Fatty Acid Production in *Schizochytrium* sp.: Involvement of a Polyunsaturated Fatty Acid Synthase and a Type I Fatty Acid Synthase

A. Hauvermale[a,b], J. Kuner[a], B. Rosenzweig[a], D. Guerra[b,1], S. Diltz[a], and J.G. Metz[a,*]

[a]Martek Biosciences Boulder Corporation, Boulder, CO, [b]University of Colorado at Colorado Springs, Colorado Springs, CO

ABSTRACT: *Schizochytrium* sp. is a marine microalga that has been developed as a commercial source for docosahexaenoic acid (DHA, C22:6 ω-3), enriched biomass, and oil. Previous work suggested that the DHA, as well as docosapentaenoic acid (DPA, C22:5 ω-6), that accumulate in *Schizochytrium* are products of a multi-subunit polyunsaturated fatty acid (PUFA) synthase (1). Here we show data to support this view and also provide information on other aspects of fatty acid synthesis in this organism. Three genes encoding subunits of the PUFA synthase were isolated from genomic DNA and expressed in *E. coli* along with an essential accessory gene encoding a phosphopantetheinyl transferase (PPTase). The resulting transformants accumulated both DHA and DPA. The ratio of DHA to DPA was approximately the same as that observed in *Schizochytrium*. Treatment of *Schizochytrium* cells with certain levels of cerulenin resulted in inhibition of ^{14}C acetate incorporation into short chain fatty acids without affecting labeling of PUFAs, indicating distinct biosynthetic pathways. A single large gene encoding the presumed short chain fatty acid synthase (FAS) was cloned and sequenced. Based on sequence homology and domain organization, the *Schizochytrium* FAS resembles a fusion of fungal FAS β and α subunits.

Paper no. L10047 in *Lipids 41*, 739–747 (August 2006)

DHA, along with other long chain (LC-, those fatty acids with carbon chain lengths greater than 18) PUFAs, have been identified as important dietary compounds for humans. The ω-3 LC-PUFAs, such as DHA and eicosapentaenoic acid (EPA, C20:5), are implicated in early neural and retinal development, the prevention of arteriosclerosis and coronary heart disease, alleviation of inflammation, and retarding the growth of tumor cells (2). The effects are seen as a result of the ω-3 fatty acids acting as competitive inhibitors of ω-6 fatty acids and from the beneficial effects of the ω-3 compounds in their own right (*3*).

A major dietary source of ω-3 LC-PUFAs is from marine fish or fish oil supplements. However, due to emerging concerns with sustainability of marine resources and with the levels of environmental contaminants (PCBs, dioxins, and mercury) in fish (4), major efforts have been made to identify or create alternative sources. One area that has received considerable attention is the creation of crop plants whose seed oils are enriched in LC-PUFA via the use of genetic engineering techniques (5, 6). The oils of these plants do not normally contain LC-PUFA but can contain significant amounts of the C18 PUFAs: linoleic (C18:2 ω-6) and α-linolenic (C18:3 ω-3) fatty acids. Most of the engineering strategies are based on modification of the plant's endogenously produced PUFA through introduction of foreign genes, encoding a variety of enzymes that can either elongate the fatty acid chain or insert additional *cis*-double bonds (i.e., elongases and desaturases, respectively). These efforts have achieved significant progress. Genes have been isolated that encode the entire suite of enzymes needed to convert linoleic and/or α-linolenic fatty acid precursors into a range of LC-PUFA up to and including DHA, and initial results of LC-PUFA production in plants have been published (6,7). A major challenge associated with this approach is the channeling of the products of these individual enzymes to the next desired reaction (8).

The cultivation of microbial organisms has provided an alternative commercial source of oils enriched in LC-PUFA. The fungus *Mortierella alpina* and a dinoflagellate, *Crypthecodinium cohnii*, are both currently grown via fermentation as sources of oils enriched in LC-PUFA [arachidonic acid, ARA (C20:4 ω-6), and DHA, respectively] that are added to infant formulas (9, 10). The microalga *Schizochytrium* sp., as well as other strains of Thraustochytrids, can produce large amounts of oil—up to 55% of the cell weight—in which DHA can comprise as much as 35% of the total fatty acids (11). *Schizochytrium* sp. is grown on a commercial scale via fermentation for both biomass (for animal feed) and for its oil (11).

Until recently, it was assumed that all LC-PUFA were produced by variations of the same basic pathway mentioned above. That is, an FAS system produces short chain saturated fatty acids, and these products are subsequently modified by the actions of distinct elongase and desaturase enzymes. Indi-

*To whom correspondence should be addressed at Martek Biosciences Boulder Corporation, 4909 Nautilus Court North, Suite 208, Boulder, Colorado 80301. Email: jmetz@martekbio.com

Abbreviations: ACP, acyl carrier protein; ARA, arachidonic acid (C20:4 ω-6); bp, nucleotide base pair; DHA, docosahexaenoic acid (C22:6 ω-3); DPA, docosapentaenoic acid (C22:5, this study refers exclusively to DPA ω-6); EPA, eicosapentaenoic acid (C20:5 ω-3); IPTG, isopropyl-1-thio-β-D-galactopyranoside; KAS, β-ketoacyl-ACP synthase; LC-PUFA, those PUFA with carbon chain lengths greater than 18; Orf; open reading frame; Orf B*, *Schizochytrium* PUFA synthase Orf B modified for expression in *E. coli*; PKS, polyketide synthase; PPTase, phosphopantetheinyl transferase; RBS, ribosome binding site.

[1]Current address: United States Dept. of Agriculture, Agricultural Research Service, Washington State University, Pullman, WA

cations of the existence of an alternative pathway for LC-PUFA synthesis initially came from studies of certain marine bacteria. In 1986, DeLong and Yayanos (12) reported the detection of LC-PUFA (either EPA or DHA) in several strains of psychrophilic marine bacteria. Subsequently, Yazawa (13) identified a segment of genomic DNA from an EPA-producing bacterium (*Shewanella* strain SCRC-2738) that when transformed into *E. coli* resulted in EPA accumulation in those cells. It was then determined that proteins encoded by five open reading frames (Orfs) were necessary and sufficient for this activity (14). A number of observations led to the conclusion that four of those proteins represented subunits of an enzyme complex capable of *de novo* synthesis of EPA (1). This PUFA synthase enzyme possessed multiple domains, some of which showed homology to those found in FAS systems while others were similar to those of polyketide synthase (PKS) systems. The fifth essential gene required for EPA synthesis in *E. coli* encoded a phosphopantetheinyl transferase (15) that activated the acyl carrier protein (ACP) domains present on the enzyme by attachment of a cofactor (16). Genes with homology similar to the *Shewanella* EPA gene cluster have been identified in other marine bacteria, including one which accumulates DHA (16-18). It is possible that most, or all, of the marine bacteria that synthesize LC-PUFA will be found to utilize this system.

A set of three genes encoding a PUFA synthase homologous to those found in marine bacteria was also discovered in *Schizochytrium*. A number of biochemical and physiological observations, in addition to the molecular data, indicated that the abundant LC-PUFAs that accumulate in this organism are the products of a PUFA synthase, or synthases (1, 11). Here we confirm that the three genes identified in *Schizochytrium* do indeed encode subunits of a PUFA synthase and that this synthase accounts for production of both DHA and DPA. Additionally, we provide evidence that the other abundant fatty acids present in *Schizochytrium* oil (C14:0 and C16:0) are the products of a separate FAS. With regard to both sequence homology and domain organization, the *Schizochytrium* FAS resembles those found in fungi. However, in contrast to the two subunits of the fungal FAS, all of the domains of the *Schizochytrium* FAS are found on a single large protein.

EXPERIMENTAL PROCEDURES

Strains. *Schizochytrium* sp. is available from the American Type Culture Collection (ATCC 20888) as are *Nostoc* sp. PCC7120 (ATCC 27893, deposited as *Anabaena* sp.) and *Bacillus subtilus* (ATCC 21332). The *E. coli* strain BL21 [BL21(DE3)] was obtained from Novagen.

General preparation of E. coli transformants. DNA containing Orfs encoding the *Schizochytrium* PUFA synthase subunits A, B, and C were cloned separately and together into *E. coli* expression vectors derived from pET21c (Novagen). The constructs were transformed into BL21 cells and expressed using the T7 system with IPTG (isopropyl-1-thio-β-D-galactopyranoside) induction (Novagen). Initial assembly of the Orfs, each containing modifications to facilitate cloning, was done in the pBluescript II SK(+) plasmid (Stratagene, Inc.). The fragments used to assemble the PUFA synthase subunits, as well as the *Schizochytrium* FAS, were cloned from a lambda library of genomic DNA. Standard methods were used to produce that library as well as polymerase chain reaction (PCR)-derived probes (based on sequences obtained from cDNA clones) for isolation of the respective DNA fragments. Sequences for the Orfs encoding PUFA synthase subunits are available from GenBank (accession numbers: AF378327, AF378328, AF378329 for subunits A, B, and C, respectively). The GenBank sequences provide the reference base pair (bp) positions and restriction site locations described in the cloning steps listed below. The Orf encoding the FAS gene was sequenced using a combination of subcloning and primer walking. The accession number for the *Schizochytrium* FAS Orf is EF015632.

Orf A was cloned as three pieces. Part 1 included the beginning of the Orf and extended to a unique PstI site 18 bp from the start codon. This fragment was generated using two complimentary oligonucleotides that were designed to add EcoRI and PacI restriction enzyme sites and a bacterial ribosome binding site (RBS) upstream of the start codon. Additionally, they incorporated an NdeI site at the start codon. Part 2 was a 1.2-kb fragment extending from the PstI site to a unique XmaI site at 1243 bp. Part 3 was a 7.5-kb fragment extending from the XmaI site to a SpeI site 6 bp downstream of the stop codon. Parts 2 and 3 were assembled into pBluescript II cut with PstI and SpeI. Part 1 was added to this construct as an EcoRI-PstI fragment generating the modified Orf A clone.

Orf B was cloned in four parts. Part 1 was generated by PCR and included the beginning of the Orf, and extended to a BsiWI site 237 bp from the start of the Orf. The 5′ primer included an EcoRI site followed by an RBS and an NdeI site at the start codon. The 3′ primer included a SacII site to facilitate further cloning. The PCR product was digested with EcoRI and SacII and cloned into pBluescript II digested with the same restriction enzymes. Part 2 was a 4.36-kb restriction fragment extending from the BsiWI site to a SacII site 4605 bp into the Orf. Part 3 was a gel-purified 1.54-kb restriction fragment extending from the SacII site at 4605 bp to a ClaI site 6147 bp into the Orf. Part 4 extended from the ClaI site at 6147 bp to the end of the Orf. This short (68 bp) fragment was generated using two overlapping oligonucleotides that included EcoRI and SacII sites upstream of the ClaI site and a BamHI site just downstream of the stop codon to facilitate further cloning. The modified Orf B was assembled in the following sequence. The cloned part 1 was digested with BsiWI and SacII and ligated together with the gel-purified part 2 fragment. The new clone was digested with EcoRI and SacII, and the 4.6-kb fragment gel purified. The part 4 clone was digested with SacII and ClaI, and ligated together with the gel-purified part 3 fragment. This plasmid was opened up by digestion with EcoRI and SacII and ligated together with the 4.6-kb gel-purified fragment containing parts 1 and 2.

Orf C was cloned in four parts. Part 1 contained the start codon and extended to a NarI site 255 bp into the Orf. This fragment was generated by PCR. The 5′ primer included an

EcoRI site followed by an RBC and an NdeI site at the start codon. The primer for the 3′ end contained a BamHI site just beyond the NarI site to facilitate further cloning. The PCR product was digested with EcoRI and BamHI and ligated into pBluescript II digested with the same restriction enzymes. Part 2 was a 0.7-kb restriction fragment extending from the NarI site to a BamHI site 971 bp into the Orf. Part 3 was a 3.3-kb restriction fragment extending from the BamHI site to a PflMI site at 4248 bp. Part 4 extended from the PflMI site to the end of the Orf. This fragment was generated by PCR using primers that added BamHI and EcoRI sites upstream of the PflMI site and NheI, HindIII, and XbaI sites downstream of the stop codon to facilitate further cloning. The PCR product was digested with BamHI and XbaI and ligated into pBluescript digested with the same enzymes. The modified Orf C was assembled using the following steps. The plasmid containing part 1 was digested with NarI and BamHI and ligated together with the gel-purified part 2 fragment. This plasmid was digested with EcoRI and BamHI, and the 1-kb fragment containing parts 1 and 2 was gel purified. The plasmid containing part 4 was digested with PflMI and BamHI and ligated together with the part 3 fragment. The plasmid containing parts 3 and 4 was opened up by digestion with EcoRI and BamHI and ligated together with the fragment containing parts 1 and 2.

Cloning of modified Orfs A, B, and C individually into a modified pET21c and testing for protein expression in E. coli. To facilitate cloning into pET21c, restriction sites for NheI, PacI, PmeI, and SpeI were added, in that order, to the multiple cloning region between the XhoI and BlpI sites using two adapter oligonucleotides. The modified vector (pMET21) was used for the cloning of the individual, as well as multiple, PUFA synthase Orfs. The modified PUFA synthase Orf A was cut out of pBluescript II with NdeI and SpeI. The resulting 8.7-kb fragment was gel purified and ligated into pMET21 cut with the same enzymes. The modified Orf B was cut out of pBluescript II with NdeI and BamHI. The resulting 6.2-kb fragment was gel purified and ligated into pMET21 cut with NdeI and BamHI. The modified Orf C was cut out of pBluescript II with NdeI and HindIII. The resulting 4.5-kb fragment was gel purified and ligated into pMET21 cut with NdeI and HindIII.

BL21 cells carrying the various pMET21 constructs were grown in Luria Broth plus ampicillin (100 μg/mL) at 32°C with shaking to an A_{600} of 0.5. IPTG was added to a final concentration of 1 mM, the cells grown for an additional four hours, and then harvested by centrifugation. The cell pellets were heated in an SDS sample buffer and the extracted proteins analyzed by SDS-PAGE (sample buffer, 4-20% polyacrylamide Tris-glycine gels, and Simple Blue stain, all from Novex).

Resynthesis of a serine codon TCT repeat region in Orf B and testing for protein expression. Orf B contains a stretch of 15 consecutive, identical serine codons (TCT), located at 4510 to 4554 bp. An approximately 200-bp segment, containing the region from a BspHI site at 4416 bp to the SacII site at 4605 bp, was resynthesized by use of PCR and a series of partially overlapping oligonucleotides, using the two-stage amplification protocol outlined in reference 19 as a guide. The final PCR product was cloned and sequenced and the BspH-SacII fragment gel purified for subsequent cloning. This fragment replaced the native region of Orf B (forming Orf B*) using the steps outlined for the general preparation of the Orf B clone described above. The pBluescript clone containing part 1 was cut with BsiWI and SacII. Into this was ligated a 4.2-kb BsiWI-BspHI fragment derived from the clone containing parts 1 and 2, plus the resynthesized BspHI-SacII fragment. Digestion of this clone with EcoRI and SacII yielded a 4.6-kb fragment which was gel purified and ligated into the clone of parts 3 and 4 digested with the same enzymes. This clone contained the complete, modified Orf B*, which was then transferred to pMET21, using the same cloning strategy as described above, and tested for protein expression.

Cloning of Orfs A, B, and C into pMET21, and construction of pMET21-Orf B/Orf C/Orf A.* Starting with the pMET21-Orf B clone, Orfs C and A were sequentially added downstream of the previously inserted genes. This created a polycistronic construct with the expression of all three genes driven by the single T7 promoter in pMET21. Orf B utilized the vector-encoded RBS, while a new RBS was supplied with Orfs C and A. A 4.5-kb EcoRI-NheI fragment isolated from pMet21-Orf C (containing Orf C plus an upstream RBS), was ligated into pMET21-Orf B cut with the same enzymes. The resulting pMET21-OrfB/Orf C was digested with PacI and SpeI. Into this was ligated an 8.7-kb PacI/SpeI fragment containing Orf A plus an upstream RBS, isolated from pMet21-Orf A. The resulting plasmid, pMET21-Orf B/Orf C/Orf A was used as the source of a 13.2-kb EcoRI-SpeI Orf C/Orf A fragment used to construct pMET21-Orf B*/Orf C/Orf A. This construct was generated by digesting pMET21-Orf B* with EcoRI and SpeI and ligating in the gel-purified fragment.

Cloning of sfp and HetI into pACYC184. Two PPTases were tested in the *E. coli* PUFA synthase expression system: sfp, from *Bacillus subtilis* (20), and HetI, from the cyanobacterium *Nostoc* strain 7120 (21). An expression vector for sfp was built by cloning the coding region along with defined upstream- and downstream-flanking DNA into pACYC184 (New England Biolabs, Inc.). The oligonucleotides: CGGGGTACCCGGGAGCCGCCTTGGCTTTGT (forward); and AAACTGCAGCCCGGGTCCAGCTGGCAGGCACCCTG (reverse), were used to amplify this region from genomic *B. subtilus* DNA. Restriction enzyme sites were included in the oligonucleotides to facilitate cloning into the EcoRV site of pACYC184. This pACYC184-sfp clone was used to create a HetI expression cassette. The sfp coding region (along with three nucleotides immediately upstream of the ATG) was replaced with a section of DNA containing NdeI and XhoI restriction sites. The ATG sequence embedded in the NdeI site was utilized as the initiation methionine codon for the introduced HetI Orf. There are no methionines present in the Orf, but there are several potential alternative start codons (TTG and ATT) near the 5′ end (see reference 21, Figs. 2 and 3). PCR was used to amplify the Orf with primers designed to replace the furthest 5′ TTG codon with an ATG codon as part of the above-mentioned NdeI site. PCR was also used to introduce an XhoI site just beyond the stop codon. The modified HetI

coding sequence was then inserted into the NdeI and XhoI sites of the pACYC184 vector construct containing the sfp regulatory elements. Upon introduction into BL21 cells, transformants were selected based on chloramphenicol (35 μg/mL) resistance.

Fatty acid content of E. coli cells expressing Schizochytrium PUFA synthase subunits A, B, and C plus sfp or HetI. BL21 cells containing pMET-OrfB*/OrfC/OrfA and either pACYC184-sfp or pACYC184-HetI were grown in various media supplemented with ampicillin (100 μg/mL) and chloramphenicol (35 μg/mL) to A_{600} = 0.5. IPTG was then added to a final concentration of 0.5 mM. Incubation was continued for 24 hours and the cells were harvested by centrifugation, washed in 50 mM Tris pH 7.5, and processed for FAME analysis. The following media were tested: Luria Broth, Luria Broth supplemented with 5- or 10% glycerol and 765 medium (recipe available from the ATCC) supplemented with 10% (wt/vol) glycerol. Cells were grown at 20°C or 32°C with shaking. Cells transformed with pMET-OrfC/OrfA served as the negative control.

Fatty acids present in the culture were converted to FAMEs and identified by comparison of retention times of standards using GC-FID. FAMEs were prepared by adding 2 mL of 1.5 N anhydrous HCl in methanol and 1 mL toluene to aliquots of the wet pellets and heating at 100°C for two hours with frequent mixing. After cooling to room temperature, 1 mL of NaCl-saturated water was added, the samples mixed, and the organic layer removed for analysis. Determination of carbon-carbon double bond positions was made by interpretation of MS data of dimethyloxazoline (DMOX) derivatives of the FAME samples (22). 100 μL of the FAME solution was transferred to screw top test tubes and the solvent removed with nitrogen gas. One mL of 2-amino-2-methyl propanol was added, the tubes flushed with nitrogen, capped, and heated at 150°C for 18 hours. The residue was dissolved in 5 mL of dichloromethane, 5 mL of water was added, and the aqueous layers were removed after thorough mixing. The wash step was carried out five times. The dichloromethane solution was dried over anhydrous sodium sulfate, filtered, and concentrated using nitrogen gas to 1 mL final volume. Identification of fatty acids in the DMOX derivatives was based on comparison with published data when available, and with the interpretation of MS of individual sample peaks.

Effects of cerulenin on ^{14}C acetate labeling of fatty acids in Schizochytrium cells. Schizochytrium cells were grown in a medium designed to promote oil accumulation (23), with shaking at 27ºC, until the start of logarithmic growth (A_{600} of ~3.0). Cerulenin (Sigma) dissolved in EtOH was added to glass tubes and the solvent evaporated prior to adding 1 mL aliquots of the culture. The amount of cerulenin in the tubes represented final concentrations of 0, 0.5, 1.0, 5.0, 10, 25, 50,100, or 200 μM in the cell culture. The cells were incubated for 10 minutes prior to the addition of ^{14}C acetate (39 mCi/mmol: Sigma) to a final concentration of 20 μM. Incubations were continued until >90% of the label had been removed from the medium (~45 minutes). Cells were collected by centrifugation, and FAMES were prepared essentially as described above. The solvent was removed with nitrogen gas and the residue dissolved in hexane and applied to silica gel G TLC plates (J.B. Baker) that had been treated with a 10% silver nitrate solution. Plates were developed with hexane/diethyl ether/acetic acid (70:20:2). Radioactive areas on the plates were detected using a phosphorimaging system (BioRad, FX Pro Plus). The intensities of photon emissions from the storage phosphor screens were used to measure the relative amounts of radioactive material present and are expressed as "counts" in rectangles of equivalent areas. Identities of the FAMEs on the plates were established by comparison to unlabeled FAME standards (C14:0, C16:0, DPA, and DHA: NuCheck Prep). The C14:0 and C16:0 standards co-migrated on the plates near the solvent front, while DPA and DHA ran as separate bands near the origin.

RESULTS

Expression of Schizochytrium PUFA synthase subunits A, B, and C in E. coli. We sought to confirm that the *Schizochytrium* Orfs A, B, and C encode a PUFA synthase and to verify its fatty acid products by expressing that system in *E. coli*. The PUFA synthase genes previously expressed in *E. coli* or in a cyanobacterium are from marine bacteria, and the Orfs encoding the subunits of those PUFA synthases are organized in apparent operons. Expression of those systems in *E. coli* utilized the promoter, initiation, and translation elements of the source DNA (13, 14, 16). The eukaryotic nature of the *Schizochytrium* genes precluded this approach. We utilized the Novagen pET vectors to express the genes encoding subunits A, B, and C individually and together as a synthetic operon. SDS-PAGE analyses of extracts from *E. coli* in which the individual genes were induced revealed that the genes encoding subunits A and C produced proteins of the expected sizes. In contrast, extracts from cells expressing the subunit B gene accumulated a novel protein with an apparent mass of 165 kDa instead of the 224 kDa protein predicted by translation of the coding sequence (data not shown). Examination of the Orf B nucleotide sequence revealed a region containing 15 sequential identical serine codons (TCT) at a position in the Orf that could account for generation of the truncated protein.

A 200-bp fragment of Orf B containing the serine repeat region was resynthesized so that a random mix of the three serine codons commonly used by *E. coli* replaced the TCT repeat. Some other potentially problematic codons in this fragment were also changed while maintaining the original deduced amino acid sequence (Fig. 1). The native fragment of Orf B was replaced with the modified one (creating Orf B*) and expressed in *E. coli*. In this case an apparently full-length subunit B protein was detected. Analysis of cells expressing the construct containing all three Orfs (A, B*, and C) revealed the presence of three new proteins of the expected sizes. Analysis of the fatty acids in these cells did not reveal any LC-PUFA.

Selection and cloning of a PPTase-HetI. Subunit A of the *Schizochytrium* PUFA synthase contains nine adjacent, highly homologous, ACP domains. ACPs require conversion from an inactive apo-form to a functional holo-form by attachment of a

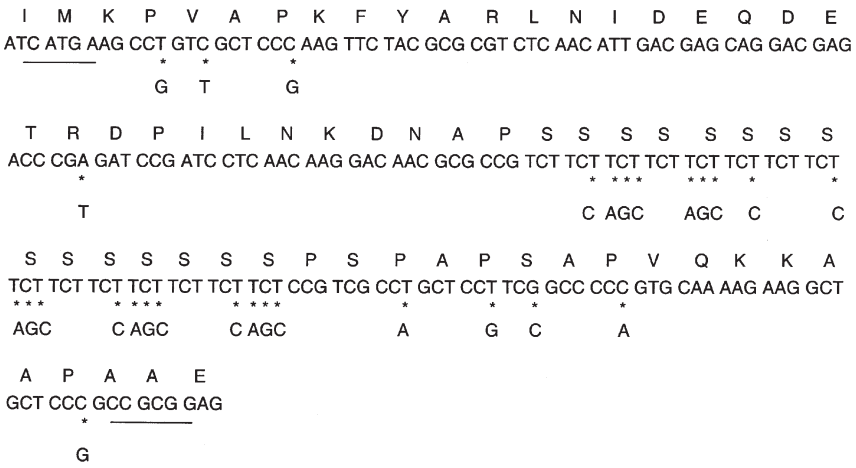

FIG. 1. Modification of the *Schizochytrium* PUFA synthase Orf B nucleotide sequence for expression in *E. coli*. The sequence of codons associated with a BspHI-SacII fragment (restriction enzyme recognition sites are underlined) is shown, along with the encoded amino acids. Nucleotides changed in the resynthesized fragment are indicated by *, with the altered nucleotide shown below. The first and last nucleotides shown correspond to positions 4414 and 4608, respectively, of the complete *Schizochytrium* Orf B sequence (Accession number, A378328).

phosphopantetheine prosthetic group. Enzymes of the PPTase superfamily catalyze this post-translational modification (15). Previous work suggested that *E. coli's* endogenous PPTases would be incapable of activation of the ACPs of PUFA synthases (16). A PPTase that was likely to meet our requirements was identified by searching for ones that may activate ACP domains with significant homology to those of the PUFA synthases. The *hetI* gene is presumed to encode a PPTase (HetI) and is found in a cluster of genes in *Nostoc* sp. responsible for the synthesis of long-chain hydroxy-fatty acids (21, 24). It is likely that HetI activates the two ACP domains of the protein encoded by the *hglE* gene present in that same cluster. The ACP domains of HglE have a high degree of sequence homology to each other and to the ACP domains found in *Schizochytrium* Orf A as well as those of other PUFA synthases (Fig. 2). The amino acids in the region immediately surrounding the phosphopantetheine attachment site are identical in these ACP domains, and there is significant homology in the extended sequence. In contrast, the *E. coli* ACP has relatively few residues that are identical to those in this alignment. To our knowledge, HetI has not been tested previously in a heterologous system, and the endogenous start codon of HetI has not been identified. There are no ATG codons in the potential coding region, but there are several alternative bacterial start codons (TTG and ATT) near the 5′ end of the Orf (21). A HetI expression construct was made by replacing the farthest 5′ alternative start codon (TTG) of the Orf with a methionine codon. Regulatory elements obtained from the upstream region of the *sfp* gene of *B. subtilus* were used to drive expression of the gene. SDS-PAGE analysis of extracts from *E. coli* transformed with this plasmid revealed the appearance of a new protein of the appropriate size (not shown). The gene-encoding sfp has been used extensively for activation of ACP domains of PKS systems as well as the peptidyl carrier protein domains of non-ribosomal peptide synthase systems during heterologous expression (15). We also tested the effect of expression of sfp with the *Schizochytrium* PUFA synthase genes in *E. coli*.

Accumulation of DHA and DPA in E. coli expressing the Schizochytrium Subunits A, B, and C plus HetI or sfp. E. coli cells were transformed with various plasmids containing some or all of the putative *Schizochytrium* PUFA synthase genes and with a second plasmid containing a PPTase gene. Transcription of the genes dependent on the T7 polymerase was induced by addition of IPTG, the cultures incubated overnight, and the fatty acids analyzed. A typical GC profile of FAMEs derived from control cells (lacking at least one component of the PUFA synthase system) is shown in Fig. 3A. The major peaks on the chromatogram were identified as 16:0, 16:1, 18:0, and 18:1 (*cis*-vaccenic acid), as well as two cylcopropane fatty acids (17:0 and 19:0). No peaks corresponding to DHA or DPA, or any other PUFA, were detected in these samples. In contrast,

FIG. 2. Alignment of the highly conserved central regions of the first ACP domains from a heterocyst glycolipid synthase of *Nostoc* sp. PCC7120 (HglE), the PUFA synthase from *Shewanella* SCRC3728 (Orf 5), and the PUFA synthase of *Schizochytrium* (Orf A). Identical amino-acid residues are indicated by the grey background. Underlined amino acids indicate residues of the *E. coli* ACP that are identical to those in this alignment. An * marks the conserved serine pantetheinylation site. The accession numbers and the range of amino acid sequences used are the following: HglE, BAB77050, 1320–1353; Orf 5, AAB81123, 1269–1302; Orf A, AAK72879, 1135–1168; *E. coli*, AAC74178, 15–48.

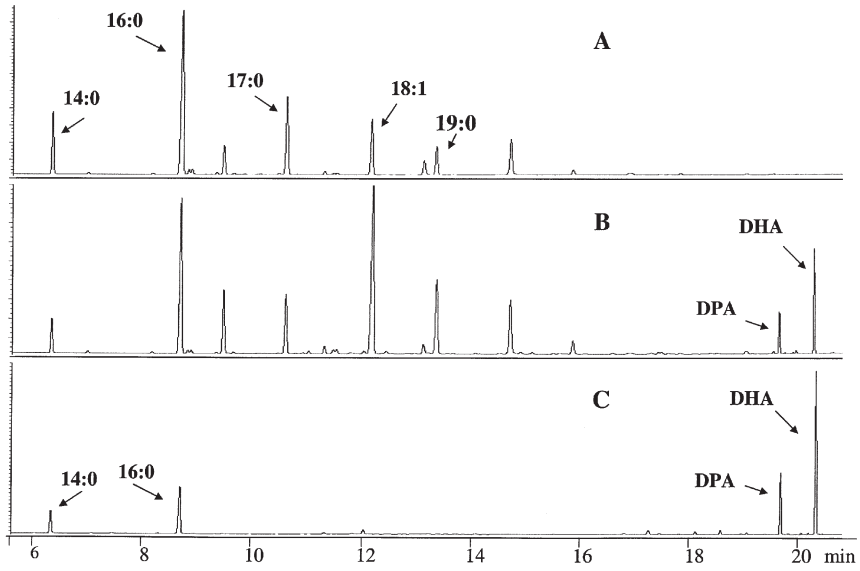

FIG. 3. GC profiles of FAMEs prepared from the following: A) control *E. coli*; B) *E. coli* expressing Orfs A, B, and C of the *Schizochytrium* PUFA synthase plus HetI; and C) oil extracted from *Schizochytrium*. The horizontal axis represents retention time on the column, and the vertical axis the relative response signal of the detector. Identities of the labeled FAME peaks have been confirmed by GC-MS. The 18:1 peak is *cis*-vaccenic acid. The *Schizochytrium* oil FAME profile is shown for comparison to the novel fatty acids produced in the *E. coli* expressing the *Schizochytrium* PUFA synthase system.

expression of *Schizochytrium's* Orfs A, B, and C, along with either HetI or sfp—under several conditions—resulted in the accumulation of DHA and DPA in those cells. The highest level of PUFA (DHA plus DPA), often representing greater than 10% by weight of the total FAME, was found in cells with HetI grown at 32°C in 765 medium supplemented with 10% (wt/vol) glycerol (Fig. 3B). In all of the experiments in which the two PPTases, sfp and HetI, were compared, significantly more DHA and DPA accumulated in the cells containing HetI than in cells containing sfp. SDS-PAGE analysis suggested approximately equivalent amounts of the proteins were being produced, but we did not attempt further analysis of the relative enzyme activities. Fig. 3C shows a typical GC separation of FAMEs prepared from purified *Schizochytrium* oil. There are four prominent peaks in the chromatogram, representing C14:0, C16:0, DPA, and DHA. The ratio of DHA to DPA observed in *E. coli* expressing this PUFA synthase system is about 2.3 to 1. This approximates that of the endogenous DHA and DPA production observed in *Schizochytrium* itself.

Cerulenin differentially affects fatty acid synthesis in Schizochytrium. Rapidly growing cells were treated with various amounts of the β-ketoacyl-ACP synthase inhibitor, cerulenin, for ten minutes prior to the addition of and incubation with ^{14}C-labeled acetate. Radiolabeled fatty acids present in those cells were converted to FAMEs, separated by TLC, and quantitated using a phosphorimaging system. Three major bands, identified as DHA, DPA, and short-chain-saturated (C16:0 and C14:0) FAMEs, were evident in the image, and all were affected by the cerulenin at some concentration. The inhibitor affected incorporation of labeled acetate into DHA and DPA equally and required high concentrations (>5 μM), while labeling of the short-chain-saturated fatty acids was blocked at lower concentrations (Fig. 4). In some cases, for example, at 5, 10, 25, and 50 μM, there is significant labeling of the PUFAs, while labeling of the short-chain saturates has been severely inhibited. The results indicate a separate pathway for synthesis

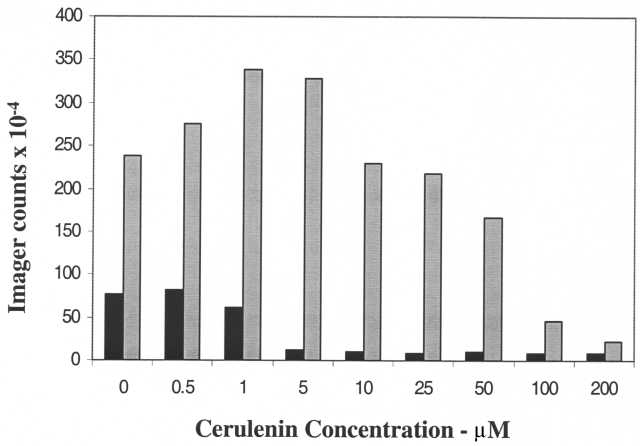

FIG. 4. Relative incorporation of radioactivity into PUFAs (DHA plus DPA - grey bars) or short-chain saturated fatty acids (C14:0 plus C16:0 - black bars) from ^{14}C acetate by *Schizochytrium* cells preincubated with various concentrations of the FAS inhibitor, cerulenin. Imager counts are values in equal-sized areas of the phosphorimager data collected from a scan of a TLC plate on which FAMEs derived from the cells were separated.

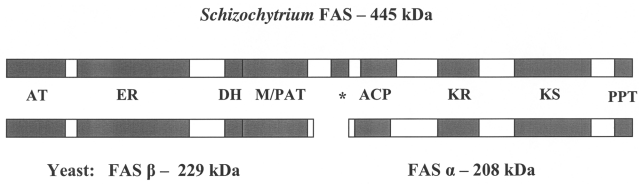

FIG. 5. The apparent domain organization of the *Schizochytrium* FAS (all on a single protein) with that of *Saccharomyces cerevisiae* FAS (domains distributed on the β and α subunits). Relative sizes of the domains reflect the results of Pfam profiling. Domain indicators refer to both *Schizochytrium* and yeast FAS proteins. The * indicates the location in the *Schizochytrium* FAS of a putative ACP domain not present in the yeast FAS. The following are abbreviations used: AT, acetyltransacylase; ER enoyl-reductase; DH, dehydratase; M/PAT, malonyl/palmitoyl transacylase; ACP, acyl carrier protein; KR, β-ketoacyl reductase; KS, β-ketoacyl synthase; PPT, phosphopantetheinyl transferase.

of the two major types of fatty acids that accumulate in *Schizochytrium*.

Cloning of the gene encoding the Schizochytrium FAS. Homologs of bacterial PUFA synthase genes were initially detected in *Schizochytrium* by analysis of sequences obtained from randomly picked cDNA library clones (1). Further analysis of the cDNA library identified many sequences (43 out of approximately 8,500) that showed homology to proteins of FAS systems. The best matches were, in most cases, to the FAS subunits of fungal organisms. Using the sequence tags as a guide, we cloned the corresponding regions of genomic DNA encoding those cDNA-derived sequences. A single large Orf was found in the genome that contained all of the EST sequences. The *Schizochytrium* FAS Orf contains 12,408 bp and encodes a protein with a deduced molecular mass of 444,884 Daltons. No evidence for introns could be found, either by analysis of the genomic sequence itself or by comparison to available cDNA sequences. Blast and Pfam results plus motif analyses were used to establish a preliminary structure and functional identification of the domains of this Type I protein. Both in terms of amino acid sequence and in the sequential organization, the *Schizochytrium* FAS resembles a fusion of the head (N-terminus) of the fungal FAS α subunit to the tail (C-terminus) of the β subunit (Fig. 5, the well-characterized *Saccharomyces cerevisiae* FAS shown for comparison). This similarity includes regions corresponding to an acetyltransferase initiation domain, a malonyl/palmitoyl acyltransferase domain, and a terminal PPTase domain. One other difference is the presence of what may be a second ACP domain at the fusion junction.

DISCUSSION

Schizochytrium sp. is a member of the Thraustochytrids, a group of marine microorganisms taxonomically related to the heterokont algae (25). Organisms of this group have attracted attention recently due to their ability to produce large amounts of oil enriched in DHA (11). Additionally, several genes encoding key enzymes of the standard pathway of PUFA synthesis, including a Δ-4 desaturase, have been cloned from the Thraustochytrid species closely related to *Schizochytrium* (26). It was somewhat surprising, therefore, that homologs of PUFA synthase genes, previously identified only in marine bacteria were found in *Schizochytrium*. Although there are some differences between the bacterial systems and the one found in *Schizochytrium*, all of the domains known to be required for PUFA synthesis in bacteria have been identified (1). One scenario that could account for the appearance of these genes in this eukaryote is a lateral gene transfer from the bacteria. Thraustochytrids are known to be phagocytotic; bacteria can be used as a food source, and this may have facilitated such a process. The isolation of genes encoding enzymes involved in the synthesis of LC-PUFA from shorter-chain, less unsaturated fatty acid precursors suggests that Thraustochytrids are, or were, capable of DHA synthesis via that pathway. When cultures of *Schizochytrium* sp. were supplied with radiolabeled 16:0, 18:1, or 18:3 fatty acids, none of the label appeared in LC-PUFA (1), indicating that at least some elements of the standard pathway have been lost in this particular organism. As more Thraustochytrids are studied, research may reveal a range of expression of these distinct LC-PUFA synthesis pathways.

The PUFA synthases characterized in marine bacteria appear to produce predominantly a single product: either EPA or DHA. In contrast, the *Schizochytrium* fatty acid profile contains significant amounts of two LC-PUFA: DHA and DPA (Fig. 3C). Based on the lack of evidence for more than one set of PUFA synthase genes, we predicted that both fatty acids would be the product of one enzyme system. This prediction was confirmed by the accumulation of both DHA and DPA in *E. coli* expressing the *Schizochytrium* genes (Fig. 3B).

The PPTases present in *E. coli* are apparently unable to modify the apo-ACP domains of the PUFA synthases. It is now recognized that the detection by Yazawa of EPA in *E. coli* transformed with a genomic fragment of *Shewanella* SCR2738 DNA was made possible because the gene encoding that PUFA-synthase-specific PPTase was also present on the fragment (15, 16). We have not identified a gene in *Schizochytrium* that encodes a PUFA-synthase-specific PPTase. While the *Shewanella* PPTase has been shown to activate other PUFA synthase enzymes (16), the psychrophilic nature of that organism suggested it might not be appropriate for our purposes. The nitrogen-fixing cyanobacterium, *Nostoc* PCC7120 (formerly called *Anabaena*), grows well at relatively high temperatures (e.g., 30ºC) and contains a cluster of genes associated with the synthesis of a long-chain hydroxyl-fatty acid that forms part of an oxygen-excluding glycolipid layer of the heterocysts (21). Many of the domains present in the proteins encoded by these genes, including 2 ACP domains, are highly homologous to those of PUFA synthases. The *hetI* gene, encoding a presumed PPTase, is adjacent to this fatty acid synthesis cluster. As in the case of the *Shewanella* PPTase, it is not clear from the nucleotide sequence what the endogenous start codon is for *hetI*. We found that an active enzyme could be produced by replacing the first potential alternative (bacterial) start codon in the HetI Orf with an ATG in an expression construct. HetI, like sfp, may have utility in activating ACP domains of a variety of PKS

systems in addition to those of the PUFA synthases when they are expressed in heterologous hosts.

While it is clear that the LC-PUFAs in *Schizochytrium* are the products of a PUFA synthase, little is known about the source of the short-chain saturated fatty acids (C16:0 and C14:0) that are also prominent components of the TAG fraction. To our knowledge, no characterization of FAS systems in Thraustochytrids has been published by others. Cerulenin is a fungal antibiotic that irreversibly inhibits β-ketoacyl-ACP synthases (KAS) of FAS systems (27). The sensitivity of specific KAS enzymes to this antibiotic varies widely. Cerulenin has been used to enhance PUFA accumulation in marine bacteria that contain PUFA synthase systems (28), presumably by selectively inhibiting the Type II FAS in those bacteria. The differential effect of cerulenin on incorporation of ^{14}C-acetate into short-chain fatty acids versus PUFA in whole cells of *Schizochytrium* (Fig. 4) confirmed our expectation that these fatty acids are products of separate pathways. EST data suggested the presence of a Type I FAS in *Schizochytrium*. In all cases, the best matches to the EST sequences were to FAS proteins found in fungi. In these organisms the Type I FAS is composed of two subunits (α and β), each carrying a distinct set of enzymatic domains (29). In contrast, we found that the *Schizochytrium* FAS contains all of these domains on one large protein. Animal FASs contain all of their catalytic domains on one large protein, but the domain arrangement and both the loading and termination mechanisms differ from that found in the fungal FAS (29). As was the case for the PUFA synthase, the EST data indicate that there is a single copy of the FAS gene in *Schizochytrium* and that messages derived from that gene are highly expressed in oil-producing cells.

In summary, our data confirm that the DHA and DPA present in the oil of *Schizochytrium* are the products of a single PUFA synthase enzyme. It is likely, though yet to be confirmed by detailed molecular analyses, that the genes encoding this enzyme were acquired at some time in the past from an LC–PUFA-producing marine bacterium. The data suggest that the two short-chain saturated fatty acids (C14:0 and C16:0) that are the other major components of the *Schizochytrium* oil, are the products of an FAS enzyme encoded by the gene we have cloned and characterized. The FAS encoded by this gene shows strong resemblance to many other fungal FAS proteins, but has the unique feature of having all of its domains present on one large protein.

REFERENCES

1. Metz, J.G., Roessler, P., Facciotti, D., Levering, C., Dittrich, F., Lassner, M., Valentine, R., Lardizabal, K., Domergue, F., Yamada, A., Yazawa, K., Knauf, V., and Browse, J. (2001) Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes. *Science 293*, 290–293.
2. Simopoulos, A.P. (1999) Essential fatty acids in health and chronic disease. *Am. J. Clin. Nutri. 70*, 560S–569S.
3. Jump, D.B. (2002) The biochemistry of n-3 polyunsaturated fatty acids. *J. Biol. Chem. 277*, 8755–8758.
4. Jacobs, N.M., Covaci, A., Gheorghe, A., and 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.
5. Napier, J.A., Sayanova, O., Qi, B. and Lazarus, C.M. (2004) Progress toward the production of long-chain polyunsaturated fatty acids in transgenic plants. *Lipids 39*, 1067–1075.
6. Wu, G., Truksa, M., Datla, N., Vrinten, P., Bauer, J., Zank, T., Cirpus, P., Heinz, E., and Qiu, X. (2005) Stepwise engineering to produce high yields of very long-chain polyunsaturated fatty acids in plants. *Nat. Biotechnol*. 23, 1013–1017.
7. Kinney, A.J., Cahoon, E.B., Damude, H.G., Hitz, W.D., Liu, Z.-B., and Kolar, C.W. (2004) Production of very long chain polyunsaturated fatty acids in oilseed plants. *United States Patent Application Publication*. Pub. No.: US 2004/0172682 A1.
8. Domergue, F., Abbadi, A., Ott, C., Zank, T.K., Zahringer, U., and Heinz, E. (2003) Acyl carriers used as substrates by the desaturases and elongases involved in very long-chain polyunsaturated fatty acids biosynthesis reconstituted in yeast. *J. Biol. Chem. 278*, 35115–35126.
9. Streekstra, H. (2005) Arachidonic acid: fermentative production by *Mortierella* fungi, in *Single Cell Oils*, Cohen, Z. and Ratledge, C., eds., pp. 73–85, AOCS Press.
10. Wynn, J., Behrens, P., Sundararajan, A., Hansen, J., and Apt, K. (2005) Production of single cell oils by Dinoflagellates, in *Single Cell Oils*, Cohen, Z. and Ratledge, C., eds., pp. 86–98, AOCS Press.
11. Barclay, W., Weaver, C., and Metz, J. (2005) Development of a docosahexaenoic acid production technology using Schizochytrium: A historical perspective, in *Single Cell Oils*, Cohen, Z. and Ratledge, C., eds., pp. 36–52, AOCS Press.
12. DeLong, E. and Yayanos, A.A. (1986) Biochemical function and ecological significance of novel bacterial lipids in deep-sea prokaryotes. *Applied and Environmental Microbiology 51*, 730–737.
13. Yazawa, K. (1996) Production of eicosapentaenoic acid from marine bacteria. *Lipids 31*, S297–S300.
14. Yu, R., Yamada, A., Watanabe, K., Yazawa, K., Takeyama, H., Matsunaga, T., and Kurane, R. (2000) Production of eicosapentaenoic acid by recombinant marine cyanobacterium, *Synechococcus* sp. *Lipids 35*, 1061–1064.
15. Lambalot, R.H., Gehring, A.H., Flugel, R.S., Zuber, P., LaCele, M., Marahiel, M.A., Khosla, C., and Walsh, C.T. (1996) A new enzyme superfamily – the phosphopantetheinyl transferases. *Chemistry & Biology 3*, 923–936.
16. Facciotti, D., Metz, J.G., and Lassner, M. (2000) Production of polyunsaturated fatty acids by expression of polyketide-like synthesis genes in plants. U.S. Patent 6,140,486.
17. Kaulmann, U. and Hertweck, C. (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew. Chem. Int. Ed. 41*, 1866–1869.
18. Orikasa, Y., Nishida, T., Hase, A., Watanabe, K., Morita, N., and Okuyama, H. (2006) A phosphopantetheinyl transferase gene essential for biosynthesis of *n*-3 polyunsaturated fatty acids from *Moritella marina* strain MP-1. *FEBS Letters 580*, 4423–4429.
19. Pikaart, M.J. and Felsenfeld, G. (1996) Expression and codon usage optimization of the erythroid-specific transcription factor cGATA-1 in baculoviral and bacterial systems. *Protein Expr. Purif. 8*, 469–475.
20. Nakano, M.M., Corbell, N., Besson, J., and Zuber, P. (1992) Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet. 232*, 313–321.
21. Black, T.A. and Wolk, C.P. (1994) Analysis of a Het$^-$ mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary

metabolite in the regulation of heterocyst spacing. *Journal of Bacteriology 176*, 2282–2292.
22. Yang, Y.M., Zhang, J.Y., and Huang, Z.H. (1989) Combined in-beam impact-B/E-linked scan mass spectrometry of oxazoline derivatives for the structure determination of long-chain unsaturated fatty acids. *J. of Lipid Research 30*, 127–133.
23. Ashford, A., Barclay, W.R., Weaver, C.A., Giddings, T.H., and Zeller, S. (2000) Electron microscopy may reveal structure of docosahexaenoic acid-rich oil within Schizochytrium sp. *Lipids 35*, 1377–1386.
24. Campbell, E.L., Cohen, M.F., and Meeks, J.C. (1997) A polyketide-synthase-like gene is involved in the synthesis of heterocyst glycolipids in *Nostoc punctiforme* strain ATCC 29133. *Arch. Microbiol. 167*, 251–258.
25. Cavalier-Smith, T., Allsopp, M.T.E.P., and Chao, E.E (1994) Thraustochytrids are chromists not fungi: 18sRNA signatures of Heterokonta. *Philos. Trans. R. Soc. London B: Biol. Sci. 346*, 387–397.
26. Qiu, X., Hong, H., and MacKenzie, S.L. (2001) Identification of a Δ4 fatty acid desaturase from *Thraustochytrium* sp. involved in the biosynthesis of docosahexaenoic acid by heterologous expression in *Saccharomyces cerevisiae* and *Brassica juncea*. *J. Biol. Chem. 276*, 31561–31566.
27. Omura, S. (1976) The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. *Bacteriological reviews 40*, 681–697.
28. Morita, N., Nishida, T., Tanaka, M., Yano, Y., and Okuyama, H. (2005) Enhancement of polyunsaturated fatty acid production by cerulenin treatment in polyunsaturated fatty acid-producing bacteria. *Biotechnol. Lett. 27*, 389–393.
29. Schweizer, E. and Hofmann, J. (2004) Microbial type I fatty acid synthases (FAS): Major players in a network of cellular FAS systems. *Microbio. Mol. Biol. Reviews 68*, 501–517.

[Received July 28, 2006; accepted September 15, 2006]