Use of Limited Proteolysis and Mutagenesis To Identify Folding Domains and Sequence Motifs Critical for Wax Ester Synthase/Acyl Coenzyme A:Diacylglycerol Acyltransferase Activity

Juan A. Villa,* Matilde Cabezas, Fernando de la Cruz, Gabriel Moncalián

Departamento de Biología Molecular e Instituto de Biomedicina y Biotecnología de Cantabria, Universidad de Cantabria-Consejo Superior de Investigaciones Científicas-SODERCAN, Santander, Spain

Triacylglycerols and wax esters are synthesized as energy storage molecules by some proteobacteria and actinobacteria under stress. The enzyme responsible for neutral lipid accumulation is the bifunctional wax ester synthase/acyl-coenzyme A (CoA): diacylglycerol acyltransferase (WS/DGAT). Structural modeling of WS/DGAT suggests that it can adopt an acyl-CoA-dependent acyltransferase fold with the N-terminal and C-terminal domains connected by a helical linker, an architecture demonstrated experimentally by limited proteolysis. Moreover, we found that both domains form an active complex when coexpressed as independent polypeptides. The structural prediction and sequence alignment of different WS/DGAT proteins indicated catalytically important motifs in the enzyme. Their role was probed by measuring the activities of a series of alanine scanning mutants. Our study underscores the structural understanding of this protein family and paves the way for their modification to improve the production of neutral lipids.

Many microbes have the capacity to accumulate intracellular neutral lipids (mainly single cell oils [SCO]) up to a large percentage of their biomass (1). This fact has a great biotechnological potential, since there is a growing interest in the production of sustainable oils for their use as biodiesel or as commodity oils (2). These SCO have been found to be produced by different oleaginous microorganisms, such as microalgae, yeast, fungi, and bacteria (3).

SCO are very appropriate for use as the source of biodiesel or commodity oils, because the producing microorganisms are able to grow using a great variety of substrates, have a short life cycle, and are easy to modify by genetic engineering.

Regarding bacteria, the accumulation of the neutral lipids triacylglycerols (TAGs), wax esters (WEs), and polyhydroxyalkonates (PHAs) has been reported. The main purpose of this accumulation is the storage of carbon and energy under growth-limiting conditions. While PHAs are synthesized in a wide variety of bacteria (4), the accumulation of TAGs has been described only for a few bacteria belonging to different genera of the actinobacterium group, like *Mycobacterium* (5), *Streptomyces* (6), or *Rhodococcus* (7; for a review, see reference 8). The accumulation of TAGs is remarkably high in *Rhodococcus* and *Gordonia*, with up to 80% of the cellular dry weight in the form of neutral lipids (9, 10). Besides, in proteobacteria of the genus *Acinetobacter*, *Pseudomonas*, or *Marinobacter*, WEs are accumulated as energy storage components under growth-limiting conditions (11, 12).

In both TAG synthesis in actinobacteria and WE synthesis in proteobacteria, the enzyme responsible for the accumulation is the wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase (WS/DGAT) (reviewed in reference 13). This promiscuous enzyme was first characterized in the bacterium *Acinetobacter baylyi*, where it was shown to act as both a DGAT and a WS (14). The WS/DGAT of *A. baylyi* uses *in vitro* a wide range of substrates as acyl acceptors, including alcohols of many chain lengths, 1,2- and 1,3-diacylglycerols, monoacylglycerol, diols, and thiols (14–17).

WS/DGATs were also found in some of the above-mentioned bacteria able to accumulate neutral lipids, such as the actinobacteria *Mycobacterium* (18), *Streptomyces*, and *Rhodococcus* (19) and the proteobacteria *Marinobacter* (20) and *Alcanivorax* (21).

Moreover, genes encoding proteins of the WS/DGAT family were also found in plants. The sequenced genomes of *Arabidopsis thaliana* and *Oryza sativa* contain 11 and 3 homologs of the *Acinetobacter* WS/DGAT, respectively, while a homolog was cloned and characterized in *Petunia* (22).

This WE production was exploited for the production of fatty acid ethyl esters (FAEEs) (microdiesel) in an ethanologenic *Escherichia coli* strain by the esterification of ethanol with acyl-CoA (23). This approach was improved, avoiding the β-oxidation of the fatty acids to increase FAEE production and expressing hemicellulases for the use of hemicellulase as the carbon source in microdiesel-producing *E. coli* (24).

WS/DGAT proteins catalyze the final step of TAG or wax ester biosynthesis by the condensation of acyl-CoA and fatty alcohol or diacylglycerol, respectively (25). All the studied WS/DGATs contain the conserved catalytic motif HHxxxDG, which has been proven to be involved in the protein activity (26). A similar HHxxxDG motif is conserved in a large number of acyl-CoA-dependent acyltransferases also involved in the transfer of fatty acyl groups by the nucleophilic attack of a hydroxyl group on the thioester bond of the fatty acyl-CoA (13, 27–30). In this motif, the

second conserved His has been proposed to be a general base that promotes deprotonation of the hydroxyl group of the alcohol to catalyze the transfer of the acyl group. The aspartic acid present in the motif has also been shown to be critical for activity, and it appears to play a structural role in the organization of the active site.

In this work, we have modeled the structure of WS/DGAT proteins. The model suggests that WS/DGAT proteins can adopt a three-dimensional (3D) structure with two domains (N-terminal and C-terminal domains) connected by a helical linker, similar to several other HHxxxDG acyltransferases whose structures have been solved. To confirm the model prediction, limited proteolysis and mutagenesis were carried out. Limited proteolysis validated the presence of two structural domains within the protein. Moreover, we have found that both domains are active when coexpressed as independent polypeptides. The structural prediction and sequence alignment of different WS/DGAT proteins allowed us to find some of the catalytically important residues of the enzyme. The role of these residues was tested by checking the activities of the corresponding mutants. Thus, the structural characterization of the family of proteins described in this article would facilitate the modification of the WS/DGAT enzymes to enhance the production of TAGs or wax esters.

MATERIALS AND METHODS

Sequence analysis software and public gene expression data sets. All protein and nucleic acid sequences were obtained from the public databases at NCBI (http://www.ncbi.nlm.nih.gov/). Protein database searches were performed using BLASTP (31), available on the NCBI web site (http://www.ncbi.nlm.nih.gov). Multiple sequence alignments were made using the software program MUSCLE (32) and plotted using the program ESPript 2.2 (33). Neighbor-joining phylogenetic analysis (34) with bootstrap values (500 replicates) (35) was carried out using the software program MEGA5 (36). The evolutionary distances were computed using the Poisson correction method (37).

Structural modeling. The Ma2 3D structure was predicted by homology modeling using the Phyre server (38) (http://www.sbg.bio.ic.ac.uk/phyre2) or the robetta server (39) (http://robetta.bakerlab.org/). The crystal structure of tyrocidine synthetase TycC (PDB identifier [ID] 2JGP) was used as the template (40). An image of the resulting 3D model was generated using the program Pymol (DeLano Scientific, Palo Alto, CA, USA) (http://www.pymol.org).

Plasmid construction. Plasmids used in this work (described in Table 1) were constructed by cloning the relevant PCR-generated fragments into pET29c (Km^r) or pET3a (Ap^r) expression vectors (Novagen). Desired DNA fragments were amplified using oligonucleotides containing NdeI and XhoI restriction sites and *Marinobacter hydrocarbonoclasticus* VT8 (DSMZ collection, Braunschweig, Germany) genomic DNA as the template. After PCR amplification, the resulting products were digested with NdeI and XhoI and ligated to an NdeI-XhoI-digested pET29c plasmid. The fragments to be inserted in pET3a were amplified and cloned using oligonucleotides containing NdeI and BamHI restriction sites. Plasmids were introduced by electroporation into DH5α cells (41). The identity of constructed plasmids was checked by DNA sequencing. Plasmid DNAs were then transferred to strain C41(DE3) (42) for protein overexpression.

Site-directed mutagenesis. Site-directed mutagenesis of the Ma2 gene within the pET29c plasmid was performed using the QuickChange kit (Stratagene, California) using the procedure specified by the manufacturer. A total of 3 colonies were selected, and their plasmids were isolated by using the Miniprep kit (Qiagen, The Netherlands). Positive mutants were selected by DNA sequencing.

TABLE 1 Plasmids used in this work

Plasmid	Description	Phenotype	Size (kb)	Reference or source
pET29c	Expression vector	Km^r	5.4	Novagen
pET3a	Expression vector	Ap^r	4.7	Novagen
pETMa2	pET29c::Ma2	Km^r	7	This work
pETMa2Δ297C	pET3a::Ma2(1–176)	Ap^r	6	This work
pETMa2Δ178N	pET29c::Ma2(179–473)	Km^r	6.3	This work
pETMa2H140A	pET29c::Ma2H140A	Km^r	7	This work
pETMa2H141A	pET29c::Ma2H141A	Km^r	7	This work
pETMa2D145A	pET29c::Ma2D145A	Km^r	7	This work
pETMa2P118A	pET29c::Ma2P118A	Km^r	7	This work
pETMa2L119A	pET29c::Ma2L119A	Km^r	7	This work
pETMa2W120A	pET29c::Ma2W120A	Km^r	7	This work
pETMa2N270A	pET29c::Ma2N270A	Km^r	7	This work
pETMa2D271A	pET29c::Ma2D271A	Km^r	7	This work
pETMa2R305A	pET29c::Ma2R305A	Km^r	7	This work

Protein purification. Ma2 protein containing a C-terminal His tag (Ma2HT) and its corresponding mutants were purified as follows. An overnight culture of *Escherichia coli* C41(DE3) cells harboring the appropriate pET29c-derived plasmid was diluted 20-fold in 1 liter of LB medium containing kanamycin (50 μg/ml) and incubated at 37°C with shaking until an A_{600} of 0.6 was reached. Then, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM. After 3 h of further incubation, cells were pelleted and then frozen at −80°C. This pellet was resuspended in 20 ml of buffer A (50 mM Tris–HCl [pH 7.5] and 1 M NaCl) containing 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and then sonicated. Cellular debris was eliminated after centrifugation at 40,000 rpm for 20 min. Supernatant was loaded onto a 5-ml His-Trap column (GE Healthcare) equilibrated with buffer A. Bound proteins were eluted with a linear gradient of buffer B (50 mM Tris–HCl [pH 7.5], 1 M NaCl, and 0.5 M imidazole) in 20 column volumes. The protein sample was then loaded onto a Superdex 200 HR 10/30 size exclusion column (GE Healthcare, United Kingdom) equilibrated with buffer C (50 mM Tris–HCl [pH 7.5] and 600 mM NaCl) and eluted with the same buffer. Glycerol was added to a 20% final concentration, and proteins were stored at −80°C. Protein molecular weights were checked by matrix-assisted laser desorption–ionization time of flight (MALDI-TOF) mass spectrometry as described below.

Limited trypsin digestion. The Ma2 protein at a concentration of 0.8 mg/ml in buffer B was incubated with different amounts of trypsin at 37°C. After 30 min of trypsin treatment, loading buffer blue (2×) (0.5 M Tris-HCl [pH 6.8], 4.4% [wt/vol] SDS, 20% [vol/vol] glycerol, 2% [vol/vol] 2-mercaptoethanol, and bromophenol blue) was added to stop the reaction. After boiling the samples for 5 min, limited trypsin digestion was verified by 12% SDS-PAGE and Coomassie brilliant blue staining.

Mass spectrometry analysis. Molecular masses of intact and protease-digested Ma2 proteins were determined by MALDI-TOF mass spectrometry. Both preparations were desalted using ZipTip C$_4$ microcolumns (Millipore) (2-μl sample of Ma2 at 0.8 mg/ml) with elution using a 0.5-μl SA (sinapinic acid [10 mg/ml] in [70:30] acetonitrile-trifluoroacetic acid [0.1%]) matrix on a GroundSteel massive 384 target (Bruker Daltonics). An Autoflex III MALDI-TOF/TOF spectrometer (Bruker Daltonics) was used in linear mode with the following settings: 5,000- to 50,000-Da window; linear positive mode; ion source 1, 20 kV; ion source 2, 18.5 kV; lens, 9 kV; pulsed ion extraction of 120 ns; and high gating ion suppression up to 1,000 M_r. Mass calibration was performed externally with Bruker's Protein 1 standard calibration mixture (Bruker Daltonics) in the same range as the samples. Data acquisition was performed using the FlexControl 3.0 software program (Bruker Daltonics), and peak peaking and subsequent spectral analysis were performed using FlexAnalysis 3.0 software (Bruker Daltonics).

Pulldown assay. Pulldown assay with fragments Ma2(1–176) and Ma2(179–473) and the His tag was carried out as follows. *E. coli* C41(DE3) cells were transformed with plasmids pETMa2Δ178N (Km^r) and pETMa2Δ297C (Ap^r). LB medium containing kanamycin (50 μg/ml) and ampicillin (100 μg/ml) was used to ensure that both plasmids were maintained in the cell. Twenty-five milliliters of these cultures were induced as described above, resuspended in 1 ml of buffer A, and then sonicated. After sonication, the pellet was centrifuged at 40,000 rpm for 10 min. Supernatant was incubated for 30 min with nickel-nitrilotriacetic acid (Ni-NTA) agarose resin (Qiagen). The resin was washed with buffer A, and then proteins were eluted with 100 μl of buffer B. All the fractions were applied to SDS-PAGE. The gels were stained with Coomassie brilliant blue.

Activity assay. A spectrophotometric assay was developed to determine the kinetic properties of Ma2 and the two-domain complex. The concentration of sulfhydryl groups of CoA released during the condensation reaction between fatty/isoprenoid CoA-activated acids and the group alcohol was determined by using Ellman's reagent (5,5′-dithio-bis-[2-nitrobenzoic acid] [DTNB]) (43), essentially as described previously (20). Final concentrations of reagents in the assay were as follows: 250 μM palmitoyl-CoA, 250 μM 1,2-dipalmitoyl glycerol or other alcohols (hexadecanol, glycerol, butanol, ethanol, or methanol), and 250 μM DTNB. Five-hundred-microliter reaction mixtures were incubated for 12 min at 37°C with 0.5 μg/ml of Ma2 or the complex formed by the N- and C-terminal domains. Released CoA was measured using a ε value (extinction coefficient) of 14,150 M^{-1}cm^{-1} for TNB (2-nitro-5-thiobenzoic acid) (44). At least 3 reactions per sample were measured.

RESULTS

Structural modeling of WS/DGAT proteins predicts the CoA-dependent acyltransferase fold. Despite the growing number of WS/DGAT proteins that have been characterized for their biotechnological interest, little is known about the structural features and the role of the conserved residues of this family of proteins.

A BLAST search against current databases using the sequence of one of the best-characterized WS/DGAT enzymes, *M. hydrocarbonoclasticus* Ma2 (UniProt accession number A1U572) (20), revealed more than 1,000 putative WS/DGAT homologs in bacteria and plants (Fig. 1A). A phylogenetic analysis with 51 representative enzymes is shown in Fig. 1B. The alignment of the proteins used in the phylogenetic analysis is shown in Fig. S1 in the supplemental material. Plant WS/DGAT forms a well-supported monophyletic subclade that shares a common origin with bacterial WS/DGAT. Within bacteria, most proteobacterial WS/DGATs are phylogenetically related. However, the different actinobacterial WS/DGAT enzymes do not form a monophyletic group. Each actinobacterium strain shows a high number of WS/DGAT paralogs (up to 15) dispersed along the bacterial branches of the phylogenetic tree. Some of them seem to be recent, such as the YP_702929 and YP_707847 proteins of *R. jostii*.

To further characterize this WS/DGAT family of proteins, we carried out structural prediction and modeling using the Protein Homology/analogY Recognition Engine (PHYRE) Web server (38).We have launched a PHYRE batch job using the primary sequences of seven representative bacterial and plant WS/DGAT proteins selected from phylogenetic analysis (Fig. 1B). Three different actinobacterial WS/DGATs (Ms, *Mycobacterium* sp., YP_001073143; Rj, *Rhodococcus jostii*, YP_701572; Nf, *Nocardia farcinica*, YP_117375), three gammaproteobacterial WS/DGATs (Ma, *M. Hydrocarbonoclasticus* Ma2; Ab, *Alcanivorax borkumensis*, YP_693524; As, *Acinetobacter* sp., YP_045555), and one plant WS/DGAT (At, *Arabidopsis thaliana*, NP_568547) were used. The alignment of these selected proteins is shown in Fig. 2A. For the seven studied proteins, the same predicted structural homologs were found, including surfactin A synthetase (PDB ID 2VSQ [45]), tyrocidine synthetase TycC (PDB ID 2JGP [40]), VibH, a nonribosomal peptide synthetase (NRPS) condensation enzyme (PDB ID 1L5A [46]), polyketide synthase-associated protein 5 (PDB ID 1Q9J [47]), trichothecene 3-*o*-acetyltransferase (PDB ID 3B30 [48]), or vinorine synthase (PDB ID 2BGH [49]). All the found homologs were CoA-dependent acyltransferases containing the HHxxxDG motif also present in WS/DGAT. The confidence of the prediction was 100% for all seven proteins. For clarity, we have chosen the structural prediction of Ma2 as the WS/DGAT structural model. Ma2 is the protein we have used for the subsequent *in vitro* experiments because it was previously shown to be soluble and active when purified after overexpression in *E. coli* (20). A 3D model of Ma2 was generated by PHYRE homology modeling using tyrocidine synthetase (TycC) acyltransferase (PDB ID 2JGP) as the template (Fig. 2B). A similar Ma2 model was obtained using robetta as the modeling software program (data not shown). Four hundred fifty-eight residues (97%) of Ma2 could be modeled at >90% accuracy, although the sequence identity between Ma2 and 2JPG is only 11%. The structural homology between Ma1 and 1Q9J was also recently reported (50).

The structural model suggests that Ma2 can adopt a 3D structure with two domains (N-terminal and C-terminal domains) connected by a helical linker. The N domain contains a core of five-stranded mixed sheet (β2, β5, β6, β7, and β12) surrounded by three alpha-helices (α2, α3, and α4). The core of the C domain is composed by a six-stranded mixed sheet (β8, β9, β10, β11, β13, and β14) and four alpha-helices covering the external face of the sheet (α8, α9, α10, and α12). The internal face of the C sheet is covered by the N domain. Both domains are connected by a helical linker formed by alpha-helices α6 and α7.

Limited proteolysis validates two structural domains. The predicted tertiary structure of WS/DGAT enzymes was probed *in vitro* by limited proteolysis of the purified protein Ma2HT. Ma2HT was treated with trypsin as described in Materials and Methods. We observed, after trypsin treatment, two discrete bands with lower molecular masses than the full-length protein (Fig. 3A). Moreover, MALDI-TOF spectrometry of the digested protein gave two peaks of molecular masses 20,174.3 and 28,556.9 Da, respectively. The sizes perfectly match the M1-R177 and L209-R468 fragments of the Ma2HT protein, respectively. The result suggests that indeed there are two connected domains in this family of proteins. The predicted connecting segment (70 amino acids [aa]) is very exposed to the protease action and was rapidly degraded by the trypsin treatment, and thus, after the trypsin treatment, only the two N and C domains were observed. In Ma2, this helical link is located approximately between residues 170 and 240 (shown in green in Fig. 2A and 3B).

Ma2 domains interact and are functional together when coexpressed. To further demonstrate the existence of this structural organization of two interacting domains, both domains were simultaneously coexpressed in different plasmids contained in the same cell.

According to the structural prediction and the trypsin treatment results, we decided to separately express amino acids 1 to 176 and 246 to 473. The 1–176 fragment (N domain) was cloned into pET3a (plasmid pETMa2Δ297C) and was effectively expressed in
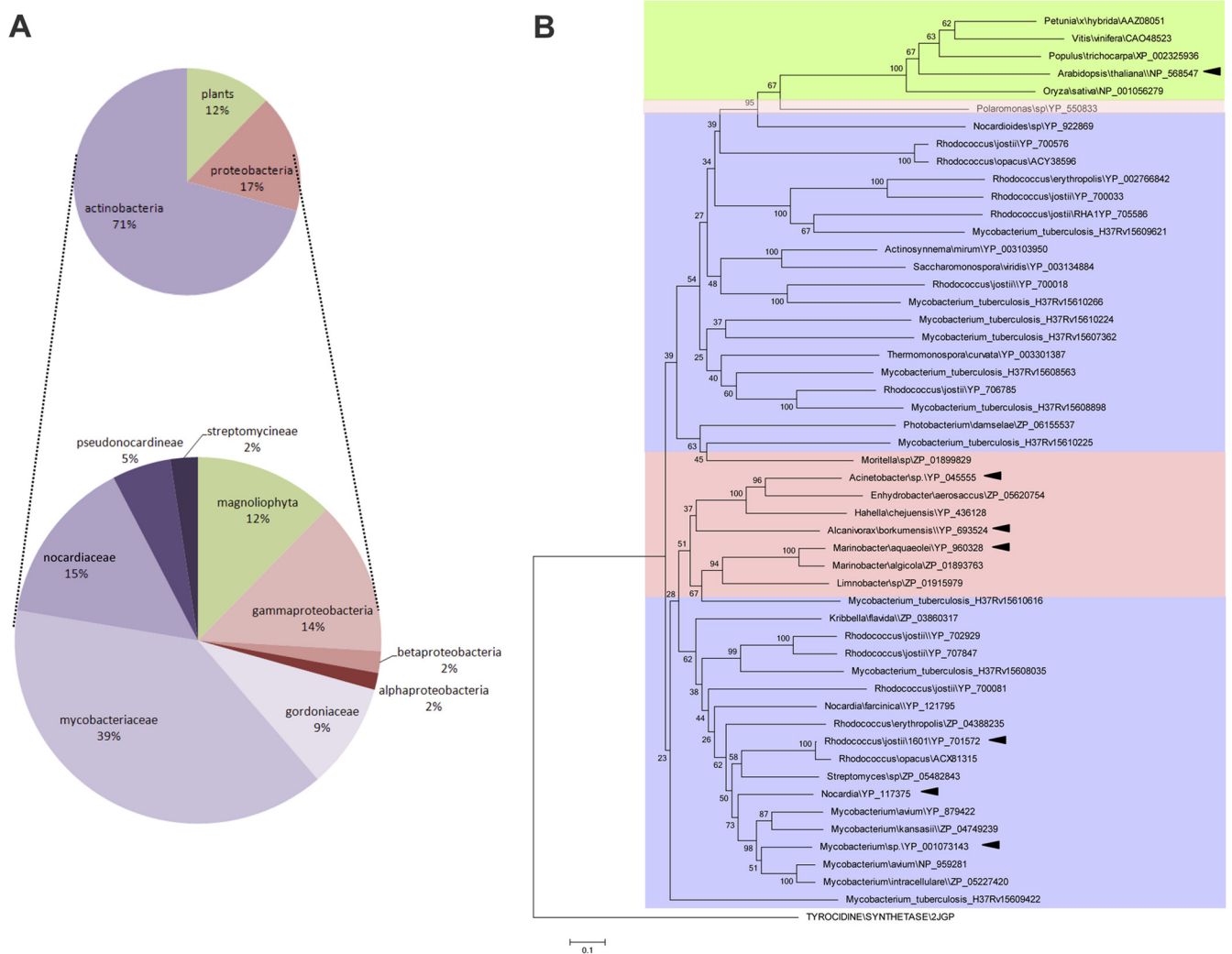
FIG 1 WS/DGAT family of CoA-dependent acyltransferases. (A) Distribution of the WS/DGAT proteins found by BLASTP. (B) Phylogeny of the WS/DGAT family of acyltransferases. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches. The evolutionary distances are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. Plant enzymes are highlighted in green, actinobacterial enzymes in blue, and proteobacterial enzymes in red. The accession number of each enzyme is shown. The sequence of the unrelated acyltransferase TycC was used to root the tree. Sequences used in the multiple alignment shown in Fig. 2 are marked by black arrowheads.

C41(DE3) cells. However, we were not able to express the 246–473 fragment cloned into pET29c. Four other constructions containing the predicted C-terminal domain were tested (178–465, 178–473, 216–465, and 216–473), and we were able to overexpressed only the 178–473 fragment (plasmid pETMa2Δ178N) (data not shown). Although according to the structural prediction both domains should fold independently, removal of the helical linker somehow affects the stability of the C-terminal domain.

As shown in Fig. 4A, both domains are soluble when coexpressed and could be purified using a Ni-NTA column. The resultant complex formed by the 1–176 and 178–473 fragments eluted as a single complex with an elution volume corresponding to a molecular mass of 100 kDa in S200 gel filtration chromatography (data not shown). The elution volume of the purified Ma2HT protein also corresponded to a molecular mass of 100 kDa. This value is in agreement with the molecular mass of a homodimer Ma2 protein as described for other WS/DGAT acyltransferases (16).

The activity of the reconstituted enzyme was checked *in vitro* using a spectrophotometric coupled assay with the Ellman's reagent (DTNB) (see Materials and Methods). The specific activities were measured with palmitoyl-CoA as the acyl-CoA and hexadecanol or dipalmitoyl glycerol (DPG) as the alcohol. Hexadecanol was used for the measurement of WS activity, and dipalmitoyl glycerol was used for DGAT activity. As shown in Fig. 4B, the activity of the purified N and C complex was in the same range as the activity of wild-type (wt) Ma2, which we have used as a control.

As mentioned in the introduction, there is a growing interest in finding natural or modified proteins able to produce intracellular biodiesel (microdiesel) in an effective manner. Biodiesel (FAEEs) is produced by the esterification of a small alcohol (ethanol or methanol) with a fatty acid. Thus, we have also checked the activity of wt Ma2 with different small alcohols (ethanol, methanol, butanol, and glycerol). Very poor activity of wt Ma2 was found in any case (see Fig. S2 in the supplemental material). The same poor
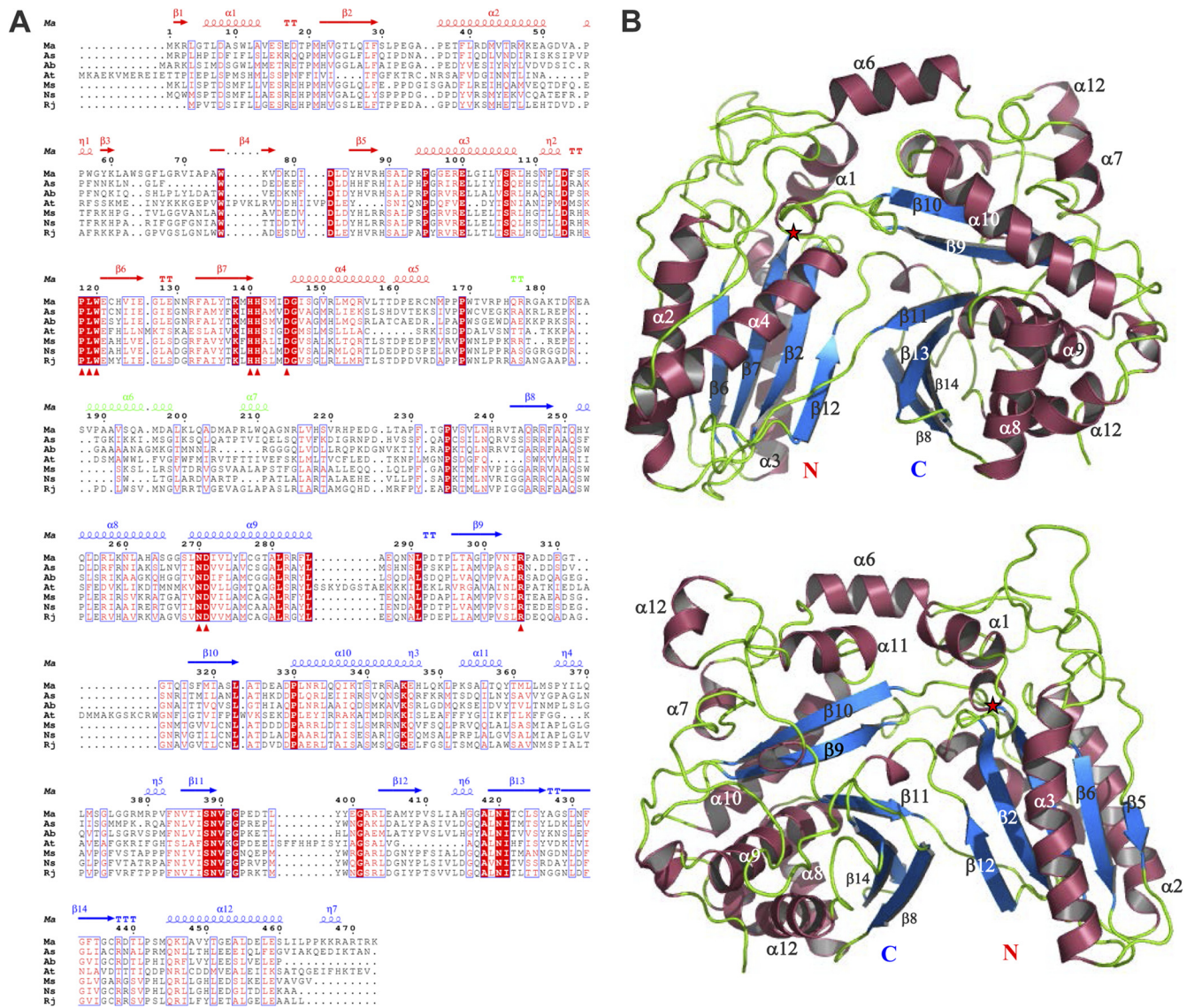
FIG 2 (A) Multiple sequence alignment of representative WS/DGAT proteins from different organisms as described in Results. Identical residues are shown in white on a red background, while similar residues are shown in red. The secondary structure elements of the modeled Ma2 protein are shown above the alignment. Secondary structure representation is colored red for the N-terminal domain, blue for the C-terminal domain, and green for the connecting helices. (B) Structural prediction analysis of the Ma2 protein revealed a monomer with a two-domain structure: beta sheets are shown in blue, alpha helices in magenta, and connecting loops in green. The position of the active site motif HHxxxDG is shown by a red star. The acyl-CoA binding face or front face (top) and the fatty alcohol/DAG binding face or back face (bottom) are shown.

activity for the production of FAEEs was found for the reconstituted complex (Fig. 4B).

The WS/DGAT HHxxxDG catalytic motif is required for *in vitro* acyltransferase activity. The main conserved motif in WS/DGAT enzymes is the HHxxxDG motif, also present in other acyl–CoA-dependent acyltransferases (Fig. 2A) (26).

To study the requirement of the active site residues of WS/DGAT enzymes in the catalytic reaction, we constructed H140A, H141A, and D145A Ma2 mutants by site-directed mutagenesis of the pETMa2 plasmid (see Materials and Methods). We compared the activities of the mutants and the wild-type enzyme by using a spectrophotometric coupled assay of the purified proteins (Fig. 5).

Our results showed that the mutation H141A resulted in a defective enzyme: Ma2 H141A is 60-fold less active than wild-type

Ma2 with hexadecanol and 5-fold less active with dipalmitoyl glycerol (Fig. 5).

H140A and D145A Ma2 mutants have also been shown to be critical for activity. The activity of the purified D145A mutant was similar to the activity of H141A mutant, while the activity of the H140A mutant was slightly higher but in any case at least 10 times lower than the activity of wt Ma2.

WS/DGAT motifs involved in the proper conformation of the active site are also essential for acyltransferase activity. Together with the conserved residues HHxxxDG, involved in the direct catalysis of the acyltransferase reaction, there are other conserved motifs characteristic of WS/DGAT enzymes (51). The conserved residues ^{118}PLW^{120} form motif I (Fig. 2A). Mutation of the residue P118 to alanine reduced catalysis to the same extent as the
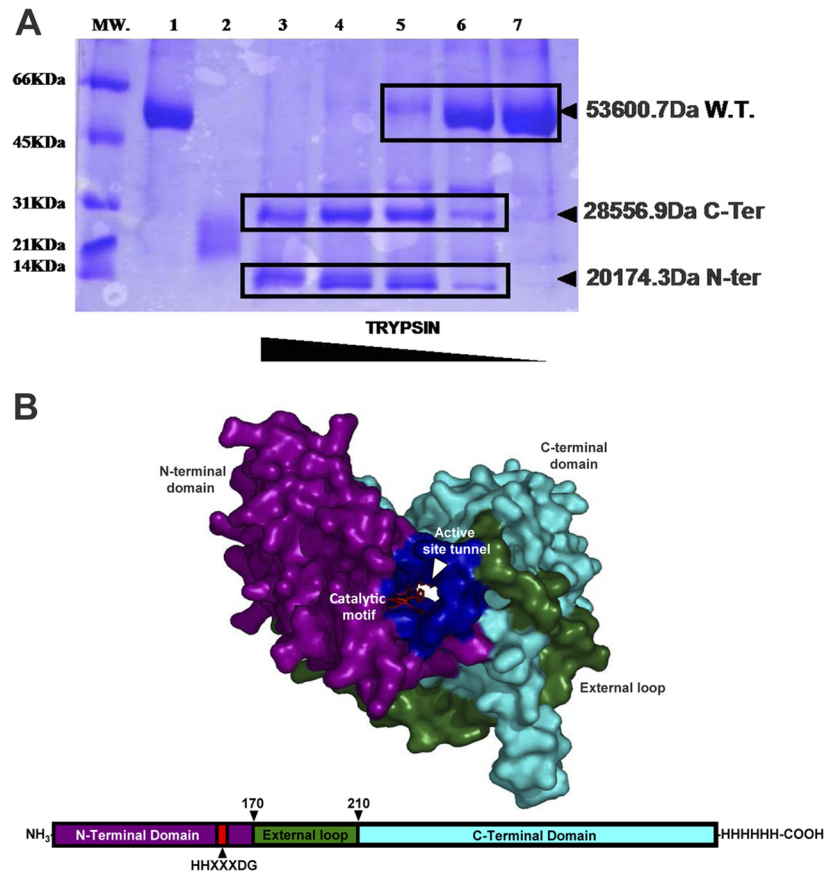
FIG 3 Limited trypsin digestion of Ma2. (A) Twelve percent SDS-PAGE gel containing 15 μM samples of Ma2 treated for 30 min at 37°C with different amounts of trypsin and subjected to electrophoresis at 200 V for 45 min. Lane 1, Ma2 without trypsin treatment; lane 2, 130 μM trypsin without Ma2; lanes 3 to 7, Ma2 treated with 130 nM, 65 nM, 13 nM, 6.5 nM, or 1.3 nM trypsin, respectively. Positions of bands corresponding to the wt Ma2 protein and the C and N-terminal domains are shown at left. The corresponding molecular masses of the fragments, as determined by MALDI-TOF mass spectrometry, are also indicated. (B) Locations of the C-terminal (magenta) and N-terminal (cyan) domains in Ma2 modeled structure and in the primary sequence. The position of the active site motif HHxxxDG is shown.

essential catalytic histidine H140 (Fig. 5). Purified Ma2 L119A mutant activity was also reduced compared to wild-type Ma2 activity. The Ma2 W120A mutant was overexpressed in C41(DE3) cells, but the resulting protein was insoluble.

Residues Asn270 and Asp271 form motif II. Mutation of either N270 or D271 to Ala significantly reduced the activity of the protein.

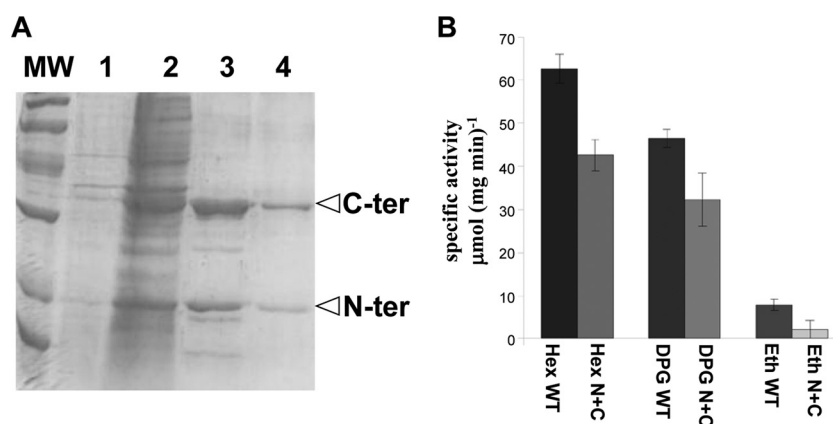In most of the solved structures of HHxxxDG acyltransferases,

FIG 4 Interaction of Ma2 N- and C-terminal domains. (A) Twelve percent SDS-PAGE gel showing the induction and binding assay in a Ni-NTA agarose resin of C41 cells transformed with both the pETMa2Δ178N and pETMa2Δ297C plasmids. Lane MW, molecular mass standards; lane 1, no induced fraction; lane 2, induced culture; lane 3, fraction after His-trap Ni agarose resin; lane 4, fraction after a Superdex 200 HR 10/30 column chromatography. (B) Activity of Ma2 wild type (WT) and the reconstituted complex (N+C) with palmitoyl-CoA and hexadecanol (Hex), 1,2-dipalmitoyl glycerol (DPG), or ethanol (Eth). Values are averages for three experiments.
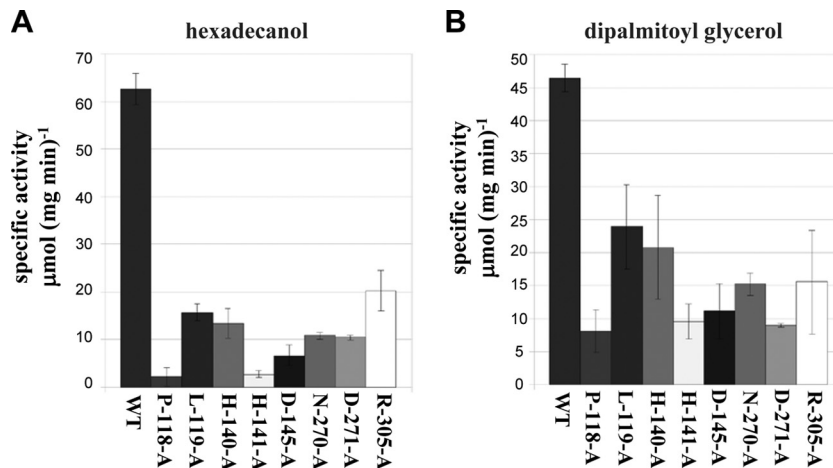
FIG 5 Activities of different Ma2 mutants for ester production using palmitoyl-CoA and hexadecanol (A) or DAG production using palmitoyl-CoA and dipalmitoyl glycerol (B). Values are averages for three experiments.

there is a conserved arginine interacting with the aspartic acid of the consensus motif. In the alignment of the WS/DGAT acyltransferases (Fig. 2A), we have found that there is a conserved arginine (R305) in the loop between β9 and β10. R305 was mutated to alanine to check the role of this residue in protein activity. As shown in Fig. 5, the protein Ma2R305A has reduced activity *in vitro*.

To analyze the position and role in catalysis of these conserved residues, we have compared the model of Ma2 from *M. hydrocarbonoclasticus* with the structure of hydroxycinnamoyltransferase (HCT) from sorghum bound both to shikimate (acyl acceptor) and *p*-coumaroyl-CoA (acyl donor) (Fig. 6). HCT is the closest homolog to WS/DGAT, whose structure has been solved, bound to both the acyl acceptor and the acyl donor (52). Thus, we have been able to model the acyl-CoA molecule bound to Ma2 (Fig. 6).

DISCUSSION

Phylogenetic analysis of the WS/DGAT family, found in bacteria and plants, showed that bacterial and plant WS/DGATs are located in two monophyletic subclades (Fig. 1B). Within the bacterial WS/DGAT enzymes, two groups have also been found: acti-

nobacteria and proteobacteria. Proteobacterial WS/DGATs show evolutionary proximity. However, actinobacterial WS/DGAT paralogs are dispersed along the bacterial clusters (53). It was recently found that there are differences in the substrate selectivities of two different *R. jostii* WS/DGAT enzymes (50). The finding of these different paralogs in actinobacteria could be due to differences in enzyme selectivity. Structural analysis of the family could help in the elucidation of the specificity factors for accepting a fatty alcohol or a diacylglycerol (DAG) molecule as a substrate.

Molecular threading and limited proteolysis showed that proteins of the WS/DGAT family structurally belong to the CoA-dependent acyltransferases family. WS/DGAT enzymes, like other CoA-dependent acyltransferases, are composed of two structural domains (N- and C-terminal domains) with similar fold (Scop: CoA-dependent acyltransferase fold) and size. Using a spectrophotometric coupled assay, we have found that the complex formed by the two domains coexpressed as independent polypeptides is active *in vitro* (Fig. 4). The slightly lower activity of the reconstituted complex could be due to some dissociation.

Not only WS activity but also significant DGAT activity was found using the coupled assay for purified wild-type Ma2 and for
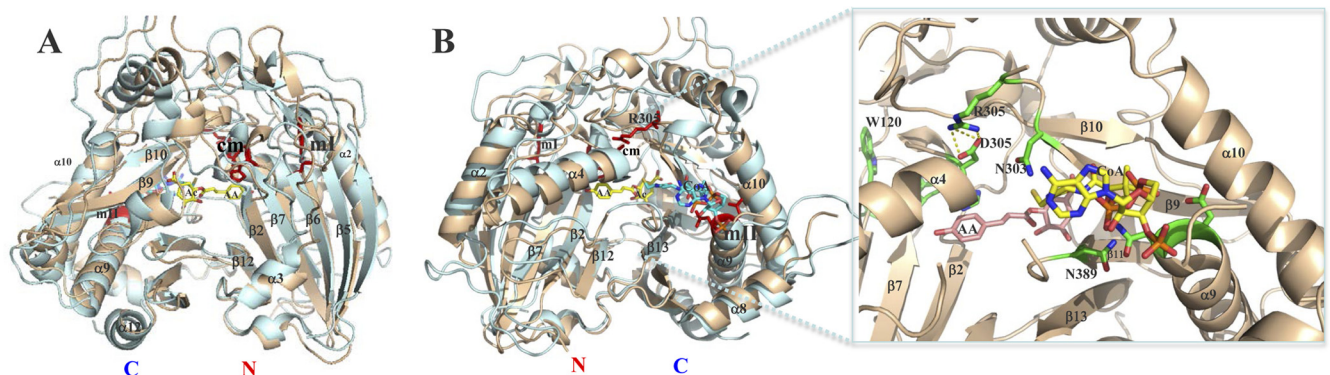

FIG 6 Structural alignment of Ma2 and HCT. Back (A) and front (B) views of the alignment are shown. The Ma2 model is shown in wheat color, and the HCT structure (4KEC.pdb) is shown in cyan. Important secondary structural elements are highlighted, including the catalytic motif (cm), motifs I (mI) and II (mII), CoA, acyl group (Ac), and acyl acceptor group (AA). A zoomed view of the Ma2 catalytic site modeled with CoA and the acyl and acyl acceptor groups is also shown.

the reconstituted complex (Fig. 4B). This DGAT activity of Ma2 was not observed before using a less-sensitive thin-layer chromatography (TLC) assay (20).

We have also found conserved residues essential for Ma2 activity (Fig. 5) that could be responsible for the WS/DGAT activity. In the HHxxxDG acyltransferases, the second histidine (His141 in Ma2) has been proposed to be the only catalytic residue that acts as a general base to extract the proton from the hydroxyl group of the alcohol to catalyze the nucleophilic attack (30), although both histidines were shown to be essential for wax ester synthase/acyl-CoA:diacylglycerol acyltransferase catalysis in *A. baylyi* (26). In Ma2, alanine scanning not only shows that the H141A mutant is poorly active *in vitro* but that it is one of the most deleterious mutants of all the mutants tested, thus confirming the essential catalytic role of the second histidine in WS/DGAT. H140 was also shown to be important for catalysis, since the H140A mutant is 10-fold less active than wt Ma2.

Although it had been proven that the aspartic acid of the HHxxxDG motif is essential for enzyme activity in acyl-CoA-dependent acyltransferases, playing a structural role in the organization of the active site (54), the mutation of the aspartic residue in acinetobacter WS/DGAT did not affect the activity of the protein (26). However, Asp145 in Ma2 again seems to be essential for both wax and TAG synthesis *in vitro* (Fig. 5). Thus, despite having the same fold, some differences in the catalytic residues could appear between different homologs of the WS/DGAT family.

According to these results, we propose that, as in some other acyltransferases, the second histidine in WS/DGAT is the catalytic histidine, while the aspartic acid and the first histidine play a structural role in the organization of the active site.

Together with the catalytic HHxxx**D**G motif, we have found some other conserved motifs involved in WS/DGAT activity. The conserved residues [118]PLW[120] that form motif I have been found to be involved in protein activity according to the alanine scanning results. However, we do not think the role of these residues is specific to the diacylglycerol acyltransferase reaction, since similar residues are also present in other acyl-CoA-dependent acyltransferases, such as HCT ([140]PLL[142]), TycC 2JPG ([202]PLV[204]), or VibH 1L5A ([104]PIT[106]). Motif I is located at the N-terminal end of β6, proximal to the C-terminal end of β7, where the catalytic His resides (Fig. 6A). This suggests a structural function of motif I in the right folding of the catalytic center.

The essential role of R305 (Fig. 5) is also consistent with the CoA-dependent acyltransferase fold, where a conserved arginine was always found in the corresponding loop connecting β9 and β10. This conserved arginine binds the aspartic acid of the active site to position the HHxxxDG catalytic loop (48, 52). However, this important role of the conserved arginine had not been demonstrated by mutagenesis. In our study, the poor acyltransferase activity of Ma2R305A is similar to the activity of Ma2D144A, as expected if the main role of R305A is the right positioning of D144.

Finally, motif II is located at the N-terminal end of α9. Comparing our WS/DGAT structure with the structure of HCT bound to acyl-CoA, we observed that motif II Asn270 and Asp271 in Ma2 could be interacting with the phosphate groups of the acyl-CoA molecule (Fig. 6) in the same way as S268 and Thr269 in HCT.

Moreover, WS/DGAT alignment with HCT also showed that

the oxyanion of the modeled CoA is close to the side chain hydroxyl of the conserved residues Ser388 and N389 at the C terminus of β11 (Fig. 6). Thus, S388 and N389 could be the residues interacting with the acyl-CoA oxyanion for a proper binding of acyl-CoA. In fact, a point mutation in the equivalent serine of WS/DGAT AtfA (S374P) resulted in significant reduction of activity in *A. baylyi* (51).

The essential role of R305 and the residues forming motifs I and II is consistent with the acyl-CoA-dependent acyltransferase fold. However, despite the essential role of the found WS/DGAT motifs in the catalysis of the acyltransferase reaction, none of them seems to be involved in selectivity. WS/DGAT motifs are conserved in all the different subfamilies, and identical or similar residues have been found in other acyltransferases not belonging to the WS/DGAT family.

Thus, the structural characterization presented in this article identifies two folding domains and sequence motifs critical for WS/DGAT acyltransferase activity. This identification facilitates the understanding of this family of proteins with a wide biotechnological potential for the production of TAGs or wax esters.

REFERENCES

1. **Shi S, Valle-Rodríguez JO, Siewers V, Nielsen J.** 2011. Prospects for microbial biodiesel production. Biotechnol. J. **6:**277–285. http://dx.doi.org/10.1002/biot.201000117.
2. **Ratledge C, Cohen Z.** 2008. Microbial and algal oils: do they have a future for biodiesel or as commodity oils? Lipid Technol. **20:**155–160. http://dx.doi.org/10.1002/lite.200800044.
3. **Cohen Z, Ratledge C.** 2010. Single cell oils: microbial and algal oils, 2nd ed. AOCS Publishing, Urbana, IL.
4. **Steinbüchel A, Hustede E, Liebergesell M, Pieper U, Timm A, Valentin H.** 1992. Molecular basis for biosynthesis and accumulation of polyhydroxyalkanoic acids in bacteria. FEMS Microbiol. Lett. **103:**217–230. http://dx.doi.org/10.1111/j.1574-6968.1992.tb05841.x.
5. **Daniel J, Deb C, Dubey VS, Sirakova TD, Abomoelak B, Morbidoni HR, Kolattukudy PE.** 2004. Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in *Mycobacterium tuberculosis* as it goes into a dormancy-like state in culture. J. Bacteriol. **186:**5017–5030. http://dx.doi.org/10.1128/JB.186.15.5017-5030.2004.
6. **Kaddor C, Biermann K, Kalscheuer R, Steinbüchel A.** 2009. Analysis of neutral lipid biosynthesis in *Streptomyces avermitilis* MA-4680 and characterization of an acyltransferase involved herein. Appl. Microbiol. Biotechnol. **84:**143–155. http://dx.doi.org/10.1007/s00253-009-2018-4.
7. **Alvarez HM, Kalscheuer R, Steinbüchel A.** 2000. Accumulation and mobilization of storage lipids by *Rhodococcus opacus* PD630 and *Rhodococcus ruber* NCIMB 40126. Appl. Microbiol. Biotechnol. **54:**218–223. http://dx.doi.org/10.1007/s002530000395.
8. **Alvarez HM, Steinbüchel A.** 2002. Triacylglycerols in prokaryotic microorganisms. Appl. Microbiol. Biotechnol. **60:**367–376. http://dx.doi.org/10.1007/s00253-002-1135-0.
9. **Alvarez HM, Mayer F, Fabritius D, Steinbüchel A.** 1996. Formation of intracytoplasmic lipid inclusions by *Rhodococcus opacus* strain PD630. Arch. Microbiol. **165:**377–386. http://dx.doi.org/10.1007/s002030050341.
10. **Gouda MK, Omar SH, Aouad LM.** 2008. Single cell oil production by *Gordonia* sp. DG using agro-industrial wastes. World J. Microbiol. Biotechnol. **24:**1703–1711. http://dx.doi.org/10.1007/s11274-008-9664-z.
11. **Alvarez HM, Pucci OH, Steinbüchel A.** 1997. Lipid storage compounds

in marine bacteria. Appl. Microbiol. Biotechnol. **47:**132–139. http://dx .doi.org/10.1007/s002530050901.

12. **Rontani J-F, Bonin PC, Volkman JK.** 1999. Production of wax esters during aerobic growth of marine bacteria on isoprenoid compounds. Appl. Environ. Microbiol. **65:**221–230.

13. **Röttig A, Steinbüchel A.** 2013. Acyltransferases in bacteria. Microbiol. Mol. Biol. Rev. **77:**277–321. http://dx.doi.org/10.1128/MMBR.00010-13.

14. **Kalscheuer R, Steinbüchel A.** 2003. A novel bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase mediates wax ester and tri-acylglycerol biosynthesis in *Acinetobacter calcoaceticus* ADP1. J. Biol. Chem. **278:**8075–8082. http://dx.doi.org/10.1074/jbc.M210533200.

15. **Kalscheuer R, Luftmann H, Steinbüchel A.** 2004. Synthesis of novel lipids in *Saccharomyces cerevisiae* by heterologous expression of an unspecific bacterial acyltransferase. Appl. Environ. Microbiol. **70:**7119–7125. http://dx.doi.org/10.1128/AEM.70.12.7119-7125.2004.

16. **Stöveken T, Kalscheuer R, Malkus U, Reichelt R, Steinbüchel A.** 2005. The wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase from *Acinetobacter* sp. strain ADP1: characterization of a novel type of acyltransferase. J. Bacteriol. **187:**1369–1376. http://dx.doi.org/10.1128/JB .187.4.1369-1376.2005.

17. **Uthoff S, Stöveken T, Weber N, Vosmann K, Klein E, Kalscheuer R, Steinbüchel A.** 2005. Thio wax ester biosynthesis utilizing the unspecific bifunctional wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase of *Acinetobacter sp.* strain ADP1. Appl. Environ. Microbiol. **71:**790–796. http://dx.doi.org/10.1128/AEM.71.2.790-796.2005.

18. **Sirakova TD, Dubey VS, Deb C, Daniel J, Korotkova TA, Abomoelak B, Kolattukudy PE.** 2006. Identification of a diacylglycerol acyltransferase gene involved in accumulation of triacylglycerol in *Mycobacterium tuberculosis* under stress. Microbiology **152:**2717–2725. http://dx.doi.org/10 .1099/mic.0.28993-0.

19. **Holder JW, Ulrich JC, DeBono AC, Godfrey PA, Desjardins CA, Zucker J, Zeng Q, Leach ALB, Ghiviriga I, Dancel C, Abeel T, Gevers D, Kodira CD, Desany B, Affourtit JP, Birren BW, Sinskey AJ.** 2011. Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. PLoS Genet. **7:**e1002219. http://dx.doi.org/10.1371/journal.pgen .1002219.

20. **Holtzapple E, Schmidt-Dannert C.** 2007. Biosynthesis of isoprenoid wax ester in *Marinobacter hydrocarbonoclasticus* DSM 8798: identification and characterization of isoprenoid coenzyme A synthetase and wax ester synthases. J. Bacteriol. **189:**3804–3812. http://dx.doi.org/10.1128/JB.01932-06.

21. **Kalscheuer R, Stöveken T, Malkus U, Reichelt R, Golyshin PN, Sabirova JS, Ferrer M, Timmis KN, Steinbüchel A.** 2007. Analysis of storage lipid accumulation in *Alcanivorax borkumensis*: evidence for alternative triacylglycerol biosynthesis routes in bacteria. J. Bacteriol. **189:**918–928. http://dx.doi.org/10.1128/JB.01292-06.

22. **King A, Nam J-W, Han J, Hilliard J, Jaworski JG.** 2007. Cuticular wax biosynthesis in petunia petals: cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. Planta **226:**381–394. http://dx .doi.org/10.1007/s00425-007-0489-z.

23. **Kalscheuer R, Stölting T, Steinbüchel A.** 2006. Microdiesel: *Escherichia coli* engineered for fuel production. Microbiology **152:**2529–2536. http: //dx.doi.org/10.1099/mic.0.29028-0.

24. **Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD.** 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature **463:**559–562. http://dx.doi.org/10.1038/nature08721.

25. **Lehner R, Kuksis A.** 1996. Biosynthesis of triacylglycerols. Prog. Lipid Res. **35:**169–201. http://dx.doi.org/10.1016/0163-7827(96)00005-7.

26. **Stöveken T, Kalscheuer R, Steinbüchel A.** 2009. Both histidine residues of the conserved HHXXXDG motif are essential for wax ester synthase/acyl-CoA:diacylglycerol acyltransferase catalysis. Eur. J. Lipid Sci. Technol. **111:**112–119. http://dx.doi.org/10.1002/ejlt.200800167.

27. **Stachelhaus T, Mootz HD, Bergendahl V, Marahiel MA.** 1998. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. J. Biol. Chem. **273:**22773–22781. http://dx.doi .org/10.1074/jbc.273.35.22773.

28. **D'Auria JC.** 2006. Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol. **9:**331–340. http://dx.doi.org/10.1016/j.pbi.2006 .03.016.

29. **Jogl G, Hsiao Y-S, Tong L.** 2004. Structure and function of carnitine acyltransferases. Ann. N. Y. Acad. Sci. **1033:**17–29. http://dx.doi.org/10 .1196/annals.1320.002.

30. **Lewendon A, Murray IA, Shaw WV, Gibbs MR, Leslie AG.** 1994.

Replacement of catalytic histidine-195 of chloramphenicol acetyltransferase: evidence for a general base role for glutamate. Biochemistry **33:** 1944–1950. http://dx.doi.org/10.1021/bi00173a043.

31. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402. http: //dx.doi.org/10.1093/nar/25.17.3389.

32. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792–1797. http://dx.doi .org/10.1093/nar/gkh340.

33. **Gouet P, Courcelle E, Stuart DI, Métoz F.** 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics **15:**305–308. http://dx.doi.org/10.1093/bioinformatics/15.4.305.

34. **Saitou N, Nei M.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.

35. **Felsenstein J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783. http://dx.doi.org/10.2307/2408678.

36. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731–2739. http://dx.doi.org/10.1093 /molbev/msr121.

37. **Zuckerkandl E, Pauling L.** 1965. Evolutionary divergence and convergence in proteins, p 97–166. *In* Bryson, V, Vogel, H (ed), Evolving genes and proteins. Academic Press, New York, NY.

38. **Kelley LA, Sternberg MJE.** 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. **4:**363–371. http://dx .doi.org/10.1038/nprot.2009.2.

39. **Kim DE, Chivian D, Baker D.** 2004. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. **32:**W526–W531. http://dx.doi.org/10.1093/nar/gkh468.

40. **Samel SA, Schoenafinger G, Knappe TA, Marahiel MA, Essen L-O.** 2007. Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase. Structure **15:**781–792. http://dx.doi.org/10.1016/j.str.2007.05.008.

41. **Sambrook J, Russell DW.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

42. **Miroux B, Walker JE.** 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Mol. Biol. **260:**289–298. http://dx.doi .org/10.1006/jmbi.1996.0399.

43. **Ellman GL.** 1959. Tissue sulfhydryl groups. Arch. Biochem. Biophys. **82:** 70–77. http://dx.doi.org/10.1016/0003-9861(59)90090-6.

44. **Riddles PW, Blakeley RL, Zerner B.** 1983. Reassessment of Ellman's reagent. Methods Enzymol. **91:**49–60.

45. **Tanovic A, Samel SA, Essen L-O, Marahiel MA.** 2008. Crystal structure of the termination module of a nonribosomal peptide synthetase. Science **321:**659–663. http://dx.doi.org/10.1126/science.1159850.

46. **Keating TA, Marshall CG, Walsh CT, Keating AE.** 2002. The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. Nat. Struct. Biol. **9:**522–526. http://dx .doi.org/10.1038/nsb810.

47. **Buglino J, Onwueme KC, Ferreras JA, Quadri LEN, Lima CD.** 2004. Crystal structure of PapA5, a phthiocerol dimycocerosyl transferase from *Mycobacterium tuberculosis*. J. Biol. Chem. **279:**30634–30642. http://dx .doi.org/10.1074/jbc.M404011200.

48. **Garvey GS, McCormick SP, Rayment I.** 2008. Structural and functional characterization of the TRI101 trichothecene 3-O-acetyltransferase from *Fusarium sporotrichioides* and *Fusarium graminearum*: kinetic insights to combating *Fusarium* head blight. J. Biol. Chem. **283:**1660–1669. http://dx .doi.org/10.1074/jbc.M705752200.

49. **Ma X, Koepke J, Panjikar S, Fritzsch G, Stöckigt J.** 2005. Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. J. Biol. Chem. **280:**13576–13583. http://dx.doi.org/10.1074/jbc .M414508200.

50. **Barney BM, Mann RL, Ohlert JM.** 2013. Identification of a residue affecting fatty alcohol selectivity in wax ester synthase. Appl. Environ. Microbiol. **79:**396–399. http://dx.doi.org/10.1128/AEM.02523-12.

51. **Röttig A, Steinbüchel A.** 2013. Random mutagenesis of atfA and screening for *Acinetobacter baylyi* mutants with an altered lipid accumulation. Eur. J. Lipid Sci. Technol. **115:**394–404. http://dx.doi.org/10.1002/ejlt .201200401.

52. **Walker AM, Hayes RP, Youn B, Vermerris W, Sattler SE, Kang C.** 2013.

Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme A-dependent transferases and synthases. Plant Physiol. **162:**640–651. http://dx.doi.org/10.1104/pp.113.217836.

53. **Barney BM, Wahlen BD, Garner E, Wei J, Seefeldt LC.** 2012. Differences in substrate specificities of five bacterial wax ester synthases.

Appl. Environ. Microbiol. **78:**5734–5745. http://dx.doi.org/10.1128/AEM.00534-12.

54. **Brown NF, Anderson RC, Caplan SL, Foster DW, McGarry JD.** 1994. Catalytically important domains of rat carnitine palmitoyltransferase II as determined by site-directed mutagenesis and chemical modification. Evidence for a critical histidine residue. J. Biol. Chem. **269:**19157–19162.