

**(12) PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

**(11) Application No. AU 199869616 B2**  
**(10) Patent No. 726807**

(54) Title  
**Methods and compositions for synthesis of long chain polyunsaturated fatty acids**

(51)<sup>7</sup> International Patent Classification(s)  
**C12N 015/53 C12N 005/10**  
**A23L 001/30 C12N 009/02**  
**A61K 031/20 C12N 015/81**  
**C11B 001/00 C12P 007/64**  
**C12N 001/19**

(21) Application No: **199869616**

(22) Application Date: **1998.04.10**

(87) WIPO No: **WO98/46763**

(30) Priority Data

(31) Number	(32) Date	(33) Country
<b>08/834655</b>	<b>1997.04.11</b>	<b>US</b>

(43) Publication Date : **1998.11.11**

(43) Publication Journal Date : **1998.12.24**

(44) Accepted Journal Date : **2000.11.23**

(71) Applicant(s)  
**Calgene LLC; Abbott Laboratories**

(72) Inventor(s)  
**Deborah Knutzon; Pradip Mukerji; Yung-Sheng Huang; Jennifer Thurmond;**  
**Sunita Chaudhary; Amanda Eun-Yeong Leonard**

(74) Agent/Attorney  
**DAVIES COLLISON CAVE,GPO Box 3876,SYDNEY NSW 2001**

(56) Related Art  
**EP 561569**  
**WO 96/21022**  
**WO 94/18337**



69616/98

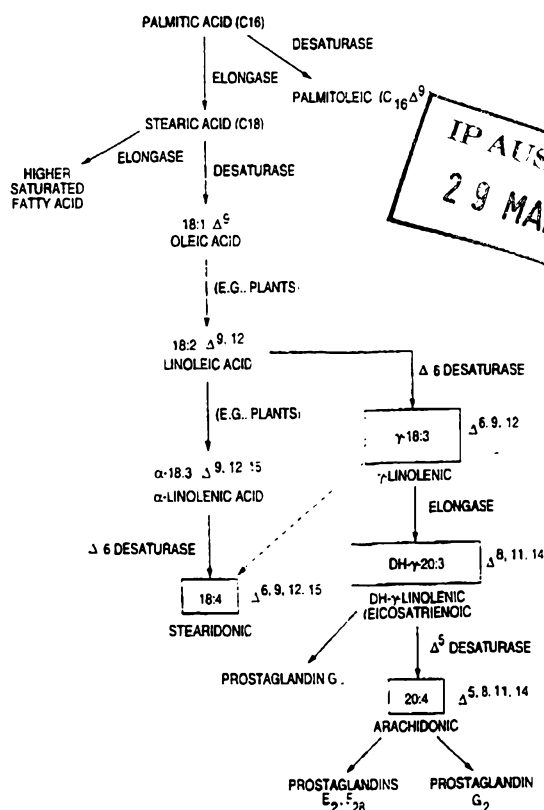
pages 1/17-17/17, drawings, replaced by new pages 1/20-20/20; due to  
INTERNATIONAL late transmittal by the receiving Office

OPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup>:</b> <b>C12N 15/53, 15/81, 9/02, 5/10, 1/19,</b> <b>C12P 7/64, C11B 1/00, A61K 31/20,</b> <b>A23L 1/30</b>		<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/46763</b>
			<b>(43) International Publication Date:</b> 22 October 1998 (22.10.98)
<b>(21) International Application Number:</b> PCT/US98/07126		<b>(74) Agents:</b> WARD, Michael, R. et al.; Limbach & Limbach L.L.P., 2001 Ferry Building, San Francisco, CA 94111-4262 (US).	
<b>(22) International Filing Date:</b> 10 April 1998 (10.04.98)		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
<b>(30) Priority Data:</b> 08/834,655      11 April 1997 (11.04.97)      US			
<b>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application</b> US      08/834,655 (CIP) Filed on      11 April 1997 (11.04.97)			
<b>(71) Applicants (for all designated States except US):</b> CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US). ABBOTT LABORATORIES [US/US]; 100 Abbott Park Road, Abbott Park, IL 60064-3500 (US).			
<b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> KNUTZON, Deborah [US/US]; 6110 Rockhurst Way, Granite Bay, CA 95746 (US). MUKERJI, Pradip [US/US]; 1069 Arcaro Drive, Ga- hanna, OH 43230 (US). HUANG, Yung-Sheng [CA/US]; 2462 Danvers Court, Upper Arlington, OH 43220 (US). THURMOND, Jennifer [US/US]; 3702 Adirondack, Colum-			

**Published***With international search report.**Before the expiration of the time limit for amending the  
claims and to be republished in the event of the receipt of  
amendments.***(54) Title:** METHODS AND COMPOSITIONS FOR SYNTHESIS OF LONG CHAIN POLYUNSATURATED FATTY ACIDS**(57) Abstract**

The present invention relates to fatty acid desaturases able to catalyze the conversion of oleic acid to linoleic acid, linoleic acid to  $\gamma$ -linolenic acid, or of  $\alpha$ -linolenic acid to stearidonic acid. Nucleic acid sequences encoding desaturases, nucleic acid sequences which hybridize thereto, DNA constructs comprising a desaturase gene, and recombinant host microorganism or animal expressing increased levels of a desaturase are described. Methods for desaturating a fatty acid and for producing a desaturated fatty acid by expressing increased levels of a desaturase are disclosed. Fatty acids, and oils containing them, which have been desaturated by a desaturase produced by recombinant host microorganisms or animals are provided. Pharmaceutical compositions, infant formulas or dietary supplements containing fatty acids which have been desaturated by a desaturase produced by a recombinant host microorganism or animal also are described.



## METHODS AND COMPOSITIONS FOR SYNTHESIS OF LONG CHAIN POLYUNSATURATED FATTY ACIDS

### RELATED APPLICATIONS

5           This application is a continuation-in-part application of United States  
Patent Application Serial No. 08/834,655 filed April 11, 1997.

### INTRODUCTION

#### Field of the Invention

10           This invention relates to modulating levels of enzymes and/or enzyme  
components relating to production of long chain poly-unsaturated fatty acids  
(PUFAs) in a microorganism or animal.

#### Background

15           Two main families of polyunsaturated fatty acids (PUFAs) are the  $\omega$ 3  
fatty acids, exemplified by eicosapentaenoic acid (EPA), and the  $\omega$ 6 fatty acids,  
exemplified by arachidonic acid (ARA). PUFAs are important components of  
the plasma membrane of the cell, where they may be found in such forms as  
phospholipids. PUFAs are necessary for proper development, particularly in the  
developing infant brain, and for tissue formation and repair. PUFAs also serve  
as precursors to other molecules of importance in human beings and animals,  
20           including the prostacyclins, eicosanoids, leukotrienes and prostaglandins. Four  
major long chain PUFAs of importance include docosahexaenoic acid (DHA)  
and EPA, which are primarily found in different types of fish oil,  $\gamma$ -linolenic  
acid (GLA), which is found in the seeds of a number of plants, including  
evening primrose (*Oenothera biennis*), borage (*Borago officinalis*) and black  
25           currants (*Ribes nigrum*), and stearidonic acid (SDA), which is found in marine  
oils and plant seeds. Both GLA and another important long chain PUFA,  
arachidonic acid (ARA), are found in filamentous fungi. ARA can be purified  
from animal tissues including liver and adrenal gland. GLA, ARA, EPA and

SDA are themselves, or are dietary precursors to, important long chain fatty acids involved in prostaglandin synthesis, in treatment of heart disease, and in development of brain tissue.

For DHA, a number of sources exist for commercial production including a variety of marine organisms, oils obtained from cold water marine fish, and egg yolk fractions. For ARA, microorganisms including the genera *Mortierella*, *Entomophthora*, *Phytium* and *Porphyridium* can be used for commercial production. Commercial sources of SDA include the genera *Trichodesma* and *Echium*. Commercial sources of GLA include evening primrose, black currants and borage. However, there are several disadvantages associated with commercial production of PUFAs from natural sources. Natural sources of PUFAs, such as animals and plants, tend to have highly heterogeneous oil compositions. The oils obtained from these sources therefore can require extensive purification to separate out one or more desired PUFAs or to produce an oil which is enriched in one or more PUFA. Natural sources also are subject to uncontrollable fluctuations in availability. Fish stocks may undergo natural variation or may be depleted by overfishing. Fish oils have unpleasant tastes and odors, which may be impossible to economically separate from the desired product, and can render such products unacceptable as food supplements. Animal oils, and particularly fish oils, can accumulate environmental pollutants. Weather and disease can cause fluctuation in yields from both fish and plant sources. Cropland available for production of alternate oil-producing crops is subject to competition from the steady expansion of human populations and the associated increased need for food production on the remaining arable land. Crops which do produce PUFAs, such as borage, have not been adapted to commercial growth and may not perform well in monoculture. Growth of such crops is thus not economically competitive where more profitable and better established crops can be grown. Large scale fermentation of organisms such as *Mortierella* is also expensive. Natural animal tissues contain low amounts of ARA and are difficult to process. Microorganisms such as *Porphyridium* and *Mortierella* are difficult to cultivate on a commercial scale.



Dietary supplements and pharmaceutical formulations containing PUFAs can retain the disadvantages of the PUFA source. Supplements such as fish oil capsules can contain low levels of the particular desired component and thus require large dosages. High dosages result in ingestion of high levels of  
5 undesired components, including contaminants. Unpleasant tastes and odors of the supplements can make such regimens undesirable, and may inhibit compliance by the patient. Care must be taken in providing fatty acid supplements, as overaddition may result in suppression of endogenous biosynthetic pathways and lead to competition with other necessary fatty acids  
10 in various lipid fractions *in vivo*, leading to undesirable results. For example, Eskimos having a diet high in  $\omega$ 3 fatty acids have an increased tendency to bleed (U.S. Pat. No. 4,874,603).

A number of enzymes are involved in PUFA biosynthesis. Linoleic acid (LA, 18:2  $\Delta$ 9, 12) is produced from oleic acid (18:1  $\Delta$ 9) by a  $\Delta$ 12-desaturase.  
15 GLA (18:3  $\Delta$ 6, 9, 12) is produced from linoleic acid (LA, 18:2  $\Delta$ 9, 12) by a  $\Delta$ 6-desaturase. ARA (20:4  $\Delta$ 5, 8, 11, 14) production from dihomo- $\gamma$ -linolenic acid (DGLA, 20:3  $\Delta$ 8, 11, 14) is catalyzed by a  $\Delta$ 5-desaturase. However, animals cannot desaturate beyond the  $\Delta$ 9 position and therefore cannot convert oleic acid (18:1  $\Delta$ 9) into linoleic acid (18:2  $\Delta$ 9, 12). Likewise,  $\alpha$ -linolenic acid  
20 (ALA, 18:3  $\Delta$ 9, 12, 15) cannot be synthesized by mammals. Other eukaryotes, including fungi and plants, have enzymes which desaturate at positions  $\Delta$ 12 and  $\Delta$ 15. The major poly-unsaturated fatty acids of animals therefore are either derived from diet and/or from desaturation and elongation of linoleic acid (18:2  $\Delta$ 9, 12) or  $\alpha$ -linolenic acid (18:3  $\Delta$ 9, 12, 15). Therefore it is of interest to obtain  
25 genetic material involved in PUFA biosynthesis from species that naturally produce these fatty acids and to express the isolated material in a microbial or animal system which can be manipulated to provide production of commercial quantities of one or more PUFAs. Thus there is a need for fatty acid desaturases, genes encoding them, and recombinant methods of producing them.  
30 A need further exists for oils containing higher relative proportions of and/or

enriched in specific PUFAs. A need also exists for reliable economical methods of producing specific PUFAs.

### **Relevant Literature**

Production of  $\gamma$ -linolenic acid by a  $\Delta 6$ -desaturase is described in USPN 5,552,306. Production of 8, 11-eicosadienoic acid using *Mortierella alpina* is disclosed in USPN 5,376,541. Production of docosaehxaenoic acid by dinoflagellates is described in USPN 5,407,957. Cloning of a  $\Delta 6$ -palmitoyl-acyl carrier protein desaturase is described in PCT publication WO 96/13591 and USPN 5,614,400. Cloning of a  $\Delta 6$ -desaturase from borage is described in PCT publication WO 96/21022. Cloning of  $\Delta 9$ -desaturases is described in the published patent applications PCT WO 91/13972, EP 0 550 162 A1, EP 0 561 569 A2, EP 0 644 263 A2, and EP 0 736 598 A1, and in USPN 5,057,419. Cloning of  $\Delta 12$ -desaturases from various organisms is described in PCT publication WO 94/11516 and USPN 5,443,974. Cloning of  $\Delta 15$ -desaturases from various organisms is described in PCT publication WO 93/11245. All publications and U.S. patents or applications referred to herein are hereby incorporated in their entirety by reference.



## STATEMENT OF INVENTION

In one aspect of the present invention there is provided a purified or isolated polypeptide  
5 which is capable of desaturating a fatty acid molecule at carbon 6 or 12 from the carboxyl  
end of said fatty acid, said polypeptide having an amino acid sequence which has at least  
60% homology to the 457 amino acid sequence of SEQ ID NO:2 or the 399 amino acid  
sequence of SEQ ID NO:4.

10

Preferably the polypeptide has an amino acid sequence which has at least 80% homology to  
the 457 amino acid sequence of SEQ ID NO: 2.

15 Alternatively, the polypeptide has an amino acid sequence which has at least 80% homology  
to the 399 amino acid sequence of SEQ ID NO: 4.

More preferably the said polypeptide has an amino acid sequence which has at least 90%  
20 homology to the 457 amino acid sequence of SEQ ID NO: 2.

Alternatively, said polypeptide has an amino acid sequence which has at least 90%  
homology to the 399 amino acid sequence of SEQ ID NO: 4.

25

More preferably, said polypeptide includes an amino acid motif selected from the group  
consisting of residues 50-53, 39-43, 172-176, 204-213 and 390-402 of SEQ ID NO: 2.

30

More preferably, said polypeptide comprises residues 50-53, 39-43, 172-176, 204-213 and  
390-402 of SEQ ID NO: 2.



More preferably, said polypeptide comprises SEQ ID NO: 2.

5 In another aspect of the present invention there is provided an isolated nucleic acid encoding a polypeptide as herein described.

Preferably, said isolated nucleic acid comprises SEQ ID NO: 1 or SEQ ID NO: 3.

10

In a further aspect of the present invention there is provided a nucleic acid construct comprising said nucleic acid as defined herein operably linked to a promoter.

15

In yet a further aspect of the present invention there is provided a host cell transformed with said construct.

20 Preferably said host cell is a microbial host cell.

More preferably said host cell is a yeast cell.

25

In a further aspect of the present invention there is provided a method for the production of fatty acid gamma linolenic acid which method comprises:

30 growing a culture of said host cells which cells produce a delta 6 desaturase in the presence of linoleic acid, under conditions wherein said acid is converted to gamma linolenic acid by the expression of the polypeptide of said desaturase; and recovering the fatty acid gamma linolenic acid from the culture.



Alternatively, the present invention provides a method for the production of fatty acid  
linolenic acid which method comprises:

5

Growing a culture of said host cells which cells produce a delta 12 desaturase in the  
presence of oleic acid, under conditions wherein said acid is converted to linolenic acid by  
the expression of the polypeptide of said desaturase; and recovering the fatty acid linolenic  
10 acid from the culture.

Preferably said methods of production of the fatty acids gamma linolenic acid for linolenic  
acid further comprise formulating the fatty acid into a product selected from the group:

15

a pharmaceutical composition comprising said oil and a pharmaceutically  
acceptable carrier;  
a nutritional formula;  
an infant formula;  
20 a dietary supplement;  
a dietary substitute;  
a cosmetic; and  
an animal feed

25

Preferably said infant formula, dietary supplement or dietary substitute is in the form of a  
liquid or a solid.

30 Preferably, the nutritional formula, infant formula, dietary supplement or dietary substitute  
contains at least one macronutrient selected from the group consisting of coconut oil, soy  
oil, canola oil, mono- and di-glycerides, glucose, edible lactose, electrodialysed whey,  
electrodialysed skimmed milk, milk whey, soy protein, and other protein hydrolysates.



More preferably, said nutritional formula, infant formula, dietary supplement or dietary substitute further contains at least one vitamin selected from the group consisting of A, C, D,  
5 E and B complex; and at least one mineral selected from the group consisting of calcium, magnesium, zinc, manganese, sodium, potassium, phosphorus, copper, chloride, iodine, selenium and iron.

10 In yet a further aspect of the present invention there is provided the fatty acid gamma linolenic acid derived from said host cells as herein described.

In a further aspect of the present invention there is provided the fatty acid linoleic acid  
15 derived from said host cells as herein described.

In yet a further aspect of the present invention there is provided a product comprising the fatty acid gamma linolenic acid or the fatty acid linoleic acid selected from the group:

20

a pharmaceutical composition comprising said oil and a pharmaceutically acceptable carrier;  
a nutritional formula;  
an infant formula;  
25 a dietary supplement;  
a dietary substitute;  
a cosmetic; and  
an animal feed.

30 Preferably said infant formula, dietary supplement or dietary substitute is in the form of a liquid or a solid.



Preferably, the nutritional formula, infant formula, dietary supplement or dietary substitute contains at least one macronutrient selected from the group consisting of coconut oil, soy oil, canola oil. Mono- and di-glycerides, glucose, edible lactose, electrolysed whey, electrolysed skimmed milk, milk whey, soy protein, and other protein hydrolysates.

5

More preferably said nutritional formula, infant formula, dietary supplement or dietary substitute further contains at least one vitamin selected from the group consisting of vitamin A, C, D, E and b complex; and at least one mineral selected from the group consisting of  
10 calcium, magnesium, zinc, manganese, sodium, potassium, phosphorus, copper, chloride, iodine, selenium and iron.

In yet a further aspect of the present invention there is provided the use of a said microbial  
15 host cell as herein described for the production of the fatty acid gamma linolenic acid or the fatty acid linoleic acid.

In a further aspect of the present invention there is provided the use of a microbial host cell  
20 for the production of a product comprising said fatty acid, said product being selected from the group:

- a pharmaceutical composition comprising said oil and a pharmaceutically acceptable carrier;
- 25 a nutritional formula;
- an infant formula;
- a dietary supplement;
- a dietary substitute;
- a cosmetic; and
- 30 an animal feed.





### **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 shows possible pathways for the synthesis of arachidonic acid (20:4  $\Delta$ 5, 8, 11, 14) and stearidonic acid (18:4  $\Delta$ 6, 9, 12, 15) from palmitic acid ( $C_{16}$ ) from a variety of organisms, including algae, *Mortierella* and humans. These PUFAs can serve as precursors to other molecules important for humans and other animals, including prostacyclins, leukotrienes, and prostaglandins, some of which are shown.

Figure 2 shows possible pathways for production of PUFAs in addition to ARA, including EPA and DHA, again compiled from a variety of organisms.

Figure 3A-E shows the DNA sequence of the *Mortierella alpina*  $\Delta$ 6-desaturase and the deduced amino acid sequence:

Figure 3A-E (SEQ ID NO 1  $\Delta$ 6 DESATURASE cDNA)





Figure 3A-E (SEQ ID NO 2  $\Delta 6$  DESATURASE AMINO ACID)

Figure 4 shows an alignment of a portion of the *Mortierella alpina*  $\Delta 6$ -desaturase amino acid sequence with other related sequences.

5 Figure 5A-D shows the DNA sequence of the *Mortierella alpina*  $\Delta 12$ -desaturase and the deduced amino acid sequence:

Figure 5A-D (SEQ ID NO 3  $\Delta 12$  DESATURASE cDNA)

Figure 5A-D (SEQ ID NO 4  $\Delta 12$  DESATURASE AMINO ACID).

Figures 6A and 6B show the effect of different expression constructs on expression of GLA in yeast.

10 Figures 7A and 7B show the effect of host strain on GLA production.

Figures 8A and 8B show the effect of temperature on GLA production in *S. cerevisiae* strain SC334.

Figure 9 shows alignments of the protein sequence of the Ma 29 and contig 253538a.

15 Figure 10 shows alignments of the protein sequence of Ma 524 and contig 253538a.

### **BRIEF DESCRIPTION OF THE SEQUENCE LISTINGS**

SEQ ID NO:1 shows the DNA sequence of the *Mortierella alpina*  $\Delta 6$ -desaturase.

20 SEQ ID NO:2 shows the protein sequence of the *Mortierella alpina*  $\Delta 6$ -desaturase.

SEQ ID NO:3 shows the DNA sequence of the *Mortierella alpina*  $\Delta 12$ -desaturase.

25 SEQ ID NO:4 shows the protein sequence of the *Mortierella alpina*  $\Delta 12$ -desaturase.

SEQ ID NO:5-11 show various desaturase sequences.

SEQ ID NO:13-18 show various PCR primer sequences.

SEQ ID NO:19 and SEQ ID NO:20 show the nucleotide and amino acid sequence of a *Dictyostelium discoideum* desaturase.

5 SEQ ID NO:21 and SEQ ID NO:22 show the nucleotide and amino acid sequence of a *Phaeodactylum tricornutum* desaturase.

SEQ ID NO:23-26 show the nucleotide and deduced amino acid sequence of a *Schizochytrium* cDNA clone.

SEQ ID NO: 27-33 show nucleotide sequences for human desaturases.

10 SEQ ID NO:34 - SEQ ID NO:40 show peptide sequences for human desaturases.

### **DESCRIPTION OF THE PREFERRED EMBODIMENTS**

In order to ensure a complete understanding of the invention, the following definitions are provided:

15  **$\Delta$ 5-Desaturase:**  $\Delta$ 5 desaturase is an enzyme which introduces a double bond between carbons 5 and 6 from the carboxyl end of a fatty acid molecule.

**$\Delta$ 6-Desaturase:**  $\Delta$ 6-desaturase is an enzyme which introduces a double bond between carbons 6 and 7 from the carboxyl end of a fatty acid molecule.

**$\Delta$ 9-Desaturase:**  $\Delta$ 9-desaturase is an enzyme which introduces a double bond between carbons 9 and 10 from the carboxyl end of a fatty acid molecule.

20  **$\Delta$ 12-Desaturase:**  $\Delta$ 12-desaturase is an enzyme which introduces a double bond between carbons 12 and 13 from the carboxyl end of a fatty acid molecule.

25 **Fatty Acids:** Fatty acids are a class of compounds containing a long hydrocarbon chain and a terminal carboxylate group. Fatty acids include the following:

Fatty Acid		
12:0	lauric acid	
16:0	palmitic acid	

Fatty Acid		
16:1	palmitoleic acid	
18:0	stearic acid	
18:1	oleic acid	$\Delta 9$ -18:1
18:2 $\Delta 5,9$	taxoleic acid	$\Delta 5,9$ -18:2
18:2 $\Delta 6,9$	6,9-octadecadienoic acid	$\Delta 6,9$ -18:2
18:2	Linolenic acid	$\Delta 9,12$ -18:2 (LA)
18:3 $\Delta 6,9,12$	Gamma-linolenic acid	$\Delta 6,9,12$ -18:3 (GLA)
18:3 $\Delta 5,9,12$	Pinolenic acid	$\Delta 5,9,12$ -18:3
18:3	alpha-linoleic acid	$\Delta 9,12,15$ -18:3 (ALA)
18:4	stearidonic acid	$\Delta 6,9,12,15$ -18:4 (SDA)
20:0	Arachidic acid	
20:1	Eicosenic Acid	
22:0	behehic acid	
22:1	erucic acid	
22:2	docasadienoic acid	
20:4 $\omega 6$	arachidonic acid	$\Delta 5,8,11,14$ -20:4 (ARA)
20:3 $\omega 6$	$\omega 6$ -eicosatrienoic dihomo-gamma linolenic	$\Delta 8,11,14$ -20:3 (DGLA)
20:5 $\omega 3$	Eicosapentanoic (Timnodonic acid)	$\Delta 5,8,11,14,17$ -20:5 (EPA)
20:3 $\omega 3$	$\omega 3$ -eicosatrienoic	$\Delta 11,16,17$ -20:3
20:4 $\omega 3$	$\omega 3$ -eicosatetraenoic	$\Delta 8,11,14,17$ -20:4
22:5 $\omega 3$	Docosapentaenoic	$\Delta 7,10,13,16,19$ -22:5 ( $\omega 3$ DPA)
22:6 $\omega 3$	Docosahexaenoic (cervonic acid)	$\Delta 4,7,10,13,16,19$ -22:6 (DHA)
24:0	Lignoceric acid	

5 Taking into account these definitions, the present invention is directed to novel DNA sequences, DNA constructs, methods and compositions are provided which permit modification of the poly-unsaturated long chain fatty acid content of, for example, microbial cells or animals. Host cells are manipulated to express a sense or antisense transcript of a DNA encoding a polypeptide(s) which catalyzes the desaturation of a fatty acid. The substrate(s) for the expressed enzyme may be produced by the host cell or may be exogenously supplied. To achieve expression, the transformed DNA is

operably associated with transcriptional and translational initiation and termination regulatory regions that are functional in the host cell. Constructs comprising the gene to be expressed can provide for integration into the genome of the host cell or can autonomously replicate in the host cell. For production of linoleic acid (LA), the expression cassettes generally used include a cassette which provides for  $\Delta 12$ -desaturase activity, particularly in a host cell which produces or can take up oleic acid (U.S. Patent No. 5,443,974). Production of LA also can be increased by providing an expression cassette for a  $\Delta 9$ -desaturase where that enzymatic activity is limiting. For production of ALA, the expression cassettes generally used include a cassette which provides for  $\Delta 15$ - or  $\omega 3$ -desaturase activity, particularly in a host cell which produces or can take up LA. For production of GLA or SDA, the expression cassettes generally used include a cassette which provides for  $\Delta 6$ -desaturase activity, particularly in a host cell which produces or can take up LA or ALA, respectively. Production of  $\omega 6$ -type unsaturated fatty acids, such as LA or GLA, is favored in a host microorganism or animal which is incapable of producing ALA. The host ALA production can be removed, reduced and/or inhibited by inhibiting the activity of a  $\Delta 15$ - or  $\omega 3$ - type desaturase (see Figure 2). This can be accomplished by standard selection, providing an expression cassette for an antisense  $\Delta 15$  or  $\omega 3$  transcript, by disrupting a target  $\Delta 15$ - or  $\omega 3$ -desaturase gene through insertion, deletion, substitution of part or all of the target gene, or by adding an inhibitor of  $\Delta 15$ - or  $\omega 3$ -desaturase. Similarly, production of LA or ALA is favored in a microorganism or animal having  $\Delta 6$ -desaturase activity by providing an expression cassette for an antisense  $\Delta 6$  transcript, by disrupting a  $\Delta 6$ -desaturase gene, or by use of a  $\Delta 6$ -desaturase inhibitor.

### MICROBIAL PRODUCTION OF FATTY ACIDS

Microbial production of fatty acids has several advantages over purification from natural sources such as fish or plants. Many microbes are known with greatly simplified oil compositions compared with those of higher organisms, making purification of desired components easier. Microbial production is not subject to fluctuations caused by external variables such as

weather and food supply. Microbially produced oil is substantially free of contamination by environmental pollutants. Additionally, microbes can provide PUFAs in particular forms which may have specific uses. For example, *Spirulina* can provide PUFAs predominantly at the first and third positions of triglycerides; digestion by pancreatic lipases preferentially releases fatty acids from these positions. Following human or animal ingestion of triglycerides derived from *Spirulina*, these PUFAs are released by pancreatic lipases as free fatty acids and thus are directly available, for example, for infant brain development. Additionally, microbial oil production can be manipulated by controlling culture conditions, notably by providing particular substrates for microbially expressed enzymes, or by addition of compounds which suppress undesired biochemical pathways. In addition to these advantages, production of fatty acids from recombinant microbes provides the ability to alter the naturally occurring microbial fatty acid profile by providing new synthetic pathways in the host or by suppressing undesired pathways, thereby increasing levels of desired PUFAs, or conjugated forms thereof, and decreasing levels of undesired PUFAs.

### PRODUCTION OF FATTY ACIDS IN ANIMALS

Production of fatty acids in animals also presents several advantages. Expression of desaturase genes in animals can produce greatly increased levels of desired PUFAs in animal tissues, making recovery from those tissues more economical. For example, where the desired PUFAs are expressed in the breast milk of animals, methods of isolating PUFAs from animal milk are well established. In addition to providing a source for purification of desired PUFAs, animal breast milk can be manipulated through expression of desaturase genes, either alone or in combination with other human genes, to provide animal milks substantially similar to human breast milk during the different stages of infant development. Humanized animal milks could serve as infant formulas where human nursing is impossible or undesired, or in cases of malnourishment or disease.

Depending upon the host cell, the availability of substrate, and the desired end product(s), several polypeptides, particularly desaturases, are of

interest. By "desaturase" is intended a polypeptide which can desaturate one or more fatty acids to produce a mono- or poly-unsaturated fatty acid or precursor thereof of interest. Of particular interest are polypeptides which can catalyze the conversion of stearic acid to oleic acid, of oleic acid to LA, of LA to ALA, of LA to GLA, or of ALA to SDA, which includes enzymes which desaturate at the  $\Delta 9$ ,  $\Delta 12$ , ( $\omega 6$ ),  $\Delta 15$ , ( $\omega 3$ ) or  $\Delta 6$  positions. By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification, for example, glycosylation or phosphorylation. Considerations for choosing a specific polypeptide having desaturase activity include the pH optimum of the polypeptide, whether the polypeptide is a rate limiting enzyme or a component thereof, whether the desaturase used is essential for synthesis of a desired poly-unsaturated fatty acid, and/or co-factors required by the polypeptide. The expressed polypeptide preferably has parameters compatible with the biochemical environment of its location in the host cell. For example, the polypeptide may have to compete for substrate with other enzymes in the host cell. Analyses of the  $K_m$  and specific activity of the polypeptide in question therefore are considered in determining the suitability of a given polypeptide for modifying PUFA production in a given host cell. The polypeptide used in a particular situation is one which can function under the conditions present in the intended host cell but otherwise can be any polypeptide having desaturase activity which has the desired characteristic of being capable of modifying the relative production of a desired PUFA.

For production of linoleic acid from oleic acid, the DNA sequence used encodes a polypeptide having  $\Delta 12$ -desaturase activity. For production of GLA from linoleic acid, the DNA sequence used encodes a polypeptide having  $\Delta 6$ -desaturase activity. In particular instances, expression of  $\Delta 6$ -desaturase activity can be coupled with expression of  $\Delta 12$ -desaturase activity and the host cell can optionally be depleted of any  $\Delta 15$ -desaturase activity present, for example by providing a transcription cassette for production of antisense sequences to the  $\Delta 15$ -desaturase transcription product, by disrupting the  $\Delta 15$ -desaturase gene, or by using a host cell which naturally has, or has been mutated to have, low  $\Delta 15$ -desaturase activity. Inhibition of undesired desaturase pathways also can be

accomplished through the use of specific desaturase inhibitors such as those described in U.S. Patent No. 4,778,630. Also, a host cell for  $\Delta 6$ -desaturase expression may have, or have been mutated to have, high  $\Delta 12$ -desaturase activity. The choice of combination of cassettes used depends in part on the PUFA profile and/or desaturase profile of the host cell. Where the host cell expresses  $\Delta 12$ -desaturase activity and lacks or is depleted in  $\Delta 15$ -desaturase activity, overexpression of  $\Delta 6$ -desaturase alone generally is sufficient to provide for enhanced GLA production. Where the host cell expresses  $\Delta 9$ -desaturase activity, expression of a  $\Delta 12$ - and a  $\Delta 6$ -desaturase can provide for enhanced GLA production. When  $\Delta 9$ -desaturase activity is absent or limiting, an expression cassette for  $\Delta 9$ -desaturase can be used. A scheme for the synthesis of arachidonic acid ( $20:4 \Delta^{5,8,11,14}$ ) from stearic acid ( $18:0$ ) is shown in Figure 2. A key enzyme in this pathway is a  $\Delta 6$ -desaturase which converts the linoleic acid into  $\gamma$ -linolenic acid. Conversion of  $\alpha$ -linolenic acid (ALA) to stearidonic acid by a  $\Delta 6$ -desaturase also is shown.

#### SOURCES OF POLYPEPTIDES HAVING DESATURASE ACTIVITY

A source of polypeptides having desaturase activity and oligonucleotides encoding such polypeptides are organisms which produce a desired poly-unsaturated fatty acid. As an example, microorganisms having an ability to produce GLA or ARA can be used as a source of  $\Delta 6$ - or  $\Delta 12$ - desaturase activity. Such microorganisms include, for example, those belonging to the genera *Mortierella*, *Conidiobolus*, *Pythium*, *Phytophthora*, *Penicillium*, *Porphyridium*, *Coidosporium*, *Mucor*, *Fusarium*, *Aspergillus*, *Rhodotorula*, and *Entomophthora*. Within the genus *Porphyridium*, of particular interest is *Porphyridium cruentum*. Within the genus *Mortierella*, of particular interest are *Mortierella elongata*, *Mortierella exigua*, *Mortierella hygrophila*, *Mortierella ramanniana*, var. *angulispora*, and *Mortierella alpina*. Within the genus *Mucor*, of particular interest are *Mucor circinelloides* and *Mucor javanicus*.

DNAs encoding desired desaturases can be identified in a variety of ways. As an example, a source of the desired desaturase, for example genomic

or cDNA libraries from *Mortierella*, is screened with detectable enzymatically- or chemically-synthesized probes, which can be made from DNA, RNA, or non-naturally occurring nucleotides, or mixtures thereof. Probes may be enzymatically synthesized from DNAs of known desaturases for normal or reduced-stringency hybridization methods. Oligonucleotide probes also can be used to screen sources and can be based on sequences of known desaturases, including sequences conserved among known desaturases, or on peptide sequences obtained from the desired purified protein. Oligonucleotide probes based on amino acid sequences can be degenerate to encompass the degeneracy of the genetic code, or can be biased in favor of the preferred codons of the source organism. Oligonucleotides also can be used as primers for PCR from reverse transcribed mRNA from a known or suspected source; the PCR product can be the full length cDNA or can be used to generate a probe to obtain the desired full length cDNA. Alternatively, a desired protein can be entirely sequenced and total synthesis of a DNA encoding that polypeptide performed.

Once the desired genomic or cDNA has been isolated, it can be sequenced by known methods. It is recognized in the art that such methods are subject to errors, such that multiple sequencing of the same region is routine and is still expected to lead to measurable rates of mistakes in the resulting deduced sequence, particularly in regions having repeated domains, extensive secondary structure, or unusual base compositions, such as regions with high GC base content. When discrepancies arise, resequencing can be done and can employ special methods. Special methods can include altering sequencing conditions by using: different temperatures; different enzymes; proteins which alter the ability of oligonucleotides to form higher order structures; altered nucleotides such as ITP or methylated dGTP; different gel compositions, for example adding formamide; different primers or primers located at different distances from the problem region; or different templates such as single stranded DNAs. Sequencing of mRNA also can be employed.

For the most part, some or all of the coding sequence for the polypeptide having desaturase activity is from a natural source. In some situations, however, it is desirable to modify all or a portion of the codons, for example, to



enhance expression, by employing host preferred codons. Host preferred codons can be determined from the codons of highest frequency in the proteins expressed in the largest amount in a particular host species of interest. Thus, the coding sequence for a polypeptide having desaturase activity can be synthesized in whole or in part. All or portions of the DNA also can be synthesized to remove any destabilizing sequences or regions of secondary structure which would be present in the transcribed mRNA. All or portions of the DNA also can be synthesized to alter the base composition to one more preferable in the desired host cell. Methods for synthesizing sequences and bringing sequences together are well established in the literature. *In vitro* mutagenesis and selection, site-directed mutagenesis, or other means can be employed to obtain mutations of naturally occurring desaturase genes to produce a polypeptide having desaturase activity *in vivo* with more desirable physical and kinetic parameters for function in the host cell, such as a longer half-life or a higher rate of production of a desired polyunsaturated fatty acid.

#### *Mortierella alpina* Desaturase

Of particular interest is the *Mortierella alpina*  $\Delta 6$ -desaturase, which has 457 amino acids and a predicted molecular weight of 51.8 kD; the amino acid sequence is shown in Figure 3. The gene encoding the *Mortierella alpina*  $\Delta 6$ -desaturase can be expressed in transgenic microorganisms or animals to effect greater synthesis of GLA from linoleic acid or of stearidonic acid from ALA. Other DNAs which are substantially identical to the *Mortierella alpina*  $\Delta 6$ -desaturase DNA, or which encode polypeptides which are substantially identical to the *Mortierella alpina*  $\Delta 6$ -desaturase polypeptide, also can be used. By substantially identical is intended an amino acid sequence or nucleic acid sequence exhibiting in order of increasing preference at least 60%, 80%, 90% or 95% homology to the *Mortierella alpina*  $\Delta 6$ -desaturase amino acid sequence or nucleic acid sequence encoding the amino acid sequence. For polypeptides, the length of comparison sequences generally is at least 16 amino acids, preferably at least 20 amino acids, or most preferably 35 amino acids. For nucleic acids, the length of comparison sequences generally is at least 50 nucleotides,

preferably at least 60 nucleotides, and more preferably at least 75 nucleotides, and most preferably, 110 nucleotides. Homology typically is measured using sequence analysis software, for example, the Sequence Analysis software package of the Genetics Computer Group, University of Wisconsin  
5 Biotechnology Center, 1710 University Avenue, Madison, Wisconsin 53705, MEGAlign (DNASTar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), and MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other  
10 modifications. Conservative substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine, and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (Kyte and  
15 Doolittle, *J. Mol. Biol.* 157: 105-132, 1982), or on the basis of the ability to assume similar polypeptide secondary structure (Chou and Fasman, *Adv. Enzymol.* 47: 45-148, 1978).

Also of interest is the *Mortierella alpina*  $\Delta 12$ -desaturase, the nucleotide and amino acid sequence of which is shown in Figure 5. The gene encoding the  
20 *Mortierella alpina*  $\Delta 12$ -desaturase can be expressed in transgenic microorganisms or animals to effect greater synthesis of LA from oleic acid. Other DNAs which are substantially identical to the *Mortierella alpina*  $\Delta 12$ -desaturase DNA, or which encode polypeptides which are substantially identical to the *Mortierella alpina*  $\Delta 12$ -desaturase polypeptide, also can be used.

### 25 Other Desaturases

Encompassed by the present invention are related desaturases from the same or other organisms. Such related desaturases include variants of the disclosed  $\Delta 6$ - or  $\Delta 12$ -desaturase naturally occurring within the same or different  
species of *Mortierella*, as well as homologues of the disclosed  $\Delta 6$ - or  $\Delta 12$ -  
30 desaturase from other species. Also included are desaturases which, although

not substantially identical to the *Mortierella alpina*  $\Delta 6$ - or  $\Delta 12$ -desaturase, desaturate a fatty acid molecule at carbon 6 or 12, respectively, from the carboxyl end of a fatty acid molecule, or at carbon 12 or 6 from the terminal methyl carbon in an 18 carbon fatty acid molecule. Related desaturases can be identified by their ability to function substantially the same as the disclosed desaturases; that is, are still able to effectively convert LA to GLA, ALA to SDA or oleic acid to LA. Related desaturases also can be identified by screening sequence databases for sequences homologous to the disclosed desaturases, by hybridization of a probe based on the disclosed desaturases to a library constructed from the source organism, or by RT-PCR using mRNA from the source organism and primers based on the disclosed desaturases. Such desaturases include those from humans, *Dictyostelium discoideum* and *Phaeodactylum tricornutum*.

The regions of a desaturase polypeptide important for desaturase activity can be determined through routine mutagenesis, expression of the resulting mutant polypeptides and determination of their activities. Mutants may include deletions, insertions and point mutations, or combinations thereof. A typical functional analysis begins with deletion mutagenesis to determine the N- and C-terminal limits of the protein necessary for function, and then internal deletions, insertions or point mutants are made to further determine regions necessary for function. Other techniques such as cassette mutagenesis or total synthesis also can be used. Deletion mutagenesis is accomplished, for example, by using exonucleases to sequentially remove the 5' or 3' coding regions. Kits are available for such techniques. After deletion, the coding region is completed by ligating oligonucleotides containing start or stop codons to the deleted coding region after 5' or 3' deletion, respectively. Alternatively, oligonucleotides encoding start or stop codons are inserted into the coding region by a variety of methods including site-directed mutagenesis, mutagenic PCR or by ligation onto DNA digested at existing restriction sites. Internal deletions can similarly be made through a variety of methods including the use of existing restriction sites in the DNA, by use of mutagenic primers via site directed mutagenesis or mutagenic PCR. Insertions are made through methods such as linker-scanning

mutagenesis, site-directed mutagenesis or mutagenic PCR. Point mutations are made through techniques such as site-directed mutagenesis or mutagenic PCR.

Chemical mutagenesis also can be used for identifying regions of a desaturase polypeptide important for activity. A mutated construct is expressed, and the ability of the resulting altered protein to function as a desaturase is assayed. Such structure-function analysis can determine which regions may be deleted, which regions tolerate insertions, and which point mutations allow the mutant protein to function in substantially the same way as the native desaturase. All such mutant proteins and nucleotide sequences encoding them are within the scope of the present invention.

### EXPRESSION OF DESATURASE GENES

Once the DNA encoding a desaturase polypeptide has been obtained, it is placed in a vector capable of replication in a host cell, or is propagated *in vitro* by means of techniques such as PCR or long PCR. Replicating vectors can include plasmids, phage, viruses, cosmids and the like. Desirable vectors include those useful for mutagenesis of the gene of interest or for expression of the gene of interest in host cells. The technique of long PCR has made *in vitro* propagation of large constructs possible, so that modifications to the gene of interest, such as mutagenesis or addition of expression signals, and propagation of the resulting constructs can occur entirely *in vitro* without the use of a replicating vector or a host cell.

For expression of a desaturase polypeptide, functional transcriptional and translational initiation and termination regions are operably linked to the DNA encoding the desaturase polypeptide. Expression of the polypeptide coding region can take place *in vitro* or in a host cell. 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.

### **Expression In Vitro**

*In vitro* expression can be accomplished, for example, by placing the coding region for the desaturase polypeptide in an expression vector designed for *in vitro* use and adding rabbit reticulocyte lysate and cofactors; labeled amino acids can be incorporated if desired. Such *in vitro* expression vectors may provide some or all of the expression signals necessary in the system used. These methods are well known in the art and the components of the system are commercially available. The reaction mixture can then be assayed directly for the polypeptide, for example by determining its activity, or the synthesized polypeptide can be purified and then assayed.

### **Expression In A Host Cell**

Expression in a host cell can be accomplished in a transient or stable fashion. Transient expression can occur 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. Transient expression also can be accomplished by inducing the activity of a regulatable promoter operably linked to the gene of interest, although such inducible systems frequently exhibit a low basal level of expression. Stable expression can be achieved by introduction of a construct that can integrate into the host genome or that autonomously replicates in the host cell. Stable expression of the gene of interest can be 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. When stable expression results from integration, integration of constructs can occur randomly within the host genome or can be targeted through the use of constructs containing regions of homology with the host genome sufficient to target recombination with the host locus. Where constructs are targeted to an endogenous locus, all or some of the transcriptional and translational regulatory regions can be provided by the endogenous locus.

When increased expression of the desaturase polypeptide in the source organism is desired, several methods can be employed. Additional genes encoding the desaturase polypeptide can be introduced into the host organism. Expression from the native desaturase locus also can be increased through homologous recombination, for example by inserting a stronger promoter into the host genome to cause increased expression, by removing destabilizing sequences from either the mRNA or the encoded protein by deleting that information from the host genome, or by adding stabilizing sequences to the mRNA (USPN 4,910,141).

When it is desirable to express more than one different gene, appropriate regulatory regions and expression methods, introduced genes can be propagated in the host cell through use of replicating vectors or by integration into the host genome. Where two or more genes are expressed from separate replicating vectors, it is desirable that each vector has a different means of replication. Each introduced construct, whether integrated or not, should have a different means of selection and should lack homology to the other constructs to maintain stable expression and prevent reassortment of elements among constructs. Judicious choices of regulatory regions, selection means and method of propagation of the introduced construct can be experimentally determined so that all introduced genes are expressed at the necessary levels to provide for synthesis of the desired products.

As an example, where the host cell is a yeast, transcriptional and translational regions functional in yeast cells are provided, particularly from the host species. The transcriptional initiation regulatory regions can be obtained, for example from genes in the glycolytic pathway, such as alcohol dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase (GPD), phosphoglucoisomerase, phosphoglycerate kinase, etc. or regulatable genes such as acid phosphatase, lactase, metallothionein, glucoamylase, etc. Any one of a number of regulatory sequences can be used in a particular situation, depending upon whether constitutive or induced transcription is desired, the particular efficiency of the promoter in conjunction with the open-reading frame of interest, the ability to join a strong promoter with a control region from a

different promoter which allows for inducible transcription, ease of construction, and the like. Of particular interest are promoters which are activated in the presence of galactose. Galactose-inducible promoters (GAL1, GAL7, and GAL10) have been extensively utilized for high level and regulated expression of protein in yeast (Lue *et al.*, *Mol. Cell. Biol.* Vol. 7, p. 3446, 1987; Johnston, *Microbiol. Rev.* Vol. 51, p. 458, 1987). Transcription from the GAL promoters is activated by the GAL4 protein, which binds to the promoter region and activates transcription when galactose is present. In the absence of galactose, the antagonist GAL80 binds to GAL4 and prevents GAL4 from activating transcription. Addition of galactose prevents GAL80 from inhibiting activation by GAL4.

Nucleotide sequences surrounding the translational initiation codon ATG have been found to affect expression in yeast cells. If the desired polypeptide is poorly expressed in yeast, the nucleotide sequences of exogenous genes can be modified to include an efficient yeast translation initiation sequence to obtain optimal gene expression. For expression in *Saccharomyces*, this can be done by site-directed mutagenesis of an inefficiently expressed gene by fusing it in-frame to an endogenous *Saccharomyces* gene, preferably a highly expressed gene, such as the lactase gene.

The termination region can be derived from the 3' region of the gene from which the initiation region was obtained or from a different gene. A large number of termination regions are known to and have been found to be satisfactory in a variety of hosts from the same and different genera and species. The termination region usually is selected more as a matter of convenience rather than because of any particular property. Preferably, the termination region is derived from a yeast gene, particularly *Saccharomyces*, *Schizosaccharomyces*, *Candida* or *Kluyveromyces*. The 3' regions of two mammalian genes,  $\gamma$  interferon and  $\alpha 2$  interferon, are also known to function in yeast.

## INTRODUCTION OF CONSTRUCTS INTO HOST CELLS

Constructs comprising the gene of interest may be introduced into a host cell by standard techniques. These techniques include transformation, protoplast fusion, lipofection, transfection, transduction, conjugation, infection, bolistic impact, electroporation, microinjection, scraping, or any other method which introduces the gene of interest into the host cell. Methods of transformation which are used include lithium acetate transformation (*Methods in Enzymology*, Vol. 194, p. 186-187, 1991). For convenience, a host cell which has been manipulated by any method to take up a DNA sequence or construct will be referred to as "transformed" or "recombinant" herein.

The subject host will have at least have one copy of the expression construct and may have two or more, depending upon whether the gene is integrated into the genome, amplified, or is present on an extrachromosomal element having multiple copy numbers. Where the subject host is a yeast, four principal types of yeast plasmid vectors can be used: Yeast Integrating plasmids (YIps), Yeast Replicating plasmids (YRps), Yeast Centromere plasmids (YCps), and Yeast Episomal plasmids (YEps). YIps lack a yeast replication origin and must be propagated as integrated elements in the yeast genome. YRps have a chromosomally derived autonomously replicating sequence and are propagated as medium copy number (20 to 40), autonomously replicating, unstably segregating plasmids. YCps have both a replication origin and a centromere sequence and propagate as low copy number (10-20), autonomously replicating, stably segregating plasmids. YEps have an origin of replication from the yeast 2 $\mu$ m plasmid and are propagated as high copy number, autonomously replicating, irregularly segregating plasmids. The presence of the plasmids in yeast can be ensured by maintaining selection for a marker on the plasmid. Of particular interest are the yeast vectors pYES2 (a YEp plasmid available from Invitrogen, confers uracil prototrophy and a GAL1 galactose-inducible promoter for expression), pRS425-pG1 (a YEp plasmid obtained from Dr. T. H. Chang, Ass. Professor of Molecular Genetics, Ohio State University, containing a constitutive GPD promoter and conferring leucine prototrophy), and pYX424 (a YEp plasmid having a constitutive TP1 promoter and conferring



leucine prototrophy; Alber, T. and Kawasaki, G. (1982). *J. Mol. & Appl. Genetics* 1: 419).

The transformed host cell can be identified by selection for a marker contained on the introduced construct. Alternatively, a separate marker  
5 construct may be introduced with the desired construct, as many transformation techniques introduce many DNA molecules into host cells. Typically, transformed hosts are selected for their ability to grow on selective media. Selective media may incorporate an antibiotic or lack a factor necessary for growth of the untransformed host, such as a nutrient or growth factor. An  
10 introduced marker gene therefor may confer antibiotic resistance, or encode an essential growth factor or enzyme, and permit growth on selective media when expressed in the transformed host. Selection of a transformed host can also occur when the expressed marker protein can be detected, either directly or indirectly. The marker protein may be expressed alone or as a fusion to another  
15 protein. The marker protein can be detected by its enzymatic activity; for example  $\beta$  galactosidase can convert the substrate X-gal to a colored product, and luciferase can convert luciferin to a light-emitting product. The marker protein can be detected by its light-producing or modifying characteristics; for example, the green fluorescent protein of *Aequorea victoria* fluoresces when  
20 illuminated with blue light. Antibodies can be used to detect the marker protein or a molecular tag on, for example, a protein of interest. Cells expressing the marker protein or tag can be selected, for example, visually, or by techniques such as FACS or panning using antibodies. For selection of yeast transformants, any marker that functions in yeast may be used. Desirably,  
25 resistance to kanamycin and the amino glycoside G418 are of interest, as well as ability to grow on media lacking uracil, leucine, lysine or tryptophan.

Of particular interest is the  $\Delta 6$ - and  $\Delta 12$ -desaturase-mediated production of PUFAs in prokaryotic and eukaryotic host cells. Prokaryotic cells of interest include *Eschericia*, *Bacillus*, *Lactobacillus*, *cyanobacteria* and the like.  
30 Eukaryotic cells include mammalian cells such as those of lactating animals, avian cells such as of chickens, and other cells amenable to genetic manipulation including insect, fungal, and algae cells. The cells may be

cultured or formed as part or all of a host organism including an animal. Viruses and bacteriophage also may be used with the cells in the production of PUFAs, particularly for gene transfer, cellular targeting and selection. In a preferred embodiment, the host is any microorganism or animal which produces  
5 and/or can assimilate exogenously supplied substrate(s) for a  $\Delta 6$ - and/or  $\Delta 12$ -desaturase, and preferably produces large amounts of one or more of the substrates. Examples of host animals include mice, rats, rabbits, chickens, quail, turkeys, bovines, sheep, pigs, goats, yaks, etc., which are amenable to genetic manipulation and cloning for rapid expansion of the transgene expressing  
10 population. For animals, the desaturase transgene(s) can be adapted for expression in target organelles, tissues and body fluids through modification of the gene regulatory regions. Of particular interest is the production of PUFAs in the breast milk of the host animal.

### Expression In Yeast

15 Examples of host microorganisms include *Saccharomyces cerevisiae*, *Saccharomyces carlsbergensis*, or other yeast such as *Candida*, *Kluyveromyces* or other fungi, for example, filamentous fungi such as *Aspergillus*, *Neurospora*, *Penicillium*, etc. Desirable characteristics of a host microorganism are, for example, that it is genetically well characterized, can be used for high level  
20 expression of the product using ultra-high density fermentation, and is on the GRAS (generally recognized as safe) list since the proposed end product is intended for ingestion by humans. Of particular interest is use of a yeast, more particularly baker's yeast (*S. cerevisiae*), as a cell host in the subject invention. Strains of particular interest are SC334 (Mat  $\alpha$  pep4-3 prbl-1122 ura3-52 leu2-  
25 3, 112 reg1-501 gal1; *Gene* 83:57-64, 1989, Hovland P. *et al.*), YTC34 ( $\alpha$  ade2-101 his3 $\Delta$ 200 lys2-801 ura3-52; obtained from Dr. T. H. Chang, Ass. Professor of Molecular Genetics, Ohio State University), YTC41 ( $a/\alpha$  ura3-52/ura3=52 lys2-801/lys2-801 ade2-101/ade2-101 trp1- $\Delta$ 1/trp1- $\Delta$ 1 his3 $\Delta$ 200/his3 $\Delta$ 200  
leu2 $\Delta$ 1/leu2 $\Delta$ 1; obtained from Dr. T. H. Chang, Ass. Professor of Molecular  
30 Genetics, Ohio State University), BJ1995 (obtained from the Yeast Genetic

Stock Centre, 1021 Donner Laboratory, Berkeley, CA 94720), INVSC1 (Mat  $\alpha$  hiw3 $\Delta$ 1 leu2 trp1-289 ura3-52; obtained from Invitrogen, 1600 Faraday Ave., Carlsbad, CA 92008) and INVSC2 (Mat  $\alpha$  his3 $\Delta$ 200 ura3-167; obtained from Invitrogen).

5

### **Expression in Avian Species**

For producing PUFAs in avian species and cells, such as chickens, turkeys, quail and ducks, gene transfer can be performed by introducing a nucleic acid sequence encoding a  $\Delta$ 6 and/or  $\Delta$ 12-desaturase into the cells following procedures known in the art. If a transgenic animal is desired, pluripotent stem cells of embryos can be provided with a vector carrying a desaturase encoding transgene and developed into adult animal (USPN 5,162,215; Ono *et al.* (1996) *Comparative Biochemistry and Physiology A* 113(3):287-292; WO 9612793; WO 9606160). In most cases, the transgene will be modified to express high levels of the desaturase in order to increase production of PUFAs. The transgene can be modified, for example, by providing transcriptional and/or translational regulatory regions that function in avian cells, such as promoters which direct expression in particular tissues and egg parts such as yolk. The gene regulatory regions can be obtained from a variety of sources, including chicken anemia or avian leukosis viruses or avian genes such as a chicken ovalbumin gene.

20

### **Expression in Insect Cells**

Production of PUFAs in insect cells can be conducted using baculovirus expression vectors harboring one or more desaturase transgenes. Baculovirus expression vectors are available from several commercial sources such as Clonetech. Methods for producing hybrid and transgenic strains of algae, such as marine algae, which contain and express a desaturase transgene also are provided. For example, transgenic marine algae may be prepared as described in USPN 5,426,040. As with the other expression systems described above, the timing, extent of expression and activity of the desaturase transgene can be

25

regulated by fitting the polypeptide coding sequence with the appropriate transcriptional and translational regulatory regions selected for a particular use. Of particular interest are promoter regions which can be induced under preselected growth conditions. For example, introduction of temperature  
5 sensitive and/or metabolite responsive mutations into the desaturase transgene coding sequences, its regulatory regions, and/or the genome of cells into which the transgene is introduced can be used for this purpose.

The transformed host cell is grown under appropriate conditions adapted for a desired end result. For host cells grown in culture, the conditions are  
10 typically optimized to produce the greatest or most economical yield of PUFAs, which relates to the selected desaturase activity. Media conditions which may be optimized include: carbon source, nitrogen source, addition of substrate, final concentration of added substrate, form of substrate added, aerobic or anaerobic growth, growth temperature, inducing agent, induction temperature,  
15 growth phase at induction, growth phase at harvest, pH, density, and maintenance of selection. Microorganisms of interest, such as yeast are preferably grown in selected medium. For yeast, complex media such as peptone broth (YPD) or a defined media such as a minimal media (contains amino acids, yeast nitrogen base, and ammonium sulfate, and lacks a  
20 component for selection, for example uracil) are preferred. Desirably, substrates to be added are first dissolved in ethanol. Where necessary, expression of the polypeptide of interest may be induced, for example by including or adding galactose to induce expression from a GAL promoter.

### **Expression In Plants**

25 Production of PUFA's in plants can be conducted using various plant transformation systems such as the use of *Agrobacterium tumefaciens*, plant viruses, particle cell transformation and the like which are disclosed in Applicant's related applications U.S. Application Serial Nos. 08/834,033 and 08/956,985 and continuation-in-part applications filed simultaneously with this  
30 application all of which are hereby incorporated by reference.

### Expression In An Animal

Expression in cells of a host animal can likewise be accomplished in a transient or stable manner. Transient expression can be accomplished via known methods, for example infection or lipofection, and can be repeated in order to maintain desired expression levels of the introduced construct (*see* Ebert, PCT publication WO 94/05782). Stable expression can be accomplished via integration of a construct into the host genome, resulting in a transgenic animal. The construct can be introduced, for example, by microinjection of the construct into the pronuclei of a fertilized egg, or by transfection, retroviral infection or other techniques whereby the construct is introduced into a cell line which may form or be incorporated into an adult animal (U.S. Patent No. 4,873,191; U.S. Patent No. 5,530,177; U.S. Patent No. 5,565,362; U.S. Patent No. 5,366,894; Wilmut *et al* (1997) Nature 385:810). The recombinant eggs or embryos are transferred to a surrogate mother (U.S. Patent No. 4,873,191; U.S. Patent No. 5,530,177; U.S. Patent No. 5,565,362; U.S. Patent No. 5,366,894; Wilmut *et al* (*supra*)).

After birth, transgenic animals are identified, for example, by the presence of an introduced marker gene, such as for coat color, or by PCR or Southern blotting from a blood, milk or tissue sample to detect the introduced construct, or by an immunological or enzymological assay to detect the expressed protein or the products produced therefrom (U.S. Patent No. 4,873,191; U.S. Patent No. 5,530,177; U.S. Patent No. 5,565,362; U.S. Patent No. 5,366,894; Wilmut *et al* (*supra*)). The resulting transgenic animals may be entirely transgenic or may be mosaics, having the transgenes in only a subset of their cells. The advent of mammalian cloning, accomplished by fusing a nucleated cell with an enucleated egg, followed by transfer into a surrogate mother, presents the possibility of rapid, large-scale production upon obtaining a "founder" animal or cell comprising the introduced construct; prior to this, it was necessary for the transgene to be present in the germ line of the animal for propagation (Wilmut *et al* (*supra*)).

Expression in a host animal presents certain efficiencies, particularly where the host is a domesticated animal. For production of PUFAs in a fluid readily obtainable from the host animal, such as milk, the desaturase transgene can be expressed in mammary cells from a female host, and the PUFA content  
5 of the host cells altered. The desaturase transgene can be adapted for expression so that it is retained in the mammary cells, or secreted into milk, to form the PUFA reaction products localized to the milk (PCT publication WO 95/24488). Expression can be targeted for expression in mammary tissue using specific regulatory sequences, such as those of bovine  $\alpha$ -lactalbumin,  $\alpha$ -casein,  $\beta$ -  
10 casein,  $\gamma$ -casein,  $\kappa$ -casein,  $\beta$ -lactoglobulin, or whey acidic protein, and may optionally include one or more introns and/or secretory signal sequences (U.S. Patent No. 5,530,177; Rosen, U.S. Patent No. 5,565,362; Clark *et al.*, U.S. Patent No. 5,366,894; Garner *et al.*, PCT publication WO 95/23868). Expression of desaturase transgenes, or antisense desaturase transcripts, adapted  
15 in this manner can be used to alter the levels of specific PUFAs, or derivatives thereof, found in the animals milk. Additionally, the desaturase transgene(s) can be expressed either by itself or with other transgenes, in order to produce animal milk containing higher proportions of desired PUFAs or PUFA ratios and concentrations that resemble human breast milk (Prieto *et al.*, PCT  
20 publication WO 95/24494).

### PURIFICATION OF FATTY ACIDS

The desaturated fatty acids may be found in the host microorganism or animal as free fatty acids or in conjugated forms such as acylglycerols, phospholipids, sulfolipids or glycolipids, and may be extracted from the host  
25 cell through a variety of means well-known in the art. Such means may include extraction with organic solvents, sonication, supercritical fluid extraction using for example carbon dioxide, and physical means such as presses, or combinations thereof. Of particular interest is extraction with hexane or methanol and chloroform. Where desirable, the aqueous layer can be acidified  
30 to protonate negatively charged moieties and thereby increase partitioning of desired products into the organic layer. After extraction, the organic solvents can be removed by evaporation under a stream of nitrogen. When isolated in

conjugated forms, the products may be enzymatically or chemically cleaved to release the free fatty acid or a less complex conjugate of interest, and can then be subject to further manipulations to produce a desired end product. Desirably, conjugated forms of fatty acids are cleaved with potassium hydroxide.

5 If further purification is necessary, standard methods can be employed. Such methods may include extraction, treatment with urea, fractional crystallization, HPLC, fractional distillation, silica gel chromatography, high speed centrifugation or distillation, or combinations of these techniques. Protection of reactive groups, such as the acid or alkenyl groups, may be done at  
10 any step through known techniques, for example alkylation or iodination. Methods used include methylation of the fatty acids to produce methyl esters. Similarly, protecting groups may be removed at any step. Desirably, purification of fractions containing GLA, SDA, ARA, DHA and EPA may be accomplished by treatment with urea and/or fractional distillation.

## 15 USES OF FATTY ACIDS

The fatty acids of the subject invention finds many applications. Probes based on the DNAs of the present invention may find use in methods for isolating related molecules or in methods to detect organisms expressing desaturases. When used as probes, the DNAs or oligonucleotides must be detectable. This is usually accomplished by attaching a label either at an internal site, for example via incorporation of a modified residue, or at the 5' or 3' terminus. Such labels can be directly detectable, can bind to a secondary molecule that is detectably labeled, or can bind to an unlabelled secondary molecule and a detectably labeled tertiary molecule; this process can be extended as long as is practical to achieve a satisfactorily detectable signal without unacceptable levels of background signal. Secondary, tertiary, or bridging systems can include use of antibodies directed against any other molecule, including labels or other antibodies, or can involve any molecules which bind to each other, for example a biotin-streptavidin/avidin system. Detectable labels typically include radioactive isotopes, molecules which chemically or enzymatically produce or alter light, enzymes which produce

detectable reaction products, magnetic molecules, fluorescent molecules or molecules whose fluorescence or light-emitting characteristics change upon binding. Examples of labelling methods can be found in USPN 5,011,770. Alternatively, the binding of target molecules can be directly detected by measuring the change in heat of solution on binding of probe to target via isothermal titration calorimetry, or by coating the probe or target on a surface and detecting the change in scattering of light from the surface produced by binding of target or probe, respectively, as may be done with the BIAcore system.

PUFAs produced by recombinant means find applications in a wide variety of areas. Supplementation of animals or humans with PUFAs in various forms can result in increased levels not only of the added PUFAs but of their metabolic progeny as well.

#### NUTRITIONAL COMPOSITIONS

The present invention also includes nutritional compositions. Such compositions, for purposes of the present invention, include any food or preparation for human consumption including for enteral or parenteral consumption, which when taken into the body (a) serve to nourish or build up tissues or supply energy and/or (b) maintain, restore or support adequate nutritional status or metabolic function.

The nutritional composition of the present invention comprises at least one oil or acid produced in accordance with the present invention and may either be in a solid or liquid form. Additionally, the composition may include edible macronutrients, vitamins and minerals in amounts desired for a particular use. The amount of such ingredients will vary depending on whether the composition is intended for use with normal, healthy infants, children or adults having specialized needs such as those which accompany certain metabolic conditions (e.g., metabolic disorders).

Examples of macronutrients which may be added to the composition include but are not limited to edible fats, carbohydrates and proteins. Examples of such edible fats include but are not limited to coconut oil, soy oil, and mono-



and diglycerides. Examples of such carbohydrates include but are not limited to glucose, edible lactose and hydrolyzed starch. Additionally, examples of proteins which may be utilized in the nutritional composition of the invention include but are not limited to soy proteins, electrodialysed whey ,  
5 electrodialysed skim milk, milk whey, or the hydrolysates of these proteins.

With respect to vitamins and minerals, the following may be added to the nutritional compositions of the present invention: calcium, phosphorus, potassium, sodium, chloride, magnesium, manganese, iron, copper, zinc, selenium, iodine, and Vitamins A, E, D, C, and the B complex. Other such  
10 vitamins and minerals may also be added.

The components utilized in the nutritional compositions of the present invention will be of semi-purified or purified origin. By semi-purified or purified is meant a material which has been prepared by purification of a natural material or by synthesis.

15 Examples of nutritional compositions of the present invention include but are not limited to infant formulas, dietary supplements, and rehydration compositions. Nutritional compositions of particular interest include but are not limited to those utilized for enteral and parenteral supplementation for infants, specialist infant formulae, supplements for the elderly, and supplements for  
20 those with gastrointestinal difficulties and/or malabsorption.

### **Nutritional Compositions**

A typical nutritional composition of the present invention will contain edible macronutrients, vitamins and minerals in amounts desired for a particular use. The amounts of such ingredients will vary depending on whether the  
25 formulation is intended for use with normal, healthy individuals temporarily exposed to stress, or to subjects having specialized needs due to certain chronic or acute disease states (e.g., metabolic disorders). It will be understood by persons skilled in the art that the components utilized in a nutritional formulation of the present invention are of semi-purified or purified origin. By  
30 semi-purified or purified is meant a material that has been prepared by

purification of a natural material or by synthesis. These techniques are well known in the art (See, e.g., Code of Federal Regulations for Food Ingredients and Food Processing; Recommended Dietary Allowances, 10<sup>th</sup> Ed., National Academy Press, Washington, D.C., 1989).

5           In a preferred embodiment, a nutritional formulation of the present invention is an enteral nutritional product, more preferably an adult or child enteral nutritional product. Accordingly in a further aspect of the invention, a nutritional formulation is provided that is suitable for feeding adults or children, who are experiencing stress. The formula comprises, in addition to the PUFAs  
10 of the invention; macronutrients, vitamins and minerals in amounts designed to provide the daily nutritional requirements of adults.

          The macronutritional components include edible fats, carbohydrates and proteins. Exemplary edible fats are coconut oil, soy oil, and mono- and diglycerides and the PUFA oils of this invention. Exemplary carbohydrates are  
15 glucose, edible lactose and hydrolyzed cornstarch. A typical protein source would be soy protein, electrodialysed whey or electrodialysed skim milk or milk whey, or the hydrolysates of these proteins, although other protein sources are also available and may be used. These macronutrients would be added in the form of commonly accepted nutritional compounds in amount equivalent to  
20 those present in human milk or an energy basis, i.e., on a per calorie basis.

          Methods for formulating liquid and enteral nutritional formulas are well known in the art and are described in detail in the examples.

          The enteral formula can be sterilized and subsequently utilized on a ready-to-feed (RTF) basis or stored in a concentrated liquid or a powder. The  
25 powder can be prepared by spray drying the enteral formula prepared as indicated above, and the formula can be reconstituted by rehydrating the concentrate. Adult and infant nutritional formulas are well known in the art and commercially available (e.g., Similac®, Ensure®, Jevity® and Alimentum® from Ross Products Division, Abbott Laboratories). An oil or acid of the  
30 present invention can be added to any of these formulas in the amounts described below.

The energy density of the nutritional composition when in liquid form, can typically range from about 0.6 to 3.0 Kcal per ml. When in solid or powdered form, the nutritional supplement can contain from about 1.2 to more than 9 Kcals per gm, preferably 3 to 7 Kcals per gm. In general, the osmolality of a liquid product should be less than 700 mOsm and more preferably less than 660 mOsm.

The nutritional formula would typically include vitamins and minerals, in addition to the PUFAs of the invention, in order to help the individual ingest the minimum daily requirements for these substances. In addition to the PUFAs listed above, it may also be desirable to supplement the nutritional composition with zinc, copper, and folic acid in addition to antioxidants. It is believed that these substances will also provide a boost to the stressed immune system and thus will provide further benefits to the individual. The presence of zinc, copper or folic acid is optional and is not required in order to gain the beneficial effects on immune suppression. Likewise a pharmaceutical composition can be supplemented with these same substances as well.

In a more preferred embodiment, the nutritional contains, in addition to the antioxidant system and the PUFA component, a source of carbohydrate wherein at least 5 weight % of said carbohydrate is an indigestible oligosaccharide. In yet a more preferred embodiment, the nutritional composition additionally contains protein, taurine and carnitine.

The PUFAs, or derivatives thereof, made by the disclosed method can be used as dietary substitutes, or supplements, particularly infant formulas, for patients undergoing intravenous feeding or for preventing or treating malnutrition. Typically, human breast milk has a fatty acid profile comprising from about 0.15 % to about 0.36 % as DHA, from about 0.03 % to about 0.13 % as EPA, from about 0.30 % to about 0.88 % as ARA, from about 0.22 % to about 0.67 % as DGLA, and from about 0.27 % to about 1.04 % as GLA. Additionally, the predominant triglyceride in human milk has been reported to be 1,3-di-oleoyl-2-palmitoyl, with 2-palmitoyl glycerides reported as better absorbed than 2-oleoyl or 2-lineoyl glycerides (USPN 4,876,107). Thus, fatty acids such as ARA, DGLA, GLA and/or EPA produced by the invention can be

used to alter the composition of infant formulas to better replicate the PUFA composition of human breast milk. In particular, an oil composition for use in a pharmacologic or food supplement, particularly a breast milk substitute or supplement, will preferably comprise one or more of ARA, DGLA and GLA.

5 More preferably the oil will comprise from about 0.3 to 30% ARA, from about 0.2 to 30% DGLA, and from about 0.2 to about 30% GLA.

In addition to the concentration, the ratios of ARA, DGLA and GLA can be adapted for a particular given end use. When formulated as a breast milk supplement or substitute, an oil composition which contains two or more of  
10 ARA, DGLA and GLA will be provided in a ratio of about 1:19:30 to about 6:1:0.2, respectively. For example, the breast milk of animals can vary in ratios of ARA:DGLA:DGL ranging from 1:19:30 to 6:1:0.2, which includes intermediate ratios which are preferably about 1:1:1, 1:2:1, 1:1:4. When produced together in a host cell, adjusting the rate and percent of conversion of  
15 a precursor substrate such as GLA and DGLA to ARA can be used to precisely control the PUFA ratios. For example, a 5% to 10% conversion rate of DGLA to ARA can be used to produce an ARA to DGLA ratio of about 1:19, whereas a conversion rate of about 75% to 80% can be used to produce an ARA to DGLA ratio of about 6:1. Therefore, whether in a cell culture system or in a  
20 host animal, regulating the timing, extent and specificity of desaturase expression as described can be used to modulate the PUFA levels and ratios. Depending on the expression system used, e.g., cell culture or an animal expressing oil(s) in its milk, the oils also can be isolated and recombined in the desired concentrations and ratios. Amounts of oils providing these ratios of  
25 PUFA can be determined following standard protocols. PUFAs, or host cells containing them, also can be used as animal food supplements to alter an animal's tissue or milk fatty acid composition to one more desirable for human or animal consumption.

For dietary supplementation, the purified PUFAs, or derivatives thereof,  
30 may be incorporated into cooking oils, fats or margarines formulated so that in normal use the recipient would receive the desired amount. The PUFAs may

also be incorporated into infant formulas, nutritional supplements or other food products, and may find use as anti-inflammatory or cholesterol lowering agents.

### **Pharmaceutical Compositions**

5 The present invention also encompasses a pharmaceutical composition comprising one or more of the acids and/or resulting oils produced in accordance with the methods described herein. More specifically, such a pharmaceutical composition may comprise one or more of the acids and/or oils as well as a standard, well-known, non-toxic pharmaceutically acceptable carrier, adjuvant or vehicle such as, for example, phosphate buffered saline,  
10 water, ethanol, polyols, vegetable oils, a wetting agent or an emulsion such as a water/oil emulsion. The composition may be in either a liquid or solid form. For example, the composition may be in the form of a tablet, capsule, ingestible liquid or powder, injectible, or topical ointment or cream.

15 Possible routes of administration include, for example, oral, rectal and parenteral. The route of administration will, of course, depend upon the desired effect. For example, if the composition is being utilized to treat rough, dry, or aging skin, to treat injured or burned skin, or to treat skin or hair affected by a disease or condition, it may perhaps be applied topically.

20 The dosage of the composition to be administered to the patient may be determined by one of ordinary skill in the art and depends upon various factors such as weight of the patient, age of the patient, immune status of the patient, etc.

25 With respect to form, the composition may be, for example, a solution, a dispersion, a suspension, an emulsion or a sterile powder which is then reconstituted.

Additionally, the composition of the present invention may be utilized for cosmetic purposes. It may be added to pre-existing cosmetic compositions such that a mixture is formed or may be used as a sole composition.

Pharmaceutical compositions may be utilized to administer the PUFA component to an individual. Suitable pharmaceutical compositions may comprise physiologically acceptable sterile aqueous or non-aqueous solutions, dispersions, suspensions or emulsions and sterile powders for reconstitution into sterile solutions or dispersions for ingestion. Examples of suitable aqueous and non-aqueous carriers, diluents, solvents or vehicles include water, ethanol, polyols (propyleneglycol, polyethyleneglycol, glycerol, and the like), suitable mixtures thereof, vegetable oils (such as olive oil) and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the maintenance of the required particle size in the case of dispersions and by the use of surfactants. It may also be desirable to include isotonic agents, for example sugars, sodium chloride and the like. Besides such inert diluents, the composition can also include adjuvants, such as wetting agents, emulsifying and suspending agents, sweetening, flavoring and perfuming agents.

Suspensions, in addition to the active compounds, may contain suspending agents, as for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar and tragacanth or mixtures of these substances, and the like.

Solid dosage forms such as tablets and capsules can be prepared using techniques well known in the art. For example, PUFAs of the invention can be tableted with conventional tablet bases such as lactose, sucrose, and cornstarch in combination with binders such as acacia, cornstarch or gelatin, disintegrating agents such as potato starch or alginic acid and a lubricant such as stearic acid or magnesium stearate. Capsules can be prepared by incorporating these excipients into a gelatin capsule along with the antioxidants and the PUFA component. The amount of the antioxidants and PUFA component that should be incorporated into the pharmaceutical formulation should fit within the guidelines discussed above.

As used in this application, the term "treat" refers to either preventing, or reducing the incidence of, the undesired occurrence. For example, to treat immune suppression refers to either preventing the occurrence of this

suppression or reducing the amount of such suppression. The terms "patient" and "individual" are being used interchangeably and both refer to an animal. The term "animal" as used in this application refers to any warm-blooded mammal including, but not limited to, dogs, humans, monkeys, and apes. As  
5 used in the application the term "about" refers to an amount varying from the stated range or number by a reasonable amount depending upon the context of use. Any numerical number or range specified in the specification should be considered to be modified by the term about.

"Dose" and "serving" are used interchangeably and refer to the amount  
10 of the nutritional or pharmaceutical composition ingested by the patient in a single setting and designed to deliver effective amounts of the antioxidants and the structured triglyceride. As will be readily apparent to those skilled in the art, a single dose or serving of the liquid nutritional powder should supply the amount of antioxidants and PUFAs discussed above. The amount of the dose or  
15 serving should be a volume that a typical adult can consume in one sitting. This amount can vary widely depending upon the age, weight, sex or medical condition of the patient. However as a general guideline, a single serving or dose of a liquid nutritional produce should be considered as encompassing a volume from 100 to 600 ml, more preferably from 125 to 500 ml and most  
20 preferably from 125 to 300 ml.

The PUFAs of the present invention may also be added to food even when supplementation of the diet is not required. For example, the composition may be added to food of any type including but not limited to margarines, modified butters, cheeses, milk, yogurt, chocolate, candy, snacks, salad oils,  
25 cooking oils, cooking fats, meats, fish and beverages.

#### **Pharmaceutical Applications**

For pharmaceutical use (human or veterinary), the compositions are generally administered orally but can be administered by any route by which they may be successfully absorbed, e.g., parenterally (i.e. subcutaneously,  
30 intramuscularly or intravenously), rectally or vaginally or topically, for example, as a skin ointment or lotion. The PUFAs of the present invention may

be administered alone or in combination with a pharmaceutically acceptable carrier or excipient. Where available, gelatin capsules are the preferred form of oral administration. Dietary supplementation as set forth above also can provide an oral route of administration. The unsaturated acids of the present invention may be administered in conjugated forms, or as salts, esters, amides or prodrugs of the fatty acids. Any pharmaceutically acceptable salt is encompassed by the present invention; especially preferred are the sodium, potassium or lithium salts. Also encompassed are the N-alkylpolyhydroxamine salts, such as N-methyl glucamine, found in PCT publication WO 96/33155. The preferred esters are the ethyl esters. As solid salts, the PUFAs also can be administered in tablet form. For intravenous administration, the PUFAs or derivatives thereof may be incorporated into commercial formulations such as Intralipids. The typical normal adult plasma fatty acid profile comprises 6.64 to 9.46% of ARA, 1.45 to 3.11% of DGLA, and 0.02 to 0.08% of GLA. These PUFAs or their metabolic precursors can be administered, either alone or in mixtures with other PUFAs, to achieve a normal fatty acid profile in a patient. Where desired, the individual components of formulations may be individually provided in kit form, for single or multiple use. A typical dosage of a particular fatty acid is from 0.1 mg to 20 g, or even 100 g daily, and is preferably from 10 mg to 1, 2, 5 or 10 g daily as required, or molar equivalent amounts of derivative forms thereof. Parenteral nutrition compositions comprising from about 2 to about 30 weight percent fatty acids calculated as triglycerides are encompassed by the present invention; preferred is a composition having from about 1 to about 25 weight percent of the total PUFA composition as GLA (USPN 5,196,198). Other vitamins, and particularly fat-soluble vitamins such as vitamin A, D, E and L-carnitine can optionally be included. Where desired, a preservative such as  $\alpha$  tocopherol may be added, typically at about 0.1% by weight.

Suitable pharmaceutical compositions may comprise physiologically acceptable sterile aqueous or non-aqueous solutions, dispersions, suspensions or emulsions and sterile powders for reconstitution into sterile injectible solutions or dispersions. Examples of suitable aqueous and non-aqueous carriers,



diluents, solvents or vehicles include water, ethanol, polyols (propyleneglycol, polyethyleneglycol, glycerol, and the like), suitable mixtures thereof, vegetable oils (such as olive oil) and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the maintenance of the required particle size in the case of dispersions and by the use of surfactants. It may also be desirable to include isotonic agents, for example sugars, sodium chloride and the like. Besides such inert diluents, the composition can also include adjuvants, such as wetting agents, emulsifying and suspending agents, sweetening, flavoring and perfuming agents.

Suspensions in addition to the active compounds, may contain suspending agents, as for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar and tragacanth, or mixtures of these substances and the like.

An especially preferred pharmaceutical composition contains diacetyltartaric acid esters of mono- and diglycerides dissolved in an aqueous medium or solvent. Diacetyltartaric acid esters of mono- and diglycerides have an HLB value of about 9-12 and are significantly more hydrophilic than existing antimicrobial lipids that have HLB values of 2-4. Those existing hydrophobic lipids cannot be formulated into aqueous compositions. As disclosed herein, those lipids can now be solubilized into aqueous media in combination with diacetyltartaric acid esters of mono- and diglycerides. In accordance with this embodiment, diacetyltartaric acid esters of mono- and diglycerides (e.g., DATEM-C12:0) is melted with other active antimicrobial lipids (e.g., 18:2 and 12:0 monoglycerides) and mixed to obtain a homogeneous mixture. Homogeneity allows for increased antimicrobial activity. The mixture can be completely dispersed in water. This is not possible without the addition of diacetyltartaric acid esters of mono- and diglycerides and premixing with other monoglycerides prior to introduction into water. The aqueous composition can then be admixed under sterile conditions with physiologically acceptable diluents, preservatives, buffers or propellants as may be required to form a spray or inhalant.

The present invention also encompasses the treatment of numerous disorders with fatty acids. Supplementation with PUFAs of the present invention can be used to treat restenosis after angioplasty. Symptoms of inflammation, rheumatoid arthritis, and asthma and psoriasis can be treated with the PUFAs of the present invention. Evidence indicates that PUFAs may be involved in calcium metabolism, suggesting that PUFAs of the present invention may be used in the treatment or prevention of osteoporosis and of kidney or urinary tract stones.

The PUFAs of the present invention can be used in the treatment of cancer. Malignant cells have been shown to have altered fatty acid compositions; addition of fatty acids has been shown to slow their growth and cause cell death, and to increase their susceptibility to chemotherapeutic agents. GLA has been shown to cause reexpression on cancer cells of the E-cadherin cellular adhesion molecules, loss of which is associated with aggressive metastasis. Clinical testing of intravenous administration of the water soluble lithium salt of GLA to pancreatic cancer patients produced statistically significant increases in their survival. PUFA supplementation may also be useful for treating cachexia associated with cancer.

The PUFAs of the present invention can also be used to treat diabetes (USPN 4,826,877; Horrobin *et al.*, Am. J. Clin. Nutr. Vol. 57 (Suppl.), 732S-737S). Altered fatty acid metabolism and composition has been demonstrated in diabetic animals. These alterations have been suggested to be involved in some of the long-term complications resulting from diabetes, including retinopathy, neuropathy, nephropathy and reproductive system damage. Primrose oil, which contains GLA, has been shown to prevent and reverse diabetic nerve damage.

The PUFAs of the present invention can be used to treat eczema, reduce blood pressure and improve math scores. Essential fatty acid deficiency has been suggested as being involved in eczema, and studies have shown beneficial effects on eczema from treatment with GLA. GLA has also been shown to reduce increases in blood pressure associated with stress, and to improve performance on arithmetic tests. GLA and DGLA have been shown to inhibit

platelet aggregation, cause vasodilation, lower cholesterol levels and inhibit proliferation of vessel wall smooth muscle and fibrous tissue (Brenner *et al.*, Adv. Exp. Med. Biol. Vol. 83, p. 85-101, 1976). Administration of GLA or DGLA, alone or in combination with EPA, has been shown to reduce or prevent gastro-intestinal bleeding and other side effects caused by non-steroidal anti-inflammatory drugs (USPN 4,666,701). GLA and DGLA have also been shown to prevent or treat endometriosis and premenstrual syndrome (USPN 4,758,592) and to treat myalgic encephalomyelitis and chronic fatigue after viral infections (USPN 5,116,871).

Further uses of the PUFAs of this invention include use in treatment of AIDS, multiple sclerosis, acute respiratory syndrome, hypertension and inflammatory skin disorders. The PUFAs of the inventions also can be used for formulas for general health as well as for geriatric treatments.

#### Veterinary Applications

It should be noted that the above-described pharmaceutical and nutritional compositions may be utilized in connection with animals, as well as humans, as animals experience many of the same needs and conditions as human. For example, the oil or acids of the present invention may be utilized in animal feed supplements.

The following examples are presented by way of illustration, not of limitation.

#### Examples

- |           |  |
|-----------|--|
| Example 1 | Construction of a cDNA Library from <i>Mortierella alpina</i>  |
| Example 2 | Isolation of a $\Delta 6$ -desaturase Nucleotide Sequence from <i>Mortierella alpina</i>                     |
| Example 3 | Identification of $\Delta 6$ -desaturases Homologous to the <i>Mortierella alpina</i> $\Delta 6$ -desaturase |
| Example 4 | Isolation of a $\Delta 12$ -desaturase Nucleotide Sequence from <i>Mortierella Alpina</i>                    |

- 5                    Example 5      Expression of *M. alpina* Desaturase Clones in Baker's Yeast
- Example 6      Initial Optimization of Culture Conditions
- Example 7      Distribution of PUFAs in Yeast Lipid Fractions
- 10                  Example 8      Further Culture Optimization and Coexpression of  $\Delta 6$  and  $\Delta 12$ -desaturases
- Example 9      Identification of Homologues to *M. alpina*  $\Delta 5$  and  $\Delta 6$  desaturases
- 10                  Example 10     Identification of *M. alpina*  $\Delta 5$  and  $\Delta 6$  homologues in other PUFA-producing organisms
- Example 11     Identification of *M. alpina*  $\Delta 5$  and  $\Delta 6$  homologues in other PUFA-producing organisms
- Example 12     Human Desaturase Gene Sequences
- Example 13     Nutritional Compositions
- 15

### **Example 1**

#### **Construction of a cDNA Library from *Mortierella alpina***

20                    Total RNA was isolated from a 3 day old PUFA-producing culture of *Mortierella alpina* using the protocol of Hoge *et al.* (1982) *Experimental Mycology* 6:225-232. The RNA was used to prepare double-stranded cDNA using BRL's lambda-ZipLox system following the manufactures instructions. Several size fractions of the *M. alpina* cDNA were packaged separately to yield libraries with different average-sized inserts. A "full-length" library contains approximately  $3 \times 10^6$  clones with an average insert size of 1.77 kb. The

25                    "sequencing-grade" library contains approximately  $6 \times 10^5$  clones with an average insert size of 1.1 kb.

## Example 2

### Isolation of a $\Delta 6$ -desaturase Nucleotide Sequence from *Mortierella Alpina*

A nucleic acid sequence from a partial cDNA clone, Ma524, encoding a  $\Delta 6$  fatty acid desaturase from *Mortierella alpina* was obtained by random sequencing of clones from the *M. alpina* cDNA sequencing grade library described in Example 1. cDNA-containing plasmids were excised as follows:

Five  $\mu$ l of phage were combined with 100  $\mu$ l of *E. coli* DH10B(ZIP) grown in ECLB plus 10  $\mu$ g/ml kanamycin, 0.2% maltose, and 10 mM  $MgSO_4$  and incubated at 37 degrees for 15 minutes. 0.9 ml SOC was added and 100  $\mu$ l of the bacteria immediately plated on each of 10 ECLB + 50  $\mu$ g Pen plates. No 45 minute recovery time was needed. The plates were incubated overnight at 37°. Colonies were picked into ECLB + 50  $\mu$ g Pen media for overnight cultures to be used for making glycerol stocks and miniprep DNA. An aliquot of the culture used for the miniprep is stored as a glycerol stock. Plating on ECLB + 50  $\mu$ g Pen/ml resulted in more colonies and a greater proportion of colonies containing inserts than plating on 100  $\mu$ g/ml Pen.

Random colonies were picked and plasmid DNA purified using Qiagen miniprep kits. DNA sequence was obtained from the 5' end of the cDNA insert and compared to the National Center for Biotechnology Information (NCBI) nonredundant database using the BLASTX algorithm. Ma524 was identified as a putative desaturase based on DNA sequence homology to previously identified desaturases.

A full-length cDNA clone was isolated from the *M. alpina* full-length library and designed pCGN5532. The cDNA is contained as a 1617 bp insert in the vector pZL1 (BRL) and, beginning with the first ATG, contains an open reading frame encoding 457 amino acids. The three conserved "histidine boxes" known to be conserved among membrane-bound deaturases (Okuley, et al. (1994) *The Plant Cell* 6:147-158) were found to be present at amino acid positions 172-176, 209-213, and 395-399 (see Figure 3). As with other

membrane-bound  $\Delta 6$ -desaturases the final HXXHH histidine box motif was found to be QXXHH. The amino acid sequence of Ma524 was found to display significant homology to a portion of a *Caenorhabditis elegans* cosmid, WO6D2.4, a cytochrome b5/desaturase fusion protein from sunflower, and the  
5 *Synechocystis* and *Spirulina*  $\Delta 6$ -desaturases. In addition, Ma524 was shown to have homology to the borage  $\Delta 6$ -desaturase amino sequence (PCT publication W) 96/21022). Ma524 thus appears to encode a  $\Delta 6$ -desaturase that is related to the borage and algal  $\Delta 6$ -desaturases. The peptide sequences are shown as SEQ ID NO:5 - SEQ ID NO:11.

10 The amino terminus of the encoded protein was found to exhibit significant homology to cytochrome b5 proteins. The *Mortierella* cDNA clone appears to represent a fusion between a cytochrome b5 and a fatty acid desaturase. Since cytochrome b5 is believed to function as the electron donor for membrane-bound desaturase enzymes, it is possible that the N-terminal  
15 cytochrome b5 domain of this desaturase protein is involved in its function. This may be advantageous when expressing the desaturase in heterologous systems for PUFA production. However, it should be noted that, although the amino acid sequences of Ma524 and the borage  $\Delta 6$  were found to contain regions of homology, the base compositions of the cDNAs were shown to be  
20 significantly different. For example, the borage cDNA was shown to have an overall base composition of 60 % A/T, with some regions exceeding 70 %, while Ma524 was shown to have an average of 44 % A/T base composition, with no regions exceeding 60 %. This may have implications for expressing the cDNAs in microorganisms or animals which favor different base compositions.  
25 It is known that poor expression of recombinant genes can occur when the host prefers a base composition different from that of the introduced gene. Mechanisms for such poor expression include decreased stability, cryptic splice sites, and/or translatability of the mRNA and the like.

### Example 3

#### Identification of $\Delta 6$ -desaturases Homologous to the *Mortierella alpina* $\Delta 6$ -desaturase

Nucleic acid sequences that encode putative  $\Delta 6$ -desaturases were  
5 identified through a BLASTX search of the Expressed Sequence Tag ("EST")  
databases through NCBI using the Ma524 amino acid sequence. Several  
sequences showed significant homology. In particular, the deduced amino acid  
sequence of two *Arabidopsis thaliana* sequences, (accession numbers F13728  
and T42806) showed homology to two different regions of the deduced amino  
10 acid sequence of Ma524. The following PCR primers were designed:  
ATTS4723-FOR (complementary to F13728) SEQ ID NO:13  
5' CUACUACUACUAGGAGTCCTCTACGGTGTTTTG and  
T42806-REV (complementary to T42806) SEQ ID NO:14  
5' CAUCAUCAUATGATGCTCAAGCTGAACTG. Five  $\mu$ g of total  
15 RNA isolated from developing siliques of *Arabidopsis thaliana* was reverse  
transcribed using BRL Superscript RTase and the primer TSyn  
(5'-CCAAGCTTCTGCAGGAGCTCTTTTTTTTTTTTTTTT-3') and is shown as  
SEQ ID NO:12. PCR was carried out in a 50  $\mu$ l volume containing: template  
derived from 25 ng total RNA, 2 pM each primer, 200  $\mu$ M each  
20 deoxyribonucleotide triphosphate, 60 mM Tris-Cl, pH 8.5, 15 mM  $(\text{NH}_4)_2\text{SO}_4$ ,  
2 mM  $\text{MgCl}_2$ , 0.2 U Taq Polymerase. Thermocycler conditions were as  
follows: 94 degrees for 30 sec., 50 degrees for 30 sec., 72 degrees for 30 sec.  
PCR was continued for 35 cycles followed by an additional extension at 72  
degrees for 7 minutes. PCR resulted in a fragment of approximately ~750 base  
25 pairs which was subcloned, named 12-5, and sequenced. Each end of this  
fragment was formed to correspond to the *Arabidopsis* ESTs from which the  
PCR primers were designed. The putative amino acid sequence of 12-5 was  
compared to that of Ma524, and ESTs from human (W28140), mouse  
(W53753), and *C. elegans* (R05219) (see Figure 4). Homology patterns with  
30 the *Mortierella*  $\Delta 6$ - desaturase indicate that these sequences represent putative

desaturase polypeptides. Based on this experiment approach, it is likely that the full-length genes can be cloned using probes based on the EST sequences. Following the cloning, the genes can then be placed into expression vectors, expressed in host cells, and their specific  $\Delta 6$ - or other desaturase activity can be determined as described below.

#### Example 4

##### Isolation of a $\Delta 12$ -desaturase Nucleotide Sequence from *Mortierella alpina*

Based on the fatty acids it accumulates, it seemed probable that *Mortierella alpina* has an  $\omega 6$  type desaturase. The  $\omega 6$ -desaturase is responsible for the production of linoleic acid (18:2) from oleic acid (18:1). Linoleic acid (18:2) is a substrate for a  $\Delta 6$ -desaturase. This experiment was designed to determine if *Mortierella alpina* has a  $\Delta 12$ -desaturase polypeptide, and if so, to identify the corresponding nucleotide sequence.

A random colony from the *M. alpina* sequencing grade library, Ma648, was sequenced and identified as a putative desaturase based on DNA sequence homology to previously identified desaturases, as described for Ma524 (*see* Example 2). The nucleotide sequence is shown in SEQ ID NO:13. The peptide sequence is shown in SEQ ID NO:4. The deduced amino acid sequence from the 5' end of the Ma648 cDNA displays significant homology to soybean microsomal  $\omega 6$  ( $\Delta 12$ ) desaturase (accession #L43921) as well as castor bean oleate 12-hydroxylase (accession #U22378). In addition, homology was observed when compared to a variety of other  $\omega 6$  ( $\Delta 12$ ) and  $\omega 3$  ( $\Delta 15$ ) fatty acid desaturase sequences.



### Example 5

#### Expression of *M. alpina* Desaturase Clones in Baker's Yeast

##### Yeast Transformation

Lithium acetate transformation of yeast was performed according to standard protocols (*Methods in Enzymology*, Vol. 194, p. 186-187, 1991). Briefly, yeast were grown in YPD at 30°C. Cells were spun down, resuspended in TE, spun down again, resuspended in TE containing 100 mM lithium acetate, spun down again, and resuspended in TE/lithium acetate. The resuspended yeast were incubated at 30°C for 60 minutes with shaking. Carrier DNA was added, and the yeast were aliquoted into tubes. Transforming DNA was added, and the tubes were incubated for 30 min. at 30°C. PEG solution (35% (w/v) PEG 4000, 100 mM lithium acetate, TE pH7.5) was added followed by a 50 min. incubation at 30°C. A 5 min. heat shock at 42°C was performed, the cells were pelleted, washed with TE, pelleted again and resuspended in TE. The resuspended cells were then plated on selective media.

##### Desaturase Expression in Transformed Yeast

cDNA clones from *Mortierella alpina* were screened for desaturase activity in baker's yeast. A canola  $\Delta 15$ -desaturase (obtained by PCR using 1<sup>st</sup> strand cDNA from *Brassica napus* cultivar 212/86 seeds using primers based on the published sequence (Arondel *et al. Science* 258:1353-1355)) was used as a positive control. The  $\Delta 15$ -desaturase gene and the gene from cDNA clones Ma524 and Ma648 were put in the expression vector pYES2 (Invitrogen), resulting in plasmids pCGR-2, pCGR-5 and pCGR-7, respectively. These plasmids were transfected into *S. cerevisiae* yeast strain 334 and expressed after induction with galactose and in the presence of substrates that allowed detection of specific desaturase activity. The control strain was *S. cerevisiae* strain 334 containing the unaltered pYES2 vector. The substrates used, the products produced and the indicated desaturase activity were: DGLA (conversion to ARA would indicate  $\Delta 5$ -desaturase activity), linoleic acid (conversion to GLA

would indicate  $\Delta 6$ -desaturase activity; conversion to ALA would indicate  $\Delta 15$ -desaturase activity), oleic acid (an endogenous substrate made by *S. cerevisiae*, conversion to linoleic acid would indicate  $\Delta 12$ -desaturase activity, which *S. cerevisiae* lacks), or ARA (conversion to EPA would indicate  $\Delta 17$ -desaturase activity).

Cultures were grown for 48-52 hours at 15°C in the presence of a particular substrate. Lipid fractions were extracted for analysis as follows: Cells were pelleted by centrifugation, washed once with sterile ddH<sub>2</sub>O, and repelleted. Pellets were vortexed with methanol; chloroform was added along with tritridecanoin (as an internal standard). The mixtures were incubated for at least one hour at room temperature or at 4°C overnight. The chloroform layer was extracted and filtered through a Whatman filter with one gram of anhydrous sodium sulfate to remove particulates and residual water. The organic solvents were evaporated at 40°C under a stream of nitrogen. The extracted lipids were then derivatized to fatty acid methyl esters (FAME) for gas chromatography analysis (GC) by adding 2 ml of 0.5 N potassium hydroxide in methanol to a closed tube. The samples were heated to 95°C to 100°C for 30 minutes and cooled to room temperature. Approximately 2 ml of 14 % boron trifluoride in methanol was added and the heating repeated. After the extracted lipid mixture cooled, 2 ml of water and 1 ml of hexane were added to extract the FAME for analysis by GC. The percent conversion was calculated by dividing the product produced by the sum of (the product produced and the substrate added) and then multiplying by 100. To calculate the oleic acid percent conversion, as no substrate was added, the total linoleic acid produced was divided by the sum of oleic acid and linoleic acid produced, then multiplying by 100. The desaturase activity results are provided in Table 1 below.

**Table 1*****M. alpina* Desaturase Expression in Baker's Yeast**

CLONE	ENZYME ACTIVITY	% CONVERSION OF SUBSTRATE
pCGR-2	$\Delta 6$	0 (18:2 to 18:3w6)
(canola $\Delta 15$ desaturase)	$\Delta 15$	16.3 (18:2 to 18:3w3)
	$\Delta 5$	2.0 (20:3 to 20:4w6)
	$\Delta 17$	2.8 (20:4 to 20:5w3)
	$\Delta 12$	1.8 (18:1 to 18:2w6)
pCGR-5	$\Delta 6$	6.0
(M. alpina Ma524)	$\Delta 15$	0
	$\Delta 5$	2.1
	$\Delta 17$	0
	$\Delta 12$	3.3
pCGR-7	$\Delta 6$	0
(M. alpina Ma648)	$\Delta 15$	3.8
	$\Delta 5$	2.2
	$\Delta 17$	0
	$\Delta 12$	63.4

5 The  $\Delta 15$ -desaturase control clone exhibited 16.3% conversion of the  
 substrate. The pCGR-5 clone expressing the Ma524 cDNA showed 6%  
 conversion of the substrate to GLA, indicating that the gene encodes a  $\Delta 6$ -  
 desaturase. The pCGR-7 clone expressing the Ma648 cDNA converted 63.4%  
 conversion of the substrate to LA, indicating that the gene encodes a  $\Delta 12$ -  
 desaturase. The background (non-specific conversion of substrate) was between  
 10 0-3% in these cases. We also found substrate inhibition of the activity by using  
 different concentrations of the substrate. When substrate was added to 100  $\mu$ M,  
 the percent conversion to product dropped compared to when substrate was added  
 to 25  $\mu$ M (see below). Additionally, by varying the substrate concentration  
 between 5  $\mu$ M and 200  $\mu$ M, conversion ratios were found to range between about

5% to about 75% greater. These data show that desaturases with different substrate specificities can be expressed in a heterologous system and used to produce poly-unsaturated long chain fatty acids.

Table 2 represents fatty acids of interest as a percent of the total lipid extracted from the yeast host *S. cerevisiae* 334 with the indicated plasmid. No glucose was present in the growth media. Affinity gas chromatography was used to separate the respective lipids. GC/MS was employed to verify the identity of the product(s). The expected product for the *B. napus*  $\Delta 15$ -desaturase,  $\alpha$ -linolenic acid, was detected when its substrate, linoleic acid, was added exogenously to the induced yeast culture. This finding demonstrates that yeast expression of a desaturase gene can produce functional enzyme and detectable amounts of product under the current growth conditions. Both exogenously added substrates were taken up by yeast, although slightly less of the longer chain PUFA, dihomo- $\gamma$ -linolenic acid (20:3), was incorporated into yeast than linoleic acid (18:2) when either was added in free form to the induced yeast cultures.  $\gamma$ -linolenic acid was detected when linoleic acid was present during induction and expression of *S. cerevisiae* 334 (pCGR-5). The presence of this PUFA demonstrates  $\Delta 6$ -desaturase activity from pCGR-5 (MA524). Linoleic acid, identified in the extracted lipids from expression of *S. cerevisiae* 334 (pCGR-7), classifies the cDNA MA648 from *M. alpina* as the  $\Delta 12$ -desaturase.

Table 2

## Fatty Acid as a Percentage of Total Lipid Extracted from Yeast

Plasmid in Yeast (enzyme)	18:2 Incorporated	$\alpha$ -18:3 Produced	$\gamma$ -18:3 Produced	20:3 Incorporated	20:4 Produced	18:1* Present	18:2 Produced
pYES2 (control)	66.9	0	0	58.4	0	4	0
pCGR-2 ( $\Delta$ 15)	60.1	5.7	0	50.4	0	0.7	0
pCGR-5 ( $\Delta$ 6)	62.4	0	4.0	49.9	0	2.4	0
pCGR-7 ( $\Delta$ 12)	65.6	0	0	45.7	0	7.1	12.2

100  $\mu$ M substrate added

\* 18:1 is an endogenous fatty acid in yeast

## Key To Tables

18:1=oleic acid

18:2=linoleic acid

 $\alpha$ -18:3= $\alpha$ -linolenic acid $\gamma$ -18:3= $\gamma$ -linolenic acid

18:4=stearidonic acid

20:3=dihomo- $\gamma$ -linolenic acid

20:4=arachidonic acid

### Example 6

#### Optimization of Culture Conditions

Table 3A shows the effect of exogenous free fatty acid substrate concentration on yeast uptake and conversion to fatty acid product as a percentage of the total yeast lipid extracted. In all instances, low amounts of exogenous substrate (1-10  $\mu$ M) resulted in low fatty acid substrate uptake and product formation. Between 25 and 50  $\mu$ M concentration of free fatty acid in the growth and induction media gave the highest percentage of fatty acid product formed, while the 100  $\mu$ M concentration and subsequent high uptake into yeast appeared to decrease or inhibit the desaturase activity. The amount of fatty acid substrate for yeast expressing  $\Delta$ 12-desaturase was similar under the same growth conditions, since the substrate, oleic acid, is an endogenous yeast fatty acid. The use of  $\alpha$ -linolenic acid as an additional substrate for pCGR-5 ( $\Delta$ 6) produced the expected product, stearidonic acid (Table 3A). The feedback inhibition of high fatty acid substrate concentration was well illustrated when the percent conversion rates of the respective fatty acid substrates to their respective products were compared in Table 3B. In all cases, 100  $\mu$ M substrate concentration in the growth media decreased the percent conversion to product. The uptake of  $\alpha$ -linolenic was comparable to other PUFAs added in free form, while the  $\Delta$ 6-desaturase percent conversion, 3.8-17.5%, to the product stearidonic acid was the lowest of all the substrates examined (Table 3B). The effect of media, such as YPD (rich media) versus minimal media with glucose on the conversion rate of  $\Delta$ 12-desaturase was dramatic. Not only did the conversion rate for oleic to linoleic acid drop, (Table 3B) but the percent of linoleic acid formed also decreased by 11% when rich media was used for growth and induction of yeast desaturase  $\Delta$ 12 expression (Table 3A). The effect of media composition was also evident when glucose was present in the growth media for  $\Delta$ 6-desaturase, since the percent of substrate uptake was decreased at 25  $\mu$ M (Table 3A). However, the conversion rate remained the

same and percent product formed decreased for  $\Delta 6$ -desaturase for in the presence of glucose.

**Table 3A**

5                    **Effect of Added Substrate on the Percentage of Incorporated  
Substrate and Product Formed in Yeast Extracts**

Plasmid in Yeast	pCGR-2 ( $\Delta 15$ )	PcGR-5 ( $\Delta 6$ )	pCGR-5 ( $\Delta 6$ )	pCGR-7 ( $\Delta 12$ )
Substrate/product	18:2 / $\alpha$ -18:3	18:2/ $\gamma$ -18:3	$\alpha$ -18:3/18:4	18:1*/18:2
1 $\mu$ M sub.	ND	0.9/0.7	ND	ND
10 $\mu$ M sub.	ND	4.2/2.4	10.4/2.2	ND
25 $\mu$ M sub.	ND	11/3.7	18.2/2.7	ND
25 $\mu$ M $\phi$ sub.	36.6/7.2 $\phi$	25.1/10.3 $\phi$	ND	6.6/15.8 $\phi$
50 $\mu$ M sub.	53.1/6.5 $\phi$	ND	36.2/3	10.8/13*
100 $\mu$ M sub.	60.1/5.7 $\phi$	62.4/4 $\phi$	47.7/1.9	10/24.8

**Table 3B**  
**Effect of Substrate Concentration in Media on the Percent Conversion**  
**of Fatty Acid Substrate to Product in Yeast Extracts**

Plasmid in Yeast	pCGR-2 ( $\Delta 15$ )	pCGR-5 ( $\Delta 6$ )	pCGR-5 ( $\Delta 6$ )	pCGR-7 ( $\Delta 12$ )
substrate→product	18:2 → $\alpha$ -18:3	18:2→ $\gamma$ 18:3	$\alpha$ -18:3→18:4	18:1*→18:2
1 $\mu$ M sub.	ND	43.8	ND	ND
10 $\mu$ M sub.	ND	36.4	17.5	ND
25 $\mu$ M sub.	ND	25.2	12.9	ND
25 $\mu$ M $\diamond$ sub.	16.4 $\diamond$	29.1 $\diamond$	ND	70.5 $\diamond$
50 $\mu$ M sub.	10.9 $\diamond$	ND	7.7	54.6*
100 $\mu$ M sub.	8.7 $\diamond$	6 $\diamond$	3.8	71.3

$\diamond$  no glucose in media

\* Yeast peptone broth (YPD)

\* 18:1 is an endogenous yeast lipid

sub. is substrate concentration

ND (not done)

10                      Table 4 shows the amount of fatty acid produced by a recombinant  
desaturase from induced yeast cultures when different amounts of free fatty acid  
substrate were used. Fatty acid weight was determined since the total amount of  
lipid varied dramatically when the growth conditions were changed, such as the  
presence of glucose in the yeast growth and induction media. To better  
15                      determine the conditions when the recombinant desaturase would produce the  
most PUFA product, the quantity of individual fatty acids were examined. The  
absence of glucose dramatically reduced by three fold the amount of linoleic  
acid produced by recombinant  $\Delta 12$ -desaturase. For the  $\Delta 12$ -desaturase the  
amount of total yeast lipid was decreased by almost half in the absence of  
20                      glucose. Conversely, the presence of glucose in the yeast growth media for  $\Delta 6$ -  
desaturase drops the  $\gamma$ -linolenic acid produced by almost half, while the total  
amount of yeast lipid produced was not changed by the presence/absence of



glucose. This points to a possible role for glucose as a modulator of  $\Delta 6$ -desaturase activity.

**Table 4**

**Fatty Acid Produced in  $\mu\text{g}$  from Yeast Extracts**

Plasmid in Yeast (enzyme)	pCGR-5 ( $\Delta 6$ )	pCGR-5 ( $\Delta 6$ )	pCGR-7 ( $\Delta 12$ )
product	Y-18:3	18:4	18:2*
1 $\mu\text{M}$ sub.	1.9	ND	ND
10 $\mu\text{M}$ sub.	5.3	4.4	ND
25 $\mu\text{M}$ sub.	10.3	8.7	115.7
25 $\mu\text{M}$ $\diamond$ sub.	29.6	ND	39 $\diamond$

$\diamond$  no glucose in media

sub. is substrate concentration

ND (not done)

\*18:1, the substrate, is an endogenous yeast lipid

**Example 7**

**Distribution of PUFAs in Yeast Lipid Fractions**

Table 5 illustrates the uptake of free fatty acids and their new products formed in yeast lipids as distributed in the major lipid fractions. A total lipid extract was prepared as described above. The lipid extract was separated on TLC plates, and the fractions were identified by comparison to standards. The bands were collected by scraping, and internal standards were added. The fractions were then saponified and methylated as above, and subjected to gas chromatography. The gas chromatograph calculated the amount of fatty acid by comparison to a standard. The phospholipid fraction contained the highest amount of substrate and product PUFAs for  $\Delta 6$ -desaturase activity. It would appear that the substrates are accessible in the phospholipid form to the desaturases.

Table 5

Fatty Acid Distribution in Various Yeast Lipid Fractions in  $\mu\text{g}$ 

Fatty acid fraction	Phospholipid	Diglyceride	Free Fatty Acid	Triglyceride	Cholesterol Ester
SC (pCGR-5) substrate 18:2	166.6	6.2	15	18.2	15.6
SC (pCGR-5) product $\gamma$ -18:3	61.7	1.6	4.2	5.9	1.2

SC = *S. cerevisiae* (plasmid)

5

Example 8Further Culture Optimization and Coexpression of  $\Delta 6$  and  $\Delta 12$ -desaturases

This experiment was designed to evaluate the growth and induction conditions for optimal activities of desaturases in *Saccharomyces cerevisiae*. A *Saccharomyces cerevisiae* strain (SC334) capable of producing  $\gamma$ -linolenic acid (GLA) was developed, to assess the feasibility of production of PUFA in yeast.

The genes for  $\Delta 6$  and  $\Delta 12$ -desaturases from *M. alpina* were coexpressed in SC334. Expression of  $\Delta 12$ -desaturase converted oleic acid (present in yeast) to linoleic acid. The linoleic acid was used as a substrate by the  $\Delta 6$ -desaturase to produce GLA. The quantity of GLA produced ranged between 5-8% of the total fatty acids produced in SC334 cultures and the conversion rate of linoleic acid to  $\gamma$ -linolenic acid ranged between 30% to 50%. The induction temperature was optimized, and the effect of changing host strain and upstream promoter sequences on expression of  $\Delta 6$  and  $\Delta 12$  (MA 524 and MA 648 respectively) desaturase genes was also determined.

20

**Plasmid Construction**

The cloning of pCGR5 as well as pCGR7 has been discussed above. To construct pCGR9a and pCGR9b, the  $\Delta 6$  and  $\Delta 12$ -desaturase genes were amplified using the following sets of primers. The primers pRDS1 and 3 had XhoI site and primers pRDS2 and 4 had XbaI site (indicated in bold). These primer sequences are presented as SEQ ID NO:15-18.

**I.  $\Delta 6$ -desaturase amplification primers**

a. pRDS1 TAC CAA **CTC GAG** AAA ATG GCT GCT GCT CCC  
AGT GTG AGG

10 b. pRDS2 AAC TGA **TCT AGA** TTA CTG CGC CTT ACC CAT  
CTT GGA GGC

**II.  $\Delta 12$ -desaturase amplification primers**

a. pRDS3 TAC CAA **CTC GAG** AAA ATG GCA CCT CCC  
AAC ACT ATC GAT

15 b. pRDS4 AAC TGA **TCT AGA** TTA CTT CTT GAA AAA GAC  
CAC GTC TCC

The pCGR5 and pCGR7 constructs were used as template DNA for amplification of  $\Delta 6$  and  $\Delta 12$ -desaturase genes, respectively. The amplified products were digested with XbaI and XhoI to create "sticky ends". The PCR amplified  $\Delta 6$ -desaturase with XhoI-XbaI ends as cloned into pCGR7, which was also cut with XhoI-XbaI. This procedure placed the  $\Delta 6$ -desaturase behind the  $\Delta 12$ -desaturase, under the control of an inducible promoter GAL1. This construct was designated pCGR9a. Similarly, to construct pCGR9b, the  $\Delta 12$ -desaturase with XhoI-XbaI ends was cloned in the XhoI-XbaI sites of pCGR5. In pCGR9b the  $\Delta 12$ -desaturase was behind the  $\Delta 6$ -desaturase gene, away from the GAL promoter.

To construct pCGR10, the vector pRS425, which contains the constitutive Glyceraldehyde 3-Phosphate Dehydrogenase (GPD) promoter, was digested with BamHI and pCGR5 was digested with BamHI-XhoI to release the

$\Delta 6$ -desaturase gene. This  $\Delta 6$ -desaturase fragment and BamHI cut pRS425 were filled using Klenow Polymerase to create blunt ends and ligated, resulting in pCGR10a and pCGR10b containing the  $\Delta 6$ -desaturase gene in the sense and antisense orientation, respectively. To construct pCGR11 and pCGR12, the  $\Delta 6$  and  $\Delta 12$ -desaturase genes were isolated from pCGR5 and pCGR7, respectively, using an EcoRI-XhoI double digest. The EcoRI-XhoI fragments of  $\Delta 6$  and  $\Delta 12$ -desaturases were cloned into the pYX242 vector digested with EcoRI-XhoI. The pYX242 vector has the promoter of TPI ( a yeast housekeeping gene), which allows constitutive expression.

#### 10 **Yeast Transformation and Expression**

Different combinations of pCGR5, pCGR7, pCGR9a, pCGR9b, pCGR10a, pCGR11 and pCGR12 were introduced into various host strains of *Saccharomyces cerevisiae*. Transformation was done using PEG/LiAc protocol (Methods in Enzymology Vol. 194 (1991): 186-187). Transformants were selected by plating on synthetic media lacking the appropriate amino acid. The pCGR5, pCGR7, pCGR9a and pCGR9b can be selected on media lacking uracil. The pCGR10, pCGR11 and pCGR12 constructs can be selected on media lacking leucine. Growth of cultures and fatty acid analysis was performed as in Example 5 above.

#### 20 **Production of GLA**

Production of GLA requires the expression of two enzymes ( the  $\Delta 6$  and  $\Delta 12$ -desaturases), which are absent in yeast. To express these enzymes at optimum levels the following constructs or combinations of constructs, were introduced into various host strains:

- 25 1) pCGR9a/SC334
- 2) pCGR9b/SC334
- 3) pCGR10a and pCGR7/SC334
- 4) pCGR11 and pCGR7/SC334
- 5) pCGR12 and pCGR5/SC334

6) pCGR10a and pCGR7/DBY746

7) pCGR10a and pCGR7/DBY746

The pCGR9a construct has both the  $\Delta 6$  and  $\Delta 12$ -desaturase genes under the control of an inducible GAL promoter. The SC334 host cells transformed with this construct did not show any GLA accumulation in total fatty acids (Fig. 6A and B, lane 1). However, when the  $\Delta 6$  and  $\Delta 12$ -desaturase genes were individually controlled by the GAL promoter, the control constructs were able to express  $\Delta 6$ - and  $\Delta 12$ -desaturase, as evidenced by the conversion of their respective substrates to products. The  $\Delta 12$ -desaturase gene in pCGR9a was expressed as evidenced by the conversion of 18:1 $\omega$ 9 to 18:2 $\omega$ 6 in pCGR9a/SC334, while the  $\Delta 6$ -desaturase gene was not expressed/active, because the 18:2 $\omega$ 6 was not being converted to 18:3 $\omega$ 6 (Fig. 6A and B, lane 1).

The pCGR9b construct also had both the  $\Delta 6$  and  $\Delta 12$ -desaturase genes under the control of the GAL promoter but in an inverse order compared to pCGR9a. In this case, very little GLA (<1%) was seen in pCGR9b/SC334 cultures. The expression of  $\Delta 12$ -desaturase was also very low, as evidenced by the low percentage of 18:2 $\omega$ 6 in the total fatty acids (Fig. 6A and B, lane 1).

To test if expressing both enzymes under the control of independent promoters would increase GLA production, the  $\Delta 6$ -desaturase gene was cloned into the pRS425 vector. The construct of pCGR10a has the  $\Delta 6$ -desaturase in the correct orientation, under control of constitutive GPD promoter. The pCGR10b has the  $\Delta 6$ -desaturase gene in the inverse orientation, and serves as the negative control. The pCGR10a/SC334 cells produced significantly higher levels of GLA (5% of the total fatty acids, Fig. 6, lane 3), compared to pCGR9a. Both the  $\Delta 6$  and  $\Delta 12$ -desaturase genes were expressed at high level because the conversion of 18:1 $\omega$ 9  $\rightarrow$  18:2 $\omega$ 6 was 65%, while the conversion of 18:2 $\omega$ 6  $\rightarrow$  18:3 $\omega$ 6 ( $\Delta 6$ -desaturase) was 30% (Fig. 6, lane 3). As expected, the negative control pCGR10b/SC334 did not show any GLA.

To further optimize GLA production, the  $\Delta 6$  and  $\Delta 12$  genes were introduced into the pYX242 vector, creating pCGR11 and pCGR12

respectively. The pYX242 vector allows for constitutive expression by the TP1 promoter (Alber, T. and Kawasaki, G. (1982). *J. Mol. & Appl. Genetics* 1: 419). The introduction of pCGR11 and pCGR7 in SC334 resulted in approximately 8% of GLA in total fatty acids of SC334. The rate of conversion of 18:1 $\omega$ 9  $\rightarrow$  18:2 $\omega$ 6 and 18:2 $\omega$ 6  $\rightarrow$  18:3 $\omega$ 6 was approximately 50% and 44% respectively (Fig. 6A and B, lane 4). The presence of pCGR12 and pCGR5 in SC334 resulted in 6.6% GLA in total fatty acids with a conversion rate of approximately 50% for both 18:1 $\omega$ 9 to 18:2 $\omega$ 6 and 18:2 $\omega$ 6 to 18:3 $\omega$ 6, respectively (Fig. 6A and B, lane 5). Thus although the quantity of GLA in total fatty acids was higher in the pCGR11/pCGR7 combination of constructs, the conversion rates of substrate to product were better for the pCGR12/pCGR5 combination.

To determine if changing host strain would increase GLA production, pCGR10a and pCGR7 were introduced into the host strain BJ1995 and DBY746 (obtained from the Yeast Genetic Stock Centre, 1021 Donner Laboratory, Berkeley, CA 94720. The genotype of strain DBY746 is Mat $\alpha$ , his3- $\Delta$ 1, leu2-3, leu2-112, ura3-32, trp1-289, gal). The results are shown in Fig. 7. Changing host strain to BJ1995 did not improve the GLA production, because the quantity of GLA was only 1.31% of total fatty acids and the conversion rate of 18:1 $\omega$ 9  $\rightarrow$  18:2 $\omega$ 6 was approximately 17% in BJ1995. No GLA was observed in DBY746 and the conversion of 18:1 $\omega$ 9  $\rightarrow$  18:2 $\omega$ 6 was very low (<1% in control) suggesting that a cofactor required for the expression of  $\Delta$ 12-desaturase might be missing in DB746 (Fig. 7, lane 2).

To determine the effect of temperature on GLA production, SC334 cultures containing pCGR10a and pCGR7 were grown at 15°C and 30°C. Higher levels of GLA were found in cultures grown and induced at 15°C than those in cultures grown at 30°C (4.23% vs. 1.68%). This was due to a lower conversion rate of 18:2 $\omega$ 6  $\rightarrow$  18:3 $\omega$ 6 at 30°C (11.6% vs. 29% in 15°C) cultures, despite a higher conversion of 18:1 $\omega$ 9  $\rightarrow$  18:2 $\omega$ 6 (65% vs. 60% at 30°C (Fig. 8). These results suggest that  $\Delta$ 12- and  $\Delta$ 6-desaturases may have different optimal expression temperatures.

Of the various parameters examined in this study, temperature of growth, yeast host strain and media components had the most significant impact on the expression of desaturase, while timing of substrate addition and concentration of inducer did not significantly affect desaturase expression.

5           These data show that two DNAs encoding desaturases that can convert LA to GLA or oleic acid to LA can be isolated from *Mortierella alpina* and can be expressed, either individually or in combination, in a heterologous system and used to produce poly-unsaturated long chain fatty acids. Exemplified is the production of GLA from oleic acid by expression of  $\Delta 12$ - and  $\Delta 6$ -desaturases in  
10       yeast.

#### **Example 9**

##### **Identification of Homologues to *M. alpina* $\Delta 5$ and $\Delta 6$ desaturases**

A nucleic acid sequence that encodes a putative  $\Delta 5$  desaturase was identified through a TBLASTN search of the expressed sequence tag databases  
15       through NCBI using amino acids 100-446 of Ma29 as a query. The truncated portion of the Ma29 sequence was used to avoid picking up homologies based on the cytochrome b5 portion at the N-terminus of the desaturase. The deduced amino acid sequence of an est from *Dictyostelium discoideum* (accession # C25549) shows very significant homology to Ma29 and lesser, but still  
20       significant homology to Ma524. The DNA sequence is presented as SEQ ID NO:19. The amino acid sequence is presented as SEQ ID NO:20.

#### **Example 10**

##### **Identification of *M. alpina* $\Delta 5$ and $\Delta 6$ homologues in other PUFA-producing organisms**

25           To look for desaturases involved in PUFA production, a cDNA library was constructed from total RNA isolated from *Phaeodactylum tricornutum*. A plasmid-based cDNA library was constructed in pSPORT1 (GIBCO-BRL)

following manufacturer's instructions using a commercially available kit (GIBCO-BRL). Random cDNA clones were sequenced and nucleic acid sequences that encode putative  $\Delta 5$  or  $\Delta 6$  desaturases were identified through BLAST search of the databases and comparison to Ma29 and Ma524 sequences.

5           One clone was identified from the *Phaeodactylum* library with homology to Ma29 and Ma524; it is called 144-011-B12. The DNA sequence is presented as SEQ ID NO:21. The amino acid sequence is presented as SEQ ID NO:22.

### **Example 11**

#### **10           Identification of *M. alpina* $\Delta 5$ and $\Delta 6$ homologues in other PUFA-producing organisms**

To look for desaturases involved in PUFA production, a cDNA library was constructed from total RNA isolated from *Schizochytrium* species. A plasmid-based cDNA library was constructed in pSPORT1 (GIBCO-BRL) following manufacturer's instructions using a commercially available kit (GIBCO-BRL). Random cDNA clones were sequenced and nucleic acid sequences that encode putative  $\Delta 5$  or  $\Delta 6$  desaturases were identified through BLAST search of the databases and comparison to Ma29 and Ma524 sequences.

20           One clone was identified from the *Schizochytrium* library with homology to Ma29 and Ma524; it is called 81-23-C7. This clone contains a ~1 kb insert. Partial sequence was obtained from each end of the clone using the universal forward and reverse sequencing primers. The DNA sequence from the forward primer is presented as SEQ ID NO:23. The peptide sequence is presented as SEQ ID NO:24. The DNA sequence from the reverse primer is presented as SEQ ID NO:25. The amino acid sequence from the reverse primer is presented as SEQ ID NO:26.



### Example 12

#### Human Desaturase Gene Sequences

Human desaturase gene sequences potentially involved in long chain polyunsaturated fatty acid biosynthesis were isolated based on homology  
5 between the human cDNA sequences and *Mortierella alpina* desaturase gene sequences. The three conserved "histidine boxes" known to be conserved among membrane-bound desaturases were found. As with some other membrane-bound desaturases the final HXXHH histidine box motif was found to be QXXHH. The amino acid sequence of the putative human desaturases  
10 exhibited homology to *M. alpina*  $\Delta 5$ ,  $\Delta 6$ ,  $\Delta 9$ , and  $\Delta 12$  desaturases.

The *M. alpina*  $\Delta 5$  desaturase and  $\Delta 6$  desaturase cDNA sequences were used to search the LifeSeq database of Incyte Pharmaceuticals, Inc., Palo Alto, California 94304. The  $\Delta 5$  desaturase sequence was divided into fragments; 1) amino acid no. 1-150, 2) amino acid no. 151-300, and 3) amino acid no. 301-  
15 446. The  $\Delta 6$  desaturase sequence was divided into three fragments; 1) amino acid no. 1-150, 2) amino acid no. 151-300, and 3) amino acid no. 301-457. These polypeptide fragments were searched against the database using the "tblastn" algorithm. This algorithm compares a protein query sequence against a nucleotide sequence database dynamically translated in all six reading frames  
20 (both strands).

The polypeptide fragments 2 and 3 of *M. alpina*  $\Delta 5$  and  $\Delta 6$  have homologies with the CloneID sequences as outlined in Table 6. The CloneID represents an individual sequence from the Incyte LifeSeq database. After the "tblastn" results have been reviewed, Clone Information was searched with the  
25 default settings of Stringency of  $\geq 50$ , and Productscore  $\leq 100$  for different CloneID numbers. The Clone Information Results displayed the information including the ClusterID, CloneID, Library, HitID, Hit Description. When selected, the ClusterID number displayed the clone information of all the clones that belong in that ClusterID. The Assemble command assembles all of the  
30 CloneID which comprise the ClusterID. The following default settings were

used for GCG (Genetics Computer Group, University of Wisconsin Biotechnology Center, Madison, Wisconsin 53705) Assembly:

	Word Size:	7
5	Minimum Overlap:	14
	Stringency:	0.8
	Minimum Identity:	14
	Maximum Gap:	10
	Gap Weight:	8
10	Length Weight:	2

GCG Assembly Results displayed the contigs generated on the basis of sequence information within the CloneID. A contig is an alignment of DNA sequences based on areas of homology among these sequences. A new sequence (consensus sequence) was generated based on the aligned DNA sequences within a contig. The contig containing the CloneID was identified, and the ambiguous sites of the consensus sequence was edited based on the alignment of the CloneIDs (see SEQ ID NO:27 - SEQ ID NO:32) to generate the best possible sequence. The procedure was repeated for all six CloneID listed in Table 6. This produced five unique contigs. The edited consensus sequences of the 5 contigs were imported into the Sequencher software program (Gene Codes Corporation, Ann Arbor, Michigan 48105). These consensus sequences were assembled. The contig 2511785 overlaps with contig 3506132, and this new contig was called 2535 (SEQ ID NO:33). The contigs from the Sequencher program were copied into the Sequence Analysis software package of GCG.

Each contig was translated in all six reading frames into protein sequences. The *M. alpina*  $\Delta 5$  (MA29) and  $\Delta 6$  (MA524) sequences were compared with each of the translated contigs using the FastA search (a Pearson

and Lipman search for similarity between a query sequence and a group of sequences of the same type (nucleic acid or protein)). Homology among these sequences suggest the open reading frames of each contig. The homology among the *M. alpina*  $\Delta 5$  and  $\Delta 6$  to contigs 2535 and 3854933 were utilized to create the final contig called 253538a. Figure 13 is the FastA match of the final contig 253538a and MA29, and Figure 14 is the FastA match of the final contig 253538a and MA524. The DNA sequences for the various contigs are presented in SEQ ID NO:27 -SEQ ID NO:33 The various peptide sequences are shown in SEQ ID NO:34 - SEQ ID NO: 40.

Although the open reading frame was generated by merging the two contigs, the contig 2535 shows that there is a unique sequence in the beginning of this contig which does not match with the contig 3854933. Therefore, it is possible that these contigs were generated from independent desaturase like human genes.

The contig 253538a contains an open reading frame encoding 432 amino acids. It starts with Gln (CAG) and ends with the stop codon (TGA). The contig 253538a aligns with both *M. alpina*  $\Delta 5$  and  $\Delta 6$  sequences, suggesting that it could be either of the desaturases, as well as other known desaturases which share homology with each other. The individual contigs listed in Table 18, as well as the intermediate contig 2535 and the final contig 253538a can be utilized to isolate the complete genes for human desaturases.

#### Uses of the human desaturases

These human sequences can be express in yeast and plants utilizing the procedures described in the preceding examples. For expression in mammalian cells transgenic animals, these genes may provide superior codon bias.

In addition, these sequences can be used to isolate related desaturase genes from other organisms.

**Table 6**

Sections of the Desaturases	Clone ID from LifeSeq Database	Keyword
-----------------------------	--------------------------------	---------

151-300 $\Delta 5$	3808675	fatty acid desaturase
301-446 $\Delta 5$	354535	$\Delta 6$
151-300 $\Delta 6$	3448789	$\Delta 6$
151-300 $\Delta 6$	1362863	$\Delta 6$
151-300 $\Delta 6$	2394760	$\Delta 6$
301-457 $\Delta 6$	3350263	$\Delta 6$

### Example 13

#### **I. INFANT FORMULATIONS**

##### **A. Isomil® Soy Formula with Iron.**

5            Usage: As a beverage for infants, children and adults with an allergy or sensitivity to cow's milk. A feeding for patients with disorders for which lactose should be avoided: lactase deficiency, lactose intolerance and galactosemia.

##### **Features:**

- 10            • Soy protein isolate to avoid symptoms of cow's-milk-protein allergy or sensitivity
- Lactose-free formulation to avoid lactose-associated diarrhea
- Low osmolality (240 mOsm/kg water) to reduce risk of osmotic diarrhea.
- 15            • Dual carbohydrates (corn syrup and sucrose) designed to enhance carbohydrate absorption and reduce the risk of exceeding the absorptive capacity of the damaged gut.
- 1.8 mg of Iron (as ferrous sulfate) per 100 Calories to help prevent iron deficiency.
- 20            • Recommended levels of vitamins and minerals.
- Vegetable oils to provide recommended levels of essential fatty acids.
- Milk-white color, milk-like consistency and pleasant aroma.

Ingredients: (Pareve, ®) 85% water, 4.9% corn syrup, 2.6% sugar (sucrose), 2.1% soy oil, 1.9% soy protein isolate, 1.4% coconut oil, 0.15% calcium citrate, 0.11 % calcium phosphate tribasic, potassium citrate, potassium phosphate monobasic, potassium chloride, mono- and diglycerides, soy  
5 lecithin, carrageenan, ascorbic acid, L-methionine, magnesium chloride, potassium phosphate dibasic, sodium chloride, choline chloride, taurine, ferrous sulfate, m-inositol, alpha-tocopheryl acetate, zinc sulfate, L-carnitine, niacinamide, calcium pantothenate, cupric sulfate, vitamin A palmitate, thiamine chloride hydrochloride, riboflavin, pyridoxine hydrochloride, folic  
10 acid, manganese sulfate, potassium iodide, phylloquinone, biotin, sodium selenite, vitamin D<sub>3</sub> and cyanocobalamin.

**B. Isomil® DF Soy Formula For Diarrhea.**

Usage: As a short-term feeding for the dietary management of diarrhea in infants and toddlers.

15 Features:

- First infant formula to contain added dietary fiber from soy fiber specifically for diarrhea management.
- Clinically shown to reduce the duration of loose, watery stools during mild to severe diarrhea in infants.
- 20 • Nutritionally complete to meet the nutritional needs of the infant.
- Soy protein isolate with added L-methionine meets or exceeds an infant's requirement for all essential amino acids.
- Lactose-free formulation to avoid lactose-associated diarrhea.
- Low osmolality (240 mOsm/kg water) to reduce the risk of  
25 osmotic diarrhea.
- Dual carbohydrates (corn syrup and sucrose) designed to enhance carbohydrate absorption and reduce the risk of exceeding the absorptive capacity of the damaged gut.

- Meets or exceeds the vitamin and mineral levels recommended by the Committee on Nutrition of the American Academy of Pediatrics and required by the Infant Formula Act.
- 1.8 mg of iron (as ferrous sulfate) per 100 Calories to help prevent iron deficiency.
- Vegetable oils to provide recommended levels of essential fatty acids.

Ingredients: (Pareve, ®) 86% water, 4.8% corn syrup, 2.5% sugar (sucrose), 2.1% soy oil, 2.0% soy protein isolate, 1.4% coconut oil, 0.77% soy fiber, 0.12% calcium citrate, 0.11 % calcium phosphate tribasic, 0.10% potassium citrate, potassium chloride, potassium phosphate monobasic, mono- and diglycerides, soy lecithin, carrageenan, magnesium chloride, ascorbic acid, L-methionine, potassium phosphate dibasic, sodium chloride, choline chloride, taurine, ferrous sulfate, m-inositol, alpha-tocopheryl acetate, zinc sulfate, L-carnitine, niacinamide, calcium pantothenate, cupric sulfate, vitamin A palmitate, thiamine chloride hydrochloride, riboflavin, pyridoxine hydrochloride, folic acid, manganese sulfate, potassium iodide, phylloquinone, biotin, sodium selenite, vitamin D<sub>3</sub> and cyanocobalamin.

**C. Isomil® SF Sucrose-Free Soy Formula With Iron.**

Usage: As a beverage for infants, children and adults with an allergy or sensitivity to cow's-milk protein or an intolerance to sucrose. A feeding for patients with disorders for which lactose and sucrose should be avoided.

**Features:**

- Soy protein isolate to avoid symptoms of cow's-milk-protein allergy or sensitivity.
- Lactose-free formulation to avoid lactose-associated diarrhea (carbohydrate source is Polycose® Glucose Polymers).
- Sucrose free for the patient who cannot tolerate sucrose.

- Low osmolality (180 mOsm/kg water) to reduce risk of osmotic diarrhea.
- 1.8 mg of iron (as ferrous sulfate) per 100 Calories to help prevent iron deficiency.
- 5      • Recommended levels of vitamins and minerals.
- Vegetable oils to provide recommended levels of essential fatty acids.
- Milk-white color, milk-like consistency and pleasant aroma.

Ingredients: (Pareve, ©) 75% water, 11.8% hydrolyzed cornstarch, 4.1%  
 10      soy oil, 4.1% soy protein isolate, 2.8% coconut oil, 1.0% modified cornstarch,  
 0.38% calcium phosphate tribasic, 0.17% potassium citrate, 0.13% potassium  
 chloride, mono- and diglycerides, soy lecithin, magnesium chloride, ascorbic  
 acid, L-methionine, calcium carbonate, sodium chloride, choline chloride,  
 15      carrageenan, taurine, ferrous sulfate, m-inositol, alpha-tocopheryl acetate, zinc  
 sulfate, L-carnitine, niacinamide, calcium pantothenate, cupric sulfate, vitamin  
 A palmitate, thiamine chloride hydrochloride, riboflavin, pyridoxine  
 hydrochloride, folic acid, manganese sulfate, potassium iodide, phylloquinone,  
 biotin, sodium selenite, vitamin D<sub>3</sub> and cyanocobalamin.

**D.      Isomil® 20 Soy Formula With Iron Ready To Feed,**  
 20      **20 Cal/fl oz.**

Usage: When a soy feeding is desired.

Ingredients: (Pareve, ©) 85% water, 4.9% corn syrup, 2.6% sugar  
 (sucrose), 2.1% soy oil, 1.9% soy protein isolate, 1.4% coconut oil, 0.15%  
 calcium citrate, 0.11% calcium phosphate tribasic, potassium citrate, potassium  
 25      phosphate monobasic, potassium chloride, mono- and diglycerides, soy  
 lecithin, carrageenan, ascorbic acid, L-methionine, magnesium chloride,  
 potassium phosphate dibasic, sodium chloride, choline chloride, taurine, ferrous  
 sulfate, m-inositol, alpha-tocopheryl acetate, zinc sulfate, L-carnitine,  
 niacinamide, calcium pantothenate, cupric sulfate, vitamin A palmitate,  
 30      thiamine chloride hydrochloride, riboflavin, pyridoxine hydrochloride, folic

acid, manganese sulfate, potassium iodide, phylloquinone, biotin, sodium selenite, vitamin D<sub>3</sub> and cyanocobalamin.

**E. Similac® Infant Formula**

5 Usage: When an infant formula is needed: if the decision is made to discontinue breastfeeding before age 1 year, if a supplement to breastfeeding is needed or as a routine feeding if breastfeeding is not adopted.

Features:

- 10 • Protein of appropriate quality and quantity for good growth; heat-denatured, which reduces the risk of milk-associated enteric blood loss.
- Fat from a blend of vegetable oils (doubly homogenized), providing essential linoleic acid that is easily absorbed.
- Carbohydrate as lactose in proportion similar to that of human milk.
- 15 • Low renal solute load to minimize stress on developing organs.
- Powder, Concentrated Liquid and Ready To Feed forms.

Ingredients: (®-D) Water, nonfat milk, lactose, soy oil, coconut oil, mono- and diglycerides, soy lecithin, ascorbic acid, carrageenan, choline chloride, taurine, m-inositol, alpha-tocopheryl acetate, zinc sulfate, niacinamid, 20 ferrous sulfate, calcium pantothenate, cupric sulfate, vitamin A palmitate, thiamine chloride hydrochloride, riboflavin, pyridoxine hydrochloride, folic acid, manganese sulfate, phylloquinone, biotin, sodium selenite, vitamin D<sub>3</sub> and cyanocobalamin.

**F. Similac® NeoCare Premature Infant Formula With Iron**

25 Usage: For premature infants' special nutritional needs after hospital discharge. Similac NeoCare is a nutritionally complete formula developed to provide premature infants with extra calories, protein, vitamins and minerals needed to promote catch-up growth and support development.

Features:



- Reduces the need for caloric and vitamin supplementation. More calories (22 Cal/fl oz) than standard term formulas (20 Cal/fl oz).
- Highly absorbed fat blend, with medium-chain triglycerides (MCT oil) to help meet the special digestive needs of premature infants.
- 5 • Higher levels of protein, vitamins and minerals per 100 Calories to extend the nutritional support initiated in-hospital.
- More calcium and phosphorus for improved bone mineralization.

Ingredients: @-D Corn syrup solids, nonfat milk, lactose, whey protein concentrate, soy oil, high-oleic safflower oil, fractionated coconut oil (medium-chain triglycerides), coconut oil, potassium citrate, calcium phosphate tribasic, calcium carbonate, ascorbic acid, magnesium chloride, potassium chloride, sodium chloride, taurine, ferrous sulfate, m-inositol, choline chloride, ascorbyl palmitate, L-carnitine, alpha-tocopheryl acetate, zinc sulfate, niacinamide, mixed tocopherols, sodium citrate, calcium pantothenate, cupric sulfate, thiamine chloride hydrochloride, vitamin A palmitate, beta carotene, riboflavin, pyridoxine hydrochloride, folic acid, manganese sulfate, phyloquinone, biotin, sodium selenite, vitamin D<sub>3</sub> and cyanocobalamin.

**G. Similac Natural Care Low-Iron Human Milk Fortifier Ready To Use, 24 Cal/fl oz.**

20 Usage: Designed to be mixed with human milk or to be fed alternatively with human milk to low-birth-weight infants.

Ingredients: @-D Water, nonfat milk, hydrolyzed cornstarch, lactose, fractionated coconut oil (medium-chain triglycerides), whey protein concentrate, soy oil, coconut oil, calcium phosphate tribasic, potassium citrate, magnesium chloride, sodium citrate, ascorbic acid, calcium carbonate, mono- and diglycerides, soy lecithin, carrageenan, choline chloride, m-inositol, taurine, niacinamide, L-carnitine, alpha tocopheryl acetate, zinc sulfate, potassium chloride, calcium pantothenate, ferrous sulfate, cupric sulfate, riboflavin, vitamin A palmitate, thiamine chloride hydrochloride, pyridoxine

hydrochloride, biotin, folic acid, manganese sulfate, phylloquinone, vitamin D<sub>3</sub>, sodium selenite and cyanocobalamin.

Various PUFAs of this invention can be substituted and/or added to the infant formulae described above and to other infant formulae known to those in the art..

## II. NUTRITIONAL FORMULATIONS

### A. ENSURE®

Usage: ENSURE is a low-residue liquid food designed primarily as an oral nutritional supplement to be used with or between meals or, in appropriate amounts, as a meal replacement. ENSURE is lactose- and gluten-free, and is suitable for use in modified diets, including low-cholesterol diets. Although it is primarily an oral supplement, it can be fed by tube.

#### Patient Conditions:

- For patients on modified diets
- For elderly patients at nutrition risk
- For patients with involuntary weight loss
- For patients recovering from illness or surgery
- For patients who need a low-residue diet

#### Ingredients:

©-D Water, Sugar (Sucrose), Maltodextrin (Corn), Calcium and Sodium Caseinates, High-Oleic Safflower Oil, Soy Protein Isolate, Soy Oil, Canola Oil, Potassium Citrate, Calcium Phosphate Tribasic, Sodium Citrate, Magnesium Chloride, Magnesium Phosphate Dibasic, Artificial Flavor, Sodium Chloride, Soy Lecithin, Choline Chloride, Ascorbic Acid, Carrageenan, Zinc Sulfate, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Gellan Gum, Niacinamide, Calcium Pantothenate, Manganese Sulfate, Cupric Sulfate, Vitamin A Palmitate, Thiamine Chloride Hydrochloride, Pyridoxine Hydrochloride, Riboflavin, Folic Acid, Sodium Molybdate, Chromium Chloride, Biotin, Potassium Iodide, Sodium Selenate.

## **B. ENSURE® BARS**

Usage: ENSURE BARS are complete, balanced nutrition for supplemental use between or with meals. They provide a delicious, nutrient-rich alternative to other snacks. ENSURE BARS contain <1 g lactose/bar, and  
5 Chocolate Fudge Brownie flavor is gluten-free. (Honey Graham Crunch flavor contains gluten.)

### **Patient Conditions:**

- For patients who need extra calories, protein, vitamins and minerals
- 10 • Especially useful for people who do not take in enough calories and nutrients
- For people who have the ability to chew and swallow
- Not to be used by anyone with a peanut allergy or any type of allergy to nuts.

### **Ingredients:**

Honey Graham Crunch -- High-Fructose Corn Syrup, Soy Protein Isolate, Brown Sugar, Honey, Maltodextrin (Corn), Crisp Rice (Milled Rice, Sugar [Sucrose], Salt [Sodium Chloride] and Malt), Oat Bran, Partially  
Hydrogenated Cottonseed and Soy Oils, Soy Polysaccharide, Glycerine, Whey  
20 Protein Concentrate, Polydextrose, Fructose, Calcium Caseinate, Cocoa Powder, Artificial Flavors, Canola Oil, High-Oleic Safflower Oil, Nonfat Dry Milk, Whey Powder, Soy Lecithin and Corn Oil. Manufactured in a facility that processes nuts.

### **Vitamins and Minerals:**

25 Calcium Phosphate Tribasic, Potassium Phosphate Dibasic, Magnesium Oxide, Salt (Sodium Chloride), Potassium Chloride, Ascorbic Acid, Ferric Orthophosphate, Alpha-Tocopheryl Acetate, Niacinamide, Zinc Oxide, Calcium Pantothenate, Copper Gluconate, Manganese Sulfate, Riboflavin, Beta-Carotene, Pyridoxine Hydrochloride, Thiamine Mononitrate, Folic Acid, Biotin,

Chromium Chloride, Potassium Iodide, Sodium Selenate, Sodium Molybdate, Phylloquinone, Vitamin D<sub>3</sub> and Cyanocobalamin.

**Protein:**

5 **Honey Graham Crunch** - The protein source is a blend of soy protein isolate and milk proteins.

Soy protein isolate	74%
Milk proteins	26%

**Fat:**

10 **Honey Graham Crunch** - The fat source is a blend of partially hydrogenated cottonseed and soybean, canola, high oleic safflower, and corn oils, and soy lecithin.

	Partially hydrogenated cottonseed and soybean oil	76%
	Canola oil	8%
	High-oleic safflower oil	8%
15	Corn oil	4%
	Soy lecithin	4%

**Carbohydrate:**

20 **Honey Graham Crunch** - The carbohydrate source is a combination of high-fructose corn syrup, brown sugar, maltodextrin, honey, crisp rice, glycerine, soy polysaccharide, and oat bran.

	High-fructose corn syrup	24%
	Brown sugar	21%
	Maltodextrin	12%
	Honey	11%
25	Crisp rice	9%
	Glycerine	9%
	Soy polysaccharide	7%
	Oat bran	7%

### C. ENSURE® HIGH PROTEIN

Usage: ENSURE HIGH PROTEIN is a concentrated, high-protein liquid food designed for people who require additional calories, protein, vitamins, and minerals in their diets. It can be used as an oral nutritional supplement with or between meals or, in appropriate amounts, as a meal replacement. ENSURE HIGH PROTEIN is lactose- and gluten-free, and is suitable for use by people recovering from general surgery or hip fractures and by patients at risk for pressure ulcers.

#### Patient Conditions

- For patients who require additional calories, protein, vitamins, and minerals, such as patients recovering from general surgery or hip fractures, patients at risk for pressure ulcers, and patients on low-cholesterol diets

#### Features-

- Low in saturated fat
- Contains 6 g of total fat and < 5 mg of cholesterol per serving
- Rich, creamy taste
- Excellent source of protein, calcium, and other essential vitamins and minerals
- For low-cholesterol diets
- Lactose-free, easily digested

#### Ingredients:

**Vanilla Supreme:** -D Water, Sugar (Sucrose), Maltodextrin (Corn), Calcium and Sodium Caseinates, High-Oleic Safflower Oil, Soy Protein Isolate, Soy Oil, Canola Oil, Potassium Citrate, Calcium Phosphate Tribasic, Sodium Citrate, Magnesium Chloride, Magnesium Phosphate Dibasic, Artificial Flavor, Sodium Chloride, Soy Lecithin, Choline Chloride, Ascorbic Acid, Carrageenan, Zinc Sulfate, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Gellan Gum, Niacinamide, Calcium Pantothenate, Manganese Sulfate, Cupric Sulfate, Vitamin A Palmitate, Thiamine Chloride Hydrochloride, Pyridoxine Hydrochloride,

Riboflavin, Folic Acid, Sodium Molybdate, Chromium Chloride, Biotin, Potassium Iodide, Sodium Selenate, Phylloquinone, Vitamin D<sub>3</sub> and Cyanocobalamin.

**Protein:**

- 5 The protein source is a blend of two high-biologic-value proteins: casein and soy.

Sodium and calcium caseinates	85%
Soy protein isolate	15%

**Fat:**

- 10 The fat source is a blend of three oils: high-oleic safflower, canola, and soy.

High-oleic safflower oil	40%
Canola oil	30%
Soy oil	30%

- 15 The level of fat in ENSURE HIGH PROTEIN meets American Heart Association (AHA) guidelines. The 6 grams of fat in ENSURE HIGH PROTEIN represent 24% of the total calories, with 2.6% of the fat being from saturated fatty acids and 7.9% from polyunsaturated fatty acids. These values are within the AHA guidelines of  $\leq 30\%$  of total calories from fat,  $< 10\%$  of the calories from saturated fatty acids, and  $\leq 10\%$  of total calories from
- 20 polyunsaturated fatty acids.

**Carbohydrate:**

- 25 ENSURE HIGH PROTEIN contains a combination of maltodextrin and sucrose. The mild sweetness and flavor variety (vanilla supreme, chocolate royal, wild berry, and banana), plus VARI-FLAVORSO® Flavor Pacs in pecan, cherry, strawberry, lemon, and orange, help to prevent flavor fatigue and aid in patient compliance.

**Vanilla and other nonchocolate flavors**

Sucrose	60%
---------	-----

Maltodextrin	40%
<b>Chocolate</b>	
Sucrose	70%
Maltodextrin	30%

5

**D. ENSURE® LIGHT**

Usage: ENSURE LIGHT is a low-fat liquid food designed for use as an oral nutritional supplement with or between meals. ENSURE LIGHT is lactose- and gluten-free, and is suitable for use in modified diets, including low-cholesterol diets.

10

**Patient Conditions:**

- For normal-weight or overweight patients who need extra nutrition in a supplement that contains 50% less fat and 20% fewer calories than ENSURE
- For healthy adults who don't eat right and need extra nutrition

15

**Features:**

- Low in fat and saturated fat
- Contains 3 g of total fat per serving and < 5 mg cholesterol
- Rich, creamy taste
- Excellent source of calcium and other essential vitamins and minerals
- For low-cholesterol diets
- Lactose-free, easily digested

20

**Ingredients:**

**French Vanilla:** ®-D Water, Maltodextrin (Corn), Sugar (Sucrose), Calcium Caseinate, High-Oleic Safflower Oil, Canola Oil, Magnesium Chloride, Sodium Citrate, Potassium Citrate, Potassium Phosphate Dibasic, Magnesium Phosphate Dibasic, Natural and Artificial Flavor, Calcium Phosphate Tribasic, Cellulose Gel, Choline Chloride, Soy Lecithin, Carrageenan, Salt (Sodium Chloride),

25

Ascorbic Acid, Cellulose Gum, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Zinc Sulfate, Niacinamide, Manganese Sulfate, Calcium Pantothenate, Cupric Sulfate, Thiamine Chloride Hydrochloride, Vitamin A Palmitate, Pyridoxine Hydrochloride, Riboflavin, Chromium Chloride, Folic Acid, Sodium  
 5 Molybdate, Biotin, Potassium Iodide, Sodium Selenate, Phylloquinone, Vitamin D<sub>3</sub> and Cyanocobalamin.

**Protein:**

The protein source is calcium caseinate.

10	Calcium caseinate	100%
----	-------------------	------

**Fat**

The fat source is a blend of two oils: high-oleic safflower and canola.

	High-oleic safflower oil	70%
	Canola oil	30%

15 The level of fat in ENSURE LIGHT meets American Heart Association (AHA) guidelines. The 3 grams of fat in ENSURE LIGHT represent 13.5% of the total calories, with 1.4% of the fat being from saturated fatty acids and 2.6% from polyunsaturated fatty acids. These values are within the AHA guidelines of  $\leq 30\%$  of total calories from fat,  $< 10\%$  of the calories from saturated fatty acids, and  $\leq 10\%$  of total calories from polyunsaturated fatty acids.

20 **Carbohydrate**

ENSURE LIGHT contains a combination of maltodextrin and sucrose. The chocolate flavor contains corn syrup as well. The mild sweetness and flavor variety (French vanilla, chocolate supreme, strawberry swirl), plus VARI-FLAVORS® Flavor Pacs in pecan, cherry, strawberry, lemon, and  
 25 orange, help to prevent flavor fatigue and aid in patient compliance.

**Vanilla and other nonchocolate flavors**

Sucrose	51%
Maltodextrin	49%



**Chocolate**

Sucrose	47.0%
Corn Syrup	26.5%
Maltodextrin	26.5%

5 **Vitamins and Minerals**

An 8-fl-oz serving of ENSURE LIGHT provides at least 25% of the RDIs for 24 key vitamins and minerals.

**Caffeine**

Chocolate flavor contains 2.1 mg caffeine/8 fl oz.

10

**E. ENSURE PLUS®**

Usage: ENSURE PLUS is a high-calorie, low-residue liquid food for use when extra calories and nutrients, but a normal concentration of protein, are needed. It is designed primarily as an oral nutritional supplement to be used with or between meals or, in appropriate amounts, as a meal replacement. ENSURE PLUS is lactose- and gluten-free. Although it is primarily an oral nutritional supplement, it can be fed by tube.

15

**Patient Conditions:**

- For patients who require extra calories and nutrients, but a normal concentration of protein, in a limited volume
- For patients who need to gain or maintain healthy weight

20

**Features**

- Rich, creamy taste
- Good source of essential vitamins and minerals

25 **Ingredients**

**Vanilla:** ®-D Water, Corn Syrup, Maltodextrin (Corn), Corn Oil, Sodium and Calcium Caseinates, Sugar (Sucrose), Soy Protein Isolate, Magnesium Chloride,

Potassium Citrate, Calcium Phosphate Tribasic, Soy Lecithin, Natural and Artificial Flavor, Sodium Citrate, Potassium Chloride, Choline Chloride, Ascorbic Acid, Carrageenan, Zinc Sulfate, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Niacinamide, Calcium Pantothenate, Manganese Sulfate, Cupric Sulfate, Thiamine Chloride Hydrochloride, Pyridoxine Hydrochloride, Riboflavin, Vitamin A Palmitate, Folic Acid, Biotin, Chromium Chloride, Sodium Molybdate, Potassium Iodide, Sodium Selenite, Phylloquinone, Cyanocobalamin and Vitamin D<sub>3</sub>.

#### **Protein**

The protein source is a blend of two high-biologic-value proteins: casein and soy.

Sodium and calcium caseinates	84%
Soy protein isolate	16%

#### **Fat**

The fat source is corn oil.

Corn oil	100%
----------	------

#### **Carbohydrate**

ENSURE PLUS contains a combination of maltodextrin and sucrose. The mild sweetness and flavor variety (vanilla, chocolate, strawberry, coffee, butter pecan, and eggnog), plus VARI-FLAVORS® Flavor Pacs in pecan, cherry, strawberry, lemon, and orange, help to prevent flavor fatigue and aid in patient compliance.

#### **Vanilla, strawberry, butter pecan, and coffee flavors**

Corn Syrup	39%
Maltodextrin	38%
Sucrose	23%

#### **Chocolate and eggnog flavors**

Corn Syrup	36%
------------	-----

Maltodextrin	34%
Sucrose	30%

#### **Vitamins and Minerals**

5           An 8-fl-oz serving of ENSURE PLUS provides at least 15% of the RDIs for 25 key Vitamins and minerals.

#### **Caffeine**

          Chocolate flavor contains 3.1 mg Caffeine/8 fl oz. Coffee flavor contains a trace amount of caffeine.

### 10           **F.     ENSURE PLUS® HN**

          Usage: ENSURE PLUS HN is a nutritionally complete high-calorie, high-nitrogen liquid food designed for people with higher calorie and protein needs or limited volume tolerance. It may be used for oral supplementation or for total nutritional support by tube. ENSURE PLUS HN is lactose- and gluten-free.

15

#### **Patient Conditions:**

- For patients with increased calorie and protein needs, such as following surgery or injury
- For patients with limited volume tolerance and early satiety

### 20           **Features**

- For supplemental or total nutrition
- For oral or tube feeding
- 1.5 CaV/mL
- High nitrogen
- 25   • Calorically dense

#### **Ingredients**

**Vanilla:** ®-D Water, Maltodextrin (Corn), Sodium and Calcium Caseinates, Corn Oil, Sugar (Sucrose), Soy Protein Isolate, Magnesium Chloride, Potassium Citrate, Calcium Phosphate Tribasic, Soy Lecithin, Natural and Artificial Flavor, Sodium Citrate, Choline Chloride, Ascorbic Acid, Taurine, L-Carnitine, 5 Zinc Sulfate, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Niacinamide, Carrageenan, Calcium Pantothenate, Manganese Sulfate, Cupric Sulfate, Thiamine Chloride Hydrochloride, Pyridoxine Hydrochloride, Riboflavin, Vitamin A Palmitate, Folic Acid, Biotin, Chromium Chloride, Sodium Molybdate, Potassium Iodide, Sodium Selenite, Phylloquinone, 10 Cyanocobalamin and Vitamin D<sub>3</sub>.

#### **G. ENSURE® POWDER**

Usage: ENSURE POWDER (reconstituted with water) is a low-residue liquid food designed primarily as an oral nutritional supplement to be used with 15 or between meals. ENSURE POWDER is lactose- and gluten-free, and is suitable for use in modified diets, including low-cholesterol diets.

##### **Patient Conditions:**

- For patients on modified diets
- For elderly patients at nutrition risk
- 20 • For patients recovering from illness/surgery
- For patients who need a low-residue diet

##### **Features**

- Convenient, easy to mix
- Low in saturated fat
- 25 • Contains 9 g of total fat and < 5 mg of cholesterol per serving
- High in vitamins and minerals
- For low-cholesterol diets
- Lactose-free, easily digested

**Ingredients:** @-D Corn Syrup, Maltodextrin (Corn), Sugar (Sucrose), Corn Oil, Sodium and Calcium Caseinates, Soy Protein Isolâte, Artificial Flavor, Potassium Citrate, Magnesium Chloride, Sodium Citrate, Calcium Phosphate Tribasic, Potassium Chloride, Soy Lecithin, Ascorbic Acid, Choline Chloride, Zinc Sulfate, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Niacinamide, Calcium Pantothenate, Manganese Sulfate, Thiamine Chloride Hydrochloride, Cupric Sulfate, Pyridoxine Hydrochloride, Riboflavin, Vitamin A Palmitate, Folic Acid, Biotin, Sodium Molybdate, Chromium Chloride, Potassium Iodide, Sodium Selenate, Phylloquinone, Vitamin D<sub>3</sub> and Cyanocobalamin.

#### 10 **Protein**

The protein source is a blend of two high-biologic-value proteins: casein and soy.

Sodium and calcium caseinates	84%
Soy protein isolate	16%

#### 15 **Fat**

The fat source is corn oil.

Corn oil	100%
----------	------

#### **Carbohydrate**

ENSURE POWDER contains a combination of corn syrup, maltodextrin, and sucrose. The mild sweetness of ENSURE POWDER, plus VARI-FLAVORS® Flavor Pacs in pecan, cherry, strawberry, lemon, and orange, helps to prevent flavor fatigue and aid in patient compliance.

#### **Vanilla**

Corn Syrup	35%
Maltodextrin	35%
Sucrose	30%

## H. ENSURE® PUDDING

Usage: ENSURE PUDDING is a nutrient-dense supplement providing balanced nutrition in a nonliquid form to be used with or between meals. It is appropriate for consistency-modified diets (e.g., soft, pureed, or full liquid) or for people with swallowing impairments. ENSURE PUDDING is gluten-free.

### Patient Conditions:

- For patients on consistency-modified diets (e.g., soft, pureed, or full liquid)
- For patients with swallowing impairments

### Features

- Rich and creamy, good taste
- Good source of essential vitamins and minerals    Convenient-needs no refrigeration
- Gluten-free

**Nutrient Profile per 5 oz:** Calories 250, Protein 10.9%, Total Fat 34.9%, Carbohydrate 54.2%

### Ingredients:

**Vanilla:** ®-D Nonfat Milk, Water, Sugar (Sucrose), Partially Hydrogenated Soybean Oil, Modified Food Starch, Magnesium Sulfate. Sodium Stearoyl Lactylate, Sodium Phosphate Dibasic, Artificial Flavor, Ascorbic Acid, Zinc Sulfate, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Choline Chloride, Niacinamide, Manganese Sulfate, Calcium Pantothenate, FD&C Yellow #5, Potassium Citrate, Cupric Sulfate, Vitamin A Palmitate, Thiamine Chloride Hydrochloride, Pyridoxine Hydrochloride, Riboflavin, FD&C Yellow #6, Folic Acid, Biotin, Phylloquinone, Vitamin D3 and Cyanocobalamin.

### Protein

The protein source is nonfat milk.

Nonfat milk

100%

**Fat**

The fat source is hydrogenated soybean oil.

Hydrogenated soybean oil	100%
--------------------------	------

**Carbohydrate**

5           ENSURE PUDDING contains a combination of sucrose and modified food starch. The mild sweetness and flavor variety (vanilla, chocolate, butterscotch, and tapioca) help prevent flavor fatigue. The product contains 9.2 grams of lactose per serving.

**Vanilla and other nonchocolate flavors**

10	Sucrose	56%
	Lactose	27%
	Modified food starch	17%

**Chocolate**

	Sucrose	58%
15	Lactose	26%
	Modified food starch	16%

**I. ENSURE® WITH FIBER**

20           Usage: ENSURE WITH FIBER is a fiber-containing, nutritionally complete liquid food designed for people who can benefit from increased dietary fiber and nutrients. ENSURE WITH FIBER is suitable for people who do not require a low-residue diet. It can be fed orally or by tube, and can be used as a nutritional supplement to a regular diet or, in appropriate amounts, as a meal replacement. ENSURE WITH FIBER is lactose- and gluten-free, and is

25           suitable for use in modified diets, including low-cholesterol diets.

**Patient Conditions**

- For patients who can benefit from increased dietary fiber and nutrients

**Features**

- New advanced formula-low in saturated fat, higher in vitamins and minerals
- Contains 6 g of total fat and < 5 mg of cholesterol per serving
- Rich, creamy taste
- 5      • Good source of fiber
- Excellent source of essential vitamins and minerals
- For low-cholesterol diets
- Lactose- and gluten-free

**Ingredients**

- 10      **Vanilla:** ®-D Water, Maltodextrin (Corn), Sugar (Sucrose), Sodium and Calcium Caseinates, Oat Fiber, High-Oleic Safflower Oil, Canola Oil, Soy Protein Isolate, Corn Oil, Soy Fiber, Calcium Phosphate Tribasic, Magnesium Chloride, Potassium Citrate, Cellulose Gel, Soy Lecithin, Potassium Phosphate Dibasic, Sodium Citrate, Natural and Artificial Flavors, Choline Chloride,
- 15      Magnesium Phosphate, Ascorbic Acid, Cellulose Gum, Potassium Chloride, Carrageenan, Ferrous Sulfate, Alpha-Tocopheryl Acetate, Zinc Sulfate, Niacinamide, Manganese Sulfate, Calcium Pantothenate, Cupric Sulfate, Vitamin A Palmitate, Thiamine Chloride Hydrochloride, Pyridoxine Hydrochloride, Riboflavin, Folic Acid, Chromium Chloride, Biotin, Sodium
- 20      Molybdate, Potassium Iodide, Sodium Selenate, Phylloquinone, Vitamin D<sub>3</sub> and Cyanocobalamin.

**Protein**

The protein source is a blend of two high-biologic-value proteins- casein and soy.

25	Sodium and calcium caseinates	80%
	Soy protein isolate	20%



**Fat**

The fat source is a blend of three oils: high-oleic safflower, canola, and corn.

	High-oleic safflower oil	40%
5	Canola oil	40%
	Corn oil	20%

The level of fat in ENSURE WITH FIBER meets American Heart Association (AHA) guidelines. The 6 grams of fat in ENSURE WITH FIBER represent 22% of the total calories, with 2.01 % of the fat being from saturated fatty acids and 6.7% from polyunsaturated fatty acids. These values are within the AHA guidelines of  $\leq 30\%$  of total calories from fat,  $< 10\%$  of the calories from saturated fatty acids, and  $\leq 10\%$  of total calories from polyunsaturated fatty acids.

**Carbohydrate**

ENSURE WITH FIBER contains a combination of maltodextrin and sucrose. The mild sweetness and flavor variety (vanilla, chocolate, and butter pecan), plus VARI-FLAVORS® Flavor Pacs in pecan, cherry, strawberry, lemon, and orange, help to prevent flavor fatigue and aid in patient compliance.

**Vanilla and other nonchocolate flavors**

20	Maltodextrin	66%
	Sucrose	25%
	Oat Fiber	7%
	Soy Fiber	2%

**Chocolate**

25	Maltodextrin	55%
	Sucrose	36%
	Oat Fiber	7%

Soy Fiber

2%

**Fiber**

The fiber blend used in ENSURE WITH FIBER consists of oat fiber and soy polysaccharide. This blend results in approximately 4 grams of total dietary fiber per 8-fl-oz can. The ratio of insoluble to soluble fiber is 95:5.

The various nutritional supplements described above and known to others of skill in the art can be substituted and/or supplemented with the PUFAs of this invention.

**J. Oxepa™ Nutritional Product**

Oxepa is low-carbohydrate, calorically dense enteral nutritional product designed for the dietary management of patients with or at risk for ARDS. It has a unique combination of ingredients, including a patented oil blend containing eicosapentaenoic acid (EPA from fish oil),  $\gamma$ -linolenic acid (GLA from borage oil), and elevated antioxidant levels.

**Caloric Distribution:**

- Caloric density is high at 1.5 Cal/mL (355 Cal/8 fl oz), to minimize the volume required to meet energy needs.
- The distribution of Calories in Oxepa is shown in Table 7.

Table 7. Caloric Distribution of Oxepa			
	per 8 fl oz.	per liter	% of Cal
Calories	355	1,500	---
Fat (g)	22.2	93.7	55.2
Carbohydrate (g)	25	105.5	28.1
Protein (g)	14.8	62.5	16.7
Water (g)	186	785	---

**Fat:**

- Oxepa contains 22.2 g of fat per 8-fl oz serving (93.7 g/L).
- The fat source is a oil blend of 31.8% canola oil, 25% medium-chain triglycerides (MCTs), 20% borage oil, 20% fish oil, and 3.2 % soy lecithin. The typical fatty acid profile of Oxepa is shown in Table 8.

- Oxepa provides a balanced amount of polyunsaturated, monounsaturated, and saturated fatty acids, as shown in Table 10.
- Medium-chain triglycerides (MCTs) -- 25% of the fat blend -- aid gastric emptying because they are absorbed by the intestinal tract without emulsification by bile acids.

5

The various fatty acid components of Oxepa™ nutritional product can be substituted and/or supplemented with the PUFAs of this invention.

Table 8. Typical Fatty Acid Profile			
	% Total Fatty Acids	g/8 fl oz*	g/L*
Caproic (6:0)	0.2	0.04	0.18
Caprylic (8:0)	14.69	3.1	13.07
Capric (10:0)	11.06	2.33	9.87
Palmitic (16:0)	5.59	1.18	4.98
Palmitoleic (16:1n-7)	1.82	0.38	1.62
Stearic (18:0)	1.84	0.39	1.64
Oleic (18:1n-9)	24.44	5.16	21.75
Linoleic (18:2n-6)	16.28	3.44	14.49
$\alpha$ -Linolenic (18:3n-3)	3.47	0.73	3.09
$\gamma$ -Linolenic (18:3n-6)	4.82	1.02	4.29
Eicosapentaenoic (20:5n-3)	5.11	1.08	4.55
n-3-Docosapentaenoic (22:5n-3)	0.55	0.12	0.49
Docosahexaenoic (22:6n-3)	2.27	0.48	2.02
Others	7.55	1.52	6.72

\* Fatty acids equal approximately 95% of total fat.

Table 9. Fat Profile of Oxepa.	
% of total calories from fat	55.2
Polyunsaturated fatty acids	31.44 g/L
Monounsaturated fatty acids	25.53 g/L
Saturated fatty acids	32.38 g/L
n-6 to n-3 ratio	1.75:1
Cholesterol	9.49 mg/8 fl oz 40.1 mg/L

**Carbohydrate:**

- The carbohydrate content is 25.0 g per 8-fl-oz serving (105.5 g/L).
- The carbohydrate sources are 45% maltodextrin (a complex carbohydrate) and 55% sucrose (a simple sugar), both of which are readily digested and absorbed.
- The high-fat and low-carbohydrate content of Oxepa is designed to minimize carbon dioxide (CO<sub>2</sub>) production. High CO<sub>2</sub> levels can complicate weaning in ventilator-dependent patients. The low level of carbohydrate also may be useful for those patients who have developed stress-induced hyperglycemia.
- Oxepa is lactose-free.

Dietary carbohydrate, the amino acids from protein, and the glycerol moiety of fats can be converted to glucose within the body. Throughout this process, the carbohydrate requirements of glucose-dependent tissues (such as the central nervous system and red blood cells) are met. However, a diet free of carbohydrates can lead to ketosis, excessive catabolism of tissue protein, and loss of fluid and electrolytes. These effects can be prevented by daily ingestion of 50 to 100 g of digestible carbohydrate, if caloric intake is adequate. The carbohydrate level in Oxepa is also sufficient to minimize gluconeogenesis, if energy needs are being met.

**Protein:**

- Oxepa contains 14.8 g of protein per 8-fl-oz serving (62.5 g/L).
- The total calorie/nitrogen ratio (150:1) meets the need of stressed patients.
- Oxepa provides enough protein to promote anabolism and the maintenance of lean body mass without precipitating respiratory problems. High protein intakes are a concern in patients with respiratory insufficiency. Although protein has little effect on CO<sub>2</sub> production, a high protein diet will increase ventilatory drive.

- The protein sources of Oxepa are 86.8% sodium caseinate and 13.2% calcium caseinate.
- As demonstrated in Table 11, the amino acid profile of the protein system in Oxepa meets or surpasses the standard for high quality protein set by the National Academy of Sciences.
- Oxepa is gluten-free.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

Throughout the specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The reference to any prior art in this specification is not, and should not be taken as, an acknowledgement or any form of suggestion that the prior art forms part of the common general knowledge in Australia.



## SEQUENCE LISTING

- 5 (1) GENERAL INFORMATION:
- 10 (i) APPLICANT: KNUTZON, DEBORAH  
MURKERJI, PRADIP  
HUANG, YUNG-SHENG  
THURMOND, JENNIFER  
CHAUDHARY, SUNITA  
LEONARD, AMANDA
- 15 (ii) TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS
- (iii) NUMBER OF SEQUENCES: 40
- 20 (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: LIMBACH AND LIMBACH LLP  
(B) STREET: 2001 FERRY BUILDING  
(C) CITY: SAN FRANCISCO  
(D) STATE: CA  
25 (E) COUNTRY: USA  
(F) ZIP: 94111
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
30 (B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: Microsoft Word
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
35 (B) (B) FILING DATE:  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
40 (A) NAME: WARD, MICHAEL R.  
(B) REGISTRATION NUMBER: 38,651  
(C) REFERENCE/DOCKET NUMBER: CGAB-210
- (ix) TELECOMMUNICATION INFORMATION:  
45 (A) TELEPHONE: (415) 433-4150  
(B) TELEFAX: (415) 433-8716  
(C) TELEX: N/A
- 50 (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1617 base pairs  
(B) TYPE: nucleic acid  
55 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGACACTCCT TCCTTCTTCT CACCCGTCCT AGTCCCCTTC AACCCCCCTC TTTGACAAAG

60

5 ACAACAAACC ATGGCTGCTG CTCCCAGTGT GAGGACGTTT ACTCGGGCCG AGGTTTGTAA 120  
 TGCCGAGGCT CTGAATGAGG GCAAGAAGGA TGCCGAGGCA CCCTTCTTGA TGATCATCGA 180  
 CAACAAGGTG TACGATGTCC GCGAGTTCGT CCCTGATCAT CCCGGTGGAA GTGTGATTCT 240  
 CACGCACGTT GGCAAGGACG GCACTGACGT CTTTGACACT TTTCACCCCG AGGCTGCTTG 300  
 10 GGAGACTCTT GCCAACTTTT ACGTTGGTGA TATTGACGAG AGCGACCGCG ATATCAAGAA 360  
 TGATGACTTT GCGGCCGAGG TCCGCAAGCT GCGTACCTTG TTCCAGTCTC TTGGTTACTA 420  
 CGATTCTTCC AAGGCATACT ACGCCTTCAA GGTCTCGTTC AACCTCTGCA TCTGGGGTTT 480  
 15 GTCGACGGTC ATTGTGGCCA AGTGGGGCCA GACCTCGACC CTCGCCAACG TGCTCTCGGC 540  
 TCGCCTTTTG GGTCTGTTCT GGCAGCAGTG CGGATGGTTG GCTCACGACT TTTTGCATCA 600  
 20 CCAGGTCTTC CAGGACCGTT TCTGGGGTGA TCTTTTCGGC GCCTTCTTGG GAGGTGTCTG 660  
 CCAGGCTTC TCGTCCTCGT GGTGGAAGGA CAAGCACAACT ACTCACCACG CCGCCCCCAA 720  
 CGTCCACGGC GAGGATCCCG ACATTGACAC CCACCCTCTG TTGACCTGGA GTGAGCATGC 780  
 25 GTTGGAGATG TTCTCGGATG TCCCAGATGA GGAGCTGACC CGCATGTGGT CGCGTTTCAT 840  
 GGTCTGAAC CAGACCTGGT TTTACTTCCC CATCTCTCG TTTGCCGTC TCTCTGGTG 900  
 30 CCTCCAGTCC ATTCTCTTTG TGCTGCCTAA CGGTCAGGCC CACAAGCCCT CGGGCGCGCG 960  
 TGTGCCCATC TCGTTGGTCG AGCAGCTGTC GCTTGCATG CACTGGACCT GGTACCTCGC 1020  
 CACCATGTTT CTGTTTCATCA AGGATCCCGT CAACATGCTG GTGTACTTTT TGGTGTGCGA 1080  
 35 GGCAGGTGTC GGAAACTTGT TGGCGATCGT GTTCTCGCTC AACCACAACG GTATGCCTGT 1140  
 GATCTCGAAG GAGGAGGCGG TCGATATGGA TTTCTTCACG AAGCAGATCA TCACGGGTGCG 1200  
 40 TGATGTCCAC CCGGGTCTAT TTGCCAACTG GTTACGGGT GGATTGAACT ATCAGATCGA 1260  
 GCACCACTTG TTCCCTTCGA TGCCTCGCCA CAACTTTTCA AAGATCCAGC CTGCTGTGCA 1320  
 GACCCTGTGC AAAAAGTACA ATGTCCGATA CCACACCACC GGTATGATCG AGGGAAGTGC 1380  
 45 AGAGGTCTTT AGCCGTCTGA ACGAGGTCTC CAAGGCTGCC TCCAAGATGG GTAAGGCGCA 1440  
 GTAAAAA AAAACAAGGAC GTTTTTTTTC GCCAGTGCCT GTGCCTGTGC CTGCTTCCCT 1500  
 50 TGTCAAGTCG AGCGTTTCTG GAAAGGATCG TTCAGTGCAG TATCATCATT CTCCTTTTAC 1560  
 CCCCCGCTCA TATCTCATTC ATTTCTCTTA TTAAACAACT TGTCCCCC TTCACCG 1617

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 457 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

5	Met	Ala	Ala	Ala	Pro	Ser	Val	Arg	Thr	Phe	Thr	Arg	Ala	Glu	Val	Leu	1	5	10	15
	Asn	Ala	Glu	Ala	Leu	Asn	Glu	Gly	Lys	Lys	Asp	Ala	Glu	Ala	Pro	Phe	20	25	30	
10	Leu	Met	Ile	Ile	Asp	Asn	Lys	Val	Tyr	Asp	Val	Arg	Glu	Phe	Val	Pro	35	40	45	
	Asp	His	Pro	Gly	Gly	Ser	Val	Ile	Leu	Thr	His	Val	Gly	Lys	Asp	Gly	50	55	60	
15	Thr	Asp	Val	Phe	Asp	Thr	Phe	His	Pro	Glu	Ala	Ala	Trp	Glu	Thr	Leu	65	70	75	80
	Ala	Asn	Phe	Tyr	Val	Gly	Asp	Ile	Asp	Glu	Ser	Asp	Arg	Asp	Ile	Lys	85	90	95	
20	Asn	Asp	Asp	Phe	Ala	Ala	Glu	Val	Arg	Lys	Leu	Arg	Thr	Leu	Phe	Gln	100	105	110	
25	Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Ala	Tyr	Tyr	Ala	Phe	Lys	Val	115	120	125	
	Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Thr	Val	Ile	Val	Ala	Lys	130	135	140	
30	Trp	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Val	Leu	Ser	Ala	Ala	Leu	Leu	145	150	155	160
	Gly	Leu	Phe	Trp	Gln	Gln	Cys	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His	165	170	175	
35	His	Gln	Val	Phe	Gln	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	Ala	Phe	180	185	190	
40	Leu	Gly	Gly	Val	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	Asp	Lys	195	200	205	
	His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp	210	215	220	
45	Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met	225	230	235	240
	Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe	245	250	255	
50	Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala	260	265	270	
55	Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro	Asn	Gly	275	280	285	
	Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu	290	295	300	
60	Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe	305	310	315	320
	Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Leu	Val	Tyr	Phe	Leu	Val	Ser	325	330	335	
65	Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His				



	340	345	350
	Asn Gly Met Pro Val Ile Ser Lys Glu Glu Ala Val Asp Met Asp Phe		
	355	360	365
5	Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe		
	370	375	380
10	Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu		
	385	390	395
	Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val		
	405	410	415
15	Glu Thr Leu Cys Lys Lys Tyr Asn Val Arg Tyr His Thr Thr Gly Met		
	420	425	430
	Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser Lys		
	435	440	445
20	Ala Ala Ser Lys Met Gly Lys Ala Gln		
	450	455	

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1488 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

40	GTCCCCTGTC GCTGTCGGCA CACCCCATCC TCCCTCGCTC CCTCTGCGTT TGTCTTGGC	60
	CCACCGTCTC TCCTCCACCC TCCGAGACGA CTGCAACTGT AATCAGGAAC CGACAAATAC	120
	ACGATTTCTT TTTACTCAGC ACCAACTCAA AATCCTCAAC CGCAACCCTT TTTCAGGATG	180
45	GCACCTCCCA ACACTATCGA TGCCGGTTTG ACCCAGCGTC ATATCAGCAC CTCGGCCCCA	240
	AACTCGGCCA AGCCTGCCTT CGAGCGCAAC TACCAGCTCC CCGAGTTCAC CATCAAGGAG	300
50	ATCCGAGAGT GCATCCCTGC CCACTGCTTT GAGCGCTCCG GTCTCCGTGG TCTCTGCCAC	360
	GTTGCCATCG ATCTGACTTG GGCCTCGCTC TTGTTCTCTG CTGCGACCCA GATCGACAAG	420
	TTTGAGAATC CCTTGATCCG CTATTTGGCC TGGCCTGTTT ACTGGATCAT GCAGGGTATT	480
55	GTCTGCACCG GTGTCTGGGT GCTGGCTCAC GAGTGTGGTC ATCAGTCCTT CTCGACCTCC	540
	AAGACCCTCA ACAACACAGT TGGTTGGATC TTGCACTCGA TGCTCTTGGT CCCCTACCAC	600
60	TCCTGGAGAA TCTCGCACTC GAAGCACCAC AAGGCCACTG GCCATATGAC CAAGGACCAG	660
	GTCTTTGTGC CCAAGACCCG CTCCCAGGTT GGCTTGCCTC CCAAGGAGAA CGCTGCTGCT	720
	GCCGTTTCAAG AGGAGGACAT GTCCGTGCAC CTGGATGAGG AGGCTCCCAT TGTGACTTTG	780
65	TTCTGGATGG TGATCCAGTT CTTGTTTCGA TGGCCCGCGT ACCTGATTAT GAACGCCTCT	840

GGCCAAGACT ACGGCCGCTG GACCTCGCAC TTCCACACGT ACTCGCCCAT CTTTGAGCCC 900  
 CGCAACTTTT TCGACATTAT TATCTCGGAC CTCGGTGTGT TGGCTGCCCT CGGTGCCCTG 960  
 5 ATCTATGCCT CCATGCAGTT GTCGCTCTTG ACCGTCACCA AGTACTATAT TGTCCCCTAC 1020  
 CTCTTTGTCA ACTTTTGGTT GGTCTGATC ACCTTCTTGC AGCACACCGA TCCAAGCTG 1080  
 10 CCCCATTACC GCGAGGGTGC CTGGAATTTT CAGCGTGGAG CTCTTTGCAC CGTTGACCGC 1140  
 TCGTTTGGCA AGTTCTTGGA CCATATGTTT CACGGCATTG TCCACACCCA TGTGGCCCAT 1200  
 CACTTGTCTT CGCAAATGCC GTTCTACCAT GCTGAGGAAG CTACCTATCA TCTCAAGAAA 1260  
 15 CTGCTGGGAG AGTACTATGT GTACGACCCA TCCCCGATCG TCGTTGCGGT CTGGAGGTCG 1320  
 TTCCGTGAGT GCCGATTCGT GGAGGATCAG GGAGACGTGG TCTTTTCAA GAAGTAAAAA 1380  
 AAAAGACAAT GGACCACACA CAACCTTGTC TCTACAGACC TACGTATCAT GTAGCCATAC 1440  
 20 CACTTCATAA AAGAACATGA GCTCTAGAGG CGTGTCAATC GCGCCTCC 1488

(2) INFORMATION FOR SEQ ID NO:4:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 399 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Pro Pro Asn Thr Ile Asp Ala Gly Leu Thr Gln Arg His Ile  
 1 5 10 15  
 40 Ser Thr Ser Ala Pro Asn Ser Ala Lys Pro Ala Phe Glu Arg Asn Tyr  
 20 25 30  
 Gln Leu Pro Glu Phe Thr Ile Lys Glu Ile Arg Glu Cys Ile Pro Ala  
 35 40 45  
 His Cys Phe Glu Arg Ser Gly Leu Arg Gly Leu Cys His Val Ala Ile  
 50 55 60  
 50 Asp Leu Thr Trp Ala Ser Leu Leu Phe Leu Ala Ala Thr Gln Ile Asp  
 65 70 75 80  
 Lys Phe Glu Asn Pro Leu Ile Arg Tyr Leu Ala Trp Pro Val Tyr Trp  
 85 90 95  
 55 Ile Met Gln Gly Ile Val Cys Thr Gly Val Trp Val Leu Ala His Glu  
 100 105 110  
 Cys Gly His Gln Ser Phe Ser Thr Ser Lys Thr Leu Asn Asn Thr Val  
 115 120 125  
 60 Gly Trp Ile Leu His Ser Met Leu Leu Val Pro Tyr His Ser Trp Arg  
 130 135 140  
 65 Ile Ser His Ser Lys His His Lys Ala Thr Gly His Met Thr Lys Asp  
 145 150 155 160

5           Gln Val Phe Val Pro Lys Thr Arg Ser Gln Val Gly Leu Pro Pro Lys  
                                   165                                   170                                   175  
 Glu Asn Ala Ala Ala Val Gln Glu Glu Asp Met Ser Val His Leu  
                                   180                                   185                                   190  
 10       Asp Glu Glu Ala Pro Ile Val Thr Leu Phe Trp Met Val Ile Gln Phe  
                                   195                                   200                                   205  
 Leu Phe Gly Trp Pro Ala Tyr Leu Ile Met Asn Ala Ser Gly Gln Asp  
                                   210                                   215                                   220  
 15       Tyr Gly Arg Trp Thr Ser His Phe His Thr Tyr Ser Pro Ile Phe Glu  
                                   225                                   230                                   235                                   240  
 Pro Arg Asn Phe Phe Asp Ile Ile Ile Ser Asp Leu Gly Val Leu Ala  
                                   245                                   250                                   255  
 20       Ala Leu Gly Ala Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu Leu Thr  
                                   260                                   265                                   270  
 Val Thr Lys Tyr Tyr Ile Val Pro Tyr Leu Phe Val Asn Phe Trp Leu  
                                   275                                   280                                   285  
 25       Val Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro His Tyr  
                                   290                                   295                                   300  
 30       Arg Glu Gly Ala Trp Asn Phe Gln Arg Gly Ala Leu Cys Thr Val Asp  
                                   305                                   310                                   315                                   320  
 Arg Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile Val His  
                                   325                                   330                                   335  
 35       Thr His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr His Ala  
                                   340                                   345                                   350  
 Glu Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr Tyr Val  
                                   355                                   360                                   365  
 40       Tyr Asp Pro Ser Pro Ile Val Val Ala Val Trp Arg Ser Phe Arg Glu  
                                   370                                   375                                   380  
 45       Cys Arg Phe Val Glu Asp Gln Gly Asp Val Val Phe Phe Lys Lys  
                                   385                                   390                                   395

## (2) INFORMATION FOR SEQ ID NO:5:

- 50       (i) SEQUENCE CHARACTERISTICS:  
           (A) LENGTH: 355 amino acids  
           (B) TYPE: amino acid  
           (C) STRANDEDNESS: not relevant  
           (D) TOPOLOGY: linear  
 55       (ii) MOLECULE TYPE: peptide  
 60       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:  
           Glu Val Arg Lys Leu Arg Thr Leu Phe Gln Ser Leu Gly Tyr Tyr Asp  
           1                                   5                                   10                                   15  
 65       Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val Ser Phe Asn Leu Cys Ile  
           20                                   25                                   30

Trp Gly Leu Ser Thr Val Ile Val Ala Lys Trp Gly Gln Thr Ser Thr  
                   35                                  40                                  45  
 5 Leu Ala Asn Val Leu Ser Ala Ala Leu Leu Gly Leu Phe Trp Gln Gln  
                   50                                  55                                  60  
 10 Cys Gly Trp Leu Ala His Asp Phe Leu His His Gln Val Phe Gln Asp  
       65                                  70                                  75                                  80  
 Arg Phe Trp Gly Asp Leu Phe Gly Ala Phe Leu Gly Gly Val Cys Gln  
                                   85                                  90                                  95  
 15 Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys His Asn Thr His His Ala  
                                   100                                  105                                  110  
 Ala Pro Asn Val His Gly Glu Asp Pro Asp Ile Asp Thr His Pro Leu  
                                   115                                  120                                  125  
 20 Leu Thr Trp Ser Glu His Ala Leu Glu Met Phe Ser Asp Val Pro Asp  
                   130                                  135                                  140  
 25 Glu Glu Leu Thr Arg Met Trp Ser Arg Phe Met Val Leu Asn Gln Thr  
       145                                  150                                  155                                  160  
 Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala Arg Leu Ser Trp Cys Leu  
                                   165                                  170                                  175  
 30 Gln Ser Ile Leu Phe Val Leu Pro Asn Gly Gln Ala His Lys Pro Ser  
                                   180                                  185                                  190  
 Gly Ala Arg Val Pro Ile Ser Leu Val Glu Gln Leu Ser Leu Ala Met  
                                   195                                  200                                  205  
 35 His Trp Thr Trp Tyr Leu Ala Thr Met Phe Leu Phe Ile Lys Asp Pro  
                   210                                  215                                  220  
 40 Val Asn Met Leu Val Tyr Phe Leu Val Ser Gln Ala Val Cys Gly Asn  
       225                                  230                                  235                                  240  
 Leu Leu Ala Ile Val Phe Ser Leu Asn His Asn Gly Met Pro Val Ile  
                                   245                                  250                                  255  
 45 Ser Lys Glu Glu Ala Val Asp Met Asp Phe Phe Thr Lys Gln Ile Ile  
                   260                                  265                                  270  
 Thr Gly Arg Asp Val His Pro Gly Leu Phe Ala Asn Trp Phe Thr Gly  
                   275                                  280                                  285  
 50 Gly Leu Asn Tyr Gln Ile Glu His His Leu Phe Pro Ser Met Pro Arg  
                   290                                  295                                  300  
 55 His Asn Phe Ser Lys Ile Gln Pro Ala Val Glu Thr Leu Cys Lys Lys  
       305                                  310                                  315                                  320  
 Tyr Asn Val Arg Tyr His Thr Thr Gly Met Ile Glu Gly Thr Ala Glu  
                                   325                                  330                                  335  
 60 Val Phe Ser Arg Leu Asn Glu Val Ser Lys Ala Ala Ser Lys Met Gly  
                                   340                                  345                                  350  
 Lys Ala Gln  
                   355

65 (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 104 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val Thr Leu Tyr Thr Leu Ala Phe Val Ala Ala Asn Ser Leu Gly Val  
 1 5 10 15  
 Leu Tyr Gly Val Leu Ala Cys Pro Ser Val Xaa Pro His Gln Ile Ala  
 20 25 30  
 Ala Gly Leu Leu Gly Leu Leu Trp Ile Gln Ser Ala Tyr Ile Gly Xaa  
 35 40 45  
 Asp Ser Gly His Tyr Val Ile Met Ser Asn Lys Ser Asn Asn Xaa Phe  
 50 55 60  
 Ala Gln Leu Leu Ser Gly Asn Cys Leu Thr Gly Ile Ile Ala Trp Trp  
 65 70 75 80  
 Lys Trp Thr His Asn Ala His His Leu Ala Cys Asn Ser Leu Asp Tyr  
 85 90 95  
 Gly Pro Asn Leu Gln His Ile Pro  
 100

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 252 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Val Leu Tyr Gly Val Leu Ala Cys Thr Ser Val Phe Ala His Gln  
 1 5 10 15  
 Ile Ala Ala Ala Leu Leu Gly Leu Leu Trp Ile Gln Ser Ala Tyr Ile  
 20 25 30  
 Gly His Asp Ser Gly His Tyr Val Ile Met Ser Asn Lys Ser Tyr Asn  
 35 40 45  
 Arg Phe Ala Gln Leu Leu Ser Gly Asn Cys Leu Thr Gly Ile Ser Ile  
 50 55 60  
 Ala Trp Trp Lys Trp Thr His Asn Ala His His Leu Ala Cys Asn Ser  
 65 70 75 80  
 Leu Asp Tyr Asp Pro Asp Leu Gln His Ile Pro Val Phe Ala Val Ser  
 85 90 95

Thr Lys Phe Phe Ser Ser Leu Thr Ser Arg Phe Tyr Asp Arg Lys Leu  
 100 105 110  
 5 Thr Phe Gly Pro Val Ala Arg Phe Leu Val Ser Tyr Gln His Phe Thr  
 115 120 125  
 Tyr Tyr Pro Val Asn Cys Phe Gly Arg Ile Asn Leu Phe Ile Gln Thr  
 130 135 140  
 10 Phe Leu Leu Leu Phe Ser Lys Arg Glu Val Pro Asp Arg Ala Leu Asn  
 145 150 155 160  
 Phe Ala Gly Ile Leu Val Phe Trp Thr Trp Phe Pro Leu Leu Val Ser  
 165 170 175  
 15 Cys Leu Pro Asn Trp Pro Glu Arg Phe Phe Phe Val Phe Thr Ser Phe  
 180 185 190  
 20 Thr Val Thr Ala Leu Gln His Ile Gln Phe Thr Leu Asn His Phe Ala  
 195 200 205  
 Ala Asp Val Tyr Val Gly Pro Pro Thr Gly Ser Asp Trp Phe Glu Lys  
 210 215 220  
 25 Gln Ala Ala Gly Thr Ile Asp Ile Ser Cys Arg Ser Tyr Met Asp Trp  
 225 230 235 240  
 30 Phe Phe Gly Gly Leu Gln Phe Gln Leu Glu His His  
 245 250

## (2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
 35 (A) LENGTH: 125 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: peptide

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Xaa Xaa Asn Phe Ala Gly Ile Leu Val Phe Trp Thr Trp Phe Pro  
 1 5 10 15  
 50 Leu Leu Val Ser Cys Leu Pro Asn Trp Pro Glu Arg Phe Xaa Phe Val  
 20 25 30  
 Phe Thr Gly Phe Thr Val Thr Ala Leu Gln His Ile Gln Phe Thr Leu  
 35 40 45  
 55 Asn His Phe Ala Ala Asp Val Tyr Val Gly Pro Pro Thr Gly Ser Asp  
 50 55 60  
 60 Trp Phe Glu Lys Gln Ala Ala Gly Thr Ile Asp Ile Ser Cys Arg Ser  
 65 70 75 80  
 Tyr Met Asp Trp Phe Phe Cys Gly Leu Gln Phe Gln Leu Glu His His  
 85 90 95  
 65 Leu Phe Pro Arg Leu Pro Arg Cys His Leu Arg Lys Val Ser Pro Val  
 100 105 110

Gly Gln Arg Gly Phe Gln Arg Lys Xaa Asn Leu Ser Xaa  
 115 120 125

5 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 131 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Pro Ala Thr Glu Val Gly Gly Leu Ala Trp Met Ile Thr Phe Tyr Val  
 1 5 10 15  
 Arg Phe Phe Leu Thr Tyr Val Pro Leu Leu Gly Leu Lys Ala Phe Leu  
 20 25 30  
 Gly Leu Phe Phe Ile Val Arg Phe Leu Glu Ser Asn Trp Phe Val Trp  
 35 40 45  
 Val Thr Gln Met Asn His Ile Pro Met His Ile Asp His Asp Arg Asn  
 50 55 60  
 Met Asp Trp Val Ser Thr Gln Leu Gln Ala Thr Cys Asn Val His Lys  
 65 70 75 80  
 Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu  
 85 90 95  
 His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr His Xaa Val Ala  
 100 105 110  
 Pro Leu Val Gln Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser  
 115 120 125  
 Lys Pro Leu  
 130

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 87 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Cys Ser Pro Lys Ser Ser Pro Thr Arg Asn Met Thr Pro Ser Pro Phe  
 1 5 10 15  
 Ile Asp Trp Leu Trp Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu  
 20 25 30

Phe Pro Thr Met Pro Arg Cys Asn Leu Asn Arg Cys Met Lys Tyr Val  
                   35                                  40                                  45  
 5 Lys Glu Trp Cys Ala Glu Asn Asn Leu Pro Tyr Leu Val Asp Asp Tyr  
           50                                  55                                  60  
 Phe Val Gly Tyr Asn Leu Asn Leu Gln Gln Leu Lys Asn Met Ala Glu  
 65                                  70                                  75                                  80  
 10 Leu Val Gln Ala Lys Ala Ala  
                                   85

## (2) INFORMATION FOR SEQ ID NO:11:

15 (i) SEQUENCE CHARACTERISTICS:  
           (A) LENGTH: 143 amino acids  
           (B) TYPE: amino acid  
           (C) STRANDEDNESS: not relevant  
 20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

30 Arg His Glu Ala Ala Arg Gly Gly Thr Arg Leu Ala Tyr Met Leu Val  
   1                  5                                  10                                  15  
 Cys Met Gln Trp Thr Asp Leu Leu Trp Ala Ala Ser Phe Tyr Ser Arg  
                   20                                  25                                  30  
 35 Phe Phe Leu Ser Tyr Ser Pro Phe Tyr Gly Ala Thr Gly Thr Leu Leu  
           35                                  40                                  45  
 Leu Phe Val Ala Val Arg Val Leu Glu Ser His Trp Phe Val Trp Ile  
 40          50                                  55                                  60  
 Thr Gln Met Asn His Ile Pro Lys Glu Ile Gly His Glu Lys His Arg  
 65                                  70                                  75                                  80  
 45 Asp Trp Ala Ser Ser Gln Leu Ala Ala Thr Cys Asn Val Glu Pro Ser  
                   85                                  90                                  95  
 Leu Phe Ile Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu His  
                   100                                  105                                  110  
 50 His Leu Phe Pro Thr Met Thr Arg His Asn Tyr Arg Xaa Val Ala Pro  
           115                                  120                                  125  
 Leu Val Lys Ala Phe Cys Ala Lys His Gly Leu His Tyr Glu Val  
 55          130                                  135                                  140

## (2) INFORMATION FOR SEQ ID NO:12:

60 (i) SEQUENCE CHARACTERISTICS:  
           (A) LENGTH: 35 base pairs  
           (B) TYPE: nucleic acid  
           (C) STRANDEDNESS: single  
           (D) TOPOLOGY: linear  
 65 (ii) MOLECULE TYPE: other nucleic acid



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

5 CCAAGCTTCT GCAGGAGCTC TTTT TTTT 35

(2) INFORMATION FOR SEQ ID NO:13:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: other nucleic acid

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CUACUACUAC UAGGAGTCCT CTACGGTGTT TTG 33

(2) INFORMATION FOR SEQ ID NO:14:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

40 CAUCAUAC AUATGATGCT CAAGCTGAAA CTG 33

(2) INFORMATION FOR SEQ ID NO:15:

45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TACCAACTCG AGAAAATGGC TGCTGCTCCC AGTGTGAGG 39

(2) INFORMATION FOR SEQ ID NO:16:

60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
65 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:  
 AACTGATCTA GATTACTGCG CCTTACCCAT CTTGGAGGC 39

10 (2) INFORMATION FOR SEQ ID NO:17:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 15 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: other nucleic acid

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:  
 TACCAACTCG AGAAAATGGC ACCTCCCAAC ACTATCGAT 39

25 (2) INFORMATION FOR SEQ ID NO:18:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 base pairs  
 30 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: other nucleic acid

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:  
 40 AACTGATCTA GATTACTTCT TGAAAAAGAC CACGTCTCC 39

(2) INFORMATION FOR SEQ ID NO:19:  
 (i) SEQUENCE CHARACTERISTICS:  
 45 (A) LENGTH: 746 nucleic acids  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: nucleic acid

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:  
 55 CGTATGTCAC TCCATTCCAA ACTCGTTCAT GGTATCATAA ATATCAACAC ATTTACGCTC 60  
 CACTCCTCTA TGGTATTTAC AACTCAAAT ATCGTACTCA AGATTGGGAA GCTTTTGTA 120  
 AGGATGGTAA AAATGGTGCA ATTCGTGTTA GTGTCGCCAC AAATTTCGAT AAGGCCGCTT 180  
 ACGTCATTGG TAAATTGTCT TTTGTTTTCT TCCGTTTCAT CCTTCCACTC CGTTATCATA 240  
 GCTTTACAGA TTTAATTTGT TATTTCTCTA TTGCTGAATT CGTCTTGGT TGGTATCTCA 300  
 CAATTAATTT CCAAGTTAGT CATGTCGCTG AAGATCTCAA ATTCTTTGCT ACCCTGAAA 360  
 60 GACCAGATGA ACCATCTCAA ATCAATGAAG ATTGGGCAAT CCTTCAACTT AAAACTACTC 420  
 AAGATTATGG TCATGGTTCA CTCCTTTGTA CTTTTTTAG TGGTCTTTA AATCATCAAG 480  
 TTGTTTCATCA TTTATTCCCA TCAATTGCTC AAGATTCTA CCCACAATT GTACCAATTG 540  
 TAAAGAAGT TTGTAAAGAA CATAACATTA CTTACCACAT TAAACCAAAC TTCACTGAAG 600  
 CTATTATGTC ACACATTAAT TACCTTTACA AAATGGGTAA TGATCCAGAT TATGTTAAAA 660  
 65 AACCATTAGC CTCAAAAGAT GATTAAATGA AATAACTTAA AAACCAATTA TTTACTTTTG 720

ACAAACAGTA ATATTAATAA ATACAA

746

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 227 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Tyr Val Thr Pro Phe Gln Thr Arg Ser Trp Tyr His Lys Tyr Gln
1          5          10          15
His Ile Tyr Ala Pro Leu Leu Tyr Gly Ile Tyr Thr Leu Lys Tyr
20        20        25        30
Arg Thr Gln Asp Trp Glu Ala Phe Val Lys Asp Gly Lys Asn Gly
35        40        45
Ala Ile Arg Val Ser Val Ala Thr Asn Phe Asp Lys Ala Ala Tyr
50        55        60
Val Ile Gly Lys Leu Ser Phe Val Phe Phe Arg Phe Ile Leu Pro
25        65        70        75
Leu Arg Tyr His Ser Phe Thr Asp Leu Ile Cys Tyr Phe Leu Ile
80        85        90
Ala Glu Phe Val Phe Gly Trp Tyr Leu Thr Ile Asn Phe Gln Val
95        100       105
Ser His Val Ala Glu Asp Leu Lys Phe Phe Ala Thr Pro Glu Arg
30        110       115       120
Pro Asp Glu Pro Ser Gln Ile Asn Glu Asp Trp Ala Ile Leu Gln
125       130       135
Leu Lys Thr Thr Gln Asp Tyr Gly His Gly Ser Leu Leu Cys Thr
35        140       145       150
Phe Phe Ser Gly Ser Leu Asn His Gln Val Val His His Leu Phe
155       160       165
Pro Ser Ile Ala Gln Asp Phe Tyr Pro Gln Leu Val Pro Ile Val
170       175       180
Lys Glu Val Cys Lys Glu His Asn Ile Thr Tyr His Ile Lys Pro
40        185       190       195
Asn Phe Thr Glu Ala Ile Met Ser His Ile Asn Tyr Leu Tyr Lys
200       205       210
Met Gly Asn Asp Pro Asp Tyr Val Lys Lys Pro Leu Ala Ser Lys
45        215       220       225
Asp Asp ***

```

## (2) INFORMATION FOR SEQ ID NO 21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 494 nucleic acids  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: nucleic acid

## (xi) SEQUENCE DESCRIPTION:SEQ ID NO:21:

```

TTTTGGAAGG NTCCAAGTTN ACCACGGANT NGGCAAGTTN ACGGGGCGGA AANCGGTTTT 60
CCCCCAAGC CTTTGTGCGA CTGGTCTGT GGTGGCTTCC AGTACCAAGT CGACCACCAC 120
TTATTCCCA GCCTGCCCG ACACAATCTG GCCAAGACAC ACGCACTGGT CGAATCGTTC 180
TGCAAGGAGT GGGGTGTCCA GTACCACGAA GCCGACCTCG TGGACGGGAC CATGGAAGTC 240
TTGACCATT TGGGCAGCGT GGCCGCGCAA TTCGTCGTGG ATTTGTACG CGACGGACCC 300

```

GCCATGTAAT CGTCGTTCGT GACGATGCAA GGGTTCACGC ACATCTACAC ACACTCACTC 360  
 ACACAACCTAG TGTAACCTCGT ATAGAATTCG GTGTCGACCT GGACCTTGTT TGA CTGGTTG 420  
 GGGATAGGGT AGGTAGGCGG ACGCGTGGGT CGNCCCCGGG AATTCTGTGA CCGGTACCTG 480  
 GCGCGCTNA AAGT 494

5

## (2) INFORMATION FOR SEQ ID NO:22:

10

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 87 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

15

- (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

20

Phe Trp Lys Xxx Pro Ser Xxx Pro Arg Xxx Xxx Gln Val Xxx Gly  
 1 5 10 15  
 Ala Glu Xxx Gly Phe Pro Pro Lys Pro Phe Val Asp Trp Phe Cys  
 20 25 30  
 Gly Gly Phe Gln Tyr Gln Val Asp His His Leu Phe Pro Ser Leu  
 25 35 40 45  
 Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val Glu Ser Phe  
 50 55 60  
 Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu Val Asp  
 65 70 75  
 Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly Glu  
 30 65 70 75  
 Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met  
 80 85

35

40

## (2) INFORMATION FOR SEQ ID NO:23:

45

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 520 nucleic acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: nucleic acid

50

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

55

GGATGGAGTT CGTCTGGATC GCTGTGCGCT ACGCGACGTG GTTTAAGCGT CATGGGTGCG 60  
 CTTGGGTACA CGCCGGGGCA GTCGTTGGGC ATGTACTTGT GCGCCTTTGG TCTCGGCTGC 120  
 ATTTACATTT TTCTGCAGTT CGCCGTAAGT CACACCCATT TGCCCGTGAG CAACCCGGAG 180  
 GATCAGCTGC ATTGGCTCGA GTACGCGCGG ACCACACTGT GAACATCAGC ACCAAGTCGT 240  
 GGT TTGTAC ATGGTGGATG TCGAACCTCA ACTTTTCAGAT CGAGCACCAC CTTTTCCTCA 300  
 CGGCGCCCCA GTTCCGTTTC AAGGAGATCA GCGCGCGCGT CGAGGCCCTC TTCAAGCGCC 360  
 ACGGTCTCCC TTA CTACGAC ATGCCCTACA CGAGCGCCGT CTCCACCACC TTTGCCAACC 420  
 TCTACTCCGT CGGCCATTCC GTCGGCGACG CCAAGCGCGA CTAGCCTCTT TTCCTAGACC 480  
 TTAATTCCCC ACCCCACCCC ATGTTCTGTC TTCCTCCCGC 520

60

## (2) INFORMATION FOR SEQ ID NO:24:

65

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

10

Met	Glu	Phe	Val	Trp	Ile	Ala	Val	Arg	Tyr	Ala	Thr	Trp	Phe	Lys
1				5					10					15
Arg	His	Gly	Cys	Ala	Trp	Val	His	Ala	Gly	Ala	Val	Val	Gly	His
			20						25					30
Val	Leu	Val	Arg	Leu	Trp	Ser	Arg	Leu	His	Leu	His	Phe	Ser	Ala
			35						40					45
Val	Arg	Arg	Lys	Ser	His	Pro	Phe	Ala	Arg	Glu	Gln	Pro	Gly	Gly
			50						55					60
Ser	Ala	Ala	Leu	Arg	Val	Arg	Ala	Asp	His	Thr	Val	Asn	Ile	
			65						70					75
Ser	Thr	Lys	Ser	Trp	Phe	Val	Thr	Trp	Trp	Met	Ser	Asn	Leu	Asn
			80						85					90
Phe	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Thr	Ala	Pro	Gln	Phe	Arg
			95						100					105
Phe	Lys	Glu	Ile	Ser	Pro	Arg	Val	Glu	Ala	Leu	Phe	Lys	Arg	His
			110						115					120
Gly	Leu	Pro	Tyr	Tyr	Asp	Met	Pro	Tyr	Thr	Ser	Ala	Val	Ser	Thr
			125						130					135
Thr	Phe	Ala	Asn	Leu	Tyr	Ser	Val	Gly	His	Ser	Val	Gly	Asp	Ala
			140						145					150
Lys	Arg	Asp												

35

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 420 nucleic acids  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: nucleic acid

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ACGCGTCCGC	CCACGCGTCC	GCCGCGAGCA	ACTCATCAAG	GAAGGCTACT	TTGACCCCTC	60
GCTCCCGCAC	ATGACGTACC	GCGTGGTCGA	GATTGTTGTT	CTCTTCGTGC	TTTCCTTTTG	120
GCTGATGGGT	CAGTCTTCAC	CCCTCGCGCT	CGCTCTCGGC	ATTGTCGTCA	GCGGCATCTC	180
TCAGGGTCGC	TGCGGCTGGG	TAATGCATGA	GATGGGCCAT	GGGTCGTTCA	CTGGTGTGTCAT	240
TTGGCTTGAC	GACCGGTTGT	GCGAGTTCTT	TTACGGCGTT	GGTTGTGGCA	TGAGCGGTCA	300
TTACTGGAAA	AACCAGCACA	GCAAACACCA	CGCAGCGCCA	AACCGGCTCG	AGCACGATGT	360
AGATCTCAAC	ACCTTGCCAT	TGGTGGCCTT	CAACGAGCGC	GTCGTGCGCA	AGGTCCGACC	420

55

(2) INFORMATION FOR SEQ ID NO:26:

60

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 125 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: linear

65

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

5 Arg Val Arg Pro Arg Val Arg Arg Glu Gln Leu Ile Lys Glu Gly  
 1 5 10 15  
 Tyr Phe Asp Pro Ser Leu Pro His Met Thr Tyr Arg Val Val Glu  
 20 25 30  
 10 Ile Val Val Leu Phe Val Leu Ser Phe Trp Leu Met Gly Gln Ser  
 35 40 45  
 Ser Pro Leu Ala Leu Ala Leu Gly Ile Val Val Ser Gly Ile Ser  
 50 55 60  
 Gln Gly Arg Cys Gly Trp Val Met His Glu Met Gly His Gly Ser  
 65 70 75  
 15 Phe Thr Gly Val Ile Trp Leu Asp Asp Arg Leu Cys Glu Phe Phe  
 65 70 75  
 Tyr Gly Val Gly Cys Gly Met Ser Gly His Tyr Trp Lys Asn Gln  
 80 85 90  
 20 His Ser Lys His His Ala Ala Pro Asn Arg Leu Glu His Asp Val  
 95 100 105  
 Asp Leu Asn Thr Leu Pro Leu Val Ala Phe Asn Glu Arg Val Val  
 110 115 120  
 Arg Lys Val Arg Pro  
 125

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1219 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid (Edited Contig 2692004)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

40 GCACGCCGAC CGGCGCCGGG AGATCCTGGC AAAGTATCCA GAGATAAAGT CCTTGATGAA 60  
 ACCTGATCCC AATTTGATAT GGATTATAAT TATGATGGTT CTCACCCAGT TGGGTGCATT 120  
 TTACATAGTA AAAGACTTGG ACTGGAAATG GGTCAATTTT GGGGCCTATG CGTTTGGCAG 180  
 45 TTGCATTAAC CACTCAATGA CTCTGGCTAT TCATGAGATT GCCCACAATG CTGCCTTTGG 240  
 CAACTGCAAA GCAATGTGGA ATCGCTGGTT TGGAAATGTTT GCTAATCTTC CTATTGGGAT 300  
 50 TCCATATTCA ATTTCTTTA AGAGGTATCA CATGGATCAT CATCGGTACC TTGGAGCTGA 360  
 TGGCGTCGAT GTAGATATTC CTACCGATTT TGAGGGCTGG TTCTTCTGTA CCGCTTTCAG 420  
 AAAGTTTATA TGGGTTATTC TTCAGCCTCT CTTTTATGCC TTTCGACCTC TGTTTCATCA 480  
 55 CCCCAAACCA ATTACGTATC TGGAAGTTAT CAATACCGTG GCACAGGTCA CTTTTGACAT 540  
 TTTAATTTAT TACTTTTTGG GAATTAAATC CTTAGTCTAC ATGTTGGCAG CATCTTTACT 600  
 60 TGGCCTGGGT TTGCACCCAA TTTCTGGACA TTTTATAGCT GAGCATTACA TGTTCCTAAA 660  
 GGGTCATGAA ACTTACTCAT ATTATGGGCC TCTGAATTTA CTTACCTTCA ATGTGGGTGA 720  
 TCATAATGAA CATCATGATT TCCCCAACAT TCCTGGAAAA AGTCTTCCAC TGGTGAGGAA 780  
 65 AATAGCAGCT GAATACTATG ACAACCTCCC TCACTACAAT TCCTGGATAA AAGTACTGTA 840

5 TGATTTTGTG ATGGATGATA CAATAAGTCC CTA CTACTCAAGA ATGAAGAGGC ACCAAAAAGG 900  
 AGAGATGGTG CTGGAGTAAA TATCATTAGT GCCAAAGGGA TTCTTCTCCA AAAC TT TAGA 960  
 TGATAAAATG GAATTTTTC ATTATTAAAC TTGAGACCAG TGATGCTCAG AAGCTCCCT 1020  
 GGCACAATTT CAGAGTAAGA GCTCGGTGAT ACCAAGAAGT GAATCTGGCT TTTAAACAGT 1080  
 10 CAGCCTGACT CTGTACTGCT CAGTTTCACT CACAGGAAAC TTGTGACTTG TGTATTATCG 1140  
 TCATTGAGGA TGTTCCTCTC ATGTCTGTCA TTTTATAAGC ATATCATTTA AAAAGCTTCT 1200  
 AAAAAGCTAT TTCGCCAGG 1219

(2) INFORMATION FOR SEQ ID NO:28:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 655 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 25 (ii) MOLECULE TYPE: other nucleic acid (Edited Contig 2153526)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

30 TTACCTTCTA CGTCCGCTTC TTCCTCACTT ATGTGCCACT ATTGGGGCTG AAAGCTTCCT 60  
 GGGCCTTTTC TTCATAGTCA GGTTCCTGGA AAGCAACTGG TTTGTGTGGG TGACACAGAT 120  
 35 GAACCATATT CCCATGCACA TTGATCATGA CCGGAACATG GACTGGGTTT CCACCCAGCT 180  
 CCAGGCCACA TGCAATGTCC ACAAGTCTGC CTTCAATGAC TGGTTCAGTG GACACCTCAA 240  
 40 CTTCCAGATT GAGCACCATC TTTTCCCAC GATGCCTCGA CACAATTACC ACAAAGTGGC 300  
 TCCCCTGGTG CAGTCCTTGT GTGCCAAGCA TGGCATAGAG TACCAGTCCA AGCCCCTGCT 360  
 GTCAGCCTTC GCCGACATCA TCCACTCACT AAAGGAGTCA GGGCAGCTCT GGCTAGATGC 420  
 45 CTATCTTCAC CAATAACAAC AGCCACCCTG CCCAGTCTGG AAGAAGAGGA GGAAGACTCT 480  
 GGAGCCAAGG CAGAGGGGAG CTTGAGGGAC AATGCCACTA TAGTTTAATA CTCAGAGGGG 540  
 50 GTTGGGTTTG GGGACATAAA GCCTCTGACT CAACTCCTC CCTTTTATCT TCTAGCCACA 600  
 GTTCTAAGAC CCAAAGTGGG GGGTGGACAC AGAAGTCCCT AGGAGGGAAG GAGCT 655

(2) INFORMATION FOR SEQ ID NO:29:

55 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 304 base pairs  
 (B) TYPE: nucleic acid  
 60 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: other nucleic acid (Edited Contig 3506132)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

65 GTCTTTTACT TTGGCAATGG CTGGATTCCT ACCCTCATCA CGGCCTTTGT CCTTGCTACC 60

5 TCTCAGGCCC AAGCTGGATG GCTGCAACAT GATTATGGCC ACCTGTCTGT CTACAGAAAA 120  
 CCCAAGTGGA ACCACCTTGT CCACAAATTC GTCATTGGCC ACTTAAAGGG TGCCTCTGCC 180  
 AACTGGTGGG ATCATCGCCA CTTCCAGCAC CACGCCAAGC CTAACATCTT CCACAAGGAT 240  
 CCCGATGTGA ACATGCTGCA CGTGTTTGTG CTGGGCGAAT GGCAGCCCAT CGAGTACGGC 300  
 10 AAGA 304

## (2) INFORMATION FOR SEQ ID NO:30:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 918 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 20 (ii) MOLECULE TYPE: other nucleic acid (Edited Contig 3854933)  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:  
 25 CAGGGACCTA CCCCGCGCTA CTTACCTGG GACGAGGTGG CCCAGCGCTC AGGGTGCAGAG 60  
 GAGCGGTGGC TAGTGATCGA CCGTAAGGTG TACAACATCA GCGAGTTCAC CCGCCGGCAT 120  
 30 CCAGGGGGCT CCCGGGTCAT CAGCCACTAC GCCGGGCAGG ATGCCACGGA TCCCTTTGTG 180  
 GCCTTCCACA TCAACAAGGG CCTTGTGAAG AAGTATATGA ACTCTCTCCT GATTGGAGAA 240  
 CTGTCTCCAG AGCAGCCAG CTTTGTAGCC ACCAAGAATA AAGAGCTGAC AGATGAGTTC 300  
 35 CGGGAGCTGC GGGCCACAGT GGAGCGGATG GGGCTCATGA AGGCCAACCA TGTCTTCTTC 360  
 CTGCTGTACC TGCTGCACAT CTTGCTGCTG GATGGTGCAG CCTGGCTCAC CCTTTGGGTC 420  
 40 TTTGGGACGT CCTTTTGGC CTTCTCCTC TGTGCGGTGC TGCTCAGTGC AGTTCAGGCC 480  
 CAGGCTGGCT GGCTGCAGCA TGACTTTGGG CACCTGTCGG TCTTCAGCAC CTCAAAGTGG 540  
 AACCATCTGC TACATCATTT TGTGATTGGC CACCTGAAGG GGGCCCCCGC CAGTTGGTGG 600  
 45 AACCACATGC ACTTCCAGCA CCATGCCAAG CCCAACTGCT TCCGCAAAGA CCCAGACATC 660  
 AACATGCATC CCTTCTTCTT TGCCTTGGGG AAGATCCTCT CTGTGGAGCT TGGGAAACAG 720  
 50 AAGAAAAAAT ATATGCCGTA CAACCACCAG CACARATACT TCTTCCTAAT TGGGCCCCCA 780  
 GCCTTGCTGC CTCTCTACTT CCAGTGGTAT ATTTTCTATT TTGTTATCCA GCGAAAGAAG 840  
 TGGGTGGACT TGGCCTGGAT CAGCAAACAG GAATACGATG AAGCCGGGCT TCCATTGTCC 900  
 55 ACCGCAAATG CTTCTAAA 918

## (2) INFORMATION FOR SEQ ID NO:31:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1686 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 65 (ii) MOLECULE TYPE: other nucleic acid (Edited Contig 2511785)



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

5	GCCACTTAAA GGGTGCCTCT GCCAACTGGT GGAATCATCG CCACTTCCAG CACCACGCCA	60
	AGCCTAACAT CTTCCACAAG GATCCCGATG TGAACATGCT GCACGTGTTT GTTCTGGGCG	120
	AATGGCAGCC CATCGAGTAC GGCAAGAAGA AGCTGAAATA CCTGCCCTAC AATCACCAGC	180
10	ACGAATACTT CTTCTGATT GGGCCGCCGC TGCTCATCCC CATGTATTTC CAGTACCAGA	240
	TCATCATGAC CATGATCGTC CATAAGAACT GGGTGGACCT GGCCTGGGCC GTCAGCTACT	300
15	ACATCCGGTT CTTTCATACC TACATCCCTT TCTACGGCAT CCTGGGAGCC CTCCTTTTCC	360
	TCAACTTCAT CAGGTTCCCTG GAGAGCCACT GGTTCGTGTG GGTACACAG ATGAATCACA	420
	TCGTCATGGA GATTGACCAG GAGGCCTACC GTGACTGGTT CAGTAGCCAG CTGACAGCCA	480
20	CCTGCAACGT GGAGCAGTCC TTCTTCAACG ACTGGTTCAG TGGACACCTT AACTTCCAGA	540
	TTGAGCACCA CCTCTTCCCC ACCATGCCCC GGCACAACTT ACACAAGATC GCCCCGCTGG	600
25	TGAAGTCTCT ATGTGCCAAG CATGGCATTG AATACCAGGA GAAGCCGCTA CTGAGGGCCC	660
	TGCTGGACAT CATCAGGTCC CTGAAGAAGT CTGGGAAGCT GTGGCTGGAC GCCTACCTTC	720
	ACAAATGAAG CCACAGCCCC CGGGACACCG TGGGAAGGG GTGCAGGTGG GGTGATGGCC	780
30	AGAGGAATGA TGGGCTTTTG TTCTGAGGGG TGTCCGAGAG GCTGGTGTAT GCACTGCTCA	840
	CGGACCCCAT GTTGGATCTT TCTCCCTTTC TCCTCTCCTT TTTCTCTTCA CATCTCCCCC	900
35	ATAGCACCTT GCCCTCATGG GAÇCTGCCCT CCCTCAGCCG TCAGCCATCA GCCATGGCCC	960
	TCCCAGTGCC TCCTAGCCCC TTCTTCCAAG GAGCAGAGAG GTGGCCACCG GGGGTGGCTC	1020
	TGTCCTACCT CCACTCTCTG CCCCTAAAGA TGGGAGGAGA CCAGCGGTCC ATGGGTCTGG	1080
40	CCTGTGAGTC TCCCCTTGCA GCCTGGTCAC TAGGCATCAC CCCCCTTTG GTTCTTCAGA	1140
	TGCTCTTGGG GTTCATAGGG GCAGGTCCTA GTCGGGCAGG GCCCCTGACC CTCCCGCCT	1200
45	GGCTTCACTC TCCCTGACGG CTGCCATTGG TCCACCCTTT CATAGAGAGG CCTGCTTTGT	1260
	TACAAAGCTC GGGTCTCCCT CCTGCAGCTC GGTTAAGTAC CCGAGGCCTC TCTTAAGATG	1320
	TCCAGGGCCC CAGGCCCGCG GGCACAGCCA GCCCAAACCT TGGGCCCTGG AAGAGTCCTC	1380
50	CACCCCATCA CTAGAGTGCT CTGACCCTGG GCTTTCACGG GCCCCATTCC ACCGCCTCCC	1440
	CAACTTGAGC CTGTGACCTT GGGACCAAAG GGGGAGTCCC TCGTCTCTTG TGA CTGAGCA	1500
55	GAGGCAGTGG CCACGTTTCTG GGAGGGGCCG GCTGGCCTGG AGGCTCAGCC CACCCTCCAG	1560
	CTTTTCTCTCA GGGTGTCTCTG AGGTCCAAGA TTCTGGAGCA ATCTGACCCT TCTCCAAAGG	1620
	CTCTGTTATC AGCTGGGCAG TGCCAGCCAA TCCCTGGCCA TTTGGCCCCA GGGGACGTGG	1680
60	GCCCTG	1686

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1843 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: other nucleic acid (Contig 2535)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

10

GTCTTTTACT TTGGCAATGG CTGGATTCCT ACCCTCATCA CGGCCTTTGT CCTTGCTACC 60

TCTCAGGCCC AAGCTGGATG GCTGCAACAT GATTATGGCC ACCTGTCTGT CTACAGAAAA 120

15

CCCAAGTGGA ACCACCTTGT CCACAAATTC GTCATTGGCC ACTTAAAGGG TGCCTCTGCC 180

AACTGGTGGA ATCATCGCCA CTTCCAGCAC CACGCCAAGC CTAACATCTT CCACAAGGAT 240

20

CCCGATGTGA ACATGCTGCA CGTGTGTTT CTGGGCGAAT GGCAGCCCAT CGAGTACGGC 300

AAGAAGAAGC TGAAATACCT GCCCTACAAT CACCAGCACG AATACTTCTT CCTGATTGGG 360

CCGCCGCTGC TCATCCCCAT GTATTTCCAG TACCAGATCA TCATGACCAT GATCGTCCAT 420

25

AAGAACTGGG TGGACCTGGC CTGGGCCGTC AGCTACTACA TCCGGTTCTT CATCACCTAC 480

ATCCCTTTCT ACGGCATCCT GGGAGCCCTC CTTTTCCTCA ACTTCATCAG GTTCCTGGAG 540

30

AGCCACTGGT TTGTGTGGGT CACACAGATG AATCACATCG TCATGGAGAT TGACCAGGAG 600

GCCTACCGTG ACTGGTTCAG TAGCCAGCTG ACAGCCACCT GCAACGTGGA GCAGTCCTTC 660

TTCAACGACT GGTTCAGTGG ACACCTTAAC TTCCAGATTG AGCACCACCT CTTCCCCACC 720

35

ATGCCCCGGC ACAACTTACA CAAGATCGCC CCGCTGGTGA AGTCTCTATG TGCCAAGCAT 780

GGCATTGAAT ACCAGGAGAA GCCGCTACTG AGGGCCCTGC TGGACATCAT CAGGTCCCTG 840

40

AAGAAGTCTG GGAAGCTGTG GCTGGACGCC TACCTTCACA AATGAAGCCA CAGCCCCCGG 900

GACACCGTGG GGAAGGGGTG CAGGTGGGGT GATGGCCAGA GGAATGATGG GCTTTTGTTC 960

TGAGGGGTGT CCGAGAGGCT GGTGTATGCA CTGCTCACGG ACCCATGTT GGATCTTTCT 1020

45

CCCTTTCTCC TCTCCTTTT CTCTTCACAT CTCCCCATA GCACCCTGCC CTCATGGGAC 1080

CTGCCCTCCC TCAGCCGTCA GCCATCAGCC ATGGCCCTCC CAGTGCCTCC TAGCCCCTTC 1140

50

TTCCAAGGAG CAGAGAGGTG GCCACCGGGG GTGGCTCTGT CCTACCTCCA CTCTCTGCCC 1200

CTAAAGATGG GAGGAGACCA GCGGTCCATG GGTCTGGCCT GTGAGTCTCC CTTGCAGCC 1260

TGGTCACTAG GCATCACCCC CGCTTTGGTT CTTCAGATGC TCTTGGGGTT CATAGGGGCA 1320

55

GGTCCTAGTC GGGCAGGGCC CCTGACCCTC CCGGCCTGGC TTCACTCTCC CTGACGGCTG 1380

CCATTGGTCC ACCCTTTCAT AGAGAGGCCT GCTTTGTAC AAAGCTCGGG TCTCCCTCCT 1440

60

GCAGCTCGGT TAAGTACCCG AGGCCTCTCT TAAGATGTCC AGGGCCCCAG GCCCGCGGGC 1500

ACAGCCAGCC CAAACCTTGG GCCCTGGAAG AGTCCTCCAC CCCATCACTA GAGTGCTCTG 1560

ACCCTGGGCT TTCACGGGCC CCATTCCACC GCCTCCCCAA CTTGAGCCTG TGACCTTGGG 1620

65

ACCAAAGGGG GAGTCCCTCG TCTCTTGTGA CTCAGCAGAG GCAGTGGCCA CGTTCAGGGA 1680

GGGGCCGGCT GGCCTGGAGG CTCAGCCAC CCTCCAGCTT TTCCTCAGGG TGCCTGAGG 1740  
 TCCAAGATTC TGGAGCAATC TGACCCTTCT CCAAAGGCTC TGTTATCAGC TGGGCAGTGC 1800  
 5 CAGCCAATCC CTGGCCATTT GGCCCCAGGG GACGTGGGCC CTG 1843

(2) INFORMATION FOR SEQ ID NO:33:

10

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2257 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: other nucleic acid (Edited Contig 253538a)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

20

CAGGGACCTA CCCCGCGCTA CTTACCTGG GACGAGGTGG CCCAGCGCTC AGGGTGCAG 60

GAGCGGTGGC TAGTGATCGA CCGTAAGGTG TACAACATCA GCGAGTTCAC CCGCCGGCAT 120

25

CCAGGGGGCT CCCGGGTCAT CAGCCACTAC GCCGGGCAGG ATGCCACGGA TCCCTTTGTG 180

GCCTTCACA TCAACAAGGG CCTTGTGAAG AAGTATATGA ACTCTCTCCT GATTGGAGAA 240

CTGTCTCCAG AGCAGCCCAG CTTTGAGCCC ACCAAGAATA AAGAGCTGAC AGATGAGTTC 300

30

CGGGAGCTGC GGGCCACAGT GGAGCGGATG GGGCTCATGA AGGCCAACCA TGTCTTCTTC 360

CTGCTGTACC TGCTGCACAT CTTGCTGCTG GATGGTGAG CCTGGCTCAC CCTTTGGGTC 420

35

TTTGGGACGT CCTTTTGGC CTTCTCCTC TGTGCGGTGC TGCTCAGTGC AGTTCAGCAG 480

GCCCAAGCTG GATGGCTGCA ACATGATTAT GGCCACCTGT CTGTCTACAG AAAACCCAAG 540

TGGAACCACC TTGTCCACAA ATTCGTCATT GGCCACTTAA AGGGTGCCTC TGCCAAGTGG 600

40

TGGAATCATC GCCACTTCCA GCACCAGCC AAGCCTAACA TCTTCACAA GGATCCCGAT 660

GTGAACATGC TGCACGTGTT TGTCTGGGC GAATGGCAGC CCATCGAGTA CGGCAAGAAG 720

45

AAGCTGAAAT ACCTGCCCTA CAATCACCAG CACGAATACT TCTTCCTGAT TGGGCCGCCG 780

CTGCTCATCC CCATGTATTT CCAGTACCAG ATCATCATGA CCATGATCGT CCATAAGAAC 840

TGGGTGGACC TGGCCTGGGC CGTCAGCTAC TACATCCGGT TCTTCATCAC CTACATCCCT 900

50

TTCTACGGCA TCCTGGGAGC CCTCCTTTTC CTCAACTTCA TCAGGTTCTT GGAGAGCCAC 960

TGGTTTGTGT GGGTCACACA GATGAATCAC ATCGTCATGG AGATTGACCA GGAGGCCTAC 1020

55

CGTGACTGGT TCAGTAGCCA GCTGACAGCC ACCTGCAACG TGGAGCAGTC CTTCTTCAAC 1080

GACTGGTTCA GTGGACACCT TAACTTCCAG ATTGAGCACC ACCTCTTCCC CACCATGCCC 1140

CGGCACAAC TACACAAGAT CGCCCCGCTG GTGAAGTCTC TATGTGCCAA GCATGGCATT 1200

60

GAATACCAGG AGAAGCCGCT ACTGAGGGCC CTGCTGGACA TCATCAGGTC CCTGAAGAAG 1260

TCTGGGAAGC TGTGGCTGGA CGCCTACCTT CACAAATGAA GCCACAGCCC CCGGGACACC 1320

65

GTGGGGAAGG GGTGCAGGTG GGGTGATGGC CAGAGGAATG ATGGGCTTTT GTTCTGAGGG 1380

GTGTCCGAGA GGCTGGTGTA TGCACTGCTC ACGGACCCCA TGTGGATCT TTCTCCCTTT 1440

CTCCTCTCCT TTTTCTCTTC ACATCTCCCC CATAGCACCC TGCCCTCATG GGACCTGCCC 1500  
 5 TCCCTCAGCC GTCAGCCATC AGCCATGGCC CTCCCAGTGC CTCCTAGCCC CTTCTTCCAA 1560  
 GGAGCAGAGA GGTGGCCACC GGGGGTGGCT CTGTCCTACC TCCACTCTCT GCCCCTAAAG 1620  
 ATGGGAGGAG ACCAGCGGTC CATGGGTCTG GCCTGTGAGT CTCCCCTTGC AGCCTGGTCA 1680  
 10 CTAGGCATCA CCCCCGCTTT GTTCTTTCAG ATGCTCTTGG GGTTCATAGG GGCAGGTCCT 1740  
 AGTCGGGCAG GGCCCTGAC CCTCCCGGCC TGGCTTCACT CTCCCTGACG GCTGCCATTG 1800  
 15 GTCCACCCTT TCATAGAGAG GCCTGCTTTG TTACAAAGCT CGGGTCTCCC TCCTGCAGCT 1860  
 CGGTTAAGTA CCCGAGGCTT CTCTTAAGAT GTCCAGGGCC CCAGGCCCCG GGGCACAGCC 1920  
 AGCCCAAACC TTGGGCCCTG GAAGAGTCTT CCACCCCATC ACTAGAGTGC TCTGACCCTG 1980  
 20 GGCTTTCACG GGCCCCATTC CACCGCTCC CCAACTTGAG CCTGTGACCT TGGGACCAAA 2040  
 GGGGGAGTCC CTCGTCTCTT GTGACTCAGC AGAGGCAGTG GCCACGTTC GGGAGGGGCC 2100  
 25 GGCTGGCCTG GAGGCTCAGC CCACCCTCCA GCTTTTCCTC AGGGTGTCTT GAGGTCCAAG 2160  
 ATTCTGGAGC AATCTGACCC TTCTCAAAG GCTCTGTTAT CAGCTGGGCA GTGCCAGCCA 2220  
 ATCCCTGGCC ATTTGGCCCC AGGGGACGTG GGCCCTG 2257

## (2) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 411 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: amino acid (Translation of Contig 2692004)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

His Ala Asp Arg Arg Arg Glu Ile Leu Ala Lys Tyr Pro Glu Ile  
 1 5 10 15  
 Lys Ser Leu Met Lys Pro Asp Pro Asn Leu Ile Trp Ile Ile Ile  
 20 25 30  
 Met Met Val Leu Thr Gln Leu Gly Ala Phe Tyr Ile Val Lys Asp  
 35 40 45  
 50 Leu Asp Trp Lys Trp Val Ile Phe Gly Ala Tyr Ala Phe Gly Ser  
 50 55 60  
 Cys Ile Asn His Ser Met Thr Leu Ala Ile His Glu Ile Ala His  
 65 70 75  
 Asn Ala Ala Phe Gly Asn Cys Lys Ala Met Trp Asn Arg Trp Phe  
 80 85 90  
 55 Gly Met Phe Ala Asn Leu Pro Ile Gly Ile Pro Tyr Ser Ile Ser  
 95 100 105  
 Phe Lys Arg Tyr His Met Asp His His Arg Tyr Leu Gly Ala Asp  
 110 115 120  
 60 Gly Val Asp Val Asp Ile Pro Thr Asp Phe Glu Gly Trp Phe Phe  
 125 130 135  
 Cys Thr Ala Phe Arg Lys Phe Ile Trp Val Ile Leu Gln Pro Leu  
 140 145 150  
 65 Phe Tyr Ala Phe Arg Pro Leu Phe Ile Asn Pro Lys Pro Ile Thr  
 155 160 165  
 Tyr Leu Glu Val Ile Asn Thr Val Ala Gln Val Thr Phe Asp Ile

		170		175		180
	Leu Ile Tyr Tyr	Phe Leu Gly Ile Lys	Ser Leu Val Tyr Met	Leu		
		185		190		195
5	Ala Ala Ser Leu	Leu Gly Leu Gly Leu	His Pro Ile Ser Gly	His		
		200		205		210
	Phe Ile Ala Glu	His Tyr Met Phe Leu	Lys Gly His Glu Thr	Tyr		
		215		220		225
	Ser Tyr Tyr Gly	Pro Leu Asn Leu Leu	Thr Phe Asn Val Gly	Tyr		
		230		235		240
10	His Asn Glu His	His Asp Phe Pro Asn	Ile Pro Gly Lys Ser	Leu		
		245		250		255
	Pro Leu Val Arg	Lys Ile Ala Ala Glu	Tyr Tyr Asp Asn Leu	Pro		
		260		265		270
	His Tyr Asn Ser	Trp Ile Lys Val Leu	Tyr Asp Phe Val Met	Asp		
15		275		280		285
	Asp Thr Ile Ser	Pro Tyr Ser Arg Met	Lys Arg His Gln Lys	Gly		
		290		295		300
	Glu Met Val Leu	Glu *** Ile Ser Leu	Val Pro Lys Gly Phe	Phe		
		305		310		315
20	Ser Lys Thr Leu	Asp Asp Lys Met Glu	Phe Leu His Tyr ***	Thr		
		320		325		330
	*** Asp Gln ***	Cys Ser Glu Ala Pro	Leu Ala Gln Phe Gln	Ser		
		335		340		345
25	Lys Ser Ser Val	Ile Pro Arg Ser Glu	Ser Gly Phe *** Thr	Val		
		350		355		360
	Ser Leu Thr Leu	Tyr Cys Ser Val Ser	Leu Thr Gly Asn Leu	***		
		365		370		375
	Leu Val Tyr Tyr	Arg His *** Gly Cys	Phe Thr His Val Cys	His		
		380		385		390
30	Phe Ile Ser Ile	Ser Phe Lys Lys Leu	Leu Lys Ser Tyr Phe	Ala		
		400		405		410
	Arg					

## (2) INFORMATION FOR SEQ ID NO:35:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 218 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid (Translation of Contig 2153526)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

	Tyr Leu Leu Arg Pro Leu Leu Pro His Leu Cys Ala Thr Ile Gly	
	1	15
	Ala Glu Ser Phe Leu Gly Leu Phe Phe Ile Val Arg Phe Leu Glu	
50		30
	Ser Asn Trp Phe Val Trp Val Thr Gln Met Asn His Ile Pro Met	
		45
	His Ile Asp His Asp Arg Asn Met Asp Trp Val Ser Thr Gln Leu	
		60
55	Gln Ala Thr Cys Asn Val His Lys Ser Ala Phe Asn Asp Trp Phe	
		75
	Ser Gly His Leu Asn Phe Gln Ile Glu His His Leu Phe Pro Thr	
		90
60	Met Pro Arg His Asn Tyr His Lys Val Ala Pro Leu Val Gln Ser	
		105
	Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser Lys Pro Leu Leu	
		120
	Ser Ala Phe Ala Asp Ile Ile His Ser Leu Lys Glu Ser Gly Gln	
		135
65	Leu Trp Leu Asp Ala Tyr Leu His Gln *** Gln Gln Pro Pro Cys	
		150

Pro Val Trp Lys Lys Arg Arg Lys Thr Leu Glu Pro Arg Gln Arg  
 155 160 165  
 Gly Ala \*\*\* Gly Thr Met Pro Leu \*\*\* Phe Asn Thr Gln Arg Gly  
 170 175 180  
 5 Leu Gly Leu Gly Thr \*\*\* Ser Leu \*\*\* Leu Lys Leu Leu Pro Phe  
 185 190 195  
 Ile Phe \*\*\* Pro Gln Phe \*\*\* Asp Pro Lys Trp Gly Val Asp Thr  
 200 205 210  
 10 Glu Val Pro Arg Arg Glu Gly Ala  
 215

## (2) INFORMATION FOR SEQ ID NO:36:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 86 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid (Translation of Contig 3506132)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Val Phe Tyr Phe Gly Asn Gly Trp Ile Pro Thr Leu Ile Thr Ala  
 1 5 10 15  
 30 Phe Val Leu Ala Thr Ser Gln Ala Gln Ala Gly Trp Leu Gln His  
 20 25 30  
 Asp Tyr Gly His Leu Ser Val Tyr Arg Lys Pro Lys Trp Asn His  
 35 35 40 45  
 Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala Ser Ala  
 35 50 55 60  
 Asn Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro Asn  
 65 70 75  
 Leu Gly Glu Trp Gln Pro Ile Glu Tyr Gly Lys Xxx  
 80 85  
 40

## (2) INFORMATION FOR SEQ ID NO:37:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 306 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid (Translation of Contig 3854933)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Gln Gly Pro Thr Pro Arg Tyr Phe Thr Trp Asp Glu Val Ala Gln  
 1 5 10 15  
 Arg Ser Gly Cys Glu Glu Arg Trp Leu Val Ile Asp Arg Lys Val  
 20 25 30  
 60 Tyr Asn Ile Ser Glu Phe Thr Arg Arg His Pro Gly Gly Ser Arg  
 35 40 45  
 Val Ile Ser His Tyr Ala Gly Gln Asp Ala Thr Asp Pro Phe Val  
 50 55 60  
 Ala Phe His Ile Asn Lys Gly Leu Val Lys Lys Tyr Met Asn Ser  
 65 65 70 75  
 Leu Leu Ile Gly Glu Leu Ser Pro Glu Gln Pro Ser Phe Glu Pro

		80		85		90									
	Thr	Lys	Asn	Lys	Glu	Leu	Thr	Asp	Glu	Phe	Arg	Glu	Leu	Arg	Ala
					95					100					105
5	Thr	Val	Glu	Arg	Met	Gly	Leu	Met	Lys	Ala	Asn	His	Val	Phe	Phe
					110					115					120
	Leu	Leu	Tyr	Leu	Leu	His	Ile	Leu	Leu	Leu	Asp	Gly	Ala	Ala	Trp
					125					130					135
	Leu	Thr	Leu	Trp	Val	Phe	Gly	Thr	Ser	Phe	Leu	Pro	Phe	Leu	Leu
					140					145					150
10	Cys	Ala	Val	Leu	Leu	Ser	Ala	Val	Gln	Ala	Gln	Ala	Gly	Trp	Leu
					155					160					165
	Gln	His	Asp	Phe	Gly	His	Leu	Ser	Val	Phe	Ser	Thr	Ser	Lys	Trp
					170					175					180
	Asn	His	Leu	Leu	His	His	Phe	Val	Ile	Gly	His	Leu	Lys	Gly	Ala
15					185					190					195
	Pro	Ala	Ser	Trp	Trp	Asn	His	Met	His	Phe	Gln	His	His	Ala	Lys
					200					205					210
	Pro	Asn	Cys	Phe	Arg	Lys	Asp	Pro	Asp	Ile	Asn	Met	His	Pro	Phe
					215					220					225
20	Phe	Phe	Ala	Leu	Gly	Lys	Ile	Leu	Ser	Val	Glu	Leu	Gly	Lys	Gln
					230					235					240
	Lys	Lys	Lys	Tyr	Met	Pro	Tyr	Asn	His	Gln	His	Xxx	Tyr	Phe	Phe
					245					250					255
	Leu	Ile	Gly	Pro	Pro	Ala	Leu	Leu	Pro	Leu	Tyr	Phe	Gln	Trp	Tyr
25					260					265					270
	Ile	Phe	Tyr	Phe	Val	Ile	Gln	Arg	Lys	Lys	Trp	Val	Asp	Leu	Ala
					275					280					285
	Trp	Ile	Ser	Lys	Gln	Glu	Tyr	Asp	Glu	Ala	Gly	Leu	Pro	Leu	Ser
					290					295					300
30	Thr	Ala	Asn	Ala	Ser	Lys									
					305										

## (2) INFORMATION FOR SEQ ID NO:38:

35

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 566 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

40

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid (Translation of Contig 2511785)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

45

	His	Leu	Lys	Gly	Ala	Ser	Ala	Asn	Trp	Trp	Asn	His	Arg	His	Phe
	1				5					10					15
50	Gln	His	His	Ala	Lys	Pro	Asn	Ile	Phe	His	Lys	Asp	Pro	Asp	Val
					20					25					30
	Asn	Met	Leu	His	Val	Phe	Val	Leu	Gly	Glu	Trp	Gln	Pro	Ile	Glu
					35					40					45
	Tyr	Gly	Lys	Lys	Lys	Leu	Lys	Tyr	Leu	Pro	Tyr	Asn	His	Gln	His
					50					55					60
55	Glu	Tyr	Phe	Phe	Leu	Ile	Gly	Pro	Pro	Leu	Leu	Ile	Pro	Met	Tyr
					65					70					75
	Phe	Gln	Tyr	Gln	Ile	Ile	Met	Thr	Met	Ile	Val	His	Lys	Asn	Trp
					80					85					90
	Val	Asp	Leu	Ala	Trp	Ala	Val	Ser	Tyr	Tyr	Ile	Arg	Phe	Phe	Ile
60					95					100					105
	Thr	Tyr	Ile	Pro	Phe	Tyr	Gly	Ile	Leu	Gly	Ala	Leu	Leu	Phe	Leu
					110					115					120
	Asn	Phe	Ile	Arg	Phe	Leu	Glu	Ser	His	Trp	Phe	Val	Trp	Val	Thr
					125					130					135
65	Gln	Met	Asn	His	Ile	Val	Met	Glu	Ile	Asp	Gln	Glu	Ala	Tyr	Arg
					140					145					150

	Asp Trp Phe Ser	Ser Gln Leu Thr Ala Thr Cys Asn Val Glu Gln	155	160	165
	Ser Phe Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile		170	175	180
5	Glu His His Leu Phe Pro Thr Met Pro Arg His Asn Leu His Lys		185	190	195
	Ile Ala Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Ile Glu		200	205	210
10	Tyr Gln Glu Lys Pro Leu Leu Arg Ala Leu Leu Asp Ile Ile Arg		215	220	225
	Ser Leu Lys Lys Ser Gly Lys Leu Trp Leu Asp Ala Tyr Leu His		230	235	240
	Lys *** Ser His Ser Pro Arg Asp Thr Val Gly Lys Gly Cys Arg		245	250	255
15	Trp Gly Asp Gly Gln Arg Asn Asp Gly Leu Leu Phe *** Gly Val		260	265	270
	Ser Glu Arg Leu Val Tyr Ala Leu Leu Thr Asp Pro Met Leu Asp		275	280	285
20	Leu Ser Pro Phe Leu Leu Ser Phe Phe Ser Ser His Leu Pro His		290	295	300
	Ser Thr Leu Pro Ser Trp Asp Leu Pro Ser Leu Ser Arg Gln Pro		305	310	315
	Ser Ala Met Ala Leu Pro Val Pro Pro Ser Pro Phe Phe Gln Gly		320	325	330
25	Ala Glu Arg Trp Pro Pro Gly Val Ala Leu Ser Tyr Leu His Ser		335	340	345
	Leu Pro Leu Lys Met Gly Gly Asp Gln Arg Ser Met Gly Leu Ala		350	355	360
30	Cys Glu Ser Pro Leu Ala Ala Trp Ser Leu Gly Ile Thr Pro Ala		365	370	375
	Leu Val Leu Gln Met Leu Leu Gly Phe Ile Gly Ala Gly Pro Ser		380	385	390
	Arg Ala Gly Pro Leu Thr Leu Pro Ala Trp Leu His Ser Pro ***		400	405	410
35	Arg Leu Pro Leu Val His Pro Phe Ile Glu Arg Pro Ala Leu Leu		415	420	425
	Gln Ser Ser Gly Leu Pro Pro Ala Ala Arg Leu Ser Thr Arg Gly		430	435	440
40	Leu Ser *** Asp Val Gln Gly Pro Arg Pro Ala Gly Thr Ala Ser		445	450	455
	Pro Asn Leu Gly Pro Trp Lys Ser Pro Pro Pro His His *** Ser		460	465	470
	Ala Leu Thr Leu Gly Phe His Gly Pro His Ser Thr Ala Ser Pro		475	480	485
45	Thr *** Ala Cys Asp Leu Gly Thr Lys Gly Gly Val Pro Arg Leu		490	495	500
	Leu *** Leu Ser Arg Gly Ser Gly His Val Gln Gly Gly Ala Gly		505	510	515
50	Trp Pro Gly Gly Ser Ala His Pro Pro Ala Phe Pro Gln Gly Val		520	525	530
	Leu Arg Ser Lys Ile Leu Glu Gln Ser Asp Pro Ser Pro Lys Ala		535	540	545
	Leu Leu Ser Ala Gly Gln Cys Gln Pro Ile Pro Gly His Leu Ala		550	555	560
55	Pro Gly Asp Val Gly Pro Xxx		565		

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 619 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear



(ii) MOLECULE TYPE: amino acid (Translation of Contig 2535)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

5

	Val	Phe	Tyr	Phe	Gly	Asn	Gly	Trp	Ile	Pro	Thr	Leu	Ile	Thr	Ala	
	1				5					10					15	
10	Phe	Val	Leu	Ala	Thr	Ser	Gln	Ala	Gln	Ala	Gly	Trp	Leu	Gln	His	
					20					25					30	
	Asp	Tyr	Gly	His	Leu	Ser	Val	Tyr	Arg	Lys	Pro	Lys	Trp	Asn	His	
					35					40					45	
	Leu	Val	His	Lys	Phe	Val	Ile	Gly	His	Leu	Lys	Gly	Ala	Ser	Ala	
					50					55					60	
15	Asn	Trp	Trp	Asn	His	Arg	His	Phe	Gln	His	His	Ala	Lys	Pro	Asn	
					65					70					75	
	Ile	Phe	His	Lys	Asp	Pro	Asp	Val	Asn	Met	Leu	His	Val	Phe	Val	
					80					85					90	
	Leu	Gly	Glu	Trp	Gln	Pro	Ile	Glu	Tyr	Gly	Lys	Lys	Lys	Leu	Lys	
20					95					100					105	
	Tyr	Leu	Pro	Tyr	Asn	His	Gln	His	Glu	Tyr	Phe	Phe	Leu	Ile	Gly	
					110					115					120	
	Pro	Pro	Leu	Leu	Ile	Pro	Met	Tyr	Phe	Gln	Tyr	Gln	Ile	Ile	Met	
					125					130					135	
25	Thr	Met	Ile	Val	His	Lys	Asn	Trp	Val	Asp	Leu	Ala	Trp	Ala	Val	
					140					145					150	
	Ser	Tyr	Tyr	Ile	Arg	Phe	Phe	Ile	Thr	Tyr	Ile	Pro	Phe	Tyr	Gly	
					155					160					165	
	Ile	Leu	Gly	Ala	Leu	Leu	Phe	Leu	Asn	Phe	Ile	Arg	Phe	Leu	Glu	
30					170					175					180	
	Ser	His	Trp	Phe	Val	Trp	Val	Thr	Gln	Met	Asn	His	Ile	Val	Met	
					185					190					195	
	Glu	Ile	Asp	Gln	Glu	Ala	Tyr	Arg	Asp	Trp	Phe	Ser	Ser	Gln	Leu	
					200					205					210	
35	Thr	Ala	Thr	Cys	Asn	Val	Glu	Gln	Ser	Phe	Phe	Asn	Asp	Trp	Phe	
					215					220					225	
	Ser	Gly	His	Leu	Asn	Phe	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Thr	
					230					235					240	
40	Met	Pro	Arg	His	Asn	Leu	His	Lys	Ile	Ala	Pro	Leu	Val	Lys	Ser	
					245					250					255	
	Leu	Cys	Ala	Lys	His	Gly	Ile	Glu	Tyr	Gln	Glu	Lys	Pro	Leu	Leu	
					260					265					270	
	Arg	Ala	Leu	Leu	Asp	Ile	Ile	Arg	Ser	Leu	Lys	Lys	Ser	Gly	Lys	
					275					280					285	
45	Leu	Trp	Leu	Asp	Ala	Tyr	Leu	His	Lys	***	Ser	His	Ser	Pro	Arg	
					290					295					300	
	Asp	Thr	Val	Gly	Lys	Gly	Cys	Arg	Trp	Gly	Asp	Gly	Gln	Arg	Asn	
					305					310					315	
50	Asp	Gly	Leu	Leu	Phe	***	Gly	Val	Ser	Glu	Arg	Leu	Val	Tyr	Ala	
					320					325					330	
	Leu	Leu	Thr	Asp	Pro	Met	Leu	Asp	Leu	Ser	Pro	Phe	Leu	Leu	Ser	
					335					340					345	
	Phe	Phe	Ser	Ser	His	Leu	Pro	His	Ser	Thr	Leu	Pro	Ser	Trp	Asp	
					350					355					360	
55	Leu	Pro	Ser	Leu	Ser	Arg	Gln	Pro	Ser	Ala	Met	Ala	Leu	Pro	Val	
					365					370					375	
	Pro	Pro	Ser	Pro	Phe	Phe	Gln	Gly	Ala	Glu	Arg	Trp	Pro	Pro	Gly	
					380					385					390	
	Val	Ala	Leu	Ser	Tyr	Leu	His	Ser	Leu	Pro	Leu	Lys	Met	Gly	Gly	
60					400					405					410	
	Asp	Gln	Arg	Ser	Met	Gly	Leu	Ala	Cys	Glu	Ser	Pro	Leu	Ala	Ala	
					415					420					425	
	Trp	Ser	Leu	Gly	Ile	Thr	Pro	Ala	Leu	Val	Leu	Gln	Met	Leu	Leu	
					430					435					440	
65	Gly	Phe	Ile	Gly	Ala	Gly	Pro	Ser	Arg	Ala	Gly	Pro	Leu	Thr	Leu	
					445					450					455	

5 Pro Ala Trp Leu His Ser Pro \*\*\* Arg Leu Pro Leu Val His Pro  
 460 465 470  
 Phe Ile Glu Arg Pro Ala Leu Leu Gln Ser Ser Gly Leu Pro Pro  
 475 480 485  
 10 Ala Ala Arg Leu Ser Thr Arg Gly Leu Ser \*\*\* Asp Val Gln Gly  
 490 495 500  
 Pro Arg Pro Ala Gly Thr Ala Ser Pro Asn Leu Gly Pro Trp Lys  
 505 510 515  
 Ser Pro Pro Pro His His \*\*\* Ser Ala Leu Thr Leu Gly Phe His  
 520 525 530  
 Gly Pro His Ser Thr Ala Ser Pro Thr \*\*\* Ala Cys Asp Leu Gly  
 535 540 545  
 Thr Lys Gly Gly Val Pro Arg Leu Leu \*\*\* Leu Ser Arg Gly Ser  
 550 555 560  
 15 Gly His Val Gln Gly Gly Ala Gly Trp Pro Gly Gly Ser Ala His  
 565 570 575  
 Pro Pro Ala Phe Pro Gln Gly Val Leu Arg Ser Lys Ile Leu Glu  
 580 585 590  
 Gln Ser Asp Pro Ser Pro Lys Ala Leu Leu Ser Ala Gly Gln Cys  
 595 600 605  
 20 Gln Pro Ile Pro Gly His Leu Ala Pro Gly Asp Val Gly Pro Xxx  
 610 615 620

25

(2) INFORMATION FOR SEQ ID NO:40:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 757 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: amino acid (Translation of Contig 253538a)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

40 Gln Gly Pro Thr Pro Arg Tyr Phe Thr Trp Asp Glu Val Ala Gln  
 1 5 10 15  
 Arg Ser Gly Cys Glu Glu Arg Trp Leu Val Ile Asp Arg Lys Val  
 20 25 30  
 Tyr Asn Ile Ser Glu Phe Thr Arg Arg His Pro Gly Gly Ser Arg  
 35 40 45  
 45 Val Ile Ser His Tyr Ala Gly Gln Asp Ala Thr Asp Pro Phe Val  
 50 55 60  
 Ala Phe His Ile Asn Lys Gly Leu Val Lys Lys Tyr Met Asn Ser  
 65 70 75  
 50 Leu Leu Ile Gly Glu Leu Ser Pro Glu Gln Pro Ser Phe Glu Pro  
 80 85 90  
 Thr Lys Asn Lys Glu Leu Thr Asp Glu Phe Arg Glu Leu Arg Ala  
 95 100 105  
 Thr Val Glu Arg Met Gly Leu Met Lys Ala Asn His Val Phe Phe  
 110 115 120  
 55 Leu Leu Tyr Leu Leu His Ile Leu Leu Leu Asp Gly Ala Ala Trp  
 125 130 135  
 Leu Thr Leu Trp Val Phe Gly Thr Ser Phe Leu Pro Phe Leu Leu  
 140 145 150  
 60 Cys Ala Val Leu Leu Ser Ala Val Gln Gln Ala Gln Ala Gly Trp  
 155 160 165  
 Leu Gln His Asp Tyr Gly His Leu Ser Val Tyr Arg Lys Pro Lys  
 170 175 180  
 Trp Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly  
 185 190 195  
 65 Ala Ser Ala Asn Trp Trp Asn His Arg His Phe Gln His His Ala  
 200 205 210

	Lys	Pro	Asn	Ile	Phe	His	Lys	Asp	Pro	Asp	Val	Asn	Met	Leu	His
					215					220					225
	Val	Phe	Val	Leu	Gly	Glu	Trp	Gln	Pro	Ile	Glu	Tyr	Gly	Lys	Lys
					230					235					240
5	Lys	Leu	Lys	Tyr	Leu	Pro	Tyr	Asn	His	Gln	His	Glu	Tyr	Phe	Phe
					245					250					255
	Leu	Ile	Gly	Pro	Pro	Leu	Leu	Ile	Pro	Met	Tyr	Phe	Gln	Tyr	Gln
					260					265					270
10	Ile	Ile	Met	Thr	Met	Ile	Val	His	Lys	Asn	Trp	Val	Asp	Leu	Ala
					275					280					285
	Trp	Ala	Val	Ser	Tyr	Tyr	Ile	Arg	Phe	Phe	Ile	Thr	Tyr	Ile	Pro
					290					295					300
	Phe	Tyr	Gly	Ile	Leu	Gly	Ala	Leu	Leu	Phe	Leu	Asn	Phe	Ile	Arg
					305					310					315
15	Phe	Leu	Glu	Ser	His	Trp	Phe	Val	Trp	Val	Thr	Gln	Met	Asn	His
					320					325					330
	Ile	Val	Met	Glu	Ile	Asp	Gln	Glu	Ala	Tyr	Arg	Asp	Trp	Phe	Ser
					335					340					345
20	Ser	Gln	Leu	Thr	Ala	Thr	Cys	Asn	Val	Glu	Gln	Ser	Phe	Phe	Asn
					350					355					360
	Asp	Trp	Phe	Ser	Gly	His	Leu	Asn	Phe	Gln	Ile	Glu	His	His	Leu
					365					370					375
	Phe	Pro	Thr	Met	Pro	Arg	His	Asn	Leu	His	Lys	Ile	Ala	Pro	Leu
					380					385					390
25	Val	Lys	Ser	Leu	Cys	Ala	Lys	His	Gly	Ile	Glu	Tyr	Gln	Glu	Lys
					400					405					410
	Pro	Leu	Leu	Arg	Ala	Leu	Leu	Asp	Ile	Ile	Arg	Ser	Leu	Lys	Lys
					415					420					425
30	Ser	Gly	Lys	Leu	Trp	Leu	Asp	Ala	Tyr	Leu	His	Lys	***	Ser	His
					430					435					440
	Ser	Pro	Arg	Asp	Thr	Val	Gly	Lys	Gly	Cys	Arg	Trp	Gly	Asp	Gly
					445					450					455
	Gln	Arg	Asn	Asp	Gly	Leu	Leu	Phe	***	Gly	Val	Ser	Glu	Arg	Leu
					460					465					470
35	Val	Tyr	Ala	Leu	Leu	Thr	Asp	Pro	Met	Leu	Asp	Leu	Ser	Pro	Phe
					475					480					485
	Leu	Leu	Ser	Phe	Phe	Ser	Ser	His	Leu	Pro	His	Ser	Thr	Leu	Pro
					490					495					500
40	Ser	Trp	Asp	Leu	Pro	Ser	Leu	Ser	Arg	Gln	Pro	Ser	Ala	Met	Ala
					505					510					515
	Leu	Pro	Val	Pro	Pro	Ser	Pro	Phe	Phe	Gln	Gly	Ala	Glu	Arg	Trp
					520					525					530
	Pro	Pro	Gly	Val	Ala	Leu	Ser	Tyr	Leu	His	Ser	Leu	Pro	Leu	Lys
					535					540					545
45	Met	Gly	Gly	Asp	Gln	Arg	Ser	Met	Gly	Leu	Ala	Cys	Glu	Ser	Pro
					550					555					560
	Leu	Ala	Ala	Trp	Ser	Leu	Gly	Ile	Thr	Pro	Ala	Leu	Val	Leu	Gln
					565					570					575
50	Met	Leu	Leu	Gly	Phe	Ile	Gly	Ala	Gly	Pro	Ser	Arg	Ala	Gly	Pro
					580					585					590
	Leu	Thr	Leu	Pro	Ala	Trp	Leu	His	Ser	Pro	***	Arg	Leu	Pro	Leu
					595					600					605
	Val	His	Pro	Phe	Ile	Glu	Arg	Pro	Ala	Leu	Leu	Gln	Ser	Ser	Gly
					610					615					620
55	Leu	Pro	Pro	Ala	Ala	Arg	Leu	Ser	Thr	Arg	Gly	Leu	Ser	***	Asp
					625					630					635
	Val	Gln	Gly	Pro	Arg	Pro	Ala	Gly	Thr	Ala	Ser	Pro	Asn	Leu	Gly
					640					645					650
60	Pro	Trp	Lys	Ser	Pro	Pro	Pro	His	His	***	Ser	Ala	Leu	Thr	Leu
					655					660					665
	Gly	Phe	His	Gly	Pro	His	Ser	Thr	Ala	Ser	Pro	Thr	***	Ala	Cys
					670					675					680
	Asp	Leu	Gly	Thr	Lys	Gly	Gly	Val	Pro	Arg	Leu	Leu	***	Leu	Ser
					685					690					695
65	Arg	Gly	Ser	Gly	His	Val	Gln	Gly	Gly	Ala	Gly	Trp	Pro	Gly	Gly
					700					705					710

5      Ser Ala His Pro    Pro Ala Phe Pro Gln Gly Val Leu Arg Ser Lys  
                               715                         720                         725  
Ile Leu Glu Gln Ser Asp Pro Ser Pro Lys Ala Leu Leu Ser Ala  
                               730                         735                         740  
Gly Gln Cys Gln Pro Ile Pro Gly His Leu Ala Pro Gly Asp Val  
                               745                         750                         755  
Gly Pro Xxx

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A purified or isolated polypeptide which is capable of desaturating a fatty acid molecule at carbon 6 or 12 from the carboxyl end of said fatty acid, said polypeptide  
5 having an amino acid sequence which has at least 60% homology to the 457 amino acid sequence of SEQ ID NO:2 or the 399 amino acid sequence of SEQ ID NO:4
2. A polypeptide according to claim 1 having an amino acid sequence which has at least 80% homology to the 457 amino acid sequence of SEQ ID NO: 2.
- 10 3. A polypeptide according to claim 1 having an amino acid sequence which has at least 80% homology to the 399 amino acid sequence of SEQ ID NO: 4.
4. A polypeptide according to claim 1 having an amino acid sequence which has at  
15 least 90% homology to the 457 amino acid sequence of SEQ ID NO: 2.
5. A polypeptide according to claim 1 having an amino acid sequence which has at least 90% homology to the 399 amino acid sequence of SEQ ID NO: 4.
- 20 6. A polypeptide according to claim 1, 2 or 4 wherein said polypeptide includes an amino acid motif selected from the group consisting of residues 50-53, 39-43, 172-176, 204-213 and 390-402 of SEQ ID NO: 2.
7. A polypeptide according to claim 6 wherein said polypeptide comprises residues 50-  
25 53, 39-43, 172-176, 204-213 and 390-402 of SEQ ID NO: 2.
8. A polypeptide according to any one of the claims 1, 2, 4, 6 or 7 comprising SEQ ID NO: 2.
- 30 9. An isolated nucleic acid encoding a polypeptide as defined in anyone of claims 1 to 8.



10. An isolated nucleic acid according to claim 9 which comprises SEQ ID NO: 1 or  
SEQ ID NO: 3.
11. A nucleic acid construct comprising a nucleic acid as defined in claim 9 or 10  
operably linked to a promoter.
12. A host cell transformed with the construct of claim 11.
13. A host cell according to claim 12 which is a microbial host cell.
14. A host cell according to claim 13 which is a yeast cell.
15. A method for the production of fatty acid gamma linolenic acid which method  
comprises:  
  
growing a culture of host cells of claim 12, 13 or 14 which cells produce a  
delta 6 desaturase in the presence of linoleic acid, under conditions wherein said acid  
is converted to gamma linolenic acid by the expression of the polypeptide of said  
desaturase; and recovering the fatty acid gamma linolenic acid from the culture.
16. A method for the production of the fatty acid linoleic acid which method comprises:  
  
growing a culture of host cells of claim 12, 13 or 14 which cells produce a  
delta 12 desaturase in the presence of oleic acid, under conditions wherein said acid  
is converted to linolenic acid by the expression of the polypeptide of said desaturase;  
and recovering the fatty acid linolenic acid from the culture.
17. A method according to any one of claims 15 or 16 which further comprises  
formulating the fatty acid into a product selected from the group:  
  
a pharmaceutical composition comprising said oil and a pharmaceutically  
acceptable carrier;  
a nutritional formula;



an infant formula;  
a dietary supplement;  
a dietary substitute;  
a cosmetic; and  
an animal feed

5

18. A method according to claim 17 wherein said infant formula, dietary supplement or dietary substitute is in the form of a liquid or a solid.

10 19. A method according to claim 17 or 18 wherein the nutritional formula, infant formula, dietary supplement or dietary substitute contains at least one macronutrient selected from the group consisting of coconut oil, soy oil, canola oil, mono- and di-glycerides, glucose, edible lactose, electrodialysed whey, electrodialysed skimmed milk, milk whey, soy protein, and other protein hydrolysates.

15

20. A method according to claim 19 herein said nutritional formula, infant formula, dietary supplement or dietary substitute contains at least one vitamin selected from the group consisting of A, C, D, E and B complex; and at least one mineral selected from the group consisting of calcium, magnesium, zinc, manganese, sodium, potassium, phosphorus, copper, chloride, iodine, selenium and iron.

20

21. The fatty acid gamma linolenic acid derived from the host cells of any one of claims 12, 13 or 14.

25

22. The fatty acid linoleic acid derived from the host cells of any one of claims 12, 13 or 14.

23. A product comprising the fatty acid of any one of claims 21 or 22 selected from the group:

30

a pharmaceutical composition comprising said oil and a pharmaceutically acceptable carrier;  
a nutritional formula;



an infant formula;  
a dietary supplement;  
a dietary substitute;  
a cosmetic; and  
an animal feed.

5

24. A product as claimed in claim 23 wherein said infant formula, dietary supplement or dietary substitute is in the form of a liquid or a solid.

10 25. A product as claimed in claims 23 or 24 wherein the nutritional formula, infant formula, dietary supplement or dietary substitute contains at least one macronutrient selected from the group consisting of coconut oil, soy oil, canola oil. Mono- and di-glycerides, glucose, edible lactose, electrodialysed whey, electrodialysed skimmed milk, milk whey, soy protein, and other protein hydrolysates.

15

26. A product as claimed in claim 25 wherein the said nutritional formula, infant formula, dietary supplement or dietary substitute contains at least one vitamin selected from the group consisting of vitamin A, C, D, E and b complex; and at least one mineral selected from the group consisting of calcium, magnesium, zinc, manganese, sodium, potassium, phosphorus, copper, chloride, iodine, selenium and iron.

20

27. The use of microbial host cell of claim 13, 14 or 15 for the production of the fatty acid gamma linolenic acid or the fatty acid linoleic acid.

25

28. The use according to claim 27 for the production of a product comprising said fatty acid, said product being selected from the group:

30

a pharmaceutical composition comprising said oil and a pharmaceutically acceptable carrier;  
a nutritional formula;  
an infant formula;  
a dietary supplement;





a cosmetic; and  
an animal feed.

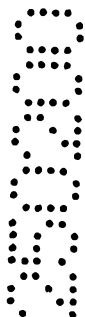
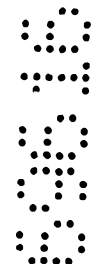
29. A polypeptide, nucleic acid construct, a host cell, a method of production of the fatty  
5 acid gamma linolenic acid or linoleic acid, the fatty acid gamma linolenic acid, the  
fatty acid linoleic acid, a product or the use of a host cell substantially as herein  
described with reference to any one of the accompanying figures or examples.

10 DATED this 24<sup>th</sup> day of February, 2000.

CALGENE LLC AND ABBOTT LABORATORIES

By its Patent Attorneys

15 DAVIES COLLISON CAVE



1/20

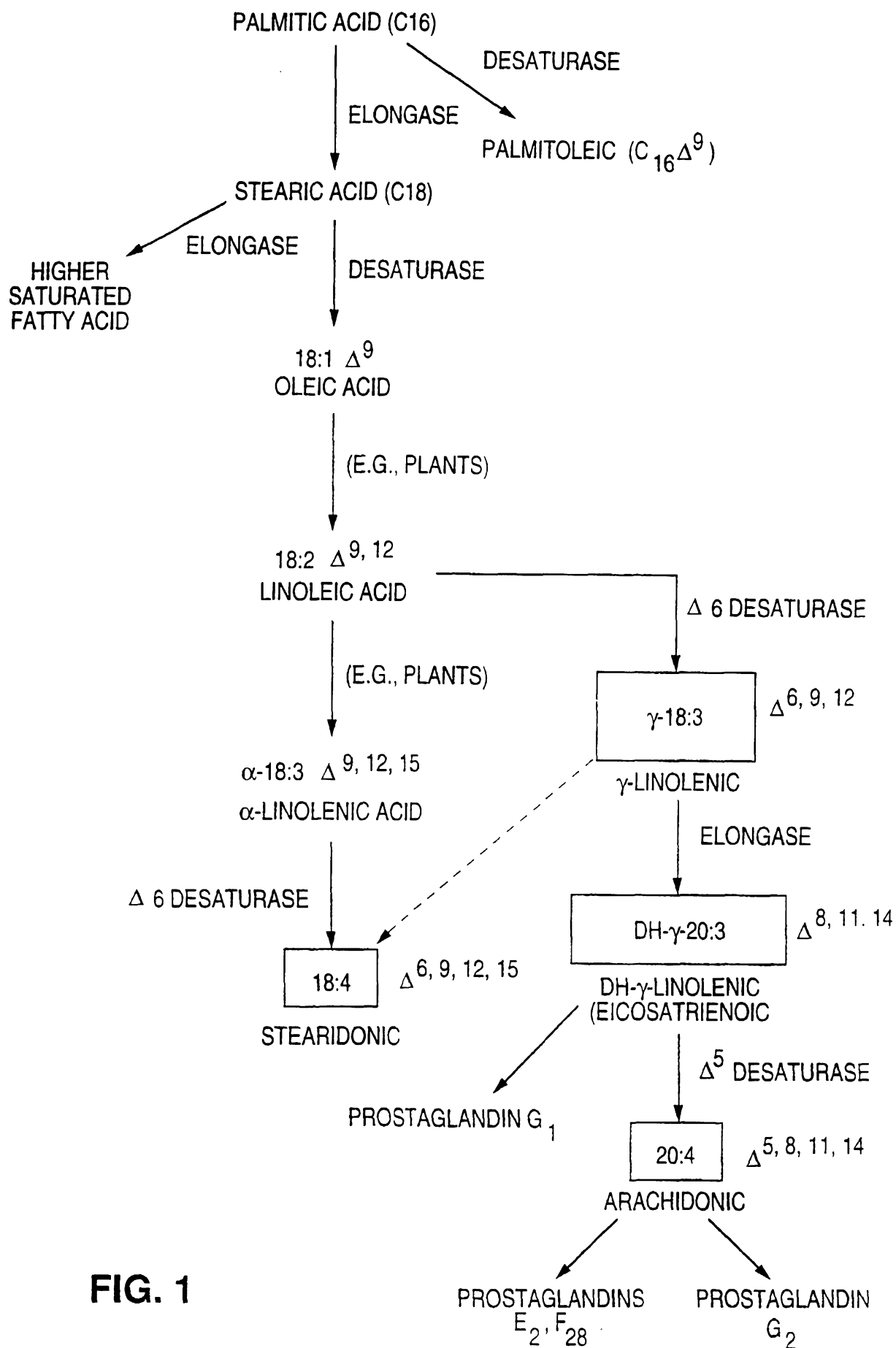


FIG. 1

2/20

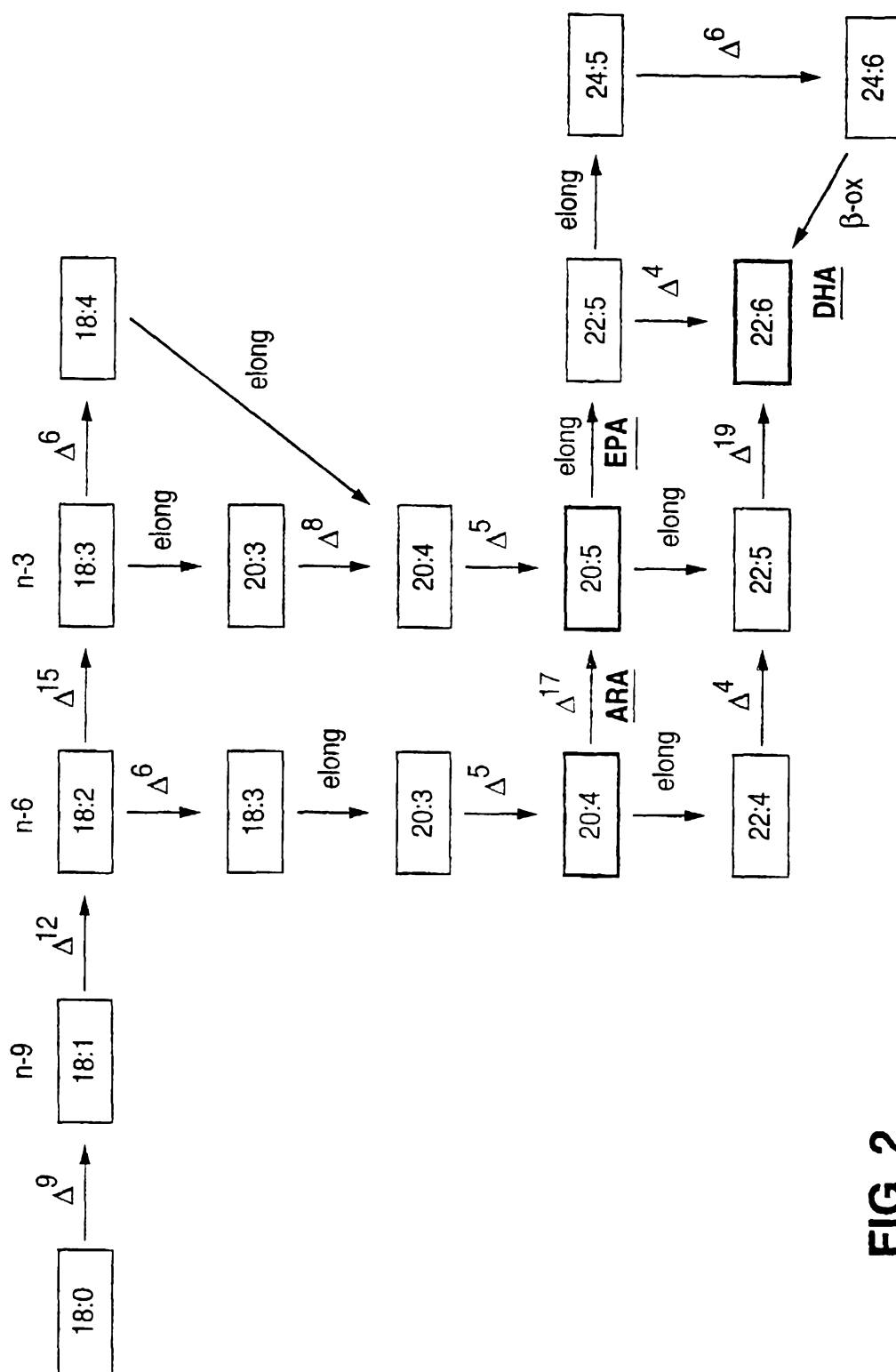


FIG. 2

																60
CGACACTCCT TCCTTCTTCT CACCCGTCCT AGTCCCCTTC AACCCCCCTC TTTGACAAAG																*
ACAACAAACC ATG GCT GCT GCT CCC AGT GTG AGG ACG TTT ACT CGG GCC GAG																
Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu																
																120
GTT TTG AAT GCC GAG GCT CTG AAT GAG GGC AAG AAG GAT GCC GAG GCA																*
Val Leu Asn Ala Glu Ala Leu Asn Glu Gly Lys Lys Asp Ala Glu Ala																
																180
CCC TTC TTG ATG ATC ATC GAC AAC AAG GTG TAC GAT GTC CGC GAG TTC																*
Pro Phe Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe																
																240
GTC CCT GAT CAT CCC GGT GGA AGT GTG ATT CTC ACG CAC GTT GGC AAG																*
Val Pro Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys																
																300
GAC GGC ACT GAC GTC TTT GAC ACT TTT CAC CCC GAG GCT GCT TGG GAG																*
Asp Gly Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu																
																360
ACT CTT GCC AAC TTT TAC GTT GGT GAT ATT GAC GAG AGC GAC CGC GAT																*
Thr Leu Ala Asn Phe Tyr Val Gly Asp Ile Asp Glu Ser Asp Arg Asp																
																360
ATC AAG AAT GAT GAC TTT GCG GCC GAG GTC CGC AAG CTG CGT ACC TTG																*
Ile Lys Asn Asp Asp Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu																

FIG. 3A

						420 *										
TTC	CAG	TCT	CTT	GGT	TAC	TAC	GAT	TCT	TCC	AAG	GCA	TAC	TAC	GCC	TTC	
Phe	Gln	Ser	Leu	Gly	Tyr	Tyr	Asp	Ser	Ser	Lys	Ala	Tyr	Tyr	Ala	Phe	
										480 *						
AAG	GTC	TCG	TTC	AAC	CTC	TGC	ATC	TGG	GGT	TTG	TCG	ACG	GTC	ATT	GTG	
Lys	Val	Ser	Phe	Asn	Leu	Cys	Ile	Trp	Gly	Leu	Ser	Thr	Val	Ile	Val	
														540 *		
GCC	AAG	TGG	GGC	CAG	ACC	TCG	ACC	CTC	GCC	AAC	GTG	CTC	TCG	GCT	GCG	
Ala	Lys	Trp	Gly	Gln	Thr	Ser	Thr	Leu	Ala	Asn	Val	Leu	Ser	Ala	Ala	
CTT	TTG	GGT	CTG	TTC	TGG	CAG	CAG	TGC	GGA	TGG	TTG	GCT	CAC	GAC	TTT	
Leu	Leu	Gly	Leu	Phe	Trp	Gln	Gln	Cys	Gly	Trp	Leu	Ala	His	Asp	Phe	
		600 *														
TTG	CAT	CAC	CAG	GTC	TTC	CAG	GAC	CGT	TTC	TGG	GGT	GAT	CTT	TTC	GGC	
Leu	His	His	Gln	Val	Phe	Gln	Asp	Arg	Phe	Trp	Gly	Asp	Leu	Phe	Gly	
						660 *										
GCC	TTC	TTG	GGA	GGT	GTC	TGC	CAG	GGC	TTC	TCG	TCC	TCG	TGG	TGG	AAG	
Ala	Phe	Leu	Gly	Gly	Val	Cys	Gln	Gly	Phe	Ser	Ser	Ser	Trp	Trp	Lys	
										720 *						
GAC	AAG	CAC	AAC	ACT	CAC	CAC	GCC	GCC	CCC	AAC	GTC	CAC	GGC	GAG	GAT	
Asp	Lys	His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	
														780 *		

FIG. 3B

CCC	GAC	ATT	GAC	ACC	CAC	CCT	CTG	TTG	ACC	TGG	AGT	GAG	CAT	GCG	TTG
Pro	Asp	Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu
GAG	ATG	TTC	TCG	GAT	GTC	CCA	GAT	GAG	GAG	CTG	ACC	CGC	ATG	TGG	TCG
Glu	Met	Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser
		840													
CGT	TTC	ATG	GTC	CTG	AAC	CAG	ACC	TGG	TTT	TAC	TTC	CCC	ATT	CTC	TCG
Arg	Phe	Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser
		*				900									
TTT	GCC	CGT	CTC	TCC	TGG	TGC	CTC	CAG	TCC	ATT	CTC	TTT	GTG	CTG	CCT
Phe	Ala	Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Leu	Phe	Val	Leu	Pro
						*				960					
AAC	GGT	CAG	GCC	CAC	AAG	CCC	TCG	GGC	GCG	CGT	GTG	CCC	ATC	TCG	TTG
Asn	Gly	Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu
										*					
GTC	GAG	CAG	CTG	TCG	CTT	GCG	ATG	CAC	TGG	ACC	TGG	TAC	CTC	GCC	ACC
Val	Glu	Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr
														1020	
ATG	TTC	CTG	TTC	ATC	AAG	GAT	CCC	GTC	AAC	ATG	CTG	GTG	TAC	TTT	TTG
Met	Phe	Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Leu	Val	Tyr	Phe	Leu
		1080				*									
GTG	TCG	CAG	GCG	GTG	TGC	GGA	AAC	TTG	TTG	GCG	ATC	GTG	TTC	TCG	CTC
Val	Ser	Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu

FIG. 3C

						1140										
						*										
AAC	CAC	AAC	GGT	ATG	CCT	GTG	ATC	TCG	AAG	GAG	GAG	GCG	GTC	GAT	ATG	
Asn	His	Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	Met	
										1200						
						*										
GAT	TTC	TTC	ACG	AAG	CAG	ATC	ATC	ACG	GGT	CGT	GAT	GTC	CAC	CCG	GGT	
Asp	Phe	Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	His	Pro	Gly	
														1260		
						*										
CTA	TTT	GCC	AAC	TGG	TTC	ACG	GGT	GGA	TTG	AAC	TAT	CAG	ATC	GAG	CAC	
Leu	Phe	Ala	Asn	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	
CAC	TTG	TTC	CCT	TCG	ATG	CCT	CGC	CAC	AAC	TTT	TCA	AAG	ATC	CAG	CCT	
His	Leu	Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	Gln	Pro	
		1320				*										
GCT	GTC	GAG	ACC	CTG	TGC	AAA	AAG	TAC	AAT	GTC	CGA	TAC	CAC	ACC	ACC	
Ala	Val	Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Asn	Val	Arg	Tyr	His	Thr	Thr	
						1380										
						*										
GGT	ATG	ATC	GAG	GGA	ACT	GCA	GAG	GTC	TTT	AGC	CGT	CTG	AAC	GAG	GTC	
Gly	Met	Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	Val	
										1440						
						*										
TCC	AAG	GCT	GCC	TCC	AAG	ATG	GGT	AAG	GCG	CAG	TAAAAAAA	AAACAAGGAC				
Ser	Lys	Ala	Ala	Ser	Lys	Met	Gly	Lys	Ala	Gln						

6/20

FIG. 3D

7/20

1500  
\*  
GTTTTTTTTTC GCCAGTGCCT GTGCCTGTGC CTGCTTCCCT TGTCAAGTCG AGCGTTTCTG  
1560  
\*  
GAAAGGATCG TTCAGTGCAG TATCATCATT CTCCTTTTAC CCCCCGCTCA TATCTCATTC  
ATTTCTCTTA TTAAACAACCT TGTTCCCCCC TTCACCG

FIG. 3E



Ma524	EVRKLRTL FQSLGYYDSSKAYYAFKVSNLCIWGLSTVIIVAKWGQTSTLANVLSAALLGL	90
ATTS4723	- - - - - VTLY - TLA FVAAMSLGVLYGVLACPSVXPHQIAAGLLGL	38
12-5	- - - - - GVL YGVLACTSVFAHQIAAALLGL	24
T42806	- - - - G X X - - - - -	4
W28140	- - - - -	1
R05219	- - - - - C - - - - -	2
W53753	- - - - -	1

Ma524	FWQQCGWLAHDFLHHQVFQDRFWGDLFGAFLGGVC - QGFSSSWWKDKHNTHHAAPNVHGE	119
ATTS4723	LWIQSAYIGXD SGHYVIMSNKSNNX - FAQLLSGNC LTGI IAWWKWTHNAHHLACNSLDY	97
12-5	LWIQSAYIGHD SGHYVIMSNKSYNR - FAQLLSGNC LTGI SIAWWKWTHNAHHLACNSLDY	83
T42806	- - - - -	4
W28140	- - - - -	1
R05219	- - - - -	2
W53753	- - - - -	1

8/20

Ma524	DPDIDTHPLL TWSEHALEMFS DVPDEELTRMWS - - - - RFMVLNQTWFYFPILSFARLSW	174
ATTS4723	GP NLQH IP - - - - -	105
12-5	DPDLQH IPVFAVSTK - - - FFS SLTSRFYDRKLTFGPVARFLVSYQHFTYYPVMCFGRINL	140
T42806	- - - - -	4
W28140	- - - - -	1
R05219	- - - - -	2
W53753	- - - - -	1

FIG. 4A

Ma524	CLQSILFVLPNGQAHKPSGARVPISLVEQLSLAM	229
ATTS4723	W W	105
12-5	F I Q T F L L L F S K R E - - - - - V P D R A L N F A G I L V - - - - - F W T W F - - P L L V S C L P N W P E R F	185
T42806	- - - - - N F A G I L V - - - - - F F T V F - - P L L V S C L P N W P E R F	29
W28140	- - - - - P A T E V G G L A W M I T - Y - R F F L T Y V P L L G L K A F L G	33
R05219	- - - - - F - S - - - - -	2
W53753	- - - - - R H E A A R G G T R L A Y M L V C M Q W T D L - - L W A A S Y R F F L S Y S P F Y G A T G T L L	48
Ma524	Y F L V S Q A V C G N L L A I V F S L N H N G M P V I S K E E A V D M D F F T K Q I I T G R D V H P G L F A N W F T G G	289
ATTS4723		105
12-5	F F V F T S F T V T A L Q H I Q F T L N H F A A D V Y V - G P P T G S D W F E K Q A A G T I D I S C R S Y M D W F F G G	244
T42806	X F V F T G F T V T A L Q H I Q F T L N H F A A D V Y V - G P P T G S D W F E K Q A A G T I D I S C R S Y M D W F F G G	88
W28140	L F F I V R F L E S N W F V W V T Q M N H - - - I P M H I D H D R N M D W V S T Q L Q A T C N V H K S A F N I D W F S G H	90
R05219	- - - - - S P K S S P T R N M T P S P F I D W L W G G	23
W53753	L F V A V R V L E S H W F V W I T Q M N H - - - I P K E I G H E K H R D W A S S Q L A A T C N V E P S L F D W F S G H	105
Ma524	L N Y Q I E H H L F P S M P R H N F S K I Q P A V E T L C K K Y N V R Y H T T G M I E G T A E V E S R L N E V S K A A S	349
ATTS4723		105
12-5	L Q F Q L E H H	252
T42806	L Q F Q L E H H L F P R L P R C H L R K V S P V G Q R G F Q R K X N L S X	125
W28140	L N F Q I E H H L F P T M P R H N Y H X V A P L V Q S L C A K H G I E Y Q S K P L	131
R05219	L N Y Q I E H H L F P T M P R C N L N R C M K Y V K E W C A E N N L P Y L V D D Y F V G Y N L N L Q Q L K N M A E L V Q	83
W53753	L N F Q I E H H L F P T M P R H N Y R X V A P L V K A F C A K H G L H Y E V	143
Ma524	K M G K A Q	355
ATTS4723		105
12-5		252
T42806		125
W28140		131
R05219	- - A K A A	87
W53753		148

FIG. 4B

GTCCCCTGTC GCTGTCGGCA CACCCCATCC TCCCTCGCTC CCTCTGCGTT TGTCCCTTGGC																60
CCACCGTCTC TCCTCCACCC TCCGAGACGA CTGCAACTGT AATCAGGAAC CGACAAATAC																120
ACGATTTCTT TTTACTCAGC ACCAACTCAA AATCCTCAAC CGCAACCCTT TTTCAGG																180
GCA CCT CCC AAC ACT ATC GAT GCC GGT TTG ACC CAG CGT CAT ATC AGC																Met
Ala	Pro	Pro	Asn	Thr	Ile	Asp	Ala	Gly	Leu	Thr	Gln	Arg	His	Ile	Ser	
ACC TCG GCC CCA <sup>*</sup> AAC TCG GCC AAG CCT GCC TTC GAG CGC AAC TAC CAG																
Thr	Ser	Ala	Pro	Asn	Ser	Ala	Lys	Pro	Ala	Phe	Glu	Arg	Asn	Tyr	Gln	
CTC CCC GAG TTC ACC ATC AAG GAG <sup>*</sup> ATC CGA GAG TGC ATC CCT GCC CAC																
Leu	Pro	Glu	Phe	Thr	Ile	Lys	Glu	Ile	Arg	Glu	Cys	Ile	Pro	Ala	His	
TGC TTT GAG CGC TCC GGT CTC CGT GGT CTC TGC CAC <sup>*</sup> GTT GCC ATC GAT																
Cys	Phe	Glu	Arg	Ser	Gly	Leu	Arg	Gly	Leu	Cys	His	Val	Ala	Ile	Asp	
CTG ACT TGG GCG TCG CTC TTG TTC CTG GCT GCG ACC CAG ATC GAC AAG <sup>*</sup>																
Leu	Thr	Trp	Ala	Ser	Leu	Leu	Phe	Leu	Ala	Ala	Thr	Gln	Ile	Asp	Lys	
TTT GAG AAT CCC TTG ATC CGC TAT TTG GCC TGG CCT GTT TAC TGG ATC																
Phe	Glu	Asn	Pro	Leu	Ile	Arg	Tyr	Leu	Ala	Trp	Pro	Val	Tyr	Trp	Ile	

FIG. 5A

SUBSTITUTE SHEET (RULE 26)			480													
	ATG	CAG	GGT	ATT <sup>*</sup>	GTC	TGC	ACC	GGT	GTC	TGG	GTG	CTG	GCT	CAC	GAG	TGT
	Met	Gln	Gly	Ile	Val	Cys	Thr	Gly	Val	Trp	Val	Leu	Ala	His	Glu	Cys
								540								
	GGT	CAT	CAG	TCC <sup>*</sup>	TTC	TCG	ACC	TCC <sup>*</sup>	AAG	ACC	CTC	AAC	AAC	ACA	GTT	GGT
	Gly	His	Gln	Ser	Phe	Ser	Thr	Ser	Lys	Thr	Leu	Asn	Asn	Thr	Val	Gly
												600				
	TGG	ATC	TTG	CAC	TCG	ATG	CTC	TTG	GTC	CCC	TAC	CAC <sup>*</sup>	TCC	TGG	AGA	ATC
	Trp	Ile	Leu	His	Ser	Met	Leu	Leu	Val	Pro	Tyr	His	Ser	Trp	Arg	Ile
																660
	TCG	CAC	TCG	AAG	CAC	CAC	AAG	GCC	ACT	GGC	CAT	ATG	ACC	AAG	GAC	CAG <sup>*</sup>
	Ser	His	Ser	Lys	His	His	Lys	Ala	Thr	Gly	His	Met	Thr	Lys	Asp	Gln
	GTC	TTT	GTG	CCC	AAG	ACC	CGC	TCC	CAG	GTT	GGC	TTG	CCT	CCC	AAG	GAG
	Val	Phe	Val	Pro	Lys	Thr	Arg	Ser	Gln	Val	Gly	Leu	Pro	Pro	Lys	Glu
			720													
	AAC	GCT	GCT	GCT <sup>*</sup>	GCC	GTT	CAG	GAG	GAG	GAC	ATG	TCC	GTG	CAC	CTG	GAT
	Asn	Ala	Ala	Ala	Ala	Val	Gln	Glu	Glu	Asp	Met	Ser	Val	His	Leu	Asp
								780								
	GAG	GAG	GCT	CCC <sup>*</sup>	ATT	GTG	ACT	TTG <sup>*</sup>	TTC	TGG	ATG	GTG	ATC	CAG	TTC	TTG
	Glu	Glu	Ala	Pro	Ile	Val	Thr	Leu	Phe	Trp	Met	Val	Ile	Gln	Phe	Leu
												840				
	TTC	GGA	TGG	CCC <sup>*</sup>	GCG	TAC	CTG	ATT <sup>*</sup>	ATG	AAC	GCC	TCT <sup>*</sup>	GGC	CAA	GAC	TAC
	Phe	Gly	Trp	Pro	Ala	Tyr	Leu	Ile	Met	Asn	Ala	Ser	Gly	Gln	Asp	Tyr

FIG. 5B

GGC	CGC	TGG	ACC	TCG	CAC	TTC	CAC	ACG	TAC	TCG	CCC	ATC	TTT	GAG	CCC <sup>*</sup>
Gly	Arg	Trp	Thr	Ser	His	Phe	his	Thr	Tyr	Ser	Pro	Ile	Phe	Glu	Pro
CGC	AAC	TTT	TTC	GAC	ATT	ATT	ATC	TCG	GAC	CTC	GGT	GTG	TTG	GCT	GCC
Arg	Asn	Phe	Phe	Asp	Ile	Ile	Ile	Ser	Asp	Leu	Gly	Val	Leu	Ala	Ala
			960												
CTC	GGT	GCC	CTG <sup>*</sup>	ATC	TAT	GCC	TCC	ATG	CAG	TTG	TCG	CTC	TTG	ACC	GTC
Leu	Gly	Ala	Leu	Ile	Tyr	Ala	Ser	Met	Gln	Leu	Ser	Leu	Leu	Thr	Val
							1020								
ACC	AAG	TAC	TAT	ATT	GTC	CCC	TAC <sup>*</sup>	CTC	TTT	GTC	AAC	TTT	TGG	TTG	GTC
Thr	Lys	Tyr	Tyr	Ile	Val	Pro	Tyr	Leu	Phe	Val	Asn	Phe	Trp	Lru	Val
											1080				
CTG	ATC	ACC	TTC	TTG	CAG	CAC	ACC	GAT	CCC	AAG	CTG <sup>*</sup>	CCC	CAT	TAC	CGC
Leu	Ile	Thr	Phe	Leu	Gln	His	Thr	Asp	Pro	Lys	Leu	Pro	His	Tyr	Arg
															1140
GAG	GGT	GCC	TGG	AAT	TTC	CAG	CGT	GGA	GCT	CTT	TGC	ACC	GTT	GAC	CGC <sup>*</sup>
Glu	Gly	Ala	Trp	Asn	Phe	Gln	Arg	Gly	Ala	Leu	Cys	Thr	Val	Asp	Arg
TCG	TTT	GGC	AAG	TTC	TTG	GAC	CAT	ATG	TTC	CAC	GGC	ATT	GTC	CAC	ACC
Ser	Phe	Gly	Lys	Phe	Leu	Asp	His	Met	Phe	His	Gly	Ile	Val	His	Thr
			1200												
CAT	GTG	GCC	CAT <sup>*</sup>	CAC	TTG	TTC	TCG	CAA	ATG	CCG	TTC	TAC	CAT	GCT	GAG
His	Val	Ala	His	His	Leu	Phe	Ser	Gln	Met	Pro	Phe	Tyr	His	Ala	Glu

FIG. 5C

1260  
\*  
GAA GCT ACC TAT CAT CTC AAG AAA CTG CTG GGA GAG TAC TAT GTG TAC  
Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr Tyr Val Tyr

1320  
\*  
GAC CCA TCC CCG ATC GTG GTT GCG GTC TGG AGG TCG TTC CGT GAG TGC  
Asp Pro Ser Pro Ile Val Val Ala Val Trp Arg Ser Phe Arg Glu Cys

1380  
\*  
CGA TTC GTG GAG GAT CAG GGA GAC GTG GTC TTT TTC AAG AAG TAAAAA  
Arg Phe Val Glu Asp Gln Gly Asp Val Val Phe Phe Lys Lys

1440  
\*  
AAAAGACAAT GGACCACACA CAACCTTGTC TCTACAGACC TACGTATCAT GTAGCCATAC  
CACTTCATAA AAGAACATGA GCTCTAGAGG CGTGTCATTTC GCGCCTCC

13/20

FIG. 5D

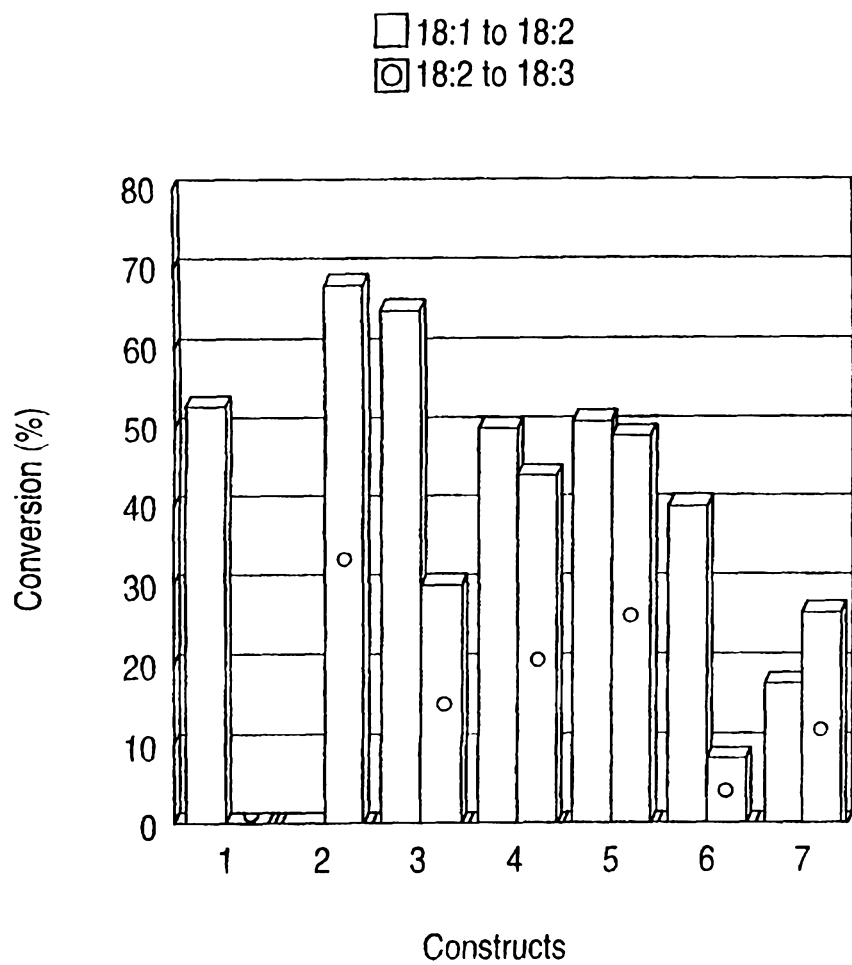


FIG. 6A

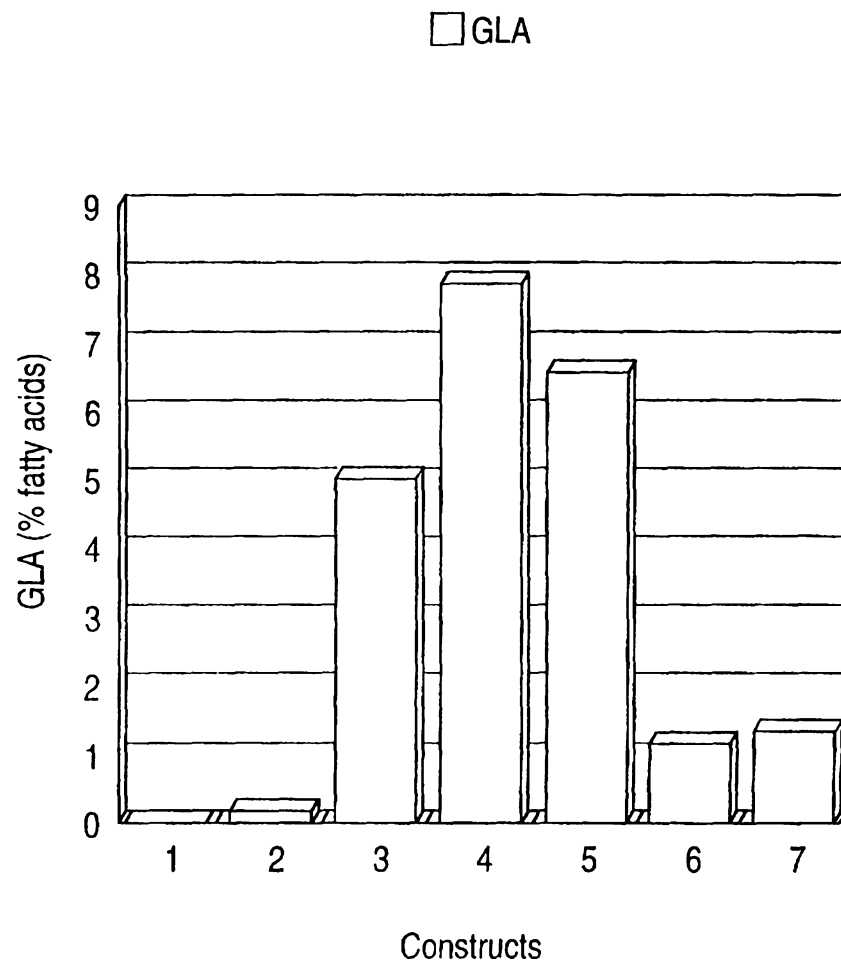


FIG. 6B

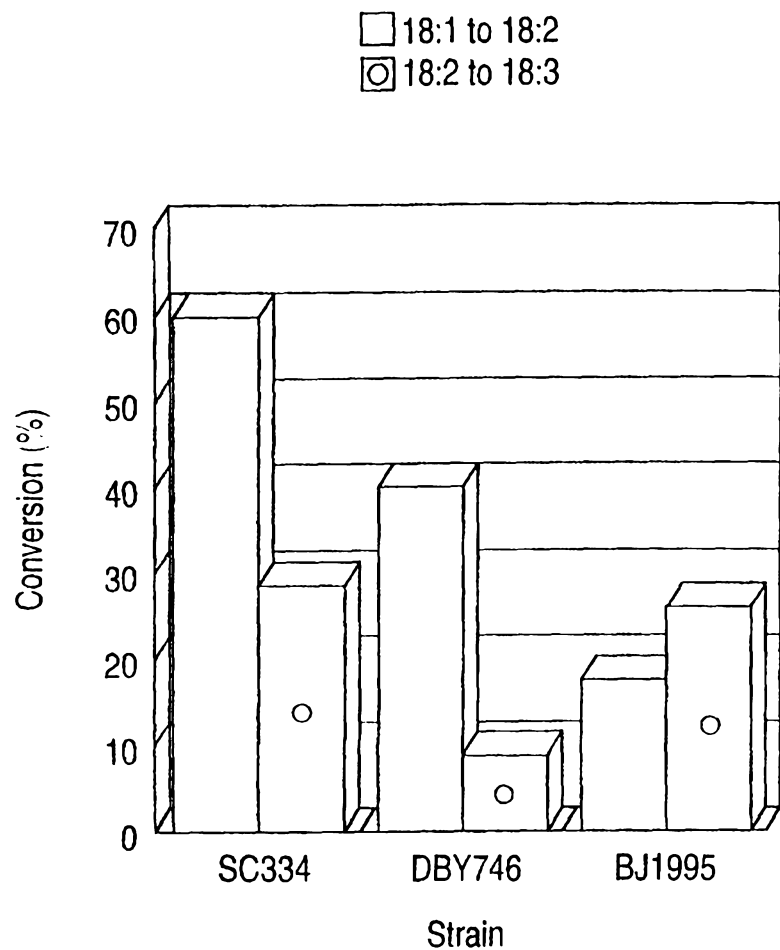


FIG. 7A

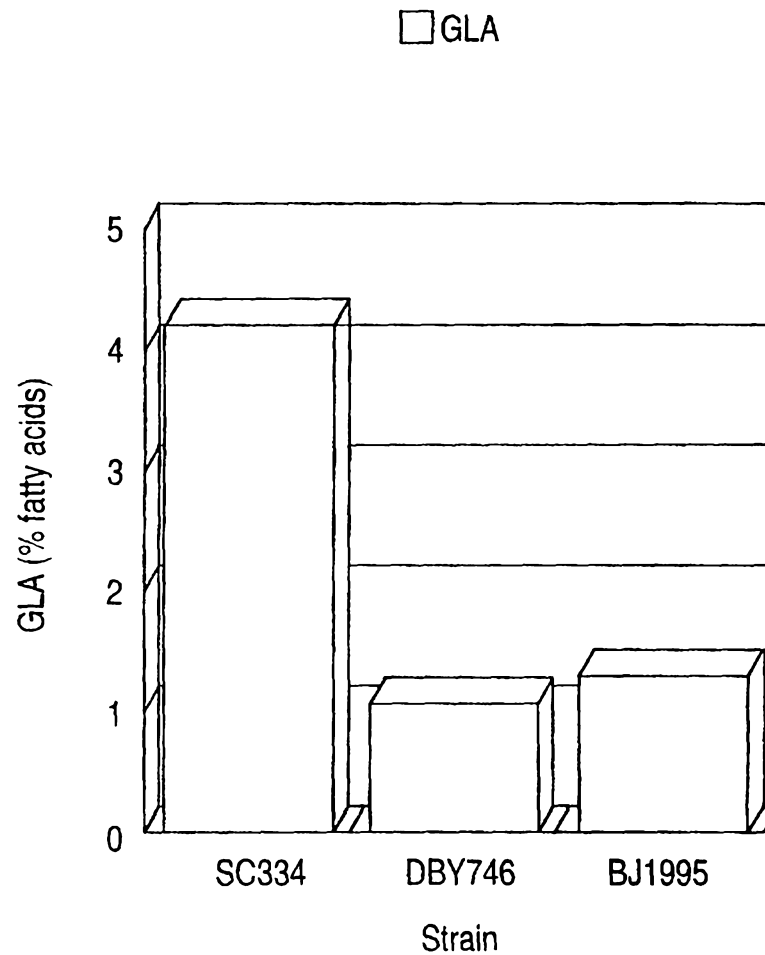


FIG. 7B



16/20

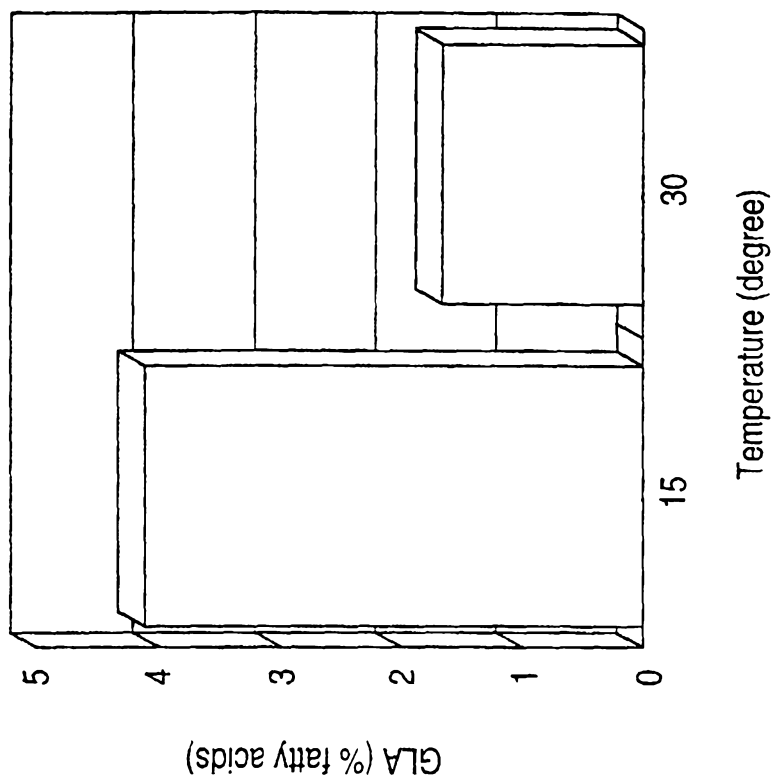


FIG. 8B

□ 18:1 to 18:2  
 ◻ 18:2 to 18:3

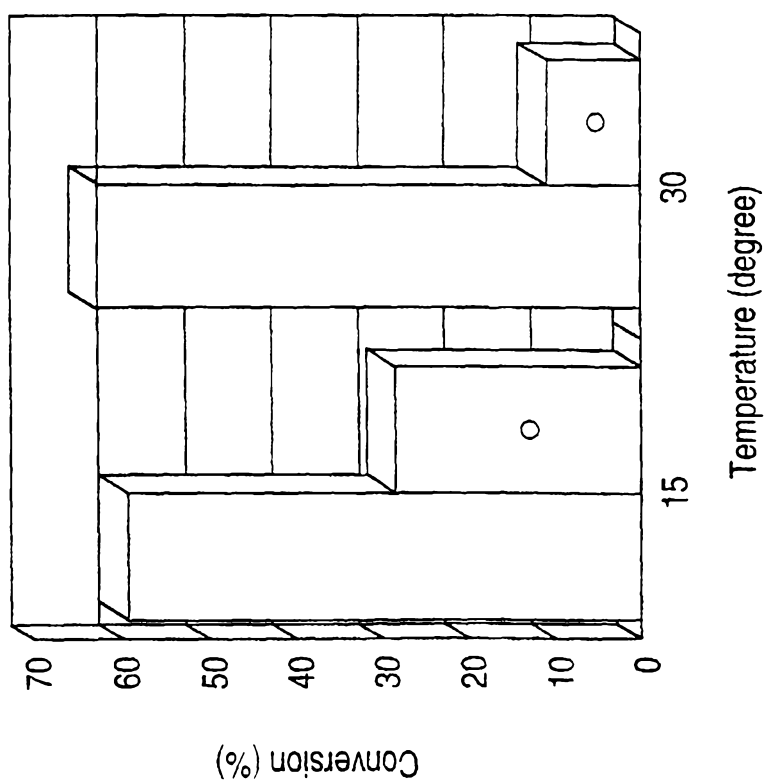


FIG. 8A

SCORES INIT1: 117 INITN: 225 OPT: 256  
SMITH-WATERMAN SCORE: 408; 27.0% IDENTITY IN 441 aa OVERLAP

SUBSTITUTE SHEET (RULE 26)

		10	20	30	40	50	
ma29gcg.pep	MGTDQGKT - - -	FTWEE	LAAHNTKDDLLA	IRGRVYDVT	KFLSRHPGGVD	TLLLGAGR	DVT
		:	: : :   :	:     : : :		: :	:
253538a		QGPTPRYFTWDE	VAQRSGCEERWL	VDRKVYNISE	FTRRHPGGSRV	ISHYAGQ	DAT
		10	20	30	40	50	
	60	70	80	90	100	110	
ma29gcg.pep	PVFEMYHAF -	GAADA	IMKKYYVGTL	VSNELP	IFPEPTVFHKT	IKTRVEGYFT	DRNIDPKN
	:	:	:		: :	: :	:
253538a	DPFVAFHINK	GLVKKYMNSLL	IGEL -	SPEQPSF -	EPTKNKELT	DEFREL	RATVERMGLMK
	60	70	80	90	100	110	
	120	130	140	150	160	170	
ma29gcg.pep	RPEIWGRYAL	IFGSLIASYYA	QLFVPFVVERT	WLQVVF -	AIIMGFACAQ	VGLNPLHD	ASH
	: : :	: : :	: : :	: :	: :   : : :	:	:
253538a	ANHVF - -	FLLYLLHILL	LDGAAWLT	LWVFGTSFL	PFLLCVLLSA	VQAQAGWLQ	- HDYGH
	120	130	140	150	160	170	
	180	190	200	210	220		
ma29gcg.pep	FSVTHNPTVWK	ILGATHDF - - -	FNGASYLVW	MYQHMLGHH	PYTNIAGAD	PDVSTSE - - -	
	:     : :	:   :	: :	: :   :		: :	:
253538a	LSVYRKPK -	WNHL - -	VHKFVIGHL	KGASANW	WNHRH -	FQHHAKPNI	FHKDPDVNMLHVFV
	180	190	200	210	220		

FIG. 9A

**WO 98/46763**

ma29gcg.pep      400                  410                  420                  430                  440  
QHHYPD I L A I I K N T C S E Y K V P Y L V K D T F W Q A F A S H L E H L R V L G L R P K E E X  
: | :        | : : | : | : : : |  
253538a        R H N L H K I A P L V K S L C A K H G I E Y Q E K P L L R A L L D I I R S L K K S G K L W L D A Y L H K X  
380                  390                  400                  410                  420                  430

18/20

PCT/US98/07126

**FIG. 9B**

SCORES INIT1: 231 INITN: 499 OPT: 401  
SMITH-WATERMAN SCORE: 620; 27.3% IDENTITY IN 455 aa OVERLAP

SUBSTITUTE SHEET (RULE 26)

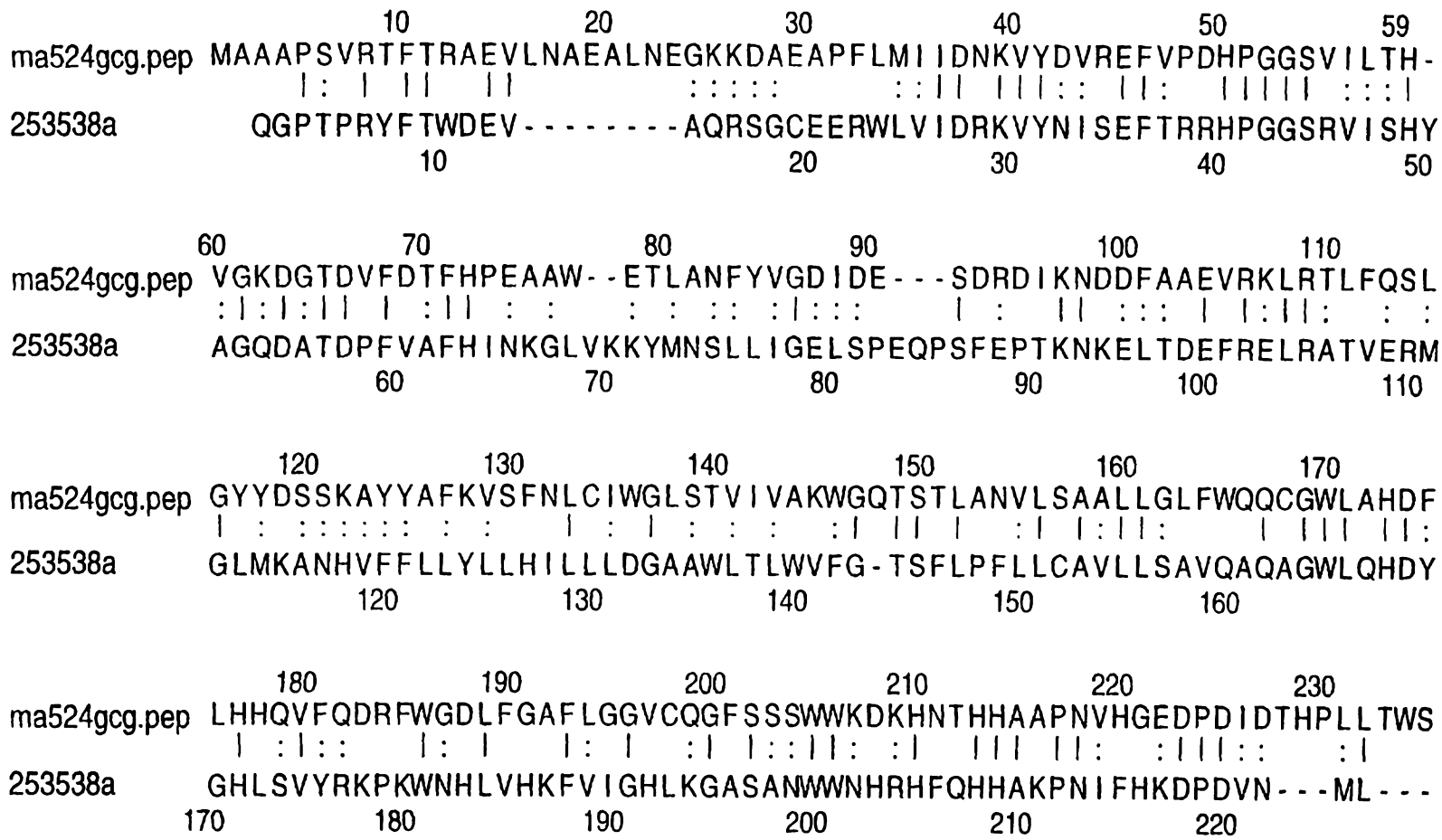


FIG. 10A

20/20

SCORES INIT1: 231 INITN: 499 OPT: 401  
SMITH-WATERMAN SCORE: 620; 27.3% IDENTITY IN 455 aa OVERLAP

SUBSTITUTE SHEET (RULE 26)

ma524gcg.pep	240	250	260	270	280	290
	EH	AL	EM	FS	DV	PD
	EL	TR	MS	RF	MV	LN
	QT	WF	YF	PI	LS	- - -
	FAR	LS	WC	LQ	SI	LF
	VL	PN	QG	AH		
253538a	-H	VF	-V	LG	EW	QP
	IE	YG	KK	KL	KY	LP
	YN	HQ	HE	YF	FL	IG
	PP	LL	IP	MY	FQ	YQ
	IM	TM	I	- - -	-VH	
	230	240	250	260	270	
ma524gcg.pep	300	310	320	330	340	349
	KP	SG	AR	VP	IS	LV
	EQ	LS	LA	MH	WT	WY
	LA	TM	FL	IK	- -	DP
	VN	ML	VY	FL	VS	QA
	VC	GN	LL	AI	VF	FS
253538a	K	- - - - -	NW	VD	LA	WA
	VS	YI	RF	FI	TY	IP
	FG	IL	GA	LL	FL	NF
	IR	FL	ES	HW	FV	WT
	Q					
	280	290	300	310	320	
ma524gcg.pep	350	360	370	380	390	400
	LN	HG	MP	VI	SK	EE
	AV	DM	DF	FT	KQ	II
	TG	RD	VH	PG	LF	AN
	WF	TG	GL	NY	QI	EH
	HL	FP	SM	PR	HN	LF
	HN					
253538a	MN	HI	VM	EI	- -	DQ
	EAY	R	-D	WF	SS	QL
	TA	TN	VE	QS	FF	ND
	WF	SG	HL	NF	QI	EH
	HL	FP	TM	PR	HN	LF
	330	340	350	360	370	380
ma524gcg.pep	410	420	430	440	450	
	SK	IQ	PA	VE	TL	CK
	KY	NV	RY	HT	TG	MI
	EG	TA	EV	FS	RL	NE
	VS	KA	AS	KM	GK	QA
	QX					
253538a	HK	IA	PL	VK	SL	CA
	KH	GI	EY	QE	KP	LL
	RA	LD	IR	SL	KK	SG
	KL	WL	DA	YL	HK	X
	390	400	410	420	430	

FIG. 10B