PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/54, 9/10, 15/63, 15/82, C12P 7/64, C12N 1/28, C11B 11/00

A2

(11) International Publication Number: **WO 00/32793**

(43) International Publication Date: 8 June 2000 (08.06.00)

(21) International Application Number: PCT/US99/28825

(22) International Filing Date: 3 December 1999 (03.12.99)

(30) Priority Data:
09/205,271 4 December 1998 (04.12.98) US

(71) Applicant *(for all designated States except US)*: CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).

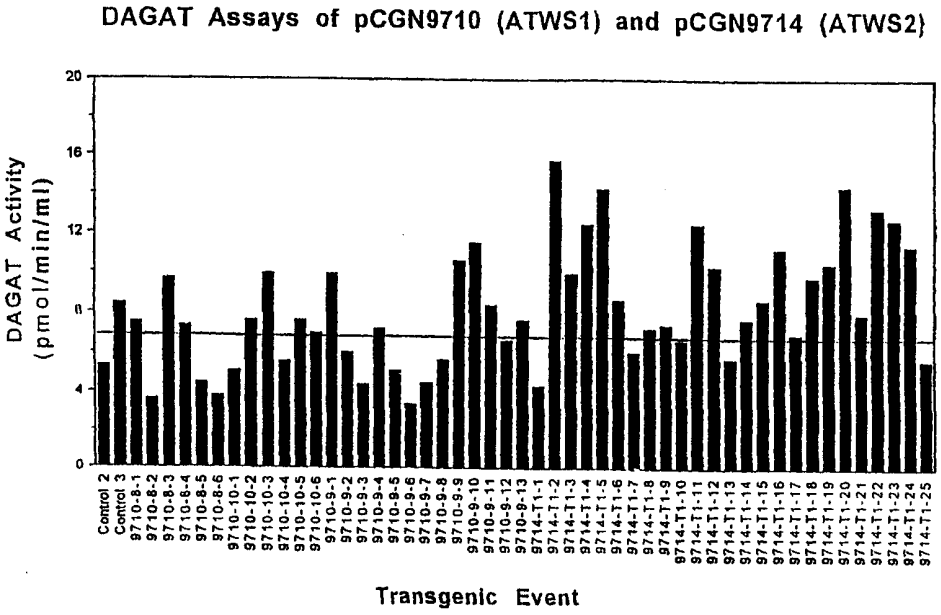
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: METZ, James, George [–/US]; 2803 Belhaven Place, Davis, CA 95616 (US). LARDIZABAL, Kathryn, Dennis [–/US]; 1546 Owens Valley Road, Woodland, CA 95616 (US).

(74) Agent: SCHWEDLER, Carl, J.; Patent Department Central, Monsanto/G.D. Searle, P.O. Box 5110, Chicago, IL 60680–5110 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, 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, SL, SZ, TZ, 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, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: DIACYLGLYCEROL ACYL TRANSFERASE PROTEINS



(57) Abstract

By this invention, nucleic acid sequences encoding for enzymes with diacylglycerol acyltransferase (DAGAT) activity are provided, wherein said DAGAT is active in the formation of triacylglycerol from fatty acyl-CoA and *sn*1,2 diacylglycerol substrates. Of special interest are nucleic acid sequences from jojoba which encode a protein capable of producing wax esters from fatty alcohol and fatty acyl substrates, as well as producing triacylglycerol from *sn*1,2 diacylglycerol and fatty acid substrates. Also of interest are sequences related to the jojoba embryo wax synthase from *Arabidopsis*. Also considered are amino acid and nucleic acid sequences endocing DAGAT proteins and the use of such sequences to provide transgenic host cells capable of producing modified triacylglycerol compositions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bułgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Diacylglycerol Acyl Transferase Proteins

This application is a continuation in part of U.S. Pat. Application No. 09/092,583, and a continuation in part of PCT/US98/11575, which is a continuation in part of U.S. Pat. Application No. 60/048,625.

Technical Field

The present invention is directed to enzymes, methods to purify, and obtain such enzymes, amino acid and nucleic acid sequences related thereto, and methods of use for such compositions in genetic engineering applications.

INTRODUCTION

Background

Plant oils are used in a variety of industrial and edible uses. Novel vegetable oils compositions and/or improved means to obtain oils compositions, from biosynthetic or natural plant sources, are needed. Depending upon the intended oil use, various different fatty acid compositions are desired.

For example, in some instances having an oilseed with a higher ratio of oil to seed meal would be useful to obtain a desired oil at lower cost. This would be typical of a high value oil product. Or such an oilseed might constitute a superior feed for animals. In some instances having an oilseed with a lower ratio of oil to seed meal would be useful to lower caloric content. In other uses, edible plant oils with a higher percentage of unsaturated fatty acids are desired for cardiovascular health reasons. And alternatively, temperate substitutes for high saturate tropical oils such as palm, coconut, or cocoa would also find uses in a variety of industrial and food applications.

One means postulated to obtain such oils and/or modified fatty acid compositions is through the genetic engineering of plants. However, in order to genetically engineer plants one must have in place the means to transfer genetic material to

the plant in a stable and heritable manner. Additionally, one must have nucleic acid sequences capable of producing the desired phenotypic result, regulatory regions capable of directing the correct application of such sequences, and the like. Moreover, it should be appreciated that in order to produce a desired phenotype requires that the so-called Kennedy Pathway for glycerolipid synthesis is modified to the extent that the ratios of reactants and metabolic flux through the pathway are modulated or changed.

Higher plants appear to synthesize oils via a common metabolic pathway. Fatty acids are made in plastids from acetyl-CoA through a series of reactions catalyzed by enzymes known collectively as Fatty Acid Synthetase (FAS). The fatty acids produced in plastids are exported to the cytosolic compartment of the cell, and are esterified to coenzyme A. These acyl-CoAs are the substrates for glycerolipid synthesis in the endoplasmic reticulum (ER). Glycerolipid synthesis itself is a series of reactions leading first to phosphatidic acid (PA) and diacylglycerol (DAG). Either of these metabolic intermediates may be directed to membrane phospholipids such as phosphatidylglycerol (PG), phosphatidylethanolamine (PE) or phosphatidylcholine (PC), or they may be directed on to form neutral triacylglycerol (TAG) the primary component of vegetable oil used by the seed as a stored form of energy to be used during seed germination.

Diacylglycerol (DAG) is synthesized from glycerol-3-phosphate and fatty acyl-CoAs in three steps catalyzed sequentially by glycerol-3-phosphate acyltransferase (G3PAT), lysophosphatidic acid acyltransferase (LPAAT) to make PA, and a further hydrolytic step catalyzed by phosphatidic acid phosphatase (PAP) to make DAG. In most cells, DAG is used to make membrane phospholipids, the first step being the synthesis of PC catalyzed by CTP-phosphocholine cytidylyltransferase. In cells producing storage oils, DAG is acylated with a third fatty acid in a reaction catalyzed by diacylglycerol acyltransferase (DAGAT). Collectively, the reactions make up part of what is commonly referred to as the Kennedy Pathway.

The structure of the TAG, as far as positional specificity of fatty acids, is determined by the specificity of each of the three acyltransferases for the fatty acyl CoA and the glycerol backbone substrates. Thus, for example, there is a tendency for the acyltransferases from many temperate zone species of seeds to allow either a saturated or an unsaturated fatty acid at the sn-1 or the sn-3 position, but only an unsaturated fatty acid at the sn-2. The absolute specificity for an unsaturated fatty acid at sn-2 is determined by the substrate preference of LPAAT enzyme. In some species such as cocoa, TAG compositions suggest that this tendency is carried further in that there is an apparent preference for acylation of the sn-3 position with a saturated fatty acid, if the sn-1 position is esterified to a saturated fatty acid. Thus, there is a higher percentage of structured TAG of the form SUS (where S = saturated fatty acid and U = unsaturated fatty acid), than would be expected from a random distribution based on the overall fatty acid composition with the sn-2 position fixed with an unsaturated fatty acid. This suggests that DAGAT plays an important role in the regulation of TAG structure, as well as in the control of TAG synthesis.

The reaction catalyzed by DAGAT is at a critical branchpoint in glycerolipid biosynthesis. Enzymes at such branchpoints are considered prime candidates for sites of metabolic regulation. Up through the synthesis of diacylglycerol, TAG and membrane lipid synthesis share in common G3PAT, LPAAT, and PAP. Since all cells have membranes, they must have these enzymes. What makes oil synthesis unique is the DAGAT reaction. The presence of DAGAT activity provides an alternative fate for DAG than going into membranes. It is logical to think that what drives the synthesis of TAG is the presence of DAGAT enzyme, and that either directly or indirectly through a regulatory cascade, DAGAT activity and/or diacylglycerol concentrations, plays a role in controlling flux into glycerolipids.

Obtaining nucleic acid sequences capable of producing a phenotypic result in the incorporation of fatty acids into a glycerol backbone to produce an oil is subject to various

obstacles including but not limited to the identification of metabolic factors of interest, choice and characterization of a protein source with useful kinetic properties, purification of the protein of interest to a level which will allow for its amino acid sequencing, utilizing amino acid sequence data to obtain a nucleic acid sequence capable of use as a probe to retrieve the desired DNA sequence, and the preparation of constructs, transformation and analysis of the resulting plants.

Thus, the identification of enzyme targets and useful tissue sources for nucleic acid sequences of such enzyme targets capable of modifying oil structure and quantity are needed. Ideally an enzyme target will be amenable to one or more applications alone or in combination with other nucleic acid sequences relating to increased/decreased oil production, TAG structure, the ratio of saturated to unsaturated fatty acids in the fatty acid pool, and/or to other novel oils compositions as a result of the modifications to the fatty acid pool. Once enzyme targets(s) are identified and qualified, quantities of protein and purification protocols are needed for sequencing. Ultimately, useful nucleic acid constructs having the necessary elements to provide a phenotypic modification and plants containing such constructs are needed.

Several putative isolation procedures have been published for DAGAT. Polokoff and Bell (1980) reported solubilization and partial purification of DAGAT from rat liver microsomes. This preparation was insufficiently pure to identify a specific protein factor responsible for the activity. Kwanyuen and Wilson (1986, 1990) reported purification and characterization of the enzyme from soybean cotyledons. However, the molecular mass (1843 kDa) suggests that this preparation was not extensively solubilized and any DAGAT protein contained therein was part of a large aggregate of many proteins. Little et al (1993) reported solubilization of DAGAT from microspore-derived embryos from rapeseed, but as with Kwanyuen and Wilson, the molecular mass of the material that was associated with activity was so high, that complete

solubilization is unlikely. Andersson et al (1994) reported solubilization and a 415-fold purification of DAGAT from rat liver using immunoaffinity chromatography. However, there is no evidence that the antibodies they used recognize DAGAT epitopes, nor that the protein that they purified is truly DAGAT. Indeed, as with Kwanyuen and Wilson, the DAGAT activity in their preparations exhibited a molecular mass typical of aggregated membrane proteins. Finally, Kamisaka et al (1993, 1994, 1996, 1997) report solubilization of DAGAT from *Mortierella ramanniana* and subsequent purification to homogeniety. They show evidence that DAGAT from this fungal species has a molecular mass of 53 kDa, which is the first public report of a DAGAT that may actually have been solubilized. We attempted to reproduce their work in our laboratory (see Example 4), but were unable to obtain a homogeneous preparation, or to associate the enzyme activity with a 53 kDa polypeptide. Indeed, we were able to show that an abundant 53-kDa polypeptide likely to be the one claimed by Kamisaka et al, does not correlate with DAGAT activity in fractions obtained using their protocol.

Relevant Literature

Cell-free homogenates from developing jojoba embryos were reported to have acyl-CoA fatty alcohol acyl transferase activity. The activity was associated with a floating wax pad which formed upon differential centrifugation (Pollard *et al.* (1979) *supra*; Wu *et al.* (1981) *supra*).

Solubilization of a multienzyme complex from *Euglena gracilis* having fatty acyl-SCoA transacylase activity is reported by Wildner and Hallick (Abstract from *The Southwest Consortium Fifth Annual Meeting*, April 22-24, 1990, Las Cruces, NM.).

Ten-fold purification of jojoba acyl-CoA: alcohol transacylase protein is reported by Pushnik *et al.* (Abstract from *The Southwest Consortium Fourth Annual Meeting*, February 7, 1989, Riverside, Ca.).

An assay for jojoba acyl-CoA:alcohol transacylase activity was reported by Garver *et al.* (*Analytical Biochemistry* (1992) *207*:335-340).

WO 93/10241 is directed to plant fatty acyl-CoA:fatty alcohol 0-acyltransferases. A jojoba 57kD protein is identified as the jojoba fatty acyl-CoA:fatty alcohol 0-acyltransferase (wax synthase). The present inventors later reported that the 57kD protein from jojoba is a ß-ketoacyl-CoA synthase involved in the biosynthesis of very long chain fatty acids (Lassner *et al.* (*The Plant Cell* (1996) *8*:281-292).

Photoaffinity labeling of a 57 kD jojoba seed polypeptide postulated to be an acyl-CoA:fatty alcohol acyltransferase was also reported by Shockey *et al.* (*Plant Phys.* (1995) *107*:155-160).

Kamisaka and Nakahara, "Characterization of the Diacylglycerol Acyltransferase Activity in the Lipid Body Fraction from an Oleaginous Fungus", J. Biochem. (1994) 116:1295-1301.

Kamisaka and Nakahara, "Activation of Detergent-Solubilized Diacylglycerol Acyltransferase by Anionic Phospholipids", J. Biochem. (1996) 119:520-523.

Kamisaka et al., "Purification and Characterization of Diacylglycerol Acyltransferase Activity from the Lipid Body Fraction from an Oleaginous Fungus", J. Biochem. (1997) 121:1107-1114.

SUMMARY OF THE INVENTION

By this invention, compositions and methods of use related to diacylglycerol acyltransferase, hereinafter also referred to as DAGAT, are provided. Also of interest are methods and compositions of amino acid sequences, as well as nucleic acid sequences encoding such amino acid sequences, related to biologically active DAGAT(s).

In particular, DAGAT protein preparations which have relatively high specific activity are of interest for use in a variety of applications, in vitro and in vivo. Especially, protein preparations having DAGAT activities are contemplated

hereunder. Of special interest are the DAGATs obtainable from *Mortierella ramanniana* and *Arabidopsis*.

Also of particular interest, is the discovery that the jojoba wax synthase of the present invention also demonstrates diacylglycerol (DAGAT) activity. TAG is not naturally produced in jojoba and thus the activity of the wax synthase enzyme with DAG substrates suggests that the wax synthase is related to DAGAT, an enzyme responsible for production of TAG in most plant species, particularly in oilseed crop plants whose seeds contain high levels of storage TAG. Thus, the use of the jojoba wax synthase protein and/or its encoding sequence for isolation of plant genes encoding DAGAT is considered in the present invention.

The exemplified jojoba DAGAT is purified away from the membrane (*i.e.* solubilized), and the solubilized DAGAT preparation is subjected to various chromatographic analyses to identify a protein associated with the DAGAT activity. In this manner a protein having a molecular weight of approximately 33 kDA based on SDS-PAGE_analysis is identified as associated with DAGAT activity. Further purification methods, such as column chromatography and polyacrylamide gel electrophoresis are utilized to obtain the DAGAT protein in sufficient purity for amino acid sequence analysis.

Peptide fragments from the DAGAT proteins are used as a template in designing various synthetic oligonucleotides which are used to obtain nucleic acid sequences encoding all or a portion of the DAGAT protein. Using the DAGAT encoding sequences so obtained, it is also possible to isolate other DAGAT genes which encode DAGAT proteins. As demonstrated herein, nucleic acid sequences coding for DAGAT are obtained from *Arabidopsis*.

Thus, this invention encompasses DAGAT peptides and the corresponding amino acid sequences of those peptides. Such sequences find particular use in the preparation of oligonucleotides containing DAGAT encoding sequences for analysis and recovery of DAGAT gene sequences. The DAGAT encoding sequence may encode a complete or partial sequence

depending upon the intended use. All or a portion of the genomic sequence, or cDNA sequence, is intended.

Of special interest are recombinant DNA constructs which provide for transcription or transcription and translation (expression) of the DAGAT sequences. In particular, constructs which are capable of transcription or transcription and translation in plant host cells are preferred. Such constructs may contain a variety of regulatory regions including transcriptional initiation regions obtained from genes preferentially expressed in plant seed tissue.

In yet a different aspect, this invention relates to a method for producing a DAGAT in a host cell or progeny thereof via the expression of a construct in the cell. Cells containing a DAGAT as a result of the production of the DAGAT encoding sequence are also contemplated herein.

In addition, this invention relates to methods of using DNA sequences encoding DAGAT for the modification of the ratio of oil (TAG) to other constituents as well as the composition and structure of the triglyceride molecules, especially in the seed oil of plant oilseed crops. Plant cells having such a modified triglyceride are also contemplated herein.

Also considered in this invention are the modified plants, seeds and oils obtained by expression of the plant DAGAT proteins of this invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 presents results of analysis of jojoba wax synthase activity in column fractions from a first wax synthase purification protocol. Figure 1A provides results of Blue A agarose chromatography. Figure 1B provides results of ceramic hydroxyapatite chromatography. Figure 1C provides results of sephracryl S-100 size exclusion chromatography. Figure 1D provides results of hydroxyapatite chromatography.

Figure 2 presents results of analysis of jojoba wax synthase activity in column fractions from a second wax synthase purification protocol. Figure 2A provides results of Blue A agarose chromatography. Figure 2B provides results of

hydroxyapatite chromatography. Figure 2C provides results of Superdex 75 size exclusion chromatography.

Figure 3 presents results of jojoba wax synthase and DAGAT activity in fractions from a purified wax synthase preparation according to the wax synthase purification represented in Figure 1. Results are from the fractions obtained following the hydroxyapatite chromatography step.

Figure 4 presents results of analysis of DAGAT activity in column fractions from a *Mortierella ramanniana* DAGAT purification protocol utilizing Yellow 86-Agarose chromatography.

Figure 5A presents results of further chromatography of *Mortierella ramanniana* DAGAT activity in column fractions from Heparin-Sepharose CL-6B chromatography of pooled DAGAT-containing fractions from Yellow 86-Agarose chromatography. Figure 5B provides results of SDS-PAGE analysis of the peak fractions.

Figure 6A presents results of further chromatography of *Mortierella ramanniana* DAGAT on Yellow 86-Agarose using gradient elution. The sample applied was pooled fractions from the late-eluting DAGAT peak from Heparin-Sepharose CL-6B. Figure 6B presents the SDS-PAGE analysis of the peak fractions.

Figure 7 presents results of analysis of DAGAT activity in column fractions from a second *Mortierella ramanniana* DAGAT purification protocol utilizing Yellow 86-Agarose chromatography.

Figure 8A presents results of analysis of *Mortierella ramanniana* DAGAT activity in column fractions from Hydroxylapatite chromatography of DAGAT fractions pooled from a Yellow 86-Agarose column. Figure 8B provides results of SDS-PAGE analysis of the peak fractions.

Figure 9 presents results of analysis of DAGAT activity in column fractions from a *Mortierella ramanniana* DAGAT purification protocol. Figure 9A provides results of tandem Yellow 86-Agarose/Hydroxylapatite chromatography. Figure 9B provides results of SDS-PAGE analysis of the peak fractions from the tandem chromotography.

Figure 10 provides the nucleic acid sequence for jojoba wax synthase.

Figure 11 provides the amino acid sequence derived from the nucleic acid sequence of jojoba wax synthase.

Figure 12 provides the nucleic acid sequence of AT_WS1.

Figure 13 provides the nucleic acid sequence of AT_WS2.

Figure 14 provides the nucleic acid sequence of AT_WS3.

Figure 15 provides the nucleic acid sequence of AT_WS4.

Figure 16 provides the nucleic acid sequence of AT_WS5.

Figure 17 provides the nucleic acid sequence of AT_WS6.

Figure 18 provides the nucleic acid sequence of AT_WS7.

Figure 19 provides an alignment between the amino acid sequences of the jojoba wax synthase, and *Arabidopsis* AT-WS1, AT_WS2, AT_WS3, AT_WS4, AT_WS5, AT_WS6, and AT_WS7.

Figure 20 provides a dendogram of the relationships between the amino acid sequences compared in Figure 20.

Figure 21 provides the results of DAGAT assays on leaves of *Arabidopsis* plants expressing the coding sequence of ATWS1 (pCGN9710) and ATWS2 (pCGN9714).

DETAILED DESCRIPTION OF THE INVENTION

A diacylglycerol acyltransferase (referred to herein as DAGAT) of this invention includes nucleic acid sequences coding sequences of amino acids, such as a protein, polypeptide or peptide, obtainable from a cell source, which demonstrates the ability to catalyze the production of triacylglycerol from 1,2-diacylglycerol and an acyl-CoA substrate under enzyme reactive conditions. By "enzyme reactive conditions" is meant that any necessary conditions are available in an environment (i.e., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

"Solubilization" refers to extraction of the DAGAT enzyme from the membranes in such a way that it then behaves in a manner typical of enzymes that are not membrane-associated. Because the membrane effectively links the DAGAT protein to other proteins which are also present therein, solubilization is an essential requirement for identification and

purification of the DAGAT protein as described in the following examples. In testing for solubilization of DAGAT activity, three different indications of solubilization, as described in more detail in the following examples, are considered.

1) DAGAT activity is not sedimented by very high-speed centrifugation.
2) DAGAT activity migrates on a size-exclusion chromatography column as though it had a native molecular weight typical of enzymes which are not membrane-associated.
3) Proteins present in the DAGAT preparation are at least partially separable from each other by column chromatography.

Because of potential alternative interpretations that may apply to any of the above criteria individually, it is necessary to confirm that all three of the criteria have been satisfied to confirm DAGAT solubilization. For example, the first criterion, of failure to sediment at very high *g* forces could be misleading if the density of the solution used for solubilization is similar to that of the unsolubilized membranes so that they sediment only very slowly. The second criterion, in which solubilized activity migrates more slowly through a size-exclusion column than the original membranes, may be compromised if the membranes themselves bind weakly to the column after exposure to detergent so that their migration through it is slowed. The third criterion, in which the solubilized proteins are chromatographically resolvable, is the least likely to be compromised by artifacts or unforeseen situations. However, it is possible that membranes could be partially dissociated by the solubilization procedure such that various aggregates of proteins are released. Such aggregates might then be resolved from each other chromatographically. Thus, satisfaction of all three criteria is necessary to assure that DAGAT solubilization is achieved.

Having obtained solubilized DAGAT protein from jojoba, it can be seen that further experiments to characterize the enzyme as to substrate specificity, cofactor requirements and possible activity inhibiting agents may now be conducted. For example, it has been found that the jojoba DAGAT of this invention has a broad range of acyl substrates, including acyl-ACP and acyl-CoA molecules. In addition, the acyl and fatty alcohol substrates may have a broad size range with respect to carbon chain length. For example, activity was tested using substrates having carbon chain lengths of from C8 to C24, and all were shown to be utilized by the enzyme. In addition, activity was shown with fatty acyl and fatty alcohols having varying degrees of unsaturation.

Surprisingly, the purified jojoba wax synthase is also shown herein to have activity with diacylglycerol (DAG) and fatty acyl-CoA substrates to produce triacylglycerol (TAG), even though TAG have not been reported to exist in jojoba plant tissues. Thus, the wax synthase has at least two acyltransferase activities, one in which the acceptor substrate for the acyl-CoA molecule is an alcohol (fatty alcohol acyltransferase) and another in which the acceptor subsrate is a diacylglycerol (DAG acyltransferase, or DAGAT). The presence of the DAGAT activity of the wax synthase enzyme suggests that wax synthase protein is closely related to DAGAT proteins in other plant species.

Solubilization of Jojoba and *Mortierella ramanniana* DAGATs are described in the following examples. Solubilization of DAGAT is confirmed by demonstration of each of the above criteria of solubilization.

Solubilized preparations of *Mortierella ramanniana* DAGAT are utilized in a variety of chromatographic experiments for identification and partial purification of the DAGAT protein. In this manner, a protein having a molecular weight of approximately 33 kDa is identified as associated with DAGAT activity. As described in more detail in the following examples, the 33 kDa protein is partially purified by chromatography on Yellow 86-Agarose and hydroxyapatite columns. The protein is then obtained in substantially

purified form by gel electrophoresis and, if desired, by blotting of the partially purified DAGAT protein to polyvinylidine difluoride (PVDF) or nitrocellulose membranes. The 33 kDA protein is recovered by cutting out that portion of the gel or membrane containing the identified band.

The purified protein is then digested with various enzymes to generate peptides for use in determination of amino acid sequence.

Thus, the tryptic peptide of the 33 kDa protein described herein represents a portion of a *Mortierella ramanniana* DAGAT. Other *Mortierella ramanniana* DAGAT peptides may be similarly obtained and the amino acid sequences determined.

The use of amino acid sequences from DAGAT peptides to obtain nucleic acid sequences which encode DAGAT is described herein. For example, synthetic oligonucleotides are prepared which correspond to the DAGAT peptide sequences. The oligonucleotides are used as primers in polymerase chain reaction (PCR) techniques to obtain partial DNA sequence of DAGAT genes. The partial sequences so obtained are then used as probes to obtain DAGAT clones from a gene library prepared from *Mortierella ramanniana* tissue. Alternatively, where oligonucleotides of low degeneracy can be prepared from particular DAGAT peptides, such probes may be used directly to screen gene libraries for DAGAT gene sequences. In particular, screening of cDNA libraries in phage vectors is useful in such methods due to lower levels of background hybridization.

A nucleic acid sequence of a DAGAT of this invention may be a DNA or RNA sequence, derived from genomic DNA, cDNA, mRNA, or may be synthesized in whole or in part. The gene sequences may be cloned, for example, by isolating genomic DNA from an appropriate source, and amplifying and cloning the sequence of interest using a polymerase chain reaction (PCR). Alternatively, the gene sequences may be synthesized, either completely or in part, especially where it is desirable to provide plant-preferred sequences. Thus, all or a portion of the desired structural gene (that portion of the gene which encodes the DAGAT protein) may be synthesized using codons

preferred by a selected host. Host-preferred codons may be determined, for example, from the codons used most frequently in the proteins expressed in a desired host species.

One skilled in the art will readily recognize that antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover "homologous" or "related" DAGATs from a variety of plant sources. Homologous sequences are found when there is an identity of sequence, which may be determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions between a known DAGAT and a candidate source. Conservative changes, such as Glu/Asp, Val/Ile, Ser/Thr, Arg/Lys and Gln/Asn may also be considered in determining sequence homology. Amino acid sequences are considered homologous by as little as 25% sequence identity between the two complete mature proteins. (*See* generally, Doolittle, R.F., *OF URFS and ORFS* (University Science Books, CA, 1986.)

Thus, other DAGATs may be obtained from the specific exemplified *Mortierella ramanniana* protein preparations and sequences provided herein. Furthermore, it will be apparent that one can obtain natural and synthetic DAGATs, including modified amino acid sequences and starting materials for synthetic-protein modeling from the exemplified DAGATs and from DAGATs which are obtained through the use of such exemplified sequences. Modified amino acid sequences include sequences which have been mutated, truncated, increased and the like, whether such sequences were partially or wholly synthesized. Sequences which are actually purified from plant preparations or are identical or encode identical proteins thereto, regardless of the method used to obtain the protein or sequence, are equally considered naturally derived.

Typically, a DAGAT sequence obtainable from the use of nucleic acid probes will show 60-70% sequence identity between the target DAGAT sequence and the encoding sequence used as a probe. However, lengthy sequences with as little as 50-60% sequence identity may also be obtained. The nucleic acid probes may be a lengthy fragment of the nucleic acid sequence,

or may also be a shorter, oligonucleotide probe. When longer nucleic acid fragments are employed as probes (greater than about 100 bp), one may screen at lower stringencies in order to obtain sequences from the target sample which have 20-50% deviation (i.e., 50-80% sequence homology) from the sequences used as probe. Oligonucleotide probes can be considerably shorter than the entire nucleic acid sequence encoding an DAGAT enzyme, but should be at least about 10, preferably at least about 15, and more preferably at least about 20 nucleotides. A higher degree of sequence identity is desired when shorter regions are used as opposed to longer regions. It may thus be desirable to identify regions of highly conserved amino acid sequence to design oligonucleotide probes for detecting and recovering other related DAGAT genes. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (*See*, Gould, *et al.*, *PNAS USA* (1989) *86*:1934-1938).

Furthermore, databases containing nucleic acid and amino acid sequences from various organisms may be searched with the sequences of the present invention to identify similar sequences. For example, using the jojoba DAGAT protein sequence to search a database containing DNA sequences from *Arabidopsis*, an approximately 12 Kb sequence containing at least seven repeats of an open reading frame with high similarity to the jojoba sequence are identified. In addition, one of the sequences is identified to be expressed preferentially in immature seeds of *Arabidopsis* and in 30 day post anthesis seeds of *Brassica*. The deduced amino acid sequences demonstrate a high level of homology to the jojoba wax synthase amino acid sequence. For example, amino acid sequence comparisons between jojoba wax synthase and the sequences obtained from the *Arabidopsis* homologues reveals an identity ranging from about 36% to about 44% between the amino acids. Furthermore, homologous similarity between the jojoba and *Arabidopsis* sequences is 52% for the lowest homologous similarity between jojoba and ATWS7. As used herein homologus

similarity refers to the amount of identity plus the amount of similarity between a set of sequences.

In addition to isolation of other DAGATs, it is considered that genes for other related acyltransferase proteins may also be obtained using sequence information from the DAGAT and related nucleic acid sequences. For example, other acyltransferase enzymes are involved in plant lipid biosynthesis, including plastidial DAGAT, mitochondrial DAGAT, lysophosphophatidylcholine acyltransferase (LPCAT), lysophosphophatidylserine acyltransferase (LPSAT), lysophosphophatidylethanolamine acyltransferase (LPEAT), and lysophosphophatidylinositol acyltransferase (LPIAT). While many of these enzymes catalyze acyltransferase reactions involving the *sn*-2 position of lysophospholipids, the genes encoding these sequences may also be related to the plant acyl-CoA DAGAT sequences of the instant invention and obtainable therefrom. Thus, as demonstrated herein, other related acyltransferases including fatty acyl-CoA:fatty alcohol 0-acyltransferase (wax synthase) from jojoba may be related to diacylglycerol acylransferases.

That DAGAT and wax synthase are members of a homologous family of proteins is supported by information obtained through the purification of DAGAT from a species of oleaginous fungus, *Mortierella ramanniana* (see Examples). Firstly, like jojoba wax synthase, the fungal DAGAT activity is membrane bound, and may be solubilized only through the use of detergents. Secondly, as with the jojoba wax synthase, it is necessary following solubilization to include a phospholipid (e.g., phosphatidic acid) in the assay mixture in order to restore enzyme activity of the fungal DAGAT. Thirdly, fungal DAGAT behaves very similar to jojoba wax synthase during the purification chromatography. Specifically, both enzyme species flow through a hydroxylapatite column with only a slight retardation, whereas most other protein species in the membrane preparations are bound to the column matrix (see Examples). Indeed, experience with the jojoba enzyme allowed the prediction that hydroxylapatite chromatography would be a key step in purification of DAGAT from *Mortierella ramanniana*.

The addition of this step proved essential to the successful purification of the fungal DAGAT, and led to the conclusion that the protocol of Kamisaka et al (1997) was not sufficient to identify the correct protein species associated with DAGAT activity. Finally, the apparent molecular weight of a fungal DAGAT polypeptide (about 33 kDa) as determined by SDS-PAGE is similar to that observed for jojoba wax synthase on SDS-PAGE (see Examples).

To determine if a related gene may be isolated by hybridization with a given sequence, the sequence is labeled to allow detection, typically using radioactivity, although other methods are available. The labeled probe is added to a hybridization solution, and incubated with filters containing the desired nucleic acids, such as Northern or Southern blots, or the filters containing cDNA or genomic clones to be screened. Hybridization and washing conditions may be varied to optimize the hybridization of the probe to the sequences of interest. Lower temperatures and higher salt concentrations allow for hybridization of more distantly related sequences (low stringency). If background hybridization is a problem under low stringency conditions, the temperature can be raised either in the hybridization or washing steps and/or salt content lowered to improve detection of the specific hybridizing sequence. Hybridization and washing temperatures can be adjusted based on the estimated melting temperature of the probe as discussed in Beltz, *et al.* (*Methods in Enzymology* (1983) *100*:266-285). A useful probe and appropriate hybridization and washing conditions having been identified as described above, cDNA or genomic libraries are screened using the labeled sequences and optimized conditions.

For immunological screening, antibodies to the DAGAT protein can be prepared by injecting rabbits or mice with the purified protein, such methods of preparing antibodies being well known to those in the art. Either monoclonal or polyclonal antibodies can be produced, although typically polyclonal antibodies are more useful for gene isolation. Western analysis may be conducted to determine that a related protein is present in a crude extract of the desired plant

species, as determined by cross-reaction with the antibodies to the coconut DAGAT. When cross-reactivity is observed, genes encoding the related proteins are isolated by screening expression libraries representing the desired plant species. Expression libraries can be constructed in a variety of commercially available vectors, including lambda gt11, as described in Maniatis, *et al.* (*Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

Many plants utilize DAGAT proteins in production of storage TAG in seeds, and thus any such plant species can be considered as a source of additional DAGAT proteins. Plants that have high oil contents in their seeds are preferred candidates to obtain plant DAGATS capable of increasing the ratio of oil to other constituents in seeds. Examples of such plants include but are not limited to macadamia nut, walnut, and castor.

Plants having high amounts of TAG with palmitate or stearate acyl groups at the sn-1 and sn-3 positions with oleate or linoleate at sn-2 are preferred candidates to obtain plant DAGATs capable of incorporating saturated fatty acids at the sn-3 position of TAG which show special selectivity for synthesis of structured TAG of the form S-U-S, where S=a saturated fatty acid and U=an unsaturated fatty acid. For example, oils from several tropical plants including cocoa, illipe, sal, shea, and *Garcinia* species such as kokum have been shown to accumulate high amounts of TAG in this form.

Plants having significant medium-chain fatty acids in their seed oils are preferred candidates to obtain plant DAGATs capable of incorporating medium-chain fatty acids into the *sn*-3 position of TAG. Several species in the genus *Cuphea* accumulate triglycerides containing medium-chain fatty acids in their seeds, e.g., *procumbens, lutea, hookeriana, hyssopifolia, wrightii* and *inflata*. Another natural plant source of medium-chain fatty acids are seeds of the *Lauraceae* family. In addition to the exemplified California Bay (*Umbellularia californica*), Pisa (*Actinodophne hookeri*), Sweet Bay (*Laurus nobilis*) and *Cinnamomum camphora* (camphor)

accumulate medium-chain fatty acids. Other plant sources include *Ulmaceae* (elm), *Palmae, Myristicaceae, Simarubaceae, Vochysiaceae,* and *Salvadoraceae.*

Also of particular interest are DAGATs from plant species which incorporate unusual long-chain fatty acids in the storage TAG. For example nasturtium and meadowfoam contain 22:1 acyl groups in the seed TAG.

It should also be noted that plant DAGATs from a variety of sources can be used to investigate TAG biosynthesis events of plant lipid biosynthesis in a wide variety of *in vivo* applications. Because all plants appear to synthesize lipids via a common metabolic pathway, the study and/or application of one plant DAGAT to a heterologous plant host may be readily achieved in a variety of species. In other applications, a plant DAGAT can be used outside the native plant source of the DAGAT to enhance the production and/or modify the composition of the TAG produced or synthesized *in vitro.*

The nucleic acid sequences associated with plant DAGAT proteins will find many uses. For example, recombinant constructs can be prepared which can be used as probes, or which will provide for expression of the DAGAT protein in host cells to produce a ready source of the enzyme and/or to modify the composition of triglycerides found therein. Other useful applications may be found when the host cell is a plant host cell, either *in vitro* or *in vivo.* For example, by increasing the amount of a respective medium-chain preferring DAGAT available to the plant TAG biosynthesis pathway, an increased percentage of medium-chain fatty acids may be obtained in the TAG. In a like manner, for some applications it may be desired to decrease the amount of DAGAT endogenously expressed in a plant cell by anti-sense technology. For example, to allow for more opportunity for an inserted foreign DAGAT to transfer saturated acyl groups, or medium-chain or unusual longer-chain fatty acyl groups to the *sn*-3 position, decreased expression of a native *Brassica* long-chain preferring DAGAT may be desired. Decreased endogenous DAGAT would allow for more opportunity for the seeds engineered with wax synthase to synthase long-chain liquid waxes. Indeed, decreased

endogenous DAGAT could be used to decrease synthesis of TAG so that additional carbon would be available for other seed constituents or products, including but not limited to protein and starch.

Thus, depending upon the intended use, the constructs may contain the sequence which encodes the entire DAGAT protein, or a portion thereof. For example, where antisense inhibition of a given DAGAT protein is desired, the entire DAGAT sequence is not required. Furthermore, where DAGAT constructs are intended for use as probes, it may be advantageous to prepare constructs containing only a particular portion of an DAGAT encoding sequence, for example a sequence which is discovered to encode a highly conserved DAGAT region.

As discussed above, nucleic acid sequence encoding a plant DAGAT of this invention may include genomic, cDNA or mRNA sequence. By "encoding" is meant that the sequence corresponds to a particular amino acid sequence either in a sense or anti-sense orientation. By "extrachromosomal" is meant that the sequence is outside of the plant genome of which it is naturally associated. By "recombinant" is meant that the sequence contains a genetically engineered modification through manipulation via mutagenesis, restriction enzymes, and the like.

A cDNA sequence may or may not contain pre-processing sequences, such as transit peptide sequences or targeting sequences to facilitate delivery of the DAGAT protein to a given organelle or membrane location. The use of any such precursor DAGAT DNA sequences is preferred for uses in plant cell expression. A genomic DAGAT sequence may contain the transcription and translation initiation regions, introns, and/or transcript termination regions of the plant DAGAT, which sequences may be used in a variety of DNA constructs, with or without the DAGAT structural gene. Thus, nucleic acid sequences corresponding to the plant DAGAT of this invention may also provide signal sequences useful to direct protein delivery into a particular organelle or membrane location, 5' upstream non-coding regulatory regions (promoters) having useful tissue and timing profiles, 3' downstream non-coding

regulatory region useful as transcriptional and translational regulatory regions and may lend insight into other features of the gene.

Once the desired plant DAGAT nucleic acid sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized. In the structural gene, one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient restriction site or other purpose involved with construction or expression. The structural gene may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

The nucleic acid or amino acid sequences encoding a plant DAGAT of this invention may be combined with other non-native, or "heterologous", sequences in a variety of ways. By "heterologous" sequences is meant any sequence which is not naturally found joined to the plant DAGAT, including, for example, combinations of nucleic acid sequences from the same plant which are not naturally found joined together.

The DNA sequence encoding a plant DAGAT of this invention may be employed in conjunction with all or part of the gene sequences normally associated with the DAGAT. In its component parts, a DNA sequence encoding DAGAT is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription and translation in a host cell, the DNA sequence encoding plant DAGAT and a transcription and translation termination region.

Potential host cells include both prokaryotic and eukaryotic cells. A host cell may be unicellular or found in a multicellular differentiated or undifferentiated organism depending upon the intended use. Cells of this invention may

be distinguished by having a plant DAGAT foreign to the wild-type cell present therein, for example, by having a recombinant nucleic acid construct encoding a plant DAGAT therein.

Depending upon the host, the regulatory regions will vary, including regions from viral, plasmid or chromosomal genes, or the like. For expression in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters may be employed. Expression in a microorganism can provide a ready source of the plant enzyme. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli, B. subtilis, Sacchromyces cerevisiae,* including genes such as beta-galactosidase, T7 polymerase, tryptophan E and the like.

For the most part, the constructs will involve regulatory regions functional in plants which provide for modified production of plant DAGAT, and possibly, modification of the fatty acid composition. The open reading frame, coding for the plant DAGAT or functional fragment thereof will be joined at its 5′ end to a transcription initiation regulatory region. In embodiments wherein the expression of the DAGAT protein is desired in a plant host, the use of all or part of the complete plant DAGAT gene is desired; namely all or part of the 5′ upstream non-coding regions (promoter) together with the structural gene sequence and 3′ downstream non-coding regions may be employed.

If a different promoter is desired, such as a promoter native to the plant host of interest or a modified promoter, i.e., having transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source, numerous transcription initiation regions are available which provide for a wide variety of constitutive or regulatable, e.g., inducible, transcription of the structural gene functions. The transcription/translation initiation regions corresponding to such structural genes are found immediately 5′ upstream to the respective start codons. Among transcriptional initiation regions used for plants are

such regions associated with the T-DNA structural genes such as for nopaline and mannopine synthases, the 19S and 35S promoters from CaMV, and the 5' upstream regions from other plant genes such as napin, ACP, SSU, PG, zein, phaseolin E, and the like. Enhanced promoters, such as double 35S, are also available for expression of DAGAT sequences. For such applications when 5' upstream non-coding regions are obtained from other genes regulated during seed maturation, those preferentially expressed in plant embryo tissue, such as ACP and napin-derived transcription initiation control regions, are desired. Such "seed-specific promoters" may be obtained and used in accordance with the teachings of U.S. Serial No. 07/147,781, filed 1/25/88 (now U.S. Serial No. 07/550,804, filed 7/9/90), and U.S. Serial No. 07/494,722 filed on or about March 16, 1990 having a title "Novel Sequences Preferentially Expressed In Early Seed Development and Methods Related Thereto," which references are hereby incorporated by reference. Transcription initiation regions which are preferentially expressed in seed tissue, i.e., which are undetectable in other plant parts, are considered desirable for TAG modifications in order to minimize any disruptive or adverse effects of the gene product.

Where expression of the DAGAT is to be directed in other plant tissues, additional promoters may find use in the constructs of the present invention. For example, where preferential expression of the genes in the pollen is desired, promoter regions, for example, Rop1At (Li, *et al.* (1998) *Plant Phyiol.* 118:407-417) and others described by Hamilton, *et al.* (1998) *Plant Mol Biol.* 38(4)663-669 may be employed. Where preferential expression in the epidermis is desired, promoter regions such as Ntltp1 (Canevascini, *et al.* (1996) *Plant Physiol.* 112:513-524) may be employed.

Regulatory transcript termination regions may be provided in DNA constructs of this invention as well. Transcript termination regions may be provided by the DNA sequence encoding the plant DAGAT or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally

associated with the transcript initiation region. Where the transcript termination region is from a different gene source, it will contain at least about 0.5 kb, preferably about 1-3 kb of sequence 3' to the structural gene from which the termination region is derived.

Plant expression or transcription constructs having a plant DAGAT as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life involved in the production of vegetable oils for edible and industrial uses. Most especially preferred are temperate oilseed crops. Plants of interest include, but are not limited to, rapeseed (Canola and High Erucic Acid varieties), sunflower, safflower, cotton, soybean, peanut, coconut and oil palms, and corn. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, this invention is applicable to dicotyledyons and monocotyledons species alike and will be readily applicable to new and/or improved transformation and regulation techniques.

Of particular interest, is the use of plant DAGAT constructs in plants which have been genetically engineered to produce a particular fatty acid in the plant seed oil, where TAG in the seeds of nonengineered plants of the engineered species, do not naturally contain that particular fatty acid. Thus, the expression of novel DAGAT in plants may be desirable for the incorporation of unique fatty acyl groups into the *sn*-3 position.

Further plant genetic engineering applications for DAGAT proteins of this invention include their use in preparation of structured plant lipids which contain TAG molecules having desirable fatty acyl groups incorporated into particular positions on the TAG molecules.

The method of transformation in obtaining such transgenic plants is not critical to the instant invention, and various methods of plant transformation are currently available. Furthermore, as newer methods become available to transform crops, they may also be directly applied hereunder. For

example, many plant species naturally susceptible to *Agrobacterium* infection may be successfully transformed via tripartite or binary vector methods of *Agrobacterium* mediated transformation. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation. In addition, techniques of microinjection, DNA particle bombardment, and electroporation have been developed which allow for the transformation of various monocot and dicot plant species.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of different host species the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), the latter being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cells and gall.

In some instances where *Agrobacterium* is used as the vehicle for transforming host plant cells, the expression or transcription construct bordered by the T-DNA border region(s) will be inserted into a broad host range vector capable of

replication in *E. coli* and *Agrobacterium*, there being broad host range vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, *et al.*, (*Proc. Nat. Acad. Sci., U.S.A.* (1980) 77:7347-7351) and EPA 0 120 515, which are incorporated herein by reference. Alternatively, one may insert the sequences to be expressed in plant cells into a vector containing separate replication sequences, one of which stabilizes the vector in *E. coli*, and the other in *Agrobacterium*. See, for example, McBride and Summerfelt (*Plant Mol. Biol.* (1990) *14*:269-276), wherein the pRiHRI (Jouanin, *et al.*, *Mol. Gen. Genet.* (1985) 201:370-374) origin of replication is utilized and provides for added stability of the plant expression vectors in host *Agrobacterium* cells.

Included with the expression construct and the T-DNA will be one or more markers, which allow for selection of transformed Agrobacterium and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, kanamycin, the aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

Examples

Example 1 - Wax synthase Assays

Methods to assay for wax synthase activity in microsomal membrane preparations or solubilized protein preparations are described.

A. Radiolabeled Material

The substrate generally used in the wax synthase assays, [1-^{14}C]palmitoyl-CoA, is purchased from Amersham (Arlington Heights, IL). Other chain length substrates were synthesized in order to perform chain length specification studies. Long chain [1-^{14}C] fatty acids (specific activity 51-56 Ci/mole), namely 11-cis-eicosenoic acid, 13-cis-docosenoic acid and 15-cis-tetracosenoic acid are prepared by the reaction of potassium [^{14}C]cyanide with the corresponding alcohol mesylate, followed by the base hydrolysis of the alcohol nitrile to the free fatty acid. The free fatty acids are converted to their methyl esters with ethereal diazomethane, and purified by preparative silver nitrate thin layer chromatography (TLC). The fatty acid methyl esters are hydrolyzed back to the free fatty acids. Radiochemical purity is assessed by three TLC methods: normal phase silica TLC, silver nitrate TLC, and C18 reversed phase TLC. Radiochemical purity as measured by these methods was 92-98%. Long chain [1-^{14}C] acyl-CoAs are prepared from the corresponding [1-^{14}C] free fatty acids by the method of Young and Lynen (*J. Bio. Chem.* (1969) *244*:377), to a specific activity of 10Ci/mole. [1-^{14}C]hexadecanal is prepared by the dichromate oxidation of [1-^{14}C]hexadecan-1-ol, according to a micro-scale modification of the method of Pletcher and Tate (*Tet. Lett.* (1978) 1601-1602). The product is purified by preparative silica TLC, and stored as a hexane solution at -70°C until use.

B. Assay for Wax synthase Activity in a Microsomal Membrane Preparation

Wax synthase activity in a microsomal membrane preparation is measured by incubation of 40μM [1-^{14}C]acyl-CoA (usually palmitoyl-CoA, sp. act. 5.1-5.6 mCi/mmol) and 200mM oleyl alcohol with the sample to be assayed in a total volume

of 0.25ml. The incubation mixture also contains either 25 mM HEPES (4-[2-hydroxyethyl]-1-piperazineethane-sulfonic acid), pH 7.5, as the buffering agent with 20% w/v glycerol, 1mM DTT, 0.5M NaCl or 25 mM Tricine-NaOH, pH 7.8, as the buffering agent with 0.28M NaCl, 10% glycerol, and 2mM ß-mercaptoethanol. Initial studies were performed with the first buffer system, when the pH was chosen to accomodate the preference of the acyl-CoA reductase enzyme. Membrane preparations were later changed to the second buffer system to accomodate the higher pH optimum of wax synthase.

A substrate mixture is prepared in a glass vial, with oleyl alcohol being added immediately before use, and is added to samples. Incubation is carried out at 30°C for up to one hour. The assay is terminated by placing the assay tube on ice and immediately adding 0.25ml isopropanol:acetic acid (4:1 v/v). Unlabeled wax esters (0.1mg) and oleyl alcohol (0.1mg) are added as carriers. The [^{14}C] lipids are extracted by the scaled-down protocol of Hara and Radin (*Anal. Biochem.* (1978) *90*:420). Two ml of hexane/isopropanol (3:2, v/v) is added to the terminated assay. The sample is vortexed, 1ml of aqueous sodium sulphate solution (6.6% w/v) is added, and the sample is again vortexed.

C. Assay for Solubilized Wax synthase Activity

Solubilized wax synthase is assayed using up to 50µl sample in a 250µl assay that contains 40µM 1-^{14}C-16:0 CoA (5 Ci/mol), 200µM 18:1-OH, 0.07% soybean phospholipid (Sigma, P-3644), 0.2 %CHAPS, 280 mM NaCl, 25 mM Tricine-NaOH, pH 7.8, 2mM ß-ME and 5.6% glycerol. Phospholipid (50mg/ml in 0.5% CHAPS) is added directly to the sample, which is in 1% CHAPS, then diluted by a cocktail containing the remaining assay components. Reconsitution of activity is presumed to be based on the incorporation of wax synthase into the phospholipid vesicles. Wax synthase is sensitive to detergent and requires the amount of phospholipid (PL) and detergent (CHAPS) to be balanced at 2.8/1 (CHAPS/PL, w/w) in the assay for maximal activity. Assays for wax synthase activity in samples concentrated by ultra-filtration require a readjustment of the

sample volume assayed because of the concentration of CHAPS. Introducing too much CHAPS into the assay results in inhibition of activity. If samples are concentrated by ultrafiltration, the optimum volume of sample to be assayed may be reestablished by performing a concentration curve of %CHAPS in the assay using a small amount of sample and assaying at a fixed concentration of phospholipid and NaCl. Wax synthase is less sensitive to changes in PL concentration than it is to changes in CHAPS concentration.

D. Analysis of Assay Products

For analyzing the products of either the microsomal membrane preparation wax synthase assay or the solubilized wax synthase assay, two protocols have been developed. One protocol, described below as "extensive assay" is more time-consuming, but yields more highly quantitative results. The other protocol, described below as "quick assay" also provides a measure of wax synthase activity, but is faster, more convenient and less quantitative.

1. *Extensive Analysis*: Following addition of the sodium sulphate and vortexing the sample, the upper organic phase is removed and the lower aqueous phase is washed with 4ml hexane/isopropanol (7:2 v/v). The organic phases are pooled and evaporated to dryness under nitrogen. The lipid residue is resuspended in a small volume of hexane, and an aliquot is assayed for radioactivity by liquid scintillation counting. The remainder of the sample can be used for TLC analysis of the labeled classes and thereby give a measure of total wax produced.

For lipid class analysis the sample is applied to a silica TLC plate, and the plate is developed in hexane/diethyl ether/acetic acid (80:20:1 or 70:30:2 v/v/v). The distribution of radioactivity between the lipid classes, largely wax esters, free fatty acids, fatty alcohols, and polar lipids at the origin, is measured using an AMBIS radioanalytic imaging system (AMBIS Systems Inc., San Diego, CA). If necessary the individual lipid classes can be recovered from the TLC plate for further analysis. Reversed-

phase TLC systems using C18 plates developed in methanol have also been used for the analysis.

2. *Quick Analysis*: Following addition of the sodium sulfate and vortexing the sample, a known percentage of the organic phase is removed and counted via liquid scintillation counting. This calculation is used to estimate the total counts in the organic phase. Another portion of the organic phase is then removed, dried down under nitrogen, redissolved in hexane and spotted on TLC plates and developed and scanned as described for the detailed assay. In this manner the percentage of the total counts which are incorporated into wax is determined.

Example 2 - Further Studies to Characterize Wax Synthase Activity

A. Seed Development and Wax Synthase Activity Profiles

Embryo development was tracked over two summers on five plants in Davis, CA. Embryo fresh and dry weights were found to increase at a fairly steady rate from about day 80 to about day 130. Lipid extractions reveal that when the embryo fresh weight reaches about 300mg (about day 80), the ratio of lipid weight to dry weight reaches the maximum level of 50%.

Wax synthase activity was measured in developing embryos as described in Example 1B. As the jojoba seed coats were determined to be the source of an inhibiting factor(s), the seed coats were removed prior to freezing the embryos in liquid nitrogen for storage at -70°C.

Development profiles for wax synthase activities as measured in either a cell free homogenate or a membrane fraction, indicate a large induction in activity which peaks at approximately 110-115 days after anthesis. Embryos for enzymology studies were thus harvested between about 90 to 110 days postanthesis, a period when the wax synthase activity is high, lipid deposition has not reached maximum levels, and the seed coat is easily removed. The highest rate of increase of wax synthase activity is seen between days 80 and 90 postanthesis. Embryos for cDNA library construction were thus harvested between about 80 to 90 days postanthesis when

presumably the rate of synthase of wax synthase protein would be maximal. Correspondingly, the level of mRNA encoding wax synthase would be presumed to be maximal at this stage.

B. Microsomal Membrane Preparation

Jojoba embryos are harvested at approximately 90-110 days after flowering, as estimated by measuring water content of the embryos (45-70%). The outer shells and seed coats are removed and the cotyledons quickly frozen in liquid nitrogen and stored at -70°C for future use. For initial protein preparation, frozen embryos are powdered by pounding in a steel mortar and pestle at liquid nitrogen temperature. In a typical experiment, 70g of embryos are processed.

The powder is added, at a ratio of 280ml of solution per 70g of embryos, to the following high salt solution: 3M NaCl, 0.3M sucrose, 100mM HEPES, 2mM DTT, and the protease inhibitors, 1mM EDTA, 0.7mg/ml leupeptin, 0.5mg/ml pepstatin and 17mg/ml PMSF. A cell free homogenate (CFH) is formed by dispersing the powdered embryos in the buffer with a tissue homogenizer (Kinematica, Switzerland; model PT10/35) for approximately 30 sec. and then filtering through three layers of Miracloth (CalBioChem, LaJolla, CA). The filtrate is centrifuged at 100,000 x g for one hour.

The resulting sample consists of a pellet, supernatant and a floating fat pad. The fat pad is removed and the supernatant fraction is collected and dialyzed overnight (with three changes of the buffering solution) versus a solution containing 1M NaCl, 100mM HEPES, 2mM DTT and 0.5M EDTA. The dialyzate is centrifuged at 200,000 x g for 1 1/2 hour to yield a pellet, DP2. The pellet is suspended in 25mM HEPES and 10% glycerol, at 1/20 of the original CFH volume, to yield the microsomal membrane preparation.

Activity is assayed as described in Example 1. Recovery of wax synthase activity is estimated at 34% of the original activity in the cell free homogenate. Wax synthase activity in this preparation is stable when stored at -70°C.

C. Substrate Specificity

Acyl-CoA and alcohol substrates having varying carbon chain lengths and degrees of unsaturation were added to

microsomal membrane fractions prepared as described above to determine the range of substrates recognized by the jojoba wax synthase.

Acyl-CoA and alcohol substrates having varying carbon chain lengths and degrees of unsaturation were added to microsomal membrane fractions prepared as described in Example 3A to determine the range of substrates recognized by the jojoba wax synthase. Assays were performed as described in Example 1B using the Tricine buffer system with the following change, both acyl-CoA and alcohol concentrations were 40 μM instead of the 200 μM alcohol concentration normally used. Acyl-CoA's were prepared as 2.5 mM stocks (in 1.25 mM Na Acetate buffer pH 4.8 and 1.5% CHAPS) and 4 μl of these stocks were used in a 250 μl assay making the final CHAPS concentration 0.024%. Without the addition of detergent, the long-chain saturated acyl-CoA's would not dissolve in the buffer. Alcohols were prepared as 25 mM stocks in 2-methoxyethanol and 0.4 μl of the stock was used in a 250 μl assay. To evaluate the acyl-CoA specificity, 1-^{14}C-hexadecanol (10.3 mCi/mmol, Sigma 31,326-2) was used as substrate. The purchased 1-^{14}C-hexadecanol was only 62% pure and had to be further purified by thin layer chromatography prior to use. The product was spotted onto a glass silica gel TLC plate and migrated in hexane:diethyl ether:acetic acid (70:30:2). Unlabeled alcohol was spotted in outside lanes and used to identify the migration level of the radiolabeled product. The TLC plate was briefly exposed to iodine vapors to identify the location of the alcohol. The 1-^{14}C-hexadecanol spot was scraped from the TLC plate and transferred to a new vial. The product was eluted from the silica with hexane/isopropanol and the organic extract was filtered to remove silica. The filtered solvent was transferred to a new vial where the solvent was evaporated to dryness. The final product was resuspended in 2-methoxyethanol at a concentration of 0.15 μCi/μl. The final product appeared to be 100% pure by TLC in the solvent system above. The 1-^{14}C-16:0-CoA was as described

in Example 1. Results of these experiments are presented in Table 1 below.

Table 1

Structure	Acyl Group	pmol/min/mg Alcohol Group
8:0	147.5	2656.3
10:0	197.8	2396.8
12:0	345.5	5663.4
14:0	1584.6	4919.1
16:0	1533.8	5250.6
18:0	1693.8	2557.9
20:0	1373.2	1666.4
22:0	1196.6	1555.9
24:0	1308.3	2582.0
18:1 9-c	821.9	12623.8
18:1 9-t	nd	12600.2
18:1 11-c	nd	12147.6
18:1 11-t	nd	13739.4
18:2 9-c,12-c	198.8	11344.2
18:3 9-2,12-c,15-c	516.1	nd
20:1 11-c	3880.1	6172.4
22:1 13-c	916.8	2783.0
22:1 13-t	nd	1862.0
24:1 15-c	1794.0	1576.3

For comparison purposes it is desirable to evaluate acyl-CoA's and alcohol's at equivalent concentrations, however, in practicality there is an endogenous pool of alcohol present in jojoba microsomal fractions of unknown concentration. This pool dilutes the ^{14}C labeled hexadecanol used to evaluate the acyl-CoA's rendering some of the waxes formed undetectable.

The result is an underestimation of the specific activity for all of the acyl-CoA's relative to the specific activities of the alcohols.

The above results demonstrate that the jojoba wax synthase utilizes a broad range of fatty acyl-CoA and fatty alcohol substrates.

In addition, wax synthase activity towards various acyl-thioester substrates was similarly tested using palmitoyl-CoA, palmitoyl-ACP and N-acetyl-S-palmitoyl cysteamine as acyl substrates. The greatest activity was observed with the acyl-CoA substrate. Significant activity (~10% of that with acyl-CoA) was observed with acyl-ACP, but no activity was detectable with the N-acetyl-S-palmitoyl cysteamine substrate.

D. Effectors of Activity

Various sulphydryl agents were screened for their effect on wax synthase activity. Organomercurial compounds were shown to strongly inhibit activity. Iodoacetamide and N-ethylmaleamide were much less effective. Inhibition by para-hydroxymercuribenzoate was observed, but this inhibition could be reversed by subsequent addition of DTT. These results demonstrate that inhibition by para-hydroxymercuribenzoate involves blocking of an essential sulphydryl group.

Example 3 - Purification of Jojoba Wax Synthase

Methods are described which may be used for isolation of a jojoba membrane preparation having wax synthase activity, solubilization of wax synthase activity, and further purification of the wax synthase protein.

A. Microsomal Membrane Preparation

The following modification of the method described in Example 2 is employed and provides an improved membrane fraction useful for purification of wax synthase from solubilized membranes.

Typically, 100 g of jojoba embryos are added to 400 ml of extraction buffer (40 mM Tricine-NaOH, pH 7.8, 200 mM KCl, 10 mM EDTA, 5 mM ß-mercaptoethanol), ground in a blender, and homogenized with a Polytron tissue disrupter. All subsequent steps are performed at 4°C. The blended material is filtered

through Miracloth (CalBioChem). Centrifugation (20,000 x *g*; 20 min.) of the filtrate yielded a floating wax layer, a turbid supernatant fraction and a dark green pellet. The supernatant fraction is collected and centrifuged (100,000 x *g*; 2 h) to obtain membrane pellets which are then resuspended in 40 ml of Buffer A (25 mM Tricine-NaOH, pH 7.8, 200 mM KCl, 5 mM EDTA, 5 mM ß-mercaptoethanol) containing 50% (w/v) sucrose. This homogenate is distributed into four SW28 centrifuge tubes (Beckman) and each is overlaid with 10 ml Buffer A containing 20% sucrose and then with 13 ml Buffer A. After centrifugation (28,000 rpm; 2 h), a membrane fraction is collected from the 20%/50% sucrose interface, diluted with four volumes Buffer A and collected by centrifugation (200,000 x *g*; 1 h). The membranes are then homogenized in 10 ml storage buffer [25 mM Tricine-NaOH, pH 7.8, 1 M NaCl, 10% (w/v) glycerol, 5 mM ß-mercaptoethanol)]. The protein concentration of membranes prepared via this protocol is typically between 7 and 9 mg/ml. Protein concentrations are estimated as described (Bradford, 1976) using BSA as the protein standard.

B. Solubilization of Wax synthase Protein

The membrane suspension is adjusted to approximately 0.83mg of protein per ml by dilution with storage buffer (25mM Tricine-NaOH, pH 7.8, 1M NaCl, 10% glycerol, 5 mM ß-mercaptoethanol). Solid 3-([3-cholamidopropyl]dimethyl-ammonio)-1-propanesulfate (CHAPS) is added to achieve a final concentration of 2% (w/v) and a detergent to protein ratio of 24:1. After incubation on ice for 1 hr, the sample is centrifuged (200,000g for 1 hr), and the supernatant fraction collected.

C. Purification of Wax Synthase Activity

The 200,000g supernatant fraction is diluted (with 0.57% CHAPS, 25 mM Tricine-NaOH, pH 7.8, 20% glycerol) to yield final concentrations of NaCl and CHAPS of 0.3M and 1%, respectively. The sample is loaded onto a Blue A-agarose (Amicon, Inc., Beverly, MA) column that has been equilibrated with buffer B (25 mM Tricine-NaOH, pH 7.8, 1% CHAPS, 20% glycerol,) containing 0.3M NaCl. After washing with

equilibration buffer, wax synthase activity is eluted with buffer B containing 2M NaCl. Active fractions eluted from the Blue A column are pooled (Blue Pool) and used for further chromatography.

Two purification protocols were used for band identification and further purification of the wax synthase protein. In Protocol 1 (Figure 1), the Blue Pool was concentrated 5.4 fold by ultrafiltration in a pressure cell fitted with a YM 30 membrane (Amicon, Inc., Beverly, MA). One-half of the concentrate was applied to a Ceramic Hydroxyapatite (CHT) column (Bio-Scale CHT-2; Bio-Rad, Hercules, CA) equilibrated in buffer B containing 2M NaCl. The column was washed with 6 column volumes of equilibration buffer and bound proteins were eluted with buffer B containing 0.1M dipotassium phosphate and 2M NaCl. After reequilibration of the CHT column, the second half of the Blue Pool concentrate was chromatographed in the same manner. In order to detect activity, wax synthase was assayed according to the protocol for samples concentrated by ultrafiltration. Wax synthase activity, measured on CHT-Run 1, was found in the flow through and wash. Protein profiles of the two CHT runs were identical so the CHT-run 2 was not assayed. Active fractions from the two CHT runs were pooled and concentrated 10 fold and applied to a Sephacryl S100 HR column (2.5 x 90cm) equilibrated in buffer B with 1.0 M NaCl. Protein and activity determinations were made and active fractions were selected from the retained portion of the run which maximized activity and minimized protein. The S100 pool (fractions 64-70) was applied to a crystalline hydroxylapatite (HA) column (Bio-Gel HT; Bio-Rad, Hercules, CA, 1 x 19.3cm) equilibrated in buffer B with 1 M NaCl. Again, the majority of the wax synthase activity was present in the flow through and wash. Bound proteins were eluted in buffer B with 0.1M dipotassium phosphate, and 1M NaCl. Fractions from the final HA run were examined by SDS-PAGE. A single protein migrating at 33 kD on SDS-PAGE was correlated with the presence of wax synthase activity.

In a second preparation (Protocol 2, Figure 2) the Blue Pool was applied directly to a crystalline HA column (1 x 11.7 cm), equilibrated in buffer B with 1M NaCl, without concentration. Two fractions were selected for further purification by size exclusion chromatography on a Superdex 75 HR 10/30 column (Bio-Rad, Hercules, CA; sizing range:5000 - 75,000 daltons) equilibrated with 25 mM Tricine-NaOH, pH 7.8, 1% CHAPS, 20% glycerol, 1M NaCl. Wax synthase activity was measured according to the protocol described for solubilized samples in Example 1C. One fraction eluted early in the flow through of the HA column (fraction 31) and the other eluted in the wash (fraction 67). The protein profiles of the two fractions were different based on SDS-PAGE analysis. Both Superdex 75 runs were examined by gradient SDS-PAGE and a protein of approximately 33 kD was identified that chromatographed with activity. A calibration curve was generated using molecular mass standards chromatographed under the same buffer and column conditions. Comparison of the elution volume of the peak of Wax Synthase activity to this standard curve yielded a value of 48 kDa for the molecular mass of the solubilized enzyme.

A chart representing the purification of wax synthase from Protocol 1 (Table 2) shows a 150 fold purification of the enzyme from the solubilized protein fraction.

Table 2

Purification of Jojoba Wax Synthase

Purification Step	Enzyme Activity (nmol/min)	Yield %	Protein (mg)	Specific Activity (nmol/min/mg)	Purification (fold)
Solubilized Fraction	274.4	100	415	0.7	1
Blue A Agarose	214.7	78.2	15	14.3	22
Ceramic Hydroxyapatite	176.6	64.3	6.4	27.6	42
Sephacryl S-100 (sizing)	41.3	15.1	1.2	33.1	50
Hydroxyapatite (crystalline)	18.8	6.9	0.2	99.2	150

D. DAGAT Assays of Purified Jojoba Wax synthase

The final fractions in the purification of jojoba wax synthase showed a single protein associated with that activity. Subsequent cloning of the gene and its expression in Arabidopsis (described below) confirmed the gene encoded a protein that has been demonstrated to be a membrane-associated wax synthase. The same fractions from jojoba wax synthase Protocol 1 were assayed for DAGAT activity. First, DAG substrates with acyl chain lengths from 8/8 DAG to 18:1/18:1 DAG were tested as substrates. The only DAG species which yielded a TAG product was 10/10 DAG. Next, utilizing the 10/10 DAG substrate, each column fraction was assayed for DAGAT activity. The results, shown in Figure 3, demonstrate that the observed DAGAT activity correlates with wax synthase activity over the column profile. This indicates that wax synthase was responsible for the observed DAGAT activity. The fact that 10/10 DAG was the only substrate utilized by wax synthase may be related to evolutionary changes which altered the substrate specificity. Subsequent identification of wax synthase homologs in Arabidopsis may yield further information on these evolutionary changes (see Example 17).

E. SDS PAGE Analysis

Samples from the column fractions were diluted in SDS PAGE sample buffer (1x buffer = 2% SDS, 250 mM ß-mercaptoethanol, 0.0025% bromphenol blue) and analyzed by electrophoresis. _ Polyacrylamide gradient gel electrophoresis (10-13%) was carried out according to the method of Laemmli (*Nature* (1970) *227*:680-685) with some of the modifications of Delepelaire (*Proc. Nat. Acad. Sci.* (1979) *76*:111-115). Sodium dodecyl sulfate was used in the upper reservoir buffer at 0.1% but was omitted from the lower reservoir buffer, stacking and resolving gels. The stacking gel contained 5% of a 30% acrylamide stock (29.2% acrylamide, 0.8% N,N'-bis-methyleneacrylamide, w/v), 0.06% ammonium persulfate (w/v) and 0.1% TEMED (v/v). The resolving gel contained a 10-13% linear gradient of acrylamide stock stabilized by a 0-10% linear gradient of sucrose. Electrophoresis was carried out at room temperature at 150V, constant voltage, for 9-10 hours. Proteins were visualized by staining with silver according to the method of Blum *et al.* (*Electrophoresis* (1987) *8*:93-99 or with Coomassie Blue (0.1% Coomassie Blue R-250, 50% methanol, 10% acetic acid). The 33 kDa protein identified as wax synthase does not appear as a major component of the active fraction until purification through the hydroxyapatite column. Following purification Protocol 1 (Example 3C) the only protein that correlates with activity on the final column is one at 33 kDa.

Example 4 - Preparation of Jojoba Wax synthase Protein for In-Gel Digestion

A. Preparation of Samples for SDS-PAGE by Concentration

Odd numbered fractions from the flow through/wash of the final HA column (Protocol 1) were pooled and concentrated three fold by ultrafiltration in a pressure cell fitted with a YM 30 membrane (Amicon, Inc., Beverly, MA). The sample was further concentrated using two Centricon-30 units (Amicon, Inc., Beverly, MA) to volumes of approximately 50µl. Each sample was treated with 6µl SDS Cocktail (4µl 20%SDS, 1µl

14.3M ß-metcaptoethanol, and 1µl 0.1% Bromophenol Blue). After sitting at room temperature for 15 minutes, the samples were applied to a 10-13% acrylamide gradient gel (Example 3D) (16 x 16 cm x 1mm thick) and proteins were resolved by electrophoresis at 150V, constant voltage, for 9.5 hours. The gel was stained with 0.1% Coomassie Blue in 50% methanol, 10% acetic acid for 15 minutes then destained in 50% methanol, 10% acetic acid for 2 x 20 minutes. The 33 kD Wax Synthase band was excised from the gel and destained in 50% ethanol for 3 x 20 minutes. One lane contained a streak of protein and was not used in the final digestion.

B. Preparation of Samples for SDS-PAGE by Precipitation

Aliquots (0.8 ml) of the even numbered fractions from the final HA column (Protocol 1) were pooled in groups of three over the column profile. The pools were divided equally into three, 1.5 ml vials. Protein was precipitated by the addition of 0.2ml 40% TCA. After 30 minutes on ice the samples were centrifuged (12,000 x g, 15 minutes at 4 C) to pellet the precipitated protein. The supernatants were removed and the pellets washed twice with 0.6 ml ice cold acetone. The final three pellets for each pooled set of samples were resuspended with the same 50 µl of SDS sample buffer by transfering the buffer from one vial to the next. The emptied vials, that had already been resuspended, were washed with 10µl of sample buffer for a total resuspended volume of 60 µl for each pooled sample. The samples were applied to a 12% acrylamide Tris/Glycine mini-gel (Novex, San Diego, CA,1.5mm x 10 well) and proteins were resolved by electrophoresis at 150 V, constant voltage, for 20 minutes beyond the elution of dye from the foot of the gel. The gel was stained with Coomassie Blue and destained using Gel-Clear (Novex, San Diego, CA). Wax Synthase was excised from three non-equivalent lanes on the gel representing the peak and tailing fractions from the column. The gel slices were placed in 1.5 ml vials and destained with 1 ml of 50% methanol, 10% acetic acid for 2 hours. The destain solution was removed and the gel slices were frozen in liquid nitrogen and sent on dry ice, overnight, to the W M Keck Foundation Biotechnology Resource Laboratory

at Yale University for in-gel-digestion. One gel slice from the sample concentrated by ultrafiltration and three gel slices from the samples concentrated by precipitation were pooled for in-gel tryptic digestion.

Example 5 - Determination of Amino Acid Sequence

DNA sequences encoding wax synthase peptides are obtained from jojoba using synthetic oligonucleotides designed from wax synthase peptide sequences. The wax synthase nucleic acid sequences may be obtained by amplification of DNA by polymerase chain reaction (PCR) using oligonucleotides as primers, or alternatively, by screening a cDNA or genomic DNA library by radiolabeling the oligonucleotides or previously isolated sequences for use as probes.

A. Construction of Jojoba cDNA Libraries

RNA is isolated from jojoba embryos collected at 80-90 days post-anthesis using a polyribosome isolation method, initially described by Jackson and Larkins (*Plant Physiol.* (1976) *57*:5-10), as modified by Goldberg *et al.* (*Developmental Biol.* (1981) *83*:201-217). In this procedure all steps, unless specifically stated, are carried out at 4°C. 10gm of tissue are ground in liquid nitrogen in a Waring blender until the tissue becomes a fine powder. After the liquid nitrogen has evaporated, 170ml of extraction buffer (200mM Tris pH 9.0, 160mM KC1, 25mM EGTA, 70mM MgC12, 1% Triton X-100, 05% sodium deoxycholate, 1mM spermidine, 10mM ß-mercaptoethanol, and 500mM sucrose) is added and the tissue is homogenized for about 2 minutes. The homogenate is filtered through sterile miracloth and centrifuged at 12,000 x g for 20 minutes. The supernatant is decanted into a 500ml sterile flask, and 1/19 volume of a 20% detergent solution (20% Brij 35, 20% Tween 40, 20% Noidet p-40 w/v) is added at room temperature. The solution is stirred at 4°C for 30 minutes at a moderate speed and the supernatant is then centrifuged at 12,000 x g for 30 minutes.

About 30ml of supernatant is aliquoted into sterile Ti 60 centrifuge tubes and underlaid with 7ml of a solution

containing 40mM Tris pH 9.0, 5mM EGTA, 200mM KC1, 30mM MgC12, 1.8M sucrose, 5mM ß-mercaptoethanol. The tubes are filled to the top with extraction buffer, and spun at 60,000 rpm for 4 hours at 4°C in a Ti60 rotor. Following centrifugation, the supernatant is aspirated off and 0.5ml of resuspension buffer (40mM Tris pH 9.0, 5mM EGTA, 200mM KC1, 30mM $MgC1_2$, 5mM ß-mercaptoethanol) is added to each tube. The tubes are placed on ice for 10 minutes, after which the pellets are thoroughly resuspended and pooled. The supernatant is then centrifuged at 120 x g for 10 minutes to remove insoluble material. One volume of self-digested 1mg/ml proteinase K in 20mM Tris pH 7.6, 200mM EDTA, 2% N-lauryl-sarcosinate is added to the supernatant and the mixture incubated at room temperature for 30 minutes.

RNA is precipitated by adding 1/10 volume of sodium acetate and 2 volumes of ethanol. After several hours at -20°C RNA is pelleted by centrifugation at 12,000 x g at 4°C for 30 minutes. The pellet is resuspended in 10ml of TE buffer (10mM Tris, 1mM EDTA) and extracted with an equal volume of Tris pH 7.5 saturated phenol. The phases are separated by centrifuging at 10,000 x g for 20 minutes at 4°C. The aqueous phase is removed and the organic phase is re-extracted with one volume of TE buffer. The aqueous phases are then pooled and extracted with one volume of chloroform. The phases are again separated by centrifugation and the aqueous phase ethanol precipitated as previously described, to yield the polyribosomal RNA.

Polysaccharide contaminants in the polyribosomal RNA preparation are removed by running the RNA over a cellulose column (Sigma-cell 50) in high salt buffer (0.5M NaCl, 20mM Tris pH 7.5, 1mM EDTA, 0.1% SDS). The contaminant binds to the column and the RNA is collected in the eluant. The eluant fractions are pooled and the RNA is ethanol precipitated. The precipitated total RNA is then resuspended in a smaller volume and applied to an oligo d(T) cellulose column to isolate the polyadenylated RNA.

Polyadenylated RNA is used to construct a cDNA library in the plasmid cloning vector pCGN1703, derived from the

commercial cloning vector Bluescribe M13- (Stratagene Cloning Systems; San Diego, CA), and made as follows. The polylinker of Bluescribe M13- is altered by digestion with *Bam*HI, treatment with mung bean endonuclease, and blunt-end ligation to create a *Bam*HI-deleted plasmid, pCGN1700. pCGN1700 is digested with *Eco*RI and *Sst*I (adjacent restriction sites) and annealed with a synthetic linker having restriction sites for *Bam*HI, *Pst*I, *Xba*I, *Apa*I and *Sma*I, a 5' overhang of AATT, and a 3' overhang of TCGA. The insertion of the linker into pCGN1700 eliminates the *Eco*RI site, recreates the *Sst*I (also, sometimes referred to as "*Sac*I" herein) site found in Bluescribe, and adds the new restriction sites contained on the linker. The resulting plasmid pCGN1702, is digested with *Hin*dIII and blunt-ended with Klenow enzyme; the linear DNA is partially digested with *Pvu*II and ligated with T4 DNA wax synthase in dilute solution. A transformant having the *lac* promoter region deleted is selected (pCGN1703) and is used as the plasmid cloning vector.

Briefly, the cloning method for cDNA synthesis is as follows. The plasmid cloning vector is digested with *Sst*I and homopolymer T-tails are generated on the resulting 3'-overhang stick-ends using terminal deoxynucleotidyl transferase. The tailed plasmid is separated from undigested or un-tailed plasmid by oligo(dA)-cellulose chromatography. The resultant vector serves as the primer for synthesis of cDNA first strands covalently attached to either end of the vector plasmid. The cDNA-mRNA-vector complexes are treated with terminal transferase in the presence of deoxyguanosine triphosphate, generating G-tails at the ends of the cDNA strands. The extra cDNA-mRNA complex, adjacent to the *Bam*HI site, is removed by *Bam*HI digestion, leaving a cDNA-mRNA-vector complex with a *Bam*HI stick-end at one end and a G-tail at the other. This complex is cyclized using an annealed synthetic cyclizing linker which has a 5' *Bam*HI sticky-end, recognition sequences for restriction enzymes *Not*I, *Eco*RI and *Sst*I, and a 3' C-tail end. Following ligation and repair the circular complexes are transformed into *E. coli* strain DH5a (BRL, Gaithersburg, MD) to generate the cDNA library. The

jojoba embryo cDNA bank contains between approximately $1.5x10^6$ clones with an average cDNA insert size of approximately 500 base pairs.

Additionally, jojoba polyadenylated RNA is also used to construct a cDNA library in the cloning vector lZAPII/*Eco*RI (Stratagene, San Diego, CA). The library is constructed using protocols, DNA and bacterial strains as supplied by the manufacturer. Clones are packaged using Gigapack Gold packaging extracts (Stratagene), also according to manufacturer's recommendations. The cDNA library constructed in this manner contains approximately 1 x 10^6 clones with an average cDNA insert size of approximately 400 base pairs.

B. Synthetic Oligonucleotides

In general, for use as PCR primers from single stranded DNA template reverse-transcribed from mRNA, oligonucleotides containing the sense orientation sequence corresponding to wax synthase peptide encoding sequences are prepared. These oligonucleotides are used as primers for the "forward" amplification reaction to produce sense strand DNA.

For the "reverse" reaction for amplification of the non-coding DNA strand, an oligonucleotide may be designed to be identical to a portion of a primer used to prepare DNA template for PCR. Alternatively, oligonucleotides which contain sequence complementary to wax synthase peptide encoding sequences may be used in combination with a "forward" wax synthase oligonucleotide primer as described above.

Where the wax synthase peptide sequences contain amino acids which may be encoded by a number of different codons, the forward or reverse primers may be "degenerate" oligonucleotides, i.e. containing a mixture of all or some of the possible encoding sequences for a particular peptide region. To reduce the number of different oligonucleotides present in such a mixture, it is preferable to select peptide regions which have the least number of possible encoding sequences when preparing the synthetic oligonucleotide for PCR primers. Similarly, where the synthetic oligonucleotide is to

be used to directly screen a library for wax synthase sequences, lower degeneracy oligonucleotides are preferred.

Following is an example of the sequence of peptide WSPEP33 (center line) and the forward (top line) and reverse (bottom line) DNA sequences that encode the peptide WSPEP33.

```
5' TTY GTN CCN GCN GTN GCN CCN CAY GGN GGN GCN YTN MGN 3'
    F   V   P   A   V   A   P   H   G   G   A   L   R
3' AAR CAN GGN CGN CAN CGN GGN GTR CCN CCN CGN RAN KCN 5'
```

Following is an example of the sequence of peptide WSPEP29 (center line) and the forward (top line) and reverse (bottom line) DNA sequences that encode the peptide WSPEP29.

```
5' ACN ATH GAY GAR TAY CCN GTN ATG TTY AAY TAY ACN CAR AAR 3'
    T   I   D   E   Y   P   V   M   F   N   Y   T   Q   K
3' TGN TAD CTR CTY ATR GGN CAN TAC AAR TTR ATR TGN GTY TTY 5'
```

Following is an example of the sequence of peptide WSPEP14 (center line) and the forward (top line) and reverse (bottom line) DNA sequences that encode the peptide WSPEP14.

```
5' TTY MGN GAY GAY CCN WSN AAY GAY CAY 3'
    F   R   D   D   P   S   N   D   H
3' AAR KCN CTR CTR GGN WSN TTR CTR GTR 5'
```

Following are sequences of synthetic oligonucleotides which may be used to obtain wax synthase sequences. The oligonucleotide names reflect the particular wax synthase peptide fragment numbers as listed in Example 5. The letter "F" in the oligonucleotide name designates a PCR forward reaction primer. The letter "R" designates a PCR reverse reaction primer.

WSPEP29-F1	5' TTYGTNCCNGCNGTNGC 3'
WSPEP29-F2	5' GCNCCNCAYGGNGGNGC 3'
WSPEP29-R1	5' GCNCCNCCRTGNGGNGC 3'

WSPEP29-R2	5'	GCNACNGCNGGNACRAA 3'
WSPEP33-F1	5'	ACNATHGAYGARTAYCCNGT 3'
WSPEP33-F2	5'	CCNGTNATGTTYAAYTAYAC 3'
WSPEP33-R1	5'	TTYTGNGTRTARTTRAACAT 3'
WSPEP33-R2	5'	AACATNACNGGRTAYTCRTC 3'
WSPEP14-F1	5'	GAYGAYCCNWSNAAYGAYCA
WSPEP14-R1	5'	TGRTCRTTNSWNGGRTCRTC

The nucleotide base codes for the above oligonucleotides are as follows:

A = adenine	T = thymine	Y = cytosine or thymine
C = cytosine	U = uracil	R = adenine or guanine
G = guanine	I = inosine	O = inosine or cytosine

H = adenine, cytosine or thymine
N = adenine, cytosine, guanine or thymine
W = adenine or thymine
S = guanine or cytosine
B = guanine, cytosine or thymine
K = guanine or thymine
M = adenine or cytosine

C. PCR Reactions

Poly(A)+ RNA is isolated from total RNA prepared from jojoba tissue as described above. cDNA is prepared from poly(A)+ or total RNA by reverse transcription using the Marathon cDNA Amplification Kit (Clontech Laboraties Inc according to the manufacturer's directions. The jojoba cDNA is used in PCR reactions 1-16 set forth below.

PCR is conducted in a Perkin Elmer Cetus GeneAmp PCR System 9600 PCR machine using reverse transcribed single-stranded cDNA as template. Commercially available PCR reaction and optimization reagents are used according to manufacturer's specifications

Reaction	Forward Primer	Reverse Primer
1	WSPEP14-F1	WSPEP29-R1
2	WSPEP14-F1	WSPEP29-R2
3	WSPEP14-F1	WSPEP33-R1
4	WSPEP14-F1	WSPEP33-R2
5	WSPEP29-F1	WSPEP14-R1
6	WSPEP29-F1	WSPEP33-R1
7	WSPEP29-F1	WSPEP33-R2
8	WSPEP29-F2	WSPEP14-R1
9	WSPEP29-F2	WSPEP33-R1
10	WSPEP29-F2	WSPEP33-R2
11	WSPEP33-F1	WSPEP14-R1
12	WSPEP33-F1	WSPEP29-R1
13	WSPEP33-F1	WSPEP29-R2
14	WSPEP33-F2	WSPEP14-R1
15	WSPEP33-F2	WSPEP29-R1
16	WSPEP33-F2	WSPEP29-R2

The temperature program used for PCR amplification is as follows: 1 cycle of 95 degrees C for 2 minutes; 4 cycles of 95 degrees C for 30 seconds, 60 degrees C for 1 minute, and 72 degrees C for 4 minutes; 4 cycles of 95 degrees C for 30 seconds, 57 degrees C for 1 minute, and 72 degrees C for 4 minutes; 4 cycles of 95 degrees C for 30 seconds, 54 degrees C for 1 minute, and 72 degrees C for 4 minutes; 4 cycles of 95 degrees C for 30 seconds, 51 degrees C for 1 minute, and 72 degrees C for 4 minutes; and 25 cycles of 95 degrees C for 30 seconds, 48 degrees C for 1 minute, and 72 degrees C for 4 minutes.

From reactions 3 and 4, a PCR product approximately 700 nucleotides in length was detected. The PCR product was purified using gel electrophoresis and cloned into pCR2.1 using a Topo TA cloning kit (Invitrogen Corp.). The DNA sequence of the cloned PCR product was determined and was 708 nucleotides long (Figure 3).

The entire cDNA can be amplified using 5' and 3' RACE (Frohman et al., 1988) using the Marathon cDNA Amplification Kit (Clontech Laboraties Inc.) according to the manufacturers instructions. From the sequence of the 708 nucleotide PCR fragment derived using primers WSPEP14-F1 and WSPEP33-R2 the following primers were synthesized:

WSRACEF1	GATTTGCCTCATTTTGTGATCTCGGTGCT
WSRACEF2	GACCTATACCCCCAGTTCAACGAGCCATAC
WSRACEF3	TTCAACGAGCCATACTTAGCCACCTCGCTG
WSRACER1	AACAACCACCCTCCAGTCACCATCACGAAC
WSRACER2	TTGCCTGAAACCGCCTTCTTCACCACCATC
WSRACER3	AAGATGTCTGACACCATGAGGTTCCACCTG

3'RACE reactions were set up using parimers WSRACEF1,WSRACEF2,and WSRACEF3. 5'RACE reactions were set up using parimers WSRACER1,WSRACER2,and WSRACER3. PCR reactions were performed according to the manufacturer's protocol (Clontech Laboraties Inc.). All 6 PCR reactions gave visible PCR products ranging in size from approximately 700 nucleotides to 1000 nucleotides. The PCR products were gel purified and cloned into pCR2.1 according to the manufacturer's protocol (Invitrogen Corp.). The DNA sequence of several clones from both the 5' and 3' RACE reactions and the previous PCR product derived from primers WSPEP14-F1 and WSPEP33-R2 were assembled using Sequencher software (Gene Codes Corp.). The assembled sequence of all the PCR products contains the coding region of the cDNA sequence.

To isolate a gene fragment suitable for cloning the wax synthase gene into expression cassettes for plant lipid modification, the coding region of the gene can be amplified from cDNA using the primers WAXSYNFOR and WASXYNREV. The sequence of WAXSYNFOR is GGATCCGTCGACACAATGGAGGTGGAGAAGGAGCTAAAG, and the sequence of WASXYNREV is GCATGCAGATCTCACCACCCCAACAAACCCATC. The PCR reaction is performed using the Marathon CDNA (Clontech Laboraties Inc.) according to the manufacturer's instructions. The PCR program consists of 30 cycles of 94 degrees C for 15 seconds, 60 degrees C for 1 minute, 72 degrees C for 2

minutes. The PCR products were cloned into pCR2.1 according to the manufacturers protocol (Invitrogen Corp.). The resulting plasmid was designated pCGN8538. The nucleic acid sequence and the derived amino acid sequence of the jojoba wax synthase is determine and shown in Figures 10 and 11.

Example 6 - Wax Synthase Constructs for Plant Expression

Constructs are prepared to express the jojoba wax synthase sequence in plant cells. Constructs for the expression of the wax synthase sequence alone are prepared to confirm wax synthase as well as DAGAT enzyme activities for the sequence. In addition, constructs employing the jojoba wax synthase sequence with sequences for reductase, as well as reductase and an elongase are also prepared.

The primers GGATCCGTCGACACAATGGAGGTGGAGAAGGAGCTAAAG and GCATGCAGATCTCACCACCCCAACAAACCCATC are used to PCR amplify the wax synthase open reading from and introduce BamHI and SalI restriction sites at the 5′ end of the gene and SphI and BglII site at the 3′ end of the gene. PCR amplification was carried out using the Marathon cDNA previously described and 30 cycles of PCR amplification using the cycle profile: 15 seconds at 94 degrees C, 30 seconds at 60 degrees C, and 1 minute at 72 degrees C. The PCR product was cloned into plasmid pCR2.1 to yield pCGN8538 and the sequence was determined to ensure that no errors were introduced during the PCR reaction. The insert from pCGN8538 was cloned as a SalI-BglII fragment into the napin cassette of pCGN7770 to form pCGN8553. Plasmid pCGN 8557 was constructed by cloning the napin/jojoba reductase gene fusion from pCGN7698 (USPN 5,445,947) and the napin/lunaria KCS (described in European Patent Application Publication 0731840) from pCGN7844 into the Asp718 and NotI sites of binary vector pCGN5139PASS. Plasmid pCGN8559 was constructed by cloning the napin/wax synthase gene fusion from plasmid pCGN8553 into the Sse8387I site of plasmid pCGN8557.

For constitutive expression of the wax synthase in transgenic plants, the SalI-BglII fragment from pCGN8538 was cloned into the SalI and BamHi sites of the double 35S (D35S)

expression cassette vector pCGN7787 (described below) to yield pCGN8591. The D35S/wax synthase gene fusion was cloned into the NotI site of pCGN5139 (described below) to yield pCGN8593.

The binary vectors pCGN8557, pCGN8559, and pCGN8593 are introduced into Agrobacterium tumefaciens EHA105 via electroporation. Arabidopsis plants are transformed by vacuum infiltration. Brassica plants are transformed with pCGN8557 and pCGN8559 as previously described.

Wax synthase and reductase gene sequences may be inserted into such cassettes to provide expression constructs for plant transformation methods. For example, a construct for expression of reductase in plant cells using 5' and 3' regulatory regions from a napin gene, is described in USPN 5,445,947.

Binary vector constructs are transformed into *Agrobacterium* cells, such as of strain EHA101 (Hood *et al.*, *J. Bacteriol* (1986) *168*:1291-1301), by the method of Holsters *et al.* (*Mol. Gen. Genet.* (1978) *163*:181-187) and used in plant transformation methods as described below.

Example 7- Diacylglycerol acyltransferase (DAGAT) Assays

Methods to assay for DAGAT activity in non-solubilized or solubilized protein preparations are described for *Mortierella ramanniana*.

A. Non-solubilized samples

DAGAT activity is assayed with 3.67 μM 1-^{14}C-18:1-Coenzyme A (53.5-54.5 Ci/mole, New England Nuclear, Boston, MA) and 1.5 mM 1,2-18:1 diacylglycerol (DAG) (Sigma D-0138, prepared as a 150 mM stock in 2-methoxyethanol) in a buffer containing 10 mM potassium phosphate (pH 7.0), 100-150 mM KCl, and 0.1 % TX-100 (w/v) in a total volume of 100 μl as similarly described by Kamisaka et al. (1993, 1994). Assays are performed at 30 °C for 5 min and terminated with the addition of 1.5 ml of heptane:isopropanol:0.5M H_2SO_4 (10:40:1, v/v/v). If necessary, samples may be diluted with buffer prior to assay in order to maintain a linear rate of product formation during the assay.

B. Solubilized samples

The assay is performed as described for non-solubilized samples with the following changes: the amount of 1,2-18:1 DAG is reduced to 0.5 mM, the amount of Triton X-100 is increased to 0.2%, and the KCl concentration is maintained between 100-125 mM. It is also necessary to include *L*-α-phosphatidic acid (Sigma P-9511, prepared as a 50 mM stock in 1% Triton X-100 (w/v)) to recover activity following solubilization with detergent as described by Kamisaka et al. (1996, 1997), with slight modification of the protocol. We find that using 300 μM of phosphatidic acid rather than 500 μM gives a higher stimulation of DAGAT activity following treatment by Triton X-100. We also find that DAGAT activity is sensitive to the amount of KCl introduced in the assay with the optimum level between 100-125 mM. Assays are performed at 30 °C for 5-30 minutes and terminated as described for non-solubilized samples.

C. Processing of Sample assays

After assays are stopped, they can be stored at 4 °C for processing at a later date or immediately processed by addition of 0.1 ml 1 M $NaHCO_3$ followed by 1 ml of heptane containing 15 nmoles/ml triolein as a carrier for extraction. The contents are vortexed and, after separation of aqueous and organic phases, the upper organic phase is removed to a new glass vial and washed with 1 ml 1M NaCl. Forty percent of the final organic phase is removed for liquid scintillation counting and the remaining organic phase is transferred to a clean vial and evaporated to dryness under nitrogen gas. The residue is resuspended in 45 μl hexane and spotted onto a silica gel-G, glass, thin-layer chromatography (TLC) plate with a preadsorbent loading zone (Analtech #31011, Newark, Delaware). The TLC plate is developed in hexane:diethyl ether:acetic acid (50:50:1, v/v/v) to the top then dried and scanned by a radio-image analyzer (AMBIS 3000, San Diego, CA)

to determine the portion of radioactivity incorporated into triacylglycerol. Activity is reported in units as pmole/min.

Example 8 - Growth and Harvesting of *Mortierella ramanniana* cultures.

Mortierella ramanniana is cultured by inoculating 1 liter of Defined Glucose Media (30 g glucose, 1.5 g $(NH_4)_2SO_4$, 3 g K_2HPO_4, 0.3 g $MgSO_4 \bullet 7H2O$, 0.1 g NaCl, 5g $CH_3COONa \bullet 3H_2O$, 10 mg $FeSO_4 \bullet 7H_2O$, 1.2 mg $CaCl_2 \bullet 2H_2O$, 0.2 mg $CuSO_4 \bullet 5H_2O$, 1.0 mg $ZnSO_4 \bullet 7H_2O$, 1.0 mg $MnCl_2 \bullet 4H_2O$, 2 mg thiamine-HCl and 0.02 mg biotin in 1 L of water purified by reverse osmosis (pH 5.7)) with 1.5-3 x 10^6 spores and incubating at 30 °C with shaking at 200 rpm for 9-11 days. Cultures are harvested by filtration through one layer of Miracloth (Calbiochem, La Jolla, CA). Excess liquid is removed by hand squeezing. The average yield of packed cells per liter harvested is 22.5 g.

Example 9 - Gradient Gel Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Samples from the column fractions are diluted in SDS-PAGE sample buffer (1x buffer = 2% SDS, 250 mM β-mercaptoethanol, 0.0025% bromphenol blue) and analyzed by electrophoresis. Polyacrylamide gradient gel electrophoresis (10-13%) is carried out according to the method of Laemmli (1970) with some of the modifications of Delepelaire (1979). Sodium dodecyl sulfate is used in the upper reservoir buffer at 0.1% but is omitted from the lower reservoir buffer, stacking and resolving gels. The stacking gel contains 5% of a 30% acrylamide stock (acrylamid:N,N'-Methylenacrylamid, 37.5:1, Bio-Rad, Hercules, CA), 0.06% ammonium persulfate and 0.1% TEMED (v/v). The resolving gel contains a 10-13% linear gradient of acrylamide stock stabilized by a 0-10% linear gradient of sucrose. Electrophoresis is carried out at room temperature at 150V, constant voltage, for 7-9 hours. Proteins are visualized by staining with silver according to the method of Blum *et al.* (1987) or with Coomassie Blue (0.1%

Coomassie Blue R-250, 50% methanol (v/v), 10% acetic acid (v/v)).

Example 10 - Evaluation of the Chromatography Used by Kamisaka et al. (1997) in the Purification of DAGAT from *Mortierella ramanniana*

A. Preparation of the Lipid Body Fraction

The following steps are performed at 4 °C.

Typically, 70-75 g of wet packed cells (stored at -70 °C) are used for each lipid body preparation. Just prior to use, cells are thawed on ice and resuspended in 150 ml of Buffer A (10 mM potassium phosphate (pH 7.0), 0.15 M KCl, 0.5 M sucrose, and 1 mM EDTA). The following protease inhibitors are added to reduce proteolysis: 0.1 μM Aprotinin, 1 μM Leupeptin, and 100 μM Pefabloc (all from Boehringer Mannheim, Germany). Cells are divided into five, 50-ml tubes and lysed with a Polytron Tissue Homogenizer (Kinematic GmbH, Brinkman Insruments, Switzerland) on setting #7 with a 1 cm diameter probe for 7 x 1 min. The resulting slurry is transferred to centrifuge tubes (29 x 104 mm) and solid debris made to pellet by spinning at 1500 x g (Beckman Instruments, J2-21, JA-20 rotor, 3500 rpm) for 10 min at 4 °C. The supernatant is removed and the pellets washed with another 5 ml of Buffer A. Following centrifugation, the supernatant volumes are combined. This fraction is referred to as the 'S1'. The S1 is divided into six ultracentrifuge tubes (25 x 89 mm, Beckman Instruments, Fullerton, CA) and each is overlayed with 5 ml of Buffer B (10 mM potassium phosphate pH, 7.0, 0.15 M KCl, 0.3 M sucrose, and 1 mM EDTA). Samples are centrifuged at 100,000 x g (Beckman Instruments, L8-M, SW-28 rotor, 21000 rpm) at 4 °C for 3 hours. The Lipid Body Fraction (LBF), floating on top of the overlay, is recovered with a spatula and transferred to a glass homogenizer (Potter-Elvehjem). Small amounts of LBF remaining in the centrifuge tube are recovered with a pipet by removing 4 ml of the Buffer B overlay and combining it with the LBF in the homogenizer. The final LBF is homogenized in 40 ml of Buffer B. The remaining fractions are collected as

follows: Interface fraction (the interface between the 0.3 and 0.5 M sucrose buffers), Soluble fraction (the liquid volume beneath the interface), and the Membrane fraction (a tan/brown pellet at the bottom of each tube). All are frozen and stored at -70 °C awaiting solubilization and further purification.

B. Solubilization of DAGAT Activity from the Lipid Body Fraction

The LBF is thawed on ice and solubilization is achieved by addition of Triton X-100 (Boehringer Mannheim, Mannheim, Germany) from a 10 % (w/v) stock to a final concentration of 1.3% (w/v). Solid sucrose (Mallinckrodt, Paris, Kentucky) is added to achieve a final concentration of 0.5M. The detergent-treated sample is rocked at 4 °C for one hour then divided into six ultracentrifuge tubes (25 x 89 mm, Beckman Instruments). Each tube is overlayed with 5 ml of Buffer B. Samples are centrifuged at 100,000 x g (Beckman Instruments, L8-M, SW-28 rotor, 21000 rpm) at 4 °C for 3 hours. The solubilized material, referred to as the 'Triton X-100 extract', is recovered by inserting a thin tube through the overlay to within 1 cm of the bottom of each ultracentrifuge tube and removing the lower, 0.5M sucrose, layer with gentle suction while leaving the upper 0.3M sucrose overlay (including a floating fat layer) and the pellet behind.

In the protocol described by Kamisaka et al. (1997), the Lipid Body Fraction was solubilized with 0.1% Triton X-100 and further centrifuged at 100,000 x g or filtered through a 0.2 µm filter. They found it necessary to increase the Triton X-100 concentration to 1.5% for DAGAT activity to bind the first column.

C. Chromatography used in the Purification of DAGAT from *M. ramanniana*

Buffer C, used for chromatography, contains 10 mM potassium phosphate (pH 7.0), 0.1% Triton X-100 (w/v) (Boehringer Mannheim, Mannheim, Germany), 10 % glycerol (w/v),

0.1 μM Aprotinin, 1 μM Leupeptin, 100 μM Pefabloc (all from Boehringer Mannheim, Mannheim, Germany) and varying amounts of potassium chloride (75-500 mM). This buffer differs from the corresponding column buffer used by Kamisaka et al.(1997) in that glycerol is substituted for ethylene glycol and EDTA, DTT, and PMSF are omitted while Aprotinin, Leupeptin and Pefabloc are included. Following the protocol by Kamisaka et al. (1997), a Yellow 86-Agarose (Sigma R-8504, St. Louis, MO) column is prepared (1.5 cm x 5.8 cm) and equilibrated with 150 mM KCl in Buffer C. We attempted to bind DAGAT activity, present in the Triton X-100 extract, to the Yellow 86-Agarose column under these conditions but found the majority of the DAGAT did not bind the column. We are able to bind a significant portion of the DAGAT activity to the column by diluting the KCl concentration of the applied sample to 75 mM with an equal volume of Buffer C (without KCl). In accordance, the Yellow 86-Agarose column is also equilibrated in 75 mM KCl in Buffer C. Following application of the sample at 0.56 ml/min, the column is washed with 4 column volumes of equilibration buffer. DAGAT activity and proteins bound to the column are eluted with 500 mM KCl in Buffer C (Figure 4).

DAGAT activity eluted from the Yellow 86-Agarose column (fractions 17-20) is diluted 1:3.33 with Buffer C to reduce the KCl concentration to 150 mM. The diluted pool (103 ml) is applied to a Heparin-Sepharose CL-6B column (Pharmacia, Uppsala, Sweden, 0.5 cm x 4.8 cm) equilibrated with 150 mM KCl in Buffer C at 0.2 ml/min. The column is washed with 5 volumes of equilibration buffer and DAGAT activity and protein are eluted in a 15 ml linear gradient of 150-500 mM KCl in Buffer C. DAGAT activity elutes in two overlapping peaks. The first peak elutes during the gradient, as found by Kamisaka et al. (1997) and a second peak, not found by Kamisaka et al., elutes at the end of the gradient with much less protein (Figure 5A).

A portion (250 μl) of the two peak fractions from the Heparin column are further purified by size exclusion chromatography on a Superdex-200 column (1 x 30 cm, Bio-Rad, Hercules, CA) at 0.2 ml/min equilibrated with 150 mM KCl in

Buffer C. For calibration only, the column is equilibrated with 150 mM KCl in a Modified Buffer C in which Triton X-100 is replaced with Triton X-100 R (Calbiochem, La Jolla, CA). The column is calibrated using Bio-Rad Gel Filtration Standards. The DAGAT activity from each of the two peaks from Heparin-Sepharose CL-6B elutes at an estimated molecular mass of 99 kDa.

Additional chromatography is performed on the later eluting peak from the Heparin column, which contained DAGAT at a higher specific activity. In this case, the second peak from the Heparin column (fractions 36-41) is diluted 1:6.6 with Buffer C to a volume of 46.7 ml. The sample is applied to a Yellow 86 Agarose column (1.0 cm x 6.4 cm) equilibrated with 75 mM KCl in Buffer C at 0.5 ml/min. After washing with 5 column volumes of equilibration buffer, bound proteins and all of the DAGAT activity elute in a 40 ml linear gradient of 75-500 mM KCl in Buffer C. DAGAT activity elutes as a single peak (Figure 6A).

The protein composition of the fractions containing DAGAT activity from the Heparin and second Yellow 86 columns are analyzed by gradient SDS-PAGE according to the protocol in Example 3. Protein bands are detected by silver-staining. The pattern of bands eluting from these columns is compared, fraction by fraction, to the respective DAGAT activity profile. Many protein candidates are present that correlate with the presence of DAGAT activity. It is our opinion that the purification protocol is insufficient to identify a particular protein candidate associated with DAGAT activity (Figure 5B, 6B).

Example 11 - New purification protocol for identifying DAGAT protein candidates purified from *Mortierella ramanniana*

A. Preparation of the Lipid Body Fraction

The following steps are performed at 4 °C.

Typically, 70-75 g of wet packed cells (stored at -70 °C) are used for each lipid body preparation. Just prior to use, cells are thawed on ice and resuspended in 150 ml of Buffer A (10 mM potassium phosphate (pH 7.0), 0.15 M KCl, 0.5 M

sucrose, 1 mM EDTA). The following protease inhibitors are added to reduce proteolysis: 0.1 μM Aprotinin, 1 μM Leupeptin, and 100 μM Pefabloc (all from Boehringer Mannheim, Germany). Samples are lysed with a cell disrupter (Bead-Beater, Biospec Products, Bartlesville, OK) using 0.5 mm glass beads. The sample chamber is filled with 180 ml of glass beads. Wet-packed cells are thawed on ice and resuspended in 150 ml of Buffer A. The cell slurry is poured over the glass beads. In general, an additional 40-50 ml of Buffer A are needed to fill the chamber for proper functioning. This volume is used to rinse the remains of the cell slurry from its original container so that it can be combined with the rest of the sample. Cells are ground ('Homogenize' setting) for 45-90 seconds depending on the viscosity of the sample. The cell slurry containing glass beads is divided into tubes (29 x 104 mm) and centrifuged at 500 x g (Beckman Instruments, GP centrifuge, GH 3.7 Horizontal rotor at 1500 rpm) and 4 °C. The supernatant is removed and the pellets washed with another 5 ml of Buffer A. Following centrifugation the supernatant volumes are combined. This fraction is referred to as the 'S1'. The S1 is divided into six ultracentrifuge tubes (25 x 89 mm, Beckman Instruments) and each is overlayed with 5 ml of Modified Buffer B (10 mM potassium phosphate pH, 7.0, 0.15 M KCl, and 0.3 M sucrose). EDTA is omitted from Buffer B (see Example 4) since it interferes with hydroxylapatite chromatography. Samples are centrifuged at 100,000 x g (Beckman Instruments, L8-M, SW-28 rotor, 21000 rpm) at 4 °C for 3 hours. The Lipid Body Fraction (LBF), floating on top of the overlay, is recovered with a spatula and transferred to a glass homogenizer. Small amounts of LBF remaining in the centrifuge tube are recovered with a pipet by removing 4 ml of the Buffer B overlay and combining it with the LBF in the homogenizer. The final LBF is homogenized in 40 ml of Buffer B. The remaining fractions are collected as follows: Interface fraction (the interface between the 0.3 and 0.5 M sucrose buffers), Soluble fraction (the liquid volume beneath the interface), and the Membrane fraction (a tan/brown pellet

at the bottom of each tube). All are frozen and stored at -70 °C awaiting solubilization and further purification.

B. Solubilization of DAGAT Activity from the Lipid Body Fraction

Prior to solubilization, a protein determination is made with an aliquot of the Lipid Body Fraction by the method of Bradford (Bio-Rad Reagent, Hercules, CA) using bovine serum albumin as a standard. The LBF is thawed on ice, then diluted to a concentration of 1 mg protein/ml and treated with Triton X-100 at a detergent to protein ratio of 15:1 (w/w, equivalent to 1.3% Triton X-100). Solid sucrose (Mallinckrodt, Paris, Kentucky) is added to achieve a final concentration of 0.5M. The detergent-treated sample is rocked at 4 °C for one hour then divided into six ultracentrifuge tubes (25 x 89 mm, Beckman Instruments). Each tube is overlayed with 5 ml of Modified Buffer B. Samples are centrifuged at 100,000 x g (Beckman Instruments, L-8M, SW-28 rotor, 21000 rpm) at 4 °C for 3 hours. The solubilized material, referred to as the 'Triton X-100 extract', is recovered by inserting a thin tube through the overlay to within 1 cm of the bottom of each ultracentrifuge tube and removing the lower, 0.5M sucrose, layer with gentle suction while leaving the upper 0.3M sucrose overlay (including a floating fat layer) and the pellet behind.

C. DAGAT Column Chromotography

Our previous experience purifying acyltransferase proteins from plant species is applied to the case of DAGAT purification from *Mortierella ramanniana*. We have been successful in using hydroxylapatite chromatography in the purification of wax synthase from jojoba (*Simmondsia chinensis*, WO Publication 95/15387, the entirety of which is incorporated herein by reference) and lysophosphatidic acid acyltransferase (LPAAT) from coconut (*Cocos nucifera*) (US Patent Application 08/231,196, the entirety of which is incorporated herein by reference). This purification step is

introduced after the Yellow 86-Agarose column. The purification protocol of Yellow 86-Agarose followed by hydroxylapatite is performed in two ways. In Protocol A, activity is bound to the first column and after elution, fractions are assayed for activity. The active fractions are then pooled and applied to the second column. We refer to this as a sequential run. In Protocol B, activity is bound to the first column then elutes and flows directly onto the second column without pooling and assaying in between. We refer to this as a tandem run.

In Protocol A, the Triton X-100 extract is applied to a Yellow 86-Agarose column (2.5 cm x 6.4 cm) equilibrated with 75 mM KCl in Buffer C (Example 4.C) at 2 ml/min. The column is washed with 5 column volumes of equilibration buffer then eluted with 500 mM KCl in Buffer C at 0.5 ml/min (Figure 7). The two most active fractions (64 and 65), containing 93% of the eluted activity, are pooled and loaded onto a hydroxylapatite column (Bio-Gel HT, Bio-Rad, 1 cm x 25.5 cm) equilibrated with 500 mM KCl in Buffer C at 0.5 ml/min. DAGAT activity flows through the column whereas the majority of the proteins bind the column. The column is washed with 3 volumes of equilibration buffer. Bound proteins are eluted with 100 mM dipotassium phosphate and 500 mM KCl in Buffer C at 0.5 ml/min (Figure 8A). A portion of the fractions containing the DAGAT activity peak are run on gradient gel SDS-PAGE as described in Example 3. The proteins are stained with silver and the pattern of the bands are compared, fraction by fraction, to the activity profile (Figure 8B). Several DAGAT protein candidates correlate with activity. In particular, attention is called to bands migrating at positions corresponding to 43 kD, 36.5 kD, 33 kDa, 29 kD, 28 kD and 27 kD. There does not appear to be a candidate protein in the region of 53 kD that correlates with activity.

In Protocol B, the Triton X-100 extract is applied to a Yellow 86-Agarose column (1.5 cm x 5.8 cm) equilibrated with 75 mM KCl in Buffer C at 1 ml/min. The column is washed with 5 column volumes of equilibration buffer. Then, the outlet from the Yellow 86-Agarose column is connected to the inlet of

a hydroxylapatite column (1.0 cm x 26.2 cm, Bio-Rad, Hercules, CA) equilibrated with 500 mM KCl in Buffer C. DAGAT activity bound to the Yellow 86 column is eluted with 110 ml of Buffer C containing 500 mM KCl and passes directly through the hydroxylapatite column at 0.2 ml/min. Finally, the hydroxylapatite column is disconnected from the Yellow 86-Agarose column and proteins bound to the hydroxylapatite column are eluted with 100 mM dipotassium phosphate and 500 mM KCl in Buffer C. DAGAT activity is found in fractions from the hydroxylapatite column collected during the 110-ml wash with Buffer C containing 500 mM KCl.

The majority of the protein in the Triton X-100 extract does not bind the Yellow 86-Agarose column and is discarded. A small subset of proteins, including DAGAT, do bind the Yellow 86-Agarose column and are eluted with 500 mM KCl in Buffer C. When this eluate is applied to the hydroxylapatite column, DAGAT activity flows through while most of the remaining proteins bind the column and are separated (Figure 9A). A portion of the fractions containing the DAGAT activity peak are run on gradient gel SDS-PAGE and are silver-stained. The pattern of bands eluting from these columns is compared, fraction by fraction, to the respective DAGAT activity profile. Examination of the stained protein bands indicate a protein at 33 kDa correlates best with DAGAT activity (Figure 9B).

Example 12 - Preparation of Protein for In-Gel Digestion

After a protein candidate has been identified, it is necessary to prepare sufficient amounts for sequencing. Protein sequencing can be performed using a wide variety of methods known in the art. One technique involves digestion of the protein, using enzymes such as trypsin, while still in an SDS-polyacrylamide gel. Several commercial enterprises known in the art have established protocols for obtaining peptides in this manner. Following the generation of peptides, standard techniques are employed to separate and sequence them.

In order to gel-purify a protein candidate, it is often necessary to concentrate the liquid sample first so that it can be loaded on the gel. Samples containing high amounts of detergent may pose special problems. Depending on the micelle size of the detergent, it may concentrate during ultrafiltration and pose problems during electrophoresis. An alternative method of concentrating the protein sample must then be employed.

A. Preparation of Samples for SDS-PAGE by Concentration

Fractions can be concentrated in a pressure cell fitted with a membrane of the appropriate molecular weight retention limit. Alternatively, the sample may be concentrated using filtration by centrifugation in individual units, for example a product such as Centricon-30 (Amicon, Inc., Beverly, MA), to volumes of approximately 50µl. Following concentration, samples can be treated with a loading buffer, for example, Laemmli.

B. Preparation of Samples for SDS-PAGE by Precipitation

Sometimes it is desirable to concentrate samples by precipitation. This can be achieved using acid and/or acetone. A typical protocol would be to add trichloroacetic acid (TCA) from a concentrated stock (40%-50%) to a final concentration of 7-10%. After about 10 minutes on ice the samples are centrifuged (12,000 x g, 15 minutes at 4 C) to pellet the precipitated protein. The supernatants are removed and in order to remove the precipitated detergent, the pellets are washed with ice cold acetone and centrifuged again. Precipitates can be resuspended with a sample loading buffer (ie. Laemmli or SDS Cocktail as in Example 12 A). SDS-PAGE may be performed using gels cast in the laboratory, as described in example 9 or from gels prepared by commercial sources.

C. SDS-PAGE

Heating of the samples prior to loading the gel may or may not be performed. It has been observed that some membrane proteins have a tendency to aggregate upon heating. In this case, samples are generally applied to the gel after sitting at room temperature for 15 minutes. Acrylamide gels may be purchased commercially or prepared in the laboratory. One

protocol for preparing 10-13% acrylamide gels is described in Example 9. Following electrophoresis, the gel can be stained with 0.1% Coomassie Blue in 50% methanol, 10% acetic acid then destained. Destaining can be accomplished with the use of a commercial product, such as Gel-Clear (Novex, San Diego, Ca) or in 50% methanol, 10% acetic acid. Protein candidates can then be excised from the gel and sent for In-Gel digestion with or without further destaining.

Example 13 - Determination of Amino Acid Sequence

Commercial facilities have been established which provide protein sequencing as a service. Among the techniques which are available, the generation of peptides by In-Gel Digestion using an endopeptidase, such as trypsin, followed by HPLC purification, has proved the most useful. N-terminal sequencing on PVDF, and to a lesser degree the generation of peptides by limited cyanogen bromide treatment of the PVDF proteins, has also proved successful. Procedures for in-gel digestion may include amino acid analysis of a portion (10-15%) of the gel slice for quantitation and amino acid composition, digestion of the protein with one of the proteolytic enzymes (trypsin or lysyl endopeptidase), and fractionation of the products by reverse phase HPLC. Absorbance peaks may be selected from the HPLC run and subjected to laser desorption mass spectrometry to determine the presence, amount, and mass of the peptide prior to protein sequencing. The longest peptides are selected for microsequencing.

Example 14 Isolation of *Mortierella ramanniana* DAGAT Nucleic Acid Sequences

In general, for use as PCR primers from single stranded DNA template reverse-transcribed from mRNA, oligonucleotides containing the sense orientation sequence corresponding to DAGAT peptide encoding sequences are prepared. These oligonucleotides are used as primers for the "forward" amplification reaction to produce sense strand DNA.

For the "reverse" reaction for amplification of the non-coding DNA strand, an oligonucleotide may be designed to be identical to a portion of a primer used to prepare DNA template for PCR. Alternatively, oligonucleotides which contain sequence complementary to DAGAT peptide encoding sequences may be used in combination with a "forward" DAGAT oligonucleotide primer as described above.

Where the DAGAT peptide sequences contain amino acids which may be encoded by a number of different codons, the forward or reverse primers may be "degenerate" oligonucleotides, i.e. containing a mixture of all or some of the possible encoding sequences for a particular peptide region. To reduce the number of different oligonucleotides present in such a mixture, it is preferable to select peptide regions which have the least number of possible encoding sequences when preparing the synthetic oligonucleotide for PCR primers. Similarly, where the synthetic oligonucleotide is to be used to directly screen a library for DAGAT sequences, lower degeneracy oligonucleotides are preferred.

DAGAT DNA fragments obtained by PCR are labeled and used as a probe to screen clones from cDNA libraries. Complementary DNA and DNA constructioin and library screening techniques are known to those in the art and described, for example in Maniatis *et al.* (*Molecular Cloning: A Laboratory Manual, Second Edition* (1989) Cold Spring Harbor Laboratory Press). In this manner, DAGAT nucleic acid sequences are obtained which may be analyzed for nucleic acid sequence and used for expression of DAGAT in various hosts, both procaryotic and eucaryotic.

Example 15 - *Mortierella ramanniana* DAGAT Constructs for Plant Expression

Constructs which provide for expression of DAGAT sequences in plant cells may be prepared as follows.

Expression cassettes which contain 5' and 3' regulatory regions from genes expressed preferentially in seed tissues may be prepared from napin, Bce4 and ACP genes as described, for example in WO 92/03564.

For example, a napin expression cassette, pCGN1808, which may be used for expression of wax synthase or reductase gene constructs is described in Kridl *et al.* (*Seed Science Research* (1991) *1*:209-219). An additional napin expression cassette, pCGN3223, contains an ampicillin resistance background, and essentially identical 1.725 napin 5′ and 1.265 3′ regulatory sequences as found in pCGN1808. The regulatory regions are flanked with *Hin*dIII, *Not*I and *Kpn*I restriction sites and unique *Sal*I, *Bgl*II, *Pst*I, and *Xho*I cloning sites are located between the 5′ and 3′ noncoding regions.

A cassette for cloning of sequences for transcriptional regulation under the control of 5′ and 3′ regions from an oleosin gene may also be used. Sequence of a *Brassica napus* oleosin gene was reported by Lee and Huang (*Plant Phys.* (1991) *96*:1395-1397). Sequence of an oleosin cassette, pCGN7636, is provided in Figure 4 of USPN 5,445,947. The oleosin cassette is flanked by *Bss*HII, *Kpn*I and *Xba*I restriction sites, and contains *Sal*I, *Bam*HI and *Pst*I sites for insertion of wax synthase, reductase, or other DNA sequences of interest between the 5′ and 3′ oleosin regions.

DAGAT gene sequences may be inserted into such cassettes to provide expression constructs for plant transformation methods. For example, a construct for expression of reductase in plant cells using 5′ and 3′ regulatory regions from a napin gene, is described in USPN 5,445,947.

Binary vector constructs are transformed into *Agrobacterium* cells, such as of strain EHA101 (Hood *et al.*, *J. Bacteriol* (1986) *168*:1291-1301), by the method of Holsters *et al.* (*Mol. Gen. Genet.* (1978) *163*:181-187) and used in plant transformation methods as described below.

Example 16 Identification of Additional Wax Synthase Sequences

The protein sequence of the jojoba wax synthase (Figure 11) is used to query the Arabidopsis DNA sequence database (http://genome-www.stanford.edu/Arabidopsis/). One of the accessions, P1 clone MTE17(Genbank accession AB015479), contains 7 repeats of open reading frames with similarity to

the jojoba wax synthase. The open reading frames have been designated ATWS1 to ATWS7 (Figures 12-18, respectively). They are found between nucleotides 23670 and 11479 of MTE17 using the numbering system of the Genbank entry. The inferred protein sequences are aligned with the jojoba wax synthase sequence (Figure 19) and a dendogram (Figure 20) of their relationships is constructed using the Clustal W algorithm of MacVector 6.5 (Oxford Molecular). The sequence alignment in Figure 19 shows a series of peptide sequences which are conserved between the amino acid sequences (Table 5). The percent identities and similarities are also determined, and are presented in table 6.

Table 5

	SEQUENCE	SEQ ID NO:
1	LXLF(A/S)(F/L)XX(G/E)	
2	PYL(A/S)TSL(Q/H)(D/E)FW(G/S)(R/H)RWNL(M/I)V	
3	FX(V/T)SGXXHEXX(F/Y)FYX(I/T)R	
4	P(S/T)(W/G)EV(T/A)XFF(V/L)LHG	

Table 6

SEQUENCES COMPARED	PERCENT IDENTITY	PERCENT SIMILARITY
JoWS vs ATWS1	41%	17%
JoWS vs ATWS2	37%	19%
JoWS vs ATWS3	42%	15%
JoWS vs ATWS4	42%	16%
JoWS vs ATWS5	44%	13%
JoWS vs ATWS6	41%	17%
JoWS vs ATWS7	36%	16%
ATWS3 vs ATWS6	59%	14%
ATWS3 vs ATWS1	49%	17%
ATWS3 vs ATWS4	71%	10%
ATWS3 vs ATWS7	55%	15%
ATWS3 vs ATWS2	49%	21%
ATWS3 vs ATWS5	64%	14%
ATWS6 vs ATWS1	47%	18%
ATWS6 vs ATWS4	57%	15%
ATWS6 vs ATWS7	51%	17%
ATWS6 vs ATWS2	45%	18%
ATWS6 vs ATWS5	57%	15%
ATWS1 vs ATWS4	52%	17%
ATWS1 vs ATWS7	46%	19%
ATWS1 vs ATWS2	65%	12%
ATWS1 vs ATWS5	49%	17%
ATWS4 vs ATWS7	58%	13%
ATWS4 vs ATWS2	50%	19%
ATWS4 vs ATWS5	65%	13%
ATWS7 vs ATWS2	45%	18%
ATWS7 vs ATWS5	55%	16%
ATWS2 vs ATWS5	49%	17%

Complementary DNA (cDNA) is constructed from Arabidopsis RNA isolated from immature seeds, whole seedlings (vegetative tissue), and inflorescences (flowers and flower stalks) using

the SMART PCR cDNA Library construction kit according to the manufacturer's protocol (Clontech). SMART cDNA is also constructed from RNA from Brassica napus leaves, and immature seeds harvested at 15 days after pollination (DAP), 18 DAP, and 30 DAP. The SMART cDNAs are used for virtual Northern analysis, according to the protocol in the SMART cDNA manual from Clontech, of expression of the Arabidopsis ATWS cDNAs. Two of the sequences, ATWS1 and ATWS2, are the most highly expressed. ATWS2 is most highly expressed in Arabidopsis immature seeds and Brassica 30 DAP seeds. Expression is not detected in Brassica leaves or Arabidopsis seedlings. This is an expression pattern consistent with that expected for DAGAT, since triglycerides are primarily formed in developing seeds of these plants.

Example 17 Expression Constructs for AT-WS Homologues

To characterize the *Arabidopsis* sequence which is preferentially expressed in the developing seeds, constructs are prepared to direct the expression of the sequence in host plant cells.

A plasmid containing the napin cassette derived from pCGN3223 (described in USPN 5,639,790, the entirety of which is incorporated herein by reference) was modified to make it more useful for cloning large DNA fragments containing multiple restriction sites, and to allow the cloning of multiple napin fusion genes into plant binary transformation vectors. An adapter comprised of the self annealed oligonucleotide of sequence CGCGATTTAAATGGCGCGCCCTGCAGGCGGCCGCCTGCAGGGCGCGCCATTTAAAT was ligated into the cloning vector pBC SK+ (Stratagene) after digestion with the restriction endonuclease BssHII to construct vector pCGN7765. Plamids pCGN3223 and pCGN7765 were digested with NotI and ligated together. The resultant vector, pCGN7770, contains the pCGN7765 backbone with the napin seed specific expression cassette from pCGN3223.

The cloning cassette, pCGN7787, essentially the same regulatory elements as pCGN7770, with the exception of the

napin regulatory regions of pCGN7770 have been replaced with the double CAMV 35S promoter and the tml polyadenylation and transcriptional termination region.

A binary vector for plant transformation, pCGN5139, was constructed from pCGN1558 (McBride and Summerfelt, (1990) Plant Molecular Biology, 14:269-276). The polylinker of pCGN1558 was replaced as a HindIII/Asp718 fragment with a polylinker containing unique restriction endonuclease sites, AscI, PacI, XbaI, SwaI, BamHI,and NotI. The Asp718 and HindIII restriction endonuclease sites are retained in pCGN5139.

A series of turbo binary vectors are constructed to allow for the rapid cloning of DNA sequences into binary vectors containing transcriptional initiation regions (promoters) and transcriptional termination regions.

The plasmid pCGN8618 was constructed by ligating oligonucleotides 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' and 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3'into SalI/XhoI-digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3' region was excised from pCGN8618 by digestion with Asp718I; the fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8622.

The plasmid pCGN8619 was constructed by ligating oligonucleotides 5'- TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC -3' and 5'- TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3'into SalI/XhoI-digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3' region was removed from pCGN8619 by digestion with Asp718I; the fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII

and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8623.

The plasmid pCGN8620 was constructed by ligating oligonucleotides 5'- TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGGAGCT -3' and 5'-CCTGCAGGAAGCTTGCGGCCGCGGATCC-3' into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region was removed from pCGN8620 by complete digestion with Asp718I and partial digestion with NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8624.

The plasmid pCGN8621 was constructed by ligating oligonucleotides 5'- TCGACCTGCAGGAAGCTTGCGGCCGCGGATCCAGCT -3' and 5'-GGATCCGCGGCCGCAAGCTTCCTGCAGG-3' into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region was removed from pCGN8621 by complete digestion with Asp718I and partial digestion with NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the

integrity of cloning junctions. The resulting plasmid was designated pCGN8625.

The open reading frame of AT-WS2 was PCR amplified from Arabidopsis genomic DNA using the primers 5'-GGATCCGCGGCC GCATTATGAAACAGTTAGCAACCAACAGA-3' and 5'-GGATCCCCTGCAGGTTACAT TAAAATACAGACAACGTGCC-3'. The PCR product is cloned into plasmid pCR 2.1 according to the manufacturer's protocol (Clontech) to generate the plasmid pCGN9706. To direct transcription of a sense transcript under control of a napin expression cassette in transgenic plants, plasmid pCGN9712 is constructed by cloning the NotI/Sse8387I fragment from pCGN9706 into NotI/PstI digested binary vector pCGN8622. To direct transcription of an antisense transcript under control of a napin expression cassette in transgenic plants, plasmid pCGN9713 was constructed by cloning the NotI/Sse8387I fragment from pCGN9706 into NotI/PstI digested binary vector pCGN8623. To direct transcription of a sense transcript under control of a double 35S expression cassette in transgenic plants, plasmid pCGN9714 was constructed by cloning the NotI/Sse8387I fragment from pCGN9706 into NotI/PstI digested binary vector pCGN8624. Plasmids pCGN9712, pCGN9713, and pCGN9714 were introduced into Agrobacterium tumefaciens EHA105 by electroporation, and the resultant Agrabacterium strains were used to transform Arabidopsis thaliana plants by vacuum infiltration. The vector pCGN9710 is similar to pCGN9714, only the coding sequence of ATWS2 from pCGN9714 has been replaced with the coding sequence of ATWS1.

Example 18 - Plant Transformation Methods and Analyses

A variety of methods have been developed to insert a DNA sequence of interest into the genome of a plant host to obtain the transcription or transcription and translation of the sequence to effect phenotypic changes.

High erucic acid varieties, such as cultivar Reston, or Canola-type varieties of *Brassica napus* may be transformed using *Agrobacterium* mediated transformation methods as described by Radke *et al.* (*Theor. Appl. Genet.* (1988) *75*:685-694; *Plant Cell Reports* (1992) *11*:499-505). Transgenic *Arabidopsis thaliana* plants may be obtained by *Agrobacterium*-mediated transformation as described by Valverkens *et al.*, (*Proc. Nat. Acad. Sci.* (1988) *85*:5536-5540), or as described by Bent et al. ((1994), *Science* 265:1856-1860), or Bechtold et al. ((1993), *C.R.Acad.Sci, Life Sciences* 316:1194-1199). Other plant species may be similarly transformed using related techniques.

Alternatively, microprojectile bombardment methods, such as described by Klein *et al.* (*Bio/Technology 10*:286-291) may also be used to obtain transformed plants comprising the reductase and wax synthase expression constructs described herein.

Arabidopsis seedlings transformed with pCGN9710 (ATWS1) and pCGN9714 (ATWS2), where gene expression was driven by a 35S promoter, were assayed for DAGAT activity. Seedlings were homogenized in 100 µl buffer (100 mM Tricine/NaOH, pH 7.8, 280 mM NaCl, 10 % glycerol, and protease inhibitors 0.1 µM Aprotinin, 1 µM Leupeptin, and 100 µM Pefabloc) and solids were pelleted by centrifugation at 16,000 xg, for 10 min and 4° C. The supernatant was removed and the pellet was resuspended in 100 µl homogenization buffer. DAGAT assays were performed as described in Example 7A with the following modifications. We determined that inclusion of 0.1% Triton X-100 in the assay cocktail was detrimental to DAGAT activity in this tissue therefore it was omitted. Also, since the tissue was isolated in homogenization buffer above, the assay was performed in the same buffer. 25 µl of sample was assayed in a total volume of 100 µl containing the following components: 3.67 µM

1-^{14}C-18:1-Coenzyme A (53.5-54.5 Ci/mole, New England Nuclear, Boston, MA) 100mM Tricine/NaOH, pH 7.8, 100mM NaCl, 1.5mM di-18:1-DAG). One to five control seedlings, Arabidopsis cultivar No-O, were assayed for background DAGAT activity. For comparison purposes the amount of tissue in two to three No-O seedling was equivalent to the transformed samples, therefore the background activity was determined as the average of these two samples. Of the transgenic plants analyzed, none of the pCGN9710 (ATWS1) plants demonstrated an activity at least 1.5 fold the control average. Assays of plants expressing pCGN9714 (ATWS2) showed 12/25 plants that exceeded the background level by 1.5 fold and three plants exceeded the background level by 2.0 fold or more (Figure 21). These data indicate that ATWS2 demonstrates DAGAT activity when expressed in Arabidopsis leaves.

The above results demonstrate the ability to obtain partially purified DAGAT proteins which are active in the formation of triacylglycerols from fatty acyl and diacylglycerol substrates. Methods to obtain the DAGAT proteins and amino acid sequences thereof are provided. In addition DAGAT nucleic acid sequences may also be obtained from the amino acid sequences using PCR and library screening methods provided herein. Such nucleic acid sequences may be manipulated to provide for transcription of the sequences and/or expression of DAGAT proteins in host cells, which proteins may be used for a variety of applications. Such applications include the modification of triacylglycerols levels and compositions in host cells.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be readily apparent to those of ordinary skill in the art in light of the teaching of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

CLAIMS

What is claimed is:

1. A recombinant DNA construct comprising a nucleic acid sequence which encodes at least a portion of an acyltransferase active in the formation of triacylglycerol, and a heterologous DNA sequence not naturally associated with said acyltransferase encoding sequence.

2. The construct of Claim 1 wherein said triacylglycerol is formed from a fatty acyl and a *sn*1,2-diacylglycerol substrate.

3. The construct of Claim 1 wherein said nucleic acid sequence is in an antisense orientation.

4. The construct of Claim 1 wherein said acyltransferase is active toward a fatty acyl substrate having a carbon chain selected from the group 8-24.

5. The construct of Claim 1 wherein said acyltransferase comprises the peptide sequence of LXLF(A/S)(F/L)X'X''(G/E), whereby X is an L, F, or I, and X' is a D, E, or H, and X'' is Q, K, E or L.

6. The construct of Claim 1 wherein said acyltransferase comprises the peptide sequence of PYL(A/S)TSL(Q/H)(D/E)FW(G/S)(R/H)RWNL(M/I)V.

7. The construct of Claim 1 wherein said acyltransferase comprises the peptide sequence of FX(V/T)SGX'X''HEX''X'''(F/Y)FYX''''(I/T)R whereby X is L, I, F, V, or T, and X' is L,V, or A, X''' is M, V, or A, X'''' is L, V, M, or I, and X''''' is L, M, V, I, or F.

8. The construct of Claim 1 wherein said acyltransferase comprises the peptide sequence of P(S/T)(W/G)EV(T/A)XFF(V/L)LHG whereby X is L, W, C, or G.

9. The construct of Claim 1 wherein said acyltransferase has greater than about 50% homologous similarity at the amino acid level to sequence of Figure 11.

10. The construct of Claim 1 wherein said acyltransferase encoding sequence is from a seed plant.

11. The construct of Claim 1 wherein said acyltransferase encoding sequence is from jojoba.

12. The construct of Claim 1 wherein said acyltransferase encoding sequence is from *Arabidopsis*.

13. The construct of Claim 1 further comprising a promoter which provides for transcription of said acyltransferase encoding sequence in a host cell.

14. The construct of Claim 13 wherein said promoter provides for expression of said acyltransferase encoding sequence in a plant cell.

15. The construct of Claim 14 wherein said plant cell is a plant seed cell.

16. The construct of Claim 13 wherein said promoter provides for expression of said acyltransferase encoding sequence in a bacterial cell.

17. The construct of Claim 14 wherein said promoter is from a gene preferentially expressed in a plant seed cell.

18. The construct of Claim 14 wherein said promoter is from a gene preferentially expressed in a plant seed embryo cell.

19. The construct of Claim 14 wherein said promoter is from a gene preferentially expressed in a plant pollen cell.

20. The construct of Claim 14 wherein said promoter is from a gene preferentially expressed in a leaf cell.

21. A cell comprising a construct according to Claim 1.

22. A plant cell comprising a construct according to Claim 1.

23. A method of producing a triacylglycerol in a host cell comprising the steps of

growing a host cell having a recombinant construct comprising a nucleic acid sequence which encodes at least a portion of an acyltransferase active in the formation of a triacylglycerol said acyltransferase encoding sequence under the control of regulatory elements functional in said cell, under conditions which will cause the expression of said acyltransferase,

wherein said host cell comprises a fatty acid substrate of said acyltransferase, whereby triacylglycerol is produced in said host cell.

24. The method according to Claim 23 wherein said triacylglycerol is produced from a fatty acid and a *sn*1,2-diacylglycerol substrate.

25. The method of Claim 23 wherein said host cell is a procaryote.

26. The method of Claim 23 wherein said host cell is a plant cell.

27. The method of Claim 23 wherein said regulatory elements is a promoter functional in a plant cell.

28. The method according to Claim 27 wherein said promoter is preferentially expressed in a specific plant tissue.

29. The method according to Claim 27 wherein said promoter is preferentially expressed in a plant seed cell.

30. The method according to Claim 27 wherein said promoter is preferentially expressed in a plant seed embryo cell.

31. The method according to Claim 27 wherein said promoter is preferentially expressed in a plant leaf cell.

32. The method according to Claim 27 wherein said promoter is preferentially expressed in a plant pollen cell.

33. The method of Claim 26 wherein said plant cell is from *Brassica*.

34. The method of Claim 23 wherein said acyltransferase encoding sequence is from a seed plant.

35. The method according to Claim 37 wherein said encoding sequence is in the antisense orientation.

36. The method of Claim 23, wherein triacylglycerol is produced results in an increase in oil as a component of said host cell.

37. A host cell comprising a triacylglycerol produced according to the method of Claim 23.

38. An oil obtained from a cell according to Claim 37.

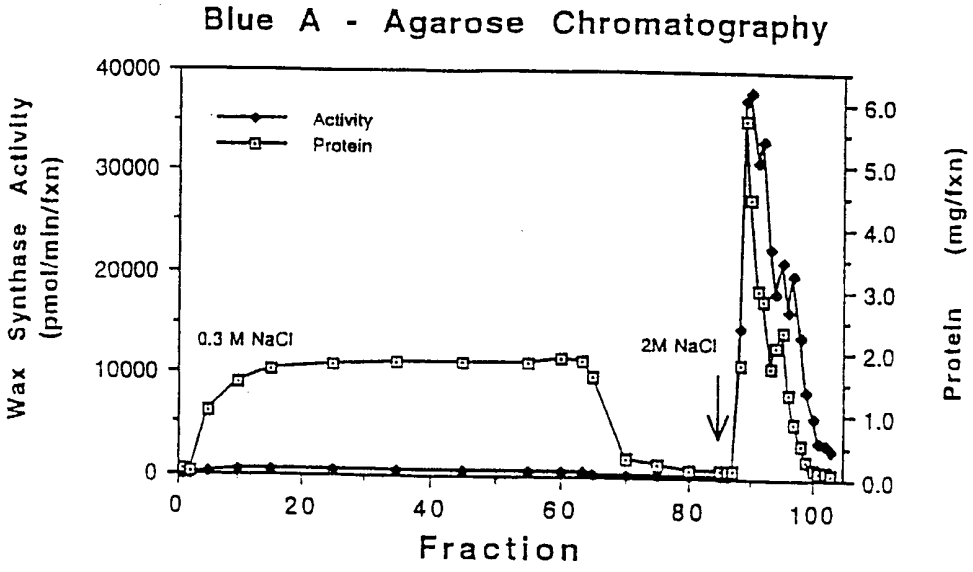


Figure 1 1/4

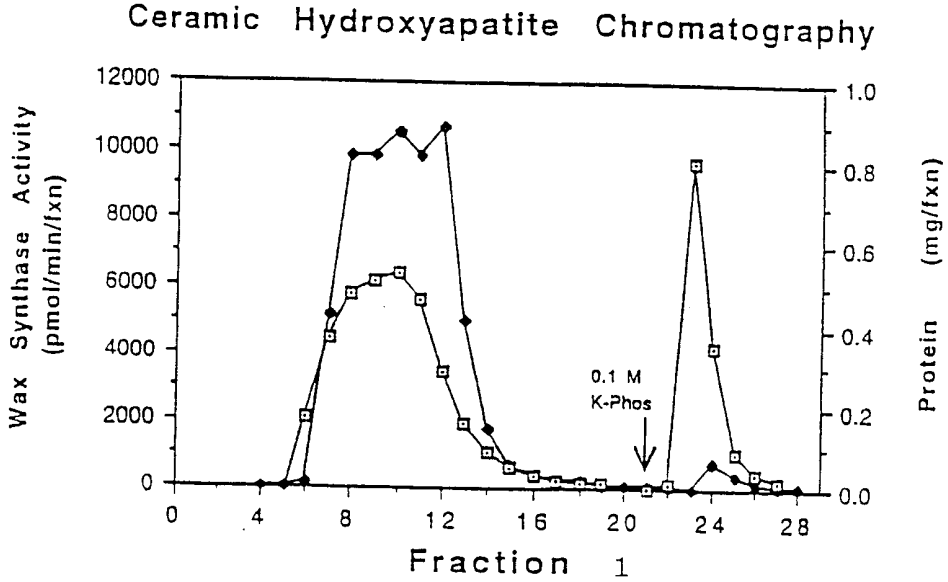


Figure 1 2/4

1/22

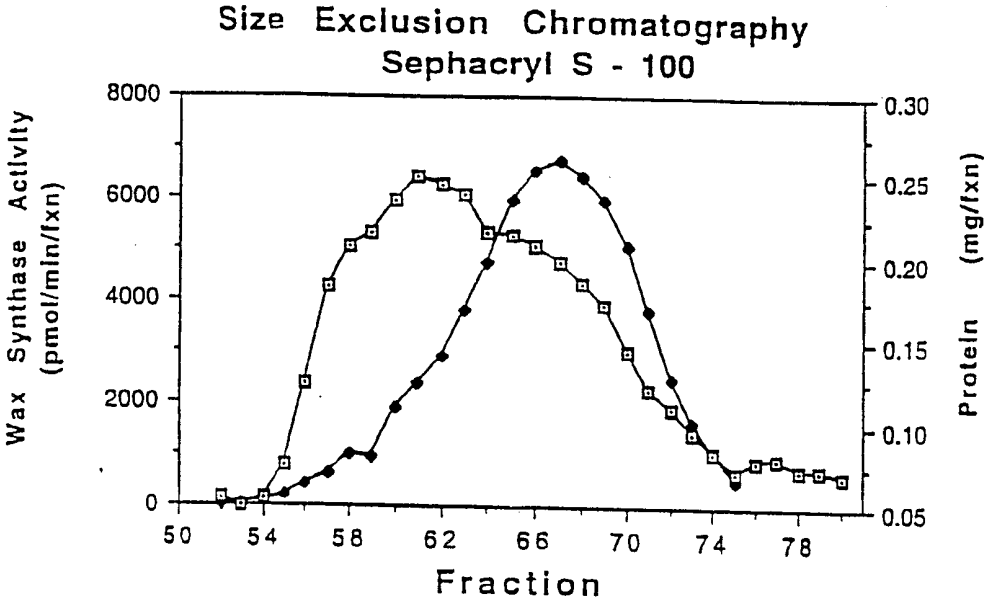


Figure 1 3/4

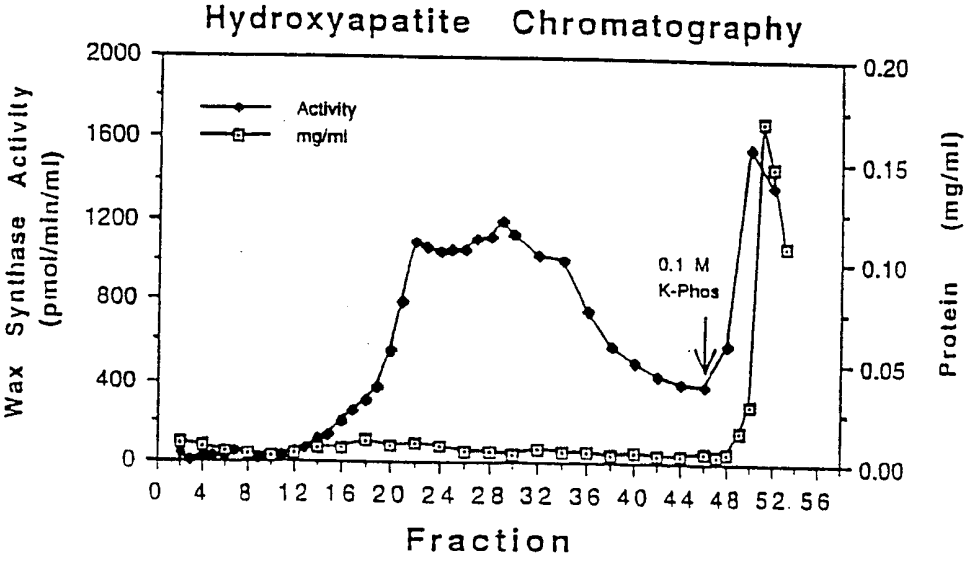


Figure 1 4/4

2/22

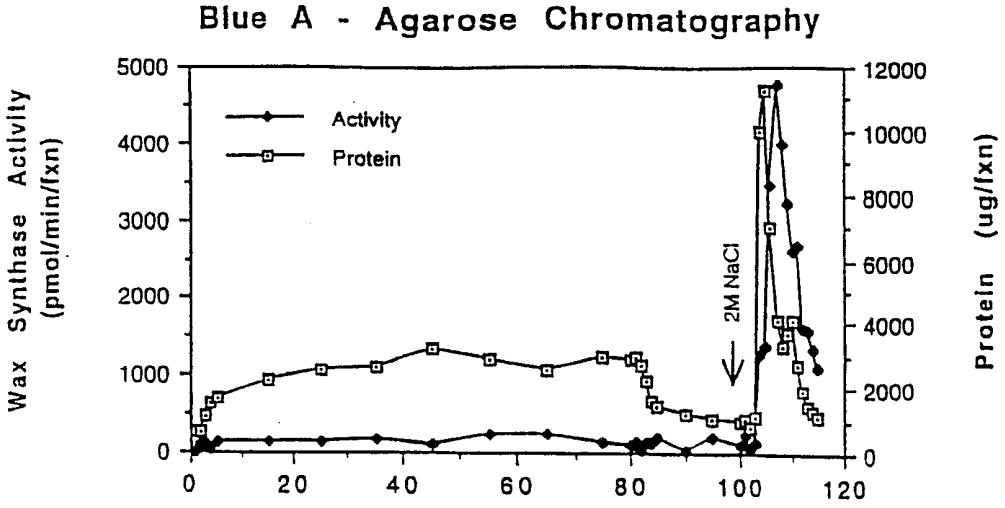


Figure 2 1/3

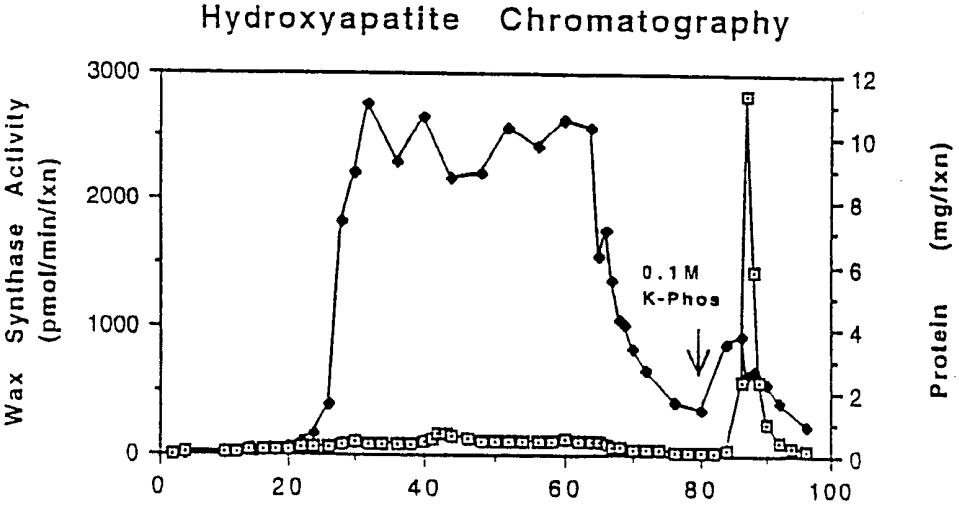
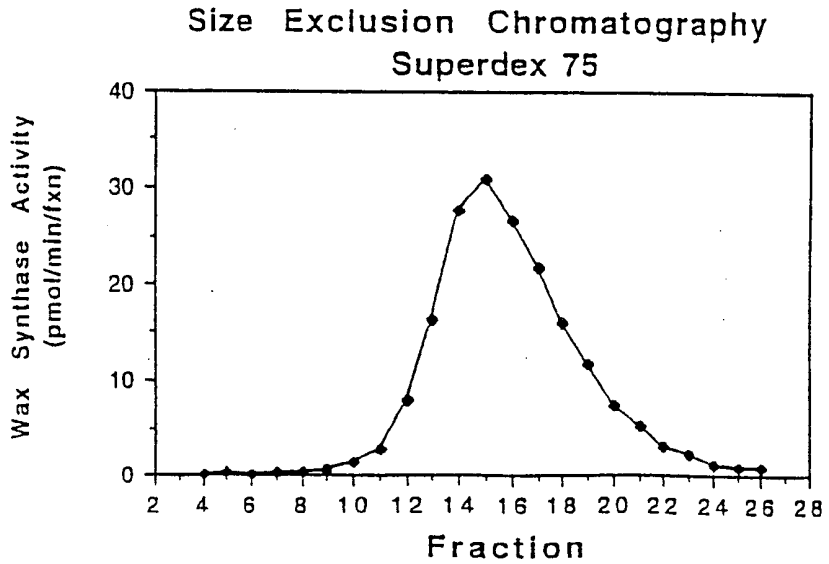


Figure 2 2/3



3/22

Figure 2 3/3

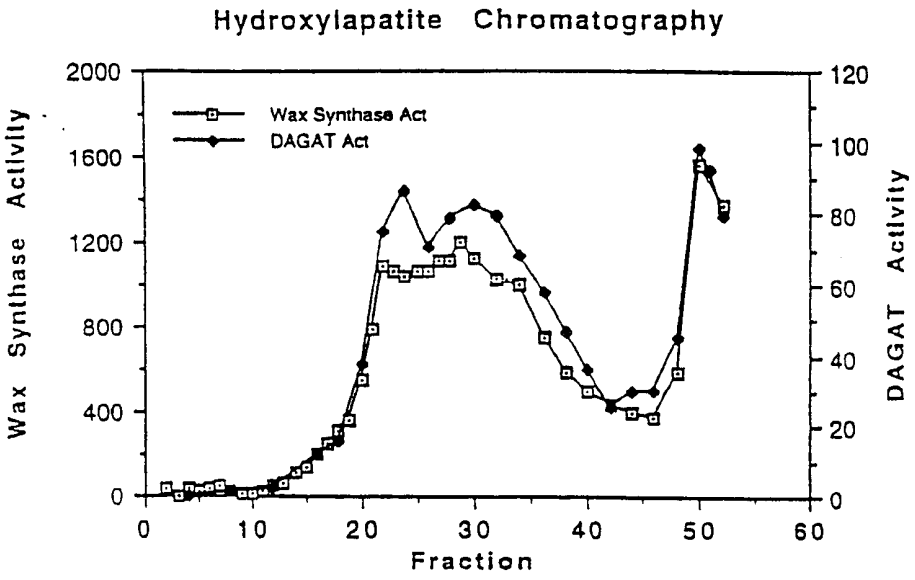

Hydroxylapatite Chromatography
Wax Synthase Act
DAGAT Act
Wax Synthase Activity
DAGAT Activity
Fraction
2000
1600
1200
800
400
0
120
100
80
60
40
20
0
0
10
20
30
40
50
60

Figure 3

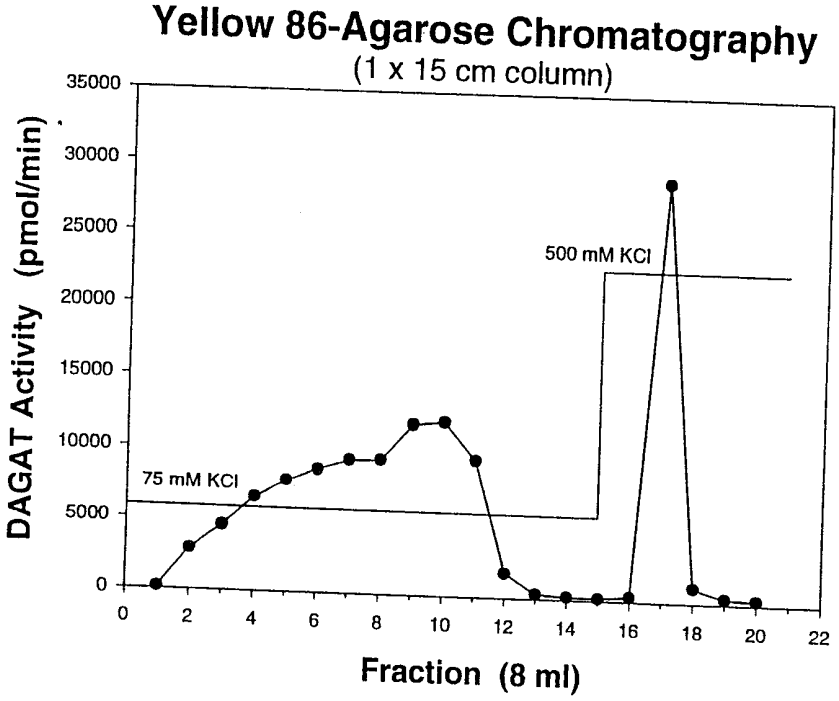

Yellow 86-Agarose Chromatography
(1 x 15 cm column)
DAGAT Activity (pmol/min)
35000
30000
25000
20000
15000
10000
5000
0
75 mM KCl
500 mM KCl
0
2
4
6
8
10
12
14
16
18
20
22
Fraction (8 ml)

FIGURE 4

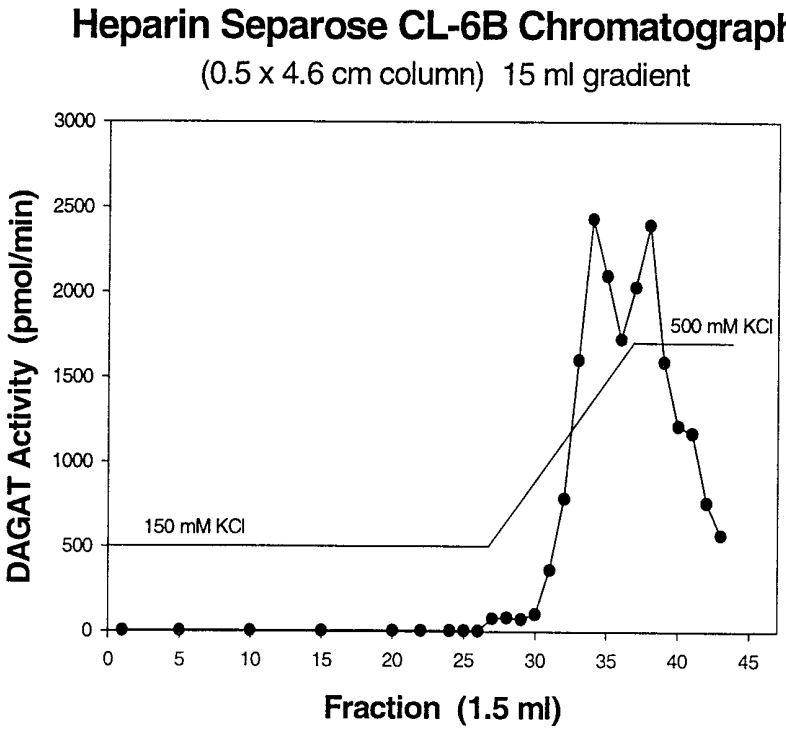


FIGURE 5 1/2

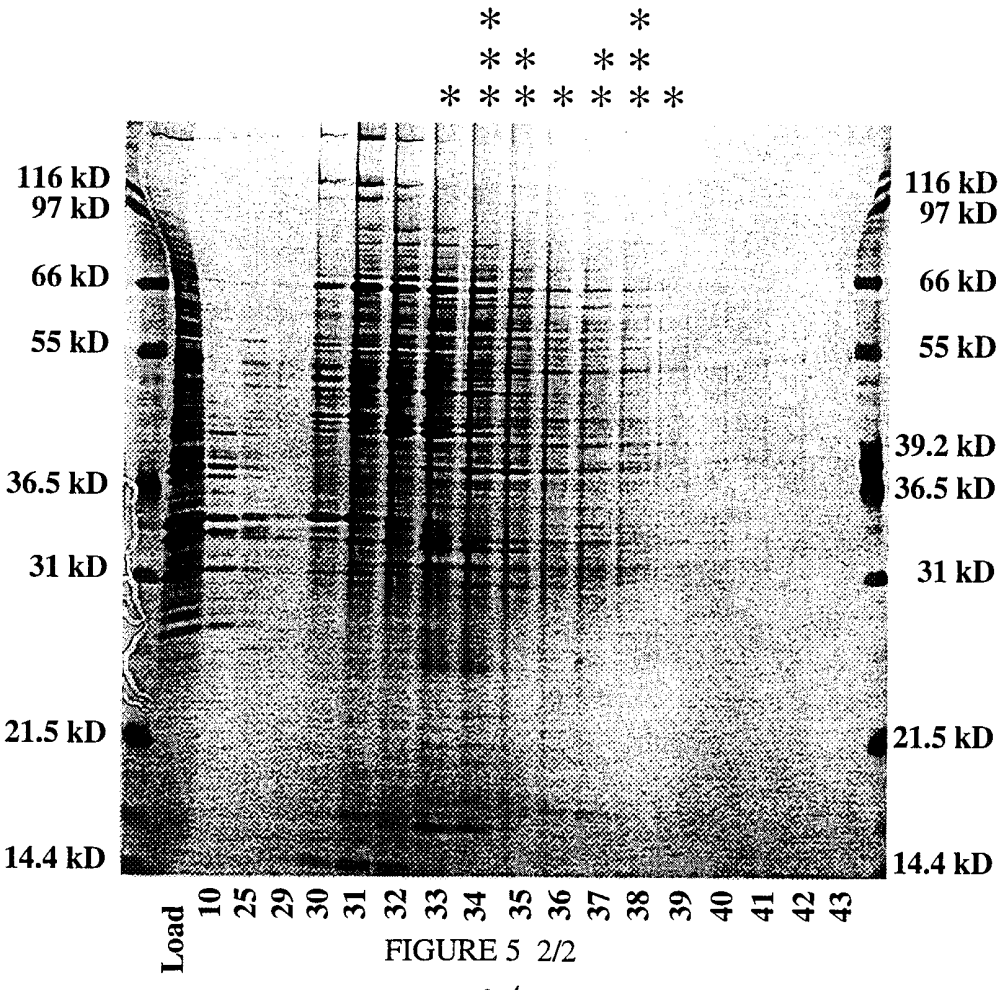


FIGURE 5 2/2

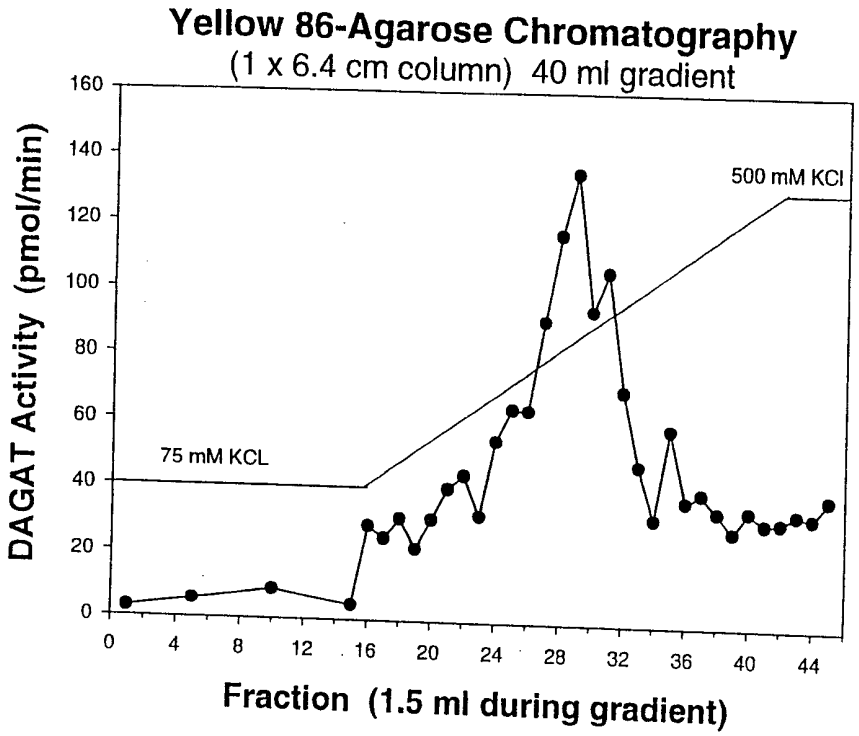


Figure 6 1/2

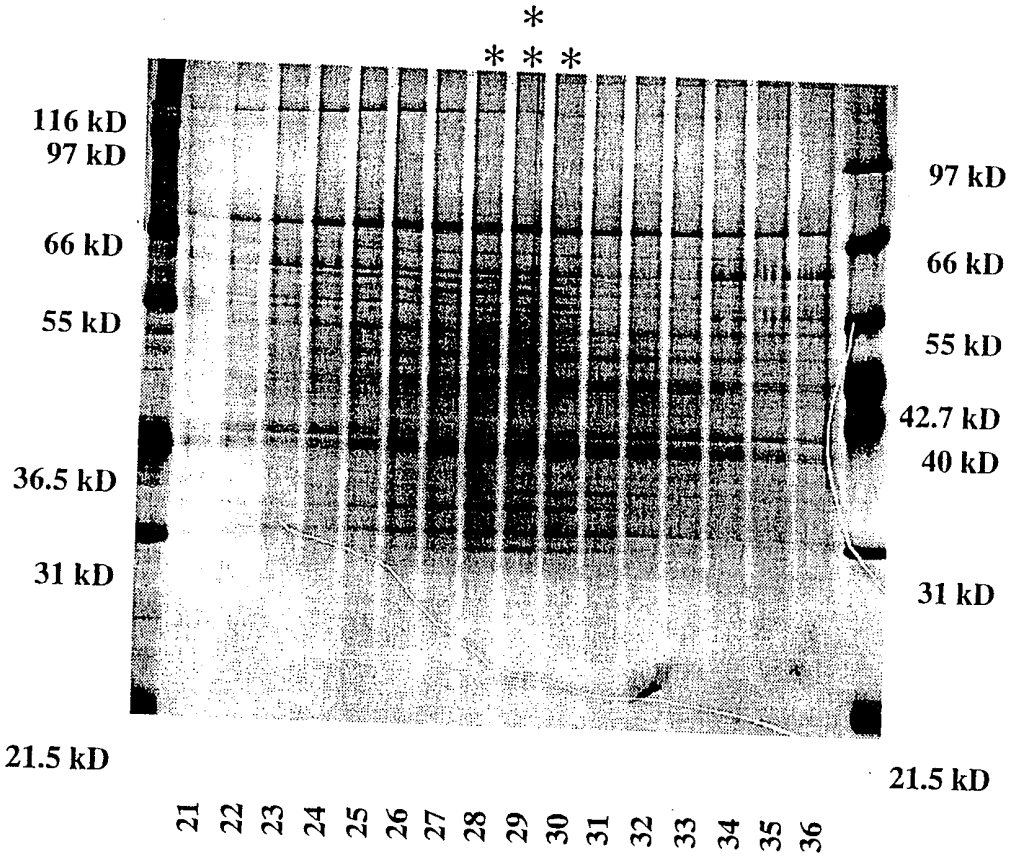


Figure 6 2/2

7/22

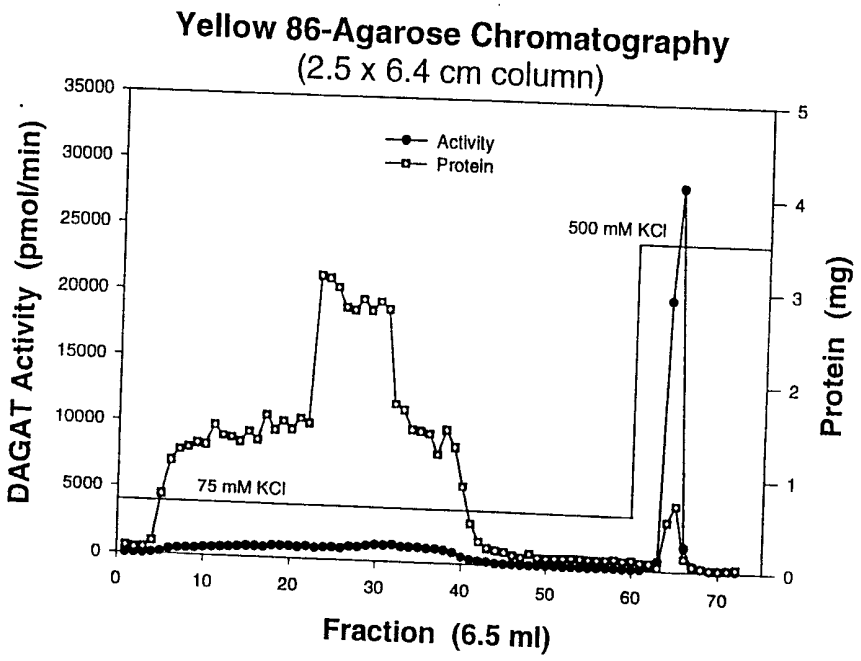

Yellow 86-Agarose Chromatography
(2.5 x 6.4 cm column)
DAGAT Activity (pmol/min)
Protein (mg)
Activity
Protein
500 mM KCl
75 mM KCl
Fraction (6.5 ml)
35000
30000
25000
20000
15000
10000
5000
0
5
4
3
2
1
0
0
10
20
30
40
50
60
70

FIGURE 7

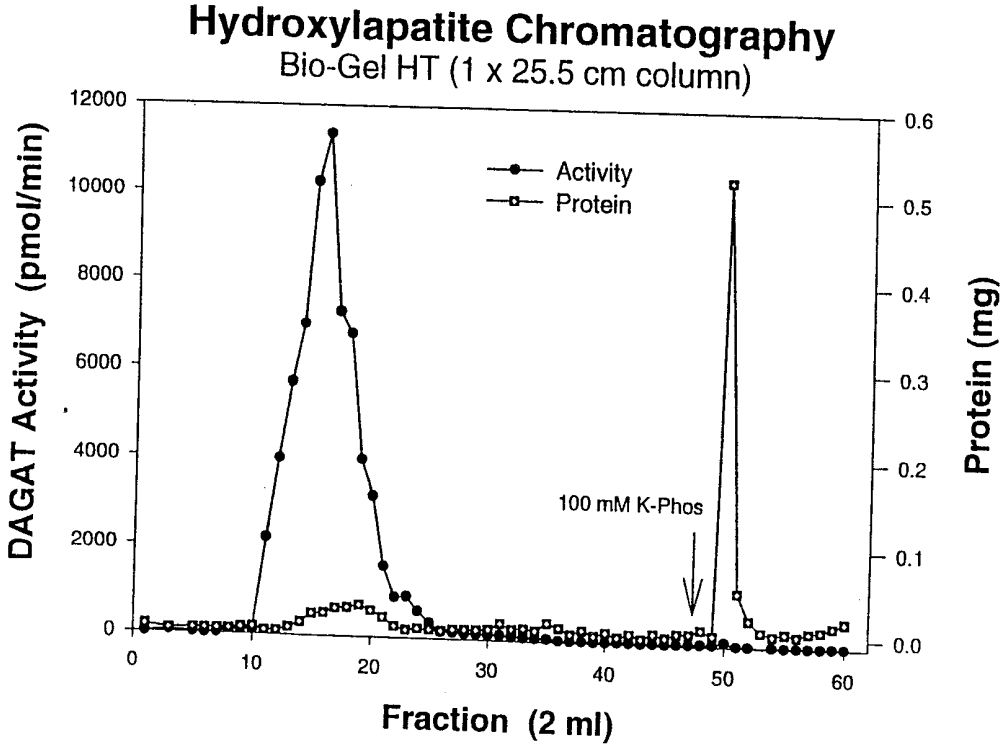


Figure 8 1/2

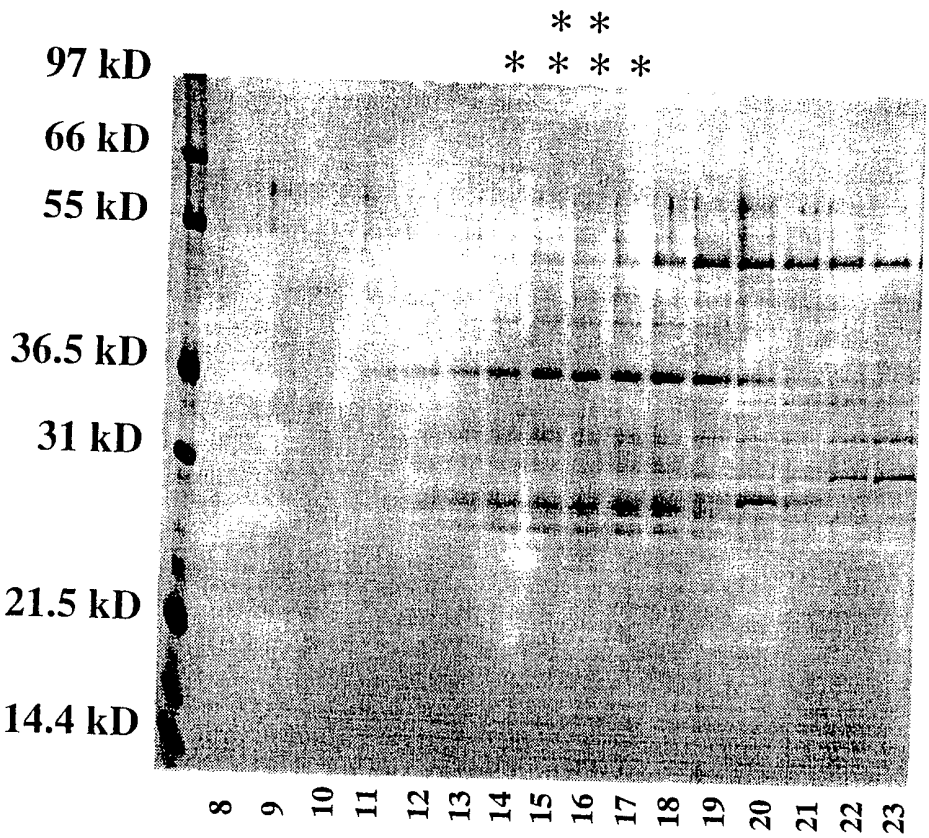


Figure 8 2/2

9/22

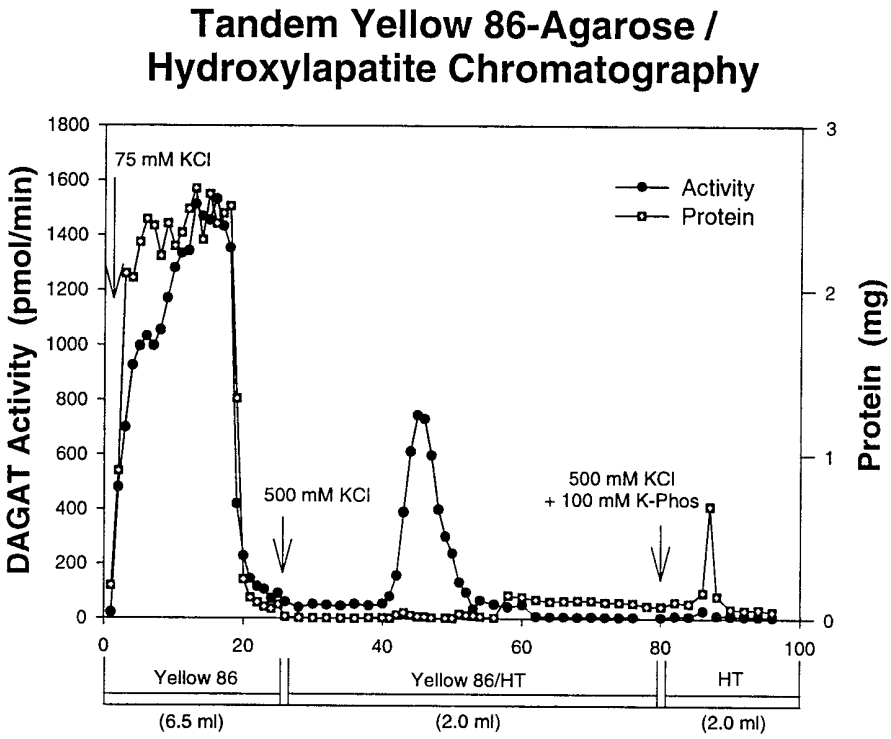


FIGURE 9 1/2

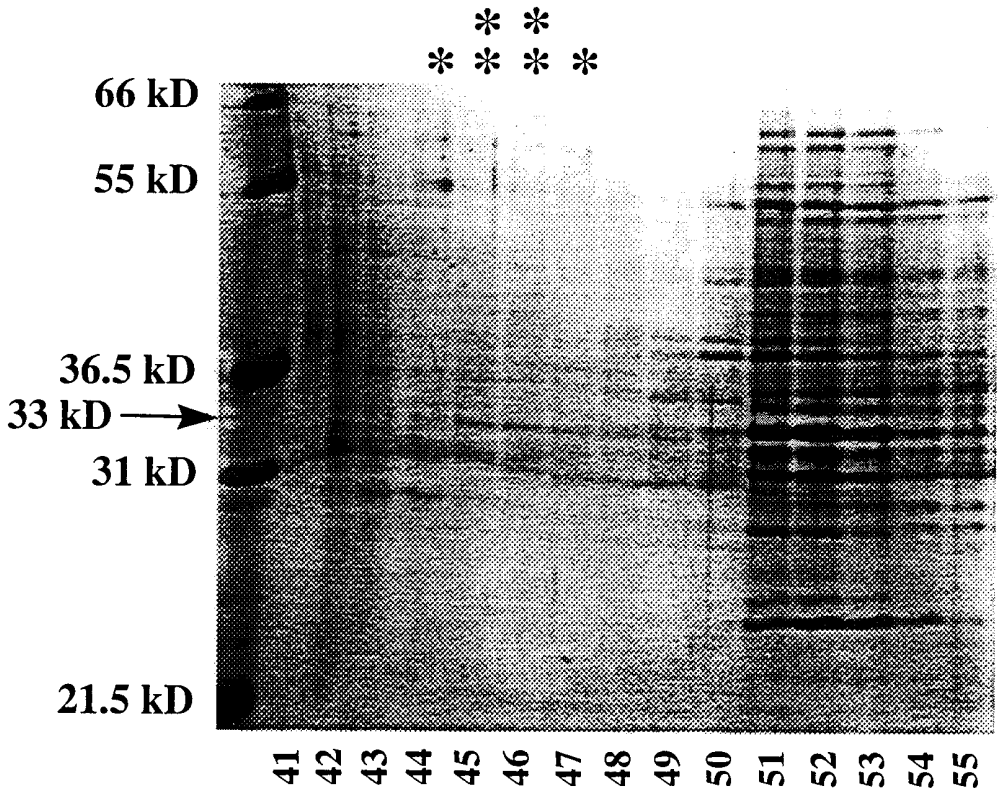


FIGURE 9 2/2

GTCTCCATTACAATGGAGGTGGAGAAGGAGCTAAAGACCTTCTCAG
AGGTATGGATCTCCGCCATAGCCGCCGCCTGCTACTGCCGCTTCG
TCCCCGCCGTTGCCCCTCACGGCGGCGCTCTCCGCCTCCTCCTCC
TCCTCCCCGTCGTCCTCCTCTTCATTTTCCTCCCCCTCCGCCTCTCC
TCCTTCCACCTCGGCGGGCCCACCGCCTTGTATCTCGTCTGGCTTG
CCAACTTCAAGCTCCTTCTCTTCGCCTTTCATCTTGGCCCTTTATCT
AACCCCTCTCTCTCTCTCCTTCACTTCATCTCCACCACCCTCCTCCC
CATCAAGTTCAGAGATGACCCATCTAATGATCATGAGAAAAACAAGA
GAACTCTGAGTTTTGAGTGGCGTAAAGTTGTTCTTTTTGTTGCTAAG
TTGGTGTTTTTTGCGGGTATTTTAAAGATTTATGAGTTTAGAAAAGA
TTTGCCTCATTTTGTGATCTCGGTGCTTTACTGTTTTCACTTCTATCT
CGGGACGGAGATCACCTTAGCAGCAAGCGCAGTCATAGCTCGAGC
CACGCTAGGGTTAGACCTATACCCCCAGTTCAACGAGCCATACTTA
GCCACCTCGCTGCAAGACTTCTGGGGGCGCAGGTGGAACCTCATG
GTGTCAGACATCTTGGGGTTGACAACATACCAGCCTGTCCGGCGT
GTCCTCTCGAGGTGGGTCAGGCTGCGGTGGGAGGTCGCCGGCGC
AATGTTGGTGGCGTTCACGGTGTCGGGGCTAATGCATGAAGTGTTT
TTCTTCTACTTAACTCGCGCGAGGCCCTCGTGGGAGGTGACGGGG
TTCTTTGTGTTGCATGGGGTTTGCACAGCCGTGGAGATGGTGGTGA
AGAAGGCGGTTTCAGGCAAGGTGCGGCTGCGCCGGGAGGTGTCA
GGGGCGCTGACGGTGGGGTTCGTGATGGTGACTGGAGGGTGGTT
GTTTTTGCCGCAGCTGGTGAGGCATGGGGTAGATTTGAAGACCATT
GATGAGTATCCTGTCATGTTTAATTATACTCAGAAGAAATTGATGGG
TTTGTTGGGGTGGTGATGAATGATGAGATGATGATCATGCATCTTC
TTTTTCGGAGATCGGTTGTACGTCACGAGGAGAACCCATGAAAAAT
GCAGATCARACGGCAAGACAGGTCGGGAAAAAAAAATGATCAATTT
TTCCTTAAGTAGCCGGCCTGCCACCCTGTCCGATTGTGGCATTTTT
GTGGTCACTTTTTCATATCGTGTAGTATTTTTGGTTTTTTGTTTTTAA
TGTTTTCTATGAATTTTGAATAATTTGTGCTTCATGAAAATTTTTTTT

FIGURE 10

MEVEKELKTFSEVWISAIAAACYCRFVPAVAPHGGALRLLLLLPVVLLFIFLPLRL
SSFHLGGPTALYLVWLANFKLLLFAFHLGPLSNPSLSLLHFISTTLLPIKFRDDPS
NDHEKNKRTLSFEWRKVVLFVAKLVFFAGILKIYEFRKDLPHFVISVLYCFHFYL
GTEITLAASAVIARATLGLDLYPQFNEPYLATSLQDFWGRRWNLMVSDILGLTT
YQPVRRVLSRWVRLRWEVAGAMLVAFTVSGLMHEVFFFYLTRARPSWEVTGF
FVLHGVCTAVEMVVKKAVSGKVRLRREVSGALTVGFVMVTGGWLFLPQLVRH
GVDLKTIDEYPVMFNYTQKKLMGLLGW

FIGURE 11

ATGGAAGAAAAGTTTAGAAACTTAATCGAGGTATGGATCTCTGCTTTAATCT
CTCTATCTTACTGTTATTACATATCGTCTAAACTCTCCAAAGGTGTTCTTCGT
CTCCTCTCTATTCTTCCAGTCTGCATTCTGTTTCTTGTTCTTCCTCTGTTCCT
CTCTTGTGTGCACTTTTGCGCCATTTCAGTTCTTTTTCTTTCATGGCTTGCAA
ACTTTAAGCTTCTTCTATTTGCCTTTGATGAGGGACCTTTGTTCCCACTTCCT
CCAAAACTCTCCCGTTTCATCTGCTTCGCTTGTTTACCCATCAAAATCAGAC
AAGACCCTTCTCCAAATGCGATACCAAATCTTCATCCTAAACCTATGCCTAA
ATGGGTTTTGGCTGTTAAAATTTTGGTCTTGGGCGTCTTGTTACATGTTTAT
GAATACAGGGATGGTTTGCCTCGGTTTGTTGTCTTGGCTCTCTATTGTCTCC
ATATTTACCTTGAGGTAGAACTTGTCTTGGTCTTTGTTGGAGCCGTGGTATC
TACTCTTCTTGGGTGTAACATCGAGCCGGTGTTCAATGAGCCCTACCTAGC
TACCTCCCTACAAGACTTCTGGAGCCGCAGATGGAACCTCATGGTTTCAGC
CGTCCTACGCTCAACCGTTCACATTCCGGTTCAGCGTTTTTTCAAACGCATA
CTCAGTCCAGACGGGGCTATGTTTGCTGGGGTCATGGCATCGTTCTTTGTC
TCAGGCTTGATGCATGAGCTGCTCTACTTTTACATGATCCGTAAGCCTCCAA
CTTGGGAAGTCACTTGTTTCTTTGTGTTGCATGGTGCTGCCACTGCCACTG
AGATAGCGGTGAAGAGAACACAATGGTTGAGGCCACCGCACCGGGCTGTC
TCTGGTCTTGTAGTTCTGACGTTTGTGAGTGTGACGGGCGTTTGGCTATTC
CTCGCTCAAGTGCTGAGAAACAATGTCCATGAGAAAGCGATTGGAGAATGT
TTATTGGTTCTTGACCTAGCCAAGTTATTCACTTCTTCATGA

Figure 12

ATGGAAGAAAAGTTTAGAAACTTAATCGAGGTATGGATCTCTGCTTT
AATCTCTCTATCTTACTGTTATTACATATCGTCTAAACTCTCCAAAGG
TGTTCTTCGTCTCCTCTCTATTCTTCCAGTCTGCATTCTGTTTCTTGT
TCTTCCTCTGTTCCTCTCTTGTGTGCACTTTTGCGCCATTTCAGTTC
TTTTTCTTTCATGGCTTGCAAACTTTAAGCTTCTTCTATTTGCCTTTG
ATGAGGGACCTTTGTTCCCACTTCCTCCAAAACTCTCCCGTTTCATC
TGCTTCGCTTGTTTACCCATCAAAATCAGACAAGACCCTTCTCCAAA
TGCGATACCAAATCTTCATCCTAAACCTATGCCTAAATGGGTTTTGG
CTGTTAAAATTTTGGTCTTGGGCGTCTTGTTACATGTTTATGAATAC
AGGGATGGTTTGCCTCGGTTTGTTGTCTTGGCTCTCTATTGTCTCC
ATATTTACCTTGAGGTAGAACTTGTCTTGGTCTTTGTTGGAGCCGTG
GTATCTACTCTTCTTGGGTGTAACATCGAGCCGGTGTTCAATGAGC
CCTACCTAGCTACCTCCCTACAAGACTTCTGGAGCCGCAGATGGAA
CCTCATGGTTTCAGCCGTCCTACGCTCAACCGTTCACATTCCGGTT
CAGCGTTTTTTCAAACGCATACTCAGTCCAGACGGGGCTATGTTTG
CTGGGGTCATGGCATCGTTCTTTGTCTCAGGCTTGATGCATGAGCT
GCTCTACTTTTACATGATCCGTAAGCCTCCAACTTGGGAAGTCACTT
GTTTCTTTGTGTTGCATGGTGCTGCCACTGCCACTGAGATAGCGGT
GAAGAGAACACAATGGTTGAGGCCACCGCACCGGGCTGTCTCTGG
TCTTGTAGTTCTGACGTTTGTGAGTGTGACGGGCGTTTGGCTATTC
CTCGCTCAAGTGCTGAGAAACAATGTCCATGAGAAAGCGATTGGAG
AATGTTTATTGGTTCTTGACCTAGCCAAGTTATTCACTTCTTCATGA

FIGURE 13

ATGAAACAGTTAGCAACCAACAGAACCAAGAGAGAAAAGATGGAAG
AAGAGTTGAGAAACCTAATCAAGGTTTGGATCTCTGCCTTAATCTCC
ATATCTTACTGTTACTACATCTCATCAAAAATCTCCAAAGGTGTTCTT
CGTCTCCTCTCTCTTCTTCCCATCTTCATCATCTTTCTTCTTCTTCCT
CTCTTCTTCTCTTCTGTCCACTTCTGCGTCATCTCAGGTTTCTTCTT
CACATGGCTCGCAAATTTCAAGCTCTTTCTCTTTGCTTTCGATCAAG
AACCTTTAAGCCCACTTCCCTCAAATCTCACCCGTTTCTTCTGCTTC
GCTTGTTTCCCCATCAAAATCAATAAAAACCCTTCTTCAAATCGAAT
CCACAACAAACCTATGTCTAAATGGGTCCTTGCTTTCAAACTTTTGA
TCTTTTCCTTCTTATTACATGTGTATAGAAACAACTATGATTCCGGTT
TATCACGGTTCGCTTTCTTGGCTCTCTTTACCATTCATGTTTACCTC
GAGGCAGAACTTATCTTAGTCTTCGTCGGTGCCTTGATGTCTATGC
TTCTTGGTTGTGAAATGGAACCGGTATTCAATGATCCTTACTTAGCC
ACTTCTTTACAAGAGTTTTGGAGCCGTAGATGGAACCTCATGGTCC
CAGCCGTACTCCGTCCAGCCGTCCACATACCGGTTCAGCGATTTTG
TGCACCGTTACTCGGTCTACACCGGGCTTTTTACGCTGGAATGTTA
GCCACGTTTATTGTCTCTGGTTTAATGCATGAGCTGATTTACTTTTA
TGTTATCCGCAAATCTCCAACTTGGGAAGTCACTTGCTTCTTTCTTT
TGCATGGAGTTGTAACTTGCCTAGAGATAGCGATGAAGAGGATGCG
GTGGCTTCCTACGCCACGTCGGGCGGTCTCGGGTCTTGCAATTAC
GGTGTTTTTGCTCGTTACAGCTGGTTGGTTGTTTTACCCTCAAATGT
TAAGAAATGATGTGCATAAGAGAGTGATAAGTGAATGTTTGTTGGTT
ATTGACGTTGTTAAAAGGCACGTTGTCTGTATTTTAATGTAA

FIGURE 14

ATGGAAGAAGAACTCAAGAACTTCATCAAGCTTTGGATTTCAGCAAT
AATCTCCATATCTTACTGTTACTACTTATCAACAGGAATCAAAGCTG
GTGTTTTTCGATTACTCTCTGTTCTTCCTGTATGTGCTCTGTTTCTTG
TTTTTCCTCTGTTTTTCTCCTATGTTCACTTCTCTGGTTGCATGGCTT
TTTTCCTCTCATGGCTCGCAAATTTCAAACTCATCCTCTTCTCCTTC
GATCAAGGTCCTCTTTCCCCACTTCCTCGAACTCTCTCCCGATTCAT
ATGCATCACTTGCTTCCCCATCAAGCCTCAACAAAACCCTAATATTC
AAAATTATAAAATCCCCATATGGCTTTTCGCCATTAAAGTTGTCATCT
TTGTTGTCTTGTTACAAATGTATGAATACAAACAATATCTGTCTCCG
GCTTTATTATTGGTTTTTAATTCTCTACATATATTCTTGGAGCTTGAG
ATTGTCTTTATGCTCGTCAAAGCATTGGTCTTTATCACTCTTGGCTG
CGATCTAGAGCCACAGTCCAATGAACCATACTTAGCCACTTCTCTTC
AAGACTTCTGGGGTCGTCGGTGGAACCTCATGGTCCCGGCGATTC
TCCGGCCGGCTGTCTACCTCCCGGCGAGACGAATGGCCTGTCGGA
AAGTTAACTCCGATCAGGCTATGTTCTTGGGAGTTTTCGCAGCGTT
TCTCGTCTCCGGTGCGGTTCATGAGATGCTCTTCTTCTATCTTACCC
GTGAGGTTCCTACAGGGGAAGTCACTTGGTTCTTTTTGTTACATGG
AGTTTGCACGGTGGCGGAAGTGGCGGTGAAGAAGAGTACATTTGT
GCGGCGATGGTGGAGAGTGAGTCCGACGGTGTCACGTCTTCTGAC
GGTCGGTTTTGTTGTTGTGACGAGTGGTTGGTTCTTTTTCCCTCTTA
TAAGGAGTGGCATCATCGAAAGACTCGCTAGCGAAGCCTTAATGTG
CATTGATTTCGTCAAGCACAAGTTTCTTCTGTTACTTTTGGGTGATT
AA

FIGURE 15

ATGGATGAAGAACTCAAGAACTTGATCAAAGTATGGGTTTCTGCAATAATCT
CGATATCTTATTGTTACTACATACCACCTAGAATCAAATCTGGTGCTCCTCG
ATTCCTCTCTGTTTCCCCTGTTCTTGCTCTGTTTCTTGTTCTTCCTCTGTTTT
TCTCCTCTCTGCATTTATCTTTAATCACAGCGTTTTTCCTCACATGGCTTGCT
AATTTCAAACTCATCCTCTTCTCCTTCGATAAAGGTCCTTTAATCCCAATTCC
AACAAATTTCCCTCGATTCTTCTGCTTCACTTGCTTCCCCATCAAGGTTCAG
CAAAACCCTAAATCTCAAAACCATTTGCCCAAATTGGTTTTCGCCATTAAACT
TGCAATCTTTGCAGTGCTATTACATTTGTATAGCTACAGACAAAATCTGTCTC
CGACTATACTATTAGGTCTCTATTTTGTGCATCTCTACTTAGAGATTGAGATT
ATATTAACGTTTGTTAAAGTTGTTGTTTTTATCTCTCTTGGCTGCGATCTTGA
GCCACAGTCCAATAAACCGTACTTAGCCACATCTCTACAAGACTTCTGGGG
TCGCCGGTGGAATCTCATGGTTCCGGCGATTCTCCGGCCAGCCGTTTACG
CACCAATGCGGCGAGTCTCTGAACGCAAAATGAGTTCCGGTTGGGCTCTGT
TTCCGGGGATTTTGGCAGCGTTTATCGTCTCCGGTTTGGTTCACGAATTGC
TCTTCTTCTATTTGATACGTGAGATGCCTACAGGAGAAGTTACTCTGTTCTT
TGTGTTACATGGCGTTTGTACTGCTGTAGAATTGGCGGTGAAGAAGAAAAC
GACGGTAGCACAGCGGTGGCGGTTGAGTCCGGGGGTGTCGCGGGTTCTC
ACGGTGGGGTTTGTGTTTGTGACTGGTGGTTGGTTGTTTACACCTCAGCTT
AAAAGGAGCGGGGTGATGGAGAGATTCACATCTGAAGCTGTGTTGCTCGTT
GAGTTCATTAAGCGATAA

Figure 16

ATGGATGAAGAACTCAAGAACTTGATCAAAGTATGGGTTTCTGCAAT
AATCTCGATATCTTATTGTTACTACATACCACCTAGAATCAAATCTG
GTGCTCCTCGATTCCTCTCTGTTTCCCCTGTTCTTGCTCTGTTTCTT
GTTCTTCCTCTGTTTTTCTCCTCTCTGCATTTATCTTTAATCACAGCG
TTTTTCCTCACATGGCTTGCTAATTTCAAACTCATCCTCTTCTCCTTC
GATAAAGGTCCTTTAATCCCAATTCCAACAAATTTCCCTCGATTCTT
CTGCTTCACTTGCTTCCCCATCAAGGTTCAGCAAAACCCTAAATCTC
AAAACCATTTGCCCAAATTGGTTTTCGCCATTAAACTTGCAATCTTT
GCAGTGCTATTACATTTGTATAGCTACAGACAAAATCTGTCTCCGAC
TATACTATTAGGTCTCTATTTTGTGCATCTCTACTTAGAGATTGAGAT
TATATTAACGTTTGTTAAAGTTGTTGTTTTTATCTCTCTTGGCTGCGA
TCTTGAGCCACAGTCCAATAAACCGTACTTAGCCACATCTCTACAAG
ACTTCTGGGGTCGCCGGTGGAATCTCATGGTTCCGGCGATTCTCC
GGCCAGCCGTTTACGCACCAATGCGGCGAGTCTCTGAACGCAAAA
TGAGTTCCGGTTGGGCTCTGTTTCCGGGGATTTTGGCAGCGTTTAT
CGTCTCCGGTTTGGTTCACGAATTGCTCTTCTTCTATTTGATACGTG
AGATGCCTACAGGAGAAGTTACTCTGTTCTTTGTGTTACATGGCGTT
TGTACTGCTGTAGAATTGGCGGTGAAGAAGAAAACGACGGTAGCAC
AGCGGTGGCGGTTGAGTCCGGGGGTGTCGCGGGTTCTCACGGTG
GGGTTTGTGTTTGTGACTGGTGGTTGGTTGTTTACACCTCAGCTTA
AAAGGAGCGGGGTGATGGAGAGATTCACATCTGAAGCTGTGTTGC
TCGTTGAGTTCATTAAGCGATAA

FIGURE 17

ATGGAGGAAGAACTCAAGTTATTCATCCAAGTATGGGTTTCTGCAAT
CATTTCAGTAACTTATTGTTACTACTTAACACCCAAAATCAAAACCA
GTCTTCTTCGATTACTATCTGTTCTTCCTGTTTGTGTTTTGTTTCTTA
TTATTCCTATCTTTTTCTCCACTGTTCATTCCTCTTTCACTATTGCAT
TTTTCCTCTCAGGTCTTGCAGTTCCAAAACTCATCCTCTTTGCATTA
GAAAAAGGTCCTCTTTTTCCACTTCCTCCTAATCTCCCTCATTTCGT
CTGCTTTGCTTGCTTCCCCATCAAGCTTCAAAAAAAACCTAACCCTG
AAAATACTAACCATTTCCCCAAATGGGTTTTTGCCCTGAAAGTTTTC
ATCTTTGGTGCCTTGTTACTACAAGCGTATCATTACAAACAATTTCT
ATCTACGAATTTTCTATTGGGTCTCTATGCTCTGCATATATACTTGG
AGCTTGAGATTTCCTTAACCTTGATAAAATTTCTCGTCAGTATCACT
CTTGGGTGTGACCTCGAGCCACAATTCAACGAACCATACTTAGCCA
CCTCTCTACATGACTTCTGGGGTCACCGATGGAACCTCATGGTCTC
GAAGATTCTCTGGCTCGCAGTGTACAACCCCATACGGCAATGGCGA
GCCAAGAGCTCCGAGTGGGATCGGTTCTTCGCGATTTTCGCCACG
TTCCTCGTCTCTGGTGTGGCTCACGAGATTCTCTACTTCTATTTGAC
ACGTGAGAAGCCTACATGGGAGGTGACTTGGTTCTTTGTGTTACAT
GGGTTTTGCATGGCGGCTGAAGTGGCACTGAAGAGGAAGACGAAG
TTGGTGCAGCGGTGGCCGGTGAATCCGGCAGTGTCGAGACTGCTT
ACGGTGGGGTTTGTGTTTGTGACTGGTGTTTGGCTATTTTCCCCCC
AGCCTATTAGGCACGGCTTGATGGAGAGGTTCATCAATGAAGACTT
GTTTCTAATTGATTTCTTTAATCGTAAGTTATATATCCTCTTAGGGTT
GTTTACGAGTCTTTAA

FIGURE 18

ClustalW Formatted Alignments

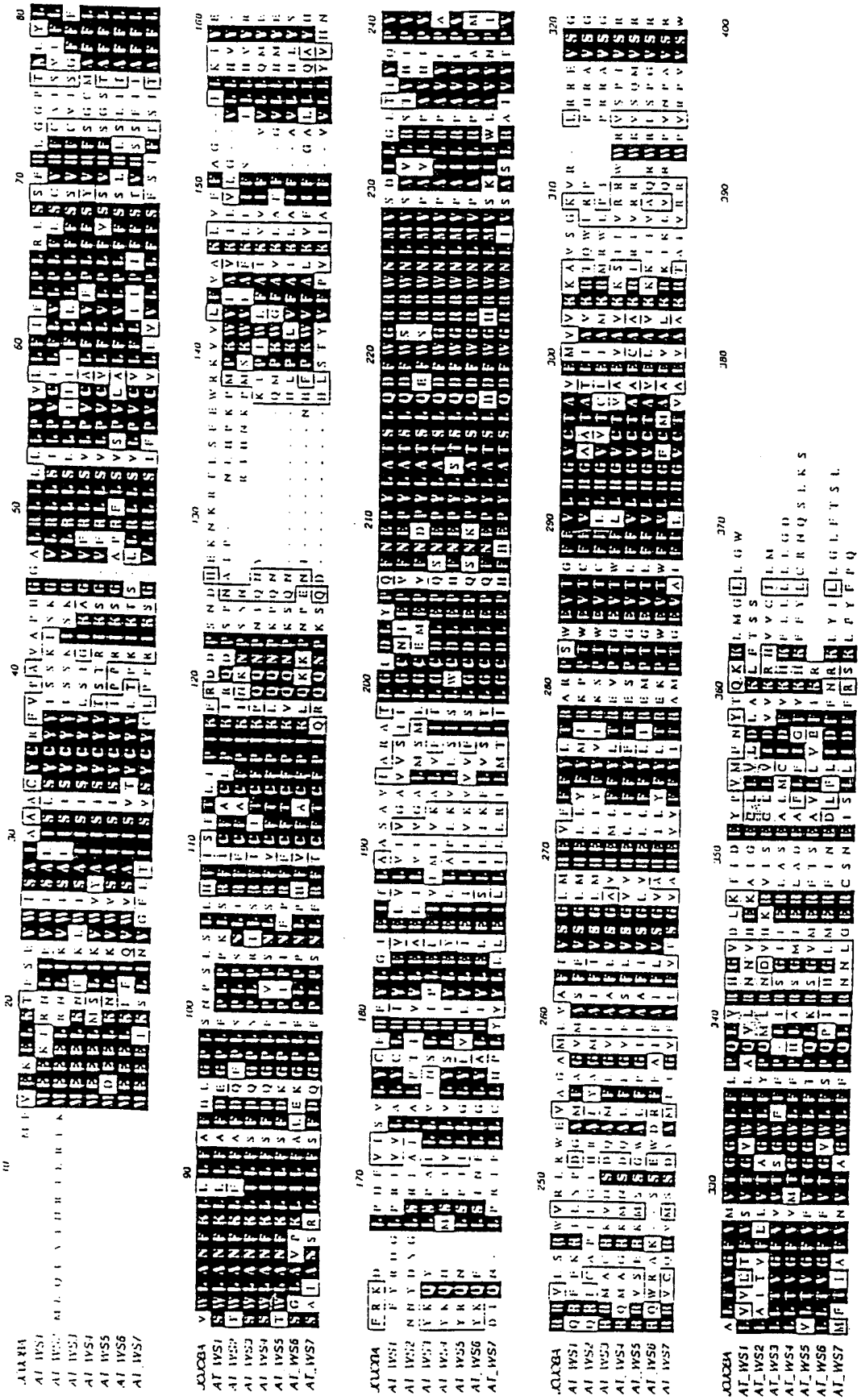


Figure 19

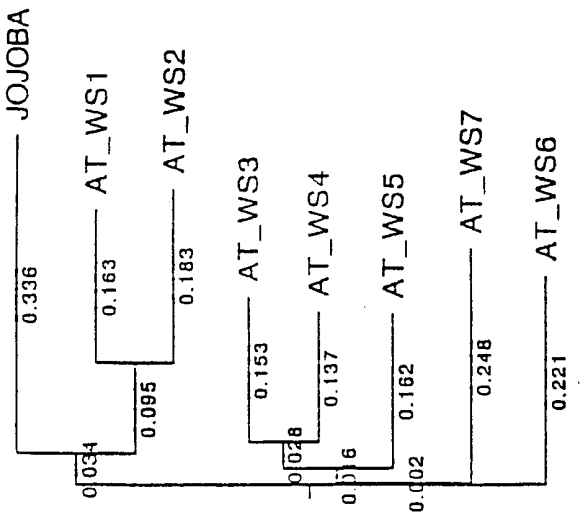

JOJOBA
AT_WS1
AT_WS2
AT_WS3
AT_WS4
AT_WS5
AT_WS7
AT_WS6
0.336
0.163
0.183
0.095
0.153
0.137
0.162
0.248
0.221

Figure 20

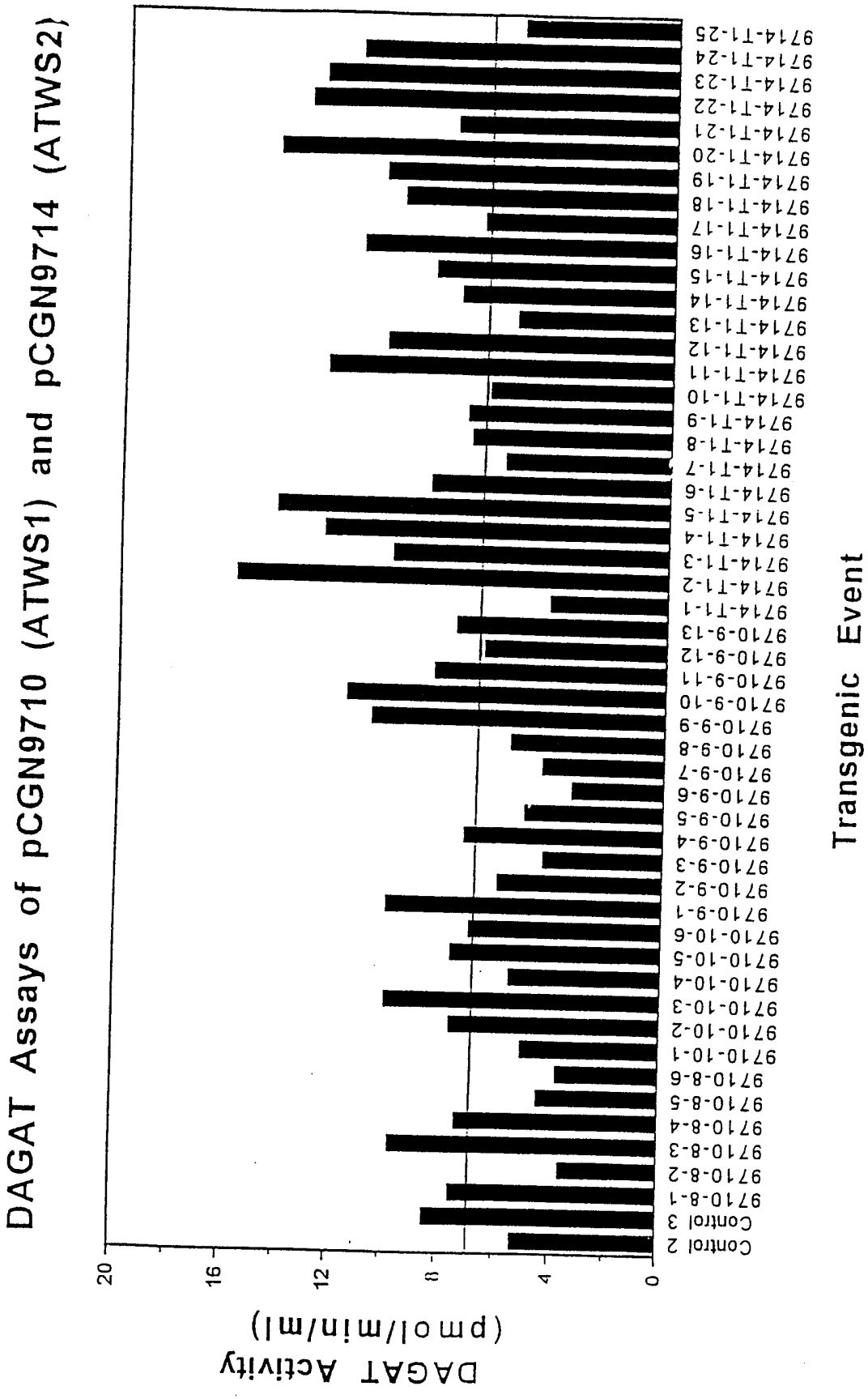


Figure 21

22/22