombinant production of DHA using pDHA3

GC analysis of the total FAMEs of the recombinant *E. coli* DH5α cells carrying pDHA3 and pET21a::*pfaE*, grown at 15 °C, showed the occurrence of a peak with a retention time of 31.2 min, which was not detected in the control *E. coli* DH5α cells carrying no vector. The retention time of this unknown component was the same as that of authentic DHA (data not shown). GC/MS analysis of the pyrrolidide derivative of the component detected an $[M+H]^+$ ion at *m*/*z* 382. A series of ions at *m*/*z* 113, 126, 138, 152, 166, 178, 192, 206, 218, 232, 246, 258, 272, 286, 298, 312, 326, 338, 352, and 366 showed the positions of double bonds at Δ4, 7, 10, 13, 16, and 19, suggesting that this fatty acid was DHA (Watanabe 2000). These results indicate that this component was DHA and that it comprised up to 3.0% (w/w) of the total fatty acids when the cells were cultivated at 15 °C in test tubes.

As shown in Figure 1, pDHA3 included the complete sequences of ORFs 5, 6, and 7 and the partial sequences of ORFs 4 and 12, in addition to *pfaA–D*. Although the functions of ORFs 5–7 have not been identified, as was the case in the biosynthesis of EPA (Orikasa et al. 2004), only the genes *pfaA–E* would be required for the biosynthesis of DHA.

Effects of growth temperature on levels of DHA

The production of DHA in *E. coli* DH5α cells carrying pDHA3 and pET21a::*pfaE* was examined at various temperatures. The levels of DHA in cells cultured in flasks at 10 °C, 15 °C, 20 °C, and 25 °C were 3.0%, 5.2%, 1.3%, and 0% of total fatty acids, respectively. This trend in the recombinant production of DHA in *E. coli* does not concur with that in *M. marina* strain MP-1, where the level of DHA increased with decreasing growth temperature (DeLong and Yayanos 1986) and much more DHA was produced: 17.9%, 8.6%, and 3.2% of total fatty acids at 2.0 °C, 15 °C, and 20 °C, respectively.

It should be noted that more DHA was produced in cells cultured in flasks (5% of total fatty acids) than in those cultured in test tubes (3% of total fatty acids) at the same temperature (15 °C). Because it is expected that the liquid medium in a flask is better aerated than that in a test tube, the concentration of dissolved oxygen could have affected the recombinant production of DHA.

According to Orikasa et al. (2004), higher levels (approximately 16%) of EPA were produced when *E. coli* cells were transformed with high-copy-number plasmid (pNEB) carrying the EPA gene cluster compared with those produced by cells transformed with a cosmid vector carrying *pfaA–E*. Much higher levels of EPA (approximately 22%) were produced by the cotransformation of a plasmid vector carrying a *pfaE*-deficient EPA gene cluster and another high-copy-number vector carrying *pfaE* (Orikasa et al. 2004). That result suggested that the relative ratio of PPTase protein to the other EPA biosynthesis enzymes is a significant factor in the production of EPA (Orikasa et al. 2004). The level of DHA in the present study was approximately 5% of the total fatty acids in the recombinant *E. coli* DH5α cells coexpressing pDHA3 and *pfaE* from MP-1 in flasks at 15 °C (Table 2). However, this value could be increased by the appropriate selection and combination of the two plasmid vectors and/or host organisms, and optimized culture conditions.

Although the domain structures of the clustered genes responsible for the biosynthesis of EPA and DHA are quite similar, there are some differences between the *pfaA–D* genes of MP-1 and those of SCRC-2738. The former has five ACP repeats in *pfaA* (Tanaka et al. 1999), whereas the latter has six repeats (Orikasa et al. 2004). However, this would not determine the final production of EPA or DHA, because the *pfaA* gene of EPA-producing *Photobacterium profundam* SS9 has five ACP repeats (Allen and Bartlett, 2002). Interestingly, *pfaB* of MP-1 has a KS domain, which is not present in the corresponding *pfaB* gene of SCRC-2738. The presence of one additional KS domain in the DHA gene cluster might be associated with the elongation of the C_{20} to the C_{22} fatty acid and ultimately with the biosynthesis of DHA.

The use of PKS genes to produce EPA or DHA in heterologous host organisms has various benefits. 1) The PKS system has a metabolic advantage. It requires less reducing equivalents, such as NADPH, to form EPA or DHA, because the double bonds formed during the growth of the acyl intermediates are left in the chains (Ratledge 2004). 2) The procedure used to produce transgenic organisms is simple. The recombinant production of EPA, at levels higher than 20% of the total fatty acids, can be achieved by introducing only one or two vectors carrying *pfa* genes into host cells (see Orikasa et al. 2004). For the production of DHA, the combination of two plasmid vectors carrying either *pfaA–D* or

pfaE is required at present. Inclusion of *pfaA–E* in one high-copy-number plasmid vector should result in much greater production of this fatty acid in prokaryotic host organisms. 3) The EPA and DHA gene clusters are exclusively involved in producing EPA and DHA, respectively. The compositional simplicity of unsaturated fatty acids is a great advantage for their commercial use.

References

Allen EE, Bartlett DH (2002) Structure and regulation of the omega-3 polyunsaturated fatty acid synthase genes from the deep-sea bacterium *Photobacterium profundum* strain SS9. Microbiology 148: 1903-1913.

Bezard J, Blond JP, Bernard A, Clouet P (1994) The metabolism and availability of essential fatty acids in animal and human tissues. Reprod Nutr Dev 34: 539−568.

DELong EF, Yayanos AA (1986) Biochemical function and ecological significance of novel bacterial lipids in deep-sea prokaryotes. Appl Environ Microbiol 51: 730–737.

Facciotti D, Metz JG, Lassner M (1998) Production of polyunsaturated fatty acids by expression of polyketide-like synthesis genes in planta. United States Patent 6,140,486.

Gentile G, Bonasera V, Amico C, Giuliano L, Yakimov MM (2003) *Shewanella* sp. GA-22, a psychrophilic hydrocarbonoclastic antarctic bacterium producing polyunsaturated fatty acids. J Appl Microbiol 95: 1124–1133.

Gill I, Valivety R (1997) Polyunsaturated fatty acids, part 1: Occurrence, biological activities and applications. Trends Biotechnol 15: 401−409.

Hirota K, Nodasaka Y, Orikasa Y, Okuyama H, Yumoto I (2005) *Shewanella pneumatophoruse* sp. nov., eicosapentanoic-acid-producing marine bacterium isolated from pacific mackerel (*Pneumatophorus japonicus*) intestine. Int J Syst Evol Microbiol 55: 2355-2359.

Kinney AJ, Cahoon EB, Damude HG, Hitz WD, Kolar CW, Liu ZB (2004) Production of very long chain polyunsaturated fatty acids in oilseed plants. EI Du Pont de Nemours WO 2004/071467 A2

Methe BA, Nelson KE, Deming JW, Momen B, Melamud E, Zhang X, Moult J, Madupu R, Nelson WC, Dodson RJ, Brinkac LM, Daugherty SC, Durkin AS, DeBoy RT, Kolonay JF, Sullivan SA, Zhou L, Davidsen TM, Wu M, Huston AL, Lewis M, Weaver B, Weidman JF, Khouri H, Utterback TR, Feldblyum TV, Fraser CM (2005) The psychrophilic lifestyle as revealed by the

genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. Proc Natl Acad Sci USA 102: 10913-10918.

Metz JG, Roessler P, Facciotti D, Levering C, Dittrich F, Lassner M, Valentine R, Lardizabal K, Domergue F, Yamada A, Yazawa K, Knauf V, Browse J (2001) Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes. Science 293: 290-293.

Nishida T, Orikasa T, Itoh Y, Yu R, Yamada A, Watanabe K, Okuyama H (2006) *Escherichia coli* engineered to produce eicosapentaenoic acid becomes resistant against oxidative damages. FEBS Lett 580: 2731-2735.

Orikasa Y, Yamada A, Yu R, Itoh Y, Nishida T, Yumoto I, Watanabe K, Okuyama H (2004) Characterization of the eicosapentaenoic acid biosynthesis gene cluster from *Shewanella* sp. strain SCRC-2738. Cell Mol Biol 50: 625–630.

Orikasa Y, Nishida T, Hase A, Watanabe K, Morita N, Okuyama H (2006) A phosphopantetheinyl transferase gene essential for biosynthesis of n-3 polyunsaturated fatty acids from *Moritella marina* strain MP-1. FEBS Lett (in press)

Ratledge C (2004) Fatty acid biosynthesis in microorganisms being used for single cell oil production. Biochimie 86: 807–815.

Robert SS (2006) Production of eicosapentaenoic and docosahexaenoic acid-containing oils in transgenic land plants for human and aquaculture. Nutrition Marine Biotechnol 8: 103–109

Takeyama H, Takeda D, Yazawa K, Yamada A, Matsunaga T. (1997) Expression of the eicosapentaenoic acid synthesis gene cluster from *Shewanella* sp. in a transgenic marine cyanobacterium, Synechococcus sp. Microbiology 143: 2725–2731

Tanaka M, Ueno A, Kawasaki K, Yumoto I, Ohgiya S, Hoshino T, Ishizaki K, Okuyama H, Morita N (1999) Isolation of clustered genes that are notably homologous to the eicosapentaenoic acid biosynthesis gene cluster from the docosahexaenoic acid-producing bacterium *Vibrio marinus* strain MP-1. Biotechnol Lett 21: 939–945.

Watanabe, K. (2000) Isolation and lipid-biochemical study of polyenoic acid-producing bacteria. Doctoral thesis, University of Tokyo, Tokyo.

Wu G, Truksa M, Datla N, Vrinten P, Bauer J, Zank T, Cirpus P, Heinz E, Qiu X (2005) Stepwise engineering to produce high yields of very long-chain polyunsaturated fatty acids in plants. Nature Biotechnol 23: 1013–1017.

Yu R, Yamada A, Watanabe K, Yazawa K, Takeyama H, Matsunaga T, Kurane R. (2000) Production of eicosapentaenoic acid by a recombinant marine cyanobacterium, *Synechococcus* sp. Lipids 35: 1061–1064

Table 1. Strains and vectors used in this study.

Strain/plasmid/cosmid	Relevant characteristics	Source
Strain		
Escherichia coli		
DH5α	*deo*R, *end*A1, *gyr*A96, *hsd*R17(rK$^-$ mK$^+$), *rec*A1 *pho*A, *rel*A1, *thi*-1, Δ(*lac* ZYA-*arg*F) U169ϕ80d*lac*ZΔM15, F$^-$, λ^-, *sup*E44	Takara Bio[a]
Moritella marina strain MP-1	Wild type	ATTC 15381
Plasmids/ Cosmids		
pCR2.1-TOPO	Cloning vector	Takara Bio
pSTV29	Cloning vector	Takara Bio
pBluescript II SK+	Cloning vector	Stratagene[b]
pET21a	Cloning/expression vector	Takara Bio
pET21a::*pfaE*	pET21a carrying *pfaE* from MP-1	Orikasa et al. 2006
p3D5	Lorist6 carrying *pfaA–D* from genome of *M. marina* strain MP-1	Tanaka et al. 1999
pDHA2	pSTV29 carrying *pfaA–D*[c]	This work
pDHA3	pSTV29 carrying *pfaA–D*[d]	This work

[a] Takara Bio Inc., Tokyo, Japan

[b] Stratagene, LA Jolla, CA, USA

[c] *pfaA* of this clone has a base replacement.

[d] *pfaA* of this clone has native sequence.

Figure legends

Fig. 1. Schematic representation of the DHA gene clusters of *M. marina* strain MP-1 and the EPA gene cluster of *S. pneumatophori* SCRC-2738. p3D5 and pDHA2 include a one-nucleotide substitution in *pfaA*, but this replacement has been corrected in pDHA3 (see text). ACP, acyl carrier protein; AT, malonyl-CoA:ACP acyltransferase; CLF, chain length factor; ER, enoyl reductase; HD, 3-hydroxydecanoyl-ACP dehydratases; KR, 3-ketoacyl-ACP reductase; KS, 3-ketoacyl-ACP synthase; PP, phosphopantetheinyl transferase.

Fig. 2. Gas chromatogram (A) of total fatty acid methyl esters prepared from *E. coli* DH5α carrying pDHA3 and pET21a::*pfaE* (upper panel) and from *E. coli* DH5α carrying no vector (lower panel). Mass spectrum (B) of a pyrrolidide derivative of the unknown peak with a retention time of 31.2 min, seen in (A). Cells were grown in test tubes at 15 °C.

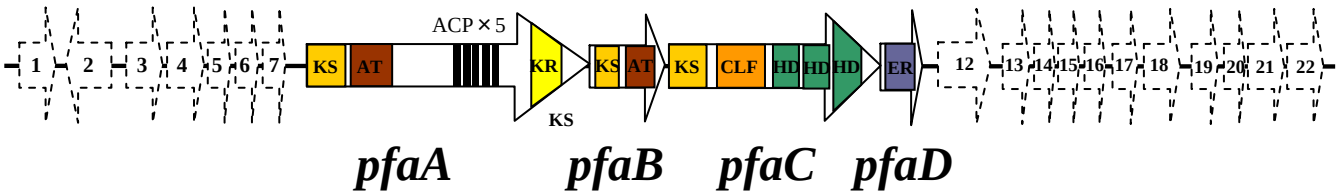
Fig. 3. Effects of growth temperature on the production of DHA by *E. coli* DH5α carrying pDHA3 and pET21a::*pfaE*. Cells were cultured in flasks at various temperatures, up to mid to late exponential phase. The data are means ± S.E. for three independent experiments.

1

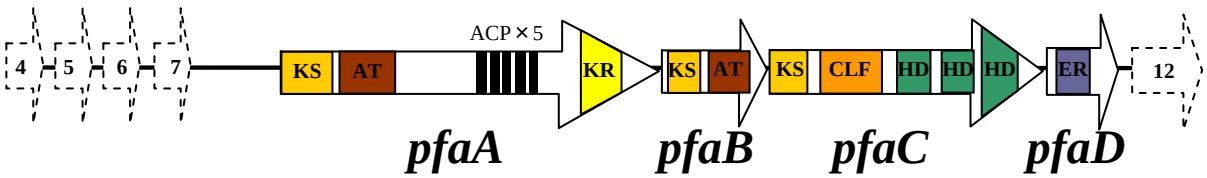
2

DHA gene cluster of *M. marina* strain MP-1

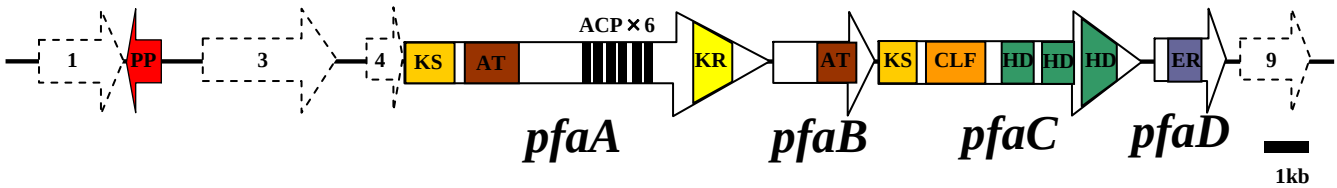
p3D5



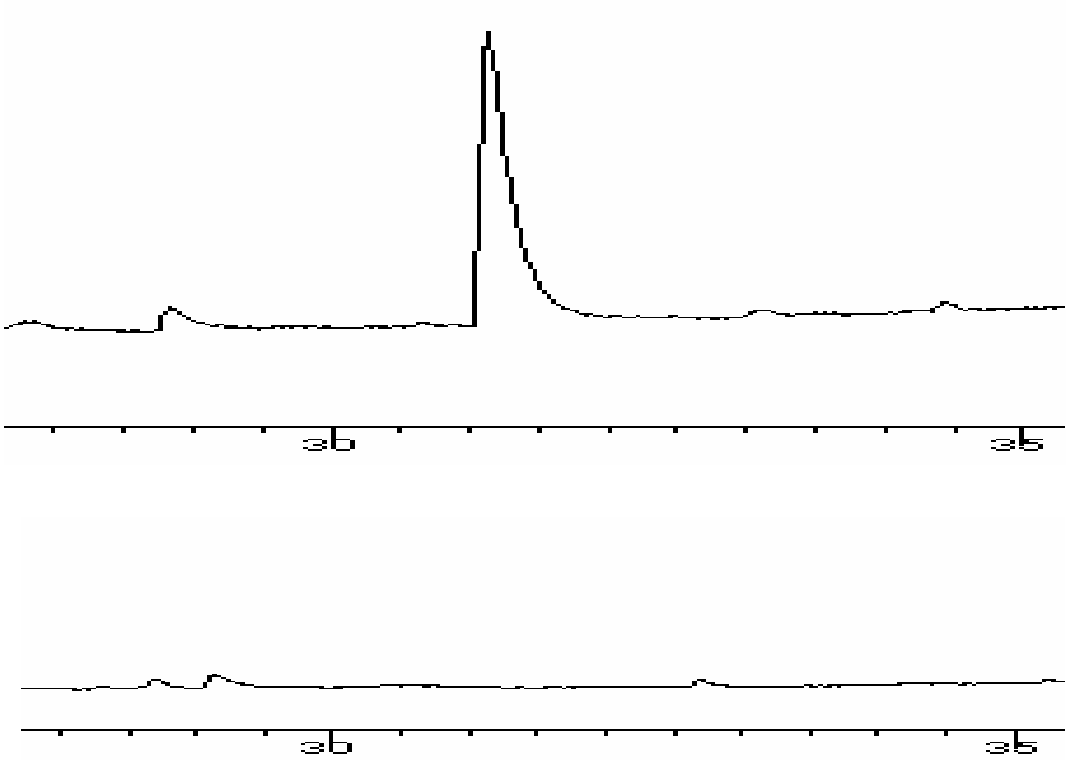
pDHA2, pDHA3



EPA gene cluster of *S. pneumatophori* SCRC-2738



A


35
35
35
35
Retention time (min)

B

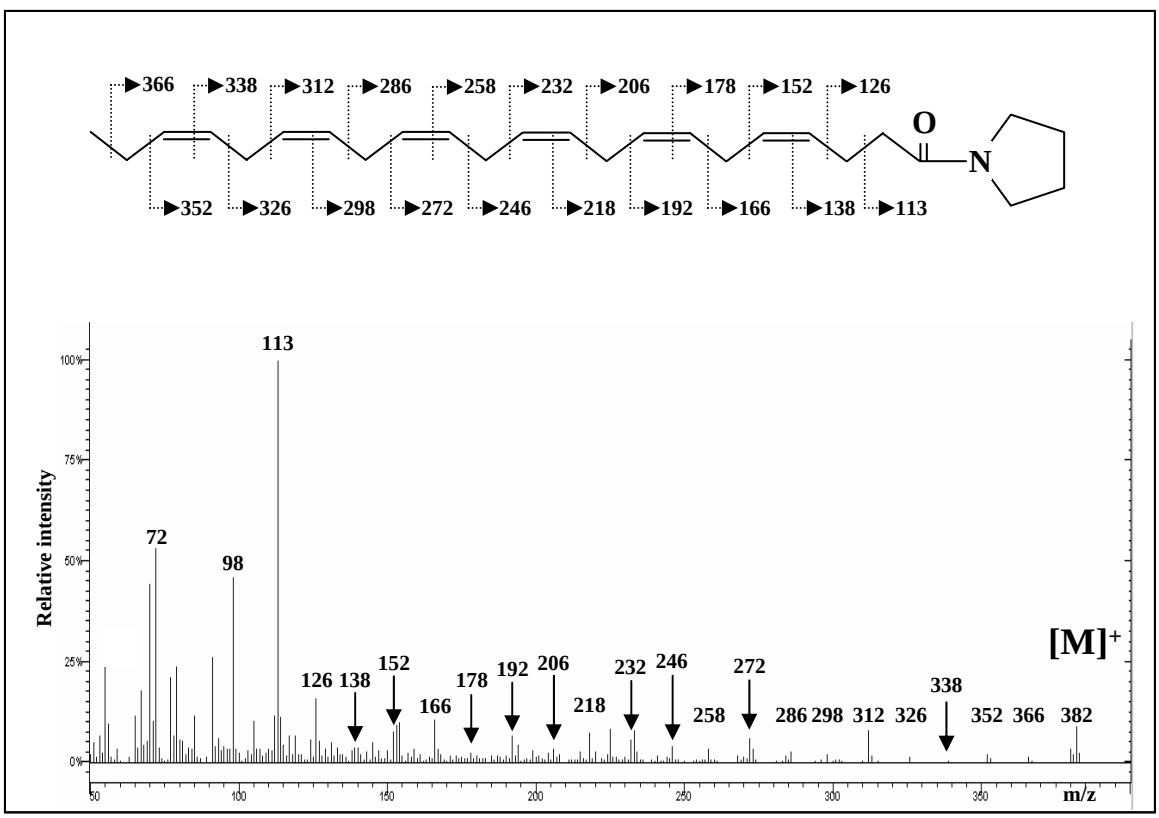

366
338
312
286
258
232
206
178
152
126
352
326
298
272
246
218
192
166
138
113
O
N
Relative intensity
113
72
98
126
138
152
166
178
192
206
218
232
246
258
272
286
298
312
326
338
352
366
382
[M]+
m/z

Fig. 3