A01H 5/00 (2006.01)

PCT/IL20 11/000863

3 November 2011 (03.11.2011)

English

614/09,964 4 November 2010 (04.1.2010) US

BEN-GURION UNIVERSITY OF THE NEGEV RESEARCH AND DEVELOPMENT AUTHORITY [IL/IL]; P.O. Box 653, 84105 Beer-Sheva (IL).

GUHENEUF, Freddy [FR/IL]; Meonot Studentim, Sde Boker Campus, 84993 Sde Boker (IL). LEU, Stefan [IL/IL]; Neve Boker 352, P.O. Box 348, 84490 Midreshet Ben Gurion (IL).

ZARKA, Aliza [IL/IL]; 69 Har Ramon Street, 84681 Beer Sheva (IL). KHOZIN-GOLDBERG, Inna [IL/IL]; Neve Boker 324, P.O. Box 136, 84990 Midreshet Ben Gurion (IL). BOUSSIBA, Sammy [IL/IL]; 16 Yaara Street, 84965 Omer (IL).

GEYRA, Assaf et al; WEBB & CO., P.O. Box 2189, 76121 Rehovot (IL).

ACYL-COA: DIACYLGLYCEROL ACYLTRANSFERASE 1-LKE GENES (PtdGATI) AND USES THEREOF

An isolated protein which is at least partially encoded by a polynucleotide sequence encoding a novel acyl-CoA: diacylglycerol acyltransferase 1-like gene (PtdGATI) from the diatom Phaeodactylum tricornutum is provided together with a composition which includes the isolated protein and a transgenic organism transformed by a polynucleotide encoding same. The invention also provides a method for producing or enhancing the production of oil or triacylglycerols with high saturated fatty acids content.
ACYL-COA: DIACYLGLYCEROL ACYLTRANSFERASE 1-LIKE GENE (PTDGAT1) AND USES THEREOF

FIELD OF INVENTION

This invention is directed to, inter alia, an isolated DNA molecule encoding acyl-CoA: diacylglycerol acyltransferase 1-like gene (PtDGAT1) from the diatom Phaeodactylum tricornutum, a protein encoded by same, and methods of making and utilizing the same.

BACKGROUND OF THE INVENTION

Triacylglycerols (TAG) are the principal carbon storage compounds in various organisms, including vertebrates, oilseed plants, oleaginous fungi, yeast and microalgae. In microalgae, TAGs are mainly accumulated in extraplastidial oil bodies, and are implied as potential resource of neutral lipids for biodiesel production.

TAG is synthesized de-novo by a sequential transfer of fatty acyl chains from acyl-CoA through the glycerol-3-phosphate (G3P) pathway, also known as the Kennedy pathway. Briefly, G3P is first acylated by the action of the acyl-CoA: glycerol-3-phosphate acyltransferase (GPAT), followed by a second acylation step catalyzed by the acyl-CoA: lysophosphatidate acyltransferase (LPAAT). The phosphatidic acid obtained is then dephosphorylated by a phosphatidate phosphatase (PAP) to generate diacylglycerol (DAG), which is finally used as substrate for the acyl-CoA: diacylglycerol acyltransferase (DGAT) to produce TAG. Consequently, DGAT catalyzes the final and committed step for TAG biosynthesis.

Three types of DGAT enzymes commonly referred to as DGAT type 1, 2, and 3, may take part in the acyl-CoA-dependent formation of TAG. Two major isoforms are encoded by DGAT1 and DGAT2 genes that have been identified to encode for distinct proteins which both function as DGAT enzymes and are responsible for the bulk of TAG synthesis in most organisms. Cytosolic DGAT3 has been discovered in peanut Arachis hypogea, but seems to play only a minor role in plant TAG formation.

DGAT1 (EC 2.3.1.20) proteins are members of the MBOAT (membrane-bound O-acyltransferase) protein superfamily and differ structurally from DGAT2 proteins. DGAT1 proteins are larger than DGAT2 and possess at least six transmembrane domains compared to the two predicted in DGAT2 (Yen et al., 2008). The emerging role of DGAT2 orthologs seems to be more important for incorporation of unusual fatty acids in the seed storage oils of
some plants. In addition, the bifunctional DGAT/wax ester synthase from Acinetobacter calcoaceticus, is also a member of the MBOAT superfamily, however, for its higher plants homologs, the DGAT activity is either absent or much lower relative to the wax synthase activity.

**SUMMARY OF THE INVENTION**

In one embodiment, the present invention provides an isolated protein comprising an amino acid sequence set forth in SEQ ID NO: 1.

In another embodiment, the present invention further provides a DNA molecule encoding a protein comprising an amino acid sequence set forth in SEQ ID NO: 1.

In another embodiment, the present invention further provides a transgenic organism or a transformed bacteria comprising an exogenous polynucleotide molecule encoding a protein comprising an amino acid sequence set forth in SEQ ID NO: 1.

In another embodiment, the present invention further provides a cell comprising an expression vector encoding a protein comprising an amino acid sequence set forth in SEQ ID NO: 1.

In another embodiment, the present invention further provides a method for enhancing a production of oil or triacylglycerols with high saturated fatty acids content in a cell, comprising the step of transforming the cell with a polynucleotide encoding a protein comprising an amino acid sequence set forth in SEQ ID NO: 1, thereby enhancing a production of oil or triacylglycerols with high saturated fatty acids content in a cell.

In another embodiment, the present invention further provides a method for increasing the proportion of a saturated fatty acid in a cell, comprising the step of transforming or transfecting the cell with a polynucleotide encoding a protein comprising an amino acid sequence set forth in SEQ ID NO: 1, thereby increasing the proportion of a saturated fatty acid in a cell.

**BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 is a set of bar graphs showing the Turbidity (A), Dry Weight (B), Chlorophyll a content (C) and Total Fatty Acid content (D) determined in Phaeodactylum tricornutum cultures at day 0, 5, 12 on full RSE medium (white columns), and under nitrogen starvation (grey columns).
Figure 2. depicts an alignment of cDNA sequences for the PrDGAT1 and PrDGAT1long, isolated from *Phaeodactylum tricornutum*, with the chromosome 2 genomic DNA (Chr2genDNA), the three EST sequences (ct887168, ct881105, ct880495short), and the putative gene sequence (XM_002177717.1). Intron I of 65 bp (light gray continuous line box), Insert I of 54 bp that does not affect the open reading frame (black dotted line box), Insert II of 63 bp that introduces a stop codon in the ORF of PrDGAT1 (black continuous line box), C-terminal extension of 33 bp (light gray dotted line box).

Figure 3. depicts the amino-acid sequence (565 AA) of PrDGAT1 gene obtained under nitrogen starvation. Underlined: N-terminal obtained from EST CT880495. Bold: ten C-terminal amino-acids (consensus C-terminal sequence of higher plant DGAT1) added to complete C-terminus (YHIDIMNRKGN (SEQ ID NO: 14-Stop). (Amino acids are not included in the gene model sequence XP_002177717.1, but were identified after cloning of the gene by sequencing of the PCR product).

Figure 4. depicts hydrophobicity plots indicating at least 8 transmembrane regions of strong hydrophobicity [created using DAS (top) and TMHMM servers (bottom)].

Figure 5. depicts a phylogram showing relationships among PrDGAT1 and diverse DGAT enzymes from algae and higher plants. The phylogram was calculated based on aligned amino acid sequences using the CLC Sequence Viewer (version 6.3, CLC bio A/S) multiple alignment program. OsDGAT3 (*Oryza sativa*, acc. no. NP001054585.1). AhDGAT3 (*Arachis hypogaea*, acc. no. AAX62735.1). AtDGAT3 (*Arabidopsis thaliana*, acc. no. NP175263.2). RcDGAT2 (*Ricinus communis*, acc. no. XP002528531.1). VfDGAT2 (*Vernicia fordii*, acc. no. CAM4973.1). AtDGAT2 (*Arabidopsis thaliana*, acc. no. NP566952.1). OsDGAT2(1) and (2) (*Oryza sativa*, acc. no. NP001057530.1. and acc. no. NP001047917.1). OtDGAT2 (1). (2) and (3) (*Ostreococcus tauri*, acc. no. CA054993.1. acc. no. CAL58088.1. and acc. no. CAL56438.1). TpDGAT (*Thalassiosira pseudonana*, acc. no. XP002287215.1). OeDGAT1 (*plea europaea*, acc. no. AAS01606.1). PdDGAT1 (*Perilla frutescens*, acc. no. AAG23696.1). NtDGAT1 (*Nicotiana tabacum*, acc. no. AAF19345.1). GmDGAT1 (1) and (2) (*Glycine max*, acc. no. AAS78662.1. and acc. no. BAE93461.1). LjDGAT1 (*Lotus japonica*, acc. no. AAW51456.1). EaDGAT1 (*Euonymus alatus*, acc. no. AAV31083.1). JcDGAT1 (*Jatropha curcas*, acc. no. ABB84383.1). ViDGAT1 (*Vernicia fordii*, acc. no. ABC94471.1). RcDGAT1 (*Ricinus communis*, acc. no. XP002514132.1). BjDGAT1 (1) and (2) (*Brassica juncea*, acc. no. AAY40784.1. and acc. no. AAY40785.1). AtDGAT1 (*Arabidopsis thaliana*, acc. no. NP179535.1). TmDGAT1 (*Tropaeolum majus*, acc. no. AAM03340.2). ZmDGAT1-2 (*Zea mays*, acc. no. ABV91586.1). and OsDGAT1 (1).
(2) and (3) (Oryza sativa, acc. no. NP001054869.2, acc. no. AAV10815.1, and acc. no. NP001054869.1).

Figure 6. depicts sequence comparison of PtDGAT1 with the DGAT hypothetical protein of the diatom Thalassiosira pseudonana and related DGAT1 enzymes from higher plants. TpDGAT (T. pseudonana, acc. no. XP002287215.1), OeDGAT1 (Olea europaea, acc. no. AAS01606.1), VfDGAT1 (Venericida fordii, acc. no. ABC94471.1), RcdDGAT1 (Ricinus communis, acc. no. XP002514132.1), AtDGAT1 (Arabidopsis thaliana, acc. no. NP179535.1), ZmDGAT1-2 (Zea mays, acc. no. ABV91586.1). Conserved motifs or putative signature (see text for details) are boxed, such as N-terminal basic motif RRR in higher plants substituted by KRS in PtDGAT1 (I), the Acyl-CoA binding signature (II), the fatty acid protein signature (III) which contains a tyrosine phosphorylation site (●), the DAG-binding (IV), and two C-terminal motifs YYHD1-like YYHDV (V) and KXXX-like NRGK (VI) as putative endoplasmic reticulum retrieval motifs ER-DIR in the C-terminus. The region containing a conserved leucine repeat (L) in higher plants coinciding with a thiolase acyl enzyme intermediate binding signature besides critical Pro and Ser residues which are marked by asterish. The conserved phenylalanine is designated by arrow.

Figure 7. is a silica-gel micrograph showing the complementation of the TAG-deficient phenotype of the yeast mutant H1246 by expression of PtDGAT1. The expression was performed for 48 h at 30°C. Lipid extracts were separated by TLC, and lipid spots visualized as described in Material and Methods. As positive control, the wild type BY742 harboring the empty plasmid was used (A). As negative control, the NL-deficient quadruple mutant strain H1246 harboring the empty plasmid (B). The NL-deficient quadruple mutant strain H1246 was not expressing PtDGAT1 because of the 63 bp insert introducing stop codons into the ORF (C). The mutant strain H1246 was expressing PtDGAT1 (D). As second positive control, the NL-deficient quadruple mutant strain H1246 was expressing the yeast DGA1 gene (E). SE: Steryl Esters. TAG: Triacylglycerol. FFA: Free Fatty Acids. 1.2-DAG: 1,2-Diacylglycerol.

Figure 8. is a set of microscopic micrographs showing that the formation of lipid bodies is restored upon expression of PtDGAT1 in the yeast strain H1246. Neutral Lipid (NL) accumulation in lipid bodies was visualized in yeast cells using the fluorescent dye Nile Red. As positive control, the wild type BY742 harboring the empty plasmid was used (A). As negative control, the NL-deficient quadruple mutant strain H1246 harboring the empty plasmid (B). The NL-deficient quadruple mutant strain H1246 was not expressing PtDGAT1.
because of the 63 bp insert introducing stop codons into the ORF (C). The mutant strain H1246 was expressing PtDGAT1 (D). As second positive control, the NL-deficient quadruple mutant strain H1246 expressing the yeast DGA1 gene (E).

Figure 9. depicts the amino-acid sequence (756 AA) of the full-length P/DGAT1 including PHdomain gene obtained under nitrogen starvation (SEQ ID NO: 4). Underlined: PH domain predicted using PROSITE-EXPASy software (http://prosite.expasy.org/).

DETAILED DESCRIPTION OF THE INVENTION

In one embodiment, the present invention provides an isolated acyl-coa: diacylglycerol acyltransferase 1-like gene (PtDGAT1) protein. In another embodiment, the present invention provides that the isolated PtDGAT1 protein is a microalga protein. In another embodiment, the present invention provides that the isolated PtDGAT1 protein is derived from a diatom microalga. In another embodiment, the present invention provides that the isolated PtDGAT1 protein is derived from Phaeodactylum tricornutum.

In another embodiment, the present invention provides a P/DGAT1 protein comprising or consisting the amino acid sequence:

DETEITPLLRFSTPSRAEHSSWIKLASECAYSETDEFLADEAARATQRALQHQERALQ MAQAMPGAKPGTLPPLLYFAPTIKRSRSFAKLQEHHGDMPVRNMRRRTKSRDFNAD KLDARSTKGYPPSKMHRSAEPSYLSADAPIQNYRGLNLGVIILIVSNFRLILGRTS NGFVLTTAVKHYKNLNLKEDPWQEFPFVSFQLQLVFVSIAGIEWMLCRKYSNFE NFGMILHHFNAHSALLPLGIVWNLIDRPAYGAILLLHATITWMKLLISMLANEDYRL SSRVVGNNPHLALTALVENLSDEANINYQPNVTLNIFYWFCAPLLTYQAIFKSPR VRYWKIADILMRMTVSIALFTFLAQIVQPALEELPSLDDETNGSYTAIAFAYELK LSIANTYWLLMFYTYFHYLNLFAELRLFRGDRVYKDWNSSEYSAYWRLWNMP VHYWLRHHYFPVCVRKLMPKVAATFVFVFFLSAVMHEVLVSVPFHPWSFNGMM QIPLVAFTKYRKFPGSIGNVLFWMTCVIGQPMALLLY (SEQ ID NO: 1).

In another embodiment, the present invention provides a PtDGAT1 protein comprising or consisting the amino acid sequence:

MDETEITPLLRFSTPSRAEHSSWIKLASECAYSETDEFLADEAARATQRALQHQERALQ QMAQAMPGAKPGTLPPLLYFAPTIKRSRSFAKLQEHHGDMPVRNMRRRTKSRDFNAD DKLDARSTKGYPPSKMHRSAEPSYLSADAPIQNYRGLNLGVIILIVSNFRLILGRTS NGFVLTTAVKHYKNLNLKEDPWQEFPFVSFQLQLVFVSIAGIEWMLCRKYSNFE NFGMILHHFNAHSALLPLGIVWNLIDRPAYGAILLLHATITWMKLLISMLANEDYRL
In another embodiment, the present invention provides a PrDGAT1 protein comprising or consisting of the amino acid sequence:

MDETEITPLLRFSTPSRAEHSSWKLAESCAYSEDFLADEAAARATQRALQHQEAL
QMAQAMPAGKPGTPLPPLYFAPTIKRSRFSAKLQEHHDGDMPRVNMRRRTKSRDFNA
DKLDARSTKGYPKPMHAAEPSYLSADAPIQNRYGFLNLGVLIVSNFRLILGTIR
SNGFVLTATAVKHYKNLNLKEDPWEQFEPFVSGFLLQLVFSIAFQGWMLCRKYFNE
NGMILHHFNAHSSALLQLQVWNLIDRPAGAillasllhtATWKMILSYMLANEDYRL
SSRVRGGNPHLATLALVENLDSDEANINYPQNVTLNRNIFYWFCAPTLYQIAFPKSPR
VRYWKKI ADILMRMTVSIALFTFLAQIVQPAEELVSDLDETNVSYTAIAFEYWLK
LSIANTYLWLLMFYTFHLYNLFAELLRGDFRZFVKDWNNSESAYWRLWNMP
VHYWLIRHVYFPVCRLKMPKVAADFVFVFLSAVMHEVLVSPFHIIPWFSFGIMM
QiPLVAFTKLYRKFPGGSIGNVLFWMTCVIGQPMAILLYYHDIMNKRGN (SEQ ID NO: 3).

In another embodiment, the present invention provides a PrDGAT1 protein comprising or consisting of the amino acid sequence:

MTTPVSSEDTATLQQKI VALQAQLLSATHALERMKNERSADHSKSAQRN
GSDPSSDPTGTAPVAAAPKSGYLFKELDRAIGWGGIKWSLRYVKLESGRISY
YGSHHDTSPRYELORLGCAVRDDGWKRNPFFKTRNEMENTL spacecraft
SVYHAPDAEKEIDEITPLLRFSTPSRAEHSSWKLAESCAYSETDEFLADE

AARATQRALQHQEALQMAQAMPAGKPGTPLPPLYFAPTIKRSRFSAKLQEHHDGDMPRVNMRRRTKSRDFNA
DKLDARSTKGYPKPMHAAEPSYLSADAPIQNRYGFLNLGVLIVSNFRLILGTIR
SNGFVLTATAVKHYKNLNLKEDPWEQFEPFVSGFLLQLVFSIAFQGWMLCRKYFNE
NGMILHHFNAHSSALLQLQVWNLIDRPAGAillasllhtATWKMILSYMLANEDYRL
SSRVRGGNPHLATLALVENLDSDEANINYPQNVTLNRNIFYWFCAPTLYQIAFPKSPR
VRYWKKI ADILMRMTVSIALFTFLAQIVQPAEELVSDLDETNVSYTAIAFEYWLK
LSIANTYLWLLMFYTFHLYNLFAELLRGDFRZFVKDWNNSESAYWRLWNMP
VHYWLIRHVYFPVCRLKMPKVAADFVFVFLSAVMHEVLVSPFHIIPWFSFGIMM
QiPLVAFTKLYRKFPGGSIGNVLFWMTCVIGQPMAILLYYHDIMNKRGN (SEQ ID NO: 3).
LIRHVYFPCVRLKMPKVAATFWFFLSAVMHEVLVSVPFHIIRPWSFIGMMM
QIPLVAFTKYLYRKFPGGSFGNVLFWMTFCVIGQPMAILLYTVDYQYGKHHS
TNMEIFDTDDCRFLW ... comprises a n amino acid
sequence that is a t least 90% identical t o the amino acid sequence o f any one o f SEQ I D

In another embodiment, the PtDGAT1 protein of the present invention comprises an
amino acid sequence that is at least 50% homologous to the amino acid sequence of any one
of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention
comprises an amino acid sequence that is at least 60% homologous to the amino acid
sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of
the present invention comprises an amino acid sequence that is at least 70% homologous to
the amino acid sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 80% homologous to the amino acid sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 95% homologous to the amino acid sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 99% homologous to the amino acid sequence of any one of SEQ ID NOs: 1-4.

In another embodiment, the PtDGAT1 protein of the present invention comprises an
amino acid sequence that is at least 50% identical to the amino acid sequence of any one
of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention
comprises an amino acid sequence that is at least 60% identical to the amino acid sequence
of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present
invention comprises an amino acid sequence that is at least 70% identical to the amino acid
sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 80% identical to the amino acid sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 85% identical to the amino acid sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 90% identical to the amino acid sequence of any one of SEQ ID
NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 95% identical to the amino acid sequence of any one of SEQ ID NOs: 1-4. In another embodiment, the PtDGAT1 protein of the present invention comprises an amino acid sequence that is at least 99% identical to the amino acid sequence of any one of SEQ ID NOs: 1-4.

In another embodiment, the PtDGAT1 protein as described herein comprises at least a portion of the amino acid shown in SEQ ID NO: 1. In another embodiment, the PtDGAT1 protein as described herein is a variant of SEQ ID NO: 1. In another embodiment, the term "variant" in relation to a certain sequence means a protein or a polypeptide which is derived from the sequence through the insertion or deletion of one or more amino acid residues or the substitution of one or more amino acid residues with amino acid residues having similar properties, e.g., the replacement of a polar amino acid residue with another polar amino acid residue, or the replacement of a non-polar amino acid residue with another non-polar amino acid residue. In all cases, variants must have a PtDGAT1 protein function as defined herein.

In another embodiment, the PtDGAT1 protein as described herein further comprises a leader peptide. In another embodiment, the leader peptide allows the polypeptide to be specifically located or targeted to a target organelle within the cell.

In another embodiment, the present invention provides a Phaeodactylum tricomutum isolated PtDGAT1 protein of any one of SEQ ID NOs: 1-4. In another embodiment, the present invention provides an isolated polypeptide comprising a functional Phaeodactylum tricomutum PtDGAT1. In another embodiment, the present invention provides that the polypeptide has the function of a PtDGAT1 protein.

In another embodiment, the present invention provides an isolated polynucleotide encoding the protein as described herein. In another embodiment, the present invention provides an isolated polynucleotide encoding a protein comprising or consisting the amino acid sequence of anyone of SEQ ID NOs: 1-4. In another embodiment, an isolated polynucleotide is an isolated DNA molecule. In another embodiment, an isolated polynucleotide is an isolated cDNA molecule. In another embodiment, the isolated polynucleotide comprises a sequence encoding the protein as described herein. In another embodiment, the isolated polynucleotide comprises a DNA sequence encoding PtDGAT1 protein as described herein. In another embodiment, the isolated polynucleotide comprises a DNA sequence encoding a polypeptide comprising a PtDGAT1 protein activity. In another
embodiment, the isolated polynucleotide comprises a DNA sequence encoding a polypeptide consisting a PtDGAT1 activity.

In another embodiment, the isolated polynucleotide comprises a DNA sequence comprising or consisting the sequence:

5  GATGAGACCGAAATTACACCTTTTGTGTGCACTCTCCCGAGGCGGACG
ACTCGTCCTGGGATAAAAGCTGGCCAGAATCTGTGGCCTAACGACGAAACGACG
AGTTTTCTCGTGCAGGCAGCCGGTCGAGCCAAACCAGCGTGTTGCAACATCAAG
AAGCGCTGCAAATGGCCCAAGCCATGCCTGGGGCAAAGCCAGGAACGCTGCCGC
CACTCTACTTCCGCCTTACAAAGCGGTCGTTCTTCTGCTAAGCTACAAAGA

10  ACATCATGGAGATGGGATGCCGTCCTGGTCAATGCTGCTTACAGCGAAACGGACG
AGTTTTCTCGCTGACGAGGCCGCTCGCGCAACCCAGCGTGCTTTGCAACATCAAG
AAGCGCTGCAAATGGCCCAAGCCATGCCTGGGGCAAAGCCAGGAACGCTGCCGC
AACTACCGAGGATTCTGAATTAGCCGTTATATTTCGATGACGATGAAACG

15  AGCAGCTAAAGAAACCTAAAATCACCTCAAGGAGATCCCCTCCAGAAATTCTTTCTT
TTGTATCAGGATTTCCTTCTCAGCCTGGCTTTGCAACGAGGCTGGTACGAA
TGGATGTGTTGACCCGAAATACCTTCAACGAAATCTCCGCGATAGTCCTCATC
TCAAATGCACCTGCTGTGACATTCTAGATTTTATGATGACGATGAAACG

20  AGCAGCTCAAAACACCTTCAAGGAGATCCCCTCCAGAATTCTTTCTT
TTGTATCAGGATTTCCTTCTCAGCCTGGCTTTGCAACGAGGCTGGTACGAA
TGGATGTGTTGACCCGAAATACCTTCAACGAAATCTCCGCGATAGTCCTCATC
TCAAATGCACCTGCTGTGACATTCTAGATTTTATGATGACGATGAAACG

25  AGCAGCTCAAAACACCTTCAAGGAGATCCCCTCCAGAATTCTTTCTT
TTGTATCAGGATTTCCTTCTCAGCCTGGCTTTGCAACGAGGCTGGTACGAA
TGGATGTGTTGACCCGAAATACCTTCAACGAAATCTCCGCGATAGTCCTCATC
TCAAATGCACCTGCTGTGACATTCTAGATTTTATGATGACGATGAAACG

30  AGCAGCTCAAAACACCTTCAAGGAGATCCCCTCCAGAATTCTTTCTT
TTGTATCAGGATTTCCTTCTCAGCCTGGCTTTGCAACGAGGCTGGTACGAA
TGGATGTGTTGACCCGAAATACCTTCAACGAAATCTCCGCGATAGTCCTCATC
TCAAATGCACCTGCTGTGACATTCTAGATTTTATGATGACGATGAAACG

GATTTGTAATGTCCATGGTTGGGAGATGACGATGAAACG

9
ATTCTTTGTACTACATCATGATATTATGAATCGAAAAGGAAATTGA (SEQ ID NO: 5).

In another embodiment, a PtDGAT1 as described herein comprises a nucleic acid sequence that is at least 50% homologous to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 60% homologous to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 70% homologous to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 85% homologous to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 90% homologous to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 95% homologous to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 99% homologous to the nucleic acid sequence of SEQ ID NO: 5.

In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 50% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 60% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid that is at least 70% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 80% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 85% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 90% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 5. In another embodiment, a PtDGAT1 of the present invention comprises a nucleic acid sequence that is at least 99% identical to the nucleic acid sequence of SEQ ID NO: 5.
In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same that enhances the production of TAGs, oil and lipid bodies. In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same that alters the fatty acid composition of TAGs, oil and lipid bodies in a cell. In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same that directs the biosynthesis of saturated fatty acids such as but not limited to 16:0 and/or 18:0 in a cell. In another embodiment, the present invention comprises a PtDGAT1 protein active within a cell. In another embodiment, the present invention comprises a PtDGAT1 protein active in an in-vitro system in the presence of lipid precursors. In another embodiment, lipid precursors comprise various polyunsaturated C18- and C20- fatty acids of both ω3 and ω6 groups. In another embodiment, lipid precursors comprise endogenous DAG and acyl-CoA pool for acylation. In another embodiment, lipid precursors comprise TAGs.

In another embodiment, the PtDGAT1 protein of the invention is active in a cell exposed to nitrogen starvation conditions. In another embodiment, the PtDGAT1 protein of the invention is active in a microalgal cell exposed to nitrogen starvation conditions. In another embodiment, the PtDGAT1 protein of the invention is active in an alga under nitrogen starvation conditions. In another embodiment, the PtDGAT1 protein of the invention is active in a microalga under nitrogen starvation conditions. In another embodiment, the PtDGAT1 protein of the invention is active in a cell comprising the lipid precursors necessary for the formation of saturated fatty acids-TAGs. In another embodiment, the PtDGAT1 protein of the invention is active in an organism comprising the lipid precursors necessary for the formation of saturated fatty acids-TAGs. In another embodiment, the PtDGAT1 protein of the invention is active in a yeast culture comprising the lipid precursors necessary for the formation of saturated fatty acids-TAGs. In another embodiment, the PtDGAT1 protein of the invention is active in-vitro in the presence of lipid precursors necessary for the formation of saturated fatty acids-TAGs.

In another embodiment, the PtDGAT1 protein of the invention directs the incorporation of various poly-unsaturated fatty acids (PUFAs) into a TAG. In another embodiment, the PtDGAT1 protein of the invention primarily directs the incorporation of saturated fatty acids into a TAG. In another embodiment, the PtDGAT1 protein of the invention has a clear preference for saturated fatty acids such as but not limited to 16:0 and 18:0 species (versus unsaturated fatty acids such as the monounsaturated 16:1 and 18:1). In another embodiment, the PtDGAT1 protein of the invention has a clear preference for
endogenous saturated fatty acids such as but not limited to 16:0 and 18:0 species (versus unsaturated fatty acids such as the monounsaturated 16:1 and 18:1). In another embodiment, active PtDGAT1 protein forms TAGs that are substantially saturated. In another embodiment, active PtDGAT1 protein forms TAGs that are substantially saturated relative to the equivalent ones formed by the yeast protein encoded by the DGA1 gene. In another embodiment, active PtDGAT1 protein enhances the accumulation of 20:5 n-3 in TAG. In another embodiment, active PtDGAT1 protein prefers n-3 C20-PUFA over n-6 C20-PUFA.

In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same combined with additional proteins and/or enzymes and/or substrates that are involved in the biosynthesis of TAGs, oil, and lipid bodies. In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same combined with additional proteins and/or enzymes and/or substrates that are involved in the biosynthesis of saturated fatty acids such as but not limited to 16:0 and/or 18:0.

In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same combined with additional proteins and/or enzymes and/or substrates that are involved in the biosynthesis of P. tricornutum TAGs. In another embodiment, the present invention comprises a PtDGAT1 protein or a nucleic acid molecule encoding the same combined with additional proteins and/or enzymes and/or substrates that are involved in the biosynthesis of P. tricornutum TAGs under nitrogen starvation conditions.

In another embodiment, a cell is a microbial cell, a microalgae cell, an alga cell, a plant cell, a cell derived from a heterotrophic microorganism, or a yeast cell. In another embodiment, alga comprises microalga. In another embodiment, a cell to be used with the methods of the invention is grown under nutrient-deplete conditions such as nitrogen starvation. In another embodiment, a cell to be used with the methods of the invention is grown under slow cell proliferation conditions.

In another embodiment, the present invention provides a method for increasing the proportion of a saturated fatty acid in a cell, comprising the step of transforming a cell with a polynucleotide encoding a PtDGAT1 protein. In another embodiment, the present invention provides a method for increasing the proportion of a saturated fatty acid in a TAG in a cell, comprising the step of transforming a cell with a polynucleotide encoding a PtDGAT1 protein. In another embodiment, increasing the proportion of a saturated fatty acid in a cell or
a TAG comprises decreasing the proportion unsaturated fatty acid in a cell or a TAG. In another embodiment, the present invention provides a method for producing or enhancing the production of oil or triacylglycerols with high saturated fatty acids content. In another embodiment, the present invention provides a method increasing the saturated fatty acids content in oil and/or triacylglycerols. In another embodiment, the present invention provides a method increasing the proportion of saturated fatty acids in oil and/or triacylglycerols. In another embodiment, increasing or enhancing comprises high expression of a polynucleotide molecule of the invention. In another embodiment, provided a method of producing cellular TGAs and/oil rich in saturated fatty acids, comprising the step of over-expressing a PtDGAT1 in the cell. In another embodiment, provided a method of producing cellular TAGs and/oil with high saturated fatty acids content or proportion, comprising the step of over-expressing a PtDGAT1 in the cell.

In another embodiment, the present invention comprises a method of producing oil, triacylglycerols, or a combination thereof in a cell comprising the step of transforming a cell with a polynucleotide encoding a PtDGAT1 protein as described herein. In another embodiment, the present invention comprises a method of producing oil rich in saturated fatty acids, triacylglycerols rich in saturated fatty acids, or a combination thereof in a cell comprising the step of transforming a cell with a polynucleotide encoding a PtDGAT1 protein as described herein. In another embodiment, the present invention comprises a method of enhancing the production of oil rich in saturated fatty acids, triacylglycerols rich in saturated fatty acids, or a combination thereof in a cell comprising the step of transforming a cell with a polynucleotide encoding a PtDGAT1 protein as described herein. In another embodiment, a polynucleotide encoding a PtDGAT1 protein as described herein is transiently or constitutively expressed in the cell.

In another embodiment, the present invention comprises a method of producing oil rich in saturated fatty acids, triacylglycerols rich in saturated fatty acids, or a combination thereof in a cell comprising the step of transforming a cell with a polynucleotide encoding a PtDGAT1 protein as described herein to be used as biodiesel.

In another embodiment, the present invention provides that PtDGAT1 prefers saturated acyl moieties 16:0 and 18:0 over monounsaturated 16:1 and 18:1. In another embodiment, the fact that PtDGAT1 prefers saturated acyl moieties 16:0 and 18:0 over monounsaturated 16:1 and 18:1 and therefore, overexpression of PtDGAT1 is of high value for increasing of production and improving the fatty acid composition of microalgal oils for biodiesel production. In another embodiment, provided a method of producing biodiesel,
comprising the step of over-expressing a PtDGAT1 in a cell. In another embodiment, the use of cells over-expressing PtDGAT1 enables the increase of production and/or improvement of the fatty acid composition of oils produced in cells for biodiesel production. In another embodiment, the use of cells over-expressing PtDGAT1 enables the increase of production and/or improvement of the fatty acid composition of microalgal oils for biodiesel production.

In another embodiment, the present invention comprises a composition comprising a PtDGAT1 protein as described herein or a nucleic acid molecule encoding the same. In another embodiment, the present invention comprises a composition comprising fatty acids and/or TAGs produced by a cell transformed or transfected with PtDGAT1 as described herein. In another embodiment, the present invention includes a composition comprising oil produced in a cell transformed or transfected with PtDGAT1 as described herein. In another embodiment, the present invention comprises a composition comprising any constituent of an oil body produced by a transgenic or transformed organism comprising a polynucleotide molecule encoding the PtDGAT1 protein as described herein. In another embodiment, the present invention comprises a composition comprising the PtDGAT1 protein as described herein or a nucleic acid molecule encoding the same combined with additional proteins and/or enzymes and/or substrates that are involved in the biosynthesis of oil, lipids, fatty acids, and/or TAGs. In another embodiment, the present invention comprises a composition comprising a cell transfected or transformed by a nucleic acid molecule encoding the PtDGAT1 protein as described herein. In another embodiment, the cell is a prokaryotic cell. In another embodiment, the cell is a eukaryotic cell. In another embodiment, the cell is a plant cell. In another embodiment, the cell is an algal cell.

In another embodiment, provided herein a transgenic organism transformed by a polynucleotide of the invention. In another embodiment, the organism is a plant. In another embodiment, the organism is yeast. In another embodiment, the organism is a seed. In another embodiment, the organism is an alga. In another embodiment, the organism is a microalga. In another embodiment, provided herein a seed or an offspring of a transgenic organism as described herein wherein the seed or offspring expresses the PtDGAT1 protein as described herein.

In another embodiment, the present invention comprises a PtDGAT1 transgenic plant or a PtDGAT1 transformed bacteria. In another embodiment, the present invention comprises a PtDGAT1 transgenic microalga or alga. In another embodiment, the present invention comprises a PtDGAT1 transgenic plant or a PtDGAT1 transformed bacteria combined with
additional enzymes and/or substrates that are involved in the biosynthesis of oil, lipids, fatty acids, and/or TAGs rich in saturated fatty acids.

In another embodiment, the present invention comprises a *P. tricornutum* over-expressing the PtDGAT1 protein as described herein. In another embodiment, the present invention comprises a cell or an organism over-expressing PtDGAT1 protein. In another embodiment, the present invention comprises a cell or an organism over-expressing an endogenic or exogenic PtDGAT1 protein. In another embodiment, the present invention provides that over-expression of PtDGAT1 protein results in hyper production of oil, lipids, fatty acids, and/or TAGs. In another embodiment, the present invention provides that over-expression of PtDGAT1 protein results in hyper production of oil, and/or TAGs characterized by high content and/or proportion of saturated fatty acids such as but not limited to: 18:0 and 16:0.

In another embodiment, a genetically engineered organism and/or *P. tricornutum* as described herein is transformed with a vector comprising a polynucleotide molecule encoding PtDGAT1 protein under the control of a constitutively active promoter. In another embodiment, a genetically engineered organism and/or *P. tricornutum* as described herein is transformed with a vector comprising a polynucleotide molecule encoding a self PtDGAT1 protein under the control of a constitutively active promoter. In another embodiment, *P. tricornutum* is transformed with a vector comprising a polynucleotide molecule encoding a self PtDGAT1 protein under the control of a constitutively active promoter. In another embodiment, *P. tricornutum* is transformed with a vector comprising the polynucleotide molecule of SEQ ID NO: 5 under the control of a constitutively active promoter. In another embodiment, *P. tricornutum* is transformed with a vector comprising a polynucleotide molecule encoding the protein comprising or consisting the amino acid sequence SEQ ID NO: 1, under the control of a constitutively active promoter. In another embodiment, yeast, an alga or microalga as described is transformed according to the methods described in EP1789530 which is hereby incorporated herein by reference in its entirety. In another embodiment, the methods described in EP1789530 are used for the introduction of an endogenous PTDGAT1 (such as the one encoded by SEQ ID NO: 2) into competent microalgae cells, thereby over-expressing PtDGAT1 protein and enhancing the biogenesis of oil and/or TAGS characterized by high saturated fatty acid content.

In another embodiment, the expression of a PtDGAT1 as described herein is controlled by a cell specific promoter. In another embodiment, the expression of a PtDGAT1 as described herein is controlled by a yeast promoter. In another embodiment, the expression
of a PtDGAT1 as described herein is controlled by a plant promoter. In another embodiment, the expression of a PtDGAT1 as described herein is controlled by an algal promoter. In another embodiment, the expression of a PtDGAT1 as described herein is controlled by a bacterial promoter. In another embodiment, the expression of a PtDGAT1 as described herein is controlled by a viral promoter.

In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a constitutive promoter. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a constitutive promoter whose expression is independent of environmental and/or developmental factors. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a constitutive promoter whose expression is independent of endogenous factors.

In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a tissue-specific or development-stage-specific promoter. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a promoter element that is expressed or affect the expression of genes in the vascular system, photosynthetic tissues, tubers, roots and/or other vegetative organs, or seeds and/or other reproductive organs.

In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by an inducible promoter. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by an inducible promoter conditioned to environmental conditions and external stimuli that can be artificially controlled. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by an inducible promoter conditioned to an abiotic factor such as light, oxygen levels, nitrogen, heat, cold and wounding. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by an inducible promoter conditioned to a chemical compound, not found naturally in the organism of interest. In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by an inducible promoter conditioned to an antibiotic, copper, alcohol, steroids, and/or herbicides, among other compounds.

In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a synthetic promoter. In another embodiment, a synthetic promoter is made by bringing together the primary elements of a promoter region from diverse origins.
In another embodiment, the expression of PtDGAT1 protein as described herein is controlled by a regulatory expression system based on transactivating proteins. In another embodiment, a regulatory expression system regulates the expression of genes of interest irrespective of their physical position to the target genes.

In another embodiment, a vector is used according to the cell or organism utilized. In another embodiment, bacterial, yeast, algal, plant, and animal cell vectors are readily available to one of average skill in the art. In another embodiment, vector control elements are used according to the cell, organism, or tissue utilized. In another embodiment, bacterial, yeast, plant, and animal cell vector control elements are readily available to one of average skill in the art. In another embodiment, vector control elements comprise an origin of replication and a promoter.

In another embodiment, the present invention provides a composition comprising a vector comprising the polynucleotide as described herein. In another embodiment, the present invention provides a composition comprising a vector comprising a polynucleotide encoding PtDGAT1 protein as described herein.

In another embodiment, one of skill in the art can readily prepare a composition as described herein. In another embodiment, one of skill in the art can readily prepare a composition comprising a polynucleotide as described herein. In another embodiment, one of skill can readily prepare a composition comprising a combination of polynucleotides, plasmids, vectors etc. as described herein. In another embodiment, the present invention provides a composition comprising PtDGAT1 protein as described herein to be used in industrial applications for the manufacturing biodiesel. In another embodiment, a composition as described herein is a kit comprising the components for the in vitro manufacturing of TAGs and/or oil comprising high saturated fatty acids content.

In another embodiment, provided herein a method of producing TAGs and/or oil comprising high saturated fatty acids content in a cell comprising the step of transforming or transfecting a cell with a polynucleotide as described herein, thereby producing TAGs and/or oil comprising high saturated fatty acids content in a cell. In another embodiment, provided herein a method for increasing the content of saturated fatty acid stored in TAGs and/or oil in a cell, comprising the step of over-expressing a polynucleotide sequence encoding PtDGAT1 protein in a cell. In another embodiment, the cell is any cell. In another embodiment, methods for over-expressing or de-novo expressing a protein encoded by a vector such as a plasmid are known to one of average skill in the art.
In another embodiment, provided herein a method for increasing the incorporation of saturated fatty acids into TAGs. In another embodiment, provided herein a method for increasing the incorporation of saturated fatty acids into TAGs in a heterotrophic microorganism, comprising the step of over-expressing a polynucleotide sequence as described herein.

In another embodiment, over-expressing a polynucleotide sequence encoding a self (endogenous) PtDGAT1 protein or an exogenous PtDGAT1 protein in a cell, results according to the methods of the present invention in enhancement of production of oil and/or TAGs with increased saturated fatty acids content or proportion. In another embodiment, cells or organisms of the invention that over express a polynucleotide sequence encoding a self (endogenous) PtDGAT1 protein or an exogenous PtDGAT1 protein are used as a source for biodiesel.

In another embodiment, enhancement or enhancing production and/or expression is measured against control cells exposed to the same conditions which do not express PtDGAT1 protein according to the present invention. In another embodiment, enhancement or enhancing production and/or expression is measured against control cells exposed to the same conditions which do not over-express PtDGAT1 protein according to the present invention.

In another embodiment, the method for enhancing a production of oil and/or TAGs with increased saturated fatty acids content or proportion in a cell further comprises the step of subjecting the cell to a condition selected from: high light, high salinity, low nutrients, or any combination thereof. In another embodiment, the method for increasing or de-novo producing oil and/or TAGs with increased saturated fatty acids content or proportion in a cell further comprises the step of subjecting the cell to temperate conditions.

In another embodiment, algae as described herein are eukaryotic organisms. In another embodiment, algae as described herein are photoautotrophic. In another embodiment, algae as described herein are mixotrophic. In another embodiment, algae as described herein are unicellular. In another embodiment, algae as described herein are multicellular. In another embodiment, algae as described herein are Excavata algae. In another embodiment, algae as described herein are Rhizaria algae. In another embodiment, algae as described herein are Chromista algae. In another embodiment, algae as described herein are Alveolata algae.
In another embodiment, algae as described herein are Chlorophyta. In another embodiment, algae as described herein are Haematococcales. In another embodiment, algae as described herein are freshwater algae or microalgae. In another embodiment, algae as described herein produce saturated fatty acids.

In another embodiment, overexpression of PtDGAT1 protein in a cell results in an increase in the saturated fatty acids content of oil and/or triacylglycerols in a cell. In another embodiment, overexpression of PtDGAT1 protein in a cell results in an increase in the proportion of saturated fatty acids in oil and/or triacylglycerols. In another embodiment, overexpression of PtDGAT1 protein in a cell results in a decrease in the proportion of unsaturated fatty acids in oil and/or triacylglycerols. In another embodiment, overexpression of PtDGAT1 protein in a cell results in a decrease in the proportion of monounsaturated fatty acids in oil and/or triacylglycerols.

In another embodiment, transforming a first alga with an algal gene derived from a second alga results in enhanced production of oil and/or TAGs with increased saturated fatty acids content. In another embodiment, transforming a first alga with an algal gene derived from a second alga results in enhanced or de-novo production of oil and/or TAGs with increased saturated fatty acids content. In another embodiment, transforming an alga with an algal gene derived from the same alga such as described herein results in enhanced production of oil and/or TAGs with increased saturated fatty acids content.

In another embodiment, increasing and/or enhancing the production of oil and/or TAGs with increased saturated fatty acids content is increasing and/or enhancing the absolute content of saturated fatty acids. In another embodiment, increasing and/or enhancing the production of oil and/or TAGs with increased saturated fatty acids content is increasing and/or enhancing the proportion of saturated fatty acids. In another embodiment, increasing and/or enhancing the production of oil and/or TAGs with increased saturated fatty acids content is increasing and/or enhancing the proportion of saturated fatty acids and decreasing the proportion of unsaturated fatty acid such as monounsaturated fatty acids.

In another embodiment, the present invention provides an expression vector comprising the polynucleotide as described herein. In another embodiment, the present invention provides a combination of expression vectors each comprising a polynucleotide as described herein. In another embodiment, the present invention provides a plant specific expression vector comprising the polynucleotide as described herein. In another embodiment,
the present invention provides an algal specific expression vector comprising the polynucleotide as described herein. In another embodiment, the present invention provides a cell comprising the expression vector/s as described herein. In another embodiment, the expression vector/s is contained within an agrobacterium. In another embodiment, a cell is a bacterial cell, an animal cell, plant cell or an algal cell.

In another embodiment, the present invention provides a transformed bacterium, a transgenic plant, a transgenic seed, or a transgenic alga transformed by a polynucleotide as described herein. In another embodiment, the present invention provides a transformed bacterium, a transgenic plant, a transgenic seed, or a transgenic alga transformed by any combination of polynucleotides as described herein. In another embodiment, the present invention provides that the transgenic plant is true-breeding for the polynucleotide/s as described herein. In another embodiment, the present invention provides a transgenic seed, produced by a transgenic plant transformed by the polynucleotide/s as described herein. In another embodiment, transformed bacteria, a transformed cell, a transgenic plant, a transgenic seed, or a transgenic alga as described herein produces oil or TAGs with high saturated fatty acids content or proportion. In another embodiment, transformed bacteria, a transformed cell, a transgenic plant, a transgenic seed, or a transgenic alga as described herein produces oil or TAGs with decreased unsaturated fatty acids content or proportion. In another embodiment, transformed bacteria, a transformed cell, a transgenic plant, a transgenic seed, or a transgenic alga as described herein produces oil or TAGs with decreased monounsaturated fatty acids content or proportion.

In another embodiment, "proportion" refers to the proportion of saturated to unsaturated fatty acids. In another embodiment, "proportion" refers to the proportion of saturated to unsaturated fatty acids in a cell. In another embodiment, "proportion" refers to the proportion of saturated to monounsaturated fatty acids. In another embodiment, "proportion" refers to the proportion of saturated to monounsaturated fatty acids in a cell.

In another embodiment, expression of the protein/s of the invention in plants or seed requires sub-cloning an ORF/s sequence encoding the protein/s into a plant expression vector, which may comprise a viral 35S promoter, and a Nos terminator. In another embodiment, a cassette or promoter/coding sequence/terminator is then be subcloned into the plant binary transformation vector, and the resulting plasmid introduced into Agrobacterium. In another embodiment, the Agrobacterium strain transforms the plant. In another
embodiment, the Agrobacterium strain transforms the plant by the vacuum-infiltration of inflorescences, and the seeds harvested and plated onto selective media containing an antibiotic. In another embodiment, the plasmid confers resistance to an antibiotic, thus only transformed plant material will grow in the presence of an antibiotic. In another embodiment, resistant lines are identified and self-fertilized to produce homozygous material. In another embodiment, leaf material is analyzed for expression of the PtDGAT1 protein. In another embodiment, transformation as described herein is a nuclear transformation. In another embodiment, transformation as described herein is an organella transformation. In another embodiment, transformation as described herein is a chloroplast transformation. In another embodiment, transformation of as described herein is a mitochondrial transformation.

In another embodiment, the present invention provides that the methods as described herein can be utilized for the de-novo production of oil and/or TAGs with increased saturated fatty acids content in a cell. In another embodiment, the present invention provides that the methods as described herein can be utilized for the production of oil and/or TAGs with increased saturated fatty acids content in cells or organisms that do not produce oil and/or TAGs with increased saturated fatty acids content endogenically (in the wild-type). In another embodiment, the present invention provides that the methods as described herein can be utilized for production of oil and/or TAGs with increased saturated fatty acids content in plant cells or a plant.

In another embodiment, the present invention provides that transforming a cell or an organism, transfecting a cell or creating a transgenic organism in accordance to the methods of the invention results in enhancing the production of oil and/or TAGs with increased saturated fatty acids content or proportion in cells that naturally produce oil and/or TAGs with increased saturated fatty acids content or proportion. In another embodiment, the present invention provides that transforming a cell or an organism, transfecting a cell, or creating a transgenic organism in accordance to the invention results in an enhancing the production of oil and/or TAGs with increased amount of saturated fatty acids content or proportion in cells that do not naturally produce oil and/or TAGs with increased amount of saturated fatty acids.

In another embodiment, the terms "enhanced production", "over expression", "increased production", and "induced production", are used interchangeably.
In another embodiment, the level of expression of PtDGATI protein correlates with the saturation index of TAGs in a cell. In another embodiment, the level of expression of PtDGATI protein correlates with the amount of saturated fatty acids in a cell subjected to conditions that favor the production of saturated fatty acids as described herein. In another embodiment, enhanced production of oil and/or TAGs with increased amount of saturated fatty acids in a cell is the result of the introduction of a vector comprising a DNA molecule as described herein under the control of a promoter as described herein into a cell. In another embodiment, enhanced production of oil and/or TAGs with increased amount of saturated fatty acids in a cell is the result of the introduction of a vector comprising a DNA molecule as described herein under the control of a promoter as described herein into a cell as described herein.

In another embodiment, enhanced expression of PtDGATI protein in a cell or an organism results in enhanced production of oil and/or TAGs with increased amount of saturated fatty acids compared to the amount of oil and/or TAGs with increased amount of saturated fatty acids present in a cell or an organism prior to the introduction of a DNA molecule as described herein.

In another embodiment, enhanced expression of PtDGATI protein as described herein results in 5% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 3-10% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 7-20% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 10-30% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 10-80% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 10-100% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 30-120% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 25-500% increase in saturated fatty acids in oil and/or TAGs present in a cell. In another embodiment, enhanced expression of PtDGATI protein as described herein results in 10-50% increase in saturated fatty acids in oil and/or TAGs present in a cell.
present in a cell. In another embodiment, enhanced expression of PtDGATl protein as
described herein results in 15-75% increase in saturated fatty acids in oil and/or TAGs
present in a cell. In another embodiment, enhanced expression of PtDGATl protein as
described herein results in 30-70% increase in saturated fatty acids in oil and/or TAGs
present in a cell. In another embodiment, enhanced expression of PtDGATl protein as
described herein results in 25-50% increase in saturated fatty acids in oil and/or TAGs
present in a cell.

In another embodiment, provided herein that the conditions for enhancing saturated
fatty acids content in oil and/or TAGs comprise environmental conditions such as but not
limited to: light intensity (170 μmol m⁻² s⁻¹), phosphate starvation and salt stress (NaCl
0.8%). In another embodiment, provided herein that the conditions for enhancing saturated
fatty acids content in oil and/or TAGs comprise conditions wherein cell growth is retarded.
In another embodiment, cell growth is retarded as reflected by a decrease in cell division
rate. In another embodiment, cell growth is retarded as reflected by an increase in cell death.
In another embodiment, provided herein that the conditions for enhancing saturated fatty
acids content in oil and/or TAGs require cell growth/algal growth under nitrogen
deprivation. In another embodiment, provided herein that the conditions for
enhancing/enriching saturated fatty acids content in oil and/or TAGs require a change in the
cell stage from biflagellate vegetative green cell to a non-motile and large resting cell. In
another embodiment, provided herein that environmental and/or nutritional stresses, which
interfere with cell division, trigger and/or enhance the production of oil and/or TAGs rich in
saturated fatty acids content. In another embodiment, provided herein that the conditions for
enhancing saturated fatty acids content in oil and/or TAGs comprise contacting the algal
cells with an inhibitor of cell division. In another embodiment, saturated fatty acids, oil
and/or TAGs content is measured both in weight (mg/g) and in cellular (pg/cell) contents.

In another embodiment, provided herein that the conditions for enhancing saturated
fatty acids content in oil and/or TAGs comprise a high dose of light. In another embodiment,
provided herein that the conditions for enhancing saturated fatty acids content in oil and/or
TAGs comprise a high dose of irradiation. In another embodiment, provided herein that the
conditions for enhancing saturated fatty acids content in oil and/or TAGs comprise a
combination of nitrogen deprivation and a salt or an ester of an acid such as acetate addition.
In another embodiment, methods for isolating oil and/or TAGs from cells, alga, microalga, a bacteria, or yeast are known to one of average skill in the art.

In another embodiment, an engineered organism is engineered to express PtDGAT1 protein as described herein. In another embodiment, an engineered organism is engineered to highly express PtDGAT1 protein as described herein. In another embodiment, an engineered plant or alga as described herein is used for the manufacturing of TAGs and/or oil with high saturated fatty acid content. In another embodiment, an engineered cell, an engineered plant or alga as described herein is used for increasing the proportion and/or amount of TAGs and/or oil comprising high saturated fatty acid content. In another embodiment, an engineered plant as described herein is used for manufacturing desired saturated fatty acid.

In another embodiment, the terms "protein", "PtDGAT1 protein", or "polypeptide" are used interchangeably. In some embodiments, the terms "protein", "PtDGAT1 protein", or "polypeptide" as used herein encompass native polypeptides (either degradation products, synthetically synthesized polypeptides or recombinant polypeptides) and peptidomimetics (typically, synthetically synthesized polypeptides), as well as peptoids and semipeptoids which are polypeptide analogs, which have, in some embodiments, modifications rendering the polypeptides/proteins even more stable while in-vivo or more capable of penetrating into cells.

In some embodiments, modifications include, but are not limited to N terminus modification, C terminus modification, polypeptide bond modification, including, but not limited to, CH2-NH, CH2-S, CH2-S=O, 0=C-NH, CH2-0, CH2-CH2, S=C-NH, CH=CH or CF=CH, backbone modifications, and residue modification. Methods for preparing peptidomimetic compounds are well known in the art and are specified, for example, in Quantitative Drug Design, C.A. Ramsden Gd., Chapter 17.2, F. Choplin Pergamon Press (1992), which is incorporated by reference as if fully set forth herein. Further details in this respect are provided hereinunder.

In some embodiments, polypeptide bonds (-CO-NH-) within the polypeptide are substituted. In some embodiments, the polypeptide bonds are substituted by N-methylated bonds (-N(CH3)-CO-). In some embodiments, the polypeptide bonds are substituted by ester bonds (-C(R)H-C-0-0-C(R)-N-). In some embodiments, the polypeptide bonds are substituted by ketomethylene bonds (-CO-CH2-). In some embodiments, the polypeptide bonds are substituted by a-aza bonds (-NH-N(R)-CO-), wherein R is any alkyl, e.g., methyl,
carbo bonds (-CH2-NH-). In some embodiments, the polypeptide bonds are substituted by hydroxyethylene bonds (-CH(OH)-CH2-). In some embodiments, the polypeptide bonds are substituted by thioamide bonds (-CS-NH-). In some embodiments, the polypeptide bonds are substituted by olefinic double bonds (-CH=CH-). In some embodiments, the polypeptide bonds are substituted by retro amide bonds (-NH-CO-). In some embodiments, the polypeptide bonds are substituted by polypeptide derivatives (-N(R)-CH2-CO-), wherein R is the "normal" side chain, naturally presented on the carbon atom. In some embodiments, these modifications occur at any of the bonds along the polypeptide chain and even at several (2-3 bonds) at the same time.

In some embodiments, natural aromatic amino acids of the polypeptide such as Trp, Tyr and Phe, be substituted for synthetic non-natural acid such as Phenylglycine, TIC, naphthylelanine (Nol), ring-methylated derivatives of Phe, halogenated derivatives of Phe or o-methyl-Tyr. In some embodiments, the polypeptides of the present invention include one or more modified amino acid or one or more non-amino acid monomers (e.g., fatty acid, complex carbohydrates, etc.).

In one embodiment, "amino acid" or "amino acid" is understood to include the 20 naturally occurring amino acid; those amino acid often modified post-translationally in vivo, including, for example, hydroxyproline, phosphoserine and phosphothreonine; and other unusual amino acid including, but not limited to, 2-aminoacidipic acid, hydroxylsine, isodesmosine, nor-valine, nor-leucine and ornithine. In one embodiment, "amino acid" includes both D- and L-amino acid.

In some embodiments, the polypeptides or proteins of the present invention are utilized in a soluble form. In some embodiments, the polypeptides or proteins of the present invention include one or more non-natural or natural polar amino acid, including but not limited to serine and threonine which are capable of increasing polypeptide or protein solubility due to their hydroxyl-containing side chain.

In some embodiments, the polypeptides or proteins of the present invention are utilized in a linear form, although it will be appreciated by one skilled in the art that in cases where cyclization does not severely interfere with polypeptide characteristics, cyclic forms of the polypeptide can also be utilized.

In some embodiments, the polypeptides or proteins of present invention are biochemically synthesized such as by using standard solid phase techniques. In some embodiments, these biochemical methods include exclusive solid phase synthesis, partial
solid phase synthesis, fragment condensation, or classical solution synthesis. In some embodiments, these methods are used when the polypeptide is relatively short (about 5-15kDa) and/or when it cannot be produced by recombinant techniques (i.e., not encoded by a nucleic acid sequence) and therefore involves different chemistry.

In some embodiments, solid phase polypeptide or protein synthesis procedures are well known to one skilled in the art and further described by John Morrow Stewart and Janis Dillaha Young, Solid Phase Polypeptide Syntheses (2nd Ed., Pierce Chemical Company, 1984). In some embodiments, synthetic polypeptides or proteins are purified by preparative high performance liquid chromatography [Creighton T. (1983) Proteins, structures and molecular principles. WH Freeman and Co. N.Y.], and the composition of which can be confirmed via amino acid sequencing by methods known to one skilled in the art.


In one embodiment, a polypeptide or protein of the present invention is synthesized using a polynucleotide encoding a polypeptide or protein of the present invention as described herein. In some embodiments, the polynucleotide encoding a polypeptide of the present invention is ligated into an expression vector, comprising a transcriptional control of a cis-regulatory sequence (e.g., promoter sequence). In some embodiments, the cis-regulatory sequence is suitable for directing constitutive expression of the polypeptide of the present invention. In some embodiments, the cis-regulatory sequence is suitable for directing tissue specific expression of the polypeptide of the present invention. In some embodiments, the cis-regulatory sequence is suitable for directing inducible expression of the polypeptide of the present invention. In another embodiment, a polypeptide is a protein comprising a ST activity as described herein.
In another embodiment, the polynucleotide comprises a genomic polynucleotide sequence. In another embodiment, the polynucleotide comprises a composite polynucleotide sequence.

In one embodiment, the phrase "a polynucleotide" refers to a single or double stranded nucleic acid sequence which is isolated and provided in the form of an RNA sequence, a complementary polynucleotide sequence (cDNA), a genomic polynucleotide sequence and/or a composite polynucleotide sequences (e.g., a combination of the above).

In one embodiment, "genomic polynucleotide sequence" refers to a sequence derived (isolated) from a chromosome and thus it represents a contiguous portion of a chromosome.

In one embodiment, "composite polynucleotide sequence" refers to a sequence, which is at least partially complementary and at least partially genomic. In one embodiment, a composite sequence can include some exonal sequences required to encode the polypeptide of the present invention, as well as some intronic sequences interposing there between. In one embodiment, the intronic sequences can be of any source, including of other genes, and typically will include conserved splicing signal sequences. In one embodiment, intronic sequences include cis acting expression regulatory elements.

In one embodiment, the polynucleotides of the present invention further comprise a signal sequence encoding a signal peptide for the secretion of the polypeptides of the present invention. In one embodiment, following expression, the signal peptides are cleaved from the precursor proteins resulting in the mature proteins.

In some embodiments, polynucleotides of the present invention are prepared using PCR techniques or any other method or procedure known to one skilled in the art. In some embodiments, the procedure involves the legation of two different DNA sequences (See, for example, "Current Protocols in Molecular Biology", eds. Ausubel et al., John Wiley & Sons, 1992).

In one embodiment, polynucleotides of the present invention are inserted into expression vectors (i.e., a nucleic acid construct) to enable expression of the recombinant polypeptide. In one embodiment, the expression vector of the present invention includes additional sequences which render this vector suitable for replication and integration in prokaryotes. In one embodiment, the expression vector of the present invention includes additional sequences which render this vector suitable for replication and integration in eukaryotes. In one embodiment, the expression vector of the present invention includes a shuttle vector which renders this vector suitable for replication and integration in both
prokaryotes and eukaryotes. In some embodiments, cloning vectors comprise transcription and translation initiation sequences (e.g., promoters, enhancers) and transcription and translation terminators (e.g., polyadenylation signals).

In one embodiment, a variety of prokaryotic or eukaryotic cells can be used as host-expression systems to express the polypeptides of the present invention. In some embodiments, these include, but are not limited to, microorganisms, such as bacteria transformed with a recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vector containing the polypeptide coding sequence; yeast transformed with recombinant yeast expression vectors containing the polypeptide coding sequence; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors, such as Ti plasmid, containing the polypeptide coding sequence.

In some embodiments, non-bacterial expression systems are used (e.g., plant expression systems) to express the polypeptide of the present invention.

In one embodiment, yeast expression systems are used. In one embodiment, algae expression systems are used. In one embodiment, plant expression systems are used. In one embodiment, a number of vectors containing constitutive or inducible promoters can be used in yeast as disclosed in U.S. Patent No: 5,932,447. In another embodiment, vectors which promote integration of foreign DNA sequences into the yeast chromosome are used.

In another embodiment, expression in a host cell can be accomplished in a transient or a stable fashion. In another embodiment, a host cell is a cell as described herein. In another embodiment, transient expression is from introduced constructs which contain expression signals functional in the host cell, but which constructs do not replicate and rarely integrate in the host cell, or where the host cell is not proliferating. In another embodiment, transient expression also can be accomplished by inducing the activity of a regulatable promoter operably linked to the gene of interest.

In another embodiment, stable expression is achieved by introduction of a construct that integrates into the host genome. In another embodiment, stable expression comprises autonomously replication within the host cell. In another embodiment, stable expression of the polynucleotide of the invention is selected for through the use of a selectable marker located on or transfected with the expression construct, followed by selection for cells expressing the marker. In another embodiment, stable expression results from integration, the site of the constructs integration can occur randomly within the host genome or can be
targeted through the use constructs containing regions of homology with the host genome sufficient to target recombination with the host locus. In another embodiment, constructs are targeted to an endogenous locus, all or some of the transcriptional and translational regulatory regions can be provided by the endogenous locus.

In another embodiment, an expression of a protein as described herein comprising PtDGAT1 activity as described herein includes functional transcriptional and translational initiation and termination regions that are operably linked to the DNA encoding the protein comprising an PtDGAT1 activity. In another embodiment, an expression of proteins as described herein comprising various PTGAT activities includes functional transcriptional and translational initiation and termination regions that are operably linked to the DNA encoding the proteins comprising PtDGAT1 activity. In another embodiment, an expression of proteins as described herein comprising PtDGAT1 activity includes functional transcriptional and translational initiation and termination regions that are operably linked to the DNA encoding the protein comprising PtDGAT1 activity. In another embodiment, transcriptional and translational initiation and termination regions are derived from a variety of nonexclusive sources, including the DNA to be expressed, genes known or suspected to be capable of expression in the desired system, expression vectors, chemical synthesis, or from an endogenous locus in a host cell. In another embodiment, expression in a plant tissue and/or plant part presents certain efficiencies, particularly where the tissue or part is one which is harvested early, such as seed, leaves, flowers, roots, etc. In another embodiment, expression can be targeted to that location in a plant by utilizing specific regulatory sequences that are known to one of skill in the art. In another embodiment, the expressed protein is an enzyme which produces a product which may be incorporated, either directly or upon further modifications, into a fluid fraction from the host plant. In another embodiment, expression of a protein of the invention, or antisense thereof, alters the levels of specific PUFAs, or derivatives thereof, found in plant parts and/or plant tissues. The PTGAT1 coding region, in some embodiments, may be expressed either by itself or with other genes, in order to produce cells, tissues, algae, and/or plant parts containing higher proportions of desired oil and/or TGA (as described herein). In another embodiment, the PTGAT1 coding region is expressed either by itself or with other genes, in order to produce cells, tissues, yeast, algae, and/or plant parts containing higher proportions of oil and/or TGA enriched with saturated fatty acids. In another embodiment, the termination region is derived from the 3' region of the gene from which the initiation region was obtained from or from a different gene. In another embodiment, the termination region usually is selected as a matter
of convenience rather than because of any particular property. In another embodiment, increasing or enhancing oil and/or TGA saturated fatty acids content or proportion is enriched oil and/or TGA with saturated fatty acids.

In another embodiment, a plant or plant tissue is utilized as a host or host cell, respectively, for expression of the protein of the invention which may, in turn, be utilized in the production of polyunsaturated fatty acids. In another embodiment, desired oil or fatty acids such as but not limited to 16:0 and 18:0 are produced in a seed. In another embodiment, methods of isolating seed oils are known in the art. In another embodiment, seed oil components are manipulated through the expression of the protein of the invention in order to provide seed oils that can be added to nutritional compositions, pharmaceutical compositions, animal feeds and cosmetics. In another embodiment, a vector which comprises a DNA sequence encoding the protein as described herein is linked to a promoter, and is introduced into the plant tissue or plant for a time and under conditions sufficient for expression of the protein.

In another embodiment, a vector as described herein comprises additional genes that encode other enzymes, involved in TGA, oil, and/or fatty acids synthesis and/or modification. In another embodiment, the bacteria, plant tissue or plant produces the relevant substrate upon which the enzymes act or a vector encoding enzymes which produce such substrates may be introduced into the plant tissue, plant cell or plant. In another embodiment, a substrate is in contact with the bacteria, or is sprayed on plant tissues expressing the appropriate enzymes. In another embodiment, the invention is directed to a transgenic plant comprising the above-described vector, wherein expression results in increased production of TGAs enriched with saturated fatty acids in, for example, the seeds of the transgenic plant.

In another embodiment, the regeneration, development, and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (for example: Weissbach and Weissbach, In: Methods for Plant Molecular Biology, (Eds.), Academic Press, Inc. San Diego, Calif., (1988)). In another embodiment, regeneration and growth process comprises the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. In another embodiment, transgenic embryos and seeds are similarly regenerated. In another embodiment, resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil. In another embodiment, regeneration and growth process of algae are known to one of skill in the art. In another embodiment, identification, selection, of transgenic algae are known to one of skill in the art.
In another embodiment, development or regeneration of plants containing an exogenous polynucleotide as described herein encodes a protein as described herein and is well known in the art. In another embodiment, development or regeneration of algae containing an exogenous polynucleotide as described herein encodes a protein as described herein and is well known in the art. In another embodiment, the regenerated plants are self-pollinated to provide homozygous transgenic plants. In another embodiment, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. In another embodiment, pollen from plants of these important lines is used to pollinate regenerated plants. In another embodiment, a transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.


In another embodiment, transformation of monocotyledons using electroporation, particle bombardment, and Agrobacterium are known. In another embodiment, transformation and plant regeneration are well established in the art. In another embodiment, assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte et al., Nature 335:454-457 (1988); Marcotte et al., Plant Cell 1:523-532 (1989); McCarty et al., Cell 66:895-905 (1991); Hattori et al., Genes Dev. 6:609-618 (1992); Goff et al., EMBO J. 9:2517-2522 (1990)).

In another embodiment, transient expression systems are used to functionally dissect the oligonucleotides constructs. In another embodiment, practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.), generation of recombinant organisms and the screening and isolating of clones, (see for example, Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor

In one embodiment, the expression vector of the present invention can further include additional polynucleotide sequences that allow, for example, the translation of several proteins from a single mRNA such as an internal ribosome entry site (IRES) and sequences for genomic integration of the promoter-chimeric polypeptide.

In some embodiments, expression vectors containing regulatory elements from eukaryotic viruses such as retroviruses are used by the present invention. In some embodiments, recombinant viral vectors are useful for in vivo expression of the polypeptides of the present invention since they offer advantages such as lateral infection and targeting specificity. In one embodiment, lateral infection is inherent in the life cycle of, for example, retrovirus, and is the process by which a single infected cell produces many progeny virions that bud off and infect neighboring cells. In one embodiment, the result is that a large area becomes rapidly infected, most of which was not initially infected by the original viral particles. In one embodiment, viral vectors are produced that are unable to spread laterally. In one embodiment, this characteristic can be useful if the desired purpose is to introduce a specified gene into only a localized number of targeted cells.


In one embodiment, plant expression vectors are used. In one embodiment, the expression of a polypeptide coding sequence is driven by a number of promoters. In some embodiments, viral promoters such as the 35S RNA and 19S RNA promoters of CaMV
[Brisson et al., Nature 310:511-514 (1984)], or the coat protein promoter to TMV
[Takamatsu et al., EMBO J. 6:307-311 (1987)] are used. In another embodiment, plant
promoters are used such as, for example, the small subunit of RUBISCO [Coruzzi et al.,
promoters, e.g., soybean hspl7.5-E or hspl7.3-B [Gurley et al., Mol. Cell. Biol. 6:559-565
(1986)]. In one embodiment, constructs are introduced into plant cells using Ti plasmid, Ri
plasmid, plant viral vectors, direct DNA transformation, microinjection, electroporation and
other techniques well known to the skilled artisan. See, for example, Weissbach &
Weissbach [Methods for Plant Molecular Biology, Academic Press, NY, Section VIII, pp
421-463 (1988)]. Other expression systems such as insects and mammalian host cell systems,
which are well known in the art, can also be used by the present invention.

It will be appreciated that other than containing the necessary elements for the
transcription and translation of the inserted coding sequence (encoding the polypeptide or
protein), the expression construct of the present invention can also include sequences
engineered to optimize stability, production, purification, yield or activity of the expressed
polypeptide or protein.

In some embodiments, transformed cells are cultured under effective conditions,
which allow for the expression of high amounts of recombinant polypeptide or protein
having PtDGATI activity. In some embodiments, effective culture conditions include, but
are not limited to, effective media, bioreactor, temperature, pH and oxygen conditions that
permit protein production. In one embodiment, an effective medium refers to any medium in
which a cell is cultured to produce the recombinant polypeptide or protein of the present
invention. In some embodiments, a medium typically includes an aqueous solution having
assimilable carbon, nitrogen and phosphate sources, and appropriate salts, minerals, metals
and other nutrients, such as vitamins. In some embodiments, cells of the present invention
can be cultured in conventional fermentation bioreactors, shake flasks, test tubes, microtiter
dishes and Petri plates. In some embodiments, culturing is carried out at a temperature, pH
and oxygen content appropriate for a recombinant cell. In some embodiments, culturing
conditions are within the expertise of one of ordinary skill in the art.

In some embodiments, depending on the vector and host system used for production,
resultant polypeptides or proteins of the present invention either remain within the
recombinant cell, secreted into the fermentation medium, secreted into a space between two
cellular membranes, or retained on the outer surface of a cell or viral membrane.
In one embodiment, following a predetermined time in culture, recovery of the recombinant polypeptide or protein is effected.

In one embodiment, the phrase "recovering the recombinant polypeptide or protein" used herein refers to collecting the whole fermentation medium containing the polypeptide or protein and need not imply additional steps of separation or purification.

In one embodiment, polypeptides or proteins of the present invention are purified using a variety of standard protein purification techniques, such as, but not limited to, affinity chromatography, ion exchange chromatography, filtration, electrophoresis, hydrophobic interaction chromatography, gel filtration chromatography, reverse phase chromatography, concanavalin A chromatography, chromatofocusing and differential solubilization.

In one embodiment, to facilitate recovery, the expressed coding sequence can be engineered to encode the polypeptide or proteins of the present invention and fused cleavable moiety. In one embodiment, a fusion protein can be designed so that the polypeptide or protein can be readily isolated by affinity chromatography; e.g., by immobilization on a column specific for the cleavable moiety. In one embodiment, a cleavage site is engineered between the polypeptide or protein and the cleavable moiety and the polypeptide or protein can be released from the chromatographic column by treatment with an appropriate enzyme or agent that specifically cleaves the fusion protein at this site [e.g., see Booth et ah, Immunol. Lett. 19:65-70 (1988); and Gardella et al, J. Biol. Chem. 265:15854-15859 (1990)].

In one embodiment, the polypeptide or protein of the present invention is retrieved in "substantially pure" form.

In one embodiment, the phrase "substantially pure" refers to a purity that allows for the effective use of the protein in the applications described herein.

In one embodiment, the polypeptide or protein of the present invention can also be synthesized using in vitro expression systems. In one embodiment, in vitro synthesis methods are well known in the art and the components of the system are commercially available.

In another embodiment, oil and/or TAGs with high saturated fatty acids or free saturated fatty acids produced by the methods as described herein are used in the cosmetic industry, biodiesel industry, the drug industry, food additives industry, baby food industry or any other applicable industry. In another embodiment, oil and/or TAGs with high saturated
fatty acids or free saturated fatty acids produced by the methods as described herein are used within a formulation.

In some embodiments, the proteins or oligonucleotides of the invention modified by the covalent attachment of water-soluble polymers such as polyethylene glycol, copolymers of polyethylene glycol and polypropylene glycol, carboxymethyl cellulose, dextran, polyvinyl alcohol, polyvinylpyrrolidone or polyproline. In another embodiment, the modified proteins or oligonucleotides of the invention exhibit substantially longer half-lives in blood following intravenous injection than do the corresponding unmodified compounds. In one embodiment, modifications also increase the proteins or oligonucleotides solubility in aqueous solution, eliminate aggregation, enhance the physical and chemical stability of the compound, and greatly reduce the immunogenicity and reactivity of the compound. In another embodiment, the desired in vivo biological activity is achieved by the administration of such polymer-compound abducts less frequently or in lower doses than with the unmodified compound.

In one embodiment, compositions of the present invention are presented in a pack or dispenser device, such as an FDA approved kit, which contain one or more unit dosage forms containing the active ingredient. In one embodiment, the pack, for example, comprise metal or plastic foil, such as a blister pack. In one embodiment, the pack or dispenser device is accompanied by instructions for administration. In one embodiment, the pack or dispenser is accommodated by a notice associated with the container in a form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals, which notice is reflective of approval by the agency of the form of the compositions or human or veterinary administration. Such notice, in one embodiment, is labeling approved by the U.S. Food and Drug Administration for prescription drugs or of an approved product insert.

Additional objects, advantages, and novel features of the present invention will become apparent to one ordinarily skilled in the art upon examination of the following examples, which are not intended to be limiting. Additionally, each of the various embodiments and aspects of the present invention as delineated hereinabove and as claimed in the claims section below finds experimental support in the following examples.

**EXAMPLES**

Generally, the nomenclature used herein and the laboratory procedures utilized in the present invention include molecular, biochemical, microbiological and recombinant DNA techniques. Such techniques are thoroughly explained in the literature. See, for example,

**Experimental Procedures**

**Materials and Methods**

**Growth conditions**

Axenic cultures of *P. tricornutum* were cultivated on RSE medium in 250-mL Erlenmeyer glass flasks in an incubator shaker under an air/C02 atmosphere (99:1, v/v) and controlled temperature (25°C) and illumination (100 μmol quanta m^-2 s^-1) at a speed of 150 rpm. RSE medium was composed of 34 g ReefSalt (Seachem™, Madison GA, USA)
complemented as following: 18.8 raM KN0₃, 0.51 raM KH₂PO₄, 10.6 mM Na₂SiO₃·9H₂O, 0.77 µM ZnSO₄·7H₂O, 0.31 µM CuSO₄·5H₂O, 1.61 µM Na₂MoO₄·2H₂O, 46.3 µM H₃BO₃, 9.15 µM MnCl₂·4H₂O, 0.172 µM Co(NO₃)₂·6H₂O, 26.8 µM C₆H₆O₂Fe·5H₂O, 46.8 µM citric acid, 50 µg Γ³ vitamin B12, 50 µg Γ³ biotin, and 0.1 mg Γ³ thiamine HCl. For the onset of nitrogen starvation conditions, daily-diluted cultures were centrifuged, cells washed twice, and resuspended in nitrogen-free RSE medium. Cultures were further grown under the same conditions before harvesting for RNA isolation. The nitrogen-free medium was prepared by omitting KN0₃ from the RSE medium. Growth parameters: turbidity, Chi a and DW content were determined as previously described (Solovchenko et al., 2008).

**RNA isolation and cDNA synthesis.**

Total RNA was isolated by the procedure described by Bekesiova et al. (1999), from 35 mL culture grown in the complete RSE medium or in nitrogen-free RSE medium for 5 days. The cells were harvested before by centrifugation at 3000 rpm during 5 min, flash-frozen in liquid nitrogen and stored at -80°C for further use. Total RNA samples were treated with RNA-free Baseline-ZERO™ DNase (Epicentre Technologies, Madison, WI, USA) before being used for cDNA synthesis. cDNA was prepared from 1 µg total RNA-template using the Verso™ cDNA kit (Thermo Fisher Scientific).

**Identification and cloning of PtDGAT1 gene.**

Detailed searches using bioinformatics tools and available databases have been employed to find putative DGAT candidates in the genome of *P. tricornutum*. A hypothetical *P. tricornutum* protein (XP_002177753.1), containing ORF of 404 amino acids, exhibited significant similarity (38% identity, 55% similarity) to the coding region of maize DGAT 1-2 protein (EU039830). A putative full-length cDNA encoding for PtDGAT1 (534 amino acids) including a putative N-terminal (130 amino acids) was further assembled from the nucleotide sequence of hypothetical protein XM_002177717.1 and the expressed sequence tag (EST CT880495) from *P. tricornutum*, which share a 400 base common overlap. The sequences of oligonucleotide primers that were used in this study are given in Table 1. The ORF coding for the putative PtDGAT1 polypeptide of 545 amino acids was amplified using the *PfuUltra* II fusion HS DNA polymerase (Stratagene, La Jolla, CA, USA) with the forward primer (*PtDGATfor-Knpl*) containing a Knpl restriction site (underlined) and a yeast translation initiation consensus followed by ATG (bold), and the reverse primer (*PtDGATrev-Xhol*) containing a C-terminal extension, a *Xhol* restriction site (underlined) and a stop codon (Bold). Since hypothetical protein XP_002177753.1 lacks at least 10 amino acids conserved
in C-terminus of higher plant DGAT1 proteins and a stop codon, ten C-terminal aminoacids, including endoplasmatic reticulum retrieval motifs (ER-DIR: YYHD) were added during amplification by incorporating a 33 base nucleotides sequence into the reverse primer encoding for amino acids YHDIMNRKGN (SEQ ID NO: 14-Stop). The PCR products of the expected size were excised, purified from the gel (Nucleospin Extract II purification kit; Macherey-Nagel, Duren, Germany), cloned into pGEM®-T Easy Vector Systems (Promega), and several clones were sequenced for each condition.

Table 1 Oligonucleotide primers used in this study

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'CGCGGTTACCATGGATGAGACCGAAATTACA C 3' (SEQ ID NO: 6)</td>
<td>PtDGATfor-Knpl</td>
</tr>
<tr>
<td>5'GGCCCTCGAGTCAATTTCCTTTTCGATCATAATATC ATGATAGTACAAGAGAGAATCGCCATTGG3' (SEQ ID NO: 7)</td>
<td>PtDGATrev-Xhol</td>
</tr>
<tr>
<td>5'GGATCCACATAATGTCAGGAACATTCATGATATA AG3' (SEQ ID NO: 8)</td>
<td>DGA1-tBamHI</td>
</tr>
<tr>
<td>5'TGCGGGCGGTTACCAACTATCCTCAATTCTGCATC 3' (SEQ ID NO: 9)</td>
<td>DGA1-rNotI</td>
</tr>
<tr>
<td>5'ATGACCCACGCTGTATCTTC3' (SEQ ID NO: 10)</td>
<td>PtDGAT1-PHfor</td>
</tr>
<tr>
<td>5'TCAACGAATCAAGCAGGAAT3' (SEQ ID NO: 11)</td>
<td>PtDGAT1-PHrev</td>
</tr>
</tbody>
</table>

As positive control in yeast expression assays, the yeast DGA1 gene encoding DGAT1 was cloned similarly to the PtDGAT1 of P. tricornutum, using the forward primer (DGA1-tBamHI) and reverse primer (DGA1-rNotI).

Further research using bioinformatics tools and available databases have been used to find the promoter upstream of the assembled putative full-length sequence encoding for PrDGAT1 (1653 bp), which was successfully expressed and functionally characterized in a S. cerevisiae neutral lipid-deficient quadruple mutant strain H1246. During the assembly (1653 bp) the start-codon was artificially added, by modifying ATT to ATG instead, after the Knpl restriction site in the forward primer PtDGATfor-Knpl. By bioinformatic searches, a new Pleckstrin Homology (PH) domain was found upstream to the artificial methionine of the catalytically important region of PtDGAT1. Consequently, new primers PtDGAT1-PH/o r and PtDGAT1-PHrev have been designed (see Table 1) to amplify a putative full-length cDNA sequence encoding for PtDGAT1 including -PHdomain (756 amino acids). The ORF for PtDGAT1 (including -PHdomain) gene is assumed to contain the genuine start-codon, and the native C-terminal sequence. This domain is similar to the PH domain recently
annotated in the diatom *Thalassiosira pseudonana* DGAT1 (ADV58933.2) and is absent in other DGAT1 proteins previously characterized.

Expression and functional characterization of PtDGAT1 cDNA by the heterologous expression in yeast *Saccharomyces cerevisiae*

*S. cerevisiae* strains used were BY742 (relevant genotype: *MATa his3A1 leu2A0 metl5A0 ura3A0*); and H1246 (relevant genotype: *MATa ADE2-1 canl-100 ura3-l arel- A::HIS3 are2-A::LEU2 dgal-A::KanMX4 Irol-A::TRPl*) containing knockouts of the *DGAl, LROI, AREJ* and *ARE2* genes, was kindly provided by Prof. S. Stymne (Scandinavian Biotechnology Research, Alnarp, Sweden). Before transformation, yeast cells were cultivated in 1% (w/v) Yeast extract, 2% (w/v) Peptone and 2% (w/v) Glucose (YPG medium) at 30°C.

Plasmids (pGEM®-T Easy Vector Systems Promega), harboring the assembled full-length sequence encoding for PtDGAT1 (1653 bp), was then restricted with *Knpl* and *XhoI* (NEB, Ipswich, MA, USA). The expected bands were purified from the gel (Nucleospin Extract II purification kit; Macherey-Nagel, Duren, Germany), ligated into a *Knpl*-*Xhol* cut pYES2 vector (Invitrogen, Carlsbad, CA, USA), and the *S. cerevisiae* strains were transformed by the PEG/lithium acetate method (Ausubel *et al.*, 1995). The yeast H1246 and BY742 cells harboring the empty pYES2 vector were used as negative and positive controls, respectively; and as second positive control the yeast H1246 cells harboring the yeast *DGAl* gene were used. Transformants were selected by uracil prototrophy on yeast synthetic medium (YSM) lacking uracil (Invitrogen).

For functional expression, a minimal selection medium containing 2% (w/v) raffinose was inoculated with the PtDGAT1 transformants and grown at 30°C for 24 h in a water bath shaker. Twenty mL of sterile YSM was inoculated with raffinose-grown cultures to obtain an OD of 0.2 at 600 nm. Expression was induced by adding galactose to a final concentration of 2% (w/v) and cultures were further grown at 30°C for 48 h. For fatty acid substrate specificity of PtDGAT1 vs. *S. cerevisiae DGAl*, supplementation of YSM cultures with α-linolenic acid (ALA, 18:3 n-3), γ-linolenic acid (GLA, 18:3 n-6), eicosatrienoic acid (ETA, 20:3 n-3), arachidonic acid (ARA, 20:4n-6), and eicosapentaenoic acid (EPA, 20:5 n-3) in some experiments was carried with 250 µM of appropriate fatty acid in the presence of 1% (w/v) Tergitol-40 (precursors). Cells were harvested by centrifugation, washed twice with 0.1% NaHCO₃, freeze-dried, and used for lipid analysis.
For fatty acid and lipid analysis of *P. tricornutum*, cells were harvested from liquid cultures. Fatty acid methyl esters (FAME) from total lipid, and TAG were obtained by transmethylation of the freeze-dried cells or TAG extracts with dry methanol containing 2% (v/v) \(\text{H}_2\text{SO}_4\) and heating at 80°C for 1.5 h while stirring under an argon atmosphere. Gas chromatographic analysis of FAME was performed on a Thermo Ultra gas chromatograph (Thermo Scientific, Italy) equipped with PTV injector, FID, and a fused silica capillary column (30 m x 0.32 mm; Supelco WAX-10, Sigma-Aldrich). FAMEs were identified by co-chromatography with authentic standards (Sigma Chemical, St Louis, MO, USA) and FAME of fish oil (Larodan Fine Chemicals, Sweden). Each sample was analyzed in triplicate.
EXAMPLE 1

Lipid content and fatty acid composition of P. tricornutum

The effects of nitrogen starvation on growth parameters and total fatty acids (TFA) content were monitored in batch cultures of P. tricornutum (Figure 1). During starvation, turbidity, DW and Chl a content decreased dramatically in comparison to nitrogen-sufficient cultures. The decreases in growth parameters were accompanied by an increase in the TFA content which reached 19% of DW after 12 days, while the sharpest increase occurred within the first 5 days. Under Nitrogen starvation, biomass lipid content of P. tricornutum increased about four-fold in comparison to the culture grown on full RSE medium.

The fatty acid composition of P. tricornutum (Table 2) was characterized by a large fraction of palmitic (16:0), palmitoleic (16:1) and eicosapentaenoic acids (EPA, 20:5 n-3). On full RSE medium, palmitoleic acid showed a substantial increase (from 21 to 30% of TFA) during growth, and represented about three-quarters of total fatty acids with palmitic acid by the end of the stationary phase. Conversely, other fatty acids decreased slightly with culture age (e.g. palmitic and stearic acids) whereas no significant change was observed in the relative proportion of eicosapentaenoic acid. Under nitrogen starvation, the main alterations in TFA profile were accounted for by relative increase of palmitic and palmitoleic acids (from 17 to 31% of TFA, and from 28 to 37% of TFA, respectively) concomitant with the decrease in the EPA proportion (from 22 to 7% of TFA).

Table 2 Fatty acid composition of P. tricornutum cultured under nitrogen starvation

<table>
<thead>
<tr>
<th>Time (days)</th>
<th>12</th>
<th>5</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-</td>
<td>4.5</td>
<td>5.9</td>
<td>4.7</td>
</tr>
<tr>
<td>N+</td>
<td>31.1</td>
<td>14.8</td>
<td>29.0</td>
</tr>
<tr>
<td>Saturated fatty acids</td>
<td>1.5</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>0.2</td>
<td>tr</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>37.4</td>
<td>22.2</td>
<td>35.6</td>
<td>22.2</td>
</tr>
</tbody>
</table>

Monounsaturated fatty acids
EXAMPLE 2

Isolation and identification of DGAT1 in P. tricornutum

Since nitrogen starvation induces TAG biosynthesis in P. tricornutum, a cDNA isolated from algae cultivated under nitrogen starvation conditions for 5 days was utilized. The PCR-amplification utilizing a forward primer designed on the basis of the EST (CT880495) and a reverse primer designed based on the nucleotide sequences of the predicted protein (XM_002177717.1), yielded a DNA fragment of 1695 bp putatively encoding for the full-length coding sequence of PtDGAT1. When cDNA was isolated from
algal cells grown on complete nutrient RSE medium, a band of 1759 bp (designated as PtDGATllong) was amplified using the same pair of primers. Importantly, both sequences appeared to be slightly longer than the DNA size of 1602 bp (PhaeoJoinedSeq), predicted from assembling the database sequences of the predicted hypothetical mRNA for protein (XM_002177717.1) with the EST (CT880495) from *P. tricornutum*.

Both DNA sequences were cloned into pGEM-T® Easy Vector Systems and sequenced. The two resulting sequences, PtDGATl and PtDGATllong differ by a single insert of 63 bases introducing stop codons into the PtDGATl open reading frame, suggesting that PtDGATllong is likely a splice variant with possible regulatory function (Insert II cDNA sequence:

\[ TAAGCCCAACGTCTGCTGCATTCAATGATTTCCGTTTTCCATGACTTACACCGCA \]

\[ TTTTGTA (SEQ ID NO: 12); \] Insert amino-acid sequence: AHVCCIQCDRFHDLHRIS (SEQ ID NO: 13)). This insert introduced stop codons in the ORF of PtDGATl.

The nucleotide sequences were aligned to the predicted gene sequence (XM_002177717.1, http://www.ebi.ac.uk/Tools/psalw2/index.html), and to three *P. tricornutum* EST sequences (ct887168, ct881105, ct880495) showing similarity to the hypothetical *P. tricornutum* protein (XP_002177753.1), and to the coding region of MBOAT family proteins. The sequences determined by us and the EST sequences differed from the predicted gene sequence (XM_002177717.1) by presence of another additional fragment of 54 bp that does not affect the open reading frame, and was apparently erroneously excluded during the predicted gene assembly (XM_002177717.1).

The sequences, PtDGATl and PtDGATllong were aligned, to the *P. tricornutum* chromosome 2 genomic DNA (Bowler et al, 2008) confirming that the 63 bp insert in PtDGATllong is an intron and in the PtDGATl is likely subject to regulated splicing during stress induction or other cellular processes (Figure 2). The other inserts (54 bp) detected in both cDNA preparations are also present in the genomic DNA sequence.

**EXAMPLE 3**

**Gene structure and phylogenetic position of PtDGATl**

The ORF for PtDGATl gene obtained from cDNA isolated from algae grown under nitrogen starvation was 1695 bp in length, coding for the corresponding predicted protein of 565 amino acids (Figure 3). The predicted amino acid sequence of *PtDGATl* is 55% identical to that of *Thalassiosira pseudonoma* putative DGAT1 (XP_002287215), while it
shares more than 35% identity with DGAT1 of higher plants (Figure 5-6) and do not share
significant homology to DGAT2 proteins from higher plants and algae. DGAT1 of diatoms
(a putative TpOGAT and PtDGAT1) form a separate branch on the phylogenetic tree from
the higher plant DGAT1 proteins. It is important to mention that no ortholog genes were
found in the green algae. DGATs in the available genomes of Chlorophyte are represented by
DGAT2.

At least 8 trans-membrane regions of strong hydrophobicity (Figure 4) were predicted
by various algorithms (Das, Phobious and TMHMM, online software) that are likely to
anchor the protein to the endoplasmic reticulum (ER) membrane, a characteristic feature of
DGAT1 proteins in agreement with the putative assessment of PtDGAT1 as DGAT1.
Moreover, the conserved motifs or putative signatures found in higher plants are slightly
different in PtDGAT1 (Figure 6). The characteristic basic motif found in amino acid N-
terminal domain, consisting of three arginine residues (RRR) in higher plants (Bouvier-Nave
et al, 2000; Xu et al., 2008; Manas-Fernandez et al, 2009), is substituted by KRS in
PtDGAT1. The previously reported Acyl-CoA binding signature A^{33}-R^{172} and the DAG-
binding motif V^{458},V^{471} involved in the active site are also found in the PtDGAT1 similarly
to other DGAT1 proteins of plants (Jako et al, 2001; Manas-Fernandez et al, 2009). There
is also a fatty acid binding protein signature spanning residues A^{427} to N^{443}, which contains a
putative tyrosine phosphorylation site: Y^{438} (Xu et al, 2008). Concerning the leucine zipper
motif besides previously described critical Pro and Ser residues coinciding with a thiolase
acyl enzyme intermediate binding signature in higher plants (Zou et al, 1999; Xu et al,
2008), only two of the six conserved leucine are present in PtDGAT1. Also relevant is the
presence of the two C-terminal motifs YYHDI-like (SEQ ID NO: 15) YYHDV (SEQ ID NO:
16) (Manas-Fernandez et al, 2009) or NRGK-KKXX-like (SEQ ID NO: 17) (WolfPSORT
prediction software) confirming the putative ER retrieval motifs in the C-terminus.

**EXAMPLE 4**

Putative intron retention alternatively splicing

Under nitrogen starvation, the nucleotide sequence of PtDGAT1 was missing the 63
bp insert (insert II) which can be a result of an intron retention alternatively splicing. This
mechanism, by which multiple forms of mature mRNAs are produced from a single
transcript, just after the transcripts synthesis, is the most common alternative splicing in
Moreover, the amino acid sequences resulting from both forms of mRNA show that the insert II introduces stop-codons into the ORF of the PtDGAT1 mRNA. Indeed, mRNAs with intron retention lead to truncated polypeptides, or are subjected to nonsense-mediated mRNA decay, as retained introns often introduce in frame stop-codons (Maquat, 2004). The production of a truncated transcript and protein product at the expense of the corresponding active enzyme form could play role in regulating the amount of active protein produced, depending on the level of correctly spliced transcript. A higher amount of the active DGAT1 protein produced could thus provide an explanation at molecular level for the TAG content increase in *P. tricornutum* and generally in microalgae induced by nitrogen starvation.

**EXAMPLE 5**

**Heterologous expression of PtDGAT1 in *S. cerevisiae***

To verify whether PtDGAT1 indeed encodes a protein with DGAT activity, the gene was expressed in a *S. cerevisiae* neutral lipid-deficient quadruple mutant strain H1246, containing knockouts of the *DGAl, LROI*, involved in the formation of TAG, *ARE1* and *ARE2* genes, involved in the formation of sterol esters. The yeast cells harboring the empty pYES2 vector were used as control, the BY742 strain and the H1246 as negative control. Moreover, the H1246 expressing the yeast *DGAl* gene was also used as positive control. The total lipid extracts of the yeast cells were prepared and then subjected to TLC. In the neutral lipid-deficient quadruple mutant strain H1246, TAG were not detected, whereas the ability to form TAG was confirmed in the wild type BY742 and the H1246 expressing the yeast *DGAl* gene strain, as positive controls (Figure 7). Upon expression of the PtDGAT1, a prominent spot corresponding to TAG appeared on the chromatograms of the lipid extract of the yeast mutant cells (H1246), showing a successful restoration of the TAG deficient phenotype by the PtDGAT1 activity (Figure 7). As expected, the expression of the PtDGAT1 long with the intron retention introducing stop-codons did not complement the mutation and did not restore the ability of yeast mutant strain to synthesize TAG. This result indicates that the mRNA with intron retention lead to production of truncated polypeptides and absence of DGAT activity.

**EXAMPLE 5**

**Formation of lipid bodies is restored upon expression of PtDGAT1**

The activity of one of the knock-out genes is sufficient to restore formation of lipid bodies in the NL-deficient quadruple mutant strain H1246 (Sandager *et al.* 2002). To
determine if the heterologous expression of PtDGAT1 was able to restore the formation of oil bodies, the fluorescent dye Nil Red was used to stain the lipid bodies in the quadruple mutant H1246 after transformation and in the BY742 strain served as control. The results have shown that expression of PtDGAT1 restored the ability to form lipid bodies, as well as the expression of the yeast DGA1 gene (Fig. 8D vs. E). As expected no TAG formation was observed following expression of PtDGAT1 long containing the 63 bp insert and harboring stop-codons (Fig. 7C), since, the intron retention introduces in frame stop-codons and leads to the loss of DGAT activity.

EXAMPLE 6

Substrate fatty acid specificity of PtDGAT1

The predominant molecular species of TAG in P. tricomutum under starvation conditions are composed of 46:1, 48:1, 48:2, and 48:3 (total of carbon chains: total of double bonds) molecular species, having palmitic (16:0), palmitoleic (16:1), and myristic (14:0) acid constituents. However the alga can accumulate a certain percentage of VLC-PUFAs, EPA (20:5 n-3) and TAG molecular species with higher degree of unsaturation containing 20:4 n-6, 20:5 n-3 and 22:6 n-3, at lesser scale are also detectable (Yu et al., 2009).

Further it was tested whether the expression of the active form of PtDGAT1, isolated at nitrogen starvation conditions, could facilitate the incorporation of various PUFAs into the yeast TAG in comparison with the endogenous DGAT1 activity encoded by DGA1. Consequently, the supplementation assay (Table 3) was carried out by using various polyunsaturated C18- and C20- fatty acids of both ω3 and ω6 groups naturally present in P. tricomutum, but not in S. cerevisiae, to investigate the possible acyl-CoA preference of PtDGAT1. P. tricomutum utilizes multiple biosynthetic pathways and fatty acid intermediates for EPA biosynthesis (Arao et al., 1994). S. cerevisiae is able to import exogenous FAs from the medium and convert them to respective acyl-CoA derivatives to be incorporated in polar lipids and TAG, thus supplementation of the medium with FA can further enhance the TAG accumulation conferred by the recombinant DGAT gene (Siloto et al., 2009). In the absence of supplemented fatty acids, the recombinant DGATs utilized available endogenous DAG and acyl-CoA pool for the acylation at the third position of glycerol. As can be seen from the Table 3, PtDGAT1 showed a clear preference for endogenous saturated 16:0 and 18:0 species relative to the yeast DGA1 which was more selective towards monounsaturated 16:1 and 18:1. Thus TAG formed by action of PtDGAT1 is substantially more saturated relative to the one formed by the yeast DGA1.
Table 3 Fatty acid composition of TAG isolated from HI246 yeast cells transformed with pYES2 containing PtDGAT1 and yeast DGAI. Cultures were supplemented or not with 250 µM ALA, GLA, ETA, ARA, and EPA. Mean values (n=3) are expressed as percentage of total fatty acids in the TAG fraction.

<table>
<thead>
<tr>
<th>SFA/ MUFA*</th>
<th>Supp. FA</th>
<th>18:1 n-9</th>
<th>18:0</th>
<th>16:1 n-7</th>
<th>16:0</th>
<th>Gene</th>
<th>Substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.53</td>
<td>-</td>
<td>31.9</td>
<td>13.8</td>
<td>28.6</td>
<td>17.7</td>
<td>PtDG AT1</td>
<td>not supplemented</td>
</tr>
<tr>
<td>0.19</td>
<td>-</td>
<td>39.6</td>
<td>7.3</td>
<td>37.0</td>
<td>7.4</td>
<td>DAG1</td>
<td>ALA</td>
</tr>
<tr>
<td>1.03</td>
<td>12.7</td>
<td>19.4</td>
<td>14.9</td>
<td>21.3</td>
<td>27.0</td>
<td>PtDG AT1</td>
<td>ALA</td>
</tr>
<tr>
<td>0.45</td>
<td>18.4</td>
<td>22.9</td>
<td>8.3</td>
<td>29.6</td>
<td>15.3</td>
<td>PtDG AT1</td>
<td>GLA</td>
</tr>
<tr>
<td>1.36</td>
<td>19.5</td>
<td>14.0</td>
<td>13.4</td>
<td>17.3</td>
<td>29.1</td>
<td>PtDG AT1</td>
<td>ETA</td>
</tr>
<tr>
<td>0.53</td>
<td>30.2</td>
<td>17.8</td>
<td>7.6</td>
<td>23.8</td>
<td>14.3</td>
<td>DAG1</td>
<td>ETA</td>
</tr>
<tr>
<td>0.70</td>
<td>5.6</td>
<td>23.0</td>
<td>13.7</td>
<td>28.2</td>
<td>22.4</td>
<td>PtDG AT1</td>
<td>GLA</td>
</tr>
<tr>
<td>0.21</td>
<td>12.4</td>
<td>28.0</td>
<td>5.3</td>
<td>40.3</td>
<td>8.9</td>
<td>DAG1</td>
<td>GLA</td>
</tr>
<tr>
<td>0.85</td>
<td>3.2</td>
<td>21.7</td>
<td>13.7</td>
<td>26.7</td>
<td>27.5</td>
<td>PtDG AT1</td>
<td>ARA</td>
</tr>
<tr>
<td>0.25</td>
<td>9.2</td>
<td>29.5</td>
<td>6.5</td>
<td>38.9</td>
<td>10.3</td>
<td>DAG1</td>
<td>ARA</td>
</tr>
<tr>
<td>0.93</td>
<td>10.6</td>
<td>20.6</td>
<td>13.8</td>
<td>22.3</td>
<td>25.9</td>
<td>PtDG AT1</td>
<td>EPA</td>
</tr>
<tr>
<td>0.35</td>
<td>11.5</td>
<td>30.6</td>
<td>7.6</td>
<td>31.4</td>
<td>13.2</td>
<td>DAG1</td>
<td>EPA</td>
</tr>
</tbody>
</table>

ratio 16:0 + 18:0 / 16:1 + 18:1

The abilities of PtDGAT1 and the yeast DGAI to incorporate PUFA into TAG were compared. The fatty acid analysis of TAG showed incorporation of all supplemented PUFAs upon expression of both PtDGAT1 and DGAI. Surprisingly, the expression of yeast DGAI resulted in higher levels of both C18 and C20 PUFA in TAG, except for similar 20:5 n-3 incorporation, even if these fatty acids are not naturally present in yeast. Similar results have been shown by Wagner et al. (2010) by comparison of the Ostreococcus tauri DGAT2 and yeast DGAI substrate specificity, where algal and yeast enzymes were shown to be promiscuous towards available acyl-CoA substrate and displayed a similar fatty acid preference. The expression of both PtDGAT1 and DGAI resulted in higher incorporation of the C18-PUFA relative to C20-PUFA into TAG. Incorporation of 18:3 n-3 and 18:3 n-6 into TAG of the recombinant yeast was associated with a corresponding decrease in the proportion of both mono-unsaturated 16:1 n-7 and 18:1 n-9 fatty acids and an increase in the proportion of 16:0. Whereas TAG of the transformed yeasts supplemented with 20:3 n-3,
20:4 n-3, and 20:5 n-3 featured only a decrease in 18:1 n-9. From these data, it can be suggested that exogenous C18-PUFA competed with 16:1 n-7 and 18:1 n-9 in yeast expressing both recombinant PrDGAT1 and DGAI. C20-PUFA were mainly incorporated at the expense of endogenous 18:1 n-9. Importantly, TAG formed by the activity of PrDGAT1 was substantially more saturated, once more indicating to the preference for saturated fatty acid species. While cells expressing DGAI incorporated n-6 C20-PUFA into TAG at higher proportions than cells expressing PrDGAT1, the similar incorporation of 20:5 n-3 was determined upon expression of both P/DGAT1 and DGAI. These results confirm the ability of PrDGAT1 to accumulate 20:5 n-3, one major PUFA in diatoms, into TAG and allow speculating that PrDGAT1 prefers n-3 C20-PUFA over n-6 C20-PUFA. That was not a case with DGAI that demonstrated a similar incorporation of C20-PUFA of both groups.

In conclusion, these results have shown a preference of PrDGAT1 to produce TAG species with high level of saturated fatty acids (16:0 and 18:0) in a heterologous system. The highly unsaturated plant oils used for biodiesel production are often prone to oxidation more rapidly than conventional diesel, resulting in formation of insoluble sediments that interfere with engine performance (Deng et al., 2009). The capability of PrDGAT1 to incorporate saturated fatty acids such as 16:0 into TAG species is a beneficial feature for biodiesel production from microalgal and even plant oils, which are generally characterized by higher than necessary unsaturation level.

EXAMPLE 7

Cloning of the full-length iYDGAT1

The ORF for the full-length PrDGAT1 gene, including Pleckstrin Homology (PH) domain recently obtained from cDNA isolated from algae grown under nitrogen starvation was 2271 bp in length, coding for the corresponding protein of 756 amino acids (Figure 9). The putative full-length cDNA sequence, which contains the PH domain, is assumed to contain the genuine translation initiation codon, and the native C-terminal sequence encoding for the endoplasmatic reticulum retrieval motif of P. tricornutum. The 100-120 amino acids of PH domains are only found in eukaryotes, share little sequence conservation and have diverse functions. They are often involved in targeting proteins to the plasma membrane, but few display strong specificity in lipid binding (Lemmon M.A., 2011. Pleckstrin Homology (PH) domains. Transduction mechanisms in cellular signaling: cell signaling collection. Eds: Edward A. D., Ralph A. B., Vol. 3, 2nd edition, Elsevier inc., 239-247).
CLAIMS

What is claimed is:

1. An isolated protein comprising an amino acid sequence set forth in SEQ ID NO: 1.
2. The protein of claim 1 comprising an amino acid sequence set forth in SEQ ID NO: 3.
3. The protein of claim 1 comprising an amino acid sequence set forth in SEQ ID NO: 4.
4. A composition comprising the protein of claim 1 and a carrier.
5. An isolated polynucleotide comprising a coding portion encoding the protein of claim 1.
6. An isolated polynucleotide comprising a coding portion encoding the protein of claim 2.
7. An isolated polynucleotide comprising a coding portion encoding the protein of claim 3.
8. A composition comprising the polynucleotide of claim 5 and a carrier.
9. An expression vector comprising the polynucleotide of claim 5.
10. A cell comprising the expression vector of claim 5.
11. A transgenic organism or a transformed bacteria transformed by a polynucleotide of claim 5.
12. The organism of claim 11, wherein said organism is a plant, a seed, a heterotrophic microorganism, a bacterium, yeast, a microalga, or an alga.
13. A transgenic seed, produced by a transgenic plant transformed by a polynucleotide of claim 5.
14. A method for enhancing the production of oil or triacylglycerols with high saturated fatty acids content in a cell, comprising the step of transforming said cell with a polynucleotide of claim 5, thereby enhancing a production of oil or triacylglycerols with high saturated fatty acids content in a cell.

15. The method of claim 12, wherein said cell is a plant cell, a seed cell, a heterotrophic microorganism cell, a bacterium, yeast, a microalgal cell, or an algal cell.

16. The method of claim 13, wherein said algal or microalgal cell is grown under nitrogen starvation conditions.

17. A method of increasing the proportion of a saturated fatty acid in a cell, comprising the step of transforming said cell with a polynucleotide of claim 5, thereby increasing the proportion of a saturated fatty acid in a cell.

18. The method of claim 15, wherein said increasing the proportion of a saturated fatty acid comprises decreasing the proportion of an unsaturated fatty acid.

19. The method of claim 15, wherein said increasing the proportion of a saturated fatty acid comprises increasing the proportion of a saturated fatty acid in a triacylglycerol or oil.

20. The method of claim 15, wherein said cell is a plant cell, a seed cell, a heterotrophic microorganism cell, a bacterium, yeast, a microalgal cell, or an algal cell.

21. The method of claim 18, wherein said algal or microalgal cell is grown under nitrogen starvation conditions.

22. The method of claim 15, wherein said saturated fatty acid is 16:0 or 18:0.

23. The method of claim 16, wherein said unsaturated fatty acid is 16:1 or 18:1.
<table>
<thead>
<tr>
<th></th>
<th>XM_002177717.1</th>
<th>Chr2genDNA</th>
<th>Pt1DGAT1long</th>
<th>Pt1DGAT1</th>
<th>CT880495short</th>
<th>CT887168</th>
<th>CT881105</th>
</tr>
</thead>
<tbody>
<tr>
<td>640</td>
<td>ATCACCTCAA</td>
<td>ATCACCTCAA</td>
<td>ATCACCTCAA</td>
<td>ATCACCTCAA</td>
<td>ATCACCTCAA</td>
<td>ATCACCTCAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGAAGATCCC</td>
<td>GGAAGATCCC</td>
<td>GGAAGATCCC</td>
<td>GGAAGATCCC</td>
<td>GGAAGATCCC</td>
<td>GGAAGATCCC</td>
<td></td>
</tr>
<tr>
<td>680</td>
<td>TGGCAGGAAT</td>
<td>TGGCAGGAAT</td>
<td>TGGCAGGAAT</td>
<td>TGGCAGGAAT</td>
<td>TGGCAGGAAT</td>
<td>TGGCAGGAAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTCTTTTGT</td>
<td>TTCTTTTGT</td>
<td>TTCTTTTGT</td>
<td>TTCTTTTGT</td>
<td>TTCTTTTGT</td>
<td>TTCTTTTGT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TACGAGATT</td>
<td>TACGAGATT</td>
<td>TACGAGATT</td>
<td>TACGAGATT</td>
<td>TACGAGATT</td>
<td>TACGAGATT</td>
<td></td>
</tr>
<tr>
<td>720</td>
<td>CTCTCCAGC</td>
<td>CTCTCCAGC</td>
<td>CTCTCCAGC</td>
<td>CTCTCCAGC</td>
<td>CTCTCCAGC</td>
<td>CTCTCCAGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TCGTCCTTGT</td>
<td>TCGTCCTTGT</td>
<td>TCGTCCTTGT</td>
<td>TCGTCCTTGT</td>
<td>TCGTCCTTGT</td>
<td>TCGTCCTTGT</td>
<td></td>
</tr>
<tr>
<td>760</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>840</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consensus

ATCACCTCAA GGAAGATCCC TGGCAGGAAT TTCTTTTGT TACGAGATT CTCTCCAGC TCGTCCTTGT