Biosynthesis and heterologous production of polyunsaturated fatty acids from myxobacteria

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes

von
Katja Gemperlein

Saarbrücken
2014

Tag des Kolloquiums:	07.01.2015
Dekan:	Prof. Dr. Dirk Bähre
Berichterstatter:	Prof. Dr. Rolf Müller
	Prof. Dr. Elmar Heinzle
	Prof. Dr. Karl-Erich Jaeger
Vorsitz:	Prof. Dr. Gert-Wieland Kohring
Akad. Mitarbeiter:	Dr. Sonja M. Keßler

DANKSAGUNG

Zuallererst gilt mein Dank Prof. Dr. Rolf Müller für die Bereitstellung des PUFA-Projekts und die Ermöglichung der Promotion in einem Laborumfeld mit einer außergewöhnlichen instrumentellen Ausstattung. Seine stete Unterstützung und Diskussionsbereitschaft in wissenschaftlichen Belangen lieferten Denkanstöße, um mich als Wissenschaftler weiterzuentwickeln.

Prof. Dr. Elmar Heinzle möchte ich vielmals für die fruchtbare Diskussion während des Treffens des Promotionskomitees und für die Übernahme des Zweitgutachtens danken.

Ganz besonders danke ich natürlich Dr. Silke Wenzel für ihr Engagement und die vielen hilfreichen Diskussionen. Ihre Ansichten und Meinungen haben meine Arbeit im positiven Sinne beeinflusst. Auch für das Korrekturlesen von Manuskripten und Vorträgen sowie der Dissertation möchte ich mich ganz herzlich bedanken.

Dr. Ronald Garcia möchte ich gerne für seine Ratschläge bei der Kultivierung von Myxobakterien danken. Bei Dr. Kirsten Harmrolfs bedanke ich mich vielmals für ihre Hilfe bei der Synthese des DHA-NAC-Thioesters. Besonders gedankt sei Dr. Thomas Hoffmann und Michael Hoffmann für die Durchführung und Auswertung von LC-MS-Messungen. Dr. Nestor Zaburannyi danke ich für Assemblierung und Analyse von Genomsequenzen. Bedanken möchte ich mich auch bei Prof. Dr. Marc Stadler vom Helmholtz-Zentrum für Infektionsforschung für seine Tipps bezüglich des heterologen Wirtsstamms und bei Dr. Hubert S. Bernauer und Gregor Zipf von ATG:biosynthetics GmbH für ihre Arbeit zu Design und Synthese der artifiziellen Gene.

Für ihre tatkräftige Unterstützung im Rahmen ihres Wahlpflichtpraktikums möchte ich herzlich Nadine Leuthner, Christina Leichner und Jennifer Rheinfrank danken.

Bei allen Mitarbeitern des Instituts für Pharmazeutische Biotechnologie/Helmholtz-Instituts für Pharmazeutische Forschung Saarland, Abteilung Mikrobielle Naturstoffe, bedanke ich mich vielmals für ihre Hilfsbereitschaft und Unterstützung. Hierbei möchte ich besonders die gute Labornachbarschaft und Freundschaft mit Stefan Müller, Philipp Jochem und Sonja Burkart hervorheben.

Nicht zuletzt gilt mein ganz besonderer Dank meinen Freunden für ihre Unterstützung und ihr offenes Ohr und meiner Familie für ihre Verlässlichkeit und ihren Rückhalt.

VORVERÖFFENTLICHUNGEN DER DISSERTATION

Teile dieser Arbeit wurden vorab mit Genehmigung der Naturwissenschaftlich-Technischen Fakultät III, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen veröffentlicht:

Patente und Publikationen*

Gemperlein K, Wenzel SC, Rachid S, & Müller R (2011) Production of fatty acids by heterologous expression of gene clusters from myxobacteria. *Int. Pat.*, WO 2011/151298.

Gemperlein K, Rachid S, Garcia RO, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

* Die vorliegende Arbeit entstand bis Juli 2012 in Kooperation mit InterMed Discovery GmbH und unterlag in dieser Zeit einer Sperre zur Publikation in wissenschaftlichen Fachzeitschriften.

Tagungsbeiträge

Gemperlein K, Stadler M, Wenzel SC, & Müller R (2012) MYXO-PUFA: Homologe und heterologe Produktion von mehrfach ungesättigten Omega-3-Fettsäuren in Myxobakterien (Vortrag). GenoMik-Transfer Statusseminar 2012, Göttingen.

Gemperlein K, Stadler M, Wenzel SC, & Müller R (2012) Heterologous production of polyunsaturated fatty acids from myxobacteria (Vortrag). VAAM International Workshop 2012 “Biology and Chemistry of Antibiotic-Producing Bacteria and Fungi”, Braunschweig.

Gemperlein K, Oßwald C, Bernauer HS, Zipf G, Müller R, & Wenzel SC (2013) Synthetic DNA platforms to produce myxobacterial metabolites (Poster). 25. Irseer Naturstofftage, Irsee.

Gemperlein K, Tang Y, Oßwald C, Zipf G, Bernauer HS, Müller R, & Wenzel SC (2013) Production of myxobacterial metabolites based on synthetic DNA (Poster). 3rd International HIPS Symposium, Saarbrücken.

Gemperlein K, Wenzel SC, & Müller R (2013) Synthetic biotechnology for recombinant production of polyunsaturated fatty acids from myxobacteria (Poster). Euroscicon event "Applications for Synthetic Biology in Industrial Biotechnology", London.

Gemperlein K, Wenzel SC, & Müller R (2013) Synthetic biotechnology for heterologous production of polyunsaturated fatty acids from myxobacteria (Vortrag). Second Status Meeting of the Helmholtz Initiative on Synthetic Biology, Heidelberg.

WEITERE PUBLIKATIONEN

Folgende Publikationen entstanden während der Promotionszeit, sind jedoch nicht Bestandteil dieser Dissertation:

Kopp M, Irschik H, **Gemperlein K**, Buntin K, Meiser P, Weissman K, Bode HB, & Müller R (2011) Insights into the complex biosynthesis of the leupyrrins in *Sorangium cellulosum* So ce690. *Mol. Biosyst.*, **7,** 1549-1563.

Kessler SM, Simon Y, **Gemperlein K**, Gianmoena K, Cadenas C, Zimmer V, Pokorny J, Barghash A, Helms V, van Rooijen N, Bohle RM, Lammert F, Hengstler JG, Müller R, Haybäck J, & Kiemer AK (2014) Fatty acid elongation in non-alcoholic steatohepatitis and hepatocellular carcinoma. *Int. J. Mol. Sci.*, **15,** 5762-5773.

Laggai S, Kessler SM, Boettcher S, Lebrun V, **Gemperlein K**, Lederer E, Leclercq IA, Müller R, Hartmann RW, Haybäck J, Kiemer AK (2014) The *IGF2* mRNA binding protein p62/IGF2BP2-2 induces fatty acid elongation as a critical feature of steatosis. *J. Lipid Res.*, **55,** 1087-1097.

Simon Y, Kessler SM, **Gemperlein K**, Bohle RM, Müller R, Haybäck J, & Kiemer AK (2014) Elevated free cholesterol as a hallmark of non-alcoholic steatohepatitis in p62/IGF2BP2-2 transgenic animals. *World J. Gastroenterol.*, in press.

TABLE OF CONTENTS

ZUSAMMENFASSUNG

Mehrfach ungesättigte Fettsäuren (PUFAs), insbesondere die Omega-3-PUFAs Eicosapentaensäure (EPA) und Docosahexaensäure (DHA), zeigen vielfältige gesundheitsfördernde Wirkungen. Da eine hohe Nachfrage nach diesen PUFAs besteht, müssen nachhaltige Quellen für die Produktion im Industriemaßstab, wie z. B. die mikrobielle Fermentation, als Alternative zu Fischöl etabliert werden. Im Rahmen der vorliegenden Arbeit wurde die Fähigkeit von terrestrischen Myxobakterien, PUFAs mit Polyketidsynthasen verwandten PUFA-Synthasen *de novo* zu produzieren, intensiv studiert. Es wurden verschiedene Typen von PUFA-Biosynthesegenclustern identifiziert und ihre Produktspektren im Detail analysiert. Da die nativen Produzentenstämme biotechnologisch und genetisch schwer zugänglich sind, wurden Methoden der Synthetischen Biotechnologie zur erfolgreichen Übertragung und Expression der myxobakteriellen PUFA-Biosynthesewege in geeigneten Wirtsstämmen entwickelt. Neben dem myxobakteriellen Modellstamm *Myxococcus xanthus* haben sich *Pseudomonas putida* und die Fetthefe *Yarrowia lipolytica* als vielversprechende Wirtsstämme erwiesen. Um Einschränkungen bei der Effizienz der heterologen Expression von nativen Biosynthesewegen zu beheben, wurden erstmals auf künstlicher DNA basierte PUFA-Gencluster für die PUFA-Produktion in den rekombinanten Stämmen eingesetzt. Die heterologe Expression mutierter oder hybrider PUFA-Gencluster lieferte tiefere Einblicke in den Synthesemechanismus der PUFA-Synthasen.

ABSTRACT

Polyunsaturated fatty acids (PUFAs), in particular the omega-3 PUFAs eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), are associated with several beneficial health effects. As there is a high demand for these PUFAs, sustainable sources for industrial-scale production, e. g. microbial fermentation, have to be established as an alternative to fish oil. In the present work, the capability of terrestrial myxobacteria to produce PUFAs *de novo* via polyketide synthase-like enzymes known as PUFA synthases was intensively studied. Different types of PUFA biosynthetic gene clusters were identified and the corresponding product spectra were analyzed in detail. Due to restricted biotechnological and genetic amenability of the native producer strains, Synthetic Biotechnology approaches have been developed to successfully transfer and express myxobacterial PUFA biosynthetic pathways into suitable host strains. In addition to the myxobacterial model strain *Myxococcus xanthus*, *Pseudomonas putida* and the oleaginous yeast *Yarrowia lipolytica* turned out as promising host strains. Limitations related to the efficacy of the heterologous expression of the native biosynthetic pathways were addressed for the first time by the usage of PUFA gene clusters based on artificial DNA for PUFA production in the recombinant strains. Heterologous expression of mutated or hybrid PUFA gene clusters provided deeper insights into the synthetic mechanism of PUFA synthases.

INTRODUCTION

Polyunsaturated fatty acids (PUFAs) – Fatty acids with health benefits whose increasing demand requires sustainable sources

Lipids are naturally occurring hydrophobic or amphiphilic small molecules that serve as energy storage, structural components of membranes, or signalling molecules [1]. Essential constituents of lipids represent the fatty acids, which are commonly aliphatic, straight-chain monocarboxylic acids with an even carbon number. However, there are some exceptions, such as ring containing fatty acids, branched-chain fatty acids with an *iso*-structure (methyl group at the penultimate carbon atom) or an *anteiso*-structure (methyl group on the third carbon from the end), odd-numbered fatty acids, hydroxy or epoxy fatty acids, to name but a few [2;3]. Fatty acids are termed saturated (SFAs) if they contain only carbon-carbon single bonds. In contrast, monounsaturated fatty acids (MUFAs) comprise an additional double bond, whereas polyunsaturated fatty acids (PUFAs) contain multiple double bonds. In the most abundant PUFAs, double bonds are in *cis* configuration and not conjugated, as they are separated by a methylene group. A simplified nomenclature specifies the chain length and number plus position of double bonds (Figure 1). Positions of double bonds are determined by the number of the first carbon atom that forms the double bond, indicated by "Δ", counting from the carboxyl group carbon as number one. Since the double bonds are usually present at three carbons intervals, PUFAs alternatively can be described simply by indicating the number of double bonds and the position of the double bond closest to the terminal ω-carbon, indicated by "*n*". The remainder of the double bond positions can then be inferred.

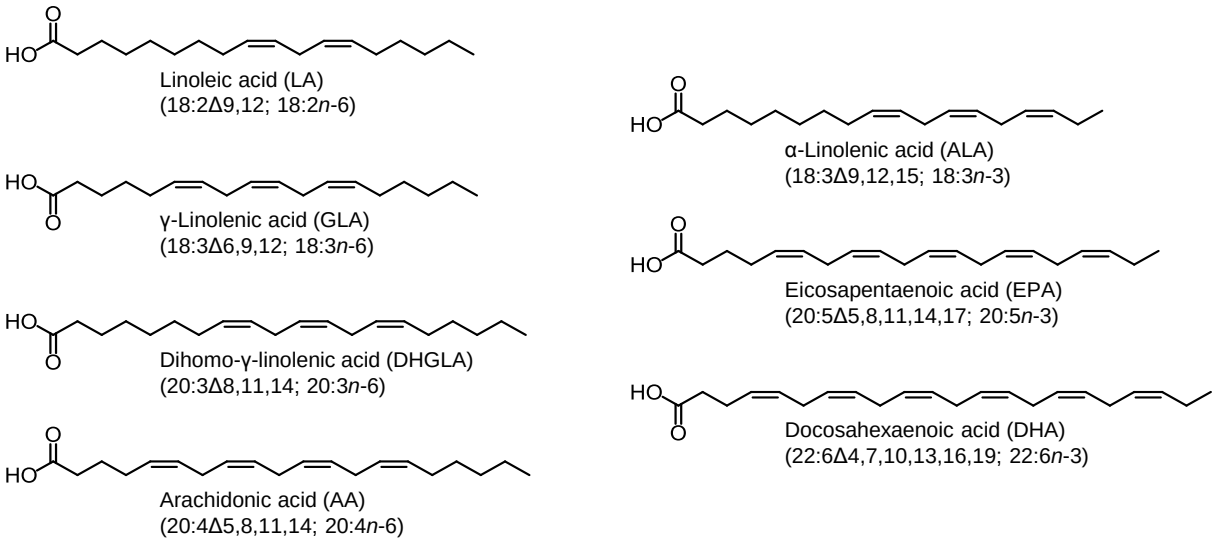


Figure 1. Structures of selected long-chain polyunsaturated fatty acids (LC-PUFAs).

The role of long-chain PUFAs (LC-PUFAs) in health is one of the most studied areas of nutrition science. The first observations were made in the 1960s when epidemiologists noticed that Greenland Eskimos had a low incidence of coronary heart disease. Compared to Danes, Greenland Eskimos had similar blood levels of total lipids but considerably higher blood levels of eicosapentaenoic acid (EPA, 20:5, *n*-3) and docosahexaenoic acid (DHA, 22:6, *n*-3) as well as lower blood levels of linoleic acid (LA, 18:2, *n*-6) and arachidonic acid (AA, 20:4, *n*-6) [4]. Based on these findings, along with observations in other populations, such as in Japan, associations between intake of *n*-3 LC-PUFAs and cardiovascular diseases have been noted [5]. Numerous studies emphasize the prominent physiological roles of *n*-3 and *n*-6 LC-PUFAs in different organs. For instance, the high content of AA, docosatetraenoic acid (DTA, 22:4, *n*-6), and DHA in the neuronal tissue implicates these fatty acids as critical for normal development and functioning of the brain [6;7]. DHA is the predominant LC-PUFA in the glycerophospholipids of the grey matter, whereas in white matter, there is a higher proportion of *n*-6 than of *n*-3 LC-PUFAs [6]. An essential biological function of LC-PUFAs is the formation of eicosanoids derived from *n*-6 and *n*-3 C20 PUFAs and of docosanoids derived from C22 PUFAs [8-15] (Figure 2).

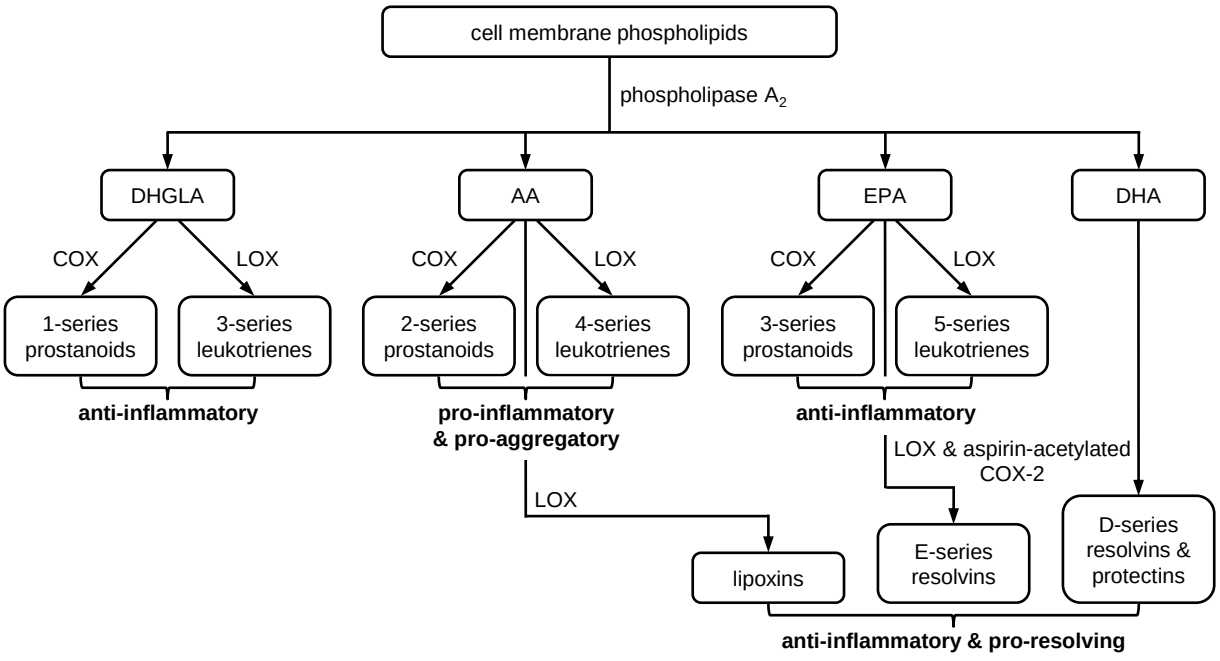


Figure 2. Formation of eicosanoids from dihomo-γ-linolenic acid (DHGLA), arachidonic acid (AA), and eicosapentaenoic acid (EPA) and docosanoids from docosahexaenoic acid (DHA) via cyclooxygenase (COX) and lipoxygenase (LOX) pathways.

Low-grade inflammation induced by pro-inflammatory mediators, such as cytokines, has been identified as a key factor in the development of metabolic syndrome. This disorder is characterized by abdominal obesity, hypertension, dyslipidemia, and elevated fasting plasma glucose or insulin resistance and increases the risk of developing cardiovascular disease, particularly heart failure, as well as diabetes mellitus type 2 [16]. It may be possible that these chronic diseases can be prevented or treated by modulating the ratio of pro- and anti-inflammatory eicosanoids [13]. In the formation of the eicosanoids and docosanoids, AA, EPA, and DHA compete for the same enzymes, COX and LOX [9] (Figure 2). As plasma lipids and cell membranes typically contain a higher proportion of AA than of EPA plus DHA [17], AA is the predominant substrate for these enzymes. However, it can be speculated that a higher concentration of EPA plus DHA could either antagonize AA metabolism or simply dilute out the concentration of AA and subsequently decrease the formation of the potent inflammatory and pro-aggregatory eicosanoids [18]. As the fatty composition in tissues reflects the fatty acid composition of the diet [19], it may be possible that the risk of chronic diseases can be reduced by a higher dietary intake of *n*-3 LC-PUFAs [9;13]. Based on these indications, it is hardly surprising that these fatty acids have gained increasing interest in nutritional research over the past decades.

Due to the potential beneficial effects of *n*-3 LC-PUFAs on human health, several worldwide government and health organizations have made some recommendations for dietary intake of *n*-3 LC-PUFAs. The Food and Agriculture Organization of the United Nations (FAO) recommends intake of 250 mg to 2 g EPA and DHA per day [20]. The intake recommendation of the European Food Safety Authority (EFSA) as well as of the German Nutrition Society (DGE) corresponds to approximately 250 mg EPA and DHA per day [21;22]. Indeed, the real daily intake of these *n*-3 LC-PUFAs among Western populations is below these recommendations: According to the National Health and Nutrition Examination Survey 2009-2010, the mean daily intake of EPA and DHA was 90 mg among US-American women and 120 mg among US-American men [23]. Using data from the German Nutrition Survey 1998, the determined mean intake of EPA and DHA was 186 mg per day for German women and 255 mg for German men [24]. Dietary sources of LC-PUFAs are ubiquitous and vary widely among the different types of fatty acids. LA occurs in almost all dietary fats and represents a main component of most plant oils. α-linolenic acid (ALA, 18:3, *n*-3) is primarily present in vegetable oils, such as flaxseed oil, canola oil, and soybean oil, although its presence is usually much lower than that of LA [25]. AA is predominantly present in meat, poultry, and eggs [26]. The major sources of EPA and DHA are marine fish, especially oily

fish, such as cod, herring, mackerel, salmon, as well as sardine, and the corresponding fish oils [27].

There has been a steady increase in fish production since 1950. In 2010, the total world production of fish and other seafood was 148.5 million tons (88.6 million tons from capture and 59.9 million tons from aquaculture), and of this, 86% were available for human consumption. The remaining 14% were destined for non-food products, mainly for the manufacture of fish meal and fish oil [28]. The total production of fish oils is approximately 1 million tons per year. In 2007, 81.3% of all fish oil was used by the aquaculture industry to produce feed, and salmon farming alone used about 37% of all fish oil produced in 2008. The remaining 18.7% were processed into products for human consumption, mainly as fish oil capsules [20;29]. The principal problem in aquaculture is the unfavourable "fish in/fish out" ratio. For salmon farming, it is around 4:1. This means it takes about four kilograms of wild-caught feeder fish to provide the fish meal or fish oil needed to raise one kilogram of salmon [30]. Intensive fishery activities, in order to supply fish for human consumption as well as to deliver feed for industrial fish farms and fish oil supplements, result in alarming effects on fish levels and on marine ecosystems [31]. Hence, neither wild nor farmed fish constitute a sustainable source of *n*-3 LC-PUFAs, and it seems improbable that this source can satisfy the keen and increasing demand for omega-3 products. These inferences are evidenced by actual market research: In 2013, global production of omega-3 products was estimated at 2.49 million tons, worth US$ 4.5 billion and is forecasted to grow substantially in the next few years [32]. Moreover, fish oil shows substantial variability in composition and quality. In contrast to vegetable oils, it is susceptible to contamination with heavy metals, such as mercury and cadmium, and lipophilic organic chemicals, such as polychlorinated biphenyls (PCBs), organochlorine pesticides (OCPs), and polybrominated diphenyl ethers (PBDEs). These contaminants are typically bioaccumulated in the marine food network and lead to a biomagnification in the lipid fraction [33;34]. Furthermore, the usage of fish oil has become unpopular due to unpleasant odour and flavour, unsuitability for vegetarians, and expensive purification [35].

In quest of alternative, sustainable sources of *n*-3 LC-PUFAs, vegetable oils produced by the reconstitution of *n*-3 LC-PUFA biosynthetic pathways in transgenic plants, are currently being developed [36]. Besides, microbial oils, also termed single cell oils, are of great commercial interest. Mainly, the primary *n*-3 LC-PUFA producers in the marine biota, microalgae, serve as uncontaminated source for the constant production of high-quality *n*-3

LC-PUFA-containing oils. An extensive number of microalgal species is capable of accumulating large amounts of *n*-3 LC-PUFAs. The heterotrophic thraustochytrids *Schizochytrium* sp., *Thraustochytrium* sp., and *Ulkenia* sp. as well as the heterotrophic dinoflagellate *Crypthecodinium cohnii* exhibit production of up to 52%, 37%, 38%, or 44% DHA of total fatty acids (TFAs), respectively, under improved cultivation conditions [37-40]. Photoautotrophic microalgae, such as the eustigmatophyte *Nannochloropsis* sp. and the diatom *Phaeodactylum tricornutum*, produce up to 30% or 36% EPA of TFAs, respectively, under improved cultivation conditions [41;42]. The commercial production of algal oil started with the large-scale fermentation of a *Schizochytrium* sp. strain by OmegaTech (USA) in the early 1990s. Since then, algae biotechnology has grown steadily with increasing numbers of companies attempting to develop production processes for a wide array of applications. Nowadays, DSM (The Netherlands) is probably the most prominent commercial producer of DHA-rich oil using both a *Schizochytrium* sp. strain and a *Crypthecodinium cohnii* strain. Similar DHA-rich oils are produced by Lonza (Switzerland) using an *Ulkenia* sp. strain, whereas Qualitas Health (Israel) uses a *Nannochloropsis oculata* strain to produce an EPA-rich oil, to name but a few [43]. At present, achieving economically viable production of *n*-3 LC-PUFAs by microalgae is still a major challenge and requires additional developments, especially in terms of improved large-scale fermentation processes ensuring optimal growth conditions for enhanced lipid biosynthesis [44]. Hence, as the ideal microorganism for industrial-scale production of single cell oils has obviously not yet been identified, further exploitation of the microbial biodiversity and screening for PUFA-producing microbes seems to be a reasonable approach.

PUFA biosynthesis – Biosynthetic diversity by aerobic and anaerobic routes

PUFAs can be biosynthesized via two principal pathways, the anaerobic route for *de novo* PUFA biosynthesis and the more common aerobic PUFA biosynthetic pathway. The latter operates in a wide range of organisms and utilises saturated fatty acids (SFAs) as biosynthetic precursors. Initially, synthesis of the SFAs is catalysed by fatty acid synthase (FAS; EC 2.3.1.85), which is found ubiquitously across all groups of organisms. The biosynthetic mechanism is strongly conserved between bacteria and eukaryotes, albeit the arrangement of the catalytic entities within the proteins differs [45]. There are two principal classes of FASs. Fungi and animals possess a type I FAS system, which consists of large, multifunctional polypeptides forming oligomers [46]. Type II FAS system is found in archaea, bacteria, and

plants and is characterized by the use of discrete, monofunctional enzymes [47]. The substrates of FAS are derived from the acetyl-coenzyme A (acetyl-CoA) pool. Typically, acetyl-CoA serves as starter molecule, whereas malonyl-CoA is required for all the elongation steps. Malonyl-CoA is formed by carboxylation of acetyl-CoA using acetyl-CoA carboxylase (EC 6.4.1.2) [48]. As a key feature of the fatty acid biosynthetic pathway, the starter molecule and the intermediates are covalently bound to a small, acidic protein called acyl carrier protein (ACP) [49]. Thereby, the carboxyl groups of the fatty acyl starter molecules or intermediates form a thioester linkage to the sulfhydryl group of the 4'-phosphopantetheine (4'-PP) prosthetic group which in turn is linked to the active site serine of ACP through a phosphodiester bond. Thus, the *apo*-form of ACP has to be posttranslationally activated to the *holo*-form by transfer of the 4'-PP moiety from CoA to an invariant serine residue. This reaction is catalysed by a 4'-phosphopantetheinyl transferase (PPTase) [50] (Figure 3).

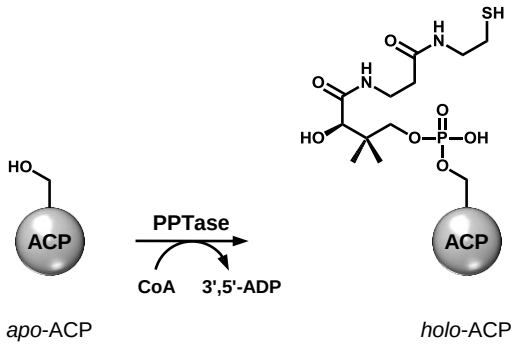


Figure 3. Posttranslational modification of an acyl carrier protein (ACP) by a 4'-phosphopantetheinyl transferase (PPTase).
The 4'-phosphopantetheine of coenzyme A (CoA) is transferred to a conserved serine of the inactive *apo*-ACP generating the active *holo*-ACP and adenosine 3',5'-bisphosphate (3',5'-ADP).

The pathway proceeds in two stages, initiation and cyclic elongation (Figure 4). At the outset, acetyl transferase/malonyl transferase transfers the acetyl/malonyl group of acetyl-CoA/malonyl-CoA to ACP. The first reaction in the elongation cycle is the decarboxylative Claisen condensation of malonyl-ACP with the acetyl group or the growing acyl chain at the active site cysteine of the β-ketoacyl-ACP synthase. This is the only irreversible step in the elongation cycle. The β-keto group is then reduced by sequential action of the NADPH (the reduced form of nicotinamide adenine dinucleotide phosphate)-dependent β-ketoacyl-ACP reductase, the β-hydroxyacyl-ACP dehydratase, and the NAD(P)H-dependent enoyl-ACP reductase to yield a saturated acyl product. This intermediate functions as a starter substrate for the next round of elongation with malonyl-ACP until the growing fatty acid chain reaches

a length of 16 to 18 carbon atoms [51;52]. The product is either released from ACP as free fatty acid by a thioesterase [46] or the acyl chain is directly transferred from ACP into glycerophospholipids by various acyltransferase systems [53].

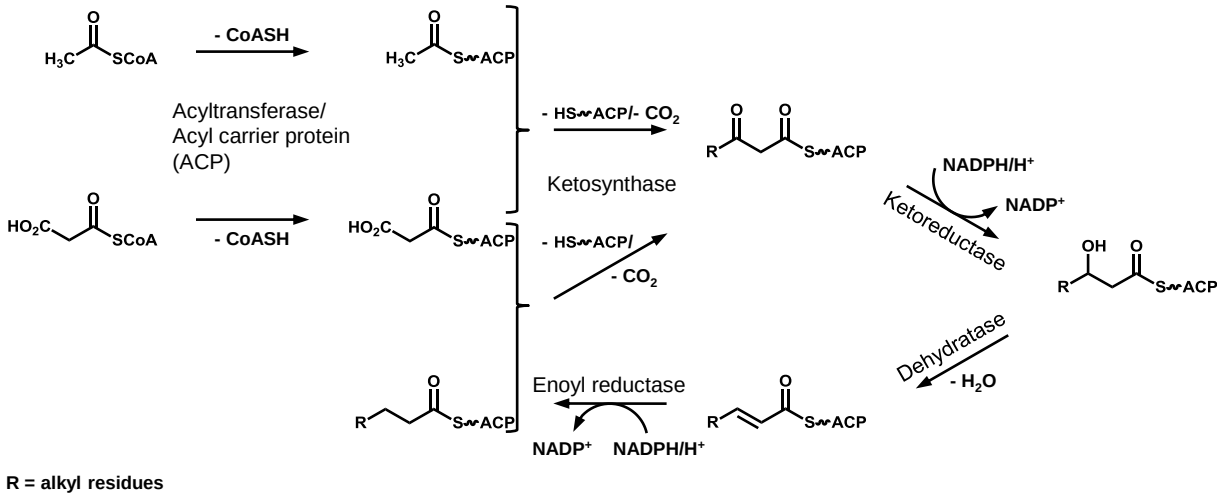


Figure 4. Biosynthesis of saturated fatty acids (SFAs) by fatty acid synthase (FAS).
The primer molecule (in the form of acetyl-CoA) undergoes several rounds of decarboxylative Claisen condensation reactions, resulting in the elongation of the fatty acyl chain by two carbons (derived from malonyl-CoA) per cycle until the growing fatty acid chain reaches its final length. Following each round of elongation the β-keto group is fully reduced.

Double bonds can be inserted into SFAs via the aerobic pathways by employing position-specific desaturases and elongases, which allow for the introduction of more than one double bond per fatty acid. Fatty acid desaturases (EC 1.14.19.-) are usually membrane-bound, non-heme iron-containing enzymes that introduce a double bond at a specific position in long-chain fatty acids. This reaction requires molecular oxygen, NAD(P)H, as well as an electron transport system (cytochrome b_5 reductase (EC 1.6.2.2) plus cytochrome b_5 or ferredoxin-$NADP^+$ reductase (EC 1.18.1.2) plus ferredoxin). The action of the desaturases alternates with the activity of fatty acid elongases. Fatty acid elongases are very-long-chain β-ketoacyl-CoA synthases (EC 2.3.1.199) and catalyse the decarboxylative Claisen condensation of malonyl-CoA with a LC-PUF acyl-CoA. The β-keto group is then reduced by sequential action of the distinct enzymes very-long-chain β-ketoacyl-CoA reductase (EC 1.1.1.330), very-long-chain β-hydroxyacyl-CoA dehydratase (EC 4.2.1.134), and very-long-chain enoyl-ACP reductase (EC 1.3.1.93) in a NADPH-dependent manner. Fatty acid desaturases and elongases are classified according to their regiospecificity: Δ-desaturases insert the double bond at a fixed position counted from the carboxyl group carbon, whereas Δ-elongases insert ethyl groups at a fixed position counted from the double bond closest to the carboxyl group carbon [54-56].

Several aerobic LC-PUFA biosynthetic pathways, utilising diverse desaturases and elongases, have been identified in eukaryotes (Figure 5). After synthesis of stearic acid (SA, 18:0) by FAS, vertebrates can readily introduce a double bond at the Δ9 position using Δ^9 desaturase to generate oleic acid (OA, 18:1, *n*-9). However, they cannot introduce additional double bonds between position Δ10 and the methyl terminal end. Thus, LA (18:2, *n*-6) and ALA (18:3, *n*-3) are essential fatty acids for vertebrates. They must be obtained from the consumption of plants, which can introduce double bonds at positions Δ12 and Δ15 using Δ^{12} and Δ^{15} desaturases [55;56]. Once LA and ALA are absorbed from the diet, they can be further processed. Thereby, the Δ^6 pathways are the conventional pathways, widely present in most eukaryotic organisms except for higher plants. These pathways need only Δ^6 and Δ^5 desaturases and Δ^6 elongase for the conversion of LA into γ-linolenic acid (GLA, 18:3, *n*-6), DHGLA (20:3, *n*-6), and finally AA (20:4, *n*-6) in the *n*-6 series or for the conversion of ALA into stearidonic acid (SDA, 18:4, *n*-3), eicosatetraenoic acid (ETA, 20:4, *n*-3), and finally EPA (20:5, *n*-3) [54] in the *n*-3 series. The two pathways are independent of each other, but there is competition for the conversions, since both pathways use the same enzymes. The predominant route for DHA biosynthesis in mammals represents the Sprecher pathway [57]. First, EPA is elongated to docosapentaenoic acid (DPA, 22:5, *n*-3) by Δ^5 elongase, and then DPA is elongated to tetracosapentaenoic acid (TPA, 24:5, *n*-3) [57-59]. TPA is desaturated to yield tetracosahexaenoic acid (THA, 24:6, *n*-3) by the Δ^6 desaturase which most likely also desaturates LA and ALA [60;61]. Finally, DHA is formed through shortening of THA by two carbon units during β-oxidation in the peroxisomes [58]. Alternative pathways for the aerobic biosynthesis of AA, EPA, and DHA (Figure 5) have been discovered and biochemically characterized in marine microalgae [62-65], in the protist *Euglena gracilis* [66;67], in oomycetes [68;69], in protozoan trypanosomes [70], and in a marine teleost fish [71].

Albeit humans have the capacity to convert ALA to EPA and DHA, the efficiency of these conversions is low. Using a stable isotope tracer of ALA, the conversion rates of ALA to EPA (estimates are from 0.2 to 21%, depending on age and gender of the experimentee) and of ALA to DHA (estimates are from 0.05 to 9%, depending on age and gender of the experimentee) was determined in the plasma [72]. Hence, the conversion of ALA to EPA, especially the rate-limiting Δ^6 desaturation, is the most restrictive step in this sequence and appears to strongly limit the production of EPA and DHA. Unlike LA, a large proportion of ingested ALA may be more readily degraded via β-oxidation then incorporated into the lipids [73;74]. The concentration of ALA in phospholipids in plasma and tissues is typically less than 0.5% of TFAs [75]. It seems most likely that this low content of ALA is not sufficient to

compete with LA for the Δ^6 desaturase. Overall, ALA appears to be a limited source for longer chain *n*-3 PUFAs in humans. Thus, adequate dietary intake of *n*-3 LC-PUFAs, in particular DHA, may have physiological significance.

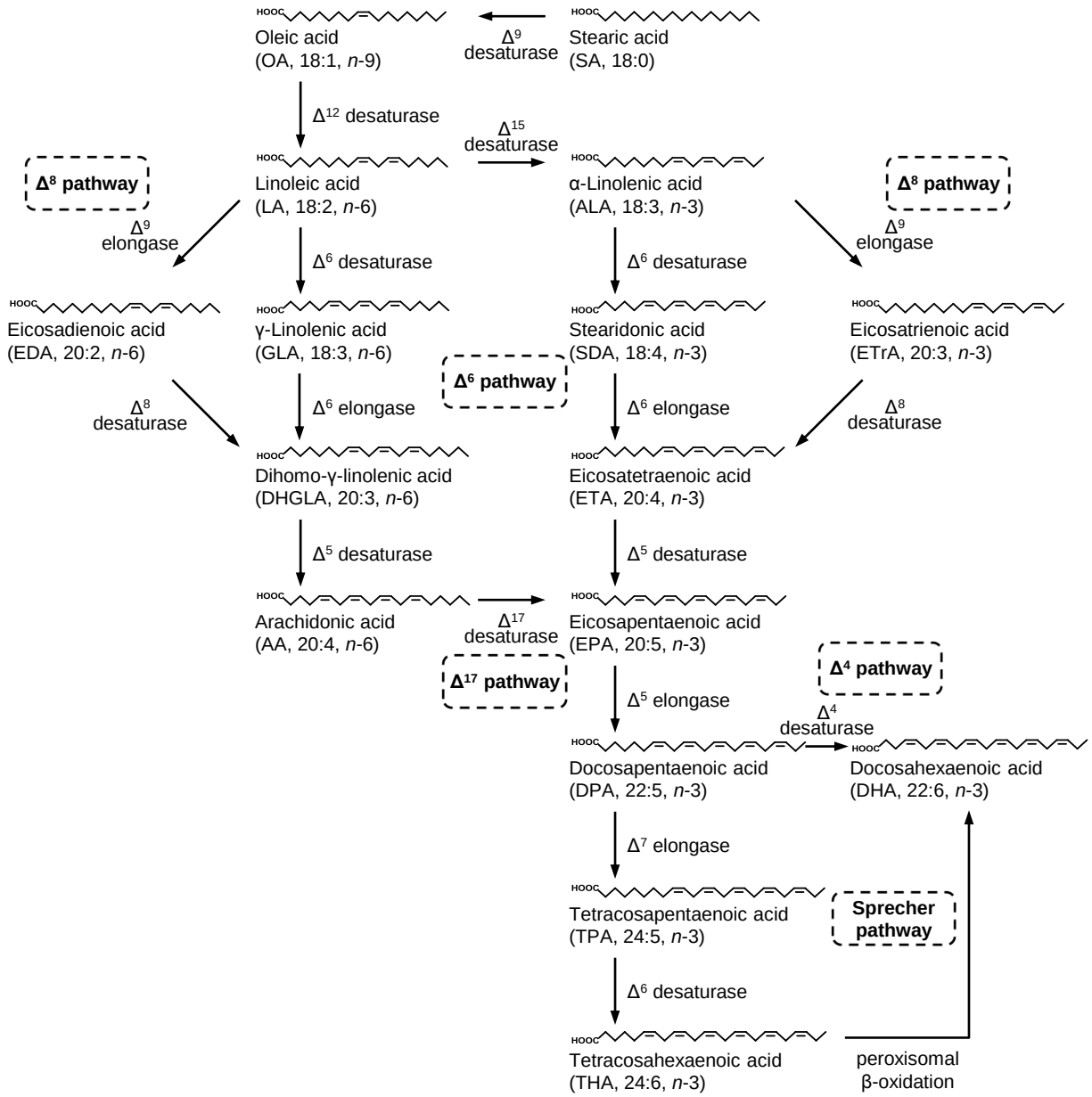


Figure 5. Aerobic biosynthesis of the polyunsaturated fatty acids (PUFAs) arachidonic acid, eicosapentaenoic acid, and docosahexaenoic acid starting from stearic acid by a combination of alternating desaturation and elongation reactions.

PUFA biosynthesis not necessarily requires the presence of fatty acid desaturases and elongases. In 1996, cloning, transfer, and expression of a gene cluster from the marine γ-Proteobacterium *Shewanella pneumatophori* SCRC-2738 in *Escherichia coli* led to heterologous EPA production [76]. Surprisingly, the cloned genes included no conserved

sequences from fatty acid desaturase genes. These findings were very important and could be regarded as the first stage of research into alternative PUFA biosynthetic pathways. Detailed sequence analysis revealed that the gene cluster encodes iterative type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like enzymes (Figure 6). Additional studies on heterologous expression of the genes demonstrated that EPA is produced via an anaerobic *de novo* synthesis from the substrates acetyl-CoA and/or malonyl-CoA and the coenzyme NADPH [77]. In the early 2000's, PUFA (*pfa*) biosynthetic gene clusters encoding these multienzyme complexes, termed PUFA synthases, have also been identified and characterized from other marine γ-Proteobacteria, such as the EPA producer *Photobacterium profundum* SS9 [78] and the DHA producer *Moritella marina* MP-1 [79], as well as from the DHA and *n*-6 DPA producing microalgae *Schizochytrium* sp. ATCC 20888 [77;80] (Figure 6).

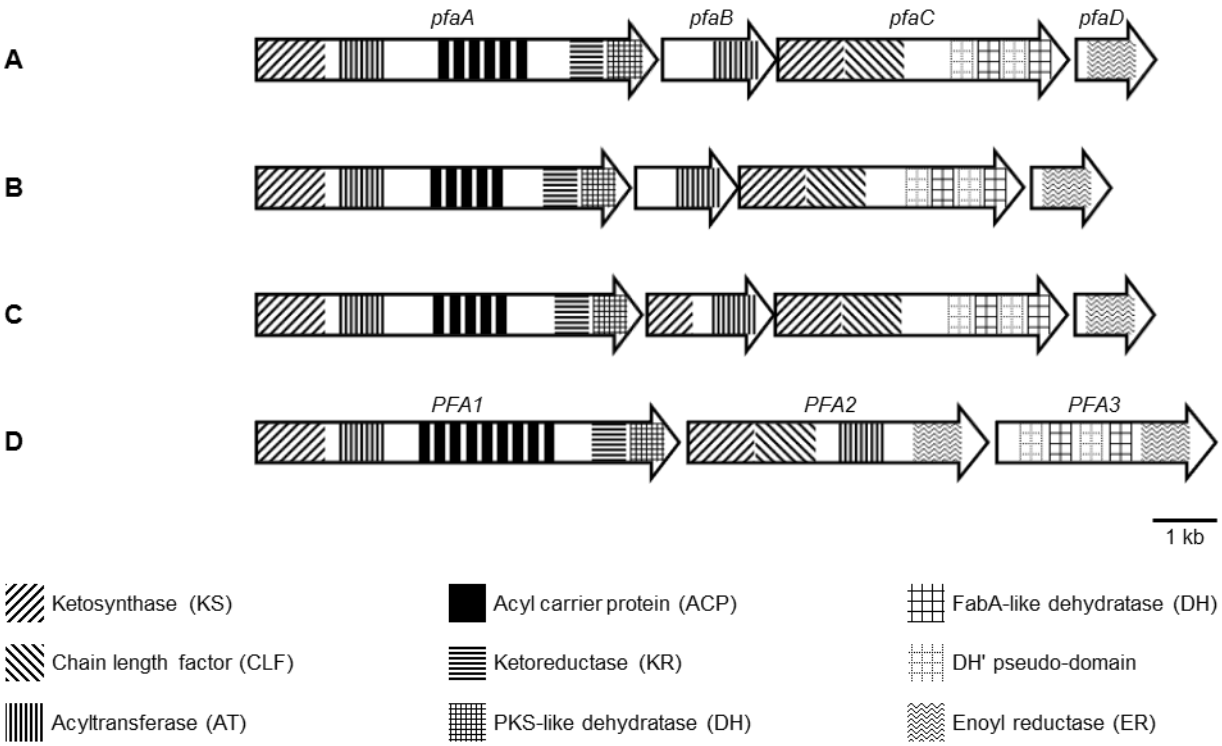


Figure 6. PUFA biosynthetic gene clusters from marine γ-Proteobacteria and microalgae.
(A) *Shewanella pneumatophori* SCRC-2738 (EPA producer), (B) *Photobacterium profundum* SS9 (EPA producer), (C) *Moritella marina* MP-1 (DHA producer), (D) *Schizochytrium* sp. ATCC 20888 (DHA and *n*-6 DPA producer).

The PUFA synthases from *Shewanella pneumatophori* SCRC-2738, *Photobacterium profundum* SS9, *Moritella marina* MP-1, and *Schizochytrium* sp. ATCC 20888 comprise enzymatic domains from FAS as well as from PKS (Figure 6). PKS is an evolutionary descendant of FAS, as both synthases share the same core of enzymatic activities [45;81].

Hence, it is hardly surprising that LC-PUFA biosynthesis by PUFA synthases proceeds in a similar way as the biosynthesis of SFAs catalysed by FAS (Figure 7). Analogous to the fatty acid assembly line, the growing fatty acid chain is linked to the 4'-PP prosthetic group of ACP via a thioester bond. Therefore, the *apo*-form of ACP has to be posttranslationally activated to the *holo*-form by transfer of the 4'-PP moiety from CoA to an invariant serine residue. This reaction is catalysed by a 4'-phosphopantetheinyl transferase (PPTase) [50] (Figure 3). For *Shewanella pneumatophori* SCRC-2738, *Photobacterium profundum* SS9, and *Moritella marina* MP-1, Sfp-type PPTases have been found to be involved in LC-PUFA biosynthesis [82;83]. During PUFA biosynthesis, acetyl-CoA serves as starter molecule, whereas malonyl-CoA, formed by carboxylation of acetyl-CoA using acetyl-CoA carboxylase (EC 6.4.1.2) [48], is required for all the elongation steps. The pathway is divided into two stages, initiation and cyclic elongation. In the initiation phase, an acyltransferase (AT) domain transfers the malonyl group of malonyl-CoA to an ACP. The first reaction in the elongation cycle is the decarboxylative Claisen condensation of malonyl-ACP with the acetyl group or the growing acyl chain at the active site cysteine of the ketosynthase (KS) domain. The β-keto group is then either fully reduced by sequential action of NADPH-dependent ketoreductase (KR) domain, PKS-like dehydratase (DH) domain, and NADPH-dependent enoyl reductase (ER) domain or only reduced to the *trans* double bond by the KR domain and a FabA-like DH domain and isomerized to the *cis* double bond by the latter domain. The intermediate then functions as a starter substrate for the next round of elongation with malonyl-ACP until the growing fatty acid chain with methylene-interrupted *cis* double bonds reaches its final length [84-86]. However, neither has the exact function and biosynthetic mechanism underlying every domain within the PUFA synthase complex yet been elucidated nor the relevant factor(s) that determine the nature of the final LC-PUFA product. The product is either released from ACP as free fatty acid by a discrete thioesterase [87;88] or the acyl chain is immediately directed from ACP into glycerophospholipids by various acyltransferase systems. Still, further studies are required to clarify the specific release mechanism.

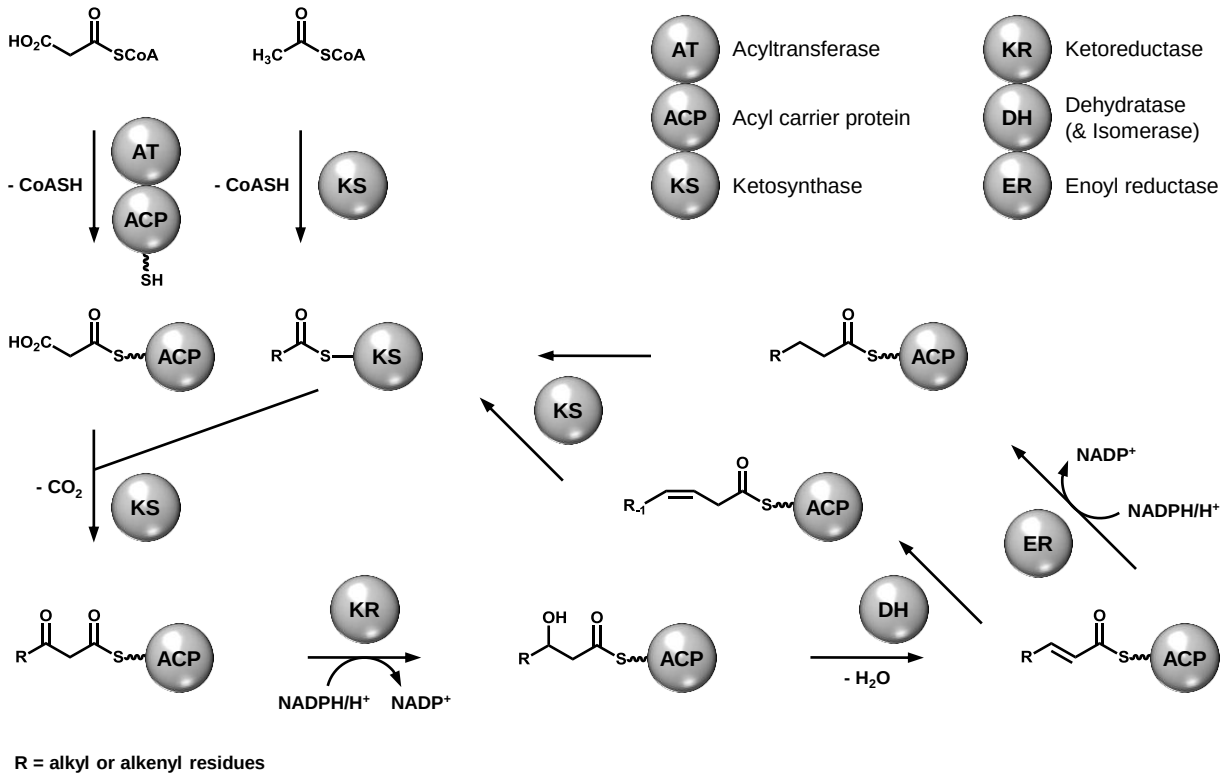


Figure 7. Anaerobic biosynthesis of polyunsaturated fatty acids (PUFAs) by iterative type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like PUFA synthases.
The primer molecule (in the form of acetyl-CoA) undergoes several rounds of decarboxylative Claisen condensation reactions, resulting in the elongation of the fatty acyl chain by two carbons (derived from malonyl-CoA) per cycle. Following each round of elongation, the β-keto group is either fully reduced or only reduced to the *trans* double bond which is then isomerized. Finally, an acyl chain with methylene-interrupted *cis* double bonds is synthesized.

Taking a closer look at the domain organization of the PUFA synthases from *Shewanella pneumatophori* SCRC-2738, *Photobacterium profundum* SS9, *Moritella marina* MP-1, and *Schizochytrium* sp. ATCC 20888 (Figure 6) discloses that proteins PfaA and PFA1 represent iterative type I PKSs. These protein systems feature the presence of multiple active sites within a single polypeptide which are used repeatedly [89-91]. In concrete terms, the catalytic domains include a KS domain, an AT domain, tandem ACP domains, a KR domain, and a DH domain. Particularly, the tandem ACP domains are characteristic for PUFA synthases. Their number ranges from five to nine, and their solution structure reveals a beads-on-a-string configuration [92]. Proteins PfaC and PFA2 contain two KS domains which resemble KS_α and KS_β of type II PKS. Type II PKSs are iteratively acting complexes of discrete monofunctional proteins [93]. In contrast to KS_α, KS_β lacks the active site cysteine for the catalysis of the Claisen condensation and has also been termed chain length factor (CLF) [94-96]. Other components of proteins PfaC and PFA3 are two DH domains which are homologous to the β-hydroxydecanoyl-ACP DH FabA from FAS of α- and γ-Proteobacteria,

exhibiting both dehydratase and isomerase activity [52]. Adjacent to the FabA-like DH domains, two inactive DH' pseudo-domains have been recently identified, which might be responsible for the proper folding and the activity of the catalytic domains [97]. Interestingly, the ER domains of PfaD, PFA2, and PFA3 are phylogenetically distinct from ERs utilized in canonical fatty acid and polyketide biosynthesis. Instead, these ERs are homologous to *trans*-ERs, such as PksE, involved in the biosynthesis of dihydrobacillaene from *Bacillus subtilis* [98;99], or TaN, involved in the biosynthesis of myxovirescin from *Myxococcus xanthus* DK1622 [100], among others.

Since the discovery of the first *pfa* gene clusters, the abundance and wide distribution of *pfa* gene homologues, primarily in the marine habitat, could be demonstrated [101]. Extensive sequencing of microbial genomes has allowed for broad and advanced genome analyses among diverse microbial lineages and was enabled by inexpensive and rapid whole genome sequencing technologies called the next generation sequencing (NGS). The main principle in NGS involves the massively parallel sequencing of DNA molecules in a flow cell. The sequencing process takes place either in a continuous real-time manner or in a stepwise iterative process. Each clonal template or single molecule is independently sequenced and can be counted among the total sequences generated [102]. In a phylogenomic approach, all sequenced microbial genomes were analysed for the presence of FAS/PKS gene clusters homologous to those involved in LC-PUFA biosynthesis. At least 20 distinct types of *pfa*-like gene clusters have already been found in 45 microbial genera representing ten phyla, demonstrating the diversity and wide pervasion of these secondary fatty acid biosynthetic pathways. However, for most of the discovered *pfa*-like gene clusters, the biosynthetic products remain uncharacterized [103].

Myxobacteria – Not only prolific producers of secondary metabolites but also of PUFAs

Myxobacteria are Gram-negative, rod-shaped slime bacteria belonging to the δ-Proteobacteria. Almost all myxobacteria are aerobic chemoorganotrophs [104]. They are commonly isolated from soil, dung, and decaying plant material but also from aquatic habitats and can inherently be found in all climate zones, vegetation belts, and altitudes [105]. A hallmark of myxobacterial behaviour is their ability to build a fruiting body by organizing the movements of thousands of cells under starvation conditions. Many of these cells then differentiate into myxospores, which are highly resistant to environmental stress factors as high temperature, desiccation, and UV radiation [106]. Myxobacteria are also noted for other

multicellular behaviours, such as swarming, rippling, elasticotaxis, and predation, which are unparalleled in the bacterial kingdom [107;108]. Moreover, they are prolific producers of novel secondary metabolites with imposing structural diversity, bioactivity, and modes-of-action [109]. The vast majority of metabolites discovered to date are polyketides (PKs), such as the anti-fungal acetyl-CoA carboxylase inhibitor soraphen [110;111], non-ribosomally biosynthesized polypeptides (NRPs), such as the immunosuppressant and proteasome inhibitor argyrin [112;113], or hybrid PK/NRP metabolites. Intriguingly, more than 50% of the myxobacterial compounds incorporate both PK and NRP elements [114]. Interesting representatives are the cytotoxic inducer of tubulin polymerization epothilone [115;116] and the inhibitor of tubulin polymerization disorazol [117;118]. A notable example for structural diversity present within the basic scaffolds of myxobacterial compounds is the leupyrrin family of metabolites. The leupyrrins not only incorporate PK and NRP building blocks but an isoprenoid unit and a dicarboxylic acid [119-121]. In addition, some myxobacterial strains produce the lipophilic carboxylic acids maracin A or maracen A, which contain an ethinyl-*trans*-vinyl ether (maracin A) or an α-chlorovinyl group (maracen A) as a very unusual structural element [122].

A sensible approach for the discovery of novel natural products is the exploitation of the myxobacterial biodiversity in order to find new producers. The phylogenetic classification of the novel myxobacterial isolates also comprises fatty acid analyses, since myxobacteria are equipped with a structural diversity in their fatty acids, which allows for their utilisation as biomarkers [123]. Thereby, it turned out that some myxobacteria are capable of producing PUFAs with a carbon chain length of ≥ C20. The strains *Plesiocystis pacifica*, *Enhygromyxa salina*, and *Pseudenhygromyxa salsuginis* from marine or estuarial environment were characterized for the production of up to 18% AA of TFAs [124-126]. In the terrestrial soil-myxobacteria *Phaselicystis flava* and *Sandaracinus amylolyticus*, even up to 36% AA of TFAs have been found [127;128]. Remarkable amounts of EPA and DHA have been discovered in the novel terrestrial myxobacterial isolates termed *Aetherobacter* [129] (Figure 8). As myxobacteria had turned out to abound in PK biosynthetic pathways, it can be speculated that PKS-like PUFA synthases might be involved in the biosynthesis of these LC-PUFAs.

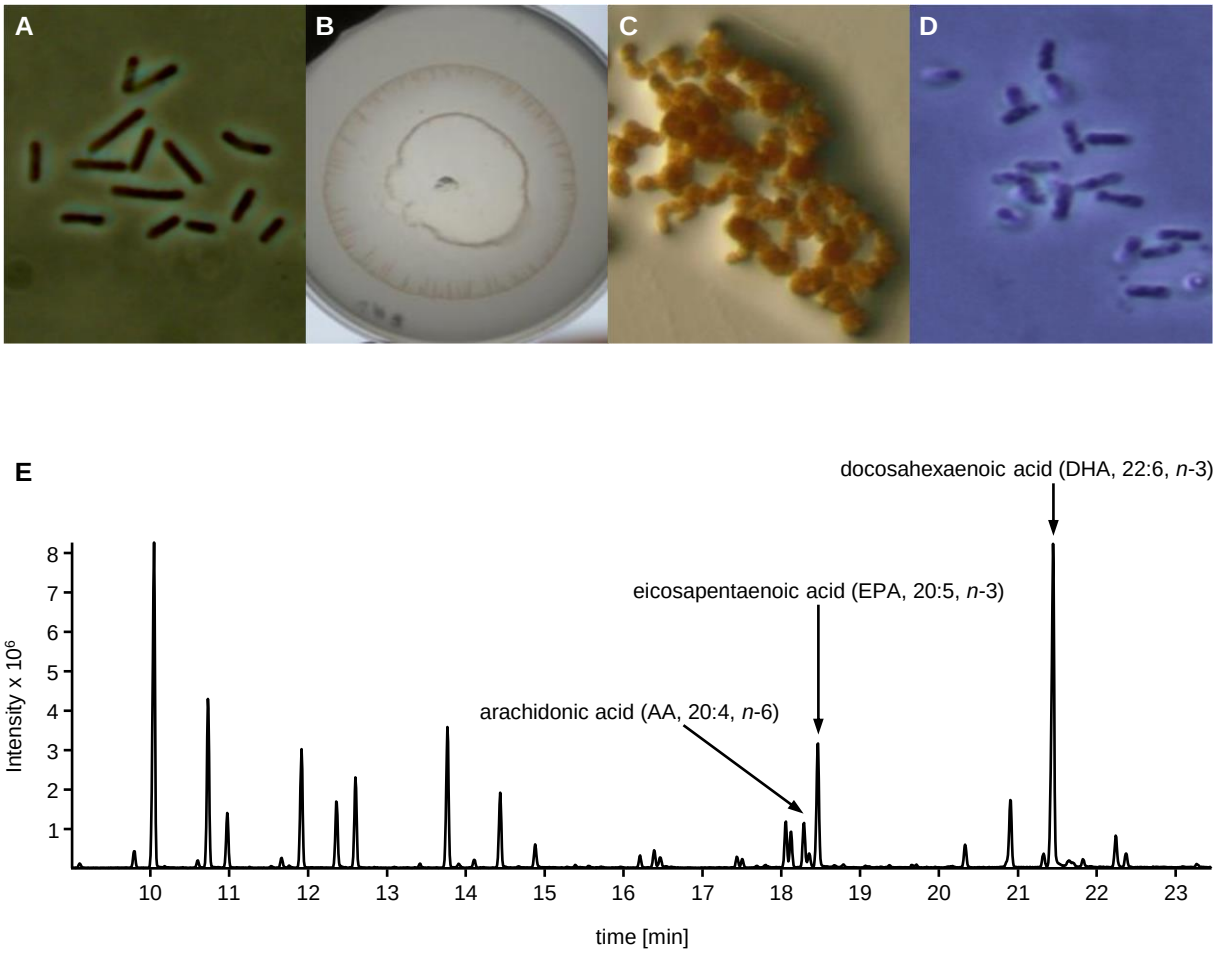


Figure 8. Growth stages and potential for production of polyunsaturated fatty acids (PUFAs) of *Aetherobacter fasciculatus* (SBSr002) [129].
(A) Phase-dark vegetative cells, (B) Dissecting photomicrograph of swarming colony on agar, (C) Dissecting photomicrographs of fruiting bodies on agar, (D) Phase contrast photomicrographs of myxospores from crushed sporangioles, (E) GC-MS chromatogram of fatty acid methyl esters. Photos by courtesy of R. O. Garcia.

Beyond the direct screening for new natural products, hidden biosynthetic potential of the genomes of both well-known and novel myxobacterial strains can be exploited via genome mining. Screening for biosynthetic gene clusters of new secondary metabolites as well as for *pfa*-like gene clusters is considerably facilitated if the complete genome sequence of the producer organism is available. The first myxobacterial genomes that were sequenced represent the genomes of *M. xanthus* DK1622, spanning 9.1 Mbp [130], and the genome of *S. cellulosum* So ce56, spanning 13.1 Mbp [131]. They are among the largest genomes yet known from bacteria. Nowadays, sequencing of the genomes from promising novel myxobacterial isolates is an established procedure, enabled by the rapid development of inexpensive and fast whole genome NGS technologies. Commonly, sequencing of myxobacterial genomes deploys 454 pyrosequencing or Illumina sequencing platforms. Illumina sequencing generates the highest throughput per run and lowest error rate of the

instruments but delivers shorter read lengths than those of the 454 sequencing [132]. However, assembly of the sequencing reads to longer contigs and scaffolds poses a challenge, since the large myxobacterial genomes comprise a plethora of repetitive sequences. *In silico* analysis of the obtained genome sequence often discloses that the myxobacterial strain encodes many more typical secondary metabolite biosynthetic gene clusters than expected based on the number of compounds previously isolated from the strain [130;131]. Different approaches have been elaborated for discovering the metabolic products of the so-called "cryptic" or "orphan" natural product biosynthetic gene clusters found within the genomes [133]. A rather generic approach for the determination of the products of cryptic biosynthetic gene clusters represents a gene knockout plus comparative metabolic profiling [134]. It involves targeted gene inactivations within the biosynthetic gene cluster hypothesized to be essential for metabolite biosynthesis. This process is followed by comparison of the metabolites in the culture supernatants or extracts of the wild type organism and the non-producing mutant using appropriate analytical techniques, such as liquid chromatography-mass spectrometry (LC-MS), and data evaluation supported by statistical tools. Metabolites present in the wild type but missing in the mutant are likely products of the cryptic gene cluster, which can be isolated and structurally characterized. Rhizopodin produced by *Stigmatella aurantiaca* Sg a15 [135] and myxoprincomide produced by *M. xanthus* DK1622 [136] are examples of metabolites successfully discovered using this strategy.

Synthetic Biotechnology approaches for the heterologous expression of polyketide biosynthetic genes – Characterization of the biosynthetic pathways and production enhancement

Many natural PK producers are not readily amenable to a complete range of genetic methodologies without considerable time-consuming effort. Most of them are slow-growing compared with commonly used laboratory strains, often requiring a few weeks for both the accumulation of reasonable amounts of biomass and for a round of genetic manipulation. Based on these findings, lateral transfer of a pathway of interest into a well-developed surrogate host has become an attractive alternative to determine the products of orphan biosynthetic gene clusters, to overproduce natural products, and/or to modify biosynthetic pathways via genetic engineering. In case of the identification of an unknown product, the metabolic profiles of the heterologous host containing and lacking the cloned cryptic biosynthetic gene cluster are compared using appropriate analytical techniques. Metabolites

present in the host containing the gene cluster but absent in the host missing the cluster are likely products of the cryptic biosynthetic pathway, which can be purified and structurally characterized as in the first approach. Flaviolin produced by a type III PKS from *S. cellulosum* So ce56 which was expressed in *Pseudomonas* sp. as heterologous host [137] is an example of a metabolite that has been identified by this method. However, reconstitution of PK biosynthesis in heterologous hosts is challenging and imposes several requirements on the surrogate host. The recombinant genes have to be maintained and stably propagated as well as efficiently expressed via a suitable promoter in the host cell. As PKS genes are often GC-rich, it is additionally important that the GC-content and codon usage bias of the foreign genes and the host genome are matched to ensure a productive heterologous gene expression. The resulting gene products are often large multienzymes that require proper folding and assembly [138] along with posttranslational modification by a PPTase [50] to become functional. Besides, the biosynthetic precursors should be abundant *in vivo* at the correct time, and the host cell has to be protected against the possible toxicity of the overexpressed proteins or biosynthetic products.

The common process of the heterologous expression of a biosynthetic gene cluster has been descriced [139;140] (Figure 9). The workflow typically starts with the isolation and cloning of the relevant genes. Traditionally, genomic libraries are constructed, using either cosmids/fosmids or bacterial artificial chromosomes (BACs), and screened for the genes of interest. Alternatively, new techniques have been developed that enable direct cloning of a gene cluster from the genomic DNA of the native producer into replicative plasmids for *E. coli*. In this case, cloning is accomplished either via transformation-associated recombination (TAR) in *Saccharomyces cerevisiae* [141;142] or via linear plus linear homologous recombination in *E. coli* [143]. The vectors can be further engineered by applying conventional cloning techniques or linear to circular homologous Red/ET recombination [144]. In doing so, partial gene clusters can be reconstituted, genes/domains can be inactivated, deleted, or replaced, and genetic elements adapted to the host organism, such as selection markers, genes for conjugation and/or integration, and heterologous promoters, can be inserted. Many different expression vectors are available for mobilisation and maintenance of the biosynthetic genes in the surrogate host. Depending on the chosen chassis, the genes can be located on autonomously replicating plasmids or on "suicide vectors". The latter ones are not able to replicate in the host and can be designed for either site-specific integration into the genome via homologous recombination at a target locus or phage-mediated recombination at a specific recognition site or for random integration into the

genome via transposition. Despite successful heterologous expression of the biosynthetic genes, production yields of the natural product might be unsatisfactory as a result of insufficient amounts of precursors channelled into the recombinant pathway. This bottleneck can be encountered by precursor feeding or by metabolic engineering of the host organism in order to introduce genes encoding enzymes for precursor biosynthesis.

The most straightforward way is the direct transfer of biosynthetic pathways from the native producers into related host organisms, such as from streptomycetes into related actinomycetes or from a myxobacterial native producer into a myxobacterial host strain. Usually, the original promoter structures, ribosome binding sites, and regulatory elements are functional in the closely related host organisms and can thus be used [145]. Nonetheless, exchange of the native promoter for a strong heterologous promoter often remarkably increases the production yield [140]. On the contrary, the genetically distant model strain *E. coli* is better studied from the physiological point of view and possesses the best developed molecular biology protocols, the most abundant genetic tool box, and well-optimized fermentation protocols using simple and inexpensive media. Additionally, pseudomonads also grow fast with simple nutrient demands, and their metabolic versatility has transformed them into promising organisms for the application in industrial biotechnology [146]. However, a drawback of these organisms is that they are not fertile producers of secondary metabolites. Hence, they often lack the intracellular machinery needed to make a wide range of natural products, such as a PPTase to activate ACPs from PKSs and/or biosynthetic pathways for the generation of special substrates for the PKSs. This might be a reason why there is a general tendency towards higher heterologous production levels of PKs when the host is phylogenetically close to the native producing strain [140]. Moreover, the high GC-content of PKS genes from actinomycetes and myxobacteria leads to an inappropriate bias in codon usage that usually causes poor translation of the proteins in host organisms with less GC-rich genomes. Ergo, engineering of natural PKS gene clusters and regulatory DNA sequences located on expression constructions via recombinant DNA technologies in order to optimally address the demands of a specific host organism is challenging and associated with several limitations.

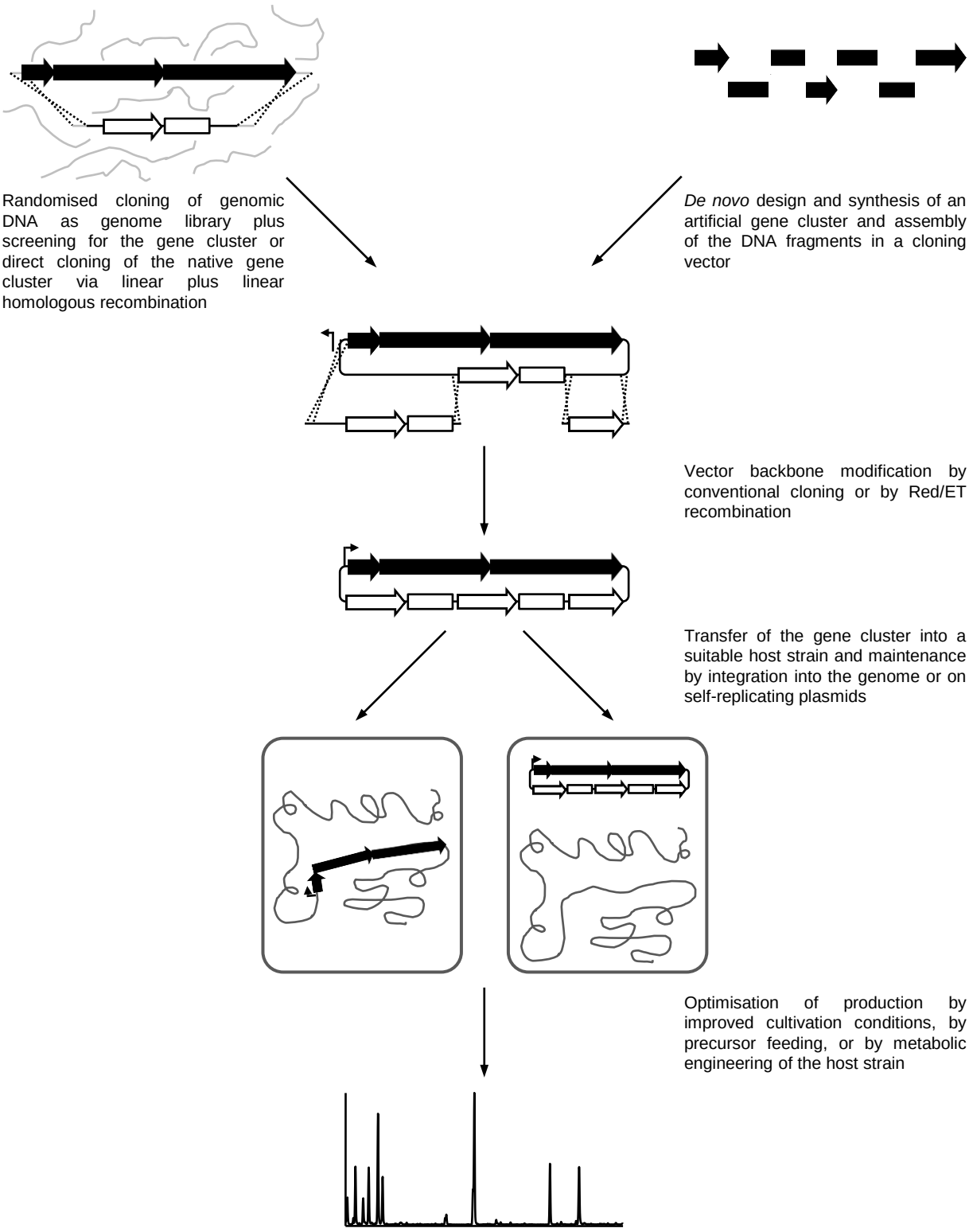


Figure 9. Simplified workflow for the heterologous expression of either a native or a *de novo* designed and synthesized biosynthetic gene cluster.

Advances in sequencing and error-free chemical synthesis of long DNA molecules enable sophisticated bottom-up engineering and construction of large biosynthetic pathways independent from physically existing DNA from natural sources. This emerging field of

research is subsumed under the term “Synthetic Biology”. By applying engineering principles to biology, new biological parts, such as enzymes, genetic circuits, and cells, can be designed as well as constructed, and existing, natural biological systems can be redesigned for useful purposes (Figure 9). The biotechnological production of huge amounts of artemisinin precursors in transgenic strains of both *E. coli* [147] and *S. cerevisiae* [148;149] is the most prominent example of how this approach improves manufacturing of natural products. Means of Synthetic Biology can also be applied to PK biosynthetic pathways. In practical terms, DNA encoding various sections of PKS genes can be synthesized inclusive of desired manipulations of the codon bias plus deletion or insertion of restriction sites without changing the encoded protein sequence. In this way, the complete erythromycin biosynthetic pathway, comprising 32 kb of contiguous DNA, was synthesized and assembled using conventional cloning techniques. Eventually, the artificial gene cluster was functionally expressed in *E. coli* [150;151]. Similarly, synthetic 58 kb epothilone biosynthetic pathways were constructed and functionally expressed in *E. coli* [138] and *M. xanthus* [152]. Unfortunately, the achieved production titres were found very low [152], even after engineering of the *E. coli* host strain [153]. Nevertheless, this method is very powerful and provides the possibility to generate a toolbox of several artificial PKS genes from different biosynthetic gene clusters. The combination of the synthetic genes or intragenic regions can then generate novel hybrid pathways to produce new compounds.

Outline of the dissertation

The aim of this thesis was to increase the body of knowledge of myxobacterial PUFA biosynthesis and to enhance the PUFA production yield by heterologous expression of the PUFA biosynthetic pathways. In doing so, PUFA biosynthetic pathways from different myxobacteria were identified and characterized. Either the native pathways were cloned by genetic engineering or artificial pathways were synthesised and assembled prior to their transfer and expression into various host organisms. The PUFA production rate could be further improved by metabolic engineering of the heterologous host.

PUFA biosynthetic pathways from several terrestrial myxobacteria were deciphered and analysed *in silico* in comparison to PUFA biosynthetic pathways described from marine γ-Proteobacteria and microalgae. Two distinct types of PUFA biosynthetic gene clusters encoding PUFA synthases were identified, originating from linoleic acid producing

myxobacteria of the genus *Sorangium* as well as from species of the recently discovered myxobacterial genus *Aetherobacter*, that turned out to be prolific producers of EPA and DHA. This most likely represents the first detailed characterization of LC-PUFA production based on *pfa* gene clusters of non-marine origin. Evaluation of PUFA production profiles included the successful establishment of heterologous expression systems using *M. xanthus* as host strain. The insights gained by this study are presented in Chapter 1.

Additionally, *S. cellulosum* strains not only producing LA but also PUFAs with a longer chain, mainly EPA, were identified. The corresponding novel PUFA biosynthetic pathway from terrestrial myxobacteria was deciphered and analysed *in silico* in comparison to the formerly known myxobacterial PUFA pathways. Evaluation of LC-PUFA production profiles, including application of the formerly established heterologous expression system using *M. xanthus* as host strain for selected pathways and hybrids thereof, provided indications how myxobacteria employ three different types of PUFA synthases to generate different fatty acid profiles. These findings are summarised in Chapter 2.

Heterologous expression of the *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002) was accomplished in the phylogenetically distant hosts *E. coli* and *Pseudomonas putida*. Using the same initial expression constructs, *P. putida* turned out to be the host strain with the higher heterologous production rate of *n*-6 DPA and DHA *per se*. Consequently, efforts were undertaken to further improve the yield of LC-PUFAs in *P. putida*. Modification of the expression constructs to increase the expression rate of the *pfa* genes as well as genetic and metabolic engineering of the strain then led to success, as illustrated in Chapter 3.

Since advantages and potential of Synthetic Biology approaches for heterologous expression systems appear attractive and promising, they were applied to the optimization of recombinant LC-PUFA production. Hence, artificial *pfa* gene clusters originating from *A. fasciculatus* (SBSr002), responsible for the biosynthesis of mainly DHA, were redesigned, synthesized, and assembled for *P. putida*. Successful transfer of the expression constructs and heterologous expression accomplished in the engineered host strain. This instance most probably represents the first polyketide biosynthetic gene cluster adapted and synthesized for this microbe and is described in Chapter 4.

Modern techniques connected with Synthetic Biology provide the opportunity to transfer and express entire gene clusters from a certain organism into entirely unrelated organisms with very different evolutionary histories. For this purpose, an artificial *pfa* gene cluster plus a

PPTase gene originating from *A. fasciculatus* (SBSr002) responsible for the biosynthesis of mainly DHA were redesigned, synthesized, and assembled for the oleaginous yeast *Yarrowia lipolytica*. Successful transfer of the expression construct and heterologous expression was accomplished in the host strain. It is highly likely that this example, reported in Chapter 5, represents the first polyketide biosynthetic gene cluster adapted and synthesized for this yeast.

The essentiality of selected Pfa protein domains as well as the factor determining the chain length of the produced LC-PUFAs was investigated by heterologous expression of mutated or chimeric versions of myxobacterial *pfa* gene clusters in *E. coli*, *P. putida*, and *Y. lipolytica*. In addition, efforts were undertaken to prove the function of the unique 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) domain from myxobacterial PUFA synthases *in vitro*. The results of these studies are displayed in Chapter 6.

CHAPTER 1

Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity

Gemperlein K, Rachid S, Garcia RO, Wenzel SC, & Müller R (2014)
Chem. Sci., **5,** 1733-1741.

ABSTRACT

Polyunsaturated fatty acids (PUFAs), particularly the omega-3 long-chain PUFAs (LC-PUFAs) eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), are well known for their beneficial health effects. The obvious limitation of the present EPA/DHA key source, fish oil, demands for alternative and sustainable PUFA resources and several biotechnological approaches addressing this problem are currently under development. Different marine microorganisms are known to produce PUFAs *de novo* under strictly anaerobic conditions employing polyketide synthase (PKS)-like enzymes known as PUFA synthases. Here, we report for the first time the characterization of such PUFA synthases from terrestrial origin. Two distinct types of PUFA biosynthetic gene clusters were discovered, originating from linoleic acid producing myxobacteria of the genus *Sorangium* as well as from species of the recently discovered myxobacterial genus *Aetherobacter*, that turned out to be prolific producers of EPA and DHA. The identified biosynthetic pathways differ significantly from the marine systems in terms of gene organization, catalytic domain arrangement, and sequence identity of the encoded PUFA synthases. Notably, a unique domain, which most likely acts as 1-acylglycerol-3-phosphate *O*-acyltransferase, was identified in these myxobacterial PUFA synthases. As the native producer strains grow slowly, are difficult to handle, and genetic modification has proven difficult, synthetic biotechnology approaches were applied to establish a heterologous production platform in the myxobacterial model strain *Myxococcus xanthus*.

INTRODUCTION

Long-chain polyunsaturated fatty acids (LC-PUFAs), including *n*-3 and *n*-6 long-chain fatty acids, are inherent constituents of nutritional research and have gained increasing scientific interest over the past decade. Numerous studies address beneficial effects of LC-PUFAs on human health. In particular, the *n*-3 LC-PUFAs eicosapentaenoic acid (EPA, 20:5, *n*-3) and docosahexaenoic acid (DHA, 22:6, *n*-3) are associated with prevention and treatment of cardiovascular diseases, obesity, and diabetes due to their blood pressure-lowering and anti-inflammatory properties [1]. They have been shown to inhibit the formation of pro-inflammatory eicosanoids derived from *n*-6 long-chain fatty acids, such as arachidonic acid (AA, 20:4, *n*-6), and to promote the formation of endogenous anti-inflammatory and pro-resolving lipid mediators [2;3].

As awareness towards beneficial properties of *n*-3 LC-PUFAs increases, the demand for dietary PUFA supplements has risen intensely over the past few years and will most likely continue to grow significantly. Fish oil is the most abundant and widely used natural source for EPA and DHA these days. However, since the supply of high quality fish oil is significantly declining, it is necessary to establish alternative and sustainable biological sources [4]. Of great interest are the prokaryotic and eukaryotic microorganisms capable of producing LC-PUFAs in high amounts [5]. To make use of their biosynthetic capacity and to develop suitable biotechnological production processes, the characterization of the involved biosynthetic pathways is of crucial importance. The most prominent and predominant pathways in eukaryotic PUFA producers represent the aerobic pathways, that combine several oxygen-dependent desaturases and elongases for the conversion of saturated fatty acids into PUFAs [6]. An alternative pathway was discovered in marine microbes that employs iterative type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like enzymes (PUFA synthases) for *de novo* biosynthesis of LC-PUFAs from acyl-CoA precursors under strictly anaerobic conditions [7-9] (Figure 2A). The biosynthetic machineries are encoded by PUFA (*pfa*) biosynthetic gene clusters, initially identified and characterized from marine γ-Proteobacteria and microalgae: *Shewanella pneumatophori* SCRC-2738 and *Photobacterium profundum* SS9 both produce EPA [10;11], *Moritella marina* MP-1 is known as DHA producer [12], whereas *Schizochytrium* sp. ATCC 20888 produces both DHA and *n*-6 docosapentaenoic acid (DPA, 22:5) [10;13]. The increasing availability of genome sequence data facilitates the screening for *pfa*-like gene clusters in other microbes. In a phylogenomic approach, the abundance and wide distribution of *pfa* gene homologues, especially in the marine habitat, could be demonstrated [14]. As a consequence, at least 20 distinct types of *pfa*-like gene clusters have already been found in numerous microbial genera. However, for most of them, the biosynthetic products remain uncharacterized [15].

A promising strategy to correlate biosynthetic pathway sequences with products is their transfer and expression in suitable heterologous host organisms, which was also applied on several marine PUFA pathways. Heterologous expression of *pfa* gene clusters from *Shewanella pneumatophori* SCRC-2738, *M. marina* MP-1, or *Schizochytrium* sp. in *Escherichia coli* and *Synechococcus* sp. led to the recombinant production of EPA, *n*-6 DPA, and/or DHA and hereby enabled the correlation of the *pfa* gene cluster with the respective LC-PUFA product [13;16-18]. In addition, the established expression systems provide a valuable basis for a more detailed characterization of the biosynthetic pathways and the

opportunity to further exploit these systems for the establishment of biotechnological PUFA production processes [14].

Here we discuss our studies on PUFA biosynthesis in myxobacteria, which represent, to the best of our knowledge, the first detailed characterization of LC-PUFA production based on *pfa* gene clusters of non-marine origin. PUFA biosynthetic pathways from several terrestrial myxobacteria were deciphered and analysed *in silico* in comparison to the described marine PUFA pathways. Evaluation of PUFA production profiles including the establishment of heterologous expression systems for selected pathways revealed insights in how myxobacteria employ two different types of PUFA synthases to generate different fatty acid profiles.

EXPERIMENTAL PROCEDURES

Culture conditions

Aetherobacter sp. (SBSr001), *Aetherobacter fasciculatus* (SBSr002), *Aetherobacter rufus* (SBSr003), and *Aetherobacter* sp. (SBSr008) were obtained from the Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Germany. They were grown in liquid MD1G-5gC-medium containing 0.5% casitone, 0.05% $CaCl_2$ x 2 H_2O, 0.2% $MgSO_4$ x 7 H_2O, and 0.35% glucose. The medium was adjusted to pH 7.0 with KOH. The cultures were incubated on a rotary shaker for 5-10 days at 30 °C and 200 rpm.

S. cellulosum So ce56, So ce10, So ce1525, So ce377, So ce38, So ce487, So ce836, and So ceGT47 were generously provided by the Department of Microbial Drugs, Helmholtz Centre for Infection Research, Braunschweig, Germany. They were grown in liquid M-medium containing 1% phytone peptone, 1% maltose, 0.1% $CaCl_2$ x 2 H_2O, 0.1% $MgSO_4$ x 7 H_2O, and 50 mM HEPES. The medium was adjusted to pH 7.2 with NaOH and supplemented with 8 mg/l NaFe-EDTA after autoclaving. The cultures were incubated on a rotary shaker for 5-10 days at 30 °C and 200 rpm.

Myxococcus xanthus DK1622 [19] was grown in CTT-medium or on CTT-agar containing 1% casitone, 10 mM Tris, 1 mM KPO_4 buffer, 8 mM $MgSO_4$ (and 1.5% agar). The medium was adjusted to pH 7.6. The cultures were incubated for 2-3 days at 30°C (and 200 rpm).

Escherichia coli DH10B [20] was used for cloning experiments. *E. coli* GB05-dir [21] was used for direct cloning experiments. *E. coli* HS996/pSC101-BAD-gbaA (tet^R) [22] was used for modification of a plasmid using Red/ET recombination. The cells were grown in LB-medium or on LB-agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl (1.5% agar)) at 30-37 °C (and 200 rpm) overnight. Antibiotics were used at the following concentrations: 100 µg/ml ampicillin, 80 µg/ml spectinomycin, 50 µg/ml kanamycin, 20 µg/ml zeocin, and 6 µg/ml tetracycline.

Isolation of genomic DNA from myxobacteria

100 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. After discarding the supernatant completely, the cells were washed once with 15 ml SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM Tris-Cl pH 7.5), pelleted at 8,000 rpm for 10 min at room temperature, and resuspended in 5 ml SET buffer afterwards. 50 µl RNase A stock solution (10 mg/ml in sterile ddH_2O) and 300 µl Proteinase K solution (10 mg/ml in 50 mM Tris-Cl pH 8.0, 1 mM $CaCl_2$) were added, and the tube was inverted several times. Subsequently, 600 µl 10% SDS were added, and the tube was incubated first at 41 °C for 1.5 h and afterwards at 55 °C for 1 h under rotation in a hybridization oven. Thereafter, one volume phenol:chloroform:isoamyl alcohol (25:24:1) was added, and the mixture was incubated at 5 rpm for 1 h on a tube rotator. The tube was centrifuged at 8,000 rpm for 5 min at room temperature, and the upper phase was transferred into a new tube by using an end-cut 1 ml-tip. The extraction step with phenol:chloroform:isoamyl alcohol (25:24:1) was repeated once. Then, one volume chloroform:isoamyl alcohol (24:1) was added to the supernatant, and the mixture was incubated at 5 rpm for 1 h on a tube rotator. The tube was centrifuged at 14,000 rpm for 10 min at room temperature, and 4 ml of the upper phase were transferred into a new tube by using an end-cut 1 ml-tip. 440 µl 3 M sodium acetate pH 7.5 were added, and the tube was inverted several times. After adding 11 ml ice-cold ethanol, the tube was inverted until the appearance of cotton-like DNA. The DNA was collected by wrapping around the tip of a Pasteur pipette. In order to wash the DNA, the tip of the pipette with DNA was immersed into a 2 ml tube containing 70% ethanol. The DNA pellet on the pipette was

dried by carefully attaching the inner site of a fresh 2 ml tube to remove the ethanol drops and suspended in 0.5-1 ml of 10 mM Tris-Cl pH 8.0.

General molecular biology techniques

Routine handling of nucleic acids, such as isolation of genomic and plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, were performed according to standard protocols [23]. All the enzymes were purchased from Thermo Scientific or New England Biolabs.

Cloning of expression constructs for heterologous PUFA production using linear plus linear homologous recombination and linear plus circular homologous recombination

For direct cloning of the *pfa* gene cluster from *S. cellulosum* So ce56, a 3.3 kb vector backbone containing the colE1 origin of replication, an ampicillin resistance gene, a tetracycline promoter plus a gene encoding the TetR transcriptional regulator, as well as the appropriate homology arms were amplified from plasmid pGB-amp-P_{tet} [21] using primers So ce-PKS-ET-up and So ce-PKS-ET-dn (Table 1). PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 20 s at 60 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. For linear plus linear homologous recombination, 1.4 ml LB-medium in a 1.5 ml tube were inoculated with a GB05-dir overnight culture and incubated in a heating block at 37 °C with shaking at 1,000 rpm until OD_{600} ~0.2. Afterwards, expression of recETgA was induced by adding 20 µl of 20% L-arabinose, and the incubation was continued at 37 °C and 1,000 rpm until OD_{600} ~0.5. The cells were harvested at 9,000 rpm for 30 s at 2 °C. After discarding the supernatant, the cells were resuspended in 1 ml ice-cooled sterile ddH_2O on ice. The cells were spun down at 9,000 rpm for 30 s at 2 °C, and the supernatant was discarded. The washing step was repeated once. The supernatant was discarded leaving around 20-30 µl sterile ddH_2O in the tube, and the cells were resuspended with it. 0.5-1 µg of the PCR product plus 3-6 µg of genomic DNA of So ce56, completely digested with *Nhe*I and *Sca*I, were added. The cells were electroporated in an ice-cooled electroporation cuvette (1 mm) using an Eppendorf electroporator at 1,300 V. 1 ml

LB-medium was added into the cuvette, and the cell suspension was transferred back into the tube. After incubation in a heating block at 37 °C with shaking at 1,000 rpm for about 70 min, the cells were plated onto LB-agar containing 100 µg/ml ampicillin. The plates were incubated at 37 °C overnight. The colonies that grew under selection for the antibiotic resistance gene were examined for the intended direct cloning product pGB-P_{tet}-Soce-PKS (Table 2) by restriction analysis with a set of different enzymes. Further modification of the vector backbone was performed by Red/ET recombineering using a DNA fragment containing the mariner transposase gene with corresponding inverted repeats, a blasticidin resistance gene, an origin of transfer, the p15A origin of replication, a neomycin resistance gene driven by Tn5 promoter, as well as the appropriate homology arms derived from p15A-epo-IR-Tps-bsd-oriT-IR-kan [24] digested with *Bam*HI. For linear plus circular homologous recombination, 1.4 ml LB-medium in a 1.5 ml tube were inoculated with a HS996/pSC101-BAD-gbaA overnight culture harbouring the plasmid to be modified and incubated in a heating block at 30 °C with shaking at 1,000 rpm until OD_{600} ~0.2. Afterwards, expression of the recombinases was induced by adding 14 µl of 20% L-arabinose, and the incubation was continued at 37 °C and 1,000 rpm until OD_{600} ~0.4. The cells were harvested at 11,000 rpm for 30 s at 2 °C. After discarding the supernatant, the cells were resuspended in 1 ml ice-cooled sterile ddH_2O on ice. The cells were spun down at 11,000 rpm for 30 s at 2 °C, and the supernatant was discarded. The washing step was repeated once. The supernatant was discarded leaving around 20-30 µl sterile ddH_2O in the tube, and the cells were resuspended with it. 1.5 µl of the linear fragment with homology arms were added. The cells were electroporated in an ice-cooled electroporation cuvette (1 mm) using the Eppendorf electroporator at 1,350 V. 1 ml LB-medium was added into the cuvette, and the cell suspension was transferred back into the tube. The tube was incubated in a heating block at 37 °C with shaking at 1,000 rpm for 70 min. Thereafter, the cells were plated onto LB-agar containing 100 µg/ml ampicillin and 50 µg/ml kanamycin. The plates were incubated at 37 °C overnight. The colonies that grew under selection for the antibiotic resistance genes were examined for the intended Red/ET cloning product pTps-p15A-Soce-PKS (Table 2) by restriction analysis and sequencing.

Direct cloning of the *pfa* gene clusters from *Aetherobacter fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008) was carried out via a two-step, double recombination 'fishing' strategy. For the first linear plus linear homologous recombination, a 2.1 kb vector backbone containing the p15A origin of replication, an ampicillin resistance gene, as well as the appropriate homology arms were amplified from plasmid pACYC177 (New England Biolabs) using primers HA+pACYC177_fwd and HA+pACYC177_rev (Table 1). PCR was performed

with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. For the succeeding linear plus circular homologous recombination, a 1.2 kb DNA fragment containing a spectinomycin resistance gene with *Pac*I restriction sites as well as the appropriate homology arms was amplified from plasmid pR6K-amp-spec (Gene Bridges) by PCR using primers HA+*Pac*I+spectinomycin_fwd and HA+*Pac*I+spectinomycin_rev (Table 1). PCR was performed with the PCR extender system (5 Prime) and standard conditions according to the manufacturer's protocol. The reactions were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 33 cycles consisting of denaturation for 1 min at 95 °C, annealing for 1 min at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. Linear plus linear homologous recombination was performed as described previously using 0.5-1 µg of the PCR product plus 3-6 µg of genomic DNA of *Aetherobacter fasciculatus* (SBSr002) or *Aetherobacter sp.* (SBSr008), completely digested with *Sca*I. After electroporation the cells were cultivated in 1 ml LB-medium in a heating block at 37 °C with shaking at 1,000 rpm for 70 min. Thereafter, 100 µg/ml ampicillin were added, and the cells were incubated at 37 °C with shaking at 1,000 rpm overnight. For the succeeding linear plus circular homologous recombination the next day, the cells from the first cloning step were used as preculture. Cell cultivation and preparation of electrocompetent cells were performed as described above. 0.5-1 µg of the PCR product was added to the electrocompetent cells. The electroporation conditions were the same as described previously. After incubation in a heating block at 37 °C with shaking at 1,000 rpm for about 70 min, the cells were plated onto LB-agar containing 100 µg/ml ampicillin and 80 µg/ml spectinomycin. The plates were incubated at 37 °C overnight. The colonies that grew under selection for the antibiotic resistance genes were examined for the intended direct cloning products pPfaAf1 and pPfaAs1 (Table 2) by restriction analysis with a set of different enzymes. The spectinomycin resistance gene was subsequently removed via the introduced *Pac*I restriction sites, generating plasmids pPfaAf2 and pPfaAs2 (Table 2). In order to introduce a heterologous promoter upstream of the *pfa* gene cluster in the expression construct pPfaAf2 and to exchange the GTG start codon of gene *pfa1* for an ATG start codon, a 1.2 kb DNA fragment containing a tetracycline promoter downstream of a neomycin resistance gene driven by Tn5 promoter as well as the appropriate homology arms was amplified from

plasmid pTps-p15A-Soce-PKS using primers HA+*neo*^{R}_fwd and HA+P_{tet}_rev (Table 1). PCR was performed with the PCR extender system (5 Prime) and standard conditions according to the manufacturer's protocol. The reactions were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 33 cycles consisting of denaturation for 1 min at 95 °C, annealing for 1 min at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. The linear plus circular homologous recombination was performed as described previously using 1.5 µl of the PCR product, resulting in plasmid pPfaAf-P_{tet} (Table 2). Plasmid pPfaAs-P_{tet} (Table 2) was constructed by exchange of the *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002) located on plasmid pPfaAf-P_{tet} for the *pfa* gene cluster from *Aetherobacter sp.* (SBSr008) located on plasmid pPfaAs2 via *Eco*91I and *Pac*I restriction sites. Afterwards, a zeocin resistance gene and the gene encoding the Mx9 integrase with *attP* site, both driven by Tn5 promoter derived from pMyx-zeo (S. C. Wenzel, unpublished), were inserted via *Pac*I restriction sites, generating plasmids pPfaAf-P_{tet}-mx9.2 and pPfaAs-P_{tet}-mx9.2 (Table 2).

Table 1. Primers used for linear plus linear homologous recombination or linear plus circular homologous recombination.
For each primer, the homology arm is shown in italics and the introduced restriction site is underlined.

Primer	Sequence (5'→3')
So ce-PKS-ET-up	*CCGCTGCGCCAGATGCCGATCGGCTGGAGGCCGGCGCTGCGGATGACCT GCGGATCCTCAATCAT*GCTGGGTGCCTCCTAGATTATTGTCGAACTATT
So ce-PKS-ET-dn	*TGTAGGCATCAAAATCGTCGAGAAATCCCTGTCCGTCCGCCCGCCTACA GCCGCAGATGGGCCGCAAGTATGGCTGTA*GTAAGCTTGACCTGTGAAGT GAAAAAT
HA+pACYC177_fwd	*CAATATTTCGACGACCAGCGCTGCGCGGCGAAATATGGCGCGCTCTGGG G*CATCAGAAGGGCACTGGTGCAACGG
HA+pACYC177_rev	*CGGGCGCCGCGGGATCCCGTACATCTACTTCGACGTGCTCGACGAGGAG A*TCGCTCACTGACTCGCTACGCTCGG
HA+*Pac*I+spectinomycin_fwd	*CTAAGGCCCCTTGTGCCCATTGAGGCACACGTCGAACGATTTCACCGTC T*TTAATTAAGGGAACAGAAAACGGATACCAAGGCG
HA+*Pac*I+spectinomycin_rev	*ATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCG A*TTAATTAACTGCATCCGATGCAAGTGTGTCGC
HA+*neo*^{R}_fwd	*CCCCAGAGCGCGCCATATTTCGCCGCGCAGCGCTGGTCGTCGAAATATT G*GACAGCAAGCGAACCGGAATTGCC
HA+P_{tet}_rev	*GGCGAGCCCCCTTGAGCGTTCCATCTACCGATTGCAGACA*TGCTGGGTG CCTCCTAGATTATTGTCG

Table 2. Plasmids and expression strains constructed in this study.
oriV = origin of replication, *oriT* = origin of transfer, *bla* = ampicillin resistance gene, *BSD* = blasticidin resistance gene, *nptII* = neomycin resistance gene, *aad9* = spectinomycin resistance gene, *ble* = zeocin resistance gene, P_{tet} = tetracycline promoter, *tetR* = gene encoding the tetracycline transcriptional regulator, PTn5 = Tn5 promoter, *tps* = mariner transposase gene, IRs = inverted repeats, Mx9 *int* = Mx9 integrase gene with phage attachment site *attP*.

Plasmid / Strain	Characteristics
pGB-P_{tet}-Soce-PKS	Directly cloned *pfa* gene cluster from *Sorangium cellulosum* So ce56 driven by P_{tet} plus *tetR*. colE1 *oriV*, *bla*
pTps-p15A-Soce-PKS	Derivative of pGB-P_{tet}-Soce-PKS in which colE1 *oriV* was replaced by *tps* plus IRs, *BSD*, *oriT*, p15A *oriV*, and PTn5-*nptII* by Red/ET recombineering
pPfaAf1	Directly cloned *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002). p15A *oriV*, *bla*, *aad9*
pPfaAf2	Derivative of pPfaAf1 in which *aad9* was removed via *Pac*I
pPfaAf-P_{tet}	Derivative of pPfaAf2 in which PTn5-*nptII*-P_{tet} was inserted upstream of the *pfa* gene cluster by Red/ET recombineering
pPfaAf-P_{tet}-mx9.2	Derivative of pPfaAf-P_{tet} in which PTn5-*ble*-Mx9 *int* was inserted via *Pac*I
pPfaAs1	Directly cloned *pfa* gene cluster from *Aetherobacter* sp. (SBSr008). p15A *oriV*, *bla*, *aad9*
pPfaAs2	Derivative of pPfaAf-P_{tet} in which *aad9* was removed via *Pac*I
pPfaAs-P_{tet}	Derivative of pPfaAf-P_{tet} in which the *pfa* gene cluster from *A. fasciculatus* (SBSr002) was replaced by the *pfa* gene cluster from *Aetherobacter sp.* (SBSr008) located on plasmid pPfaAs2 via *Eco*91I and *Pac*I
pPfaAs-P_{tet}-mx9.2	Derivative of pPfaAs-P_{tet} in which PTn5-*ble*-Mx9 *int* was inserted via *Pac*I
M. xanthus DK1622::pTps-p15A-Soce-PKS	*Myxococcus xanthus* DK1622 with pTps-p15A-Soce-PKS randomly integrated into the genome
M. xanthus DK1622::pPfaAf-P_{tet}-mx9.2	*M. xanthus* DK1622 with pPfaAf-P_{tet}-mx9.2 site-specifically integrated at Mx9 *attB2* site of the genome
M. xanthus DK1622::pPfaAs-P_{tet}-mx9.2	*M. xanthus* DK1622 with pPfaAs-P_{tet}-mx9.2 site-specifically integrated at Mx9 *attB2* site of the genome

Transformation of *M. xanthus* by electroporation and heterologous expression

1.5×10^9 cells from an overnight culture of *M. xanthus* DK1622 were harvested at 8,000 rpm for 10 min. After discarding the supernatant, the cells were resuspended in 20 ml sterile ddH_2O. The cells were spun down at 8,000 rpm for 10 min, and the supernatant was discarded. The cells were resuspended in 40 µl sterile ddH_2O and transferred into a 2 ml tube. 2 µl of the expression plasmid were added. The cells were electroporated in an ice-cooled electroporation cuvette (1 mm) using a Bio-Rad electroporator at 650 V, 400 Ω, and 25 µF. 1 ml CTT-medium was added into the cuvette, and the cell suspension was transferred back into the tube. The tube

was incubated in a heating block at 30 °C with shaking at 1,000 rpm for about 5-8 h. Thereafter, the cells were mixed with 2-3 ml CTT-soft agar (CTT-medium plus 0.75% agar) containing 60 µg/ml kanamycin plus 20 µg/ml zeocin and plated onto two CTT-agar plates containing 60 µg/ml kanamycin plus 20 µg/ml zeocin. The plates were incubated at 30 °C for about 5 days. Thereafter, six colonies were transferred onto new CTT-agar plates containing 60 µg/ml kanamycin and 20 µg/ml zeocin. In order to verify the integration of the expression construct into the genome of *M. xanthus* via Mx9 *attB* sites, DNA from selected clones was isolated and used for PCR analyses [25]. The primers used are listed in Table 3. PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 56 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. Expression was carried out in 50 ml CTT-medium containing 60 µg/ml kanamycin and 20 µg/ml zeocin at 30 °C. The cells were cultivated for 48 h and then harvested at 4,000 rpm for 10 min.

Table 3. Primers used for the verification of the integration of expression constructs into the genome of *M. xanthus* DK1622 via Mx9 *attB* sites.

Primer	Sequence (5'→3')
Mx9*attB1*_up	TGCCAGGGCTTACGGCTTC
Mx9*attB1*_down	CAGCACGGGTGCAGCAAC
Mx9*attB2*_up	TATCCCAGCAACCGCCGGAG
Mx9*attB2*_down	CGAGGTCCGGGACGCGCGCA
Mx9*attP*_up	GCGCCGAACTTAACAAGTTG
Mx9*attP*_down	TCCAGGTCCTCACGCTTGAC

Extraction of cellular fatty acids

The cellular fatty acids were extracted using the FAME method [26]. For this purpose, 50-100 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. The cell pellet was transferred to a glass vial and dried in a vacuum concentrator. Subsequently, the cell dry weight was determined. 5 µl (50 µg) of *n*-3 DPA (Sigma-Aldrich) and 500 µl of a mixture of methanol, toluene, and sulphuric acid (50:50:2, v/v/v) were added. The vial was capped with a teflon-lined screw cap and incubated at 80 °C for 24-48 h. After the mixture was cooled to room temperature, 400 µl of an aqueous solution consisting of 0.5 M NH_4HCO_3 and 2 M KCl were added, and the sample was vortexed for 30 s. Phase separation was

achieved by centrifugation at 4,000 rpm for 5 min at room temperature. 75 µl of the upper phase were mixed with 25 µl *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) and incubated at 37 °C for 30 min. Subsequently, the sample was used for GC-MS analysis.

Extraction and fractionation of lipids

Extraction of lipids from microbial cells was carried out using the method of Bligh and Dyer [27], modified by Lewis *et al.* [28], on a small scale. In the first step, the cell pellet from a 100 ml culture was transferred into a polypropylene tube. Successively, 4 ml chloroform, 8 ml methanol, and 3.2 ml 1% NaCl were added, and the tube was vortexed at high speed for 15 s after every addition. The sample was agitated on a tube rotator at 30 rpm overnight. 4 ml chloroform and 4 ml 1% NaCl were then added, and the tube was inverted 30 times. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. The bottom layer containing the lipid extract was evaporated to dryness under a gentle stream of nitrogen and dissolved in 1 ml of a chloroform:methanol mixture (2:1, v/v). The lipid solution was fractionated by solid-phase extraction on a Strata SI-1 silica column with 100 mg sorbent mass (Phenomenex) equilibrated with chloroform + 1% acetic acid. Neutral lipids and free fatty acids were eluted from the column with 1 ml of chloroform + 1% acetic acid, glycolipids were eluted with 1.5 ml of an acetone:methanol mixture (9:1, v/v), and phospholipids were eluted with 1 ml methanol. For GC-MS analysis, the fractions were dried in a vacuum concentrator and further processed according to the FAME method described previously.

Analysis of fatty acid methyl esters by GC-MS

GC-MS was carried out on an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a 7683B split/splitless injector with autosampler (Agilent Technologies) and coupled to a 5973 electron impact mass selective detector (Agilent Technologies). Helium was used as the carrier gas at a flow rate of 1 ml/min. 1 µl of the sample was injected in split mode (split ratio, 10:1). The analytical column was a (5% phenyl)-methylpolysiloxane capillary column (Agilent J&W DB-5ht; 30 m x 0.25 mm i.d. x 0.1 µm film thickness, maximum temperature 400 °C; Agilent Technologies). The column temperature was kept at 130 °C for 2.5 min, increased to 240 °C at a rate of 5 °C/min, then ramped to 300 °C at 30 °C/min, and held at 300 °C for 5 min. Other temperatures were as follows: inlet, 275 °C; GC-MS transfer line, 280 °C; ion source, 230 °C; and quadrupole, 150 °C. The mass selective

detector was operated in scan mode, scanning the mass range from *m*/*z* 40 to 700. Scan control, data acquisition, and processing were performed by MSD ChemStation and AMDIS software, version 2.68, based on the fragmentation patterns and retention times, in comparison with Supelco 37 Component FAME Mix and LC-PUFAs (all Sigma-Aldrich) as reference standards, and NIST 08 library. Absolute amounts of PUFAs were quantified by integration of the peaks using MSD ChemStation and by subsequent calculation in relation to the integral of *n*-3 DPA and to cell dry weight. In case *n*-3 DPA was produced by the culture itself, the sample was spiked with *n*-3 DPA for the calculation of the absolute amounts of PUFAs.

Bioinformatic analysis of DNA and protein sequences

Prediction of open reading frames was performed with FramePlot 4.0 [29], and functional annotation of proteins was based on BlastP [30] and Pfam [31] searches. The Geneious software suite (Biomatters) [32] was used for *in silico* analyses of DNA and protein sequences. All sequence alignments were performed using Geneious and ClustalW [33].

RESULTS AND DISCUSSION

During continuous efforts to extend our myxobacterial strain collection with novel isolates, which optionally includes the analysis of their fatty acid profiles, PUFA producing myxobacteria were discovered [34] (Figure 1). This finding motivated us to screen available genome sequence data for putative PUFA biosynthetic pathways and to start further genome sequencing projects on selected myxobacterial isolates. By applying such a genome-mining approach, a putative *pfa* gene cluster could be identified in the genome sequence of *Sorangium cellulosum* So ce56 (genes sce0818, sce0819, and sce0820) [35]. The same type of *pfa* gene cluster could also be detected in the yet unpublished genome sequence data of the *S. cellulosum* strains So ce10 [36], So ce1525 [37], So ce377 [38], So ce38 [39], So ce487, So ce836, and So ceGT47, as well as in the recently published genome sequence of the alkaline-adaptive *S. cellulosum* strain So0157-2 [40]. GC-MS analysis of the fatty acid methyl esters (FAMEs) was carried out for all *S. cellulosum* strains mentioned above except for So0157-2. Interestingly, PUFA production was shown to be restricted to linoleic acid (LA, 18:2, *n*-6; ca. 10% of total FAMEs) and eicosadienoic acid (EDA, 20:2, *n*-6) in trace amounts (< 1% of total FAMEs) in these strains. In addition to these identified producer strains, the

novel myxobacterial isolates of the genus *Aetherobacter*, belonging to the same bacterial family *Polyangiaceae* as the genus *Sorangium* (Figure 3), were recently described as prolific producers of LC-PUFAs [34;41]. FAME analysis carried out for *Aetherobacter* sp. (SBSr001) nov. ined., *Aetherobacter fasciculatus* sp. nov. ined. (SBSr002), *Aetherobacter rufus* sp. nov. ined. (SBSr003) [42], and *Aetherobacter* sp. nov. ined. (SBSr008) revealed the production of AA (ca. 1% of total FAMEs), EPA (ca. 7-10% of total FAMEs), and DHA (ca. 10-19% of total FAMEs) as well as of LA, γ-linolenic acid (GLA, 18:3, *n*-6), stearidonic acid (SDA, 18:4, *n*-3), *n*-6 DPA, and *n*-3 DPA in trace amounts (< 1% of total FAMEs). Genome sequence data of all four *Aetherobacter* strains were generated allowing for the identification of a putative *pfa* gene cluster in each strain, which turned out to be similar to the *pfa* gene cluster from *S. cellulosum* strains.

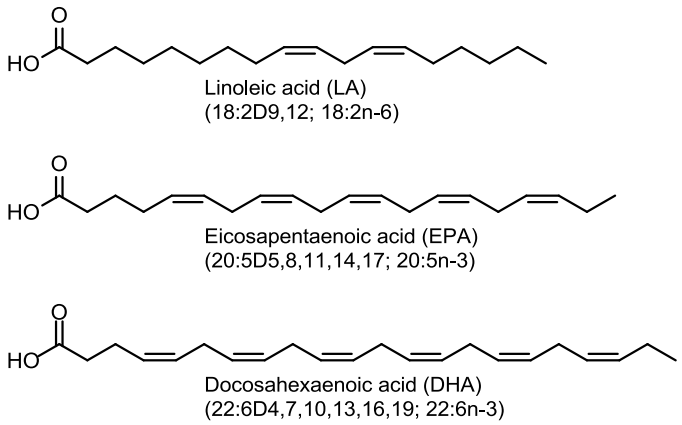


Figure 1. Structures of main polyunsaturated fatty acids (PUFAs) produced by myxobacteria.

However, these identified myxobacterial PUFA biosynthetic gene clusters differ from the previously characterized *pfa* gene clusters from *Shewanella pneumatophori* SCRC-2738, *P. profundum* SS9, *M. marina* MP-1, and *Schizochytrium* sp. ATCC 20888. Unlike the PUFA pathways from the marine γ-Proteobacteria, the two types of myxobacterial *pfa* gene clusters only consist of three genes (*pfa1*, *pfa2*, and *pfa3*) arranged in a different order and encoding multifunctional PUFA synthases that also differ in their catalytic domain composition (Figure 2). The *pfaD* homologue *pfa1*, encoding an enoyl reductase (ER), is located at the 5' end of the myxobacterial gene cluster. The downstream gene, *pfa2*, and its homologues *pfaA* and *PFA1* encode multifunctional PKS enzymes that harbour a variable number of tandem acyl carrier protein (ACP) domains, which actually also differs among the two myxobacterial pathways. Another remarkable variation in the domain architecture is the location of the acyltransferase (AT) domain which is separately encoded by *pfaB* or *PFA2* in PUFA

biosynthetic gene clusters from marine γ-Proteobacteria and microalgae but integrated in *pfa3* from *Aetherobacter* spp.. Interestingly, a corresponding domain could not be identified in the *S. cellulosum* pathway. The probably most significant deviation from the marine PUFA synthases can be observed at the carboxyl-terminus of Pfa3. *In silico* analyses revealed that it most likely contains a 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) (EC 2.3.1.51), which represents a unique and common characteristic of terrestrial myxobacterial PUFA synthases. The function of this enzyme is the acylation of the 2-position of 1-acylglycerol-3-phosphate during glycerophospholipid biosynthesis, utilising either acyl-ACP or acyl-CoA thioesters as substrate [43] (Figure 2B). In addition to the three clustered *pfa* biosynthetic genes, a 4'-phosphopantetheinyl transferase (PPTase) for posttranslational modification of the ACP domains is required. A homologue of the Sfp-type PPTase encoding gene *pfaE,* which is assumed to be involved in the phosphopantetheinylation of PUFA synthases [44], was found at a separate locus within the genome of *Sorangium cellulosum* (sce5058) and *Aetherobacter* spp..

Detailed comparison on the sequence level [45] revealed that the Pfa proteins from strains of the same genus, e.g. *S. cellulosum* So ce56 and So ce377 or *A. fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008), show a pairwise sequence identity of > 90%. Considering the fact that *S. cellulosum* So ce56 and *A. fasciculatus* (SBSr002) belong to two different bacterial genera with distinct PUFA production profiles (Figure 3), it is hardly surprising that the similarity between those strains is significantly lower (pairwise sequence identity of ≤ 60%). One of the most remarkable differences between the two types of myxobacterial *pfa* gene clusters is the presence of an AT domain within Pfa3 from *Aetherobacter* spp., which could not be identified in the *S. cellulosum* pathway (Figure 4). This finding possibly explains the lack of EPA and DHA production in *Sorangium*, a speculation supported by a recent study suggesting that the homologous AT domain encoded by *pfaB* from marine γ-Proteobacteria may play a role in determining the final PUFA product [46]. Comparison of myxobacterial Pfa proteins from *S. cellulosum* So ce56 or *A. fasciculatus* SBSr002 with homologous Pfa protein regions from *Shewanella pneumatophori* SCRC-2738, *P. profundum* SS9, *M. marina* MP-1, or *Schizochytrium* sp. ATCC 20888 disclosed a considerably low pairwise sequence identity of ≤ 45%. These values indicate that the function and the biosynthetic mechanism of the domains within the PUFA synthases of different origins is more conserved than their sequence.

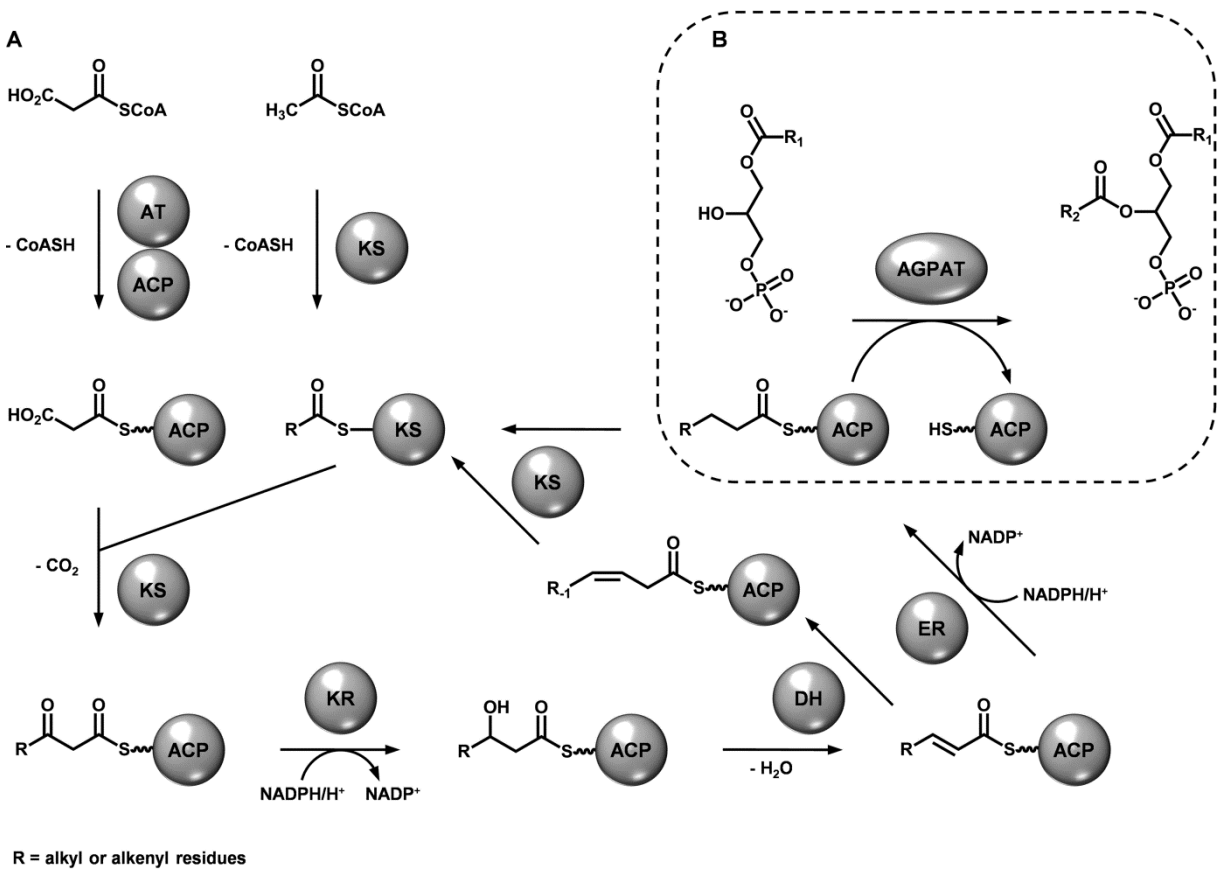


Figure 2. Biosynthetic scheme for the anaerobic biosynthesis of polyunsaturated fatty acids (PUFAs) by iterative type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like PUFA synthases.
(A) The primer molecule acetyl-CoA undergoes several rounds of decarboxylative Claisen condensation reactions, resulting in the elongation of the fatty acyl chain by two carbons derived from malonyl-CoA per cycle. Following each round of elongation, the β-keto group is either fully reduced or only reduced to the *trans* double bond which is then isomerized. Finally, an acyl chain with methylene-interrupted *cis* double bonds is synthesized. (B) Reaction scheme for the acylation of the 2-position of 1-acylglycerol-3-phosphate catalysed by the unique 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) domain from myxobacteria using acyl-ACP as substrate during glycerophospholipid biosynthesis: AT = acyltransferase, ACP = acyl carrier protein, KS = ketosynthase, KR = ketoreductase, DH = dehydratase/isomerase, ER = enoyl reductase.

However, as already highlighted before, a unique characteristic of terrestrial myxobacterial PUFA synthases is the presence of a putative AGPAT domain. This domain might catalyse the direct transfer of the synthesized fatty acid chains from the PUFA synthase into lipids (Figure 2B). Based on this, the majority of the synthesized PUFAs was expected to be bound in the lipid fraction. Therefore, extraction and fractionation of the lipids from *A. fasciculatus* (SBSr002), subsequent preparation of FAMEs by transesterification, and GC-MS analysis was carried out. Indeed, most of the synthesized DHA (54%) was incorporated into the phospholipids. Forty-three per cent of synthesized DHA were present in the fraction containing neutral lipids plus free fatty acids and three per cent in the glycolipid fraction. In contrast, anaerobic PUFA biosynthesis in marine organisms seems to involve the hydrolysis of the synthesized PUFA chains from the biosynthetic complex to produce free fatty acids as

primary products, which can then be further metabolized. For *Schizochytrium* sp. ATCC 20888, it has been postulated that the AT domain encoded by *PFA2* is involved in the release of the synthesized LC-PUFAs from the assembly line as free fatty acids [47]. For *P. profundum* SS9, a thioesterase, encoded in close proximity to the *pfa* gene cluster, was identified and shown to be capable of hydrolysing long-chain fatty acyl-CoA thioester bonds [48]. According to text book biochemistry, free fatty acid chains (released from marine PUFA synthases) have to be activated by conversion into the corresponding acyl-CoA thioesters for incorporation into lipids by acyl transferases like AGPAT. The presence of an integrated AGPAT domain in myxobacterial PUFA synthases, which is assumed to allow for the direct transfer of the biosynthetic products into lipids, seems to be a more efficient way compared to the marine systems.

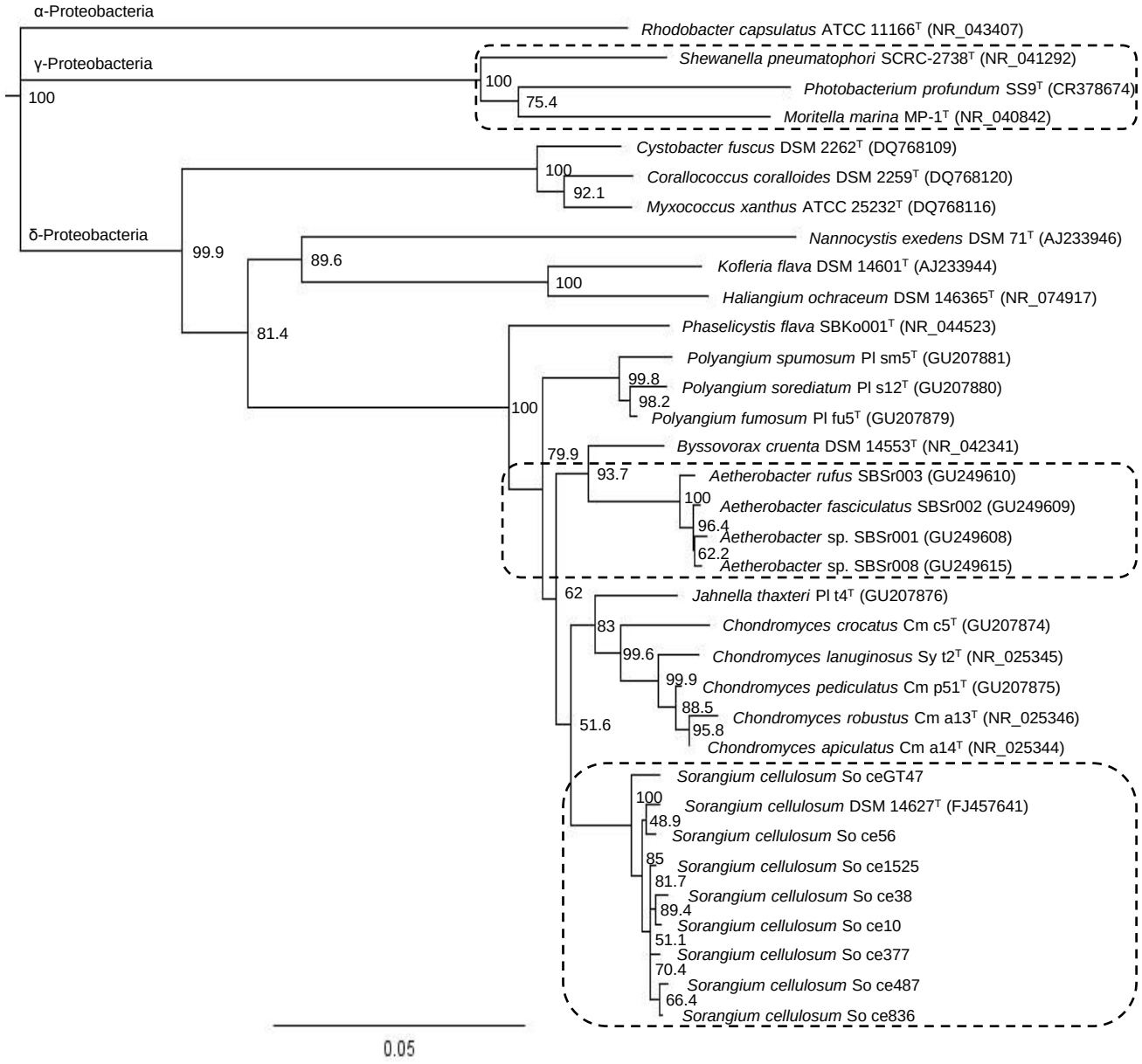


Figure 3. Neighbour-joining tree of myxobacteria and selected marine γ-Proteobacteria inferred from 16S rRNA gene sequences showing the clades of PUFA producing strains with *pfa* gene clusters (encircled by dashed lines).

The α-Proteobacterium *Rhodobacter capsulatus* ATCC 11166T was used as an outgroup to root the tree. GenBank accession number is indicated in parenthesis. The numbers at branch points indicate the per cent bootstrap support based on 1000 replicates. Bar, 0.05 substitutions per nucleotide position.

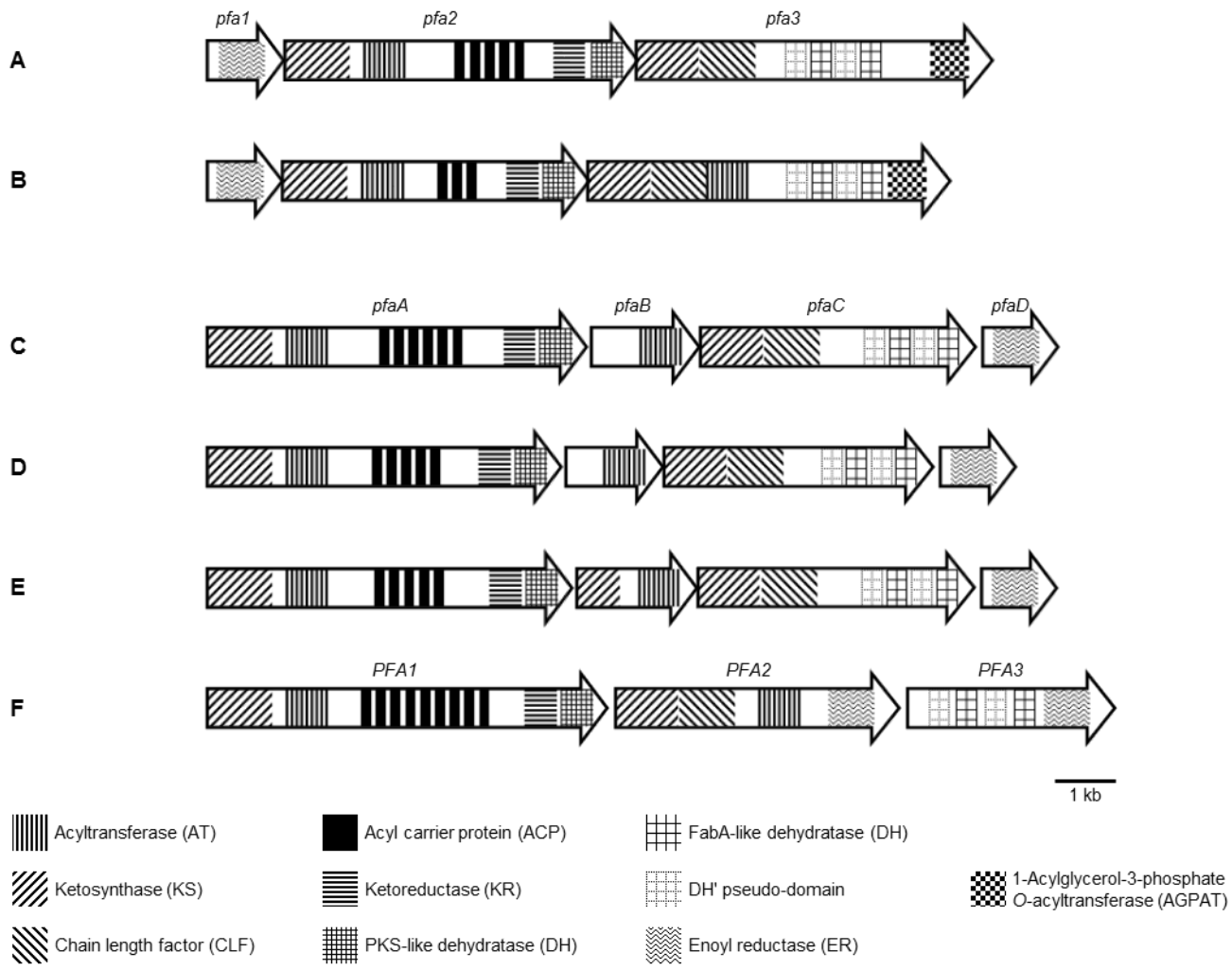


Figure 4. PUFA biosynthetic gene clusters from terrestrial myxobacteria and marine γ-Proteobacteria or microalgae.
(A) *Sorangium cellulosum* So ce56 and So ce377 (produce mainly LA), (B) *Aetherobacter* sp. (SBSr001), *Aetherobacter fasciculatus* (SBSr002), *Aetherobacter rufus* (SBSr003), and *Aetherobacter* sp. (SBSr008) (produce mainly AA, EPA and DHA), (C) *Shewanella pneumatophori* SCRC-2738 (EPA producer), (D) *Photobacterium profundum* SS9 (EPA producer), (E) *Moritella marina* MP-1 (DHA producer), (F) *Schizochytrium* sp. ATCC 20888 (DHA and *n*-6 DPA producer). Two inactive DH' pseudo-domains, which have been recently discovered in *pfaC* of *P. profundum* SS9 [49], were identified adjacent to the FabA-like DH domains in every depicted gene cluster.

In order to prove the functionality of the identified myxobacterial *pfa* gene clusters and to further exploit these systems for biotechnological PUFA production, we aimed to establish suitable synthetic biotechnology approaches. As the native producers, especially the strains of the recently discovered novel genus *Aetherobacter*, are not easy to handle, grow slowly, and are not accessible for genetic manipulation, we intended to express some of the identified pathways in a suitable host strain. As discussed in more detail below, heterologous expression of the *pfa* gene clusters from *S. cellulosum* So ce56, *A. fasciculatus* (SBSr002), and *Aetherobacter* sp. (SBSr008) was established in the myxobacterial model strain *M. xanthus*

DK1622, which required their initial cloning from chromosomal DNA of the native producer strains into plasmids and subsequent engineering of the constructs (Figure 5)**.** For that reason, a recently described linear plus linear homologous recombination technique was applied, which allows for the direct cloning of a target fragment from a mixture of chromosomal DNA fragments [21]. By using this technology, the 17.2 kb *pfa* gene cluster from *S. cellulosum* So ce56 was cloned into a suitable vector backbone. During this step, the native promoter upstream of *pfa1* was replaced with a tetracycline promoter. Further modification of the obtained plasmid pGB-P_{tet}-Soce-PKS was performed by linear plus circular homologous Red/ET recombination [50]. Thereby, the mariner transposase gene with corresponding inverted repeats and an additional Tn5 promoter located upstream of an adjacent neomycin resistance gene were inserted. The final expression construct pTps-p15A-Soce-PKS can be integrated randomly into the host chromosome by transposition. In parallel, the 16.2 kb *pfa* gene clusters from *A. fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008) were cloned applying a modified strategy: a two-step, double recombination 'fishing' approach [21]. The *pfa* gene clusters were cloned by linear plus linear homologous recombination from chromosomal DNA into a minimal vector backbone, in a similar manner as described before. However, immediately after this first cloning step, a second linear plus circular homologous recombination step was carried out to 'fish' for correct direct cloning products via a specific recombination event with a second resistance gene cassette. This modified procedure extremely facilitates the identification of correct recombinant constructs. After isolation and verification of the obtained plasmids pPfaAf1 (containing the *pfa* gene cluster from SBSr002) and pPfaAs1 (containing the *pfa* gene cluster from SBSr008), the spectinomycin resistance cassette was deleted via restriction digest and religation to generate the derivatives pPfaAf2 and pPfaAs2. The construct pPfaAf2 was modified by an additional linear plus circular homologous recombination step to introduce the tetracycline promoter downstream of the neomycin resistance gene driven by Tn5 promoter from pTps-p15A-Soce-PKS upstream of *pfa1*. At the same time the GTG start codon of *pfa1* was exchanged for the more common ATG start codon, resulting in plasmid pPfaAf-P_{tet}. A similar plasmid, pPfaAs-P_{tet}, was constructed by applying conventional cloning techniques to exchange the *pfa* gene cluster from *A. fasciculatus* (SBSr002) located on plasmid pPfaAf-P_{tet} for the *pfa* gene cluster from *Aetherobacter* sp. (SBSr008) located on plasmid pPfaAs2. In a final engineering step, a zeocin resistance gene and a gene encoding the Mx9 integrase with *attP* site [51], allowing for the site-specific integration into the host genome via the Mx9 phage attachment site, were

ligated into both constructs, yielding the expression plasmids pPfaAf-P_{tet}-mx9.2 and pPfaAs-P_{tet}-mx9.2.

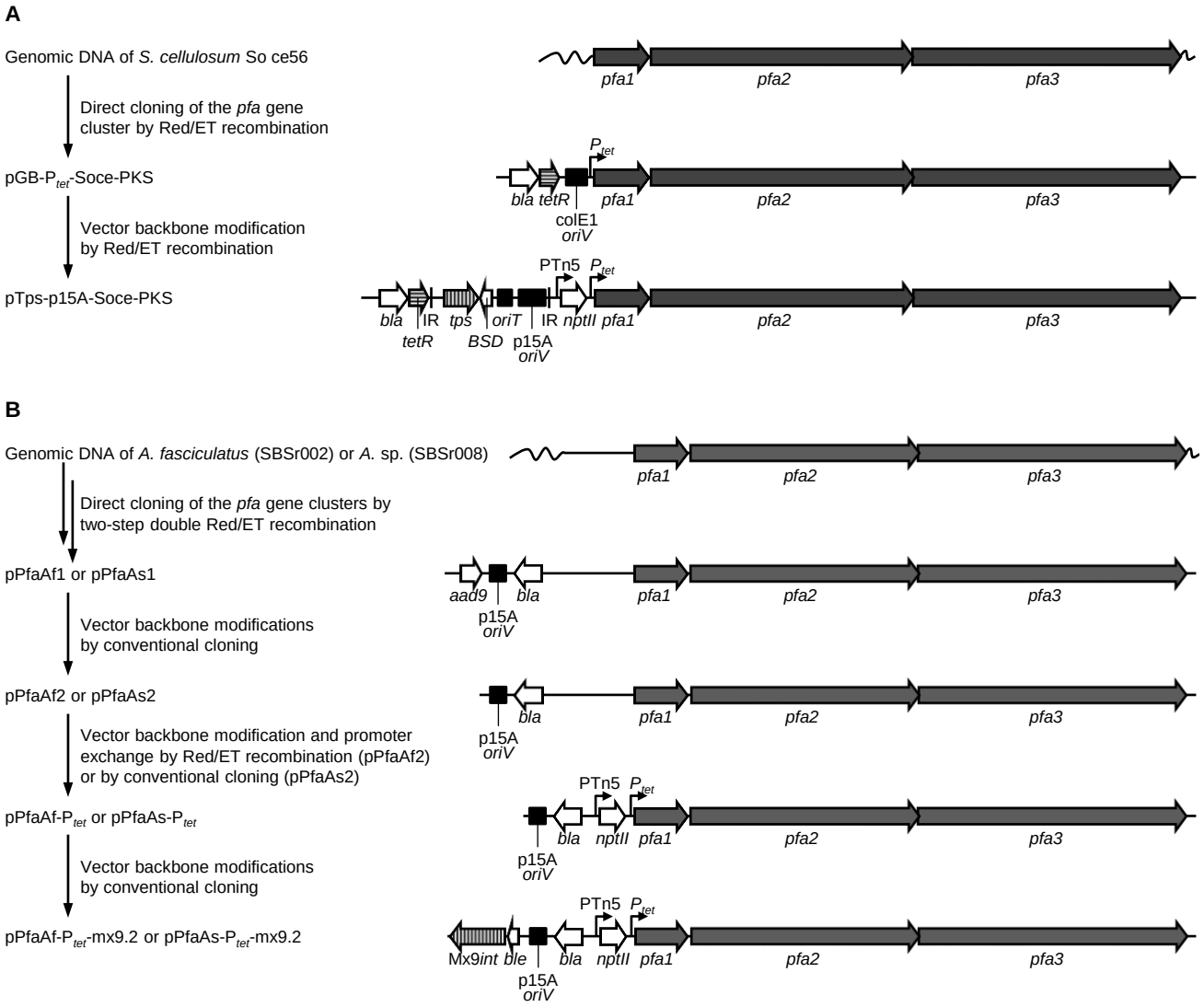


Figure 5. Cloning strategy for the construction of expression plasmids for recombinant PUFA production in *Myxococcus xanthus* DK1622.
(A) Cloning of the *pfa* gene cluster from *Sorangium cellulosum* So ce56, (B) Cloning of the *pfa* gene clusters from *Aetherobacter fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008). *oriV* = origin of replication, *oriT* = origin of transfer, *bla* = ampicillin resistance gene, *BSD* = blasticidin resistance gene, *nptII* = neomycin resistance gene, *aad9* = spectinomycin resistance gene, *ble* = zeocin resistance gene, P_{tet} = tetracycline promoter, *tetR* = gene encoding the tetracycline transcriptional regulator, PTn5 = Tn5 promoter, *tps* = mariner transposase gene, IRs = inverted repeats, Mx9*int* = Mx9 integrase gene with phage attachment site *attP*.

The generated expression constructs pTps-p15A-Soce-PKS, pPfaAf-P_{tet}-mx9.2, and pPfaAs-P_{tet}-mx9.2 were transferred into *M. xanthus* DK1622 by electroporation. In case of pPfaAf-P_{tet}-mx9.2 or pPfaAs-P_{tet}-mx9.2, six resistant transformants were randomly selected to verify the correct integration via Mx9 *attB* sites by PCR analysis.

For every transformant, integration of the expression construct at the *attB2* site [51] could be confirmed. Six transgenic clones of *M. xanthus* DK1622::pTps-p15A-Soce-PKS, *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2, and *M. xanthus* DK1622::pPfaAs-P_{tet}-mx9.2, respectively, were cultivated in 50 ml scale at 30 °C. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [26] and analysed by GC-MS (Figure 6 and Table 4). All transgenic clones of *M. xanthus* DK1622::pTps-p15A-Soce-PKS produced LA but also GLA and oleic acid (OA, 18:1, *n*-9), a monounsaturated fatty acid (MUFA), as well as EDA and dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6) in trace amounts. These fatty acids are not present in *M. xanthus* DK1622 wild type. Production maximum was reached after two days. In contrast to this, *S. cellulosum* So ce56 produced significantly lower amounts of OA and LA after seven days of cultivation. All the six transgenic clones of *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2 and *M. xanthus* DK1622::pPfaAs-P_{tet}-mx9.2, respectively, produced EPA and DHA, as well as AA, *n*-6 DPA, and *n*-3 DPA in trace amounts after two days of cultivation. In this case, production yields of AA, EPA, and DHA, achieved with the native producer strains *A. fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008) after seven days of cultivation surpass the PUFA production rate of the corresponding heterologous host strains. Nonetheless, *M. xanthus* DK1622 turned out to be a reliable expression host for PUFA synthases in our study. Contrary to other host strains, heterologous expression of an exogenous PPTase was not necessary since one of the two Sfp-type PPTase from *M. xanthus* DK1622 with broad substrate specificity, MxPpt1 or MxPpt2 [52], was obviously able to activate the tandem ACP domains of the PUFA synthases by posttranslational modification.

Interestingly, when comparing the PUFA production profiles of the native producer *S. cellulosum* So ce56 and the corresponding transgenic host *M. xanthus* DK1622::pTps-p15A-Soce-PKS remarkable differences can be observed. The host strain produces not only OA and LA but also GLA. Feeding of 5 mg LA to a 50 ml culture of *M. xanthus* DK1622 and cultivation for two days showed that 22% of LA was converted into GLA probably catalysed by a linoleoyl-CoA desaturase (EC 1.14.19.3). On top of that, 65% of 5 mg OA fed to a 50 ml culture of *M. xanthus* DK1622 were converted after two days of cultivation into LA (44% of original OA) most likely by the action of an oleoyl-CoA-Δ^{12} desaturase (EC 1.14.19.6) and again subsequently converted into GLA (21% of original OA). Candidates for these desaturases could be identified by *in silico* analysis of the genome of *M. xanthus*

DK1622 [53], which comprises 16 genes encoding mostly uncharacterized fatty acid desaturases. The most obvious interpretation of these data is the assumption that the PUFA synthase of the analysed *S. cellulosum* strains seems to be responsible for the production of both OA plus LA or of simply OA which could be converted into LA by an oleoyl-CoA-Δ^{12} desaturase. The genome sequence of *S. cellulosum* So ce56 encodes at least five uncharacterized fatty acid desaturases [35]. Apparently, *S. cellulosum* does not express any active linoleoyl-CoA desaturase under the tested culture conditions.

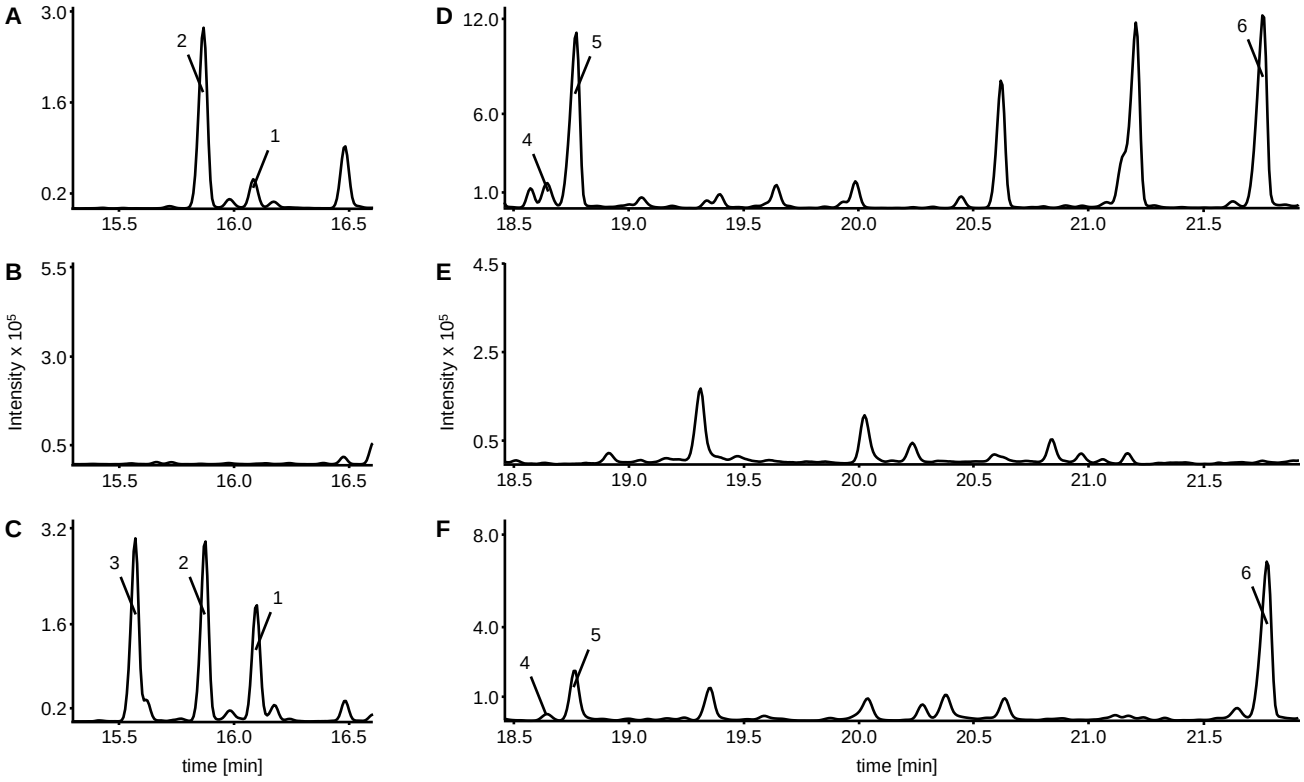


Figure 6. GC-MS chromatograms of fatty acid methyl esters from myxobacterial PUFA producers and heterologous hosts.
(A) *Sorangium cellulosum* So ce56, (B + E) *Myxococcus xanthus* DK1622, (C) *M. xanthus* DK1622::pTps-p15A-Soce-PKS, (D) *Aetherobacter fasciculatus* (SBSr002), (F) *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2. Most prominent MUFA and PUFA methyl esters: (1) oleic acid methyl ester (18:1, *n*-9), (2) linoleic acid methyl ester (18:2, *n*-6), (3) γ-linolenic acid methyl ester (18:3, *n*-6), (4) arachidonic acid methyl ester (20:4, *n*-6), (5) eicosapentaenoic acid methyl ester (20:5, *n*-3), (6) docosahexaenoic acid methyl ester (22:6, *n*-3).

Table 4. Production of MUFAs and PUFAs with a carbon chain length of ≥ C18 by *Sorangium cellulosum* So ce56, *Myxococcus xanthus* DK1622::pTps-p15A-Soce-PKS, *Aetherobacter fasciculatus* (SBSr002), *Aetherobacter* sp. (SBSr008), *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2, and *M. xanthus* DK1622::pPfaAs-P_{tet}-mx9.2.
Cultivation was carried out for 7 days (for *S. cellulosum* and for *Aetherobacter* spp.) or for 2 days (for *M. xanthus*). The indicated values for *M. xanthus* are means of three samples of the best producing clones.

	Absolute amount [mg/g cell dry weight]			
MUFA or PUFA	***S. cellulosum* So ce56**	***M. xanthus* DK1622::pTps-p15A-Soce-PKS**	***A. fasciculatus* (SBSr002)/ *Aetherobacter* sp. (SBSr008)**	***M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2/ *M. xanthus* DK1622::pPfaAs-P_{tet}-mx9.2**
oleic acid (OA, 18:1, *n*-9)	0.2	0.5	0.2	–
linoleic acid (LA, 18:2, *n*-6)	3.7	5.5	0.4	–
γ-linolenic acid (GLA, 18:3, *n*-6)	–	5.8	< 0.1	–
stearidonic acid (SDA, 18:4, *n*-3)	–	–	< 0.1	–
eicosadienoic acid (EDA, 20:2, *n*-6)	< 0.1	< 0.1	–	–
dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6)	–	< 0.1	–	–
arachidonic acid (AA, 20:4, *n*-6)	–	–	0.5-1.0	< 0.1
eicosapentaenoic acid (EPA, 20:5, *n*-3)	–	–	3.8	0.2
docosapentaenoic acid (DPA, 22:5, *n*-6)	–	–	0.1	< 0.1
docosapentaenoic acid (DPA, 22:5, *n*-3)	–	–	< 0.1	< 0.1
docosahexaenoic acid (DHA, 22:6, *n*-3)	–	–	4.7-6.2	0.8

In the case of the *pfa* pathways from *Aetherobacter* spp., it seems at first glance that the encoded PUFA synthase is able to synthesize a broad range of PUFAs since both the native producers *A. fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008) and the transgenic hosts *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2 and *M. xanthus* DK1622::pPfaAs-P_{tet}-mx9.2 produce at least AA, EPA, and DHA. However, it has been shown that the EPA/DHA ratio rises in *Aetherobacter* spp. with increasing cultivation time [41]. In addition to this, feeding of 5 mg DHA to a 50 ml culture of *M. xanthus* DK1622 and cultivation for two days showed that > 95% of DHA was

converted mostly into EPA. From this data, it can be hypothesized that the PUFA synthase from *Aetherobacter* spp. may only produce DHA and some minor amounts of *n*-6 DPA and *n*-3 DPA as by-products. These C22 PUFAs then could be converted into the corresponding C20 PUFAs EPA or AA via β-oxidation plus 2,4-dienoyl-CoA reductase (EC 1.3.1.34) and Δ^3-Δ^2-enoyl-CoA isomerase (EC 5.3.3.8), needed for the oxidation of unsaturated fatty acids. However, further studies are required to clarify this issue.

CONCLUSIONS

In conclusion, for the first time PUFA synthases from terrestrial origin catalysing the biosynthesis of LC-PUFAs could be identified and characterized. They are encoded by two types of *pfa* gene clusters present in the genome of the closely related myxobacterial genera *Sorangium* and *Aetherobacter*. Although the organization of these gene clusters is very similar, the profiles of the produced fatty acids vary significantly. The presence of different types of PUFA biosynthetic pathways with such a product diversity within the same bacterial family is an outstanding aspect in the field of research on PUFA biosynthesis. By construction of functional hybrid PUFA pathways their exact biosynthetic mechanism and factors determining product specificity can now be studied. Consequently, the insights expected from such work would allow for the design of an enzymatic system for the targeted production of defined PUFAs. Another important aspect of the present work is the successful establishment of an expression platform for the heterologous PUFA production using the myxobacterial *pfa* gene clusters in the promising host *Myxococcus xanthus* via synthetic biotechnology approaches. These achievements pave the way for the heterologous PUFA production in host strains with fast growth characteristics, amenability to genetic modifications, and vast potential for industrial applications, like the GRAS certified microbe *Pseudomonas putida* KT2440 or oleaginous yeasts.

ACKNOWLEDGEMENTS

The authors would like to thank Marc Stadler (Helmholtz Centre for Infection Research) for fruitful discussions, Andrea Rademacher (previously InterMed Discovery GmbH) for strain cultivation, Susanne Schneiker-Bekel (CeBiTec) for genome sequencing, and Nestor Zaburannyi (Helmholtz Institute for Pharmaceutical Research Saarland) for assembly of the sequencing data. This work was generously supported by a grant from the Bundesministerium für Bildung und Forschung (FKZ 0315798).

NOTES

The GenBank accession numbers for the *pfa* biosynthetic gene cluster sequences from *Aetherobacter fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008) reported in this paper are KF977699 and KF977700.

K. Gemperlein wrote the manuscript in collaboration with S. C. Wenzel and R. Müller and performed all experiments except for cloning of the plasmids pGB-P_{tet}-Soce-PKS and pTps-p15A-Soce-PKS (performed by S. Rachid) and construction of the phylogenetic tree (performed by R. O. Garcia).

REFERENCES

[1] Lorente-Cebrian S, Costa AGV, Navas-Carretero S, Zabala M, Martinez JA, & Moreno-Aliaga MJ (2013) Role of omega-3 fatty acids in obesity, metabolic syndrome, and cardiovascular diseases: a review of the evidence. *J. Physiol. Biochem.*, **69,** 633-651.

[2] Serhan CN, Yacoubian S, & Yang R (2008) Anti-inflammatory and proresolving lipid mediators. *Annu. Rev. Pathol.*, **3,** 279-312.

[3] Simopoulos AP (2008) The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp. Biol. Med.*, **233,** 674-688.

[4] Lenihan-Geels G, Bishop KS, & Ferguson LR (2013) Alternative sources of omega-3 fats: can we find a sustainable substitute for fish? *Nutrients*, **5,** 1301-1315.

[5] Ratledge C (2004) Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie*, **86,** 807-815.

[6] Warude D, Joshi K, & Harsulkar A (2006) Polyunsaturated fatty acids: Biotechnology. *Crit. Rev. Biotechnol.*, **26,** 83-93.

[7] Kaulmann U & Hertweck C (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew. Chem. Int. Ed.*, **41,** 1866-1869.

[8] Napier JA (2002) Plumbing the depths of PUFA biosynthesis: a novel polyketide synthase-like pathway from marine organisms. *Trends Plant Sci.*, **7,** 51-54.

[9] Wallis JG, Watts JL, & Browse J (2002) Polyunsaturated fatty acid synthesis: what will they think of next? *Trends Biochem. Sci.*, **27,** 467-473.

[10] 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.

[11] 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.

[12] Morita N, Tanaka M, & Okuyama H (2000) Biosynthesis of fatty acids in the docosahexaenoic acid-producing bacterium *Moritella marina* strain MP-1. *Biochem. Soc. Trans.*, **28,** 943-945.

[13] Hauvermale A, Kuner J, Rosenzweig B, Guerra D, Diltz S, & Metz JG (2006) Fatty acid production in *Schizochytrium* sp.: Involvement of a polyunsaturated fatty acid synthase and a type I fatty acid synthase. *Lipids*, **41,** 739-747.

[14] Okuyama H, Orikasa Y, Nishida T, Watanabe K, & Morita N (2007) Bacterial genes responsible for the biosynthesis of eicosapentaenoic and docosahexaenoic acids and their heterologous expression. *Appl. Environ. Microbiol.*, **73,** 665-670.

[15] Shulse CN & Allen EE (2011) Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. *PLOS One*, **6,** e20146.

[16] Yazawa K (1996) Production of eicosapentaenoic acid from marine bacteria. *Lipids*, **31 Suppl,** S297-S300.

[17] 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 (Pt 8),** 2725-2731.

[18] Orikasa Y, Nishida T, Yamada A, Yu R, Watanabe K, Hase A, Morita N, & Okuyama H (2006) Recombinant production of docosahexaenoic acid in a polyketide biosynthesis mode in *Escherichia coli*. *Biotechnol. Lett.*, **28,** 1841-1847.

[19] Chen H, Keseler IM, & Shimkets LJ (1990) Genome size of *Myxococcus xanthus* determined by pulsed-field gel electrophoresis. *J. Bacteriol.*, **172,** 4206-4213.

[20] Grant SG, Jessee J, Bloom FR, & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U S A*, **87,** 4645-4649.

[21] Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Müller R, Stewart AF, & Zhang Y (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.*, **30,** 440-446.

[22] Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF, & Zhang Y (2006) An improved recombineering approach by adding RecA to λ Red recombination. *Mol. Biotechnol.*, **32,** 43-53.

[23] Sambrook J & Russell DW (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[24] Fu J, Wenzel SC, Perlova O, Wang J, Gross F, Tang Z, Yin Y, Stewart AF, Müller R, & Zhang Y (2008) Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res.*, **36,** e113.

[25] Oßwald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, & Wenzel SC (2012) Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.*, Epub ahead of print.

[26] Bode HB, Ring MW, Schwär G, Kroppenstedt RM, Kaiser D, & Müller R (2006) 3-Hydroxy-3-methylglutaryl-coenzyme A (CoA) synthase is involved in the biosynthesis of isovaleryl-CoA in the myxobacterium *Myxococcus xanthus* during fruiting body formation. *J. Bacteriol.*, **188,** 6524-6528.

[27] Bligh EG & Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can. J. Biochem. and Phys.*, **37,** 911-917.

[28] Lewis T, Nichols PD, & McMeekin TA (2000) Evaluation of extraction methods for recovery of fatty acids from lipid-producing microheterotrophs. *J. Microbiol. Methods*, **43,** 107-116.

[29] Ishikawa J & Hotta K (1999) FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol. Lett.*, **174,** 251-253.

[30] Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215,** 403-410.

[31] Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, & Finn RD (2012) The Pfam protein families database. *Nucleic Acids Res.*, **40,** D290-D301.

[32] Drummond AJ, Ashton B, Buxton S, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, & Wilson A (2011) *Geneious v5.6*.

[33] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, & Higgins DG (2007) Clustal W and clustal X version 2.0. *Bioinformatics*, **23,** 2947-2948.

[34] Garcia R, Pistorius D, Stadler M, & Müller R (2011) Fatty acid-related phylogeny of myxobacteria as an approach to discover polyunsaturated omega-3/6 fatty acids. *J. Bacteriol.*, **193,** 1930-1942.

[35] Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez-Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Munoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Rückert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorhölter FJ, Weissman KJ, Welch RD, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Blöcker H, Pühler A, & Müller R (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat. Biotechnol.*, **25,** 1281-1289.

[36] Gerth K, Washausen P, Höfle G, Irschik H, & Reichenbach H (1996) The jerangolids: A family of new antifungal compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties of jerangolid A. *J. Antibiot.*, **49,** 71-75.

[37] Gerth K, Steinmetz H, Höfle G, & Jansen R (2008) Chlorotonil A, a macrolide with a unique gem-dichloro-1,3-dione functionality from *Sorangium cellulosum*, So ce1525. *Angew. Chem. Int. Ed.*, **47,** 600-602.

[38] Irschik H, Augustiniak H, Gerth K, Höfle G, & Reichenbach H (1995) The ripostatins, novel inhibitors of eubacterial RNA polymerase isolated from myxobacteria. *J. Antibiot.*, **48,** 787-792.

[39] Jahns C, Hoffmann T, Müller S, Gerth K, Washausen P, Höfle G, Reichenbach H, Kalesse M, & Müller R (2012) Pellasoren: structure elucidation, biosynthesis, and total synthesis of a cytotoxic secondary metabolite from *Sorangium cellulosum*. *Angew. Chem. Int. Ed.*, **51,** 5239-5243.

[40] Han K, Li ZF, Peng R, Zhu LP, Zhou T, Wang LG, Li SG, Zhang XB, Hu W, Wu ZH, Qin N, & Li YZ (2013) Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci. Rep.*, **3,** 2101.

[41] Stadler M, Roemer E, Müller R, Garcia RO, Pistorius D, & Brachmann A (2010) Production of omega-3 fatty acids by myxobacteria. *Int. Pat.*, WO 2010/063451.

[42] Plaza A, Garcia R, Bifulco G, Martinez JP, Huttel S, Sasse F, Meyerhans A, Stadler M, & Müller R (2012) Aetheramides A and B, potent HIV-inhibitory depsipeptides from a myxobacterium of the new genus *"Aetherobacter"*. *Org. Lett.*, **14,** 2854-2857.

[43] Zhang YM & Rock CO (2008) Thematic review series: Glycerolipids - Acyltransferases in bacterial glycerophospholipid synthesis. *J. Lipid Res.*, **49,** 1867-1874.

[44] 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.*, **580,** 4423-4429.

[45] Gemperlein K, Wenzel SC, Rachid S, & Müller R (2011) Production of fatty acids by heterologous expression of gene clusters from myxobacteria. *Int. Pat.*, WO 2011/151298.

[46] Orikasa Y, Tanaka M, Sugihara S, Hori R, Nishida T, Ueno A, Morita N, Yano Y, Yamamoto K, Shibahara A, Hayashi H, Yamada Y, Yamada A, Yu RK, Watanabe K, & Okuyama H (2009) *pfaB* products determine the molecular species produced in bacterial polyunsaturated fatty acid biosynthesis. *FEMS Microbiol. Lett.*, **295,** 170-176.

[47] Metz JG, Kuner J, Rosenzweig B, Lippmeier JC, Roessler P, & Zirkle R (2009) Biochemical characterization of polyunsaturated fatty acid synthesis in *Schizochytrium*: release of the products as free fatty acids. *Plant Physiol Biochem.*, **47,** 472-478.

[48] Rodriguez-Guilbe M, Oyola-Robles D, Schreiter ER, & Baerga-Ortiz A (2013) Structure, Activity, and Substrate Selectivity of the Orf6 Thioesterase from *Photobacterium profundum*. *J. Biol. Chem.*, **288,** 10841-10848.

[49] Oyola-Robles D, Gay DC, Trujillo U, Sanchez-Pares JM, Bermudez ML, Rivera-Diaz M, Carballeira NM, & Baerga-Ortiz A (2013) Identification of novel protein domains required for the expression of an active dehydratase fragment from a polyunsaturated fatty acid synthase. *Protein Sci.*, **22,** 954-963.

[50] Zhang Y, Muyrers JPP, Testa G, & Stewart AF (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18,** 1314-1317.

[51] Julien B (2003) Characterization of the integrase gene and attachment site for the *Myxococcus xanthus* bacteriophage Mx9. *J. Bacteriol.*, **185,** 6325-6330.

[52] Meiser P & Müller R (2008) Two functionally redundant Sfp-type 4'-phosphopantetheinyl transferases differentially activate biosynthetic pathways in *Myxococcus xanthus*. *ChemBioChem*, **9,** 1549-1553.

[53] Goldman BS, Nierman WC, Kaiser D, Slater SC, Durkin AS, Eisen J, Ronning CM, Barbazuk WB, Blanchard M, Field C, Halling C, Hinkle G, Iartchuk O, Kim HS, Mackenzie C, Madupu R, Miller N, Shvartsbeyn A, Sullivan SA, Vaudin M, Wiegand R, & Kaplan HB (2006) Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Natl. Acad. Sci. U S A*, **103,** 15200-15205.

APPENDIX

DNA sequences of the 16S ribosomal RNA from myxobacterial strains which are not deposited in any database and which were used for the construction of the phylogenetic tree by the neighbour-joining method

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ce10

AGAGTTTGATCCTGGCTCAGAACGAACGTTAGCGGCGCGCTTAACACATGCAAGTCGAGCGAGAAAGGGCTTCGG
CCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGGA
GCGCTAATACAGCATGAGACCACGTCCTCGAAAGGGGATGAGGTCAAAGCCGGCCTCTTCATGAAAGCTGGCGCC
AGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCTG
AGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGC
AATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGGG
ACGAATAAGGGTTGGCTAACATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGCA
GCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCATGTAGGCGGTTCGTAAAGT
CAGATGTGAAAGCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTCACGAACTTGAGTCCCGGAGAGGAAGGCG
GAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCTTCTGGACGGTG
ACTGACGCTGAGATGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGG
TGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGGCC
GCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACGC
GCAGAACCTTACCTGGGCTAGAAAATGCAGGGACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAGTT
AGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCG
TTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTTAAATCGGAGGAAGGTGGGGATGACGT
CAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAAACGGTTGCGAACTCGCGA
GGGGAAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATCGC
TAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAG
TCGATTGCTCCAGAAGTGGCTACGCCAACCCGCAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGTGA
AGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ce1525

AGAGTTTGATCCTGGCTCAGAACGAACGTTAGCGGCGCGCTTAACACATGCAAGTCGAGCGAGAAAGGGCTTCGG
CCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGGA
GCGCTAATACAGCATGAGACCACGTCCTCGAGAGGGGATGAGGTCAAAGCCGGCCTCTTCACGAAAGCTGGCGCC
AGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCTG
AGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGC
AATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGGG
ACGAATAAGGGTTGGCTAACATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGCA
GCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCATGTAGGCGGTTCGTAAAGT
CAGATGTGAAAGCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTCACGAACTTGAGTCCCGGAGAGGAAGGCG
GAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCTTCTGGACGGTG
ACTGACGCTGAGATGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGG
TGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGGCC
GCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACGC
GCAGAACCTTACCTGGGCTAGAAAATGCAGGGACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAGTT
AGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCG
TTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTTAAATCGGAGGAAGGTGGGGATGACGT
CAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAAACGGTCGCGAACCCGCGA
GGGGAAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATCGC
TAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAG
TCGATTGCTCCAGAAGTGGCTGCGCCAACCCGCAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGTGA
AGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ce377

AGAGTTTGATCCTGGCTCAGAACGAACGTTAGCGGCGCGCTTAACACATGCAAGTCGAGCGAGAAAGGGCTTCGG
CCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGGA
GCGCTAATACAGCATGAGACCACGCCTTCGAAAGAGGGTGAGGTCAAAGCCGGCCTCTTCACGAAAGCTGGCGCC
AGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCTG
AGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGC
AATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGGG
ACGAATAAGGGTTGGCTAACATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGCA
GCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCATGTAGGCGGTTCGTAAAGT
CAGATGTGAAAGCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTCACGAACTTGAGTCCCGGAGAGGAAGGCG
GAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCTTCTGGACGGTG
ACTGACGCTGAGATGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGG
TGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGGCC
GCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACGC
GCAGAACCTTACCTGGGCTAGAAAATGCAGGAACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAGTT
AGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCG
TTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTTAAATCGGAGGAAGGTGGGGATGACGT
CAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAGACGGTCGCGAACCCGCGA
GGGGGAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATCGC
TAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAG
TCGATTGCTCCAGAAGTGGCTGCGCCAACCCGCAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGTGA
AGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ce38

AGAGTTTGATCCTGGCTCAGAACGAACGTTAGCGGCGCGCTTAACACATGCAAGTCGAGCGAGAAAGGGCTTCGG
CCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGGA
GCGCTAATACAGCATGAGACCACGTCCTCGAAAGGGGATGAGGTCAAAGCCGGCCTCTTCATGAAAGCTGGCGCC
AGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCTG
AGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGC
AATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGGG
ACGAATAAGGGTTGGCTAACATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGCA
GCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCATGTAGGCGGTTCGTAAAGT
CAGATGTGAAAGCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTCACGAACTTGAGTCCCGGAGAGGAAGGCG
GAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCTTCTGGACGGTG
ACTGACGCTGAGATGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGG
TGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGGCC
GCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACGC
GCAGAACCTTACCTGGGCTAGAAAATGCAGGAACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAGTT
AGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCG
TTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTTAAATCGGAGGAAGGTGGGGATGACGT
CAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAAACGGTCGCAAACTCGCGA
GAGCAAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATCGC
TAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAG
TCGATTGCTCCAGAAGTGGCTACGCCAACCCGCAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGTGA
AGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ce836

TGAGTTTGATTCCTGGCTCAGGACGAACGTTAGCGGCGCGCTTAACACATGCAAGTCGAGCGAGAAAGGGCTTCG
GCCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGG
AGCGCTAATACAGCATGAGACCACGTCTTCGAAAGAGGATGAGGTCAAAGCCGGCCTCTTCACGAAAGCTGGCGC
CAGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCT
GAGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCG
CAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGG
GACGAATAAGGGTTGGCTAACATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGC
AGCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCATGTAGGCGGTTCGTAAAG

TCAGATGTGAAAGCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTCACGAACTTGAGTCCCGGAGAGGAAGGC
GGAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCTTCTGGACGGT
GACTGACGCTGAGATGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGG
GTGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGGC
CGCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACG
CGCAGAACCTTACCTGGGCTAGAAAATGCAGGGACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAGT
TAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATC
GTTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTTAAATCGGAGGAAGGTGGGGATGACG
TCAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAGACGGTCGCGAACCCGCG
AGGGGAAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATCG
CTAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGA
GTCGATTGCTCCAGAAGTGGCTGCGCCAACCCGCAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGTG
AAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTA

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ceGT47

AGAGTTTGATCCTGGCTCAGAACGAACGTTAGCGGCGCGCCTAACACATGCAAGTCGAGCGAGAAAGGGCTTCGG
CCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGGA
GCGCTAATACAGCATGAGACCACGTCCTCGAAAGGGGATGAGGTCAAAGCCGGCCTCTTCATGAAAGCTGGCGCC
AGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCTG
AGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGC
AATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGGG
ACGAATAAGGGTTGGCTAATATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGCA
GCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCGTGTAGGCGGTCCGTAAAGT
CAGGTGTGAAAGCCCTGGGCTCAACCCAGGAAGTGCACTTGAAACTCACGGACTCGAGTCCCGGAGAGGAAGGCG
GAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCTTCTGGACGGTG
ACTGACGCTGAGACGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGG
TGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGGCC
GCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACGC
GCAGAACCTTACCTGGGCTAGAAAATGCAGGGACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAGTT
AGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCG
TTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTCAAATCGGAGGAAGGTGGGGATGACGT
CAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAAACGGTCGCGAACCCGCGA
GGGGGAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATCGC
TAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAG
TCGATTGCTCCAGAAGTGGCTGCGCCAACCCGAAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGTGA
AGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

>16S ribosomal RNA gene sequence from *Sorangium cellulosum* So ce487

TGAGTTTGATTCCTGGCTCAGGACGAACGTTAGCGGCGCGCTTAACACATGCAAGTCGAGCGAGAAAGGGCTTCG
GCCCCGGTAAAGCGGCGCACGGGTGAGTAACACGTAGGTAATCTGCCCCCAGGTGGTGGATAACGTTCCGAAAGG
AGCGCTAATACAGCATGAGACCACGCCTTCGAAAGAGGGTGAGGTCAAAGCCGGCCTCTTCACGAAAGCTGGCGC
CAGGGGATGAGCCTGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGAAGACGGGTAGCTGGTCT
GAGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCG
CAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCTTCGGGTTGTAAAGCTCTGTGGAGGGG
GACGAATAAGGGTTGGCTAACATCCAGCTCGATGACGGTACCCCTTTAGCAAGCACCGGCTAACTCTGTGCCAGC
AGCCGCGGTAAGACAGAGGGTGCAAACGTTGTTCGGAATTACTGGGCGTAAAGCGCATGTAGGCGGTTCGTAAAG
TCAGATGTGAAAGCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTCACGAACTTGAGTCCCGGAGAGGAAGGC
GGAATTCTCGGTGTAGAGGTGAAATTCGTAGATATCGAGAGGAACATCGGTGGCGAAGGCGGCCCTTCTGGACGG
TGACTGACGCTGAGATGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATG
GGTGCTAGGTGTCGCGGGCTTTGACTCCTGCGGTGCCGTAGCTAACGCATTAAGCACCCCGCCTGGGGAGTACGG
CCGCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAAC
GCGCAGAACCTTACCTGGGCTAGAAAATGCAGGAACCTGGTTGAAAGATCGGGGTGCTCTTCGGAGAACCTGTAG
TTAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTAT
CGTTAGTTGCCAGCGGTTCGGCCGGGCACTCTAGCGAGACTGCCGATATTTAAATCGGAGGAAGGTGGGGATGAC
GTCAAGTCCTCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGGTACAGACGGTCGCGAACCCGC
GAGGGGAAGCCAATCCGAAAAAACCGTCCTCAGTACGGATAAGAGTCTGCAACTCGACTCTTTGAAGTTGGAATC

GCTAGTAATCCCTGATCAGCAGGCAGGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGG
AGTCGATTGCTCCAGAAGTGGCTGCGCCAACCCGCAAGGGAGGCAGGCCCCCAAGGAGTGGTTGGTAACTGGGGT
GAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGCTCACCTCCTTAA

CHAPTER 2

Discovery and characterization of a second type of PUFA synthase from *Sorangium cellulosum* with a new product spectrum

INTRODUCTION

Various biosynthetic mechanisms are known for the production of long-chain polyunsaturated fatty acids (LC-PUFAs). The predominant aerobic pathways combine several oxygen-dependent desaturases and elongases for the conversion of saturated fatty acids into unsaturated fatty acids [1]. However, LC-PUFAs can also be synthesized in an anaerobic manner, employing iterative type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like enzymes (PUFA synthases) for *de novo* biosynthesis from acyl-CoA precursors [2-4]. These multienzyme complexes are encoded by PUFA (*pfa*) biosynthetic gene clusters, initially identified and characterized from marine γ-Proteobacteria and microalgae: *Shewanella pneumatophori* SCRC-2738 and *Photobacterium profundum* SS9 both produce eicosapentaenoic acid (EPA, 20:5, *n*-3) [5;6], *Moritella marina* MP-1 is known as docosahexaenoic acid (DHA, 22:6, *n*-3) producer [7], whereas *Schizochytrium* sp. ATCC 20888 produces both DHA and *n*-6 docosapentaenoic acid (DPA, 22:5) [5;8]. The increasing availability of genome sequence data enabled the identification of at least 20 distinct types of *pfa*-like gene clusters, which are widely distributed especially within the marine habitat [9]. However, most of the encoded PUFA synthases and their biosynthetic products remain uncharacterized [10].

Interestingly, the first characterized PUFA synthases from terrestrial origin catalysing the biosynthesis of LC-PUFAs could be discovered in myxobacteria [11]. They are encoded by two types of *pfa* gene clusters found in the genome of the novel isolates *Aetherobacter* spp., identified as prolific producers of EPA and DHA [12;13], and of the linoleic acid (LA, 18:2, *n*-6) producing *Sorangium cellulosum* strains So ce56 (genes sce0818, sce0819, and sce0820) [14], So ce10 [15], So ce1525 [16], So ce307 [17], So ce377 [18], So ce38 [19], So ce487, So ce836, So ce969, and So ceGT47. Although the profiles of the produced fatty acids vary significantly among *Aetherobacter* and *Sorangium*, the structural organization of these gene clusters is very similar. In addition to the *S. cellulosum* strains mentioned above, So ce strains not only producing LA but also PUFAs with a longer chain, mainly EPA, were identified during screening of an in-house myxobacterial strain collection. Remarkably, these strains harbour an additional, novel PUFA biosynthetic pathway for the production of EPA and other LC-PUFAs, which was deciphered and analysed *in silico* in comparison to the formerly known myxobacterial PUFA pathways. Evaluation of LC-PUFA production profiles, including application of the established heterologous expression system using *Myxococcus xanthus* as host strain for selected pathways and hybrids thereof, provided indications how

myxobacteria employ three distinct types of PUFA synthases to generate different fatty acid profiles.

EXPERIMENTAL PROCEDURES

Culture conditions

Sorangium cellulosum SBSo026 was isolated by the Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Germany. It was grown in liquid HS-medium containing 0.15% casitone, 0.2% HEPES, 0.1% KNO_3, and 0.1% $MgSO_4$ x 7 H_2O. The medium was adjusted to pH 7.2 with NaOH and supplemented with 0.00625% K_2HPO_4, 0.4% glucose, 0.0075% $CaCl_2$ x 2 H_2O, and 8 mg/l NaFe-EDTA after autoclaving. Alternatively, the strain could also be cultivated in liquid M-medium containing 1% phytone peptone, 1% maltose, 0.1% $CaCl_2$ x 2 H_2O, 0.1% $MgSO_4$ x 7 H_2O, and 50 mM HEPES. The medium was adjusted to pH 7.2 with NaOH and supplemented with 8 mg/l NaFe-EDTA after autoclaving. *S. cellulosum* So ce1128, So ce340 and So ce427 were isolated by the Department of Microbial Drugs, Helmholtz Centre for Infection Research, Braunschweig, Germany, and were grown in liquid M-medium. The cultures were incubated on a rotary shaker for 5-10 days at 30 °C and 200 rpm.

Escherichia coli DH10B [20] was used for cloning experiments. *E. coli* HS996/pSC101-BAD-gbaA (tet^R) [21] and GB05-red [22] were used for modification of a plasmid using Red/ET recombination. The cells were grown in LB-medium / on LB-agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) or in 2xYT-medium / on 2xYT-agar (1.6% tryptone, 1% yeast extract, 0.5% NaCl, (1.5% agar)) at 30-37 °C (and 200 rpm) overnight. Antibiotics were used at the following concentrations: 100 µg/ml ampicillin, 50 µg/ml kanamycin, 20 µg/ml zeocin, 12.5 µg/ml chloramphenicol, and 6-12.5 µg/ml tetracycline.

Myxococcus xanthus DK1622 [23] was grown in CTT-medium or on CTT-agar containing 1% casitone, 10 mM Tris, 1 mM KPO_4 buffer, 8 mM $MgSO_4$ (and 1.5% agar). The medium was adjusted to pH 7.6. The cultures were incubated for 2-3 days at 30°C (and 200 rpm).

General molecular biology techniques

Routine handling of nucleic acids, such as isolation of genomic and plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, were performed according to standard protocols [24]. All the enzymes were purchased from Thermo Scientific or New England Biolabs.

Isolation of genomic DNA from myxobacteria

100 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. After discarding the supernatant completely, the cells were washed once with 15 ml SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM Tris-Cl pH 7.5), pelleted at 8,000 rpm for 10 min at room temperature, and resuspended in 5 ml SET buffer afterwards. 50 µl RNase A stock solution (10 mg/ml in sterile ddH_2O) and 300 µl Proteinase K solution (10 mg/ml in 50 mM Tris-Cl pH 8.0, 1 mM $CaCl_2$) were added, and the tube was inverted several times. Subsequently, 600 µl 10% SDS were added, and the tube was incubated first at 41 °C for 1.5 h and afterwards at 55 °C for 1 h under rotation in a hybridization oven. Thereafter, one volume phenol:chloroform:isoamyl alcohol (25:24:1) was added, and the mixture was incubated at 5 rpm for 1 h on a tube rotator. The tube was centrifuged at 8,000 rpm for 5 min at room temperature, and the upper phase was transferred into a new tube by using an end-cut 1 ml-tip. The extraction step with phenol:chloroform:isoamyl alcohol (25:24:1) was repeated once. Then, one volume chloroform:isoamyl alcohol (24:1) was added to the supernatant, and the mixture was incubated at 5 rpm for 1 h on a tube rotator. The tube was centrifuged at 14,000 rpm for 10 min at room temperature, and 4 ml of the upper phase were transferred into a new tube by using an end-cut 1 ml-tip. 440 µl 3 M sodium acetate pH 7.5 were added, and the tube was inverted several times. After adding 11 ml ice-cold ethanol, the tube was inverted until the appearance of cotton-like DNA. The DNA was collected by wrapping around the tip of a Pasteur pipette. In order to wash the DNA, the tip of the pipette with DNA was immersed into a 2 ml tube containing 70% ethanol. The DNA pellet on the pipette was dried by carefully attaching the inner site of a fresh 2 ml tube to remove the ethanol drops and suspended in 0.5-1 ml of 10 mM Tris-Cl pH 8.0.

Replicating of the fosmid library of *S. cellulosum* SBSo026 and spotting onto nylon membrane filters

A fosmid library of *S. cellulosum* SBSo026 was obtained from CeBiTec, Bielefeld, and all the *E. coli* clones were transferred from 36 96-well microtitre plates into nine 384-well microtitre plates containing 2xYT-medium plus 10% 10x Hogness Modified Freezing Medium (4 mM $MgSO_4$, 16 mM trisodium citrate, 68 mM $(NH_4)_2SO_4$, 35% (w/v) glycerol, 260 mM K_2HPO_4, and 132 mM KH_2PO_4) using a QPix2 system (Molecular Devices). After growth at 37 °C for 2 days, the clones were spotted onto two 222 x 222 mm positively charged nylon membrane filters (Performa II, Molecular Devices) with the QPix2 system. In order to enable clear-cut assignment of the transferred colonies to the corresponding clone position in the nine 384-well microtitre plates, each clone from the microtitre plates was spotted twice in a defined 4x4 array. The filters were placed onto 2xYT-agar containing 12.5 µg/ml chloramphenicol plus 0.1% L-arabinose (to increase the copy number of fosmids per cell) and incubated at 30 °C overnight. For subsequent lysis of the cells, the filters were processed as follows: Whatman 3MM Chr paper was soaked with denaturation solution (0.5 M NaOH, 1.5 M NaCl), neutralization solution (1.5 M NaCl, 1 M Tris-Cl pH 7.5) and 2x SSC solution (0.3 M NaCl, 30 mM trisodium citrate; pH 7.0). The filters (bottom side) were transferred onto the Whatman 3MM Chr paper, initially soaked with denaturation solution and incubated for 15 min. Subsequent incubation on neutralization solution (15 min) and 2x SSC (5 min) was carried out. The colonies were fixed on the filters through incubation at 80 °C for 1 h. The filters were equilibrated for 5 min in 2x SSC and treated with 25 ml Proteinase K solution (2 mg/ml in 2x SSC) for 2 h at 37 °C. In order to remove cell debris, the filters were washed in ddH_2O, treated with ddH_2O soaked tissues, then dried and stored at room temperature.

Screening of the fosmid library of *S. cellulosum* SBSo026 for clones containing the *pfa* gene clusters by non-radioactive hybridization

The 0.5 kb DNA probes 1, 2, and 3 were amplified from genomic DNA of *S. cellulosum* SBSo026 using primers So026_probe 1_fwd / So026_probe 1_rev, So026_probe 2_fwd / So026_probe 2_rev, and So026_probe 3_fwd / So026_probe 3_rev (Table 1). These primers were designed on the basis of sequences of homologous *pfa* gene clusters from the *S. cellulosum* strains So ce56 [14], So ce1525 [16], So ce377 [18], So ce38 [19], So ceGT47, and So ce427. PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5%

DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 ºC, annealing for 20 s at 72ºC, and extension for 10 s/kb at 72ºC; and a final extension for 10 min at 72ºC. The amplified probes were cleaned up, denatured at 95 °C for 10 min and quickly chilled on ice. DIG-High Prime (Roche Applied Science) was mixed thoroughly, and 4 µl were added to 16 µl denatured DNA. The samples were incubated at 37 °C overnight. The reactions were stopped by adding 0.7 µl 0.5 M EDTA pH 8.0. The DNA was precipitated after addition of 80 µl ddH_2O, 11 µl 3 M sodium acetate pH 7.5 plus 500 µl ice-cold ethanol at -80 °C for 30 min and subsequently spun down at 13,000 rpm for 1 min. After discarding the supernatant, the DNA pellets were washed with 700 µl 70% ethanol, dried, and resuspended in 30 µl ddH_2O.

The filter was pre-hybridized in 25 ml DIG Easy Hyb solution (Roche Applied Science) at 42 °C for 1 h under rotation in a hybridization oven. The DIG-labelled probe was denatured at 95 °C for 10 min, directly cooled on ice, and added to 25 ml DIG Easy Hyb solution to obtain a hybridization solution with a final concentration of min. 7.5 ng/ml DIG-labelled probe. Hybridization of the filter was then performed in the fresh hybridization solution at 42 °C overnight under rotation. Afterwards, the filter was washed twice for 20 min with 25 ml 2x SSC washing buffer (2x SSC, 0.1% (v/v) SDS) at room temperature, followed by two rounds of stringent washing for 20 min with 25 ml 0.5x SSC (0.5x SSC, 0.1% (v/v) SDS) at 58 °C. The filter was subsequently incubated with 25 ml blocking solution (1% Blocking Reagent (Roche Applied Science) in maleic acid buffer (150 mM NaCl, 100 mM maleic acid; pH 7.5)) at room temperature for 1 h under rotation. Fab fragments from polyclonal anti-digoxigenin antibodies conjugated to alkaline phosphatase (Roche Applied Science) were added to 25 ml fresh blocking solution (1:10,000), and the filter was incubated in this solution at room temperature for 30 min under rotation. The filter was then washed twice with maleic acid buffer plus 0.3% (v/v) Tween 20 at room temperature for 15 min and briefly rinsed with maleic acid buffer. After incubation in detection buffer (100 mM NaCl, 100 mM Tris-Cl pH 9.5) for 2 min, the filter was covered with CDP-*Star*, ready-to-use (Roche Applied Science) and exposed to the gel documentation system to detect chemiluminescence. Prior to stripping, the filter was washed in ddH_2O for 10 min. Thereafter, the filter was treated twice with stripping buffer (0.2 M NaOH, 0.1% (v/v) SDS) for 30 min at room temperature, washed twice in ddH_2O for 5 min, dried, and stored at room temperature.

Cloning of expression constructs containing the *pfa* gene clusters from *S. cellulosum* SBSo026 for heterologous LC-PUFA production in *M. xanthus*

For subcloning of the OA/LA-type *pfa* gene cluster of *S. cellulosum* SBSo026 from fosmid pCC1FOS B:I22 (Table 2) into a suitable expression construct for *M. xanthus* DK1622, a 3.3 kb vector backbone containing the p15A origin of replication, an ampicillin resistance gene, a neomycin resistance gene driven by Tn5 promoter, a tetracycline promoter, as well as the appropriate homology arms was amplified from plasmid pPfaAf-P_{tet} [11] by PCR using primers HA+*Pac*I+pACYC177-neo^R-P_{tet}_fwd and HA+pACYC177-neo^R-P_{tet}_rev (Table 1). PCR was performed with the PCR extender system (5 Prime) and standard conditions according to the manufacturer's protocol. The reactions were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 94 °C; 30 cycles consisting of denaturation for 20 s at 94 °C, annealing for 20 s at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. The linear plus circular homologous Red/ET recombineering was performed as described previously [11] using 1.5 µl of the PCR product, resulting in plasmid pPfaSo026-P_{tet} (Table 2). Afterwards, a zeocin resistance gene and the gene encoding the Mx9 integrase with *attP* site, both driven by Tn5 promoter derived from pMyx-zeo (S. C. Wenzel, unpublished), were inserted via *Pac*I restriction sites, generating plasmid pPfaSo026-P_{tet}-mx9.2 (Table 2).

For subcloning of the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026 from fosmid pCC1FOS C:C03 (Table 2) into a suitable expression construct for *M. xanthus* DK1622, a 2.1 kb vector backbone containing the p15A origin of replication, a tetracycline resistance gene, as well as the appropriate homology arms was amplified from plasmid pACYC184 (New England Biolabs) by PCR using primers HA+*Pac*I+pACYC184_fwd and HA+*Eco*81I+pACYC184_rev (Table 1). PCR was performed with the PCR extender system (5 Prime) and standard conditions according to the manufacturer's protocol. The reactions were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 94 °C; 30 cycles consisting of denaturation for 20 s at 94 °C, annealing for 20 s at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. The linear plus circular homologous Red/ET recombineering was performed as described previously [11] using 1.5 µl of the PCR product, resulting in plasmid pPfa2So026_1 (Table 2). Plasmid pHybPfa1 (Table 2) was constructed by exchange of the genes *pfa2* and *pfa3* of the OA/LA-type *pfa* gene cluster located on plasmid pPfaSo026-P_{tet} for the AA/EPA-type *pfa* gene cluster located on plasmid pPfa2So026_1 via *Eco*81I and *Pac*I restriction sites. Afterwards, a zeocin resistance gene and the gene encoding the Mx9

integrase with *attP* site, both driven by Tn5 promoter derived from pMyx-zeo (S. C. Wenzel, unpublished), were inserted via *Pac*I restriction sites, generating plasmid pHybPfa1-mx9.2 (Table 2). In order to eliminate *pfa1* for the construction of a plasmid exclusively comprising the AA/EPA-type *pfa* gene cluster (pPfa2So026-P_{tet}-mx9.2) via overlap extension PCR, a 1.3 kb DNA fragment comprising the neomycin resistance gene plus the tetracycline promoter and 0.8 kb of gene *pfa2* were amplified from plasmid pHybPfa1 by PCR using primers neo^R-P_{tet}_*Aan*I_fwd / neo^R-P_{tet}+overlap_rev and *pfa2_2*_So026+overlap_fwd / *pfa2_2*_So026_*Psc*I_rev (Table 1). PCR amplification of the two fragments was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 67 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The subsequent overlap extension PCR of the 2 kb fragment was performed as described for the amplification of the two fragments to be spliced using primers neo^R-P_{tet}_*Aan*I_fwd and *pfa2_2*_So026_*Psc*I_rev plus the two amplified fragments as templates. For the construction of plasmid pPfa2So026-P_{tet}-mx9.2 (Table 2), the neomycin resistance gene, the tetracycline promoter, gene *pfa1*, as well as parts of gene *pfa2* of plasmid pHybPfa1-mx9.2 were replaced by the overlap extension PCR amplicon neo^R-P_{tet}-5'*pfa2* by conventional restriction digest and ligation using *Aan*I and *Psc*I restriction sites.

Table 1. Primers used for amplification of probes and for cloning of expression plasmids in this study.
The homology arms are shown in italics and the introduced restriction sites are underlined.

Primer	Sequence (5'→3')
So026_probe 1_fwd	TGCACGTGGTGCGCGAGACGCCG
So026_probe 1_rev	CCTCCGGGCGCGAGATCTTCGCG
So026_probe 2_fwd	CATCATCGCGGGGCTCGGGGCC
So026_probe 2_rev	AGCTTGCGCAGCCAGTCCGCCCG
So026_probe 3_fwd	CCACCGAGCTCGTCGGCGAGTTCGC
So026_probe 3_rev	CCGTGCTCCAGGAACACGCGCACGC
HA+*Pac*I+pACYC177-neo^R-P$_{tet}$_fwd	*AGGCCCTCGTCCAGCGCTTGCTCTGGCGTTCTGCCTGGCCCGGTCAGCTTAATTAA*TCGCTCACTGACTCGC
HA+pACYC177-neo^R-P$_{tet}$_rev	*GAAGGCCGGCGCTGCGGAGGGCGTGCGGATTCTCTATCAT*TGCTGGGTGCCTCCTAG
HA+*Pac*I+pACYC184_fwd	*TCAGTAATCCCGATTCCGGCAGAACGCCTTGTACTTCTCCTTCACGAGCTTAATTAA*TCCGTTAGCGAGGTGCCG
HA+*Eco81*I+pACYC184_rev	*CGCCCACGATGGCAACGGGAGTGTGCGGCATGTGGTACCC*CTCAGGGCGGAAATGGCTTACGAACG
neo^R-P$_{tet}$_*Aan*I_fwd	GAATTTTATAAACCGTGGAGCGGGCAATACTG
neo^R-P$_{tet}$+overlap_rev	GATGGCAACGGGAGTGTGCGGCATTGCTGGGTGCCTCCTAGATTATTGTC
*pfa2_2*_So026+overlap_fwd	GACAATAATCTAGGAGGCACCCAGCAATGCCGCACACTCCCGTTGCCATC
*pfa2_2*_So026_*Psc*I_rev	GAAGCACATGTACATGAAGATGTCGTTCATGG

Table 2. Plasmids and expression strains constructed in this study.
oriV = origin of replication, *bla* = ampicillin resistance gene, *nptII* = neomycin resistance gene, *ble* = zeocin resistance gene, *tetA* = tetracycline resistance gene, P_{tet} = tetracycline promoter, PTn5 = Tn5 promoter, Mx9 *int* = Mx9 integrase gene with phage attachment site *attP*.

Plasmid / Strain	Characteristics
pCC1FOS B:I22	Derivative of pCC1FOS (Epicentre) with 38.7 kb of genomic DNA of *Sorangium cellulosum* SBSo026 including the OA/LA-type *pfa* gene cluster as insert
pPfaSo026-P_{tet}	Derivative of pPfaAf-P_{tet} [11] in which the *pfa* gene cluster of *Aetherobacter fasciculatus* (SBSr002) was replaced by the OA/LA-type *pfa* gene cluster of *S. cellulosum* SBSo026 located on fosmid pCC1FOS B:I22 by Red/ET recombineering. p15A *oriV, bla, nptII*
pPfaSo026-P_{tet}-mx9.2	Derivative of pPfaSo026-P_{tet} in which PTn5-*ble*-Mx9 *int* was inserted via *Pac*I
pCC1FOS C:C03	Derivative of pCC1FOS (Epicentre) with 44 kb of genomic DNA of *S. cellulosum* SBSo026 including the AA/EPA-type *pfa* gene cluster as insert
pPfa2So026_1	AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026 subcloned from pCC1FOS C:C03 by Red/ET recombineering. p15A *oriV, tetA*
pHybPfa1	Derivative of pPfaSo026-P_{tet} in which the genes *pfa2* and *pfa3* of the OA/LA-type *pfa* gene cluster were replaced by the AA/EPA-type *pfa* gene cluster located on plasmid pPfa2So026_1 via *Eco*81I and *Pac*I
pHybPfa1-mx9.2	Derivative of pHybPfa1 in which PTn5-*ble*-Mx9 *int* was inserted via *Pac*I
pPfa2So026-P_{tet}-mx9.2	Derivative of pHybPfa1-mx9.2 in which a fragment comprising *pfa1* was replaced by a corresponding fragment lacking *pfa1* via *Aan*I and *Psc*I
M. xanthus DK1622:: pPfaSo026-P_{tet}-mx9.2	*Myxococcus xanthus* DK1622 with pPfaSo026-P_{tet}-mx9.2 site-specifically integrated at Mx9 *attB2* site of the genome
M. xanthus DK1622::pHybPfa1-mx9.2	*M. xanthus* DK1622 with pHybPfa1-mx9.2 site-specifically integrated at Mx9 *attB2* site of the genome
M. xanthus DK1622:: pPfa2So026-P_{tet}-mx9.2	*M. xanthus* DK1622 with pPfa2So026-P_{tet}-mx9.2 site-specifically integrated at Mx9 *attB2* site of the genome

Transformation of *M. xanthus* by electroporation and heterologous expression

1.5 x 10^9 cells from an overnight culture of *M. xanthus* DK1622 were harvested at 8,000 rpm for 10 min. After discarding the supernatant, the cells were resuspended in 20 ml sterile ddH_2O. The cells were spun down at 8,000 rpm for 10 min, and the supernatant was discarded. The cells were resuspended in 40 µl sterile ddH_2O and transferred into a 2 ml tube. 2 µl of the expression plasmid were added. The cells were electroporated in an ice-cooled electroporation cuvette (1 mm) using a Bio-Rad electroporator at 650 V, 400 Ω, and 25 µF. 1 ml CTT-medium was added into the cuvette, and the cell suspension was transferred back into the tube. The tube was incubated in a heating block at 30 °C with shaking at 1,000 rpm for about 5-8 h. Thereafter,

the cells were mixed with 2-3 ml CTT-soft agar (CTT-medium plus 0.75% agar) containing 60 µg/ml kanamycin plus 20 µg/ml zeocin and plated onto two CTT-agar plates containing 60 µg/ml kanamycin plus 20 µg/ml zeocin. The plates were incubated at 30 °C for about 5 days. Thereafter, six colonies were transferred onto new CTT-agar plates containing 60 µg/ml kanamycin and 20 µg/ml zeocin. In order to verify the integration of the expression construct into the genome of *M. xanthus* via Mx9 *attB* sites, DNA from selected clones was isolated and used for PCR analyses [25]. The primers used are listed in Table 3. PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 56 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. The constructed strains are listed in Table 2. Expression was carried out in 50 ml CTT-medium containing 60 µg/ml kanamycin and 20 µg/ml zeocin at 30 °C. The cells were cultivated for 48 h and then harvested at 4,000 rpm for 10 min.

Table 3. Primers used for the verification of the integration of expression constructs into the genome of *M. xanthus* DK1622 via Mx9 *attB* sites.

Primer	Sequence (5'→3')
Mx9*attB1*_up	TGCCAGGGCTTACGGCTTC
Mx9*attB1*_down	CAGCACGGGTGCAGCAAC
Mx9*attB2*_up	TATCCCAGCAACCGCCGGAG
Mx9*attB2*_down	CGAGGTCCGGGACGCGCGCA
Mx9*attP*_up	GCGCCGAACTTAACAAGTTG
Mx9*attP*_down	TCCAGGTCCTCACGCTTGAC

Extraction of cellular fatty acids

The cellular fatty acids were extracted using the FAME method [26]. For this purpose, 50 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. The cell pellet was transferred to a glass vial and dried in a vacuum concentrator. Subsequently, the cell dry weight was determined. 5 µl (50 µg) of DHA (Sigma-Aldrich) and 500 µl of a mixture of methanol, toluene, and sulphuric acid (50:50:2, v/v/v) were added. The vial was capped with a teflon-lined screw cap and incubated at 80 °C for 24-48 h. After the mixture was cooled to room temperature, 400 µl of an aqueous solution consisting of 0.5 M NH_4HCO_3 and 2 M KCl were added, and the sample was vortexed for 30 s. Phase separation was achieved by

centrifugation at 4,000 rpm for 5 min at room temperature. 75 µl of the upper phase were mixed with 25 µl *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) and incubated at 37 °C for 30 min. Subsequently, the sample was used for GC-MS analysis.

Extraction and fractionation of lipids

Extraction of lipids from microbial cells was carried out using the method of Bligh and Dyer [27], modified by Lewis *et al.* [28], on a small scale. In the first step, the cell pellet from a 100 ml culture was transferred into a polypropylene tube. Successively, 4 ml chloroform, 8 ml methanol, and 3.2 ml 1% NaCl were added, and the tube was vortexed at high speed for 15 s after every addition. The sample was agitated on a tube rotator at 30 rpm overnight. 4 ml chloroform and 4 ml 1% NaCl were then added, and the tube was inverted 30 times. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. The bottom layer containing the lipid extract was evaporated to dryness under a gentle stream of nitrogen and dissolved in 1 ml of a chloroform:methanol mixture (2:1, v/v). The lipid solution was fractionated by solid-phase extraction on a Strata SI-1 silica column with 100 mg sorbent mass (Phenomenex) equilibrated with chloroform + 1% acetic acid. Neutral lipids and free fatty acids were eluted from the column with 1 ml of chloroform + 1% acetic acid, glycolipids were eluted with 1.5 ml of an acetone:methanol mixture (9:1, v/v), and phospholipids were eluted with 1 ml methanol. For GC-MS analysis, the fractions were dried in a vacuum concentrator and further processed according to the FAME method described previously.

Analysis of fatty acid methyl esters by GC-MS

GC-MS was carried out on an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a 7683B split/splitless injector with autosampler (Agilent Technologies) and coupled to a 5973 electron impact mass selective detector (Agilent Technologies). Helium was used as the carrier gas at a flow rate of 1 ml/min. 1 µl of the sample was injected in split mode (split ratio, 10:1). The analytical column was a (5% phenyl)-methylpolysiloxane capillary column (Agilent J&W DB-5ht; 30 m x 0.25 mm i.d. x 0.1 µm film thickness, maximum temperature 400 °C; Agilent Technologies). The column temperature was kept at 130 °C for 2.5 min, increased to 240 °C at a rate of 5 °C/min, then ramped to 300 °C at 30 °C/min, and held at 300 °C for 5 min. Other temperatures were as follows: inlet, 275 °C; GC-MS transfer line, 280 °C; ion source, 230 °C; and quadrupole, 150 °C. The mass selective

detector was operated in scan mode, scanning the mass range from *m*/*z* 40 to 700. Scan control, data acquisition, and processing were performed by MSD ChemStation and AMDIS software, version 2.68, based on the fragmentation patterns and retention times, in comparison with Supelco 37 Component FAME Mix and LC-PUFAs (all Sigma-Aldrich) as reference standards, and NIST 08 library. Absolute amounts of PUFAs were quantified by integration of the peaks using MSD ChemStation and by subsequent calculation in relation to the integral of DHA and to cell dry weight.

RESULTS AND DISCUSSION

Among various myxobacterial species, the two genera *Sorangium* and *Aetherobacter*, belonging to the same bacterial family *Polyangiaceae*, were found to produce diverse types of PUFAs *de novo*, employing PUFA synthases encoded by *pfa* biosynthetic gene clusters [11;12]. *Sorangium cellulosum* So ce56 [29], So ce10 [15], So ce1525 [16], So ce307 [17], So ce377 [18], So ce38 [19], So ce487, So ce836, So ce969, and So ceGT47 are producers of LA (18:2, *n*-6; ca. 10% of total fatty acid methyl esters (FAMEs)) and eicosadienoic acid (EDA, 20:2, *n*-6) in trace amounts (< 1% of total FAMEs). The *Aetherobacter* species *Aetherobacter* sp. (SBSr001) nov. ined., *Aetherobacter fasciculatus* sp. nov. ined. (SBSr002), *Aetherobacter rufus* sp. nov. ined. (SBSr003) [30], and *Aetherobacter* sp. nov. ined. (SBSr008) produce arachidonic acid (AA, 20:4, *n*-6; ca. 1% of total FAMEs), EPA (20:5, *n*-3; ca. 7-10% of total FAMEs), and DHA (22:6, *n*-3; ca. 10-19% of total FAMEs), as well as LA, γ-linolenic acid (GLA, 18:3, *n*-6), stearidonic acid (SDA, 18:4, *n*-3), *n*-6/*n*-3 DPA (22:5) in trace amounts (< 1% of total FAMEs). Genome sequence data revealed the presence of a distinct type of *pfa* gene cluster in the analysed *S. cellulosum* strains, responsible for the production of LA and/or oleic acid (OA, 18:1; *n*-9) and of another type of *pfa* gene cluster in *Aetherobacter* strains, most likely responsible for the production of DHA, *n*-6 DPA, and *n*-3 DPA (Figure 1), as proven by heterologous expression in *M. xanthus*. As a consequence, these gene clusters can be designated as OA/LA-type *pfa* gene cluster and DPA/DHA-type *pfa* gene cluster, respectively.

During continuous efforts to discover novel myxobacterial isolates, which optionally includes the analysis of their fatty acid profiles, the prolific PUFA producer *S. cellulosum* strain SBSo026 was found. GC-MS analysis of the FAMEs carried out for this strain disclosed the production of LA (6.2% of total FAMEs) and EDA (1.3% of total FAMEs).

Additionally, EPA (7.1% of total FAMEs), as well as GLA, SDA, dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6), AA, eicosatetraenoic acid (ETA, 20:4, *n*-3), and *n*-3 DPA in trace amounts (< 1% of total FAMEs) could be detected. This high capability of *S. cellulosum* SBSo026 for production of various PUFA species stimulated sequencing and screening of the genome of this strain, which had been cloned in a fosmid library, for putative PUFA biosynthetic pathways. **T**he fosmid library was screened for the OA/LA-type *pfa* gene cluster at low stringency using probe 1, which hybridizes at the 5' end of *pfa1*, and probe 2, which hybridizes at the 3' end of *pfa3*. This resulted in the identification of fosmid pCC1FOS B:I22, comprising a *pfa* gene cluster homologous to the OA/LA-type *pfa* gene cluster of *S. cellulosum* So ce56, So ce10, So ce1525, So ce307, So ce377, So ce38, So ce487, So ce836, So ce969, and So ceGT47 [11], plus flanking genes. The 17.2 kb OA/LA-type *pfa* gene cluster of *S. cellulosum* SBSo026 (Figure 1) was completely sequenced by primer walking or subcloning of fragments prior to Sanger sequencing. On top of that, an additional, second *pfa* gene cluster, only comprising genes *pfa2* and *pfa3* (Figure 1), was discovered in the genome sequence of *S. cellulosum* SBSo026. In order to detect fosmids harbouring the second PUFA biosynthetic pathway, probe 3, targeting the unique AT domain of *pfa3*, was used for hybridization at low stringency. Fosmids pCC1FOS C:C03, C:O01, A:I24, B:J03, A:L03, B:E08, E:O20, and A:F04, containing the 15.6 kb second *pfa* gene cluster (or parts thereof) plus adjacent genes, could be identified. Fosmid pCC1FOS C:C03 was completely sequenced by next generation sequencing. This second type of *pfa* gene cluster of *S. cellulosum* SBSo026 differs from the OA/LA-type *pfa* gene cluster in terms of number of the catalytic domains (Figure 1), which might be indicative of an extended genetic capability of the strain for production of diverse PUFA species. In order to check for the presence of the novel PUFA biosynthetic pathway among other *S. cellulosum* strains, all genome sequence data available from *S. cellulosum* strains (including the So ce strains mentioned previously plus So ce1128 [31], So ce340, So ce427, So ce690 [32], and So0157-2 [33]) was screened for *pfa* gene clusters. The formerly characterized OA/LA-type *pfa* gene cluster could be detected in every So ce strain already mentioned, whereas the second type of *pfa* gene cluster could exclusively be identified in the genome sequences of *S. cellulosum* SBSo026, So ce1128, So ce340, So ce427, and So ce690. FAME analysis unveiled that all So ce strains are capable of producing LA, but EPA production could only be observed for the strains harbouring two types of *pfa* gene clusters. Thus, the potential for EPA production is most likely coupled to the presence of this second *pfa* gene cluster, which seems to encode a novel PUFA synthase.

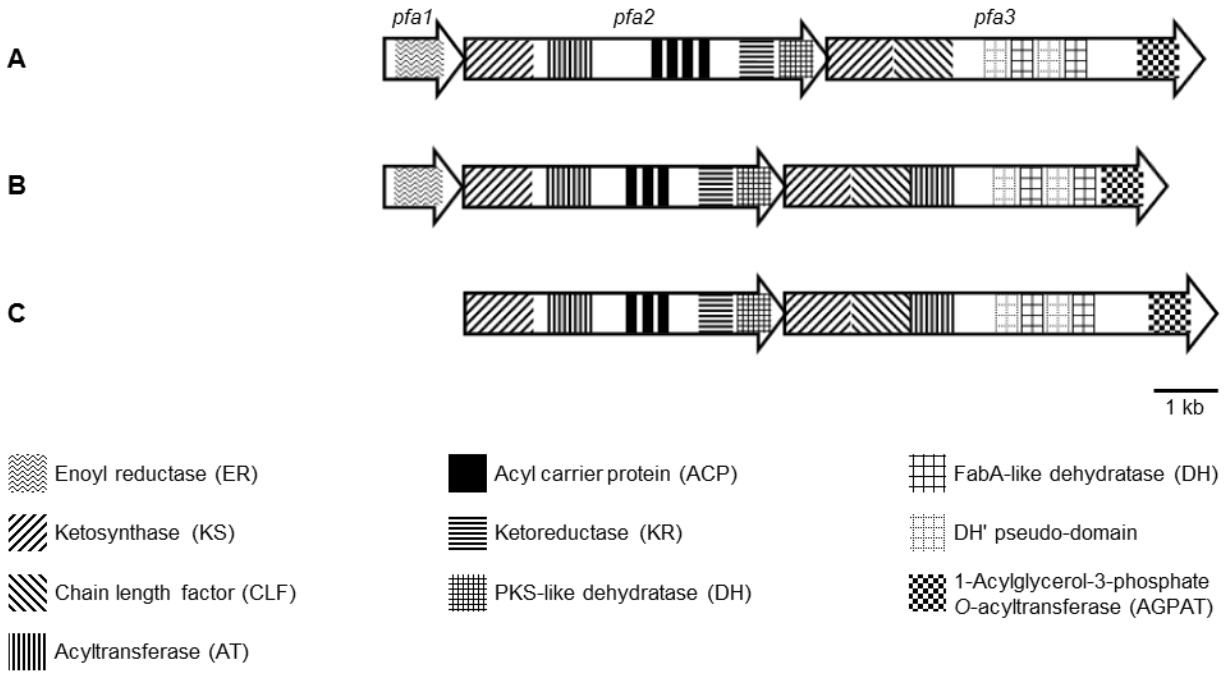


Figure 1. PUFA biosynthetic gene clusters from terrestrial myxobacteria of the *Polyangiaceae* family.
(A) *Sorangium cellulosum* (OA/LA-type *pfa* gene cluster; the number of ACP domains is ranging from four to five), (B) *Aetherobacter* sp. (DPA/DHA-type *pfa* gene cluster), (C) *Sorangium cellulosum* (AA/EPA-type *pfa* gene cluster).

Comparison of the two types of myxobacterial PUFA biosynthetic gene clusters of *S. cellulosum* and the DPA/DHA-type *pfa* gene cluster of *Aetherobacter* spp. disclosed that they share most of their functional elements (Figure 1). The most remarkable differences are the absence of gene *pfa1* encoding an enoyl reductase (ER) in the second *pfa* gene cluster of *S. cellulosum*, the variable number of tandem acyl carrier protein (ACP) domains, and the lack of an acyltransferase (AT) domain within Pfa3 encoded by the OA/LA-type *pfa* gene cluster of *Sorangium*. As the ER domain is supposed to represent a crucial catalytic activity of PUFA synthases, Pfa1 encoded by the OA/LA-type *pfa* gene cluster is assumed to complement the activity of the PUFA synthase encoded by genes *pfa2* and *pfa3* of the second PUFA biosynthetic pathway identified in some *S. cellulosum* strains. Notably, an AT domain is present within Pfa3 encoded by the DPA/DHA-type *pfa* gene cluster of *Aetherobacter* spp. and the second *pfa* gene cluster of *S. cellulosum* which could not be identified in the OA/LA-type PUFA synthase. This might explain the inability of the latter PUFA biosynthetic pathway to synthesize LC-PUFAs like EPA or DHA, a speculation supported by a recent study suggesting that the homologous AT domain from marine γ-Proteobacteria may play a role in determining the final LC-PUFA product [34]. Detailed analyses on the sequence level revealed that the Pfa proteins of the second PUFA synthases from *S. cellulosum* SBSo026,

So ce1128, So ce340, and So ce427 show a pairwise sequence identity of > 90% among the different producer strains. Comparison of the second PUFA synthase and the putative OA/LA-type PUFA synthase of *S. cellulosum* SBSo026 unveiled considerably lower similarity (pairwise sequence identity of ≤ 53%). This value is in the range of the observed low pairwise sequence identity of ≤ 63% when comparing the second PUFA synthase of *S. cellulosum* SBSo026 with the DPA/DHA-type PUFA synthase of *A. fasciculatus* (SBSr002). These data indicate that the sequences of myxobacterial PUFA synthases correlate significantly more with their product specificity than with their source, which implies that the second *pfa* gene cluster rather originates from horizontal gene transfer into a *S. cellulosum* ancestor already harbouring the OA/LA-type *pfa* gene cluster than results from a duplication event of this gene cluster.

In order to prove the functionality and the product spectrum of the *pfa* gene clusters from the novel myxobacterial isolate *S. cellulosum* SBSo026, their heterologous expression was performed in the myxobacterial model strain *M. xanthus* DK1622, established for recombinant LC-PUFA production [11]. The practical application initially required the subcloning of the gene clusters from fosmids into suitable plasmids and subsequent engineering of the constructs**.** Cloning of the OA/LA-type *pfa* gene cluster from fosmid pCC1FOS B:I22 into a plasmid containing the p15A origin of replication and an ampicillin resistance gene was achieved by linear plus circular homologous Red/ET recombination [35] (Figure 2A). During this step, the native promoter upstream of *pfa1* was replaced with a tetracycline promoter downstream of a neomycin resistance gene driven by Tn5 promoter. In a final engineering step, a zeocin resistance gene and a gene encoding the Mx9 integrase with *attP* site for the site-specific integration into the host genome via the Mx9 phage attachment site [36] were ligated into the obtained plasmid pPfaSo026-P_{tet}, yielding expression plasmid pPfaSo026-P_{tet}-mx9.2. In parallel, the second *pfa* gene cluster was cloned from fosmid pCC1FOS C:C03 into a plasmid containing the p15A origin of replication and a tetracycline resistance gene by linear plus circular homologous Red/ET recombination (Figure 2B). The backbone of the obtained plasmid pPfa2So026_1 was exchanged for the backbone of plasmid pPfaSo026-P_{tet} containing the p15A origin of replication and an ampicillin resistance gene by applying conventional cloning techniques. During this step, gene *pfa1* of the OA/LA-type *pfa* gene cluster together with the upstream located tetracycline promoter and the neomycin resistance gene driven by Tn5 promoter was inserted in front of genes *pfa2* and *pfa3*. Afterwards, a zeocin resistance gene and a gene encoding the Mx9 integrase with *attP* site were ligated into the obtained plasmid pHybPfa1, yielding expression plasmid

pHybPfa1-mx9.2. Gene *pfa1* was deleted via restriction digest and religation to construct the derivative pPfa2So026-P_{tet}-mx9.2.

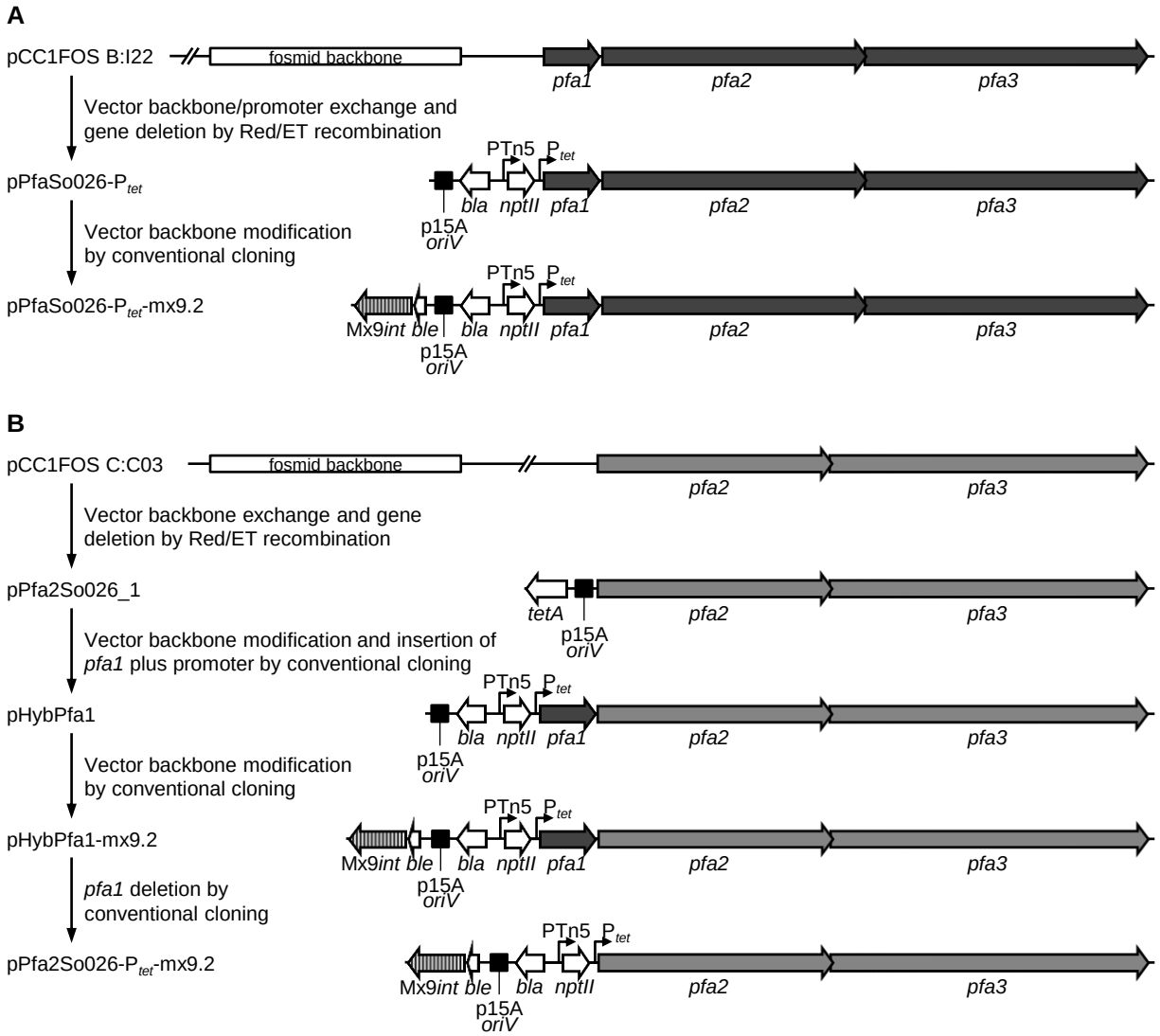


Figure 2. Cloning strategy for the construction of expression plasmids for recombinant LC-PUFA production in *Myxococcus xanthus* DK1622.
(A) Cloning of the OA/LA-type *pfa* gene cluster from *Sorangium cellulosum* SBSo026, (B) Cloning of the AA/EPA-type *pfa* gene cluster from *S. cellulosum* SBSo026 and construction of a hybrid *pfa* gene cluster. *oriV* = origin of replication, *bla* = ampicillin resistance gene, *nptII* = neomycin resistance gene, *ble* = zeocin resistance gene, *tetA* = tetracycline resistance gene, P_{tet} = tetracycline promoter, PTn5 = Tn5 promoter, Mx9*int* = Mx9 integrase gene with phage attachment site *attP*. Flanking genes on the fosmids are not illustrated.

The generated expression constructs pPfaSo026-P_{tet}-mx9.2, pPfa2So026-P_{tet}-mx9.2, and pHybPfa1-mx9.2 were transferred into *M. xanthus* DK1622 by electroporation. Six kanamycin and zeocin resistant transformants of *M. xanthus* DK1622::pPfaSo026-P_{tet}-mx9.2, *M. xanthus* DK1622::pPfa2So026-P_{tet}-mx9.2, or *M. xanthus* DK1622::pHybPfa1-mx9.2 were randomly selected to verify the correct integration via

Mx9 *attB* sites by PCR analysis. For every transformant, integration of the expression construct at the *attB2* site [36] could be confirmed. The transgenic clones were cultivated in 50 ml scale at 30 °C for two days. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [26] and analysed by GC-MS (Table 4). All transgenic clones of *M. xanthus* DK1622::pPfaSo026-P_{tet}-mx9.2 were shown to produce OA, isomers of octadecenoic acid, LA, and GLA, but also EDA as well as DHGLA in trace amounts. These fatty acids are not present in *M. xanthus* DK1622 wild type, which therefore provided functional proof that the designated OA/LA-type PUFA pathway can indeed direct the biosynthesis towards C18 MUFAs/PUFAs (and also C20 derivatives). However, in the native producer strain *S. cellulosum* SBSo026 production of these fatty acids is clearly shifted to LA and EDA and only low amounts of OA, isomers of octadecenoic acid, GLA, as well as DHGLA can be detected after seven days of cultivation. When analysing the fatty acid profiles of the six transgenic clones of *M. xanthus* DK1622::pPfa2So026-P_{tet}-mx9.2, expressing genes *pfa2* and *pfa3* of the second PUFA pathway, no deviation from the fatty acid profile of the expression host *M. xanthus* DK1622 wild type could be observed. This indicates that the PUFA synthase encoded by *pfa2* and *pfa3* is not fully functional most likely due to the lack of the ER activity encoded by *pfa1* homologues. Indeed, when coexpressing Pfa1 by construction of a hybrid pathway (including *pfa1* of the OA/LA-type PUFA gene cluster), production of PUFAs could be observed. All transgenic clones of *M. xanthus* DK1622::pHybPfa1-mx9.2 were shown to produce EPA and AA, as well as LA, GLA, SDA, DHGLA, ETA, and *n*-3 DPA in trace amounts. However, production yields of EPA achieved with the native producer strain after seven days of cultivation surpass the production rate of the heterologous host strain.

Comparison of LC-PUFA production profiles of the transgenic hosts *M. xanthus* DK1622::pPfaSo026-P_{tet}-mx9.2, *M. xanthus* DK1622::pPfa2So026-P_{tet}-mx9.2, or *M. xanthus* DK1622::pHybPfa1-mx9.2 with the native producer *S. cellulosum* SBSo026 allows for drawing of several conclusions. The observed changes in the fatty acid profile after heterologous expression of the OA/LA-type *pfa* gene cluster from *S. cellulosum* SBSo026 are consistent with the results of the previously reported heterologous expression of the homologous *pfa* gene cluster from *S. cellulosum* So ce56 [11]. Together with the results of previously performed feeding experiments [11], it can be deduced that the OA/LA-type *pfa* gene cluster of *Sorangium* is most likely responsible for the production of EDA, isomers of octadecenoic acid, and either of both OA and LA or of merely OA, which could be converted

into LA by an oleoyl-CoA-Δ^{12} desaturase (EC 1.14.19.6). LA is then most likely converted into GLA in *M. xanthus* DK1622, probably catalysed by a linoleoyl-CoA desaturase (EC 1.14.19.3), whereas a Δ^8 desaturase (EC 1.14.19.4) could convert EDA into DHGLA in *M. xanthus* DK1622. As expected, the PUFA synthase encoded by the incomplete, second *pfa* gene cluster of *Sorangium* turned out to be not functional in the absence of the third subunit encoding the required ER activity. However, Pfa1 encoded by the OA/LA-type *pfa* gene cluster was shown to effectively complement the multienzyme complex to reconstitute a functional PUFA synthase. This hybrid PUFA synthase most likely produces EPA and some minor amounts of AA, ETA, DHGLA, and *n*-3 DPA as by-products and can therefore be designated as AA/EPA-type PUFA synthase. The generated C20 PUFAs ETA and DHGLA could then be catabolized to the corresponding C18 PUFAs SDA and GLA via β-oxidation. However, further studies are required to obtain deeper insights into the cellular metabolism of the fatty acid products in the heterologous host as well as in the natural producer.

Table 4. Production of MUFAs and PUFAs with a carbon chain length of ≥ C18 by *Sorangium cellulosum* SBSo026, *Myxococcus xanthus* DK1622::pPfaSo026-P_{tet}-mx9.2, and *M. xanthus* DK1622::pHybPfa1-mx9.2. Cultivation was carried out in 50 ml HS-medium at 30 °C and 200 rpm for 7 days (for *S. cellulosum*) or in 50 ml CTT-medium at 30 °C and 200 rpm for 2 days (for *M. xanthus*). The indicated values for *M. xanthus* are means of three biological samples.

MUFA or PUFA	Absolute amount [mg/g cell dry weight]		
	***S. cellulosum* SBSo026**	***M. xanthus* DK1622:: pPfaSo026-P_{tet}-mx9.2**	***M. xanthus* DK1622:: pHybPfa1-mx9.2**
oleic acid (OA, 18:1, *n*-9)	0.1	0.1	–
linoleic acid (LA, 18:2, *n*-6)	3.3	0.9	< 0.1
γ-linolenic acid (GLA, 18:3, *n*-6)	< 0.1	1.1	< 0.1
stearidonic acid (SDA, 18:4, *n*-3)	0.2	–	< 0.1
eicosadienoic acid (EDA, 20:2, *n*-6)	0.7	< 0.1	–
dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6)	< 0.1	< 0.1	< 0.1
arachidonic acid (AA, 20:4, *n*-6)	0.2	–	0.2
eicosatetraenoic acid (ETA, 20:4, *n*-3)	< 0.1	–	< 0.1
eicosapentaenoic acid (EPA, 20:5, *n*-3)	3.7	–	1.2
docosapentaenoic acid (DPA, 22:5, *n*-3)	0.1	–	< 0.1

Another very interesting feature associated with the presence of the AA/EPA-type *pfa* gene cluster was obtained when screening our database for the secondary metabolite profiles of the analysed *Sorangium* species. In contrast to the strains that only harbour the OA/LA-type *pfa* gene cluster, all *S. cellulosum* strains additionally comprising the AA/EPA-type *pfa* gene cluster were shown to produce maracin A or maracen A (Figure 3A). These secondary metabolites show a good activity against *Mycobacterium tuberculosis* and have a low cytotoxicity [31]. They represent lipophilic carboxylic acids with 19 carbon atoms, four to five *trans* double bonds, and one *cis* double bond. Maracin A contains an ethinyl-*trans*-vinyl ether as a very unusual structural element, whereas the alkinyl group is replaced by an α-chlorovinyl group in maracen A. Feeding experiments revealed the biosynthetic origin of this compound family containing nine complete acetate units plus a carboxyl group originating from another acetate unit [31]. When analysing the flanking regions of AA/EPA-type *pfa* gene clusters, 18 genes could be identified, which are exclusively present in the genomes of the EPA producing So ce strains. *In silico* analyses based on homology searches indicate that these genes encode several proteins, possibly involved in the conversion of EPA into maracen/maracin. These include putative homologues of a Δ^{12}-fatty acid dehydrogenase, a zinc-dependent alcohol dehydrogenase, a phospholipid/glycerol acyltransferase, four phospholipases, an enoyl-CoA hydratase, a cytochrome P450, a lipoxygenase, a lipase maturation factor, two PUFA isomerases, and five proteins of unknown function. By means of the identified putative enzyme activities, a hypothesis for the biosynthesis of maracin/maracen starting from EPA was developed: The 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) domain encoded by the AA/EPA-type *pfa* gene cluster most likely catalyses the incorporation of EPA from the assembly line into the phospholipids. The phospholipases encoded adjacently to the *pfa* genes could hydrolyse the ester bond between EPA and the phospholipids and thereby release the free LC-PUFA as precursor for maracin/maracen biosynthesis. Interestingly, extraction and fractionation of the lipids from *S. cellulosum* SBSo026, subsequent preparation of FAMEs by transesterification, and GC-MS analysis showed that most of the synthesized EPA (61%) was present in the fraction containing neutral lipids plus free fatty acids. 36% of produced EPA were incorporated into the phospholipids and 3% were present in the glycolipid fraction. The required isomerizations of the double bonds to enable maracin/maracen from EPA are assumed to be carried out by the PUFA isomerases. Incorporation of an oxygen atom to form the internal ether moiety via a hydroperoxide intermediate may be catalysed by the lipoxygenase and the cytochrome P450 [37-39]. The enoyl-CoA hydratase and the alcohol dehydrogenase might participate in

α-oxidation of the C20 fatty acid derivative to generate a C19 chain, whereas the Δ^{12}-fatty acid dehydrogenase might be involved in the oxidation of the *cis* double bond next to the ether group to a triple bond. However, further biochemical and genetic studies need to be performed to validate this theory and to elucidate the biosynthesis of maracin/maracen.

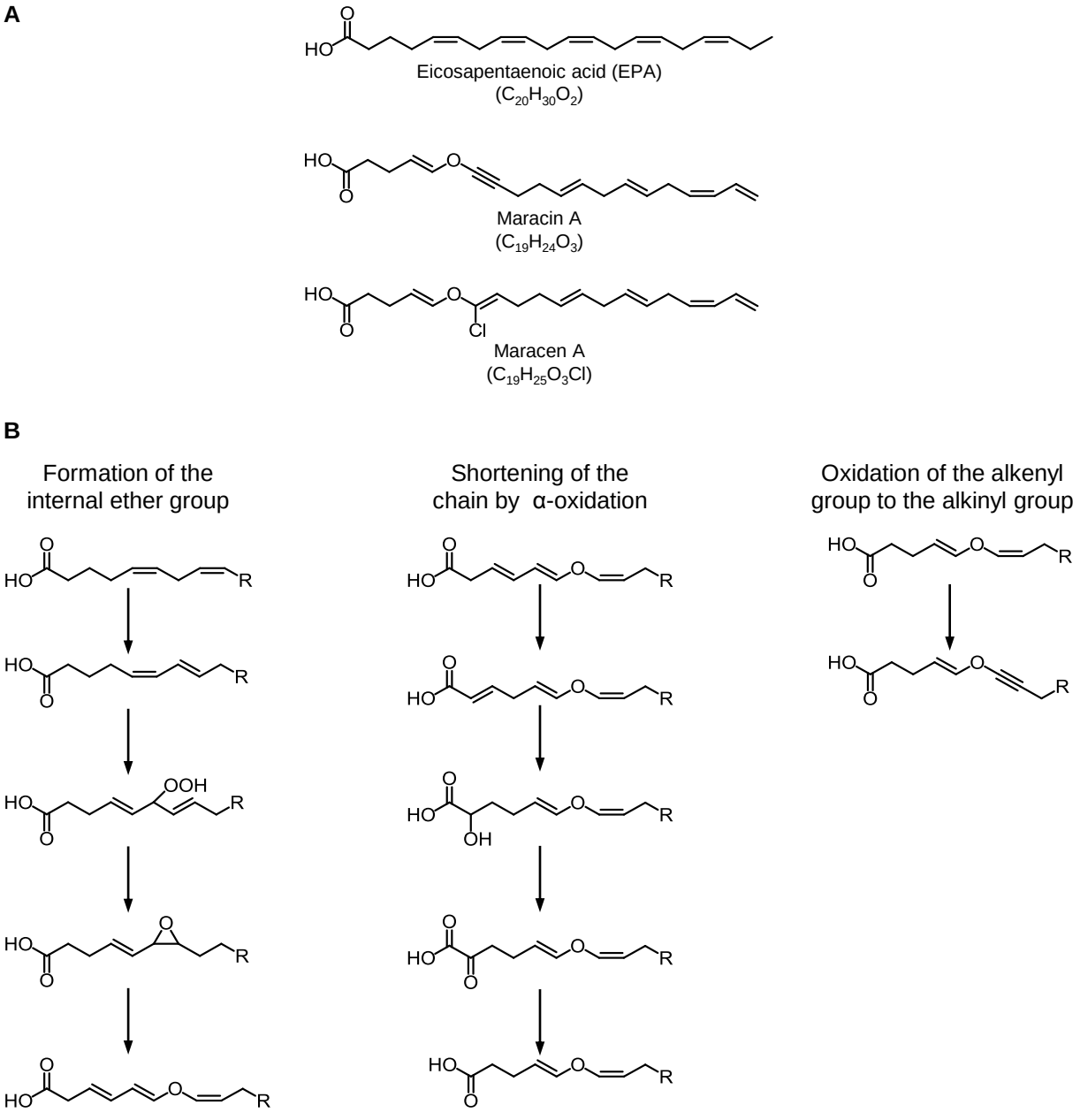


Figure 3. (A) Structures of eicosapentaenoic acid (EPA), maracin A, and maracen A and (B) proposed biosynthesis of structural elements of maracin A/maracen A starting from EPA.
Incorporation of an oxygen atom to form the internal ether bond via a hydroperoxide intermediate may be catalysed by the lipoxygenase and the cytochrome P450; α-oxidation of the C20 fatty acid derivative to generate a C19 chain might be catalysed with the involvement of the enoyl-CoA hydratase and the alcohol dehydrogenase, among others; the Δ^{12}-fatty acid dehydrogenase might be involved in the oxidation of the *cis* double bond next to the ether group to a triple bond. The required isomerizations of the double bonds are assumed to be carried out by the PUFA isomerases. It is uncertain if all these reactions occur on the free carboxylic acids or on the corresponding acyl-CoA thioesters.

REFERENCES

[1] Warude D, Joshi K, & Harsulkar A (2006) Polyunsaturated fatty acids: Biotechnology. *Crit. Rev. Biotechnol.*, **26,** 83-93.

[2] Kaulmann U & Hertweck C (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew. Chem. Int. Ed.*, **41,** 1866-1869.

[3] Napier JA (2002) Plumbing the depths of PUFA biosynthesis: a novel polyketide synthase-like pathway from marine organisms. *Trends Plant Sci.*, **7,** 51-54.

[4] Wallis JG, Watts JL, & Browse J (2002) Polyunsaturated fatty acid synthesis: what will they think of next? *Trends Biochem. Sci.*, **27,** 467-473.

[5] 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.

[6] 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.

[7] Morita N, Tanaka M, & Okuyama H (2000) Biosynthesis of fatty acids in the docosahexaenoic acid-producing bacterium *Moritella marina* strain MP-1. *Biochem. Soc. Trans.*, **28,** 943-945.

[8] Hauvermale A, Kuner J, Rosenzweig B, Guerra D, Diltz S, & Metz JG (2006) Fatty acid production in *Schizochytrium* sp.: Involvement of a polyunsaturated fatty acid synthase and a type I fatty acid synthase. *Lipids*, **41,** 739-747.

[9] Okuyama H, Orikasa Y, Nishida T, Watanabe K, & 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] Shulse CN & Allen EE (2011) Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. *PLOS One*, **6,** e20146.

[11] Gemperlein K, Rachid S, Garcia R, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

[12] Garcia R, Pistorius D, Stadler M, & Müller R (2011) Fatty acid-related phylogeny of myxobacteria as an approach to discover polyunsaturated omega-3/6 fatty acids. *J. Bacteriol.*, **193,** 1930-1942.

[13] Stadler M, Roemer E, Müller R, Garcia RO, Pistorius D, & Brachmann A (2010) Production of omega-3 fatty acids by myxobacteria. *Int. Pat.*, WO 2010/063451.

[14] Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez-Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Munoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Rückert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorhölter FJ, Weissman KJ, Welch RD, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Blöcker H, Pühler A, & Müller R (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat. Biotechnol.*, **25,** 1281-1289.

[15] Hopf NW, Yonsel S, & Deckwer WD (1990) Ambruticin S production in air-lift and stirred-tank bioreactors. *Appl. Microbiol. Biotechnol.*, **34,** 350-353.

[16] Gerth K, Steinmetz H, Höfle G, & Jansen R (2008) Chlorotonil A, a macrolide with a unique gem-dichloro-1,3-dione functionality from *Sorangium cellulosum*, So ce1525. *Angew. Chem. Int. Ed.*, **47,** 600-602.

[17] Julien B, Tian Z-Q, Reid R, & Reeves CD (2006) Analysis of the ambruticin and jerangolid gene clusters of *Sorangium cellulosum* reveals unusual mechanisms of polyketide biosynthesis. *Chem. Biol.*, **13,** 1277-1286.

[18] Irschik H, Augustiniak H, Gerth K, Höfle G, & Reichenbach H (1995) The ripostatins, novel inhibitors of eubacterial RNA polymerase isolated from myxobacteria. *J. Antibiot.*, **48,** 787-792.

[19] Jahns C, Hoffmann T, Müller S, Gerth K, Washausen P, Höfle G, Reichenbach H, Kalesse M, & Müller R (2012) Pellasoren: structure elucidation, biosynthesis, and total synthesis of a cytotoxic secondary metabolite from *Sorangium cellulosum*. *Angew. Chem. Int. Ed.*, **51,** 5239-5243.

[20] Grant SG, Jessee J, Bloom FR, & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U S A*, **87,** 4645-4649.

[21] Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF, & Zhang Y (2006) An improved recombineering approach by adding RecA to λ Red recombination. *Mol. Biotechnol.*, **32,** 43-53.

[22] Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Müller R, Stewart AF, & Zhang Y (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.*, **30,** 440-446.

[23] Chen H, Keseler IM, & Shimkets LJ (1990) Genome size of *Myxococcus xanthus* determined by pulsed-field gel electrophoresis. *J. Bacteriol.*, **172,** 4206-4213.

[24] Sambrook J & Russell DW (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[25] Oßwald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, & Wenzel SC (2012) Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.* Epub ahead of print.

[26] Bode HB, Ring MW, Schwär G, Kroppenstedt RM, Kaiser D, & Müller R (2006) 3-Hydroxy-3-methylglutaryl-coenzyme A (CoA) synthase is involved in the biosynthesis of isovaleryl-CoA in the myxobacterium *Myxococcus xanthus* during fruiting body formation. *J. Bacteriol.*, **188,** 6524-6528.

[27] Bligh EG & Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can. J. Biochem. and Phys.*, **37,** 911-917.

[28] Lewis T, Nichols PD, & McMeekin TA (2000) Evaluation of extraction methods for recovery of fatty acids from lipid-producing microheterotrophs. *J. Microbiol. Methods*, **43,** 107-116.

[29] Pradella S, Hans A, Sproer C, Reichenbach H, Gerth K, & Beyer S (2002) Characterisation, genome size and genetic manipulation of the myxobacterium *Sorangium cellulosum* So ce56. *Arch. Microbiol.*, **178,** 484-492.

[30] Plaza A, Garcia R, Bifulco G, Martinez JP, Huttel S, Sasse F, Meyerhans A, Stadler M, & Müller R (2012) Aetheramides A and B, potent HIV-inhibitory depsipeptides from a myxobacterium of the new genus "Aetherobacter". *Org. Lett.*, **14,** 2854-2857.

[31] Herrmann M, Böhlendorf B, Irschik H, Reichenbach H, & Höfle G (1998) Maracin und Maracen: Neuartige Ethinyl-vinylether- bzw. α-Chlordivinylether-Antibiotica aus *Sorangium cellulosum* mit spezifischer Wirkung gegen Mycobakterien. *Angew. Chem.*, **110,** 1313-1315.

[32] Bode HB, Irschik H, Wenzel SC, Reichenbach H, Müller R, & Höfle G (2003) The leupyrrins: A structurally unique family of secondary metabolites from the myxobacterium *Sorangium cellulosum*. *J. Nat. Prod.*, **66,** 1203-1206.

[33] Han K, Li ZF, Peng R, Zhu LP, Zhou T, Wang LG, Li SG, Zhang XB, Hu W, Wu ZH, Qin N, & Li YZ (2013) Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci. Rep.*, **3,** 2101.

[34] Orikasa Y, Tanaka M, Sugihara S, Hori R, Nishida T, Ueno A, Morita N, Yano Y, Yamamoto K, Shibahara A, Hayashi H, Yamada Y, Yamada A, Yu RK, Watanabe K, & Okuyama H (2009) *pfaB* products determine the molecular species produced in bacterial polyunsaturated fatty acid biosynthesis. *FEMS Microbiol. Lett.*, **295,** 170-176.

[35] Zhang Y, Muyrers JPP, Testa G, & Stewart AF (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18,** 1314-1317.

[36] Julien B (2003) Characterization of the integrase gene and attachment site for the *Myxococcus xanthus* bacteriophage Mx9. *J. Bacteriol.*, **185,** 6325-6330.

[37] Fahlstadius P & Hamberg M (1990) Stereospecific removal of the pro-R hydrogen at C-8 of (9S)-hydroperoxyoctadecadienoic acid in the biosynthesis of colneleic acid. *J. Chem. Soc., Perkin Trans. 1*, **1990,** 2027-2030.

[38] Jiang ZD & Gerwick WH (1997) Novel oxylipins from the temperate red alga *Polyneura latissima*: Evidence for an arachidonate 9(S)-lipoxygenase. *Lipids*, **32,** 231-235.

[39] Grechkin AN (2002) Hydroperoxide lyase and divinyl ether synthase. *Prostaglandins Other Lipid Mediat.*, **68-69,** 457-470.

CHAPTER 3

Heterologous expression of a myxobacterial PUFA biosynthetic pathway in *Escherichia coli* and *Pseudomonas putida* and production enhancement by genetic and metabolic engineering

INTRODUCTION

Long-chain polyunsaturated fatty acids (LC-PUFAs), especially eicosapentaenoic acid (EPA, 20:5, *n*-3) and docosahexaenoic acid (DHA, 22:6, *n*-3), show advantageous effects on human health, like prevention and treatment of cardiovascular diseases, obesity, and diabetes [1]. As more and more people strive to benefit from these positive effects, the demand of *n*-3 LC-PUFAs as dietary supplements has increased intensely over the past years - with upward tendency for the future. These days, most of EPA and DHA are obtained from oceanic fish and fish oil, but these natural sources are depleting and often contaminated with environmental toxins. In order to satisfy the demand for high-quality LC-PUFAs, the quest for alternative, sustainable sources is indispensable [2]. Fermentation of prokaryotic and eukaryotic microorganisms capable of producing EPA or DHA in high amounts might have the potential to permit industrial-scale production of LC-PUFAs [3]. Until recently, it was thought that PUFA-producing microbes are from marine origin. Some of them synthesize LC-PUFAs *de novo* from acyl-CoA precursors by iteratively acting polyketide synthase (PKS)-like enzymes known as PUFA synthases. These multienzyme complexes are encoded by PUFA (*pfa*) biosynthetic gene clusters [4-6]. Initially, the occurrence of *pfa* gene clusters was described in the marine γ-Proteobacteria *Shewanella pneumatophori* SCRC-2738 and *Photobacterium profundum* SS9 (EPA producers) [7;8], *Moritella marina* MP-1 (DHA producer) [9], and in the marine microalga *Schizochytrium* sp. ATCC 20888 (DHA and *n*-6 docosapentaenoic acid (DPA, 22:5) producer) [7;10]. Establishment of optimal fermentation conditions for these marine microorganisms [11-13], treatment with the fatty acid synthase inhibitor cerulenin [14;15], or transposon mutagenesis [16] gave rise to an improved production of LC-PUFAs under laboratory conditions. In order to correlate the biosynthetic pathways with their products, to reduce the cultivation time, and/or to study LC-PUFA biosynthesis, several marine PUFA biosynthetic gene clusters were transferred and expressed into suitable host organisms. Thereby, recombinant production of EPA, *n*-6 DPA, and/or DHA could be accomplished with the *pfa* gene clusters from *Shewanella* sp., *M. marina* MP-1, or *Schizochytrium* sp. by use of *Synechococcus* sp., *Lactococcus lactis* ssp. *cremoris*, and the well-established host *Escherichia coli* [10;17-20]. Besides of optimizing the cultivation conditions for the latter host harbouring the *pfa* gene cluster from *Shewanella* strains [21;22], recombinant EPA production could be enhanced by substitution of native promoters with heterologous promoters to increase *pfa* gene expression levels [23] or by intracellular accumulation of a high-performance catalase [24].

Recently, PUFA synthases catalysing the biosynthesis of LC-PUFAs could be identified and characterized from several terrestrial myxobacterial species. They were found to be encoded by two types of *pfa* gene clusters identified in the genomes of the linoleic acid (LA, 18:2, *n*-6) producing myxobacteria *Sorangium cellulosum* as well as of the novel myxobacterial isolates *Aetherobacter* spp., described as prolific producers of EPA and DHA [25]. However, these native producer strains grow very slowly, are not easy to handle, and genetic modifications are difficult to implement. Cloning, transfer and heterologous expression of the *pfa* genes into the myxobacterial model strain *Myxococcus xanthus* has been demonstrated as a promising strategy to improve LC-PUFA production and to reduce cultivation times [25]. Another attractive host for recombinant LC-PUFA production represents *Pseudomonas putida* KT2440. The strain is generally recognized as safe (GRAS), highly robust in extreme environmental conditions, genetically well accessible, and grows rapidly. Compared to *E. coli*, GC-content of the genome and codon usage are more related to myxobacterial genes, and *P. putida* has already been successfully used for heterologous expression of complex natural product pathways from myxobacteria [26-29]. These studies also included metabolic engineering approaches to optimize the precursor supply for target compound biosynthesis [30].

In the present work, heterologous expression of the *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002) in *E. coli* and *P. putida* was established. Using the same initial expression constructs, *P. putida* turned out to be the host strain with the higher production rate of recombinant *n*-6 DPA and DHA *per se*. Consequently, efforts to further improve production yields were focussed on *P. putida*. Modification of the expression constructs to increase the expression rate of the *pfa* genes combined with genetic and metabolic engineering of the strain then led to a significant increase in heterologous LC-PUFA production in *P. putida*.

EXPERIMENTAL PROCEDURES

Culture conditions

Escherichia coli DH10B [31] or SCS110 (Stratagene) were used for cloning experiments. *E. coli* HB101/pRK2013 [32] was used as helper strain for conjugation experiments. *E. coli* BL21(DE3) [33] was used for heterologous expression experiments. *E. coli* HS996/pSC101-BAD-gbaA (tet^R) [34] and GB05-red [35] were used for modification of a plasmid using Red/ET recombination. The cells were grown in LB-medium or on LB-agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) at 30-37 °C (and 200 rpm) overnight. Antibiotics were used at the following concentrations: 100 µg/ml ampicillin, 50 µg/ml kanamycin, 34 µg/ml chloramphenicol, 20 µg/ml gentamicin, and 6-12.5 µg/ml tetracycline.

Pseudomonas putida KT2440 [36] was cultivated in liquid LB-medium at 30 °C and 200 rpm overnight. Conjugation was carried out onto PMM-agar plates containing 0.61% K_2HPO_4, 0.5% KH_2PO_4, 0.1% $(NH_4)_2SO_4$, 0.66% disodium succinate, and 1.5% agar. The medium was adjusted to pH 7.0 and supplemented with 0.12% $MgSO_4$ after autoclaving. Antibiotics were used at the following concentrations: 60 µg/ml kanamycin, 30 µg/ml tetracycline, and 20 µg/ml gentamicin.

General molecular biology techniques

Routine handling of nucleic acids, such as isolation of genomic and plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, were performed according to standard protocols [37]. All the enzymes were purchased from Thermo Scientific or New England Biolabs.

Cloning of expression constructs containing the *pfa* gene cluster from *A. fasciculatus* (SBSr002) for heterologous LC-PUFA production in *E. coli* and *P. putida*

In order to enable the insertion of heterologous promoters upstream of the *pfa* gene cluster, which was previously subcloned from chromosomal DNA of SBSr002, the expression construct pPfaAf2 [25] was modified via Red/ET recombineering. Thereby, the GTG start codon of gene *pfa1* was exchanged for the more frequent ATG start codon. A chloramphenicol resistance gene flanked by *Swa*I and *Nde*I restriction sites as well as the appropriate homology arms (1.1 kb) were amplified from plasmid pACYC184 (New England

Biolabs) using primers HA+*Swa*I+chloramphenicol_fwd and HA+*Nde*I+chloramphenicol_rev (Table 1). PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. The linear plus circular homologous Red/ET recombineering was performed as described previously [25] using 1.5 µl of the PCR product, resulting in plasmid pPfaAf6 (Table 2).

The heterologous inducible promoter systems *xylS-Pm*, *xylS-Pm**, and *araC*-P_{BAD} were introduced upstream of the *pfa* gene cluster into plasmid pPfaAf6 via *Swa*I and *Nde*I restriction sites, generating plasmids pPfaAf7.1, pPfaAf7.1*, and pPfaAf7.4 (Table 2). The 1.8 kb sequence of the *xylS-Pm* or the *xylS-Pm** promoter system was amplified from plasmid pJB861 [38] using primers *Pm_Swa*I_fwd and *Pm_Nde*I_rev or *Pm_Nde*I_rev_2 (Table 1). The 1.3 kb sequence of the *araC*-P_{BAD} promoter system was amplified from plasmid pBMTBX-2 [39] using primers P_{BAD}_*Swa*I_fwd and P_{BAD}_*Nde*I_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 56-60 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. Plasmid pPfaAf7.1** (Table 2) was created by introducing a *xylS-Pm*** promoter fragment into plasmid pPfaAf7.1 via *Mun*I and *Nde*I restriction sites. The 1.3 kb sequence of the *xylS-Pm*** promoter fragment was amplified from plasmid pPfaAf7.1* using primers ML2-2+LII-11_*Mun*I_fwd and ML2-2+LII-11_*Nde*I_rev_2 (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C.

The backbone of plasmids pPfaAf7.1, pPfaAf7.1*, pPfaAf7.1**, and pPfaAf7.4 was exchanged for the backbone of the *E. coli*/*Pseudomonas* shuttle vector pJB861-*Swa*I-cm^R-*Pac*I (Table 2) via *Swa*I and *Pac*I restriction sites, yielding plasmids pJBPfaAf1, pJBPfaAf1*,

pJBPfaAf1**, and pJBPfaAf4 (Table 2). For the construction of plasmid pJB861-*Swa*I-cm^R-*Pac*I by modification of plasmid pJB861, a chloramphenicol resistance gene with *Swa*I and *Pac*I restriction sites as well as the appropriate homology arms (1.1 kb) was amplified from plasmid pACYC184 using primers HA+*Swa*I+cm^R for pJB861_fwd and HA+*Pac*I+cm^R for pJB861_rev (Table 1). PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95°C; 30 cycles consisting of denaturation for 30 s at 95ºC, annealing for 30 s at 62ºC, and extension for 10 s/kb at 72ºC; and a final extension for 10 min at 72ºC. The linear plus circular homologous Red/ET recombineering was performed as described previously [25] using 1.5 µl of the PCR product.

Additionally, the backbone of plasmid pPfaAf7.1** was exchanged for the backbone of plasmid pJB861-*Swa*I-cm^R-*Pac*I-*cop271C* (Table 2) via *Swa*I and *Pac*I restriction sites, resulting in plasmid pJB*PfaAf1** (Table 2). For the construction of a version of plasmid pJB861-*Swa*I-cm^R-*Pac*I with a mutated *trfA* gene, a fragment of *trfA* was amplified from plasmid pJB861-*Swa*I-cm^R-*Pac*I by overlap extension PCR using primers *cop271C_Nde*I_fwd_2 / *cop271C*_G to A_rev and *cop271C*_G to A_fwd / *cop271C_Ehe*I_rev_2 (Table 1). Thereby, the R271C copy-up point mutation, designated as *cop271C*, was introduced. PCR amplification of the 0.1 kb and 1 kb fragments to be spliced was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The *reactions* contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. For the subsequent overlap extension PCR using primers *cop271C_Nde*I_fwd_2 and *cop271C_Ehe*I_rev_2, the two amplified fragments were used as templates. PCR of the 1.1 kb DNA fragment was performed as described for the amplification of the two fragments to be spliced. The wild type version of *trfA* in plasmid pJB861-*Swa*I-cm^R-*Pac*I was exchanged for the mutated version of the gene via *Nde*I and *Ehe*I restriction sites, yielding plasmid pJB861-*Swa*I-cm^R-*Pac*I-*cop271C*.

Cloning of expression constructs containing the gene encoding the 4'-phosphopantetheinyl transferase (PPTase) AfPpt from *A. fasciculatus* (SBSr002) for heterologous LC-PUFA production in *E. coli* and *P. putida*

The gene of the 4'-phosphopantetheinyl transferase (PPTase) AfPpt (0.8 kb) was amplified from genomic DNA of *A. fasciculatus* (SBSr002) using primers *Afppt*_fwd_*Nde*I and *Afppt*_rev_*Pac*I (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 ºC, annealing for 20 s at 72 ºC, and extension for 10 s/kb at 72 ºC; and a final extension for 10 min at 72 ºC. Plasmids pPptAfA and pPptAfD (Table 2) were constructed by exchange of the *pfa* gene cluster from *A. fasciculatus* (SBSr002) located on plasmid pPfaAf7.1 or pPfaAf7.4, respectively, for *Afppt* via *Nde*I and *Pac*I restriction sites. For the transfer and integration of the constructs into *P. putida*, an origin of transfer, a gene encoding the TetR transcriptional regulator, the tetracycline resistance gene *tetA*, and gene *trpE* encoding the anthranilate synthase component 1 from *P. putida* (4.2 kb) were amplified from cosmid CMch37 [26] using primers PseuInt_*Pac*I_fwd and PseuInt_*Pac*I_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 ºC, annealing for 20 s at 72 ºC, and extension for 10 s/kb at 72 ºC; and a final extension for 10 min at 72 ºC. The PCR amplicon was inserted into plasmids pPptAfA and pPptAfD via *Pac*I restriction sites, generating plasmids pPptAf1 and pPptAf4, respectively (Table 2).

Cloning of a mutated and truncated version of the 2,4-dienoyl-CoA reductase gene *fadH* for inactivation in *P. putida*

A mutated and truncated version of the 2,4-dienoyl-CoA reductase gene *fadH* from *P. putida* KT2440 (1.2 kb) was amplified from genomic DNA using primers *fadH*_KO_*Swa*I_fwd and *fadH*_KO_*Sac*I_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for

15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. Plasmid p*fadH*_KO-tet^R (Table 2) was constructed by exchange of *xylS-Pm-Afppt-trpE* located on plasmid pPfaPpt1 for the *fadH*_KO fragment via *Swa*I and *Sac*I restriction sites. For replacement of the antibiotic resistance marker, the gentamicin resistance gene (0.8 kb) was amplified from plasmid pMycoMarHyg [40] using primers gm^R_*Sal*I_fwd and gm^R_*Sac*I_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The tetracycline resistance gene of plasmid p*fadH*_KO-tet^R was exchanged for the gentamicin resistance gene via *Sal*I and *Sac*I restriction sites, yielding plasmid p*fadH*_KO-gm^R (Table 2).

Cloning of constructs for the overexpression of endogenous glucose-6-phosphate dehydrogenase Zwf-1 and acetyl-CoA carboxylase ACC from *P. putida*

The heterologous promoter P_{tac} (0.1 kb) was amplified from plasmid pGEX-6P-1 (GE Healthcare) using primers P_{tac}_REs_fwd and P_{tac}_*Nde*I_rev_2 (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98°C; 30 cycles consisting of denaturation for 15 s at 98 ºC, annealing for 20 s at 63 ºC, and extension for 10 s/kb at 72 ºC; and a final extension for 10 min at 72 ºC. The PCR amplicon was ligated into plasmid pJET1.2 (Thermo Scientific), generating plasmid pJET1.2-P_{tac} (Table 2). The genes *zwf-1* (1.5 kb), *accA* (1 kb), *accD* (0.9 kb), and *accB-accC-1* (1.8 kb) from *P. putida* KT2440 were amplified from genomic DNA using primers *zwf-1_Nde*I_fwd / *tonB*+*zwf-1*_REs_rev, *accA_Nde*I_fwd / *accA_Kpn*I+*Xba*I_rev, *accD_Nde*I_fwd / *tonB*+*accD*_REs_rev, and *accB_Nde*I_fwd / *accC-1_Xba*I_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 65-66 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The genes were

introduced downstream of P_{tac} via *Nde*I and *Xba*I restriction sites into plasmid pJET1.2-P_{tac}, generating plasmids pJET1.2-P_{tac}-*zwf-1*-T_{tonB}, pJET1.2-P_{tac}-*accA*, pJET1.2-P_{tac}-*accD*-T_{tonB}, and pJET1.2-P_{tac}-*accB-accC-1* (Table 2).

Plasmid pME3 (Table 2) was constructed by insertion of P_{tac}-*zwf-1*-T_{tonB} hydrolysed with *Swa*I and *Aan*I into plasmid pPptAf1 linearized with *Aan*I. P_{tac}-*accA* was inserted into plasmid pUC18 (Thermo Scientific) via *Bam*HI and *Kpn*I restriction sites, yielding plasmid pUC18-P_{tac}-*accA* (Table 2). Plasmid pUC18-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} (Table 2) was generated by insertion of P_{tac}-*accD*-T_{tonB} into plasmid pUC18-P_{tac}-*accA* via *Kpn*I and *Eco*RI restriction sites. In order to exchange of the *Bam*HI restriction site for the *Bgl*II restriction site in plasmid pUC18-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB}, a 0.6 kb fragment of P_{tac}-*accA* was amplified from plasmid pUC18-P_{tac}-*accA* using primers ME_*Bam*HI to *Bgl*II_fwd and ME_*Bam*HI to *Bgl*II_rev (Table 1). Thereby, the *Bam*HI to *Bgl*II point mutations were introduced. PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 ºC, annealing for 20 s at 62 ºC, and extension for 10 s/kb at 72 ºC; and a final extension for 10 min at 72 ºC. The PCR amplicon was inserted into plasmid pUC18-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} via *Xba*I and *Eco*88I restriction sites. Plasmid pUC18-P_{tac}-*accB-accC-1*-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} (Table 2) was then constructed by insertion of P_{tac}-*accB-accC-1* into plasmid pUC18-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} via *Sda*I and *Bgl*II restriction sites. The fragment P_{tac}-*accB-accC-1*-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} was then excised with *Swa*I and *Pac*I, blunted, and inserted into plasmid pPptAf1 linearized with *Aan*I, resulting in plasmid pME2 (Table 2).

Table 1. Primers used for cloning of expression or knockout plasmids in this study.
The homology arms are shown in italics and the introduced restriction sites are underlined.

Primer	Sequence (5'→3')
HA+*Swa*I+chloramphenicol_fwd	*CCCCAGAGCGCGCCATATTTCGCCGCGCAGCGCTGGTCGTCGAAATATTG*ATTTAAATTTAACGACCCTGCCCTGAACCGACG
HA+*Nde*I+chloramphenicol_rev	*GGCGAGCCCCCTTGAGCGTTCCATCTACCGATTGCAGACA*TATGAGTAAGTTGGCAGCATCACCCGACGC
*Pm_Swa*I_fwd	CCAGTGATTTAAATCCTTCTTCGGCTACGTTCG
*Pm_Nde*I_rev	GCAATTCATATGTTCATGACTCC
*Pm_Nde*I_rev_2	CATATGTTCATTACTCCATTATTGTGTATTCT
P_{BAD}_*Swa*I_fwd	CTGCTATTTAAATGGAGGTCAGG
P_{BAD}_*Nde*I_rev	TTACCCATATGGGTATATCTCCTTC
ML2-2+LII-11_*Mun*I_fwd	CAATTGTTGTCACTGCATGCCCGATCAAGGACC
ML2-2+LII-11_*Nde*I_rev_2	CATATGTTCATTACTCCATTATTGTGTATTCTGTGCTAAACTATAAGGTGTAGGC
HA+*Swa*I+cm^R for pJB861_fwd	*AGATCCACATCCTTGAAGGCCGCAGCGACGAGCAGAAGGAAACCCTCAT*TTAAATAGTAAGTTGGCAGCATCACCCGACGC
HA+*Pac*I+cm^R for pJB861_rev	*ATGGTGACGTCACCGGTTCTAGATACCTAGGTGAGCTCTGGTACCGGCG*GTTAATTAACGACCCTGCCCTGAACCGACG
*cop271C_Nde*I_fwd_2	TCCCATATGAATTTCGTGTAGTGGTCGCCAGC
*cop271C*_G to A_rev	GTCGCTGCTGCACTGCTTCCGCGTCC
*cop271C*_G to A_fwd	GGACGCGGAAGCAGTGCAGCAGCGAC
*cop271C_Ehe*I_rev_2	CTGGGGCGCCCTCTGGTAAGGTTGG
*Afppt*_fwd_*Nde*I	AGGCCCATATGGCGCTCCTCGACCTCCC
*Afppt*_rev_*Pac*I	CCATCTTAATTAATTACACGGCCAGGGTCGCGGG
PseuInt_*Pac*I_fwd	CGTTACTTAATTAAGTGGATCCGCTTACGGCCAGCC
PseuInt_*Pac*I_rev	GAGGTGTTAATTAACGGGGCTGGTGCTGGATCC
*fadH*_KO_*Swa*I_fwd	ATTTAAATGAGGTGCCCGCATTCCGTGACACGC
*fadH*_KO_*Sac*I_rev	GAGCTCGGTCTCGTGGCAGGCC
gm^R_*Sal*I_fwd	CACCGTCGACACGGATGAAGGCACGAACCC
gm^R_*Sac*I_rev	GCCGAGCTCGGCTTGAACGAATTGTTAGGTGGC
P_{tac}_REs_fwd	CCTGCAGGATTTAAATGGTACCGGATCCAAATGAGCTGTTGACAATTAATCATCGG
P_{tac}_*Nde*I_rev_2	CATATGTACTGTTTCCTGTGTGAAATTGTTATCC
*zwf-1_Nde*I_fwd	ACATCATATGGCCGCAATCAGTGTCG
tonB+*zwf-1*_REs_rev	TCAATCTAGAAGCTTATAAGTCAAAAGCCTCCGGTCGGAGGCTTTTGACTTCAGATATCCCCATACCACGC
*accA_Nde*I_fwd	TGACCATATGAACCCGAATTTCCTCGACTTCG
*accA_Kpn*I+*Xba*I_rev	CAGAATCTAGAGGTACCTTAGAGGCCGTAGCTCATCAGG
*accD_Nde*I_fwd	AAAGCATATGAGCAACTGGTTAGTCGAC
tonB+*accD*_REs_rev	TTTCATCTAGAGAATTCTTAATTAAAGTCAAAAGCCTCCGACCGGAGGCTTTTGACTTCACGCGACGGCAGCC
*accB_Nde*I_fwd	GATTACATATGGATATCCGTAAAGTCAAGAAACTG
*accC-1_Xba*I_rev	CACGATCTAGATCACTCCTGGTTGGCC
ME_*Bam*HI to *Bgl*II_fwd	TCTAGAAGATCTAAATGAGCTGTTGAC
ME_*Bam*HI to *Bgl*II_rev	CGTCGATGCCCGGGTAGG

Table 2. Plasmids constructed in this study.
oriV = origin of replication, *trfA* = gene encoding replication initiation protein in RK2 replicons, *oriT* = origin of transfer, *bla* = ampicillin resistance gene, *cat* = chloramphenicol resistance gene, *aph(3')-Ia* = kanamycin resistance gene, *tetA* = tetracycline resistance gene, *aac(3)-Ia* = gentamicin resistance gene, *tetR* = gene encoding the transcriptional regulator, *trpE* = gene encoding the anthranilate synthase component 1 from *P. putida* KT2440, P_{tac} = *tac* promoter, T_{tonB} = *tonB* transcription terminator.

Plasmid	Characteristics
pPfaAf6	Derivative of pPfaAf2 [25] in which *cat* was inserted upstream of the *pfa* gene cluster of *Aetherobacter fasciculatus* (SBSr002) by Red/ET recombineering. p15A *oriV*, *bla*, *cat*
pPfaAf7.1	Derivative of pPfaAf6 in which *cat* was replaced by the *xylS-Pm* promoter system via *Swa*I and *Nde*I
pPfaAf7.1*	Derivative of pPfaAf6 in which *cat* was replaced by the *xylS-Pm** promoter system (containing the mutated DNA-region LII-11 [41]) via *Swa*I and *Nde*I
pPfaAf7.1**	Derivative of pPfaAf7.1 in which the *xylS-Pm* promoter system was replaced by the *xylS-Pm*** promoter system (containing the mutated DNA-regions LII-11 [41] plus ML2-2 [42]) via *Mun*I and *Nde*I
pPfaAf7.4	Derivative of pPfaAf6 in which *cat* was replaced by the *araC*-P_{BAD} promoter system via *Swa*I and *Nde*I
pJB861-*Swa*I-cm^R-*Pac*I	Derivative of pJB861 [38] in which *xylS-Pm* was exchanged for *cat* flanked by *Swa*I and *Pac*I restriction sites by Red/ET recombineering. RK2 *oriT*, RK2 *oriV*, *trfA*, *aph(3')-Ia*, *cat*
pJB861-*Swa*I-cm^R-*Pac*I-*cop271C*	Derivative of pJB861-*Swa*I-cm^R-*Pac*I in which *trfA* was replaced with *trfA** (containing the R→C copy-up point mutation *cop271C*) via *Nde*I and *Ehe*I
pJBPfaAf1	Derivative of pPfaAf7.1 in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I via *Swa*I and *Pac*I
pJBPfaAf1*	Derivative of pPfaAf7.1* in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I via *Swa*I and *Pac*I
pJBPfaAf1**	Derivative of pPfaAf7.1** in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I via *Swa*I and *Pac*I
pJBPfaAf4	Derivative of pPfaAf7.4 in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I via *Swa*I and *Pac*I
pJB*PfaAf1**	Derivative of pPfaAf7.1** in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I-*cop271C* via *Swa*I and *Pac*I
pPptAfA	Derivative of pPfaAf7.1 in which the *pfa* gene cluster was replaced by the gene encoding the PPTase AfPpt from *A. fasciculatus* (SBSr002) via *Nde*I and *Pac*I
pPptAfD	Derivative of pPfaAf7.4 in which the *pfa* gene cluster was replaced by the gene encoding the PPTase AfPpt from *A. fasciculatus* (SBSr002) via *Nde*I and *Pac*I
pPptAf1	Derivative of pPptAfA in which *oriT-tetR-tetA-trpE* was inserted via *Pac*I
pPptAf4	Derivative of pPptAfD in which *oriT-tetR-tetA-trpE* was inserted via *Pac*I
p*fadH*_KO-tet^R	Derivative of pPptAf1 in which *xylS-Pm-Afppt-trpE* was replaced by *fadH*_KO (mutated and truncated version of the 2,4-dienoyl-CoA reductase gene *fadH* from *P. putida* KT2440) via *Swa*I and *Sac*I

Continuation of Table 2.

Plasmid	Characteristics
p*fadH*_KO-*gm*^{R}	Derivative of p*fadH*_KO-*tet*^{R} in which *tetR-tetA* was exchanged for *aac(3)-Ia* via *Sal*I and *Sac*I
pJET1.2-P_{tac}	Derivative of pJET1.2 (Thermo Scientific) in which P_{tac} was inserted
pJET1.2-P_{tac}-*zwf-1*-T_{tonB}	Derivative of pJET1.2-P_{tac} in which *zwf-1* (encoding the glucose-6-phosphate dehydrogenase from *P. putida* KT2440) and T_{tonB} were inserted downstream of P_{tac} via *Nde*I and *Xba*I
pJET1.2-P_{tac}-*accA*	Derivative of pJET1.2-P_{tac} in which *accD* (encoding the carboxyl transferase subunit alpha of the acetyl-CoA carboxylase from *P. putida* KT2440) was inserted downstream of P_{tac} via *Nde*I and *Xba*I
pJET1.2-P_{tac}-*accD*-T_{tonB}	Derivative of pJET1.2-P_{tac} in which *accD* (encoding the carboxyl transferase subunit beta of the acetyl-CoA carboxylase from *P. putida* KT2440) and T_{tonB} were inserted downstream of P_{tac} via *Nde*I and *Xba*I
pJET1.2-P_{tac}-*accB-accC-1*	Derivative of pJET1.2-P_{tac} in which *accB-accC-1* (encoding the biotin carboxyl carrier protein subunit and the biotin carboxylase subunit of the acetyl-CoA carboxylase from *P. putida* KT2440) was inserted downstream of P_{tac} via *Nde*I and *Xba*I
pUC18-P_{tac}-*accA*	Derivative of pUC18 (Thermo Scientific) in which P_{tac}-*accA* from pJET1.2-P_{tac}-*accA* was inserted via *Bam*HI and *Kpn*I
pUC18-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB}	Derivative of pUC18-P_{tac}-*accA* in which P_{tac}-*accD*-T_{tonB} from pJET1.2-P_{tac}-*accD*-T_{tonB} was inserted via *Kpn*I and *Eco*RI
pUC18-P_{tac}-*accB-accC-1*-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB}	Derivative of pUC18-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} in which the *Bam*HI restriction site was replaced by a *Bgl*II restriction site and P_{tac}-*accB-accC-1* from pJET1.2-P_{tac}-*accB-accC-1* was inserted via *Sda*I and *Bgl*II
pME3	Derivative of pPptAf1 in which P_{tac}-*zwf-1*-T_{tonB} (excised from pJET1.2-P_{tac}-*zwf-1*-T_{tonB} via *Swa*I and *Aan*I) was inserted via *Aan*I
pME2	Derivative of pPptAf1 in which P_{tac}-*accB-accC-1*-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} (excised from pUC18-P_{tac}-*accB-accC-1*-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} via *Swa*I and *Pac*I plus blunted) was inserted via *Aan*I

Transformation of *P. putida* by triparental conjugation

Expression plasmids and knockout plasmids were transferred into *P. putida* KT2440 by triparental conjugation. 1 ml of an overnight culture of the acceptor strain *P. putida* KT2440, the donor strain *E. coli* DH10B containing the final expression construct, and the helper strain *E. coli* HB101/pRK2013 were harvested at 13,000 rpm for 1 min. After discarding the supernatant, the cells were resuspended in 1 ml LB-medium in each case. The cells were spun down at 13,000 rpm for 1 min, and the supernatant was discarded. This wash step was repeated once. The cells were resuspended in 300 µl LB-medium in each case. The triparental mating was performed by combining 50 µl of each 300 µl suspension on a small area of a LB-agar plate [43]. After incubation at 37 °C for 4 h, the plate was transferred to 30 °C and incubated

overnight. The cells were scraped from the plate and resuspended in 1 ml sterile ddH_2O. 10-100 µl of this suspension was then plated out on PMM-agar containing 60 µg/ml kanamycin, and if necessary 30 µg/ml tetracycline, and 20 µg/ml gentamicin. The *E. coli* strains were additionally counterselected by 150 µg/ml ampicillin or 100 µg/ml hygromycin B, which *P. putida* is resistant to. The plates were incubated at 30 °C for 1-2 days. Transconjugants were verified by plasmid isolation and restriction analysis for replicative plasmids. The constructed strains are listed in Table 4.

In order to verify the integration of the expression construct into the genome of *P. putida* via *trpE,* DNA from selected clones was isolated and used for PCR analyses. The primers used are listed in Table 3. PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 56 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C.

In order verify the knockout of *fadH* in the genome of *P. putida* via single crossover, DNA from selected clones was isolated and used for PCR analyses. The primers used are listed in Table 3. PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 56 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C.

Table 3. Primers used for the verification of the integration of expression constructs into the genome of *P. putida* KT2440 via *trpE* or for the verification of the knockout of *fadH* in the genome of *P. putida* via single crossover.

Primer	Sequence (5'→3')
*trpE*_genome_fwd	CCGTCCGTTTGCCTGACC
*trpE*_genome_rev	CTCCTACGACGGCCTCG
*trpE*_integration_fwd	GCCGATCGCTGATCGTGG
*trpE*_integration_rev	GGCTGGTGCTGGATCCC
*fadH*_genome_fwd	AGCCTTCCCGACAGTCGC
*fadH*_genome_rev	GGTGGCTCGACACAGGC
*fadH*_integration_fwd	CCGTTGCACCAGTGCCC
*fadH*_integration_rev	GACCCAAGTACCGCCACC

Heterologous LC-PUFA production in *E. coli* and *P. putida*

E. coli BL21(DE3) was cotransformed with the expression constructs pJBPfaAf1 plus pPptAfA, pJBPfaAf1* plus pPptAfA, or pJBPfaAf4 plus pPptAfD. The constructed strains are listed in Table 4. For heterologous LC-PUFA production, cultivation of the strains was carried out in 50 ml LB-medium containing 50 µg/ml kanamycin and 100 µg/ml ampicillin. The medium was inoculated with an overnight culture (1:100) and incubated at 37 °C. Expression of the Pfa proteins and AfPpt was induced at OD_{600} of 0.8 by addition of *m*-toluic acid to a final concentration of 2 mM in the case of pJBPfaAf1 plus pPptAfA and pJBPfaAf1* plus pPptAfA or of L-arabinose to a final concentration of 2% in the case of pJBPfaAf4 plus pPptAfD. After induction, the cells were cultivated at 16 °C and 200 rpm for 24 h and then harvested at 8,000 rpm for 5 min.

For heterologous LC-PUFA production in *P. putida*, cultivation was carried out in 50 ml LB-medium containing 60 µg/ml kanamycin, and if necessary 30 µg/ml tetracycline, and 20 µg/ml gentamicin. The medium was inoculated with an overnight culture (1:100) and incubated at 30 °C for 4 h. Expression of the Pfa proteins or AfPpt was induced by addition of either *m*-toluic acid to a final concentration of 2 mM or of L-arabinose to a final concentration of 2%, depending on the heterologous promoter system used. After induction, the cells were cultivated at 16 °C and 200 rpm for 24 h and then harvested at 8,000 rpm for 5 min.

Table 4. Expression strains constructed in this study.

Strain	Characteristics
E. coli BL21(DE3)/ pJBPfaAf1/pPptAfA	*Escherichia coli* BL21(DE3) with replicative plasmids pJBPfaAf1 and pPptAfA
E. coli BL21(DE3)/ pJBPfaAf1*/pPptAfA	*E. coli* BL21(DE3) with replicative plasmids pJBPfaAf1* and pPptAfA
E. coli BL21(DE3)/ pJBPfaAf4/pPptAfD	*E. coli* BL21(DE3) with replicative plasmids pJBPfaAf4 and pPptAfD
P. putida KT2440+ pJBPfaAf1	*Pseudomonas putida* KT2440 with replicative plasmid pJBPfaAf1
P. putida KT2440::pPptAf1	*P. putida* KT2440 with pPptAf1 site-specifically integrated into *trpE* within the genome
P. putida KT2440::pPptAf1 + pJBPfaAf1	*P. putida* KT2440::pPptAf1 with replicative plasmid pJBPfaAf1
P. putida KT2440::pPptAf1 + pJBPfaAf1*	*P. putida* KT2440::pPptAf1 with replicative plasmid pJBPfaAf1*
P. putida KT2440::pPptAf1 + pJBPfaAf1**	*P. putida* KT2440::pPptAf1 with replicative plasmid pJBPfaAf1**
P. putida KT2440::pPptAf1 + pJB*PfaAf1**	*P. putida* KT2440::pPptAf1 with replicative plasmid pJB*PfaAf1**
P. putida KT2440::pfadH_KO	*P. putida* KT2440 with knockout of *fadH* within the genome via single crossover
P. putida KT2440::pfadH_KO + pJB*PfaAf1**	*P. putida* KT2440::pfadH_KO with replicative plasmid pJB*PfaAf1**
P. putida KT2440::pfadH_KO/pPptAf1 + pJB*PfaAf1**	*P. putida* KT2440::pfadH_KO + pJB*PfaAf1** with pPptAf1 site-specifically integrated into *trpE* within the genome
P. putida KT2440::pfadH_KO/pME3 + pJB*PfaAf1**	*P. putida* KT2440::pfadH_KO + pJB*PfaAf1** with pME3 site-specifically integrated into *trpE* within the genome
P. putida KT2440::pfadH_KO/pME2 + pJB*PfaAf1**	*P. putida* KT2440::pfadH_KO + pJB*PfaAf1** with pME2 site-specifically integrated into *trpE* within the genome
P. putida KT2440::pfadH_KO/pME2	*P. putida* KT2440::pfadH_KO/pME2 + pJB*PfaAf1** cured of plasmid pJB*PfaAf1**

Extraction of cellular fatty acids

The cellular fatty acids were extracted using the FAME method [44]. For this purpose, 50 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. The cell pellet was transferred to a glass vial and dried in a vacuum concentrator. Subsequently, the cell dry weight was determined. 5 µl (50 µg) of *n*-3 DPA (Sigma-Aldrich) and 500 µl of a mixture of methanol, toluene, and sulphuric acid (50:50:2, v/v/v) were added. The vial was capped with a teflon-lined screw cap and incubated at 80 °C for 24-48 h. After the mixture was cooled to room temperature, 400 µl of an aqueous solution consisting of 0.5 M NH_4HCO_3 and 2 M KCl

were added, and the sample was vortexed for 30 s. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. 100 µl of the upper phase were used for GC-MS analysis.

Analysis of fatty acid methyl esters by GC-MS

GC-MS was carried out on an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a 7683B split/splitless injector with autosampler (Agilent Technologies) and coupled to a 5973 electron impact mass selective detector (Agilent Technologies). Helium was used as the carrier gas at a flow rate of 1 ml/min. 1 µl of the sample was injected in split mode (split ratio, 10:1). The analytical column was a (5% phenyl)-methylpolysiloxane capillary column (Agilent J&W DB-5ht; 30 m x 0.25 mm i.d. x 0.1 µm film thickness, maximum temperature 400 °C; Agilent Technologies). The column temperature was kept at 130 °C for 2.5 min, increased to 240 °C at a rate of 5 °C/min, then ramped to 300 °C at 30 °C/min, and held at 300 °C for 5 min. Other temperatures were as follows: inlet, 275 °C; GC-MS transfer line, 280 °C; ion source, 230 °C; and quadrupole, 150 °C. The mass selective detector was operated in scan mode, scanning the mass range from *m*/*z* 40 to 700. Scan control, data acquisition, and processing were performed by MSD ChemStation and AMDIS software, version 2.68, based on the fragmentation patterns and retention times, in comparison with Supelco 37 Component FAME Mix and LC-PUFAs (all Sigma-Aldrich) as reference standards, and NIST 08 library. Absolute amounts of PUFAs were quantified by integration of the peaks using MSD ChemStation and by subsequent calculation in relation to the integral of *n*-3 DPA and to cell dry weight.

RESULTS AND DISCUSSION

Recombinant LC-PUFA production using the DPA (22:5, *n*-6 or *n*-3)/DHA (22:6, *n*-3)-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) has already been accomplished by cloning, transfer, and heterologous expression of the gene cluster in the myxobacterial model strain *M. xanthus* [25]. After this successful first proof of principle, approaches towards alternative heterologous expression platforms were initiated to move the pathway into hosts which show better growth characteristics and for which a considerable set of genetic tools is available. Consequently, two additional, more commonly used heterologous hosts, *E. coli* as well as the

GRAS strain *P. putida* KT2440, were investigated as expression strains for the myxobacterial DPA/DHA-type PUFA synthase.

The initial strategy to achieve heterologous expression of the DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) in *Pseudomonas* was similar to a previously described approach applied for heterologous expression of a myxobacterial lipopeptide pathway [26]. It included the chromosomal integration of the expression construct into the *trpE* locus of the genome, encoding the anthranilate synthase component 1, via homologous recombination and the expression of the pathway genes under control of the heterologous *xylS-Pm* promoter system (details not shown). Unfortunately, this procedure did not result in a detectable heterologous LC-PUFA production in *P. putida*. On the supposition that the *pfa* gene dosage might be a bottleneck, an expression construct based on a medium-copy number replicative plasmid for *P. putida*, analogous to the construct used for successful heterologous expression of a chalcone synthase from *S. cellulosum* [27], was generated. The cloning procedure started from plasmid pPfaAf2, constructed previously [25], which contains the DPA/DHA-type *pfa* gene cluster of *A. fasciculatus* (SBSr002) (Figure 1A). This plasmid was modified by linear plus circular homologous Red/ET recombination [45] to replace dispensable DNA regions, inclusive of the native promoter upstream of *pfa1*, by a chloramphenicol resistance gene, resulting in plasmid pPfaAf6. By applying conventional cloning techniques, the chloramphenicol resistance gene was exchanged for the heterologous *xylS-Pm* promoter system. In a final engineering step, the backbone of the obtained plasmid pPfaAf7.1 was replaced by a backbone on the basis of the broad-host-range expression vector pJB861 [38], yielding expression plasmid pJBPfaAf1. This expression construct was transferred into *P. putida* KT2440 by triparental conjugation. The transgenic clones were cultivated in 50 ml LB-medium at 30 °C for 4 h. Expression of the Pfa proteins was induced by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C for 24 h. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [44] and analysed by GC-MS. The transgenic strain *P. putida* KT2440 + pJBPfaAf1 was shown to produce 0.03 mg *n*-6 DPA/g cell dry weight (CDW) and 0.2 mg DHA/g CDW (Figure 2), demonstrating the applicability of *P. putida* as expression host for PUFA synthases. The only PPTase from *P. putida* KT2440, PP1183, exhibits broad substrate specificity [46] and was obviously able to (at least partially) activate the tandem ACP domains of the PUFA synthase by posttranslational modification.

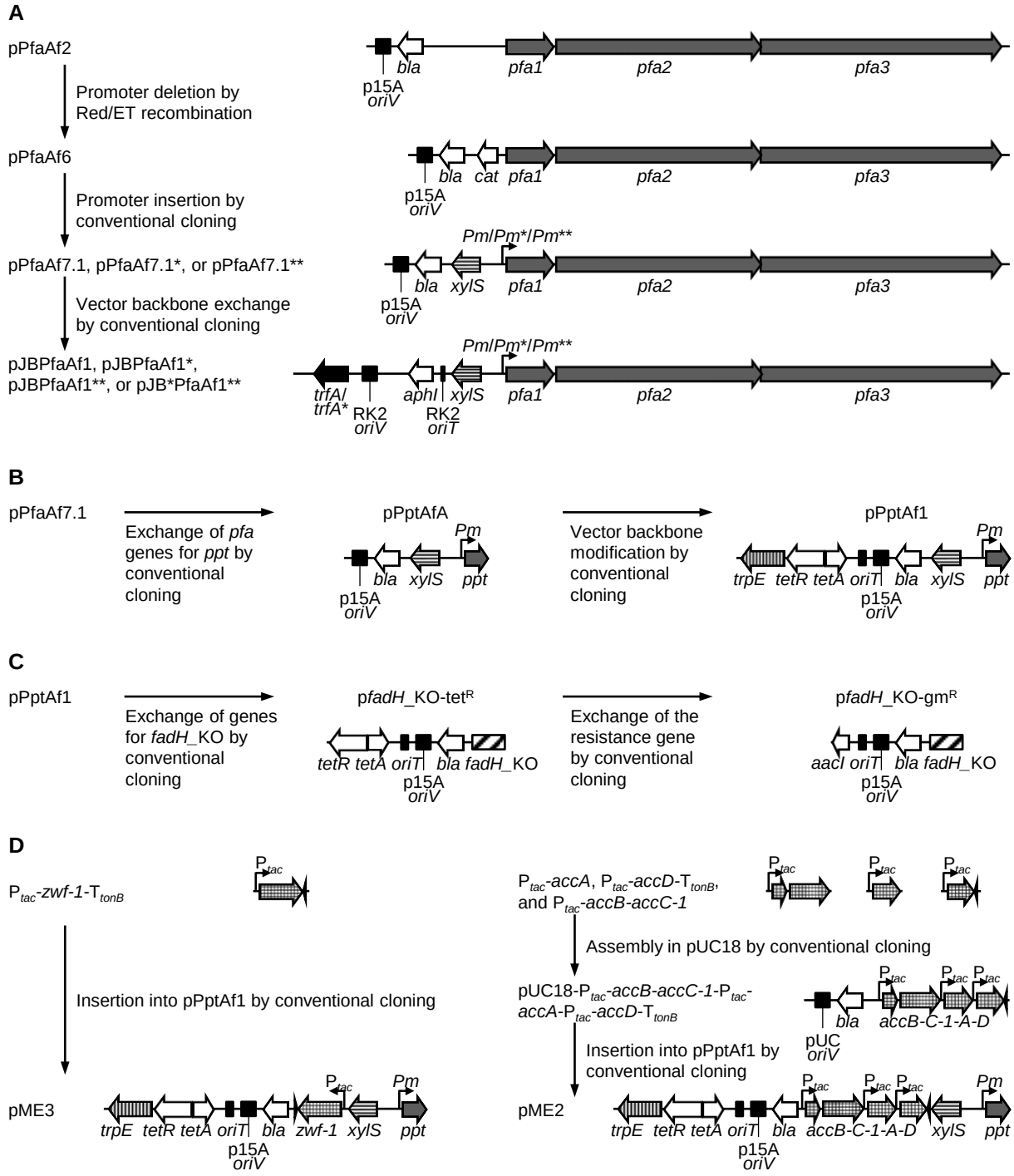


Figure 1. Cloning strategy for the construction of plasmids for recombinant LC-PUFA production in *Pseudomonas putida* KT2440 and strain improvement by genetic/metabolic engineering.
(A) Cloning of the *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002) into replicative plasmids, (B) Cloning of the gene encoding the PPTase AfPpt from *A. fasciculatus* (SBSr002), into an integrative plasmid, (C) Cloning of *fadH*-KO, a mutated and truncated version of the 2,4-dienoyl-CoA reductase gene *fadH* from *P. putida* KT2440, into an integrative plasmid, (D) Cloning of *zwf-1*, *accA*, *accB-C-1*, and *accD*, encoding the glucose-6-phosphate dehydrogenase and the acetyl-CoA carboxylase from *P. putida* KT2440, into integrative plasmids. *oriV* = origin of replication, *trfA* = gene encoding replication initiation protein in RK2 replicons (*trfA** = *cop271C* mutation included), *oriT* = origin of transfer, *bla* = ampicillin resistance gene, *cat* = chloramphenicol resistance gene, *aphI* = kanamycin resistance gene, *tetA* = tetracycline resistance gene, *aacI* = gentamicin resistance gene, *Pm* = *Pm* promoter (*Pm** = mutated 5'-UTR of version LII-11 [41] included; *Pm*** = mutated 5'-UTR of version LII-11 [41] plus the mutated core promoter region of version ML2-2 [42] included), *xylS* = gene encoding the transcriptional regulator for *Pm* promoter, *tetR* = gene encoding the tetracycline transcriptional regulator, *trpE* = gene encoding the anthranilate synthase component 1 from *P. putida* KT2440, P_{tac} = *tac* promoter.

However, as LC-PUFA production by strain *P. putida* KT2440 + pJBPfaAf1 was relatively low, it was assumed that posttranslational activation of the PUFA synthase might be not efficient and could be improved by coexpression of the native PPTase from *A. fasciculatus* (SBSr002). A corresponding gene, *Afppt*, which encodes a Sfp-type PPTase that likely catalyses phosphopantetheinylation of PUFA synthases, was identified in the myxobacterial PUFA producer strain and cloned into an integrative expression plasmid under control of the *xylS-Pm* promoter system (Figure 1B). In the first instance, the *pfa* gene cluster located on plasmid pPfaAf7.1 was replaced with *Afppt* by conventional cloning, generating plasmid pPptAfA. Afterwards, an origin of transfer (*oriT*) for conjugation purposes, a tetracycline resistance cassette for selection in *P. putida*, and *trpE* for chromosomal integration into *P. putida* via homologous recombination were inserted by conventional cloning, yielding plasmid pPptAf1 (Figure 1B). This plasmid was transferred into *P. putida* KT2440. Three out of many obtained tetracycline resistant clones were chosen, and the correct integration of pPptAf1 could be confirmed by PCR. After transformation of pJBPfaAf1 into *P. putida* KT2440::pPptAf1, heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. Transgenic strain *P. putida* KT2440::pPptAf1 + pJBPfaAf1, additionally overexpressing the Sfp-type PPTase from *A. fasciculatus* (SBSr002), produces significantly higher amounts of LC-PUFAs (0.07 mg *n*-6 DPA/g CDW and 0.3 mg DHA/g CDW) than *P. putida* KT2440 + pJBPfaAf1 (Figure 2). This result indicates that the activity of the intrinsic PPTase PP1183 from *P. putida* KT2440 is not sufficient to phosphopantetheinylate quantitatively all the acyl carrier protein (ACP) domains of the overexpressed PUFA synthases.

In order to further optimize heterologous PUFA production, different cultivation conditions were evaluated, including the use of several complex or minimal liquid media (details not shown), at which cultivation in standard LB-medium served as reference. Bacterial growth was boosted in the rich, complex media, but the yield of LC-PUFAs was lower than in LB-medium. Using minimal media, growth was significantly impaired and absolute amounts of produced LC-PUFAs were not elevated. All in all, LB-medium seems to be the best medium for heterologous LC-PUFA production in *P. putida*, providing good growth characteristics for the bacteria and considerable LC-PUFA production titres. Feeding of fatty acid synthase inhibitors was reported as promising approach to increase the concentration of LC-PUFAs produced by the anaerobic biosynthetic pathway in marine bacteria [14;15]. Unfortunately, no clear, reproducible effect could be attained by feeding of cerulenin or thiolactomycin to the transgenic *P. putida* strains. The cultivation temperatures of

16 °C and 30 °C were tested for heterologous expression of AfPpt and/or the Pfa proteins in *P. putida* strains, with the result that LC-PUFA production could only be observed at a cultivation temperature of 16 °C. Investigations on the LC-PUFA production kinetics revealed that *n*-6 DPA and DHA continually accumulated within 24 h of cultivation at 16 °C and that the yield of these LC-PUFAs per 50 ml culture volume was only slightly increased after additional 24 h of cultivation.

In another approach to further increase the LC-PUFA production rate, engineering of the expression constructs was selected as starting point. In the first place, the heterologous promoter system was chosen, as a promoter driving the expression of heterologous genes represents a determining factor for recombinant protein production. As a consequence, expression plasmids pPfaAf7.4, pJBPfaAf4, pPptAfD, and pPptAf4, all harbouring the *araC*-P_{BAD} promoter system instead of the *xylS-Pm* promoter system, were constructed analogous to plasmids pPfaAf7.1, pJBPfaAf1, pPptAfA, and pPptAf1 mentioned previously. The expression constructs pJBPfaAf4 and pPptAf4 were transferred into *P. putida* KT2440. Heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described previously, except for the use of 2% L-arabinose instead of 2 mM *m*-toluic acid as inducer. *P. putida* KT2440::pPptAf4 + pJBPfaAf4 produces *n*-6 DPA and DHA in amounts comparable to the yield obtained with *P. putida* KT2440::pPptAf1 + pJBPfaAf1 (details not shown). This indicates that the *araC*-P_{BAD} promoter system was not superior to the *xylS-Pm* promoter system in *P. putida*. However, any downstream bottlenecks which would mask an effect in either direction cannot be excluded. In case of the *xylS-Pm* promoter system, large mutant libraries exist in *E. coli*, that were obtained by random mutagenesis of either the 5'-untranslated region (UTR) [41] or the core region [42] of *Pm* promoter. It has been shown that the mutated 5'-UTR of *Pm* promoter version LII-11 is related to an increase of reporter protein activity up to about 20-fold and to an increase of reporter gene transcript amount of up to about seven-fold in *E. coli* [41]. These results encouraged introduction of the modified *xylS-Pm** promoter system comprising the mutated 5'-UTR of version LII-11 upstream of the *pfa* gene cluster. Thereby, plasmids pPfaAf7.1* and pJBPfaAf1* were constructed (Figure 1A). Expression construct pJBPfaAf1* was then transferred into *P. putida* KT2440::pPptAf1. Heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. *P. putida* KT2440::pPptAf1 + pJBPfaAf1* produces significantly higher amounts of LC-PUFAs (0.09 mg *n*-6 DPA/g CDW and 0.5 mg DHA/g CDW) than *P. putida* KT2440::pPptAf1 + pJBPfaAf1 (Figure 2), indicating a similar effect of

the mutated 5'-UTR of version LII-11 on gene expression for *E. coli* and *P. putida*. Additional exchange of the wild type *Pm* promoter upstream of *Afppt* for *Pm** promoter did not result in further production enhancement (details not shown). Nonetheless, these results motivated further engineering of *Pm** promoter, resulting in the construction of *Pm*** promoter by combining the mutated 5'-UTR of version LII-11 and the mutated core promoter region of version ML2-2. For *Pm* promoter version ML2-2, an increase of reporter protein activity up to about 13.5-fold and an increase of reporter gene transcript amount of up to about ten-fold in *E. coli* has been demonstrated [42]. Plasmids pPfaAf7.1** and pJBPfaAf1**, containing the *xylS-Pm*** promoter system upstream of the *pfa* gene cluster, were constructed (Figure 1A). Expression construct pJBPfaAf1** was then transferred into *P. putida* KT2440::pPptAf1. Heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. Compared to *P. putida* KT2440::pPptAf1 + pJBPfaAf1*, the amounts of LC-PUFAs produced by *P. putida* KT2440::pPptAf1 + pJBPfaAf1** (0.1 mg *n*-6 DPA/g CDW and 0.5 mg DHA/g CDW) are only slightly higher (Figure 2). Thus, either the core promoter region of version ML2-2 has only a minor effect on gene expression in *P. putida* or the combination with the mutated 5'-UTR of version LII-11 blocks the positive effect of this mutation.

Apart from targeting the promoter region, genetic engineering was extended to the RK2 replicon, which consists of an origin of replication (*oriV*) and gene *trfA*, whose gene product is required for initiation of replication [47]. It has been shown that the copy number of RK2 based vectors can easily be raised by introducing an R271C point mutation into *trfA* [38]. Assuming that this copy-up mutation increases the *pfa* gene dosage which correlates with LC-PUFA production, the vector backbone of expression construct pJBPfaAf1** was exchanged for an identical vector backbone except for the R271C point mutation within *trfA* (Figure 1A). The resulting expression construct pJB*PfaAf1** was then transferred into *P. putida* KT2440::pPptAf1. Heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. Compared to *P. putida* KT2440::pPptAf1 + pJBPfaAf1**, only slight improvements of LC-PUFA production can be attained with *P. putida* KT2440::pPptAf1 + pJB*PfaAf1** (0.1 mg *n*-6 DPA/g CDW and 0.6 mg DHA/g CDW) (Figure 2). A conclusive explanation for this result can hardly be found, as the copy numbers of plasmids pJBPfAf1** and pJB*PfaAf1** in *P. putida* have not been checked and as it is still unclear if the gene dose represents a bottleneck or if there are any other limitations in downstream cellular processes.

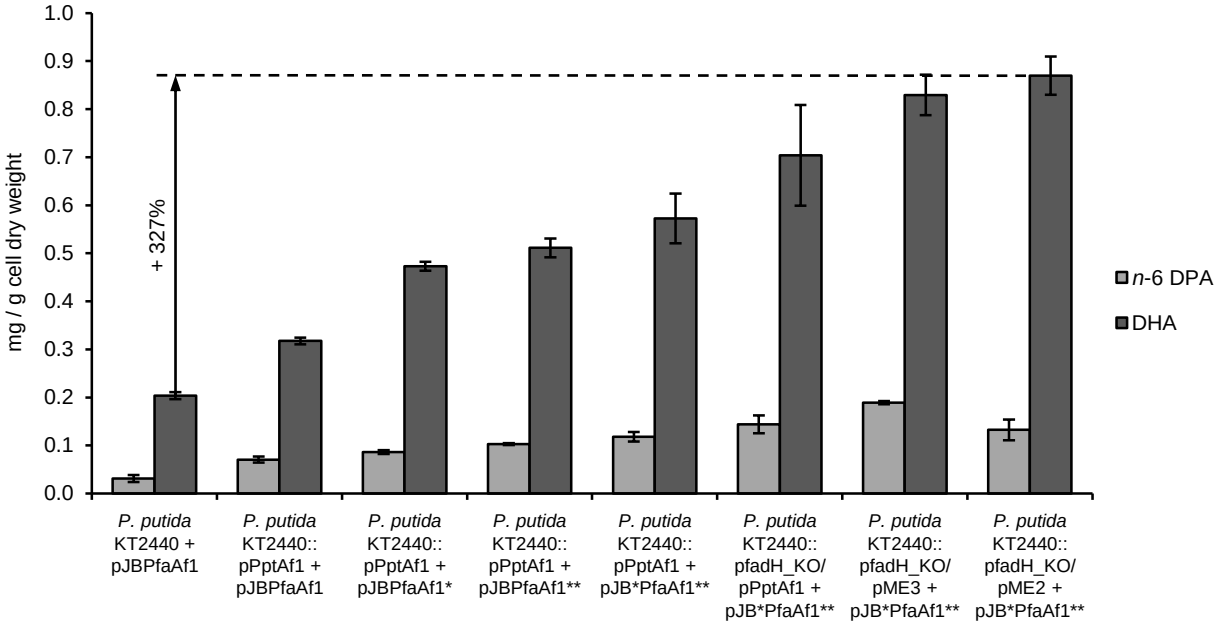


Figure 2. Recombinant production of *n*-6 DPA and DHA with the *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002) by wild type and genetically/metabolically engineered *Pseudomonas putida* KT2440.
Initial cultivation was carried out in 50 ml LB-medium at 30 °C for 4 h. After induction of gene expression, the cultivation was continued at 16 °C and 200 rpm for 24 h. The indicated values are means and standard deviations of three biological samples.

Besides *P. putida*, *E. coli* was established as host for heterologous expression of the DPA/DHA-type Pfa proteins plus AfPpt from *A. fasciculatus* (SBSr002). It could be shown that expression of the Pfa proteins without coexpression of AfPpt did not lead to any detectable LC-PUFA production in *E. coli* (details not shown). Therefore, cotransformation of the expression constructs pJBPfaAf1 plus pPptAfA, pJBPfaAf1* plus pPptAfA, or pJBPfaAf4 plus pPptAfD in *E. coli* BL21(DE3) was performed by electroporation. The transgenic clones were cultivated in 50 ml LB-medium at 37 °C. Expression of the Pfa proteins and AfPpt was induced at OD_{600} of 0.8 by addition of *m*-toluic acid to a final concentration of 2 mM in the case of pJBPfaAf1 plus pPptAfA and pJBPfaAf1* plus pPptAfA or of L-arabinose to a final concentration of 2% in the case of pJBPfaAf4 plus pPptAfD. After induction, the cells were cultivated at 16 °C for 24 h. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [44] and analysed by GC-MS. Best production was obtained with *E. coli* BL21(DE3)/pJBPfaAf4/pPptAfD, which produces only 0.1 mg DHA/g CDW and trace amounts of *n*-6 DPA. Thus, the *araC*-P_{BAD} promoter system was superior to the *xylS-Pm* promoter system in *E. coli* and/or L-arabinose more conducive to cell growth than *m*-toluic acid. Surprisingly, unlike for *P. putida*, no improvement of DHA production could be achieved with *Pm** promoter compared to *Pm* promoter. This observation indicates that the

transcription rate of the *pfa* genes is probably not the bottleneck for recombinant LC-PUFA production in *E. coli* under the tested conditions. It is more likely that efficient translation of the GC-rich *pfa*- and *Afppt*-mRNAs is impaired in *E. coli*, whose codon usage differs significantly from those of myxobacteria. In contrast, the higher GC-content of the genome sequence and the codon usage of *P. putida* certainly make it more suitable for the expression of the GC-rich myxobacterial genes, which might explain the observed higher LC-PUFA production rates. However, additional factors might also play a significant role, e.g. mRNA stability, translational helper factors, precursor supply, product stability, to name but a few. *P. putida* can be clearly regarded as the better host strain, especially because of its higher heterologous production rate of LC-PUFAs *per se*. Accordingly, subsequent engineering approaches were targeted to *P. putida* and extended towards optimization of product stability and precursor supply.

Interestingly, in some cultures of the transgenic strains of *P. putida* or *E. coli* expressing the recombinant PUFA biosynthetic pathway not only *n*-6 DPA and DHA were detected but also some trace amounts of arachidonic acid (AA, 20:4, *n*-6) and EPA (20:5, *n*-3). As it is hypothesized that the PUFA synthase of *Aetherobacter* spp. may only produce DHA and some minor amounts of DPA as by-products, AA and EPA could result from the degradation of these C22 PUFAs [25]. This speculation is supported by the finding that feeding of 5 mg DHA to a 50 ml culture of *P. putida* KT2440 and cultivation for 4 h at 30 °C plus 24 h at 16 °C resulted in a conversion of > 74% of DHA into EPA. In order to convert DHA into EPA or *n*-6 DPA into AA via β-oxidation, the activity of the additional enzymes 2,4-dienoyl-CoA reductase (EC 1.3.1.34) and Δ^3-Δ^2-enoyl-CoA isomerase (EC 5.3.3.8) is needed. As a consequence, gene *fadH*, encoding 2,4-dienoyl-CoA reductase, was inactivated by single crossover in the genome of *P. putida* to prevent the produced LC-PUFAs *n*-6 DPA and DHA from degradation. For this purpose, the *xylS-Pm* promoter system, *Afppt*, and *trpE* located on plasmid pPptAf1 were replaced by a version of *fadH* with frameshift mutations at the 5' end and with a truncation at the 3' end by conventional cloning, yielding plasmid p*fadH*_KO-tet^R (Figure 1C). The tetracycline resistance gene of plasmid p*fadH*_KO-tet^R was exchanged for the gentamicin resistance gene (Figure 1C). The resulting plasmid p*fadH*_KO-gm^R was then transferred into *P. putida* KT2440. Three out of many obtained gentamicin resistant clones were chosen, and the correct knockout of *fadH* by single crossover could be confirmed by PCR. Repetition of the feeding experiment using DHA and the strain *P. putida* KT2440::pfadH_KO under the same cultivation conditions described previously resulted in a conversion of 31% of DHA into EPA. This suggests some residual activity of the mutated

gene or complementation of its function by other enzymes. After transformation of pPptAf1 and pJB*PfaAf1** into *P. putida* KT2440::pfadH_KO, verification of the correct integration of pPptAf1, heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described above. Compared to *P. putida* KT2440::pPptAf1 + pJB*PfaAf1**, the mutated strain *P. putida* KT2440::pfadH_KO/pPptAf1 + pJB*PfaAf1** produces higher amounts of LC-PUFAs (0.1 mg *n*-6 DPA/g CDW and 0.7 mg DHA/g CDW) (Figure 2).

In addition to the reduction of PUFA catabolism, optimization of the precursor supply for LC-PUFA biosynthesis via PUFA synthases was also aimed at. These multifunctional enzyme systems assemble their products based on the substrate malonyl-CoA and reduce its β-keto group with the essential reductant NADPH after each chain elongation step. Insufficient supply of these substrates could be a bottleneck in LC-PUFA biosynthesis and explain the rather moderate LC-PUFA yield in the transgenic *P. putida* strains. In order to overcome this problem, metabolic engineering approaches were applied to the host strain. Thereby, additional copies of the genes encoding either intrinsic acetyl-CoA carboxylase (EC 6.4.1.2) or glucose-6-phosphate dehydrogenase (EC 1.1.1.49) were integrated into the genome of *P. putida* KT2440 under control of the constitutive *tac* promoter. Acetyl-CoA carboxylase catalyses the first committing step in fatty acid biosynthesis by a two-step mechanism. In the first reaction, biotin carboxylase is involved in the ATP-dependent carboxylation of biotin with bicarbonate. The carboxyl group is then transferred to acetyl-CoA by carboxyl transferase, yielding malonyl-CoA in the second reaction [48]. Glucose-6-phosphate dehydrogenase represents an enzyme of the Entner-Doudoroff pathway, the predominant pathway for glucose catabolism in *P. putida* [49], and catalyses the oxidation of D-glucose-6-phosphate into 6-phospho-D-glucono-1,5-lactone under consumption of one equivalent of $NADP^+$. Gene *zwf-1* encoding the glucose-6-phosphate dehydrogenase from *P. putida* was cloned downstream of the *tac* promoter plus upstream of the *tonB* transcription terminator and inserted into plasmid pPptAf1 by applying conventional cloning techniques (Figure 1D). The resulting plasmid pME3 was then transferred into *P. putida* KT2440::pfadH_KO + pJB*PfaAf1**. The correct integration of pME3 could be confirmed for selected clones by PCR. Heterologous expression of the Pfa proteins, AfPpt, plus Zwf-1 and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. The engineered strain *P. putida* KT2440::pfadH_KO/pME3 + pJBPfaAf1** produces higher amounts of LC-PUFAs (0.2 mg *n*-6 DPA/g CDW and 0.8 mg DHA/g CDW) than *P. putida* KT2440::pfadH_KO/pPptAf1 + pJB*PfaAf1** (Figure 2). The acetyl-CoA

carboxylase from *P. putida* is encoded by the genes *accA* and *accD* plus an operon comprising genes *accB-accC-1*. The genes *accA* and *accD* encode the carboxyl transferase subunits α and β, whereas gene *accC-1* encodes the biotin carboxylase and gene *accB* the biotin carboxyl carrier protein of the multienzyme complex. Each gene or operon was cloned downstream of the *tac* promoter plus upstream of the *tonB* transcription terminator in case of *accD* and assembled into cloning vector pUC18 by conventional cloning (Figure 1D). P_{tac}-*accB-accC-1*-P_{tac}-*accA*-P_{tac}-*accD*-T_{tonB} was inserted into plasmid pPptAf1 by applying conventional cloning techniques (Figure 1D). The resulting plasmid pME2 was then transferred into *P. putida* KT2440::pfadH_KO + pJB*PfaAf1**. The correct integration of pME2 could be confirmed for selected clones by PCR. Heterologous expression of the Pfa proteins, AfPpt, AccB, AccC-1, AccA, plus AccD and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. Compared to *P. putida* KT2440::pfadH_KO/pPptAf1 + pJB*PfaAf1** and *P. putida* KT2440::pfadH_KO/pME3 + pJBPfaAf1**, the engineered strain *P. putida* KT2440::pfadH_KO/pME2 + pJBPfaAf1** produces higher amounts of DHA (0.1 mg *n*-6 DPA/g CDW and 0.9 mg DHA/g CDW) (Figure 2). Hence, it can be deduced that limited supply of malonyl-CoA or NADPH represented a bottleneck for LC-PUFA biosynthesis, which could, at least partially, be eliminated by overexpression of acetyl-CoA carboxylase or glucose-6-phosphate dehydrogenase. As rational consequence of this result, collective overexpression of genes *accB-accC-1*, *accA*, and *accD* plus gene *zwf-1* was attempted by assembly of the entire gene set onto one expression plasmid, which was transferred into *P. putida* (details not shown). Unfortunately, the resulting expression strain produces unexpectedly low amounts of LC-PUFAs (0.1 mg *n*-6 DPA/g CDW and 0.7 mg DHA/g CDW). The most obvious reason for this finding is probably the instability of the expression plasmid due to constitutive overexpression of five genes in *E. coli* and *P. putida* during cloning and transformation.

In conclusion, recombinant LC-PUFA production could be successfully accomplished in *E. coli* and *P. putida* with the DPA/DHA-type *pfa* gene cluster from the myxobacterium *A. fasciculatus* (SBSr002). The latter host strain produces 0.05 mg *n*-6 DPA/l and 0.3 mg DHA/l with the initial expression construct. These yields could be increased to 0.2 mg *n*-6 DPA/l and 1.3 mg DHA/l by stepwise optimization of the expression system over seven transgenic *P. putida* strains (Figure 2). Total DHA production yields are comparable to the LC-PUFA production rate achieved with the native producer strain *A. fasciculatus* (SBSr002), which produces 1.2-1.9 mg DHA/l as major PUFA species, but not until after seven days of cultivation. For the first time, extensive genetic engineering of expression plasmids

containing a *pfa* gene cluster combined with metabolic engineering of a host strain heterologously expressing a PUFA synthase were performed in a systematic manner. All these procedures eventually led to an increase in recombinant LC-PUFA production in *P. putida* of more than 300%.

REFERENCES

[1] Lorente-Cebrian S, Costa AGV, Navas-Carretero S, Zabala M, Martinez JA, & Moreno-Aliaga MJ (2013) Role of omega-3 fatty acids in obesity, metabolic syndrome, and cardiovascular diseases: a review of the evidence. *J. Physiol. Biochem.*, **69,** 633-651.

[2] Lenihan-Geels G, Bishop KS, & Ferguson LR (2013) Alternative sources of omega-3 fats: can we find a sustainable substitute for fish? *Nutrients*, **5,** 1301-1315.

[3] Ratledge C (2004) Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie*, **86,** 807-815.

[4] Kaulmann U & Hertweck C (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew. Chem. Int. Ed.*, **41,** 1866-1869.

[5] Napier JA (2002) Plumbing the depths of PUFA biosynthesis: a novel polyketide synthase-like pathway from marine organisms. *Trends Plant Sci.*, **7,** 51-54.

[6] Wallis JG, Watts JL, & Browse J (2002) Polyunsaturated fatty acid synthesis: what will they think of next? *Trends Biochem. Sci.*, **27,** 467-473.

[7] 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.

[8] 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.

[9] Morita N, Tanaka M, & Okuyama H (2000) Biosynthesis of fatty acids in the docosahexaenoic acid-producing bacterium *Moritella marina* strain MP-1. *Biochem. Soc. Trans.*, **28,** 943-945.

[10] Hauvermale A, Kuner J, Rosenzweig B, Guerra D, Diltz S, & Metz JG (2006) Fatty acid production in *Schizochytrium* sp.: Involvement of a polyunsaturated fatty acid synthase and a type I fatty acid synthase. *Lipids*, **41,** 739-747.

[11] Ren LJ, Huang H, Xiao AH, Lian M, Jin LJ, & Ji XJ (2009) Enhanced docosahexaenoic acid production by reinforcing acetyl-CoA and NADPH supply in *Schizochytrium* sp. HX-308. *Bioprocess. Biosyst. Eng.*, **32,** 837-843.

[12] Chang G, Gao N, Tian G, Wu Q, Chang M, & Wang X (2013) Improvement of docosahexaenoic acid production on glycerol by *Schizochytrium* sp. S31 with constantly high oxygen transfer coefficient. *Bioresour. Technol.*, **142,** 400-406.

[13] Ren LJ, Sun LN, Zhuang XY, Qu L, Ji XJ, & Huang H (2014) Regulation of docosahexaenoic acid production by *Schizochytrium* sp.: effect of nitrogen addition. *Bioprocess. Biosyst. Eng*., **37,** 865-872.

[14] Fang J, Kato C, Sato T, Chan O, & McKay D (2004) Biosynthesis and dietary uptake of polyunsaturated fatty acids by piezophilic bacteria. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, **137,** 455-461.

[15] Morita N, Nishida T, Tanaka M, Yano Y, & Okuyama H (2005) Enhancement of polyunsaturated fatty acid production by cerulenin treatment in polyunsaturated fatty acid-producing bacteria. *Biotechnol. Lett.*, **27,** 389-393.

[16] Amiri-Jami M, Wang H, Kakuda Y, & Griffiths MW (2006) Enhancement of polyunsaturated fatty acid production by Tn5 transposon in *Shewanella baltica*. *Biotechnol. Lett.*, **28,** 1187-1192.

[17] Yazawa K (1996) Production of eicosapentaenoic acid from marine bacteria. *Lipids*, **31 Suppl,** S297-S300.

[18] 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 (Pt 8),** 2725-2731.

[19] Orikasa Y, Nishida T, Yamada A, Yu R, Watanabe K, Hase A, Morita N, & Okuyama H (2006) Recombinant production of docosahexaenoic acid in a polyketide biosynthesis mode in *Escherichia coli*. *Biotechnol. Lett.*, **28,** 1841-1847.

[20] Amiri-Jami M, Lapointe G, & Griffiths MW (2014) Engineering of EPA/DHA omega-3 fatty acid production by *Lactococcus lactis* subsp. *cremoris* MG1363. *Appl. Microbiol. Biotechnol.*, **98,** 3071-3080.

[21] Lee SJ, Jeong YS, Kim DU, Se JW, & Hur BK (2006) Eicosapentaenoic acid (EPA) biosynthetic gene cluster of *Shewanella oneidensis* MR-1: Cloning, heterologous expression, and effects of temperature and glucose on the production of EPA in *Escherichia coli*. *Biotechnology and Bioprocess Engineering*, **11,** 510-515.

[22] Amiri-Jami M & Griffiths MW (2010) Recombinant production of omega-3 fatty acids in *Escherichia coli* using a gene cluster isolated from *Shewanella baltica* MAC1. *J. Appl. Microbiol.*, **109,** 1897-1905.

[23] Lee SJ, Kim CH, Seo PS, Kwon O, Hur BK, & Seo JW (2008) Enhancement of heterologous production of eicosapentaenoic acid in *Escherichia coli* by substitution of promoter sequences within the biosynthesis gene cluster. *Biotechnol. Lett.*, **30,** 2139-2142.

[24] Orikasa Y, Ito Y, Nishida T, Watanabe K, Morita N, Ohwada T, Yumoto I, & Okuyama H (2007) Enhanced heterologous production of eicosapentaenoic acid in *Escherichia coli* cells that co-express eicosapentaenoic acid biosynthesis *pfa* genes and foreign DNA fragments including a high-performance catalase gene, *vktA*. *Biotechnol. Lett.*, **29,** 803-812.

[25] Gemperlein K, Rachid S, Garcia R, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

[26] Wenzel SC, Gross F, Zhang Y, Fu J, Stewart FA, & Müller R (2005) Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via red/ET recombineering. *Chem. Biol.*, **12,** 349-356.

[27] Gross F, Luniak N, Perlova O, Gaitatzis N, Jenke-Kodama H, Gerth K, Gottschalk D, Dittmann E, & Müller R (2006) Bacterial type III polyketide synthases: Phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expession in pseudomonads. *Arch. Microbiol.*, **185,** 28-38.

[28] Fu J, Wenzel SC, Perlova O, Wang J, Gross F, Tang Z, Yin Y, Stewart AF, Müller R, & Zhang Y (2008) Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res.*, **36,** e113.

[29] Chai Y, Shan S, Weissman KJ, Hu S, Zhang Y, & Müller R (2012) Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis. *Chem. Biol.*, **19,** 361-371.

[30] Gross F, Ring MW, Perlova O, Fu J, Schneider S, Gerth K, Kuhlmann S, Stewart AF, Zhang Y, & Müller R (2006) Metabolic engineering of *Pseudomonas putida* for methylmalonyl-CoA biosynthesis to enable complex heterologous secondary metabolite formation. *Chem. Biol.*, **13,** 1253-1264.

[31] Grant SG, Jessee J, Bloom FR, & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U S A*, **87,** 4645-4649.

[32] Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc. Natl. Acad. Sci. U S A*, **76,** 1648-1652.

[33] Studier FW & Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **189,** 113-130.

[34] Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF, & Zhang Y (2006) An improved recombineering approach by adding RecA to λ Red recombination. *Mol. Biotechnol.*, **32,** 43-53.

[35] Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Müller R, Stewart AF, & Zhang Y (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.*, **30,** 440-446.

[36] Bagdasarian M, Lurz R, Ruckert B, Franklin FC, Bagdasarian MM, Frey J, & Timmis KN (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene*, **16,** 237-247.

[37] Sambrook J & Russell DW (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[38] Blatny JM, Brautaset T, Winther-Larsen HC, Karunakaran P, & Valla S (1997) Improved broad-host-range RK2 vectors useful for high and low regulated gene expression levels in Gram-negative bacteria. *Plasmid*, **38,** 35-51.

[39] Prior JE, Lynch MD, & Gill RT (2010) Broad-host-range vectors for protein expression across Gram negative hosts. *Biotechnol. Bioeng.*, **106,** 326-332.

[40] Kopp M, Irschik H, Gross F, Perlova O, Sandmann A, Gerth K, & Müller R (2004) Critical variations of conjugational DNA transfer into secondary metabolite multiproducing *Sorangium cellulosum* strains So ce12 and So ce56: development of a mariner-based transposon mutagenesis system. *J. Biotechnol.*, **107,** 29-40.

[41] Berg L, Lale R, Bakke I, Burroughs N, & Valla S (2009) The expression of recombinant genes in *Escherichia coli* can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5'-untranslated part of mRNA. *Microb. Biotechnol.*, **2,** 379-389.

[42] Bakke I, Berg L, Aune TE, Brautaset T, Sletta H, Tondervik A, & Valla S (2009) Random mutagenesis of the *Pm* promoter as a powerful strategy for improvement of recombinant-gene expression. *Appl. Environ. Microbiol.*, **75,** 2002-2011.

[43] Hill DS, Stein JI, Torkewitz NR, Morse AM, Howell CR, Pachlatko JP, Becker JO, & Ligon JM (1994) Cloning of genes involved in the synthesis of pyrrolnitrin from *Pseudomonas fluorescens* and role of pyrrolnitrin synthesis in biological control of plant disease. *Appl. Environ. Microbiol.*, **60,** 78-85.

[44] Bode HB, Ring MW, Schwär G, Kroppenstedt RM, Kaiser D, & Müller R (2006) 3-Hydroxy-3-methylglutaryl-coenzyme A (CoA) synthase is involved in the biosynthesis of isovaleryl-CoA in the myxobacterium *Myxococcus xanthus* during fruiting body formation. *J. Bacteriol.*, **188,** 6524-6528.

[45] Zhang Y, Muyrers JPP, Testa G, & Stewart AF (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18,** 1314-1317.

[46] Gross F, Gottschalk D, & Müller R (2005) Posttranslational modification of myxobacterial carrier protein domains in *Pseudomonas* sp. by an intrinsic phosphopantetheinyl transferase. *Appl. Microbiol. Biotechnol.*, **68,** 66-74.

[47] Durland RH & Helinski DR (1987) The sequence encoding the 43-kilodalton *trfA* protein is required for efficient replication or maintenance of minimal RK2 replicons in *Pseudomonas aeruginosa*. *Plasmid*, **18,** 164-169.

[48] Cronan JE, Jr. & Waldrop GL (2002) Multi-subunit acetyl-CoA carboxylases. *Prog. Lipid Res.*, **41,** 407-435.

[49] del Castillo T, Ramos JL, Rodriguez-Herva JJ, Fuhrer T, Sauer U, & Duque E (2007) Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida*: genomic and flux analysis. *J. Bacteriol.*, **189,** 5142-5152.

CHAPTER 4

PUFA production in *Pseudomonas putida* based on expression of artificial gene clusters encoding a myxobacterial PUFA synthase

INTRODUCTION

In quest of a suitable microbial chassis for recombinant production of long-chain polyunsaturated fatty acids (LC-PUFAs) using a native myxobacterial *pfa* gene cluster, the Gram-negative and generally recognized as safe (*GRAS)* bacterium *Pseudomonas putida* KT2440 turned out to be a fast-growing, reliable, and robust host organism (Chapter 3)**.** Recombinant production of docosahexaenoic acid (DHA, 22:6, *n*-3) and *n*-6 docosapentaenoic acid (DPA, 22:5) could be successfully accomplished in *P. putida* with the DPA/DHA-type *pfa* gene cluster from the myxobacterium *Aetherobacter fasciculatus* (SBSr002)**.** Moreover, the strain has proven to be accessible for genetic and metabolic engineering, which led to further improvements of the production yield of LC-PUFAs. Compared to *Escherichia coli*, *P. putida* turned out to be a better host strain for the expression of the native PUFA biosynthetic pathway from *A. fasciculatus* (SBSr002). However, there are some significant differences in the codon bias of the *pfa* gene cluster from *A. fasciculatus* (SBSr002) and the coding sequences in the genome of *P. putida*, which appear to interfere with optimal heterologous gene expression.

Total synthesis of large DNA fragments to reconstitute complete gene cluster offers opportunities to holistically optimize DNA sequences of biosynthetic pathways for the chosen production host including codon bias adaptation and optimization of several other sequence parameters important for functional expression. On top of this, sophisticated constructional design of artificial pathways, e.g. including engineering of restriction sites, enables facile and flexible genetic manipulation of pathway components. Only three examples for using this modern technique on complex multimodular assembly lines are described in literature so far. The 32 kb erythromycin biosynthetic pathway has been redesigned, synthesized, and functionally expressed in *E. coli* [1;2]. Synthetic 58 kb epothilone biosynthetic pathways were constructed and functionally expressed in *E. coli* [3] and *Myxococcus xanthus* [4]. Unfortunately, the achieved production titres were found very low [4], even after engineering of the *E. coli* host strain [5]. Nonetheless, advantages and potential of Synthetic Biology approaches for heterologous expression systems appear attractive as well as promising and were therefore applied to the optimization of recombinant LC-PUFA production. Hence, artificial *pfa* gene clusters originating from *A. fasciculatus* (SBSr002) PUFA biosynthetic pathway [6] were redesigned, synthesized, and assembled for *P. putida*. To the best of our knowledge, this represents the first example for the application of gene synthesis on multifunctional PUFA synthases and the usage of *Pseudomonas* as expression host for complex, artificial polyketide synthase-like pathways. Different *pfa* pathway versions were

constructed and successful transfer of the expression plasmids and heterologous PUFA production was accomplished in the engineered host strain.

EXPERIMENTAL PROCEDURES

Sequence analysis and design of the synthetic gene clusters

The sequence of the *pfa* gene cluster of *A. fasciculatus* (SBSr002) was analysed and compared to the genome sequence of *P. putida* KT2440 [7] retrieved from NCBI Genome RefSeq NC_002947 in collaboration with ATG:biosynthetics. Based on this, relevant parameters for constructional and functional sequence design were defined to generate artificial pathway versions using the proprietary evoMAG-GeneDesign software [4]. The sequence design process included engineering of restriction sites, adaptation of the codon usage, elimination of sequence repeats as well as rare codon clusters, engineering of Shine-Dalgarno (SD)–anti-SD interactions, and introduction of hidden stop codons in unused frames.

Culture conditions

Escherichia coli DH10B [8] was used for cloning experiments. *E. coli* HB101/pRK2013 [9] was used as helper strain for conjugation experiments. The cells were grown in LB-medium or on LB-agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) at 37 °C (and 200 rpm) overnight. Kanamycin was used at the concentration of 50 µg/ml.

Pseudomonas putida KT2440 [10] was cultivated in liquid LB-medium at 30 °C and 200 rpm overnight. Conjugation was carried out onto PMM-agar plates containing 0.61% K_2HPO_4, 0.5% KH_2PO_4, 0.1% $(NH_4)_2SO_4$, 0.66% disodium succinate, and 1.5% agar. The medium was adjusted to pH 7.0 and supplemented with 0.12% $MgSO_4$ after autoclaving. Antibiotics were used at the following concentrations: 60 µg/ml kanamycin, 30 µg/ml tetracycline, and 20 µg/ml gentamicin.

General molecular biology techniques

Routine handling of nucleic acids, such as isolation of genomic and plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, were performed

according to standard protocols [11]. All the enzymes were purchased from Thermo Scientific or New England Biolabs.

Cloning of expression constructs containing the synthetic *pfa* gene clusters originating from *A. fasciculatus* (SBSr002) for heterologous LC-PUFA production in *P. putida*

For heterologous expression of artificial PUFA biosynthetic pathways in *P. putida* KT2440, three synthetic versions of the *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) were designed. The synthetic DNA fragments were supplied by gene synthesis companies. For each version, the gene cluster is composed of seven building blocks, BB1-BB7, flanked by unique restriction sites (see appendix for sequences). Version 1 consists of the native DNA sequence of the *pfa* gene cluster; only constructional design was implemented. Therefore, *Swa*I, *Nhe*I, *Mun*I, *Eco*RV, *Mlu*I, *Sca*I, *Sgr*DI, *Bgl*II, *Fse*I, *Acl*I, *Apa*LI, *Fsp*I, *Sna*BI, *Bsp*EI, and *Pac*I restriction sites were introduced for cloning purposes as well as for exchangeability of genes/domains and were excluded from any other unwanted position within the gene cluster. The resulting artificial version 1 of the pathway was further modified to generate version 2 by applying functional design approaches with special focus on sequence features affecting *pfa* gene translation (adaptation of the codon usage, elimination of sequence repeats as well as rare codon clusters, engineering of Shine-Dalgarno (SD)–anti-SD interactions, and introduction of hidden stop codons in unused frames). In version 2a, a derivative of version 2, the overlap of the stop codon of *pfa2* and the start codon of *pfa3* was eliminated by introducing a 40 bp intergenic region, TR-*pfa2-pfa3*_V2, comprising a Shine-Dalgarno sequence (see appendix for sequence).

A cloning plasmid with kanamycin resistance gene, p15A origin of replication, and restriction sites for assembly of the seven DNA building blocks was constructed. Therefore, a multiple cloning site comprising *Swa*I, *Mun*I, *Mlu*I, *Bgl*II, *Acl*I, *Apa*LI, *Sna*BI, and *Pac*I restriction sites (0.1 kb) was amplified by PCR using overlapping primers MCS for pACYC_fwd_2 and MCS for pACYC_rev_2 (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The MCS was phosphorylated and ligated into plasmid pACYC177 (New England Biolabs), digested with

*Dra*I and *Aan*I and dephosphorylated, yielding plasmid pACYC_assembly_2 (Table 2). The seven DNA building blocks of version 1 and version 2 were stitched together in plasmid pACYC_assembly_2, respectively. The assembly procedure is explained in more detail below: BB1_V1 or BB1_V2 was inserted into pACYC_assembly_2 via *Swa*I and *Mun*I restriction sites, generating plasmid pACYC_BB1_V1 or pACYC_BB1_V2 (Table 2); BB2_V1 or BB2_V2 was inserted into pACYC_BB1_V1 or pACYC_BB1_V2 via *Mun*I and *Mlu*I restriction sites, generating plasmid pACYC_BB1+2_V1 or pACYC_BB1+2_V2 (Table 2); BB3_V1 or BB3_V2 was inserted into pACYC_BB1+2_V1 or pACYC_BB1+2_V2 via *Mlu*I and *Bgl*II restriction sites, generating plasmid pACYC_BB1-3_V1 or pACYC_BB1-3_V2 (Table 2); BB4_V1 or BB4_V2 was inserted into pACYC_BB1-3_V1 or pACYC_BB1-3_V2 via *Bgl*II and *Acl*I restriction sites, generating plasmid pACYC_BB1-4_V1 or pACYC_BB1-4_V2 (Table 2); BB5_V1 or BB5_V2 was inserted into pACYC_BB1-4_V1 or pACYC_BB1-4_V2 via *Acl*I and *Apa*LI restriction sites, generating plasmid pACYC_BB1-5_V1 or pACYC_BB1-5_V2 (Table 2); BB6_V1 or BB6_V2 was inserted into pACYC_BB1-5_V1 or pACYC_BB1-5_V2 via *Apa*LI and *Sna*BI restriction sites, generating plasmid pACYC_BB1-6_V1 or pACYC_BB1-6_V2 (Table 2); BB7_V1 or BB7_V2 was inserted into pACYC_BB1-6_V1 or pACYC_BB1-6_V2 via *Sna*BI and *Pac*I restriction sites, generating plasmid pACYC_BB1-7_V1 or pACYC_BB1-7_V2 (Table 2). TR-*pfa2-pfa3*_V2 was inserted into pACYC_BB1-7_V2 via *Fse*I and *Acl*I restriction sites, yielding plasmid pACYC_BB1-7_V2a (Table 2). The vector backbones of pACYC_BB1-7_V1, pACYC_BB1-7_V2, and pACYC_BB1-7_V2a were exchanged for the backbone of plasmid pJB*PfaAf1** (Chapter 3) via *Pml*I and *Pac*I restriction sites, resulting in plasmids pPm**SynPfaAf1, pPm**SynPfaAf2, and pPm**SynPfaAf2a (Table 2).

Table 1. Primers used for cloning of the assembly plasmid in this study.
The introduced restriction sites are underlined.

Primer	Sequence (5'→3')
MCS for pACYC_fwd_2	GCCCTTCCGGCATTTAAATGCGCAATTGGCGACGCGTCGCAGATCTCGCAACGTTGCGC
MCS for pACYC_rev_2	GGATCCTTAATTAACGCTACGTAGCGCGTGCACGCGCAACGTTGCGAGATCTGCGACGC

Table 2. Plasmids and expression strains constructed in this study.
oriV = origin of replication, *trfA* = gene encoding replication initiation protein in RK2 replicons, *oriT* = origin of transfer, *aph(3')-Ia* = kanamycin resistance gene.

Plasmid / Strain	Characteristics
pACYC_assembly_2	Derivative of pACYC177 (New England Biolabs) in which a multiple cloning site comprising *Swa*I, *Mun*I, *Mlu*I, *Bgl*II, *Acl*I, *Apa*LI, *Sna*BI, and *Pac*I restriction sites was inserted via *Dra*I and *Aan*I
pACYC_BB1_V1	Derivative of pACYC_assembly_2 in which BB1_V1 was inserted via *Swa*I and *Mun*I
pACYC_BB1+2_V1	Derivative of pACYC_BB1_V1 in which BB2_V1 was inserted via *Mun*I and *Mlu*I
pACYC_BB1-3_V1	Derivative of pACYC_BB1+2_V1 in which BB3_V1 was inserted via *Mlu*I and *Bgl*II
pACYC_BB1-4_V1	Derivative of pACYC_BB1-3_V1 in which BB4_V1 was inserted via *Bgl*II and *Acl*I
pACYC_BB1-5_V1	Derivative of pACYC_BB1-4_V1 in which BB5_V1 was inserted via *Acl*I and *Apa*LI
pACYC_BB1-6_V1	Derivative of pACYC_BB1-5_V1 in which BB6_V1 was inserted via *Apa*LI and *Sna*BI
pACYC_BB1-7_V1	Derivative of pACYC_BB1-6_V1 in which BB7_V1 was inserted via *Sna*BI and *Pac*I
pACYC_BB1_V2	Derivative of pACYC_assembly_2 in which BB1_V2 was inserted via *Swa*I and *Mun*I
pACYC_BB1+2_V2	Derivative of pACYC_BB1_V2 in which BB2_V2 was inserted via *Mun*I and *Mlu*I
pACYC_BB1-3_V2	Derivative of pACYC_BB1+2_V2 in which BB3_V2 was inserted via *Mlu*I and *Bgl*II
pACYC_BB1-4_V2	Derivative of pACYC_BB1-3_V2 in which BB4_V2 was inserted via *Bgl*II and *Acl*I
pACYC_BB1-5_V2	Derivative of pACYC_BB1-4_V2 in which BB5_V2 was inserted via *Acl*I and *Apa*LI
pACYC_BB1-6_V2	Derivative of pACYC_BB1-5_V2 in which BB6_V2 was inserted via *Apa*LI and *Sna*BI
pACYC_BB1-7_V2	Derivative of pACYC_BB1-6_V2 in which BB7_V2 was inserted via *Sna*BI and *Pac*I
pACYC_BB1-7_V2a	Derivative of pACYC_BB1-7_V2 in which TR-*pfa2-pfa3*_V2 was inserted via *Fse*I and *Acl*I
pPm**SynPfaAf1	Derivative of pACYC_BB1-7_V1 in which the vector backbone was exchanged for the vector backbone of pJB*PfaAf1** (Chapter 3) via *Pml*I and *Pac*I. RK2 *oriT*, RK2 *oriV*, *trfA**, *aph(3')-Ia*
pPm**SynPfaAf2	Derivative of pACYC_BB1-7_V2 in which the vector backbone was exchanged for the vector backbone of pJB*PfaAf1** (Chapter 3) via *Pml*I and *Pac*I
pPm**SynPfaAf2a	Derivative of pACYC_BB1-7_V2a in which the vector backbone was exchanged for the vector backbone of pJB*PfaAf1** (Chapter 3) via *Pml*I and *Pac*I
P. putida KT2440::pfadH_KO/pME2 + pPm**SynPfaAf1	*P. putida* KT2440::pfadH_KO/pME2 (Chapter 3) with replicative plasmid pPm**SynPfaAf1

Continuation of Table 2.

Plasmid / Strain	Characteristics
P. putida KT2440::pfadH_KO/pME2 + pPm**SynPfaAf2	*P. putida* KT2440::pfadH_KO/pME2 (Chapter 3) with replicative plasmid pPm**SynPfaAf2
P. putida KT2440::pfadH_KO/pME2 + pPm**SynPfaAf2a	*P. putida* KT2440::pfadH_KO/pME2 (Chapter 3) with replicative plasmid pPm**SynPfaAf2a

Transformation of *P. putida* by triparental conjugation and heterologous expression

Expression plasmids were transferred into *P. putida* KT2440 by triparental conjugation. 1 ml of an overnight culture of the acceptor strain *P. putida* KT2440::pfadH_KO/pME2, the donor strain *E. coli* DH10B containing the final expression construct, and the helper strain *E. coli* HB101/pRK2013 were harvested at 13,000 rpm for 1 min. After discarding the supernatant, the cells were resuspended in 1 ml LB-medium in each case. The cells were spun down at 13,000 rpm for 1 min, and the supernatant was discarded. This wash step was repeated once. The cells were resuspended in 300 µl LB-medium in each case. The triparental mating was performed by combining 50 µl of each 300 µl suspension on a small area of a LB-agar plate [12]. After incubation at 37 °C for 4 h, the plate was transferred to 30 °C and incubated overnight. The cells were scraped from the plate and resuspended in 1 ml sterile ddH_2O. 10-100 µl of this suspension was then plated out on PMM-agar containing 60 µg/ml kanamycin, 30 µg/ml tetracycline, and 20 µg/ml gentamicin. The *E. coli* strains were additionally counterselected by 150 µg/ml ampicillin, which *P. putida* is resistant to. The plates were incubated at 30 °C for 1-2 days. Transconjugants were verified by plasmid isolation and restriction analysis of the replicative plasmids. The constructed strains are listed in Table 2.

For heterologous LC-PUFA production in *P. putida*, cultivation was carried out in 50 ml LB-medium containing 60 µg/ml kanamycin, 30 µg/ml tetracycline, and 20 µg/ml gentamicin. The medium was inoculated with an overnight culture (1:100) and incubated at 30 °C for 4 h. Expression of the Pfa proteins and AfPpt was induced by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C and 200 rpm for 24 h and then harvested at 8,000 rpm for 5 min.

Extraction of cellular fatty acids

The cellular fatty acids were extracted using the FAME method [13]. For this purpose, 50 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. The cell pellet was transferred to a glass vial and dried in a vacuum concentrator. Subsequently, the cell dry weight was determined. 5 µl (50 µg) of *n*-3 DPA (Sigma-Aldrich) and 500 µl of a mixture of methanol, toluene, and sulphuric acid (50:50:2, v/v/v) were added. The vial was capped with a teflon-lined screw cap and incubated at 80 °C for 24-48 h. After the mixture was cooled to room temperature, 400 µl of an aqueous solution consisting of 0.5 M NH_4HCO_3 and 2 M KCl were added, and the sample was vortexed for 30 s. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. 100 µl of the upper phase were used for GC-MS analysis.

Analysis of fatty acid methyl esters by GC-MS

GC-MS was carried out on an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a 7683B split/splitless injector with autosampler (Agilent Technologies) and coupled to a 5973 electron impact mass selective detector (Agilent Technologies). Helium was used as the carrier gas at a flow rate of 1 ml/min. 1 µl of the sample was injected in split mode (split ratio, 10:1). The analytical column was a (5% phenyl)-methylpolysiloxane capillary column (Agilent J&W DB-5ht; 30 m x 0.25 mm i.d. x 0.1 µm film thickness, maximum temperature 400 °C; Agilent Technologies). The column temperature was kept at 130 °C for 2.5 min, increased to 240 °C at a rate of 5 °C/min, then ramped to 300 °C at 30 °C/min, and held at 300 °C for 5 min. Other temperatures were as follows: inlet, 275 °C; GC-MS transfer line, 280 °C; ion source, 230 °C; and quadrupole, 150 °C. The mass selective detector was operated in scan mode, scanning the mass range from *m*/*z* 40 to 700. Scan control, data acquisition, and processing were performed by MSD ChemStation and AMDIS software, version 2.68, based on the fragmentation patterns and retention times, in comparison with Supelco 37 Component FAME Mix and LC-PUFAs (all Sigma-Aldrich) as reference standards, and NIST 08 library. Absolute amounts of PUFAs were quantified by integration of the peaks using MSD ChemStation and by subsequent calculation in relation to the integral of *n*-3 DPA and to cell dry weight.

RESULTS AND DISCUSSION

Recombinant LC-PUFA production using the DPA (22:5, *n*-6 or *n*-3)/DHA (22:6, *n*-3)-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) has already been successfully accomplished after cloning, transfer, and expression of the gene cluster in an engineered *P. putida* strain (Chapter 3). Thereby, production of *n*-6 DPA (0.1 mg/g CDW; 0.2 mg/l) and DHA (0.9 mg/g CDW; 1.3 mg/l) could finally be achieved after 28 h of cultivation. Here, in a new approach, three versions of a synthetic DPA/DHA-type *pfa* gene cluster based on the sequence of *A. fasciculatus* (SBSr002) were created for *P. putida* in order to achieve further yield improvements. Design and synthesis of the artificial DNA sequences were restricted to the three *pfa* genes, which were intended to be cloned into the same vector backbone used for expression of the native *pfa* gene cluster (pJB*PfaAf1**, containing engineered genetic elements for improved expression rates; Chapter 3). During sequence modulation processes, the *pfa* genes were subjected to the algorithms of the proprietary evoMAG-GeneDesign software by ATG:biosynthetics. The software applies concepts of evolutionary algorithms [14;15] to generate artificial sequences according to predefined multivariate sequence parameter values. Degeneracy of the genetic code allows for the substitution of synonymous codons by silent mutations in order to generate the artificial sequence based on the native template without altering the amino acid sequence. The resulting artificial sequence of the 16.2 kb *pfa* gene cluster from *A. fasciculatus* (SBSr002) was dissected into seven DNA fragments, which could be supplied by gene synthesis companies despite their high GC-content (Figure 1).

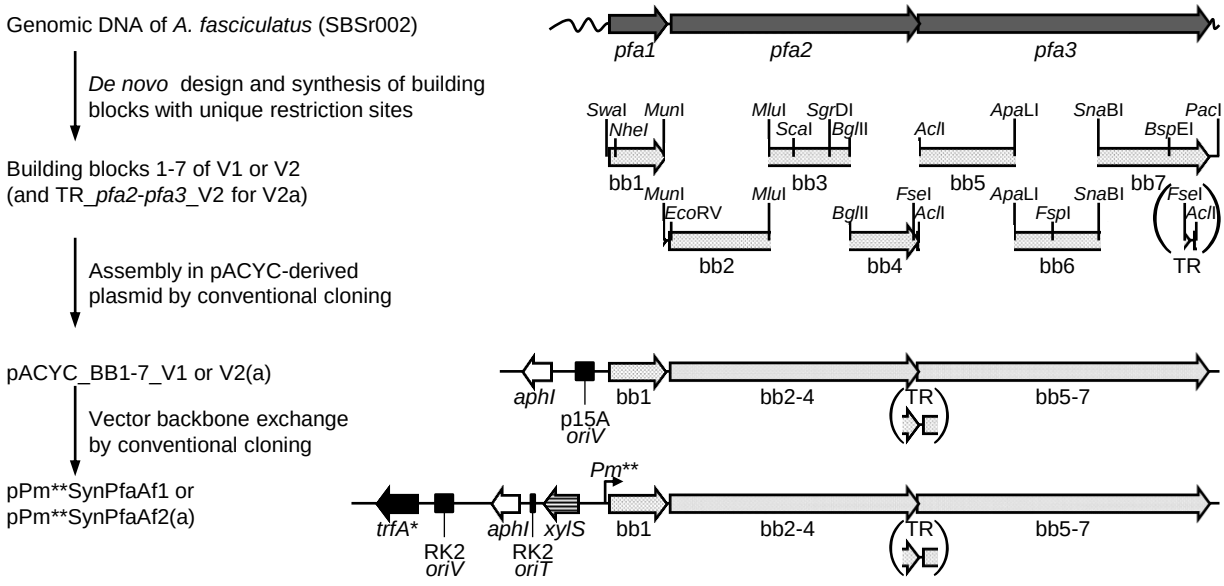


Figure 1. Design, synthesis, and cloning of a synthetic *pfa* gene cluster originating from *Aetherobacter fasciculatus* (SBSr002) for recombinant LC-PUFA production in *Pseudomonas putida* KT2440.
oriV = origin of replication, *trfA** = gene encoding replication initiation protein in RK2 replicons, *cop271C* mutation included, *oriT* = origin of transfer, *aphI* = kanamycin resistance gene, *Pm*** = *Pm* promoter, mutated 5'-UTR of version LII-11 [16] plus the mutated core promoter region of version ML2-2 [17] included, *xylS* = gene encoding the transcriptional regulator for *Pm* promoter.

The sequence design process involved engineering of restriction sites to allow for pathway assembly and interchangeability of inter- and intragenic regions. During this procedure, 69 codons were affected by synonymous substitutions to generate the artificial gene cluster version 1, which is still highly similar to the native *pfa* gene cluster from *A. fasciculatus* (SBSr002). Based on the same constructional sequence design, yet another artificial *pfa* gene cluster sequence (version 2) was generated to additionally consider several functional sequence parameters, especially those affecting *pfa* gene translation (Figure 2), as well as some general aspects of sequence optimization (e.g. avoidance of sequence repeats, mRNA stabilities). Comparison of the codon usage of the original cluster sequence and the codon table of *P. putida* led to the application of a 10% exclusion threshold (Figure 2A). The resulting artificial codon usage table only contains 48 instead of 61 sense codons. Renormalization of the synonymous codon fractions from the genome sequence of *P. putida* had to be performed to ensure that the sum of each subset of codons for the same amino acid equals 1. Eventually, high adaptation of the codon usage of the artificial cluster sequence version 2 to the reduced and renormalized table could be achieved. Another important parameter affecting translational elongation represents the local codon adaptation index (CAI), which is a measure for the deviation of the synonymous codon usage bias of a given coding sequence with respect to a reference set of genes [18]. After sequence adaptation, the local CAI shows a clearly smoothed shape without any distinct peaks as found for the native sequence (Figure 2B). In the modified sequence version 2, the gradient between start and stop codon of each coding DNA sequence was 0.05, which is intended to enhance the ribosome occupancy along the mRNA to shield the transcript from degradation [19]. In general, the speed of ribosomal protein synthesis and the resulting translational elongation rate depends on several factors, including the availability of aminoacyl-tRNAs, but also on certain sequence features of the mRNA template, e.g. rare codon clusters [20] or internal Shine-Dalgarno sequences [21]. These sequence features are often assigned to temporal separation of translation of segments within the peptide chain and proper cotranslational folding of proteins with direct impact on their solubility and activity [22]. However, the search for relevant rare codon clusters within the native *pfa* genes did not reveal any conspicuous candidates. Thus, rare codon clusters were eliminated (Figure 2C) and Shine-Dalgarno (SD)–anti-SD interactions were attenuated (Figure 2D) during the codon adaptation procedure to design the synthetic gene cluster version 2. Another important factor to be considered is the metabolic cost of translational miselongation upon a frameshift [23]. In order to cause early termination of translation of incorrect reading frames, the ribosome must encounter an off-frame stop

codon. Hence, stop codons were accumulated in the two unused frames of the artificial cluster sequence version 2 (Figure 2E).

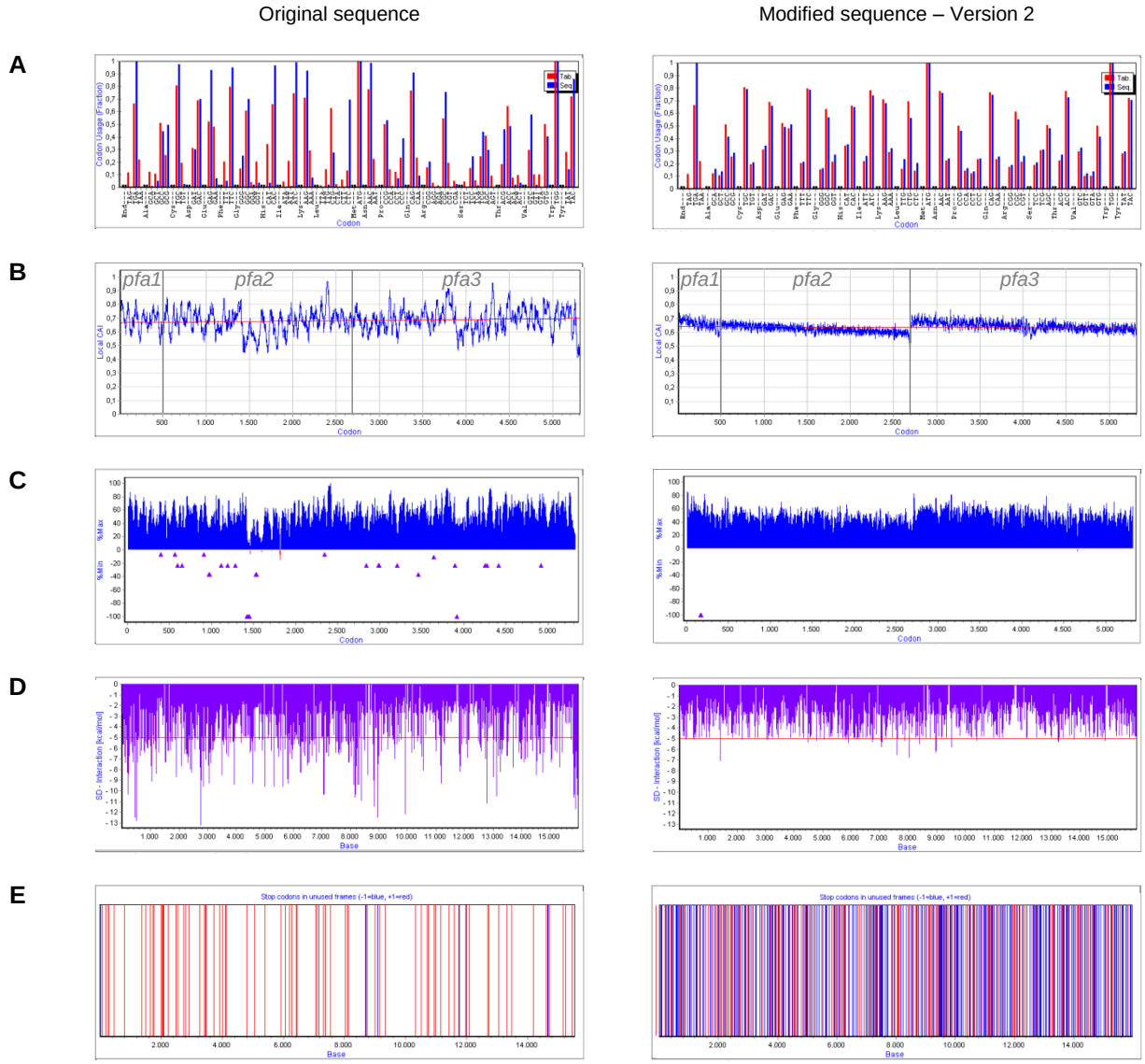


Figure 2. Comparison of the native *pfa* gene cluster sequence from *Aetherobacter fasciculatus* (SBSr002) with the artificial sequence (version 2) adapted for expression in *Pseudomonas putida*.
(A) Codon usage of the original cluster sequence (left blue bars) compared to the codon table of *Pseudomonas putida* (left red bars) and codon usage of the modified cluster sequence version 2 (right blue bars) compared to the reduced and renormalized codon table of *Pseudomonas putida* (right red bars; the cut-off for the synonymous codon fractions was 0.1), (B) Course of the local codon adaptation index (CAI); in the modified sequence version 2, the gradient between start and stop codon of each coding DNA sequence (CDS) was 0.05, (C) Rare codon clusters according to Clarke and Clark [20]; additional violet triangles: tandem-repeats of codons below average usage frequency, (D) Shine-Dalgarno (SD)–anti-SD interactions, (E) Hidden stop codons in unused frames. Untranslated regions between the protein coding regions were excluded from the graphs in (B) and (C).

The myxobacterial *pfa* gene clusters are characterized by an overlap of the 3' end of gene *pfa2* and the 5' end of gene *pfa3* by four bases, indicating a translational coupling (Figure 3). A putative Shine-Dalgarno sequence for *pfa3* is located 8 bp upstream of the stop codon of *pfa2*. For the synthetic gene cluster version 2, the minimum free energy of the mRNA secondary structure comprising 60 bp upstream and 37 bp downstream of the start codon of *pfa3* is calculated as -31.7 kcal/mol, based on predictions from the RNAfold web server [24]. In the artificial gene cluster version 2a, a derivative of version 2, the overlap of *pfa2* and *pfa3* was eliminated by introducing a 40 bp intergenic transition region, comprising a Shine-Dalgarno sequence 7 bp upstream of the start codon of *pfa3* (Figure 3). The minimum free energy of the predicted mRNA secondary structure ranging from 17 bp upstream of the stop codon of *pfa2* to 37 bp downstream of the start codon of *pfa3* is calculated as -11.7 kcal/mol. This should result in an optimization of the translation initiation rate.

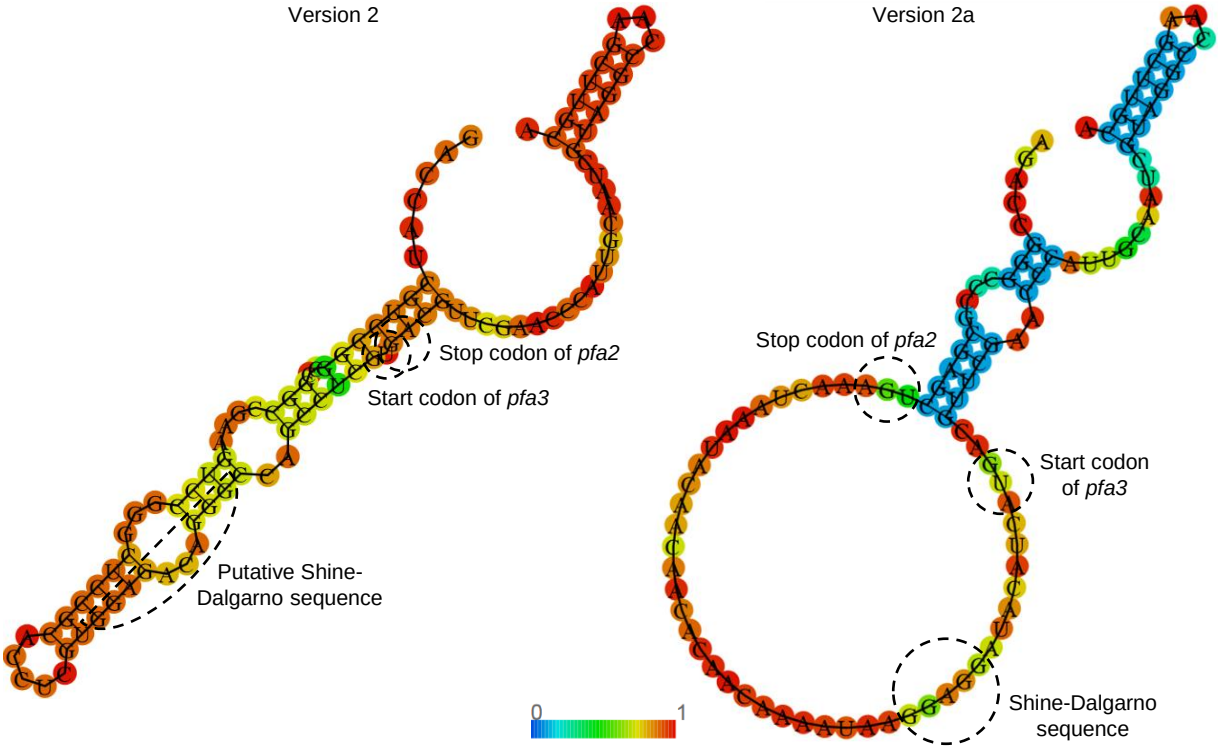


Figure 3. Minimum free energy (MFE) structures of the mRNAs transcribed from the 3' end of synthetic gene *pfa2* and the 5' end of synthetic gene *pfa3*, as predicted by RNAfold web server [24].
In version 2 the genes *pfa2* and *pfa3* overlap by four bases, whereas in version 2a the genes are separated by a 40 bp intergenic region. The structures are coloured by base-pairing probabilities. For unpaired regions, the colour denotes the probability of being unpaired.

Practical implementation of the heterologous expression of the artificial *pfa* gene clusters in *P. putida* started with reconstitution of the entire pathway by assembly of the DNA building blocks obtained from gene synthesis companies (Figure 1). For this purpose, a pACYC-derived assembly plasmid was constructed by insertion of a suitable multiple cloning site. The DNA building blocks of version 1 and version 2, respectively, were assembled stepwise by conventional cloning techniques using unique restriction sites engineered during the constructional sequence design process. Gene cluster version 2 was further modified to incorporate an artificial intergenic transition region between genes *pfa2* and *pfa3* by replacing the respective fragment for a modified version in DNA building block 4 to yield version 2a. The backbones of the resulting plasmids pACYC_BB1-7_V1, pACYC_BB1-7_V2, or pACYC_BB1-7_V2a were then exchanged for the vector backbone of pJB*PfaAf1** (Chapter 3). It is derived from the broad-host-range expression vector pJB861 with a copy-up point mutation [25]. Heterologous expression of the *pfa* genes is driven by the *xylS-Pm* promoter system with mutations in the 5'-untranslated region [16] plus the core region [17] of *Pm* promoter. The resulting plasmids pPm**SynPfaAf1, pPm**SynPfaAf2, or pPm**SynPfaAf2a were transferred into *P. putida* KT2440::pfadH_KO/pME2 by triparental conjugation. This strain has been engineered for improved LC-PUFA production (Chapter 3). Gene *fadH* encoding 2,4-dienoyl-CoA reductase was inactivated in the genome of *P. putida* KT2440::pfadH_KO/pME2 to prevent the produced LC-PUFAs from degradation; 4'-phosphopantetheinyl transferase AfPpt from *A. fasciculatus* (SBSr002) and glucose-6-phosphate dehydrogenase from *P. putida* were overexpressed to increase the phosphopantetheinylation of PUFA synthases and the supply of the substrate malonyl-CoA, respectively. The transgenic clones were cultivated in 50 ml LB-medium at 30 °C for 4 h. Expression of the Pfa proteins plus AfPpt was induced by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C for 24 h. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [13] and analysed by GC-MS. PUFA production was analysed in comparison to strain *P. putida* KT2440::pfadH_KO/pME2 + pJB*PfaAf1** (Chapter 3) expressing the native *pfa* gene cluster from *A. fasciculatus* (SBSr002) under the same conditions to yield 0.1 mg *n*-6 DPA/g cell dry weight (CDW) and 0.9 mg DHA/g CDW (Figure 4). Surprisingly, strain *P. putida KT2440::pfadH_KO/pME2* + pPm**SynPfaAf1, expressing artificial *pfa* gene cluster version 1, turned out to produce LC-PUFAs in significantly lower amounts (0.1 mg *n*-6 DPA/g CDW and 0.5 mg DHA/g CDW) (Figure 4). The only differences between these two strains are silent mutations in 69 codons within the *pfa* genes modified by silent mutations during

constructional sequence design, which was not expected to have a negative impact on their functional expression. However, reproducible lower production yields from several clones of *P. putida KT2440::pfadH_KO/pME2* + pPm**SynPfaAf1 indicate that one or more of these point mutations affect *pfa* gene expression on the transcriptional and/or translational level. On the contrary, after applying additional changes to the DNA sequence to better adapt the *pfa* genes for expression in *P. putida* with special emphasis on optimization of translational elongation, a clearly positive effect on recombinant LC-PUFA production was observed. Strain *P. putida* KT2440::pfadH_KO/pME2 + pPm**SynPfaAf2 containing the artificial *pfa* gene cluster version 2 produces 0.3 mg *n*-6 DPA/g CDW and 1.2 mg DHA/g CDW. Further engineering of the synthetic gene cluster version 2 into version 2a included the optimization of the translation initiation rate on *pfa3*-mRNA. The corresponding strain *P. putida* KT2440::pfadH_KO/pME2 + pPm**SynPfaAf2a shows only a slight but not significant improvement of DHA production (0.2 mg *n*-6 DPA/g CDW and 1.4 mg DHA/g CDW). From all these data, it can be deduced that the outcome of heterologous expression of artificial gene cluster sequences rationally designed by introducing silent point mutations is still far from being easily predictable, as many cellular processes affecting gene expression and stability of foreign DNA are not yet sufficiently studied. However, the application of additional, functional sequence design approaches on the first synthetic gene cluster (version 1) led to the development of optimized pathways (version 2/2a) that allow for much higher PUFA production with yield improvements of around 180%. However, all the strains with artificial *pfa* gene clusters show a high standard deviation between distinct clones, which is not the case for different clones of strains with the native *pfa* gene cluster (Figure 4).

In conclusion, recombinant LC-PUFA production could be successfully accomplished in *P. putida* by design, construction and heterologous expression of artificial DPA/DHA-type *pfa* gene clusters originating from the myxobacterium *A. fasciculatus* (SBSr002). Sophisticated functional sequence design for *P. putida* led to a production of 0.4 mg *n*-6 DPA/l and 3 mg DHA/l, which surpass the LC-PUFA production rate achieved with the native myxobacterial producer (1.2-1.9 mg DHA/l as major PUFA species after seven days of cultivation). Using the first polyketide biosynthetic gene cluster modulated and synthesized for this microbe, an increase in LC-PUFA production of more than 50% could be attained compared to heterologous expression of the native *pfa* gene cluster.

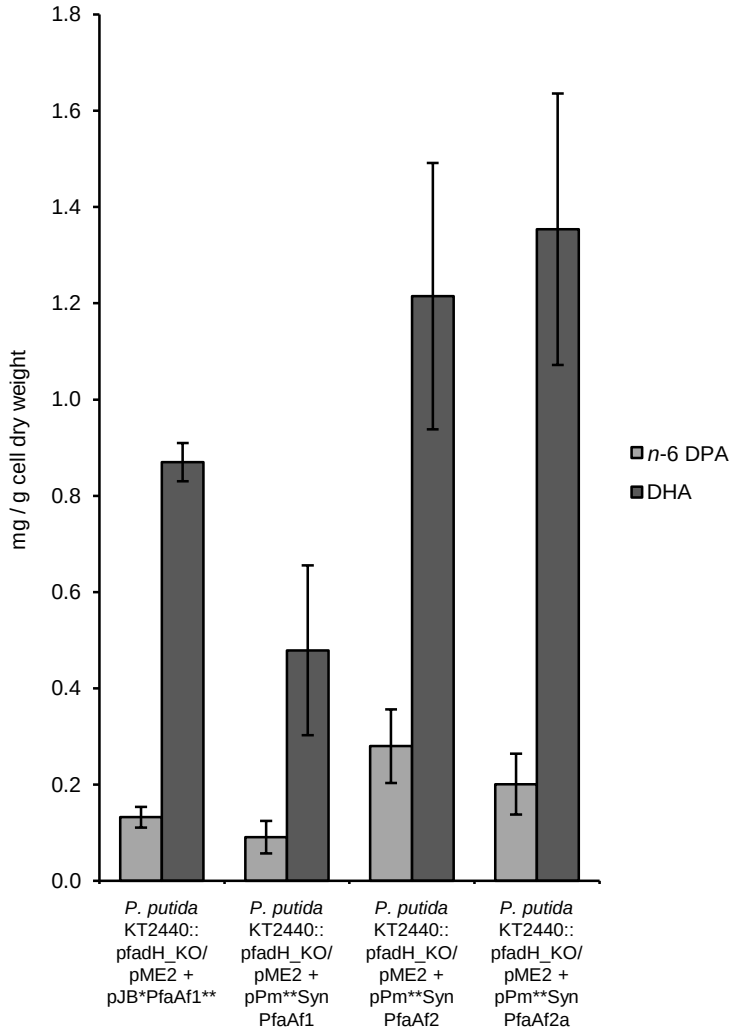


Figure 4. Recombinant production of *n*-6 DPA and DHA by metabolically engineered *Pseudomonas putida* KT2440 with native or synthetic *pfa* gene clusters originating from *Aetherobacter fasciculatus* (SBSr002). Initial cultivation was carried out in 50 ml LB-medium at 30 °C for 4 h. After induction of gene expression, the cultivation was continued at 16 °C and 200 rpm for 24 h. The indicated values are means and standard deviations of three biological samples.

REFERENCES

[1] Kodumal SJ, Patel KG, Reid R, Menzella HG, Welch M, & Santi DV (2004) Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl. Acad. Sci. U S A*, **101,** 15573-15578.

[2] Menzella HG, Reisinger SJ, Welch M, Kealey JT, Kennedy J, Reid R, Tran CQ, & Santi DV (2006) Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster. *J. Ind. Microbiol. Biotechnol.*, **33,** 22-28.

[3] Mutka SC, Carney JR, Liu Y, & Kennedy J (2006) Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry*, **45,** 1321-1330.

[4] Oßwald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, & Wenzel SC (2012) Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.* Epub ahead of print.

[5] Gao X, Wang P, & Tang Y (2010) Engineered polyketide biosynthesis and biocatalysis in *Escherichia coli*. *Appl. Microbiol. Biotechnol.*, **88,** 1233-1242.

[6] Gemperlein K, Rachid S, Garcia R, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

[7] Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, Martins Dos Santos VA, Fouts DE, Gill SR, Pop M, Holmes M, Brinkac L, Beanan M, DeBoy RT, Daugherty S, Kolonay J, Madupu R, Nelson W, White O, Peterson J, Khouri H, Hance I, Chris Lee P, Holtzapple E, Scanlan D, Tran K, Moazzez A, Utterback T, Rizzo M, Lee K, Kosack D, Moestl D, Wedler H, Lauber J, Stjepandic D, Hoheisel J, Straetz M, Heim S, Kiewitz C, Eisen J, Timmis KN, Dusterhoft A, Tummler B, & Fraser CM (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.*, **4,** 799-808.

[8] Grant SG, Jessee J, Bloom FR, & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U S A*, **87,** 4645-4649.

[9] Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc. Natl. Acad. Sci. U S A*, **76,** 1648-1652.

[10] Bagdasarian M, Lurz R, Ruckert B, Franklin FC, Bagdasarian MM, Frey J, & Timmis KN (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene*, **16,** 237-247.

[11] Sambrook J & Russell DW (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[12] Hill DS, Stein JI, Torkewitz NR, Morse AM, Howell CR, Pachlatko JP, Becker JO, & Ligon JM (1994) Cloning of genes involved in the synthesis of pyrrolnitrin from *Pseudomonas fluorescens* and role of pyrrolnitrin synthesis in biological control of plant disease. *Appl. Environ. Microbiol.*, **60,** 78-85.

[13] Bode HB, Ring MW, Schwär G, Kroppenstedt RM, Kaiser D, & Müller R (2006) 3-Hydroxy-3-methylglutaryl-coenzyme A (CoA) synthase is involved in the biosynthesis of isovaleryl-CoA in the myxobacterium *Myxococcus xanthus* during fruiting body formation. *J. Bacteriol.*, **188,** 6524-6528.

[14] Goldberg DE (1989) *Genetic algorithms in search, optimization and machine learning.* Addison-Wesley Longman Publishing Co., Inc. Boston, MA.

[15] Bäck T & Schwefel HP (1993) An overview of evolutionary algorithms for parameter optimization. *Evol. Comput.*, **1,** 1-23.

[16] Berg L, Lale R, Bakke I, Burroughs N, & Valla S (2009) The expression of recombinant genes in *Escherichia coli* can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5'-untranslated part of mRNA. *Microb. Biotechnol.*, **2,** 379-389.

[17] Bakke I, Berg L, Aune TE, Brautaset T, Sletta H, Tondervik A, & Valla S (2009) Random mutagenesis of the *Pm* promoter as a powerful strategy for improvement of recombinant-gene expression. *Appl. Environ. Microbiol.*, **75,** 2002-2011.

[18] Sharp PM & Li WH (1987) The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.*, **15,** 1281-1295.

[19] Pedersen M, Nissen S, Mitarai N, Lo SS, Sneppen K, & Pedersen S (2011) The functional half-life of an mRNA depends on the ribosome spacing in an early coding region. *J. Mol. Biol.*, **407,** 35-44.

[20] Clarke TF & Clark PL (2008) Rare codons cluster. *PLOS One*, **3,** e3412.

[21] Li GW, Oh E, & Weissman JS (2012) The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature*, **484,** 538-541.

[22] Zhang G, Hubalewska M, & Ignatova Z (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.*, **16,** 274-280.

[23] Itzkovitz S & Alon U (2007) The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Res.*, **17,** 405-412.

[24] Gruber AR, Lorenz R, Bernhart SH, Neubock R, & Hofacker IL (2008) The Vienna RNA websuite. *Nucleic Acids Res.*, **36,** W70-W74.

[25] Blatny JM, Brautaset T, Winther-Larsen HC, Karunakaran P, & Valla S (1997) Improved broad-host-range RK2 vectors useful for high and low regulated gene expression levels in Gram-negative bacteria. *Plasmid*, **38,** 35-51.

APPENDIX

DNA sequences of the synthetic building blocks

>BB1_V1

ATTTAAATATGTCTGCAATCGGTAGATGGAACGCTCAAGGGGGCTCGCCGGAGCTCACCCCTGCGGCGATGGTCG
CGGCGATAGCGCGTTGTCGAGAGTCGTTGCACGTGGTGCGCGACGGGGAGTCCGGGCGGATCGGCCTGGGCGTGG
GGGGCGAGCTGCTCCCCCCCGGCGAGCGCCGCGCGGGCGAGTACACGCTTGTGGCTAGCCTGCCGCCGCTGTTCC
CCGAGTGGCTCGGTGACCGCTCGTTCTGCGAGGTACACGGGGTCCGCTTCCCCTACGTCGCCGGCGAGATGGCCA
ACGGCATTGCCACCGTGGAGATGGTCGTGGCCATGAGCCGGAGCGGGATGCTCGGCTTCTTCGGCGCCGCGGGCC
TGGGCTTCTCCCGGGTCGAGGCCGCGGTGGAGACGCTCCGCGCCACGCTGGGGGACGAGGCGCCGTGGGGGGTGA
ACCTCATCCACTCACCGAACGAGGCCTCGCTGGAGAACCGCGTCGCCGATCTGCTGATCGAGAGCGGGGTGCGGC
GCATCTCGGTGTCGGCCTTCATGGCGCTGACGCCGGCGGTGGTGCGCTGCGCGGCGCGCGGGCTGCGGCTCGACG
CGGCGGGGCGGATCGTGCGGCCGCGGCACATCTTCGCCAAGCTCTCGCGGCGCGAGGTGGCGCTGCCGTTCCTCT
CCCCCGCGCCGGAGGAGATGCTGCGCGCCCTCGTCGCCAGCGGCGGGTTGACGGAGGAAGAGGCCCGCCTCGCCG
CGAAGATCCCGGTGGCCGAGGACATCACCGTCGAGGCCGACTCCGGCGGTCACACCGACAACCAGGCGCTGGTGG
CGCTGCTCCCGACGATCTTGCAGCTCCGCGACGAGCTGGCGGACAAGTACGACTACACGCGGCCGATCCGCGTCG
GCGCGGCGGGTGGGCTCGGGACGCCGACGGCGGTCGCCGCGGCGTTCTCGCTCGGCGCGGCCTACGTCGTCACCG
GCTCGGTCAACCAGGCGGCGCTGGAGTCCGGCCTCTCCGACGAGGGCAAGCGGATGCTGGCGCAGGCCGAGGTCG
CCGACGTGATCATGGCGCCCGCGGCCGATATGTTCGAGCTCGGGGTCAAGCTCCAGGTGCTGCGGCGCGGGAGCA
TGTTCGGCGTGCGCGCGGCCAAGCTCTACGAGGCGTACCACGCGCACCCCTCGCTCGAGGCCATCCCCGACGCGC
TGCGGGCGCGGCTCGAGCGCGAGGTCTTGCACGCGACCTTCGACGGCATCTGGGCGGAGACGCGGGCCTTCTGGG
AGGCGCGCGATCCCGAGCAGGTCGCGCTGGCGGCGCGGGATCCCAAGCACCGGATGGCGCTGGTCTTCCGCTGGT
ACCTCGGCATGGCGAGCCGCTGGGCCATCGCCGGCGAGGCGTCGCGGCGGGCCGACTACCAAATCTGGTGCGGGC
CGGCGATGGGGGCCTTCAACTCCTGGGTCCGCGGCTCGTTCCTCGAAGATCCCAAGCAGCGCGGCGTGGTGGCGA
TCGCGCTCAACCTGCTCGAGGGCGCGGCCGTCGTCACCCGCGCCCACCAATTG

>BB2_V1

CAATTGCGGACCTACGGCGTGCCCGTGCCGGCCGCCGCGTTCGACTTTCGACCTCGTCCGCTCGCTTGAGCGCTC
GGCTTCACCTTCACTCGACAGGTCTTCGCATGACTCAGGTTCCGGTCGCCATCGTCGGCGTGGGCGCGCTCTTTC
CGGGCTCGCCCGACGGTGCAGGCTTCTGGCGCGATATCGTCGCGTCACGCGATCTCGTCTCCGACGTCCCCCCGG
GGCACTGGCTGATCGAGGACTACTACGATCCCGATCCCTCGGCGCCGGACAAGACCTACTCCAAGCGCGGGGCCT
TCCTCGGCGAGACGGCCTTCGATCCGCTGGAGTTCGGCGTGGTGCCGAGCTCGCTCCCGACCACCGACACCGCGC
AGCTGCTGGCGCTGATCGTGGCGCGGCGGGTGCTCGACGACGCGACCCAGGGGCGCTTCGCTTCCATCGATCGAT
CGCGCATCGGCGTCGTCCTCGGCGTGACCTCGGGGCAAGAGCTCCTCGGCACGATGGTGAGCCGGCTCCAGCGGC
CGGTGTGGCTCAAGGCGCTGCGCGACGACGGCGTGCCCGAGGCCGACGCGCAGCGGATCTGCGATCGGATCACCA
ACGAGTACGTGCCCTGGAAGGAGTCGACGTTCCCCGGGCTCCTCGGCAACGTGGTCGCCGGGCGCATCGCCAACC
GCTTCGATCTCGGCGGCACCAACTGCGTCACCGACGCGGCGTGCGCGAGCTCGCTCGCGGCGCTCTCGATCGCCA
TGAACGAGCTGGCGCTCGGGCAGTCGGATCTCGTGATCACCGGCGGGGTGGACACGATGAACGACATCCTGATGT
ATACCTGCTTCAGCAAGACCCCGGCGCTCTCGAAGTCGGGCGACTGCCGCCCCTTCTCCGACGCCGCCGACGGGA
CGCTGCTCGGCGAGGGCCTGGCGATGTTCGCGCTCAAGCGCCTCACCGACGCCGAGCGCGACGGGGATCGCGTCT
ACGCGGTGATCCGCGGCCTGGGATCGTCGTCGGACGGGCGGAGCAAGAGCGTGTACGCGCCGGTGCCGGAAGGGC
AGGCCAAGGCGCTGGTGCGGGCCTACGAGGCGGCCGGCTACGGCGCGGACACGGTGGAGCTGGTCGAGGCGCACG
GCACCGCGACCAAGGCCGGCGACGCGGCGGAGTTCGGCGGCCTCAAGCTGGCGTTCGAGGCGAGCGGGCGCGCGG
ATCACCAGTGGTGCGCGCTCGGATCGGTGAAATCGCAGATCGGGCACACCAAGTCGGCGGCCGGCGCGGCGGGCC
TCTTCAAGGCGGTGATGGCGCTGCACCACAAGGTCCTGCCGCCGACGATCAAGGTCGATCGCCCCAACCCCGAGC
TGGCGATCGAGGCGAGCCCGTTCTACCTCAACACGCAGGCGCGGCCCTGGGTACGCAACGCGCAGCACCTGCGGC
GGGCCTCGGTCTCGAGCTTCGGCTTCGGCGGGAGCAACTTCCACGTAACGCTGGAGGAGTACGCCGGCACGGCGG
AGCCGGCGGCGCGCCTGCGGACCGCGCCGACGGAGCTCTTCGTGATCTCGGCGCCCTCGGCGGCGGAGGTCCTCG
CGCGCGCCGGCAAGCTCGCCGAGCTCGCGGACAAAGAGGGCACGACGACGGCGGCGCGGACCTCGCAGCGCGACT
TCGTGACGGCGGATCACGCGCGGCTCTCCATCGTGGCGAGCGACGCCGCGGATCTGCGCGCCAAGCTCGAGCAAG
CGCGCTCGCTCCTGTCGGCGCGCCCCGAGGAGGCGTTCTCGACGCCGACCGGCGTCCACTTCGCGGTGGGCGCGG
CGGCGGGCGGCGTGGCGTTCCTCTTCCCCGGGCAGGGGAGCCAGTACGTGGGCATGGGCGCCGACGTGGCCATGG
AGCACGGCGTCGCGCTCGACGCCTGGGACGCGGCGGCGGGCCTGGTCGAGACCGCGGGGCTGCGCGACGTGGTCT
TCCCGCGCCCGGCCTTCACCGACGAGGCGCGAGCCGAGCAGGCGCAGCTCTTGACGGCGACCGAGTGGGCGCAGC

CGGCGATCGCCACGGCGAGCCTCGCGCTCCTCGCGGTGCTGCGGGCGGTCGGCGTCCGCCCGTCGAGCGTCGCCG
GTCACAGCCTGGGCGAGCTCACCGCGCTCGCCGCGGCCGATGCTTTGAGCACCAACGATCTCGTGGCCACGGCGC
GCAAGCGCGGCCAGCTCATGAACGCGGCGTCGAGCACGCCGGGGGCGATGACCGCGGTGGCGCGGCCCCTCGCCG
AGGTGCAGGCCCTCCTCGAGCGCCTCCAGAGCGAGGTGGTGGTGGCCAACCACAACCACCCGACCCAGGTCGTCC
TCGCCGGCGAGACCGCCGCGATCGAGGCGGTCGAGGCCGAGCTCGCGCGCCAGGGGATCACCGCCAAGCGGCTCC
CGGTCGCCACGGCGTTCCACAGCCAGGTCGTGGCCGGGGCGACGGCGCCCTTCGCCGAGCACCTCGCCGGGATCG
CGGTGAAGGCGCCGGCCATCGACGTGCTCTCCAACAGCGAGGCCGCGGCCTACCCGTCGGATCCCGACGCGGTGC
GCGCGCTCCTCGCCGGGCAGATCACGCGCTCGGTCCGCTTCGTCGAGGTGGTCGAGGCGATGTGGGCGCGCGGGG
CGCGCATCTTCGTGGAGGTCGGCCCCTCGGCGGTGCTGACCGAGCTCTGCGCGCGGATCCTCGGCGAGCGGCCGC
ACCTGGCGATCGCCATGGATCGCAAGGGCAAGCACGGCGTGACCGCGCTCCAGGACGCGCTCGGGCGCCTCGCCG
CGGCCGGCGTGGCCATCGACTTCGCCCCGCTGTGGGCGCCGTTCGGCCCGCCCGCGGCGCCGAAACGCGT

>BB3_V1

ACGCGTGCCGGCGATGTCGATGATGATCAACGGCGCGAACTACGGGAAGGTCTACCCGCCCAAGGGCGGCGCCGC
GGCGCTGCCGCCGCCGAACCCGCCGCGCCCGGTCGTTGCTGCTGCAACGATCGCCGCGGCGCCGAGCTCGGCGGC
CAAGGCTCCGGTCGTTGCTGCTGCAACGATCGCGGCGGCGCCGATCCCGGCGGCCAAGGCTCCGGTCGTTGCTGC
TGCAACGATCGCGGCGCGCCCCGCGGCGGCGCCCGAGTCGCGGCCGGTCGTCGCGGCTGCCCCGGCCCCTGCGCC
GGTTGTGGCGGCGCCCGCGTCGGCGGCGGCGGCGCTGTCGGCGCCGGGTCAGCTCGGCTGGATCGAGGCGTACCA
GGAGTCGCAGCGGCAGCTCACCGAGGCGCACGCGAGCTACGTGCGGACGATGGCGTCGGGGCACGAGGCGTTCCT
CACGACGATGGAGGCCTCGTTCATCGGGCTCGGCGGCCTGCTCGGCGCGCCGGCTTCGGTGTCCGCGGTCGCCCC
GCGGAGCGCAGCGGCTCCGGTGCTCGCGCAGCCGGTCGCGGTCGCTCCGGCTCCGGTGCTCGCGCCGGCTCCGGT
GCCCGCGCCCGCTCCGACTCCGGTGCTCGCTCCGGCTCCCGCTCCAGTACTGGCGCCTGCGCCTGTCGCCGCGGT
GGCGAAGGCGGCTCCCGCGCCGGTGATGGCGCCCGCGGCGCCGTCAATCGATCTCCAGGCGCTGCTGCTCGACGT
GGTCGCGGACAAGACCGGCTACCCCGCCGAGATGCTCGGTATGCACATGGAGCTCGAGGCCGATCTCGGCATCGA
CTCCATCAAGCGCGTCGAGATCCTCTCCGCCATGCGGAAGCGCGCGCCCAACCTCCCGGAGGTCGTCGCCTCCGA
GATGGCCACGCTCCGCACCCTCGGGCAGATCGTCGAGTACATGCGCGAGCGGTCCGGTGCGCTGACGGCTGCGAC
TCACGAGACGCTGGTTCCCGCGACGGCAGAGGCTGCGACCCAGGCGGCTCCGGCGCTCGGCGATCTCCAGGCGCT
GCTGCTCGACGTGGTCGCGGACAAGACCGGCTACCCCGCCGAGATGCTCGGGATGCACATGGAGCTCGAGGCCGA
TCTCGGCATCGACTCCATCAAGCGCGTCGAGATCCTCTCCGCCATGCGTAAACGCGCGCCCAACCTCCCGGAGGT
CGTCGCCTCCGAGATGGCCACGCTCCGCACGCTCGGGCAGATCGTCGAGTACATGCGCGACCGCTCCGGCGCGAC
GGCGGCTCCGGTGAAGGCAGCGCCTGCGCCTGCGCCTGCGCCTGCGGCGGCTCCGACGAAGGAAGCGCCGACGCT
GGCGGATCTGCAAACGCTGCTGCTCAGCGTCGTCGCGGACAAGACCGGCTACCCCGCCGAGATGCTCGGTATGCA
CATGGAGCTCGAGGCCGATCTCGGCATCGACTCCATCAAGCGCGTCGAGATCCTCTCCGCCATGCGTAAACGCGC
GCCCAACCTCCCGGAGGTCGTCGCCTCCGAGATGGCCACGCTCCGTACGCTCGGGCAGATCGTCGAGTACATGCG
CGACCGCTCCGGCGGGCACAGCGTGGCGGCCACGACACCGGCACCGGCTCCGGCGTCGACGCTCGCTCCGGCGGC
GAGCCGGATCGGTCGCTACGTGGTGGGCACGCGCGAGGAGGCGGCGAGCGGGCTGGCGATGGGCGGCGTGCTCGG
CGCGCGGCGCTTCGTCATCACCAACGAAGGGACGGGCATCGCGCAGGCGCTGGTGCGGGCGCTGCAAGAGCGCGG
CGTCAGCAGCGAGGTCGTGACCGAGGTGCCGGCGGACGCCGACGCGGTGGTGTTCCTCGGCGGGTTGCGGACCCT
CGTGGACGTCGACACGGCCATCGCGATCAACCGCGAGGCGTTCCTCGCGGCGCGCGCTGTGGCCAGTGTGTTCTC
GCGCGACGGCGGCGCCTTCGTGACGGTCCAGGACACGGGCGGCGATCTCGGGATCGAGGCGCGCGAGCCGCGGGC
AGCGTGGCTCGCGGGCGTGGCCGCGCTGGCGCGGACGGCGCAGCTCGAGTGGCCGAAGGCCTCGGTGAAGGCAGT
GGACGTCGAGCGCGGCGGGCGCAGCTCCGAGGAGATCGCCGCGGCGATCGCCGAGGAGATCT

>BB4_V1

AGATCTTGACCGGCGGCGGCGCGCTCGAGGTCGGCCTGACGGCGGCCGGCGCGCGGCGCGTCCCCGAGGTCGTGC
AGCGCGCGGTGGGCGAGCGGACCCCGGTCCTCGCTCGCAGCTCGGTCGTCGTCGCCTCCGGCGGGGCCCGCGGCG
TCACCGCCACGGCGCTGATCGCGCTGGCGCGCCACGCGCAGCCGAAGCTGGTGATCCTCGGGCGATCCGCGCTCG
AGGCGGAGCCGGCGGCCTGCCTCGGCGTGCGCGGCGACGCGGCGCTCAAGCGCGCCTTGCTCGAGGACGCGCGCG
CGCAGGGCCGGACGGCGACCATCGCCGAGATCGGCGCCCACGTAGCGCGCATCTCCGCGAGCCGCGAGGTGCTGG
AGACGCTCGCGGCGATCAAGGCCGCGGGCGGCGAGGCGCGCTACCTCGCAGTCGACGTCACCGACGCCGATCGCC
TCCGCCTCGCGCTGGAAGAGGTCCGGCGCGACTGGGGCCCGATCACCGCCGTCGTCCACGGCGCCGGCGTCCTCG
CCGACAAGCGGATCGAGGACAAAACGCCCGAGCAGTTCGATCGGGTCTTCGACACCAAGGTCCTCGGGCTGCGCG
CGCTGCTCGCCGCGACCGAGGGCGATCCGCTGCGCGCGATCATCCTCTTCTCCTCGGTCGCCGCGCGCGCGGGCA
ACGTCGGCCAGTGCGACTACGCCATGGCCAACGAGATCCTCAACAAGGTGGCCATCGCCGAGCGGCGGCGGCGCG
GCGGCGCTTGCACCGTCAAGTCCCTCGGCTGGGGCCCCTGGGAGGGCGGGATGGTGACGCCCGCGCTCAAGGGTC
ACTTCGAATCGATGGGCGTGGCGCTGATCCCGCTGGCCGAGGGCGCCGAGCACTTCATCGCCGAGATCGAGGGCT
CCGCCGACGAGATCGAGCTGGTGATCGGCGGCGGCGCCGACGGCAGCGCCATCCGCGGGGACGCCGAGCCGACGA
TCCACTTCGGGGGTCTCGGTCTCGGATCGCACGAGCCCGTGGCTCGCCGATCACCGCGTCGCTGGTCGCCGCGTCA

TCCCTGTCGCCGTCGTGATCGAGTGGTTCACCCGCGCGGTCACGGCGCTCCAGGGCCGCGCCGCGGTGAGCCTCC
GCGACGTCAAGGTGCTGCGCGGGATCAAGCTCCCCGCCGGCGACGCCGCGGCCTCGTTCTCGATCGGCTGCCGCA
AGCTCGCCGAGGGCAAGATCGAGCTCACGCTCCACGGGGCCGACGGCGCCCTGCACTACCGCGCGAGCGCCGAGC
CCCGGGGCGCGGCGCTGAAGCCCGAGGGCGCCGCGCCTCCGCTCGCGGCGTGGGGCGATCGGCCGATCTACGACG
GCGGGCTGCTCTTCCACGGGCCGCGCTTCCACGCGATCCACAGCATCGAGGGCGCCTCGGCCGAAGGGATCCGCG
CCACGCTCACCGGCGCGCGCGAGCTCGGCTGGCCGGGCGAAGCGCAGCGCACCGATCCGGCGATTGTCGACGGCG
CGCTCCAGCTCGCGCTGCTCTGGACGCGCGAGCGCCTCGGTCAGGCCTCGCTCCCGATGGCCATCGCGGCCTTCG
ACTGGCAGGCCTCGGGGCCGGTCGAAGGTCCAGTGCGCGCGGTCGTTCAAGGGCGCGAGGCCGACAGCGCTCATG
CGGTCTGCGACGTGTGGCTCGAGGACGCGCGCGGCGCGATCGTCGGCGCGCTCCGCGGCGTCGAGACCATCGTGC
GGCCGGCCGAGGTCCGCGCGCCGCACCTCGTGGAGACAGGGCCAGCCTCGTGACCTTCGAACCGATCGCGATCGT
CGGGCAGGCCTGCACGCTCCCCGGCGCGCTGACGCCGGAGGCGCTGTGGGCCAACGTT

>BB5_V1

AACGTTCTCCGCGGCGAGAGCGCCGTGACCTCGCCGCCGGAGGGCCGCTGGCGCCTCCCGGCGTGGCACGCGATG
GGGAAGCCGGGCGCGCAGGAAGAGCGCGCCTGGTCCGACGCCGGAGGCTATGTCCGCGGCTTCGAATCGGCCTTC
GATCCCGGCGGCCTGCGCCTCCCGGAGGCCGAGCTCCGCGGCCTCGATCCGGTGTTCCAGTGGACGCTCCACACC
GCGCGCGCGGCCCTGCGCTCCGCGGGGATGGAGGGGCCGTCGGCGCGCGCCGGGCTCGTCCTCGGCAACCTCTCG
TTCCCGACGGCGTCGATGTCGCGCTACGCCGAGAGCGTCTGGCTCGAGGCGCTGGGCAAGGAGCTCCTCGGCGGG
CGCGCCGCGGCGCAGGCCGGCGTCACCCACCCCGACGCGCACAACCGCTTCATGTCCGGGCTCCCCGCGCACCTC
GCCGCGCAGGCCCTCGGGCTCGGCGGGGGCGCCTTCGCGCTCGACGCCGCCTGCGCCTCGTCGCTCTACGCGATC
AAGCTCGCCTGCGACCGGCTCCAGGATCGCGCTACTGATCTGATGCTGGCCGGCGCCGTCAACGCCGCCGACAGC
CTCTTCCTCCACATCGGGTTCTCGGCCCTCTCGGCGATGAGCCGCACCGGCCAGTCGCGCCCGTTTCACCGCGAG
GCCGACGGCCTCGTCCCCTCGGAGGGGGCCGTCTTCTTCGCGCTCCAGCGCCTCGACGACGCGCGCGCCGCCGGC
CGCCGCGTCCTCGGGATCATCCGCGGCGTCGGCCTCTCCAACGACGGGCGCGGCCGCGGCCTCCTCGCGCCCTCG
GAGGAGGGACAGATCCGCGCCATGCGCCTGGCCTACGCCGCGGCGGGGCTCACGCCCTCCGACGTGAGCTTCGTC
GAGTGCCACGCCACCGGCACGCCCGTCGGTGACGCCACCGAGATTCGGAGCCTCAGCGAGGTCTTCGCCGGGTGC
CGCGACGTGCCCATCGGCTCGCTCAAGTCCAACCTCGGCCACCTCATCACCGCCGCGGGCGGCGCCGGCTTGATC
AAGGTCCTCGGATCGTTCACCGCCGAAACGATCGCGCCGACGATCCACGCCGGCGCTTCGATCCCCGAGCTCGCC
GGGTCGCCGGTCCGCCTCGTCACGCAGGCCGAGCCCTGGCGCGCGGCGGGCCGCCGCATCGCCGCGATCAGCGCC
TTCGGGTTCGGCGGCAACAACGCCCACCTGATCGTCGAGCAGGACGACGGCCGCGCCACGGCGCCGCGGGCCGCA
GCGCCGCGATCCGCGCCGGCGCCGATCGCCGTCATCAGCCTCGGCGCCAGGATCGGCGACACCACCGGCCGCGAC
GAGGCCGCGCGCGCGCTGCTCTCCGGCGCGGGCTGGGGCAGCCGCCGCGAGGCGGTCACGGTCGAGCTCGAGGGG
CTGCGCTTCCCGCCGCGCGATCTCGAGCAGACGCTCCCGCAGCAGCTCCTCGTGCTCGAAGCGGGGCGCGAGGCG
ATCCGCGGGCTCTCGCTCCCGCGCGATCGCACCGCCGTGCTCATCGGCATGGGCGCCGATCCCGAGGTCGCGCGC
TACGGCGCTCGCTGGCGCCTGCCCGCGTTCGCCGACGCCTGGGCGGCCGCCGGGCTCCCCGTGTCCAGCGCCTGG
ACCGAGGCCGCGCGCGACGCCGTGCAAGAGCAGCACGGCGCGGCCGGCGTCGTCGGCGCCATGCCCAACATCCCG
GCCAACCGCCTCAGCAGCCAGCTCGATCTCGCCGGGCCGAGCTTCACCATCTCCTCGGAGGAGCTCTCCGGCGTC
GTCGCCCTGCAGCTCGCCGCGCGCGCGCTGCGCTCGGGCGAGATCGACGCGGCCCTCGTCGGCGCCGTCGATCTC
GCCGATCAGATCGTCCACCGCAAGGCGCTGGGCGAGCTCGGCCTCGAGCCCCGCGCCGGCGACGCCGCGGTCGCG
CTCGTGCTCAAGCGCCTCACCGACGTGCGGAAGAGCGAGAAGGTCCTGGCGATCCTCGACGAGGGCATCGTGGAT
GCGTCCAGCGATCCCGCGCGCCTCCGCGTCGGCGACGGCGGCGCCGATCTCGATCCCTTCGCCACCACGGGCAAG
CCGCACGCGGCGGCGGGCCTCGTCCACGTCGCCGCCGCCGTGCTCGCGCTGCACCACGGCGCGCGCCCCGCCCCC
GGGCGCGCCTCCATGCCCTGGGTCGGCCCCCGCCACGCCGAGGTGACGACGCAGGCCCTCGGCGATCAGGCGGCG
ACGATCCACCTCCGCGGCGGCGGCCTCGCCGCGCCGCTGCTGCTCGACGCCACGCCGCAGATCCATGTCTTCGCC
GGCGCCGATCGCGAGGCCGTGCTCCGCGCCCTCAGCGACGGGCGCGAGGGCCTCGACGGCCCGGCGCGCGCGGTG
ATCGTGGCCGGCTCGCGCGAGGAGCTCGAGGCGCGCCGCGCCGACGCCCTCCGCGCCGCCGAGCGCAACGCCCCC
ATGCCCGACGGCGTCTTCCTCCGCGATCGCCCCGTCGCGGGCGAGCTCGCCTTCGTCTTCGCAGGCGCCGCCGCG
GCGTACCCGAGCATGGGCCGTGCAC

>BB6_V1

GTGCACTGGCGCTCGCGCTCCCCGACGTGATGGCGCGCCTCGGCAACCGGCTCTCGACGCTGCCCGCGGTGATCG
ACTGGATCCACGGCGCGCCCCGCGAGCCGCGCCACCCGCTCGAGCAGCTCTGGGCGTCGTCGTTCCTCTGCCAGC
TCCACGCCGAGATCACCCGCGTCGACCTCGGGCTCCGCCCCAGCGCGACCATCGGCTACTCCTCCGGCGAGAGCA
ATGCGCTCTTCGCCATGGGCGCCTGGCGCGACCTCGACGCCATGATGCACGAGGCCTGGACCGAGCCGCTGCTGA
CCACGGAGCTCGTCGGCGACTACGCCGCGGCGCGCCGCCTCTGGGCGCGCTCCGGCGTCACCAGCCCCGCGGCGA
CGTGGGCCGCGTGGGTCATCGCCGCGCCCGTGGACGACGTGCGCCGCGCCCTCACGGGCGAGCCGCTCGCGCACC

TCACCATCATCAACACCGCCGAGGAGTGCGTGATCGGCGGCGAGGCCTCCGCGTGCGAACGTGTCGTCGCGCGGC
TCGGCAAGAACCTCGCGCTGCCGCTCAGCTACGCCATGGCCTCGCACTGCCCCGAGGTCGCCGAGGCGCGCGAGG
CGTGGATCAAGATGCACCGCCGCGCCACCTTCGACGTCCCCGGCGTTCGCATCTACTCCGCCGGCGCCGAGCAGG
CCTTCGCCCCCACCGCCGACGCCGCCGCGATGGCCATCACCGGCCAGGCCGTCGACACGCTCGATTTCCCCCGCG
TGATCGAGCGCGCCTACGCCGACGGCGTCCGCGTCTTCGTCGAGCACGGCCCGCGCGGCCTCTGCAGCCGCTGGA
TCGCGCGCATCCTCGGCGCGCGCGAGCACCTCGTCGTCCCGCTCGACGTCGCCGGGCAGCAGCCGGTCCGCCAGC
TCGTGAGCGCCGTGGCCCAGCTCATCGCCGCCGGCGTGCCCGTCGCCGCCGACGCCCTCCTCGCCAGGCTCGCGC
CGCCTCCCCCGGCGCGCGCCTCGGGCCCCACGCTCACCGTGCGTGCGCACCCTCCTGACATTCACTTCCCGACGG
CCCCTGCCATGCAACTGATGCCGCCCGCGCCCCTTCTCCCGCCGGTGCTCGCCGACAACCCCGCGCCGCGCACGG
AGCGGAGCTTCACCGGGCGGTCGGCGCTGCGAGAGGACGCGCCCGACGGCGTCCCCTCGCCTCGCCTCCCCGTGT
TCGCGAGGCTCGCGCCGGTGTTCACCCACACCGAGCCGGTGTTCACGCACGTCGAGCCGGCGCTCGCGAGCCCTG
AGCCGGCGGTCACGCACCCCGAGCCGGTGATCGCGCACGCCGAGCCGGCGCTCACACACGCCGAGCCCGCGCTCG
CGAGCCCCGCGGATGGCCAGCACGCCTCCATCGCCGCGCGCGGCGCCGCCGTGAGCGCGGGGATCGCCGCCGCCC
ATCGCGCCTACGTCGAGGCGCAAGGCGCCGTTCACGCCCGCTTCCTCGCGCTCTCGGGCTCCTTGCTCTCGCGCC
TCGCCGCCGCCCCGACAGCCGCCGCCTTCCCGGCGCCTGCTCCCTCGCTCGCGCCGCAGCCCGCCGCCCTCCCGG
CGCCCGCTCCGCAGGCCGCCCCGCCGCAGCGCCTCACCGCCCCCACCCTCGTGGTCGCTCGGCCGGTCGCGCCGC
CCGCCCCGACGCCGGTGCCCGCCGCCGTGGCCGCGTTCGCTCCCGCGCCGATCTCGGCTCCGGCCTCCACCCTGC
CCGCGCCGTCCCTGCTCCGCTCCACGCTCCCCGGGCCCAAGTTCTCGCGCGCGGATCTCGAGGTCCTGGCCTCGG
GCTCGATCTCCTCGATCTTCGGCCCCCAGTTCGCCGGCCAGGATGGCTATGCCAAACAATGTCGGATGCCGGAGC
CCCCGCTCCTGCTCGCCGATCGCGTCACTGGCATCGACGCTGTCCCCGGCTCGATGGACACGGGCACCATCTGGA
CCGAGACCGACGTCAAGCACGACAGCTGGTATCTCCACCAGGGCCGGATGCCGACTGGCATCGTGATCGAGTCCG
GCCAGGCGGATCTCCTCCTGATCTCCTGGCTCGGTATCGACGCCCTCAACAAGGGGGAGCGCGTCTATCGCCTCC
TCGGCTGCGAGGCCACCTATCACGGCGAGCTGCCAAAGCCGGGCGACACGCTCGTCTACGAGATCCACATCGACG
GCCACGCCGAGCAGGGCGACGTGCGGCTGTTCTTCTTCCACTACGACTGCCGCATCAACGGCGAGCTGCGGATCT
CGGTACGTA

>BB7_V1

TACGTAATGGCCAGGCCGGGTTCTTCACCTATCAAGAGCTCGCCAACTCCGGCGGAGTGCTCTGGGACGCCGACG
GCGCGCCCCACGAGGCGACAGGCCCCCTCGATCCGCCTCAAGTGATCTGCGAGAAGCAGTCCTTCACCTCGGCCG
AGGTCCTCGCCTTCTCCGAGGGGGACGCCTACGCCTGCTTCGGGCGCGGCTTCGAGCGCGCGCAGACGCACGTTC
GCACGCCGCGCATCCAGGACGGGCCGATGCGGCTCCTCGATCGCGTCACCCACTTCGATCCTCGCGGCGGGCCCT
GGGGGCGCGGCTACCTGCGCGCCGAGGCCAAGCTCTCGCCCGACGACTGGTTCTTGCGAGGGCACTTCAAGAACG
ATCCCTGTATGCCAGGGACGCTGATCCTCGAGGCCTGCGTGCAGGCCATGGCCTTCTATCTCTGCGCCATGGGCT
ACTCCATCGATCGCGACGGCTGGCGCTTCGAGCCTGTCCCCGGCCAGAAGAGCCCGATGCGCTGCCGCGGCCAGG
CCAGCGTGACCTCGCGGGAGATCGTCTACGAAATCTTCGTCCACGAGGTCACGAGCGGGCCTATCCCCAGCGTCA
AGGTCGACATGCTCTGCACGGTCGACGGCCTCAAGGCCTTCCACGGCGAGCGCGTCGGCCTGCGCCTCGTCCCCG
CCTGGCCCCTCGACGACTGGCGACAGGCCCCCGTCGCCCCGCAGAGCCTGCACCGCCGCCCCAAGCAGATCGCCC
CTCTCGGCGGGATCCTCGGCCACGTCGAGACCCGGCCGGTGCCCTCGGTCAACGGCTTTGCCTTCGACTACGCCT
CGCTCATCGCCTGCGCCTGGGGCAAGCCCTCGCAGGCCTTCGGGCCGATGTACGCGCCCTTCGACGGCACTCGCC
GCGTCGCCCGCCTGCCCGGGCCGCCTTATCACTTCATCACGCGCATCATCTCGGTCGAGGGCCCGATCGGCGTCC
CCCAGCCCGGCGCCGAGGTCGAGGTCGAGTACGACGTCCCCCACGACGCCTGGTATTTCGCGGAGAACGGCGCGC
GCACCATGCCGTTCTGCGTCCTGCTCGAGGCCGCGCTCCAGCCCTGCGGCTGGCTGGCGAGCTACATCGGCAGCG
CCCTCACCACCGAGGACGATCTGCTCTTCCGCAACCTCGACGGGACTGGCACCGTCCTCGCGGAGATCACCCCCG
AGTCCGGCACGCTGCGGACCCACGTCAAGGTGGTCGAGGTCTCGCAGTCCGCGGGGATGATCATCGAGACCTTCG
AAGTCGAGTGCTTCGTCGGCGAAGCCCACGTTTACACGCTGCGTACCGTCTTTGGCTTCTTCCCCAAATCGGCCT
TCGACAACCAGGTCGGACTCCCCGTCTCCGCCGAGGCCCGCGCCCGGATCGAGCGTCAAGCCACCTCGACCATCG
ATCTGGCTCCAGAGCCGCCGCGCTACTTCGCCGGCGCCCCTCGCCTCCCCGAAGCGATGCTGTGCATGTTGGATC
GCATCACCGCGTTCGAGCCGACCGGCGGCCTCAAGGGCCTTGGCTATGGCCGCGCCGAGAAGACGGTCAAACCCG
GGGAGTGGTCCTTCAAAGCCCACTTCTTCCAGGATCCCGTCCAGCCGGGCTCCCTCGGCCTCGAGGCCCTGCTCC
AGCTCCTTCAGGCGTGGATGATCGACGCCGGTGACGCCGAGGGTATCCCGGGCGCGCGCTTCGAACCCATCGCCG
TCGGCCGCCCGATGACCTGGAAATACCGCGGCCAGGTCGTCCCCAAGAACCTCCTCGTGACGGTGGAGATCGACA
TCACCGAGCGCGGCACCGACGATCGCGGGCGCTTCGCCCTGGCCGAGGGCTCGCTCTGGGTCGACGGCAAGCGCA
TCTACGGCGCGAAGAACCTCGGGATGCGCGTCGTCGCCGGTCCGGAACGCCCGCCGCCCCGGCTCGCCGCCTTCGC
CCGCTTCGCCGACCGCGCCGACCGCGCCCCCTTCGCCGCCTGCCCCGACGGTGCCGGCTGCGCTGGCTGCGCCGC
CTTCGCGCGATCTCGCCTCGTTTCGAAGCTTCTGGCGCGCCTCGATCGGGGTCGGGCCCTGGCCGATCGAGGACC
TCTACATCGCCCTGGTCGAGCGCTTCCTCGGCCGGATCGTCATCGCCGATCCGGTCGCCTTCTCGCGCGTCCACG
GCCGCGGCTGCCTCTACGTCGCCAACCACCAGGTCGCGCTGGAGTCGCTGATCTTCATCATGGTCGCCTCGGCGA
TCTCCGGCACGCGGACGATGGCGCTCGCCAAGGCCGAACATCAGGAGTCGTGGCTCGGGCGCCTGATCACCCGCG
GCATCGCCTACCCCGGCGTCACCGATCCAGGCCTCATCACCTTCTTCGACCGCGAGGATCACGCCTCCTTCGCCC

GCATCGCCGAGCAGGCCACGCGCGATCTCGCGGCCTCGACCAAGAGCGTGCTCGTCCACGTCGAGGGCACGCGCG
CCATCTCGTGCCGCACCCCCGTCACCCGCCTCGGCGCCTCGGTGATCGACATGGCGCTGGCCGCCGGCGCGCCCA
TCGTCCCGGTCCGCTTCGCGCGCGGCCTGCCCGAGAGCGACTCCCCCGAGCGCCTCGATTTCCCGCTCGGCTACG
GCCGCCAGGACATCTACTTCGGCGCCCCGATCGAGCCCGCCGAGCTCGCGCGCCTCCCGCTCAAGGACCGCAAGG
ACGCCTTGCTCACGGCCCTCAACACGCTCGGCCCCGGCGCGAGCGACACGCCCTCCGACCCCGACCCGAGCTTCG
CCGCCCAGGTCGAGGCCTGGCGAGCCCGCACCGGCTGCGCCATCGAAGATGCCGTGTTCTATGCCACCCTCGCCG
GAGCGTCAGCCCCGGAGTCGGAGGGCGTGCAGCGGTTACTGGAGGGCGCGAGAACGGGGAAGCTGGTTATCGGGA
GTGGGGAGAAGGATCGCTGGCTGGGGGAGTTTGCGCGGAGCCTGTTTGGATCGGCGGATGCGACGCTCTTGGTGG
AGGGCGGCTGAGAGCGTGTTTACGGCGCGAATACGTGCCAGGTTCGGGACCCTCACCCCAACCCCTCTCCCGCGG
GGAGAGGGGCTCGGATCTGCTGCGAGGATCTTCCCCCTCTACCCTCGGGAGAGGGGGACCGAGGGGGTGAGGGAG
CCACGTAACGGCCTGTCAAATACGCGCTAAGGCCCCTTGTGCCCATTGAGGCACACGTCGAACGATTTCACCGTC
TTTAATTAA

>BB1_V2

ATTTAAATATGTCTGCAATCGGTAGATGGAACGCTCAAGGGGGCTCGCCGGAGCTCACCCCTGCGGCGATGGTCG
CGGCGATAGCGCGTTGTCGAGAGTCGTTGCACGTGGTGCGCGACGGGGAGTCCGGGCGGATCGGCCTGGGCGTGG
GGGGCGAGCTGCTCCCCCCCGGCGAGCGCCGCGCGGGCGAGTACACGCTTGTGGCTAGCCTGCCGCCCCTGTTCC
CAGAGTGGCTGGGCGACCGCAGCTTCTGTGAAGTGCATGGCGTCCGCTTCCCGTATGTTGCAGGCGAGATGGCTA
ACGGCATCGCGACCGTAGAGATGGTCGTGGCCATGAGCCGCAGCGGGATGCTCGGGTTCTTCGGTGCCGCCGGGC
TGGGCTTCAGCCGGGTCGAAGCCGCCGTCGAGACCCTGCGCGCCACCTTGGGTGATGAGGCGCCCTGGGGCGTCA
ATCTGATCCACAGCCCAAATGAGGCCTCGCTGGAAAACCGCGTCGCCGACCTGTTGATTGAGTCGGGCGTGCGTC
GTATTAGCGTGTCGGCCTTCATGGCGCTGACCCCGGCCGTTGTCCGCTGCGCCGCACGCGGGTTGCGCCTGGACG
CCGCAGGCCGCATCGTCCGCCCACGCCACATTTTCGCTAAGCTGAGCCGCCGTGAAGTCGCGCTGCCGTTCTTGA
GCCCCGCCCCTGAAGAGATGCTGCGTGCGCTGGTGGCCTCGGGTGGCTTGACCGAAGAAGAGGCGCGGCTGGCGG
CCAAGATTCCCGTCGCTGAAGACATCACCGTGGAAGCTGACAGCGGCGGCCATACGGATAACCAGGCCCTCGTAG
CGCTGCTGCCGACGATCCTGCAGCTGCGTGACGAGCTGGCCGATAAGTATGACTACACGCGGCCGATCCGCGTAG
GCGCTGCGGGTGGCCTGGGTACCCCGACCGCTGTGGCCGCCGCATTCTCCCTGGGCGCGGCCTACGTTGTGACCG
GTAGCGTGAACCAGGCGGCCCTCGAGTCGGGTCTGTCCGACGAAGGTAAACGGATGCTGGCCCAGGCCGAAGTCG
CAGATGTGATCATGGCGCCGGCAGCGGATATGTTTGAGCTGGGCGTGAAACTGCAGGTGTTGCGGCGCGGCTCCA
TGTTCGGCGTCCGTGCCGCCAAGCTGTATGAGGCTTACCATGCACACCCGTCCCTCGAAGCCATCCCGGACGCCC
TGCGCGCACGCCTCGAGCGTGAAGTGCTCCACGCAACCTTCGACGGGATTTGGGCCGAGACCCGGGCGTTCTGGG
AAGCCCGGGACCCTGAGCAGGTAGCGCTGGCCGCTCGCGATCCCAAGCACCGCATGGCGCTGGTGTTTCGTTGGT
ACTTGGGCATGGCTTCCCGCTGGGCGATCGCCGGTGAAGCCTCGCGCCGCGCAGATTACCAAATCTGGTGCGGGC
CGGCAATGGGCGCGTTTAACAGCTGGGTCCGTGGCTCGTTCCTCGAAGACCCTAAGCAGCGCGGCGTGGTAGCCA
TCGCGCTGAACTTGCTCGAAGGCGCCGCCGTTGTGACGCGGGCCCACCAATTG

>BB2_V2

CAATTGCGTACCTACGGCGTGCCAGTGCCGGCCGCCGCGTTTGACTTCCGCCCGCGCCCCCTGGCGTGAAAGACA
GAGTAGACAGAACACAGGGAGGTAATTAAATGACGCAGGTACCGGTAGCAATCGTAGGGGTAGGGGCCCTGTTCC
CGGGCAGCCCTGACGGCGCCGGCTTCTGGCGTGATATCGTGGCGAGCCGCGATCTGGTTTCGGACGTTCCCCCTG
GCCACTGGTTGATCGAAGATTACTACGACCCGGACCCTAGCGCGCCCGACAAAACCTACAGCAAGCGTGGTGCGT
TCCTCGGCGAGACGGCGTTCGACCCGCTCGAGTTCGGTGTGGTGCCGAGCTCGCTCCCGACCACGGACACCGCCC
AGCTCCTGGCGTTGATCGTGGCGCGTCGCGTTCTGGATGACGCCACCCAGGGTCGCTTCGCAAGCATTGACCGCT
CCCGCATCGGCGTAGTGCTGGGCGTCACGTCCGGCCAAGAGCTCCTGGGCACCATGGTCAGCCGCCTCCAACGCC
CGGTGTGGTTGAAGGCCCTCCGCGACGACGGCGTACCTGAGGCCGATGCCCAACGGATCTGCGACCGCATCACGA
ACGAGTACGTTCCGTGGAAAGAGTCCACCTTCCCAGGGCTGCTGGGCAACGTGGTAGCAGGCCGCATTGCCAACC
GCTTTGACTTGGGCGGCACGAACTGCGTCACGGACGCCGCGTGCGCCAGCAGCCTGGCAGCGTTGAGCATCGCCA
TGAATGAGCTGGCTTTGGGTCAGTCGGACCTGGTCATCACGGGCGGCGTCGACACGATGAACGACATCCTCATGT
ATACCTGCTTCTCCAAGACGCCAGCTCTGAGCAAGAGCGGTGATTGCCGGCCGTTCAGCGACGCCGCAGACGGCA
CCTTGCTGGGCGAAGGTTTGGCGATGTTCGCGTTGAAGCGCCTCACCGATGCCGAGCGTGACGGCGACCGCGTTT
ACGCGGTGATCCGCGGCCTCGGCTCGTCGTCCGACGGGCGTAGCAAGTCGGTCTATGCCCCCGTGCCCGAAGGCC
AGGCGAAAGCCCTCGTGCGCGCGTACGAAGCGGCCGGGTATGGCGCCGACACGGTCGAGCTGGTAGAAGCACACG
GCACCGCGACCAAGGCTGGTGACGCCGCAGAGTTCGGCGGCTTGAAACTGGCGTTCGAAGCCTCCGGGCGCGCAG
ACCACCAGTGGTGCGCTCTGGGTAGCGTGAAGTCGCAGATCGGGCACACGAAATCGGCCGCGGGCGCAGCTGGCC
TGTTCAAGGCTGTGATGGCTCTGCACCATAAGGTCCTGCCGCCGACGATCAAGGTCGATCGGCCGAACCCCGAGT
TGGCGATTGAGGCGTCGCCCTTCTACTTGAACACCCAAGCCCGCCCGTGGGTGCGTAACGCGCAGCATCTGCGGC

GCGCGAGCGTTTCGTCGTTTGGGTTCGGCGGCTCGAACTTCCATGTGACCCTCGAAGAGTACGCCGGGACCGCTG
AGCCGGCTGCACGCCTGCGTACCGCGCCCACCGAGCTGTTCGTAATCTCCGCCCCGAGCGCAGCCGAAGTGTTGG
CCCGCGCAGGTAAGTTGGCCGAGCTGGCTGACAAAGAAGGCACCACCACCGCAGCGCGTACCTCCCAGCGCGATT
TCGTCACCGCAGATCACGCGCGGCTGAGCATCGTTGCCAGCGACGCGGCCGACCTGCGTGCTAAGCTCGAGCAGG
CACGCAGCTTGCTGAGCGCGCGCCCCGAAGAGGCCTTTAGCACGCCCACCGGCGTGCATTTTGCCGTGGGTGCAG
CGGCCGGCGGTGTGGCCTTCCTCTTCCCCGGCCAGGGTTCGCAGTACGTCGGTATGGGCGCTGACGTAGCCATGG
AACACGGTGTCGCGCTGGATGCATGGGACGCGGCCGCCGGGCTGGTCGAAACCGCGGGCTTGCGTGACGTCGTGT
TCCCTCGCCCAGCCTTTACCGATGAGGCCCGCGCAGAGCAGGCACAGCTCCTGACCGCCACCGAGTGGGCTCAGC
CCGCCATCGCAACGGCGTCGCTGGCGCTGCTCGCGGTGCTGCGTGCCGTCGGCGTGCGTCCGTCGAGCGTCGCCG
GGCATAGCCTGGGCGAGCTCACGGCCCTGGCGGCTGCGGACGCGCTGTCCACCAACGATCTGGTGGCGACCGCGC
GCAAGCGTGGGCAGCTGATGAATGCGGCGTCGTCGACCCCTGGCGCGATGACCGCCGTGGCACGGCCGCTGGCTG
AAGTCCAGGCGCTGCTGGAACGGCTGCAGTCCGAAGTAGTCGTGGCAAACCACAATCACCCCACCCAAGTGGTGC
TCGCCGGTGAGACCGCCGCGATCGAAGCCGTCGAAGCAGAGCTCGCGCGCCAGGGTATCACGGCTAAGCGCCTGC
CAGTGGCAACCGCCTTCCATTCGCAGGTCGTCGCCGGTGCCACCGCGCCCTTCGCTGAGCACCTCGCCGGGATCG
CGGTGAAAGCCCCCGCGATCGACGTGTTGAGCAATAGCGAAGCCGCGGCGTACCCTAGCGACCCGGACGCTGTCC
GGGCCCTCCTGGCCGGGCAGATCACCCGCAGCGTTCGGTTCGTGGAAGTCGTTGAGGCGATGTGGGCTCGCGGCG
CCCGGATCTTCGTTGAAGTCGGGCCCAGCGCGGTGCTGACCGAGTTGTGCGCGCGCATTCTGGGCGAGCGGCCTC
ACCTGGCGATCGCTATGGACCGCAAAGGTAAACACGGTGTGACCGCGCTGCAGGACGCCCTCGGCCGTTTGGCCG
CAGCCGGCGTAGCCATCGATTTCGCACCGCTGTGGGCCCCATTCGGCCCTCCCGCCGCTCCGAAACGCGT

>BB3_V2

ACGCGTGCCGGCTATGTCCATGATGATTAACGGCGCTAACTACGGTAAGGTCTACCCGCCTAAGGGCGGTGCAGC
CGCCTTGCCGCCTCCGAACCCGCCGCGTCCTGTCGTGGCCGCTGCCACCATCGCTGCCGCCCCCAGCTCCGCCGC
GAAAGCCCCCGTGGTAGCCGCAGCCACCATCGCCGCTGCGCCGATCCCGGCGGCTAAGGCGCCGGTCGTCGCCGC
AGCCACCATTGCAGCCCGCCCGGCAGCCGCCCCTGAAAGCCGCCCGGTCGTAGCCGCCGCACCAGCCCCAGCCCC
GGTCGTGGCAGCCCCAGCCAGCGCAGCCGCAGCCTTGTCGGCCCCGGGCCAACTGGGCTGGATTGAGGCGTACCA
AGAGTCGCAACGCCAGCTGACGGAAGCCCACGCAAGCTATGTTCGGACCATGGCGTCGGGCCATGAGGCGTTTCT
GACCACGATGGAAGCCAGCTTCATTGGGCTGGGCGGCTTGCTGGGTGCCCCTGCCTCCGTGAGCGCAGTGGCCCC
TCGTAGCGCCGCCGCACCAGTGCTGGCCCAGCCGGTAGCCGTGGCTCCGGCCCCGGTTCTGGCACCGGCACCGGT
GCCAGCCCCAGCCCCGACCCCAGTGTTGGCCCCGGCCCCGGCACCAGTACTGGCCCCGGCACCGGTGGCAGCCGT
TGCCAAGGCCGCGCCGGCACCGGTAATGGCCCCCGCCGCCCCTAGCATCGATTTGCAGGCCCTGCTCCTGGACGT
CGTCGCGGATAAGACGGGCTACCCCGCAGAGATGCTCGGGATGCATATGGAACTGGAAGCCGATTTGGGCATTGA
CTCGATCAAGCGCGTTGAGATTCTGAGCGCAATGCGTAAGCGCGCGCCCAACTTGCCCGAAGTTGTGGCCTCCGA
GATGGCCACCCTGCGGACCCTCGGGCAAATCGTGGAATATATGCGCGAGCGCAGCGGTGCTTTGACCGCGGCCAC
CCATGAGACCCTCGTCCCGGCCACGGCCGAAGCGGCTACCCAGGCTGCCCCGGCGTTGGGCGACCTCCAGGCGCT
GCTGCTGGATGTAGTGGCAGACAAGACCGGCTATCCGGCTGAGATGCTCGGTATGCACATGGAATTGGAAGCAGA
CCTGGGTATCGATTCCATCAAGCGGGTCGAAATCTTGAGCGCCATGCGGAAGCGCGCTCCCAACCTGCCTGAAGT
GGTCGCGTCGGAAATGGCAACGCTGCGTACCTTGGGTCAGATCGTCGAGTACATGCGTGATCGTTCGGGCGCCAC
CGCCGCCCCGGTAAAGGCCGCGCCAGCCCCAGCACCAGCACCGGCCGCCGCACCGACCAAAGAAGCCCCGACCCT
CGCCGACCTGCAGACGTTGCTGCTGTCCGTGGTAGCTGACAAGACCGGGTACCCAGCGGAAATGTTGGGTATGCA
CATGGAACTGGAAGCGGATCTCGGCATCGACAGCATTAAACGTGTCGAGATCCTCTCGGCTATGCGTAAACGCGC
CCCGAACCTGCCGGAAGTGGTAGCAAGCGAGATGGCTACCCTCCGCACCCTCGGCCAGATTGTAGAGTATATGCG
GGACCGCAGCGGCGGCCATAGCGTGGCTGCCACCACCCCAGCCCCAGCGCCAGCGTCGACGCTGGCACCGGCGGC
GAGCCGCATTGGCCGCTACGTCGTCGGGACCCGCGAAGAGGCGGCCAGCGGGTTGGCAATGGGCGGCGTTTTGGG
CGCGCGCCGGTTCGTGATCACCAATGAAGGGACCGGCATCGCGCAGGCCTTGGTCCGGGCGCTGCAGGAACGTGG
CGTTTCGTCCGAAGTTGTGACCGAAGTGCCGGCTGACGCGGACGCGGTAGTGTTCCTGGGCGGTTTGCGTACCCT
GGTTGACGTCGACACCGCGATTGCTATCAACCGCGAAGCGTTCCTGGCAGCTCGCGCCGTCGCGTCGGTCTTCTC
CCGTGACGGCGGCGCATTCGTGACGGTTCAAGATACCGGCGGCGACCTGGGCATTGAGGCACGCGAGCCACGCGC
CGCATGGTTGGCCGGCGTTGCGGCCCTGGCCCGTACGGCACAGCTGGAATGGCCGAAAGCGAGCGTTAAGGCCGT
TGACGTAGAGCGTGGCGGTCGCTCGTCCGAAGAGATCGCGGCAGCCATCGCTGAAGAGATCT

>BB4_V2

AGATCTTGACGGGCGGCGGCGCGTTGGAAGTGGGCCTGACGGCAGCCGGCGCTCGTCGTGTGCCGGAAGTGGTTC
AGCGCGCGGTCGGTGAGCGGACCCCCGTCCTGGCCCGGTCCAGCGTGGTCGTCGCTTCGGGCGGTGCCCGCGGTG

TGACCGCGACGGCCTTGATCGCGCTCGCCCGCCACGCGCAGCCCAAGTTGGTCATCTTGGGCCGCAGCGCCCTCG
AAGCTGAGCCGGCGGCCTGTTTGGGCGTGCGTGGTGACGCCGCGCTGAAGCGGGCTCTGCTGGAAGATGCGCGCG
CGCAGGGTCGTACCGCCACCATTGCAGAGATCGGCGCACACGTCGCCCGTATCAGCGCTTCGCGTGAAGTGTTGG
AAACCCTCGCCGCTATCAAAGCGGCCGGCGGCGAAGCGCGCTACTTGGCTGTGGACGTGACCGATGCGGATCGTC
TGCGGTTGGCCCTCGAAGAGGTCCGGCGCGACTGGGGCCCGATCACGGCCGTAGTGCATGGCGCTGGCGTGCTCG
CGGATAAGCGGATTGAAGACAAGACGCCCGAGCAGTTCGACCGCGTTTTCGATACCAAAGTGCTCGGCTTGCGGG
CGCTGCTGGCGGCTACCGAAGGCGACCCCCTCCGCGCGATCATCCTGTTCAGCTCCGTAGCGGCCCGCGCGGGTA
ATGTCGGGCAGTGCGATTATGCGATGGCGAACGAGATTCTGAACAAGGTCGCGATCGCTGAGCGTCGCCGTCGCG
GTGGCGCTTGCACCGTGAAATCGTTGGGTTGGGGTCCCTGGGAAGGCGGTATGGTCACCCCCGCGCTGAAAGGGC
ACTTTGAGAGCATGGGCGTTGCCCTGATTCCACTGGCAGAAGGCGCAGAGCACTTTATTGCCGAGATCGAAGGCT
CCGCAGATGAGATTGAGCTGGTGATTGGCGGCGGCGCAGATGGGTCGGCGATCCGCGGTGATGCGGAACCCACCA
TCCACTTTGGTGTGAGCGTCTCCGACCGCACGTCGCCCTGGCTGGCTGACCATCGTGTCGCCGGGCGCCGTGTGA
TCCCGGTCGCTGTCGTCATCGAGTGGTTTACGCGGGCCGTGACCGCGTTGCAGGGTCGCGCCGCAGTGAGCTTGC
GTGACGTCAAGGTGCTCCGCGGCATCAAGCTCCCAGCCGGTGACGCGGCGGCCTCGTTCTCGATCGGGTGTCGCA
AGTTGGCTGAAGGTAAGATCGAGCTGACCCTGCATGGCGCTGACGGTGCATTGCACTACCGTGCAAGCGCGGAAC
CGCGCGGCGCCGCTCTGAAGCCCGAAGGTGCAGCCCCTCCCCTGGCGGCGTGGGGTGACCGGCCTATCTACGATG
GCGGTCTGCTGTTCCACGGGCCTCGCTTCCACGCAATCCACTCCATTGAAGGCGCTTCGGCTGAAGGCATTCGCG
CCACCTTGACCGGTGCCCGCGAGCTCGGTTGGCCCGGCGAAGCCCAACGCACGGACCCCGCTATCGTCGATGGCG
CGCTCCAGCTGGCGCTGCTCTGGACGCGCGAGCGGCTGGGCCAAGCCTCCCTGCCCATGGCTATCGCGGCCTTTG
ACTGGCAAGCGTCGGGCCCGGTAGAAGGGCCGGTTCGCGCCGTGGTCCAAGGGCGCGAAGCCGATAGCGCTCACG
CGGTCTGCGATGTGTGGCTCGAAGATGCCCGTGGTGCGATCGTAGGCGCCCTCCGCGGTGTCGAGACCATCGTCC
GGCCGGCCGAAGTCCGGGCTCCGCACCTCGTGGAGACAGGGCCAGCCTCGTGACGTTCGAACCCATTGCAATCGT
AGGCCAAGCTTGCACCCTGCCGGGCGCTCTGACCCCGGAAGCCCTGTGGGCCAACGTT

>BB5_V2

AACGTTCTGCGCGGTGAGAGCGCTGTGACCTCCCCTCCCGAAGGCCGTTGGCGCCTCCCGGCCTGGCACGCGATG
GGTAAGCCGGGCGCGCAAGAAGAGCGTGCCTGGTCCGATGCCGGCGGCTACGTCCGTGGGTTCGAATCCGCCTTC
GACCCAGGCGGCTTGCGCCTGCCGGAAGCTGAGTTGCGTGGCCTGGACCCGGTCTTCCAGTGGACCTTGCACACC
GCACGCGCCGCACTGCGCTCGGCTGGCATGGAAGGCCCGAGCGCACGGGCCGGTCTGGTCCTGGGTAACCTGAGC
TTCCCAACCGCGTCCATGTCGCGCTACGCGGAATCGGTGTGGCTCGAAGCCCTGGGTAAAGAGCTGCTGGGCGGT
CGCGCAGCCGCACAGGCTGGCGTGACCCACCCCGACGCCCATAACCGCTTTATGAGCGGTCTGCCGGCGCATTTG
GCCGCCCAGGCCCTGGGCTTGGGTGGCGGCGCCTTCGCTCTGGACGCTGCCTGCGCGAGCAGCCTCTACGCCATC
AAGCTCGCCTGTGACCGCCTGCAAGATCGCGCCACCGATCTGATGCTCGCGGGCGCCGTGAACGCAGCCGATAGC
CTGTTTCTCCACATCGGCTTCTCGGCCCTGAGCGCTATGAGCCGGACCGGGCAGTCGCGCCCCTTCCATCGCGAA
GCTGATGGCCTGGTGCCAAGCGAAGGCGCGGTGTTCTTCGCACTGCAGCGGCTCGATGACGCCCGTGCCGCCGGT
CGCCGCGTTCTGGGCATCATCCGCGGTGTGGGCCTGAGCAATGACGGTCGCGGCCGCGGGTTGCTGGCACCGAGC
GAAGAGGGCCAGATCCGGGCGATGCGCTTGGCGTACGCGGCCGCGGGCTTGACCCCGTCGGACGTTAGCTTCGTC
GAGTGCCATGCGACGGGCACCCCGGTTGGCGACGCGACGGAAATCCGCAGCCTGTCCGAAGTCTTCGCCGGGTGC
CGTGACGTCCCCATCGGCTCGCTGAAATCGAACCTCGGCCACCTCATCACCGCCGCGGGCGGTGCCGGCCTGATC
AAGGTCTTGGGTAGCTTTACCGCCGAGACCATTGCCCCAACCATCCACGCGGGTGCCTCGATCCCTGAGCTGGCC
GGGTCGCCCGTGCGCCTGGTAACCCAGGCCGAACCGTGGCGTGCTGCGGGCCGCCGCATTGCCGCAATCTCCGCG
TTTGGCTTCGGCGGCAACAACGCCCACCTGATCGTTGAGCAAGACGATGGCCGCGCAACCGCCCCACGCGCTGCA
GCTCCGCGTAGCGCCCCAGCCCCGATCGCCGTGATCTCCCTGGGCGCGCGCATCGGCGATACCACCGGCCGTGAT
GAAGCGGCCCGGGCACTGTTGAGCGGCGCGGGTTGGGGTAGCCGCCGTGAAGCCGTGACCGTCGAGCTCGAAGGC
CTGCGCTTCCCTCCGCGTGACCTCGAGCAGACGCTGCCGCAGCAGCTCCTGGTGCTCGAAGCAGGCCGCGAAGCG
ATCCGTGGCCTGAGCCTGCCCCGTGACCGCACCGCGGTCTTGATTGGCATGGGCGCGGACCCTGAAGTGGCACGC
TACGGTGCGCGCTGGCGCCTGCCGGCCTTTGCCGACGCCTGGGCAGCGGCCGGGTTGCCCGTGAGCAGCGCCTGG
ACCGAAGCTGCGCGTGACGCGGTCCAGGAACAGCATGGCGCTGCGGGCGTGGTCGGCGCGATGCCTAACATCCCC
GCCAACCGGCTGAGCTCCCAGCTGGATCTCGCTGGCCCGAGCTTTACCATCTCCAGCGAAGAGCTGAGCGGCGTC
GTGGCGCTGCAACTCGCCGCCCGCGCCTTGCGCTCCGGTGAGATCGACGCCGCCCTCGTAGGCGCCGTCGACCTG
GCTGACCAGATCGTCCACCGTAAGGCGCTGGGTGAGCTGGGCCTCGAGCCGCGCGCGGGTGATGCGGCTGTGGCC
CTGGTGTTGAAGCGCCTGACGGATGTGCGGAAGTCGGAAAAGGTGCTGGCGATCCTCGATGAAGGCATCGTGGAT
GCTTCGTCGGACCCGGCGCGCCTGCGCGTAGGTGATGGCGGCGCGGACCTGGACCCGTTCGCTACCACGGGTAAA
CCGCACGCCGCCGCCGGGCTGGTGCATGTAGCCGCCGCCGTGTTGGCCCTGCACCACGGCGCACGCCCAGCACCG
GGTCGTGCCTCGATGCCGTGGGTCGGGCCTCGCCACGCCGAAGTGACGACCCAGGCGCTGGGTGACCAGGCGGCC
ACGATCCATCTGCGTGGCGGTGGCTTGGCCGCCCCGCTGCTGTTGGACGCCACGCCGCAGATCCACGTTTTTGCC
GGTGCTGACCGCGAAGCCGTGCTCCGCGCGCTGAGCGACGGCCGCGAAGGCCTCGACGGCCCAGCGCGCGCAGTG
ATCGTGGCCGGTTCCCGCGAAGAGTTGGAAGCGCGCCGCGCTGATGCCCTGCGCGCAGCCGAGCGGAACGCCCCG

ATGCCTGACGGCGTCTTCCTCCGCGACCGGCCGGTAGCGGGCGAGCTGGCCTTTGTCTTCGCGGGCGCAGCCGCC
GCATACCCGAGCATGGGCCGTGCAC

>BB6_V2

GTGCACTGGCCCTGGCCCTGCCTGACGTGATGGCTCGGCTCGGCAACCGGCTGTCCACCCTGCCCGCGGTGATCG
ATTGGATCCATGGCGCGCCCCGTGAACCGCGGCATCCGCTCGAACAGCTGTGGGCCTCCTCGTTCCTGTGCCAGT
TGCACGCTGAGATCACGCGCGTAGACTTGGGCCTGCGCCCGAGCGCGACGATCGGGTACAGCAGCGGTGAGAGCA
ATGCGCTGTTCGCGATGGGCGCATGGCGTGATCTGGACGCGATGATGCATGAGGCATGGACCGAGCCGTTGCTCA
CCACCGAACTGGTTGGCGACTATGCAGCCGCCCGTCGTCTGTGGGCGCGTAGCGGCGTGACCAGCCCGGCTGCGA
CCTGGGCGGCATGGGTGATTGCGGCCCCCGTGGATGACGTGCGCCGCGCCTTGACCGGCGAGCCCTTGGCGCACT
TGACGATCATCAACACCGCTGAAGAGTGCGTGATTGGCGGTGAAGCCTCCGCCTGCGAGCGCGTCGTCGCGCGTT
TGGGTAAGAATCTGGCCCTGCCGCTGAGCTATGCTATGGCGAGCCACTGCCCGGAAGTAGCTGAGGCCCGCGAAG
CATGGATCAAAATGCACCGTCGGGCGACCTTCGATGTGCCCGGTGTCCGCATCTATTCGGCCGGCGCTGAGCAGG
CCTTTGCCCCCACCGCCGATGCGGCCGCGATGGCGATCACGGGCCAAGCCGTAGATACCCTGGATTTCCCGCGTG
TGATTGAGCGCGCGTACGCCGACGGTGTGCGTGTCTTCGTCGAGCACGGGCCACGCGGGCTCTGCAGCCGCTGGA
TCGCCCGGATCCTGGGCGCTCGTGAGCACCTGGTAGTGCCTCTGGACGTCGCGGGCCAGCAGCCAGTTCGCCAGC
TGGTTAGCGCGGTCGCGCAGCTGATCGCCGCTGGCGTGCCGGTCGCCGCTGACGCACTGCTGGCACGCCTGGCCC
CGCCTCCACCTGCCCGTGCCAGCGGTCCCACCCTGACCGTGCGTGCGCACCCGCCTGACATCCACTTCCCTACCG
CGCCCGCGATGCAGCTGATGCCCCCAGCGCCCTTGCTGCCGCCTGTCCTGGCCGATAACCCGGCACCGCGCACCG
AGCGCTCCTTCACCGGTCGGTCGGCCCTGCGTGAAGACGCGCCTGACGGCGTTCCTAGCCCGCGCTTGCCGGTCT
TCGCGCGGCTGGCCCCGGTGTTCACCCACACGGAACCGGTCTTCACGCATGTAGAGCCGGCCTTGGCGAGCCCGG
AACCCGCGGTCACGCATCCCGAGCCCGTGATTGCCCACGCCGAACCGGCCCTGACCCACGCTGAACCAGCACTCG
CTTCGCCGGCAGACGGCCAGCACGCAAGCATCGCCGCCCGTGGCGCCGCTGTAAGCGCCGGCATCGCAGCCGCCC
ATCGCGCGTACGTTGAGGCGCAAGGCGCCGTCCATGCGCGCTTCTTGGCCTTGAGCGGCTCGCTGCTGAGCCGTC
TGGCCGCTGCACCGACCGCCGCCGCATTCCCGGCCCCAGCACCTAGCCTGGCGCCCCAACCGGCAGCCCTGCCAG
CCCCAGCCCCACAGGCAGCCCCGCCACAGCGCTTGACCGCCCCGACCCTGGTCGTGGCACGTCCAGTGGCACCGC
CGGCCCCAACCCCGGTTCCGGCAGCCGTAGCCGCCTTCGCCCCGGCACCGATTAGCGCGCCGGCAAGCACGCTCC
CGGCACCGAGCCTGCTGCGCTCCACCCTCCCCGGCCCGAAGTTCTCCCGCGCGGACTTGGAAGTGCTGGCTTCGG
GCTCCATTAGCTCGATCTTCGGCCCACAGTTCGCGGGTCAAGATGGCTACGCGAAGCAGTGCCGGATGCCCGAGC
CGCCTCTGCTGCTGGCAGACCGGGTGACCGGGATTGACGCCGTGCCCGGTAGCATGGATACGGGTACCATCTGGA
CCGAGACGGACGTCAAGCATGACAGCTGGTATCTGCATCAAGGGCGTATGCCGACCGGGATCGTAATCGAGAGCG
GCCAGGCTGATCTCCTGCTGATCTCGTGGCTGGGTATCGACGCTCTGAACAAGGGTGAGCGCGTCTATCGCTTGC
TCGGGTGCGAAGCTACCTACCACGGCGAGCTGCCTAAGCCCGGCGACACGCTGGTGTATGAGATCCACATTGATG
GGCACGCAGAGCAAGGCGACGTGCGCTTGTTCTTCTTCCATTACGACTGTCGGATCAACGGCGAGCTCCGCATCA
GCGTACGTA

>BB7_V2

TACGTAATGGCCAGGCCGGGTTTTTCACCTACCAGGAACTGGCGAACTCGGGCGGTGTCCTCTGGGACGCAGACG
GCGCGCCGCATGAAGCAACCGGCCCCCTGGATCCCCCGCAAGTAATCTGCGAAAAGCAGTCCTTCACCTCGGCAG
AAGTGCTGGCTTTCAGCGAAGGTGATGCGTATGCCTGCTTCGGCCGCGGTTTTGAGCGGGCGCAGACGCACGTCC
GCACCCCGCGGATCCAAGATGGCCCGATGCGGCTGTTGGACCGCGTGACCCATTTCGATCCACGCGGTGGCCCCT
GGGGTCGTGGCTACCTCCGCGCTGAGGCCAAACTGAGCCCCGACGACTGGTTCCTCCGCGGGCACTTCAAGAATG
ACCCGTGTATGCCTGGCACCTTGATTCTGGAAGCGTGCGTGCAGGCGATGGCGTTCTATTTGTGCGCGATGGGCT
ATTCGATCGATCGCGACGGGTGGCGCTTTGAGCCCGTGCCTGGCCAGAAATCGCCCATGCGGTGCCGCGGCCAGG
CGAGCGTAACCTCCCGTGAGATCGTTTACGAGATTTTCGTCCACGAAGTCACCTCCGGGCCGATCCCGAGCGTGA
AAGTCGACATGCTCTGCACCGTTGACGGTTTGAAAGCGTTCCACGGTGAGCGCGTCGGTCTGCGCCTCGTGCCGG
CCTGGCCGCTGGATGACTGGCGCCAGGCTCCCGTAGCTCCGCAGTCGTTGCATCGTCGCCCGAAGCAAATCGCCC
CCTTGGGCGGCATCCTGGGCCACGTCGAGACGCGCCCCGTCCCTAGCGTGAACGGGTTTGCCTTTGACTACGCGA
GCCTCATCGCGTGCGCCTGGGGCAAGCCTAGCCAGGCGTTTGGGCCGATGTACGCCCCTTTCGACGGGACCCGCC
GGGTAGCCCGTTTGCCGGGCCCACCGTACCACTTCATCACCCGGATCATTAGCGTCGAAGGGCCTATCGGCGTCC
CCCAGCCCGGCGCAGAAGTCGAAGTAGAGTATGACGTGCCCCACGATGCGTGGTACTTCGCTGAGAACGGTGCGC
GCACGATGCCCTTTTGCGTCCTGCTGGAAGCGGCGCTCCAGCCGTGTGGCTGGCTGGCGTCGTACATCGGCTCGG
CGTTGACGACCGAAGATGACCTGTTGTTCCGTAATCTCGACGGCACCGGGACCGTGCTGGCCGAGATCACCCCAG
AATCCGGGACCCTCCGCACCCATGTGAAGGTCGTTGAAGTCTCGCAGTCCGCCGGGATGATTATCGAGACCTTCG
AAGTTGAGTGTTTCGTAGGCGAAGCGCACGTCTATACCCTGCGGACCGTGTTCGGGTTCTTCCCCAAGTCCGCGT
TCGACAATCAGGTCGGGCTGCCTGTGAGCGCTGAGGCGCGTGCCCGCATTGAACGCCAGGCCACCTCCACCATCG

ACTTGGCGCCTGAGCCACCGCGGTACTTCGCTGGCGCCCCGCGCTTGCCCGAAGCGATGCTGTGTATGCTGGATC
GTATCACCGCCTTTGAGCCGACGGGCGGTTTGAAGGGTCTGGGCTACGGGCGCGCCGAGAAGACGGTGAAACCAG
GTGAGTGGTCCTTCAAAGCCCACTTTTTCCAAGACCCGGTGCAACCGGGCTCCCTCGGCCTCGAAGCGCTGCTCC
AGCTGCTGCAGGCGTGGATGATTGACGCAGGTGACGCCGAAGGCATTCCCGGCGCGCGCTTTGAGCCAATCGCTG
TGGGCCGTCCCATGACCTGGAAGTACCGTGGCCAGGTTGTGCCTAAGAACCTCCTGGTGACGGTCGAGATTGACA
TTACCGAGCGTGGCACCGACGATCGCGGGCGGTTCGCGCTGGCCGAAGGCTCCTTGTGGGTCGACGGTAAACGCA
TCTACGGTGCGAAGAACCTGGGTATGCGCGTAGTTGCGGGTCCGGAACGCCCGCCACCGGGTAGCCCACCGAGCC
CTGCCTCGCCGACCGCCCCAACCGCCCCTCCAAGCCCGCCCGCTCCAACCGTGCCTGCCGCCCTGGCAGCTCCGC
CGTCGCGTGACCTGGCCTCGTTCCGGTCGTTCTGGCGCGCCTCGATTGGCGTAGGGCCCTGGCCGATTGAAGACC
TCTACATCGCATTGGTCGAGCGCTTCCTGGGTCGCATCGTGATCGCAGACCCCGTGGCATTCAGCCGGGTCCACG
GGCGCGGGTGCCTGTATGTGGCTAACCATCAGGTCGCCCTCGAGAGCTTGATCTTCATTATGGTCGCCTCGGCGA
TCTCCGGCACCCGCACGATGGCACTCGCCAAGGCTGAGCACCAAGAGTCCTGGCTGGGCCGGCTGATTACCCGCG
GGATCGCCTACCCAGGCGTGACCGATCCAGGGTTGATCACCTTCTTCGACCGTGAAGATCATGCCTCGTTCGCAC
GTATCGCTGAGCAGGCCACGCGCGACTTGGCGGCGTCGACCAAAAGCGTGCTGGTCCACGTTGAAGGTACCCGGG
CCATCAGCTGCCGGACGCCCGTGACCCGGCTGGGTGCATCGGTGATTGACATGGCGCTGGCCGCAGGCGCCCCGA
TCGTGCCAGTTCGCTTCGCACGCGGGCTGCCCGAGTCGGATAGCCCCGAGCGGCTCGACTTCCCGTTGGGCTACG
GGCGCCAGGACATCTACTTCGGTGCGCCAATCGAGCCGGCTGAATTGGCCCGCCTGCCCTTGAAGGATCGTAAGG
ACGCACTGTTGACCGCCCTGAACACCTTGGGTCCGGGCGCCTCCGACACGCCCTCGGACCCCGATCCGTCGTTCG
CGGCGCAGGTTGAGGCATGGCGTGCCCGCACCGGGTGCGCCATTGAAGACGCGGTGTTCTACGCGACCTTGGCTG
GTGCCTCGGCGCCAGAGAGCGAAGGTGTGCAACGGCTGCTCGAAGGCGCCCGCACCGGCAAACTGGTAATCGGTA
GCGGTGAGAAGGATCGCTGGCTCGGTGAGTTCGCCCGGTCGCTGTTTGGGTCGGCCGACGCTACCCTGCTGGTCG
AAGGCGGCTGAGAGCGTGTTTACGGCGCGAATACGTGCCAGGTTCGGGACCCTCACCCCAACCCCTCTCCCGCGG
GGAGAGGGGCTCGGATCTGCTGCGAGGATCTTCCCCCTCTACCCTCGGGAGAGGGGGACCGAGGGGGTGAGGGAG
CCACGTAACGGCCTGTCAAATACGCGCTAAGGCCCCTTGTGCCCATTGAGGCACACGTCGAACGATTTCACCGTC
TTTAATTAA

>TR-*pfa2-pfa3*_V2

GGCCGGCCGAAGTCCGGGCTCCGCATCTGGTTGAGACCGGGCCCGCGAGCTGAAACTAAATACAACAACACAACA
AAATAAGGAGGATACATCATGACGTTCGAACCCATTGCAATCGTAGGCCAAGCTTGCACCCTGCCGGGCGCTCTG
ACCCCGGAAGCCCTGTGGGCCAACGTT

CHAPTER 5

PUFA production in the oleaginous yeast *Yarrowia lipolytica* with an artificial PUFA biosynthetic gene cluster originating from myxobacteria

INTRODUCTION

Application of the hemiascomycetous yeast *Yarrowia lipolytica* has awakened a strong industrial interest due to its capacity to grow efficiently on hydrophobic substrates (e.g. alkanes, fatty acids, and oils) as a sole carbon source and to produce high amounts of organic acids, especially citric acid [1]. Developments of genetic and cellular tools have contributed to the establishment of *Y. lipolytica* as an amenable host for heterologous protein production [2-4]. Since this yeast is considered as non-pathogenic, several processes based on *Y. lipolytica* were classified as generally recognized as safe (GRAS) by the food and drug administration (FDA).

Y. lipolytica is classified among the oleaginous yeasts because of its ability to accumulate large amounts of lipids (up to 50% of its cell dry weight, depending on the growth conditions) [5]. Beyond the essential amount of lipids required for membrane composition of cells and organelles, oleaginous microorganisms produce lipids above about 20% of their cell mass to accumulate them as reserve storage [6]. The fatty acid biosynthetic pathway in most oleaginous microorganisms is principally the same as found in non-oleaginous species, as exemplified by *Saccharomyces cerevisiae*. The crucial difference between oleaginous and non-oleaginous yeasts is unveiled during cultivation in a medium with an excessive carbon to nitrogen ratio [7;8]: For oleaginous yeasts, the catalytic growth rate slows down quickly, whereas carbon assimilation slows down more gradually. The carbon flux is then preferentially channelled toward lipid biosynthesis, leading to an accumulation of triacylglycerols within discrete lipid bodies in the cells. On the contrary, non-oleaginous yeasts tend to cease further cell proliferation under nitrogen-limited conditions, and the carbon flux into the cell is diverted into synthesis of various polysaccharides.

A hypothesis for the biochemical basis of lipid accumulation in oleaginous yeasts is represented by the following metabolic pathway: Nitrogen limitation leads to an activation of AMP deaminase (EC 3.5.4.6), resulting in a decrease of AMP concentration and an increase of ammonium concentration in mitochondria [9]. Consequently, isocitrate dehydrogenase (EC 1.1.1.41) is stopped from working due to its strict dependence on the presence of AMP in oleaginous yeasts [10]. As isocitrate cannot be metabolized, it accumulates and is rapidly interconverted with citrate by aconitase (EC 4.2.1.3). Citrate is exported from mitochondria into the cytosol by the citrate shuttle and metabolized into acetyl-CoA plus oxaloacetate via a cytosolic ATP:citrate lyase (EC 2.3.3.8). This enzyme has not yet been found in any non-oleaginous microorganism [11]. Oxaloacetate is then reduced by malate dehydrogenase

(EC 1.1.1.37) to malate, which is next converted into pyruvate via oxidative decarboxylation catalysed be the $NADP^+$-malic enzyme (EC 1.1.1.40). In addition to this enzyme, glucose-6-phosphate dehydrogenase (EC 1.1.1.49) and 6-phosphogluconate dehydrogenase (EC 1.1.1.44) of the pentose phosphate pathway are considered as major suppliers of NADPH for fatty acid biosynthesis [12]. Acetyl-CoA and NADPH are directly channelled into synthesis of fatty acids, which are esterified with glycerol to triacylglycerols and incorporated into fatty acid droplets.

Obviously, *Y. lipolytica* seems to be a promising host strain for recombinant production of polyunsaturated fatty acids (PUFAs), which could be already demonstrated in previous studies. By expression of a bifunctional $\Delta^{12}/\omega3$ desaturase from *Fusarium moniliforme* in *Y. lipolytica*, 28% α-linolenic acid (ALA, 18:3, *n*-3) of total fatty acids (TFAs) could be produced [13]. Production of 20% γ-linolenic acid (GLA, 18:3, *n*-6) of TFAs was obtained by overexpression of Δ^6 and Δ^{12} desaturases from *Mortierella alpina* [14]. Similarly, optimal fermentation conditions yielded production of 44% *trans*-10,*cis*-12-conjugated linoleic acid (CLA) of TFAs by overexpression of Δ^{12} desaturase from *M. alpina* and a linoleic acid isomerase gene from *Propionibacterium acnes*, which was codon-optimized for *Y. lipolytica* [15]. Recently, an engineered *Y. lipolytica* strain was published that produces 56.6% eicosapentaenoic acid (EPA, 20:5, *n*-3) and less than 5% saturated fatty acids of TFAs, which are the highest and the lowest percentages, respectively, among known EPA sources. These levels were achieved by four gene knockouts in the yeast genome and by overexpression of twenty desaturase genes, eight elongase genes, and two cholinephosphotransferase genes of different origins, which were codon-optimized for *Y. lipolytica* [16]. All these examples made use of enzymes from the aerobic PUFA biosynthetic pathways, which, in contrast to anaerobically acting PUFA synthase complexes, are not encoded by a gene cluster and do not enable *de novo* PUFA biosynthesis.

Modern techniques connected with Synthetic Biology provide the opportunity to transfer and express not only single genes but entire gene clusters from a certain organism into entirely unrelated organisms with very different evolutionary histories. For this purpose, complete biosynthetic pathways can be synthesized and assembled from DNA sequences that were holistically optimized for the chosen production host including codon bias adaptation. On top of this, sophisticated design of artificial pathways allows for facilitated genetic manipulation of pathway elements, such as the interchange of DNA sections. Hence, an artificial long-chain PUFA (LC-PUFA) biosynthetic gene cluster encoding iteratively acting

polyketide synthase (PKS)-like enzymes plus a 4'-phosphopantetheinyl transferase (PPTase) gene originating from *A. fasciculatus* (SBSr002) for the biosynthesis of docosapentaenoic acid (DPA, 22:5) and docosahexaenoic acid (DHA, 22:6, *n*-3) **[17]** were redesigned, synthesized, and assembled for *Y. lipolytica*. Successful transfer of the expression construct and heterologous expression was accomplished in the host strain. To the best of our knowledge, this example represents the first polyketide-like biosynthetic gene cluster adapted and synthesized for this yeast.

EXPERIMENTAL PROCEDURES

Sequence analysis and design of the synthetic gene clusters

The sequences of the *pfa* gene cluster plus the gene encoding the PPTase AfPpt from *A. fasciculatus* (SBSr002) were analysed and compared to the genome sequence of *Y. lipolytica* CLIB 122 [18] retrieved from NCBI Genome RefSeq NC_006067, NC_006068, NC_006069, NC_006070, NC_006071, and NC_006072 in collaboration with ATG:biosynthetics. Based on this, relevant parameters for constructional and functional sequence design were defined to generate artificial pathway versions using the proprietary evoMAG-GeneDesign software [19]. The sequence design process included engineering of restriction sites, adaptation of the codon usage, elimination of sequence repeats as well as rare codon clusters, elimination of potential donor splice signals, and introduction of hidden stop codons in unused frames.

Culture conditions

Escherichia coli DH10B [20] was used for cloning experiments. *E. coli* HS996/pSC101-BAD-gbaA (tet^R) [21] were used for modification of a plasmid using Red/ET recombination. The cells were grown in LB-medium or on LB-agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) at 30-37 °C (and 200 rpm) overnight. Antibiotics were used at the following concentrations: 100 µg/ml ampicillin, 50 µg/ml kanamycin, 34 µg/ml chloramphenicol, and 6 µg/ml tetracycline.

Auxotrophic *Yarrowia lipolytica* Po1g (CLIB 725) [3] and Po1h (CLIB 882) [2] were obtained from Centre International de Ressources Microbiennes (CIRM)-Levures, Institut

National de la Recherche Agronomique (INRA), AgroParisTech, Thiverval-Grignon, France. They were grown in YPD-medium or on YPD-agar containing 1% yeast extract, 2% peptone, and 2% dextrose. Prototrophic transformants were grown in minimal medium YNB containing 0.67% yeast nitrogen base (with 75% $(NH_4)_2SO_4$ and without amino acids), and 50 mM KPO_4 buffer pH 6.8. Different carbon sources were used: YNB-N_{5000} contains 1% glucose, YNBD contains 2% glucose, $YNBD_8$ contains 8% glucose, YNBG contains 2% (w/v) glycerol, and YNBO contains 3% (w/v) oleic acid (oleic acid was emulsified by sonication in the presence of 0.1% Tween 80). The cultures were incubated for 1-7 days at 16-30 °C (and 200 rpm).

Isolation of genomic DNA from *Y. lipolytica*

Isolation of genomic DNA from yeast cells was carried out using the method of Hoffman and Winston [22]. In the first step, 50 ml of a culture of *Y. lipolytica* grown at 28 °C for 3 days were harvested by centrifugation at 8,000 rpm for 5 min. The cell pellet was resuspended in 500 µl sterile ddH_2O and transferred into a 2 ml tube. The sample was centrifuged at 15,000 rpm for 1 min, the supernatant was decanted, and the tube was briefly vortexed to resuspend the pellet in the residual liquid. 200 µl lysis buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 1 mM EDTA, and 10 mM Tris-Cl pH 8.0), 200 µl phenol:chloroform:isoamyl alcohol (25:24:1), plus 0.3 g acid-washed glass beads (425-600 µm) were added to the cells. The tube was vortexed for 3 min. Afterwards, 200 µl TE buffer (10 mM Tris-Cl pH 8.0 and 1 mM EDTA) were added, and the sample was centrifuged at 15,000 rpm for 5 min. The upper aqueous phase was transferred to a tube containing 1 ml ice-cold ethanol and mixed by inversion. The sample was centrifuged at 15,000 rpm for 2 min, and the supernatant was discarded. The pellet was resuspended in 400 µl TE buffer plus 30 µg RNase A. After incubation at 37 °C for 15 min, 44 µl 4 M ammonium acetate and 1 ml ice-cold ethanol were added, and the tube was inverted to mix. Genomic DNA was precipitated at -20 °C for 2 min to increase yield. The sample was centrifuged at 15,000 rpm for 2 min, and the supernatant was discarded. The pellet was washed with 700 µl 70% ethanol, centrifuged at 15,000 rpm for 1 min, and the supernatant was discarded. The dried DNA was resuspended in 50 µl 10 mM Tris-Cl pH 8.0.

General molecular biology techniques

Routine handling of nucleic acids, such as isolation of genomic and plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, were performed according to standard protocols [23]. All the enzymes were purchased from Thermo Scientific or New England Biolabs.

Cloning of expression constructs containing the synthetic *pfa* gene cluster and the 4'-phosphopantetheinyl transferase (PPTase) gene originating from *A. fasciculatus* (SBSr002) for heterologous LC-PUFA production in *Y. lipolytica*

For the heterologous expression of artificial PUFA biosynthetic pathways in *Y. lipolytica*, four building blocks containing the three *pfa* genes as well as the gene encoding the PPTase AfPpt originating from *A. fasciculatus* (SBSr002) were designed and supplied by gene synthesis companies. Each building block contains one of the four genes flanked by the strong hybrid hp4d promoter [3] plus *LIP2* terminator for *Y. lipolytica*. Moreover, non-coding sequences were attached to each building block allowing for the connection of the transcription units by 200 bp intergenic linkers (see appendix for sequences). In the course of constructional design, the unique *Sda*I, *Apa*LI, *Nco*I, *Sal*I, *Acl*I, *Kpn*I, *Fse*I, *Xma*JI, and *Pac*I restriction sites were introduced for cloning purposes as well as for exchangeability of genes/domains and were excluded from any other unwanted position within the gene cluster. In addition to this, the four coding sequences were optimized by applying functional design approaches with special focus on sequence features affecting translational elongation (adaptation of the codon usage, elimination of sequence repeats as well as rare codon clusters, engineering of Shine-Dalgarno (SD)–anti-SD interactions, and introduction of hidden stop codons in unused frames).

A cloning plasmid with kanamycin resistance gene, p15A origin of replication, and restriction sites for assembly of the four DNA building blocks was constructed. Therefore, a multiple cloning site comprising *Dra*I, *Sda*I, *Apa*LI, *Nco*I, *Sal*I, *Acl*I, *Aat*II, *Ngo*MIV, *Xma*JI, *Pac*I, and *Bam*HI restriction sites (0.1 kb) was amplified by PCR using overlapping primers MCS for pACYC_fwd and MCS for pACYC_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The MCS was

inserted into plasmid pACYC177 via *Dra*I and *Bam*HI restriction sites, yielding plasmid pACYC_assembly (Table 2). The four building blocks were stitched together in plasmid pACYC_assembly. The assembly procedure is explained in more detail below: building block 1 was inserted into pACYC_assembly via *Sda*I and *Apa*LI restriction sites, generating plasmid pACYC_building_block_1 (Table 2); building block 4 was inserted into pACYC_building_block_1 via *Xma*JI and *Pac*I restriction sites, generating plasmid pACYC_building_block_1+4 (Table 2); building block 2 was inserted into pACYC_building_block_1+4 via *Apa*LI and *Acl*I restriction sites, generating plasmid pACYC_building_block_1+2+4 (Table 2); building block 3 was inserted into pACYC_building_block_1+2+4 via *Acl*I and *Xma*JI restriction sites, generating plasmid pACYC_building_block_1-4 (Table 2). The backbone of plasmid pACYC_building_block_1-4 was exchanged for the backbones of plasmids pINA1269-*Sda*I-cm^R-*Pac*I, pINA1312-*Sda*I-cm^R-*Pac*I, and pINA1292-*Sda*I-cm^R-*Pac*I (Table 2) via *Sda*I and *Pac*I restriction sites, yielding plasmids pSynPfaPptAf1, pSynPfaPptAf2, and pSynPfaPptAf3 (Table 2). For the construction of plasmids pINA1269-*Sda*I-cm^R-*Pac*I, pINA1312-*Sda*I-cm^R-*Pac*I, and pINA1292-*Sda*I-cm^R-*Pac*I by modification of plasmids pINA1269 [3], pINA1312, and pINA1292 [4], a 1.2 kb DNA fragment containing a chloramphenicol resistance gene with *Sda*I and *Pac*I restriction sites as well as the appropriate homology arms was amplified from plasmid pACYC184 (New England Biolabs) using primers HA+*Sda*I+cm^R for pINA1269_fwd / HA+*Pac*I+cm^R for pINA1269_rev, HA+*Sda*I+cm^R for pINA1312_fwd / HA+*Pac*I+cm^R for pINA1292+1312_rev, and HA+*Sda*I+cm^R for pINA1292_fwd / HA+*Pac*I+cm^R for pINA1292+1312_rev (Table 1). PCR was performed with *Taq* DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 8% glycerol and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 3 min at 95 °C; 30 cycles consisting of denaturation for 30 s at 95 °C, annealing for 30 s at 62 °C, and extension for 1 min/kb at 72 °C; and a final extension for 10 min at 72 °C. The linear plus circular homologous Red/ET recombineering was performed as described previously [17] using 1.5 µl of the PCR product.

A synthetic version of gene *mcherry* (0.7 kb), serving as reporter gene, was amplified from plasmid pPm1-*mcherry-rppA* (K. Gemperlein, unpublished) using primers *mcherry*_A_fwd and *mcherry_Bam*HI_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the

following conditions: initial denaturation for 2 min at 98°C; 30 cycles consisting of denaturation for 15 s at 98ºC, annealing for 20 s at 67ºC, and extension for 10 s/kb at 72ºC; and a final extension for 10 min at 72ºC. The PCR amplicon was inserted into plasmid pINA1312 via *Pml*I and *Bam*HI restriction sites, yielding plasmid pINA1312-*mcherry* (Table 2). For the fusion of hp4d-*mcherry* and *LIP2*t, the strong hybrid promoter hp4d in front of *mcherry* plus overlapping sequence was amplified as the first fragment (1.3 kb) from plasmid pINA1312-*mcherry* using primers UAS1B_*Psp*1406I_fwd and *mcherry*+*LIP2*t overlap_rev (Table 1). The second fragment (0.1 kb) contained the *LIP2* terminator from *Y. lipolytica* plus overlapping sequence and was amplified from plasmid pACYC_building_block_1 using primers *LIP2*t+*mcherry* overlap_fwd and *LIP2*t_*Psp*1406I_rev (Table 1). PCR amplification of the two fragments to be spliced was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 20 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 64 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. For the subsequent overlap extension PCR using primers UAS1B_*Psp*1406I_fwd and *LIP2*t_*Psp*1406I_rev, the two amplified fragments were used as templates. PCR of the 1.4 kb fragment was performed as described for the amplification of the two fragments to be spliced. Plasmid pSynPfaPptAf2mc (Table 2) was constructed by insertion of hp4d-*mcherry*-*LIP2*t into plasmid pSynPfaPptAf2 via *Acl*I restriction site.

Table 1. Primers used for cloning of assembly and expression plasmids in this study.
The homology arms are shown in italics and the introduced restriction sites are underlined.

Primer	Sequence (5'→3')
MCS for pACYC_fwd	TTTAAACCTGCAGGGCGGTGCACGCGCCATGGCGCGTCGACCGCAACGTTGCGCGACG
MCS for pACYC_rev	GGATCCTTAATTAACGCCTAGGGCGCGCCGGCGCGCGACGTCGCGCAACGTTGCGGTCGACGC
HA+*Sda*I+cm^R for pINA1269_fwd	*CAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGG*CCTGCAGGATTTAACGACCCTGCCCTGAACCGACG
HA+*Pac*I+cm^R for pINA1269_rev	*GAATTCTCATGTTTGACAGCTTATCATCGATGATAAGCTGTCAAACATGA*TTAATTAAGTAAGTTGGCAGCATCACCCGACGC
HA+*Sda*I+cm^R for pINA1312_fwd	*GAGACTGAAATAAATTTAGTCTGCAGCCCAAGCTAGCTTATCGATACGCG*CCTGCAGGATTTAACGACCCTGCCCTGAACCGACG
HA+*Sda*I+cm^R for pINA1292_fwd	*CGAGCTTCGTAGGAGGGCATTTTGGTATCAAGCTAGCTTATCGATACGCG*CCTGCAGGATTTAACGACCCTGCCCTGAACCGACG
HA+*Pac*I+cm^R for pINA1292+1312_rev	*TACCCTGTCGGATGACTAACTCTCCAGAGCGAGTGTTACACATGGAATTC*TTAATTAAGTAAGTTGGCAGCATCACCCGACGC
*mcherry*_A_fwd	AATGGTGAGCAAGGGCGAGGAGG
*mcherry_Bam*HI_rev	CCTTAGGATCCTCACTTGTACAGCTCGTCCATGC
UAS1B_*Psp*1406I_fwd	AAGAACGTTAGCTTATCGATACGCGTGC
mcherry+*LIP2*t overlap_rev	GGTAGAAGTTGTAAAGAGTGATAAATAGCTCACTTGTACAGCTCGTCCATG
*LIP2*t+*mcherry* overlap_fwd	CATGGACGAGCTGTACAAGTGAGCTATTTATCACTCTTTACAACTTCTACC
*LIP2*t_*Psp*1406I_rev	ATCCAACGTTGGTTTCGATTTGTCTTAGAGG

Table 2. Plasmids and expression strains constructed in this study.
oriV = origin of replication, *bla* = ampicillin resistance gene, *cat* = chloramphenicol resistance gene, *aph(3')-IIa* = kanamycin resistance gene, *LEU2* = 3-isopropylmalate dehydrogenase from *Y. lipolytica*, *URA3* = orotidine 5'-phosphate decarboxylase from *Y. lipolytica*, zeta = long terminal repeats (LTRs) of the retrotransposon Ylt1 from *Y. lipolytica*, hp4d = hybrid promoter comprising four copies of the distal upstream activating sequence from the *XPR2* promoter in front of a minimal *LEU2* promoter from *Y. lipolytica*, *XPR2*t = terminator of gene *XPR2* encoding an alkaline extracellular protease from *Y. lipolytica*, *LIP2*t = terminator of gene *LIP2* encoding an extracellular lipase from *Y. lipolytica*.

Plasmid / Strain	Characteristics
pACYC_assembly	Derivative of pACYC177 (New England Biolabs) in which a multiple cloning site comprising *Sda*I, *Apa*LI, *Nco*I, *Sal*I, *Acl*I, *Aat*II, *Ngo*MIV, *Xma*JI, and *Pac*I restriction sites was inserted via *Dra*I and *Bam*HI
pACYC_building_block_1	Derivative of pACYC_assembly in which building block 1 was inserted via *Sda*I and *Apa*LI
pACYC_building_block_1+4	Derivative of pACYC_building_block_1 in which building block 4 was inserted via *Xma*JI and *Pac*I
pACYC_building_block_1+2+4	Derivative of pACYC_building_block_1+4 in which building block 2 was inserted via *Apa*LI and *Acl*I
pACYC_building_block_1-4	Derivative of pACYC_building_block_1+2+4 in which building block 3 was inserted via *Acl*I and *Xma*JI
pINA1269-*Sda*I-cm^R-*Pac*I	Derivative of pINA1269 [3] in which hp4d and *XPR2*t were exchanged for *cat* flanked by *Sda*I and *Pac*I restriction sites by Red/ET recombineering. pMB1 *oriV*, *bla*, *cat*, *LEU2*.
pINA1312-*Sda*I-cm^R-*Pac*I	Derivative of pINA1312 [4] in which hp4d and *XPR2*t were exchanged for *cat* flanked by *Sda*I and *Pac*I restriction sites by Red/ET recombineering. pMB1 *oriV*, *aph(3')-IIa*, *cat*, *URA3*, zeta.
pINA1292-*Sda*I-cm^R-*Pac*I	Derivative of pINA1292 [4] in which hp4d and *XPR2*t were exchanged for *cat* flanked by *Sda*I and *Pac*I restriction sites by Red/ET recombineering. pMB1 *oriV*, *aph(3')-IIa*, *cat*, *URA3*, zeta.
pINA1312-*mcherry*	Derivative of pINA1312 [4] in which *mcherry* was inserted via *Pml*I and *Bam*HI
pSynPfaPptAf1	Derivative of pACYC_building_block_1-4 in which the vector backbone was exchanged for the vector backbone of pINA1269-*Sda*I-cm^R-*Pac*I via *Sda*I and *Pac*I
pSynPfaPptAf2	Derivative of pACYC_building_block_1-4 in which the vector backbone was exchanged for the vector backbone of pINA1312-*Sda*I-cm^R-*Pac*I via *Sda*I and *Pac*I
pSynPfaPptAf3	Derivative of pACYC_building_block_1-4 in which the vector backbone was exchanged for the vector backbone of pINA1292-*Sda*I-cm^R-*Pac*I via *Sda*I and *Pac*I
pSynPfaPptAf2mc	Derivative of pSynPfaPptAf2 in which hp4d-*mcherry*-*LIP2*t was inserted via *Acl*I
Y. lipolytica Po1g::pSynPfaPptAf1 clones	*Y. lipolytica* Po1g with pSynPfaPptAf1 randomly integrated in the genome
Y. lipolytica Po1h::pSynPfaPptAf2 clones	*Y. lipolytica* Po1h with pSynPfaPptAf2 randomly integrated in the genome
Y. lipolytica Po1h::pSynPfaPptAf2mc clones	*Y. lipolytica* Po1h with pSynPfaPptAf2mc randomly integrated in the genome

Transformation of *Y. lipolytica* by lithium acetate-mediated heat shock transformation and heterologous expression

Transformation of *Y. lipolytica* was carried out using a protocol developed by M.-T. Le Dall, modified by C. Madzak (Laboratoire de Microbiologie de l'Alimentation au Service de la Santé (MICALIS), AgroParisTech, Thiverval-Grignon, France; personal communication). One loopful of *Y. lipolytica* cells from a YPD-agar plate grown at 30 °C overnight were resuspended in 1 ml TE buffer in a sterile tube. The cells were centrifuged at 10,000 rpm for 1 min, and the supernatant was discarded. After resuspension in 600 µl 0.1 M lithium acetate pH 6.0, the cells were incubated at 28 °C for 1 h in a water bath. The samples were centrifuged at 3,000 rpm for 2 min, the supernatant was discarded, and the cells were softly resuspended in 80 µl 0.1 M lithium acetate pH 6.0. 40 µl of competent cells were mixed with 2 µl herring testes carrier DNA (10 mg/ml in TE buffer, denatured) and 3 µl linearized plasmid DNA to be transformed. The samples were incubated at 28 °C for 15 min in a water bath, and 350 µl 40% PEG 4000 in 0.1 M lithium acetate pH 6.0 plus 16 µl 1 M dithiothreitol (DTT) were added. After incubation of the cells at 28 °C for 1 h in a water bath, 40 µl DMSO were added. Subsequently, heat shock was carried out at 39 °C for 10 min in a heating block. 600 µl 0.1 M lithium acetate pH 6.0 were then added, and the cells were plated onto YNB-N_{5000}-agar. The plates were incubated at 30 °C for 2 days to four weeks. Thereafter, selected colonies were transferred onto new YNB-N_{5000}-agar plates. The constructed strains are listed in Table 2. Expression was carried out in 50 ml YNB-medium inoculated with an overnight culture starting from OD_{600} = 0.1. The cells were cultivated at 28 °C and 200 rpm for 90-168 h and then harvested at 4,000 rpm for 5 min.

Extraction of cellular fatty acids

The cellular fatty acids were extracted using the FAME method [24]. For this purpose, 50 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. The cell pellet was transferred to a glass vial and dried in a vacuum concentrator. Subsequently, the cell dry weight was determined. 5 µl (50 µg) of *n*-3 DPA (Sigma-Aldrich) and 500 µl of a mixture of methanol, toluene, and sulphuric acid (50:50:2, v/v/v) were added. The vial was capped with a teflon-lined screw cap and incubated at 80 °C for 24-48 h. After the mixture was cooled to room temperature, 400 µl of an aqueous solution consisting of 0.5 M NH_4HCO_3 and 2 M KCl were added, and the sample was vortexed for 30 s. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. 75 µl of the upper phase were

mixed with 25 µl *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) and incubated at 37 °C for 30 min. Subsequently, the sample was used for GC-MS analysis.

Extraction and fractionation of lipids

Extraction of lipids from microbial cells was carried out using the method of Bligh and Dyer [25], modified by Lewis *et al.* [26], on a small scale. In the first step, the cell pellet from a 15 ml culture was transferred into a polypropylene tube. Successively, 4 ml chloroform, 8 ml methanol, and 3.2 ml 1% NaCl were added, and the tube was vortexed at high speed for 15 s after every addition. The sample was agitated on a tube rotator at 30 rpm overnight. 4 ml chloroform and 4 ml 1% NaCl were then added, and the tube was inverted 30 times. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. The bottom layer containing the lipid extract was evaporated to dryness under a gentle stream of nitrogen and dissolved in 1 ml of a chloroform:methanol mixture (2:1, v/v). The lipid solution was fractionated by solid-phase extraction on a Strata SI-1 silica column with 100 mg sorbent mass (Phenomenex) equilibrated with chloroform + 1% acetic acid. Neutral lipids and free fatty acids were eluted from the column with 1 ml of chloroform + 1% acetic acid, glycolipids were eluted with 1.5 ml of an acetone:methanol mixture (9:1, v/v), and phospholipids were eluted with 1 ml methanol. For GC-MS analysis, the fractions were dried in a vacuum concentrator and further processed according to the FAME method described previously.

Analysis of fatty acid methyl esters by GC-MS

GC-MS was carried out on an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a 7683B split/splitless injector with autosampler (Agilent Technologies) and coupled to a 5973 electron impact mass selective detector (Agilent Technologies). Helium was used as the carrier gas at a flow rate of 1 ml/min. 1 µl of the sample was injected in split mode (split ratio, 10:1). The analytical column was a (5% phenyl)-methylpolysiloxane capillary column (Agilent J&W DB-5ht; 30 m x 0.25 mm i.d. x 0.1 µm film thickness, maximum temperature 400 °C; Agilent Technologies). The column temperature was kept at 130 °C for 2.5 min, increased to 240 °C at a rate of 5 °C/min, then ramped to 300 °C at 30 °C/min, and held at 300 °C for 5 min. Other temperatures were as follows: inlet, 275 °C; GC-MS transfer line, 280 °C; ion source, 230 °C; and quadrupole, 150 °C. The mass selective detector was operated in scan mode, scanning the mass range from *m*/*z* 40 to 700. Scan

control, data acquisition, and processing were performed by MSD ChemStation and AMDIS software, version 2.68, based on the fragmentation patterns and retention times, in comparison with Supelco 37 Component FAME Mix and LC-PUFAs (all Sigma-Aldrich) as reference standards, and NIST 08 library. Absolute amounts of PUFAs were quantified by integration of the peaks using MSD ChemStation and by subsequent calculation in relation to the integral of *n*-3 DPA and to cell dry weight. In case *n*-3 DPA was produced by the culture itself, the sample was spiked with *n*-3 DPA for the calculation of the absolute amounts of PUFAs.

Fluorescence measurements

The fluorescence of mCherry produced by yeast cells grown on YNB-N_{5000}-agar plates was determined using a Molecular Dynamics Typhoon 9410 Variable Mode Imager (Amersham Biosciences/GE Healthcare). An excitation wavelength of 532 nm ± 8 nm and an emission wavelength of 580 nm ± 15 nm were used.

The fluorescence of mCherry produced by yeast cells grown in liquid YNBD-medium was determined using an Infinite 200 PRO microplate reader (Tecan). Ahead of the measurements, 20 µl of the cultures were mixed with 180 µl YNBD-medium in a 96-well microtitre plate. An excitation wavelength of 587 nm and an emission wavelength of 610 nm were used. For data analysis, the obtained values were normalized by autofluorescence background subtraction.

RESULTS AND DISCUSSION

Development of heterologous expression platforms for recombinant LC-PUFA production using the DPA (22:5, *n*-6 or *n*-3)/DHA (22:6, *n*-3)-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) ([17]; Chapter 3 and 4) served as proof of principle and paved the way for the establishment of a more promising host organism. By exploiting the potential of Synthetic Biology, myxobacterial PUFA biosynthetic pathways can thereby be introduced and expressed into an entirely unrelated organism with a very different evolutionary history. In this context, the oleaginous yeast *Y. lipolytica* was focussed on as expression host of choice, since it features GRAS designation, fast growth characteristics, accessibility for genetic manipulation, and ability to accumulate large amounts of lipids.

For the heterologous expression of the DPA/DHA-type Pfa proteins and the PPTase AfPpt from *A. fasciculatus* (SBSr002) in the oleaginous yeast *Y. lipolytica*, synthetic versions of the three *pfa* genes and *Afppt* were created. Each gene is flanked by the strong hybrid hp4d promoter [3] and the *LIP2* terminator for *Y. lipolytica* to construct single transcription units (Figure 1). Moreover, non-coding sequences were attached allowing for the connection of each transcription unit by 200 bp intergenic linkers. During sequence modulation processes, the myxobacterial genes were subjected to the algorithms of the proprietary evoMAG-GeneDesign software by ATG:biosynthetics. The software applies concepts of evolutionary algorithms [27;28] to generate sequences which take into account predefined multivariate sequence parameter values. Degeneracy of the genetic code allows for the substitution of synonymous codons by silent mutations in order to generate the artificial sequence based on the native template without altering the amino acid sequence. For practical reasons, the 20.2 kb gene cluster was dissected into smaller DNA fragments, which were supplied by gene synthesis companies (Figure 1). Constructive sequence requirements for pathway assembly from the synthetic DNA building blocks and for future interchangeability of inter- and intragenic regions were specified and implemented by insertion of restriction enzyme sites at any required position within the sequence by silent point mutations. In parallel, interfering restriction sites had to be eliminated from the sequences.

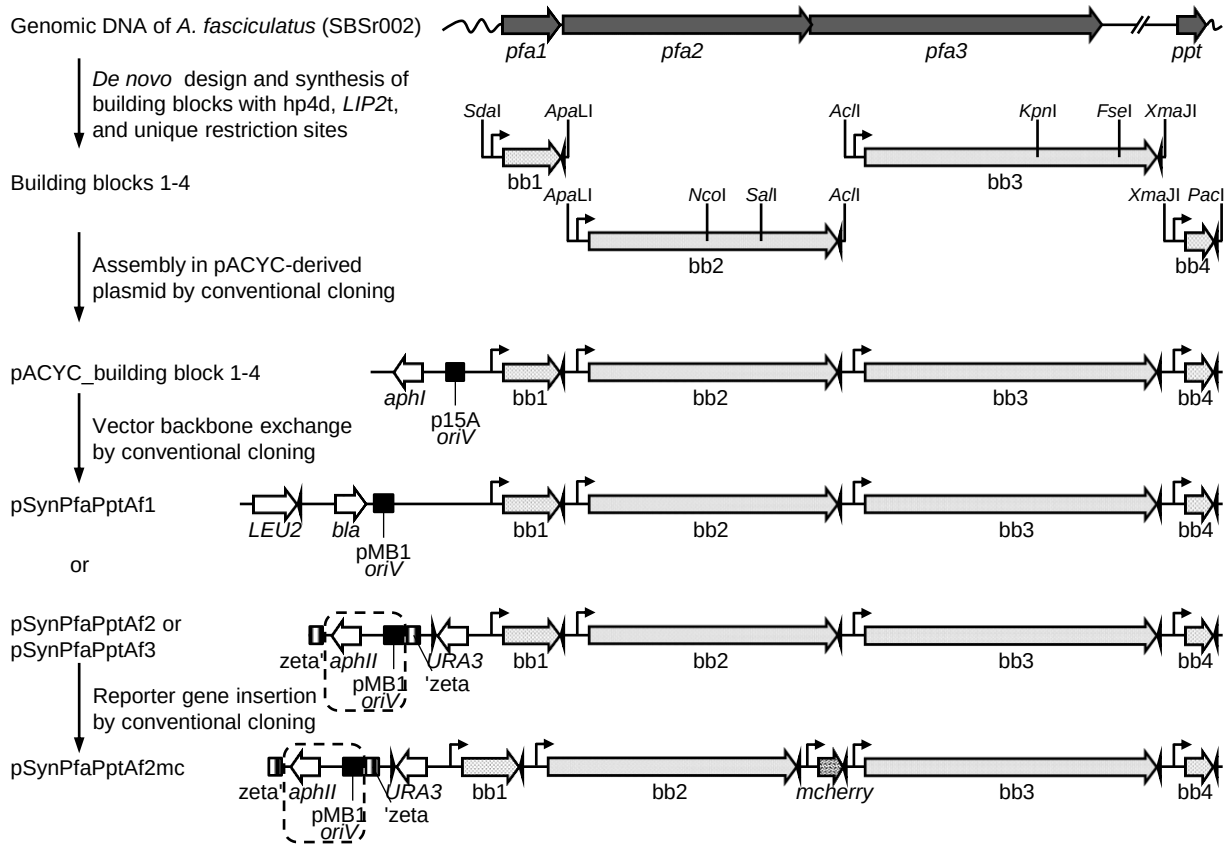


Figure 1. Design, synthesis, and cloning of a synthetic *pfa* gene cluster originating from *Aetherobacter fasciculatus* (SBSr002) for recombinant LC-PUFA production in *Yarrowia lipolytica* Po1g or Po1h.
The region encircled by the dashed line was excised from pSynPfaPptAf2 and pSynPfaPptAf3 prior to transformation into *Y. lipolytica*. *oriV* = origin of replication, *bla* = ampicillin resistance gene, *aphI* and *aphII* = kanamycin resistance genes, *LEU2* = 3-isopropylmalate dehydrogenase from *Y. lipolytica*, *URA3* = orotidine 5'-phosphate decarboxylase from *Y. lipolytica*, zeta = long terminal repeats (LTRs) of the retrotransposon Ylt1 from *Y. lipolytica*, hp4d = hybrid promoter comprising four copies of the distal upstream activating sequence from the *XPR2* promoter in front of a minimal *LEU2* promoter from *Y. lipolytica*, *LIP2*t = terminator of gene *LIP2* encoding an extracellular lipase from *Y. lipolytica*.

In order to satisfy constructional demands, the complete sequence was adapted for specific requirements of the host *Y. lipolytica* in the context of functional sequence design. Moreover, general sequence requirements for gene design and gene synthesis (e.g. avoidance of sequence repeats, mRNA stabilities) were considered in the construction process of the synthetic gene cluster. In the process of sequence optimization, different parameters were addressed in detail (Figure 2). Comparison of the codon usage of the original cluster sequence and the codon table of *Y. lipolytica* led to the application of a 20% exclusion threshold (Figure 2A). The resulting artificial codon usage table only contains 39 instead of 61 sense codons. Renormalization of the synonymous codon fractions from the genome sequence of *Y. lipolytica* had to be performed to ensure that the sum of each subset of codons for the same amino acid equals 1. Eventually, high adaptation of the codon usage of the artificial genes to

the reduced and renormalized table could be achieved. As *Y. lipolytica* shows a higher preference for AT-rich codons than myxobacteria, the mean GC-content of the coding sequences was lowered from 73% to 61% (Figure 2B). Another important parameter affecting translational elongation represents the local codon adaptation index (CAI), which is a measure for the deviation of the synonymous codon usage bias of a given coding sequence with respect to a reference set of genes [29]. After sequence adaptation, the local CAI shows a clearly smoothed shape without any distinct peaks as found for the native sequence (Figure 2C). In the artificial sequences, the gradient between start and stop codon of each coding DNA sequence was 0.05, which is intended to enhance the ribosome occupancy along the mRNA to shield the transcript from degradation [30]. In general, the speed of the ribosomes and the resulting translational elongation rate depends on several factors, including the availability of aminoacyl-tRNAs, but also on certain sequence features of the mRNA template, e.g. rare codon clusters [20] and tandem-repeats of codons below average usage frequency. These sequence features are often assigned to temporally separate translation of segments within the peptide chain and enable proper cotranslational folding of proteins with direct impact on their solubility and activity [31]. However, the search for relevant rare codon clusters within the native genes did not reveal any conspicuous candidates. Thus, they were eliminated from the synthetic genes (Figure 2D). Moreover, quantity and strength of specific sequences within the artificial genes which could be recognized as donor splice sites by *Y. lipolytica* [32] were reduced to avoid splicing of the intronless pre-mRNAs (Figure 2E). Another important factor to be considered is the metabolic cost of translational miselongation upon a frameshift [33]. In order to cause early termination of translation of incorrect reading frames, the ribosome must encounter an off-frame stop codon. Hence, stop codons were accumulated in the unused frames of the artificial cluster sequence (Figure 2F).

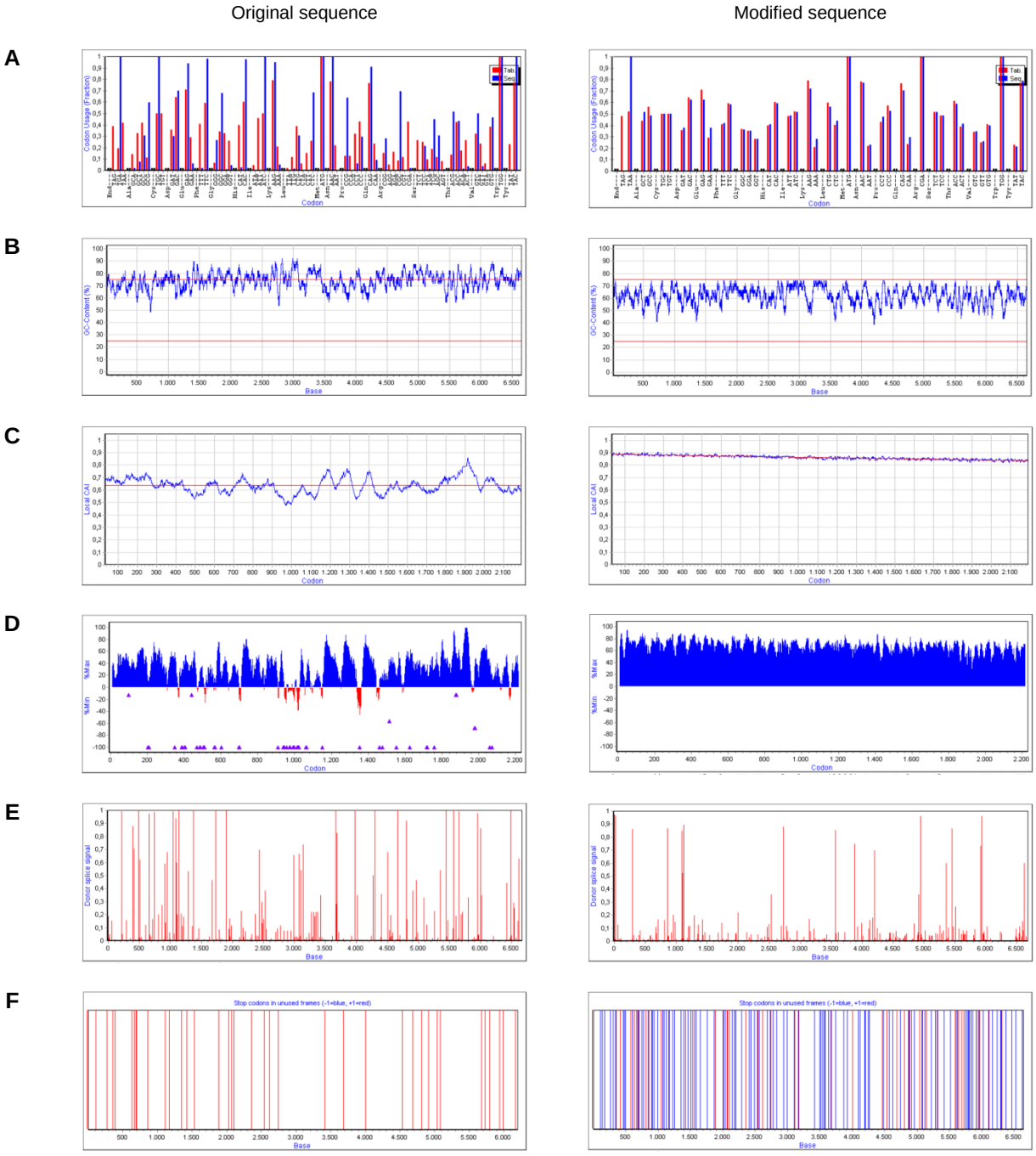


Figure 2. Comparison of the native and the artificial sequence of gene *pfa2* for *Yarrowia lipolytica* originating from *Aetherobacter fasciculatus* (SBSr002).
(A) Codon usage of the original sequence (left blue bars) compared to the codon table of *Yarrowia lipolytica* (left red bars) and codon usage of the modified sequence (right blue bars) compared to the reduced and renormalized codon table of *Yarrowia lipolytica* (right red bars; the cut-off for the synonymous codon fractions was 0.2), (B) Course of the GC-content, (C) Course of the local codon adaptation index (CAI); in the modified coding DNA sequence (CDS), the gradient between start and stop codon was 0.05, (D) Rare codon clusters according to Clarke and Clark [34]; additional violet triangles: tandem-repeats of codons below average usage frequency, (E) Potential donor splice signals (higher values reflect higher probabilities of cryptic splice sites), (F) Hidden stop codons in unused frames.

Practical implementation of the heterologous expression of the artificial *pfa* gene cluster in *Y. lipolytica* started with reconstitution of the entire pathway by assembly of the DNA

building blocks supplied by a gene synthesis company (Figure 1). For this purpose, a pACYC-derived assembly plasmid was constructed by insertion of a suitable multiple cloning site, and the DNA building blocks were assembled stepwise by conventional cloning techniques using unique restriction sites engineered during the constructional sequence design process. The backbone of the resulting plasmid pACYC_building block 1-4 was then exchanged for a backbone on the basis of the *Y. lipolytica* shuttle vectors pINA1269 [3], pINA1312, and pINA1292 [4]. The resulting plasmid pSynPfaPptAf1 was linearized with *Not*I and transferred into *Y. lipolytica* Po1g [3] by lithium acetate-mediated heat shock transformation, whereas the "yeast cassette" of plasmid pSynPfaPptAf2 or pSynPfaPptAf3, generated via hydrolysis with *Not*I, was transferred into *Y. lipolytica* Po1h [2]. Numerous clones were obtained for the monocopy vector pSynPfaPptAf1, designed to be integrated at the pBR322 docking platform into the genome of Po1g strain, and for the monocopy auto-cloning vector pSynPfaPptAf2, able to promote random integration into the genome of Po1h strain. In case of the multicopy auto-cloning vector pSynPfaPptAf3, able to promote ectopic integration of multiple copies of the integration construct into the genome of Po1h strain, no transformants were attained due to unknown reasons. In an initial screening, 17 transformants of *Y. lipolytica* Po1g::pSynPfaPptAf1 and 38 transformants of *Y. lipolytica* Po1h::pSynPfaPptAf2 were cultivated in 50 ml YNB-N_{5000}-medium at 28 °C for 90 h. hp4d-driven expression of the synthetic genes was supposed to start at the entry into stationary phase. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [24] and analysed by GC-MS. Three *Y. lipolytica* Po1g::pSynPfaPptAf1 clones (clones A-C) and eleven *Y. lipolytica* Po1h::pSynPfaPptAf2 clones (clones A-K) were all shown to produce DHA. However, significant deviations in DHA production yields (up to 25-fold difference) among different transformants could be observed (Figure 3).

In order to investigate the extreme variability among the DHA producing clones and the occurrence of a large number of non-producing transformants (14 out of 17 and 27 out of 38 analysed clones) on the genotype level, next generation sequencing approaches were applied. The genomes of the 14 DHA producers as well as of two non-producing clones of *Y. lipolytica* Po1g::pSynPfaPptAf1 and of *Y. lipolytica* Po1h::pSynPfaPptAf2, respectively, were sequenced. The integration sites of the constructs were mapped according to the genome sequence of *Y. lipolytica* CLIB 122 [18] (Table 3). In all DHA producers that have been sequenced, the expression constructs are integrated into different loci, which are widely distributed among the six chromosomes of *Y. lipolytica*. In case of the four non-producers, the expression constructs are either incompletely integrated into the genome or contain a

missense mutation. Interestingly, for none of the *Y. lipolytica* Po1g::pSynPfaPptAf1 clones, the presumed integration of the construct into the pBR322 docking platform located at chromosome C of Po1g strain was observed. With a few exceptions, either the integration site of the 5' end or of the 3' end of the linear expression construct could be identified, whereas the other end of the construct appears to be integrated into repetitive sequences which cannot be mapped due to the limited length of the sequencing reads. Only in case of *Y. lipolytica* Po1g::pSynPfaPptAf1 clone C and *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, the integration sites for both ends of the linear expression construct could be mapped. Unfortunately, for *Y. lipolytica* Po1h::pSynPfaPptAf2 clones H and I, both ends of the expression construct were integrated into repetitive sequences within the genome which cannot be mapped. The copy number of the transgenes integrated in the genome was determined by analysis of the integration site as well as by quantification of the coverage with sequence reads compared to surrounding genomic regions. For *Y. lipolytica* Po1g::pSynPfaPptAf1 clone A and *Y. lipolytica* Po1h::pSynPfaPptAf2 clones D plus F, duplications and inversions of parts of the expression constructs could be observed, but an obvious correlation between gene duplication and enhanced recombinant LC-PUFA production rate is not discernible. As gene expression in eukaryotes is a complex process regulated at multiple levels, transcription of transgenes is not only influenced by recombinant promoters and regulatory DNA elements but also by the local action of enhancers and other regulatory elements at the integration site [35]. Moreover, epigenetic effects, such as compartmentalization of the chromosomes into various types of chromatin domains, affect gene transcription levels depending on the spatial positioning of the transgenes within the genome [36]. Thus, the high variability of DHA production by the different transformants most likely results from position effects due to random integration of the expression constructs into various loci.

The transgenic strain *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, that showed the highest DHA yields in the initial analysis (Figure 3), was selected for studies on production optimization starting with the evaluation of different cultivation conditions. Fermentation studies using several complex and minimal liquid media (details not shown) revealed that growth of the yeast cells and lipid accumulation were boosted in rich, complex media, but DHA production was strongly reduced or even undetectable. Using minimal YNB-media, growth rate was inferior, but the amounts of recombinant produced LC-PUFAs were elevated depending on the carbon source: Glucose or glycerol at a concentration of 2% (w/v) led to an increase in heterologous LC-PUFA production. Raising the glucose concentration to 8% or feeding of 3% (w/v) oleic acid resulted in an accumulation of lipids up to 19% of the cell dry

weight but impaired recombinant LC-PUFA production. Cultivation at different temperatures (16 °C, 25 °C, 28 °C, and 30 °C) revealed that heterologous LC-PUFA production at 28 °C clearly surpassed the production rate at the other temperatures. Investigations on the production kinetics showed that DHA continually accumulated with progressive cultivation time. In conclusion, initial approaches for production optimization resulted in the cultivation of *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C in 50 ml YNBD-medium (containing 2% glucose) at 28 °C for 168 h and in an improved recombinant LC-PUFA production (Table 4). Remarkable amounts of DHA and some minor amounts of *n*-6 DPA, *n*-3 DPA, EPA (20:5, *n*-3), arachidonic acid (AA, 20:4, *n*-6), and eicosatetraenoic acid (ETA, 20:4, *n*-3) could be achieved. Production yields clearly surpass the heterologous LC-PUFA production rate achieved with the native producer strain *A. fasciculatus* (SBSr002), which produces ca. 7-10% EPA of TFAs / 3.8-4.3 mg EPA/g cell dry weight (CDW) / 1 mg EPA/l and ca. 10-19% DHA of TFAs / 4.7-6.2 mg DHA/g CDW / 1.2-1.9 mg DHA/l as major PUFA species after seven days of cultivation. The total amount of fatty acids accounts for 11% of the cell dry weight of *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C under the preferred cultivation conditions. Extraction and fractionation of the lipids, subsequent preparation of fatty acid methyl esters (FAMEs) by transesterification, and GC-MS analysis showed that most of the synthesized DHA (83%) is incorporated into the phospholipids. 15% of synthesized DHA are present in the fraction containing neutral lipids plus free fatty acids and 2% in the glycolipid fraction.

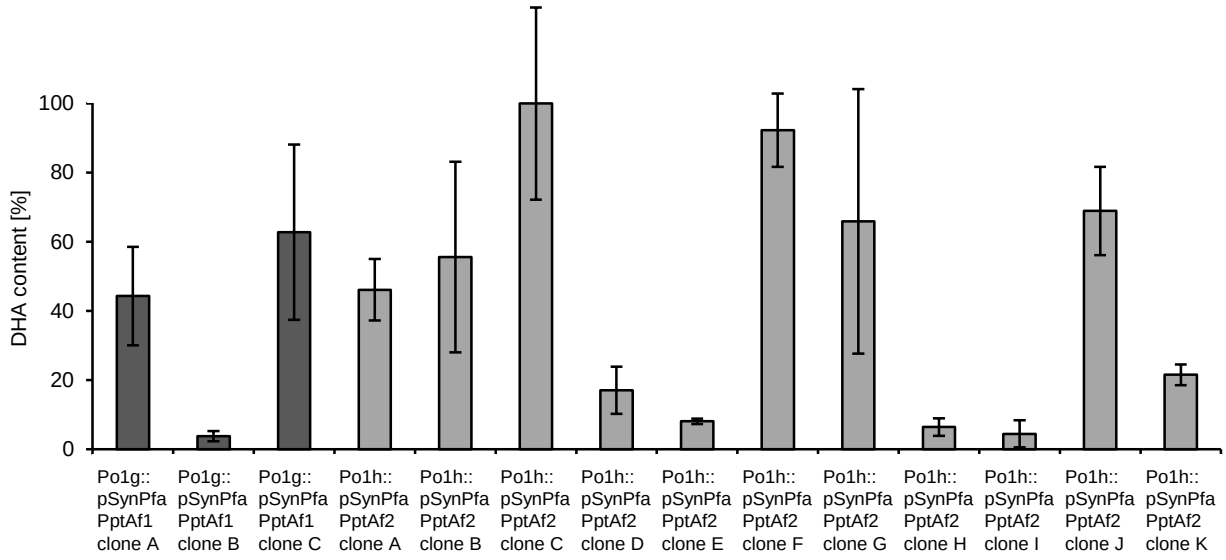


Figure 3. Recombinant DHA production by *Yarrowia lipolytica* Po1g or Po1h with the synthetic *pfa* gene clusters originating from *Aetherobacter fasciculatus* (SBSr002).
Cultivation was carried out in 50 ml YNB-N_{5000}-medium at 28 °C for 90 h. The DHA content is presented as a relative value and was set at 100% (≙ 0.4 mg/g CDW) for the best producing clone Po1h::pSynPfaPptAf2 clone C. The indicated values are means and standard deviations of three biological samples.

Table 3. Characteristics of the integration of the synthetic *pfa* gene cluster originating from *Aetherobacter fasciculatus* (SBSr002) into the genome of *Yarrowia lipolytica* Po1g or Po1h, as examined by genome sequencing.

Transgenic strain	Integration site		Copy number of the integrated genes
	Chromosome	Gene locus	
Po1g::pSynPfaPptAf1 clone A	F	within YALI0F09537g	1x: *pfa1*, *pfa2*, and first half of *pfa3*; 2x: second half of *pfa3* and *ppt* due to duplication and inversion within the construct
Po1g::pSynPfaPptAf1 clone B	E	near YALI0E08822g	1x
Po1g::pSynPfaPptAf1 clone C	E	between YALI0E05005g and YALI0E05027g	1x
Po1h::pSynPfaPptAf2 clone A	E	within YALI0E26763g	1x
Po1h::pSynPfaPptAf2 clone B	A	within YALI0A03597g	1x
Po1h::pSynPfaPptAf2 clone C	C	within YALI0C05907g	1x
Po1h::pSynPfaPptAf2 clone D	C	between YALI0C10208g and YALI0C10230g	1x: *pfa1* and first third of *pfa2*; 2x: two-thirds of *pfa2*, *pfa3*, and *ppt* due to duplication and inversion within the construct
Po1h::pSynPfaPptAf2 clone E	F	near YALI0F17270g	1x
Po1h::pSynPfaPptAf2 clone F	A	near YALI0A08448g	1x: *pfa1*, *pfa2*, *pfa3*, and first half of *ppt*; 2x second half of *ppt* due to duplication and inversion within the construct
Po1h::pSynPfaPptAf2 clone G	E	within YALI0E35200g	1x
Po1h::pSynPfaPptAf2 clone H	repetitive sequence which cannot be mapped		probably 1x
Po1h::pSynPfaPptAf2 clone I	repetitive sequence which cannot be mapped		probably 1x
Po1h::pSynPfaPptAf2 clone J	F	between YALI0F06974g and YALI0F07018g	1x
Po1h::pSynPfaPptAf2 clone K	E	within YALI0E26433g	1x

Table 4. Fatty acid composition of *Yarrowia lipolytica* Po1h::pSynPfaPptAf2 clone C.
Cultivation was carried out in 50 ml YNBD-medium + 50 mM potassium phosphate buffer pH 6.8 at 28 °C and 200 rpm for 168 h. The indicated values are means of three biological samples. The recombinant produced PUFAs are shown in bold.

Fatty acid	Relative amount [% of total fatty acids]	Absolute amount [mg/g cell dry weight]	Yield [mg/l]
14:0	0.02	0.02	0.13
15:0	0.10	0.10	0.76
16:0	9.22	10.15	79.15
16:1	5.59	6.05	46.60
16:0 2-OH	0.72	0.77	5.80
17:0	0.17	0.17	1.23
17:1	0.37	0.38	2.75
18:0	4.29	4.84	38.40
18:1	55.51	60.90	473.26
18:2	8.73	9.72	76.35
20:0	0.09	0.10	0.75
20:1	0.16	0.17	1.22
20:2	0.02	0.02	0.19
20:4 *n*-6	**0.06**	**0.06**	**0.48**
20:4 *n*-3	**0.01**	**0.01**	**0.08**
20:5	**0.14**	**0.14**	**1.03**
22:0	0.15	0.16	1.24
22:5 *n*-6	**0.21**	**0.21**	**1.56**
22:5 *n*-3	**0.10**	**0.10**	**0.70**
22:6	**8.89**	**9.35**	**70.41**
22:0 2-OH	0.03	0.03	0.22
24:0	1.11	1.23	9.60
24:0 2-OH	4.15	4.39	33.23
26:0	0.16	0.19	1.46

As shown earlier in the chapter, the low rate of recombinant LC-PUFA producers among the *Y. lipolytica* transformants and the high variability of DHA production between the producing clones were caused by truncated expression constructs still comprising the selection marker or by position effects associated with random integration, respectively. A powerful traditional approach to study the influence of the local environment on gene expression represents the integration of reporter genes into different chromosomal sites [37]. Following this strategy, incorporation of a fluorescence gene into the LC-PUFA expression construct was aimed at. A reporter gene expression cassette consisting of *mcherry* flanked by the same promoter and transcriptional terminator as used for the *pfa* genes and *Afppt* (strong hybrid hp4d promoter [3] plus *LIP2* terminator) was constructed and inserted into the linker region between genes *pfa2* and *pfa3* of plasmid pSynPfaPptAf2 by conventional cloning (Figure 1). By positioning of the reporter gene in the centre of plasmid pSynPfaPptAf2mc, a

linkage between integration of not only the outer regions comprising the auxotrophic marker but even the middle parts of the "yeast cassette" and mCherry fluorescence was attempted to be established. The "yeast cassette" of plasmid pSynPfaPptAf2mc was transferred into *Y. lipolytica* Po1h [2] by lithium acetate-mediated heat shock transformation. Ten transformants of *Y. lipolytica* Po1h::pSynPfaPptAf2mc which exhibited intensive mCherry fluorescence on agar plate plus *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, as a reference, were cultivated in 1 ml YNBD-medium at 28 °C for 92 h. The mCherry fluorescence signal from the liquid cultures was quantified by photometry; the cellular fatty acids were extracted via acidic methanolysis using the FAME method [24] and analysed by GC-MS. Six *Y. lipolytica* Po1h::pSynPfaPptAf2mc clones produce DHA in similar amounts as *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C with deviations of ± 30%, as determined by ratios of DHA and endogenous fatty acids. Two clones produce 48% or 6% of the DHA amount synthesized by the reference strain, whereas two transformants do not show any production of recombinant LC-PUFAs. DHA production and mCherry fluorescence correlate for almost all of the transformants, except for those clones exhibiting a very low or no DHA production. In these cases, mCherry fluorescence reaches elevated values falsely indicating a superior recombinant LC-PUFA production. The discrepancy between the anticipated high LC-PUFA production yield based on the reporter signal and the actual low amounts of detected DHA in some clones might be due to duplication of *mcherry*, which would result in a higher fluorescence as consequence of the increased gene dosage. Regarding the clones which do not produce any recombinant LC-PUFAs, the integrated PUFA biosynthetic pathway might be either truncated or mutated, but without affecting the functionality of the *mcherry* reporter gene. All in all, the use of the reporter gene within the LC-PUFA expression construct enabled the identification of DHA producing transformants with a probability of 80%. Three quarters of the producing clones synthesize half or more of the DHA amount produced by the best transformant. In comparison, the use of the parent expression construct pSynPfaPptAf2 led to an identification rate for a DHA producing transformant of only 29%, and merely 40% of the producing clones synthesize half or more of the DHA amount produced by the best transformant (Figure 3). The best producing strain harbouring the *mcherry* construct, *Y. lipolytica* Po1h::pSynPfaPptAf2mc clone A, was cultivated in 50 ml YNBD-medium at 28 °C for 168 h, analogous to the reference strain *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C. Production of 7.1% DHA of TFAs / 6.0 mg DHA/g CDW / 41.0 mg DHA/l and some minor amounts of *n*-6 DPA, *n*-3 DPA, EPA, AA, and ETA could be achieved, as determined in triplicates. The relative amounts of recombinant produced LC-PUFAs are comparable to the recombinant LC-

PUFA production by *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, considering the standard deviations, but the absolute amounts and the yields are inferior (Table 4) due to a faster and denser growth and a higher lipid content of the latter strain.

In conclusion, recombinant LC-PUFA production could be successfully accomplished in *Y. lipolytica* with an artificial DPA/DHA-type *pfa* gene cluster plus PPTase gene originating from the myxobacterium *A. fasciculatus* (SBSr002). Using the first polyketide biosynthetic gene cluster modulated and synthesized for this yeast, a 40-fold increase in DHA yield per litre of fermentation broth compared to the native myxobacterial producer could be attained. Twelve different integration sites of the LC-PUFA expression construct could be identified within the genomes of the *Y. lipolytica* transformants. They confer a high variability of DHA production with up to 25-fold differences. The identified integration site for the best producing clone allows for stable propagation and high transcription rates of the LC-PUFA genes and, most probably, of any heterologous gene. However, site-specific integration of a (large) expression construct into the genome of *Yarrowia* turned out to be challenging. Therefore, inclusion of a reporter gene into the artificial LC-PUFA expression construct to be ectopically integrated enables a facilitated and rapid identification of proficient recombinant LC-PUFA producers. Application of this system might be useful for the heterologous expression of different recombinant genes in the future.

REFERENCES

[1] Madzak C, Gaillardin C, & Beckerich JM (2004) Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review. *J. Biotechnol.*, **109,** 63-81.

[2] Madzak C (2003) New tools for heterologous protein production in the yeast *Yarrowia lipolytica* . *Rec. Res. Dev. Microbiol.*, **7,** 453-479.

[3] Madzak C, Treton B, & Blanchin-Roland S (2000) Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J. Mol. Microbiol. Biotechnol.*, **2,** 207-216.

[4] Nicaud JM, Madzak C, van den Broek P, Gysler C, Duboc P, Niederberger P, & Gaillardin C (2002) Protein expression and secretion in the yeast *Yarrowia lipolytica*. *FEMS Yeast Res.*, **2,** 371-379.

[5] Papanikolaou S, Chevalot I, Komaitis M, Marc I, & Aggelis G (2002) Single cell oil production by *Yarrowia lipolytica* growing on an industrial derivative of animal fat in batch cultures. *Appl. Microbiol. Biotechnol.*, **58,** 308-312.

[6] Thorpe RF & Ratledge C (1972) Fatty acid distribution in triglycerides of yeasts grown on glucose or n-alkanes. *Microbiology*, **72,** 151-163.

[7] Ratledge C (2004) Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie*, **86,** 807-815.

[8] Beopoulos A, Cescut J, Haddouche R, Uribelarrea JL, Molina-Jouve C, & Nicaud JM (2009) *Yarrowia lipolytica* as a model for bio-oil production. *Prog. Lipid Res.*, **48,** 375-387.

[9] Morin N, Cescut J, Beopoulos A, Lelandais G, Le Berre V, Uribelarrea JL, Molina-Jouve C, & Nicaud JM (2011) Transcriptomic Analyses during the Transition from Biomass Production to Lipid Accumulation in the Oleaginous Yeast *Yarrowia lipolytica*. *PLOS One*, **6,** e27966.

[10] Morgunov IG, Solodovnikova NY, Sharyshev AA, Kamzolova SV, & Finogenova TV (2004) Regulation of NAD(+)-dependent isocitrate dehydrogenase in the citrate producing yeast *Yarrowia lipolytica*. *Biochemistry (Mosc.)*, **69,** 1391-1398.

[11] Evans CT, Scragg AH, & Ratledge C (1983) A comparative study of citrate efflux from mitochondria of oleaginous and non-oleaginous yeasts. *Eur. J. Biochem.*, **130,** 195-204.

[12] Zhang H, Zhang L, Chen H, Chen YQ, Ratledge C, Song Y, & Chen W (2013) Regulatory properties of malic enzyme in the oleaginous yeast, *Yarrowia lipolytica*, and its non-involvement in lipid accumulation. *Biotechnol. Lett.*, **35,** 2091-2098.

[13] Damude HG, Zhang H, Farrall L, Ripp KG, Tomb JF, Hollerbach D, & Yadav NS (2006) Identification of bifunctional Δ12/ω3 fatty acid desaturases for improving the ratio of ω3 to ω6 fatty acids in microbes and plants. *Proc. Natl. Acad. Sci. U S A*, **103,** 9446-9451.

[14] Chuang LT, Chen DC, Chen YH, Nicaud JM, Madzak C, & Huang YS (2009) Production of functional γ-linolenic acid (GLA) by expression of fungal Δ12- and Δ6-desaturase genes in the oleaginous yeast *Yarrowia lipolytica*. In *Biocatalysis and Agricultural Biotechnology* (Hou CT & Shaw JF, eds), pp. 163-180. CRC Press Inc.

[15] Zhang B, Chen H, Li M, Gu Z, Song Y, Ratledge C, Chen YQ, Zhang H, & Chen W (2013) Genetic engineering of *Yarrowia lipolytica* for enhanced production of *trans*-10, *cis*-12 conjugated linoleic acid. *Microb. Cell Fact.*, **12,** 70.

[16] Xue ZX, Sharpe PL, Hong SP, Yadav NS, Xie DM, Short DR, Damude HG, Rupert RA, Seip JE, Wang J, Pollak DW, Bostick MW, Bosak MD, Macool DJ, Hollerbach DH, Zhang HX, Arcilla DM, Bledsoe SA, Croker K, Mccord EF, Tyreus BD, Jackson EN, & Zhu Q (2013) Production of omega-3 eicosapentaenoic acid by metabolic engineering of *Yarrowia lipolytica*. *Nat. Biotechnol.*, **31,** 734-740.

[17] Gemperlein K, Rachid S, Garcia R, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

[18] Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuveglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisrame A, Boyer J, Cattolico L, Confanioleri F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud JM, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard GF, Straub ML, Suleau A, Swennen D, Tekaia F, Wesolowski-Louvel M, Westhof E, Wirth B, Zeniou-Meyer M, Zivanovic I, Bolotin-Fukuhara M, Thierry A, Bouchier C, Caudron B, Scarpelli C, Gaillardin C, Weissenbach J, Wincker P, & Souciet JL (2004) Genome evolution in yeasts. *Nature*, **430,** 35-44.

[19] Oßwald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, & Wenzel SC (2012) Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.* Epub ahead of print.

[20] Grant SG, Jessee J, Bloom FR, & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U S A*, **87,** 4645-4649.

[21] Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF, & Zhang Y (2006) An improved recombineering approach by adding RecA to λ Red recombination. *Mol. Biotechnol.*, **32,** 43-53.

[22] Hoffman CS & Winston F (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*, **57,** 267-272.

[23] Sambrook J & Russell DW (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[24] Bode HB, Ring MW, Schwär G, Kroppenstedt RM, Kaiser D, & Müller R (2006) 3-Hydroxy-3-methylglutaryl-coenzyme A (CoA) synthase is involved in the biosynthesis of isovaleryl-CoA in the myxobacterium *Myxococcus xanthus* during fruiting body formation. *J. Bacteriol.*, **188,** 6524-6528.

[25] Bligh EG & Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can. J. Biochem. and Phys.*, **37,** 911-917.

[26] Lewis T, Nichols PD, & McMeekin TA (2000) Evaluation of extraction methods for recovery of fatty acids from lipid-producing microheterotrophs. *J. Microbiol. Methods*, **43,** 107-116.

[27] Goldberg DE (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA.

[28] Bäck T & Schwefel HP (1993) An overview of evolutionary algorithms for parameter optimization. *Evol. Comput.*, **1,** 1-23.

[29] Sharp PM & Li WH (1987) The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.*, **15,** 1281-1295.

[30] Pedersen M, Nissen S, Mitarai N, Lo SS, Sneppen K, & Pedersen S (2011) The functional half-life of an mRNA depends on the ribosome spacing in an early coding region. *J. Mol. Biol.*, **407,** 35-44.

[31] Zhang G, Hubalewska M, & Ignatova Z (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.*, **16,** 274-280.

[32] Neuveglise C, Marck C, & Gaillardin C (2011) The intronome of budding yeasts. *C. R. Biol.*, **334,** 662-670.

[33] Itzkovitz S & Alon U (2007) The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Res.*, **17,** 405-412.

[34] Clarke TF & Clark PL (2008) Rare Codons Cluster. *PLOS One*, **3,** e3412.

[35] Flagfeldt DB, Siewers V, Huang L, & Nielsen J (2009) Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast*, **26,** 545-551.

[36] Chen M, Licon K, Otsuka R, Pillus L, & Ideker T (2013) Decoupling epigenetic and genetic effects through systematic analysis of gene position. *Cell Rep.*, **3,** 128-137.

[37] Thompson A & Gasson MJ (2001) Location effects of a reporter gene on expression levels and on native protein synthesis in *Lactococcus lactis* and *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.*, **67,** 3434-3439.

APPENDIX

DNA sequences of the synthetic building blocks

>building block 1

CCTGCAGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTA
CGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTC
CCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCG
GCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACAT
TATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCC
CACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCA
CTGATCACGGGCAAAAGTGCGTATATATACAAGAGCGTTTGCCAGCCACAGATTTTCACTCCACACACCACATCA
CACATACAACCACACACATCCACAATGTCCGCTATTGGCCGATGGAATGCTCAAGGAGGATCTCCTGAGCTGACC
CCCGCTGCTATGGTGGCCGCTATCGCTCGATGTCGAGAATCTCTGCACGTCGTGCGAGACGGCGAATCTGGCCGA
ATCGGCCTGGGCGTGGGCGGAGAACTGCTGCCCCCTGGAGAGCGACGAGCCGGAGAATACACCCTGGTGGCCTCT
CTGCCTCCCCTCTTCCCTGAATGGCTGGGCGACCGATCTTTCTGCGAGGTCCACGGCGTTCGATTTCCCTACGTT
GCTGGAGAGATGGCTAACGGAATCGCTACTGTCGAGATGGTTGTGGCCATGTCCCGATCCGGCATGCTGGGTTTC
TTCGGTGCTGCTGGTCTCGGCTTCTCCCGAGTCGAGGCTGCCGTCGAGACTCTGCGAGCCACTCTGGGTGACGAG
GCCCCTTGGGGCGTGAACCTGATCCACTCTCCCAATGAGGCTTCCCTGGAGAACCGAGTCGCCGATCTCCTCATT
GAATCTGGCGTGCGACGAATCTCCGTGTCCGCCTTCATGGCCCTGACTCCTGCCGTTGTCCGATGCGCTGCTCGA
GGCCTCCGACTGGACGCTGCCGGACGAATTGTTCGACCCCGACATATTTTTGCTAAGCTGTCCCGACGAGAGGTT
GCCCTGCCCTTTCTCTCCCCCGCTCCCGAGGAGATGCTCCGAGCTCTCGTTGCTTCCGGCGGACTGACTGAGGAG
GAAGCCCGACTCGCCGCCAAGATCCCCGTGGCTGAGGATATCACCGTCGAGGCTGACTCTGGTGGTCATACCGAC
AACCAGGCTCTGGTCGCCCTGCTCCCCACTATCCTCCAGCTGCGAGACGAGCTGGCTGACAAATACGACTACACC
CGACCTATCCGAGTTGGAGCCGCTGGTGGTCTCGGCACCCCTACCGCCGTCGCTGCTGCTTTTTCCCTGGGTGCC
GCCTACGTCGTGACCGGATCTGTGAACCAAGCCGCTCTGGAATCTGGCCTGTCTGATGAGGGAAAGCGAATGCTC
GCTCAGGCTGAGGTCGCCGACGTTATCATGGCCCCTGCTGCTGATATGTTTGAGCTCGGCGTGAAGCTGCAGGTC
CTGCGACGAGGTTCTATGTTTGGTGTTCGAGCCGCCAAGCTCTATGAGGCCTACCACGCCCATCCCTCTCTCGAG
GCCATTCCCGATGCTCTGCGAGCTCGACTGGAACGAGAGGTCCTCCATGCTACCTTTGACGGAATTTGGGCTGAG
ACTCGAGCTTTCTGGGAGGCCCGAGATCCTGAACAGGTCGCCCTGGCTGCTCGAGACCCCAAGCACCGAATGGCT
CTCGTGTTCCGATGGTATCTCGGAATGGCCTCCCGATGGGCTATTGCTGGAGAAGCTTCTCGACGAGCTGACTAC
CAAATTTGGTGTGGTCCTGCTATGGGAGCCTTTAACTCCTGGGTCCGAGGATCCTTCCTGGAAGATCCTAAACAG
CGAGGCGTGGTGGCCATTGCCCTGAACCTGCTCGAGGGTGCCGCCGTTGTCACTCGAGCTCACCAGCTCCGAACC
TACGGCGTTCCTGTTCCTGCCGCCGCCTTTGATTTCCGACCTCGACCCCTCGCCTAAGCTATTTATCACTCTTTA
CAACTTCTACCTCAACTATCTACTTTAATAAATGAATATCGTTTATTCTCTATGATTACTGTATATGCGTTCCTC
TAAGACAAATCGAAACCGGATCCGGCGCGCCGATCCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTA
TTTCACACCGCATATATGGTGCAC

>building block 2

GTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGC
GCCCTGACGGGCTTGCGCTCGGCTTCACCTTCACTCGACAGGTCTTCGCCATGCTGAGGTGTCTCACAAGTGCCG
TGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCT
CGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTA
CGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTC
CCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCG
GCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACAT
TATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCACTGATCACGGGCAAAAGTGCGTATATATACAAGA
GCGTTTGCCAGCCACAGATTTTCACTCCACACACCACATCACACATACAACCACACACATCCACAATGACCCAAG
TCCCCGTTGCTATTGTCGGAGTCGGCGCCCTGTTCCCCGGATCTCCCGACGGTGCCGGATTCTGGCGAGATATTG
TGGCTTCCCGAGATCTCGTCTCTGATGTCCCTCCCGGACACTGGCTGATTGAGGACTACTACGACCCCGACCCCT
CCGCCCCTGACAAGACTTATTCCAAGCGAGGCGCCTTCCTCGGAGAGACTGCCTTCGACCCCCTGGAATTCGGCG
TCGTGCCCTCTTCCCTGCCCACCACTGACACCGCTCAGCTCCTCGCTCTGATTGTGGCTCGACGAGTCCTGGATG
ACGCTACTCAGGGTCGATTCGCTTCCATTGACCGATCCCGAATTGGCGTCGTTCTCGGCGTGACCTCCGGACAGG
AGCTGCTGGGCACTATGGTTTCCCGACTGCAGCGACCTGTTTGGCTGAAGGCCCTCCGAGACGACGGCGTCCCCG
AAGCTGACGCTCAGCGAATCTGTGACCGAATCACTAACGAATACGTCCCCTGGAAGGAATCTACCTTCCCTGGAC

TCCTGGGAAACGTCGTGGCTGGTCGAATTGCCAACCGATTCGACCTCGGTGGAACCAACTGCGTGACTGATGCCG
CCTGCGCCTCTTCCCTCGCTGCCCTGTCCATCGCCATGAACGAGCTGGCCCTCGGCCAATCTGACCTGGTCATCA
CCGGTGGCGTGGACACTATGAATGACATTCTGATGTACACCTGCTTCTCCAAGACTCCCGCTCTGTCTAAATCTG
GCGATTGCCGACCCTTCTCTGACGCCGCCGACGGCACTCTGCTGGGCGAAGGCCTGGCTATGTTCGCCCTGAAGC
GACTCACCGACGCTGAGCGAGATGGAGACCGAGTCTACGCCGTGATCCGAGGTCTCGGATCTTCTTCTGACGGCC
GATCTAAATCTGTCTATGCCCCCGTGCCTGAGGGCCAGGCCAAAGCCCTCGTGCGAGCCTACGAGGCCGCTGGAT
ACGGTGCTGACACCGTCGAGCTGGTCGAGGCCCACGGCACCGCTACCAAGGCTGGAGACGCCGCCGAATTTGGCG
GACTGAAGCTGGCCTTTGAGGCCTCTGGTCGAGCTGACCATCAGTGGTGTGCTCTGGGTTCCGTTAAGTCTCAAA
TCGGCCACACCAAATCTGCCGCCGGAGCCGCCGGACTCTTTAAGGCCGTGATGGCTCTCCACCACAAGGTCCTGC
CTCCCACCATCAAGGTTGACCGACCCAACCCCGAGCTCGCCATTGAGGCCTCCCCCTTTTACCTGAACACTCAAG
CCCGACCTTGGGTCCGAAACGCTCAGCATCTGCGACGAGCTTCCGTCTCTTCTTTTGGTTTCGGCGGTTCTAACT
TTCATGTGACCCTGGAGGAATACGCTGGCACTGCTGAGCCTGCTGCCCGACTGCGAACCGCTCCCACCGAGCTGT
TTGTGATTTCCGCCCCTTCCGCTGCTGAGGTCCTGGCTCGAGCCGGAAAGCTCGCCGAACTGGCCGATAAGGAGG
GAACCACTACCGCCGCCCGAACTTCCCAGCGAGACTTCGTGACTGCTGACCATGCCCGACTCTCTATCGTTGCCT
CTGATGCCGCCGATCTCCGAGCCAAGCTCGAGCAGGCTCGATCTCTCCTCTCTGCCCGACCTGAGGAGGCTTTCT
CTACCCCTACCGGCGTTCACTTTGCTGTGGGCGCTGCCGCCGGAGGCGTTGCCTTTCTGTTCCCTGGCCAGGGAT
CTCAATACGTGGGAATGGGAGCTGACGTTGCTATGGAGCACGGCGTGGCCCTGGATGCCTGGGACGCTGCTGCTG
GCCTGGTCGAAACCGCTGGACTCCGAGATGTGGTGTTCCCCCGACCCGCCTTTACTGATGAGGCCCGAGCCGAGC
AGGCCCAGCTGCTGACTGCCACCGAATGGGCCCAGCCCGCTATCGCTACTGCCTCTCTGGCCCTCCTCGCTGTGC
TGCGAGCCGTCGGTGTCCGACCTTCCTCCGTGGCCGGACACTCCCTCGGTGAACTGACCGCTCTGGCTGCTGCTG
ACGCCCTCTCCACCAACGATCTCGTGGCCACTGCCCGAAAGCGAGGCCAGCTGATGAATGCCGCCTCTTCTACCC
CCGGAGCCATGACCGCCGTTGCTCGACCTCTGGCTGAGGTCCAGGCTCTGCTCGAGCGACTGCAATCTGAGGTCG
TGGTGGCCAACCATAACCACCCCACTCAGGTCGTGCTCGCCGGAGAAACCGCCGCCATCGAAGCTGTGGAGGCCG
AACTGGCCCGACAGGGAATCACCGCTAAGCGACTCCCCGTTGCCACCGCCTTCCATTCTCAGGTCGTGGCTGGAG
CCACCGCCCCCTTTGCCGAACACCTCGCCGGAATCGCTGTGAAGGCTCCTGCCATCGATGTGCTGTCTAACTCTG
AAGCTGCCGCTTACCCCTCCGATCCTGACGCTGTCCGAGCTCTCCTCGCTGGACAGATCACCCGATCTGTGCGAT
TTGTTGAGGTCGTGGAGGCTATGTGGGCTCGAGGCGCTCGAATTTTCGTTGAGGTCGGACCCTCTGCCGTGCTGA
CTGAACTCTGTGCCCGAATCCTCGGCGAACGACCCCACCTCGCTATTGCTATGGATCGAAAGGGAAAGCACGGCG
TGACCGCCCTCCAAGATGCTCTGGGCCGACTGGCCGCCGCTGGCGTGGCCATTGATTTTGCTCCCCTGTGGGCCC
CTTTCGGACCCCCTGCCGCTCCTAAGCGAGTCCCCGCTATGTCCATGATGATCAATGGCGCTAACTACGGCAAGG
TCTACCCTCCCAAAGGCGGTGCCGCTGCCCTGCCTCCCCCTAACCCTCCTCGACCTGTGGTCGCTGCCGCTACCA
TCGCCGCTGCTCCCTCTTCCGCTGCCAAGGCTCCCGTCGTTGCCGCTGCCACCATTGCTGCCGCCCCCATCCCTG
CCGCTAAAGCCCCCGTGGTGGCTGCCGCCACTATCGCTGCTCGACCCGCCGCTGCTCCTGAATCTCGACCCGTCG
TGGCCGCTGCCCCTGCTCCTGCTCCTGTTGTGGCTGCCCCTGCTTCCGCTGCCGCTGCTCTGTCCGCTCCTGGCC
AGCTGGGCTGGATTGAGGCCTACCAGGAATCCCAGCGACAGCTGACCGAAGCTCATGCTTCTTACGTTCGAACCA
TGGCCTCTGGTCATGAGGCTTTCCTGACCACTATGGAAGCCTCCTTTATCGGACTCGGCGGACTGCTCGGTGCTC
CCGCCTCTGTGTCCGCCGTTGCCCCTCGATCTGCTGCCGCCCCCGTTCTGGCTCAGCCTGTGGCCGTGGCTCCTG
CTCCTGTCCTCGCCCCTGCTCCCGTTCCCGCCCCTGCTCCTACTCCTGTGCTGGCCCCCGCTCCTGCTCCTGTGC
TGGCTCCTGCTCCTGTGGCTGCCGTGGCTAAGGCTGCCCCTGCCCCCGTGATGGCTCCTGCTGCCCCTTCTATCG
ACCTGCAAGCTCTGCTGCTCGACGTGGTCGCCGATAAGACCGGCTATCCCGCTGAGATGCTGGGCATGCACATGG
AGCTCGAAGCTGACCTGGGCATCGACTCTATTAAGCGAGTTGAAATCCTGTCTGCTATGCGAAAGCGAGCCCCCA
ATCTCCCCGAGGTTGTCGCTTCTGAGATGGCCACCCTCCGAACCCTGGGCCAGATTGTTGAATATATGCGAGAGC
GATCCGGAGCCCTGACCGCCGCTACCCACGAGACTCTGGTTCCCGCTACCGCTGAGGCCGCTACTCAGGCCGCCC
CTGCTCTGGGAGATCTGCAGGCCCTCCTCCTGGACGTTGTGGCCGACAAAACCGGATACCCTGCCGAGATGCTCG
GAATGCACATGGAACTCGAGGCCGACCTCGGAATTGACTCCATCAAGCGAGTCGAAATCCTCTCCGCCATGCGAA
AACGAGCCCCTAACCTGCCCGAGGTCGTTGCTTCCGAGATGGCTACCCTGCGAACTCTCGGACAGATCGTGGAAT
ACATGCGAGACCGATCTGGTGCTACTGCCGCCCCTGTGAAGGCTGCTCCTGCTCCTGCTCCTGCCCCTGCCGCTG
CTCCCACCAAGGAGGCTCCCACCCTCGCTGACCTCCAGACTCTCCTGCTGTCTGTCGTTGCTGACAAGACTGGTT
ACCCCGCCGAAATGCTGGGAATGCATATGGAGCTGGAGGCCGATCTGGGTATTGATTCCATTAAACGAGTCGAGA
TTCTGTCTGCCATGCGAAAGCGAGCTCCCAACCTCCCTGAGGTGGTGGCCTCTGAAATGGCCACTCTGCGAACCC
TCGGTCAAATTGTCGAATACATGCGAGATCGATCTGGCGGACACTCTGTGGCTGCCACCACCCCTGCTCCCGCCC
CCGCTTCCACTCTGGCTCCTGCTGCTTCCCGAATTGGTCGATACGTCGTCGGAACTCGAGAGGAAGCCGCCTCCG
GACTCGCTATGGGTGGTGTTCTGGGTGCTCGACGATTCGTGATCACTAACGAAGGCACTGGCATCGCTCAGGCTC
TCGTCCGAGCTCTCCAGGAGCGAGGCGTTTCCTCCGAGGTCGTGACCGAGGTCCCCGCTGACGCCGATGCTGTGG
TTTTCCTGGGAGGACTCCGAACTCTGGTCGACGTTGATACTGCCATCGCCATCAATCGAGAGGCTTTCCTCGCTG
CCCGAGCCGTGGCTTCCGTGTTTTCCCGAGATGGTGGCGCTTTCGTGACTGTCCAGGACACCGGTGGCGACCTCG
GCATTGAAGCTCGAGAGCCCCGAGCCGCTTGGCTGGCTGGCGTTGCTGCTCTGGCCCGAACCGCCCAACTGGAAT
GGCCCAAGGCCTCTGTGAAAGCCGTCGATGTGGAGCGAGGTGGTCGATCCTCTGAGGAGATTGCTGCTGCCATCG
CTGAGGAAATCCTGACTGGTGGTGGTGCCCTCGAGGTCGGACTGACCGCTGCTGGAGCCCGACGAGTCCCTGAGG
TTGTGCAGCGAGCTGTTGGAGAGCGAACTCCTGTTCTCGCTCGATCCTCTGTGGTTGTGGCTTCCGGTGGTGCTC
GAGGCGTTACCGCCACTGCCCTGATCGCCCTCGCTCGACATGCTCAGCCTAAGCTCGTGATCCTCGGTCGATCCG
CCCTCGAGGCTGAACCCGCCGCCTGTCTGGGTGTCCGAGGAGATGCCGCCCTGAAACGAGCCCTGCTGGAGGACG
CTCGAGCCCAAGGTCGAACCGCCACTATTGCTGAGATTGGTGCCCATGTTGCCCGAATTTCCGCTTCCCGAGAGG

TTCTGGAGACTCTCGCTGCCATTAAGGCCGCCGGTGGCGAAGCCCGATACCTCGCTGTTGACGTGACCGATGCCG
ACCGACTCCGACTCGCCCTCGAGGAGGTCCGACGAGACTGGGGTCCTATTACTGCTGTTGTGCATGGTGCTGGCG
TCCTCGCCGATAAGCGAATCGAGGACAAGACCCCCGAACAATTCGACCGAGTCTTCGATACCAAGGTCCTGGGAC
TCCGAGCCCTCCTGGCTGCTACCGAGGGTGACCCCCTCCGAGCTATTATTCTCTTTTCTTCTGTGGCCGCCCGAG
CTGGCAACGTCGGCCAATGCGACTATGCTATGGCCAACGAAATCCTGAATAAGGTCGCCATCGCTGAGCGACGAC
GACGAGGTGGCGCTTGTACCGTGAAATCTCTGGGCTGGGGTCCCTGGGAAGGCGGAATGGTCACCCCCGCCCTGA
AAGGCCACTTCGAATCTATGGGCGTGGCCCTGATCCCCCTGGCTGAAGGAGCCGAACACTTCATCGCTGAGATCG
AAGGATCCGCTGACGAAATTGAGCTCGTGATCGGTGGCGGTGCTGATGGATCCGCCATTCGAGGCGACGCTGAGC
CTACCATTCACTTTGGCGTTTCCGTTTCCGATCGAACTTCCCCTTGGCTCGCTGACCATCGAGTCGCCGGACGAC
GAGTCATCCCCGTCGCTGTTGTGATTGAATGGTTTACCCGAGCTGTGACTGCTCTCCAGGGTCGAGCTGCTGTTT
CTCTGCGAGATGTTAAGGTCCTGCGAGGAATCAAGCTCCCCGCTGGCGACGCTGCCGCTTCCTTTTCTATCGGTT
GTCGAAAACTGGCCGAGGGCAAAATTGAACTGACCCTGCACGGCGCCGATGGAGCCCTCCATTACCGAGCTTCCG
CTGAGCCCCGAGGCGCTGCCCTGAAACCCGAGGGAGCTGCTCCTCCCCTGGCTGCTTGGGGCGACCGACCCATCT
ATGACGGTGGCCTCCTCTTCCACGGTCCTCGATTTCATGCTATTCATTCCATTGAGGGAGCTTCCGCTGAAGGAA
TTCGAGCTACTCTGACCGGAGCTCGAGAGCTCGGATGGCCCGGAGAGGCCCAACGAACTGATCCCGCTATTGTGG
ACGGAGCTCTGCAGCTGGCCCTGCTGTGGACTCGAGAACGACTCGGCCAAGCCTCTCTGCCTATGGCCATTGCCG
CCTTTGATTGGCAAGCCTCCGGTCCCGTGGAAGGTCCCGTCCGAGCCGTGGTTCAGGGTCGAGAGGCTGACTCCG
CCCACGCTGTTTGCGATGTTTGGCTCGAGGACGCTCGAGGAGCTATCGTTGGAGCTCTCCGAGGTGTCGAGACCA
TTGTTCGACCTGCTGAGGTCCGAGCTCCTCACCTCGTGGAGACCGGTCCCGCTTCTTAAGCTATTTATCACTCTT
TACAACTTCTACCTCAACTATCTACTTTAATAAATGAATATCGTTTATTCTCTATGATTACTGTATATGCGTTCC
TCTAAGACAAATCGAAACCACGCCAGCAAGACGTAGCCCAGCGCGTCGGCCGCCATGCGGGCGATAATGGCCTGC
TTCTCGCCGAAACGTT

>building block 3

AACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGC
CGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCCATGCTGAGGTGTCTC
ACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTC
CCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGT
GTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTG
CCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCA
CCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGT
GTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCACTGATCACGGGCAAAAGTGCGTAT
ATATACAAGAGCGTTTGCCAGCCACAGATTTTCACTCCACACACCACATCACACATACAACCACACACATCCACA
ATGACCTTTGAACCTATTGCTATCGTCGGACAAGCCTGTACCCTGCCCGGAGCTCTCACCCCTGAGGCTCTGTGG
GCCAACGTCCTGCGAGGCGAATCTGCCGTGACCTCTCCTCCTGAGGGACGATGGCGACTGCCTGCTTGGCACGCT
ATGGGAAAGCCCGGCGCCCAGGAGGAGCGAGCTTGGTCTGATGCTGGAGGATACGTCCGAGGATTCGAATCTGCT
TTCGACCCTGGTGGCCTCCGACTGCCCGAAGCTGAGCTGCGAGGCCTCGACCCCGTTTTTCAGTGGACTCTGCAT
ACCGCCCGAGCCGCTCTGCGATCTGCCGGAATGGAGGGCCCCTCTGCCCGAGCTGGCCTCGTCCTCGGCAACCTG
TCCTTCCCCACCGCTTCCATGTCCCGATACGCTGAATCCGTCTGGCTCGAGGCTCTGGGAAAGGAGCTGCTCGGT
GGACGAGCCGCCGCCCAAGCTGGCGTGACTCATCCCGACGCCCATAACCGATTCATGTCTGGCCTGCCCGCCCAT
CTGGCTGCTCAGGCTCTGGGACTGGGCGGTGGAGCCTTCGCCCTGGATGCTGCTTGTGCCTCCTCTCTCTACGCT
ATTAAGCTCGCCTGCGACCGACTGCAGGATCGAGCCACCGATCTGATGCTGGCTGGTGCCGTGAACGCCGCTGAT
TCTCTGTTTCTCCATATCGGATTTTCTGCCCTGTCCGCCATGTCCCGAACCGGACAGTCTCGACCTTTCCACCGA
GAAGCCGACGGCCTCGTGCCCTCTGAAGGAGCTGTGTTCTTTGCCCTCCAGCGACTCGACGATGCCCGAGCCGCC
GGACGACGAGTCCTGGGAATTATCCGAGGCGTTGGTCTCTCTAACGACGGTCGAGGACGAGGACTGCTCGCTCCT
TCTGAGGAAGGCCAGATCCGAGCTATGCGACTCGCTTACGCTGCCGCCGGACTGACCCCCTCTGACGTGTCCTTC
GTCGAATGCCATGCCACCGGAACTCCCGTTGGAGACGCCACTGAGATCCGATCCCTGTCCGAGGTCTTTGCTGGT
TGTCGAGACGTGCCCATCGGCTCTCTGAAGTCTAATCTGGGCCATCTCATCACCGCCGCTGGTGGTGCTGGACTG
ATTAAGGTCCTGGGTTCCTTCACCGCTGAGACTATCGCCCCTACCATCCATGCCGGAGCTTCCATCCCTGAGCTG
GCCGGATCCCCTGTGCGACTCGTGACTCAGGCTGAACCCTGGCGAGCTGCTGGCCGACGAATTGCCGCCATCTCC
GCCTTTGGTTTCGGAGGCAACAACGCCCATCTGATTGTTGAGCAAGACGACGGACGAGCTACTGCTCCCCGAGCT
GCTGCCCCTCGATCTGCTCCCGCCCCTATTGCCGTGATCTCCCTGGGTGCCCGAATCGGAGATACCACCGGACGA
GATGAGGCTGCCCGAGCTCTCCTGTCTGGCGCTGGCTGGGGTTCCCGACGAGAGGCTGTGACCGTTGAGCTGGAG
GGTCTCCGATTTCCCCCTCGAGATCTGGAGCAAACCCTGCCCCAGCAGCTGCTGGTCCTCGAGGCTGGTCGAGAG
GCTATCCGAGGCCTCTCCCTGCCCCGAGATCGAACCGCCGTTCTGATCGGAATGGGTGCTGACCCCGAGGTCGCT
CGATACGGCGCTCGATGGCGACTCCCCGCCTTTGCTGATGCCTGGGCCGCTGCTGGTCTCCCCGTGTCCTCTGCC
TGGACTGAGGCTGCTCGAGACGCTGTGCAGGAACAGCATGGAGCCGCCGGTGTCGTCGGCGCCATGCCCAACATC
CCCGCTAACCGACTCTCCTCTCAGCTCGACCTCGCTGGCCCCTCCTTTACTATTTCCTCTGAGGAGCTCTCCGGT
GTCGTGGCTCTCCAGCTCGCCGCTCGAGCCCTGCGATCCGGAGAGATCGATGCCGCCCTGGTCGGAGCTGTTGAT

CTGGCCGATCAGATCGTCCACCGAAAGGCCCTGGGAGAGCTCGGCCTGGAGCCTCGAGCTGGAGACGCTGCCGTG
GCCCTGGTGCTGAAACGACTGACCGATGTCCGAAAATCTGAGAAGGTCCTGGCCATTCTGGATGAGGGCATCGTG
GACGCCTCCTCCGATCCCGCCCGACTCCGAGTCGGAGACGGTGGCGCTGACCTGGATCCTTTCGCTACCACCGGC
AAGCCTCACGCCGCTGCTGGCCTCGTGCATGTGGCTGCCGCTGTCCTGGCTCTGCATCACGGAGCTCGACCTGCT
CCTGGTCGAGCTTCTATGCCTTGGGTCGGACCCCGACACGCTGAGGTGACCACTCAGGCCCTCGGCGATCAGGCT
GCCACCATCCACCTCCGAGGCGGCGGTCTCGCTGCTCCTCTCCTGCTGGACGCTACCCCTCAGATCCACGTCTTC
GCCGGAGCCGACCGAGAAGCCGTTCTGCGAGCCCTCTCTGATGGCCGAGAAGGCCTGGACGGTCCCGCCCGAGCT
GTGATTGTGGCTGGTTCTCGAGAGGAGCTCGAAGCTCGACGAGCCGACGCTCTGCGAGCCGCTGAGCGAAACGCT
CCCATGCCCGATGGCGTGTTTCTCCGAGATCGACCCGTGGCCGGAGAGCTGGCTTTTGTGTTTGCTGGCGCCGCC
GCCGCTTATCCCTCTATGGGTCGAGCCCTCGCCCTGGCCCTCCCTGACGTGATGGCTCGACTGGGCAATCGACTG
TCTACCCTGCCTGCTGTGATTGACTGGATTCATGGAGCCCCTCGAGAGCCTCGACACCCCCTCGAGCAGCTCTGG
GCTTCTTCTTTCCTCTGTCAGCTCCACGCCGAAATTACCCGAGTTGACCTCGGCCTGCGACCCTCTGCCACTATC
GGCTATTCTTCTGGCGAATCTAACGCCCTCTTCGCTATGGGAGCCTGGCGAGACCTCGACGCCATGATGCATGAA
GCCTGGACCGAGCCTCTGCTGACCACCGAGCTCGTGGGAGATTACGCCGCTGCTCGACGACTCTGGGCTCGATCC
GGTGTTACCTCCCCCGCCGCCACTTGGGCTGCTTGGGTTATCGCCGCCCCTGTTGACGATGTCCGACGAGCCCTG
ACTGGAGAGCCCCTCGCCCACCTGACCATTATTAACACTGCTGAAGAGTGTGTGATTGGAGGAGAGGCCTCCGCT
TGCGAGCGAGTCGTCGCCCGACTCGGAAAGAATCTGGCCCTCCCCCTCTCCTACGCTATGGCTTCCCATTGCCCT
GAGGTTGCCGAGGCTCGAGAGGCTTGGATTAAGATGCATCGACGAGCTACCTTTGACGTTCCCGGTGTCCGAATT
TACTCCGCTGGAGCTGAGCAGGCCTTTGCCCCTACTGCCGACGCCGCCGCTATGGCCATTACTGGTCAAGCCGTG
GACACCCTGGACTTCCCTCGAGTGATTGAGCGAGCCTATGCCGATGGCGTTCGAGTCTTCGTTGAGCACGGACCC
CGAGGACTCTGCTCCCGATGGATTGCCCGAATTCTCGGAGCCCGAGAGCATCTCGTGGTGCCCCTCGACGTTGCT
GGCCAGCAACCCGTCCGACAGCTCGTCTCTGCCGTGGCTCAGCTGATTGCCGCTGGCGTGCCCGTTGCTGCTGAC
GCTCTCCTCGCTCGACTCGCCCCTCCCCCTCCTGCTCGAGCCTCTGGACCCACCCTGACCGTCCGAGCCCACCCT
CCTGACATCCATTTTCCTACCGCTCCTGCTATGCAGCTGATGCCCCCTGCTCCTCTCCTGCCCCCTGTTCTGGCC
GACAACCCCGCCCCTCGAACTGAACGATCTTTCACTGGTCGATCCGCTCTGCGAGAGGACGCTCCTGACGGTGTT
CCCTCTCCCCGACTGCCTGTTTTCGCCCGACTGGCCCCCGTGTTTACCCATACCGAACCTGTGTTCACTCACGTG
GAGCCCGCTCTCGCCTCTCCCGAACCTGCCGTGACCCACCCTGAGCCCGTTATCGCTCATGCTGAGCCCGCCCTG
ACCCACGCCGAACCTGCTCTGGCTTCCCCCGCCGATGGCCAACACGCTTCTATCGCCGCTCGAGGAGCCGCCGTC
TCCGCCGGAATTGCTGCTGCCCACCGAGCTTACGTTGAGGCTCAGGGTGCCGTTCACGCCCGATTCCTCGCTCTC
TCTGGTTCTCTGCTGTCCCGACTGGCTGCCGCTCCCACCGCTGCCGCTTTCCCTGCTCCTGCTCCCTCTCTGGCT
CCTCAACCTGCCGCTCTGCCTGCCCCCGCTCCTCAGGCTGCTCCTCCTCAGCGACTGACCGCTCCCACTCTGGTG
GTGGCCCGACCCGTTGCTCCTCCCGCTCCTACTCCCGTGCCTGCCGCTGTGGCTGCCTTTGCCCCTGCTCCCATT
TCTGCCCCCGCCTCTACCCTGCCCGCCCCTTCTCTGCTCCGATCTACTCTCCCCGGTCCTAAATTCTCCCGAGCC
GATCTGGAGGTTCTGGCCTCTGGTTCCATCTCCTCCATCTTCGGCCCCCAATTCGCCGGACAAGACGGCTACGCC
AAGCAATGTCGAATGCCCGAGCCCCCTCTGCTCCTGGCTGACCGAGTGACTGGAATCGACGCCGTCCCCGGTTCT
ATGGATACTGGAACCATCTGGACTGAGACTGACGTCAAGCATGACTCTTGGTACCTCCACCAAGGCCGAATGCCT
ACCGGCATCGTGATCGAATCTGGCCAGGCCGATCTCCTGCTCATTTCTTGGCTGGGCATTGACGCCCTGAACAAA
GGCGAACGAGTCTACCGACTCCTGGGATGCGAGGCCACCTATCACGGCGAGCTCCCCAAGCCTGGAGACACCCTG
GTTTACGAGATTCACATTGACGGACATGCTGAGCAAGGAGACGTTCGACTGTTCTTTTTCCACTACGACTGTCGA
ATTAACGGAGAACTGCGAATTTCTGTGCGAAACGGCCAAGCTGGTTTCTTCACTTACCAGGAACTGGCCAACTCT
GGCGGTGTCCTCTGGGATGCTGACGGCGCTCCCCACGAGGCTACTGGTCCCCTCGACCCTCCCCAGGTGATCTGC
GAGAAGCAATCCTTTACCTCCGCTGAGGTCCTGGCCTTTTCTGAGGGAGACGCTTACGCCTGCTTCGGACGAGGT
TTTGAACGAGCCCAAACCCACGTCCGAACTCCCCGAATTCAGGACGGTCCTATGCGACTCCTGGATCGAGTGACT
CACTTTGATCCCCGAGGTGGACCTTGGGGTCGAGGATACCTCCGAGCTGAGGCCAAGCTCTCCCCCGATGACTGG
TTTCTGCGAGGCCATTTTAAGAACGACCCCTGCATGCCTGGCACTCTGATTCTGGAAGCCTGTGTTCAGGCTATG
GCCTTCTATCTGTGTGCTATGGGCTACTCTATCGATCGAGATGGCTGGCGATTTGAGCCTGTGCCCGGCCAAAAA
TCTCCCATGCGATGCCGAGGCCAGGCTTCCGTGACCTCCCGAGAAATTGTCTACGAGATCTTTGTCCACGAGGTC
ACCTCCGGCCCTATTCCTTCTGTTAAGGTCGATATGCTCTGTACCGTGGATGGACTGAAAGCCTTCCACGGAGAG
CGAGTTGGACTGCGACTCGTTCCCGCCTGGCCCCTGGACGACTGGCGACAAGCCCCCGTTGCCCCTCAATCTCTG
CACCGACGACCTAAGCAGATCGCTCCCCTCGGTGGCATTCTCGGACACGTTGAGACTCGACCCGTCCCCTCCGTG
AACGGTTTTGCCTTCGATTATGCCTCCCTGATCGCCTGTGCCTGGGGCAAACCCTCTCAGGCCTTCGGTCCTATG
TACGCTCCCTTTGACGGAACTCGACGAGTCGCCCGACTCCCTGGTCCCCCTTACCACTTTATCACTCGAATTATC
TCCGTCGAAGGTCCTATCGGCGTTCCCCAGCCCGGCGCCGAGGTTGAGGTCGAATACGACGTTCCTCACGATGCT
TGGTACTTTGCCGAGAACGGTGCCCGAACCATGCCCTTTTGCGTGCTCCTGGAAGCTGCCCTCCAGCCCTGCGGA
TGGCTCGCCTCCTACATTGGTTCCGCCCTGACCACTGAAGATGACCTGCTGTTCCGAAATCTCGACGGCACCGGC
ACCGTGCTGGCTGAAATCACCCCCGAATCTGGCACCCTCCGAACCCACGTGAAGGTCGTGGAGGTCTCCCAATCT
GCTGGCATGATCATCGAAACCTTCGAGGTTGAATGTTTCGTGGGCGAGGCCCACGTCTATACCCTGCGAACCGTG
TTCGGCTTCTTCCCCAAATCTGCCTTTGACAATCAGGTCGGACTGCCTGTCTCCGCCGAGGCCCGAGCTCGAATT
GAGCGACAGGCTACTTCTACCATTGATCTCGCTCCCGAACCCCCTCGATACTTTGCCGGAGCTCCTCGACTCCCT
GAGGCTATGCTGTGCATGCTCGATCGAATTACCGCTTTCGAGCCCACTGGTGGACTGAAGGGTCTCGGCTACGGT
CGAGCTGAAAAGACCGTTAAGCCCGGAGAATGGTCCTTCAAGGCCCATTTCTTTCAGGATCCCGTTCAGCCCGGA
TCCCTCGGCCTCGAAGCTCTGCTCCAGCTGCTGCAGGCTTGGATGATTGACGCTGGAGATGCCGAGGGCATCCCT
GGTGCTCGATTCGAACCCATCGCCGTGGGCCGACCCATGACCTGGAAATATCGAGGCCAGGTCGTGCCTAAGAAC

CTCCTCGTGACCGTGGAGATTGATATCACTGAGCGAGGAACTGATGACCGAGGTCGATTTGCCCTCGCTGAGGGT
TCTCTGTGGGTGGACGGCAAACGAATCTACGGCGCCAAGAATCTCGGCATGCGAGTTGTGGCCGGCCCTGAGCGA
CCTCCCCCTGGTTCTCCCCCTTCCCCTGCTTCCCCTACTGCCCCTACCGCCCCTCCCTCTCCCCCTGCTCCCACT
GTTCCCGCTGCCCTCGCCGCTCCTCCTTCCCGAGATCTGGCTTCCTTTCGATCCTTCTGGCGAGCTTCCATTGGT
GTCGGTCCTTGGCCTATTGAAGACCTCTATATCGCCCTGGTGGAGCGATTCCTCGGCCGAATCGTGATCGCTGAT
CCTGTCGCCTTCTCCCGAGTCCACGGTCGAGGCTGCCTCTATGTCGCTAACCACCAGGTCGCCCTGGAATCCCTG
ATTTTTATTATGGTCGCTTCTGCCATTTCCGGAACTCGAACTATGGCTCTCGCTAAAGCCGAGCACCAGGAATCC
TGGCTGGGCCGACTGATCACTCGAGGAATCGCCTACCCTGGTGTCACTGATCCCGGACTGATCACTTTCTTTGAC
CGAGAGGACCATGCTTCCTTTGCTCGAATTGCTGAACAGGCTACCCGAGACCTCGCCGCTTCTACCAAATCCGTG
CTGGTCCATGTTGAGGGAACTCGAGCTATTTCCTGTCGAACTCCTGTGACCCGACTGGGTGCTTCCGTTATCGAC
ATGGCTCTCGCTGCCGGAGCTCCCATCGTTCCTGTCCGATTCGCTCGAGGACTGCCTGAATCTGACTCTCCCGAA
CGACTCGACTTCCCCCTGGGTTACGGCCGACAAGACATTTACTTCGGTGCCCCTATTGAACCCGCTGAACTGGCC
CGACTGCCCCTGAAGGACCGAAAGGATGCTCTGCTCACTGCCCTGAATACCCTCGGTCCCGGAGCCTCCGATACC
CCTTCTGACCCCGATCCTTCCTTTGCCGCCCAGGTCGAGGCTTGGCGAGCTCGAACTGGTTGTGCTATCGAGGAC
GCCGTTTTCTACGCTACTCTCGCTGGCGCCTCTGCTCCCGAATCTGAAGGCGTGCAGCGACTGCTCGAGGGAGCT
CGAACTGGCAAACTCGTGATCGGCTCCGGCGAAAAGGATCGATGGCTGGGAGAATTCGCCCGATCCCTCTTCGGT
TCTGCTGACGCCACCCTGCTCGTCGAGGGAGGTTAAGCTATTTATCACTCTTTACAACTTCTACCTCAACTATCT
ACTTTAATAAATGAATATCGTTTATTCTCTATGATTACTGTATATGCGTTCCTCTAAGACAAATCGAAACCGGCA
CCTGTCCTAGG

>building block 4

CCTAGGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAG
CTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCCACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCC
AGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAACATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCC
CGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGG
CATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATT
ATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCC
ACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTG
AGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGA
CCGTTGTTCCCGCCCACCTCGATCCGGCATGCACTGATCACGGGCAAAAGTGCGTATATATACAAGAGCGTTTGC
CAGCCACAGATTTTCACTCCACACACCACATCACACATACAACCACACACATCCACAATGGCCCTGCTGGACCTG
CCCCGAGGAGAAGCTCACCTGTACTATGCTCTCGAGGACCAGATCCGAGAGCCCGCTCTGCTCGCTGCCTACGAG
GCCCTCCTGACCCCTGAGGAGCGAGCCCGAAAGCAGCGATACTATTTTGAGAAGAACCGACATGAATACCTGATT
ACTCGAGCCCTGGTCCGATCCGTGCTGTCTCGATACGCCCGAGTCGTGCCCGCTGCTTGGGCTTTCTCTGCTAAT
CAATGGGGATGCCCCGCCATTGCCTCTCCCGAGGGAACCGGTCTCCGATTTAACCTGTCCAACACCCGAGGTCTC
GTTTGCTGTCTCGTCGCTCGAGACCGAGATGTCGGCGTTGACGTGGAAGACGTGGAGCGAGATGGCGAGACCGTG
GCCATTGCCGATCGATTTTTCTCTCCCACTGAGGTTGAGGAACTCCGAGCTCAGCCTGAGCCCCGACGACGACCT
CGATTTTTCGACTACTGGACTCTGAAGGAAGCCTACATTAAAGCCCGAGGAATGGGTCTCGCCATCCCCCTCGAC
CAATTCTCTTTCCAGCTCTCTCCCGGCCGACCTATTGGCATCGCCTTTGATCCCGTTCTGGAGGATGACCCTGCC
TCCTGGCAATTCGAACAGCTGCGACCTACCCCTTCCCACCTGGTCTCCCTCGCTATCCGACGACGAGACGAAGCT
GATGTTCGAGTCCTGGCCCGACGAGTCACCCCTCTCCTCCCTGCCGCTGATGTCGTGATCGACCTGCCCGCCACT
CTCGCCGTGTAAGCTATTTATCACTCTTTACAACTTCTACCTCAACTATCTACTTTAATAAATGAATATCGTTTA
TTCTCTATGATTACTGTATATGCGTTCCTCTAAGACAAATCGAAACCTTAATTAA

CHAPTER 6

Elucidation of biosynthetic mechanisms of myxobacterial PUFA synthases

INTRODUCTION

Various microorganisms from the marine ecosystem [1-3] but also some genera of terrestrial myxobacteria [4] have been found to produce long-chain polyunsaturated fatty acids (LC-PUFAs) *de novo* from acyl-CoA precursors by iteratively acting polyketide synthase (PKS)-like PUFA synthases. These multienzyme complexes are encoded by PUFA (*pfa*) biosynthetic gene clusters, which were described for the first time from marine microbes, including *Shewanella pneumatophori* SCRC-2738 and *Photobacterium profundum* SS9 [5;6], and *Moritella marina* MP-1 [7] (Figure 1). Some of these γ-Proteobacteria produce mainly eicosapentaenoic acid (EPA, 20:5, *n*-3) (*Shewanella* and *Photobacterium*), whereas others synthesize mainly docosahexaenoic acid (DHA, 22:6, *n*-3) (*Moritella*). The correlation of the distinct biosynthetic pathways with their products was supported by heterologous expression of these *pfa* gene clusters in suitable host strains, resulting in the same LC-PUFA production profiles as in the corresponding native producer [8-10]. Similarly, the three different types of myxobacterial *pfa* gene clusters, that have been recently discovered, contribute to individual LC-PUFA production profiles, as confirmed by heterologous expression described previously ([4]; Chapter 2). Consequently, it can be concluded that the LC-PUFA products primarily synthesized by a specific PUFA synthase is determined by one or more of the *pfa* gene products. Deeper understanding of the working principles of PUFA synthases is crucial to fully exploit their potential for directed biotechnological production of selected LC-PUFAs. Up to now, several genetic and biochemical experiments have been performed with marine *pfa* gene clusters. Unfortunately, neither the exact biosynthetic function and underlying biochemistry of every catalytic domain from the PUFA synthase complex has been elucidated, nor the relevant factor(s) that determine the nature of the final LC-PUFA product have been identified.

The function of the enoyl reductase (ER) domain encoded by *pfaD* from *Shewanella oneidensis* MR-1 has been demonstrated *in vitro* by *trans*-enoyl reduction of crotonic acid tethered to the acyl carrier protein (ACP) domains of PfaA in the presence of NADPH [11]. The consecutive ACP domains are encoded in a variable number by each *pfa* gene cluster (Figure 1+2), and all of these so called tandem ACP domains of a particular PUFA synthase share a striking degree of sequence similarity (> 95% identical). Site-directed mutagenesis of the six ACPs encoded by *pfaA* from *Shewanella japonica* to abolish the substrate binding capability of the respective ACP revealed that LC-PUFA biosynthesis is an ACP-dependent process. Moreover, it could be shown that each of the tandem ACPs is functionally equivalent for LC-PUFA biosynthesis regardless of its physical location within the PfaA subunit and that

the overall LC-PUFA titre directly depends on total number of active ACPs [12]. On the basis of a hybrid PUFA synthase in which *pfaA* from *S. japonica* was replaced by its homologue from *M. marina*, it could be concluded that the LC-PUFA product distribution is regulated independent of the ACPs and the other domains of PfaA, as the product spectrum was not significantly altered [12]. However, there are several hypotheses with respect to the role of the multiple tandem ACP domains. As they may provide a scaffold for parallel LC-PUFA biosynthesis, the overall pathway flow could be increased (additive effect). Furthermore, the extra ACP domains could make the extension step (catalysed by the ketosynthase (KS) domain) a biosynthetic bottleneck, thus controlling the complete processing of PUFA intermediates.

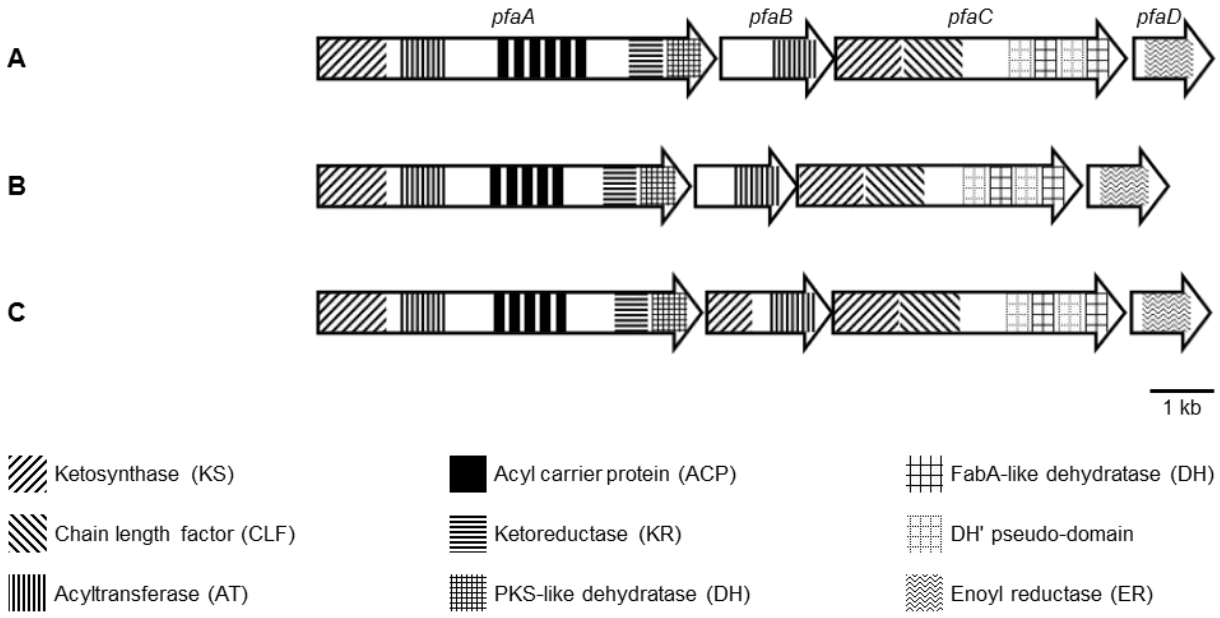


Figure 1. PUFA biosynthetic gene clusters of marine γ-Proteobacteria.
(A) *Shewanella pneumatophori* SCRC-2738 (EPA producer), (B) *Photobacterium profundum* SS9 (EPA producer), (C) *Moritella marina* MP-1 (DHA producer).

In order to obtain deeper insights into the product specificity of different PUFA synthases and to identify relevant factors directing the biosynthesis toward specific LC-PUFA products, *pfa* genes from different pathways were coexpressed followed by analysis of PUFA production profiles. These combinatorial studies were based on the advantage that PUFA synthase subunits from different origins seem to be compatible to each other allowing for functional interaction. Coexpression of the genes *pfaBCD* from *S. pneumatophori* SCRC-2738 (EPA producer) plus the complete *pfa* gene cluster from *M. marina* MP-1 (DHA producer) in *E. coli* yielded the production of both EPA and DHA [13]. Similarly,

coexpression of *pfaB* from *S. pneumatophori* SCRC-2738 plus the complete *pfa* gene cluster from *M. marina* MP-1 or coexpression of *pfaACD* from *S. pneumatophori* SCRC-2738 plus *pfaB* from *M. marina* MP-1 in *E. coli* also resulted in the production of both EPA and DHA [14]. These results suggest that *pfaB* from marine γ-Proteobacteria may play a role in determining the final LC-PUFA product. In addition to the acyltransferase (AT) domain, *pfaB* from *M. marina* MP-1 encodes an extra KS domain (Figure 1), which was found to lack the active site cysteine [15]. The question whether the KS or the AT domain (or both) determine the final LC-PUFA product in these studies could not yet be resolved. Interestingly, the three types of myxobacterial PUFA synthases, catalysing the biosynthesis of linoleic acid, EPA, or DHA as main product, do not contain any domain homologues to the inactive KS domain of PfaB. It is even more astonishing that the *pfa* gene cluster responsible for the production of oleic acid (OA, 18:1, *n*-9)/linoleic acid (LA, 18:2, *n*-6) from *Sorangium cellulosum* lacks the AT domain homologous to the AT domain encoded by *pfaB*, whereas this domain is present in the *pfa* gene cluster responsible for the production of EPA or DHA from *S. cellulosum* or *Aetherobacter* spp. (Chapter 2).

The discovery of three different types of myxobacterial PUFA pathways with different production profiles and the availability of heterologous production platforms established in previous studies ([4]; Chapter 2, 3, and 5) provided a convenient basis for experimental functional studies. The impact of selected catalytic domains and their relevance for chain length control was investigated by heterologous expression of mutated or chimeric versions of myxobacterial *pfa* gene clusters in *E. coli*, *P. putida*, and *Y. lipolytica*. In addition, efforts were undertaken to prove the function of the unique 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) domain from myxobacterial PUFA synthases *in vitro*.

EXPERIMENTAL PROCEDURES

Culture conditions

Escherichia coli DH10B [16] was used for cloning experiments. *E. coli* BL21(DE3) [17] and Rosetta 2(DE3)pLysS (Novagen) were used for heterologous expression experiments. *E. coli* HS996/pSC101-BAD-gbaA (tet^R) [18] was used for modification of a plasmid using Red/ET recombination. *E. coli* HB101/pRK2013 [19] was used as helper strain for conjugation experiments. The cells were grown in LB-medium or on LB-agar (1% tryptone, 0.5% yeast

extract, 0.5% NaCl, (1.5% agar)) at 37 °C (and 200 rpm) overnight. Antibiotics were used at the following concentrations: 100 µg/ml ampicillin, 50 µg/ml kanamycin, 20 µg/ml zeocin, and 6 µg/ml tetracycline.

Pseudomonas putida KT2440 [20] was cultivated in liquid LB-medium at 30 °C and 200 rpm overnight. Conjugation was carried out onto PMM-agar plates containing 0.61% K_2HPO_4, 0.5% KH_2PO_4, 0.1% $(NH_4)_2SO_4$, 0.66% disodium succinate, and 1.5% agar. The medium was adjusted to pH 7.0 and supplemented with 0.12% $MgSO_4$ after autoclaving. Antibiotics were used at the following concentrations: 60 µg/ml kanamycin and 30 µg/ml tetracycline.

Auxotrophic *Yarrowia lipolytica* Po1h (CLIB 882) [21] was obtained from Centre International de Ressources Microbiennes (CIRM)-Levures, Institut National de la Recherche Agronomique (INRA), AgroParisTech, Thiverval-Grignon, France. It was grown in YPD-medium or on YPD-agar containing 1% yeast extract, 2% peptone, and 2% dextrose. Prototrophic transformants were grown in minimal medium YNB containing 0.67% yeast nitrogen base (with 75% $(NH_4)_2SO_4$ and without amino acids), and 50 mM KPO_4 buffer pH 6.8. Different carbon sources were used: YNB-N_{5000} contains 1% glucose and YNBD contains 2% glucose. The cultures were incubated for 1-7 days at 20-30 °C (and 200 rpm).

General molecular biology techniques

Routine handling of nucleic acids, such as isolation of genomic and plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, were performed according to standard protocols [22]. All the enzymes were purchased from Thermo Scientific or New England Biolabs.

Cloning of expression constructs containing the *pfa* gene cluster of *A. fasciculatus* (SBSr002) with active site mutations in the target domains of *pfa3* for heterologous LC-PUFA production in *P. putida*

Five protein domains encoded by gene *pfa3* of the directly cloned *pfa* gene cluster from *A. fasciculatus* (SBSr002) were targeted for inactivation. Therefore, the wild type sequences were exchanged for mutated sequences either by conventional cloning or by linear plus circular homologous Red/ET recombineering. In the latter case, a fragment of *pfa3* with active

site mutation in the target domain as well as a zeocin resistance gene flanked by unique restriction sites were amplified by PCR. Both fragments were then fused by overlap extension PCR. The zeocin resistance gene allowed for selection after the Red/ET recombineering step and could subsequently be removed via the flanking restriction sites. In the case of C194S, a zeocin resistance gene with *Fse*I restriction sites as well as the appropriate homology arm and overlap, as the first fragment (0.6 kb), was amplified from plasmid pMyx-zeo (S. C. Wenzel, unpublished) using primers zeo^R+HA_*Fse*I_fwd and zeo^R+overlap_*Fse*I_rev (Table 1). The second fragment (0.6 kb) contained the overlapping sequence with the *Fse*I restriction site plus the introduced T to A point mutation, which was amplified from plasmid pJB*PfaAf1** (Chapter 3) using primers C194S+overlap_*Fse*I_fwd and C194S_T to A_rev (Table 1). In the case of S953A, a zeocin resistance gene with *Sca*I restriction sites (introduced by silent mutation) as well as the appropriate homology arm and overlap, as the first fragment (0.6 kb), was amplified from plasmid pMyx-zeo using primers zeo^R+HA_*Sca*I_fwd and zeo^R+overlap_*Sca*I_rev (Table 1). The second fragment (0.2 kb) contained the overlapping sequence with the *Sca*I restriction site (introduced by silent mutation) plus the introduced T to G point mutation, which was amplified from plasmid pJB*PfaAf1** using primers S953A+overlap_*Sca*I_fwd and S953A_T to G_rev (Table 1). In the case of H1755Q, a zeocin resistance gene with *Spe*I restriction sites (introduced by silent mutation) as well as the appropriate homology arm and overlap, as the first fragment (0.6 kb), was amplified from plasmid pMyx-zeo using primers zeo^R+overlap_*Spe*I_fwd and zeo^R+HA_*Spe*I_rev (Table 1). The second fragment (0.2 kb) contained the overlapping sequence with the *Spe*I restriction site (introduced by silent mutation) plus the introduced C to G point mutation, which was amplified from plasmid pJB*PfaAf1** using primers H1755Q_C to G_fwd and H1755Q+overlap_*Spe*I_rev (Table 1). In the case of H2166Q, a zeocin resistance gene with *Hind*III restriction sites as well as the appropriate homology arm and overlap, as the first fragment (0.6 kb), was amplified from plasmid pMyx-zeo using primers zeo^R+overlap_*Hind*III_fwd and zeo^R+HA_*Hind*III_rev (Table 1). The second fragment (0.5 kb) contained the overlapping sequence with the *Hind*III restriction site plus the introduced C to G point mutation, which was amplified from plasmid pJB*PfaAf1** using primers H2166Q_C to G_fwd and H2166Q+overlap_*Hind*III_rev (Table 1). In the case of E2455R, a 0.4 kb and a 0.8 kb fragment were amplified from plasmid pJB*PfaAf1** using primers E2455R_*Hind*III_fwd / E2455R_GA to CG_rev and E2455R_GA to CG_fwd / E2455R_*Pac*I_rev (Table 1). Thereby, the GA to CG mutations as well as *Hind*III and *Pac*I restriction sites were introduced. PCR amplification of the two fragments to be spliced was

performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 66 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. For the subsequent overlap extension PCR, the two outer primers without overlap and the two amplified fragments as templates were used in each case. PCR was performed as described for the amplification of the two fragments to be spliced.

In the case of C194S, S953A, and H1755Q, linear plus circular homologous Red/ET recombineering was performed as described previously [4] on the basis of plasmid pJB*PfaAf1** using 1.5 µl of the PCR product, resulting in plasmid pJB*PfaAf1**-C194S-zeo^R, pJB*PfaAf1**-S953A-zeo^R, and pJB*PfaAf1**-H1755Q-zeo^R (Table 2). In the case of H2166Q, linear plus circular homologous Red/ET recombineering was performed on the basis of plasmid pPfaAf7.1** (Chapter 3) using 1.5 µl of the PCR product, yielding plasmid pPfaAf7.1**-H2166Q-zeo^R (Table 1). The zeocin resistance gene was subsequently removed via *Fse*I restriction sites (in the case of C194S), the introduced *Sca*I restriction sites (in the case of S953A), the introduced *Spe*I restriction sites (in the case of H1755Q), or *Hind*III restriction sites (in the case of H2166Q), generating plasmids pJB*PfaAf1**-C194S, pJB*PfaAf1**-S953A, pJB*PfaAf1**-H1755Q, or pPfaAf7.1**-H2166Q (Table 2). In the case of E2455R, 1170 bp of the wild type version of *pfa3* in plasmid pPfaAf7.1** were exchanged for the PCR amplicon with the point mutations via *Hind*III and *Pac*I restriction sites, resulting in plasmid pPfaAf7.1**-E2455R (Table 2). Additionally, the backbones of plasmids pPfaAf7.1**-H2166Q and pPfaAf7.1**-E2455R were exchanged for the backbone of plasmid pJB861-*Swa*I-cm^R-*Pac*I-*cop271C* (Chapter 3) via *Swa*I and *Pac*I restriction sites, yielding plasmids pJB*PfaAf1**-H2166Q and pJB*PfaAf1**-E2455R (Table 2).

Table 1. Primers used for cloning of expression plasmids or for active site mutagenesis in this study.
The homology arms are shown in italics, the introduced restriction sites are underlined, and the introduced nucleotide exchanges are shown in bold.

Primer	Sequence (5'→3')
zeo^R+HA_*Fse*I_fwd	*CGGCCCGCCGAGGTCCGCGCGCCGCACCTCGTGGAGACAGGGCCGGCC*T CAGTCCTGCTCCTCGGCCAC
zeo^R+overlap_*Fse*I_rev	CGAAGGTCACGAGGCCGGCCGAATTCTGGACAGCAAGCGAACCG
C194S+overlap_*Fse*I_fwd	CGGTTCGCTTGCTGTCCAGAATTCGGCCGGCCTCGTGACCTTCG
C194S_T to A_rev	ACGAGGCGC**T**GGCGGCGTCG
zeo^R+HA_*Sca*I_fwd	*CGCGCTCCCCGACGTGATGGCGCGCCTCGGCAACCGGCTC*AGTACTCAG TCCTGCTCCTCGGCCAC
zeo^R+overlap_*Sca*I_rev	GTCGATCACCGCGGGCAGAGTACTGAATTCTGGACAGCAAGCGAACCG
S953A+overlap_*Sca*I_fwd	CGGTTCGCTTGCTGTCCAGAATTCAGTACTCTGCCCGCGGTGATCGAC
S953A_T to G_rev	CGGAGG**C**GTAGCCGATGGTCG
zeo^R+overlap_*Spe*I_fwd	CGGCCAGGCCAGCGTGACTAGTGAATTCTGGACAGCAAGCGAACCG
zeo^R+HA_*Spe*I_rev	*TCGTGACCTCGTGGACGAAGATCTCGTAGACGATCTCCCG*ACTAGTCAG TCCTGCTCCTCGGCCAC
H1755Q_C to G_fwd	GCGAGGGCA**G**TTCAAGAACGATCC
H1755Q+overlap_*Spe*I_rev	CGGTTCGCTTGCTGTCCAGAATTCACTAGTCACGCTGGCCTGGCCG
zeo^R+overlap_*Hind*III_fwd	CGATCTCGCCTCGTTTCGAAGCTTGAATTCTGGACAGCAAGCGAACCG
zeo^R+HA_*Hind*III_rev	*CTCGATCGGCCAGGGCCCGACCCCGATCGAGGCGCGCCAG*AAGCTTCAG TCCTGCTCCTCGGCCAC
H2166Q_C to G_fwd	AAGCCCA**G**TTCTTCCAGGATCCCG
H2166Q+overlap_*Hind*III_rev	CGGTTCGCTTGCTGTCCAGAATTCAAGCTTCGAAACGAGGCGAGATCG
E2455R_*Hind*III_fwd	AAGCTTCTGGCGCGCCTCGATCGGG
E2455R_GA to CG_rev	GGCGCGCGTGCCC**CG**GACGTGGACGAGC
E2455R_GA to CG_fwd	GCTCGTCCACGTC**CG**GGGCACGCGCGCC
E2455R_*Pac*I_rev	TTAATTAAAGACGGTGAAATCGTTCGACGTGTGCCTC
AT_So026_2_*Apa*LI_fwd	GGGCGTGCACTCGCGCTGGCGCTG
AT_So026_2_*Fsp*I_rev	GCGGATGCGCAGGGATCTTCAGCGTGG
*AfAGPAT_Bam*HI_fwd_2	GCGTCGGATCCGTCGCCGGCCCCGAGC
*AfAGPAT_Not*I_rev_2	CACGCTGCGGCCGCTCAGCCGCCCTCCACC

Table 2. Plasmids constructed in this study.
oriV = origin of replication, *trfA* = gene encoding replication initiation protein in RK2 replicons, *oriT* = origin of transfer, *aph(3')-Ia* = kanamycin resistance gene, *ble* = zeocin resistance gene, *bla* = ampicillin resistance gene, PTn5 = Tn5 promoter.

Plasmid	Characteristics
pJB*PfaAf1**-C194S-zeo^R	Derivative of pJB*PfaAf1** (Chapter 3) in which PTn5-*ble* plus a C194S mutation were introduced into *pfa3* by Red/ET recombineering. RK2 *oriT*, RK2 *oriV*, *trfA**, *aph(3')-Ia*
pJB*PfaAf1**-C194S	Derivative of pJB*PfaAf1**-C194S-zeo^R in which PTn5-*ble* was removed via *Fse*I
pJB*PfaAf1**-S953A-zeo^R	Derivative of pJB*PfaAf1** in which PTn5-*ble* plus a S953A mutation flanked by *Sca*I restriction sites were introduced into *pfa3* by Red/ET recombineering
pJB*PfaAf1**-S953A	Derivative of pJB*PfaAf1**-S953A-zeo^R in which PTn5-*ble* was removed via *Sca*I
pJB*PfaAf1**-H1755Q-zeo^R	Derivative of pJB*PfaAf1** in which PTn5-*ble* plus an H1755Q mutation flanked by *Spe*I restriction sites were introduced into *pfa3* by Red/ET recombineering
pJB*PfaAf1**-H1755Q	Derivative of pJB*PfaAf1**-H1755Q-zeo^R in which PTn5-*ble* was removed via *Spe*I
pPfaAf7.1**-H2166Q-zeo^R	Derivative of pPfaAf7.1** (Chapter 3) in which PTn5-*ble* plus an H2166Q mutation were introduced into *pfa3* by Red/ET recombineering. p15A *oriV*, *bla*, *ble*
pPfaAf7.1**-H2166Q	Derivative of pPfaAf7.1**-H2166Q-zeo^R in which PTn5-*ble* was removed via *Hind*III
pJB*PfaAf1**-H2166Q	Derivative of pPfaAf7.1**-H2166Q in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I-*cop271C* (Chapter 3) via *Swa*I and *Pac*I
pPfaAf7.1**-E2455R	Derivative of pPfaAf7.1** in which 1170 bp of *pfa3* were replaced with the same 1170 bp of gene *pfa3* except for an E2455R mutation via *Hind*III and *Pac*I
pJB*PfaAf1**-E2455R	Derivative of pPfaAf7.1**-E2455R in which the vector backbone was exchanged for the vector backbone of pJB861-*Swa*I-cm^R-*Pac*I-*cop271C* (Chapter 3) via *Swa*I and *Pac*I
pSynHybPfa1	Derivative of pPm**SynPfaAf2 (Chapter 4) in which the KS2 plus the CLF domain of gene *pfa3* originating from *A. fasciculatus* (SBSr002) were replaced by KS/CLF_So026_2 via *Fse*I and *Apa*LI. RK2 *oriT*, RK2 *oriV*, *trfA**, *aph(3')-Ia*
pSynHybPfa2	Derivative of pPm**SynPfaAf2 in which the AT2 domain of gene *pfa3* originating from *A. fasciculatus* (SBSr002) was replaced by the AT2 domain of gene *pfa3* from *S. cellulosum* SBSo026 via *Apa*LI and *Fsp*I
pSynHybPfa3	Derivative of pSynHybPfa2 in which the KS2 plus the CLF domain of gene *pfa3* originating from *A. fasciculatus* (SBSr002) were replaced by KS/CLF_So026_2 via *Fse*I and *Apa*LI
pSynPfaPptAf2-hyb*pfa3*	Derivative of pSynPfaPptAf2 (Chapter 5) in which hp4d, the KS2, the CLF, plus the AT2 domain of gene *pfa3* originating from *A. fasciculatus* (SBSr002) was replaced by building block 5 via *Acl*I and *Kpn*I
pSynPfaPptAf2mc- hyb*pfa3*	Derivative of pSynPfaPptAf2-hyb*pfa3* in which hp4d-*mcherry-LIP2*t was inserted via *Acl*I
pGEX-6P-1-*AfAGPAT*+linker	Derivative of pGEX-6P-1 (GE Healthcare) in which *AfAGPAT* flanked by 129 bp and 135 bp linker sequences was inserted via *Bam*HI and *Not*I

Sequence analysis and design of the synthetic gene clusters

The KS2 and chain length factor (CLF) domains of gene *pfa3* from the arachidonic acid (AA, 20:4, *n*-6)/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026 were analysed and compared to the genome sequence of *P. putida* KT2440 [23] retrieved from NCBI Genome RefSeq NC_002947 in collaboration with ATG:biosynthetics. Based on this, relevant parameters for constructional and functional sequence design were defined to generate an artificial DNA sequence comprising the KS2/CLF domains analogous to the artificial docosapentaenoic acid (DPA, 22:5)/DHA-type *pfa* gene cluster version 2a (Chapter 4) using the proprietary evoMAG-GeneDesign software [24].

The KS2, CLF, and AT2 domains of gene *pfa3* from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026 were analysed and compared to the genome sequence of *Y. lipolytica* CLIB 122 [25] retrieved from NCBI Genome RefSeq NC_006067, NC_006068, NC_006069, NC_006070, NC_006071, and NC_006072. Based on this, relevant parameters for constructional and functional sequence design were defined to generate an artificial DNA sequence comprising the KS2/CLF/AT2 domains analogous to the artificial DPA/DHA-type *pfa* gene cluster (Chapter 5) using the proprietary evoMAG-GeneDesign software [24].

Cloning of expression constructs containing hybrid *pfa* gene clusters for heterologous LC-PUFA production in *P. putida, E. coli,* or *Y. lipolytica*

For the heterologous expression of artificial and hybrid PUFA biosynthetic pathways in *P. putida* KT2440, building block KS/CLF_So026_2, comprising the KS2 plus CLF domain of gene *pfa3* originating from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026 and adapted to *P. putida* KT2440, was designed and supplied by gene synthesis companies. The building block contains a 40 bp intergenic region comprising a Shine-Dalgarno sequence at the 5' end and is flanked by the unique *Fse*I and *Apa*LI restriction sites for cloning purposes, which were excluded from any other unwanted position within the sequence (see appendix for sequence). Plasmid pSynHybPfa1 (Table 2) was constructed by exchange of the KS2 plus CLF domain of gene *pfa3* of the artificial *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) located on plasmid pPm**SynPfaAf2 (Chapter 4) for the building block KS/CLF_So026_2 via *Fse*I and *Apa*LI restriction sites.

The AT2 domain of gene *pfa3* from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026 (1 kb) was amplified from plasmid pHybPfa1-mx9.2 (Chapter 2) using primers

AT_So026_2_*Apa*LI_fwd and AT_So026_2_*Fsp*I_rev (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The PCR amplicon served for replacement of the AT2 domain of gene *pfa3* from the *pfa* gene cluster of *A. fasciculatus* (SBSr002) located on plasmid pPm**SynPfaAf2 (Chapter 4) via *Apa*LI and *Fsp*I restriction sites, yielding plasmid pSynHybPfa2 (Table 2). Plasmid pSynHybPfa3 (Table 2) was constructed by exchange of the KS2 plus the CLF domain of gene *pfa3* from the artificial *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) located on plasmid pSynHybPfa2 for the building block KS/CLF_So026_2 via *Fse*I and *Apa*LI restriction sites.

For the heterologous expression of an artificial and hybrid PUFA biosynthetic pathway in *Y. lipolytica* Po1h, building block 5, comprising the KS2, the CLF, plus the AT2 domain of gene *pfa3* originating from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026, was designed and supplied by gene synthesis companies. The building block additionally contains a 134 bp non-coding intergenic linker at the 5' end, the strong hybrid hp4d promoter [26] for *Y. lipolytica* and is flanked by the unique *Acl*I and *Kpn*I restriction sites for cloning purposes, which were excluded from any other unwanted position within the sequence (see appendix for sequence). Plasmid pSynPfaPptAf2-hyb*pfa3* (Table 2) was constructed by exchange of hp4d, domains KS2, CLF, plus AT2 of gene *pfa3* of the artificial *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) located on plasmid pSynPfaPptAf2 (Chapter 5) for the building block 5 via *Acl*I and *Kpn*I restriction sites. Plasmid pSynPfaPptAf2mc-hyb*pfa3* (Table 2) was constructed by insertion of hp4d-*mcherry-LIP2*t from plasmid pSynPfaPptAf2mc (Chapter 5) into plasmid pSynPfaPptAf2-hyb*pfa3* via *Acl*I restriction sites.

Transformation of *P. putida* by triparental conjugation and heterologous expression

Expression plasmids were transferred into *P. putida* KT2440 by triparental conjugation. 1 ml of an overnight culture of the acceptor strain *P. putida* KT2440::pPptAf1, the donor strain *E. coli* DH10B containing the final expression construct, and the helper strain *E. coli* HB101/pRK2013 were harvested at 13,000 rpm for 1 min. After discarding the supernatant, the cells were resuspended in 1 ml LB-medium in each case. The cells were spun down at 13,000 rpm for 1 min, and the supernatant was discarded. This wash step was repeated once.

The cells were resuspended in 300 µl LB-medium in each case. The triparental mating was performed by combining 50 µl of each 300 µl suspension on a small area of a LB-agar plate [27]. After incubation at 37 °C for 4 h, the plate was transferred to 30 °C and incubated overnight. The cells were scraped from the plate and resuspended in 1 ml sterile ddH_2O. 10-100 µl of this suspension was then plated out on PMM-agar containing 60 µg/ml kanamycin and 30 µg/ml tetracycline. The *E. coli* strains were additionally counterselected by 150 µg/ml ampicillin, which *P. putida* is resistant to. The plates were incubated at 30 °C for 1-2 days. Transconjugants were verified by plasmid isolation and restriction analysis of the replicative plasmids. The constructed strains are listed in Table 3.

Transformation of *Y. lipolytica* by lithium acetate-mediated heat shock transformation and heterologous expression

Transformation of *Y. lipolytica* was carried out using a protocol developed by M.-T. Le Dall, modified by C. Madzak (Laboratoire de Microbiologie de l'Alimentation au Service de la Santé (MICALIS), AgroParisTech, Thiverval-Grignon, France; personal communication). One loopful of *Y. lipolytica* cells from a YPD-agar plate grown at 30 °C overnight were resuspended in 1 ml TE buffer in a sterile tube. The cells were centrifuged at 10,000 rpm for 1 min, and the supernatant was discarded. After resuspension in 600 µl 0.1 M lithium acetate pH 6.0, the cells were incubated at 28 °C for 1 h in a water bath. The samples were centrifuged at 3,000 rpm for 2 min, the supernatant was discarded, and the cells were softly resuspended in 80 µl 0.1 M lithium acetate pH 6.0. 40 µl of competent cells were mixed with 2 µl herring testes carrier DNA (10 mg/ml in TE buffer, denatured) and 3 µl linearized plasmid DNA to be transformed. The samples were incubated at 28 °C for 15 min in a water bath, and 350 µl 40% PEG 4000 in 0.1 M lithium acetate pH 6.0 plus 16 µl 1 M dithiothreitol (DTT) were added. After incubation of the cells at 28 °C for 1 h in a water bath, 40 µl DMSO were added. Subsequently, heat shock was carried out at 39 °C for 10 min in a heating block. 600 µl 0.1 M lithium acetate pH 6.0 were then added, and the cells were plated onto YNB-N_{5000}-agar. The plates were incubated at 30 °C for 3 days. Thereafter, selected colonies were transferred onto new YNB-N_{5000}-agar plates. The constructed strains are listed in Table 3. Expression was carried out in 50 ml YNBD-medium inoculated with an overnight culture starting from OD_{600} = 0.1. The cells were cultivated at 28 °C and 200 rpm for 76-168 h and then harvested at 4,000 rpm for 5 min.

Heterologous LC-PUFA production in *E. coli* and *P. putida*

E. coli BL21(DE3) was cotransformed with the expression constructs pSynHybPfa1 plus pPptAfA (Chapter 3), pSynHybPfa2 plus pPptAfA, or pSynHybPfa3 plus pPptAfA. The constructed strains are listed in Table 3. For heterologous LC-PUFA production, cultivation of the strains was carried out in 50 ml LB-medium containing 50 µg/ml kanamycin and 100 µg/ml ampicillin. The medium was inoculated with an overnight culture (1:100) and incubated at 37 °C. Expression of the Pfa proteins and AfPpt was induced at OD_{600} of 0.4-0.8 by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C and 200 rpm for 24 h and then harvested at 8,000 rpm for 5 min.

For heterologous LC-PUFA production in *P. putida*, cultivation was carried out in 50 ml LB-medium containing 60 µg/ml kanamycin and 30 µg/ml tetracycline. The medium was inoculated with an overnight culture (1:100) and incubated at 30 °C for 4 h. Expression of the Pfa proteins and AfPpt was induced by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C and 200 rpm for 24 h and then harvested at 8,000 rpm for 5 min.

Cloning of an expression construct containing the AGPAT domain encoded by the *pfa* gene cluster of *A. fasciculatus* (SBSr002), heterologous expression in *E. coli*, and protein purification

The AGPAT domain flanked by 129 bp and 135 bp linker sequences of gene *pfa3* from the *pfa* gene cluster of *A. fasciculatus* (SBSr002) (1.1 kb) was amplified from plasmid pPfaAf7.1* (Chapter 3) using primers *AfAGPAT_Bam*HI_fwd_2 and *AfAGPAT_Not*I_rev_2 (Table 1). PCR was performed with Phusion DNA polymerase (Thermo Scientific) and standard conditions according to the manufacturer's protocol. The reactions contained 5% DMSO and were carried out in an Eppendorf Mastercycler under the following conditions: initial denaturation for 2 min at 98 °C; 30 cycles consisting of denaturation for 15 s at 98 °C, annealing for 20 s at 72 °C, and extension for 10 s/kb at 72 °C; and a final extension for 10 min at 72 °C. The PCR amplicon was inserted into plasmid pGEX-6P-1 (GE Healthcare) via *Bam*HI and *Not*I restriction sites, yielding plasmid pGEX-6P-1-*AfAGPAT*+linker (Table 2).

E. coli Rosetta 2(DE3)pLysS was transformed with the expression construct pGEX-6P-1-*AfAGPAT+linker*. The constructed strain is listed in Table 3. For heterologous

protein production, cultivation of the strain was carried out in 50 ml LB-medium containing 34 µg/ml chloramphenicol and 100 µg/ml ampicillin. The medium was inoculated with an overnight culture (1:100) and incubated at 37 °C. Expression of the GST-tagged protein AfAGPAT flanked by 43 aa and 45 aa linker sequences was induced at OD_{600} of 1.0 by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After induction, the cells were cultivated at 16 °C and 200 rpm for 19 h and then harvested at 8,000 rpm for 5 min.

Purification of GST-AfAGPAT+linker was carried out at 4 °C using the GST SpinTrap Purification Module (GE Healthcare). The cell pellet was resuspended in 1.5 ml PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 1 mM DTT; pH 7.3), and the cells were broken by three passes through a French Press (700 psi). The insoluble material was removed from the lysate by centrifugation at 14,000 rpm for 45 min at 4 °C. After loading 600 µl lysate onto a SpinTrap column (GE Healthcare) equilibrated with PBS, the sample was agitated on a tube rotator at 15 rpm for 1 h at 4 °C. The column was centrifuged at 3,500 rpm for 1 min, the flow-through was discarded, and the loading step was repeated. Afterwards, the column was washed by adding 600 µl PBS, centrifuging at 3,500 rpm for 1 min, and discarding the flow-through. The wash step was repeated twice. Removal of the GST-tag was performed by on-column enzymatic cleavage with 20 U PreScission Protease (GE Healthcare) in 150 ml cleavage buffer (50 mM Tris-Cl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT) at 4 °C overnight. The eluate was collected by centrifugation at 3,500 rpm for 1 min and analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Purified protein AfAGPAT+linker was then mixed with glycerol to a final concentration of 10% (v/v), flash frozen in liquid nitrogen, and stored at -80 °C. Determination of the protein concentration using the Bradford assay (Bio-Rad) revealed that 180 mg of purified protein AfAGPAT+linker were obtained from 50 ml cell culture. The protein identity was confirmed by ESI-qToF-MS analysis.

Table 3. Expression strains constructed in this study.

Strain	Characteristics
P. putida KT2440::pPptAf1 + pJB*PfaAf1**-C194S	*P. putida* KT2440::pPptAf1 (Chapter 3) with replicative plasmid pJB*PfaAf1**-C194S
P. putida KT2440::pPptAf1 + pJB*PfaAf1**-S953A	*P. putida* KT2440::pPptAf1 with replicative plasmid pJB*PfaAf1**-S953A
P. putida KT2440::pPptAf1 + pJB*PfaAf1**-H1755Q	*P. putida* KT2440::pPptAf1 with replicative plasmid pJB*PfaAf1**-H1755Q
P. putida KT2440::pPptAf1 + pJB*PfaAf1**-H2166Q	*P. putida* KT2440::pPptAf1 with replicative plasmid pJB*PfaAf1**-H2166Q
P. putida KT2440::pPptAf1 + pJB*PfaAf1**-E2455R	*P. putida* KT2440::pPptAf1 with replicative plasmid pJB*PfaAf1**-E2455R
E. coli BL21(DE3)/ pSynHybPfa1/pPptAfA	*Escherichia coli* BL21(DE3) with replicative plasmids pSynHybPfa1 and pPptAfA (Chapter 3)
E. coli BL21(DE3)/ pSynHybPfa2/pPptAfA	*E. coli* BL21(DE3) with replicative plasmids pSynHybPfa2 and pPptAfA
E. coli BL21(DE3)/ pSynHybPfa3/pPptAfA	*E. coli* BL21(DE3) with replicative plasmids pSynHybPfa3 and pPptAfA
E. coli Rosetta 2(DE3)pLysS/ pGEX-6P-1-*AfAGPAT*+linker	*E. coli* Rosetta 2(DE3)pLysS with replicative plasmid pGEX-6P-1-*AfAGPAT*+linker
Y. lipolytica Po1h::pSynPfaPptAf2mc-hyb*pfa3* clones	*Y. lipolytica* Po1h with pSynPfaPptAf2mc-hyb*pfa3* randomly integrated in the genome

Extraction of cellular fatty acids

The cellular fatty acids were extracted using the FAME method [28]. For this purpose, 50 ml of a culture were harvested at 8,000 rpm for 10 min at room temperature. The cell pellet was transferred to a glass vial and dried in a vacuum concentrator. Subsequently, the cell dry weight was determined. 50 µg of *n*-6 DPA or *n*-3 DPA (Sigma-Aldrich) and 500 µl of a mixture of methanol, toluene, and sulphuric acid (50:50:2, v/v/v) were added. The vial was capped with a teflon-lined screw cap and incubated at 80 °C for 24-48 h. After the mixture was cooled to room temperature, 400 µl of an aqueous solution consisting of 0.5 M NH_4HCO_3 and 2 M KCl were added, and the sample was vortexed for 30 s. Phase separation was achieved by centrifugation at 4,000 rpm for 5 min at room temperature. 100 µl of the upper phase were used for GC-MS analysis.

Analysis of fatty acid methyl esters by GC-MS

GC-MS was carried out on an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a 7683B split/splitless injector with autosampler (Agilent Technologies) and

coupled to a 5973 electron impact mass selective detector (Agilent Technologies). Helium was used as the carrier gas at a flow rate of 1 ml/min. 1 µl of the sample was injected in split mode (split ratio, 10:1). The analytical column was a (5% phenyl)-methylpolysiloxane capillary column (Agilent J&W DB-5ht; 30 m x 0.25 mm i.d. x 0.1 µm film thickness, maximum temperature 400 °C; Agilent Technologies). The column temperature was kept at 130 °C for 2.5 min, increased to 240 °C at a rate of 5 °C/min, then ramped to 300 °C at 30 °C/min, and held at 300 °C for 5 min. Other temperatures were as follows: inlet, 275 °C; GC-MS transfer line, 280 °C; ion source, 230 °C; and quadrupole, 150 °C. The mass selective detector was operated in scan mode, scanning the mass range from *m*/*z* 40 to 700. Scan control, data acquisition, and processing were performed by MSD ChemStation and AMDIS software, version 2.68, based on the fragmentation patterns and retention times, in comparison with Supelco 37 Component FAME Mix and LC-PUFAs (all Sigma-Aldrich) as reference standards, and NIST 08 library. Absolute amounts of PUFAs were quantified by integration of the peaks using MSD ChemStation and by subsequent calculation in relation to the integral of DPA and to cell dry weight.

Fluorescence measurements

The fluorescence of mCherry produced by yeast cells grown on YNB-N_{5000}-agar plates was determined using a Molecular Dynamics Typhoon 9410 Variable Mode Imager (Amersham Biosciences/GE Healthcare). An excitation wavelength of 532 nm ± 8 nm and an emission wavelength of 580 nm ± 15 nm were used.

Synthesis of a docosahexaenoic acid *N*-acetylcysteamine thioester (DHA-NAC thioester)

DHA methyl ester (1 mg, 0.003 mmol, 1 equiv.) dissolved in 100 µl heptane was mixed with LiOH (3.6 mg, 0.15 mmol, 50 equiv.) dissolved in 300 µl of a mixture of methanol and ddH_2O (50:50, v/v) and stirred at room temperature for 4 days. Thereafter, 300 µl 3 M potassium acetate pH 5.5 were added, and the reaction mixture was extracted with 100 µl ethyl acetate for three times. The organic phase was evaporated by using a rotary evaporator, followed by a gentle stream of nitrogen, and concentrated to dryness *in vacuo*. The residue was dissolved in 1 ml dichloromethane. Dimethylaminopyridine (1 crystal, cat.), *N*,*N'*-diisopropylcarbodiimide (1.6 mg, 0.012 mmol, 4 equiv.), and *N*-acetylcysteamine (NAC; 2.2 mg, 0.018 mmol, 6 equiv.) were added, and the reaction mixture was stirred, protected

from light, under a gentle stream of nitrogen for 3 h. Afterwards, 2 ml ethyl acetate and 2 ml 10% HCl were added. The organic phase was mixed with 2 ml of a saturated solution of $NaHCO_3$. After phase separation, the organic phase was evaporated to dryness under a gentle stream of nitrogen, and the residue was dissolved in 60 µl methanol. The crude product was purified by semi-preparative HPLC. Purified all-*cis* DHA-NAC thioester (0.14 mg, 0.0003 mmol, 10% yield) was dissolved in methanol and stored at -20 °C.

In vitro enzyme assay for the AGPAT domain encoded by the *pfa* gene cluster of *A. fasciculatus* (SBSr002)

In order to assay for transfer of DHA bound onto an ACP (mimicked by NAC thioester) to the 2-position of 1-acyl-*sn*-glycerol-3-phosphate catalysed by AGPAT, reactions (50 µl) contained 50 mM Tris-Cl or HEPES pH 7.0, 7.4, or 8.0, (5 mM $MgCl_2$), 1 mg/ml bovine serum albumin (BSA; essentially fatty acid free), 20-100 µM oleoyl-L-α-lysophosphatidic acid, 20 µM DHA-NAC thioester, and 2 µM purified protein AfAGPAT+linker. The reactions were incubated at 30 °C for 30-120 min or overnight, quenched by the addition of 50 µl methanol, and then analysed by high resolution nano-ESI-FT-Orbitrap-MS analysis. Measurements were performed on a Dionex Ultimate 3000 RSLC system (Thermo Scientific) coupled to a Thermo Fisher Orbitrap (Thermo Scientific) supported by an Advion Triversa Nanomate nano-ESI system (Advion). 1 µl of the sample was injected. The analytical column was a BEH RP-C18 column (50 x 2.1 mm, 1.7 µm particle size; Waters). Separation was achieved by a linear gradient with (A) ddH_2O containing 0.1% formic acid to (B) acetonitrile containing 0.1% formic acid at a flow rate of 600 µl/min and 45 °C. The gradient was initiated by a 0.33 min isocratic step at 5% B, followed by an increase to 95% B in 9 min to end up with a 4 min flush step at 95% B before reequilibration with initial conditions. UV and MS detection were performed simultaneously. Mass spectra were acquired in centroid mode ranging from *m*/*z* 200 to 2000 at a resolution of R = 30000.

RESULTS AND DISCUSSION

PUFA synthases are multienzyme complexes for *de novo* production of LC-PUFAs. They consist of several catalytic domains, many of which have not yet been investigated in detail in

terms of their exact function and biochemistry during PUFA assembly. Thus, selected myxobacterial protein domains were subjected to further analyses (Figure 2-4).

A powerful traditional approach for studying essentiality and redundancy of catalytic protein domains represents their targeted inactivation by site-specific mutagenesis of the active site residues [29]. Hence, an active site amino acid of the KS2 domain, the AT2 domain, the dehydratase (DH) domains DH2 and DH3, or the AGPAT domain was substituted by a catalytically inactive amino acid via introduction of point mutations by overlap extension PCR and subsequent Red/ET recombineering. In case of the KS2 domain, the active site cysteine was replaced by a serine (C194S) [30]; for the AT2 domain, the active site serine was replaced by an alanine (S953A) [31]; for the FabA-like DH domains, the active site histidine was replaced by a glutamine (H1755Q, H2166Q) [32]; and for the AGPAT domain, the active site glutamate was replaced by an arginine (E2455R) [33] (Figure 2). *P. putida* KT2440 is known as a fast-growing, reliable, and robust host organism for recombinant LC-PUFA production (Chapter 3) and was therefore chosen as chassis for expression of the *pfa* genes encoding the mutated versions of PUFA synthases.

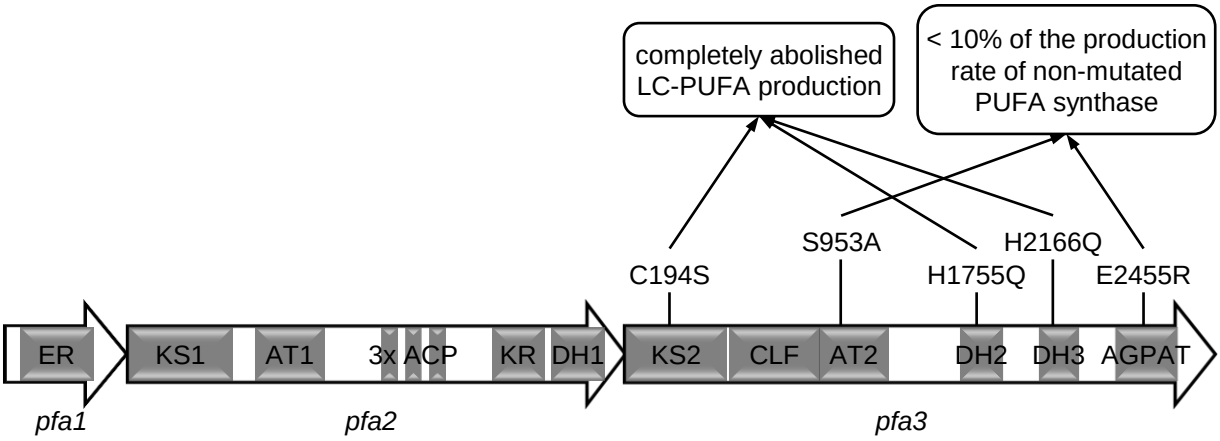


Figure 2. Active site mutations introduced into selected domains encoded by the *pfa* gene cluster of *Aetherobacter fasciculatus* (SBSr002) and their impact on heterologous LC-PUFA production in *Pseudomonas putida* KT2440.
ER = Enoyl reductase, KS = Ketosynthase, AT = Acyltransferase, ACP = Acyl carrier protein, KR = Ketoreductase, DH = Dehydratase, CLF = Chain length factor, AGPAT = 1-Acylglycerol-3-phosphate *O*-acyltransferase.

Due to restriction site requirements during cloning procedures, site-directed mutagenesis of *pfa3* was performed on two different plasmids constructed previously (Chapter 3), which both contain the native DPA (22:5, *n*-6 or *n*-3)/DHA (22:6, *n*-3)-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) but differ in their vector backbone. Plasmid pJB*PfaAf1** was

used in case of C194S, S953A, as well as H1755Q and plasmid pPfaAf7.1** in case of H2166Q and E2455R, which contain the directly cloned DPA (22:5, *n*-6 or *n*-3)/DHA (22:6, *n*-3)-type *pfa* gene cluster of *A. fasciculatus* (SBSr002). The plasmids were modified by linear plus circular homologous Red/ET recombination [34] to exchange base pairs in the catalytic core regions of *pfa3* and to introduce a zeocin resistance gene, resulting in plasmids pJB*PfaAf1**-C194S-zeo^R, pJB*PfaAf1**-S953A-zeo^R, pJB*PfaAf1**-H1755Q-zeo^R, or pPfaAf7.1**-H2166Q-zeo^R. The zeocin resistance cassette was deleted via restriction digest and religation to generate the derivatives pJB*PfaAf1**-C194S, pJB*PfaAf1**-S953A, pJB*PfaAf1**-H1755Q, and pPfaAf7.1**-H2166Q. By applying conventional cloning techniques, base pairs were exchanged in the active site of the AGPAT domain of *pfa3*, yielding plasmid pPfaAf7.1**-E2455R. In a final engineering step, the backbones of the obtained plasmids pPfaAf7.1**-H2166Q and pPfaAf7.1**-E2455R were replaced by the vector backbone of pJB*PfaAf1**. It is derived from the broad-host-range expression vector pJB861 with a copy-up point mutation [35]. Heterologous expression of the *pfa* genes is driven by the *xylS-Pm* promoter system with mutations in the 5'-untranslated region [36] plus the core region [37] of *Pm* promoter. The resulting plasmids pJB*PfaAf1**-C194S, pJB*PfaAf1**-S953A, pJB*PfaAf1**-H1755Q, pJB*PfaAf1**-H2166Q, or pJB*PfaAf1**-E2455R were transferred into *P. putida* KT2440::pPptAf1 by triparental conjugation. This strain overexpresses the 4'-phosphopantetheinyl transferase (PPTase) AfPpt from *A. fasciculatus* (SBSr002) to increase the phosphopantetheinylation of PUFA synthases (Chapter 3). The transgenic clones were cultivated in 50 ml LB-medium at 30 °C for 4 h. Expression of the Pfa proteins plus AfPpt was induced by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C for 24 h. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [28] and analysed by GC-MS. Strain *P. putida* KT2440::pPptAf1 + pJB*PfaAf1** (Chapter 3) harbours the non-mutated DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002), produces 0.1 mg *n*-6 DPA/g cell dry weight (CDW) plus 0.6 mg DHA/g CDW, and was used as a reference. Recombinant LC-PUFA production is completely abrogated for the strains *P. putida* KT2440::pPptAf1 + pJB*PfaAf1**-C194S, *P. putida* KT2440::pPptAf1 + pJB*PfaAf1**-H1755Q, and *P. putida* KT2440::pPptAf1 + pJB*PfaAf1**-H2166Q, and, in addition, no short-chain biosynthetic intermediate can be detected via GC-MS (Figure 2). In the other two expression strains, PUFA production can still be observed but is significantly reduced compared to the experiment with the non-mutated PUFA synthase (Figure 2). Strain *P. putida* KT2440::pPptAf1 + pJB*PfaAf1**-S953A produces 0.007 mg *n*-6 DPA/g CDW

plus 0.05 mg DHA/g CDW, and strain *P. putida* KT2440::pPptAf1 + pJB*PfaAf1**-E2455R produces 0.008 mg *n*-6 DPA/g CDW plus 0.04 mg DHA/g CDW. The profile of the intrinsic fatty acids does not alter among the recombinant *P. putida* strains with different mutated expression constructs. In conclusion, the KS2 domain and the FabA-like DH domains DH2 and DH3 within Pfa3 appear to be essential and not substitutable by the homologous domains of Pfa2. Moreover, domains DH2 and DH3 of Pfa3 are obviously not able to substitute each other. After inactivation of the AT2 domain or the AGPAT domain, there is still some very low LC-PUFA production (< 10% of the production rate of non-mutated PUFA synthase). This could result from either some residual activity of the mutated domains or from complementation of their function by other enzymes. In case of the mutated APGAT domain, it is also conceivable that the fully processed LC-PUFA chains are released from the assembly line by a less efficient, spontaneous hydrolysis.

The factor determining the chain length of the LC-PUFAs within the PUFA synthase has not yet been elucidated. A useful approach for dissecting the molecular basis of the specificity of PKS-catalysed reactions as well as for generation of new promising polyketide derivatives represents the expression of hybrid PKS. Generation of chimeric genes by swapping of homologous or heterologous domains has been applied to both modular polyketide synthase [38-40] and iterative polyketide synthases [41;42]. By taking a closer look at the catalytic domains of myxobacterial PUFA synthases (Figure 3), the KS2/CLF pair of Pfa3 proved to be a good candidate for the determinant of the final LC-PUFA molecular species, especially because it is well known that the CLF of type II PKS systems is the primary determinant of carbon chain length of the type II PKS products [43-45]. Besides, a recent study suggests that the AT domain encoded by *pfaB* from marine γ-Proteobacteria, which is homologous to domain AT2 of myxobacterial PUFA synthases, may play a role in determining the length of the final PUFA product [14]. In order to study the impact of KS2/CLF and AT2 domains on the PUFA production profile, chimeric Pfa3 versions were constructed by combining domains from two different myxobacterial PUFA synthases. Thus, the KS2/CLF domains, the AT2 domain, or the KS2/CLF domains plus the AT2 domain of the DPA/DHA-type PUFA synthase from *A. fasciculatus* (SBSr002) were replaced for the homologous domains of the AA (20:4, *n*-6)/EPA (20:5, *n*-3)-type PUFA synthase from *S. cellulosum* SBSo026 (Figure 3). The resulting constructs were then transferred and expressed into *P. putida*, *E. coli*, and *Y. lipolytica*.

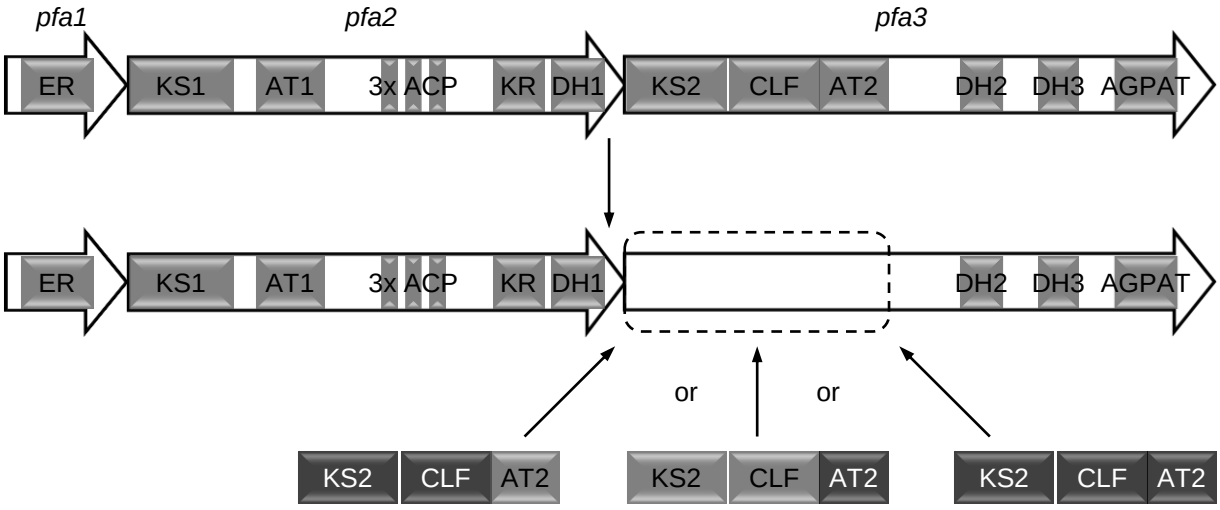


Figure 3. Construction of three hybrid PUFA synthases by exchange of the domains KS2-CLF, AT2, or KS2-CLF-AT2 of the DPA/DHA-type *pfa* gene cluster from *Aetherobacter fasciculatus* (SBSr002) (grey) for the corresponding domains of the AA/EPA-type *pfa* gene cluster from *Sorangium cellulosum* SBSo026 (black).
ER = Enoyl reductase, KS = Ketosynthase, AT = Acyltransferase, ACP = Acyl carrier protein, KR = Ketoreductase, DH = Dehydratase, CLF = Chain length factor, AGPAT = 1-Acylglycerol-3-phosphate *O*-acyltransferase.

The cloning procedure for the heterologous expression of hybrid *pfa* gene clusters in *P. putida* and *E. coli* started from plasmid pPm**SynPfaAf2, constructed previously (Chapter 4), which contains an artificial DPA/DHA-type *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) with adapted DNA sequences for *P. putida*. The vector backbone is derived from the broad-host-range expression vector pJB861 with a copy-up point mutation [35]. Heterologous expression of the *pfa* genes is driven by the *xylS-Pm* promoter system with mutations in the 5'-untranslated region [36] plus the core region [37] of *Pm* promoter. During the cloning procedure, it was taken advantage of unique restriction sites introduced for exchangeability of genes/domains (Chapter 4). The KS2/CLF domains were exchanged for the corresponding region of the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026, which was adapted for *P. putida* and equipped with a 40 bp intergenic region between *pfa2* and *pfa3* comprising a Shine-Dalgarno sequence, yielding plasmid pSynHybPfa1. Plasmid pSynHybPfa2 was constructed by replacement of the AT2 domain by the native AT2 domain from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026. Exchange of the KS2/CLF domains plus the AT2 domain resulted in plasmid pSynHybPfa3. The plasmids were transferred into *P. putida* KT2440::pPptAf1 by triparental conjugation (details not shown). Heterologous expression of the Pfa proteins plus AfPpt and extraction as well as GC-MS analysis of the fatty acids were performed as described previously. Unfortunately, this procedure did not lead to any detectable LC-PUFA production in *P. putida*. Besides *P. putida*,

E. coli was tested as host for heterologous expression of the chimeric Pfa proteins in presence of the appropriate PPTase AfPpt from *A. fasciculatus* (SBSr002). Thereby, the expression constructs pSynHybPfa1, pSynHybPfa2, or pSynHynPfa3 were cotransformed with the compatible plasmid pPptAfA, which contains the gene encoding the PPTase AfPpt from *A. fasciculatus* (SBSr002) (Chapter 3), into *E. coli* BL21(DE3) by electroporation. The transgenic clones were cultivated in 50 ml LB-medium at 37 °C. Expression of the Pfa proteins and AfPpt was induced at OD_{600} of 0.4-0.8 by addition of *m*-toluic acid to a final concentration of 2 mM. After induction, the cells were cultivated at 16 °C for 24 h. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [28] and analysed by GC-MS. In contrast to *P. putida*, the recombinant *E. coli* strains produce various LC-PUFAs: AA, EPA, dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6), docosatetraenoic acid (DTA, 22:4, *n*-6), and *n*-3 DPA are produced by all the three strains *E. coli* BL21(DE3)/pSynHybPfa1/pPptAfA, *E. coli* BL21(DE3)/pSynHybPfa2/pPptAfA, and *E. coli* BL21(DE3)/pSynHybPfa3/pPptAfA. DHA production can only be observed for strains *E. coli* BL21(DE3)/pSynHybPfa1/pPptAfA and *E. coli* BL21(DE3)/pSynHybPfa2/pPptAfA. Hence, compared to *P. putida*, *E. coli* seems to be less fastidious regarding the non-optimal arrangement of the protein domains and their poor ability to cooperate. However, only trace amounts of these LC-PUFAs were produced in the three recombinant *E. coli* strains, which cannot be exactly and reliably quantified. Merely the approximate ratios of the LC-PUFA species produced by two or three biological samples of each strain could be calculated. Thereby, it turned out that the exchange of both the KS2/CLF domains and the AT2 domain in the artificial DPA/DHA-type *pfa* gene cluster for the corresponding domains of the AA/EPA-type *pfa* gene cluster on plasmid pSynHybPfa3 led to the highest LC-PUFA production and to the lowest ratio of C22 PUFAs/C20 PUFAs.

In order to obtain more informative results, a more productive system for the heterologous expression of the DPA/DHA-type *pfa* gene cluster with exchange of both the KS2/CLF domains and the AT2 domain for the corresponding domains of the AA/EPA-type *pfa* gene cluster was required. Consequently, a synthetic chimeric *pfa* gene cluster was generated for *Y. lipolytica*, which has proven to be the best heterologous LC-PUFA producer employing myxobacterial PUFA synthases (Chapter 5). The cloning procedure started from the monocopy auto-cloning vector pSynPfaPptAf2, constructed previously (Chapter 5), which is able to promote random integration into the genome of Po1h strain. It contains four single transcription units with the coding sequences of the DPA/DHA-type *pfa* gene cluster plus the PPTase AfPpt originating from *A. fasciculatus* (SBSr002) and adapted for *Y. lipolytica*,

flanked by the strong hybrid hp4d promoter [26] plus *LIP2* terminator. During the cloning procedure, it was taken advantage of unique restriction sites introduced for exchangeability of genes/domains (Chapter 5). The KS2/CLF domains plus the AT2 domain of gene *pfa3* were exchanged for the corresponding region of the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026, which was adapted for *Y. lipolytica* and equipped with a 134 bp non-coding intergenic linker plus the strong hybrid hp4d promoter [26] located at the 5' end, yielding plasmid pSynPfaPptAf2-hyb*pfa3*. A reporter gene expression cassette consisting of *mcherry* flanked by the strong hybrid hp4d promoter [26] plus *LIP2* terminator (Chapter 5) was inserted into the linker region between genes *pfa2* and *pfa3* of plasmid pSynPfaPptAf2-hyb*pfa3* by conventional cloning for facilitated and rapid identification of good recombinant LC-PUFA producers. The "yeast cassette" of the resulting plasmid pSynPfaPptAf2mc-hyb*pfa3* was generated via hydrolysis with *Not*I and transferred into *Y. lipolytica* Po1h [21] by lithium acetate-mediated heat shock transformation. Ten transformants of *Y. lipolytica* Po1h::pSynPfaPptAf2mc-hyb*pfa3*, which exhibited intensive mCherry fluorescence on agar plate, were cultivated in 1 ml YNBD-medium at 28 °C for 76 h. The cellular fatty acids were extracted via acidic methanolysis using the FAME method [28] plus analysed by GC-MS, and the best producing strain, *Y. lipolytica* Po1h::pSynPfaPptAf2mc-hyb*pfa3* clone E, was cultivated in 50 ml YNBD-medium at 28 °C for 168 h, analogous to the reference strain *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, which expresses the unmodified DPA/DHA-type PUFA synthase and the PPTase AfPpt from *A. fasciculatus* (SBSr002) (Chapter5). Interestingly, an evident deviation from the production profile of the native DPA/DHA-type PUFA synthase can be observed for the hybrid PUFA synthase (Table 4). Exchange of the KS2/CLF domains plus the AT2 domain of gene *pfa3* in the artificial DPA/DHA-type *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) for the corresponding domains derived from the AA/EPA-type *pfa* gene cluster from *S. cellulosum* SBSo026 leads to the production of EPA and *n*-3 DPA as major LC-PUFA species. In contrast, DHA is the main product of the DPA/DHA-type PUFA synthase. On top of that, the high ratio of C22 PUFAs/C20 PUFAs observed for the native PUFA synthase is significantly decreased in the case of the chimeric PUFA synthase. Indeed, this indicates a change in the product specificity of the PUFA synthases and the identification of the KS2/CLF pair plus domain AT2 as factors determining the chain length of the produced LC-PUFAs. The low yields achieved with chimeric PUFA synthases (approx. 50-fold reduction of heterologous LC-PUFA production obtained with *Y. lipolytica* Po1h::pSynPfaPptAf2mc-hyb*pfa3* compared to *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C) might result from perturbations in the overall structure of the

multienzyme complex that translate into reduced catalytic efficiency, as the heterologous domains of the hybrid PUFA synthases have not been evolutionary optimized for cooperative work with the native domains of the Pfa enzymes.

Table 4. Recombinant LC-PUFA production by *Yarrowia lipolytica* Po1h::pSynPfaPptAf2 clone C and *Y. lipolytica* Po1h::pSynPfaPptAf2mc-hyb*pfa3* clone E.
Cultivation was carried out in 50 ml YNBD-medium + 50 mM potassium phosphate buffer pH 6.8 at 28 °C and 200 rpm for 168 h. The indicated values are means of three or four biological samples.

PUFA	Yield [mg/l]	
	***Y. lipolytica* Po1h::pSynPfaPptAf2 clone C**	***Y. lipolytica* Po1h:: pSynPfaPptAf2mc-hyb*pfa3* clone E**
eicosatrienoic acid (ETrA, 20:3, *n*-6 or *n*-9)	–	< 0.10
arachidonic acid (AA, 20:4, *n*-6)	0.48	< 0.10
eicosatetraenoic acid (ETA, 20:4, *n*-3)	< 0.10	< 0.10
eicosapentaenoic acid (EPA, 20:5, *n*-3)	1.03	0.70
docosapentaenoic acid (DPA, 22:5, *n*-6)	1.56	–
docosapentaenoic acid (DPA, 22:5, *n*-3)	0.70	0.61
docosahexaenoic acid (DHA, 22:6, *n*-3)	70.41	0.35
Ratio C22/C20	**45.72**	**1.19**

The probably most significant deviation of the PUFA synthases of terrestrial myxobacteria from the marine PUFA synthases described previously can be observed at the carboxyl-terminus of Pfa3. *In silico* analyses revealed that it most likely contains a 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) (EC 2.3.1.51) domain, which represents a unique and common characteristic of terrestrial myxobacterial PUFA synthases. The known function of such enzyme is the acylation of the 2-position of 1-acyl-*sn*-glycerol-3-phosphate during glycerophospholipid biosynthesis, utilising either acyl-ACP or acyl-CoA thioesters as substrate [46] (Figure 4A), which was to be proven via an *in vitro* enzyme assay. The AGPAT domain of gene *pfa3* from the *pfa* gene cluster of *A. fasciculatus* (SBSr002), flanked by 129 bp and 135 bp linker sequences to increase the protein solubility, was cloned into expression plasmid pGEX-6P-1. The resulting plasmid pGEX-6P-1-*AfAGPAT+linker* was transformed into *E. coli* Rosetta 2(DE3)pLysS by electroporation, and the strain was

cultivated in 50 ml LB-medium at 37 °C. Expression of the GST-tagged protein domain AfAGPAT flanked by 43 aa and 45 aa linker sequences was induced at OD_{600} of 1.0 by addition of IPTG to a final concentration of 1 mM. After induction, the cells were cultivated at 16 °C and 200 rpm for 19 h. Purification of GST-AfAGPAT+linker was carried out at 4 °C using a GST SpinTrap column. Removal of the GST-tag was performed by on-column enzymatic cleavage with a suitable protease, and the purified 39.1 kDa protein domain AfAGPAT+linker could be eluted from the column, as analysed by SDS-PAGE (Figure 4B). The protein identity was confirmed by ESI-qToF-MS analysis. In order to assay for transfer of DHA bound onto an ACP to the 2-position of 1-acyl-*sn*-glycerol-3-phosphate catalysed by AGPAT *in vitro*, a DHA *N*-acetylcysteamine (NAC) thioester was synthesized. NAC is identical to the end of the 4'-phosphopantetheine prosthetic group of carrier proteins. Therefore, DHA-NAC is supposed to mimic the ACP bound DHA. The 1-acyl-*sn*-glycerol-3-phosphate species 1-oleoyl-*sn*-glycerol-3-phosphate was used as second substrate. The assay conditions were varied in terms of the utilized buffer in the reaction, pH value, presence of $MgCl_2$, substrate concentration, as well as incubation time. Unfortunately, the expected product 1-oleoyl-2-docosahexaenoyl-*sn*-glycerol-3-phosphate could never be detected by high resolution nano-ESI-FT-Orbitrap-MS analysis. Possible explanations for the lack of the conversion of the substrates into the product are non-optimal assay conditions, lacking acceptance of the DHA-NAC thioester as substrate instead of ACP bound DHA, lacking acceptance of 1-oleoyl-*sn*-glycerol-3-phosphate as the second substrate, or expression and purification of AGPAT in an inactive form. Conceivably, enzyme activity of the isolated protein domain is not observed in case the active site is situated at an interface of Pfa domains or subunits, and electrostatic plus hydrophobic interactions between the subunits might be required for the functionality of AGPAT.

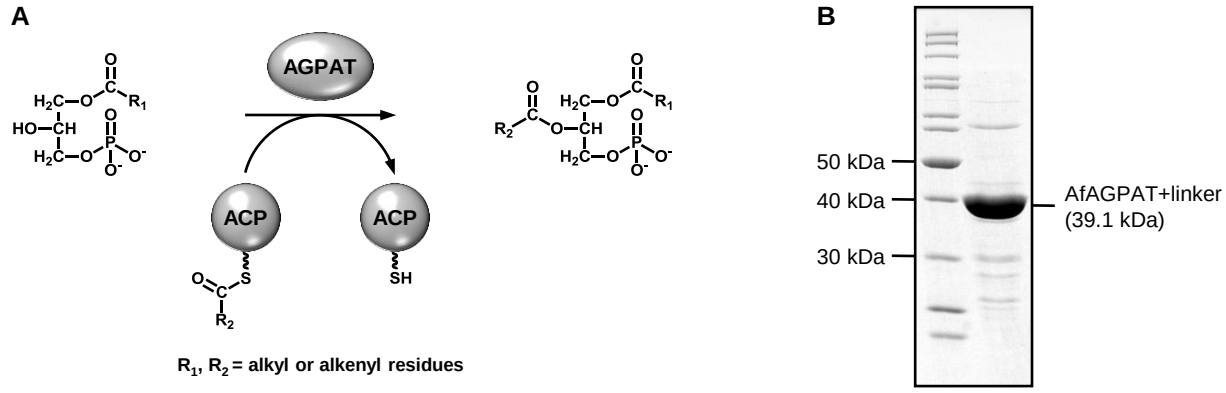


Figure 4. *In vitro* enzyme assay to prove the function of the 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) domain in LC-PUFA biosynthesis.
(A) Reaction scheme for the acylation of the 2-position of 1-acyl-*sn*-glycerol-3-phosphate catalysed by AGPAT using acyl-ACP as substrate during glycerophospholipid biosynthesis, (B) SDS-PAGE of heterologously produced and purified AGPAT domain flanked by 87 aa linker sequences from *Aetherobacter fasciculatus* (SBSr002).

REFERENCES

[1] Kaulmann U & Hertweck C (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew. Chem. Int. Ed.*, **41,** 1866-1869.

[2] Napier JA (2002) Plumbing the depths of PUFA biosynthesis: a novel polyketide synthase-like pathway from marine organisms. *Trends Plant Sci.*, **7,** 51-54.

[3] Wallis JG, Watts JL, & Browse J (2002) Polyunsaturated fatty acid synthesis: what will they think of next? *Trends Biochem. Sci.*, **27,** 467-473.

[4] Gemperlein K, Rachid S, Garcia R, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

[5] 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.

[6] 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.

[7] Morita N, Tanaka M, & Okuyama H (2000) Biosynthesis of fatty acids in the docosahexaenoic acid-producing bacterium *Moritella marina* strain MP-1. *Biochem. Soc. Trans.*, **28,** 943-945.

[8] Yazawa K (1996) Production of eicosapentaenoic acid from marine bacteria. *Lipids*, **31 Suppl,** S297-S300.

[9] 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 (Pt 8),** 2725-2731.

[10] Orikasa Y, Nishida T, Yamada A, Yu R, Watanabe K, Hase A, Morita N, & Okuyama H (2006) Recombinant production of docosahexaenoic acid in a polyketide biosynthesis mode in *Escherichia coli*. *Biotechnol. Lett.*, **28,** 1841-1847.

[11] Bumpus SB, Magarvey NA, Kelleher NL, Walsh CT, & Calderone CT (2008) Polyunsaturated fatty-acid-like trans-enoyl reductases utilized in polyketide biosynthesis. *J. Am. Chem. Soc.*, **130,** 11614-+.

[12] Jiang H, Zirkle R, Metz JG, Braun L, Richter L, Van Lanen SG, & Shen B (2008) The role of tandem acyl carrier protein domains in polyunsaturated fatty acid biosynthesis. *J. Am. Chem. Soc.*, **130,** 6336-6337.

[13] Okuyama H, Orikasa Y, Nishida T, Watanabe K, & Morita N (2007) Bacterial genes responsible for the biosynthesis of eicosapentaenoic and docosahexaenoic acids and their heterologous expression. *Appl. Environ. Microbiol.*, **73,** 665-670.

[14] Orikasa Y, Tanaka M, Sugihara S, Hori R, Nishida T, Ueno A, Morita N, Yano Y, Yamamoto K, Shibahara A, Hayashi H, Yamada Y, Yamada A, Yu RK, Watanabe K, & Okuyama H (2009) *pfaB* products determine the molecular species produced in bacterial polyunsaturated fatty acid biosynthesis. *FEMS Microbiol. Lett.*, **295,** 170-176.

[15] Shulse CN & Allen EE (2011) Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. *PLOS One.*, **6,** e20146.

[16] Grant SG, Jessee J, Bloom FR, & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U S A*, **87,** 4645-4649.

[17] Studier FW & Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **189,** 113-130.

[18] Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF, & Zhang Y (2006) An improved recombineering approach by adding RecA to λ Red recombination. *Mol. Biotechnol.*, **32,** 43-53.

[19] Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc. Natl. Acad. Sci. U S A*, **76,** 1648-1652.

[20] Bagdasarian M, Lurz R, Ruckert B, Franklin FC, Bagdasarian MM, Frey J, & Timmis KN (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene*, **16,** 237-247.

[21] Madzak C (2003) New tools for heterologous protein production in the yeast *Yarrowia lipolytica* . *Rec. Res. Dev. Microbiol.*, **7,** 453-479.

[22] Sambrook J & Russell DW (2001) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[23] Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, Martins Dos Santos VA, Fouts DE, Gill SR, Pop M, Holmes M, Brinkac L, Beanan M, DeBoy RT, Daugherty S, Kolonay J, Madupu R, Nelson W, White O, Peterson J, Khouri H, Hance I, Chris Lee P, Holtzapple E, Scanlan D, Tran K, Moazzez A, Utterback T, Rizzo M, Lee K, Kosack D, Moestl D, Wedler H, Lauber J, Stjepandic D, Hoheisel J, Straetz M, Heim S, Kiewitz C, Eisen J, Timmis KN, Dusterhoft A, Tummler B, & Fraser CM (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.*, **4,** 799-808.

[24] Oßwald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, & Wenzel SC (2012) Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.* Epub ahead of print.

[25] Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuveglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisrame A, Boyer J, Cattolico L, Confanioleri F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud JM, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard GF, Straub ML, Suleau

A, Swennen D, Tekaia F, Wesolowski-Louvel M, Westhof E, Wirth B, Zeniou-Meyer M, Zivanovic I, Bolotin-Fukuhara M, Thierry A, Bouchier C, Caudron B, Scarpelli C, Gaillardin C, Weissenbach J, Wincker P, & Souciet JL (2004) Genome evolution in yeasts. *Nature*, **430,** 35-44.

[26] Madzak C, Treton B, & Blanchin-Roland S (2000) Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J. Mol. Microbiol. Biotechnol.*, **2,** 207-216.

[27] Hill DS, Stein JI, Torkewitz NR, Morse AM, Howell CR, Pachlatko JP, Becker JO, & Ligon JM (1994) Cloning of genes involved in the synthesis of pyrrolnitrin from *Pseudomonas fluorescens* and role of pyrrolnitrin synthesis in biological control of plant disease. *Appl. Environ. Microbiol.*, **60,** 78-85.

[28] Bode HB, Ring MW, Schwär G, Kroppenstedt RM, Kaiser D, & Müller R (2006) 3-Hydroxy-3-methylglutaryl-coenzyme A (CoA) synthase is involved in the biosynthesis of isovaleryl-CoA in the myxobacterium *Myxococcus xanthus* during fruiting body formation. *J. Bacteriol.*, **188,** 6524-6528.

[29] Kao CM, Pieper R, Cane DE, & Khosla C (1996) Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. *Biochemistry*, **35,** 12363-12368.

[30] Dreier J & Khosla C (2000) Mechanistic analysis of a type II polyketide synthase. Role of conserved residues in the beta-ketoacyl synthase-chain length factor heterodimer. *Biochemistry*, **39,** 2088-2095.

[31] Dreier J, Shah AN, & Khosla C (1999) Kinetic analysis of the actinorhodin aromatic polyketide synthase. *J. Biol. Chem.*, **274,** 25108-25112.

[32] Leesong M, Henderson BS, Gillig JR, Schwab JM, & Smith JL (1996) Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: Two catalytic activities in one active site. *Structure*, **4,** 253-264.

[33] Lewin TM, Wang P, & Coleman RA (1999) Analysis of amino acid motifs diagnostic for the sn-glycerol-3-phosphate acyltransferase reaction. *Biochemistry*, **38,** 5764-5771.

[34] Zhang Y, Muyrers JPP, Testa G, & Stewart AF (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18,** 1314-1317.

[35] Blatny JM, Brautaset T, Winther-Larsen HC, Karunakaran P, & Valla S (1997) Improved broad-host-range RK2 vectors useful for high and low regulated gene expression levels in Gram-negative bacteria. *Plasmid*, **38,** 35-51.

[36] Berg L, Lale R, Bakke I, Burroughs N, & Valla S (2009) The expression of recombinant genes in *Escherichia coli* can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5'-untranslated part of mRNA. *Microb. Biotechnol.*, **2,** 379-389.

[37] Bakke I, Berg L, Aune TE, Brautaset T, Sletta H, Tondervik A, & Valla S (2009) Random mutagenesis of the Pm promoter as a powerful strategy for improvement of recombinant-gene expression. *Appl. Environ. Microbiol.*, **75,** 2002-2011.

[38] Oliynyk M, Brown MJ, Cortes J, Staunton J, & Leadlay PF (1996) A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.*, **3,** 833-839.

[39] Kuhstoss S, Huber M, Turner JR, Paschal JW, & Rao RN (1996) Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene*, **183,** 231-236.

[40] Bedford D, Jacobsen JR, Luo G, Cane DE, & Khosla C (1996) A functional chimeric modular polyketide synthase generated via domain replacement. *Chem. Biol.*, **3,** 827-831.

[41] Khosla C, McDaniel R, Ebert-Khosla S, Torres R, Sherman DH, Bibb MJ, & Hopwood DA (1993) Genetic construction and functional analysis of hybrid polyketide synthases containing heterologous acyl carrier proteins. *J. Bacteriol.*, **175,** 2197-2204.

[42] Zhu X, Yu F, Li XC, & Du L (2007) Production of dihydroisocoumarins in *Fusarium verticillioides* by swapping ketosynthase domain of the fungal iterative polyketide synthase Fum1p with that of lovastatin diketide synthase. *J. Am. Chem. Soc.*, **129,** 36-37.

[43] McDaniel R, Ebert-Khosla S, Hopwood DA, & Khosla C (1993) Engineered biosynthesis of novel polyketides. *Science*, **262,** 1546-1550.

[44] Tang Y, Tsai SC, & Khosla C (2003) Polyketide chain length control by chain length factor. *J. Am. Chem. Soc.*, **125,** 12708-12709.

[45] Keatinge-Clay AT, Maltby DA, Medzihradszky KF, Khosla C, & Stroud RM (2004) An antibiotic factory caught in action. *Nat. Struct. Mol. Biol.*, **11,** 888-893.

[46] Zhang YM & Rock CO (2008) Thematic review series: Glycerolipids - Acyltransferases in bacterial glycerophospholipid synthesis. *J. Lipid Res.*, **49,** 1867-1874.

APPENDIX

DNA sequence of the synthetic building block

>KS/CLF_So026_2

GGCCGGCCGAAGTCCGGGCTCCGCATCTGGTTGAGACCGGGCCCGCGAGCTGAAACTAAATACAACAACACAACA
AAATAAGGAGGATACATCATGACGTTCGAACCCATTGCAATCGTAGGCCAAGGCTGCGTGCTGCCAGGTGCATTG
AGCCCGCGCGCCCTGAACGAAGCAGTGCGCGCACGTCGCTGCGCAGTAACCGGCGCACCCGAAGGTCGTCTGCGC
TTGAGCGCCGCCCATGCCATGGGTCCAGCTGACCAGGCTGGCGACCGGATGTGGTCCGATGCGGGTGGCTACGTC
GAAGGGTTCGAGCCCGCGTTCGATGCGGCCGGGTTCCAGCTCGATGAAGGCGTCGTCCGCGCACTCGATCCGTCC
CTGAAATGGGTGCTGCATGCCGGTCGCGAAGCGTTGCGCCCGTTGGGCCATGACCGTGCAAGCGCACGCGCTGGC
TTGGCCTTGGGCAACCTGAGCTTTCCCACCCCCGCCATGGCTCGCTACGCTGAAGGCGTCTGGCTCGACGCCCAA
GGTGCCGCCTTCCTCGATGGTGCAGCCCGCGGTTTCGCAGCCTCGGAACGCCCTGATCCGCGCTCGCGGTTCATG
AGCGGTCTGCCTGCCTCGCTGGCTGCAGAAGCCCTGGGTTTGGGTGGCGGCGGCTTTGCCTTGGACGCAGCGTGC
GCGAGCTCGCTGTACGCGATTAAGCTGGCGTGCGACCGGCTCCACGATCGGACCGCTGACCTGATGCTCGCCGGT
GCGGTGAATGCCGCGGATCCGCTGTTCCTCCACATGGGCTTCTGTGCGTTGGCGGCGATGAGCCGCAGCGGTGCA
TCGCGGCCTTTCCATCGCGACGCTGACGGCCTGGTGCCAGCTGAAGGTGCCGCCCTGGTCGCCCTGAAACGGTTG
GCAGATGCCCGCGCAGCTGGCGATCGTGTGCTGGGCGTGATCCGCGGGATTGGCCTGAGCAATGACGGTCGTGGT
CGTGGGCTGCTGTCGCCGTCCGAAGAAGGGCAGATCCGGGCCATGCGCCTGGCATACGCTGCGGCTGGGCTGACC
CCAGCTGACGTGTCCATGATCGAGTGCCACGCGACCGGGACCTTGGTCGGTGACGCGACCGAAGTCCGGTCCACC
GCTGCGGTCTTCGAAGGGCAGCGTGACGTGCCGATCGGGTCCCTGAAGTCGAACCTCGGGCACCTGATCACCGCG
GCTGGTGCCGCAGGGCTGATGAAGGTCCTGGGCGCTCTGGAAGCGGGTGTACGCCCGGCTACCCTGCACGCTGAA
GCGCCCATCGAAGCGCTGCGTGGGTCCCCGTTTCGGCTGCTGGCCGAAGAAGAGCCCTGGCCTAGCGATCGTCCG
CGCGTAGCTGCGGTGAGCGCTTTCGGGTTTGGCGGCAACAACGCGCACCTCCTGGTCTCGCAAGACGACGCGCAC
GATGACCTGGGCACGGCTCCGGCTAGCTTCGTGCCTGCACGCCCACGTCCAGCCGTAGCCATCATCGGGATCGGC
GCGATGCTGGGCGATACCACCGGTGCCCCTGAGGCAGCTCGTGCAGTGCTGGGCGGTGAACCATGGACCCCTCGC
CGTGCAGAAGTGGCCGTAGCCCAAGAAGGCTTGCGCTTCCCGCCCCGGGATCTCGAGCAGACGCTGCCCCAGCAG
CTGCTCGTCTTCGAAGCGGCGCGTGAAGCGATCGCGGGTATCTCCGTCCCCCGTGAGCGCACGGGTGTGTTGGTC
GGCATGGGCTGCGACCCGGAAGTAGCGCGGTACGGGCTGCGTTGGCGTCTCGCGGACCTGGCCGACGCTTGGTCG
TCGTCGTCGCCTGCCCGTGGTGACCACAAGGCTCCGCCCGATTGGCTGGGTCAGGCTCGCGACGCGGTCTTGCCC
AAGCTGACCGCTGTCGGCGTGCTCGGCACGATGCCCAACATCCCCGCGAACCGGATTAGCTCGCAGTTGGACCTG
GGCGGTCCCGGGTATACGGTGAGCGCTGAACAGGCGAGCGGCGTCGTGGCGCTGGAAATTGCAGCGCGGGCCTTG
CGCGAAGGCGAACTGGATGCGGCTGTGGTCGCGGCTGTCGACCTGAGCGATGAGCCAGTGCATCGCGCGGCGTTG
TCGGCGTTGGGCATCGAGACCCCGACCGGTGACGCAGCCGTAGCTCTGGTGCTCAAGCGCCTCGATGACGCGCGT
CGTGACGGCGACCACGTCTTGGCCGAGCTCGACGAAGAAGGCGCCCCAGCACTCCGTGTGGGTGACGGCGACGGT
GCAGTGGACCCGTTGGCCCGTGGTGCTGCACATGCAGCCGCAGGCCTGCTGCACGTAGCCGCAGCCGCATGGTCG
GTCCACCATGGTGCCCGTCCGGCACGTAGCGCCGTAGGTAGCGGCGCACGTGGTGCAGCCCCATGGTTCGGTGCT
CGTACCGCCGAGACCCGCACCCGTGTGTTGGGCGGTGCAGACGCAAGCGTGCGTCTGAGCGCCCGTGGTCCAGCA
GCCCCATTGGGCTTTGAGCCGCCGCCGCGGTTGTTCGTGTACGGCGGCGCAGATCGCGCGAGCGTGTTGCGCGCG
TTGGCAGAAGGTCGCCCTAGCGATGCCGGCCCAGCACGTCTGGTGCTGGTAGCCGCCAGCGAAGAGCAACGTGCT
GCCCGTGCCGCACAGGCACGTCTGCACCTGGAAAAAGGCGGCCCGGCACCTGAAGGTGTGGCATATCGCGATGCC
CCAGCCGGCGGCCAGCTGGCCTTTGTCTTCGCGGGCGCAGCCGCCGCATACCCGAGCATGGGCCGTGCAC

>building block 5

AACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGC
CGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCCATGCTGAGGTGTCTC
ACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTC
CCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGT
GTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCTGAGGTGTCTCACAAGTG
CCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGTGTACGTACATTATCGAGACCGTTGTTCCCGCCCA
CCTCGATCCGGCATGCTGAGGTGTCTCACAAGTGCCGTGCAGTCCCGCCCCCACTTGCTTCTCTTTGTGTGTAGT
GTACGTACATTATCGAGACCGTTGTTCCCGCCCACCTCGATCCGGCATGCACTGATCACGGGCAAAAGTGCGTAT
ATATACAAGAGCGTTTGCCAGCCACAGATTTTCACTCCACACACCACATCACACATACAACCACACACATCCACA
ATGCGATTCGAACCCATTGCTATCGTTGGTCAAGGCTGCGTGCTGCCTGGAGCCCTGTCCCCTCGAGCCCTGAAC
GAGGCTGTTCGAGCTCGACGATGTGCTGTGACTGGCGCCCCTGAGGGTCGACTCCGACTGTCTGCTGCCCATGCT

ATGGGTCCTGCCGACCAGGCTGGAGACCGAATGTGGTCTGACGCCGGCGGCTATGTGGAGGGATTTGAGCCTGCT
TTTGACGCCGCTGGCTTCCAGCTGGACGAAGGCGTTGTCCGAGCCCTGGACCCCTCTCTGAAATGGGTGCTGCAC
GCCGGACGAGAGGCTCTCCGACCCCTGGGTCACGATCGAGCTTCCGCTCGAGCCGGACTCGCCCTGGGAAACCTG
TCCTTCCCTACCCCCGCCATGGCCCGATATGCTGAGGGCGTCTGGCTCGATGCCCAGGGTGCCGCCTTTCTCGAC
GGCGCTGCTCGAGGATTTGCTGCTTCTGAGCGACCCGACCCCCGATCCCGATTTATGTCTGGACTCCCCGCCTCT
CTCGCTGCCGAAGCCCTGGGACTGGGAGGAGGCGGCTTCGCTCTGGACGCCGCTTGCGCCTCTTCTCTCTATGCC
ATTAAACTGGCCTGCGACCGACTGCATGACCGAACTGCTGACCTGATGCTCGCCGGCGCCGTTAACGCTGCTGAC
CCTCTGTTCCTGCACATGGGCTTCTGTGCTCTCGCTGCTATGTCCCGATCTGGCGCTTCCCGACCCTTTCATCGA
GACGCCGACGGTCTCGTCCCTGCTGAGGGAGCCGCCCTGGTTGCCCTGAAACGACTGGCTGACGCCCGAGCTGCT
GGCGATCGAGTCCTGGGCGTGATTCGAGGCATCGGTCTCTCTAACGATGGCCGAGGTCGAGGACTGCTGTCCCCT
TCCGAAGAGGGACAGATCCGAGCCATGCGACTCGCCTACGCTGCTGCTGGCCTGACCCCTGCCGATGTCTCTATG
ATTGAATGTCATGCTACCGGAACCCTCGTTGGCGACGCCACCGAGGTCCGATCCACCGCTGCTGTCTTCGAGGGA
CAGCGAGATGTCCCCATCGGATCTCTGAAATCTAACCTCGGACACCTGATTACCGCTGCTGGAGCTGCCGGACTG
ATGAAGGTCCTGGGTGCTCTCGAGGCTGGCGTTCGACCCGCTACCCTGCACGCTGAAGCCCCCATTGAAGCTCTG
CGAGGCTCCCCCTTCCGACTGCTCGCTGAGGAGGAGCCCTGGCCTTCTGATCGACCTCGAGTCGCCGCTGTGTCT
GCCTTCGGTTTCGGTGGCAACAATGCCCACCTGCTCGTGTCTCAGGATGACGCCCACGATGACCTCGGAACCGCC
CCTGCTTCTTTTGTGCCCGCTCGACCTCGACCTGCTGTGGCTATCATTGGCATTGGTGCTATGCTGGGAGACACT
ACTGGCGCTCCTGAGGCCGCTCGAGCTGTGCTGGGTGGAGAGCCTTGGACTCCCCGACGAGCTGAGGTTGCCGTC
GCTCAGGAGGGCCTGCGATTTCCTCCCCGAGATCTGGAACAGACCCTGCCCCAGCAACTGCTGGTCTTCGAGGCC
GCTCGAGAGGCTATCGCTGGCATCTCTGTCCCTCGAGAACGAACCGGCGTGCTCGTGGGAATGGGCTGTGATCCT
GAGGTTGCCCGATACGGTCTCCGATGGCGACTGGCTGATCTCGCTGACGCCTGGTCCTCCTCCTCCCCTGCTCGA
GGAGACCATAAGGCTCCTCCTGACTGGCTCGGCCAGGCTCGAGACGCTGTTCTGCCCAAGCTGACTGCCGTTGGC
GTTCTCGGCACCATGCCCAACATTCCCGCTAACCGAATTTCCTCCCAACTGGACCTGGGTGGTCCCGGCTACACC
GTTTCCGCTGAACAGGCTTCCGGCGTGGTGGCCCTCGAGATTGCTGCTCGAGCTCTGCGAGAGGGAGAGCTGGAT
GCCGCCGTCGTTGCCGCTGTCGACCTCTCTGATGAGCCCGTCCATCGAGCCGCCCTCTCTGCCCTGGGAATTGAG
ACTCCCACTGGCGATGCCGCCGTTGCCCTCGTGCTGAAGCGACTCGACGATGCTCGACGAGATGGAGACCACGTG
CTCGCTGAGCTGGACGAGGAGGGTGCTCCCGCTCTGCGAGTCGGAGACGGCGATGGTGCCGTTGATCCCCTGGCT
CGAGGAGCTGCCCATGCTGCTGCCGGACTCCTCCACGTGGCCGCTGCTGCTTGGTCTGTCCACCATGGAGCTCGA
CCCGCTCGATCCGCTGTTGGTTCCGGAGCCCGAGGCGCTGCTCCTTGGTTTGGCGCCCGAACCGCCGAAACCCGA
ACTCGAGTCCTGGGCGGAGCTGACGCTTCCGTTCGACTCTCCGCCCGAGGACCTGCTGCTCCTCTGGGTTTTGAG
CCCCCTCCCCGACTCTTTGTTTACGGTGGTGCCGACCGAGCTTCTGTGCTGCGAGCCCTCGCCGAAGGACGACCC
TCTGACGCTGGACCCGCTCGACTCGTTCTGGTGGCCGCCTCTGAAGAGCAGCGAGCTGCCCGAGCTGCCCAAGCT
CGACTGCACCTGGAGAAGGGAGGTCCCGCTCCCGAAGGCGTTGCCTACCGAGACGCCCCTGCTGGTGGCCAACTG
GCCTTCGTGTTCACTGGAGCCGCCGCCTCCTACGCTGGCATGGGTCGAGAACTGGCCCTGGCCCTCCCTGATGTG
ACCGCTCGACTCGGAGAGCGATTCGCCTCCATGGACCAAGCCGCCCGATGGATCTTTGATCCTCCTGAGGCTCCC
TCCCACCCTCTCGACCAGCTCTGGGCCTCCGCTTTTCTCTGTCAGCTGCATGCTGAGGTCTCCCGACGAGTCCTG
GGACTGACTCCTGACGCTACTGTCGGATACTCCTCCGGAGAATCCAACGCCCTCTTTGCCACCGGCGCTTGGCGA
GACCTGGATGCCATGATCCGAGATTGCTCTGAATCCCCCGTGTTCACCACTGAACTCGTGGGCGAATTCGCCGCC
GCTCGACGAGCCTGGCGAAAGCTCGGTGGTGGAGCCCTCGAGGCTTGGAAGGCTTGGTCTGTCGCTGCCCCCGTG
GAATCTGTGCGAGAAGCCCTGCGAGGCGAGCCCCTCGCTCATCTCACCATCGTGAACACCCCTGAGGATTGCGTG
CTGGGTGGCGAAGCCTCCGCTTGCGAACGAGTGATCGAGCGACTCGGTCGAGGTCGAGCTATGCCTCTCGGTTAC
GAGATGGCCGCCCACTGTCCTGAAATCGAAGAGATTCGAGGAGCTTGGTACGACCTCCACCACCGAGCTACTTGG
GAGGTCCCCGGCGTTCGACATTACTCTGCTGGTCGAGCCACCGCCTTCCAAGCCACTTCTGAGGCCGCCGCTGAG
GCTATCACTGCCCAAGCCCTGGGCACTCTGGACTTCCCTCGAATGATCGAACGAGCCTGGGCTGATGGCGTCCGA
GTCTTCCTCGAACACGGACCCCGAGGCCTCTGCTCCGGATGGATTCGACGAATTCTCGGAGACCGAGAGCACGTT
GCCGTTCCTCTCGACGTGGCTGGTCGATCTGGCGTGCAACAGCTGGCCAATGCCGCTGCTTGGCTGATTGCCGCT
GGCGTCCCCGTTGACCGAGGAGCCCTCGAACGATCCCTGGGTGCTATTGGCGTGCCCGCTGCTCCTCCTCGAGCT
ACCCTGAAGATCCCTGCCCATCCTCCCGCCATCCGACTGCCCCCTCTCGAGCAATGGATGGCCCCCGCTCCTCGA
CTGGTTCCCGTTCTGGACGATGCTACCGCTGCTCCTGTGAAGGCTCCCGCCCCTTCTCCCGTCCATTTCGCCGCT
CCTGCTCAAGCTCGAGCCGCCGTGCCCTCTCTGGTCCACGCTGCTCCCCCTTCCCTCGTGCATGCTGCTCCCCCC
TCCCTCGCTCACGCCGCTCCTCCCTCTCTGGTGCATGCCGCCCCTCCTTCTCTGGCCCAAGCTGCTCCTCCTGCT
CCCGCCCGACCTTCCACTCGATCTCCTGCCATCTCTATCGGTTCTCCCGCTGGTGAAATCGTGGCCCGAGCCACC
GAACAACAGGCCCGACTGGGCGCCCTGCATCGAGAATTCCTGGCTTCTCAGGCCACCGCCCACGCCCAATTCCTG
GCCCTCCAGCAGGGACTGCTCGGTCGATTTCTCCGAGCCTACGAACACCAGGCTTCTCCCCCTCCTCCCCCTCCT
CTGCCCTCTCTGCCCCCTGTTGCTCCTCCTCTCCCCTCTCTCCCCCGAGCTGGTGAAGACGCCGGACTGCCTGCT
CCTGCTCCCGTCTCCCCCGAGGCTTCTACCGACGCTTCCTCCGTCGATGCCCCCGTCGGAGGTCGACCCCTGCTC
CCCGGTCCTAAATTCTCCCGAGCCGATCTGGAGGTTCTGGCCTCTGGTTCCATCTCCTCCATCTTCGGCCCCCAA
TTCGCCGGACAAGACGGCTACGCCAAGCAATGTCGAATGCCCGAGCCCCCTCTGCTCCTGGCTGACCGAGTGACT
GGAATCGACGCCGTCCCCGGTTCTATGGATACTGGAACCATCTGGACTGAGACTGACGTCAAGCATGACTCTTGG
TACC

DISCUSSION

Myxobacteria have been proven to be proficient producers of not only secondary metabolites with impressive structural diversity, bioactivity, and diversity of mode-of-action [109] but additionally of LC-PUFAs, such as AA, EPA, or DHA [127-129]. In the present work, for the first time, iteratively acting type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like enzymes, termed PUFA synthases, responsible for the production of LC-PUFAs have been identified and characterized in detail from both terrestrial origin and myxobacteria.

Discovery and detailed comparisons of characteristic features of myxobacterial *pfa* gene clusters at the molecular level

Extensive shotgun sequencing of myxobacterial genomes by next generation sequencing (NGS) technologies has tremendously simplified the detection of biosynthetic pathways. In the present studies, five types of PUFA biosynthetic pathways employing PUFA synthases could be identified in a variety of PUFA-producing myxobacteria, belonging to the suborders *Sorangiineae* and *Nannocystineae*, whose genomes have been sequenced (Table 1). Among all the catalytic domains of PUFA synthases, the tandem acyl carrier protein (ACP) domains are obviously most remarkable. Five to nine consecutive domains are present in these multienzyme complexes from the marine γ-Proteobacteria *Shewanella pneumatophori* SCRC-2738, *Photobacterium profundum* SS9, and *Moritella marina* MP-1, as well as from the marine microalgae *Schizochytrium* sp. ATCC 20888 [77-79]. The DNA sequence of the region encoding the tandem ACP domains of myxobacterial PUFA synthases was determined in case of the OA/LA-type *pfa* gene clusters from *S. cellulosum* So ce56, So ce377, and SBSo026, in case of the arachidonic acid (AA, 20:4, *n*-6)/EPA-type *pfa* gene clusters from *S. cellulosum* SBSo026 and So ce1128, and in case of the docosapentaenoic acid (DPA, 22:5)/DHA-type *pfa* gene clusters from the *Aetherobacter* species *Aetherobacter* sp. (SBSr001) nov. ined., *Aetherobacter fasciculatus* sp. nov. ined. (SBSr002), *Aetherobacter rufus* sp. nov. ined. (SBSr003), and *Aetherobacter* sp. nov. ined. (SBSr008). The striking degree of sequence similarity (> 90% pairwise identity at DNA level) between the ACP domains encoded by a particular *pfa* gene cluster considerably impeded proper assembly of the short sequencing reads from NGS to completely cover this highly repetitive region. Therefore, it is hardly surprising that for most of the myxobacterial *pfa* gene clusters mentioned previously, the exact DNA sequence of this section had to be ascertained by PCR

amplification or subcloning from a chromosomal gene library prior to Sanger sequencing (details not shown). The number of tandem ACP domains detected in the PUFA synthases from myxobacteria of the suborder *Sorangiineae* ranges from three to five (Figure 10). It had been demonstrated that all tandem ACPs encoded by *pfaA* from *Shewanella japonica* are functionally equivalent for LC-PUFA biosynthesis and that an increase in the number of ACPs increases the throughput of LC-PUFA biosynthesis [154]. In order to make a valid statement about the impact of the quantity of ACPs on the production rate of myxobacterial PUFA synthases, the amounts of octadecenoic acid, LA, and eicosadienoic acid (EDA, 20:2, *n*-6) synthesized by *S. cellulosum* So ce56, harbouring an OA/LA-type PUFA synthase with five tandem ACPs, and by *S. cellulosum* SBSo026, comprising an OA/LA-type PUFA synthase with only four tandem ACPs, were directly compared. Actually, no correlation can be deduced, as both strains produce comparable amounts of these fatty acids ([155]; Chapter 2). Thus, it remains unresolved whether additional copies of ACP domains within the (myxobacterial) PUFA synthases would lead to enhanced LC-PUFA production.

So far, an OA/LA-type *pfa* gene cluster could be detected in virtually all of the *S. cellulosum* strains whose genome sequence is available (Table 1). From this data, it can be concluded that this gene cluster is most likely almost ubiquitous in *S. cellulosum*. Intriguingly, among the *S. cellulosum* strains mentioned previously, the five strains SBSo026, So ce1128, So ce340, So ce427, and So ce690 additionally contain an AA/EPA-type *pfa* gene cluster embedded in 18 putative maracin/maracen biosynthetic genes, exclusively present in the genomes of these five So ce strains. The OA/LA-type PUFA synthase and the AA/EPA-type PUFA synthase from *S. cellulosum* SBSo026 only show a pairwise sequence identity of ≤ 64% at DNA level and of ≤ 53% at protein level. These observations indicate that the genomic island comprising the AA/EPA-type *pfa* gene cluster rather originates from horizontal gene transfer into a *S. cellulosum* ancestor already harbouring the OA/LA-type *pfa* gene cluster than from duplication of this gene cluster. Remarkably, the AA/EPA-type *pfa* gene clusters miss gene *pfa1* (Figure 10), but the point in time of its loss is uncertain. A common characteristic of both types of PUFA synthases from *Sorangium* represents the long linker sequence between the final FabA-like dehydratase (DH) domain and the 1-acylglycerol-3-phosphate *O*-acyltransferase (AGPAT) domain (Figure 10). It exhibits a length of more than 400 aa in case of the OA/LA-type PUFA synthase as well as in the AA/EPA-type PUFA synthase. The corresponding linker sequences in the DPA/DHA-type PUFA synthases from *Aetherobacter* spp. only have an average length of 40 aa. All these non-catalytic regions are rich in alanine, proline, and charged amino acids. They are

principally flexible, but the presence of alanine-proline peptide bonds introduces a degree of rigidity if they exist in the *trans* form [156]. These features ensure that tethered domains are kept apart while allowing them to make essential interactions with other catalytic entities. Thus, the different length of these linkers might affect the performance of adjacent domains to a varying extent. However, this is purely speculative and an experimental proof is pending.

Apart from the described *pfa* gene clusters from *Sorangium* and *Aetherobacter*, gene clusters encoding PUFA synthases have also been identified in other myxobacterial species in the course of the present studies (Table 1). *Minicystis rosea* (SBNa008) nov. ined. (R. O. Garcia et al., submitted), which belongs to the suborder *Sorangiineae*, produces AA as major PUFA (> 20% of TFAs) and minor amounts of EPA, γ-linolenic acid (GLA, 18:3, *n*-6), dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6), and docosatetraenoic acid (DTA, 22:4, *n*-6). Sequencing of the genome of this strain by NGS and *in silico* analysis of the genome data revealed the presence of an incompletely sequenced *pfa* gene cluster. It comprises the genes *pfa1*, *pfa2*, and *pfa3*. Missing sequences within the gene cluster were amplified by PCR prior to Sanger sequencing in order to complete the cluster sequence. Catalytic domain arrangements and sequences of the proteins Pfa2 and Pfa3 resemble the homologous proteins of the AA/EPA-type PUFA synthase from *S. cellulosum* the most (> 73% pairwise identity). Parts of a *pfa* gene cluster could also be identified in the unpublished genome sequence of the terrestrial AA-producing myxobacterium *Sandaracinus amylolyticus* NOSO-4, also belonging to the suborder *Sorangiineae* [128]. A fragment of gene *pfa2* encoding a ketoreductase (KR) domain and a PKS-like DH domain and fragments of gene *pfa3* encoding a ketosynthase (KS) domain, parts of a chain length factor (CLF) domain, two FabA-like DH domains separated by a DH' pseudo-domain, an additional PKS-like DH domain, and an AGPAT domain could be detected on two small scaffolds. Even the marine AA-producing myxobacterium *Plesiocystis pacifica* SIR-1, belonging to the suborder *Nannocystineae* [124], possess a *pfa* gene cluster (the genome was sequenced by the J. Craig Venter Institute in the context of the Gordon and Betty Moore Foundation Marine Microbial Genome Sequencing Project and deposited into the public domain). This finding obviously shows that the existence of PUFA synthases in myxobacteria is not only restricted to the suborder *Sorangiineae*. The PUFA synthase from *P. pacifica* SIR-1 contains FAS/PKS domains in the same number and order as the homologous enzymes from *Aetherobacter* spp. except that only two tandem ACP domains could be found. However, as there seems to be a misassembly of the sequencing reads in this genomic region, a larger number of ACPs might be available. Completely unexpected, an

AGPAT domain neither was encoded by the *pfa* gene cluster nor could be identified at any other locus within the genome.

Analyses of the PUFA synthases with focus on the AGPAT domains encoded by the OA/LA-type *pfa* gene cluster from *S. cellulosum* So ce56, the AA/EPA-type *pfa* gene cluster from *S. cellulosum* SBSo026, and the DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) using BlastP [157] and Pfam [158] searches revealed occurrence of PUFA synthase homologues including an AGPAT domain. They were found to be encoded by the genomes of *Azospirillum lipoferum* 4B, a bacterium associated with roots of terrestrial plants [159], of the soil bacterium *Paenibacillus curdlanolyticus* YK9 (the genome was sequenced by the DOE Joint Genome Institute and deposited into the public domain), and of the actinobacterium *Nakamurella multipartita* Y-104, isolated from activated sludge [160]. The biosynthetic products of all these PUFA synthase-like proteins, however, have not yet been determined. Nevertheless, this clearly indicates that the presence of an AGPAT domain, which is assumed to allow for direct transfer of the produced PUFAs into lipids, is a characteristic of PUFA synthases from non-marine origin. Interestingly, exceptions to the rule constitute the homologous, uncharacterized Pfa proteins from terrestrial streptomycetes, such as *Streptomyces coelicolor* A3(2) [161] and *Streptomyces lividans* 1326 [162], which are lacking an AGPAT domain, and the putative PUFA synthase form the marine γ-Proteobacterium *Saccharophagus degradans* 2-40 [163], which contains an AGPAT domain.

Myxobacterial *pfa* gene clusters have been shown to differ from the well-known PUFA biosynthetic pathways of marine γ-Proteobacteria and microalgae. They only consist of three genes (*pfa1*, *pfa2*, and *pfa3*) arranged in a different order and encoding multifunctional PUFA synthases that also differ in their catalytic domain composition [155]. Although the organization of the *pfa* gene clusters is very similar among different myxobacterial species, the profiles of the produced fatty acids vary significantly (Table 1 and Figure 10). The presence of different types of PUFA biosynthetic pathways with such product diversity within the same bacterial suborder (*Sorangiineae*) is an outstanding aspect in the field of research on PUFA biosynthesis and serves as unique selling point with the potential to arouse economic interests.

Table 1. Overview of the myxobacterial strains with identified and characterized *pfa* gene clusters and correlating MUFA/PUFA production profiles, as determined in the present work.
OA = oleic acid (18:1, *n*-9), LA = linoleic acid (18:2, *n*-6), GLA = γ-linolenic acid (18:3, *n*-6), SDA = stearidonic acid (18:4, *n*-3), EDA = eicosadienoic acid (20:2, *n*-6), DHGLA = dihomo-γ-linolenic acid (20:3, *n*-6), AA = arachidonic acid (20:4, *n*-6), ETA = eicosatetraenoic acid (20:4, *n*-3), EPA = eicosapentaenoic acid (20:5, *n*-3), DTA = docosatetraenoic acid (22:4, *n*-6), DPA = docosapentaenoic acid (22:5, *n*-6 or *n*-3), and DHA = docosahexaenoic acid (22:6, *n*-3).

Suborder	Genus	Species	Strain	Type of *pfa* gene cluster	MUFA/PUFA production profile
Sorangiineae	*Sorangium*	*cellulosum*	So ce56	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So0157-2	OA/LA-type *pfa* gene cluster	n.d.
Sorangiineae	*Sorangium*	*cellulosum*	So ce10	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce1525	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce307	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce377	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce38	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce487	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce836	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ceGT47	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	So ce969	OA/LA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA
Sorangiineae	*Sorangium*	*cellulosum*	SBSo026	OA/LA-type *pfa* gene cluster + AA/EPA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA, GLA, SDA, DHGLA, AA, ETA, EPA, *n*-3 DPA

Continuation of Table 1.

Suborder	Genus	Species	Strain	Type of *pfa* gene cluster	MUFA/PUFA production profile
Sorangiineae	*Sorangium*	*cellulosum*	So ce1128	OA/LA-type *pfa* gene cluster + AA/EPA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA, GLA, SDA, DHGLA, AA, ETA, EPA, *n*-3 DPA
Sorangiineae	*Sorangium*	*cellulosum*	So ce340	OA/LA-type *pfa* gene cluster + AA/EPA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA, GLA, SDA, DHGLA, AA, ETA, EPA, *n*-3 DPA
Sorangiineae	*Sorangium*	*cellulosum*	So ce427	OA/LA-type *pfa* gene cluster + AA/EPA-type *pfa* gene cluster	OA, isomers of octadecenoic acid, LA, EDA, GLA, SDA, DHGLA, AA, ETA, EPA, *n*-3 DPA
Sorangiineae	*Sorangium*	*cellulosum*	So ce690	OA/LA-type *pfa* gene cluster + AA/EPA-type *pfa* gene cluster	n.d.
Sorangiineae	*Sorangium*	*cellulosum*	So ce26	AA-type *pfa* gene cluster	AA, DHGLA, DTA, EPA
Sorangiineae	*Aetherobacter*	sp.	SBSr001	DPA/DHA-type *pfa* gene cluster	AA, EPA, *n*-6 DPA, *n*-3 DPA, and DHA
Sorangiineae	*Aetherobacter*	*fasciculatus*	SBSr002	DPA/DHA-type *pfa* gene cluster	AA, EPA, *n*-6 DPA, *n*-3 DPA, and DHA
Sorangiineae	*Aetherobacter*	*rufus*	SBSr003	DPA/DHA-type *pfa* gene cluster	AA, EPA, *n*-6 DPA, *n*-3 DPA, and DHA
Sorangiineae	*Aetherobacter*	sp.	SBSr008	DPA/DHA-type *pfa* gene cluster	AA, EPA, *n*-6 DPA, *n*-3 DPA, and DHA
Sorangiineae	*Minicystis*	*rosea*	SBNa008	EPA/AA-type *pfa* gene cluster	AA, EPA, GLA, DHGLA, DTA
Sorangiineae	*Sandaracinus*	*amylolyticus*	NOSO-4	AA-type *pfa* gene cluster	AA [128]
Nannocystineae	*Plesiocystis*	*pacifica*	SIR-1	AA-type *pfa* gene cluster	AA [124]

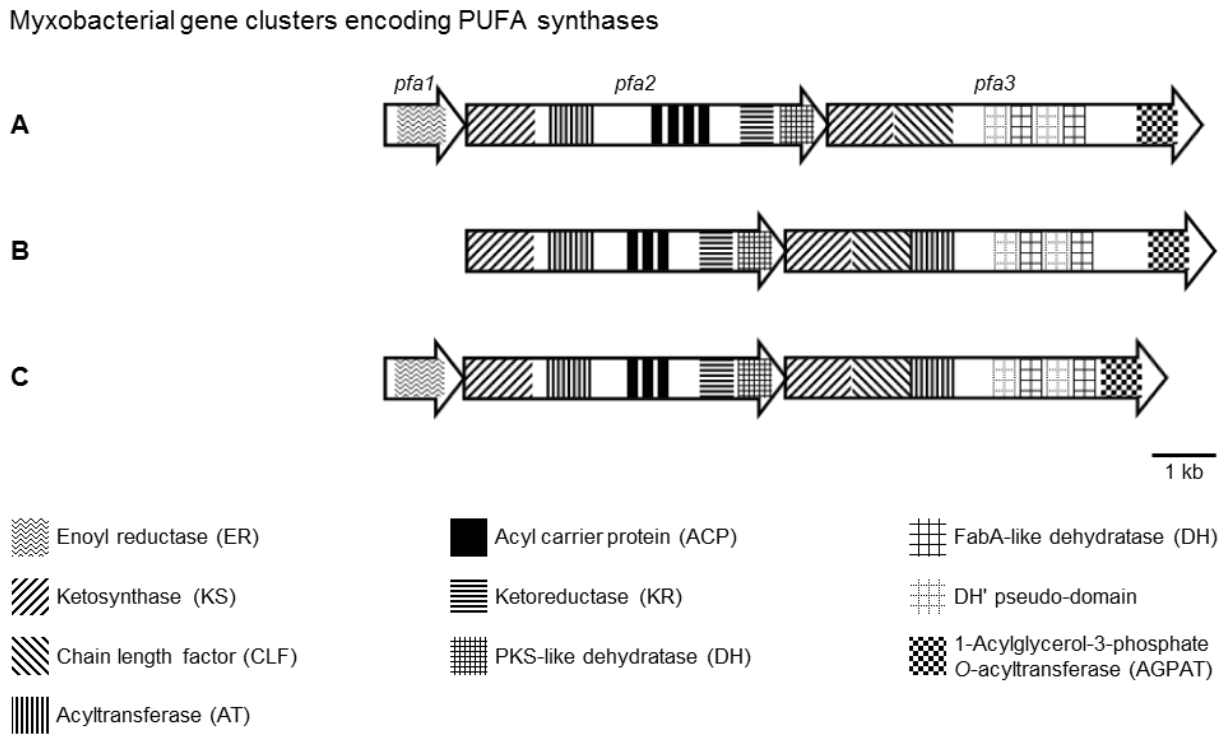
Physiological roles of PUFAs in microbes

The exact function of PUFA biosynthesis by microorganisms remains obscure. However, it is known that a dynamic structure of biological membranes, in which the individual components are free to spin, wobble, and diffuse laterally within the membrane leaflet, are essential requirements for the cells of all organisms. Hence, this lamellar, liquid crystalline phase of the membrane is essential for proper functioning and must be preserved in the face of altered environmental conditions [164]. Adjustments in various aspects of acyl chain composition are among the most thoroughly documented of membrane adaptations to temperature change [165]. A nearly ubiquitous response to low temperatures or high hydrostatic pressures is a reduction in the proportion of saturated fatty acids (SFAs) and a corresponding increase in the proportion of unsaturated fatty acids (UFAs). The introduction of a *cis* double bond generates a bend in the acyl chain and can therefore have a pronounced effect on both the molecular shape and physical properties of a phospholipid. Consequently, UFAs decrease the molecular order of the acyl chains, promote less compact packing of phospholipids, and possess lower melting points than their saturated homologues [165]. Numerous psychrotolerant piezophilic deep-sea bacterial isolates were found to contain substantial quantities of *n*-3 long-chain polyunsaturated fatty acids (LC-PUFAs), namely eicosapentaenoic acid (EPA, 20:5, *n*-3) and docosahexaenoic acid (DHA, 22:6, *n*-3), whereas mesophilic and pressure-sensitive bacteria only produce scant amounts thereof. This preponderance of PUFA producers in the deep-sea environment led to the speculation that such polyenoic fatty acids are specifically involved in the adaptation of bacteria to the high-pressure, low-temperature conditions [166-168]. Moreover, it could be demonstrated for *Shewanella oneidensis* MR-1 that *pfa* gene products serve as precursors for the production of olefinic hydrocarbons, which are proposed to aid in adapting cells to a rapid drop in temperature [169]. However, oleic acid (OA, 18:1, *n*-9) or linoleic acid (LA, 18:2, *n*-6) located at position sn-2 of phosphatidylcholine lower the gel-to-liquid-crystalline phase transition temperature considerably more effectively than DHA [170]. In line with this, it has been shown that OA rather than EPA is required for growth of the psychrotolerant piezophilic deep-sea bacterium *Photobacterium profundum* SS9 at high pressure and low temperature [171].

In general, PUFAs are among the molecules most susceptible to oxygen and reactive oxygen species (ROS), because they have many bis-allylic carbon atoms [172]. However, in the aqueous system, *n*-3 LC-PUFAs are rather stable against oxidative stresses caused by ROS [173]. An antioxidative function of EPA could be demonstrated *in vivo* using EPA-producing *Escherichia coli* strains that had been transformed with the *pfa* gene cluster from

the marine γ-Proteobacterium *Shewanella pneumatophori* SCRC-2738. In the recombinant *E. coli* strains grown in the presence of H_2O_2, cellular EPA seemed to be stable and protected cells against harmful effects of H_2O_2, namely, growth inhibition, carbonylation of cellular proteins, and breakage of the cell structure [174;175]. Analogous results have been obtained using the marine EPA-producer *Shewanella marinintestina* IK-1 and its EPA-deficient mutant IK-1Δ8 [176]. In bacteria, and probably in microalgae, membrane phospholipids with EPA and/or DHA are supposed to function as shield molecules against exogenous and endogenous ROS. The molecular mechanism of the membrane-shielding effects of LC-PUFAs has not yet been elucidated. Possibly, lipid membranes consisting of phospholipids with saturated fatty acids plus LC-PUFAs may form more hydrophobic interfaces between the phospholipid bilayers, which could prevent the entry of hydrophilic ROS molecules [177]. Another explanation for the protection against oxidative stress might be in terms of a hydroxyl radical scavenging capacity of the double bonds of PUFAs [178].

Certain members of mesophilic terrestrial myxobacteria, especially of the genus *Aetherobacter*, are capable of synthesizing up to 20% LC-PUFAs of total fatty acids (TFAs). This implies important physiological roles of LC-PUFAs in these myxobacteria, which have not yet been investigated. Nonetheless, an antioxidative function of LC-PUFAs to protect the cells against ROS induced by ultraviolet radiation appears likelier than an adaptation mechanism regarding modulation of membrane fluidity. Additionally, it is hypothesized that there is a connection between LC-PUFA biosynthesis and secondary metabolism in myxobacteria. EPA produced by the *Sorangium cellulosum* strains SBSo026, So ce1128, So ce340, So ce427, and So ce690 may serve as biosynthetic precursor for the secondary metabolites maracin/maracen (Chapter 2). Also in case of the antibiotic zeamine produced by *Serratia plymuthica* RVH1, *pfa*-like genes have been identified in the corresponding biosynthetic gene cluster. It is postulated that the Pfa-like proteins and hybrid non-ribosomal polypeptide synthetases (NRPSs)/PKSs each synthesize parts of the backbone, which are linked together post-assembly [179].



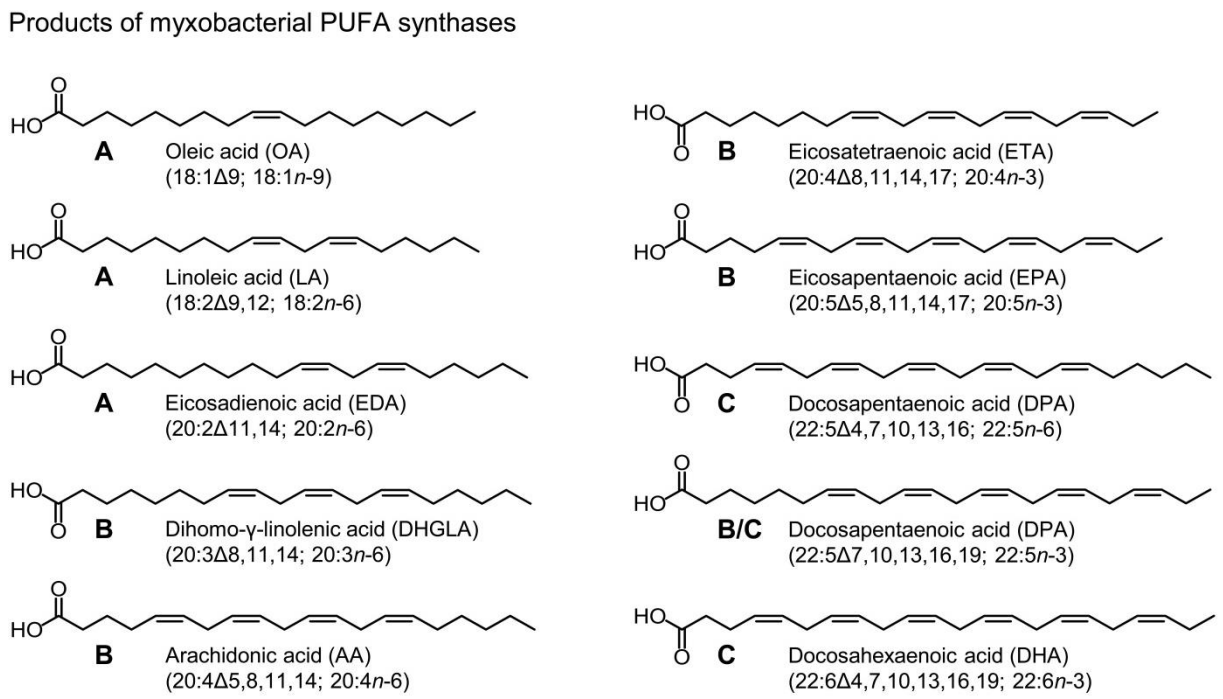


Figure 10. PUFA biosynthetic gene clusters from terrestrial myxobacteria of the *Polyangiaceae* family and their corresponding fatty acid products.
(A) OA/LA-type *pfa* gene cluster from *Sorangium cellulosum* (the number of ACP domains is ranging from four to five), production of LA > OA > EDA, (B) AA/EPA-type *pfa* gene cluster from *Sorangium cellulosum*, production of EPA > AA > *n*-3 DPA > ETA > DHGLA, (C) DPA/DHA-type *pfa* gene cluster from *Aetherobacter* sp., production of DHA > *n*-6 DPA > *n*-3 DPA.

Profound analyses of the product spectra of myxobacterial PUFA synthases

In the present work, selected myxobacterial PUFA biosynthetic pathways have been cloned, engineered, and transferred as well as expressed in the heterologous hosts *Myxococcus xanthus* DK1622, *E. coli* BL21 (DE3), *Pseudomonas putida* KT2440, and *Yarrowia lipolytica* Po1h. This approach enabled the characterisation of the products of a certain PUFA synthase by comparison of the fatty acid production profiles of the myxobacterial producers with those of corresponding transgenic host strains. However, recombinant LC-PUFA production revealed remarkable differences in the production profile between the native producers and the heterologous hosts as well as among the host strains themselves, which might be due to the activity of strain-specific housekeeping genes. Nonetheless, rational considerations enable the correlation between the different types of myxobacterial *pfa* gene clusters and the synthesized PUFA species (Figure 10).

In accordance with the results from feeding experiments [155], it can be deduced from heterologous expression in *M. xanthus* DK1622 ([155]; Chapter 2) that the OA/LA-type PUFA synthase from *S. cellulosum* is most likely responsible for the production of isomers of octadecenoic acid, EDA, and either of both OA plus LA or of simply OA, which can be converted into LA by an oleoyl-coenzyme A (oleoyl-CoA)-Δ^{12} desaturase (EC 1.14.19.6). Candidates for this desaturase could be identified by *in silico* analysis of the genome of *M. xanthus* DK1622 [130], which comprises 16 genes encoding mostly uncharacterized fatty acid desaturases. The genome sequence of *S. cellulosum* So ce56 encodes at least five uncharacterized fatty acid desaturases [131]. In contrast to *S. cellulosum*, *M. xanthus* DK1622 converts LA into GLA, probably catalysed by a linoleoyl-CoA desaturase (EC 1.14.19.3), and EDA into DHGLA, possibly via a Δ^{8} desaturase (EC 1.14.19.4).

Besides LA and EDA, the LC-PUFAs GLA, stearidonic acid (SDA, 18:4, *n*-3), DHGLA, AA, eicosatetraenoic acid (ETA, 20:4, *n*-3), EPA, and *n*-3 DPA are produced by *S. cellulosum* SBSo026 as well as by the transgenic host strain *M. xanthus* DK1622::pHybPfa1-mx9.2, harbouring a hybrid *pfa* gene cluster comprising gene *pfa1* from the OA/LA-type *pfa* gene cluster plus genes *pfa2* and *pfa3* from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026, in a similar ratio (Table 2). According to this, the proteins Pfa2 and Pfa3 from the AA/EPA-type PUFA synthase from *S. cellulosum* in collaboration with Pfa1 from the OA/LA-type PUFA synthase catalyse the synthesis of EPA as major product and of minor amounts of other C20 PUFAs with fewer double bonds (AA, ETA, and DHGLA). The biosynthesis of *n*-3 DPA can be accomplished via the elongation of EPA with malonyl-CoA

catalysed either by the hybrid PUFA synthase or, less likely, by a separate fatty acid elongase. Additional shunt products of the AA/EPA-type PUFA synthase may be GLA and SDA. However, these C18 PUFAs could also easily result from degradation of the C20 PUFAs ETA and DHGLA via β-oxidation.

Table 2. Ratios of LC-PUFAs synthesized by the native producer *S. cellulosum* SBSo026 or the heterologous host *M. xanthus* DK1622::pHybPfa1-mx9.2, comprising gene *pfa1* from the OA/LA-type *pfa* gene cluster plus genes *pfa2* and *pfa3* from the AA/EPA-type *pfa* gene cluster of *S. cellulosum* SBSo026.
In each case, the least produced LC-PUFA is given as 1.0.

PUFA	Ratio of produced LC-PUFAs	
	***S. cellulosum* SBSo026**	***M. xanthus* DK1622::pHybPfa1-mx9.2**
γ-linolenic acid (GLA, 18:3, *n*-6)	1.0	7.0
stearidonic acid (SDA, 18:4, *n*-3)	4.5	1.0
dihomo-γ-linolenic acid (DHGLA, 20:3, *n*-6)	1.2	2.5
arachidonic acid (AA, 20:4, *n*-6)	6.5	43.5
eicosatetraenoic acid (ETA, 20:4, *n*-3)	2.0	7.0
eicosapentaenoic acid (EPA, 20:5, *n*-3)	105.5	295.0
docosapentaenoic acid (DPA, 22:5, *n*-3)	3.5	2.5

Aetherobacter spp. have proven to be prolific produces of a broad range of LC-PUFAs because these strains synthesize AA, EPA, *n*-6 DPA, *n*-3 DPA, and DHA. The DPA/DHA-type PUFA synthase encoded by the *pfa* gene cluster of *A. fasciculatus* (SBSr002) was expressed in various heterologous hosts including *M. xanthus* DK1622, *P. putida* KT2440, and *Y. lipolytica* Po1h. The transgenic host strain *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2 produces the same LC-PUFAs in a similar ratio like *A. fasciculatus* (SBSr002) (Table 3). As with the native producer, both EPA and DHA are the predominant LC-PUFAs. Surprisingly, the transgenic hosts strains of *P. putida* KT2440 harbouring the native or a synthetic *pfa* gene cluster originating from *A. fasciculatus* (SBSr002) produce only *n*-6 DPA and DHA (Table 3). Trace amounts of AA, EPA, and *n*-3 DPA can be detected merely in potent overproducing strains of *P. putida* KT2440, but synthesis of the latter LC-PUFAs is not stably reproducible. The recombinant strain *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, which contains an artificial *pfa* gene cluster originating from *A. fasciculatus* (SBSr002), heterologously produces

AA, ETA, EPA, *n*-6 DPA, *n*-3 DPA, and DHA, with a clear preference for DHA (Table 3). From this data, in keeping with the results from feeding experiments ([155]; Chapter 3) and studies on production kinetics [129], it can be inferred that the PUFA synthase from *Aetherobacter* spp. may only produce DHA and minor amounts of the other C22 PUFAs *n*-6 DPA and *n*-3 DPA. The C20 PUFA EPA is likely a degradation product resulting from the β-oxidation of *n*-3 DPA and/or is derived from DHA via β-oxidation plus 2,4-dienoyl-CoA reductase (EC 1.3.1.34) and Δ^3-Δ^2-enoyl-CoA isomerase (EC 5.3.3.8), needed for the oxidation of unsaturated fatty acids. In the same degradation pathway, *n*-6 DPA could be converted into AA. Furthermore, there are strain-dependent differences in the synthesis ratios of *n*-6 DPA, *n*-3 DPA, and DHA, ranging from 2 (*n*-6 DPA) : 1 (*n*-3 DPA) : 97 (DHA) in the case of *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C to 1 (*n*-6 DPA) : 0 (*n*-3 DPA) : 5.5 (DHA) in the case of transgenic host strains of *P. putida* KT2440. Conceivably, the host-related intracellular milieu may have an influence on the functioning and consequently on the product spectrum of the PUFA synthase.

Table 3. Ratios of LC-PUFAs synthesized by the native producer *A. fasciculatus* (SBSr002) or the heterologous hosts *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2, *P. putida* KT2440 strains harbouring native or synthetic *pfa* gene clusters, and *Y. lipolytica* Po1h::pSynPfaPptAf2 clone C, comprising the DPA/DHA-type *pfa* gene cluster derived from *A. fasciculatus* (SBSr002).
In each case, the least produced LC-PUFA is given as 1.0.

PUFA	Ratio of produced LC-PUFAs			
	***A. fasciculatus* (SBSr002)**	***M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2**	***P. putida* KT2440 strains harbouring native or synthetic *pfa* gene clusters**	***Y. lipolytica* Po1h::pSynPfaPptAf2 clone C**
arachidonic acid (AA, 20:4, *n*-6)	5.5	1.0	–	6.0
eicosatetraenoic acid (ETA, 20:4, *n*-3)	–	–	–	1.0
eicosapentaenoic acid (EPA, 20:5, *n*-3)	42.5	2.5	–	13.0
docosapentaenoic acid (DPA, 22:5, *n*-6)	1.5	2.5	1.0	19.5
docosapentaenoic acid (DPA, 22:5, *n*-3)	1.0	1.1	–	9.0
docosahexaenoic acid (DHA, 22:6, *n*-3)	52.0	32.5	5.5	851.5

In-depth analyses of the biosynthetic mechanisms of myxobacterial PUFA synthases and the contributing 4'-phosphopantetheinyl transferases (PPTases)

Deeper understanding of the working principles of PUFA synthases is crucial to fully exploit their potential for directed biotechnological production of selected LC-PUFAs. Up to now, several genetic and biochemical experiments have been performed with marine *pfa* gene clusters [101;103;154;180]. Unfortunately, these results neither elucidated the exact biosynthetic function and underlying biochemistry of every catalytic domain from the PUFA synthase complex, nor identified the relevant factor(s) that determine the nature of the final LC-PUFA product.

LC-PUFA biosynthesis catalysed by iteratively acting PUFA synthases is closely related to the biosynthesis of SFAs or monounsaturated fatty acids (MUFAs) catalysed by FAS. In both assembly lines, the *apo*-form of ACP must initially be activated to the *holo*-form by transfer of the 4'-phosphopanthetheine (4'-PP) moiety from CoA to an invariant serine residue, as the growing fatty acid chain is linked to the 4'-PP prosthetic group of ACP via a thioester bond during biosynthesis. This posttranslational modification of ACP is catalysed by 4'-phosphopantetheinyl transferases (PPTases), which are divided into several types [50]. Sfp-type PPTases have been shown to be involved in LC-PUFA biosynthesis of *Shewanella pneumatophori* SCRC-2738, *Photobacterium profundum* SS9, and *Moritella marina* MP-1 [82;83]. A homologue of gene *pfaE* from *S. pneumatophori* SCRC-2738 and *M. marina* MP-1, encoding the PPTase involved in the 4'-phosphopantetheinylation of PUFA synthases, was found in the genome of *S. cellulosum* (sce5058) and *Aetherobacter* spp.. Indeed, it could be demonstrated via heterologous expression that the PPTase from *A. fasciculatus* (SBSr002) activates the tandem ACPs encoded by the *pfa* gene cluster of this strain *in vivo* (Chapter 3). Sfp-type PPTases involved in the biosynthesis of polyketides and non-ribosomal peptides incorporate the three highly conserved motifs P1, P2, and P3 [181]. Additively, the P0 motif could be identified at the amino-terminal region, and motif P1 can be separately recognized as P1a and P1b [82]. Conserved residues of P2 and P3 participate in Mg^{2+} binding, whereas conserved residues of P1 and P3 are involved in binding of the substrate CoA and in catalysis [182]. Consensus sequences of the motifs P0, P1b, P2, and P3 of the Sfp-type PPTases with broad substrate specificity Sfp from *Bacillus subtilis* [183] and MtaA from *Stigmatella aurantiaca* DW4/3-1 [184], of the PUFA synthase-specific Sfp-type PPTases PfaE from *S. pneumatophori* SCRC-2738 and *M. marina* MP-1, as well as of the PPTases from *S. cellulosum* and *Aetherobacter* spp. were defined by alignment of the proteins (Table 4). In terms of motif P1a, these PPTases can be divided into two groups. For Sfp, MtaA, and the

PPTases of *S. cellulosum* and *Aetherobacter* spp., the consensus sequence is Y/W/FGXP, while KXKP represents the consensus sequence of the PfaE proteins. A possible explanation for these deviations might be linked to the different substrates. The identified PPTases are the only Sfp-type PPTases encoded by the genomes of *S. cellulosum* and *Aetherobacter* spp., but these strains contain a number of NRP and PK biosynthetic pathways. Hence, like Sfp and MtaA, these PPTases may have the ability to posttranslationally activate a variety of carrier proteins. In contrast, the activity of PfaE proteins might be restricted to tandem ACPs of PUFA synthases.

Table 4. Conserved motifs in Sfp-type 4'-phosphopantetheinyl transferases (PPTases) involved in the biosynthesis of non-ribosomal peptides, polyketides, and polyunsaturated fatty acids.

PPTase	P0	P1		P2	P3
		P1a	P1b		
Sfp (*Bacillus subtilis*)	VRSVIS	YGKP	FNISHS	GIDIE	WSMKESFIK
MtaA (*Stigmatella aurantiaca* DW4/3-1)	VRLTLS	YGRP	FNLSHT	GADVE	WTLKEAYIK
PfaE (*Shewanella pneumatophori* SCRC-2738)	LRALLS	KGKP	FNVSHS	GVDIE	WALKESYIK
PfaE (*Moritella marina* MP-1)	VRDLLS	KDKP	FNISHT	GCDVE	WTLKESYIK
PPTase (*Sorangium cellulosum*)	VRXV/TLS	F/YGRP	FNLSNT	GVDVE	WTLKESYIK
PPTase (*Aetherobacter* spp.)	VRSVLS	WGCP	FNLSNT	GVDVE	WTLKEAYIK
Consensus	**L/VRXXL/IS**	**XGXP KXKP**	**FNXSH/NT/S**	**GXDV/IE**	**WXL/MKES/AY/FIK**

In the biosynthesis of LC-PUFAs catalysed by PUFA synthases, acetyl-CoA serves as starter molecule, whereas malonyl-CoA, formed by carboxylation of acetyl-CoA using acetyl-CoA carboxylase (EC 6.4.1.2) [48], is required for all elongation steps (Figure 12). The pathway proceeds in two stages, initiation and cyclic elongation. In the initiation phase, an acyltransferase (AT) domain transfers the malonyl group of malonyl-CoA to an ACP. The AA/EPA-type *pfa* gene cluster from *S. cellulosum* and the DPA/DHA-type *pfa* gene cluster from *Aetherobacter* spp. encode domain AT1 within gene *pfa2* and domain AT2 in gene *pfa3* (Figure 10 and 11). Interestingly, the AT2 domain is missing in the OA/LA-type PUFA synthase from *S. cellulosum* (Figure 10). Thus, in conformity with the results obtained from

mutated or chimeric PUFA synthases (Chapter 6), it can be speculated that domain AT1 prevails in the early cycles of the elongation, approximately until the fatty acid chain consists of 18 carbon atoms. In the later rounds of elongation, the activity of domain AT2 appears to be dominating over the AT1 domain and might contribute to the determination of the final chain length of the produced LC-PUFAs (Figure 12).

At the onset of the elongation cycle, the decarboxylative Claisen condensation of malonyl-ACP with the acetyl group or the growing acyl chain takes places at the active site cysteine of the KS domain. Again, redundancy of domains in PUFA synthases emerges, as all *pfa* gene clusters encode domain KS1 within gene *pfa2* and domain KS2 in gene *pfa3* (Figure 10). Functionality of domain KS2 is supposed to depend on direct interaction with the CLF domain, a KS domain that lacks the active site cysteine and is characteristic for type II PKS systems [94]. The results from mutated PUFA synthases (Chapter 6) are indicative that domain KS1 is only active, if at all, in the early cycles of elongation. In the later rounds of elongation, the KS2 domain is expected to catalyse the decarboxylative Claisen condensation and to determine the chain length of the produced LC-PUFAs in cooperation with the CLF domain (Figure 12), as deciphered by rational domain swaps (Chapter 6). Similar results, that indicate the proposed function of the KS/CLF pair, have been obtained for the common type II PKS systems from streptomycetes synthesizing the octaketide actinorhodin from *S. coelicolor* A3(2) [185], the decaketide tetracenomycin from *S. glaucescens* [186], the dodecaketide WhiE pigment from *S. coelicolor* A3(2) [187], or the pentadecaketide fredericamycin from *S. griseus* [188]. *In vivo* experiments with various PKS hybrids suggested that the CLF controls the number of Claisen condensations during polyketide assembly and thereby the chain length [94]. Moreover, chimeric KS and CLF genes designed from sequences of actinorhodin and tetracenomycin biosynthetic genes reinforced the importance of CLF in defining chain length and proposed that the regions of KS and CLF which are most important for activity and specificity are located at the interface of this dimer [189]. The model of a protein cavity which controls the size of the growing chain is strongly supported by the finding that specific residues in the CLF of the actinorhodin biosynthetic pathway serve as gatekeepers in the polyketide tunnel. Reducing the size of these residues by site-directed mutagenesis lengthened the channel and allowed for two more elongation cycles [95]. In the context of KS/CLF structural studies, the number of the residues forming a gate at the interface of the dimer was extended, and it was proposed that novel chain lengths could be engineered by opening or closing the gates [96]. Regrettably, *in silico* analyses of the protein sequences unveiled that the consensus of these CLFs from streptomycetes and the CLF

domains of myxobacterial PUFA synthases are insufficient to assign the gatekeeping residues in the latter protein species. However, detailed sequence comparisons of the DPA/DHA-type PUFA synthase from *Aetherobacter* spp., the OA/LA-type PUFA synthase and the AA/EPA-type PUFA synthase from *S. cellulosum* strains revealed that the KS2/CLF domains of the particular types of PUFA synthases mainly differ from each other in several insertions at specific positions (Figure 11). These observations indicate that conversion of a chain length determinant specific for PUFAs with a longer chain length into a factor specific for a shorter chain length requires the sum of the differences in several regions within the KS2/CLF domains rather than simple single amino acid substitutions.

After elongation of the fatty acyl chain by two carbons, the β-keto group of the biosynthetic intermediate is either fully reduced by sequential action of NADPH-dependent KR domain, PKS-like DH domain, and NADPH-dependent enoyl reductase (ER) domain or only reduced to the *trans* double bond by the KR domain and a FabA-like DH domain and isomerized to the *cis* double bond by the latter domain. Both FabA-like DH domains of the PUFA synthases, DH2 and DH3, turned out to be essential for LC-PUFA biosynthesis, as shown by inactivation studies (Chapter 6). Based on the biosynthetic logic, the *trans*-2 biosynthetic intermediates have to be alternately isomerized to the *cis*-3 intermediate, similar to the reaction catalysed by the β-hydroxydecanoyl-ACP dehydratase FabA from α- and γ-Proteobacteria [52], or converted into the *cis*-2 intermediate (Figure 12). It seems plausible to assume that each of the domains DH2 and DH3 may catalyse only one of these different isomerization reactions. Unfortunately, as no short-chain biosynthetic intermediate could be detected for the heterologous host strains harbouring the corresponding mutated *pfa* gene cluster (Chapter 6), the two individual FabA-like DH domains cannot be assigned to a distinct isomerization. Likewise, *in silico* analyses of protein sequences are not able to clarify this issue, as the outcomes from sequence alignments of FabA from *E. coli* and either DH2 or DH3 are almost indistinguishable (29% versus 26% pairwise identity / 46% versus 44% pairwise similarity).

Figure 11. Sequence alignment of the KS2/CLF domains of the OA/LA-type PUFA synthase from *Sorangium cellulosum* So ce56, So ce377, So ce38, So ce1525, and So ceGT47, of the DPA/DHA-type PUFA synthase from *Aetherobacter* spp. Sr001, Sr002, Sr003, and Sr008, and of the AA/EPA-type PUFA synthase from *S. cellulosum* So026 and So ce1128.

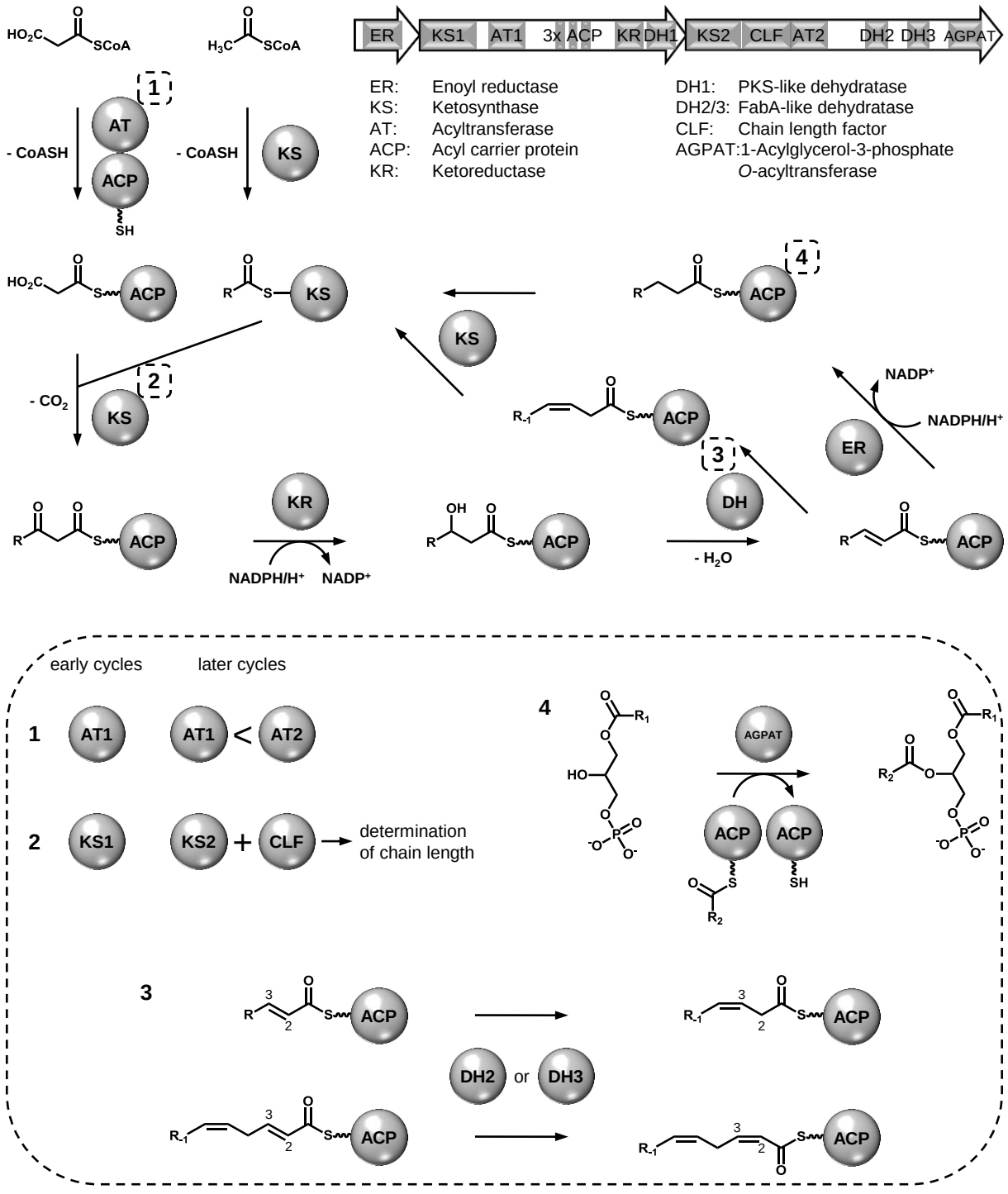


Figure 12. Anaerobic biosynthesis of polyunsaturated fatty acids (PUFAs) by myxobacterial iterative type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like PUFA synthases encoded by a *pfa* gene cluster. The primer molecule (in the form of acetyl-CoA) undergoes several rounds of decarboxylative Claisen condensation reactions, resulting in the elongation of the fatty acyl chain by two carbons (derived from malonyl-CoA) per cycle. Following each round of elongation, the β-keto group is either fully reduced or only reduced to the *trans* double bond which is then isomerized with or without shifting the position (3). The intermediate then functions as a starter substrate for the next round of elongation with malonyl-ACP, and an acyl chain with methylene-interrupted *cis* double bonds is synthesized. The chain length is probably determined by the action of AT2 and/or CLF (1+2). The fatty acyl-ACP can then serve as substrate for the acylation of the 2-position of 1-acyl-*sn*-glycerol-3-phosphate catalysed by AGPAT during glycerophospholipid biosynthesis (4).

When the fatty acid chain with methylene-interrupted *cis* double bonds reaches its final length, the product is either released from ACP as free fatty acid by a discrete thioesterase, as presumed for *Schizochytrium* sp. ATCC 20888 [87] and *P. profundum* SS9 [88], or the acyl chain is directly transferred from ACP into glycerophospholipids. In the latter case, the acylation of the 2-position of 1-acyl-*sn*-glycerol-3-phosphate utilising either acyl-ACP or acyl-CoA thioesters as substrate can be performed by AGPAT [190] (Figure 12). An AGPAT domain was also discovered *in silico* at the carboxyl-terminus of Pfa3. Since its activity could not been verified via an *in vitro* enzyme assay, detailed sequence analyses may help to provide some information about its functionality. Four highly conserved motifs were found to be incorporated into AGPATs, such as PlsC from *E. coli* [191] and human AGPAT1 [192]. Unfortunately, basic information regarding topology of PlsC and the key residues involved in catalysis are lacking. However, for the human homologue, site-directed mutagenesis revealed the importance of all the motifs for AGPAT catalysis, and important catalytic amino acid residues were identified [192] (Table 5). These results suggest that conserved residues of motifs 2 and 3 are involved in binding of 1-acyl-*sn*-glycerol-3-phosphate and motifs 1 and 4 are involved in acyl-CoA binding and catalysis [191;192]. Sequence alignments of PlsC, AGPAT1, and AGPAT domains of the PUFA synthases from *S. cellulosum* and *Aetherobacter* spp. were performed, disclosing the conserved motifs within the myxobacterial AGPAT domains (Table 5). Amino acid residues identical or at least similar to those of PlsC and AGPAT1 could be found at putative catalytic sites, which serve as indicator for a proper, proposed functioning of the AGPAT domains.

Table 5. Conserved motifs in 1-acylglycerol-3-phosphate *O*-acyltransferases (AGPATs) involved in the glycerophospholipid biosynthesis.
The most important catalytic amino acid residues determined for AGPAT1 [192] as well as the glutamic acid residue exchanged by site-directed mutagenesis of the AGPAT domain from *Aetherobacter fasciculatus* (SBSr002), leading to a very low LC-PUFA production (< 10% of the production rate of non-mutated PUFA synthase) (Chapter 6), are shown in italics.

AGPAT	Motif 1	Motif 2	Motif 3	Motif 4
PlsC (*Escherichia coli*)	NHQNNYD	LIDR	MFPEGTR	VPIIPV
AGPAT1 (*Homo sapiens*)	N*H*QSSL*D*	*F*ID*R*	VFP*EGTR*	VPIV*PI*
AGPAT domain of OA/LA-type PUFA synthase (*Sorangium cellulosum*)	NHQVG/AVE	YFDR	VHVEGTR	V/APVVPV
AGPAT domain of AA/EPA-type PUFA synthase (*Sorangium cellulosum*)	NHQVAVE	FFDR	VHVEGTR	APIVPV
AGPAT domain of DPA/DHA-type PUFA synthase (*Aetherobacter* spp.)	NHQVALE	FFDR	VHV*E*GTR	A/TP/LIVPV
consensus	**NHQXXXD/E**	**XXDR**	**V/MXXEGTR**	**XP/LI/VV/IPV/I**

In conclusion, it can be stated that the identification of different types of myxobacterial PUFA synthases and the establishment of heterologous expression platforms, which allow genetic engineering of the myxobacterial PUFA biosynthetic pathways, provided the basis for studying the exact working principles of PUFA synthases and thus, for full exploitation of their potential for directed biotechnological production of selected LC-PUFAs.

Exploring different host strains for efficient heterologous PUFA production employing myxobacterial PUFA synthases

Myxobacteria have been characterized as producers of a wide range of different LC-PUFAs, such as AA, EPA, or DHA [127-129], employing unique biosynthetic gene clusters [155]. Regretfully, most myxobacteria are characterized by slow growth behaviour, and particularly for novel myxobacterial isolates, tools for the genetic modification are not readily available. Hence, Synthetic Biotechnology approaches for lateral transfer of the PUFA biosynthetic

pathways into a well-developed surrogate host and for further engineering of the host strain are an attractive alternative to overproduce the LC-PUFAs of interest.

In the present work, the OA/LA-type *pfa* gene clusters from *S. cellulosum* So ce56 and SBSo026, the AA/EPA-type *pfa* gene cluster from *S. cellulosum* SBSo026, and the DPA/DHA-type *pfa* gene clusters from *A. fasciculatus* (SBSr002) and *Aetherobacter* sp. (SBSr008) have been transferred and expressed into the myxobacterial model strain *M. xanthus* DK1622 ([155]; Chapter 2). Therefore, the gene clusters were directly cloned from the genome of the producing strain into cloning vectors via linear plus linear homologous recombination in *E. coli* [143]. The obtained plasmids were further engineered by applying conventional cloning techniques or linear to circular homologous Red/ET recombination [144] in order to insert genetic elements adapted to the host organism, such as genes for the integration into the genome and heterologous promoters (Tn5 promoter plus tetracycline promoter). The yields of DHA produced by the native producer *A. fasciculatus* (SBSr002) and the heterologous host (1.6 mg/l versus 1.2 mg/l) are comparable. However, using *M. xanthus* DK1622::pPfaAf-P_{tet}-mx9.2, harbouring the DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) site-specifically integrated at Mx9 *attB2* site of the genome, the fermentation time is 3.5-fold reduced (two versus seven days) (Figure 13). In general, *M. xanthus* is readily utilised for the heterologous expression of gene clusters from slower growing myxobacteria. Common compounds produced in this manner are epothilone, myxochromide S, and myxothiazol. While the yield of recombinant epothilone was disappointing (0.1 mg/l in *M. xanthus* DK1622 [193]), production could be doubled for myxothiazol (20 mg/l in *M. xanthus* DZF1 versus 10 mg/l in the native producer *S. aurantiaca* DW4/3-1 [194]) and enormously raised in case of myxochromide S (1 g/l in *M. xanthus* DK1622 versus 8 mg/l in the native producer *S. aurantiaca* DW4/3-1 [193]). According to this, *M. xanthus* has been proven to be a popular host for the heterologous expression of myxobacterial gene clusters, albeit the current unpredictability of the yield.

Successful heterologous expression of a biosynthetic pathway in genetically distant microbes poses a major challenge, as codon usage bias and functionality of regulatory elements and promoter structures in both native producer and host are unlikely. The DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) was cloned into replicative plasmids under control of inducible promoter systems (*xylS-Pm* or *araC*-P_{BAD}) suitable for the easy to handle host strains *E. coli*, the best studied bacterial model strain, as well as *P. putida* KT2440, a genetically amenable and metabolically versatile microbe [146], which is generally

recognized as safe (GRAS) (Chapter 3). Using initial expression constructs, *P. putida* KT2440 turned out to be a fast-growing and reliable host organism for heterologous LC-PUFA production with a higher and more robust production rate of DHA compared to *E. coli* (0.3 mg/l after 28 h of cultivation for *P. putida* KT2440 + pJBPfaAf1 versus 0.2 mg/l after 27 h of cultivation for *E. coli* BL21(DE3)/pJBPfaAf4/pPptAfD) (Figure 13). The low GC-content of the genome sequence from *E. coli* and its codon usage make it apparently rather unsuitable for the expression of myxobacterial genes. Using the Graphical Codon Usage Analyser software tool [195], the fraction of usage of each codon in the sequences of the DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) was computed and plotted against the fraction of usage of the codon in the hosts. In fact, a mean difference of 17% was obtained for *P. putida*, whereas a value of 48% was acquired for *E. coli*. The assertion, that deviations in the codon usage of the native producer and *E. coli* are the most reasonable explanation for the low production rate, is supported by the outcome of an additional comparison: The mean difference of codon usage in the coding sequences from *E. coli* and in the EPA-type *pfa* gene cluster from the marine γ-Proteobacterium *S. pneumatophori* SCRC-2738 is only 27%, and in fact, heterologous expression of this gene cluster located on a cosmid under control of its native promoters in *E. coli* yielded 3.3 mg/g cell dry weight (CDW) after 72 h of cultivation [196]. In contrast, *E. coli* BL21(DE3)/pJBPfaAf4/pPptAfD only synthesizes 0.1 mg DHA/g CDW after 27 h of cultivation (Chapter 3). The produced amounts of EPA plus DHA among the corresponding native producers, however, are comparable [76;155].

In the follow-up work, *P. putida* was chosen as preferred chassis for expression of the *pfa* gene cluster, and efforts were undertaken to further improve the production yield. Modification of the expression constructs to increase the expression rate of the *pfa* genes, genetic and metabolic engineering of the strain (Chapter 3), and adaptation of the sequences of the *pfa* genes to the host (Chapter 4) gave rise to a ten-fold increase in DHA production rate (3 mg/l after 28 h of cultivation for *P. putida* KT2440::pfadH_KO/pME2 + pPm**SynPfaAf2a). This gain significantly surpasses the LC-PUFA production rate achieved with the native myxobacterial producer, even in a six-fold reduced fermentation time (Figure 13). Despite the obvious potential of pseudomonads in biotechnological processes, only a few examples for the heterologous expression of myxobacterial natural product biosynthetic pathways in these bacteria are described in literature. Prominent instances for this involve gene clusters from *S. aurantiaca* DW4/3-1: An enormous increase in the production rate was achieved in case of myxochromide S in the heterologous host *P. putida* KT2440

(production of up to 40 mg/l in two to three days) compared to the native producer (production of 8 mg/l in six days) [197]. However, a drastic reduction in the yield of myxothiazol (production of 10 mg/l in the native producer) could be observed using the heterologous host *P. putida* KT2440 (production of 0.6 mg/l in two to three days) [198].

Techniques and methodologies of Synthetic Biology were exploited to implement redesign and subsequent synthesis of the DPA/DHA-type *pfa* gene cluster from *A. fasciculatus* (SBSr002) adapted to the hemiascomycetous, oleaginous yeast *Y. lipolytica* under control of the strong hybrid hp4d promoter [199] (Chapter 5). This yeast has emerged as a promising host for heterologous protein production [199-201], and several processes based on *Y. lipolytica* were classified as GRAS by the food and drug administration (FDA). The transgenic strain *Y. lipolytica* Po1h::pSynPfaAf2 clone C is by far the best heterologous LC-PUFA producer utilizing the DPA/DHA-type myxobacterial PUFA synthase (70.4 mg/l after seven days of cultivation) and thus clearly beats the native producer with a 44-fold increase in DHA yield (Figure 13).

Over the course of studies on the heterologous expression of myxobacterial *pfa* gene clusters, it became obvious that in case of integrative expression constructs the random integration into the genome of a host yielded highly productive clones. This has been demonstrated for transgenic strains of *Y. lipolytica* Po1h (Chapter 5) or *M. xanthus* DK1622 [155] and can be illustrated well by an example: In regard to production yields of octadecenoic acid, LA, GLA plus EDA, integration of the OA/LA-type *pfa* gene cluster from *S. cellulosum* So ce56 into the genome of *M. xanthus* DK1622 via transposition and screening for a good producer resulted in a yield of 11.3 mg/l [155]. Hence, this strategy was superior to the site-specific integration of the homologous gene cluster from *S. cellulosum* SBSo026 at Mx9 *attB2* site of *M. xanthus* DK1622, which generated 3.4 mg/l (Chapter 2). The produced amounts of these fatty acids among the corresponding native producers, however, are comparable ([155]; Chapter 2).

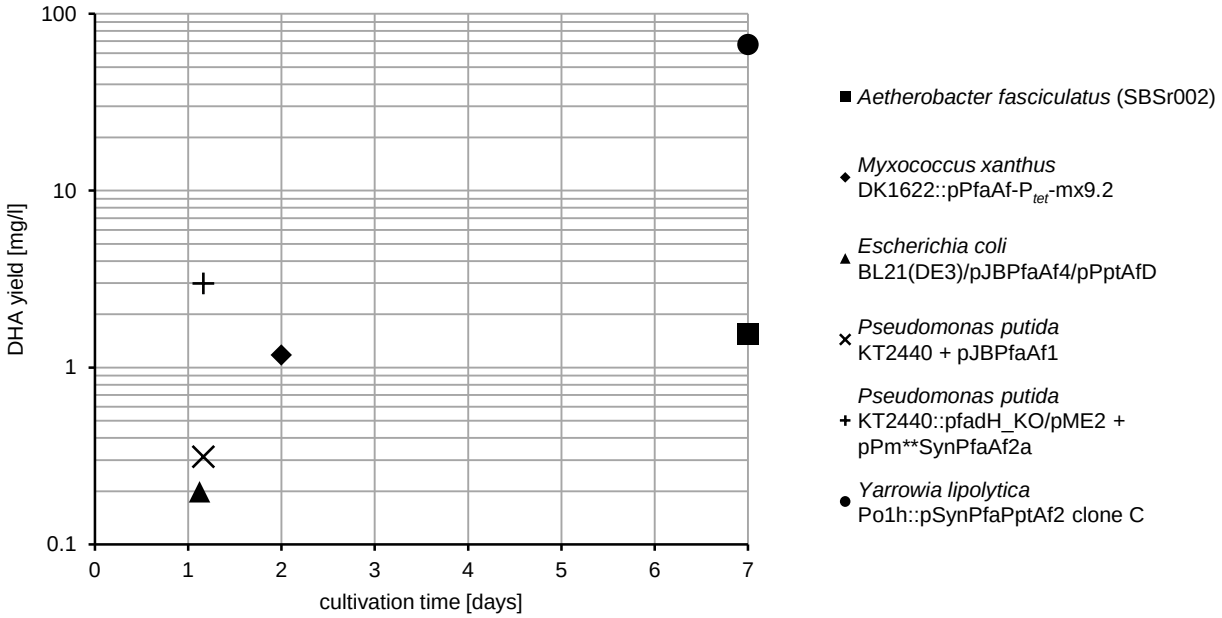


Figure 13. Yields of DHA produced by *Aetherobacter fasciculatus* (SBSr002) and different host strains with native or synthetic DPA/DHA-type *pfa* gene clusters after appropriate cultivation times in a shake flask.

Comparison of different recombinant enzyme systems for LC-PUFA production in the oleaginous yeast *Yarrowia lipolytica*

As already mentioned, the yeast *Y. lipolytica* has turned out to be especially suitable for heterologous LC-PUFA production not least because of the ability to accumulate large amounts of lipids (up to 50% of its cell dry weight, depending on the growth conditions) [202]. As a consequence, DuPont (USA) invested in the engineering of a *Y. lipolytica* strain capable of producing large amounts of EPA [203]. Contrary to *Y. lipolytica* Po1h::pSynPfaAf2 clone C, which employs a DPA/DHA-type myxobacterial PUFA synthase for heterologous LC-PUFA production, DuPont made use of alternatingly acting position-specific desaturases and elongases (Figure 14). These aerobic enzyme systems allow for the introduction of double bonds into SFAs and MUFAs synthesized by FAS and their elongation. In order to ensure efficient expression of the heterologous genes in *Y. lipolytica*, integration vectors with codon-optimized coding sequences driven by strong *Y. lipolytica* promoters were constructed [203]. In total, a C16/18 elongase gene originating from *Mortierella alpina*, two copies of a Δ^{12} desaturase gene originating from *Fusarium moniliforme*, three copies of a Δ^{9} elongase gene and two copies of a mutant Δ^{8} desaturase gene originating from *Euglena gracilis*, three copies of Δ^{5} desaturase genes originating from *E. gracilis* and *Peridinium* sp., and three copies of a Δ^{17} desaturase gene originating from

Pythium aphanidermatum have been integrated into the genome and expressed. In addition to the fatty acids produced by the corresponding wild type, the resulting strain *Y. lipolytica* Y4086 synthesizes 6.9% α-linolenic acid (ALA, 18:3, *n*-3), 7.6% EDA, 1% DHGLA, 2% ETA, and 9.8% EPA of TFAs [204]. As opposed to this, simply four synthetic genes encoding the DPA/DHA-type myxobacterial PUFA synthase and the PPTase originating from *A. fasciculatus* (SBSr002) were integrated into the genome of *Y. lipolytica* Po1h::pSynPfaAf2 clone C and expressed (Chapter 5). This strain produces 8.9% DHA of TFAs, thus amounts comparable to those of EPA produced by Y4086, but only 0.5% of TFAs of the unpreferred LC-PUFA shunt or degradation products. Substantial advantages of LC-PUFA biosynthesis via PUFA synthases as against exploitation of the aerobic pathways are the low consumption of NAD(P)H and the independence from endogenous fatty acids as biosynthetic precursors. For instance, *de novo* synthesis of EPA catalysed by PUFA synthases merely relies on ten molecules of acetyl-CoA/malonyl-CoA and 13 molecules of NADPH. However, using the aerobic route in which palmitic acid (synthesized by FAS using eight molecules of acetyl-CoA/malonyl-CoA and 14 molecules of NADPH) is converted into EPA via diverse fatty acid desaturases and elongases, additionally two molecules malonyl-CoA plus nine molecules NAD(P)H are consumed.

Beyond the integration of an aerobic EPA biosynthetic pathway, the transgenic strain Y4086 was further engineered by DuPont to improve the flow of fatty acids into EPA biosynthesis [203]. Accordingly, one additive copy of the Δ^{12} desaturase gene as well as of the Δ^{9} elongase gene plus two additional copies of mutant Δ^{8} desaturase genes have been integrated and expressed. Remarkably, the generated strain Y4128 produces 38% EPA of TFAs. Taking a closer look at the genome sequence of this strain disclosed that the integration cassette of the expression plasmid inadvertently interrupted the coding region of the peroxisomal biogenesis factor 10 (*PEX10*) gene. Pex10p is necessary for normal peroxisome morphology and functions in the peroxisomal protein import by interacting with other Pex proteins [205]. It can be deduced from further experiments that the loss-of-function mutation causes a defect in β-oxidation and therefore has a considerable impact on lipid accumulation [203]. Subsequent strain engineering involved the integration and expression of two additive copies of the Δ^{12} desaturase gene, three copies of Δ^{9} elongase genes originating from *E. gracilis* and *Eutreptiella* sp., three additional copies of mutant Δ^{8} desaturase genes, two extra copies of the Δ^{5} desaturase gene originating from *E. gracilis*, and two copies of the cholinephosphotransferase 1 gene from *Y. lipolytica*. Other open reading frames of interest which were accidentally disrupted by integration events are the lipase 1 gene *LIP1* and gene

SCP2, encoding a sterol carrier protein. The final strain Y4305 contains 30 copies of nine different genes and produces EPA at 56.6% of TFAs (mainly incorporated into triacylglycerols) and less than 5% SFAs of TFAs. The resulting yeast based production process has already been translated into an industrial process as sustainable source of EPA, and two commercial products have been developed: NewHarvest EPA-rich oil (a human nutritional supplement) and Verlasso sustainably farmed salmon (where the yeast serves as aquaculture feed) [203].

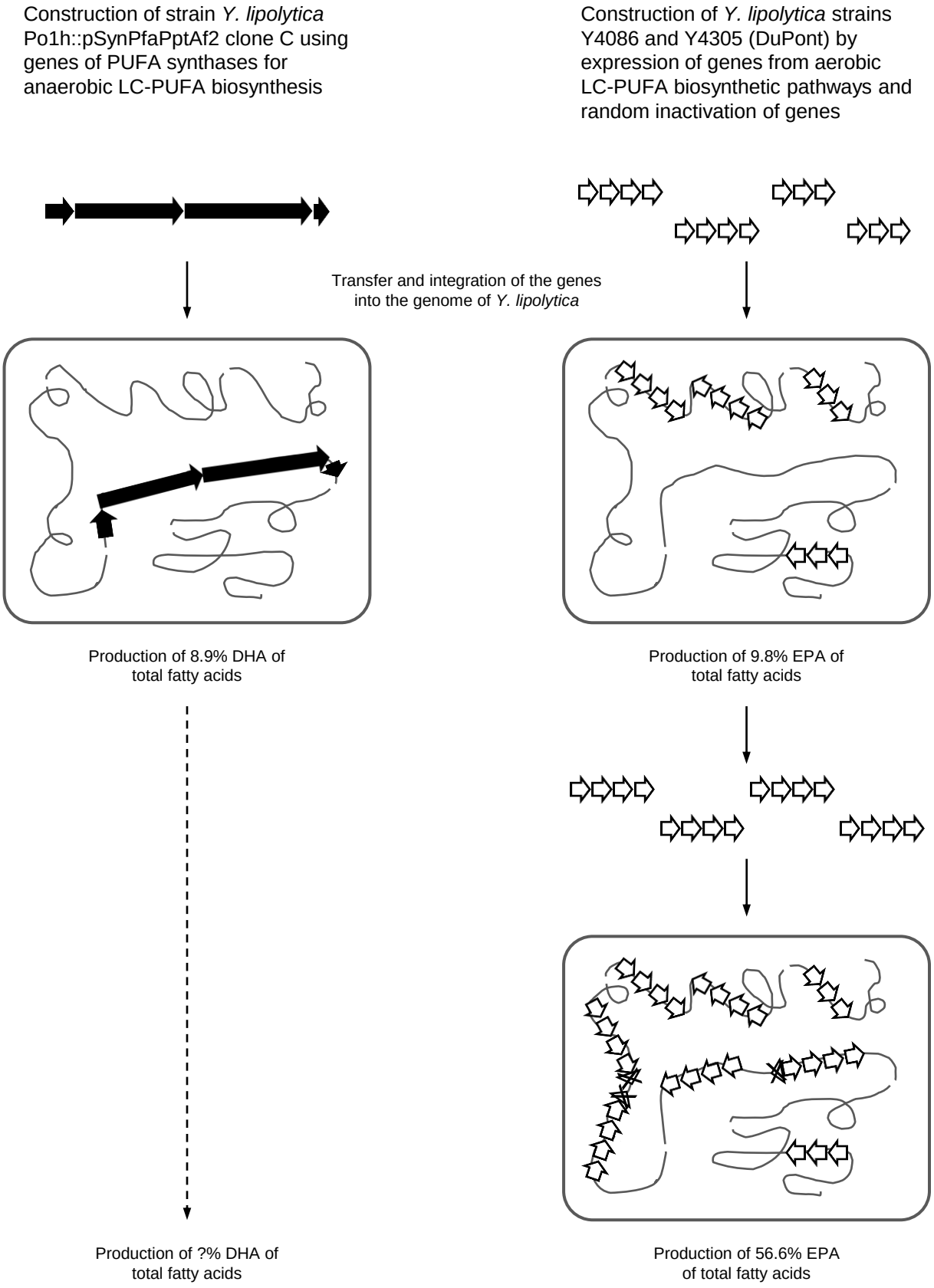


Figure 14. Simplified scheme for the comparison of LC-PUFA production in recombinant strains of *Yarrowia lipolytica* engineered in the present work and by DuPont.

OUTLOOK

In the present work, three different types of myxobacterial polyunsaturated fatty acid (PUFA) synthases with distinct product spectra have been identified and characterized from *S. cellulosum* and *Aetherobacter* spp.. These kinds of iteratively acting type I fatty acid synthase (FAS)/polyketide synthase (PKS)-like enzymes are specific for *de novo* biosynthesis of the major products linoleic acid (LA, 18:2, *n*-6), eicosapentaenoic acid (EPA, 20:5, *n*-3), or docosahexaenoic acid (DHA, 22:6, *n*-3). Moreover, some myxobacteria, such as *Plesiocystis pacifica*, *Phaselicystis flava*, and *Minicystis rosea*, have been found to produce considerable amounts of arachidonic acid (AA, 20:4, *n*-6) [124;127]. There is evidence that these myxobacteria synthesize this long-chain PUFA (LC-PUFA) by use of AA-type PUFA synthases. Characterization and analysis of the putative fourth type of myxobacterial PUFA synthase in comparison with the already described myxobacterial *de novo* PUFA biosynthetic pathways would contribute to the understanding of the mechanisms underlying determination of the main LC-PUFA product.

Establishment of an efficient expression system for heterologous PUFA production employing myxobacterial PUFA synthases is a major challenge. Successful DHA production could be implemented in the host *Pseudomonas putida* KT2440 by use of the PUFA biosynthetic pathway from *Aetherobacter fasciculatus* (SBSr002). Expression of the biosynthetic genes was significantly improved using mutated versions of a heterologous promoter. Therefore, construction of a synthetic promoter library, identification of strong promoter variants prior to further tuning expression of the recombinant PUFA genes appears to be the obvious choice for future experiments. Nevertheless, the oleaginous yeast *Yarrowia lipolytica* expressing the PUFA synthase from *A. fasciculatus* (SBSr002) turned out to be by far the best of the generated DHA producers. Hence, most efforts should be directed towards further increase of the DHA production rate in the transgenic strain. Obviously, this will involve metabolic engineering in order to achieve accumulation of recombinantly produced LC-PUFAs (e.g. by preventing degradation of DHA via β-oxidation and/or by diverting allocation of incorporated DHA from the phospholipid fraction toward triacylglycerols) in combination with the establishment of tailored fermentation and downstream processes. Besides, exploitation of the diversity of myxobacterial PUFA biosynthetic pathways is targeted to create chimeric PUFA synthases for *Y. lipolytica* with deliberately directed product spectra.

REFERENCES FOR INTRODUCTION AND DISCUSSION

[1] Santos CR & Schulze A (2012) Lipid metabolism in cancer. *FEBS J.*, **279,** 2610-2623.

[2] Leray C (2013) Cyberlipid Center, http://www.cyberlipid.org/fa/acid0001.htm#top (access on 23.01.2014).

[3] American Oil Chemists' Society (2014) AOCS Lipid Library, http://lipidlibrary.aocs.org/Lipids/fa_eic.html (access on 23.01.2014).

[4] Dyerberg J, Bang HO, & Hjorne N (1975) Fatty acid composition of the plasma lipids in Greenland Eskimos. *Am. J. Clin. Nutr.*, **28,** 958-966.

[5] Okuda N, Ueshima H, Okayama A, Saitoh S, Nakagawa H, Rodriguez BL, Sakata K, Choudhury SR, Curb JD, & Stamler J (2005) Relation of long chain n-3 polyunsaturated fatty acid intake to serum high density lipoprotein cholesterol among Japanese men in Japan and Japanese-American men in Hawaii: the INTERLIPID study. *Atherosclerosis*, **178,** 371-379.

[6] Svennerholm L (1968) Distribution and fatty acid composition of phosphoglycerides in normal human brain. *J. Lipid Res.*, **9,** 570-579.

[7] Janssen CI & Kiliaan AJ (2013) Long-chain polyunsaturated fatty acids (LCPUFA) from genesis to senescence: The influence of LCPUFA on neural development, aging, and neurodegeneration. *Prog. Lipid Res.*, **53C,** 1-17.

[8] Smith WL & Murphy RC (2002) The eicosanoids: cyclooxygenase, lipoxygenase, and epoxygenase pathways. In *Biochemistry of Lipids. Lipoproteins and Membranes (4th Edn.)* (Vance DE & Vance JE, eds), pp. 341-371. Elsevier Science B.V.

[9] Simopoulos AP (2008) The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp. Biol. Med.*, **233,** 674-688.

[10] Lorente-Cebrian S, Costa AGV, Navas-Carretero S, Zabala M, Martinez JA, & Moreno-Aliaga MJ (2013) Role of omega-3 fatty acids in obesity, metabolic syndrome, and cardiovascular diseases: a review of the evidence. *J. Physiol. Biochem.*, **69,** 633-651.

[11] Smith WL, Dewitt DL, & Garavito RM (2000) Cyclooxygenases: structural, cellular, and molecular biology. *Annu. Rev. Biochem.*, **69,** 145-182.

[12] Das UN (2006) Essential fatty acids: biochemistry, physiology and pathology. *Biotechnol. J.*, **1,** 420-439.

[13] Calder PC (2001) Polyunsaturated fatty acids, inflammation, and immunity. *Lipids*, **36,** 1007-1024.

[14] Serhan CN, Yacoubian S, & Yang R (2008) Anti-inflammatory and proresolving lipid mediators. *Annu. Rev. Pathol.*, **3,** 279-312.

[15] Serhan CN, Hong S, Gronert K, Colgan SP, Devchand PR, Mirick G, & Moussignac RL (2002) Resolvins: a family of bioactive products of omega-3 fatty acid transformation circuits initiated by aspirin treatment that counter proinflammation signals. *J. Exp. Med.*, **196,** 1025-1037.

[16] Emanuela F, Grazia M, Marco dR, Maria PL, Giorgio F, & Marco B (2012) Inflammation as a link between obesity and metabolic syndrome. *J. Nutr. Metab*, **2012,** 476380.

[17] Manku MS, Horrobin DF, Huang YS, & Morse N (1983) Fatty acids in plasma and red cell membranes in normal humans. *Lipids*, **18,** 906-908.

[18] Sears B & Ricordi C (2011) Anti-inflammatory nutrition as a pharmacological approach to treat obesity. *J. Obes.*, **2011**.

[19] Andersson A, Nalsen C, Tengblad S, & Vessby B (2002) Fatty acid composition of skeletal muscle reflects dietary fat composition in humans. *Am. J. Clin. Nutr.*, **76,** 1222-1229.

[20] Food and Agriculture Organization of the United Nations (2010) Fats and fatty acids in human nutrition. Report of an expert consultation. *FAO Food Nutr. Pap.*, **91,** 1-166.

[21] European Food Safety Authority Panel on Dietetic Products, Nutrition, and Allergies (NDA) (2010) Scientific opinion on dietary reference values for fats, including saturated fatty acids, polyunsaturated fatty acids, monounsaturated fatty acids, *trans* fatty acids, and cholesterol. *EFSA J.*, **8,** 1461.

[22] German Society of Nutrition (2009) D-A-CH-Referenzwerte für die Nährstoffzufuhr, http://www.dge.de/modules.php?name=News&file=print&sid=920 (access on 23.01.2014).

[23] *What we eat in America*, NHANES 2009-2010, U.S. Department of Agriculture, Agricultural Research service (2012) nutrient intakes from food: Mean amounts consumed per individual, by gender and age.

[24] Bauch A, Lindtner O, Mensink GB, & Niemann B (2006) Dietary intake and sources of long-chain n-3 PUFAs in German adults. *Eur. J. Clin. Nutr.*, **60,** 810-812.

[25] Caligiuri SP, Love K, Winter T, Gauthier J, Taylor CG, Blydt-Hansen T, Zahradka P, & Aukema HM (2013) Dietary linoleic acid and alpha-linolenic acid differentially affect renal oxylipins and phospholipid fatty acids in diet-induced obese rats. *J. Nutr.*, **143,** 1421-1431.

[26] Mann NJ, Johnson LG, Warrick GE, & Sinclair AJ (1995) The arachidonic acid content of the Australian diet is lower than previously estimated. *J. Nutr.*, **125,** 2528-2535.

[27] Strobel C, Jahreis G, & Kuhnt K (2012) Survey of *n*-3 and *n*-6 polyunsaturated fatty acids in fish and fish products. *Lipids Health Dis.*, **11,** 144.

[28] Food and Agriculture Organization of the United Nations (2012) FAO yearbook. Fishery and Aquaculture Statistics. 2010.

[29] Tacon AGJ, Hassan MR, & Metian M (2011) Demand and supply of feed ingredients for farmed fish and crustaceans. Trends and prospects. *FAO Fish. Tech. Pap.*, **564,** 1-87.

[30] Tacon AGJ & Metian M (2008) Global overview on the use of fish meal and fish oil in industrially compounded aquafeeds: Trends and future prospects. *Aquaculture*, **285,** 146-158.

[31] Reynolds JD, Dulvy NK, Goodwin NB, & Hutchings JA (2005) Biology of extinction risk in marine fishes. *Proc. Biol. Sci.*, **272,** 2337-2344.

[32] Amadee & Company, Inc. (2013) Omega-3 marine and plant products: Global markets, competitors and opportunities–2012-2018 analysis and forecasts.

[33] Jacobs MN, Covaci A, Gheorghe A, & Schepens P (2004) Time trend investigation of PCBs, PBDEs, and organochlorine pesticides in selected *n*-3 polyunsaturated fatty acid rich dietary fish oil and vegetable oil supplements; nutritional relevance for human essential *n*-3 fatty acid requirements. *J. Agric. Food Chem.*, **52,** 1780-1788.

[34] Braune BM, Outridge PM, Fisk AT, Muir DC, Helm PA, Hobbs K, Hoekstra PF, Kuzyk ZA, Kwan M, Letcher RJ, Lockhart WL, Norstrom RJ, Stern GA, & Stirling I (2005) Persistent organic pollutants and mercury in marine biota of the Canadian Arctic: an overview of spatial and temporal trends. *Sci. Total Environ.*, **351-352,** 4-56.

[35] Ortiz X, Carabellido L, Marti M, Marti R, Tomas X, & Diaz-Ferrero J (2011) Elimination of persistent organic pollutants from fish oil with solid adsorbents. *Chemosphere*, **82,** 1301-1307.

[36] Haslam RP, Ruiz-Lopez N, Eastmond P, Moloney M, Sayanova O, & Napier JA (2013) The modification of plant oil composition via metabolic engineering – better nutrition by design. *Plant Biotechnol. J.*, **11,** 157-168.

[37] Ren LJ, Ji XJ, Huang H, Qu L, Feng Y, Tong QQ, & Ouyang PK (2010) Development of a stepwise aeration control strategy for efficient docosahexaenoic acid production by *Schizochytrium* sp. *Appl. Microbiol. Biotechnol.*, **87,** 1649-1656.

[38] Taoka Y, Nagano N, Okita Y, Izumida H, Sugimoto S, & Hayashi M (2011) Effect of Tween 80 on the growth, lipid accumulation and fatty acid composition of *Thraustochytrium aureum* ATCC 34304. *J. Biosci. Bioeng.*, **111,** 420-424.

[39] Lee Chang KJ, Dunstan GA, Abell GC, Clementson LA, Blackburn SI, Nichols PD, & Koutoulis A (2012) Biodiscovery of new Australian thraustochytrids for production of biodiesel and long-chain omega-3 oils. *Appl. Microbiol. Biotechnol.*, **93,** 2215-2231.

[40] de Swaaf ME, Pronk JT, & Sijtsma L (2003) Fed-batch cultivation of the docosahexaenoic-acid-producing marine alga *Crypthecodinium cohnii* on ethanol. *Appl. Microbiol. Biotechnol.*, **61,** 40-43.

[41] Pal D, Khozin-Goldberg I, Didi-Cohen S, Solovchenko A, Batushansky A, Kaye Y, Sikron N, Samani T, Fait A, & Boussiba S (2013) Growth, lipid production and metabolic adjustments in the euryhaline eustigmatophyte *Nannochloropsis oceanica* CCALA 804 in response to osmotic downshift. *Appl. Microbiol. Biotechnol.*, **97,** 8291-8306.

[42] Perez-Lopez P, Gonzalez-Garcia S, Allewaert C, Verween A, Murray P, Feijoo G, & Moreira MT (2014) Environmental evaluation of eicosapentaenoic acid production by *Phaeodactylum tricornutum*. *Sci. Total Environ.*, **466-467,** 991-1002.

[43] Martins DA, Custodio L, Barreira L, Pereira H, Ben-Hamadou R, Varela J, & Abu-Salah KM (2013) Alternative sources of *n*-3 long-chain polyunsaturated fatty acids in marine microalgae. *Mar. Drugs*, **11,** 2259-2281.

[44] Adarme-Vega TC, Lim DK, Timmins M, Vernen F, Li Y, & Schenk PM (2012) Microalgal biofactories: a promising approach towards sustainable omega-3 fatty acid production. *Microb. Cell Fact.*, **11,** 96.

[45] Jenke-Kodama H, Sandmann A, Müller R, & Dittmann E (2005) Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.*, **22,** 2027-2039.

[46] Smith S & Tsai SC (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat. Prod. Rep.*, **24,** 1041-1072.

[47] White SW, Zheng J, Zhang YM, & Rock (2005) The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.*, **74,** 791-831.

[48] Cronan JE, Jr. & Waldrop GL (2002) Multi-subunit acetyl-CoA carboxylases. *Prog. Lipid Res.*, **41,** 407-435.

[49] Prescott DJ & Vagelos PR (1972) Acyl carrier protein. *Adv. Enzymol. Relat Areas Mol. Biol.*, **36,** 269-311.

[50] Lambalot RH, Gehring AM, Flugel RS, Zuber P, LaCelle M, Marahiel MA, Reid R, Khosla C, & Walsh CT (1996) A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem. Biol.*, **3,** 923-936.

[51] Maier T, Leibundgut M, & Ban N (2008) The crystal structure of a mammalian fatty acid synthase. *Science*, **321,** 1315-1322.

[52] Cronan JE & Thomas J (2009) Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *Methods Enzymol.*, **459,** 395-433.

[53] Zhang YM & Rock CO (2008) Thematic review series: Glycerolipids. Acyltransferases in bacterial glycerophospholipid synthesis. *J. Lipid Res.*, **49,** 1867-1874.

[54] Sayanova OV & Napier JA (2004) Eicosapentaenoic acid: biosynthetic routes and the potential for synthesis in transgenic plants. *Phytochemistry*, **65,** 147-158.

[55] Warude D, Joshi K, & Harsulkar A (2006) Polyunsaturated fatty acids: biotechnology. *Crit Rev. Biotechnol.*, **26,** 83-93.

[56] Jiao J & Zhang Y (2013) Transgenic biosynthesis of polyunsaturated fatty acids: a sustainable biochemical engineering approach for making essential fatty acids in plants and animals. *Chem. Rev.*, **113,** 3799-3814.

[57] Leonard AE, Kelder B, Bobik EG, Chuang LT, Lewis CJ, Kopchick JJ, Mukerji P, & Huang YS (2002) Identification and expression of mammalian long-chain PUFA elongation enzymes. *Lipids*, **37,** 733-740.

[58] Sprecher H, Luthria DL, Mohammed BS, & Baykousheva SP (1995) Reevaluation of the pathways for the biosynthesis of polyunsaturated fatty acids. *J. Lipid Res.*, **36,** 2471-2477.

[59] Sprecher H (2000) Metabolism of highly unsaturated *n*-3 and *n*-6 fatty acids. *Biochim. Biophys. Acta*, **1486,** 219-231.

[60] de Antueno RJ, Knickle LC, Smith H, Elliot ML, Allen SJ, Nwaka S, & Winther MD (2001) Activity of human Delta5 and Delta6 desaturases on multiple *n*-3 and *n*-6 polyunsaturated fatty acids. *FEBS Lett.*, **509,** 77-80.

[61] D'Andrea S, Guillou H, Jan S, Catheline D, Thibault JN, Bouriel M, Rioux V, & Legrand P (2002) The same rat Delta6-desaturase not only acts on 18- but also on 24-carbon fatty acids in very-long-chain polyunsaturated fatty acid biosynthesis. *Biochem. J.*, **364,** 49-55.

[62] Qi B, Beaudoin F, Fraser T, Stobart AK, Napier JA, & Lazarus CM (2002) Identification of a cDNA encoding a novel C18-Delta(9) polyunsaturated fatty acid-specific elongating activity from the docosahexaenoic acid (DHA)-producing microalga, *Isochrysis galbana*. *FEBS Lett.*, **510,** 159-165.

[63] Qiu X, Hong H, & MacKenzie SL (2001) Identification of a Delta 4 fatty acid desaturase from *Thraustochytrium* sp. involved in the biosynthesis of docosahexanoic acid by heterologous expression in *Saccharomyces cerevisiae* and *Brassica juncea*. *J. Biol. Chem.*, **276,** 31561-31566.

[64] Tonon T, Harvey D, Larson TR, & Graham IA (2003) Identification of a very long chain polyunsaturated fatty acid Delta4-desaturase from the microalga *Pavlova lutheri*. *FEBS Lett.*, **553,** 440-444.

[65] Pereira SL, Leonard AE, Huang YS, Chuang LT, & Mukerji P (2004) Identification of two novel microalgal enzymes involved in the conversion of the omega3-fatty acid, eicosapentaenoic acid, into docosahexaenoic acid. *Biochem. J.*, **384,** 357-366.

[66] Wallis JG & Browse J (1999) The Delta8-desaturase of *Euglena gracilis*: an alternate pathway for synthesis of 20-carbon polyunsaturated fatty acids. *Arch. Biochem. Biophys.*, **365,** 307-316.

[67] Meyer A, Cirpus P, Ott C, Schlecker R, Zahringer U, & Heinz E (2003) Biosynthesis of docosahexaenoic acid in *Euglena gracilis*: biochemical and molecular evidence for the involvement of a Delta4-fatty acyl group desaturase. *Biochemistry*, **42,** 9779-9788.

[68] Pereira SL, Huang YS, Bobik EG, Kinney AJ, Stecca KL, Packer JC, & Mukerji P (2004) A novel omega3-fatty acid desaturase involved in the biosynthesis of eicosapentaenoic acid. *Biochem. J.*, **378,** 665-671.

[69] Xue Z, He H, Hollerbach D, Macool DJ, Yadav NS, Zhang H, Szostek B, & Zhu Q (2013) Identification and characterization of new Delta-17 fatty acid desaturases. *Appl. Microbiol. Biotechnol.*, **97,** 1973-1985.

[70] Tripodi KE, Buttigliero LV, Altabe SG, & Uttaro AD (2006) Functional characterization of front-end desaturases from trypanosomatids depicts the first polyunsaturated fatty acid biosynthetic pathway from a parasitic protozoan. *FEBS J.*, **273,** 271-280.

[71] Li Y, Monroig O, Zhang L, Wang S, Zheng X, Dick JR, You C, & Tocher DR (2010) Vertebrate fatty acyl desaturase with Delta4 activity. *Proc. Natl. Acad. Sci. U S A*, **107,** 16840-16845.

[72] Burdge G (2004) Alpha-linolenic acid metabolism in men and women: nutritional and biological implications. *Curr. Opin. Clin. Nutr. Metab Care*, **7,** 137-144.

[73] Nettleton JA (1991) Omega-3 fatty acids: comparison of plant and seafood sources in human nutrition. *J. Am. Diet. Assoc.*, **91,** 331-337.

[74] DeLany JP, Windhauser MM, Champagne CM, & Bray GA (2000) Differential oxidation of individual dietary fatty acids in humans. *Am. J. Clin. Nutr.*, **72,** 905-911.

[75] Burdge GC & Calder PC (2005) Conversion of alpha-linolenic acid to longer-chain polyunsaturated fatty acids in human adults. *Reprod. Nutr. Dev.*, **45,** 581-597.

[76] Yazawa K (1996) Production of eicosapentaenoic acid from marine bacteria. *Lipids*, **31 Suppl,** S297-S300.

[77] 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.

[78] 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.

[79] Morita N, Tanaka M, & Okuyama H (2000) Biosynthesis of fatty acids in the docosahexaenoic acid-producing bacterium *Moritella marina* strain MP-1. *Biochem. Soc. Trans.*, **28,** 943-945.

[80] Hauvermale A, Kuner J, Rosenzweig B, Guerra D, Diltz S, & Metz JG (2006) Fatty acid production in *Schizochytrium* sp.: Involvement of a polyunsaturated fatty acid synthase and a type I fatty acid synthase. *Lipids*, **41,** 739-747.

[81] Smith JL & Sherman DH (2008) Biochemistry. An enzyme assembly line. *Science*, **321,** 1304-1305.

[82] 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.*, **580,** 4423-4429.

[83] Sugihara S, Orikasa Y, & Okuyama H (2008) An EntD-like phosphopantetheinyl transferase gene from *Photobacterium profundum* SS9 complements *pfa* genes of *Moritella marina* strain MP-1 involved in biosynthesis of docosahexaenoic acid. *Biotechnol. Lett.*, **30,** 411-414.

[84] Kaulmann U & Hertweck C (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew. Chem. Int. Ed.*, **41,** 1866-1869.

[85] Napier JA (2002) Plumbing the depths of PUFA biosynthesis: a novel polyketide synthase-like pathway from marine organisms. *Trends Plant Sci.*, **7,** 51-54.

[86] Wallis JG, Watts JL, & Browse J (2002) Polyunsaturated fatty acid synthesis: what will they think of next? *Trends Biochem. Sci.*, **27,** 467-473.

[87] Metz JG, Kuner J, Rosenzweig B, Lippmeier JC, Roessler P, & Zirkle R (2009) Biochemical characterization of polyunsaturated fatty acid synthesis in *Schizochytrium*: release of the products as free fatty acids. *Plant Physiol Biochem.*, **47,** 472-478.

[88] Rodriguez-Guilbe M, Oyola-Robles D, Schreiter ER, & Baerga-Ortiz A (2013) Structure, activity, and substrate selectivity of the Orf6 thioesterase from *Photobacterium profundum*. *J. Biol. Chem.*, **288,** 10841-10848.

[89] Hopwood DA & Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.*, **24,** 37-66.

[90] Hopwood DA (1997) Genetic contributions to understanding polyketide synthases. *Chem. Rev.*, **97,** 2465-2498.

[91] Weissman KJ (2005) Polyketide synthases: mechanisms and models. *Ernst. Schering. Res. Found. Workshop*, 43-78.

[92] Trujillo U, Vazquez-Rosa E, Oyola-Robles D, Stagg LJ, Vassallo DA, Vega IE, Arold ST, & Baerga-Ortiz A (2013) Solution structure of the tandem acyl carrier protein domains from a polyunsaturated fatty acid synthase reveals beads-on-a-string configuration. *PLOS One.*, **8,** e57859.

[93] Hertweck C, Luzhetskyy A, Rebets Y, & Bechthold A (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.*, **24,** 162-190.

[94] McDaniel R, Ebert-Khosla S, Hopwood DA, & Khosla C (1993) Engineered biosynthesis of novel polyketides. *Science*, **262,** 1546-1550.

[95] Tang Y, Tsai SC, & Khosla C (2003) Polyketide chain length control by chain length factor. *J. Am. Chem. Soc.*, **125,** 12708-12709.

[96] Keatinge-Clay AT, Maltby DA, Medzihradszky KF, Khosla C, & Stroud RM (2004) An antibiotic factory caught in action. *Nat. Struct. Mol. Biol.*, **11,** 888-893.

[97] Oyola-Robles D, Gay DC, Trujillo U, Sanchez-Pares JM, Bermudez ML, Rivera-Diaz M, Carballeira NM, & Baerga-Ortiz A (2013) Identification of novel protein domains required for the expression of an active dehydratase fragment from a polyunsaturated fatty acid synthase. *Protein Sci.*, **22,** 954-963.

[98] Butcher RA, Schroeder FC, Fischbach MA, Straight PD, Kolter R, Walsh CT, & Clardy J (2007) The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U S A*, **104,** 1506-1509.

[99] Bumpus SB, Magarvey NA, Kelleher NL, Walsh CT, & Calderone CT (2008) Polyunsaturated fatty-acid-like *trans*-enoyl reductases utilized in polyketide biosynthesis. *J. Am. Chem. Soc.*, **130,** 11614-11616.

[100] Simunovic V, Zapp J, Rachid S, Krug D, Meiser P, & Müller R (2006) Myxovirescin A biosynthesis is directed by hybrid polyketide synthases/nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl-CoA synthases, and *trans*-acting acyltransferases. *Chembiochem.*, **7,** 1206-1220.

[101] Okuyama H, Orikasa Y, Nishida T, Watanabe K, & Morita N (2007) Bacterial genes responsible for the biosynthesis of eicosapentaenoic and docosahexaenoic acids and their heterologous expression. *Appl. Environ. Microbiol.*, **73,** 665-670.

[102] Pareek CS, Smoczynski R, & Tretyn A (2011) Sequencing technologies and genome sequencing. *J. Appl. Genet.*, **52,** 413-435.

[103] Shulse CN & Allen EE (2011) Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. *PLOS One.*, **6,** e20146.

[104] Shimkets L, Dworkin M, & Reichenbach H (2006) The Myxobacteria. In *The Prokaryotes* (Dworkin M, ed), pp. 31-115. Springer, Berlin.

[105] Dawid W (2000) Biology and global distribution of myxobacteria in soils. *FEMS Microbiol. Rev.*, **24,** 403-427.

[106] Kaiser D (2001) Building a multicellular organism. *Annu. Rev. Genet.*, **35,** 103-123.

[107] Yu R & Kaiser D (2007) Gliding motility and polarized slime secretion. *Mol. Microbiol.*, **63,** 454-467.

[108] Zhang H, Vaksman Z, Litwin DB, Shi P, Kaplan HB, & Igoshin OA (2012) The mechanistic basis of *Myxococcus xanthus* rippling behavior and its physiological role during predation. *PLOS Comput. Biol.*, **8,** e1002715.

[109] Weissman KJ & Müller R (2010) Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Nat. Prod. Rep.*, **27,** 1276-1295.

[110] Gerth K, Bedorf N, Irschik H, Hofle G, & Reichenbach H (1994) The soraphens: a family of novel antifungal compounds from *Sorangium cellulosum* (Myxobacteria). I. Soraphen A1 alpha: fermentation, isolation, biological properties. *J. Antibiot. (Tokyo)*, **47,** 23-31.

[111] Shen Y, Volrath SL, Weatherly SC, Elich TD, & Tong L (2004) A mechanism for the potent inhibition of eukaryotic acetyl-coenzyme A carboxylase by soraphen A, a macrocyclic polyketide natural product. *Mol. Cell*, **16,** 881-891.

[112] Sasse F, Steinmetz H, Schupp T, Petersen F, Memmert K, Hofmann H, Heusser C, Brinkmann V, von MP, Hofle G, & Reichenbach H (2002) Argyrins, immunosuppressive cyclic peptides from myxobacteria. I. Production, isolation, physico-chemical and biological properties. *J. Antibiot. (Tokyo)*, **55,** 543-551.

[113] Nickeleit I, Zender S, Sasse F, Geffers R, Brandes G, Sorensen I, Steinmetz H, Kubicka S, Carlomagno T, Menche D, Gutgemann I, Buer J, Gossler A, Manns MP, Kalesse M, Frank R, & Malek NP (2008) Argyrin a reveals a critical role for the tumor suppressor protein p27(kip1) in mediating antitumor activities in response to proteasome inhibition. *Cancer Cell*, **14,** 23-35.

[114] Weissman KJ & Müller R (2009) A brief tour of myxobacterial secondary metabolism. *Bioorganic & Medicinal Chemistry*, **17,** 2121-2136.

[115] Bollag DM, McQueney PA, Zhu J, Hensens O, Koupal L, Liesch J, Goetz M, Lazarides E, & Woods CM (1995) Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. *Cancer Res.*, **55,** 2325-2333.

[116] Gerth K, Bedorf N, Hofle G, Irschik H, & Reichenbach H (1996) Epothilons A and B: antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties. *J. Antibiot. (Tokyo)*, **49,** 560-563.

[117] Irschik H, Jansen R, Gerth K, Hofle G, & Reichenbach H (1995) Disorazol A, an efficient inhibitor of eukaryotic organisms isolated from myxobacteria. *J. Antibiot. (Tokyo)*, **48,** 31-35.

[118] Elnakady YA, Sasse F, Lunsdorf H, & Reichenbach H (2004) Disorazol A1, a highly effective antimitotic agent acting on tubulin polymerization and inducing apoptosis in mammalian cells. *Biochem. Pharmacol.*, **67,** 927-935.

[119] Bode HB, Irschik H, Wenzel SC, Reichenbach H, Müller R, & Hofle G (2003) The leupyrrins: a structurally unique family of secondary metabolites from the myxobacterium *Sorangium cellulosum*. *J. Nat. Prod.*, **66,** 1203-1206.

[120] Bode HB, Wenzel SC, Irschik H, Hofle G, & Müller R (2004) Unusual biosynthesis of leupyrrins in the myxobacterium *Sorangium cellulosum*. *Angew. Chem. Int. Ed.*, **43,** 4163-4167.

[121] Kopp M, Irschik H, Gemperlein K, Buntin K, Meiser P, Weissman KJ, Bode HB, & Müller R (2011) Insights into the complex biosynthesis of the leupyrrins in *Sorangium cellulosum* So ce690. *Mol. Biosyst.*, **7,** 1549-1563.

[122] Herrmann M, Böhlendorf B, Irschik H, Reichenbach H, & Höfle G (1998) Maracin und Maracen: Neuartige Ethinyl-vinylether- bzw. a-Chlordivinylether-Antibiotica aus *Sorangium cellulosum* mit spezifischer Wirkung gegen Mycobakterien. *Angew. Chem.*, **110,** 1313-1315.

[123] Garcia R, Pistorius D, Stadler M, & Müller R (2011) Fatty acid-related phylogeny of myxobacteria as an approach to discover polyunsaturated omega-3/6 Fatty acids. *J. Bacteriol.*, **193,** 1930-1942.

[124] Iizuka T, Jojima Y, Fudou R, Hiraishi A, Ahn JW, & Yamanaka S (2003) *Plesiocystis pacifica* gen. nov., sp. nov., a marine myxobacterium that contains dihydrogenated menaquinone, isolated from the Pacific coasts of Japan. *Int. J. Syst. Evol. Microbiol.*, **53,** 189-195.

[125] Iizuka T, Jojima Y, Fudou R, Tokura M, Hiraishi A, & Yamanaka S (2003) *Enhygromyxa salina* gen. nov., sp. nov., a slightly halophilic myxobacterium isolated from the coastal areas of Japan. *Syst. Appl. Microbiol.*, **26,** 189-196.

[126] Iizuka T, Jojima Y, Hayakawa A, Fujii T, Yamanaka S, & Fudou R (2013) *Pseudenhygromyxa salsuginis* gen. nov., sp. nov., a myxobacterium isolated from an estuarine marsh. *Int. J. Syst. Evol. Microbiol.*, **63,** 1360-1369.

[127] Garcia RO, Reichenbach H, Ring MW, & Müller R (2009) *Phaselicystis flava* gen. nov., sp. nov., an arachidonic acid-containing soil myxobacterium, and the description of *Phaselicystidaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.*, **59,** 1524-1530.

[128] Mohr KI, Garcia RO, Gerth K, Irschik H, & Müller R (2012) *Sandaracinus amylolyticus* gen. nov., sp. nov., a starch-degrading soil myxobacterium, and description of *Sandaracinaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.*, **62,** 1191-1198.

[129] Stadler M, Roemer E, Müller R, Garcia RO, Pistorius D, & Brachmann A (2010) Production of omega-3 fatty acids by myxobacteria. *Int. Pat.*, WO 2010/063451.

[130] Goldman BS, Nierman WC, Kaiser D, Slater SC, Durkin AS, Eisen JA, Ronning CM, Barbazuk WB, Blanchard M, Field C, Halling C, Hinkle G, Iartchuk O, Kim HS, Mackenzie C, Madupu R, Miller N, Shvartsbeyn A, Sullivan SA, Vaudin M, Wiegand R, & Kaplan HB (2006) Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Natl. Acad. Sci. U S A*, **103,** 15200-15205.

[131] Schneiker S, Perlova O, Kaiser O, Gerth K, Alici A, Altmeyer MO, Bartels D, Bekel T, Beyer S, Bode E, Bode HB, Bolten CJ, Choudhuri JV, Doss S, Elnakady YA, Frank B, Gaigalat L, Goesmann A, Groeger C, Gross F, Jelsbak L, Jelsbak L, Kalinowski J, Kegler C, Knauber T, Konietzny S, Kopp M, Krause L, Krug D, Linke B, Mahmud T, Martinez-Arias R, McHardy AC, Merai M, Meyer F, Mormann S, Munoz-Dorado J, Perez J, Pradella S, Rachid S, Raddatz G, Rosenau F, Ruckert C, Sasse F, Scharfe M, Schuster SC, Suen G, Treuner-Lange A, Velicer GJ, Vorholter FJ, Weissman KJ, Welch RD, Wenzel SC, Whitworth DE, Wilhelm S, Wittmann C, Blocker H, Puhler A, & Müller R (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat. Biotechnol.*, **25,** 1281-1289.

[132] Nikolaki S & Tsiamis G (2013) Microbial diversity in the era of omic technologies. *Biomed. Res. Int.*, **2013,** 958719.

[133] Challis GL (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology*, **154,** 1555-1569.

[134] Corre C & Challis GL (2007) Heavy tools for genome mining. *Chem. Biol.*, **14,** 7-9.

[135] Pistorius D & Müller R (2012) Discovery of the rhizopodin biosynthetic gene cluster in *Stigmatella aurantiaca* Sg a15 by genome mining. *Chembiochem.*, **13,** 416-426.

[136] Cortina NS, Krug D, Plaza A, Revermann O, & Müller R (2012) Myxoprincomide: a natural product from *Myxococcus xanthus* discovered by comprehensive analysis of the secondary metabolome. *Angew. Chem. Int. Ed.*, **51,** 811-816.

[137] Gross F, Luniak N, Perlova O, Gaitatzis N, Jenke-Kodama H, Gerth K, Gottschalk D, Dittmann E, & Müller R (2006) Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. *Arch. Microbiol.*, **185,** 28-38.

[138] Mutka SC, Carney JR, Liu Y, & Kennedy J (2006) Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry*, **45,** 1321-1330.

[139] Wenzel SC & Müller R (2005) Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways. *Curr. Opin. Biotechnol.*, **16,** 594-606.

[140] Ongley SE, Bian X, Neilan BA, & Müller R (2013) Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat. Prod. Rep.*, **30,** 1121-1138.

[141] Noskov VN, Kouprina N, Leem SH, Ouspenski I, Barrett JC, & Larionov V (2003) A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC. Genomics*, **4,** 16.

[142] Kim JH, Feng Z, Bauer JD, Kallifidas D, Calle PY, & Brady SF (2010) Cloning large natural product gene clusters from the environment: piecing environmental DNA gene clusters back together with TAR. *Biopolymers*, **93,** 833-844.

[143] Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Müller R, Stewart AF, & Zhang Y (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.*, **30,** 440-446.

[144] Zhang Y, Muyrers JP, Testa G, & Stewart AF (2000) DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.*, **18,** 1314-1317.

[145] Kieser T, Bibb M, Buttner MJ, Chater KF, & Hopwood DA (2000) *Practical Streptomyces Genetics*. The John Innes Foundation, Norwich, England.

[146] Poblete-Castro I, Becker J, Dohnt K, dos Santos VM, & Wittmann C (2012) Industrial biotechnology of *Pseudomonas putida* and related species. *Appl. Microbiol. Biotechnol.*, **93,** 2279-2290.

[147] Anthony JR, Anthony LC, Nowroozi F, Kwon G, Newman JD, & Keasling JD (2009) Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metab Eng*, **11,** 13-19.

[148] Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MC, Withers ST, Shiba Y, Sarpong R, & Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440,** 940-943.

[149] Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Jiang H, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievense J, Covello PS, Keasling JD, Reiling KK, Renninger NS, & Newman JD (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, **496,** 528-532.

[150] Kodumal SJ, Patel KG, Reid R, Menzella HG, Welch M, & Santi DV (2004) Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster. *P. Natl. Acad. Sci. U S A*, **101,** 15573-15578.

[151] Menzella HG, Reisinger SJ, Welch M, Kealey JT, Kennedy J, Reid R, Tran CQ, & Santi DV (2006) Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster. *J. Ind. Microbiol. Biotechnol.*, **33,** 22-28.

[152] Oßwald C, Zipf G, Schmidt G, Maier J, Bernauer HS, Müller R, & Wenzel SC (2012) Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.*, Epub ahead of print.

[153] Gao X, Wang P, & Tang Y (2010) Engineered polyketide biosynthesis and biocatalysis in *Escherichia coli*. *Appl. Microbiol. Biotechnol.*, **88,** 1233-1242.

[154] Jiang H, Zirkle R, Metz JG, Braun L, Richter L, Van Lanen SG, & Shen B (2008) The role of tandem acyl carrier protein domains in polyunsaturated fatty acid biosynthesis. *J. Am. Chem. Soc.*, **130,** 6336-6337.

[155] Gemperlein K, Rachid S, Garcia R, Wenzel SC, & Müller R (2014) Polyunsaturated fatty acid biosynthesis in myxobacteria: Different PUFA synthases and their product diversity. *Chem. Sci.*, **5,** 1733-1741.

[156] Radford SE, Laue ED, Perham RN, Martin SR, & Appella E (1989) Conformational flexibility and folding of synthetic peptides representing an interdomain segment of polypeptide chain in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Biol. Chem.*, **264,** 767-775.

[157] Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215,** 403-410.

[158] Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, & Finn RD (2012) The Pfam protein families database. *Nucleic Acids Res.*, **40,** D290-D301.

[159] Wisniewski-Dye F, Borziak K, Khalsa-Moyers G, Alexandre G, Sukharnikov LO, Wuichet K, Hurst GB, McDonald WH, Robertson JS, Barbe V, Calteau A, Rouy Z, Mangenot S, Prigent-Combaret C, Normand P, Boyer M, Siguier P, Dessaux Y, Elmerich C, Condemine G, Krishnen G, Kennedy I, Paterson AH, Gonzalez V, Mavingui P, & Zhulin IB (2011) *Azospirillum* genomes reveal transition of bacteria from aquatic to terrestrial environments. *PLOS Genet.*, **7,** e1002430.

[160] Tice H, Mayilraj S, Sims D, Lapidus A, Nolan M, Lucas S, Glavina Del RT, Copeland A, Cheng JF, Meincke L, Bruce D, Goodwin L, Pitluck S, Ivanova N, Mavromatis K, Ovchinnikova G, Pati A, Chen A, Palaniappan K, Land M, Hauser L, Chang YJ, Jeffries CD, Detter JC, Brettin T, Rohde M, Goker M, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Kyrpides NC, Klenk HP, & Chen F (2010) Complete genome sequence of *Nakamurella multipartita* type strain (Y-104). *Stand. Genomic. Sci.*, **2,** 168-175.

[161] Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, & Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, **417,** 141-147.

[162] Cruz-Morales P, Vijgenboom E, Iruegas-Bocardo F, Girard G, Yanez-Guerra LA, Ramos-Aboites HE, Pernodet JL, Anne J, van Wezel GP, & Barona-Gomez F (2013) The genome sequence of *Streptomyces lividans* 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island. *Genome Biol. Evol.*, **5,** 1165-1175.

[163] Weiner RM, Taylor LE, Henrissat B, Hauser L, Land M, Coutinho PM, Rancurel C, Saunders EH, Longmire AG, Zhang H, Bayer EA, Gilbert HJ, Larimer F, Zhulin IB, Ekborg NA, Lamed R, Richardson PM, Borovok I, & Hutcheson S (2008) Complete genome sequence of the complex carbohydrate-degrading marine bacterium, *Saccharophagus degradans* strain 2-40 T. *PLOS Genet.*, **4,** e1000087.

[164] Singer SJ & Nicolson GL (1972) The fluid mosaic model of the structure of cell membranes. *Science*, **175,** 720-731.

[165] Hazel JR & Williams EE (1990) The role of alterations in membrane lipid composition in enabling physiological adaptation of organisms to their physical environment. *Prog. Lipid Res.*, **29,** 167-227.

[166] Delong EF & Yayanos AA (1986) Biochemical function and ecological significance of novel bacterial lipids in deep-sea procaryotes. *Appl. Environ. Microbiol.*, **51,** 730-737.

[167] Yano Y, Nakayama A, & Yoshida K (1997) Distribution of polyunsaturated fatty acids in bacteria present in intestines of deep-sea fish and shallow-sea poikilothermic animals. *Appl. Environ. Microbiol.*, **63,** 2572-2577.

[168] Kato C & Nogi Y (2001) Correlation between phylogenetic structure and function: examples from deep-sea *Shewanella*. *FEMS Microbiol. Ecol.*, **35,** 223-230.

[169] Sukovich DJ, Seffernick JL, Richman JE, Hunt KA, Gralnick JA, & Wackett LP (2010) Structure, function, and insights into the biosynthesis of a head-to-head hydrocarbon in *Shewanella oneidensis* strain MR-1. *Appl. Environ. Microbiol.*, **76,** 3842-3849.

[170] Holte LL, Peter SA, Sinnwell TM, & Gawrisch K (1995) 2H nuclear magnetic resonance order parameter profiles suggest a change of molecular shape for phosphatidylcholines containing a polyunsaturated acyl chain. *Biophys. J.*, **68,** 2396-2403.

[171] Allen EE, Facciotti D, & Bartlett DH (1999) Monounsaturated but not polyunsaturated fatty acids are required for growth of the deep-sea bacterium *Photobacterium profundum* SS9 at high pressure and low temperature. *Appl. Environ. Microbiol.*, **65,** 1710-1720.

[172] Niki E (2009) Lipid peroxidation: physiological levels and dual biological effects. *Free Radic. Biol. Med.*, **47,** 469-484.

[173] Yazu K, Yamamoto Y, Ukegawa K, & Niki E (1996) Mechanism of lower oxidizability of eicosapentaenoate than linoleate in aqueous micelles. *Lipids*, **31,** 337-340.

[174] Nishida T, Orikasa Y, Ito 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.

[175] Nishida T, Orikasa Y, Watanabe K, & Okuyama H (2006) The cell membrane-shielding function of eicosapentaenoic acid for *Escherichia coli* against exogenously added hydrogen peroxide. *FEBS Lett.*, **580,** 6690-6694.

[176] Nishida T, Morita N, Yano Y, Orikasa Y, & Okuyama H (2007) The antioxidative function of eicosapentaenoic acid in a marine bacterium, *Shewanella marinintestina* IK-1. *FEBS Lett.*, **581,** 4212-4216.

[177] Okuyama H, Orikasa Y, & Nishida T (2008) Significance of antioxidative functions of eicosapentaenoic and docosahexaenoic acids in marine microorganisms. *Appl. Environ. Microbiol.*, **74,** 570-574.

[178] Lipinski B (2011) Hydroxyl radical and its scavengers in health and disease. *Oxid. Med. Cell Longev.*, **2011,** 809696.

[179] Masschelein J, Mattheus W, Gao LJ, Moons P, Van HR, Uytterhoeven B, Lamberigts C, Lescrinier E, Rozenski J, Herdewijn P, Aertsen A, Michiels C, & Lavigne R (2013) A PKS/NRPS/FAS hybrid gene cluster from *Serratia plymuthica* RVH1 encoding the biosynthesis of three broad spectrum, zeamine-related antibiotics. *PLOS One.*, **8,** e54143.

[180] Orikasa Y, Tanaka M, Sugihara S, Hori R, Nishida T, Ueno A, Morita N, Yano Y, Yamamoto K, Shibahara A, Hayashi H, Yamada Y, Yamada A, Yu RK, Watanabe K, & Okuyama H (2009) *pfaB* products determine the molecular species produced in bacterial polyunsaturated fatty acid biosynthesis. *FEMS Microbiol. Lett.*, **295,** 170-176.

[181] Weissman KJ, Hong H, Oliynyk M, Siskos AP, & Leadlay PF (2004) Identification of a phosphopantetheinyl transferase for erythromycin biosynthesis in *Saccharopolyspora erythraea*. *Chembiochem.*, **5,** 116-125.

[182] Reuter K, Mofid MR, Marahiel MA, & Ficner R (1999) Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. *EMBO J.*, **18,** 6823-6831.

[183] Quadri LE, Weinreb PH, Lei M, Nakano MM, Zuber P, & Walsh CT (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry*, **37,** 1585-1595.

[184] Gaitatzis N, Hans A, Müller R, & Beyer S (2001) The mtaA gene of the myxothiazol biosynthetic gene cluster from *Stigmatella aurantiaca* DW4/3-1 encodes a phosphopantetheinyl transferase that activates polyketide synthases and polypeptide synthetases. *J. Biochem.*, **129,** 119-124.

[185] Wright LF & Hopwood DA (1976) Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.*, **96,** 289-297.

[186] Weber W, Zahner H, Siebers J, Schroder K, & Zeeck A (1979) [Metabolic products of microorganisms. 175. Tetracenomycin C (author's transl)]. *Arch. Microbiol.*, **121,** 111-116.

[187] Davis NK & Chater KF (1990) Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Mol. Microbiol.*, **4,** 1679-1691.

[188] Pandey RC, Toussaint MW, Stroshane RM, Kalita CC, Aszalos AA, Garretson AL, Wei TT, Byrne KM, Geoghegan RF, Jr., & White RJ (1981) Fredericamycin A, a new antitumor antibiotic. I. Production, isolation and physicochemical properties. *J. Antibiot. (Tokyo)*, **34,** 1389-1401.

[189] Burson KK & Khosla C (2000) Dissecting the chain length specificity in bacterial aromatic polyketide synthases using chimeric genes. *Tetrahedron*, **56,** 9401-9408.

[190] Zhang YM & Rock CO (2008) Thematic review series: Glycerolipids - Acyltransferases in bacterial glycerophospholipid synthesis. *J. Lipid Res.*, **49,** 1867-1874.

[191] Lewin TM, Wang P, & Coleman RA (1999) Analysis of amino acid motifs diagnostic for the sn-glycerol-3-phosphate acyltransferase reaction. *Biochemistry*, **38,** 5764-5771.

[192] Yamashita A, Nakanishi H, Suzuki H, Kamata R, Tanaka K, Waku K, & Sugiura T (2007) Topology of acyltransferase motifs and substrate specificity and accessibility in 1-acyl-sn-glycero-3-phosphate acyltransferase 1. *Biochim. Biophys. Acta*, **1771,** 1202-1215.

[193] Fu J, Wenzel SC, Perlova O, Wang J, Gross F, Tang Z, Yin Y, Stewart AF, Müller R, & Zhang Y (2008) Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res.*, **36,** e113.

[194] Perlova O, Fu J, Kuhlmann S, Krug D, Stewart AF, Zhang Y, & Müller R (2006) Reconstitution of the myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*. *Appl. Environ. Microbiol.*, **72,** 7485-7494.

[195] Fuhrmann M, Hausherr A, Ferbitz L, Schodl T, Heitzer M, & Hegemann P (2004) Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Mol. Biol.*, **55,** 869-881.

[196] 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 (Pt 8),** 2725-2731.

[197] Wenzel SC, Gross F, Zhang Y, Fu J, Stewart AF, & Müller R (2005) Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via red/ET recombineering. *Chem. Biol.*, **12,** 349-356.

[198] Gross F, Ring MW, Perlova O, Fu J, Schneider S, Gerth K, Kuhlmann S, Stewart AF, Zhang Y, & Müller R (2006) Metabolic engineering of *Pseudomonas putida* for methylmalonyl-CoA biosynthesis to enable complex heterologous secondary metabolite formation. *Chem. Biol.*, **13,** 1253-1264.

[199] Madzak C, Treton B, & Blanchin-Roland S (2000) Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J. Mol. Microbiol. Biotechnol.*, **2,** 207-216.

[200] Madzak C (2003) New tools for heterologous protein production in the yeast *Yarrowia lipolytica* . *Rec. Res. Dev. Microbiol.*, **7,** 453-479.

[201] Nicaud JM, Madzak C, van den Broek P, Gysler C, Duboc P, Niederberger P, & Gaillardin C (2002) Protein expression and secretion in the yeast *Yarrowia lipolytica*. *FEMS Yeast Res.*, **2,** 371-379.

[202] Papanikolaou S, Chevalot I, Komaitis M, Marc I, & Aggelis G (2002) Single cell oil production by *Yarrowia lipolytica* growing on an industrial derivative of animal fat in batch cultures. *Appl. Microbiol. Biotechnol.*, **58,** 308-312.

[203] Xue ZX, Sharpe PL, Hong SP, Yadav NS, Xie DM, Short DR, Damude HG, Rupert RA, Seip JE, Wang J, Pollak DW, Bostick MW, Bosak MD, Macool DJ, Hollerbach DH, Zhang HX, Arcilla DM, Bledsoe SA, Croker K, Mccord EF, Tyreus BD, Jackson EN, & Zhu Q (2013) Production of omega-3 eicosapentaenoic acid by metabolic engineering of *Yarrowia lipolytica*. *Nat. Biotechnol.*, **31,** 734-740.

[204] Xue Z, Yadav NS, & Zhu QQ (2009) Optimized strains of *Yarrowia lipolytica* for high eicosapentaenoic acid production. *Int. Pat.*, US 2009/0093543

[205] Prestele J, Hierl G, Scherling C, Hetkamp S, Schwechheimer C, Isono E, Weckwerth W, Wanner G, & Gietl C (2010) Different functions of the C3HC4 zinc RING finger peroxins PEX10, PEX2, and PEX12 in peroxisome formation and matrix protein import. *Proc. Natl. Acad. Sci. U S A*, **107,** 14915-14920.