Title of Invention: Candida Antarctica Lipase A and Orthologues and Their Applications

[0001] The present application claims priority to Singapore patent application number 10202301529Y filed on 30 May 2023 titled “LIPASES” which is incorporated by reference herein in its entirety. A sequence listing of the amino acid sequences and nucleotide sequences described herein in xml format is provided with the present application and is incorporated herein by reference.

Technical Field

[0002] The present invention relates to Candida Antarctica Lipase A and orthologues and the use of these lipases.

Background

[0003] Structured/modified lipids are specifically created by altering the composition/positional distribution of the fatty acids (Berčíková et al., 2020). To synthesize structured lipids, enzymes, which offer selectivity, are preferred over chemical means (Osborn & Akoh, 2002). In addition, enzymatic synthesis also possesses greater potential for large-scale production in a greener manner, especially if the alternative chemical synthesis involves hazardous solvents / harsh conditions (Arslan et al., 2020; Zhang et al., 2021). Some examples of structured lipids are the estolides and human milk fat mimic (Cermak et al., 2017; Wei et al., 2020). To synthesize these lipids, promiscuous yet selective lipases are commonly used (Kapoor & Gupta, 2012).

[0004] A peculiar group of lipases, namely the Candida antarctica Lipase A (CalA) and CalA-like lipases, has garnered attention in the biotechnological setting due to its unique selectivity and marked thermostability (Monteiro et al., 2021). However, the applications and catalytic mechanism of the CalA and CalA-like lipases are not fully understood.

Summary

[0005] In an aspect, there is provided a method of producing a fatty acid ester. The method comprises providing a sn-2 selective lipase comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 22 (CL20), SEQ ID NO:

82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19),SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), SEQ ID NO: 101 (CL39), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 40 (CL38), SEQ ID NO: 100 (CL38), SEQ ID NO: 46 (CL44), SEQ ID NO: 106 (CL44), SEQ ID NO: 26 (CL24), SEQ ID NO: 86 (CL24), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 17 (CL15), SEQ ID NO: 77 (CL15), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 51 (CL49), SEQ ID NO: 111 (CL49), SEQ ID NO: 56 (CL54), and SEQ ID NO: 116 (CL54); and contacting the sn-2 selective lipase with a fatty acid and a secondary hydroxy fatty acid to produce the fatty acid ester. The fatty acid and the secondary hydroxy fatty acid may be the same or different.

[0006] In an embodiment, where the fatty acid and the secondary hydroxy fatty acid is the same, this may produce polymeric type fatty acid esters. Advantageously, this may be used as a lubricant.

[0007] In another embodiment, the fatty acid and the secondary hydroxy fatty acid are different and may be used to produce fatty acid ester at the secondary hydroxyl position. Palmitic esters of hydroxy stearic acids are believed to possess anti-inflammatory and anti-diabetic effects and these esters may be linked to disorders like diabetes and chronic inflammation. Similar other fatty acid hydroxy esters may have similar properties.

[0008] In an embodiment, the sn-2 selective lipase consists essentially of the amino acid sequence, preferably the sn-2 selective lipase consists of the amino acid sequence.

[0009] In an embodiment, the amino acid sequence is selected from the group consisting of SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19),SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), and SEQ ID NO: 101 (CL39).

[0010] In an embodiment, the fatty acid and the secondary hydroxy fatty acid are each independently a C14 to C22 fatty acid.

[0011] In an embodiment, the hydroxyl group in the secondary fatty acid is located at one of the following positions: C2, C10, C12, C15, and any combinations thereof. The

carboxylic acid in the fatty acid is C1 and the secondary hydroxyl group position is counted based on it.

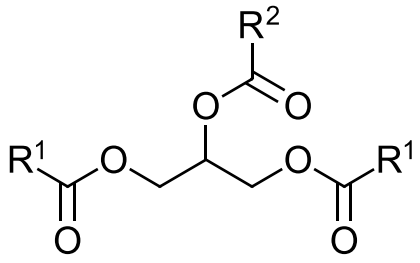
[0012] In an embodiment, the secondary hydroxy fatty acid is a C16 hydroxy fatty acid or a C18 hydroxy fatty acid, preferably the secondary hydroxy fatty acid is hydroxystearic acid.

[0013] In an embodiment, the fatty acid is a C16 fatty acid or a C18 fatty acid, preferably the fatty acid is palmitic acid.

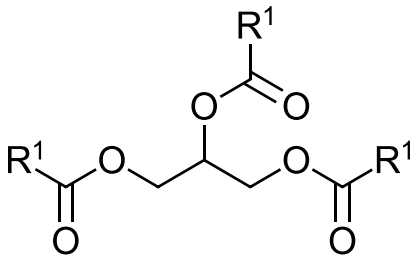
[0014] In an embodiment, providing the sn-2 selective lipase comprises expressing a corresponding nucleotide sequence in a host. Advantageously, the sn-2 selective lipase may be used directly in the bioreactor after the expression, or with minimal separation of the other components present in the bioreactor. The sn-2 selective lipase may be purified if required.

[0015] Advantageously, the method described allows a large variety of fatty acid hydroxy esters to be synthesized and used in a variety of different applications.

[0016] In an aspect, there is provided a method of producing a compound of Formula I. The method comprises providing a sn-2 selective lipase; and contacting the sn-2 selective lipase with a compound of Formula II and a long chain fatty acid to produce the compound of Formula I under suitable conditions. Formula I is



and Formula II is



, R^1 is a second alkyl tail of a second long chain fatty acid, R^2 is an alkyl tail of the long chain fatty acid, the long chain fatty acid and the second long chain fatty acid being different long chain fatty acids.

[0017] Advantageously, the method allows a fatty acid to be selectively attached to the middle hydroxyl group (2 position) in the triacylglycerides. The method provides alternative production pathways to produce a wide range of triacylglycerides and possibly new unusual triacylglycerides.

[0018] In an embodiment, the sn-2 selective lipase comprises an amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 22

(CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 2 (UMLA), SEQ ID NO: 62 (UMLA), SEQ ID NO: 1 (CalA), SEQ ID NO: 61 (CalA), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 26 (CL24), SEQ ID NO: 86 (CL24), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 41 (CL39), SEQ ID NO: 101 (CL39), SEQ ID NO: 17 (CL15), SEQ ID NO: 77 (CL15), SEQ ID NO: 46 (CL44), SEQ ID NO: 106 (CL44), SEQ ID NO: 51 (CL49), SEQ ID NO: 111 (CL49), SEQ ID NO: 33 (CL31), and SEQ ID NO: 93 (CL31). Preferably, the amino acid sequence is selected from the group consisting of SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 47 (CL45), and SEQ ID NO: 107 (CL45).

[0019] In an embodiment, the sn-2 selective lipase comprises the amino acid sequence. Preferably, the sn-2 selective lipase consists essentially of the amino acid sequence, more preferably the sn-2 selective lipase consists of the amino acid sequence.

[0020] In an embodiment, the long chain fatty acid and the second long chain fatty acid is a C14 to C22 fatty acid. Preferably, the long chain fatty acid and the second long chain fatty acid is a C16 fatty acid or a C18 fatty acid.

[0021] In an embodiment, the long chain fatty acid is palmitic acid, and the second long chain fatty acid is oleic acid, and the compound of Formula I is 1,3-olein-2-palmitin. In an embodiment, the method further comprises providing a sn-1,3 selective lipase; and contacting the sn-1,3 selective lipase with tripalmitin and oleic acid, wherein contacting the sn-1,3 selective lipase with tripalmitin and oleic acid and contacting the sn-2 selective lipase with triolein and palmitic acid is performed in a reactor. This allows the oleic acid and palmitic acid produced in their corresponding reactions to be used as the reactant in the other reaction and reduces the amount of waste products, thereby allowing for increased overall efficiency. The two reactions may preferably be done concurrently in a single reactor. Alternatively, it may be done in a telescoped sequence in the same reactor or in separate reactions.

[0022] In an embodiment, the long chain fatty acid is oleic acid, and the second long chain fatty acid is stearic acid, and the compound of Formula I is 1,3-distearoyl-2-oleoyl-glycerol.

[0023] In an embodiment, providing the sn-2 selective lipase comprises expressing a corresponding nucleotide sequence in a host. Advantageously, the sn-2 selective lipase may be used directly in the bioreactor after the expression, or with minimal separation of the other components present in the bioreactor. The sn-2 selective lipase may be purified if required.

[0024] In an aspect, there is provided an isolated sn-2 selective lipase comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), SEQ ID NO: 101 (CL39), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 40 (CL38), SEQ ID NO: 100 (CL38), SEQ ID NO: 46 (CL44), SEQ ID NO: 106 (CL44), SEQ ID NO: 26 (CL24), SEQ ID NO: 86 (CL24), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 17 (CL15), SEQ ID NO: 77 (CL15), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 51 (CL49), SEQ ID NO: 111 (CL49), SEQ ID NO: 56 (CL54), and SEQ ID NO: 116 (CL54).

[0025] Preferably, the amino acid sequence is selected from the group consisting of SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), and SEQ ID NO: 101 (CL39).

[0026] Preferably, the isolated sn-2 selective lipase of claim 20 consists essentially of the amino acid sequence, more preferably the lipase consists of the amino acid sequence.

[0027] In an embodiment, the lipase further comprises a purification tag bonded to either N-terminal or C-terminal of the amino acid sequence, preferably the N-terminal. Non-limiting examples of suitable purification tags include the 6x His tag sequence, like VMHHHHHH (SEQ ID NO: 183) and the Strep II tag.

[0028] In another aspect, there is described a method of separating, detecting and/or quantifying a triglyceride using a liquid chromatography-mass spectrometry (LC-MS) system comprising eluting a mixture comprising the triglyceride through a CHIRALPAK IG-3 column with a mobile phase comprising isopropanol and acetonitrile is used; and

detecting a molecular weight of each separated peak in the mixture. In an embodiment, the triglyceride is selected from the group consisting of triolein (OOO), 1,3-diolein-2-palmitin (OPO) isomers, 1,3-dipalmitoyl-2-olein (POP) isomers, tripalmitin (PPP), tristearin (SSS), 1,3-distearoyl-2-oleoylglycerol (SOS), 1,2-distearoyl-3-oleoylglycerol (SSO/OSS), 1,3-dioleoyl-2-stearoylglycerol (OSO),1,2-dioleoyl-3-stearoylglycerol (SOO/OOS) and trilaurin (LLL). In an embodiment, the mixture comprises at least two different triglycerides.

[0029] In another aspect, there is provided a method of separating, detecting and/or quantifying a fatty acid ester of hydroxy fatty acid (FAHFA) using a liquid chromatography-mass spectrometry (LC-MS) system comprising eluting a mixture comprising the fatty acid ester of hydroxy fatty acid through a C8-bonded silica column with a mobile phase comprising isopropanol and acetonitrile and a gradient of decreasing acetonitrile concentration over time when eluting the mixture; and detecting a molecular weight of each separated peak in the mixture. In an embodiment, the FAHFA is 12-hexadecanoyloxyoctadecanoic acid (12-PAHSA).

Description of Figures (FIGS.)

[0030] FIG. 1 shows the SSN of Pfam PF03585. On the left, the orthologues were separated into prokaryote/eukaryotes, at an alignment score of 75 and CalA is in the small cluster that was boxed up. On the right, the cluster of fungal lipases where CalA resides was further separated at an alignment score of 140. The Uniprot number for CalA is W3VKA4 while UMLA is A0A0D1DXG7.

[0031] FIG. 2 shows the expression vector for the expression of the CalA orthologues in *Pichia pastoris*. The RFP gene will be substituted with a combination of genetic elements including the AOX1 promoter, an aMF secretion tag, a gene of interest, and an AOX1 terminator. The antibiotic, Zeocin, was utilized as the selection marker for both *E. coli* and *Pichia pastoris*. The HIS4 gene encodes histidinol dehydrogenase, which allows the plasmid to serve as an auxotrophic marker in *Pichia pastoris* if the AOX1 promoter/terminator is not utilized for plasmid insertion.

[0032] FIG. 3 shows the hydrolytic activity of CalA and the orthologues against triolein as seen by the presence of dioleins on TLC. The numbers above the TLC plates indicate the lipase tested. The 16 orthologues with hydrolytic activity are indicated by the boxes in FIG. 3.

[0033] FIG. 4a shows the quantification of the sn-2 selectivity (hydrolysis) based on

LC-MS analysis with the ratio of the 1,2-diolefin to 1,3-diolefin. The higher the ratio the greater the hydrolysis of the 1,3-position.

[0034] FIG. 4b shows a comparison of the peaks in the LC corresponding to RML, GCL1, and CalA.

[0035] FIG. 5 shows the Tm and Tagg of CalA and selected orthologues. It can be seen that CalA has a high T_m and T_{agg} which corresponds to its good thermostability.

[0036] FIG. 6 shows a reaction scheme for the transesterification of a triacylglyceride at the 2-position.

[0037] FIG. 7a shows the OPO production using immobilized CalA and RML enzymes after 4 hours of reaction. For RML, most of the product is POO instead of OPO, indicating its high sn1,3 selectivity. Neg is the negative control which has no reaction. FIG. 7b shows the OPO production using immobilized CalA, CL20, CL23, and CL31. FIG. 7c shows the OPO to OOP/POO ratio of CL20, CL23 and CL31 relative to CalA.

[0038] FIG. 8 shows the LC chromatograms of TAGs and separation of OPO, POP, SOS and SSO isomers using Chiralpak-IG. Panel A on the left shows the separation of OOO, OPO isomers, POP isomers and PPP. Panel B on the right shows the separation of OOO, OSO isomers, SOS isomers and SSS.

[0039] FIG. 9a shows the esterification reaction between palmitic acid and hydroxystearic acid.

[0040] FIG. 9b shows the LC-MS detection of various compounds in less than 6 mins and the total ion chromatograms of the different analytes.

[0041] FIG. 10 shows the screening of selected CalA orthologues with n-hydroxylated stearic acid and palmitic acid at 50˚C, 70˚C and 90˚C.

[0042] FIG. 11 shows the conversion for 12-PAHSA over various time periods. For each lipase tested, the bar chart from left to right shows the percentage conversion at 5 mins, 10 mins, 20 mins, 40 mins, 80 mins, 160 mins, and 320 mins. Lifetech™ ECR1030M and Lifetech™ ECR8806M were used to immobilize the different enzymes.

Detailed Description

[0043] In the description herein, numerous specific details are set forth in order to provide a thorough understanding of various illustrative embodiments of the invention. It will be understood, however, to one skilled in the art, that embodiments of the invention may be practiced without some or all of these specific details. Embodiments described in the context of one of the methods or devices are analogously valid for the

other methods or devices. Similarly, embodiments described in the context of a method are analogously valid for a device, and vice versa.

[0044] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs.

[0045] As used herein, the articles “a”, “an” and “the” as used with regard to a feature or element include a reference to one or more of the features or elements. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items. As used herein, the terms “first,” “second,” and “third,” etc. are used merely as labels, and are not intended to impose numerical requirements on their objects.

[0046] Where a combination is disclosed, each subcombination of the elements of that combination is also specifically disclosed and is within the scope of the invention. Conversely, where different elements or groups of elements are disclosed, combinations thereof are also disclosed. Where any element of an invention is disclosed as having a plurality of alternatives, examples of that invention in which each alternative is excluded singly or in any combination with the other alternatives are also hereby disclosed; more than one element of an invention can have such exclusions, and all combinations of elements having such exclusions are hereby disclosed.

[0047] Where a range of values is recited, it is to be understood that each intervening integer value, and each fraction thereof, between the recited upper and lower limits of that range is also specifically disclosed, along with each subrange between such values. The upper and lower limits of any range can independently be included in or excluded from the range, and each range where either, neither or both limits are included is also encompassed within the invention. Where a value being discussed has inherent limits, for example where a component can be present at a concentration of from 0 to 100%, or where the pH of an aqueous solution can range from 1 to 14, those inherent limits are specifically disclosed. Where a value is explicitly recited, it is to be understood that values which are about the same quantity or amount as the recited value are also within the scope of the invention, as are ranges based thereon.

[0048] Although each of these terms has a distinct meaning, the terms “comprising”, “consisting of” and “consisting essentially of” may be interchanged for one another throughout the instant application. The term “having” has the same meaning as

“comprising” and may be replaced with either the term “consisting of” or “consisting essentially of”.

[0049] The terms “polypeptide” and “protein”, used interchangeably herein, refer to a polymer of amino acids without regard to the length of the polymer; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not specify or exclude chemical or post-expression modifications of the polypeptides of the invention, although chemical or post-expression modifications of these polypeptides may be included or excluded as specific embodiments. Therefore, for example, modifications to polypeptides that include the covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like are expressly encompassed by the term polypeptide. Further, polypeptides with these modifications may be specified as individual species to be included or excluded from the present invention. The natural or other chemical modifications, such as those listed in examples above can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Also included within the definition are polypeptides which contain one or more analogs of an amino acid (including, for example, non-naturally occurring amino acids, amino acids which only occur naturally in an unrelated biological system, modified amino acids from mammalian systems, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

[0050] The terms “sequence similarity”, “percentage of sequence identity” and “percentage homology” are used interchangeably herein to refer to comparisons among polynucleotides and polypeptides, and are determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Identity is evaluated using any of the variety of sequence comparison algorithms and programs known in the art. Such algorithms and programs include, but are by no means limited to, TBLASTN, BLASTP, FASTA, TFASTA, CLUSTAL W, FASTDB [Pearson and Lipman, (1988), Proc. Natl. Acad. Sci. USA 85(8):2444-2448; Altschul et al., (1990), J. Mol. Biol. 215(3):403-410; Thompson et al. (1994), Nucleic Acids Res. 22(2):4673-4680; Higgins et al., (1996), Meth. Enzymol. 266:383-402; Altschul et al., (1993), Nature Genetics 3:266-272; Brutlag et al. (1990) Comp. App. Biosci. 6:237-24], the disclosures of which are incorporated by reference in their entireties. Percentage sequence identity, also known as “% sequence identity” or “Seq ID to CalA (%) Without Met”, is retrieved by aligning the amino acid sequence without Methionine and without the innate secretion tag.

[0051] The term “alkyl” used alone or as part of a larger moiety, such as alkoxy, haloalkyl, arylalkyl, alkylamine, cycloalkyl, dialkyamine, alkylamino, dialkyamino alkylcarbonyl, alkoxycarbonyl and the like, includes as used herein means saturated straight-chain, cyclic or branched aliphatic group. The term “alkyl tail” refers to the alkyl portion of the fatty acid excluding the carboxyl group.

[0052] A long chain fatty acid as used herein refers to a carboxylic acid with an alkyl tail with 13 to 21 carbon atoms, in other words the fatty acid has 14 to 22 carbon atoms including the carboxylic acid. The long chain fatty acid may be saturated and unsaturated, odd or even numbered, branched or unbranched, and may be substituted or unsubstituted. Some non-limiting examples of long chain fatty acids include myristic acid (C14, 14:0), pentadecanoic acid (C15, 15:0), palmitic acid (C16, 16:0), heptadecanoic acid (C17, 17:0), stearic acid (C18, 18:0), arachidic acid (C20, 20:0),

behenic acid (C22, 22:0), myristoleic acid (C14, 14:1(9)), palmtoleic acd (C16, 16:1(9)), sapienic acid (C16, 16:1(6)), oleic acid (C18, !8:1(9)), elaidic acid (C18, 18:1(9t)), vaccenic acid (C18, 18:1(11t), linoleic acid (C18, 18:2(9, 12)), linoelaidic acid (C18, 18:2(9t, 12t)), α-linolenic acid (C18, 18:3(9, 12, 15)), arachidonic acid (C20, 20:4(5, 8, 11, 14)), eicosapentaenoic acid (C20, 20:5(5, 8, 11, 14, 17)), erucic acid (C22, 22:1(13), docosahexaenoic acid (C22, 22:6(4, 7, 10, 13, 16, 19)), and hydoxyfatty acids like 2-hydroxysteraric acid, 10-hydroxysteraric acid, 12-hydroxysteraric acid, 15-hydroxysteraric acid, and 2-hydroxypalmitiic acid. The long chain fatty acid may be referred to as CX fatty acid where X is the number of carbons in the fatty acid chain including the carboxylic acid.

[0053] A sn-2 selective lipase herein refers to a lipase which has a preference to react at the 2-position of a triacylglyceride or glycerol over the 1 and 3 positions, or more generally to form or cleave (hydorlyze) a fatty acid ester with a secondary hydroxyl group. Based on probability alone without considering steric or electronic factors, a non-selective lipase will have twice the probability to react at the 1 and 3 positions over the 2-position. Furthermore, due to steric the secondary hydroxyl group is considerably less reactive than the primary hydroxyl group. Hence, herein a sn-2 selective lipase is a lipase in which a ratio of a reaction (e.g. hydrolysis or esterification) at a primary hydroxyl group to a secondary hydroxyl group is 4:1 or lower, and specifically for a triacylglyceride a ratio of a reaction (e.g. hydrolysis or esterification) at the 1,3-hydroxyl group to 2-hydroxyl group is 8:1 or lower. Preferably, the ratios are 3:1 or lower and 6:1 or lower respectively, more preferably the ratios are 2:1 or lower and 4:1 or lower respectively, even more preferably the ratios are 1:1 or lower and 2:1 or lower respectively.

[0054] The following detailed description is merely exemplary in nature and is not intended to limit the invention or the application and uses of the invention. Furthermore, there is no intention to be bound by any theory presented in the preceding background of the invention or the following detailed description. Other technical advantages may become readily apparent to one of ordinary skill in the art after review of the following figures and description.

[0055] The biotechnological potential of CalA and its orthologues was investigated and is described herein. Using an enzyme similarity tool (EFI-EST), a group of 58 orthologues (CL1 to CL58) of fungal lipases were identified and characterized. The sequences of CalA and the 58 orthologues are shown in Table 3. SEQ ID NO: 1 and

SEQ ID NO: 61 is the peptide sequence of CalA without and with the signal peptide respectively. SEQ ID NO: 2 and 62 is the peptide sequence of *Ustilago maydis* lipase A (UMLA) without and with the signal peptide respectively. The peptide sequences of the CalA orthologues (CL1 to CL58) are SEQ ID NO: 3 to SEQ ID NO: 60 and SEQ ID NO: 63 to SEQ ID NO: 120 without and with the signal peptide respectively. There are similarities in terms of sequence and function among the CalA orthologues of the present embodiments. Of the 58 orthologues, 16 are capable of accepting long chain triacylglycerides and most of them are thermostable. A notable feature of CalA is its sn-2 selectivity, though not prominent (Monteiro et al., 2021). Investigation of the sn-2 selectivity of the orthologues revealed six are more selective than CalA. Six of the 16 orthologues identified have better sn-2 selectivity than CalA. A liquid chromatography mass spectrometry (LC-MS) method was developed to measure the selectivity based on the ratio of the 1,2-diolein and 1,3-diolein upon hydrolysis of triolein. The sn-2 selectivity of the lipase would potentially have implication as an alternative/complementary reaction for the production of 1,3-olein-2-palmitin (OPO) (also known as 1,3-Dioleoyl-2-palmitoylglycerol), which is a human milk fat mimic.

[0056] A conventional OPO synthesis requires an sn-1,3 selective lipase to carry out acidolysis reaction using tripalmitin (in palm stearin) and oleic acid (Ferreira & Tonetto, 2017). Palmitic acid is the waste product as it is replaced by oleic acid during the reaction (Ferreira & Tonetto, 2017). With an ideal sn-2 selective lipase, an acidolysis reaction using triolein (in high oleic acid sunflower oil) and palmitic acid, which is the waste product from the conventional synthesis, would also result in production of OPO. In other words, the sn-2 selective lipase would selectively perform the transesterification on the 2-position of the triacylglyceride over the 1,3-position of the triacylglyceride as shown in **FIG. 6**. Thus, with these two methods, circular bioeconomy can be achieved by minimizing waste products (waste of both methods becomes the feed of the alternate method) (Leong et al., 2021). If both reaction methods can be done in a single reactor (one pot) the hydrolyzed fatty acid from one reaction can be used as the reagent for the other transesterification reaction, which greatly reduces the amount of waste products and improves the efficiency of the synthesis. Alternatively, the two methods may be done in a telescoped reaction without purification of the intermediates. To quantify the levels of OPO and other triacylglycerides in the reaction, an LC-MS method was developed using a chiral column which effectively separated the various triacylglycerides in less than 40 mins.

[0057] Another interesting property of the CalA lipase is its ability to accept sterically hindered alcohols (Zhang et al. 2021). This remarkable characteristic has allowed CalA to be the only commercially available lipase that can effectively synthesize fatty acid esters of hydroxy fatty acids (FAHFAs), which belong to the family of estolides (Riecan et al., 2022; Zhang et al., 2021). FAHFA is a new class of lipids discovered in 2014 (Yore et al., 2014) and show promising effects in the treatment of inflammation and diabetes mellitus (Yore et al., 2014). The discovery of CalA catalyzing such a reaction had greatly advanced the synthesis of such lipids (Zhang et al., 2021). This advancement is necessary since the functional role of the variation of such lipids remains unelucidated. Other than CalA, the orthologues of CalA may be capable of FAHFA synthesis and may open up possibilities for the synthesis of FAHFA of other variations. Two orthologues of CalA were found to be able to produce 12-hexadecanoyloxyoctadecanoic acid (12-PAHSA) in the presence of palmitic acid (PA) and 12-hydroxylstearic acid (12-HSA). An LC-MS method was developed to detect the various compounds in 5 mins as compared to the existing method in literature, which takes up to 40 mins using the gas chromatography mass spectrometry (GC-MS)(Zhang et al., 2021), and is thus more cost effective with the reduction of time spent and reagents used.

[0058] CL20, CL23, and CL31 each has less than 80% sequence identity to CalA. The expression of CL20 and CL23 is higher than that of CalA in *Pichia pastoris*. In addition, CL20 and CL31 exhibited good thermotolerance as they remained functional after heat treatment at 90ºC.

[0059] **Materials and Methods**

[0060] **Chemicals**

[0061] Triolein, tripalmitin, 1,3-dioleoyl-2-palmitoylglycerol, 1,2-diolein and 1,3-diolein were purchased from Sigma Aldrich (Steinheim, Germany). 1,2-dioleoyl-3-palmitoylglycerol, 1,3-dipalmitoyl-2-oleoylglycerol and 1,3-dipalmitoyl-2-oleoylglycerol were purchased from Larodan (Sweden). Acetonitrile (Hypergrade for LC-MS) and isopropanol (Hypergrade for LC-MS) were from Fisher Scientific (Schwerte, Germany). Formic acid for LC-MS was from Merck (Darmstadt, Germany). Nucleodur Gravity-SB C18 column (150 mm x 3 mm, 3 µm) was from Macherey-Nagel (Düren, Germany), Luna 3 µm C8(2) 100A, LC column 100 x 2 mm) was from Phenomenex (California, United States) and Chiralpak IG-3 (4.6 mm i.d. x 250mm ,3 µm) was from Daicel Corporation (Osaka, Japan).

[0062] Generating sequence similarity network (SSN) and target selection

[0063] EFI-EST (https://efi.igb.illinois.edu/efi-est/) was used to generate the SSN map and visualized using Cytoscape v3.7.2 (Zallot et al., 2019). Pfam PF03585 was used to generate the SSN as CalA (Lipase A from *Candida antarctica* / *Pseudozyma aphidis* (uniprot: W3VKA4), SEQ ID NO: 61) is from this protein family. The targets were between 107 to 653 residues long and were segregated into clusters with shared sequence identities of less than 35% among one another. The nodes were separated into their respective kingdoms (prokaryotes, eukaryotes or archaea), and the alignment score was reduced to 75 such that a smaller cluster of sequences from the same kingdom can be further explored. The genes in the cluster where CalA were codon optimized for expression in *Pichia pastoris* and synthesized and cloned into pFAi2 vector with AOX1 promoter, α-mating factor pre-region (aMFpre), and an AOX1 terminator. All genes were synthesized without the innate signal sequence. The nucleotide sequences for preparing the peptides are shown in SEQ ID NO: 121 to 180 corresponding to CalA, UMLA, and CL1 to CL58 respectively and is shown in Table 3.

[0064] Screening of active CalA orthologues

[0065] **FIG. 2** shows the expression vector for the expression of the CalA orthologues in *Pichia pastoris*. Proteins were expressed under an AOX1 promoter and secreted via the α-mating factor pre-region (aMFpre) with an AOX1 terminator. After sequencing, the expression vector was first linearized with *Sac* I restriction enzyme before being integrated to the AOX1 promoter region in the *P. pastoris* (GS115) genome by electroporation and plated on yeast peptone dextrose (YPD) agar with zeocin at 100 µg/mL.

[0066] To express the proteins, *Pichia* colonies were grown in buffered glycerol-complex (BMGY) medium overnight and transferred into buffered methanol-complex media (BMMY), with 10x more volume, the next day. The cultures were subsequently induced with 1% methanol every day for five consecutive days with ampicillin (100 ug/mL) added every other day to prevent bacterial contamination. The spent media with the secreted proteins were subsequently concentrated 10 times with a 10 kDa Amicon® Ultra Centrifugal Filter (Merck, Germany), and used for thin layer chromatography (TLC) assay reactions.

[0067] *Pichia pastoris* is generally regarded as safe, and being eukaryotes allow expressed protein to form disulfide bonds. Other host cells such as *Yarrowia lipolytica* and *Saccharomyces cerevisiae* are possible.

[0068] TLC offers a quick test to screen for lipases active against long chain triacylglycerides (TAG). 2 µL of triolein and 1 µL of concentrated media with the enzymes at 37°C for 2 hours was used in the screening assay. Thereafter, 24 µL of isopropanol was added and vortexed to stop and dilute the reaction. Subsequently, 1 µL of sample was spotted on the TLC plate and ran with hexane:diethyl ether:formic acid (80:20:2 v/v) solution and stained with iodine for visualization.

[0069] Expression and purification of secreted lipases capable of accepting long chain TAG

[0070] The signal peptides are usually processed (cleaved) by the host (*P. pastoris*) by signal peptidases. However, there are some instances where the cleavage is not entirely complete. The signal peptide generally does not affect the activity of the lipases.

[0071] To express and purify the active enzymes, a 6x His tag (VMHHHHHH, SEQ ID NO: 183) was added at the N-terminal of the lipase sequence by golden gate assembly into a pFAi2 expression vector. Likewise, proteins are secreted due to the aMFpre secretion tag. Any other suitable purification tags may be used. Expression of the proteins were carried out in 800 mL of BMMY culture to obtain sufficient purified proteins. To adjust the pH of the spent media, 40 mL of 2 M Tris solution (unadjusted pH) was added to reach a pH of 7-8. The medium was centrifuged at 6000 rpm with JLA 10.500 rotor in Beckman coulter centrifuge for 30 mins at 4°C and subsequently centrifuged using a JA25.5 rotor in a Beckman Coulter centrifuge at 20,000 rpm for 15 mins at 4°C. The supernatant was then filtered with 0.2 µm filter paper before being purified using the Akta pure 25 FPLC using a HisTrap HP, 5 mL column. The equilibration and/or wash buffer was made up of 20 mM Tris, pH 8, 100 mM NaCl, 40 mM imidazole and the elution buffer is made up of 20 mM Tris, pH 8, 100 mM NaCl and 250 mM imidazole. Proteins eluted were desalted with gel filtration. The concentration of the proteins was measured from the absorbance reading at 280 nm.

[0072] Any other affinity tags may be used such as the Strep II tag. However, in industrial production, the 6xHis tag may not be needed as the enzyme will be produced in the bioreactor, and subsequent purification of the protein is not needed/required.

[0073] The proteins used for characterization contain a Methionine (ATG) after *Pichia pastoris* cleave the secretion tag. The proteins used for sn-2 selectivity, thermostability, OPO and FAHFA production contains a 6x His tag sequence, VMHHHHHH (SEQ ID NO: 183) (GTGATGcatcatcaccaccatcac, SEQ ID NO: 184), at the N-terminal.

[0074] Determination of thermostability

[0075] The melting temperature (T_m) was used to measure the thermostability of the enzymes. Using differential scanning fluorimetry, protein unfolding due to increasing temperature can be detected based upon a fluorescence shift of the intrinsic tryptophan/tyrosine residues. Using UNcle (Unchained Labs), 9 µL of purified proteins (>1 mg/ml) were loaded into the quartz wells in triplicates. The wells were then loaded after being sealed within an aluminum thermal transfer holder. Samples were heated from 20°C to 95°C at a rate of 0.3°C/min. A fluorescence spectrum was obtained 0.8 times/s from 250 to 720 nm. The UNcle analysis software was used to calculate the T_m using the barycentric mean method.

[0076] *Sn*-2 selectivity and Detection

[0077] The regioselectivity of a lipase can be determined by hydrolysis with triolein (Robles-Machuca et al., 2019). If a lipase is non-regioselective, they would cleave the ester bonds of triolein regardless of the position, resulting in the formation of two molecules of 1,2-diolein and one molecule of 1,3-diolein (Cambon et al., 2008). However, an *sn*-2 lipase would result in a higher ratio of 1,3-diolein to 1,2-diolein upon hydrolyzing triolein. Similarly, an *sn*-1,3 lipase would produce a higher ratio of 1,2-diolein to 1,3-diolein. In existing methods, the diolein isomers were quantified by measuring the intensities of the TLC bands of the compounds after staining with copper acetate in water and phosphoric acid (50:50 v/v)(Robles-Machuca et al., 2019; Velasco-Lozano et al., 2012). Even though TLC offers a quick screen for the lipases' activity, it is less dependable than liquid chromatography mass spectrometry (LC-MS) because the latter offers higher sensitivity and reliability for the identification of individual lipid species (Cajka & Fiehn, 2014).

[0078] However, the separation of the 1,2-diolein and 1,3-diolein with LC-MS has been a challenge without a published or widely applicable method that enables the efficient separation of diolein isomers in LC-MS. A method from a recent publication that suggested the possibility of separating the diolein isomers was tested, but to no avail (Schick et al., 2021). Hence, a method was developed and described herein to effectively separate the diolein isomers in LC-MS. To measure the hydrolytic activity of the triolein-active lipases, 2 µL of 10 mg/mL triolein in hexane, 50 µL of 100 mM MOPS buffer, pH 7.2, 38 µL of water and 10 µL of enzymes at a concentration which result in less than 20% of the reaction was carried out in 15 mins at 40°C and agitated at 800 rpm. After the incubation, 200 µL of 0.9% aqueous sodium chloride was added,

followed by 200 µL of n-hexane, which was then subjected to vortexing and centrifugation at 15000 g for 2 mins. The top layer (in n-hexane) was extracted and transferred into a new tube. The process of adding hexane was repeated for another two more times before drying the hexane using a SpeedVac (Concentrator plus, Eppendorf, Hamburg, Germany). The content was then diluted in isopropanol:chloroform (49:1 v/v) for measurement.

[0079] The amount of diolein was measured using a liquid chromatography-mass spectrometry (LC-MS) system. The diolein may be a 1,2-diolein or 1,3-diolein. The parameters and elution gradient of the LC allows for more efficient elution of the diolein isomers and is set out in Table 1 below. The reversed phase chromatography on a Nucleodur Gravity-SB C18 column (150 mm x 3 mm, 3 µm) was performed using 0.1% formic acid in acetonitrile as solvent A and 1 mM ammonium formate in isopropanol as solvent B. The flow rate was 0.4 ml/mL throughout, and the temperature of the column was adjusted to 50°C. From 0 to 1 min, the column was equilibrated with 100% solvent A and then dropped to 85% solvent A from 1 to 1.2 min. From 1.2 min to 10 min, solvent A decreased to 83% and further to 82% after 2.5 mins. From 12.5 mins to 14 mins, solvent A decreased from 82% to 40%. Solvent A decreased further to 20% for 3 mins. From 17 to 18 mins, solvent A decreased to 10% and was maintained for another 2 mins. Finally, from 20 to 20.5 mins, solvent A increased back to 100%. The injection volume was 1-2 µL and measurement was performed in positive mode using DualAJS ESI. The gas temperature was at 280°C, drying gas rate at 15 l/min, nebulizer at 40 psi, sheath gas temperature at 300°C and sheath gas flow rate at 12 l/min.

Table 1. LC Parameters for Separation of 1,2 and 1,3-dioleins

	Method of purifying diolein
Solvent	(Solvent A) Acetonitrile with 0.1% formic acid, (Solvent B) 1 mM ammonium formate with isopropanol
Flow rate (50°C to reduce the pressure on the column)	0.4 ml/min, 50°C
Gradient	100% A (0-1 min) 85% A (1-1.2 min) 83% A (1.2-10 min)

	82% (10-12.5 min) 40% A (12.5-14 min) 20% A (14-17 min) 10% A (17-18 min) 10% A (18-20 min) 100% A (20-20.5 min)
Diolein eluted (detected)	Below 13 mins

[0080] Detection of Triglycerides using Liquid chromatography mass spectrometry (LC-MS)

[0081] The TAG solutions (2 µL) were injected into an Agilent 1290 Infinity UHPLC (Agilent Technologies, Waldbronn, Germany) with a CHIRALPAK IG-3 column (4.6 mm i.d. x 250mm ,3 µm) (Daicel Corporation) that was kept at 30°C during the entire analysis. The TAGs were separated isocratically using 98% of solvent A (acetonitrile) and 2 % of solvent B (90% isopropanol, 10% acetonitrile, 0.1% formic acid and 1 mM ammonium formate) at a flow rate of 0.5 mL/min for 35 mins. The run was extended to 60 mins for batches with SSS, SOS, SSO/OSS, OOS/SOO, and OSO. The LC system was coupled to an Agilent 6550 iFunnel Q-TOF (Agilent Technologies) mass spectrometer equipped with a DualAJS ESI (Agilent Technologies, Waldbronn, Germany) for the detection and quantitation of the TAGs. The MS was set to positive mode for detection and quantification. The drying gas temperature was set at 280°C, drying gas rate at 15 L/min, nebulizer at 40 psi, sheath gas temperature at 300°C and sheath gas flow rate at 12 L/min. To quantify, m/z of PPP, OPP/PPO/POP, OPO/POO/OOP, OOO, OSO/OOS/SOO, SOS/SSO/OSS, and SSS are, 824.7702, 850.7858, 876.8015, 902.8171, 904.8328, 906.8484 and 908.8641 respectively.

[0082] The use of the ChiralPak IG-3 column provides high accuracy of the quantification of OOP, OPO and POO. Advantageously, the TAG LC-MS method of the present embodiments result in better peak separation compared to existing methods.

[0083] Screening of triolein-active CalA-like orthologues for 2-, 10-, 12- and 15-PAHSA

[0084] To screen for lipases that can produce FAHFAs, the orthologues were expressed under the constitutive GAP promoter with aMFPre secretion tag in BMGY Zeo at 30°C, 800 rpm with 70 pulse/sec in a deep-well block in a Glascol microplate shaker for four days. The supernatant containing the enzyme was pre-heated at 50°C, 60°C or 70°C for one hour before the reaction at the respective temperature for one hour at 800 rpm using the Eppendorf SmartBlock DWP 1000 (Eppendorf, Germany). Following pre-heating, 2.5 µL of 10 mg/mL palmitic acid and 5 mg/mL n-hydroxystearic acid (HSA) (n= 2, 10, 12 or 15) dissolved in DMSO were added to 50 µL of supernatant. After the reaction, the deep-well block was centrifuged, and heptadecanoic acid in isopropanol was added as internal standard. It was subsequently placed in the epMotion® 5070 (Eppendorf, Germany) for automated extraction. It was observed that 2-HSA exhibited limited solubility in commonly used organic solvents such as acetonitrile, dichloromethane, toluene, chloroform, and ethyl acetate with solubility not exceeding 1 mg/mL. However, it is soluble in methanol. Hence, the BUME method for the extraction process was used with some modifications (L. Löfgren et al., 2012). The BUME method enabled liquid-liquid extraction, and this allows the separation of the compounds of interest from the supernatant containing the enzymes. For each reaction, 150 µL of butanol:methanol (3:1) mixture was added to the 1 mL deep well block and agitated at 1000 rpm for 5 minutes. Subsequently, 150 µL of heptane:ethyl acetate (3:1) was added and was similarly agitated at 1000 rpm for 5 minutes. Next, 150 µL of 0.9% sodium chloride was added and the mixture was mixed by pipetting before sealing and centrifuging at 3000 x g for 1 min. After the centrifugation, a distinct bottom aqueous and top organic layer is observed. The deep-well block is then placed back into the epMotion® 5070 (Eppendorf, Germany), and 20 µL of the organic layer was added to the Greiner 96 well polypropylene plate (Greiner, Austria). 180 µL of isopropanol was added, thoroughly mixed, and prepared for detection using liquid chromatography mass spectrometry (LC-MS).

[0085] The FAFHA LC-MS method uses the C8 column, which can result in a runtime of 5 minutes or less. In comparison, the use of C18 column for the separation of lipids results in runtimes of up to 30 minutes. Therefore, the use of C8 column in the FAHFA LC-MS method of the present embodiments increases the throughput of experiments and is suitable for use in the field of enzyme engineering.

[0086] Lipase immobilization and OPO Biosynthesis

[0087] Lifetech™ ECR8806M was purchased from Purolite® ECR Enzyme Immobilization Resins (Wales, UK). A small-scale immobilization was carried out using 5 mg of the support in 50 µL of 40 mM potassium phosphate buffer at pH 7. Subsequently, 50 µL of the purified enzymes which were desalted was added, thus the final concentration of the potassium phosphate is 20 mM. The suspension was continuously agitated in a thermomixer (Eppendorf, Germany) at 18°C, 800 rpm overnight. The supernatant was removed, and protein concentration of the supernatant was determined to calculate the amount of protein immobilized. The semi-dried support was then lyophilized before reaction.

[0088] To synthesize OPO, a small-scale reaction of 10 uL (mixture of 4 uL of triolein: 1 mg palmitic acid) was added to the immobilized enzyme. The reaction was kept at 60ºC and continuously agitated at 800 rpm in a thermomixer (Eppendorf, Germany) for four hours. Subsequently, 250 uL of hexane was added to solubilize the substrate and products, which was then centrifuged at 10,000 g for 2 mins for the immobilization beads to settle. Then, 100 uL of solution was removed from the reaction tube to prevent further reaction. The content of the reaction was then diluted in isopropanol:chloroform (49:1 v/v) for measurement using the LC-MS

[0089] Bioreactor-scale Lipase expression

[0090] CalA, together with CL20, CL23 and CL31 with high expression or ability to tolerate high temperature were selected to be expressed in a 5 L Biostat® B bioreactor (Sartorius, Germany). The protocol developed by Invitrogen Corporation for protein expression in *P. pastoris* was adopted with some modifications (Invitrogen Corporation, 2022). For the glycerol fed batch, 100 mL of 50% glycerol was fed per 1 L of starter culture over six hours. When the dissolved oxygen spike, 1 L of methanol was fed per 1 L of starting fermentation volume for four days. As basal salts medium (BSM) was used, the linearized plasmid integrated in GS115 (histidine deficient type) consist of a functional HIS4 to ensure cell growth.

[0091] Lipase immobilization and 12-PAHSA Biosynthesis reaction

[0092] Lifetech™ ECR8806M and Lifetech™ ECR1030M were purchased from Purolite® ECR Enzyme Immobilization Resins (Wales, UK). The supernatant containing the enzymes from the bioreactor were concentrated and buffer-exchanged with 20 mM sodium phosphate pH 7, 100 mM NaCl. Immobilization was carried out using 80 ug of protein (measured by bradford assay) to a gram of immobilization beads. The suspension was continuously agitated at 4°C for a day. Subsequently, the

supernatant was removed, and the beads were washed with the buffer before lyophilization. After lyophilization, 20.1 (±0.17) mg of the beads were added to 62.5 mg of 12-hydroxystearic acid and 125 mg of palmitic acid. The mixture was then agitated at 70°C, 1000 rpm and approximately 1 uL of the mixture was taken out at various time points, diluted in hexane:isopropanol (1:1 v/v) and further diluted in isopropanol:chloroform (49:1 v/v), and injected into the LC-TQ for analysis.

[0093] Detection of Fatty Acyl Esters of Hydroxy Fatty Acid (FAHFA) using Liquid chromatography mass spectrometry (LC-MS)

[0094] A liquid chromatography multisampler, Agilent 1290 Infinity II Multisampler (G7167B) (Agilent Technologies, Waldbronn, Germany), coupled with an Agilent 6495 Triple Quadrupole (Agilent Technologies) mass spectrometer was used for the detection and quantitation of various compounds. AJS ESI (Agilent Technologies, Waldbronn, Germany) was used as the source to ionize the compounds. The MS was set to negative mode to detect the compounds. To separate the compounds, a Phenomenex Luna C8(2) 100A column (3um, 100 x 2 mm) was used. The mobile phases A and B are water:methanol (1:4 v/v) and isopropanol with 0.1% formic acid, respectively, and 1 µL of sample is injected. The flow rate was set at 0.4 mL/min throughout at 50°C. From 0 to 1 min, the column was equilibrated with 100% A before going down to 30% A thereafter from 1 to 5 mins, and then maintained for another 3 mins for the elution. The drying gas temperature was at 250°C, drying gas rate at 15 l/min, nebulizer at 50 psi, sheath gas temperature at 300°C and sheath gas flow rate at 11 l/min. The multiple reaction monitoring mode was used to measure the respective compounds as shown in Table 2 below.

[0095] Table 2.

Compound	Precursor ion	Product ion	Collision energy
HSAHSA	581.5	581.5	0
12-SAHSA	565.5	299.1	30
12-SAHSA	565.5	283.4	30
n-PAHSA	537.5	537.5	0
n-PAHSA	537.5	281	30
n-PAHSA	537.5	255.3	30
nHSA	299.3	281	20

nHSA	299.3	253	20
C17	269.2	269.2	0
C16	255.1	255.1	0

[0096] Results and discussion

[0097] Sequence similarity network (SSN) of fungal CalA-like lipases

[0098] *Candida antarctica* Lipase A (CalA) (SEQ ID NO: 61) has a highly divergent sequence as compared to other hydrolases and is often described as a "peculiar" enzyme (Monteiro et al., 2021; Robles-Machuca et al., 2019; Widmann et al., 2010). They are structurally different from the common lipases and only has 15% in sequence identity with the closest lipase family from *Bacillus subtilis* (Monteiro et al., 2021; Widmann et al., 2010). Interestingly, CalA has demonstrated selectivity and activity not commonly present in other hydrolases. For example, CalA is thermostable, exhibit a capability to accept highly sterically hindered alcohols (both secondary and tertiary), and *sn*-2 selectivity, though not pronounced (Monteiro et al., 2021; Robles-Machuca et al., 2019). These interesting properties made CalA relevant for use in the industries. However, the group of lipases which CalA belonged to is not well-characterized and understood. Characterization of CalA orthologues would be very useful to gain mechanistic insights for meaningful engineering of this peculiar family of lipases. Recently, *Ustilago maydis* lipase A (UMLA) SEQ ID NO: 62), which is similar to CalA was characterized (Robles-Machuca et al., 2019). Both lipases come from Pfam PF03585. To characterize and investigate the biotechnological potential of CalA orthologues, EFI-EST was used to create an SSN. Templates close to CalA and UMLA were selected as described in the methodology, and all of which are of fungal origin as shown in **FIG. 1**. There are 58 orthologues and is shown in Table 3.

[0099] Activity and Expression of CalA-like lipases

[0100] The 58 orthologues were expressed under the AOX1 promoter, aMFpre secretion tag and AOX1 terminator in a pFAi2 vector as shown in **FIG. 2**. Of the 58 orthologues tested, 16 orthologues (identified by the boxes in FIG. 3 and the presence of the diolein on the TLC) showed hydrolytic activity against triolein (long-chain triacylglyceride) as shown in **FIG. 3**. None of the lipases showed prominent *sn*-2 preference (single band of 1,3-diolein). Most active lipases, except for some interesting singletons such as CL13 and CL15, come from the largest cluster which is

similar to CalA and UMLA as shown in **FIG. 1**. The active lipases, with sufficient protein expression, were purified using a 6xHis tag at the N-terminal as the β-strands of CalA at the C-terminal could be a lid controlling substrate access in the lipase (Ericsson et al., 2008; Widmann et al., 2010). Out of 15 lipases, CL20 showed exceptionally higher lipase expression as compared to CalA, whereas CL21, 23 and 28 showed similar protein expression profiles. Purified lipases were verified for their activity against triolein and their *sn*-2 selectivity, measured using LC-MS. As shown in **FIG. 4a**, six of the orthologues (CL45, CL13, CL42, CL28, CL20, and CL21) showed enhanced *sn*-2 selectivity as compared to CalA. CL38 and CL54 were not tested. The LC-MS method effectively separated the 1,2-diolein and 1,3-diolein. The difference between the dioleins profile of the known enzymes, namely CalA (*sn*-2 selective), *Geotrichum candidum* lipase 1 (GCL1) (SEQ ID NO: 182) (non-regioselective) and *Rhizomucor miehei* Lipase (RML) (SEQ ID NO: 181) (*sn*-1,3 selective) is shown in **FIG. 4b** (Laguerre et al., 2017; Palla et al., 2012; Robles-Machuca et al., 2019).

[0101] Thermostability of CalA-like lipases

[0102] In industrial processes, a thermostable enzyme is imperative because enzymatic biosynthesis at high temperatures often result in better effectiveness and economic performance (Domínguez de María et al., 2005). To determine the thermostability of the lipases, the melting temperature (T_m) and the temperature at which the lipase aggregates (T_{agg}) were measured using fluorescence spectroscopy. Comparison of the parameters would allow us to assess the thermostability of the orthologues. The parameters used to measure thermostability of the CalA orthologues are shown in **FIG. 5**. All the lipases were less thermostable than CalA and two lipases which exhibit similar T_m and T_{agg} as CalA, namely CL21 and CL28, has high sequence identity to CalA (96.8% and 98.2% respectively) as compared to all the other tested orthologues. The T_m and T_{agg} of these orthologues ranged from 40ºC to 77ºC and 39ºC to 69ºC respectively.

[0103] **OPO Biosynthesis**

[0104] CalA was able to catalyse the reaction between palmitic acid and triolein at 60°C when immobilized. However, a mixture of triacylglycerides were produced as a result of the lack of selectivity as shown in **FIG. 7**a. However, as compared to a *sn*1,3-lipase (RML), CalA does show potential in the making of OPO if the selectivity can be improved (**FIG. 7a**). Our LC-MS method efficiently separated the various triglycerides involved in OPO biosynthesis using triolein and palmitic acid as a starting material as

shown in **FIG. 7b** and **FIG. 8 (panel A)**. Immobilized CL20, CL23, and CL31 were also able to catalyse the reaction between palmitic acid and triolein at 60°C as shown in **FIG. 7b** and **FIG. 7c**. In **FIG. 7c**, it may be observed that CL20 produces OPO in a higher ratio than CalA whilst CL23 and CL31 produces OPO in a lower ratio which is consistent with the data in **FIG. 4a**. In addition, the capability of the column to separate TAGs was extended to the enzymatic synthesis of cocoa butter substitutes, specifically 1,3-distearoyl-2-oleoyl-glycerol and similarly, good separation resolution was observed. The ability to effectively separate these TAG isomers and enantiomers provides a valuable tool for exploring the regioselectivity and enantioselectivity of lipases, offering significant utility for both academic research and industrial applications in the bioproduction of structured lipids.

[0105] FAHFA Biosynthesis (Screening)

[0106] Biosynthesis of FAHFA from the esterification reaction between free fatty acid and hydroxy fatty acid is believed to be optimal in the absence of water because esterification is likely a thermodynamically controlled process (Zhang et al., 2021). Hence, immobilization of lipases on supports, coupled with subsequent lyophilization would be necessary to remove the water. However, immobilization requires sufficient amount of enzymes, which is not easily attainable, at least in our study, for proteins expressed in a shake flask condition. As such, we have developed an assay to evaluate lipase suitability for FAHFA biosynthesis, taking into consideration factors such as expression yield, substrate solubility and availability, and the throughput as described in the material and method. To diversify the type of FAHFA which CalA and its orthologues can synthesize, we have chemically synthesized several variations of hydroxylated stearic acid for the screening as shown in **FIG. 9a**. To solubilize the substrates, we have opted to use dimethyl sulfoxide (DMSO) as it can effectively dissolve all the relevant substrates and is suitable to be used at higher temperatures (Byrne et al., 2016). Here, the presence of activity would suggest the lipase's potential to catalyze the biosynthesis of FAHFA qualitatively. Using this assay, we were also able to test a range of temperatures. Here, we tested the enzymes in three temperatures (50 °C, 70 °C, and 90 °C). This choice reflects the known optimal activity temperature of CalA (50 °C), the melting point of the substrates (palmitic acid and hydroxylated stearic acid) (70 °C), and the goal of identifying enzymes with potential for future applications involving substrates with even higher melting points (90 °C) (Domínguez de María et al., 2005; Monteiro et al., 2021; Robles-Machuca et al., 2019).

To evaluate the synthesis of FAHFAs, we have developed an LC-MS method to detect and quantify n-FAHFAs and the corresponding reactants in less than 8 mins as shown in **FIG. 9b**. This is an improvement from previous method using gas chromatography mass spectrometry (GC-MS) which takes more than 40 mins per run (Zhang et al., 2021).

[0107] In the screening for FAHFA biosynthesis, we have included all the CalA-like orthologues which showed activity against triolein, as well as CL29, CL37 and CL40 which were randomly chosen. In addition, GCL1 and RML, which are non-regioselective and *sn*1,3-regioselective lipase respectively are included. RFP is used as negative control. As shown in **FIG. 10**, all CalA orthologues which are active against triolein, as well as GCL1, were able to synthesize at least 1 type of FAHFA. CL29, CL37 and CL40 were unable to synthesize FAHFA, possibly due to their limited ability to accept long-chain substrates, insufficient thermostability or low enzyme expression. GCL1 was only active at 50°C and not at 70°C and 90°C, possibly due to its inability to withstand high temperature. On the other hand, RML was not able to synthesize FAHFA as it is a *sn*1,3-regioselective lipase, and thus, unable to target the secondary alcohol on the hydroxystearic acid.

[0108] 12-PAHSA Biosynthesis

[0109] CalA was the only lipase discovered to biosynthesize 12-PAHSA from 12-HSA and palmitic acid with high efficiency in the presence of organic solvent when immobilized (Zhang et al., 2021). Here, 12-PAHSA biosynthesis was performed in the absence of solvent using immobilized enzymes and substrates (12-hydroxystearic acid and palmitic acid) which melt at 70˚C. CL20, 23 and 31 were selected as they are either expressed at high levels or exhibited thermostability (active above 70˚C) based on the FAHFA screening. In terms of expression, CL20 and 23 showed higher expression levels compared to CalA and CL31 from the bioreactor, based on bradford protein quantification assay. The enzymes were processed as described in the method section and immobilized in Lifetech™ ECR1030M and Lifetech™ ECR8806. As shown in **FIG. 11**, CL20, 23 and 31 were also able to synthesize 12-PAHSA, like CalA, when immobilized. **FIG. 11** shows the percentage conversion to form 12-PAHSA over time. Each group of bars is for the identified lipase and the resin used to immobilize the lipase. The bars in each group from left to right is for the percentage conversion after 5 minutes (min), 10 minutes, 20 minutes, 40 minutes, 80 minutes, 160 minutes, and 320 minutes. The tested lipases can achieve more than 85% 12-PAHSA conversion

in a solvent-free environment. Generally, ECR8806M was better than ECR1030M in terms of initial rate of conversion, or the time taken to reach 50% of the conversion, when the same enzyme (on different immobilization beads) was compared.

[0110] To synthesize OPO using triolein and palmitic acid as substrates, several important factors must be considered, such as the ability to perform an acidolysis reaction of long chain triglyceride with fatty acid, high selectivity, and thermostability as the substrates need to melt at a high reaction temperature. While CalA has potential as a candidate for OPO synthesis, its selectivity needs improvement.

[0111] Additionally, CalA and its orthologues have shown promising results in the synthesis of FAHFA. These enzymes have demonstrated high thermostability and the ability to esterify a long chain fatty acid with a secondary alcohol group (on 12-HSA), which is rare among other hydrolases. Future research will focus on determining the rate of FAHFA synthesis and exploring the substrate specificity of these enzymes.

Table 3. Sequences without signal peptide

SEQ ID NO: (without signal peptide)	**SEQ ID NO: (with signal peptide)**	**SEQ ID NO: (gene sequence)**	**CalA-ID**	**Uniprot ID**	**Seq ID to CalA (%) Without Met**
1	61	121	CalA	W3VKA4	100
2	62	122	UMLA	A0A0D1DXG7	68.89
3	63	123	CL1	A0A5C3F0N6	44.04
4	64	124	CL2	A0A316W5L4	33.02
5	65	125	CL3	A0A316W4B5	36.78
6	66	126	CL4	A0A507R6B2	37.8
7	67	127	CL5	A0A511KN21	35.86
8	68	128	CL6	A0A316V586	40.33
9	69	129	CL7	A0A316ZLT3	36.24
10	70	130	CL8	A0A316UJ04	41.81
11	71	131	CL9	A0A316YVL4	53.77
12	72	132	CL10	A0A511KP43	33.73
13	73	133	CL11	A0A316YP38	42.89
14	74	134	CL12	A0A6A6Q743	41.3
15	75	135	CL13	A0A316UJA2	44.42
16	76	136	CL14	A0A316W751	35.46
17	77	137	CL15	A0A061H6R3	48.58
18	78	138	CL16	A0A109FLN7	34.1
19	79	139	CL17	A0A316W1U4	31.88
20	80	140	CL18	A0A5C5FX20	32.06
21	81	141	CL19	Q8NIP2	77.19
22	82	142	CL20	A0A5C3EAQ1	76.55
23	83	143	CL21	A0A081CG82	96.83
24	84	144	CL22	A0A109FLE5	34.53

25	85	145	CL23	V5F2U3	68.37
26	86	146	CL24	A0A4U7KKA1	67.82
27	87	147	CL25	A0A4U0VNC5	34.1
28	88	148	CL26	A0A0K3CAV5	32.4
29	89	149	CL27	V5F249	56.8
30	90	150	CL28	A0A2Z5XE53	98.19
31	91	151	CL29	A0A316VVP8	33.33
32	92	152	CL30	A0A0K3CAK9	35.11
33	93	153	CL31	R9NZ15	70.49
34	94	154	CL32	A0A2N8UHB4	59.12
35	95	155	CL33	A0A2S5B370	35.48
36	96	156	CL34	A0A061BDY4	35.38
37	97	157	CL35	I2FWQ0	68.88
38	98	158	CL36	A0A1K0G6P3	70.31
39	99	159	CL37	A0A194SC18	33.17
40	100	160	CL38	E7A3D1	67.82
41	101	161	CL39	A0A0F7S6Y3	69.43
42	102	162	CL40	A0A0P1BAY9	32.82
43	103	163	CL41	E6ZNA3	58.89
44	104	164	CL42	A0A077R396	69.91
45	105	165	CL43	A0A0P1BAG4	35.22
46	106	166	CL44	A0A5C3EQP3	71.03
47	107	167	CL45	A0A061BF66	36.55
48	108	168	CL46	A0A0N7LB57	31.88
49	109	169	CL47	A0A4U7KRI5	54.25
50	110	170	CL48	A0A6A6PIP8	42.23
51	111	171	CL49	A0A0K3CCS6	36.72
52	112	172	CL50	A0A0P1BNE1	32.55
53	113	173	CL51	A0A061BC61	32.65
54	114	174	CL52	A0A511KRF5	32.78
55	115	175	CL53	A0A4V5NAI5	34.07
56	116	176	CL54	A0A316U942	43.1
57	117	177	CL55	A0A316U023	42.21
58	118	178	CL56	A0A2S5B3F6	33.68
59	119	179	CL57	A0A4U7KQZ9	66.67
60	120	180	CL58	A0A2N8UFQ4	68.05

[0112] References

[0113] Arslan, A., Rancke-Madsen, A., & Brask, J. (2020). Enzymatic Synthesis of Estolides from Castor Oil. *Catalysts*, *10*(8), 835. https://doi.org/10.3390/catal10080835

[0114] Berčíková, M., Šimková, A., Hudecová, K., Kyselka, J., Filip, V., & Hrádková, I. (2020). Physical Properties of Structured Fats and Fat Blends Based on the Long Chain Fatty Acids. *Food Biophysics*, *15*(1), 143–152. https://doi.org/10.1007/s11483-019-09609-8

[0115] Byrne, F. P., Jin, S., Paggiola, G., Petchey, T. H. M., Clark, J. H., Farmer, T. J., Hunt, A. J., Robert McElroy, C., & Sherwood, J. (2016). Tools and techniques for solvent selection: Green solvent selection guides. *Sustainable Chemical Processes*, *4*(1), 7. https://doi.org/10.1186/s40508-016-0051-z

[0116] Cajka, T., & Fiehn, O. (2014). Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. *TrAC Trends in Analytical Chemistry*, *61*, 192–206. https://doi.org/10.1016/j.trac.2014.04.017

[0117] Cambon, E., Rodriguez, J. A., Pina, M., Arondel, V., Carriere, F., Turon, F., Ruales, J., & Villeneuve, P. (2008). Characterization of typo-, regio-, and stereo-selectivities of babaco latex lipase in aqueous and organic media. *Biotechnology Letters*, *30*(4), 769–774. https://doi.org/10.1007/s10529-007-9603-3

[0118] Cermak, S. C., Isbell, T. A., Bredsguard, J. W., & Thompson, T. D. (2017). Estolides. In *Fatty Acids* (pp. 431–475). Elsevier. https://doi.org/10.1016/B978-0-12-809521-8.00015-5

[0119] Domínguez de María, P., Carboni-Oerlemans, C., Tuin, B., Bargeman, G., van der Meer, A., & van Gemert, R. (2005). Biotechnological applications of Candida antarctica lipase A: State-of-the-art. *Journal of Molecular Catalysis B: Enzymatic*, *37*(1–6), 36–46. https://doi.org/10.1016/j.molcatb.2005.09.001

[0120] Ericsson, D. J., Kasrayan, A., Johansson, P., Bergfors, T., Sandström, A. G., Bäckvall, J.-E., & Mowbray, S. L. (2008). X-ray Structure of Candida antarctica Lipase A Shows a Novel Lid Structure and a Likely Mode of Interfacial Activation. *Journal of Molecular Biology*, *376*(1), 109–119. https://doi.org/10.1016/j.jmb.2007.10.079

[0121] Ferreira, M. L., & Tonetto, G. M. (2017). What Is The Importance of Structured Triglycerides and Diglycerides? In *Enzymatic Synthesis of Structured Triglycerides*. Springer International Publishing. https://doi.org/10.1007/978-3-319-51574-8

[0122] Invitrogen Corporation. (2022). *Pichia Fermentation Process Guidelines*.

[0123] Kapoor, M., & Gupta, M. N. (2012). Lipase promiscuity and its biochemical applications. *Process Biochemistry*, *47*(4), 555–569. https://doi.org/10.1016/j.procbio.2012.01.011

[0124] Laguerre, M., Nlandu Mputu, M., Brïys, B., Lopez, M., Villeneuve, P., & Dubreucq, E. (2017). Regioselectivity and fatty acid specificity of crude lipase extracts from *Pseudozyma tsukubaensis* , *Geotrichum candidum,* and *Candida*

rugosa. *European Journal of Lipid Science and Technology*, *119*(8), 1600302. https://doi.org/10.1002/ejlt.201600302

[0125] Leong, H. Y., Chang, C.-K., Khoo, K. S., Chew, K. W., Chia, S. R., Lim, J. W., Chang, J.-S., & Show, P. L. (2021). Waste biorefinery towards a sustainable circular bioeconomy: A solution to global issues. *Biotechnology for Biofuels*, *14*(1), 87. https://doi.org/10.1186/s13068-021-01939-5

[0126] Löfgren, J., Görbe, T., Oschmann, M., Svedendahl Humble, M., & Bäckvall, J. (2019). Transesterification of a Tertiary Alcohol by Engineered *Candida antarctica* Lipase A. *ChemBioChem*, *20*(11), 1438–1443. https://doi.org/10.1002/cbic.201800792

[0127] Löfgren, L., Ståhlman, M., Forsberg, G.-B., Saarinen, S., Nilsson, R., & Hansson, G. I. (2012). The BUME method: A novel automated chloroform-free 96-well total lipid extraction method for blood plasma. *Journal of Lipid Research*, *53*(8), 1690–1700. https://doi.org/10.1194/jlr.D023036

[0128] Monteiro, R. R. C., Virgen-Ortiz, J. J., Berenguer-Murcia, Á., da Rocha, T. N., dos Santos, J. C. S., Alcántara, A. R., & Fernandez-Lafuente, R. (2021). Biotechnological relevance of the lipase A from Candida antarctica. *Catalysis Today*, *362*, 141–154. https://doi.org/10.1016/j.cattod.2020.03.026

[0129] Osborn, H. T., & Akoh, C. C. (2002). Structured Lipids-Novel Fats with Medical, Nutraceutical, and Food Applications. *Comprehensive Reviews in Food Science and Food Safety*, *1*(3), 110–120. https://doi.org/10.1111/j.1541-4337.2002.tb00010.x

[0130] Palla, C. A., Pacheco, C., & Carrín, M. E. (2012). Production of structured lipids by acidolysis with immobilized Rhizomucor miehei lipases: Selection of suitable reaction conditions. *Journal of Molecular Catalysis B: Enzymatic*, *76*, 106–115. https://doi.org/10.1016/j.molcatb.2011.11.022

[0131] Riecan, M., Paluchova, V., Lopes, M., Brejchova, K., & Kuda, O. (2022). Branched and linear fatty acid esters of hydroxy fatty acids (FAHFA) relevant to human health. *Pharmacology & Therapeutics*, *231*, 107972. https://doi.org/10.1016/j.pharmthera.2021.107972

[0132] Robles-Machuca, M., del Campo, M. M., Camacho-Ruiz, M. Á., Ordaz, E., Zamora-González, E. O., Müller-Santos, M., & Rodríguez, J. A. (2019). Comparative features between recombinant lipases CALA-like from U. maydis and CALA from C.

antarctica in thermal stability and selectivity. *Biotechnology Letters*, *41*(2), 241–252. https://doi.org/10.1007/s10529-018-2630-4

[0133] Schick, D., Link, K., Schwack, W., Granvogl, M., & Oellig, C. (2021). Analysis of mono-, di-, triacylglycerols, and fatty acids in food emulsifiers by high-performance liquid chromatography–mass spectrometry. *European Food Research and Technology*, *247*(4), 1023–1034. https://doi.org/10.1007/s00217-021-03684-3

[0134] Velasco-Lozano, S., Rodríguez-González, J. A., Mateos-Díaz, J. C., Reyes-Duarte, D., & Favela-Torres, E. (2012). Catalytic profiles of lipolytic biocatalysts produced by filamentous fungi. *Biocatalysis and Biotransformation*, *30*(5–6), 459–468. https://doi.org/10.3109/10242422.2012.740475

[0135] Wei, W., Sun, C., Wang, X., Jin, Q., Xu, X., Akoh, C. C., & Wang, X. (2020). Lipase-Catalyzed Synthesis of Sn-2 Palmitate: A Review. *Engineering*, *6*(4), 406–414. https://doi.org/10.1016/j.eng.2020.02.008

[0136] Widmann, M., Juhl, P. B., & Pleiss, J. (2010). Structural classification by the Lipase Engineering Database: A case study of Candida antarctica lipase A. *BMC Genomics*, *11*(1), 123. https://doi.org/10.1186/1471-2164-11-123

[0137] Yore, M. M., Syed, I., Moraes-Vieira, P. M., Zhang, T., Herman, M. A., Homan, E. A., Patel, R. T., Lee, J., Chen, S., Peroni, O. D., Dhaneshwar, A. S., Hammarstedt, A., Smith, U., McGraw, T. E., Saghatelian, A., & Kahn, B. B. (2014). Discovery of a Class of Endogenous Mammalian Lipids with Anti-Diabetic and Anti-inflammatory Effects. *Cell*, *159*(2), 318–332. https://doi.org/10.1016/j.cell.2014.09.035

[0138] Zallot, R., Oberg, N., & Gerlt, J. A. (2019). The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry*, *58*(41), 4169–4182. https://doi.org/10.1021/acs.biochem.9b00735

[0139] Zhang, Y., Eser, B. E., & Guo, Z. (2021). A Bi-Enzymatic Cascade Pathway towards Optically Pure FAHFAs**. *ChemBioChem*, *22*(12), 2146–2153. https://doi.org/10.1002/cbic.202100070

Claims

[Claim 1] A method of producing a fatty acid ester, the method comprising:

providing a sn-2 selective lipase comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), SEQ ID NO: 101 (CL39), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 40 (CL38), SEQ ID NO: 100 (CL38), SEQ ID NO: 46 (CL44), SEQ ID NO: 106 (CL44), SEQ ID NO: 26 (CL24), SEQ ID NO: 86 (CL24), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 17 (CL15), SEQ ID NO: 77 (CL15), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 51 (CL49), SEQ ID NO: 111 (CL49), SEQ ID NO: 56 (CL54), and SEQ ID NO: 116 (CL54); and

contacting the sn-2 selective lipase with a fatty acid and a secondary hydroxy fatty acid to produce the fatty acid ester.

[Claim 2] The method of claim 1, wherein the sn-2 selective lipase consists essentially of the amino acid sequence, preferably the sn-2 selective lipase consists of the amino acid sequence.

[Claim 3] The method of claims 1 or 2, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), and SEQ ID NO: 101 (CL39).

[Claim 4] The method of any one of claims 1 to 3, wherein the fatty acid and the secondary hydroxy fatty acid are each independently a C14 to C22 fatty acid.

[Claim 5] The method of any one of claims 1 to 4, wherein the hydroxyl group in the secondary fatty acid is located at one of the following positions: C2, C10, C12, C15, and any combinations thereof.

[Claim 6] The method of any one of claims 1 to 5, wherein the secondary hydroxy fatty acid is a C16 hydroxy fatty acid or a C18 hydroxy fatty acid, preferably the secondary hydroxy fatty acid is hydroxystearic acid.

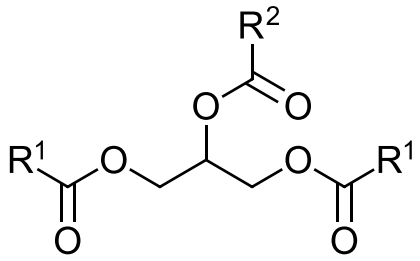
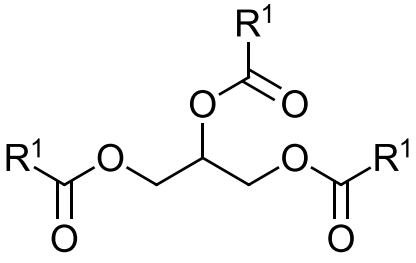
[Claim 7] The method of any one of claims 1 to 6, wherein the fatty acid is a C16 fatty acid or a C18 fatty acid, preferably the fatty acid is palmitic acid.

[Claim 8] The method of any one of claims 1 to 7, wherein providing the sn-2 selective lipase comprises expressing a corresponding nucleotide sequence in a host.

[Claim 9] A method of producing a compound of Formula I, the method comprising:

providing a sn-2 selective lipase; and

contacting the sn-2 selective lipase with a compound of Formula II and a long chain fatty acid to produce the compound of Formula I under suitable conditions,

wherein Formula I is , Formula II is ,

R^1 is a second alkyl tail of a second long chain fatty acid, R^2 is a first alkyl tail of the long chain fatty acid, the long chain fatty acid and the second long chain fatty acid being different long chain fatty acids.

[Claim 10] The method of claim 9, wherein the sn-2 selective lipase comprises an amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28),

SEQ ID NO: 90 (CL28), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 2 (UMLA), SEQ ID NO: 62 (UMLA), SEQ ID NO: 1 (CalA), SEQ ID NO: 61 (CalA), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 26 (CL24), SEQ ID NO: 86 (CL24), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 41 (CL39), SEQ ID NO: 101 (CL39), SEQ ID NO: 17 (CL15), SEQ ID NO: 77 (CL15), SEQ ID NO: 46 (CL44), SEQ ID NO: 106 (CL44), SEQ ID NO: 51 (CL49), SEQ ID NO: 111 (CL49), SEQ ID NO: 33 (CL31), and SEQ ID NO: 93 (CL31).

[Claim 11] The method of claim 10, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 47 (CL45), and SEQ ID NO: 107 (CL45).

[Claim 12] The method of claims 10 or 11 wherein the sn-2 selective lipase comprises the amino acid sequence.

[Claim 13] The method of claims 10 or 11, wherein the sn-2 selective lipase consists essentially of the amino acid sequence, preferably the sn-2 selective lipase consists of the amino acid sequence.

[Claim 14] The method of claims 9 to 13, wherein the long chain fatty acid and the second long chain fatty acid is a C14 to C22 fatty acid.

[Claim 15] The method of claim 14, wherein the long chain fatty acid and the second long chain fatty acid is a C16 fatty acid or a C18 fatty acid.

[Claim 16] The method of claim 14, wherein the long chain fatty acid is palmitic acid, and the second long chain fatty acid is oleic acid, and the compound of Formula I is 1,3-olein-2-palmitin.

[Claim 17] The method of claim 16 further comprising providing a sn-1,3 selective lipase; and contacting the sn-1,3 selective lipase with tripalmitin and oleic acid, wherein contacting the sn-1,3 selective lipase with tripalmitin and oleic acid and contacting the sn-2 selective lipase with triolein and palmitic acid is performed in a reactor.

[Claim 18] The method of claim 14, wherein the long chain fatty acid is oleic acid, and the second long chain fatty acid is stearic acid, and the compound of Formula I is 1,3-distearoyl-2-oleoyl-glycerol.

[Claim 19] The method of any one of claims 9 to 18, wherein providing the sn-2 selective lipase comprises expressing a corresponding nucleotide sequence in a host.

[Claim 20] An isolated sn-2 selective lipase comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), SEQ ID NO: 101 (CL39), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 40 (CL38), SEQ ID NO: 100 (CL38), SEQ ID NO: 46 (CL44), SEQ ID NO: 106 (CL44), SEQ ID NO: 26 (CL24), SEQ ID NO: 86 (CL24), SEQ ID NO: 47 (CL45), SEQ ID NO: 107 (CL45), SEQ ID NO: 17 (CL15), SEQ ID NO: 77 (CL15), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 51 (CL49), SEQ ID NO: 111 (CL49), SEQ ID NO: 56 (CL54), and SEQ ID NO: 116 (CL54).

[Claim 21] The isolated sn-2 selective lipase of claim 20 consisting essentially of the amino acid sequence, preferably the lipase consists of the amino acid sequence.

[Claim 22] The isolated sn-2 selective lipase of claims 20 or 21, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 23 (CL21), SEQ ID NO: 83 (CL21), SEQ ID NO: 22 (CL20), SEQ ID NO: 82 (CL20), SEQ ID NO: 30 (CL28), SEQ ID NO: 90 (CL28), SEQ ID NO: 44 (CL42), SEQ ID NO: 104 (CL42), SEQ ID NO: 15 (CL13), SEQ ID NO: 75 (CL13), SEQ ID NO: 47 (CL45), SEQ ID NO:

107 (CL45), SEQ ID NO: 33 (CL31), SEQ ID NO: 93 (CL31), SEQ ID NO: 21 (CL19), SEQ ID NO: 81 (CL19), SEQ ID NO: 25 (CL23), SEQ ID NO: 85 (CL23), SEQ ID NO: 41 (CL39), and SEQ ID NO: 101 (CL39).

[Claim 23] The lipase of any one of claims 20 to 22 comprising a purification tag bonded to either N-terminal or C-terminal of the amino acid sequence, preferably the N-terminal.