(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(10) International Publication Number
WO 2021/209482 A1

(43) International Publication Date
21 October 2021 (21.10.2021)

(51) International Patent Classification:
C12P 17/18 (2006.01) *C07D 493/08* (2006.01) *C12P 17/02* (2006.01) *C11B 9/00* (2006.01) *C07D 313/08* (2006.01)

(21) International Application Number: PCT/EP2021/059618

(22) International Filing Date: 14 April 2021 (14.04.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2005468.0 15 April 2020 (15.04.2020) GB

(71) Applicant: **GIVAUDAN SA** [CH/CH]; Chemin de la Parfumerie 5, 1214 Vernier (CH).

(72) Inventors: **EICHHORN, Eric**; Kronenstrasse 39, 8006 Zürich (CH). **FLACHSMANN, Felix**; Langhagweg 2, 8600 Duebendorf (CH). **GOEKE, Andreas**; Casinostrasse 8, 8600 Duebendorf (CH).

(74) Agent: **GLOBAL PATENTS**; Givaudan SA, Grafenaustrasse 7, 6300 Zug (CH).

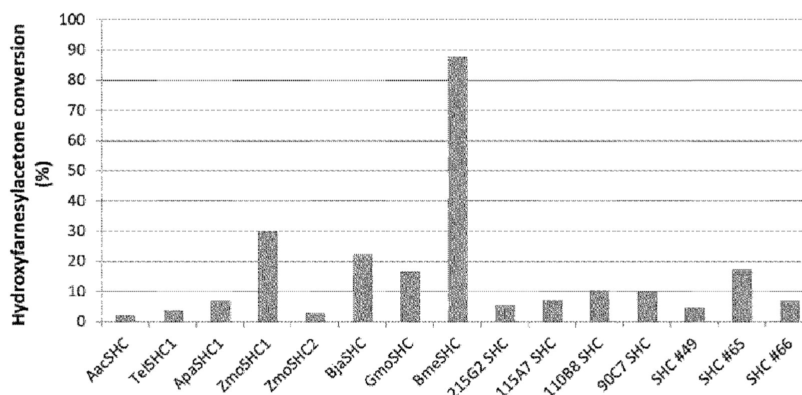
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: ENZYME-MEDIATED PROCESS FOR MAKING AMBERKETAL AND AMBERKETAL HOMOLOGUES

Figure 2



(57) Abstract: An enzyme-mediated method for the production of Amberketal and Amberketal homologues, the products of said method, and uses of said products.

WO 2021/209482 A1

ENZYME-MEDIATED PROCESS FOR MAKING AMBERKETAL AND AMBERKETAL HOMOLOGUES

TECHNICAL FIELD

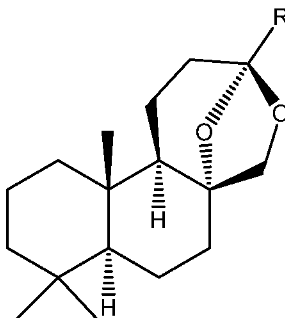
The present invention generally relates to a method of making Amberketal and
5 Amberketal homologues using a squalene-hopene cyclase (SHC) enzyme or enzyme variant. The invention further relates to compositions made by said method, the various uses of said compositions, and consumer products comprising said compositions.

BACKGROUND

10 Amberketal provides a powerful and tenacious ambery and woody odour that is useful in fragrance compositions alone or in combination with other woody or ambery ingredients. Amberketal is traditionally prepared from Manool via a number of chemical transformations. However, the supply of natural Manool is limited. It is therefore desirable to provide a new efficient and cost effective synthetic route to obtain
15 Amberketal and amberketal homologues.

SUMMARY

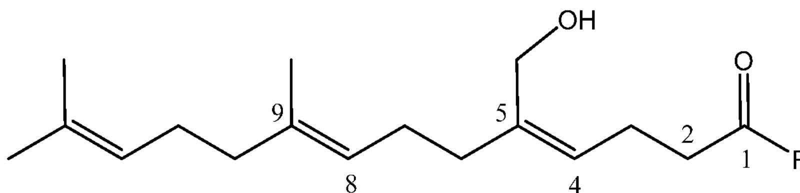
In accordance with a first aspect of the present invention there is provided a method for making a compound of formula (I),



20

Formula (I)

wherein the method comprises contacting a compound of formula (II) with a squalene-hopene cyclase (SHC) enzyme or enzyme variant,



25 Formula (II)

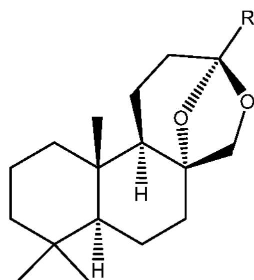
wherein R is H, methyl, or ethyl.

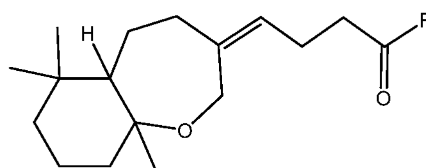
In certain embodiments, the method comprises contacting a compound of formula (I) wherein the double bond beteween C-8 and C-9 is in *E*-configuration and the double bond between C-4 and C-5 is in Z-configuration (*E,Z*-compound of formula (II)) with a squalene-hopene cyclase (SHC) enzyme (wild-type or variant enzyme).

In certain embodiments, the method comprises contacting a mixture comprising a compound of formula (II) wherein both double bonds are in *E*-configuration (*E,E*-compound) and a compound of formula (II) wherein the double bond between C-8 and C-9 is in *E*-configuration and the double bond between C-4 and C-5 is in *Z*-configuration (*E,Z*-compound) with a squalene-hopene cyclase (SHC) enzyme or enzyme variant.

In certain embodiments, the weight ratio of the *E,Z*-compound to the *E,E*-compound ranges from about 99:1 to about 10:90. For example, the weight ratio of the *E,Z* compound of formula (II) to the *E,E*-compound of formula (II) may range from about 95:5 to about 50:50, or from about 80:20 to about 50:50 or from about 80:20 to about 60:40.

In accordance with a second aspect of the present invention there is provided a composition comprising, consisting essentially of, or consisting of a compound of formula (I) and a compound of formula (III),



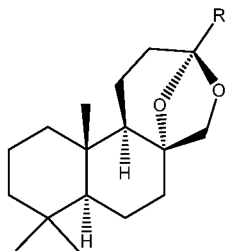


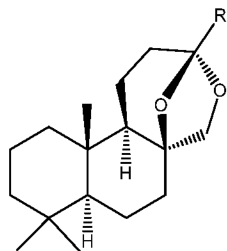
Formula (I) Formula (III)

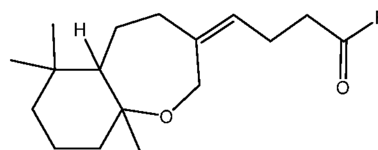
wherein R is H, methyl, or ethyl.

In accordance with a third aspect of the present invention there is provided a composition comprising, consisting essentially of, or consisting of a compound of

formula (I), a compound of formula (IV), and a compound of formula (III),







Formula (I) Formula (IV) Formula (III)

wherein R is H, methyl, or ethyl.

5

In accordance with a fourth aspect of the present invention there is provided a compound or composition obtained by or obtainable by the method of the first aspect of the present invention. The composition may, for example, be as defined in the second aspect of the present invention, including any embodiment thereof.

10

In accordance with a fifths aspect of the present invention there is provided a use of a composition of the second or third aspect of the present invention as or in a fragrance composition.

15 In accordance with a sixth aspect of the present invention there is provided a consumer product comprising a composition of the second or third aspect of the present invention.

In certain embodiments of any aspect of the present invention R is methyl. The
20 compound of formula (I) may be referred to as (+)-Amberketal when R is methyl.

Certain embodiments of the present invention may provide one or more of the following advantages:

- biocatalytic route to produce (+)-Amberketal and homologues of
25 (+)-Amberketal;
- milder reaction conditions (e.g. lower temperatures);
- high selectivity;
- use of alternative feed stock (e.g. alternative feed stock to that used for the chemical synthesis).

30

The details, examples and preferences provided in relation to any particular one or more of the stated aspects of the present invention will be further described herein and apply equally to all aspects of the present invention. Any combination of the embodiments, examples and preferences described herein in all possible variations
5 thereof is encompassed by the present invention unless otherwise indicated herein, or otherwise clearly contradicted by context.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the reaction scheme for the production of a compound of formula (II).
10 For the the compounds R is H, methyl or ethyl.

Figure 2 shows the conversion [%] of Hydroxyfarnesylacetone to (+)-Amberketal with wild type and variant SHC enzymes and gives an overview of the performance of the tested SHC enzymes under their optimal reaction conditions as set out in Table 2.

15

SUMMARY OF THE SEQUENCES

SEQ ID NO: 1 is the amino acid sequence of wild-type *Alicyclobacillus acidocaldarius* AacSHC.

SEQ ID NO: 2 is the nucleotide sequence of wild-type *Alicyclobacillus acidocaldarius*
20 AacSHC.

SEQ ID NO: 3 is SEQ ID NO: 1 with the substitutions M132R, A224V, I432T, A557T and R613S and may be referred to as SHC enzyme variant #65 herein.

SEQ ID NO: 4 is the nucleotide sequence of SHC variant #65.

SEQ ID NO: 5 corresponds to SEQ ID NO: 1 with the substitutions M132R, A224V,
25 I432T, Y81H, A557T and R613S and may be referred to as SHC enzyme variant #66 herein.

SEQ ID NO: 6 is the nucleotide sequence of SHC variant SHC #66.

SEQ ID NO: 7 corresponds to SEQ ID NO: 1 with the substitutions M132R, A224V, I432T, T90A and R613S and may be referred to as SHC enzyme variant #90C7 herein.

30 SEQ ID NO: 8 is the nucleotide sequence of SHC variant #90C7.

SEQ ID NO: 9 corresponds to SEQ ID NO: 1 with the substitutions M132R, A224V, I432T, Y81H, H431L and A557T and may be referred to as SHC enzyme variant #110B8 herein.

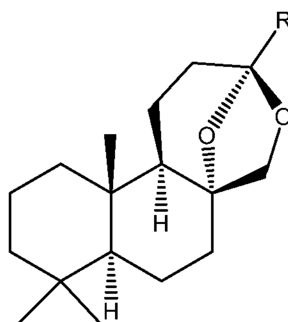
SEQ ID NO: 10 is the nucleotide sequence of SHC variant #110B8.

SEQ ID NO: 11 corresponds to SEQ ID NO: 1 with the substitutions M132R, A224V, I432T, A172T and M277K and may be referred to as SHC enzyme variant #115A7 herein.
SEQ ID NO: 12 is the nucleotide sequence of SHC variant #115A7.
5 SEQ ID NO: 13 corresponds to SEQ ID NO: 1 with the mutations M132R, A224V and I432T and may be referred to as SHC enzyme variant 215G2 herein.
SEQ ID NO: 14 is the nucleotide sequence of SHC variant #215G2.
SEQ ID NO: 15 is the amino acid sequence of wild-type *Zymomonas mobilis* ZmoSHC1.
10 SEQ ID NO: 16 is the amino acid sequence of wild-type *Zymomonas mobilis* ZmoSHC2.
SEQ ID NO: 17 is the amino acid sequence of wild-type *Bradyrhizobium japonicum* BjaSHC.
SEQ ID NO: 18 is the amino acid sequence of wild-type *Thermosynechococcus*
15 *elongatus* TelSHC.
SEQ ID NO: 19 is the amino acid sequence of wild-type *Acetobacter pasteurianus* ApaSHC1.
SEQ ID NO: 20 is the amino acid sequence of wild-type *Gluconobacter morbifer* GmoSHC.
20 SEQ ID NO: 21 is the amino acid sequence of wild-type *Bacillus megaterium* BmeSHC
SEQ ID NO: 22 corresponds to SEQ ID NO: 1 with the substitutions M132R, A224V, I432T, A557T and H431L and may be referred to as SHC enzyme variant #49 herein.

DETAILED DESCRIPTION

25 The present invention is based, at least in part, on the surprising finding that squalene-hopene cyclase (SHC) enzymes and enzyme variants can be used to make (+)-Amberketal and Amberketal homologues from polyunsaturated alcohols of formula (II). It is particular surprising that a substrate wherein the alkenyl chain is substituted with a hydroxymethyl group as defined by formula (II) herein undergoes an enzymatic
30 polycyclisation reaction terminated by internal ketalisation.

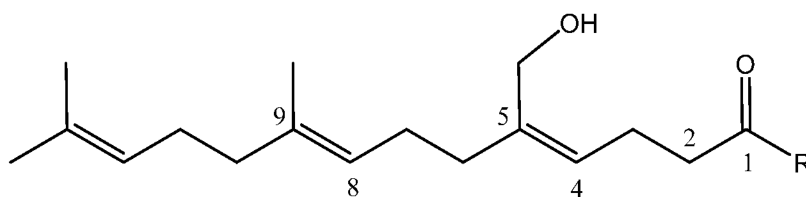
Thus, there is provided herein in a first aspect a method of making a compound of formula (I),



Formula (I)

wherein the method comprises contacting a compound of formula (II) with a squalene-hopene cyclase (SHC) enzyme or enzyme variant,

5



Formula (II)

wherein R is H, methyl, or ethyl.

In certain embodiments both double bonds are in *E*-configuration (*E,E*-compound of
10 formula (II)).

In certain embodiments the double between C-8 and C-9 is in *E*-configuration and the double bond between C-4 and C-5 is in *Z*-configuration (*E,Z*-compound of formula (II)).

15 In particular embodiments, R is methyl.

The methods provided herein enzymatically convert a compound of formula (II) to a compound of formula (I) using an SHC enzyme or enzyme variant (bioconversion reaction).

20

COMPOUND OF FORMULA (II)

The compound of formula (II) exists in the form of four different stereoisomers, for example, as a compound of formula (II) having an *E,E*- or *E,Z*-configuration.

25 In certain embodiments, the method comprises contacting an *E,Z*- compound of

formula (II) with a squalene-hopene cyclase (SHC) enzyme or enzyme variant in the absence of any other stereoisomers of formula (II)

In other embodiments, the compound of formula (II) may, for example, be a mixture of stereoisomers. In certain embodiments, the mixture comprises the *E,E*-compound of formula (II) and one or more other stereoisomers of formula (II). In certain embodiments, the mixture comprises the *E,Z*-compound of formula (II) and one or more other stereoisomers of formula (II).

In certain embodiments, the method comprises contacting a mixture comprising, consisting essentially of, or consisting of *E,E*-compound of formula (II) and *E,Z*-compound of formula (II) with a SHC enzyme or enzyme variant. In certain embodiments, the composition does not comprise any other stereoisomers of formula (II).

The weight ratio of the *E,Z*-compound of formula (II) to total other stereoisomers of formula (II) may, for example, be equal to or greater than about 10:90. For example, the weight ratio of the *E,Z*-compound of formula (II) to total other stereoisomers of formula (II) may be equal to or greater than about 20:80 or equal to or greater than about 30:70 or equal to or greater than about 40:60 or equal to or greater than about 50:50 or equal to or greater than about 60:40 or equal to or greater than about 70:30 or equal to or greater than about 80:20 or equal to or greater than about 90:10 or equal to or greater than about 95:5, or equal to or greater than about 99:1.

The weight ratio of the *E,Z*-compound of formula (II) to total other stereoisomers of formula (II) may, for example, be equal to or less than about 99:1. For example, the weight ratio of the *E,Z*-compound of formula (II) to other stereoisomers of formula (II) may be equal to or less than about 95:5 or equal to or less than about 90:10 or equal to or less than about 85:15 or equal to or less than about 80:20 or equal to or less than about 60:40.

For example, the weight ratio of the *E,Z*-compound of formula (II) to total other stereoisomers of formula (II) may range from about 10:90 to about 99:1 or from about

10:90 to about 90:10 or from about 20:80 to about 80:20 or from about 50:50 to about 80:20 or from about 60:40 to about 80:20.

The weight ratio of the *E,Z*-compound of formula (II) to the *E,E*-compound of formula
5 (II) may, for example, be equal to or greater than about 10:90. For example, the weight ratio of the *E,Z*-compound of formula (II) to the *E,E*-compound of formula (II) may be equal to or greater than about 20:80 or equal to or greater than about 30:70 or equal to or greater than about 40:60 or equal to or greater than about 50:50 or equal to or greater than about 60:40 or equal to or greater than about 70:30 or equal to or greater
10 than about 80:20 or equal to or greater than about 90:10 or equal to or greater than about 95:5, or equal to or greater than about 99:1.

The weight ratio of the *E,Z*-compound of formula (II) to the *E,E*-compound of formula (II) may, for example, be equal to or less than about 99:1. For example, the weight ratio
15 of the *E,Z*-compound of formula (II) to the *E,E*-compound of formula (II) may be equal to or less than about 95:5 or equal to or less than about 90:10 or equal to or less than about 85:15 or equal to or less than about 80:20 or equal to or less than about 70:30 or equal to or less than 60:40.

20 For example, the weight ratio of the *E,Z*- compound of formula (II) to the *E,E*-compound of formula (II) may range from about 10:90 to about 99:1 or from about 10:90 to about 90:10 or from about 20:80 to about 80:20 or from about 50:50 to about 80:20 or from about 60:40 to about 80:20.

25 The amount of each stereoisomer in a mixture of stereoisomers may, for example, be identified by gas chromatography or NMR spectroscopy analysis.

When R is methyl, the compound of formula (II) may be referred to as hydroxyfarnesylacetone, encompassing *E,E*-hydroxyfarnesylacetone, *E,Z*-
30 hydroxyfarnesylacetone, *Z,E*-hydroxyfarnesylacetone and *Z,Z*-hydroxyfarnesylacetone, and mixtures thereof.

In certain embodiments, not all of the compound of formula (II) is converted to a compound of formula (I) or a by-product of the reaction. Thus, in certain embodiments,

the compositions described herein, for example the compositions obtained by or obtainable by the methods described herein, may comprise a compound of formula (I) and a compound of formula (II).

5 In certain embodiments, any non-converted compound of formula (II) in the mixture made by the methods described herein may be separated from the other reaction products such that the compositions do not comprise any compounds of formula (II).

In alternative embodiments, all compound of formula (II) is converted to a compound of
10 formula (I) or a by-product of the reaction by the methods described herein.

The number of stereoisomers of the compound of formula (II) present may influence the reaction rate. A SHC enzyme or enzyme variant may be capable of converting an *E,Z*-compound of formula (II) to a compound of formula (I) from a complex mixture of
15 stereoisomers of the compound of formula (II) (said mixture may include only two of the stereoisomers, for example, *E,Z*-compound and *E,E*-compound of formula (II), or it may comprise three (i.e. *E,Z*-, and *E,E*- and *Z,E*- or *Z,Z*- compound of formula (II), or even all four stereoisomers). However, a lower conversion rate may be observed, which is consistent with the view that the other stereoisomers may compete with the *E,Z*-
20 compound of formula (II) for access to the SHC enzyme or enzyme variant and thus may act as competitive inhibitors for the conversion of the *E,Z*-compound of formula (II) to the compound of formula (I) and/or also act as alternative substrates. Accordingly, the compound of formula (II) substrate may comprise an isomeric mixture of 2-4 isomers, preferably two isomers. Preferably the compound of formula (II) substrate
25 comprises, consists essentially of or consists of an isomeric mixture of *E,Z*- and *E,E*-compound of formula (II).

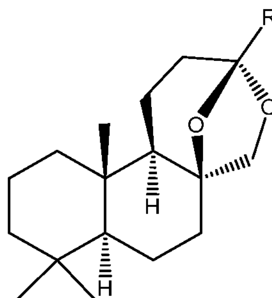
The compound of formula (II) may be synthesized following the general procedure depicted by Fujiwara et al. (Tetrahedron Letters, 1995 Vol 36(46), 8435-8438).
30 Alternatively the compounds of formula (II) may be obtained as briefly demonstrated in Figure 1, wherein R is H, methyl or ethyl.

<u>COMPOUND OF FORMULA (I)</u>

The compound of formula (I) contains a number of chiral carbon atoms and thus one or

more stereoisomers of the compound of formula (I) may also exist, including enantiomers and diastereomers. In addition to the compound of formula (I), the products made by the methods described herein may include one or more of the stereoisomers of the compound of formula (I). The stereoisomers obtained may depend on the stereoisomers of the compound of formula (II) that are used.

For example, a compound of formula (IV) may also be made in addition to the compound of formula (I),



Formula (IV)

wherein R is H, methyl or ethyl.

The compound of formula (IV) wherein R is methyl is also known as (-)-*epi*-8-Amberketal.

In certain embodiments, no other stereoisomers of the compound of formula (I) are made by the method or are present in the product of the method, e.g. in certain embodiments a compound of formula (IV) is not made by the method or are present in the product of the method.

The methods described herein may, for example, make a compound of formula (I) and one or more other stereoisomers of the compound of formula (I) (e.g. a compound of formula (IV)). Thus, the compositions described herein, for example the compositions obtained by or obtainable by the methods described herein may comprise a compound of formula (I) and one or more stereoisomers of the compound of formula (I) (e.g. a compound of formula (IV)).

The weight ratio of the compound of formula (I) to total other stereoisomers of formula (I) may, for example be equal to or greater than about 50:50. For example, the weight

ratio of the compound of formula (I) to total other stereoisomers of formula (I) may be equal to or greater than about 55:45 or equal to or greater than about 60:40 or equal to or greater than about 65:35 or equal to or greater than about 70:30 or equal to or greater than about 75:25 or equal to or greater than about 80:20 or equal to or greater than about 85:15 or equal to or greater than about 90:10 or equal to or greater than about 95:5, or equal to or greater than about 99:1.

For example, the compound of formula (I) may be the sole stereoisomeric form of formula (I) made or present in the composition. In other words, the composition may comprise 100 wt% of the compound of formula (I) based on the total weight of all stereoisomeric forms of the compound of formula (I) or have a weight ratio of compound of formula (I) to total other stereoisomers of formula (I) of 100:0. Alternatively, the weight ratio of the compound of formula (I) to total other stereoisomers of formula (I) may be less than about 100 wt%. For example, the weight ratio of the compound of formula (I) to total other stereoisomers of formula (I) may be equal to or less than about 99:1 or equal to or less than about 98:2 or equal to or less than about 97:3.

For example, the weight ratio of the compound of formula (I) to total other stereoisomers of formula (I) may be from about 50:50 to about 100:0 or from about 60:40 to about 99:1 or from about 70:30 to about 98:2 or from about 80:20 to about 97:3 or from about 90:10 to about 97:3.

The amount of each stereoisomer of the compound of formula (I) in a mixture of stereoisomers may, for example, be identified by gas chromatography on chiral columns or NMR spectroscopy in the presence of shift reagents.

The compound of formula (I) as obtained by the method described herein may, for example, be in amorphous form or in crystalline form.

The compound of formula (I) produced by the methods described herein may be isolated by steam extraction/distillation or organic solvent extraction using a non-water miscible solvent (to separate the reaction products and unreacted substrate from the biocatalyst which stays in the aqueous phase) followed by subsequent evaporation of

the solvent to obtain a crude reaction product as determined by gas chromatography (GC) analysis. The steam extraction/distillation and organic solvent extraction methods are known to those skilled in the art.

5 By way of example, the resulting compound of formula (I) may be extracted from the whole reaction mixture using an organic solvent such as a non-water miscible solvent (for example toluene). Alternatively, the resulting compound of formula (I) may be extracted from the solid phase of the reaction mixture (obtained by, for example, centrifugation or filtration) using a water miscible solvent (for example ethanol) or a
10 non-water miscible solvent (for example toluene). By way of further example, the compound of formula (I) may be present in the solid phase as crystals or in amorphous form and can be separated from the remaining solid phase (cell material or debris thereof) and the liquid phase also by means of filtration. By way of further example, at a temperature above the melting point of the compound of formula (I), the compound of
15 formula (I) may form an oil layer on top of aqueous phase, which oil layer can be removed and collected. In order to ensure a complete recovery of compound of formula (I) after the oil layer is removed, an organic solvent may be added to the aqueous phase containing the biomass in order to extract any residual compound of formula (I) (e.g. (+)-Amberketal) contained in, or on or about the biomass. The organic layer can
20 be combined with the oil layer, before the whole is further processed to isolate and purify the compound of formula (I). The compound of formula (I) may be further selectively crystallised to remove by-products and any unreacted compound of formula (II) from the final product. The term "selective crystallization" refers to a process step whereby the compound of formula (I) is caused to crystallise from a solvent whilst the
25 by-products remain dissolved in the crystallising solvent to such an extent that isolated crystalline material contains only the compound of formula (I), or if it contains any by-products, then they are present only in olfactory acceptable amounts. The compound of formula (I), for example, is free or substantially free of by-products such as the compound of formula (III) or (IIIa). The selective crystallisation step may use a water
30 miscible solvent such as ethanol or the like. The selective crystallisation of the compound of formula (I) may be influenced by the presence of unreacted compound of formula (II) and also the ratio of compound of formula (I) to the other detectable by-products. Even if only 10% conversion of the compound of formula (II) to compound of

formula (I) is obtained, the selective crystallisation of the compound of formula (I) may be still possible.

The olfactive purity of the final compound of formula (I) product may be determined
5 using a 10% ethanol extract in water or by testing the crystalline material. The final compound of formula (I) product is tested against a commercially available reference of compound of formula (I) for its olfactive purity, quality and its sensory profile. The compound of formula (I) material is also tested in application studies by experts in order to determine if the material meets the specifications with respect to its
10 organoleptic profile.

Examples of suitable water miscible and non-water miscible organic solvents suitable for use in the extraction and/or selective crystallization of compound of formula (I) include but are not limited to aliphatic hydrocarbons, preferably those having 5 to 8
15 carbon atoms, such as pentane, cyclopentane, hexane, cyclohexane, heptane, octane or cyclooctane, halogenated aliphatic hydrocarbons, preferably those having one or two carbon atoms, such as dichloromethane, chloroform, carbon tetrachloride, dichloroethane or tetrachloroethane, aromatic hydrocarbons, such as benzene, toluene, the xylenes, chlorobenzene or dichlorobenzene, aliphatic acyclic and cyclic
20 ethers or alcohols, preferably those having 4 to 8 carbon atoms, such as ethanol, isopropanol, diethyl ether, methyl tert.-butyl ether, ethyl tert.-butyl ether, dipropyl ether, diisopropyl ether, dibutyl ether, tetrahydrofuran or esters such as ethyl acetate or n-butyl acetate or ketones such as methyl isobutyl ketone or dioxane or mixtures of these. The solvents which are especially preferably used are the abovementioned
25 heptane, Methyl tert-butyl ether (also known as MTBE, tert-butyl methyl ether, tertiary butyl methyl ether, and tBME), diisopropyl ether, tetrahydrofuran, ethyl acetate and/or mixtures thereof. Preferably, a water miscible solvent such as ethanol is used for the extraction of the compound of formula (I) from the solid phase of the reaction mixture. The use of ethanol is advantageous because it is easy to handle, it is non-toxic and it is
30 environmentally friendly.

The term "isolated" as used herein refers to a bioconversion product such as the compound of formula (I) which has been separated or purified from components which accompany it. An entity that is produced in a cellular system different from the source

from which it naturally originates is "isolated", because it will necessarily be free of components which naturally accompany it. The degree of isolation or purity can be measured by any appropriate method, e.g. gas chromatography (GC), HPLC or NMR analysis.

5

In some embodiments, the compound of formula (I) (e.g. (+)-Amberketal) is isolated and purified from the obtained crude product (e.g. to a purity of at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%).

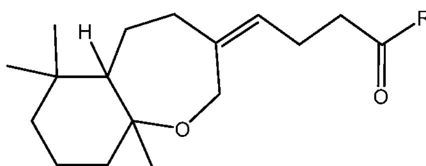
10

Desirably, the concentration of compound of formula (I) in the reaction broth obtained by the method described herein can be from about 1 mg/l to about 20,000 mg/l (20g/l) or higher such as from about 20g/l to about 200g/l or from 100 - 500g/l (including 150g/l, 250g/l, 300g/l, 350g/l, 400g/l or 450g/l).

15

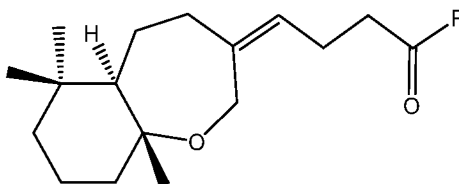
COMPOUND OF FORMULA (III)

The methods described herein may, for example, make a compound of formula (III) as a by-product, wherein R is H, methyl, or ethyl:



20 Formula (III)

The compound of formula (III) may, for example, be a compound having the relative configuration of formula (IIIa), wherein R is H, methyl, or ethyl:



25

Formula (IIIa)

The methods described herein may, for example, make one or more stereoisomer(s) of

compound of formula (III). The composition described herein may include one or more stereoisomer(s) of the compound of formula (III). The methods described herein may, for example, make a compound of formula (III) having the relative configuration shown in formula (IIIa). The composition described herein may therefore include a compound of formula (III) having the relative configuration shown in formula (IIIa). In certain embodiments, the only compound of formula (III) made by the methods described herein and therefore present in the compositions described herein are compounds having the relative configuration shown in formula (IIIa).

In certain embodiments, at least about 50 wt% of the total compounds of formula (III) have the relative configuration shown in formula (IIIa). For example, at least about 60 wt% or at least about 70 wt% or at least about 80 wt% or at least about 90 wt% of the total compounds of formula (III) may have the relative configuration shown in formula (IIIa).

In certain embodiment, compounds having the configuration shown in formula (IIIa) are the only stereoisomers of formula (III). For example, 100 wt% of the total compounds of formula (III) have the relative configuration shown in formula (IIIa). In certain embodiments, equal to or less than about 99 wt% or equal to or less than about 95 wt% or equal to or less than about 90 wt% or equal to or less than about 85 wt% or equal to or less than about 80 wt% or equal to or less than about 75 wt% of the total compounds of formula (III) have the relative configuration shown in formula (IIIa).

For example, from about 50 wt% to about 100 wt% or from about 60 wt% to about 99 wt% or from about 70 wt% to about 95 wt% of the compounds of formula (III) have the relative configuration shown in formula (IIIa).

The amount of the different isomers of the compound of formula (III) in a mixture of stereoisomers may, for example, be identified by gas chromatography on chiral columns or NMR spectroscopy in the presence of shift reagents.

PRODUCTS OBTAINED BY THE METHODS DESCRIBED HEREIN

There is also provided herein the products of the methods described herein. Thus,

there is also provided herein a composition obtained by or obtainable by the method described herein, including all embodiments thereof.

There is provided herein a composition comprising, consisting essentially of or
5 consisting of a compound of formula (I) and a compound of formula (III). The composition may, for example, further comprise one or more other stereoisomers of formula (I), for example a compound of formula (IV). The composition may, for example, comprise one or more stereoisomers of formula (III). For example, the composition may comprise a compound having the relative configuration of formula
10 (IIIa). The composition may, for example, further comprise any unreacted compound of formula (II). In certain embodiments, R is methyl.

There is provided in a particular embodiment a composition comprising, consisting essentially of or consisting of a compound of formula (I), a compound of formula (IV),
15 and a compound of formula (III).

There is also provided herein a composition comprising the compound of formula (I) and one or more stereoisomers of the compound of formula (I), for example a compound of formula (IV). The composition may, for example, further comprise a
20 compound of formula (III). The composition may further comprise any unreacted compound of formula (II). In certain embodiments, R is methyl.

The weight ratio of the compound of formula (I) to the compound of formula (III) in the compositions described herein may, for example, range from about 60:40 to about
25 99:1. For example, the weight ratio of the compound of formula (I) to the compound of formula (III) may range from about 65:35 to about 99:1 or from about 70:30 to about 99:1 or from about 75:25 to about 99:1 or from about 80:20 to about 99:1 or from about 85:15 to about 99:1 or from about 90:10 to about 99:1 or from about 95:5 to about 99:1. For example, the weight ratio of the compound of formula (I) to the compound of
30 formula (III) may range from about 65:35 to about 98:2 or from about 70:30 to about 97:3 or from about 75:25 to about 96:4 or from about 80:20 to about 95:5 or from about 85:15 to about 90:10.

The weight ratio of the compound of formula (I) to the compound of formula (II) in the

crude reaction product described herein may, for example, range from about 90:10 to about 100:0. For example, the weight ratio of the compound of formula (I) to the compound of formula (II) in the compositions described herein may range from about 92:8 to about 100:0 or from about 94:6 to about 100:0 or from about 95:5 to about 100:0 or from about 96:4 to about 99.5:0.5 or from about 97:3 to about 99.0:1.0 or from about 98:2 to about 99.0:1.0.

FRAGRANCE COMPOSITIONS

There is further provided herein the use of the reaction products described herein in or as a fragrance composition.

Thus, there is also provided herein a fragrance composition comprising one or more compounds of formula (I). By "fragrance composition" is meant any composition comprising one or more compounds of formula (I) and a base material.

As used herein, the "base material" includes all known fragrance ingredients selected from the extensive range of natural products, and synthetic molecules currently available, such as essential oils, alcohols, aldehydes and ketones, ethers and acetals, esters and lactones, macrocycles and heterocycles, and/or in admixture with one or more ingredients or excipients conventionally used in conjunction with odorants in fragrance compositions, for example, carrier materials, diluents, and other auxiliary agents commonly used in the art.

Fragrance ingredients known to the art are readily available commercially from the major fragrance manufacturers. Non-limiting examples of such ingredients include:

- essential oils and extracts, e.g. castoreum, costus root oil, oak moss absolute, geranium oil, tree moss absolute, basil oil, fruit oils, such as bergamot oil and mandarine oil, myrtle oil, palmarose oil, patchouli oil, petitgrain oil, jasmine oil, rose oil, sandalwood oil, wormwood oil, lavender oil and/ or ylang-ylang oil;
- alcohols, e.g. cinnamic alcohol ((*E*)-3-phenylprop-2-en-1-ol); cis-3-hexenol ((*Z*)-hex-3-en-1-ol); citronellol (3,7-dimethyloct-6-en-1-ol); dihydro myrcenol (2,6-dimethyloct-7-en-2-ol); Ebanol™ ((*E*)-3-methyl-5-(2,2,3-trimethylcyclopent-3-en-1-yl)pent-4-en-2-ol); eugenol (4-allyl-2-methoxyphenol); ethyl linalool ((*E*)-3,7-

dimethylnona-1,6-dien-3-ol); farnesol ((2*E*,6*Z*)-3,7,11-trimethyldodeca-2,6,10-trien-1-ol); geraniol ((*E*)-3,7-dimethylocta-2,6-dien-1-ol); Super Muguet™ ((*E*)-6-ethyl-3-methyloct-6-en-1-ol); linalool (3,7-dimethylocta-1,6-dien-3-ol); menthol (2-isopropyl-5-methylcyclohexanol); Nerol (3,7-dimethyl-2,6-octadien-1-ol); phenyl ethyl alcohol (2-phenylethanol); Rhodinol™ (3,7-dimethyloct-6-en-1-ol); Sandalore™ (3-methyl-5-(2,2,3-trimethylcyclopent-3-en-1-yl)pentan-2-ol); terpineol (2-(4-methylcyclohex-3-en-1-yl)propan-2-ol); or Timberol™ (1-(2,2,6-trimethylcyclohexyl)hexan-3-ol); 2,4,7-trimethylocta-2,6-dien-1-ol, and/or [1-methyl-2(5-methylhex-4-en-2-yl)cyclopropyl]-methanol;

- aldehydes and ketones, e.g. anisaldehyde (4-methoxybenzaldehyde); alpha amyl cinnamic aldehyde (2-benzylideneheptanal); Georgywood™ (1-(1,2,8,8-tetramethyl-1,2,3,4,5,6,7,8-octahydronaphthalen-2-yl)ethanone); Hydroxycitronellal (7-hydroxy-3,7-dimethyloctanal); Iso E Super® (1-(2,3,8,8-tetramethyl-1,2,3,4,5,6,7,8-octahydronaphthalen-2-yl)ethanone); Isoraldeine® ((*E*)-3-methyl-4-(2,6,6-trimethylcyclohex-2-en-1-yl)but-3-en-2-one); 3-(4-isobutyl-2-methylphenyl)propanal; maltol; methyl cedryl ketone; methylionone; verbenone; and/or vanillin;
- ether and acetals, e.g. Ambrox® (3a,6,6,9a-tetramethyl-2,4,5,5a,7,8,9,9b-octahydro-1*H*-benzo[*e*][1]benzofuran); geranyl methyl ether ((2*E*)-1-methoxy-3,7-dimethylocta-2,6-diene); rose oxide (4-methyl-2-(2-methylprop-1-en-1-yl)tetrahydro-2*H*-pyran); and/ or Spirambrene® (2',2',3,7,7-pentamethylspiro[bicyclo[4.1.0]heptane-2,5'-[1,3]dioxane]) ;
- macrocycles, e.g. Ambrettolide ((*Z*)-oxacycloheptadec-10-en-2-one); ethylene brassylate (1,4-dioxacycloheptadecane-5,17-dione); and / or Exaltolide® (16-oxacyclohexadecan-1-one); and
- heterocycles, e.g. isobutylquinoline (2-isobutylquinoline).

As used herein, "carrier material" means a material which is practically neutral from an odorant point of view, i.e. a material that does not significantly alter the organoleptic properties of odorants.

By “diluents” is meant any diluent conventionally used in conjunction with odorants, such as diethyl phthalate (DEP), dipropylene glycol (DPG), isopropyl myristate (IPM), triethyl citrate (TEC) and alcohol (e.g. ethanol).

The term “auxiliary agent" refers to ingredients that might be employed in a fragrance composition for reasons not specifically related to the olfactive performance of said composition. For example, an auxiliary agent may be an ingredient that acts as an aid to processing a fragrance ingredient or ingredients, or a composition containing said ingredient(s), or it may improve handling or storage of a fragrance ingredient or composition containing same, such as anti-oxidant adjuvant. Said anti-oxidant may be selected, for example, from Tinogard® TT (BASF), Tinogard® Q (BASF), Tocopherol (including its isomers, CAS 59-02-9; 364-49-8; 18920-62-2; 121854-78-2), 2,6-bis(1,1-dimethylethyl)-4-methylphenol (BHT, CAS 128-37-0) and related phenols, hydroquinones (CAS 121-31-9).

It might also be an ingredient that provides additional benefits such as imparting colour or texture. It might also be an ingredient that imparts light resistance or chemical stability to one or more ingredients contained in a fragrance composition.

A detailed description of the nature and type of auxiliary agent commonly used in fragrance compositions containing same cannot be exhaustive, but it has to be mentioned that said ingredients are well known to a person skilled in the art.

Various applications for the compound of formula (I) include but are not limited to a fine fragrance or a consumer product such as fabric care, toiletries, beauty care and cleaning products, detergent products, and soap products, including essentially all products where the currently available (+)-Amberketal ingredients are used commercially.

There is also provided herein a consumer product comprising a composition or fragrance composition as described herein, including any embodiment thereof. The consumer product may, for example, be a cosmetic product (e.g. a eau de parfum or eau de toilette), a cleaning product, a detergent product, or a soap product.

<u>INTERMEDIATES AND STARTING MATERIALS</u>

There is also provided herein the intermediates and starting materials used in the methods described herein.

There is also provided herein a mixture comprising, consisting essentially of, or consisting of a compound of formula (II). For example, mixture may comprise, consist essentially of, or consist of a compound of formula (II) wherein both double bonds are in *E*-configuration (*E,E*-compound) and a compound of formula (II) wherein the double bond between C-8 and C-9 is in *E*-configuration and the double bond between C-4 and C-5 is in Z-configuration (*E,Z*-compound). In one embodiment, said mixture may comprise, consit essentialy of, or consist of three of the stereoisomers of formula (II) (i.e. *E,Z*-, *E,E*- and *Z,E*- or *Z,Z*- compound of formula (II)), or even all four stereoisomers of formula (II).

The weight ratio of the *E,Z*-compound to the *E,E*-compound of formula (II) may, for example, be equal to or greater than about 10:90. For example, the weight ratio of the *E,Z*-compound to the *E,E*-compound may be equal to or greater than about 20:80 or equal to or greater than about 30:70 or equal to or greater than about 40:60 or equal to or greater than about 50:50 or equal to or greater than about 60:40 or equal to or greater than about 70:30 or equal to or greater than about 80:20 or equal to or greater than about 90:10 or equal to or greater than about 95:5, or equal to or greater than 99:1.

The weight ratio of the *E,Z*-compound to the *E,E*-compound of formula (II) may, for example, be equal to or less than about 99:1. For example, the weight ratio of the *E,Z*-compound to the *E,E*-compound of formula (II) may be equal to or less than about 95:5 or equal to or less than about 90:10 or equal to or less than about 85:15 or equal to or less than about 80:20 or equal to or less than about 70:30 or equal to or less than 60:40.

For example, the weight ratio of the *E,Z*- compound of formula (II) to the *E,E*-compound of formula (II) may range from about 10:90 to about 99:1 or from about 10:90 to about 90:10 or from about 20:80 to about 80:20 or from about 50:50 to about 80:20 or from about 60:40 to about 80:20.

SHC ENZYME OR ENZYME VARIANT

The methods described herein use an SHC enzyme or variant enzyme to enzymatically

convert a compound of formula (II) to a compound of formula (I).

As used herein, the term “SHC enzyme” means a wild-type Squalene Hopene Cyclase enzyme that is naturally occurring in, for example thermophilic bacteria such as
5 *Alicyclobacillus acidocaldarius*.

As used herein, the term "variant" is to be understood as a polypeptide which differs in comparison to the polypeptide from which it is derived by one or more changes in the amino acid sequence. The polypeptide from which a variant is derived is also known as
10 the parent or reference polypeptide. Typically a variant is produced artificially, preferably by gene-technological means. Typically, the polypeptide from which the variant is derived is a wild-type enzyme or wild-type enzyme. However, the variants usable in the present disclosure may also be derived from homologs, orthologs, or paralogs of the parent polypeptide or from artificially constructed variants. The changes
15 in the amino acid sequence may be amino acid exchanges (substitutions), insertions, deletions, *N*-terminal truncations, or *C*-terminal truncations, or any combination of these changes, which may occur at one or several sites.

As used herein, the term “SHC enzyme variant” means an enzyme that is derived from
20 a wild-type SHC enzyme but has one or more amino acid alterations compared to the wild-type SHC enzyme and is therefore not naturally occurring. The one or more amino acid alterations may, for example, modify (e.g. increase) the enzymatic activity for a substrate (e.g. compound of formula (II)). Alternatively, a variant SHC enzyme can be derived from an already exisiting SHC enzyme variant.

25

Assays for determining and quantifying SHC enzyme and/or SHC enzyme variant activity are described herein and are known in the art. By way of example, SHC enzyme and/or SHC enzyme variant activity can be determined by incubating purified SHC enzyme or enzyme variant or extracts from host cells or a complete recombinant
30 host organism that has produced the SHC enzyme or enzyme variant with an appropriate substrate under appropriate conditions and carrying out an analysis of the substrate and reaction products (e.g. by gas chromatography (GC) or HPLC analysis). Further details on SHC enzyme and/or SHC enzyme variant activity assays and

analysis of the reaction products are provided in the Examples. These assays may include producing the SHC enzyme variant in recombinant host cells (e.g. *E. coli*).

As used herein, the term "activity" means the ability of an enzyme to react with a
5 substrate to provide a target product. The activity can be determined in what is known as an activity test via the increase of the target product, the decrease of the substrate (or starting material) or via a combination of these parameters as a function of time. The SHC enzymes of the present disclosure may be characterized by their ability to convert a compound of formula (II) (e.g. hydroxyfarnesylacetone) to a compound of
10 formula (I) (e.g. (+)-Amberketal).

A "biological activity" as used herein, refers to any activity a polypeptide may exhibit, including without limitation: enzymatic activity; binding activity to another compound (e.g. binding to another polypeptide, in particular binding to a receptor, or binding to a
15 nucleic acid); inhibitory activity (e.g. enzyme inhibitory activity); activating activity (e.g. enzyme-activating activity); or toxic effects. It is not required that the variant exhibits such an activity to the same extent as the parent or wild-type polypeptide. In other embodiments, the SHC enzyme variants used herein show a better substrate conversion yield than the reference SHC enzyme (e.g. a wild-type SHC enzyme or a
20 known SHC enzyme variant). In additional embodiments, the SHC enzyme variants used herein may show a modified (e.g. increased) productivity relative to the reference SHC enzyme (e.g. 215G2 AacSHC *vs.* wild-type AacSHC). The term "productivity" refers to the amount of recoverable product in grams per litre of reaction per hour of bioconversion time (i.e. time after the substrate was added).
25

As used herein, the term “amino acid alteration” means an insertion of one or more amino acids between two amino acids, a deletion of one or more amino acids or a substitution (which may be conservative or non-conservative) of one or more amino acids with one or more different amino acids relative to the amino acid sequence of a
30 reference amino acid sequence. Substitutions replace the amino acids of the reference sequence with the same number of amino acids in the variant sequence. Reference amino acid sequences may, for example, be a wild-type (WT) amino acid sequence (for example SEQ ID NO: 1 or SEQ ID NO: 15, or SEQ ID NO: 16, or SEQ ID NO: 17, or SEQ ID NO: 18, or SEQ ID NO: 19, or SEQ ID NO: 20, or SEQ ID NO: 21) or may, for

example, itself be an SHC enzyme variant sequence (for example the AacSHC variant 215G2 – SEQ ID NO: 13).

The amino acid alterations can be easily identified by a comparison of the amino acid
5 sequences of the SHC enzyme variant with the amino acid sequence of the reference amino acid sequence (e.g. a wild-type or variant SHC amino acid sequence).

Suitable sources of SHC enzymes include, for example, *Alicyclobacillus acidocaldarius* (Aac), *Zymomonas mobilis* (Zmo), *Bradyrhizobium japonicum* (Bja), *Gluconobacter*
10 *morbifer* (Gmo), *Burkholderia ambifaria*, *Bacillus anthracis*, *Bacillus megaterium* (Bme), *Methylococcus capsulatus*, *Frankia alni*, *Acetobacter pasteurianus* (Apa), *Thermosynechococcus elongatus* (Tel), *Streptomyces coelicolor* (Sco), *Rhodopseudomonas palustris* (Rpa), *Teredinibacter urnerae* (Ttu), *Pelobacter carbinolicus* (Pca), and *Tetrahymena pyriformis* (see, for example WO 2010/139719,
15 US 2012/01345477, WO 2012/066059, WO 2016/170099; WO 2018/157021, and JP2009060799, the contents of which are incorporated herein by reference). Suitable enzymes are also described in e.g. Neumann & Simon 1986, Biol Chem Hoppe-Seyler 367, 723-729; Seckler & Poralla 1986, Biochem Biophys Act 356-363; Ochs *et al* 1990, J Bacteriol 174, 298-302; Seitz *et al* 2012, J Molecular Catalysis B: Enzymatic 84, 72-
20 77; and Seitz 2012 PhD thesis (http://elib.uni-stuttgart.de/handle/1 1682/1400), the contents of which are incorporated herein by reference. These SHC enzymes and variants may be used in the methods described herein.

In particular, the SHC enzyme (e.g. from which the SHC enzyme variant may be
25 derived) may be the *Alicyclobacillus acidocaldarius* (Aac) SHC enzyme, a *Zymomonas mobilis* (Zmo) SHC enzyme, a *Bradyrhizobium japonicum* (Bja) SHC enzyme, a *Bacillus megaterium* (Bme) SHC enzyme, or a *Gluconobacter morbifer* (Gmo) SHC enzyme. In a certain embodiment, the SHC enzyme (e.g. from which the SHC enzyme variant may be derived) may be the *Alicyclobacillus acidocaldarius* (Aac) SHC enzyme.
30 In a certain embodiment, the SHC enzyme (e.g. from which the SHC enzyme variant may be derived) may be the *Bacillus megaterium* (Bme) SHC enzyme.

For ease of reference, the designation “AacSHC” may be used to refer to the *Alicyclobacillus acidocaldarius* (Aac) SHC enzyme, “ZmoSHC” may be used to refer to

a *Zymomonas mobilis* (Zmo) SHC enzyme, "BmeSHC" may be used to refer to the *Bacillus megaterium* (Bme) SHC enzyme, "BjaSHC" may be used to refer to the *Bradyrhizobium japonicum* (Bja) SHC enzyme and "GmoSHC" may be used to refer to the *Gluconobacter morbifer* (Gmo) SHC enzyme.

5

AacSHC, ZmoSHC and BjaSHC enzyme sequences are disclosed in BASF WO 2010/139719, US 2012/01345477A1, Seitz *et al* (as cited above) and Seitz (2012 PhD thesis as cited above). Two different sequences are disclosed for ZmoSHC, referred to as ZmoSHC1 and ZmoSHC2. The GmoSHC enzyme sequence is disclosed in WO 10 2018/157021. Table 1 discloses sources and accession numbers of wild-type SHC enzymes.

Table 1. Sources and accession numbers of wild-type SHC enzymes.

SHC Source Strain (SHC name)	Reference	Accession No.
Alicyclobacillus acidocaldarius (WT AacSHC)	JP2009-060799 Neumann *et al* Biol Chem (1986) 367; 723-729	NBRC15652
Zymomonas mobilis (WT ZmoSHC)	WO2010139719 US20120135477	ATCC31821 PF62207_2 Genpept Accession No AAV90172
Zymomonas mobilis (WT ZmoSHC)	Reipen *et al* (1995) Microbiology 141:155-161	EMBL/Genbank Accession No. X80766
Bradryhizobium japonicum (WT BjaSHC)	WO2010139719 US2012/0135477	PF62207_5
Bacillus megaterium (WT BmeSHC)	WO2017/150695 WO2019/045058 WO2015/033746	WP_016763969.1
Burkholderia ambifaria *Bacillus anthracis* *Frankia alni* *Rhodopseudomonas palustris*	WO2010139719 US2012/0135477	
Gluconobacter morbifer	WO 2018/157021	

15 The amino acid sequences of the wild-type AacSHC, wild-type ZmoSHC1, wild-type ZmoSHC2, wild-type BjaSHC, wild-type GmoSHC, wild-type TelSHC, wild-type ApaSHC1, and wild-type BmeSHC are also disclosed herein (SEQ ID NO: 1, SEQ ID

NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, and SEQ ID NO: 21 respectively).

The SHC enzymes or enzyme variants described herein and used in the methods described herein may, for example, be based on one of the wild-type amino acid sequences (SEQ ID NO 1, 13, 15, 16, 17, 18, 19, 20, 21), or a variant, homologue, mutant, derivative or fragment thereof. The SHC enzyme or enzyme variant may, for example, have an amino acid sequence with at least 30%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%,70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to one of the wild-type amino acid sequences disclosed herein.

The SHC enzymes or enzyme variants described herein and used in the methods described herein may, for example, have a selectivity equal to or greater than about 75 %. For example, the SHC enzyme or enzyme variant may have a selectivity equal to or greater than about 80 % or equal to or greater than about 85 % or equal to or greater than about 90 % or equal to or greater than about 95 %. For example, the SHC enzyme or enzyme variant may have a selectivity up to 100 %, for example less than 100 %, for example equal to or less than about 99.5 % or equal to or less than about 99.0 % or equal to or less than about 98.0 % or equal to or less than about 97.0 %.

"Percent (%) identity" with respect to the nucleotide sequence of a gene is defined as the percentage of nucleotides in a candidate DNA sequence that is identical with the nucleotides in the DNA sequence, after aligning the sequence and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent nucleotide sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full

length of the sequences being compared. The terms "polypeptide" and "protein" are used interchangeably herein and mean any peptide-linked chain of amino acids, regardless of length or post-translational modification.

5 As used herein the term "derivative" includes but is not limited to a variant. The terms "derivative" and "variant" are used interchangeably herein.

Specific SHC enzymes and enzymes variants that may be used in the methods described herein are further described below.

10

Wild-type SHC Enzymes

The methods described herein may, for example, use a SHC enzyme having 100% sequence identity to a wild-type SHC enzyme. The wild-type SHC enzyme does not have to have been obtained directly from its natural organism and may have been
15 synthesized in a laboratory, for example using recombinant DNA techniques.

The wild-type SHC enzyme may, for example, be from *Alicyclobacillus acidocaldarius* (Aac), *Zymomonas mobilis* (Zmo), *Bradyrhizobium japonicum* (Bja), *Gluconobacter morbifer* (Gmo), *Burkholderia ambifaria*, *Bacillus anthracis, Bacillus megaterium* (Bme),
20 *Methylococcus capsulatus*, *Frankia alni*, *Acetobacter pasteurianus* (Apa), *Thermosynechococcus elongatus* (Tel), *Streptomyces coelicolor* (Sco), *Rhodopseudomonas palustris* (Rpa), *Teredinibacter turnerae* (Ttu), *Pelobacter carbinolicus* (Pca), or *Tetrahymena pyriformis* (see, for example WO 2010/139719, US 2012/01345477, WO 2012/066059, the contents of which are incorporated herein by
25 reference).

In particular, the wild-type SHC enzyme may be an *Alicyclobacillus acidocaldarius* (Aac) SHC enzyme, a *Zymomonas mobilis* (Zmo) SHC enzyme, a *Bacillus megaterium* (Bme) SHC enzyme, a *Bradyrhizobium japonicum* (Bja) SHC enzyme or a
30 *Gluconobacter morbifer* (Gmo) SHC enzyme. In particular, the wild-type SHC may be the *Alicyclobacillus acidocaldarius* (Aac) SHC enzyme, or the *Bacillus megaterium* (Bme) SHC enzyme, or a *Zymomonas mobilis* (Zmo) SHC enzyme (e.g. ZmoSHC1).

For ease of reference, the designation “AacSHC” may be used to refer to the

Alicyclobacillus acidocaldarius SHC enzyme, “ZmoSHC” may be used to refer to the *Zymomonas mobilis* SHC enzymes ZmoSHC1 and ZmoSHC2, “BjaSHC” may be used to refer to the *Bradyrhizobium japonicum* SHC enzyme, “BmeSHC” may be used to refer to the *Bacillus megaterium* SHC enzyme, “GmoSHC” may be used to refer to the
5 *Gluconobacter morbifer* SHC enzyme, “TelSHC” may be used to refer to the *Thermosynechococcus elongatus* SHC enzyme and “ApaSHC1” may be used to refer to the *Acetobacter pasteurianus* SHC enzyme.

The wild-type SHC enzyme amino acid sequence may, for example, be AacSHC (SEQ
10 ID NO: 1), ZmoSHC1 (SEQ ID NO: 15), ZmoSHC2 (SEQ ID NO: 16), BjaSHC (SEQ ID NO: 17), GmoSHC (SEQ ID NO: 20), BmeSHC (SEQ ID NO: 21), TelSHC (SEQ ID NO: 18) or ApaSHC1 (SEQ ID NO: 19). For example, the wild-type SHC enzyme may be BmeSHC (SEQ ID NO: 21), ZmoSHC1 (SEQ ID NO: 15), or AacSHC (SEQ ID NO: 1).
15

SHC Variant Enzymes

The methods described herein may, for example, use an SHC enzyme variant (i.e. an SHC enzyme having less than 100% sequence identity to a wild-type SHC enzyme).

20 The methods described herein may, for example, use an SHC enzyme variant as described in WO 2016/170099 or WO 2018/157021, the contents of which are incorporated herein by reference. For example, the SHC enzyme variant used in the methods described herein may be the SHC enzyme variant 215G2, which is described in WO 2016/170099.
25

The SHC enzyme variant may, for example, have an amino acid sequence having at least about 70.0 % identity to a wild-type SHC enzyme amino acid sequence. For example, the SHC enzyme variant may have an amino acid sequence having at least about 75.0 % or at least about 80.0 % or at least about 85.0 % or at least about 90.0 %
30 or at least about 95.0 % or at least about 95.5 % or at least about 96.5 % or at least about 97.0 % or at least about 97.5 % or at least about 98.0 % or at least about 98.5 % or at least about 99.0 % identity to a wild-type SHC enzyme amino acid sequence.

The SHC enzyme variant has an amino acid sequence having less than 100 % identity,

for example equal to or less than about 99.5 % or equal to or less than about 99.0 % identity to a wild-type SHC enzyme amino acid sequence.

For example, the SHC enzyme variant may have from about 70.0 % to about 99.5 % or from about 80.0 % to about 99.0 % or from about 85.0 % to about 98.5 % or from about 90.0 % to about 98.0 % identity to a wild-type SHC enzyme amino acid sequence.

The wild-type SHC enzyme may, for example, be from *Alicyclobacillus acidocaldarius* (Aac), *Zymomonas mobilis* (Zmo), *Bradyrhizobium japonicum* (Bja), *Gluconobacter morbifer* (Gmo), *Burkholderia ambifaria*, *Bacillus anthracis*, *Bacillus megaterium* (Bme), *Methylococcus capsulatus*, *Frankia alni*, *Acetobacter pasteurianus* (Apa), *Thermosynechococcus elongatus* (Tel), *Streptomyces coelicolor* (Sco), *Rhodopseudomonas palustris* (Rpa), *Teredinibacter turnerae* (Ttu), *Pelobacter carbinolicus* (Pca), or *Tetrahymena pyriformis* (see, for example WO 2010/139719, US 2012/01345477, WO 2012/066059, the contents of which are incorporated herein by reference).

The wild-type SHC enzyme amino acid sequence may, for example, be AacSHC (SEQ ID NO: 1), ZmoSHC1 (SEQ ID NO: 15), ZmoSHC2 (SEQ ID NO: 16), BjaSHC (SEQ ID NO: 17), GmoSHC (SEQ ID NO: 20), BmeSHC (SEQ ID NO: 21), TelSHC (SEQ ID NO: 18) or ApaSHC1 (SEQ ID NO: 19). For example, the wild-type SHC enzyme may be AacSHC (SEQ ID NO: 1).

Therefore, in certain embodiments, the SHC enzyme or SHC enzyme variant may have an amino acid sequence having at least about 70.0 % identity to SEQ ID NO: 1, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 20, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 21. For example, the SHC enzyme or SHC enzyme variant has an amino acid sequence having at least about 75.0 % or at least about 80.0 % or at least about 85.0 % or at least about 90.0 % or at least about 95.0 % or at least about 95.5 % or at least about 96.5 % or at least about 97.0 % or at least about 97.5 % or at least about 98.0 % or at least about 98.5 % or at least about 99.0 % identity to SEQ ID NO: 1, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 20, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 21.

For example, the SHC enzyme variant may, for example, have an amino acid sequence having less than 100 % identity, for example equal to or less than about 99.5 % or equal to or less than about 99.0 % identity to SEQ ID NO: 1, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 20, SEQ ID NO: 18, SEQ ID NO: 19, or
5 SEQ ID NO: 21.

For example, the SHC enzyme variant may have from about 70.0 % to about 99.5 % or from about 80.0 % to about 99.0 % or from about 85.0 % to about 98.5 % or from about 90.0 % to about 98.0 % identity to SEQ ID NO: 1, SEQ ID NO: 15, SEQ ID NO: 16,
10 SEQ ID NO: 17, SEQ ID NO: 20, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 21.

"Percent (%) identity" with respect to a polypeptide or nucleotide sequence is defined respectively as the percentage of amino acids or nucleotides in a candidate sequence that are identical with the amino acids or nucleotides in the reference sequence, after
15 aligning the sequence and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software. Those skilled in the art can determine
20 appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. The terms "polypeptide" and "protein" are used interchangeably herein and mean any peptide-linked chain of amino acids, regardless of length or post-translational modification.
25

The similarity of nucleotide and amino acid sequences, i.e. the percentage of sequence identity, can be determined via sequence alignments. Such alignments can be carried out with several art-known algorithms, preferably with the mathematical algorithm of Karlin and Altschul (Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873-
30 5877), with hmmalign (HMMER package, http://hmmer.wustl.edu/) or with the CLUSTAL algorithm (Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673-80) available eg. on https://www.ebi.ac.uk/Tools/msa/clustalo/ or the GAP program (mathematical algorithm of the University of Iowa) or the mathematical algorithm of Myers and Miller (1989 - Cabios 4: 11-17). Preferred

parameters used are the default parameters as they are set on https://www.ebi.ac.uk/Tools/msa/clustalo/.

Percentage sequence identity may be calculated using, for example, BLAST, BLAT or
5 BlastZ (or BlastX). A similar algorithm is incorporated into the BLASTN and BLASTP programs of Altschul et al (1990) J. Mol. Biol. 215, 403-410. BLAST polynucleotide searches may be performed with the BLASTN program, score = 100, word length = 12, to obtain polynucleotide sequences that are homologous to those nucleic acids which encode the relevant protein. BLAST protein searches may be performed with the
10 BLASTP program, score = 50, word length = 3, to obtain amino acid sequences homologous to the polypeptide.

To obtain gapped alignments for comparative purposes, Gapped BLAST may be utilized as described in Altschul et al (1997) Nucleic Acids Res. 25, 3389-3402. When
15 utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs are used. Sequence matching analysis may be supplemented by established homology mapping techniques like Shuffle-LAGAN (Brudno M., Bioinformatics 2003b, 19 Suppl 1: 154-162) or Markov random fields. When percentages of sequence identity are referred to in the present application, these
20 percentages are calculated in relation to the full length of the longer sequence, if not specifically indicated otherwise.

In particular embodiments, % identity between two sequences is determined using CLUSTAL O (version 1.2.4).

25

The SHC enzyme variants may, for example, have increased enzymatic activity for the conversion of the compound of formula (II) to the compound of formula (I) compared to the parent SHC enzyme. Increased enzymatic activity may refer to any aspect of the enzymatic conversion of the compound of formula (II) to the compound of formula (I)
30 including, for example, increased total conversion of the compound of formula (II) to the compound of formula (I), increased rate of conversion of the compound of formula (II) (e.g. in the first 4 hours, or first 6 hours, or in the first 12 hours of reaction), increased production of the compound of formula (I), and/or decreased production of by-products. Increased enzymatic activity may be defined by increased productivity in general,

which may be defined in terms of compound of formula (I) produced per hour, per gram of biocatalyst and per litre of reaction.

The SHC enzyme variants may, for example, provide increased compound of formula (II) conversion compared to the parent SHC enzyme. Therefore, the process described herein may have an increased level of compound of formula (II) conversion compared to the process using the parent SHC enzyme. The SHC enzyme variants may, for example, provide increased rate of compound of formula (II) conversion compared to the parent SHC enzyme. Therefore, the process described herein may have an increased rate of compound of formula (II) conversion compared to the parent SHC enzyme. The SHC enzyme variant may, for example, provide increased rate of compound of formula (II) conversion over the first 4 hours or over the first 2 hours, or first 4 hours, or first 6 hours, or over the first 8 hours or over the first 12 hours or over the first 24 hours of the reaction compared to the parent SHC enzyme. Therefore, the process described herein may have an increased rate of compound of formula (II) conversion over the first 2 hour, or first 4 hours or over the first 6 hours or over the first 8 hours or over the first 12 hours or over the first 24 hours of the reaction compared to the parent SHC enzyme. This may be when compared to using both enzymes (i.e. the SHC enzyme variant and the parent SHC enzyme) under the same reaction conditions (e.g. same pH and temperature) or when compared to using each enzyme under conditions that have been individually defined as optimal for its activity (e.g. optimized pH and temperature), and which may be different to each other.

The conversion of compound of formula (II) to compound of formula (I) may, for example, be determined using an activity assay as described above and may be calculated as gram of recoverable product per gram of feedstock (which can be calculated if desired as a percent molar conversion rate).

The processes for making the compound of formula (I) disclosed herein may be carried out at the optimum temperature range or optimum temperature and/or the optimum pH range or optimum pH and/or the SDS optimum concentration range or optimum SDS concentration for the specific enzyme used, as set out in the Examples below.

NUCLEIC ACIDS AND METHODS OF MAKING NUCLEIC ACIDS

The SHC enzyme and enzyme variants described herein may be encoded by a nucleic acid sequence. The nucleic acid containing the coding sequence may, for example, be an isolated nucleic acid.

5

Thus, there is provided herein a construct comprising a nucleic acid sequence encoding an SHC enzyme or enzyme variant as described herein. As used herein, a “construct” is an artificially created segment of nucleic acid that is to be transfected into a target cell. The construct may comprise the nucleic acid sequence encoding the SHC
10 enzyme or enzyme variant and a gene expression controller (e.g. promoter).

There is further provided herein a vector comprising a construct as described herein. As used herein, a “vector” is a DNA molecule that is used as a vehicle to artificially carry foreign genetic material into a cell where it can be replicated and/or expressed.
15 The vector may, for example, be a plasmid, a viral vector, a cosmid, or an artificial chromosome.

The terms “construct” and “vector” may overlap, for example where the construct is a plasmid.

20

The term “nucleic acid” or "nucleic acid molecule" as used herein shall specifically refer to polynucleotides of the disclosure which can be DNA, cDNA, genomic DNA, synthetic DNA, or RNA, and can be double-stranded or single-stranded, the sense and/or an antisense strand. The term “nucleic acid” or "nucleic acid molecule" shall particularly
25 apply to the polynucleotide(s) as used herein, e.g. as full-length nucleotide sequence or fragments or parts thereof, which encodes a polypeptide with enzymatic activity, e.g. an enzyme of a metabolic pathway, or fragments or parts thereof, respectively.

The term also includes a separate molecule such as a cDNA where the corresponding
30 genomic DNA has introns and therefore a different sequence; a genomic fragment that lacks at least one of the flanking genes; a fragment of cDNA or genomic DNA produced by polymerase chain reaction (PCR) and that lacks at least one of the flanking genes; a restriction fragment that lacks at least one of the flanking genes; a DNA encoding a non-naturally occurring protein such as a fusion protein (e.g. a His tag), mutein, or

fragment of a given protein; and a nucleic acid which is a degenerate variant of a cDNA or a naturally occurring nucleic acid. In addition, it includes a recombinant nucleotide sequence that is part of a hybrid gene, i.e. a gene encoding a non-naturally occurring fusion protein. Fusion proteins can add one or more amino acids (such as but not limited to Histidine (His)) to a protein, usually at the *N*-terminus of the protein but also at the *C*-terminus or fused within regions of the protein. Such fusion proteins or fusion vectors encoding such proteins typically serve three purposes: (i) to increase production of recombinant proteins; (ii) to increase the solubility of the recombinant protein; and (iii) to aid in the purification of the recombinant protein by providing a ligand for affinity purification.

The term “nucleic acid sequence” also includes codon-optimised sequences suitable for expression in a particular microbial host cell (e.g. *E. coli* host cell). As used herein, the term "codon optimized" means a nucleic acid protein coding sequence which has been adapted for expression in a particular prokaryotic or a eukaryotic host cell, particularly bacterial host cells such as *E. coli* host cells considering its particular codon-usage, by e.g. substitution of one or more or preferably a significant number of codons with codons that are more frequently used in bacterial host cell genes (e.g. *E. coli genes*).

In this regard, the nucleotide sequence or gene encoding the reference amino acid sequence (e.g. SEQ ID NO: 1 or SEQ ID NO: 13) and variants/derivatives thereof may be the original one as found in the source or the gene can be codon-optimized for the selected host organisms, such as e.g. *E. coli.*

In a further aspect the nucleic acid sequence(s) of the present disclosure is/are operatively linked to expression control sequences allowing expression in prokaryotic and/or eukaryotic host cells. As used herein, "operatively linked" means incorporated into a genetic construct so that expression control sequences effectively control expression of a coding sequence or of a gene of interest. The transcriptional/translational regulatory elements referred to above include but are not limited to inducible and non-inducible, constitutive, cell cycle regulated, metabolically regulated promoters, enhancers, operators, silencers, repressors and other elements that are known to those skilled in the art and that drive or otherwise regulate gene

expression. Such regulatory elements include but are not limited to regulatory elements directing constitutive expression or which allow inducible expression like, for example, CUP-1 promoter, the tet-repressor as employed, for example, in the tet-on or tet-off systems, the lac system, the trp system regulatory elements. By way of example, Isopropyl β-D-1-thiogalactopyranoside (IPTG) is an effective inducer of gene expression in the concentration range of e.g. 100 μM to 1.0 mM. This compound is a molecular mimic of allolactose, a lactose metabolite that triggers transcription of the lac operon, and it is therefore used to induce gene expression when the gene is under the control of the lac operator. Another example of a regulatory element which induces gene expression is lactose. Similarly, the nucleic acid molecule(s) of the present disclosure can form part of a hybrid gene encoding additional polypeptide sequences, for example, a sequence that functions as a marker or reporter. Examples of marker and reporter genes including beta-lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase dihydrofolate reductase (DHFR), hygromycin-B-phosphotransferase (HPH), thymidine kinase (TK), lacZ (encoding beta- galactosidase), and xanthine guanine phosphoribosyltransferase (XGPRT). As with many of the standard procedures associated with the practice of the disclosure, skilled artisans will be aware of additional useful reagents, for example, additional sequences that can serve the function of a marker or reporter.

There is also provided herein a recombinant polynucleotide encoding the SHC enzyme or variant thereof, which may be inserted into a vector for gene expression and optional enzyme purification. One type of vector is a plasmid representing a circular double stranded DNA loop into which additional DNA segments are ligated. Certain vectors can control the expression of genes to which they are functionally linked. These vectors are called "expression vectors". Usually expression vectors suitable for DNA recombination techniques are of the plasmid type. Typically, an expression vector comprises a gene for the production of the SHC wild-type or variant enzyme or as described herein. In the present description, the terms "plasmid" and "vector" may be used interchangeably since the plasmid is the vector type most often used.

Such vectors can include DNA sequences which include but are not limited to DNA sequences that are not naturally present in the host cell, DNA sequences that are not normally transcribed into RNA or translated into a protein ("expressed") and other

genes or DNA sequences which one desires to introduce into the non-recombinant host. It will be appreciated that typically the genome of a recombinant host described herein is augmented through the stable introduction of one or more recombinant genes. However, autonomous or replicative plasmids or vectors can also be used within the
5 scope of this disclosure. Moreover, the present disclosure can be practiced using a low copy number, e.g. a single copy, or high copy number (as exemplified herein) plasmid or vector.

The vector of the present disclosure includes plasmids, phagemids, phages, cosmids,
10 artificial bacterial and artificial yeast chromosomes, knock-out or knock-in constructs, synthetic nucleic acid sequences or cassettes and subsets may be produced in the form of linear polynucleotides, plasmids, megaplasmids, synthetic or artificial chromosomes, such as plant, bacterial, mammalian or yeast artificial chromosomes.

15 It is preferred that the proteins encoded by the introduced polynucleotide are expressed within the cell upon introduction of the vector. The plasmids are often standard cloning vectors, e.g. bacterial multicopy plasmids. The substrates can be incorporated into the same or different plasmids. Often at least two different types of plasmid having different types of selectable markers are used to allow selection for
20 cells containing at least two types of vectors.

Typically bacterial or yeast cells may be transformed with any one or more nucleotide sequences as is well known in the art. For *in vivo* recombination, the gene to be recombined with the chromosome or other genes is used to transform the host using
25 standard transforming techniques. In a suitable embodiment DNA providing an origin of replication is included in the construct. The origin of replication may be suitably selected by the skilled person. Depending on the nature of the genes, a supplemental origin of replication may not be required if sequences are already present with the genes that are operable as origins of replication themselves.

30

HOST CELLS, METHODS OF MAKING HOST CELLS AND METHODS OF MAKING THE COMPOUND OF FORMULA (I) USING HOST CELLS

Recombinant host cells may be used in the methods described herein.

There is further provided herein a recombinant host cell comprising a nucleic acid

sequence or a construct or a vector as described herein. There is further provided herein a recombinant host cell that produces an SHC enzyme or enzyme variant as described herein.

5 The processes described herein for producing the compound of formula (I) may, for example, comprise culturing a recombinant host cell as described herein. As used herein, the term “culturing” refers to a process multiplying living cells such that they produce an SHC enzyme or enzyme variant as described herein that can be used in a process for producing the compound of formula (I) as described herein.

10

A bacterial or yeast cell may be transformed by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated, i.e. covalently linked into the chromosome of the cell. In prokaryotes, and yeast, for example, the transforming DNA may be maintained on an episomal element

15 such as a plasmid. With respect to eukaryotic cells, a stably transfected cell is one in which the transfected DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA.

20

Generally, the introduced DNA is not originally resident in the host that is the recipient of the DNA, but it is within the scope of the disclosure to isolate a DNA segment from a given host, and to subsequently introduce one or more additional copies of that DNA into the same host, e.g. to enhance production of the product of a gene or alter the

25 expression pattern of a gene. In some instances, the introduced DNA will modify or even replace an endogenous gene or DNA sequence, e.g. by homologous recombination or site-directed mutagenesis. Suitable recombinant hosts include microorganisms, plant cells, and plants.

30 The present disclosure also features recombinant hosts. The term "recombinant host", also referred to as a "genetically modified host cell" or a "transgenic cell" denotes a host cell that comprises a heterologous nucleic acid or the genome of which has been augmented by at least one incorporated DNA sequence. A host cell of the present

disclosure may be genetically engineered with the polynucleotide or the vector as outlined above.

The host cells that may be used for purposes of the disclosure include but are not limited to prokaryotic cells such as bacteria (for example, *E. coli* and *B. subtilis*), which may, for example, be transformed with, for example, recombinant bacteriophage DNA, plasmid DNA, bacterial artificial chromosome, or cosmid DNA expression vectors containing the polynucleotide molecules of the disclosure; simple eukaryotic cells like yeast (for example, *Saccharomyces* and *Pichia*), which may, for example, be transformed with, for example, recombinant yeast expression vectors containing the polynucleotide molecule of the disclosure. Depending on the host cell and the respective vector used to introduce the polynucleotide of the disclosure the polynucleotide can integrate, for example, into the chromosome or the mitochondrial DNA or can be maintained extrachromosomally like, for example, episomally or can be only transiently comprised in the cells.

The term "cell" as used herein in particular with reference to genetic engineering and introducing one or more genes or an assembled cluster of genes into a cell, or a production cell is understood to refer to any prokaryotic or eukaryotic cell. Prokaryotic and eukaryotic host cells are both contemplated for use according to the disclosure, including bacterial host cells like *E. coli* or *Bacillus sp.*, yeast host cells, such as *S. cerevisiae*, insect host cells, such as *Spodoptora frugiperda* or human host cells, such as HeLa and Jurkat.

Specifically, the cell is a eukaryotic cell, preferably a fungal, mammalian or plant cell, or a prokaryotic cell. Suitable eukaryotic cells include, for example, without limitation, mammalian cells, yeast cells, or insect cells (including Sf9), amphibian cells (including melanophore cells), or worm cells including cells of *Caenorhabditis* (including *Caenorhabditis elegans*). Suitable mammalian cells include, for example, without limitation, COS cells (including Cos-1 and Cos-7), CHO cells, HEK293 cells, HEK293T cells, HEK293 T-RexTM cells, or other transfectable eukaryotic cell lines. Suitable bacterial cells include without limitation *E. coli*.

Preferably prokaryotes, such as *E. coli*, *Bacillus*, *Streptomyces*, or mammalian cells,

like HeLa cells or Jurkat cells, or plant cells, like *Arabidopsis*, may be used.

The cell may, for example, be selected from prokaryotic, yeast, plant, and/or insect host cells.

5

Preferably the cell is an *Aspergillus* sp. or a fungal cell, preferably, it can be selected from the group consisting of the genera *Saccharomyces*, *Candida*, *Kluyveromyces*, *Hansenula*, *Schizosaccharomyces*, *Yarrowia*, *Pichia* and *Aspergillus*.

10 Preferably, the cell is a bacteria cell, for example, of genus selected from *Escherichia*, *Streptomyces*, *Bacillus*, *Pseudomonas*, *Lactobacillus* and *Lactococcus*. For example, the bacteria may be *E. coli*.

Preferably the *E. coli* host cell is an *E. coli* host cell which is recognized by the industry
15 and regulatory authorities (including but not limited to an *E. coli* K12 host cell or an *E. coli* BL21 host cell).

One preferred host cell to use with the present disclosure is *E. coli*, which may be recombinantly prepared as described herein. Thus, the recombinant host may be a
20 recombinant *E. coli* host cell. There are libraries of mutants, plasmids, detailed computer models of metabolism and other information available for *E. coli*, allowing for rational design of various modules to enhance product yield.

In one embodiment, the recombinant *E. coli* microorganism comprises nucleotide
25 sequences encoding an SHC enzyme or enzyme variant (e.g. the microorganism comprises a nucleodite sequence selected from SEQ ID NO: 2, 4, 6, 8, 10, 12, and 14).

Preferably, the recombinant *E. coli* microorganism comprises a vector construct as described herein. In another preferred embodiment, the recombinant *E. coli*
30 microorganism comprises nucleotide sequences encoding the SHC enzymes and or enzyme variant disclosed herein.

Another preferred host cell to use with the present disclosure is *S. cerevisiae.* There are libraries of mutants, plasmids, detailed computer models of metabolism and other

information available for *S. cerevisiae*, allowing for rational design of various modules to enhance product yield. Methods are known for making recombinant *S. cerevisiae* microorganisms.

5 Culturing of cells may be performed in a conventional manner. The culture medium may contain a carbon source, at least one nitrogen source and inorganic salts, and vitamins are added to it. The constituents of this medium can be the ones which are conventionally used for culturing the species of microorganism in question. Carbon sources of use in the instant method include any molecule that can be metabolized by 10 the recombinant host cell to facilitate growth and/or production of the SHC enzyme of interest for the conversion of a compound of formula (II) to the compound of formula (I). Examples of suitable carbon sources include, but are not limited to, sucrose (e.g. as found in molasses), fructose, xylose, glycerol, glucose, cellulose, starch, cellobiose or other glucose containing polymer.

15

In embodiments employing yeast as a host, for example, carbon sources such as sucrose, fructose, xylose, ethanol, glycerol, and glucose are suitable. The carbon source can be provided to the host organism throughout the cultivation period in batch or fed-batch, or alternatively, another energy source can be used, like e.g. protein or 20 protein hydrolysate.

The recombinant host cell microorganism for use in the methods of the present disclosure may be propagated in a rich medium (e.g. LB-medium, Bacto-tryptone yeast extract medium, nutrient medium and the like) at a pH, temperature and under reaction 25 conditions commonly used for propagation of the microorganism. In one embodiment of the present disclosure, a defined minimal medium such as M9A is used for cultivation.

The components of M9A medium comprise: 14 g/l KH_2PO_4, 16 g/l K_2HPO_4, 1 g/l $Na_3Citrate.2H_2O$, 7.5 g/l $(NH_4)_2SO_4$, 0.25 g/l $MgSO_4.7H_2O$, 0.015 g/l $CaCl_2.2H_2O$, 5 g/l 30 glucose and 1.25 g/l yeast extract).

In another embodiment of the present disclosure, nutrient rich medium such as LB was used. The components of LB medium comprise: 10 g/l tryptone, 5 g/l yeast extract, 5

g/l NaCl. Other examples of Mineral Medium and M9 Mineral Medium are disclosed, for example, in US 6524831B2 and US 2003/0092143A1.

Another example of a minimal medium may be prepared as follows: for 350 ml culture:
5 to 35 ml citric acid/phosphate stock (133 g/l KH_2PO_4, 40 g/l $(NH_4)_2HPO_4$, 17 g/l citric acid.H_2O with pH adjusted to 6.3) was added 307 ml H_2O, the pH adjusted to 6.8 with 32% NaOH as required. After autoclaving 0.850 ml 50% $MgSO_4$, 0.035 ml trace elements solution (see below) solution, 0.035 ml Thiamin solution and 7 ml 20% glucose were added.
10

Trace elements solution: 50 g/l $Na_2EDTA.2H_2O$, 20 g/l $FeSO_4.7H_2O$, 3 g/l H_3BO_3, 0.9 g/l $MnSO_4.2H_2O$, 1.1 g/l $CoCl_2$, 80 g/L $CuCl_2$, 240 g/l $NiSO_4.7H_2O$, 100 g/l KI, 1.4 g/l $(NH_4)_6Mo_7O_{24}.4H_2O$, 1 g/l $ZnSO_4.7H_2O$, in deionized water

15 Thiamin solution: 2.25 g/l Thiamin.HCl in deionized water

$MgSO_4$ solution : 50 % (w/v) $MgSO_4.7H_2O$ in deionized water

The recombinant microorganism may be grown in a batch, fed batch or continuous
20 process or combinations thereof. Typically, the recombinant microorganism is grown in a fermentor at a defined temperature in the presence of a suitable nutrient source, e.g. a carbon source, for a desired period of time to produce sufficient SHC enzyme to be able to convert the compound of formula (II) to the compound of formula (I) and to produce a desired amount of the compound of formula (I). The recombinant host cells
25 may be cultivated in any suitable manner, for example by batch cultivation or fed-batch cultivation.

As used herein, the term "batch cultivation" is a cultivation method in which culture medium is neither added nor withdrawn during the cultivation.
30

As used herein, the term "fed-batch" means a cultivation method in which culture medium is added during the cultivation but no culture medium is withdrawn.

One embodiment of the present disclosure provides a method of producing the

compound of formula (I) in a cellular system comprising producing the SHC enzyme or enzyme variant under suitable conditions in a cellular system, feeding the compound of formula (II) to the cellular system, converting the compound of formula (II) to the compound of formula (I) using the SHC enzymes or enzyme variants produced using
5 the cellular system, collecting the compound of formula (I) from cellular system and optionally isolating the compound of formula (I) from the system. Expression of other nucleotide sequences may serve to enhance the method. The expression of other nucleotide sequences may enhance the activity of the cellular system used in the bioconversion for making the compound of formula (I).
10

A further embodiment of the present disclosure is a bioconversion method of making the compound of formula (I) comprising growing host cells comprising a gene coding for an SHC enzyme or enzyme variant, producing SHC enzymes or enzyme variants in the host cells, feeding the compound of formula (II) to the host cells, incubating the
15 host cells under conditions of pH, temperature and solubilizing agent suitable to promote the conversion of the compound of formula (II) to the compound of formula (I) and collecting the compound of formula (I). The production of the SHC enzymes or enzyme variants in the host cells provides a method of making the compound of formula (I) when the compound of formula (II) is added to the host cells under suitable
20 reaction conditions. Achieved conversion may be enhanced by adding more biocatalyst and SDS to the reaction mixture.

The recombinant host cell microorganism may be cultured in a number of ways in order to provide cells in suitable amounts that have produced the SHC enzyme or enzyme
25 variant for the subsequent bioconversion step. Since the microorganisms applicable for the bioconversion step vary broadly (e.g. yeasts, bacteria and fungi), culturing conditions are, of course, adjusted to the specific requirements of each species and these conditions are well known and documented. Any of the art known methods for growing cells of recombinant host cell microorganisms may be used to produce the
30 cells utilizable in the subsequent bioconversion step of the present disclosure. Typically the cells are grown to a particular density (measurable as optical density (OD)) to produce a sufficient biomass for the bioconversion reaction.

The cultivation conditions chosen influence not only the amount of cells obtained (the

biomass) but the quality of the cultivation conditions also influences how the biomass becomes a biocatalyst. The recombinant host cell microorganism expressing the SHC enzyme or enzyme variant gene and producing the SHC enzyme or enzyme variant is termed a biocatalyst which is suitable for use in a bioconversion reaction. In some
5 embodiments the biocatalyst is a recombinant whole cell producing SHC enzymes or enzyme variants or it may be in suspension or an immobilized format. In other embodiments, the biocatalyst is a membrane fraction or a liquid fraction prepared from the recombinant whole cell producing the SHC enzyme or enzyme variant (as disclosed for example in Seitz et al 2012 - as cited above). The recombinant whole cell
10 producing SHC enzymes or enzyme variants include whole cells collected from the fermenter (for the bioconversion reaction) or the cells in the fermenter (which are then used in a one-pot reaction). The recombinant whole cell producing SHC enzymes or enzyme variants can include intact recombinant whole cell and/or cell debris. Either way, the SHC enzyme or enzyme variant is associated with a membrane (such as a
15 cell membrane) in some way in order to receive and/or interact with a substrate (e.g. compound of formula (II)), which membrane (such as a cell membrane) can be part of a whole cell (e.g. a recombinant whole cell). The SHC enzymes or enzyme variants may also be in an immobilized form (e.g. associated with an enzyme carrier) which allows the SHC enzymes or enzyme variants to interact with a substrate (e.g.
20 compound of formula (II)). The SHC enzymes or enzyme variants may also be used in a soluble form.

In one embodiment, the biocatalyst is produced in sufficient amounts (to create a sufficient biomass), harvested and washed (and optionally stored (e.g. frozen or
25 lyophilized)) before the bioconversion step.

In a further embodiment, the cells are produced in sufficient amounts (to create a sufficient biocatalyst) and the reaction conditions are then adjusted without the need to harvest and wash the biocatalyst for the bioconversion reaction. This one step (or "one
30 pot") method is advantageous as it simplifies the process. The culture medium used to grow the cells is also suitable for use in the bioconversion reaction provided that the reaction conditions are adjusted to facilitate the bioconversion reaction.

The optimum pH for growing the cells is in the range of 6.0 - 8.0. The optimum pH for

the bioconversion reaction is dependent on the type of SHC enzyme or enzyme variant used in the bioconversion reaction. The pH is regulated using techniques which are well known to the skilled person.

5 The bioconversion methods of the present disclosure are carried out under conditions of time, temperature, pH and solubilizing agent to provide for conversion of the compound of formula (II) to the compound of formula (I).

The pH of the reaction mixture may be in the range of 4-8, preferably, 4.5 to 6.5, more 10 preferably 4.5-6.5 for the SHC wild-type enzyme or SHC enzyme variant considered and can be maintained by the addition of buffers to the reaction mixture. An exemplary buffer for this purpose is a citric acid buffer, or a succinic acid buffer.

The temperature is between from about 15°C to about 60°C, for example from about 15 15°C to about 50°C or from about 15°C to about 45°C or from about 30°C to about 60°C or from about 35°C to about 55°C for the SHC wild-type enzyme or SHC enzyme variant considered. The temperature can be kept constant or can be altered during the bioconversion process.

20 The [SDS]/[cells] ratio may be in the range of about, 10:1-20:1, preferably about 15:1 - 18:1, preferably about 16:1 when the ratio of biocatalyst to compound of formula (II) is about 2:1.

The processes for making the compound of formula (I) disclosed herein may be carried 25 out at the optimum temperature, pH and surfactant concentration enabling optimal activity of each of the individual SHC enzyme (wild-type or variant) considered.

It may be useful to include a solubilizing agent (e.g. a surfactant, detergent, solubility enhancer, water miscible organic solvent and the like) in the bioconversion reaction.

30

As used herein, the term "surfactant" means a component that lowers the surface tension (or interfacial tension) between two liquids or between a liquid and a solid. Surfactants may act as detergents, wetting agents, emulsifiers, foaming agents, and dispersants. Examples of surfactants include but are not limited to Triton X-100, Tween

80, taurodeoxycholate, Sodium taurodeoxycholate, Sodium dodecyl sulfate (SDS), and/or sodium lauryl sulfate (SLS).

Whilst Triton X-100 may be used to partially purify the SHC enzyme or enzyme variant (in soluble or membrane fraction / suspension form), it may also be used in the bioconversion reaction (see for example the disclosure in Seitz (2012 PhD thesis as cited above) as well as the disclosure in Neumann and Simon (1986 - as cited above) and JP2009060799. SDS may be used as a solubilizing agent.

Without wishing to be bound by theory, the use of SDS with recombinant microbial host cells may be advantageous as the SDS may interact advantageously with the host cell membrane in order to make the SHC enzyme or enzyme variant (which is a membrane bound enzyme) more accessible to the compound of formula (II) substrate. In addition, the inclusion of SDS at a suitable level in the reaction mixture may improve the properties of the emulsion (e.g. compound of formula (II) in water) and/or improve the access of the compound of formula (II) substrate to the SHC enzyme within the host. The concentration of the solubilising agent (e.g. SDS) used in the bioconversion reaction is influenced by the biomass amount and the substrate concentration. That is, there is a degree of interdependency between the solubilising agent (e.g. SDS) concentration, the biomass amount and the substrate concentration. By way of example, as the concentration of compound of formula (II) substrate increases, sufficient amounts of biocatalyst and solubilising agent (e.g. SDS) are required for an efficient bioconversion reaction to take place. If, for example, the solubilising agent (e.g. SDS) concentration is too low, a suboptimal conversion of compound of formula (II) may be observed. On the other hand, if, for example, the solubilising agent (e.g. SDS) concentration is too high, then there may be a risk that the biocatalyst is affected through either the disruption of the intact microbial cell and/or denaturation/inactivation of the SHC enzyme or enzyme variant. The selection of a suitable concentration of SDS in the context of the biomass amount and, substrate must be carefully investigated.

In some embodiments, the compound of formula (I) is produced using a biocatalyst to which the compound of formula (II) substrate is added.

It is possible to add the substrate by feeding using known means (e.g. peristaltic pump, infusion syringe and the like). The compound of formula (II) may be oil soluble and provided in an oil format. Given that the biocatalyst (microbial cells such as intact recombinant whole cell and/or cell debris and/or immobilised enzyme) is present in an
5 aqueous phase, the bioconversion reaction may be regarded as a three phase system (comprising an aqueous phase, a solid phase and an oil phase) when compound of formula (II) is added to the bioconversion reaction mixture. This is the case even when SDS is present.

10 A fermenter may be used to grow recombinant host cells producing the SHC enzyme or enzyme variant gene and producing active SHC enzymes or enzyme variants to a sufficient biomass concentration suitable for use as a biocatalyst in the same fermenter vessel which is used to convert the compound of formula (II) to the compound of formula (I), for example in admixture with one or more of the by-product of formula (III).
15

The skilled person will understand that higher cumulative production titers can be achieved by implementing a continuous process, such as product removal, substrate feed, and biomass addition or (partial) replacement. Preferably the bioconversion of compound of formula (II) into compound of formula (I) in the presence of a recombinant
20 host cell comprising an SHC enzyme or enzyme variant generates a compound of formula (I) yield of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93,
25 94, 95, 96, 97, 98, 99, 100, given in mol percent and based on the mols of compound of formula (II) employed; especially preferably, the yield is between 5 and 100, 10 and 100, and 100, 25 and 100, 30 and 100, 35 and 100, in particular between 40 and 100, 45 and 100, 50 and 100, 60 and 100, 70 and 100 mol percent.

30 The activity of the SHC enzyme or enzyme variant is defined via the reaction rate ((amount of product/(amount of product + amount of remaining starting material)) x 100) in percent. Preferably, the bioconversion of compound of formula (II) into compound of formula (I) in the presence of an SHC enzyme or enzyme variant produces compound of formula (I) yield of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,

18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 given in mol percent and based on the mols of compound of formula (II) employed; especially preferably, the yield is between 5 and 100, 10 and 100, 20 and 100, 25 and 100, 30 and 100, 35 and 100, in particular between 40 and 100, 45 and 100, 50 and 100, 60 and 100, 70 and 100.

In a preferred embodiment of the invention, the yield and/or the reaction rate are determined over a defined time period of, for example, 4, 6, 8, 10, 12, 16, 20, 24, 36 48, or 72 hours, during which compound of formula (II) is converted into compound of formula (I) by a recombinant host cell comprising a nucleotide sequence encoding an SHC enzyme or enzyme variant, and which has produced the SHC enzyme or enzyme variant.

In a further embodiment, the reaction is carried out under precisely defined conditions of, for example, 25°C, 30°C, 40°C, 50°C or 60°C. In particular, the yield and/or the reaction rate are determined by carrying out the reaction of converting compound of formula (II) into compound of formula (I) by the SHC enzyme or enzyme variant according to the invention at a temperature range from about 35°C to about 55°C over a period of 24-72 hours.

In a further embodiment of the present invention, a recombinant host cell comprising a nucleotide sequence encoding an SHC enzyme variant is characterized in that it shows a 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, 14-, 15-, 16-, 17-, 18-, 19-, 20-, 21-, 22-, 23-, 24-, 25-, 26-, 27-, 28-, 29-, 30-, 31-, 32-, 33-, 34-, 35-, 36-, 37-, 38-, 39-, 40-, 41-, 42-, 43-, 44-, 45-, 46-, 47-, 48-, 49-, 50-, 51-, 52-, 53-, 54-, 55-, 56-, 57-, 58-, 59-, 60-, 61-, 62-, 63-, 64-, 65-, 66-, 67-, 68-, 69-, 70-, 71-, 72-, 73-, 74-, 75-, 76-, 77-, 78-, 79-, 80-, 81-, 82-, 83-, 84-, 85-, 86-, 87-, 88-, 89-, 90-, 91-, 92-, 93-, 94-, 95-, 96-, 97-, 98-, 99-, 100-, 200-, 500-, 1000-fold or higher yield and/or reaction rates in the reaction of compound of formula (II) to give compound of formula (I) in comparison with the parent wild-type or variant SHC enzyme under the same conditions, preferably under conditions that have been individually defined as being optimal for the activity of the SHC enzyme considered. Here, the term condition relates but are not limited to

reaction conditions such as e.g., pH, temperature and concentration of sulibilizing agent (e.g. SDS).

The successful development of a bioconversion process for making compound of
5 formula (I) from compound of formula (II) in a recombinant strain of *E. coli* comprising a nucleotide sequence encoding a wild-type/reference SHC or a SHC variant can offer a low cost and industrially economical process for compound of formula (I) production.

Throughout this specification and the claims which follow, unless the context requires
10 otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integer or step. The term "comprising" also means "including" as well as "consisting" e.g. a composition "comprising" X may consist exclusively of X or may include something additional e.g. X
15 + Y. It must be noted also that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. By way of example, a reference to "a gene" or "an enzyme" is a reference to "one or more genes" or "one or more enzymes".

20 It is to be understood that this disclosure is not limited to the particular methodology, protocols and reagents described herein as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present disclosure which will be limited only by the appended claims. Unless defined otherwise, all
25 technical and scientific terms used herein have the same meanings as commonly understood by the person skilled in the art. In accordance with the present disclosure there may be conventional molecular biology, microbiology, and recombinant DNA techniques employed which are within the skill of the art.

30 This disclosure is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The disclosure is capable of other embodiments and of being practiced or of being carried out in various ways. Also, the phraseology and terminology used herein is for the purpose of description and should not be regarded as limiting. Preferably, the

terms used herein are defined as described in "A multilingual glossary of biotechnological terms: (IUPAC Recommendations)", Leuenberger, H.G.W, Nagel, B. and Kolbl, H. eds. (1995), Helvetica Chimica Acta, CH-4010 Basel, Switzerland).

5 Several documents are cited throughout the text of this specification. Each of the documents cited herein (including all patents, patent applications, scientific publications, manufacturer's specifications, instructions, GenBank Accession Number sequence submissions etc.), whether supra or infra, is hereby incorporated by reference in its entirety.

10

The examples described herein are illustrative of the present disclosure and are not intended to be limitations thereon. Different embodiments of the present disclosure have been described according to the present disclosure. Many modifications and variations may be made to the techniques described and illustrated herein without 15 departing from the spirit and scope of the disclosure. Accordingly, it should be understood that the examples are illustrative only and are not limiting upon the scope of the disclosure.

EXAMPLES

20 **EXAMPLE 1 - Production of SHC enzyme**

SHC Plasmid Preparation:

The gene encoding a wild-type or variant squalene hopene cyclase (SHC) enzyme was inserted into plasmid pET-28a(+), where it is under the control of an IPTG inducible T7-promotor for protein production in *Escherichia coli*. The plasmid was transformed into 25 *E. coli* strain BL21(DE3) using a standard heat-shock transformation protocol.

Media Preparation:

The minimal medium chosen as default was prepared as follows for 350 ml culture : to 35 ml citric acid/phosphate stock (133 g/l KH_2PO_4, 40 g/l $(NH_4)_2HPO_4$, 17 g/l citric 30 acid.H_2O with pH adjusted to 6.3) was added 307 ml H_2O, the pH adjusted to 6.8 with 32% NaOH as required. After autoclaving 0.850 ml 50% $MgSO_4$, 0.035 ml trace elements solution (see below), 0.035 ml Thiamin solution and 7 ml 20% glucose were added.

Trace elements solution : 50 g/l $Na_2EDTA.2H_2O$, 20 g/l $FeSO_4.7H_2O$, 3 g/l H_3BO_3, 0.9 g/l $MnSO_4.2H_2O$, 1.1 g/l $CoCl_2$, 80 g/L $CuCl_2$, 240 g/l $NiSO_4.7H_2O$, 100 g/l KI, 1.4 g/l $(NH_4)_6Mo_7O_{24}.4H_2O$, 1 g/l $ZnSO_4.7H_2O$, in deionized water.

5 Thiamin solution : 2.25 g/l Thiamin.HCl in deionized water.
$MgSO_4$ solution : 50 % (w/v) $MgSO_4.7H_2O$ in deionized water.
SHC Enzyme or Enzyme Variant Production (Biocatalyst Production).

Small scale biocatalyst production (wild-type SHC or SHC variants)
10 350 ml culture (medium supplemented with 50 µg/ml kanamycin) were inoculated from a preculture of the *E. coli* strain BL21(DE3) containing the SHC production plasmid. Cells were grown to an optical density of approximately 0.5 (OD_{650nm}) at 37°C with constant agitation (250 rpm).

15 Protein production was then induced by the addition of IPTG to a concentration of 300 µM followed by incubation for a further 5-6 hours with constant shaking. The resulting biomass was finally collected by centrifugation and washed with e.g. 50 mM Tris-HCl buffer pH 7.5. The cells were stored as pellets at 4 °C or -20 °C until further use. In general 2.5 to 4 grams of cells (wet weight) were obtained from 1 liter of culture,
20 independently of the medium used.

Biocatalyst production in fermenters
Fermentations were prepared and run in 750 ml InforsHT reactors. To the fermentation vessel was added 168ml deionized water. The reaction vessel was equipped with all
25 required probes (pO_2, pH, sampling, antifoam), C + N feed and sodium hydroxide bottles and autoclaved. After autoclaving is added to the reactor

20 ml 10x phosphate/citric acid buffer
14 ml 50% glucose
0.53 ml $MgSO_4$ solution
30 2 ml $(NH_4)_2SO_4$ solution
0.020 ml trace elements solution
0.400 ml thiamine solution
0.200 ml kanamycin stock

The running parameters were set are as follows: pH = 6.95, pO_2 = 40 %, T = 30 °C, Stirring at 300 rpm. Cascade: rpm setpoint at 300, min 300, max 1000, flow (l/min) set point 0.1, min 0, max 0.6. Antifoam control: 1:9.

5 The fermenter was inoculated from a seed culture to an OD_{650nm} of 0.4-0.5. This seed culture was grown in LB medium (+ Kanamycin) at 37°C, 220 rpm for 8 h. The fermentation was run first in batch mode for 11.5 h, where after was started the C+ N feed with a feed solution (sterilized glucose solution (143 ml H_2O+ 35 g glucose) to which had been added after sterilization : 17.5 ml $(NH_4)_2SO_4$ solution, 1.8 ml $MgSO_4$ 10 solution, 0.018 ml trace elements solution, 0.360 ml Thiamine solution, 0.180 ml kanamycin stock. The feed was run at a constant flow rate of approx. 4.2 ml/h. Glucose and NH_4^+ measurements were done externally to evaluate availability of the C- and N-sources in the culture. Usually glucose levels stay very low.

15 Cultures were grown for a total of approximately 25 hours, where they reached typically an OD_{650nm} of 40-45. SHC production was then started by adding IPTG to a concentration of approx. 1 mM in the fermenter (as IPTG pulse or over a period of 3-4 hours using an infusion syringe), setting the temperature to 40°C and pO_2 to 20%. Induction of SHC production lasted for 16 h at 40 °C. At the end of induction the cells 20 were collected by centrifugation, washed with 0.1 M citric acid/sodium citrate buffer pH 5.4 and stored as pellets at 4 °C or -20 °C until further use.

EXAMPLE 2 – GC analytics

Samples were extracted with an appropriate volume of tert-butylmethyl ether 25 (MBTE/tBME) for quantification of their content in substrate and reaction products. The solvent fraction was separated from the water phase by centrifugation prior to analysis with gas chromatography. 1μl of the solvent phase was injected (split ratio 10) onto a 30 m x 0.32 mm x 0.25 μm Zebron ZB-Wax column. The column was developed at constant flow (4 ml/min H_2) with the temperature gradient: 100 °C, 15 °C/min to 200 °C, 30 120 °C/min to 240 °C, 4 min at 240 °C. Inlet temperature: 250 °C, detector temperature: 250 °C. This resulted in separation of substrate and product peaks.

Hydroxyfarnesylacetone conversion was calculated from the areas of the peaks corresponding to substrate and reaction products with the following formula:

$$\text{Conversion (\%)} = 100 \times (\text{Area}_{\text{Product Peaks}}/(\text{Area}_{\text{Product Peaks}} + \text{Area}_{\text{Substrate Peak(s)}}))$$

<u>EXAMPLE 3 – Screening for Hydroxyfarnesylacetone cyclization with SHC</u>
5 <u>enzymes and enzyme variants</u>

A collection of SHC enzymes was produced in *E. coli* as outlined in Example 1 for use in Hydroxyfarnesylacetone cyclization reactions. The reactions contained 1 g/l Hydroxyfarnesylacetone and cells to an OD_{650nm} of 10. Reaction conditions were applied as listed in Table 2. An overview of the performance of the SHC enzymes
10 tested under their optimal reaction conditions as set out in Table 2 is shown in Figure 2.

Table 2. SHC enzymes and reaction conditions.

SHC enzyme	SEQ ID NO	Temperature (°C)	pH*	[SDS] (w/w %)
AacSHC	1	55	6.0	0.05
TelSHC	18	45	6.4	0.075
ApaSHC1	19	45	5.2	0.075
ZmoSHC1	15	40	5.2	0.005
ZmoSHC2	16	35	6.4	0.005
BjaSHC	17	50	5.8	0.005
BmeSHC	21	45	5.6	0.0025
GmoSHC	20	40	5.6	0.0075
215G2 SHC	13	35	5.4	0.07
115A7 SHC	11	35	5.4	0.05
110B8SHC	9	35	5.8	0.05
90C7 SHC	7	45	5.6	0.04
SHC #49	22	45	5.6	0.05
SHC #65	3	45	5.6	0.07
SHC #66	5	45	5.6	0.06

* in citric acid/sodium phosphate buffer

15 All enzymes tested were able to cyclize Hydroxyfarnesylacetone to (+)-Amberketal. Conversion with wild-type enzymes was between 2 and approx. 90 %, and highest with BmeSHC as outlined in Figure 2. Hydroxyfarnesylacetone cyclization was increased when mutations were introduced into wild-type AacSHC as observed with SHC variants

215G2 SHC, 115A7 SHC, 110B8 SHC, 90C7 SHC, SHC #49, SHC #65, SHC #66 as outlined in Figure 2.

EXAMPLE 4 - Hydroxyfarnesylacetone cyclization with SHC enzymes

5 With cells that had produced the ZmoSHC1 or the BmeSHC enzyme was tested hydroxyfarnesylacetone cyclization in reactions containing 2 and 8 g/l hydroxyfarnesylacetone, respectively. Reactions contained cells to an OD_{600nm} of 80. Reactions were run in 50 mM succinic acid/NaOH buffer pH 5.2 and incubated for 24 hours at 35 °C. GC-FID analysis of the solvent-extracted reactions indicated 52 % and 10 90 % hydroxyfarnesylacetone conversion, respectively with ZmoSHC1 and BmeSHC in reactions run at 2 g/l substrate. At 8 g/l hydroxyfarnesylacetone was conversion 17 % and 79 %, respectively with ZmoSHC1 and BmeSHC.

EXAMPLE 5 – Cyclization of Hydroxyfarnesylacetone with a Squalene Hopene 15 Cyclase variant enzyme

SHC variant 215G2 was produced as outlined in Example 1 and used in a cyclization reaction of Hydroxyfarnesylacetone.

A typical reaction (150 g total volume) was set up as follows in 0.75 liter Infors 20 fermenters. The reaction vessel was loaded with Hydroxyfarnesylacetone (0.75 g, 2.7 mmol), 1.95 g SDS was added from a 31 % (w/w) solution prepared in deionized water. A cell suspension was prepared from *E. coli* cells that had produced the SHC variant of interest by suspending the cells in 0.1 M succinic acid/NaOH buffer pH 5.1. After determination of the cell wet weight concentration of this cell suspension by 25 centrifugation for 10 min at 10 °C and 17210 *g*, the appropriate volume of cells was added to the reaction vessel in order to introduce 37.5 g of cells into the reaction. The volume of the reaction was completed to 150 g with the required amount of reaction buffer pH 5.1. The reaction was run at 35°C and pH 5.4 under constant stirring (700 rpm). pH was set to 5.4 using 85 % H_3PO_4. pH regulation was done manually using 85 30 % phosphoric acid as required. The reaction was sampled over time (1 ml), extracted with 5 volumes of MTBE/tBME (5 ml). The substrate and product content of the reaction was determined by GC analysis after clarification of the solvent phase by centrifugation (table top centrifuge, 13000 rpm, 2 min).

After approx. 72 hours of reaction was the reaction extracted 5 times with 100 ml MTBE by vigorous shaking followed by phase separation by centrifugation (6000 *g*, 10 min, room temperature). The crude extract was filtered over silica gel and the filtrate was concentrated. The residue (2.4 g) was purified by column chromatography on
5 silica gel, eluting with a gradient of 7-70% ethyl acetate in n-heptane. From this was isolated (+)-Amberketal (compound of formula (I), 173 mg, 23%, slightly yellow solid) and (*Z*)-5-((5aR,9aR)-6,6,9a-trimethyloctahydrobenzo[b]oxepin-3(2H)-ylidene)pentan-2-one (compound of formula (III), 65 mg, 9%, off-white solid).

10 Characterization of (+)-Amberketal. $[\alpha]_D^{23}$ = +16.4° (c = 0.42, $CHCl_3$). TLC (silica gel, heptane/EtOAc 3:2): R_f = 0.54.

^{1}H-NMR (400 MHz, $CDCl_3$) 4.32 (d, *J*=6.8 Hz, 1 H), 3.37 (dd, *J*=7.1, 1.2 Hz, 1 H), 1.90 (dt, *J*=12.7, 3.4 Hz, 1 H), 1.82 (dd, *J*=13.7, 4.6 Hz, 1 H), 1.49 - 1.78 (m, 8 H), 1.43 -
15 1.47 (m, 1 H), 1.42 (s, 3 H), 1.10 - 1.24 (m, 3 H), 0.99 (dd, *J*=12.2, 2.0 Hz, 1 H), 0.90 (s, 6 H), 0.83 - 0.95 (m, 1 H), 0.82 (s, 3 H). ^{13}C-NMR (101 MHz, $CDCl_3$): 106.0 (s), 82.6 (s), 73.5 (t), 55.7 (d), 53.3 (d), 41.8 (t), 38.7 (t), 37.3 (s), 36.2 (t), 35.9 (t), 33.6 (q), 33.1 (s), 24.3 (q), 21.7 (q), 20.0 (t), 18.3 (t), 17.4 (t), 14.6 (q).
EI-MS (70 eV): 278 (M+, <1), 263 (<1), 248 (2), 236 (4), 218 (36), 203 (19), 190 (55),
20 175 (43), 162 (11), 147 (24), 137 (34), 121 (42), 109 (47), 95 (34), 79 (38), 69 (36), 55 (42), 43 (100).

Characterization of (*Z*)-5-((5aR,9aR)-6,6,9a-trimethyloctahydrobenzo[b]oxepin-3(2H)-ylidene)pentan-2-one.
25 TLC (silica gel, heptane/EtOAc 3:2): R_f = 0.29.

^{1}H-NMR (600 MHz, BENZENE-d_6) δ ppm 5.01 (br t, *J*=7.3 Hz, 1 H), 4.44 (d, *J*=16.6 Hz, 1 H), 4.29 (br d, *J*=16.2 Hz, 1 H), 2.47 (ddd, *J*=12.4, 7.9, 3.8 Hz, 1 H), 2.22 (ddd, *J*=12.7, 8.6, 4.3 Hz, 1 H), 2.08 (q, *J*=7.2 Hz, 2 H),), 1.94 (t, *J*=1.0 Hz, 2 H), 1.73 - 1.62
30 (m, 3 H), 1.61 (s, 3 H), 1.49 - 1.22 (m, 4 H), 1.21-1.14 (m, 2H), 1.19 (s, 3 H), 0.87 (s, 3 H), 0.72 (s, 3 H). ^{13}C-NMR (151 MHz, BENZENE-d_6) δ ppm 205.59, 143.75, 120.82, 78.59, 62.81, 55.39, 42.97, 42.28, 41.65, 35.76, 35.38, 33.45, 29.42, 26.21, 21.56, 21.56, 20.89, 19.64.

EI-MS (70 eV): 278 (M+, <1), 260 (<1), 245 (<1), 220 (1), 141 (16), 135 (15), 123 (24), 109 (32), 95 (25), 81 (19), 69 (12), 55 (13), 43 (100).

EXAMPLE 5A – Cyclization of Hydroxyfarnesylacetone with wild type BmeSHC

Wild type BmeSHC was produced as outlined in Example 1 and used in cyclization reactions of Hydroxyfarnesylacetone.

The reactions (4 ml volume) contained 135 g/l Hydroxyfarnesylacetone, 221 g/l cells (wet cell weight), 0.09 % SDS, in 0.2 M acetic acid/sodium acetate buffer pH 5.2. The reactions were incubated at 45°C with constant agitation (650 rpm, Radleys Carousel). One reaction served as a control, was sampled over time, samples extracted with MTBE, the solvent phase clarified by centrifugation (table top centrifuge, 13000 rpm, 2 min). and analyzed by GC analysis for its substrate and product content. Hydroxyfarnesylacetone conversion was complete or almost complete (100 %) approx. 50 h after start, expecting approx. 920 mg of (+)-Amberketal.

The pooled reactions were centrifuged (4500 *g*, 4°C, 15 min). The pellet was recovered and washed 3 times with 20 ml deionized water (vigorous shaking + centrifugation), the aqueous phases discarded. The pellet was finally resuspended in 15 ml deionized water, and extracted 3 times with 15-20 ml MTBE (vigorous shaking + centrifugation). The organic phases were collected and analyzed for their (+)-amberketal content. The pooled organic phases were filtered (silica gel), the solvent evaporated under nitrogen flow yielding approx. 720 mg dry crystalline powder. The residue (720 mg) was purified by column chromatography on silica gel, eluting with a gradient of 6-50% MTBE in n-heptane. Fractions containing pure (+)-Amberketal were pooled and solvent evaporated. 560 mg (+)-Amberketal (compound of formula (I)) was isolated (approx. 61% yield).

Characterization of (+)-Amberketal. $[\alpha]_D^{25}$ = +22.6° (c = 0.94, $CHCl_3$). TLC (silica gel, n-heptane/MTBE 3:1): R_f = 0.55. NMR : the spectroscopic data was in agreement with the data shown in Example 5.

EXAMPLE 6 – Hydroxyfarnesylacetone (see also Figure 1)

1a) Preparation of (*E*)-6,10-dimethyl-1-((tetrahydro-2H-pyran-2-yl)oxy)undeca-5,9-dien-

2-one (4): To the solution of ethyl 4-(benzyloxy)-3-oxobutanoate (**1**, 60 g, 0.25 mol, 1 equiv) in ethanol (600 mL) in a 1L autoclave was added 10% palladium on carbon (6.0 g, 10%w/w) in portions under nitrogen and the mixture was stirred under 3 atm hydrogen for 12h. The reaction mixture was filtered through a celite bed and washed
5 with a mixture of dichloromethane/ethanol 1:1. The filtrate was evaporated under vacuum to give a pale yellow residue which was dissolved in dichloromethane (600 mL). To the solution was added 3,4-dihydro-2H-pyran (42.7 g, 0.51 mol, 2 equiv) and PPTS (6.3 g, 0.025 mol; 0.1equiv), and the mixture was stirred at room temperature for 16h, then water (500 mL) was added, followed by extraction with dichloromethane (2 x
10 200mL). The combined organic layer was washed with brine, dried over Na_2SO_4 and concentrated under vacuum to yield a crude product which was purified by column chromatography over silica gel eluting with ethyl acetate (5-8%) in petrol ether to afford ethyl (*E*)-5,9-dimethyl-2-(2-((tetrahydro-2H-pyran-2-yl)oxy)acetyl)deca-4,8-dienoate (**2**) as pale-yellow liquid (45 g, 77%). This product (44 g, 0.19 mol, 1.0 equiv) was
15 dissolved in THF (440 mL) and potassium carbonate (31.6 g, 0.23 mol, 1.2 equiv) was added. The suspension was stirred at room temperature for 1h, then (E)-1-bromo-3,7-dimethylocta-2,6-diene (geranyl bromide, 37.3 g, 0.172 mol, 0.9 equiv) was added at 0°C and the mixture was allowed to stir at room temperature for 16h. The mixture was filtered and the filter cake was washed with dichloromethane. The filtrate was
20 concentrated under vacuum to yield ethyl (E)-5,9-dimethyl-2-(2-((tetrahydro-2H-pyran-2-yl)oxy)acetyl)deca-4,8-dienoate (**3**, 70 g), which was dissolved in ethanol (500 mL). The solution of KOH (35 g) in water (100 mL) was added and the mixture was heated to 80°C during 2 h. The solution was concentrated under vacuum to yield a residue which was dissolved in dichloromethane (1L) and washed with water (2 x 200 mL) and
25 brine. The organic layer was dried over Na_2SO_4 and concentrated under vacuum to yield a crude oil which was purified by column chromatography over silica gel using 5-8% ethyl acetate in petrol ether to afford (*E*)-6,10-dimethyl-1-((tetrahydro-2H-pyran-2-yl)oxy)undeca-5,9-dien-2-one (**4**, 20 g, 36%) as a pale yellow liquid.

30 1b) Preparation of (3-(2-methyl-1,3-dioxolan-2-yl)propyl)triphenylphosphonium iodide (8): To a solution of 5-Iodopentan-2-one (45 g, 0.21 mol, 1 equiv) in toluene (200 mL) was added triphenyl phosphine (66.8 g, 0.25 mol, 1.2 equiv). The mixture was stirred for 16 h at 120°C, then cooled to room temperature upon which a solid precipitated which was filtered, washed with diethyl ether (100 mL) and dried under vacuum to

obtain (4-oxopentyl)triphenylphosphonium iodide (90 g, 90%) as a pale brown solid. This product (40 g, 0.10 mol, 1 equiv.) was suspended in toluene (400 mL) and ethylene glycol (80 mL) and p-toluene sulphonic acid (1 g, 0.0167 mol, 0.16 equiv) were added and the apparatus equipped with a Dean-Stark condenser. The mixture
5 was heated to 130°C for 12h, then cooled to room temperature and the toluene layer was decanted. The remaining rubbery residue was dissolved in dichloromethane (500 mL) and washed with water (500 mL) and brine (200 mL).
The organic layer was dried over Na_2SO_4 and concentrated under vacuum to yield (3-(2-methyl-1,3-dioxolan-2-yl)propyl)triphenylphosphonium iodide (**8**, 43 g, 80%) as a
10 pale brown solid.

1c) Preparation of Hydroxyfarnesyl acetone: A suspension of (3-(2-methyl-1,3-dioxolan-2-yl)propyl)triphenylphosphonium iodide (52.8 g, 0.10 mol, 2 equiv) in THF (150 mL) was cooled at -78°C before adding 1.6 M n-BuLi in Hexane (63.8 mL, 0.10
15 mol, 2 equiv). The mixture was stirred at room temperature for 30 min, upon which an orange yellow suspension was formed, which was cooled again to -78°C. A solution of (*E*)-6,10-dimethyl-1-((tetrahydro-2H-pyran-2-yl)oxy)undeca-5,9-dien-2-one (**4**, 15 g, 0.05 mol, 1 equiv) in THF (20 mL) was added dropwise and stirring was continued at room temperature for 2h. During this period, the reaction mixture turned into a yellow
20 suspension. The reaction mixture was quenched with ice water (100 mL) and extracted with ethyl acetate (2 x 100mL). The organic layer was dried over Na_2SO_4 and concentrated under vacuum to yield a crude which was triturated with hexane (4 x 50 mL) and the precipitated triphenyl phosphine oxide was removed by filtration. After removal of the solvent in vacuo 2-(((2*Z*,5*E*)-6,10-dimethyl-2-(3-(2-methyl-1,3-dioxolan-
25 2-yl)propylidene)undeca-5,9-dien-1-yl)oxy)tetrahydro-2H-pyran (**9**, 22 g) was obtained as a pale yellow liquid, which, according to 1H-NMR, contained still traces of triphenyl phosphine oxide. The product (22 g, 0.05 mol, 1 equiv.) was dissolved in acetone (200 mL) and 1.5 N aqueous HCl-solution (220 mL) was added at 0°C. The solution was stirred at room temperature for 16 h, then water (50 mL) was added and the mixture
30 was extracted with diethyl ether (2 x 100 mL). The combined organic layer was washed with brine, dried over Na_2SO_4 and concentrated under vacuum to afford a crude oil which was purified by column chromatography over silica gel eluting with 10-15% ethyl acetate in petrolether to yield (5*Z*,9*E*)-6-(hydroxymethyl)-10,14-dimethylpentadeca-5,9,13-trien-2-one (Hydroxyfarnesyl acetone, 7.5 g, 81% purity by GC/MS), from which

a volatile impurity was removed by distillation at 85°C/1mm/Hg. The residue (5.8 g) was submitted to a second column chromatography over silica gel eluting with 10-15% ethyl acetate in petrolether to yield Hydroxyfarnesyl acetone (2.8 g, 19%) as a pale yellow liquid.

5

^{13}C-NMR (d_6-DMSO, 100 MHz): 208.4 (s), 140.2 (s), 134.7 (s), 131.1 (s), 125.0 (d), 124.6 (d), 124.6 (d), 58.5 (t), 43.6 (t), 39.7 (2t), 35.0 (t), 30.1 (q), 26.7 (t), 25.9 (q), 21.9 (t), 18.0 (q), 16.2 (q).

SEQUENCE LISTING

SEQ ID NO: 1 (amino acid sequence of wild-type *Alicyclobacillus acidocaldarius* SHC (AacSHC))

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEY
PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIER
R

SEQ ID NO: 2 (nucleotide sequence encoding wild-type AacSHC)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGG
CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATAT
CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
ACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGC
ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCAT
CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC
GACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTG
ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACCGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
AGGTGA

SEQ ID NO: 3 (amino acid sequence of AacSHC enzyme variant #65)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG

VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTTWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYSHVFPTLALGRYKQAIER
R

SEQ ID NO: 4 (nucleotide sequence encoding SHC enzyme variant #65)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGTACCCGGGTGGCCGCCGGACCTCGACACGACCATCGAGG
CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATAT
CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
ACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGC
ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCAT
CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC
GACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTG
ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACAGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
AGGTGA

SEQ ID NO: 5 (amino acid sequence of AacSHC enzyme variant #66)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTTWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYSHVFPTLALGRYKQAIER
R

SEQ ID NO: 6 (nucleotide sequence encoding SHC enzyme variant #66)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGCACCCGGGTGGCCGCCGGACCTCGACACGACCATCGAGG
CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATAT

CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
ACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGC
ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCAT
CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC
GACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTG
ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACAGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
AGGTGA

SEQ ID NO: 7 (amino acid sequence of AacSHC enzyme variant #90C7)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALYPGGPPDLDATIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYSHVFPTLALGRYKQAIER
R

SEQ ID NO: 8 (nucleotide sequence encoding SHC variant #90C7)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACGCGACCATCGAGG
CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATAT
CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
ACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGC
ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCAT
CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC

GACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
5 CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTG
ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACAGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
10 AGGTGA

SEQ ID NO: 9 (amino acid sequence of AacSHC enzyme variant #110B8)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
15 PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
20 VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTTWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIER
R

SEQ ID NO: 10 (nucleotide sequence encoding SHC enzyme variant #110B8)

25 ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGCACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGG
CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
30 TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATAT
CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
ACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGC
35 ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCAT
CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
40 TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC
GACGTCGACAACACGAGCGATCTCCCGAACCTCACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
45 CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTG
ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
50 CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACCGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
AGGTGA

SEQ ID NO: 11 (amino acid sequence of AacSHC enzyme variant #115A7)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARTTVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDKTQH
5 PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIER
10 R

SEQ ID NO: 12 (nucleotide sequence encoding SHC variant #115A7)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
15 CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGG
CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATAT
CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
20 ACGAGTTTGGCTCGTGGGCTCGGACGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGC
ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACAAGACGCAGCAT
25 CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
30 GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC
GACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
35 CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTG
ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACCGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
40 AGGTGA

SEQ ID NO: 13 (amino acid sequence of AacSHC enzyme variant 215G2)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
45 PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
50 VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIER
R

SEQ ID NO: 14 (nucleotide sequence encoding Aac 215G2 SHC enzyme variant)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCT
CCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGT
CCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCAC
GAGCAGCGCGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGG
5 CGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCAT
TCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATAT
CCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCT
ACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTT
CCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCC
10 AAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGC
ACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAG
CTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCAT
CCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGT
TTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCC
15 GGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGAC
TGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGG
ACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCG
GGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTAC
GACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGC
20 CGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGT
CATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGC
GTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGC
CGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTG
CCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTG
25 ATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGA
CGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCT
CGGCTACACCATGTACCGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGC
AGGTGA

30 SEQ ID NO: 15 (amino acid sequence of wild-type ZmoSHC1)

MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKTMGNSLNNTLVSA
CDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLRPRLGNALLEMQREDGSWGVYFGAGNGD
INATVEAYAALRSLGYSADNPVLKKAAAWIAEKGGLKNIRVFTRYWLALIGEWPWEKTPNLPPEIIWFPD
NFVFSIYNFAQWARATMVPIAILSARRPSRPLRPQDRLDELFPEGRARFDYELPKKEGIDLWSQFFRTTD
35 RGLHWVQSNLLKRNSLREAAIRHVLEWIIRHQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSA
LDDPGWRHDRGESSWIQATNSPVWDTMLALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDWSIKLPD
VEPGGWAFEYANDRYPDTDDTAVALIALSSYRDKEEWQKKGVEDAITRGVNWLIAMQSECGGWGAFDKDN
NRSILSKIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDMPVIQKAIDYVRSEQEAEGAWFGRWGVNYI
YGTGAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYME

40

SEQ ID NO: 16 (amino acid sequence of wild-type ZmoSHC2)

MTVSTSSAFHHSPLSDDVEPIIQKATRALLEKQQQDGHWVFELEADATIPAEYILLKHYLGEPEDLEIEA
KIGRYLRRIQGEHGGWSLFYGGDLDLSATVKAYFALKMIGDSPDAPHMLRARNEILARGGAMRANVFTRI
QLALFGAMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLLVLQALKPVARNRRGILVDELFVPD
45 VLPTLQESGDPIWRRFFSALDKVLHKVEPYWPKNMRAKAIHSCVHFVTERLNGEDGLGAIYPAIANSVMM
YDALGYPENHPERAIARRAVEKLMVLDGTEDQGDKEVYCQPCLSPIWDTALVAHAMLEVGGDEAEKSAIS
ALSWLKPQQILDVKGDWAWRRPDLRPGGWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSDLHDDFEESKAR
AMEWTIGMQSDNGGWGAFDANNSYTYLNNIPFADHGALLDPPTVDVSARCVSMMAQAGISITDPKMKAAV
DYLLKEQEEDGSWFGRWGVNYIYGTWSALCALNVAALPHDHLAVQKAVAWLKTIQNEDGGWGENCDSYAL
50 DYSGYEPMDSTASQTAWALLGLMAVGEANSEAVTKGINWLAQNQDEEGLWKEDYYSGGGFPRVFYLRYHG
YSKYFPLWALARYRNLKKANQPIVHYGM

SEQ ID NO: 17 (amino acid sequence of wild-type BjaSHC)

MTVTSSASARATRDPGNYQTALQSTVRAAADWLIANQKPDGHWVGRAESNACMEAQWCLALWFMGLEDHP
LRKRLGQSLLDSQRPDGAWQVYFGAPNGDINATVEAYAALRSLGFRDDEPAVRRAREWIEAKGGLRNIRV
FTRYWLALIGEWPWEKTPNIPPEVIWFPLWFPFSIYNFAQWARATLMPIAVLSARRPSRPLPPENRLDAL
FPHGRKAFDYELPVKAGAGGWDRFFRGADKVLHKLQNLGNRLNLGLFRPAATSRVLEWMIRHQDFDGAWG
GIQPPWIYGLMALYAEGYPLNHPVLAKGLDALNDPGWRVDVGDATYIQATNSPVWDTILTLLAFDDAGVL
GDYPEAVDKAVDWVLQRQVRVPGDWSMKLPHVKPGGWAFEYANNYYPDTDDTAVALIALAPLRHDPKWKA
KGIDEAIQLGVDWLIGMQSQGGGWGAFDKDNNQKILTKIPFCDYGEALDPPSVDVTAHIIEAFGKLGISR
NHPSMVQALDYIRREQEPSGPWFGRWGVNYVYGTGAVLPALAAIGEDMTQPYIGRACDWLVAHQQADGGW
GESCASYMDVSAVGRGTTTASQTAWALMALLAANRPQDKDAIERGCMWLVERQSAGTWDEPEFTGTGFPG
YGVGQTIKLNDPALSQRLMQGPELSRAFMLRYGMYRHYFPLMALGRALRPQSHS

SEQ ID NO: 18 (amino acid sequence of wild-type TelSHC)

MPTSLATAIDPKQLQQAIRASQDFLFSQQYAEGYWWAELESNVTMTAEVILLHKIWGTEQRLPLAKAEQY
LRNHQRDHGGWELFYGDGGDLSTSVEAYMGLRLLGVPETDPALVKARQFILARGGISKTRIFTKLHLALI
GCYDWRGIPSLPPWIMLLPEGSPFTIYEMSSWARSSTVPLLIVMDRKPVYGMDPPITLDELYSEGRANVV
WELPRQGDWRDVFIGLDRVFKLFETLNIHPLREQGLKAAEEWVLERQEASGDWGGIIPAMLNSLLALRAL
DYAVDDPIVQRGMAAVDRFAIETETEYRVQPCVSPVWDTALVMRAMVDSGVAPDHPALVKAGEWLLSKQI
LDYGDWHIKNKKGRPGGWAFEFENRFYPDVDDTAVVVMALHAVTLPNENLKRRAIERAVAWIASMQCRPG
GWAAFDVDNDQDWLNGIPYGDLKAMIDPNTADVTARVLEMVGRCQLAFDRVALDRALAYLRNEQEPEGCW
FGRWGVNYLYGTSGVLTALSLVAPRYDRWRIRRAAEWLMQCQNADGGWGETCWSYHDPSLKGKGDSTASQ
TAWAIIGLLAAGDATGDYATEAIERGIAYLLETQRPDGTWHEDYFTGTGFPCHFYLKYHYYQQHFPLTAL
GRYARWRNLLAT

SEQ ID NO: 19 (amino acid sequence of wild-type ApaSHC1

MNMASRFSLKKILRSGSDTQGTNVNTLIQSGTSDIVRQKPAPQEPADLSALKAMGNSLTHTLSSACEWLM
KQQKPDGHWVGSVGSNASMEAEWCLALWFLGLEDHPLRPRLGKALLEMQRPDGSWGTYYGAGSGDINATV
ESYAALRSLGYAEDDPAVSKAAAWIISKGGLKNVRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFS
IYNFAQWARATMMPLAILSARRPSRPLRPQDRLDALFPGGRANFDYELPTKEGRDVIADFFRLADKGLHW
LQSSFLKRAPSREAAIKYVLEWIIWHQDADGGWGGIQPPWVYGLMALHGEGYQFHHPVMAKALDALNDPG
WRHDKGDASWIQATNSPVWDTMLSLMALHDANAEERFTPEMDKALDWLLSRQVRVKGDWSVKLPNTEPGG
WAFEYANDRYPDTDDTAVALIAIASCRNRPEWQAKGVEEAIGRGVRWLVAMQSSCGGWGAFDKDNNKSIL
AKIPFCDFGEALDPPSVDVTAHVLEAFGLLGLPRDLPCIQRGLAYIRKEQDPTGPWFGRWGVNYLYGTGA
VLPALAALGEDMTQPYISKACDWLINCQQENGGWGESCASYMEVSSIGHGATTPSQTAWALMGLIAANRP
QDYEAIAKGCRYLIDLQEEDGSWNEEEFTGTGFPGYGVGQTIKLDDPAISKRLMQGAELSRAFMLRYDLY
RQLFPIIALSRASRLIKLGN

SEQ ID NO: 20 (amino acid sequence of wild-type GmoSHC)

MSPADISTKSSSFQRLDNMLPEAVSSACDWLIDQQKPDGHWVGPVESNACMEAQWCLALWFLGQEDHPLR
PRLAQALLEMQREDGSWGIYVGADHGDINTTVEAYAALRSMGYAADMPIMAKSAAWIQQKGGLRNVRVFT
RYWLALIGEWPWDKTPNLPPEIIWLPDNFIFSIYNFAQWARATMMPLTILSARRPSRPLLPENRLDGLFP
EGRENFDYELPVKGEEDLWGRFFRAADKGLHSLQSFPVRRFVPREAAIRHVIEWIIRHQDADGGWGGIQP
PWIYGLMALSVEGYPLHHPVLAKAMDALNDPGWRRDKGDASWIQATNSPVWDTMLAVLALHDAGAEDRYS
PQMDKAIGWLLDRQVRVKGDWSIKLPDTEPGGWAFEYANDKYPDTDDTAVALIALAGCRHRPEWRERDIE
GAISRGVNWLLAMQSSSGGWGAFDKDNNRSILTKIPFCDFGEALDPPSVDVTAHVLEAFGLLGISRNHPS
VQKALAYIRSEQERNGAWFGRWGVNYVYGTGAVLPALAAIGEDMTQPYIVRACDWLMSVQQENGGWGESC
ASYMDINAVGHGVATASQTAWALIGLLAAKRPKDREAIARGCQFLIERQEDGSWTEEEYTGTGFPGYGVG
QAIKLDDPSLPDRLLQGAELSRAFMLRYDLYRQYFPVMALSRARRMMKEDASAAA

SEQ ID NO: 21 (amino acid sequence of wild-type BmeSHC)

MIILLKEVQLEIQRRIAYLRPTQKNDGSFRYCFETGVMPDAFLIMLLRTFDLDKEVLIKQLTER
IVSLQNEDGLWTLFDDEEHNLSATIQAYTALLYSGYYQKNDRILRKAERYIIDSGGISRAHFLT

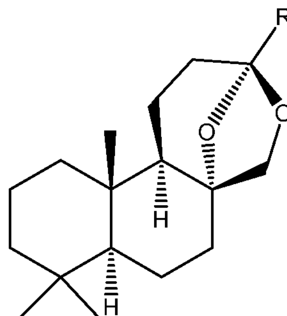
RWMLSVNGLYEWPKLFYLPLSLLLVPTYVPLNFYELSTYARIHFVPMMVAGNKKFSLTSRHTPS
LSHLDVREQKQESEETTQESRASIFLVDHLKQLASLPSYIHKLGYQAAERYMLERIEKDGTLYS
YATSTFFMIYGLLALGYKKDSFVIQKAIDGICSLLSTCSGHVHVENSTSTVWDTALLSYALQEA
GVPQQDPMIKGTTRYLKKRQHTKLGDWQFHNPNTAPGGWGFSDINTNNPDLDDTSAAIRALSRR
5 AQTDTDYLESWQRGINWLLSMQNKDGGFAAFEKNTDSILFTYLPLENAKDAATDPATADLTGRV
LECLGNFAGMNKSHPSIKAAVKWLFDHQLDNGSWYGRWGVCYIYGTWAAITGLRAVGVSASDPR
IIKAINWLKSIQQEDGGFGESCYSASLKKYVPLSFSTPSQTAWALDALMTICPLKDQSVEKGIK
FLLNPNLTEQQTHYPTGIGLPGQFYIQYHSYNDIFPLLALAHYAKKHSS

10 SEQ ID NO: 22 (amino acid sequence of AacSHC enzyme variant #49)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLH
EQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEY
PWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGA
KGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH
15 PAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGD
WAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
DVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWG
VNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTTWAL
MALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIER
20 R

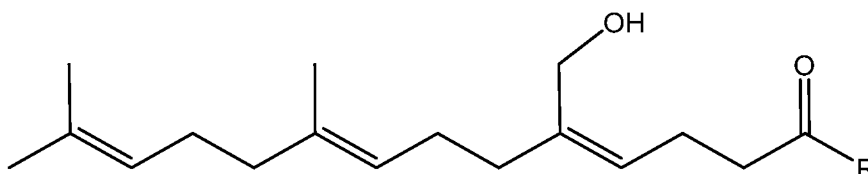
CLAIMS

1. A method for making a compound of formula (I),



5 Formula (I)

wherein the method comprises contacting a compound of formula (II) with a squalene-hopene cyclase (SHC) enzyme or enzyme variant,

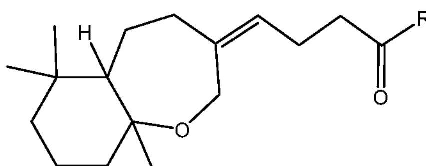


10 Formula (II)

wherein R is H, methyl, or ethyl.

2. The method of claim 1, wherein the method comprises contacting a compound of formula (II) wherein the double between C-8 and C-9 is in *E*-configuration and the
15 double bond between C-4 and C-5 is in *Z*-configuration with the SHC enzyme or enzyme variant.

3. The method of any preceding claim, wherein a compound of formula (III) is made as a by-product,

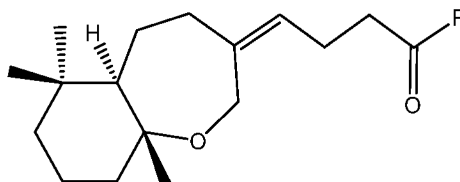


20

Formula (III)

wherein R is H, methyl or ethyl.

4. The method of claim 3, wherein a compound having the relative configuration shown in formula (IIIa) is made as a by-product,



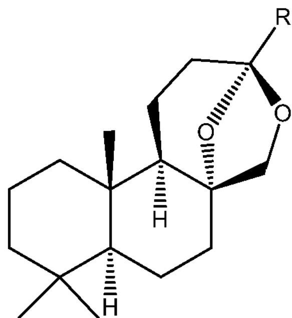
Formula (IIIa)

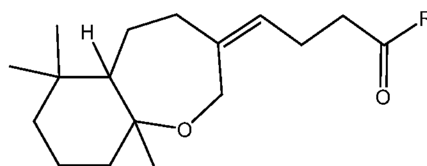
wherein R is H, methyl or ethyl.

5. The method of any preceding claim, wherein R is methyl.

6. The method of any preceding claim, wherein the method further comprises purification of the compound of formula (I).

7. A composition comprising a compound of formula (I) and a compound of formula (III):





Formula (I) Formula (III)

wherein R is H, methyl, or ethyl.

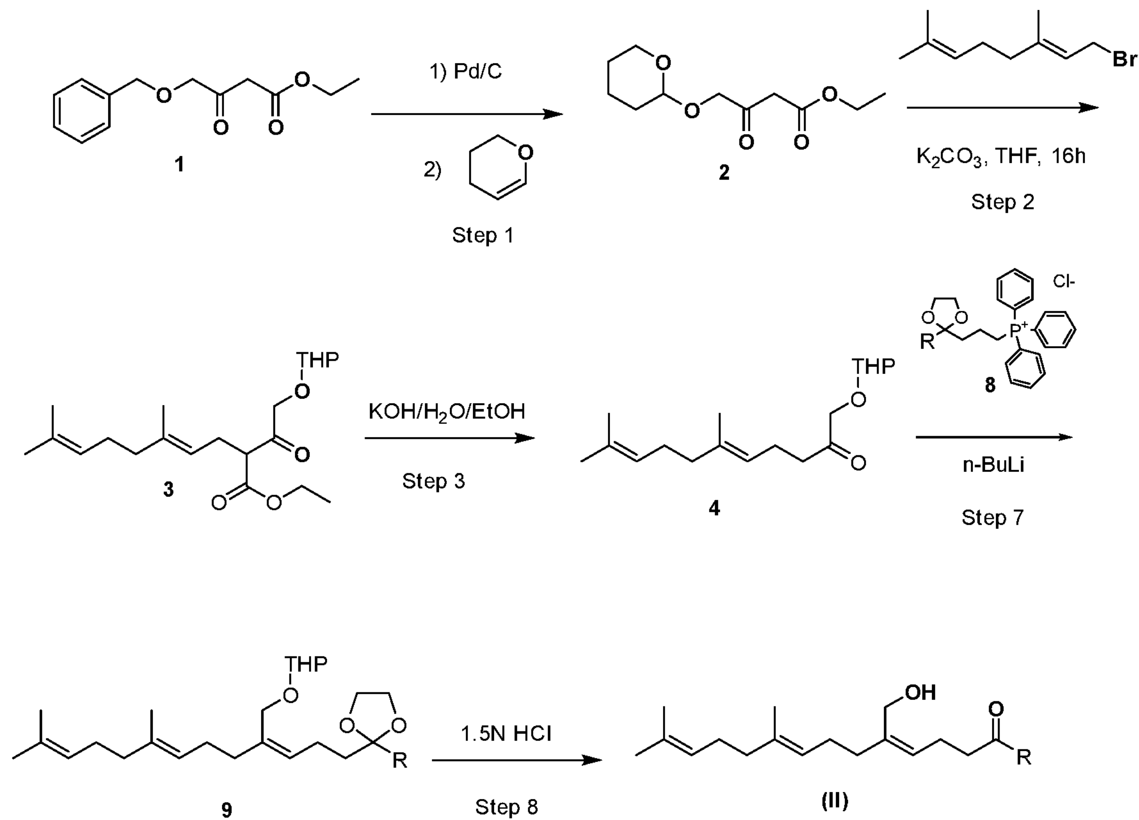
8. The composition of claims 7, wherein R is methyl.

9. A composition obtained by or obtainable by the method of any one of claims 1 to 6, for example wherein the composition is according to any of claims 7 to 8.

10. Use of a composition of any of claims 7 to 8 as or in a fragrance composition.

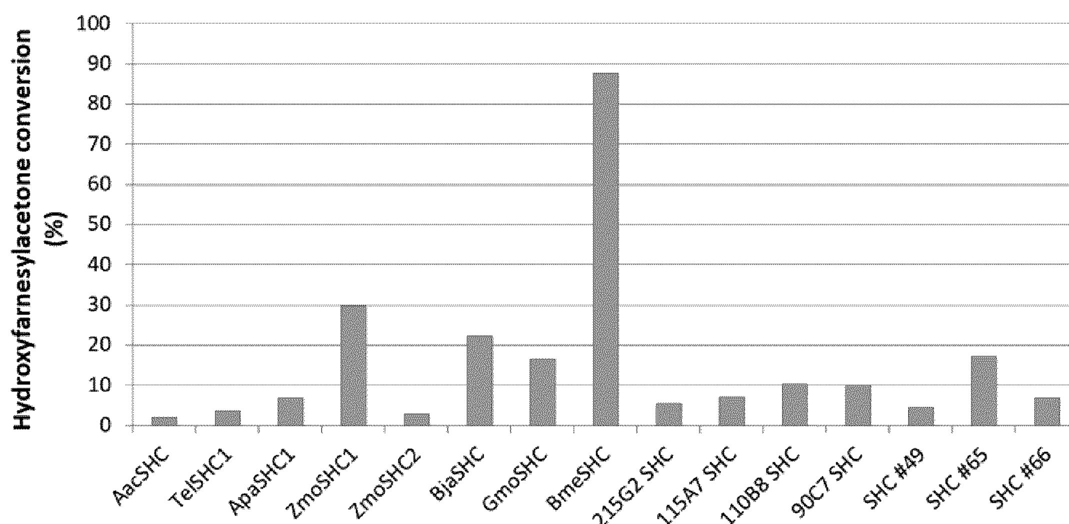
11. A consumer product comprising a composition of any of claims 7 to 8.

Figure 1



5

Figure 2



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2021/059618

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12P17/18 C12P17/02 C07D313/08 C07D493/08 C11B9/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12P C07D A23L C11B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data, BIOSIS, COMPENDEX, EMBASE, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 3 144 465 A (LEOPOLD RUZICKA ET AL) 11 August 1964 (1964-08-11)	9
A	claims 7-9; examples 9, 10; compounds XI, XII	7,8,10, 11

X	KONGKATHIP BOONSONG ET AL: "Stereospecific Total Synthesis of Amberketal and a Homologue", CHEMISTRY LETTERS, vol. 28, no. 1, 1 January 1999 (1999-01-01), pages 51-52, XP055816201, JP ISSN: 0366-7022, DOI: 10.1246/cl.1999.51 Retrieved from the Internet: URL:http://dx.doi.org/10.1246/cl.1999.51>	9
A	abstract; compound 1a scheme 2	7,8,10, 11

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 June 2021	30/06/2021

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Schröder, Gunnar

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2010/139719 A2 (BASF SE [DE]; BREUER MICHAEL [DE] ET AL.) 9 December 2010 (2010-12-09) cited in the application abstract; claim 1	1-6
A	WO 2017/140909 A1 (BASF SE [DE]) 24 August 2017 (2017-08-24) abstract; claims 1, 2	1-6
A	WO 2016/050690 A1 (BASF SE [DE]) 7 April 2016 (2016-04-07) abstract; claims 1, 4	1-6
X	CAMBIE RICHARD C ET AL: "Chemistry of the Podocarpaceae. LVII. The preparation of some 1,3-dioxans with ambergris-type odors", AUSTRALIAN JOURNAL OF CHEMISTRY, C S I R O PUBLISHING, AU, vol. 34, no. 6, 1 January 1981 (1981-01-01), pages 1265-1284, XP009121980, ISSN: 0004-9425	9
A	page 1266; compounds 1, 2 page 1274, last paragraph - page 1275, paragraph 1	7,8,10, 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2021/059618

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed:

      [X] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:

      [ ] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2021/059618

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 3144465 A	11-08-1964	NONE	
WO 2010139719 A2	09-12-2010	CN 102449158 A	09-05-2012
		EP 2438182 A2	11-04-2012
		EP 3404108 A1	21-11-2018
		ES 2703770 T3	12-03-2019
		JP 5735493 B2	17-06-2015
		JP 2012528578 A	15-11-2012
		US 2012135477 A1	31-05-2012
		WO 2010139719 A2	09-12-2010
WO 2017140909 A1	24-08-2017	BR 112018016789 A2	26-12-2018
		CN 109072265 A	21-12-2018
		EP 3417067 A1	26-12-2018
		EP 3816297 A1	05-05-2021
		JP 2019505222 A	28-02-2019
		US 2019144899 A1	16-05-2019
		WO 2017140909 A1	24-08-2017
WO 2016050690 A1	07-04-2016	CN 107002105 A	01-08-2017
		EP 3201184 A1	09-08-2017
		ES 2718753 T3	04-07-2019
		JP 6598852 B2	30-10-2019
		JP 2017529091 A	05-10-2017
		US 2017233780 A1	17-08-2017
		WO 2016050690 A1	07-04-2016