FEBS Letters 588 (2014) 4032–4036

journal homepage: www.FEBSLetters.org

Identification of a novel type of polyunsaturated fatty acid synthase involved in arachidonic acid biosynthesis

Tetsuro Ujihara [a,b,*], Megumi Nagano [a], Hajime Wada [b], Satoshi Mitsuhashi [a]

[a] Kyowa Hakko Bio Co. Ltd., Bioprocess Development Center, Tsukuba, Ibaraki 305-0841, Japan
[b] Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Komaba 3-8-1, Meguro-ku, Tokyo 153-8902, Japan

ARTICLE INFO

Article history:
Received 14 August 2014
Revised 28 August 2014
Accepted 16 September 2014
Available online 26 September 2014

Edited by Felix Wieland

Keywords:
Arachidonic acid
pfa gene
Polyunsaturated fatty acid
Polyunsaturated fatty acid synthase
Aureispira marina

ABSTRACT

Arachidonic acid (ARA) is a polyunsaturated fatty acid (PUFA) and an essential component of membrane lipids. However, the PUFA synthase required for ARA biosynthesis has not been identified in any organism. To identify the PUFA synthase producing ARA, we determined the draft genome sequence of the marine bacterium *Aureispira marina*, which produces a high level of ARA, and found a gene cluster encoding a putative PUFA synthase for ARA production. Expression of the gene cluster in *Escherichia coli* induced production of ARA, demonstrating that the gene cluster encodes a PUFA synthase required for ARA biosynthesis.

1. Introduction

Polyunsaturated fatty acids (PUFAs) such as docosahexaenoic acid [DHA, 22:6(4,7,10,13,16,19)], eicosapentaenoic acid [EPA, 20:5(5,8,11,14,17)], and arachidonic acid [ARA, 20:4(5,8,11,14)] are essential components of membrane lipids and used as biosynthetic precursors of signaling molecules [1–4]. Two major pathways are known for PUFA biosynthesis. One is the aerobic pathway, in which saturated fatty acids synthesized by fatty acid synthase are desaturated and elongated by desaturases and elongases, respectively, to PUFAs [5–7]. The other is the anaerobic pathway, in which PUFAs are anaerobically synthesized from acetyl-CoA by PUFA synthases rather than multiple desaturation and elongation reactions [8,9]. In comparison with the aerobic pathway, the anaerobic pathway has some benefits for production of PUFAs [9]. The anaerobic pathway requires fewer reducing equivalents such as NADPH and produces simple fatty acids consisting solely of a specific fatty acid such as EPA and DHA.

PUFA synthases are composed of four to five separate multidomain proteins. Each protein has one to several domains responsible for PUFA synthesis. A PUFA synthase capable of synthesizing EPA was first found in *Shewanella* sp. [8,10], and then a PUFA synthase responsible for DHA biosynthesis was identified in the marine bacterium *Moritella marina* [11]. Although bioinformatic analyses of genomic data deposited in the databases have suggested PUFA synthases that may be involved in biosynthesis of other PUFAs other than EPA and DHA [12], the PUFA synthase required for ARA biosynthesis has not been identified in any organism. Because ARA is an omega-6 PUFA different from omega-3 PUFAs such as EPA and DHA, comparing PUFA synthases that produce ARA with those that produce EPA and DHA is important to understand the mechanism of how fatty acids are elongated and double bonds are inserted during PUFA biosynthesis.

Aureispira marina is a bacterium that belongs to the phylum Bacteroides. This bacterium was isolated from a marine sponge and algae collected from the southern coastline of Thailand and produces 30–40% ARA of total fatty acids [13]. However, how this bacterium produces ARA is still uncertain.

In this study, we identified the *A. marina* draft genome sequence and found a candidate gene cluster encoding a PUFA synthase for

Abbreviations: ACP, acyl-carrier protein; ARA, arachidonic acid; AT, acyltransferase; DHA, docosahexaenoic acid; EPA, eicosapentaenoic acid; ER, enoyl-ACP reductase; GC, gas chromatography; GC–MS, GC–mass spectrometry; DH, 3-hydroxyacyl-ACP dehydratase; FAME, fatty acid methyl ester; IPTG, isopropyl β-D-1-thiogalactopyranoside; KR, 3-ketoacyl-ACP reductase; KS, 3-ketoacyl-ACP synthase; LB, Luria–Bertani; MAT, malonyl-CoA:ACP transacylase; PCR, polymerase chain reaction; ppTase, phosphopantetheinyl transferase; PUFA, polyunsaturated fatty acid; X:Y(z), fatty acid containing X carbon atoms with Y double bonds in the *cis* configuration at position Z counted from the carboxy terminus

* Corresponding author at: Kyowa Hakko Bio Co. Ltd., Bioprocess Development Center, Tsukuba, Ibaraki 305-0841, Japan. Fax: +81 29 856 4122.
E-mail address: tetsuro.ujihara@kyowa-kirin.co.jp (T. Ujihara).

http://dx.doi.org/10.1016/j.febslet.2014.09.023
0014-5793/

ARA biosynthesis. Expression of the gene cluster in *Escherichia coli* resulted in the production of ARA, demonstrating that the gene cluster is involved in ARA biosynthesis. This is the first identification of a PUFA synthase for ARA production, and the information obtained in this study will provide a clue to understanding the mechanism of how ARA is synthesized by the anaerobic pathway and catalyzed by a PUFA synthase.

2. Materials and methods

2.1. Bacterial strains

A. marina (JCM23201) was obtained from the RIKEN Bioresource Center. The *E. coli* strains BW25113 (*lacIq*, *rrnBT14*, *ΔlacZWJ16*, *hsdR514*, *ΔaraBADAH33*, *ΔrhaBADLD78*) and a derivative of BW25113 (BW25113 *ΔfadE*::FRT) were obtained from the Keio Collection [14].

2.2. Gene identification

A. marina (JCM23201) genomic DNA was extracted from cells using a Wizard Genomic DNA Extraction kit (Promega, Madison, WI, USA) according to the manufacturer's instructions.

The nucleotide sequence of the genomic DNA was determined using a 454 GS FLX sequencer [15]. In total, 609 397 reads, including up to 196 000 700 bp were obtained, which represented a 33-fold coverage of the genome. Assembly was performed using the GS de novo assembler software program 454 and Velvet ver. 0.7.63, algorithms for de novo short read assembly [16]. Finally, we obtained the *A. marina* draft genome of 5.18 Mbp distributed in 103 contigs with a GC content of 37%. A putative PUFA gene cluster was identified using the NCBI BLAST tool. A local TBLASTN search of *pfa* genes known to produce EPA and DHA was performed against the draft genome sequence. A subsequent TBLASTN search of individual domains (KS, MAT, ACP, KR, DH, and ER) was performed. An additional round of sequencing using the Sanger method was performed for gap closure. Polymerase chain reaction (PCR) products were subjected to cycle sequencing with ABI BigDye Terminator v3.1 and analyzed on a 3730XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA). A domain analysis of each *pfa* gene was carried out using the Pfam database [17]. The nucleotide sequence of the region including the gene cluster was deposited in DDBJ/EMBL/GenBank nucleotide sequence databases with accession number AB980240.

2.3. Vector construction

The *A. marina pfaABCDE* gene cluster was expressed in *E. coli* (BW25113 *ΔfadE*::FRT) using three compatible vectors; pUC19 [18] for *pfaE* expression, pSTV29 (Takara Bio, Otsu, Shiga, Japan) for *pfaA* and *pfaB*, and pMW219 (Nippon Gene, Tokyo, Japan) for *pfaC* and *pfaD*. The pUC19-*Plac-pfaE*, pSTV29-*Plac-pfaAB*, and pMW219-*Plac-pfaCD* plasmids containing one or two genes were transformed into *E. coli* cells to express the *pfa* genes. The *E. coli* cells were transformed by electroporation according to the method of Miller and Nickoloff [19]. The coding region of each gene was placed downstream of the *E. coli lac* promoter for inducible expression. The procedure to construct these expression vectors is described in the Supplemental Materials and methods.

2.4. Culture conditions

A. marina was cultivated at 25 °C on Sap2 medium [0.5 × artificial seawater (1.5% NaCl, 0.035% KCl, 0.54% $MgCl_2$, 0.27% $MgSO_4$, and 0.05% $CaCl_2$), 0.1% tryptone, 0.1% yeast extract, and 1.5% agar] [13]. Transformants of *E. coli* cells were cultivated by shaking in Luria–Bertani (LB) medium supplemented with 50 mg/L ampicillin, 10 mg/L kanamycin, and 20 mg/L chloramphenicol. A portion of the *E. coli* cells precultured at 30 °C was inoculated into fresh LB medium supplemented with antibiotics and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and then cultivated at 20 °C for 40 h. After the incubation, the *E. coli* cells were collected and used for lipid analysis.

2.5. Lipid analysis

Total lipids were extracted from *E. coli* by the method of Bligh and Dyer [20]. The extracted lipids were separated on thin-layer chromatography plates with a solvent system (chloroform/methanol:acetic acid, 65:25:5, by volume). The isolated lipids were subjected to methanolysis with 5% HCl in methanol at 85 °C for 2.5 h. The resulting fatty acid methyl esters (FAMEs) were analyzed by gas chromatography (GC; GC-2010; Shimadzu, Tokyo, Japan) with a flame-ionization detector and a capillary column (ULBON HR-SS-10, 0.25 mm in diameter and 25 m in length; Shinwa Kako, Bangkok, Thailand). Column temperature was maintained at 190 °C for 30 min. The flow rate of the carrier gas (nitrogen) was 1.35 ml min^{-1}. FAMEs were identified by comparing the retention time of each FAME with those of authentic standards and by analyzing each FAME by gas chromatography–mass spectrometry (GC–MS; GCMS-2010; Shimadzu). The positions of the double bonds in the FAMEs were determined by analyzing the pyrrolidide derivatives of FAMEs by GC–MS as described by Andersson and Holman [21]. A ULBON HR-52 capillary column (0.25 mm in diameter and 25 m in length; Shinwa Kako) was used for the GC–MS analysis. Column temperature was kept at 220 °C for 30 min. The flow rate of the carrier gas (helium) was 1.03 ml min^{-1}. Mass spectra were scanned every 0.3 s.

3. Results and discussion

3.1. Identification of a candidate PUFA synthase gene cluster involved in ARA biosynthesis

To identify a PUFA synthase gene cluster involved in ARA biosynthesis, we searched a gene cluster that shows similarities to a *S. oneidensis* MR-1 gene cluster for EPA biosynthesis [22] in the draft genome sequence of *A. marina*. We found a gene cluster containing five open reading frames (Fig. 1), which were homologous to the *pfaE*, *pfaA*, *pfaB*, *pfaC*, and *pfaD* genes of *S. oneidensis* MR-1. As described below, these genes encode components of the PUFA synthase responsible for ARA biosynthesis. Thus, we named these genes *pfaE*, *A*, *B*, *C*, and *D* (GenBank, accession number AB980240).

Domain analyses of the proteins encoded in the isolated *pfa* gene clusters revealed many enzyme and acyl-carrier protein (ACP) domains similar to PUFA synthases, which are responsible for EPA and DHA biosynthesis. As described by Lee et al. [22], the PUFA synthase of *S. oneidensis* MR-1 for EPA production is encoded in a gene cluster consisting of five genes (Fig. 1). The encoded proteins contained 10 enzyme domains and six repeated ACP domains: a 3-ketoacyl-ACP synthase (KS), a malonyl-CoA:ACP transacylase (MAT), six repeat ACP and a 3-ketoacyl-ACP reductase (KR) domains in PfaA; an acyltransferase (AT) domain in PfaB, a KS, a chain length factor (CLF), and two repeat 3-hydroxyacyl-ACP dehydratase (DH) domains in PfaC; an enoyl-ACP reductase (ER) domain in PfaD, and a phosphopantetheinyl transferase (ppTase) domain in PfaE. Although the structure of the gene cluster and domain organization of the *S. oneidensis* MR-1 encoded proteins for EPA production were similar to those of *M. marina* MP-1 for DHA production, there were some differences between the two

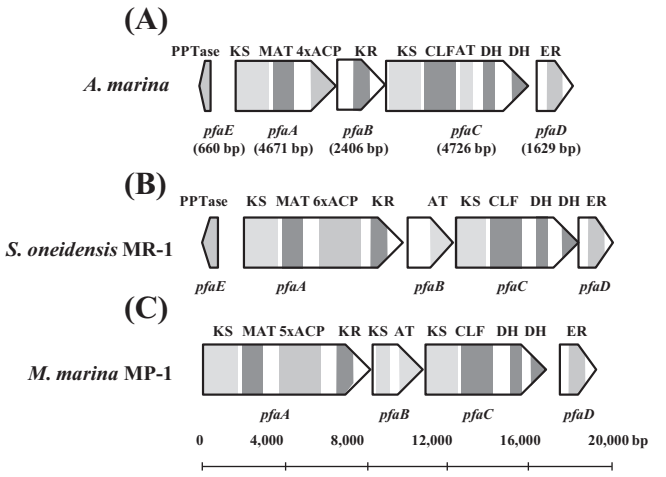

Fig. 1. Organization of the gene clusters for polyunsaturated fatty acid (PUFA) synthases. (A) A gene cluster responsible for arachidonic acid (ARA) biosynthesis identified in *Aureispira marina* in this study; (B) a *Shewanella oneidensis* MR-1 gene cluster responsible for eicosapentaenoic acid (EPA) biosynthesis [22]; (C) a *Moritella marina* MP-1 gene cluster responsible for docosahexaenoic acid (DHA) biosynthesis [9].

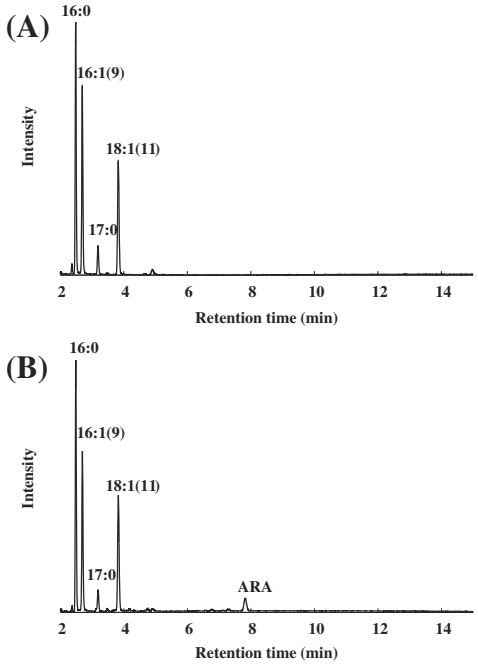

Fig. 2. Gas chromatograms of fatty acid methyl esters prepared from total lipids of *Escherichia coli* (BW25113 *ΔfadE*::FRT) transformant cells. (A) Transformant with empty vectors, pUC19, pSTV29, and pMW219; (B) transformant with pUC19-*Plac-pfaE*, pSTV29-*Plac-pfaAB*, and pMW219-*Plac-pfaCD*. After a 40-h IPTG incubation, lipids were extracted from the cells, methylated, and analyzed by gas chromatography. The number in parenthesis represents the double bond position counted from the carboxy terminus.

gene clusters of *S. oneidensis* MR-1 and *M. marina* MP-1. The gene cluster of *M. marina* MP-1 was composed of four genes, and the *pfaE* gene encoding a protein with a ppTase domain was not present in the cluster but was located in another region of the genome. Six repeat ACP domains were present in *S. oneidensis* MR-1 PfaA, whereas five repeat ACP domains were present in PfaA and an additional KS domain was present in *M. marina* MP-1 PfaB. In the *A. marina pfa* gene cluster identified in this study, the order of *pfa* genes was identical to that of *S. oneidensis* MR-1. Ten enzyme domains and four repeat ACP domains were identified within the proteins encoded in the *A. marina pfa* genes: a KS, a MAT, and four repeat ACP domains in PfaA; a KR in PfaB; a KS, a CLF, an AT, and two DH domains in PfaC; an ER in PfaD; and a ppTase in PfaE. The KR domain present in *S. oneidensis* MR-1 PfaA and *M. marina* MP-1 was present in PfaB of *A. marina*, whereas the AT domain in PfaB of *S. oneidensis* and *M. marina* was present in PfaC of *A. marina*. Four repeat ACP domains were found in the *A. marina* PfaA, although six and five repeat ACP domains were found in PfaA of *S. oneidensis* MR-1 and *M. marina* MP-1, respectively.

3.2. Expression of pfa genes isolated from A. marina in E. coli

To confirm that the gene cluster identified in *A. marina* encoded enzymes responsible for ARA biosynthesis, we expressed the *pfa* genes in the *A. marina* cluster in *E. coli*. As a host *E. coli* strain, BW25113 *ΔfadE*::FRT, which is defective in β-oxidation, was used to prevent degradation of the synthesized fatty acids. The BW25113 *ΔfadE*::FRT cells were transformed with pUC19-*Plac-pfaE*, pSTV29-*Plac-pfaAB*, and pMW219-*Plac-pfaCD* plasmids. As a control, the cells were also transformed with the pUC19, pSTV29, and pMW219 empty vectors. Expression of the *A. marina pfa* genes in the *E. coli* transformants was induced by adding IPTG at the beginning of culture. Before and after a 40-h incubation with IPTG, total lipids were extracted from the cells and analyzed.

Fig. 2 shows gas chromatograms of the FAMEs prepared from total lipids of the *E. coli* transformants with the *A. marina pfa* genes or empty vectors. In the control transformant with the empty vectors, 16:0, 16:1(9), cyclopropane fatty acid (17:0), and *cis*-vaccenic acid [18:1(11)] were detected as major fatty acids, but no peak corresponding to ARA was detected. In contrast, an additional fatty acid peak was identified in the transformant with the three plasmids carrying all *A. marina pfa* genes. As observed in the control, this additional peak was not detected even after the incubation with IPTG when only one or two of the 3 plasmids, pUC19-*plac-pfaE*, pSTV29-*Plac-pfaAB* and pMW219-*Plac-pfaCD*, were introduced into the *E. coli* cells (data not shown). The retention time of this peak was identical to that of an authentic methyl ester of ARA. These results suggest that this fatty acid is ARA, and that all *pfa* genes are required to biosynthesize this fatty acid.

To confirm that this fatty acid produced in the transformant with the *A. marina pfa* genes was ARA, the mass spectrum of the fatty acid pyrrolidide derivative was obtained, as shown in Fig. 3. In the mass spectrum, the M^+ parental ion was detected at *m/z* 357, which is identical to that of the calculated value for the ARA pyrrolidide derivative, and a series of major ions at *m/z* 113, 126, 140, 154, 166, 180, 192, 206, 220, 232, 246, 260, 272, 286, 300,

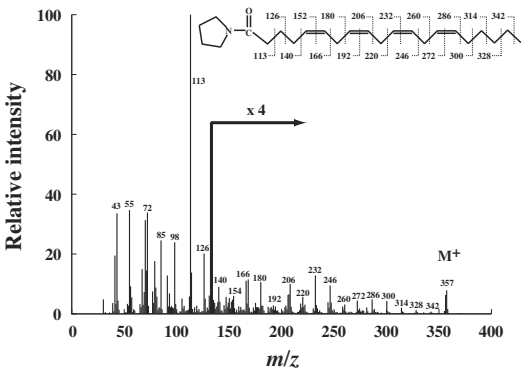

Fig. 3. Mass spectrum of a pyrrolidide derivative of the fatty acid corresponding to arachidonic acid (ARA) produced in *Escherichia coli* (BW25113 *ΔfadE*::FRT) cells transformed with pUC19-*Plac-pfaE*, pSTV29-*Plac-pfaAB*, and pMW219-*Plac-pfaCD*.

Table 1
Arachidonic acid (ARA) production in *Escherichia coli* (BW25113 *ΔfadE*::FRT) cells transformed with the *Aureispira marina pfa* genes. *E. coli* cells transformed with three empty vectors or three vectors including the *A. marina pfa* genes were incubated with IPTG for 40 h. Lipids were extracted from the cells before and after the IPTG incubation, and ARA content in total lipids and individual lipid classes were determined.

Plasmids	IPTG	ARA content (mol%)			
		Total	PE	PG	CL
Three empty vectors pUC19 + pSTV29 + pMW219	–	0	0	0	0
	+	0	0	0	0
Three vectors with *pfa* genes pUC19-*Plac-pfaE* + pSTV29-*Plac-pfaAB* + pMW219-*Plac-pfaCD*	–	0	0	0	0
	+	0.6 ± 0.0	0.2 ± 0.0	3.3 ± 0.5	0.9 ± 0.1

Values represent means ± standard deviations of independent preparations ($n > 3$).
CL, cardiolipin; PE, phosphatidylethanolamine; PG, phosphatidylglycerol.

314, 328, and 342 were detected. Gaps of 12 amu between *m/z* 180 and *m/z* 192, 220 and 232 as well as *m/z* 260 and 272 indicated double bonds at positions of 8, 11, and 14, respectively. The major ion at *m/z* 154 was probably due to rearrangement of the double bond at the Δ5 position under electron impact to the conjugated system (Δ5, 8 to Δ6, 8) causing an increase in the intensity of the 5:0 fragment (*m/z* 154) at the expense of the 5:1 (*m/z* 152), as reported by Valicenti et al. [23]. These results demonstrate that the fatty acid produced in the *E. coli* transformant with *A. marina pfa* genes was 20:4(5,8,11,14), namely ARA, and that the proteins encoded in the *A. marina pfa* genes are the components of a PUFA synthase required for ARA production. This is the first identification of a PUFA synthase that produces ARA.

Table 1 shows the ARA content in the lipid classes of the transformants with the empty vectors or *A. marina pfa* genes (Table 1). ARA was detected in all lipid classes of the transformant with the three plasmids carrying all of the *A. marina pfa* genes after the IPTG incubation but not in those of the transformant with empty vectors after the IPTG incubation. ARA content in phosphatidylglycerol was higher than that in phosphatidylethanolamine and cardiolipin.

S. oneidensis MR-1 [22] synthesizes EPA but not DHA, whereas *M. marina* MP-1 [24,25] synthesizes DHA but not EPA, although they have PUFA synthases with highly similar structures, as shown in Fig. 1. The PUFA synthase identified in this study is responsible for ARA biosynthesis and has a similar structure to those of *S. oneidensis* MR-1 and *M. marina* MP-1. The exact sequence of reactions involved in the biosynthesis of ARA has not been determined. However, a possible pathway for ARA biosynthesis can be proposed as shown in supplemental Fig. 1. Four double bonds could be inserted at the 4 elongation steps indicated by asterisks in the figure through the action of bifunctional DH, which catalyzes a dehydration reaction and an isomerization of double bond. At the elongation steps, the reduction of double bond by ER was skipped and unreduced intermediates are further converted to longer chain fatty acids, finally to ARA, 20:4(5,8,11,14).

How the PUFA synthases with high structural similarity preferentially synthesize either EPA, DHA, or ARA has not been clarified. However, detailed analysis of the PUFA synthases for EPA, DHA, and ARA production, such as the exchange of enzyme and ACP domains among PUFA synthases and their expression in *E. coli*, will suggest how a particular fatty acid is synthesized by each PUFA synthase.

4. Conclusion

We identified a PUFA synthase *pfa* gene cluster in *A. marina* and functionally expressed the *pfa* genes in an *E. coli* cluster. The *E. coli* cells expressing *A. marina pfa* genes synthesized a significant amount of ARA, demonstrating that the *pfa* genes in the cluster encode a PUFA synthase for ARA biosynthesis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2014.09.023.

References

[1] Ratledge, C. (2004) Fatty acid biosynthesis in microorganisms being used for single cell oil production. Biochimie 86, 807–815.
[2] Berge, J.P. and Barnathan, G. (2005) Fatty acids from lipids of marine organisms: molecular biodiversity, roles as biomarkers, biologically active compounds, and economical aspects. Adv. Biochem. Eng. Biotechnol. 96, 49–125.
[3] Ward, O.P. and Singh, A. (2005) Omega-3/6 fatty acids: alternative sources of production. Proc. Biochem. 40, 3627–3652.
[4] La Guardia, M., Giammanco, S., Di Majo, D., Tabacchi, G., Tripoli, E. and Giammanco, M. (2005) Omega 3 fatty acids: biological activity and effects on human health. Panminerva Med. 47, 245–257.
[5] Wallis, J.G., Watts, J.L. and Browse, J. (2002) Polyunsaturated fatty acid synthesis: what will they think of next? Trends Biochem. Sci. 27, 467–473.
[6] Uttaro, A.D. (2006) Biosynthesis of polyunsaturated fatty acids in lower eukaryotes. IUBMG Life 58, 563–571.
[7] Warude, D., Joshi, K. and Harsulkar, A. (2006) Polyunsaturated fatty acids: biotechnology. Crit. Rev. Biotechnol. 26, 83–93.
[8] Metz, J.G., Roessler, P., Facciotti, D., Levering, C., Dittrich, F., Lassner, M., Valentine, R., Lardizabal, K., Domergue, F., Yamada, A., Yazawa, K., Knauf, V. and Browse, J. (2001) Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes. Science 293, 290–293.
[9] Okuyama, H., Orikasa, Y., Nishida, T., Watanabe, K. and Morita, N. (2007) Bacterial genes responsible for the biosynthesis of eicosapentaenoic and docosahexaenoic acids and their heterologous expression. Appl. Environ. Microbiol. 73, 665–670.
[10] Yazawa, K. (1996) Production of eicosapentaenoic acid from marine bacteria. Lipids 31, S297–S300.
[11] Orikasa, Y., Nishida, T., Yamada, A., Yu, R., Watanabe, K., Hase, A., Morita, N. and Okuyama, H. (2006) Recombinant production of docosahexaenoic acid in a polyketide biosynthesis mode in *Escherichia coli*. Biotechnol. Lett. 28, 1841–1847.
[12] Shulse, C.N. and Allen, E.E. (2011) Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. PLoS ONE 6, e20146.
[13] Hosoya, S., Arunpairojana, V., Suwannachart, C., Kanjana-Opas, A. and Yokota, A. (2006) *Aureispira marina* gen. nov., sp. nov., a gliding, arachidonic acid-containing bacterium isolated from the southern coastline of Thailand. Int. J. Syst. Evol. Microbiol. 56, 2931–2935.
[14] Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio Collection. Mol. Syst. Biol. 2, 0008.
[15] Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. and Rothberg, J.M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.
[16] Zerbino, D.R. and Birney, E. (2008) Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. 18, 821–829.
[17] Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M. and Sonnhammer, E.L. (2002) The Pfam protein families database. Nucleic Acids Res. 30, 276–280.

[18] Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33, 103–119.
[19] Miller, E.M. and Nickoloff, J.A. (1996) *Escherichia coli* electrotransformation in: Electroporation Protocols for Microorganisms (Nickoloff, J.A., Ed.), Humana Press, Totowa, NJ.
[20] Bligh, E.G. and Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol. 37, 911–917.
[21] Andersson, B. and Holman, R. (1974) Pyrrolidides for mass spectrometric determination of the position of the double bond in monounsaturated fatty acids. Lipids 9, 185–190.
[22] Lee, S.J., Jeong, Y.S., Kim, D.U., Seo, J.W. and Hur, B.K. (2006) Eicosapentaenoic acid (EPA) biosynthetic gene cluster of *Shewanella oneidensis* MR-1: cloning, heterologous expression, and effect of temperature and glucose on the production of EPA in *Escherichia coli*. Biotechnol. Bioprocess Eng. 11, 510–515.
[23] Valicenti, A.J., Chapman, C.J., Holman, R.T. and Chipault, J.R. (1978) Mass spectrometry identification of C20 fatty acids in bovine lens using the pyrrolidide derivative. Lipids 13, 190–194.
[24] DeLong, E.F. and Yanos, A.A. (1986) Biochemical function and ecological significance of novel bacterial lipids in deep-sea procaryotes. Appl. Environ. Microbiol. 51, 730–737.
[25] Morita, N., Nishida, T., Tanaka, M., Yano, Y. and Okuyama, H. (2005) Enhancement of polyunsaturated fatty acid production by cerulenin treatment in polyunsaturated fatty acid-producing bacteria. Biotechnol. Lett. 27, 389–393.