The present invention relates to processes and intermediates useful for the manufacture of spiroindolone compounds such as (R,3S)-5,7-dichloro-6-fluoro-3-methyl-2',3',4',9'-tetrahydrospiro[indoline-3,3'-pyrrolo[3,4-b]indol]-2-one and salts and hydrates and solvates thereof.
CHEMICAL PROCESS FOR PREPARING
SPIROINDOLONES AND INTERMEDIATES THEREOF

REFERENCE TO SEQUENCE LISTING, TABLE OR COMPUTER PROGRAM

The official copy of the Sequence Listing is submitted concurrently with the specification as an ASCII formatted text file via EFS-Web, with a file name of "PAT055051_seql2.txt", a creation date of March 22, 2013, and a size of 447 kilobytes. The Sequence Listing filed via EFS-Web is part of the specification and is incorporated in its entirety by reference herein.

Prior art:

(1'R,3'S)-5, 7'-dichloro-6'-fluoro-3'-methyl-2', 3',4',9'-tetrahydrospiro[indoline-3, 1'-pyrido[3,4-b]indol]-2-one (e.g. a compound of formula (IV), which comprises a spiroindolone moiety) and a 6-steps synthetic method for preparing, including known chiral amine intermediate compound (MA) are known (WO 2009/132921):

Invention:

The present invention is directed to an improved method of synthesizing spiroindolone compounds, in particular, (1'R,3'S)-5, 7'-dichloro-6'-fluoro-3'-methyl-2',3',4',9'-tetrahydrospiro[indoline-3, 1'-pyrido[3,4-b]indol]-2-one, and intermediates used in the improved method.
In a first embodiment, the invention is a process for preparing a compound of formula (II), or a salt or solvate or hydrate thereof,

\[
\text{(II)}
\]

comprising converting a compound of formula (I) to compound of formula (II), or a salt, solvate or hydrate thereof,

\[
\text{(I)}
\]

wherein: the dashed line is a bond or absent; A is selected from C=0 and C=NH; or when the dashed line is a double bond A-R^8 is:
R^4 and R^7 are each, independently, H or -Cl; R^5 is H, -OH, -CH_3, -OCH_3, -F, -Cl, -CF_3 or -CN; R^6 is H, -OH, -OCH_3, -F or -Cl; R^8 is H, -CH_3, -CH_2CH_3, -CH_2OH, -CO_2H, -CO_2CH_3, -C_6H_4CH_2CH_3 or -CF_3; and n is 1 or 2.

In a second embodiment, the invention is a process for preparing a compound of formula (IIA) or a salt or hydrate or solvate thereof,

comprising enzymatically transaminating a compound of formula (IA) or a salt or solvate or hydrate thereof,
to provide the compound of formula (IIA).

In a third embodiment, the invention is a process for preparing a compound of formula (IV), or a salt or hydrate or solvate thereof,

comprising, reacting a compound of formula (III), or a salt or hydrate or solvate thereof,

with a compound of formula (II), or a salt or hydrate or solvate thereof,
wherein the compound of formula (II), or a salt or hydrate or solvate thereof, is prepared from a compound of formula (I), or a salt or hydrate or solvate thereof,

wherein: the dashed line is a bond or absent; A is selected from C=0 and C=NH; or when the dashed line is a double bond A-R^8 is:

wherein: A is selected from C=0 and C=NH; or when the dashed line is a double bond A-R^8 is:

In a fourth embodiment, the invention is a process for preparing a compound of formula (IVA), or a salt or solvate or hydrate thereof,

comprising: reacting a compound of formula (IIIA) or a salt or solvate or hydrate thereof,
with a compound of formula (II) or a salt or solvate or hydrate thereof,

\[
\text{(IIA)}
\]

to provide the compound of formula (IVA), or a salt or solvate or hydrate thereof,

wherein the compound of formula (IIA) is prepared from the compound of formula (IA), or a salt or hydrate or solvate thereof,

\[
\text{(IV)}
\]

In a fifth embodiment, the invention is a compound of formula (IC):

\[
\text{(IC)}
\]

wherein: \(R^1\) is \((C_{1-6})\)alkyi, optionally substituted with an amino, \((\text{C}-C_{6})\)alkyi amino, \((C_{1-6})\)alkyi di-alkyi amino or \((C_{1-6})\)alkyi C(0)NH \((C_{1-6})\)alkyi; \(R^5\) and \(R^6\) are each, independently hydrogen, halo, hydroxyl. \((C_{1-6})\)alkyi, trihalo \((C_1)\)alkyi, cyano or \((C_{1-6})\)alkyi.
alkoxy; R is (C,-C₃) alkyl, optionally substituted with a hydroxyl; n is 1 or 2; or a pharmaceutically acceptable salt or hydrate or solvate thereof.

In a sixth embodiment the invention is a compound selected from:

or

, or a pharmaceutically acceptable salt, solvate or hydrate thereof.

Field of the invention

The invention relates to novel processes a novel process step and a novel intermediate useful for the preparation of spiroindolone compounds useful for the treatment of parasitic diseases comprising e.g. a spiroindolone moiety, such as (1′R,3′S)-5, 7′-dichloro-6′fluoro-3′-methyl-2′3′4′9′-tetrahydrospiro[indoline-3, 1′-pyrido[3,4-b]indol]-2-one.

Background of the invention

The present invention relates to processes for the preparation of spiroindolone compounds, such as (1′R,3′S)-5, 7′-dichloro-6′fluoro-3′-methyl-2′,3′,4′,9′-tetrahydrospiro[indoline-3, 1′-pyrido[3,4-b]indol]-2-one. (1′R,3′S)-5, 7′-dichloro-6′fluoro-3′-methyl-2′, 3′,4′,9′-tetrahydrospiro[indoline-3, 1′-pyrido[3,4-b]indol]-2-one is useful in the treatment and/or prevention of infections such as those caused by Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae,
Plasmodium ovale, Trypanosoma cruzi and parasites of the Leishmania genus such as, for example, Leishmania donovani, and it has the following structure:

\[
\text{(IVA)}
\]

\((1'R,3'S)-5,7'-\text{dichloro-6'-fluoro-3'-methyl-2'; 3',4',9'-tetrahydrospiro[indoline-3, 1'-pyrido}[3, 4-b]indol]-2\text{-one)}\]

and a synthesis thereof are described in WO 2009/132921 A1 in particular in Example 49 therein.

There is a need to provide new process for the preparation of \((1'R,3'S)-5,7'-\text{dichloro-6'-fluoro-3'-methyl-2'; 3',4',9'-tetrahydrospiro[indoline-3, 1'-pyrido}[3, 4-b]indol]-2\text{-one)}\) in order to improve the overall efficiency of the synthesis to make it suitable for manufacturing. In particular, there is a need to increase the efficiency of synthesizing the chiral amine intermediate (MA):

\[
\text{(IIA)}
\]

Detailed Description of the Invention

The process(es), according to the present invention, for producing spiroindolone compounds, such as compounds according to formula (IV), or salt or hydrate or solvate thereof, and intermediates, as defined herein, are summarized in Scheme 1.
Namely, a compound of formula (I), or salt or hydrate or solvate thereof, is converted into a compound of formula (II), or salt or hydrate or solvate thereof, according to methods 1, 2, 3, 4, 5 or 6 wherein

- method 1 comprises

  a) Enzymatic transamination to convert a compound of formula (I), or salt or hydrate or solvate thereof, into a compound of formula (II), or salt or hydrate or solvate thereof

- method 2 comprises;

  a) Chemical asymmetric catalysis to convert a compound of formula (I), or salt or hydrate or solvate thereof, into a compound of formula (II), or salt thereof;
- method 3 comprises

a) Chemical asymmetric reduction to convert a compound of formula (I), or salt or hydrate or solvate thereof, into a compound of formula (II), or salt or hydrate or solvate thereof;
method 4 comprises

a) Reduction followed by chiral resolution to convert a compound of formula (I), or salt or hydrate or solvate thereof, into a compound of formula (II), or salt or hydrate or solvate thereof;

method 5 comprises

- method 5 comprises
a) Lipase resolution to convert a racemate of a compound of formula (II), or salt or hydrate or solvate thereof, into a single enantiomer compound of formula (II), or salt or hydrate or solvate thereof:

\[
\begin{align*}
\text{Racemate} & \\
\text{ee}>99 & \\
\text{ee}=85 & 
\end{align*}
\]

Novozym 435: *Candida antarctica* Lipase B immobilized on an acrylic resin
- method 6 comprises
  
  a) A combination of two or more of methods 1-5 to convert a compound of formula (I), or salt or hydrate or solvate thereof, into a compound of formula (II), or salt or hydrate or solvate thereof.

A compound of formula (II), or salt thereof, may be converted into a compound formula (IV), or salt thereof, for example, as described in WO 2009/132921 in particular as described in the relevant claims and examples, which are incorporated by reference herein.
The invention specially relates to the processes described in each section. The invention likewise relates, independently, to every single step described in a process sequence within the corresponding section. Therefore, each and every single step of any process, consisting of a sequence of steps, described herein is itself a preferred embodiment of the present invention. Thus, the invention also relates to those embodiments of the process, according to which a compound obtainable as an intermediate in any step of the process is used as a starting material.

The invention likewise relates to novel starting materials which have been specifically developed for the preparation of the compounds according to the invention, to their use and to processes for their preparation.

The invention also relates to intermediates which have been specifically developed for the preparation of the compounds according to the invention, to their use and to processes for their preparation.

It is noted that in the present application usually explanations made in one section are also applicable for other sections, unless otherwise stated. For example, the explanations for the residue R₁ in formula (I) given in section A also apply if formula (I) occurs in other sections, such as Section B, unless otherwise stated.

**Section A: Preparation of a compound of formula (I)**

A compound of formula (I), or salt thereof, or hydrate or solvate thereof, may be prepared as described below and/or according to Examples 1-3 contained herein.
Section B: Conversion of a compound of formula (I) into a compound of formula (II).
In a first embodiment, the invention is a process for preparing a compound of formula (II), or a salt or solvate or hydrate thereof,

\[ \text{NH}_2 \]
\[ R^4 \]
\[ (\text{CH}_2)_n \]
\[ R^5 \]
\[ R^6 \]
\[ R^7 \]
\[ R^1 \]
\[ \text{(II)} \]

comprising converting a compound of formula (I) to compound of formula (II), or a salt, solvate or hydrate thereof,

\[ \text{A} \]
\[ R^8 \]
\[ \text{(I)} \]

wherein: the dashed line is a bond or absent; A is selected from C=0 and C=NH; or when the dashed line is a double bond A-R^8 is:
R\textsuperscript{4} and R\textsuperscript{7} are each, independently, H or -CI; R\textsuperscript{5} is H, -OH, -CH\textsubscript{3}, -OCH\textsubscript{3}, -F, -Cl, -CF\textsubscript{3} or -CN; R\textsuperscript{6} is H, -OH, -OCH\textsubscript{3}, -F or -Cl; R\textsuperscript{8} is H, -CH\textsubscript{3}, -CH\textsubscript{2}CH\textsubscript{3}, -CH\textsubscript{2}OH, -C\textsubscript{6}H\textsubscript{4}H, -C\textsubscript{6}H\textsubscript{5}CH\textsubscript{3}, -C\textsubscript{6}H\textsubscript{4}COCH\textsubscript{3} or -CF\textsubscript{3}; and n is 1 or 2.

In a first alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, R\textsuperscript{5} and R\textsuperscript{6} are fluoro when: R\textsuperscript{8} is -CH\textsubscript{3}, and n is 1.

In a second alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, R\textsuperscript{5} and R\textsuperscript{6} are fluoro and chloro, when: R\textsuperscript{8} is -CH\textsubscript{3}, and n is 1.

In a third alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, R\textsuperscript{5} and R\textsuperscript{6} are hydrogen when: R\textsuperscript{3} is -CH\textsubscript{3} and n is 1.

In a fourth alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, R\textsuperscript{5} is fluoro when: n is 1, and R\textsuperscript{6} is hydrogen.
In a sixth alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, the compound of formula (II) is of formula (HA), or a salt or solvate or hydrate thereof,

\[
\text{(IIA).}
\]

In a seventh alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, the compound of formula (II) is converted from the compound of formula (I) under a condition selected from enzymatic transamination, chemical asymmetric catalysis, asymmetric reduction and chiral resolution, or a combination of two or more conditions.

In an eighth alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, A is C=O.

In a ninth alternative embodiment, the invention is a process for converting a compound of formula (I) into a compound of formula (II), wherein, the compound of formula (I) is a compound of formula (IA) or a salt or hydrate or solvate thereof,

\[
\text{(IA).}
\]

In an exemplary embodiment, the enzyme is SEQ ID NO: 134.

In a second embodiment, the invention is a process for preparing a compound of formula (HA) or a salt or hydrate or solvate thereof,
comprising enzymatically transaminating a compound of formula (IA) or a salt or solvate or hydrate thereof,

<table>
<thead>
<tr>
<th>F</th>
<th>N</th>
<th>H</th>
</tr>
</thead>
</table>
| Cl | (S) | NH$_2$

(HA)

to provide the compound of formula (HA).

Typically, the ketone, compound (I) is dissolved in an organic solvent, e.g. glycol, and added to an aqueous mixture of isopropyl amine HCL and pyridoxal phosphate, followed by the addition of TEA buffer. The pH is then adjusted to neutral with an appropriate base, e.g. NaOH followed by warming and addition of the transaminase. The reaction is allowed to stir at temperature for approximately 24 hrs. The solid chiral amine product (compound of formula (II))) is isolated by means known to one of skill in the art, and/or according to examples 10-12.

In an exemplary embodiment, the enzyme is SEQ ID NO: 134.

Section C: Conversion of a compound of formula (II) into a compound of formula (IV)

In a third embodiment, the invention is a process for preparing a compound of formula (IV), or a salt or hydrate or solvate thereof,
comprising, reacting a compound of formula (III), or a salt or hydrate or solvate thereof,

with a compound of formula (II), or a salt or hydrate or solvate thereof,

wherein the compound of formula (II), or a salt or hydrate or solvate thereof, is prepared from a compound of formula (I), or a salt or hydrate or solvate thereof,
wherein: the dashed line is a bond or absent; A is selected from C=0 and C=NH; or when the dashed line is a double bond A-R^8 is:

\[ \text{R}^8 \text{N} \ \text{R}^a \]

; R^a is C_{1-8} alkyl; R^1 is (d-d) alkyl, optionally substituted with an amino, (d-d) alkyl amino, (d-d) alkyl di-alkyl amino or (d-d) alkyl C(0)NH (d-d) alkyl; R^4 and R^7 are each, independently, H or halo; R^5 and R^6 are each, independently hydrogen, halo, hydroxyl, (d-d) alkyl, trihalo (d) alkyl, cyano or (d-d) alkoxy; R^8 is (d-d) alkyl, optionally substituted with a hydroxyl; n is 1 or 2, R^7 is hydrogen or (d-d)alkyl; and R^4, R^5, R^6 and R^7, are each, independently hydrogen, halo, hydroxy, amino, alkylamino, dialkylamino, (d-d)alkyl, and (C_{1-6})alkyloxy.

In a tenth alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein: R^5 and R^6 are fluoro when: R^8 is -CH_3, and n is 1.

In an eleventh alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein, R^5 and R^6 are fluoro and chloro when: R^8 is -CH_3, and n is 1.

In a twelve alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein, R^5 and R^6 are hydrogen when: R^8 is -CH_3 and n is 1.

In a fourteenth alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein, R^5 is fluoro when: n is 1, and R^6 is hydrogen.

R^1 is H, -CH_3, or
R⁴ and R⁷ are each, independently, H or -Cl; R⁵ is H, -OH, -CH₃, -OCH₃, -F, -Cl, -CF₃ or -CN; R⁶ is H, -OH, -OCH₃, -F or -Cl; R⁸ is H, -CH₃, -CH₂CH₃, -CH₂OH, -C0₂H, -C0₂CH₃, -C0₂CH₂CH₃ or -CF₃; and n is 1 or 2.

In a fifteenth alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein, the compound of formula (II) is of formula (IIA), or a salt or solvate or hydrate thereof,

\[
\begin{array}{c}
\text{F} \\
\text{Cl} \\
\text{(S)} \\
\text{NH₂} \\
\end{array}
\]  

(IIA).

In a sixteenth alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein, wherein R⁵ is a halo.

In a seventeenth alternative embodiment, the invention is a process for preparing a compound of formula (IV) wherein, R⁵ is a chloro, and R⁴, R⁵, R⁶ and R⁷ are each hydrogen.

In an exemplary embodiment, the enzyme is SEQ ID NO: 134.

In a fourth embodiment, the invention is a process for preparing a compound of formula (IVA), or a salt or solvate or hydrate thereof,
comprising: reacting a compound of formula (IMA) or a salt or solvate or hydrate thereof,

with a compound of formula (II) or a salt or solvate or hydrate thereof,

to provide the compound of formula (IVA), or a salt or solvate or hydrate thereof,

wherein the compound of formula (IIA) is prepared from the compound of formula (IA), or a salt or hydrate or solvate thereof,

In a nineteenth alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the compound of formula (IIA) is converted to a salt of formula (IIB):
prior to reaction with the compound of formula (IIIA).

In an exemplary embodiment, the enzyme is SEQ ID NO: 34.

In a twentieth alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the reaction is carried out under basic conditions.

In a twenty-first alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the reaction is carried out in the presence of triethyl amine.

In a twenty-second alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the compound of formula (IVA) is isolated as a salt of formula (IVB):

In a twenty-third alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the salt of formula (IVB) is converted to the compound of formula (IVA) in free base.

In a twenty-fourth alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the salt of formula (IVB) is converted to the compound of formula (IVA) in free base with sodium carbonate.

In a twenty-fifth alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the compound of formula (IVA) is a hydrate.
In a twenty-sixth alternative embodiment, the invention is a process for preparing a compound of formula (IVA) wherein, the compound of formula (IVA) is a $\frac{1}{2}$ hydrate.

In a twenty-seventh alternative embodiment, the invention is a process for preparing a compound of formula (IVA) $\frac{1}{2}$ hydrate wherein, the compound of formula (IVA) $\frac{1}{2}$ hydrate is milled after isolation.

Section D: Use of the Novel and Inventive Compounds of Formula (IC).

In a fifth embodiment, the invention is a compound of formula (IC):

wherein: $R^1$ is ($C_1$-$C_6$) alkyi, optionally substituted with an amino, ($C_1$-$C_6$) alkyi amino, ($C_1$-$C_6$) alkyi di-alkyl amino or ($C_1$-$Cs$) alkyi C(0)NH ($C_1$-$6$) alkyi; $R^5$ and $R^6$ are each, independently hydrogen, halo, hydroxyl, ($C_1$-$C_6$) alkyi, trihalo ($C_i$-$C_6$) alkyi, cyano or ($C_i$-$C_6$) alkoxy; $R^8$ is ($C_1$-$C_6$) alkyl, optionally substituted with a hydroxyl; $n$ is 1 or 2; or a pharmaceutically acceptable salt or hydrate or solvate thereof.

In a twenty-eighth alternative embodiment, the invention is a compound of formula (IC) wherein, $R^5$ and $R^6$ are fluoro when: $R^8$ is -CH$_3$, and $n$ is 1.

In a twenty-ninth alternative embodiment, the invention is a compound of formula (IC) wherein, $R^5$ and $R^6$ are fluoro and chloro, when: $R^8$ is -CH$_3$, and $n$ is 1.

In a thirtieth alternative embodiment, the invention is a compound of formula (IC) wherein, $R^5$ and $R^6$ are hydrogen when: $R^8$ is -CH$_3$ and $n$ is 1.

In a thirty-first alternative embodiment, the invention is a compound of formula (IC) wherein, $R^5$ is fluoro when: $n$ is 1, and $R^8$ is hydrogen.
In a thirty-second alternative embodiment, the invention is a compound of formula (IC)

\[
\begin{align*}
R^1 & \text{ is } H, -\text{CH}_3, \\
R^5 & \text{ is } H, -\text{OH}, -\text{CH}_3, -\text{OCH}_3, -\text{F}, -\text{CF}_3 \text{ or } -\text{CN}; \\
R^8 & \text{ is } H, -\text{OH}, -\text{OCH}_3, -\text{F} \text{ or } -\text{Cl}; \\
R^6 & \text{ is } H, -\text{CH}_3, -\text{CH}_2\text{CH}_3, -\text{CH}_2\text{OH}, -\text{C}_2\text{H}_5, -\text{C}_2\text{H}_5\text{CH}_3, -\text{C}_2\text{H}_5\text{CH}_2\text{CH}_3 \text{ or } -\text{CF}_3; \text{ and } n \text{ is } 1 \text{ or } 2.
\end{align*}
\]

In a thirty-third alternative embodiment, the invention is a compound of formula (I)

wherein, the compound is of formula (IA):

or a pharmaceutically acceptable salt or hydrate or solvate thereof.

In a sixth embodiment the invention is a compound selected from:
or a pharmaceutically acceptable salt, solvate or hydrate thereof.

Section I: General terms

Listed below are definitions of various terms used to describe the novel intermediates and synthetic steps of the present invention. These definitions, either by replacing one, more than one or all general expressions or symbols used in the present disclosure and thus yielding embodiments of the invention, in particular apply to the terms as they are used throughout the specification unless they are otherwise defined in specific instances either individually or as part of a larger group. Thus, the general definitions used above and below, unless defined differently, have the following meanings:

The term "Cl-C_{20}" defines a moiety with up to and including maximally 20, especially up to and including maximally 7 carbon atoms, said moiety being branched (one or more times) or straight-chained and bound via a terminal or a non-terminal carbon.

Alkyl being a radical or part of a radical is a straight or branched (one or, if desired and possible, more times) carbon chain, and is especially C_r-C_{s} alkyl, such as C^aC^alkyl, in particular branched C_r-C_{s} alkyl, such as isopropyl. The term "lower" or "C_{1}-C_{r}" defines a moiety with up to and including maximally 7, especially up to and including maximally 4, carbon atoms, said moiety being branched (one or more times) or straight-chained and
bound via a terminal or a non-terminal carbon. Lower or CVCy-alkyl, for example, is n-pentyl, n-hexyl or n-heptyl or preferably C4-C8-alkyl, especially as methyl, ethyl, n-propyl, iso-propyl, n-butyl, isobutyl, sec-butyl, tert-butyl, in particular methyl, ethyl, n-propyl, iso-propyl, n-butyl, isobutyl, sec-butyl, tert-butyl; preferably methyl.

Alkylamino and dialkylamino refer to alkyl-NH- and (alkyl)2N-, respectively, wherein alkyl may be linear or branched. The alkyl group for example comprises 1 to 7 and in particular 1 to 4 C atoms. Some examples are methylamino, dimethylamino, ethylamino, and diethylamino; preferably methyamino.

Halo or halogen is preferably fluoro, chloro, bromo or iodo, preferably fluoro or chloro; where halo is mentioned as a substituent, where possible, one or more (e.g. up to three or one) halogen atoms may be present, e.g. in halo-C4-C8-alkyl, such as trifluoromethyl, 2,2,2-trifluoroethyl or 2,2,2-trifluorooethyl.

Halo-C4-alkyl may be linear or branched and in particular comprises 1 to 4 C atoms, for example 1 or 2 C atoms. Examples are fluoromethyl, difluoromethyl, trifluoromethyl, chloromethyl, dichloromethyl, trichloromethyl, 2-chloroethyl and 2,2,2-trifluoroorothyl; preferably trifluoromethyl.

Alkoxy, being a radical or part of a radical, refers to alkyl-O-, wherein the term alkyl is as defined herein, and includes, for example, CrC3-alkoxy (-O-C1-C8-alkyl), preferably C4-C8-alkoxy (-O-C4-C8-alkyl). In particular, alkoxy includes, for example, methoxy, ethoxy, n-propyloxy, isopropyloxy, n-butyloxy, isobutyloxy, sec-butyloxy, tert-butyloxy, pentyloxy, hexyloxy and heptyloxy radicals; preferably methoxy.

The term "optically active base" describes, for example, chiral amines, preferably chiral tertiary amines, more preferably cinchona alkaloids, such as quinidine and quinine, most preferably modified cinchona alkaloids. Examples of such modified cinchona alkaloids are detailed, for example, in Tian, S.-K.; Chen, Y.; Hang, J.; Tang, L.; McDiad, P.; Deng, L. Acc. Chem. Res. 2004, 37, 621-631 and references cited therein.

The term "phase transfer catalyst" as used herein refers to a catalytic amount of a chemical agent that enhances the rate of a reaction between chemical species located in different phases (e.g. immiscible liquids or solid and liquid) by extracting one of the reactants, most commonly an anion, across the interface into the other phase. These catalysts include quaternary ammonium or phosphonium salts (e.g. tetraalkylammonium...
salts, wherein alkyl can be same or different), or agents that complex inorganic cations (e.g. crown ethers or other cryptands). The catalyst cation is not consumed in the reaction although an anion exchange does occur. In particular, suitable phase transfer catalysts to be used according to the present invention are quaternary ammonium salts, for example of the formula $R_mR_sR_iR_kNX$, wherein $R_mR_sR_iR_k$ are, either the same or different, alkyl, and X is halo (eg. chloride, bromide, iodide) or hydroxide, for example, tetra-n-butylammonium hydroxide.

A "heterogeneous" catalyst as used herein refers to a catalyst supported on a carrier, typically although not necessarily a substrate comprised of an inorganic material, for example, a porous material such as carbon, silicon and/or aluminum oxide.

A "homogeneous" catalyst as used herein refers to a catalyst that is not supported on a carrier.

The term "chiral" refers to molecules which have the property of non-superimposability on their mirror image partner, while the term "achiral" refers to molecules which are superimposable on their mirror image partner.

The term "catalyst" means any substance that affects the rate of a chemical reaction by lowering the activation energy for the chemical reaction.

The term "powder catalyst" means a catalyst with a water contain of from 0 to 30 mass%.

The term "substrate to catalyst ratio" (S/C) refers to the molar ratio of starting compounds, or salts thereof, to "transition metal catalyst".

The term "work-up" means the work of isolation and/or purification which is carried out once the reaction is finished.

As used herein, unless specified otherwise, the term "room temperature" or "ambient temperature" means a temperature of from 15 to 30°C, such as of from 20 to 30°C, such as of from 20 to 25°C.

The term "inert" as used throughout this application, means unreactive with any of the reactants, solvents, or other components of the reaction mixture. Such inert conditions are generally accomplished by using inert gas such as carbon dioxide, helium, nitrogen, argon, among other gases.

Bonds with the asterisk (*) denote point of binding to the rest of the molecule.
The compounds of the present invention can possess one or more asymmetric centers. The preferred absolute configurations are as indicated herein specifically. However, any possible pure enantiomer, pure diastereoisomer, or mixtures thereof, e.g., mixtures of enantiomers, such as racemates, are encompassed by the present invention.

In the formulae of the present application the term "\( \text{C}^{s} \text{p}^{3} \)" or "--|--" on a C-sp\(^{3}\) represents a covalent bond wherein the stereochemistry of the bond is not defined. This means that the term "\( \text{J} \rightarrow \text{J} \)" or "---" on a C-sp\(^{3}\) comprises an (S) configuration as well as an (R) configuration of the respective chiral centre. Furthermore, mixtures are also encompassed, e.g., mixtures of enantiomers, such as racemates, are encompassed by the present invention.

In the formulae of the present application the term "\( \text{C}^{s} \text{p}^{2} \)" or "\( \text{C}^{s} \text{p}^{3} \)" on a C-sp\(^{2}\) represents a covalent bond, wherein the stereochemistry or the geometry of the bond is not defined. This means that the term "\( \text{J} \rightarrow \text{J} \)" on a C-sp\(^{2}\) comprises a cis (Z) configuration as well as a trans (E) configuration of the respective double bond. Furthermore, mixtures are also encompassed, e.g., mixtures of double bond isomers are encompassed by the present invention.

In the formulae of the present application, the term "\( \text{C}^{s} \text{p}^{3} \)" indicates a Csp\(^{3}\)-Csp \(^{3}\) bond or a Csp\(^{2}\)-Csp\(^{2}\) bond.

The compounds of the present invention can possess one or more asymmetric centers. The preferred absolute configurations are as indicated herein specifically.

In the formulae of the present application the term "\( \text{C}^{s} \text{p}^{3} \)" indicates the absolute stereochemistry, either (R) or (S).

In the formulae of the present application, the term "\( \text{C}^{s} \text{p}^{3} \)" indicates the absolute stereochemistry, either (R) or (S).

The term "stereomeric purity" at a given percentage means that the designated stereoisomer predominates at that given percentage in a mixture of stereoisomers.

The term "stereoisomer" means one of the absolute configurations of a single organic molecule having at least one asymmetric carbon. Included within the definition of a stereoisomer are enantiomers and diastereomers.
The term "resolution" refers to the separation or concentration or depletion of one of the stereoisomers of a molecule.

Salts are especially pharmaceutically acceptable salts or generally salts of any of the intermediates mentioned herein, where salts are not excluded for chemical reasons the skilled person will readily understand. They can be formed where salt forming groups, such as basic or acidic groups, are present that can exist in dissociated form at least partially, e.g. in a pH range from 4 to 10 in aqueous solutions, or can be isolated especially in solid, especially crystalline, form.

Such salts are formed, for example, as acid addition salts, preferably with organic or inorganic acids, from compounds or any of the intermediates mentioned herein with a basic nitrogen atom (e.g. imino or amino), especially the pharmaceutically acceptable salts. Suitable inorganic acids are, for example, halogen acids, such as hydrochloric acid, sulfuric acid, or phosphoric acid. Suitable organic acids are, for example, carboxylic, phosphonic, sulfonic or sulfamic acids, for example acetic acid, propionic acid, lactic acid, fumaric acid, succinic acid, citric acid, amino acids, such as glutamic acid or aspartic acid, maleic acid, hydroxymaleic acid, methylmaleic acid, benzoic acid, methane- or ethane-sulfonic acid, ethane-1,2-disulfonic acid, benzenesulfonic acid, 2-naphthalenesulfonic acid, 1,5-naphthalene-disulfonic acid, N-cyclohexylsulfamic acid, N-methyl-, N-ethyl- or N-propyl-sulfamic acid, or other organic protonic acids, such as ascorbic acid.

In the presence of negatively charged radicals, such as carboxy or sulfo, salts may also be formed with bases, e.g. metal or ammonium salts, such as alkali metal or alkaline earth metal salts, for example sodium, potassium, magnesium or calcium salts, or ammonium salts with ammonia or suitable organic amines for example triethylamine or tri(2-hydroxyethyl)amine, N-ethyl-piperidine, N,N'-dimethylpiperazine, t-butylamine, n-butylamine, phenylethylamine, dicyclohexylamine or cyclohexylamine.

When a basic group and an acid group are present in the same molecule, any of the intermediates mentioned herein may also form internal salts.

For isolation or purification purposes of any of the intermediates mentioned herein it is also possible to use pharmaceutically unacceptable salts, for example picrates or perchlorates.
Preferred salts forms include, for example, acid addition salts. The compounds having at least one acid group (e.g., COOH or 5-tetrazolyl) can also form salts with bases. Suitable salts with bases are, e.g., metal salts, such as alkali metal or alkaline earth metal salts, e.g., sodium, potassium, calcium or magnesium salts, or salts with ammonia or an organic amine, such as morpholine, thiomorpholine, piperidine, pyrrolidine, a mono-, di- or tri-lower alkylamine, e.g., ethyl-, tert-butyl-, diethyl-, diisopropyl-, triethyl-, tributyl- or dimethylpropylamine, or a mono-, di- or trihydroxy lower alkylamine, e.g., mono-, di- or tri-ethanolamine. Corresponding internal salts may furthermore be formed. Salts which are unsuitable for pharmaceutical uses but which can be employed, e.g., for the isolation or purification of free compounds or their pharmaceutically acceptable salts, are also included. Most preferably the salt of formula (IV) is the camphorsulfonic acid salt.

In particular, the term "salt of a compound of formula (IV)" refers, for example, to an amine salt thereof, an alkali salt thereof or an earth alkali metal salt thereof (eg. sodium salt, potassium salt, calcium salt, magnesium salt, etc). In particular, the term "amine" in the expression "amine salt thereof," for example when referring to an amine salt of the compound of formula (IV), means tertiary amine of formula NR9R10R11, secondary amine of formula NHR9R10R or primary amine of formula NH2R9, wherein R9, R10 and R11 are, independently from one another, alkyl, aryl, cycloalkyl or heterocyclyl, as defined herein, preferably alkyl or cycloalkyl. The term "amine" is, for example, diphenylamine, diisopropylamine, dimethylamine, triethylamine, diisopropylethylamine, dicyclohexylamine, t-butylamine, n-butylamine or cyclohexylamine, in particular, t-butylamine, n-butylamine or cyclohexylamine, more preferably n-butylamine or cyclohexylamine.

As used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the context clearly indicates otherwise. Thus, for example, reference to "a polypeptide" includes more than one polypeptide.

Similarly, "comprise," "comprises," "comprising" "include," "includes," and "including" are interchangeable and not intended to be limiting.

It is to be further understood that where descriptions of various embodiments use the term "comprising," those skilled in the art would understand that in some specific instances, an embodiment can be alternatively described using language "consisting essentially of" or "consisting of."
It is to be understood that both the foregoing general description, including the drawings, and the following detailed description are exemplary and explanatory only and are not restrictive of this disclosure.

The section headings used herein are for organizational purposes only and not to be construed as limiting the subject matter described.

The abbreviations used for the genetically encoded amino acids are conventional and are as follows:

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Three-Letter</th>
<th>One-Letter Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
</tr>
<tr>
<td>Aspartate</td>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
</tr>
<tr>
<td>Glutamate</td>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gin</td>
<td>Q</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
</tr>
</tbody>
</table>

When the three-letter abbreviations are used, unless specifically preceded by an "L" or a "D" or clear from the context in which the abbreviation is used, the amino acid may be in either the L- or D-configuration about α-carbon (C_α). For example, whereas "Ala" designates alanine without specifying the configuration about the α-carbon, "D-Ala" and "L-Ala" designate D-alanine and L-alanine, respectively. When the one-letter
abbreviations are used, upper case letters designate amino acids in the L-configuration about the \( \alpha \)-carbon and lower case letters designate amino acids in the D-configuration about the \( \alpha \)-carbon. For example, "A" designates L-alanine and "a" designates D-alanine. When polypeptide sequences are presented as a string of one-letter or three-letter abbreviations (or mixtures thereof), the sequences are presented in the amino (N) to carboxy (C) direction in accordance with common convention.

The abbreviations used for the genetically encoding nucleosides are conventional and are as follows: adenosine (A); guanosine (G); cytidine (C); thymidine (T); and uridine (U). Unless specifically delineated, the abbreviated nucleotides may be either ribonucleosides or 2'-deoxyribonucleosides. The nucleosides may be specified as being either ribonucleosides or 2'-deoxyribonucleosides on an individual basis or on an aggregate basis. When nucleic acid sequences are presented as a string of one-letter abbreviations, the sequences are presented in the 5' to 3' direction in accordance with common convention, and the phosphates are not indicated.

In reference to the present disclosure, the technical and scientific terms used in the descriptions herein will have the meanings commonly understood by one of ordinary skill in the art, unless specifically defined otherwise. Accordingly, the following terms are intended to have the following meanings:

"Protein", "polypeptide," and "peptide" are used interchangeably herein to denote a polymer of at least two amino acids covalently linked by an amide bond, regardless of length or post-translational modification (e.g., glycosylation, phosphorylation, lipidation, myristilation, ubiquitination, etc.). Included within this definition are D- and L-amino acids, and mixtures of D- and L-amino acids.

"Polynucleotide" or "nucleic acid" refers to two or more nucleosides that are covalently linked together. The polynucleotide may be wholly comprised ribonucleosides (i.e., an RNA), wholly comprised of 2' deoxyribonucleotides (i.e., a DNA) or mixtures of ribo- and 2' deoxyribonucleosides. While the nucleosides will typically be linked together via standard phosphodiester linkages, the polynucleotides may include one or more non-standard linkages. The polynucleotide may be single-stranded or double-stranded, or may include both single-stranded regions and double-stranded regions. Moreover, while a polynucleotide will typically be composed of the naturally occurring encoding nucleobases (i.e., adenine, guanine, uracil, thymine and cytosine), it may include one or
more modified and/or synthetic nucleobases, such as, for example, inosine, xanthine, hypoxanthine, etc. Preferably, such modified or synthetic nucleobases will be encoding nucleobases.

"Aminotransferase" and "transaminase" are used interchangeably herein to refer to a polypeptide having an enzymatic capability of transferring an amino group (NH₂) from a primary amine to a carbonyl group (C=O) of an acceptor molecule. Transaminases as used herein include naturally occurring (wild type) transaminase as well as non-naturally occurring engineered polypeptides generated by human manipulation.

"Amino acceptor" and "amine acceptor," "keto substrate," "keto," and "ketone" are used interchangeably herein to refer to a carbonyl (keto, or ketone) compound which accepts an amino group from a donor amine. In some embodiments, amino acceptors are molecules of the following general formula,

\[
\begin{align*}
\text{amino acceptor} & = R^\alpha \text{O} \text{R}^\beta \\
\end{align*}
\]

in which each of \(R^\alpha\) and \(R^\beta\), when taken independently, is an alkyl, cycloalkyl, heterocycloalkyl, aryl, or heteroaryl, which can be unsubstituted or substituted with one or more enzymatically acceptable groups. \(R^\alpha\) may be the same or different from \(R^\beta\) in structure or chirality. In some embodiments, \(R^\alpha\) and \(R^\beta\), taken together, may form a ring that is unsubstituted, substituted, or fused to other rings. Amino acceptors include keto carboxylic acids and alkanones (ketones). Typical keto carboxylic acids are a-keto carboxylic acids such as glyoxalic acid, pyruvic acid, oxaloacetic acid, and the like, as well as salts of these acids. Amino acceptors also include substances which are converted to an amino acceptor by other enzymes or whole cell processes, such as fumaric acid (which can be converted to oxaloacetic acid), glucose (which can be converted to pyruvate), lactate, maleic acid, and others. Amino acceptors that can be used include, by way of example and not limitation, 3,4-dihyronaphthalen-1 (2H)-one, 1-phenylbutan-2-one, 3,3-dimethylbutan-2-one, octan-2-one, ethyl 3-oxobutanoate, 4-phenylbutan-2-one, 1-(4-bromophenyl)ethanone, 2-methyl-cyclohexamone, 7-methoxy-2-tetralone, 1-hydroxybutan-2-one, pyruvic acid, acetophenone, 3'-
hydroxyacetophenone, 2-methoxy-5-fluoroacetophenone, levulinic acid, 1-phenylpropan-1-one, 1-(4-bromophenyl)propan-1-one, 1-(4-nitrophenyl)propan-1-one, 1-phenylpropan-2-one, 2-oxo-3-methylbutanoic acid, 1-(3-trifluoromethylphenyl)propan-1-one, hydroxypropanone, methoxyoxypropanone, 1-phenylbutan-1-one, 1-(2,5-dimethoxy-4-methylphenyl)butan-2-one, 1-(4-hydroxyphenyl)butan-3-one, 2-acetylnaphthalene, phenylpyruvic acid, 2-ketoglutaric acid, and 2-ketosuccinic acid, including both (R) and (S) single isomers where possible.

"Amino donor" or "amine donor" refers to an amino compound which donates an amino group to the amino acceptor, thereby becoming a carbonyl species. In some embodiments, amino donors are molecules of the following general formula,

\[
\begin{align*}
\text{R}^3 & \quad \text{R}^6 \\
\text{R}^4 & \quad \text{NH}_2
\end{align*}
\]

amino donor

in which each of R³ and R⁶, when taken independently, is an alkyl, cycloalkyl, heterocycloalkyl, aryl, or heteroaryl, which is unsubstituted or substituted with one or more enzymatically non-inhibiting groups. R⁴ can be the same or different from R⁵ in structure or chirality. In some embodiments, R⁴ and R⁵, taken together, may form a ring that is unsubstituted, substituted, or fused to other rings. Typical amino donors that can be used include chiral and achiral amino acids, and chiral and achiral amines. Amino donors that can be used include, by way of example and not limitation, isopropylamine (also referred to as 2-aminopropane), α-phenethylamine (also termed 1-phenylethanolamine), and its enantiomers (S)-1-phenylethanolamine and (R)-1-phenylethanolamine, 2-amino-4-phenylbutane, glycine, L-glutamic acid, L-glutamate, monosodium glutamate, L-alanine, D-alanine, D,L-alanine, L-aspartic acid, L-lysine, D,L-ornithine, β-alanine, taurine, n-octylamine, cyclohexylamine, 1,4-butanediamine (also referred to as putrescine), 1,6-hexanediamine, 6-aminohexanoic acid, 4-aminobutyric acid, tyramine, and benzyl amine, 2-amino-2-butanone, 2-amino-1-butanol, 1-amino-1-phenylethane, 1-amino-1-(2-methoxy-5-fluorophenyl)ethane, 1-amino-1-phenylpropane, 1-amino-1-(4-hydroxyphenyl)propane, 1-amino-1-(4-bromophenyl)propane, 1-amino-1-(4-nitrophenyl)propane, 1-phenyl-2-amino propane, 1-(3-trifluoromethylphenyl)-2-aminopropane, 2-aminopropanol, 1-amino-1-phenylbutane, 1-phenyl-2-aminobutane, 1-
(2,5-dimethoxy-4-methylphenyl)-2-aminobutane, 1-phenyl-3-aminobutane, 1-(4-
hydroxyphenyl)-3-aminobutane, 1-amino-2-methylcyclopentane, 1-amino-3-
methylcyclopentane, 1-amino-2-methylcyclohexane, 1-amino-1-(2-naphthyl)ethane, 3-
methylcyclopentylamine, 2-methylcyclopentylamine, 2-ethylcyclopentylamine, 2-
methylcyclohexylamine, 3-methylcyclohexylamine, 1-aminotetralin, 2-aminotetralin, 2-
amino-5-methoxytetralin, and 1-aminoindan, including both (R) and (S) single isomers
where possible and including all possible salts of the amines.

"Chiral amine" refers to amines of general formula R1-CH(NH2)- R1 and is employed
herein in its broadest sense, including a wide variety of aliphatic and alicyclic
compounds of different, and mixed, functional types, characterized by the presence of a
primary amino group bound to a secondary carbon atom which, in addition to a hydrogen
atom, carries either (i) a divalent group forming a chiral cyclic structure, or (ii) two
substituents (other than hydrogen) differing from each other in structure or chirality.
Divalent groups forming a chiral cyclic structure include, for example, 2-methylbutane-
1,4-diyl, pentane-1,4-diyl, hexane-1,4-diyl, hexane-1,5-diyl, 2-methylpentane-1,5-diyl.
The two different substituents on the secondary carbon atom (R1 and R2 above) also can
vary widely and include alkyl, aralkyl, aryl, halo, hydroxy, lower alkyl, lower alkoxy, lower
alkylthio, cycloalkyl, carboxy, carbalkoxy, carbamoyl, mono- and di-(lower alkyl)
substituted carbamoyl, trifluoromethyl, phenyl, nitro, amino, mono- and di-(lower alkyl)
substituted amino, alkylsulfonyl, arylsulfonyl, alkylcarboxamido, arylcarboxamido, etc.,
as well as alkyl, aralkyl, or aryl substituted by the foregoing.

"Pyridoxal-phosphate, " "PLP," "pyridoxal-5'-phosphate, " "PYP," and "P5P" are used
interchangeably herein to refer to the compound that acts as a coenzyme in
transaminase reactions. In some embodiments, pyridoxal phosphate is defined by the
structure 1-(4'-formyl-3'-hydroxy-2'-methyl-5'-pyridyl)methoxyphosphonic acid, CAS
number [54-47-7]. Pyridoxal-5'-phosphate can be produced in vivo by phosphorylation
and oxidation of pyridoxol (also known as Vitamin B6). In transamination reactions using
transaminase enzymes, the amine group of the amino donor is transferred to the
coenzyme to produce a keto byproduct, while pyridoxal-5'-phosphate is converted to
pyridoxamine phosphate. Pyridoxal-5'-phosphate is regenerated by reaction with a
different keto compound (the amino acceptor). The transfer of the amine group from
pyridoxamine phosphate to the amino acceptor produces a chiral amine and regenerates
the coenzyme. In some embodiments, the pyridoxal-5'-phosphate can be replaced by
other members of the vitamin B₆ family, including pyridoxine (PN), pyridoxal (PL), pyridoxamine (PM), and their phosphorylated counterparts; pyridoxine phosphate (PNP), and pyridoxamine phosphate (PMP).

"Coding sequence" refers to that portion of a nucleic acid (e.g., a gene) that encodes an amino acid sequence of a protein.

"Naturally-occurring" or "wild-type" refers to the form found in nature. For example, a naturally occurring or wild-type polypeptide or polynucleotide sequence is a sequence present in an organism that can be isolated from a source in nature and which has not been intentionally modified by human manipulation.

"Recombinant" or "engineered" or "non-naturally occurring" when used with reference to, e.g., a cell, nucleic acid, or polypeptide, refers to a material, or a material corresponding to the natural or native form of the material, that has been modified in a manner that would not otherwise exist in nature, or is identical thereto but produced or derived from synthetic materials and/or by manipulation using recombinant techniques. Non-limiting examples include, among others, recombinant cells expressing genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level.

"Percentage of sequence identity" and "percentage homology" are used interchangeably herein to refer to comparisons among polynucleotides and polypeptides, and are determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence for optimal alignment of the two sequences. The percentage may be calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Alternatively, the percentage may be calculated by determining the number of positions at which either the identical nucleic acid base or amino acid residue occurs in both sequences or a nucleic acid base or amino acid residue is aligned with a gap to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Those of skill in the art appreciate
that there are many established algorithms available to align two sequences. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman, 1981, Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch, 1970, J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the GCG Wisconsin Software Package), or by visual inspection (see generally, Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)). Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., 1990, J. Mol. Biol. 215: 403-410 and Altschul et al., 1977, Nucleic Acids Res. 3389-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information website. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as, the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, 1989, Proc Natl Acad Sci USA 89:10915).
Exemplary determination of sequence alignment and % sequence identity can employ the BESTFIT or GAP programs in the GCG Wisconsin Software package (Accelrys, Madison WI), using default parameters provided.

"Reference sequence" refers to a defined sequence used as a basis for a sequence comparison. A reference sequence may be a subset of a larger sequence, for example, a segment of a full-length gene or polypeptide sequence. Generally, a reference sequence is at least 20 nucleotide or amino acid residues in length, at least 25 residues in length, at least 50 residues in length, or the full length of the nucleic acid or polypeptide. Since two polynucleotides or polypeptides may each (1) comprise a sequence (i.e., a portion of the complete sequence) that is similar between the two sequences, and (2) may further comprise a sequence that is divergent between the two sequences, sequence comparisons between two (or more) polynucleotides or polypeptide are typically performed by comparing sequences of the two polynucleotides or polypeptides over a "comparison window" to identify and compare local regions of sequence similarity. In some embodiments, a "reference sequence" can be based on a primary amino acid sequence, where the reference sequence is a sequence that can have one or more changes in the primary sequence. For instance, a "reference sequence based on SEQ ID NO:4 having at the residue corresponding to X14 a valine" or X14V refers to a reference sequence in which the corresponding residue at X14 in SEQ ID NO:4, which is a tyrosine, has been changed to valine.

"Comparison window" refers to a conceptual segment of at least about 20 contiguous nucleotide positions or amino acids residues wherein a sequence may be compared to a reference sequence of at least 20 contiguous nucleotides or amino acids and wherein the portion of the sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The comparison window can be longer than 20 contiguous residues, and includes, optionally 30, 40, 50, 100, or longer windows.

"Substantial identity" refers to a polynucleotide or polypeptide sequence that has at least 80 percent sequence identity, at least 85 percent identity and 89 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison window of at least 20 residue positions, frequently over a window of at least 30-50 residues, wherein the percentage of sequence identity is
calculated by comparing the reference sequence to a sequence that includes deletions or additions which total 20 percent or less of the reference sequence over the window of comparison. In specific embodiments applied to polypeptides, the term "substantial identity" means that two polypeptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, share at least 80 percent sequence identity, preferably at least 89 percent sequence identity, at least 95 percent sequence identity or more (e.g., 99 percent sequence identity). Preferably, residue positions which are not identical differ by conservative amino acid substitutions.

"Corresponding to", "reference to" or "relative to" when used in the context of the numbering of a given amino acid or polynucleotide sequence refers to the numbering of the residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to the reference sequence. In other words, the residue number or residue position of a given polymer is designated with respect to the reference sequence rather than by the actual numerical position of the residue within the given amino acid or polynucleotide sequence. For example, a given amino acid sequence, such as that of an engineered transaminase, can be aligned to a reference sequence by introducing gaps to optimize residue matches between the two sequences. In these cases, although the gaps are present, the numbering of the residue in the given amino acid or polynucleotide sequence is made with respect to the reference sequence to which it has been aligned.

"Amino acid difference" or "residue difference" refers to a change in the amino acid residue at a position of a polypeptide sequence relative to the amino acid residue at a corresponding position in a reference sequence. The positions of amino acid differences generally are referred to herein as "Xn," where n refers to the corresponding position in the reference sequence upon which the residue difference is based. For example, a "residue difference at position X14 as compared to SEQ ID NO: 4" refers to a change of the amino acid residue at the polypeptide position corresponding to position 14 of SEQ ID NO: 4. Thus, if the reference polypeptide of SEQ ID NO: 4 has a tyrosine at position 14, then a "residue difference at position X14 as compared to SEQ ID NO:4" an amino acid substitution of any residue other than tyrosine at the position of the polypeptide corresponding to position 14 of SEQ ID NO: 4. In most instances herein, the specific amino acid residue difference at a position is indicated as "XnY" where "Xn" specified the corresponding position as described above, and "Y" is the single letter identifier of the
amino acid found in the engineered polypeptide (i.e., the different residue than in the reference polypeptide). In some embodiments, there more than one amino acid can appear in a specified residue position, the alternative amino acids can be listed in the form XnYZ, where Y and Z represent alternate amino acid residues. In some instances (e.g., in Table 2A and 2B), the present disclosure also provides specific amino acid differences denoted by the conventional notation “AnB”, where A is the single letter identifier of the residue in the reference sequence, “n” is the number of the residue position in the reference sequence, and B is the single letter identifier of the residue substitution in the sequence of the engineered polypeptide. Furthermore, in some instances, a polypeptide of the present disclosure can include one or more amino acid residue differences relative to a reference sequence, which is indicated by a list of the specified positions where changes are made relative to the reference sequence. The present disclosure includes engineered polypeptide sequences comprising one or more amino acid differences that include either/or both conservative and non-conservative amino acid substitutions.

"Conservative amino acid substitution" refers to a substitution of a residue with a different residue having a similar side chain, and thus typically involves substitution of the amino acid in the polypeptide with amino acids within the same or similar defined class of amino acids. By way of example and not limitation, an amino acid with an aliphatic side chain may be substituted with another aliphatic amino acid, e.g., alanine, valine, leucine, and isoleucine; an amino acid with hydroxyl side chain is substituted with another amino acid with a hydroxyl side chain, e.g., serine and threonine; an amino acid having aromatic side chains is substituted with another amino acid having an aromatic side chain, e.g., phenylalanine, tyrosine, tryptophan, and histidine; an amino acid with a basic side chain is substituted with another amino acid with a basic side chain, e.g., lysine and arginine; an amino acid with an acidic side chain is substituted with another amino acid with an acidic side chain, e.g., aspartic acid or glutamic acid; and a hydrophobic or hydrophilic amino acid is replaced with another hydrophobic or hydrophilic amino acid, respectively. Exemplary conservative substitutions are provided in Table 1 below.

Table 1

<table>
<thead>
<tr>
<th>Residue</th>
<th>Possible Conservative Substitutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, L, V, I</td>
<td>Other aliphatic (A, L, V, I)</td>
</tr>
<tr>
<td></td>
<td>Other non-polar (A, L, V, I, G, M)</td>
</tr>
</tbody>
</table>
G, M Other non-polar (A, L, V, I, G, M)
D, E Other acidic (D, E)
K, R Other basic (K, R)
N, Q, S, T Other polar
H, Y, W, F Other aromatic (H, Y, W, F)
C, P None

"Non-conservative substitution" refers to substitution of an amino acid in the polypeptide with an amino acid with significantly differing side chain properties. Non-conservative substitutions may use amino acids between, rather than within, the defined groups and affects (a) the structure of the peptide backbone in the area of the substitution (e.g., proline for glycine), (b) the charge or hydrophobicity, or (c) the bulk of the side chain. By way of example and not limitation, an exemplary non-conservative substitution can be an acidic amino acid substituted with a basic or aliphatic amino acid; an aromatic amino acid substituted with a small amino acid; and a hydrophilic amino acid substituted with a hydrophobic amino acid.

"Deletion" refers to modification to the polypeptide by removal of one or more amino acids from the reference polypeptide. Deletions can comprise removal of 1 or more amino acids, 2 or more amino acids, 5 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, or up to 20% of the total number of amino acids making up the reference enzyme while retaining enzymatic activity and/or retaining the improved properties of an engineered transaminase enzyme. Deletions can be directed to the internal portions and/or terminal portions of the polypeptide. In various embodiments, the deletion can comprise a continuous segment or can be discontinuous.

"Insertion" refers to modification to the polypeptide by addition of one or more amino acids from the reference polypeptide. In some embodiments, the improved engineered transaminase enzymes comprise insertions of one or more amino acids to the naturally occurring transaminase polypeptide as well as insertions of one or more amino acids to other improved transaminase polypeptides. Insertions can be in the internal portions of the polypeptide, or to the carboxy or amino terminus. Insertions as used herein include fusion proteins as is known in the art. The insertion can be a contiguous segment of amino acids or separated by one or more of the amino acids in the naturally occurring polypeptide.
"Fragment" as used herein refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion, but where the remaining amino acid sequence is identical to the corresponding positions in the sequence. Fragments can be at least 14 amino acids long, at least 20 amino acids long, at least 50 amino acids long or longer, and up to 70%, 80%, 90%, 95%, 98%, and 99% of the full-length transaminase polypeptide, for example the polypeptide of SEQ ID NO:2 or engineered transaminase of SEQ ID NO:34.

"Isolated polypeptide" refers to a polypeptide which is substantially separated from other contaminants that naturally accompany it, e.g., protein, lipids, and polynucleotides. The term embraces polypeptides which have been removed or purified from their naturally-occurring environment or expression system (e.g., host cell or in vitro synthesis). The improved transaminase enzymes may be present within a cell, present in the cellular medium, or prepared in various forms, such as lysates or isolated preparations. As such, in some embodiments, the improved transaminase enzyme can be an isolated polypeptide.

"Substantially pure polypeptide" refers to a composition in which the polypeptide species is the predominant species present (i.e., on a molar or weight basis it is more abundant than any other individual macromolecular species in the composition), and is generally a substantially purified composition when the object species comprises at least about 50 percent of the macromolecular species present by mole or % weight. Generally, a substantially pure transaminase composition will comprise about 60 % or more, about 70% or more, about 80% or more, about 90% or more, about 95% or more, and about 98% or more of all macromolecular species by mole or % weight present in the composition. In some embodiments, the object species is purified to essential homogeneity (i.e., contaminant species cannot be detected in the composition by conventional detection methods) wherein the composition consists essentially of a single macromolecular species. Solvent species, small molecules (<500 Daltons), and elemental ion species are not considered macromolecular species. In some embodiments, the isolated improved transaminases polypeptide is a substantially pure polypeptide composition.

"Stereoselectivity" refers to the preferential formation in a chemical or enzymatic reaction of one stereoisomer over another. Stereoselectivity can be partial, where the formation of one stereoisomer is favored over the other, or it may be complete where only one stereoisomer is formed. When the stereoisomers are enantiomers, the
stereoselectivity is referred to as enantioselectivity, the fraction (typically reported as a percentage) of one enantiomer in the sum of both. It is commonly alternatively reported in the art (typically as a percentage) as the enantiomeric excess (e.e.) calculated therefrom according to the formula \([\text{major enantiomer} - \text{minor enantiomer}] / [\text{major enantiomer} + \text{minor enantiomer}]\). Where the stereoisomers are diastereoisomers, the stereoselectivity is referred to as diastereoselectivity, the fraction (typically reported as a percentage) of one diastereomer in a mixture of two diastereomers, commonly alternatively reported as the diastereomeric excess (d.e.). Enantiomeric excess and diastereomeric excess are types of stereomeric excess.

"Highly stereoselective" refers to a chemical or enzymatic reaction that is capable of converting a substrate, \(e.g.,\) Compound (IA), to its corresponding chiral amine product, \(e.g.,\) Compound (HA), with at least about 85% stereomeric excess.

"Improved enzyme property" refers to a transaminase polypeptide that exhibits an improvement in any enzyme property as compared to a reference transaminase. For the engineered transaminase polypeptides described herein, the comparison is generally made to the wild-type transaminase enzyme, although in some embodiments, the reference transaminase can be another improved engineered transaminase. Enzyme properties for which improvement is desirable include, but are not limited to, enzymatic activity (which can be expressed in terms of percent conversion of the substrate), thermo stability, solvent stability, pH activity profile, cofactor requirements, refractoriness to inhibitors (\(e.g.,\) substrate or product inhibition), stereospecificity, and stereoselectivity (including enantioselectivity).

"Increased enzymatic activity" refers to an improved property of the engineered transaminase polypeptides, which can be represented by an increase in specific activity (\(e.g.,\) product produced/time/weight protein) or an increase in percent conversion of the substrate to the product (\(e.g.,\) percent conversion of starting amount of substrate to product in a specified time period using a specified amount of transaminase) as compared to the reference transaminase enzyme. Exemplary methods to determine enzyme activity are provided in the Examples. Any property relating to enzyme activity may be affected, including the classical enzyme properties of \(K_m, V_{max}\), or \(k_{cat}\), changes of which can lead to increased enzymatic activity. Improvements in enzyme activity can be from about 1.2 times the enzymatic activity of the corresponding wild-type transaminase enzyme, to as much as 2 times, 5 times, 10 times, 20 times, 25 times, 50 times, 75
times, 100 times, or more enzymatic activity than the naturally occurring transaminase or another engineered transaminase from which the transaminase polypeptides were derived. Transaminase activity can be measured by any one of standard assays, such as by monitoring changes in spectrophotometric properties of reactants or products. In some embodiments, the amount of products produced can be measured by High-Performance Liquid Chromatography (HPLC) separation combined with UV absorbance or fluorescent detection following derivatization, such as with o-phthalaldehyde (OPA). Comparisons of enzyme activities are made using a defined preparation of enzyme, a defined assay under a set condition, and one or more defined substrates, as further described in detail herein. Generally, when lysates are compared, the numbers of cells and the amount of protein assayed are determined as well as use of identical expression systems and identical host cells to minimize variations in amount of enzyme produced by the host cells and present in the lysates.

"Conversion" refers to the enzymatic conversion of the substrate(s) to the corresponding product(s). "Percent conversion" refers to the percent of the substrate that is converted to the product within a period of time under specified conditions. Thus, the "enzymatic activity" or "activity" of a transaminase polypeptide can be expressed as "percent conversion" of the substrate to the product.

"Thermostable" refers to a transaminase polypeptide that maintains similar activity (more than 60% to 80% for example) after exposure to elevated temperatures (e.g., 40-80°C) for a period of time (e.g., 0.5-24 hrs) compared to the wild-type enzyme.

"Solvent stable" refers to a transaminase polypeptide that maintains similar activity (more than e.g., 60% to 80%) after exposure to varying concentrations (e.g., 5-99%) of solvent (ethanol, isopropyl alcohol, dimethylsulfoxide (DMSO), tetrahydrofuran, 2-methyltetrahydrofuran, acetone, toluene, butyl acetate, methyl tert-butyl ether, etc.) for a period of time (e.g., 0.5-24 hrs) compared to the wild-type enzyme.

"Thermo- and solvent stable" refers to a transaminase polypeptide that is both thermostable and solvent stable.

"Stringent hybridization" is used herein to refer to conditions under which nucleic acid hybrids are stable. As known to those of skill in the art, the stability of hybrids is reflected in the melting temperature ($T_m$) of the hybrids. In general, the stability of a hybrid is a function of ion strength, temperature, G/C content, and the presence of chaotropic...

"Hybridization stringency" relates to hybridization conditions, such as washing conditions, in the hybridization of nucleic acids. Generally, hybridization reactions are performed under conditions of lower stringency, followed by washes of varying but higher stringency. The term "moderately stringent hybridization" refers to conditions that permit target-DNA to bind a complementary nucleic acid that has about 60% identity, preferably about 75% identity, about 85% identity to the target DNA, with greater than about 90% identity to target-polynucleotide. Exemplary moderately stringent conditions are conditions equivalent to hybridization in 50% formamide, 5x Denhart's solution, 5xSSPE, 0.2% SDS at 42°C, followed by washing in 0.2xSSPE, 0.2% SDS, at 42°C. "High stringency hybridization" refers generally to conditions that are about 10°C or less from the thermal melting temperature $T_m$ as determined under the solution condition for a defined polynucleotide sequence. In some embodiments, a high stringency condition refers to conditions that permit hybridization of only those nucleic acid sequences that form stable hybrids in 0.018M NaCl at 65°C (i.e., if a hybrid is not stable in 0.018M NaCl at 65°C, it will not be stable under high stringency conditions, as contemplated herein). High stringency conditions can be provided, for example, by hybridization in conditions equivalent to 50% formamide, 5º Denhartz's solution, 5ºSSPE, 0.2% SDS at 42°C, followed by washing in 0.1xSSPE, and 0.1% SDS at 65°C. Another high stringency condition is hybridizing in conditions equivalent to hybridizing in 5X SSC containing 0.1% (w:v) SDS at 65°C and washing in 0.1x SSC containing 0.1% SDS at 65°C. Other high
stringency hybridization conditions, as well as moderately stringent conditions, are
described in the references cited above.

"Heterologous" polynucleotide refers to any polynucleotide that is introduced into a host
cell by laboratory techniques, and includes polynucleotides that are removed from a host
cell, subjected to laboratory manipulation, and then reintroduced into a host cell.

"Codon optimized" refers to changes in the codons of the polynucleotide encoding a
protein to those preferentially used in a particular organism such that the encoded
protein is efficiently expressed in the organism of interest. Although the genetic code is
degenerate in that most amino acids are represented by several codons, called
"synonyms" or "synonymous" codons, it is well known that codon usage by particular
organisms is nonrandom and biased towards particular codon triplets. This codon usage
bias may be higher in reference to a given gene, genes of common function or ancestral
origin, highly expressed proteins versus low copy number proteins, and the aggregate
protein coding regions of an organism's genome. In some embodiments, the
polynucleotides encoding the transaminase enzymes may be codon optimized for
optimal production from the host organism selected for expression.

"Preferred, optimal, high codon usage bias codons" refers interchangeably to codons
that are used at higher frequency in the protein coding regions than other codons that
code for the same amino acid. The preferred codons may be determined in relation to
codon usage in a single gene, a set of genes of common function or origin, highly
expressed genes, the codon frequency in the aggregate protein coding regions of the
whole organism, codon frequency in the aggregate protein coding regions of related
organisms, or combinations thereof. Codons whose frequency increases with the level of
gene expression are typically optimal codons for expression. A variety of methods are
known for determining the codon frequency (e.g., codon usage, relative synonymous
codon usage) and codon preference in specific organisms, including multivariate
analysis, for example, using cluster analysis or correspondence analysis, and the
effective number of codons used in a gene (see GCG CodonPreference, Genetics
Computer Group Wisconsin Package; CodonW, John Peden, University of Nottingham:
Res. 222437-46; Wright, F., 1990, Gene 87:23-29). Codon usage tables are available for
a growing list of organisms (see for example, Wada et al., 1992, Nucleic Acids Res.
20:21 11-21 18; Nakamura et al., 2000, Nucl. Acids Res. 28:292; Duret, et al., supra;

"Control sequence" is defined herein to include all components, which are necessary or advantageous for the expression of a polynucleotide and/or polypeptide of the present disclosure. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleic acid sequence encoding a polypeptide.

Operably linked" is defined herein as a configuration in which a control sequence is appropriately placed (i.e., in a functional relationship) at a position relative to a polynucleotide of interest such that the control sequence directs or regulates the expression of the polynucleotide and/or polypeptide of interest.

"Promoter sequence" refers to a nucleic acid sequence that is recognized by a host cell for expression of a polynucleotide of interest, such as a coding sequence. The promoter sequence contains transcriptional control sequences, which mediate the expression of a polynucleotide of interest. The promoter may be any nucleic acid sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

"Suitable reaction conditions" refer to those conditions in the biocatalytic reaction solution (e.g., ranges of enzyme loading, substrate loading, cofactor loading, temperature, pH, buffers, co-solvents, etc.) under which a transaminase polypeptide of
the present disclosure is capable of converting a substrate compound to a product compound (e.g., conversion of Compound (IA) to Compound (HA)). Exemplary "suitable reaction conditions" are provided in the present disclosure and illustrated by the Examples.

"Loading", such as in "compound loading" or "enzyme loading" or "cofactor loading" refers to the concentration or amount of a component in a reaction mixture at the start of the reaction.

"Substrate" in the context of a biocatalyst mediated process refers to the compound or molecule acted on by the biocatalyst. For example, an exemplary substrate for the transaminase biocatalyst in the process disclosed herein is compound (IA).

"Product" in the context of a biocatalyst mediated process refers to the compound or molecule resulting from the action of the biocatalyst. For example, an exemplary product for the transaminase biocatalyst in the process disclosed herein is compound (IIA).

The present disclosure provides methods of using polypeptides having transaminase activity for the synthesis of (S)-1-(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine in enantiomeric excess of (R)-1-(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine. Where the description relates to polypeptides, it is to be understood that it also describes the polynucleotides encoding the polypeptides.

Aminotransferases, also known as transaminases, catalyze the transfer of an amino group from a primary amine of an amino donor substrate to the carbonyl group (e.g., a keto or aldehyde group) of an amino acceptor molecule. Aminotransferases have been identified from various organisms, such as Alcaligenes denitrificans, Bordetella bronchiseptica, Bordetella parapertussis, Brucella melitensis, Burkholderia malle, Burkholderia pseudomallei, Chromobacterium violaceum, Oceanicola granulosus HTCC251 6, Oceanobacter sp. RED65, Oceanospirillum sp. MED92, Pseudomonas putida, Ralstonia solanacearum, Rhizobium meliloti, Rhizobium sp. (strain NGR234), Vibrio fluvialis, Bacillus thuringensis, and Klebsiella pneumoniae (Shin et al., 2001, Biosci. Biotechnol. Biochem. 65:1782-1 788).

Transaminases are useful for the chiral resolution of racemic amines by exploiting the ability of the transaminases to carry out the reaction in a stereospecific manner, i.e., preferential conversion of one enantiomer to the corresponding ketone, thereby resulting in a mixture enriched in the other enantiomer (see, e.g., Koselewski et al., 2009, Org
Lett. 11(21):4810-2). The stereoselectivity of transaminases in the conversion of a ketone to the corresponding amine also make these enzymes useful in the asymmetric synthesis of optically pure amines from the corresponding keto compounds (see, e.g., Hohne et al., "Biocatalytic Routes to Optically Active Amines," Chem Cat Chem 1(1):42-51; Zua and Hua, 2009, Biotechnol J. 4(10):1420-31).


Engineered transaminases derived from the transaminase of *Vibrio fluvialis* having increased resistance to aliphatic ketones are described in Yun et al., 2005, Appl Environ Microbiol., 71(8):4220-4224) while ω-VfTs with broadened amino donor substrate specificity are described in Cho et al., 2008, Biotechnol Bioeng. 99(2):275-84. Patent publications WO201 0081 053 and US201 00209981, incorporated by reference herein, describe engineered ω-VfTs having increased stability to temperature and/or organic solvent and enzymatic activity towards structurally different amino acceptor molecules.

The present disclosure relates to methods of using engineered transaminase polypeptides derived from *V. fluvialis* that efficiently mediate conversion of 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-one, to product compound (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine in enantiomeric excess. Significantly, the disclosure identifies amino acid residue positions and corresponding mutations in the transaminase polypeptide that increase the enzymatic activity, enantioselectivity, stability and refractoriness to the product inhibition in the conversion of 1-(1H-5-fluoro-6-chloro-indol-
Accordingly, in one aspect, the present disclosure relates to methods of using polypeptides that are capable of converting the substrate converting substrate compound (IA), 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-one, to product compound (IIA), (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine, 

![Scheme 1](image)

in the presence of an amino donor, where the (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine is produced in enantiomeric excess of compound (IIC), (R)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine.

In some embodiments, the polypeptides used in the methods are non-naturally occurring transaminases engineered for improved properties as compared to the wild-type V. fluvialis polypeptide of SEQ ID NO:2, or another engineered polypeptide, for example SEQ ID NO:4. These engineered transaminase polypeptides adapted for efficient conversion of conversion of 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-one to (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine have one or more residue differences as compared to the amino acid sequence of SEQ ID NO:2 or a reference engineered transaminase polypeptide, such as the reference polypeptide of SEQ ID NO:4. The residue differences are associated with enhancements in enzyme properties, including enzymatic activity, enzyme stability, and resistance to inhibition by the product amine.

In some embodiments, the engineered transaminase polypeptides used in the instant disclosure show increased activity in the conversion of the substrate 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-one to product (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine product in enantiomeric excess in a defined time with the same amount of
enzyme as compared to the wild type or a reference engineered enzyme. In some embodiments, the engineered transaminase polypeptide has at least about 1.2 fold, 1.5 fold, 2 fold, 3 fold, 4 fold, 5 fold, 10 fold, 20 fold, 30 fold, 40 fold, or 50 fold or more the activity as compared to the polypeptide represented by SEQ ID NO:4 under suitable reaction conditions.

In some embodiments, the engineered transaminase polypeptides used in the instant disclosure have increased stability to temperature and/or solvents used in the conversion reaction as compared to the wild type or a reference engineered enzyme. In some embodiments, the engineered transaminase polypeptide has at least 1.2 fold, 1.5 fold, 2 fold, 3 fold, 4 fold, 5 fold, 10 fold or more the stability as compared to the polypeptide of SEQ ID NO:4 under suitable reaction conditions.

In some embodiments, the engineered transaminase polypeptides used in the instant disclosure have increased refractoriness or resistance to product amine (S)-1 -(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine as compared to the wild type or a reference engineered enzyme. In some embodiments, the engineered transaminase polypeptide has at least 1.2 fold, 1.5 fold, 2 fold, 3 fold, 4 fold, 5 fold, or more resistant to inhibition by 1-(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine, in particular (S)-1 -(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine, as compared to the polypeptide represented by SEQ ID NO:4 under suitable reaction conditions, as further described below.

In some embodiments, the engineered transaminase polypeptides used in the instant disclosure are capable of converting the substrate 1-(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-one to product (S)-1 -(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine in enantiomeric excess of greater than 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5 or greater over (R)-1 -(1 /-5-fluoro-6-chloro-indol-3-yl)propan-2-under suitable reaction conditions.

In some embodiments, the engineered transaminase polypeptides used in the instant disclosure are capable of converting compound (IA) to compound (MA) with increased tolerance for the presence of substrate relative to the reference polypeptide of SEQ ID NO:4 under suitable reaction conditions. Thus, in some embodiments the engineered transaminase polypeptides are capable of converting the substrate compound (IIA) to product compound (IA) in the presence of a substrate loading concentration of at least about 1 g/L, about 5 g/L, about 10 g/L, about 20 g/L, about 30 g/L, about 40 g/L, about 50 g/L, about 70 g/L, about 100 g/L, about 125 g/L, about 150 g/L, about 175 g/L or
about 200 g/L or more with a percent conversion of at least about at least about 50%, at
least about 60%, at least about 70%, at least about 80%, at least about 90%, at least
about 95%, at least about 98%, or at least about 99%, in a reaction time of about 72 h or
less, about 48 h or less, about 36 h or less, or about 24 h less, under suitable reaction
conditions.

The suitable reaction conditions under which the above-described improved properties
of the engineered polypeptides carry out the conversion can be determined with respect
to concentrations or amounts of polypeptide, substrate, cofactor, buffer, co-solvent, pH,
and/or conditions including temperature and reaction time, as further described below
and in the Examples.

The exemplary engineered polypeptides used in the instant disclosure have associated
with their improved properties for conversion of compound (IA) to compound (MA) and
which include one or more residue differences as compared to SEQ ID NO:2 at the
following residue positions: X9; X14; X18; X21; X26; X31; X33; X41; X45; X47; X57;
X70; X86; X88; X107; X 113; X132; X133; X146; X147; X148; X153; X163; X168; X173;
X177; X203; X211; X233T; X235; X244; X250; X284; X294; X314; X315; X318; X323;
X324; X324; X346; X383; X391 ;X395; X398; X400; X417; X419; X420; X423; X424;
X427; X448; and X451. The specific amino acid differences at each of these positions
that are associated with the improved properties of the exemplary polypeptides of Table
2A and 2B include: X9T; X14V; X18A; X21H; X26R; X31M; X31S; X33T; X41 L; X45H;
X47N; X57F; X57Y; X70A; X86D; X86Y; X88A; X88L; X107P; X113L; X113V; X132F;
X133R; X146L; X147K; X148Q; X148R; X153S; X163F; X163I; X163L; X163R; X163V;
X168K; X168S; X173A; X177L; X203S; X211 K; X233T; X235P; X244T; X250A; X284A;
X294V; X314N; X315G; X318D; X323T; X324G; X346L; X383V; X391A; X395P;
X398L; X398V; X398W; X400G; X417M; X419S; X420N; X423I; X424V; X424A; X427Y;
X448E; and X451D.

The residue differences as compared to the engineered transaminase represented by
SEQ ID NO:4 includes those at residue positions: X14; X26; X31 ;X33; X41; X47; X57;
X70; X86; X88; X107; X 113; X132; X148; X163; X168; X173; X203; X250; X284; X314;
X315; X324; X346; X395; X398; X400; X417; X419; X420; X423; X424; X448; and X451 .
The specific amino acid differences at these positions include: X14V; X26R; X31S;
X33T; X41L; X47N; X57F; X57Y; X70A; X86D; X88A; X88L; X107P; X113L; X113V;
X132F; X148Q; X148R; X163I; X163L; X163R; X163V; X168K; X168S; X173A; X203S;
X250A; X284A; X314N; X315G; X324H; X346L; X395P; X398L; X398V; X398W; X400G; X417M; X419S; X420N; X423I; X424V; X448E; and X451D. Although residue differences compared to SEQ ID NO:4 also occur at residue positions X153 and X383, these differences represent reversions to the amino acid residue present on the wild type sequence of SEQ ID NO:2, indicating that interconversions between amino acids S and V at residue position X153 and between amino acids A and V at residue position X383 have no significant deleterious effects on the engineered enzyme properties.

The structure and function information for exemplary non-naturally occurring (or engineered) transaminase polypeptides used in the method of the present disclosure are shown below in Table 2A and 2B. The odd numbered sequence identifiers (i.e., SEQ ID NOs) refer to the nucleotide sequence encoding the amino acid sequence provided by the even numbered SEQ ID NOs, and the sequences are provided in the electronic sequence listing file accompanying this disclosure, which is hereby incorporated by reference herein. The amino acid residue differences are based on comparison to the reference sequence of SEQ ID NO: 4, which is an engineered transaminases derived from the wild type α-Vif polypeptide having the following 24 amino acid residue differences A9T; G18A; D21H; V31M; N45H; F86Y; A133R; R146L; W147K; V153S; K163F; V177L; R211k; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y relative to SEQ ID NO:2. The activity of each engineered polypeptide relative to the reference polypeptide of SEQ ID NO: 4 was determined as conversion of the ketone substrate 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-one, to product amine compound (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine over a set time period and temperature in a high-throughput (HTP) assay, which was used as the primary screen. The HTP assay values in Table 2A were determined using E. coli clear cell lysates in 96 well-plate format of ~200 µL volume per well following assay reaction conditions as noted in the table and the Examples. In some instances, a shake-flask powder (SFP) and/or downstream processed (DSP) powder assay were used as a secondary screen to assess the properties of the engineered transaminases, the results of which are provided in Table 2B. The SFP and DSP forms provide a more purified powder preparation of the engineered polypeptides. For example, the engineered transaminase in the SFP preparations is approximately 30% of the total protein while the DSP preparations can contain the engineered transaminases that are approximately 80% of total protein.
The levels of activity (i.e., "+" "++", etc.) are defined as follows: "+" indicates 1.2 fold or greater activity as compared to that of SEQ ID NO: 4 for engineered transaminase polypeptide SEQ ID NO: 6 to 14, and 1.2 or greater activity as compared to that of SEQ ID NO: 8 for engineered transaminase polypeptide SEQ ID NO: 16 to 154; and a"++" indicates 5 fold or greater activity as compared to that of SEQ ID NO:4 for engineered transaminase polypeptides SEQ ID NO: 6-14, and 5 fold or greater activity as compared that of SEQ ID NO:8 for engineered transaminase polypeptides SEQ ID NO: 16 to 154.

The stability data is obtained by including the following amounts of product (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine in the assay and comparing the activity to that of a reference enzyme under the same conditions: 14 g/L for analysis of the engineered transaminase polypeptides of SEQ ID NO. 6 to 14, and 16 g/L for analysis of the engineered transaminases polypeptides SEQ ID NO. 16 to 154. Assessment of stability was made by comparing activities at two different temperatures, 55°C and 50°C.

**Table 2A:** Engineered Polypeptides and Relative Enzyme Improvements Using HTP Preparations

<table>
<thead>
<tr>
<th>SEQ ID NO:</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity&lt;sup&gt;a, b&lt;/sup&gt;</th>
<th>Stability&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Product Tolerance&lt;sup&gt;c, d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>n/a</td>
<td></td>
<td>Nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>3/4</td>
<td>A9T; G18A; D21H; V31M; N45H; F88Y; A133R; R146L; W147K; V153S; K163F; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>&quot;Control&quot;</td>
<td>&quot;Control&quot;</td>
<td>&quot;Control&quot;</td>
<td></td>
</tr>
<tr>
<td>5/6</td>
<td>A9T; G18A; D21H; V31M; N45H; F88Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>F163L;</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>7/8</td>
<td>A9T; G18A; D21H; V31M; N45H; F88Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>F163L;</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>9/10</td>
<td>A9T; G18A; D21H; V31M; N45H; F88Y; A133R; R146L; W147K; V153S; K163R; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>F163R;</td>
<td>+</td>
<td>nd</td>
<td></td>
</tr>
</tbody>
</table>

55
<table>
<thead>
<tr>
<th>SEQ ID NO:</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity&lt;sup&gt;a&lt;/sup&gt;,&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Stability&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Product Tolerance&lt;sup&gt;c, d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/12</td>
<td>A9T, G18A, D21H, V31M, N45H, F86D, A133R, R146L; W147K, V153S, K163F, V177L; R211K, P233T, A235P, P244T; M294V, P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y66D;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>13/14</td>
<td>A9T, G18A, D21H, V31M, N45H, F86Y, A133R, R146L; W147K, V153S, K163V, V177L; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>F163V;</td>
<td>+</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>15/16</td>
<td>A9T, Y14V, G18A, D21H, H26R; V31M, N45H, F86Y; R88L; Y113L; A133R, R146L; W147K, V153S, K163L, V177L; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V, H26R; R88L; Y113L; F163L;</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>17/18</td>
<td>A9T, Y14V, G18A, D21H, H26R; V31M, N45H, W57F; F86Y, A133R, R146L; W147K; V153S, K163L; V177L; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; F163L;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>19/20</td>
<td>A9T, Y14V, G18A, D21H; V31M; V33T, N45H, W57F; F86Y, R88L, Y113L, A133R; R146L; W147K; V153S, K163L; V177L; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G, A383V, T391A; C424A; F427Y;</td>
<td>Y14V, V33T, W57F; R88L; Y113L; F163L; V448E;</td>
<td>Nd</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>21/22</td>
<td>A9T, Y14V, G18A, D21H; H26R; V31M, N45H, F86Y, R88A; Y113L, A133R, R146L; W147K; N148Q; V153S, K163L; V177L; R203S; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G, A383V; T391A; C424A; F427Y;</td>
<td>Y14V, H26R; R88A; Y113L; N148Q; L163I; R203S;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>23/24</td>
<td>A9T, G18A, D21H, V31M; N45H, W57F; F86Y, R88L, Y113L; A133R, R146L; W147K; N148Q; V153S, K163I; V177L; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G, A383V; T391A, C424A; F427Y;</td>
<td>W57F; R88L; Y113L; N148Q; L163I; V448E;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>25/26</td>
<td>A9T, Y14V, G18A, D21H; V31M, N45H, F86Y, A133R; R146L; W147K; V153S; K163L; V177L; R211K, P233T, A235P, P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; F163L;</td>
<td>Nd</td>
<td>nd</td>
<td>+</td>
</tr>
<tr>
<td>SEQ ID NO:</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity b</td>
<td>Stability b</td>
<td>Product Tolerance c, d</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------</td>
<td>----------------------------------------</td>
<td>-----------</td>
<td>------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>27/28</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; F86Y; R88A; A133R; R146L; W147K; N148Q; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; R88A; N148Q; L163I;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>29/30</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; F86Y; R88L; Y113V; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; R88L; Y113V; F163L;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>31/32</td>
<td>A9T; Y14V; G18A; D21H; V31M; N45H; W57F; F86Y; Y113V; A133R; R146L; W147K; N148Q; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; W57F; Y113V; N148Q; F163L;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>33/34</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; W57F; F86Y; A133R; R146L; W147K; N148Q; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; N148Q; L163I;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>35/36</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; F68Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; F163L;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>37/38</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; W57F; F68Y; Y113L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; Y113L; F163L;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>39/40</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; F86Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; L163I;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>41/42</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; W57F; F68Y; R88L; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; R88L; L163I;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SEQ ID NO:</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity b</td>
<td>Stability b</td>
<td>Product Tolerance c, d</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>43/44</td>
<td>A9T; Y14V, G18A, D21H; H26R; V31M; N45H; F68Y; Y113V; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; Y113V; L163I;</td>
<td>Nd</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>45/46</td>
<td>A9T; Y14V, G18A, D21H; H26R; V31M; N45H; W57F; F68Y; R88L; Y113L; A133R; R146L; W147K; N148Q; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; R88L; Y113L; N148Q; L163I;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>47/48</td>
<td>A9T; Y14V, G18A, D21H; V31M; N45H; W57F; F68Y; R88L; A133R; R146L; W147K; N148Q; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; W57F; R88L; N148Q; L163I;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>49/50</td>
<td>A9T; Y14V, G18A, D21H; H26R; V31M; N45H; W57F; F68Y; R88L; Y113V; A133R; R146L; W147K; N148Q; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; R88L; Y113V; N148R; F163L;</td>
<td>Nd</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>51/52</td>
<td>A9T; Y14V, G18A, D21H; V31M; N45H; F68Y; R88A; Y113V; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; R88A; Y113V; L163I;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>53/54</td>
<td>A9T; Y14V, G18A, D21H; H26R; V31M; N45H; F68Y; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; F163L; D250A;</td>
<td>Nd</td>
<td>nd</td>
<td>+</td>
</tr>
<tr>
<td>55/56</td>
<td>A9T; Y14V, G18A, D21H; H26R; V31M; N45H; F68Y; R88L; Y113L; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y;</td>
<td>Y14V; H26R; R88L; Y113L; L163I;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SEQ ID NO:</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity b</td>
<td>Stability b</td>
<td>Product Tolerance c, d</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>SEQ ID NO:</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity b</td>
<td>Stability c, d</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------</td>
<td>---------------------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td></td>
</tr>
</tbody>
</table>

*a* Activity brating
*b* Stability
*c* Product Tolerance
<table>
<thead>
<tr>
<th>SEQ ID NO:</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Stability&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Product Tolerance&lt;sup&gt;c, d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>87/88</td>
<td>A9T; G18A; D21H; V31M; I41L; N45H; W57Y; F66Y; D107P; H132F; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; E315G; P318D; A323T; S324G; A383V; T391A; G395P; S398W; C424A; F427Y;</td>
<td>I41L; W57Y; D107P; H132F; F163L; E315G; G395P; S398W;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>89/90</td>
<td>A9T; G18A; D21H; V31M; I41L; N45H; F66Y; D107P; H132F; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; S398V; C424A; F427Y; E451D;</td>
<td>I41L; D107P; H132F; F163L; S388V; E451D;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>91/92</td>
<td>A9T; G18A; D21H; V31M; I41L; N45H; W57Y; F66Y; D107P; H132F; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; I314N; E315G; P318D; A323T; S324G; A383V; T391A; S398V; L423I; C424A; F427Y; E451D;</td>
<td>I41L; W57Y; D107P; H132F; F163L; E315G; S398V; L423I; E451D;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>93/94</td>
<td>A9T; G18A; D21H; V31M; I41L; N45H; W57Y; F66Y; D107P; H132F; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; E315G; P318D; A323T; S324G; A383V; T391A; S398V; L423I; C424A; F427Y;</td>
<td>I41L; W57Y; D107P; H132F; F163L; E315G; S398V; L423I;</td>
<td>Nd</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>95/96</td>
<td>A9T; G18A; D21H; V31M; I41L; N45H; W57Y; D70A; F66Y; D107P; H132F; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; I314N; P318D; A323T; S324G; A383V; T391A; G395P; S398W; Q419S; C424A; F427Y;</td>
<td>I41L; W57Y; D70A; D107P; H132F; F163L; I314N; S398W; Q419S;</td>
<td>Nd</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>97/98</td>
<td>A9T; G18A; D21H; V31S; N45H; W57F; F66Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; I314N; P318D; A323T; S324G; A383V; T391A; G395P; S398W; C424A; F427Y; E451D;</td>
<td>M31S; W57F; L163I; I314N; A346L; G395P; S398W; E451D;</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>99/100</td>
<td>A9T; G18A; D21H; V31S; N45H; W57F; F66Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; E315G; P318D; A323T; S324G; A383V; T391A; S398L; Q419S; C424A; F427Y;</td>
<td>M31S; W57F; L163I; E315G; A346L; S398L; Q419S;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>SEQ NO: 101/10</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity b</td>
<td>Stability b</td>
<td>Product Tolerance c, d</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SEQ NO: 103/10</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SEQ NO: 105/10</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SEQ NO: 107/10</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SEQ NO: 109/11</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SEQ NO: 111/11</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SEQ NO: 113/11</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Activity b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEQ NO:</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity a, b</td>
<td>Stability b</td>
<td>Product Tolerance c, d</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Nd = Not determined
<table>
<thead>
<tr>
<th>SEQ ID NO:</th>
<th>Active Mutations (relative to SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ NO ID: 4)</th>
<th>Activity $a$ b</th>
<th>Stability b</th>
<th>Product Tolerance c, d</th>
</tr>
</thead>
<tbody>
<tr>
<td>129/13 0</td>
<td>A9T, Y14V, G18A, D21H; H26R; V31M; N45H; F86Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P318D; A323T; S324G; A383V; T391A; S400G; L417M; S420N;</td>
<td>Y14V; H26R; L163I; S284A; S400G; L417M; S420N;</td>
<td>Nd</td>
<td>nd</td>
<td>+</td>
</tr>
<tr>
<td>131/13 2</td>
<td>A9T, Y14V, G18A, D21H; V31M; N45H; F86Y; A133R; R146L; W147K; V153S; K163L; L173A; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; S400G; S420N; C424A; F427Y;</td>
<td>Y14V; F163L; L173A; S400G; S420N;</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>133/13 4</td>
<td>A9T, Y14V, G18A, D21H; V31M; N45H; F86Y; Y113L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P318D; A323T; S324G; A383V; T391A; C424V; F427Y;</td>
<td>Y14V; Y113L; F163L; S284A; A424V;</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>135/13 6</td>
<td>A9T, Y14V, G18A, D21H; V31M; N45H; F86Y; R88L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324H; A383V; T391A; S400G; L417M; C424A; F427Y;</td>
<td>Y14V; R88L; F163L; G324H; S400G; L417M;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>137/13 8</td>
<td>A9T, Y14V, G18A; D21H; H26R; V31M; N45H; W57F; F86Y; Y113L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P318D; A323T; S324G; A383V; T391A; S400G; L417M; C424A; F427Y;</td>
<td>Y14V; H26R; W57F; Y113L; F163L; S284A;</td>
<td>Nd</td>
<td>nd</td>
<td>+</td>
</tr>
<tr>
<td>139/14 0</td>
<td>A9T, Y14V, G18A, D21H; H26R; V31M; N45H; F86Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P318D; A323T; S324G; T391A; S400G; C424A; F427Y;</td>
<td>Y14V; H26R; V383A; S400G;</td>
<td>Nd</td>
<td>nd</td>
<td>++</td>
</tr>
<tr>
<td>141/14 2</td>
<td>A9T, Y14V, G18A, D21H; H26R; V31M; N45H; F86Y; Y113L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; T391A; C424V; F427Y;</td>
<td>Y14V; H26R; Y113L; F163L; V383A; A424V;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>143/14 4</td>
<td>A9T, G18A, D21H; V31M; N45H; F86Y; R88L; A133R; R145K; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; R88L; F163L;</td>
<td>+</td>
<td>+</td>
<td>#VALUE!</td>
<td></td>
</tr>
<tr>
<td>SEQ ID NO:</td>
<td>Active Mutations (relative to SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Activity&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Stability&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Product Tolerance&lt;sup&gt;c, d&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>145/146</td>
<td>F427Y;</td>
<td>Y14V; H286R; N148Q; L163I; S284A; V383A; S400G; S420N;</td>
<td>Nd</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>147/148</td>
<td>A97; Y14V; G18A; D21H; V31M; V45H; W57F; F66Y; F86Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P318D; A323T; S324G; T391A; S400G; S420N; C424A; F427Y;</td>
<td>Y14V; V33T; W57F; Y113L; F163L;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>149/150</td>
<td>A97; Y14V; G18A; D21H; V31M; N45H; W57F; F66Y; R88L; Y113L; A133R; R146L; W147K; N148Q; V153S; K163L; L173A; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; T391A; S400G; C424A; F427Y;</td>
<td>Y14V; W57F; R88L; Y113L; N148Q; L163I; L173A; V383A; S400G;</td>
<td>Nd</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>151/152</td>
<td>A97; G18A; D21H; V31M; N45H; W57F; F66Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; A324G; T391A; C424A; F427Y;</td>
<td>W57F; F163L;</td>
<td>Nd</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>153/154</td>
<td>A97; Y14V; G18A; D21H; H268R; V31M; N45H; F66Y; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P318D; A323T; S324H; A383V; T391A; S400G; C424A; F427Y;</td>
<td>Y14V; H286R; F163L; S284A; G324H; S400G;</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
</tbody>
</table>

<sup>a</sup> HTP Assay Condition 1: Cells grown in 96 well plates were lysed with 200 μL of Lysis Buffer (1 mg/mL lysozyme, 0.5 mg/mL polymyxin B sulfate, 1 mM PLP, 0.1 M triethanolamine (TEA), pH 7.0). The reaction conditions comprised: 10 g/L (44.4 mM) Compound (IA); 1 mM pyridoxal-5-phosphate (PLP); 2 M isopropylamine (IPM), pH 7.0; 100 mM triethanolamine (TEA), pH 7.0; 5% PEG 200 v/v; 10 μL Lysate; and 50°C for 24 h.

<sup>b</sup> HTP Assay Condition 2: Cells grown in 96 well plates were lysed with 400 μL of Lysis Buffer (1 mg/mL lysozyme, 0.5 mg/mL polymyxin B sulfate, 1 mM PLP, 0.1 M triethanolamine (TEA), pH 7.0. The reaction conditions comprised: 10 g/L (44.4 mM) Compound (IA); 1 mM pyridoxal-5-phosphate (PLP); 2 M isopropylamine (IPM), pH 7.0; 100 mM triethanolamine (TEA), pH 7.0; 5% PEG 200 v/v; 10 μL Lysate; and
50°C for 24 h.

 Stability Assay conditions: 10 g/L (44.4 mM) Compound (IA); 1 mM pyridoxal-5-phosphate (PLP); 2 M isopropylamine (IPM), pH 7.0; 100 mM triethanolamine (TEA), pH 7.0; 5% PEG 200 v/v; 10 μL Lysate (prepared according to HTP Assay Condition 1 or 2); and 55°C for 24 h. Lysates were prepared according to HTP Assay Condition 1 (for SEQ ID NO: 4:90) or HTP Assay Condition 2 (for SEQ ID NO:92:230).

 Product Inhibition Assay Condition 1: 10 g/L (44.4 mM) Compound (IIA); 14 g/L Compound (IIA); 1 mM pyridoxal-5-phosphate (IPM), pH 7.0; 10 mM triethanolamine (TEA), pH 7.0; 5% PEG 200 v/v; 10 μL Lysate; and 50°C. Lysates were prepared according to HTP Assay Condition 1.

 Product Inhibition Assay Condition 2: 10 g/L (44.4 mM) Compound (IIA); 16 g/L Compound (IIA); 1 mM pyridoxal-5-phosphate (IPM), pH 7.0; 100 mM triethanolamine (TEA), pH 7.0; 5% PEG 200 v/v; 10 μL Lysate; and 50°C. Lysates were prepared according to HTP Assay Condition 2.

| nd: not determined |

Table 2B: Engineered Polypeptides and Relative Enzyme Improvements Using Shake Flask and DSP Preparations

<table>
<thead>
<tr>
<th>SEQ ID NO</th>
<th>Active Mutations (relative to Wild-type of SEQ ID NO: 2)</th>
<th>Active Mutations (relative to SEQ ID NO: 4)</th>
<th>Shake Flask Activity</th>
<th>DSP Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>½</td>
<td>n/a</td>
<td></td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>¾</td>
<td>A9T; G18A; D21 H; V31 M; N45H; F86Y; A133R; R146L; W147K; V153S; K163F; V177L; R21 1K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y</td>
<td>&quot;Control&quot;</td>
<td>&quot;Control&quot;</td>
<td></td>
</tr>
<tr>
<td>5/6</td>
<td>A9T; G18A; D21 H; V31 M; N45H; F86Y; A133R; R146L; W147K; V153S; K163L; V177L; R21 1K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y</td>
<td>F163L</td>
<td>99.7</td>
<td>+</td>
</tr>
<tr>
<td>7/8</td>
<td>A9T; G18A; D21 H; V31 M; N45H; F86Y; A133R; R145L; W147K; V153S; K163I; V177L; R21 1K; P233T;</td>
<td>F163L</td>
<td>99.5</td>
<td>+</td>
</tr>
<tr>
<td>SEQ ID NO.</td>
<td>Active Mutations (relative to Wild-type of SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Shake Flask Activity</td>
<td>DSP Activity</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>45/46</td>
<td>A9T; Y14V; G18A; D21H; H26R; V31M; N45H; W57F; F86Y; R88L; Y113L; A133R; R146L; W147K; N148Q; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; P316D; A323T; S324G; A335V; T391A; C424A; F427Y</td>
<td>Y14V; H26R; W57F; R88L; Y113L; N148Q; L163I</td>
<td>99.6 +</td>
<td>99.4 +</td>
</tr>
<tr>
<td>99/100</td>
<td>A9T; G18A; D21H; V31S; N45H; W57F; F86Y; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; E315G; P316D; A323T; S324G; A336L; A338V; T391A; S398L; Q419S; C424A; F427Y</td>
<td>M31S; W57F; L163I; E315G; A346L; S398L; Q419S</td>
<td>99.8 +</td>
<td>nd nd</td>
</tr>
<tr>
<td>111/112</td>
<td>A9T; Y14V; G18A; D21H; V31S; N45H; W57F; F86Y; A133R; R146L; W147K; V153S; K163I; V177L; R211K; P233T; A235P; P244T; M294V; I314N; A346L; S398V</td>
<td>Y14V; M31S; W57F; F163L; I314N; A346L; S398V</td>
<td>99.7 +</td>
<td>nd nd</td>
</tr>
<tr>
<td>133/134</td>
<td>A9T; Y14V; G18A; D21H; V31M; N45H; F86Y; Y113L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; S284A; M294V; P316D; A323T; S324G; A335V; T391A; C424V; F427Y</td>
<td>Y14V; Y113L; F163L; S284A; A424V</td>
<td>nd</td>
<td>Nd 99.8 +</td>
</tr>
<tr>
<td>143/144</td>
<td>A9T; G18A; D21H; V31M; N45H; F86Y; R88L; A133R; R146L; W147K; V153S; K163L; V177L; R211K; P233T; A235P; P244T; M294V; P316D; A323T; S324G; A335V; T391A; C424A; F427Y</td>
<td>R88L; F163L</td>
<td>99</td>
<td>Nd nd nd</td>
</tr>
<tr>
<td>SEQ ID NO:</td>
<td>Active Mutations (relative to Wild-type of SEQ ID NO: 2)</td>
<td>Active Mutations (relative to SEQ ID NO: 4)</td>
<td>Shake Flask Activity % ee</td>
<td>DSP Activity % ee</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
</tr>
</tbody>
</table>
| Active Shake Flask DSP Assay Conditions: 25 g/L (or 50 or 100 g/L) Compound (IA); 1 mM pyridoxal-5-phosphate (PLP); 2 M isopropylamine (IPM), pH 7.0; 5% v/v PEG200; 100 mM triethanolamine (TEA), pH 7.0; 2 g/L protein of transaminase-containing shake flask preparation; and 50°C for a reaction time of 24 h.

Active DSP Assay Conditions: 25 g/L (or 50 or 100 g/L) Compound (IA); 1 mM pyridoxal-5-phosphate (PLP); 2 M isopropylamine (IPM), pH 7.0; 5% v/v PEG200; 100 mM triethanolamine (TEA), pH 7.0; 2 g/L protein from transaminase-containing DSP preparation; and 50°C for a reaction time of 24 h.

From an inspection of the exemplary polypeptides useful in the methods of the present invention, improvements in enzyme properties are associated with residue differences as compared to SEQ ID NO:4 at residue positions X14; X26; X31; X33; X41; X47; X57; X70; X86; X88; X107; X113; X132; X148; X163; X168; X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X420; X423; X424; X448; and X451. The specific residue differences at each of these positions that are associated with the improved properties include: X14V; X26R; X31S; X33T; X41L; X47N; X57F; X57Y; X70A; X86D; X88A; X88L; X107P; X113L; X113V; X132F; X148Q; X148R; X163I; X163L; X163R; X163V; X168K; X168S; X173A; X203S; X250A; X284A; X314N; X315G; X324H; X346L; X395P; X398L; X398V; X398W; X400G; X417M; X419S; X420N; X423I; X424V; X448E; and X451D.

Accordingly, in some embodiments, the engineered transaminase polypeptide useful in the methods of the instant disclosure are capable of converting substrate Compound (IA), 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-one, to the product Compound (HA), (S)-1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine with improved properties as compared to SEQ ID NO:4, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to reference sequence SEQ ID NO:2 and one or more residue differences as compared to SEQ ID NO:4 at residue positions selected from: X14; X26; X31; X33; X41; X47; X57; X70; X86; X88; X107; X132; X148; X163; X168; X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X423; X448; and X451, where the residue differences at residue positions X31; X57; X86; X163; X168; X314; X324; X398; and X417 are selected from: X31S; X57Y; X86D; X163I; X133L; X163R; X163V; X168S; X314N; X324H; X398L; X398V; X398W; and X417M.
In some embodiments, the engineered transaminase polypeptide useful in the instantly disclosed methods is capable of converting substrate Compound (IA) to the product Compound (HA) with improved properties as compared to SEQ ID NO:4, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to reference sequence SEQ ID NO:4 and one or more residue differences as compared to SEQ ID NO:4 at residue positions selected from: X14; X26; X31; X33; X41; X47; X57; X70; X86; X88; X107; X132; X148; X163; X168; X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X423; X448; and X451, wherein the residue differences at residue positions X31; X57; X86; X163; X168; X314; X324; X398; and X417 are selected from: X31S; X57Y; X86D; X163I; X163L; X163R; X163V; X168S; X314N; X324H; X398L; X398V; X398W; and X417M. In some embodiments, the engineered transaminase polypeptides are capable of carrying out the conversion with the enantioselectivities described above, e.g., ≥ 90% ee.

In some embodiments, the transaminase polypeptide useful in the instantly disclosed methods is capable of converting substrate Compound (IA) to product Compound (HA) with improved properties as compared to SEQ ID NO:4, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identity to a reference sequence selected from SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154, and one or more residue differences as compared to SEQ ID NO:4 at residue positions selected from: X14; X26; X31; X33; X41; X47; X57; X70; X86; X88; X107; X132; X148; X163; X168; X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X423; X448; and X451, where the residue differences at residue positions X31; X57; X86; X163; X168; X314; X324; X398; and X417 are selected from: X31S; X57Y; X86D; X163I; X163L; X163R; X163V; X168S; X314N; X324H; X398L; X398V; X398W; and X417M. In some embodiments, the reference sequence is selected from SEQ ID NO: 4, 8, 14, 16, 132, 134, and 146. In some embodiments, the reference sequence is SEQ ID NO:4. In some embodiments, the reference sequence is SEQ ID NO:8. In some embodiments, the reference sequence is SEQ ID NO:134. In some embodiments, the reference sequence is SEQ ID NO:146.
In some embodiments, the transaminase polypeptide capable of converting substrate Compound (IA) to product Compound (HA) with improved properties as compared to SEQ ID NO:4, comprises an amino acid sequence selected from: SEQ ID NO:6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154, and having one or more residue differences as compared to SEQ ID NO:4 at residue positions selected from: X14; X26; X31, X33; X41; X47; X57; X70; X86; X88; X107; X132; X148; X163; X168; X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X423; X448; and X451, where the residue differences at residue positions X31; X57; X86; X163; X168; X314; X324; X398; and X417 are selected from: X31S; X57Y; X86D; X163I; X163L; X163R; X163V; X168S; X314N; X324H; X398L; X398V; X398W; and X417M. In some embodiments, the amino acid sequence is selected from SEQ ID NO: 4, 8, 14, 16, 132, 134, and 146. In some embodiments, the amino acid sequence is SEQ ID NO:4. In some embodiments, the amino acid sequence is SEQ ID NO:8. In some embodiments, the amino acid sequence is SEQ ID NO:134. In some embodiments, the amino acid sequence is SEQ ID NO:146.

In some embodiments, the engineered transaminase polypeptide is capable of converting the substrate Compound (IA) to the product Compound (HA) with at least 1.2 fold the activity relative to SEQ ID NO:4 comprises an amino acid sequence selected from: SEQ ID NO: 6, 8, 10, 12, 14, 16, 22, 24, 28, 30, 32, 48, 52, 56, 62, 64, 66, 68, 86, 88, 90, 92, 98, 100, 104, 106, 108, 110, 112, 114, 116, 118, 124, 126, 128, 132, 134, 136, 142, 144, 148, and 154.

In some embodiments, the engineered transaminase polypeptide is capable of converting the substrate Compound (IA) to the product Compound (IIA) with at least 5 fold the activity relative to SEQ ID NO:4 and comprises an amino acid sequence having one or more residue differences selected from: X14V, X26R; X33T; X88A/L; X163I/L; and X284A.

In some embodiments, the engineered transaminase polypeptide useful in the instantly disclosed methods is capable of converting the substrate Compound (IA) to the product Compound (IIA) with at least 5 fold the activity relative to SEQ ID NO:4 comprises an amino acid sequence selected from: SEQ ID NO: 6, 8, 10, 14, 16, 22, 24, 28, 30, 32, 48,

In some embodiments, the engineered transaminase polypeptide has at least 1.2 fold, 1.5 fold, 2 fold, 3 fold, 4 fold, 5 fold, or more refractoriness to inhibition by 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine, in particular (S)-1-(1 H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine, as compared to the polypeptide represented by SEQ ID NO:4 in the conversion of Compound (IA) to Compound (HA). Generally, the increase in refractoriness or resistance to inhibition by the product compound can be measured under HTP assay conditions in presence of 14 g/L or 16 g/L of Compound (HA), as described in Table 2A and 2B and the Examples. In some embodiments, the engineered transaminase polypeptide having at least 1.2 fold or greater refractoriness to inhibition by 1-(1H-5-fluoro-6-chloro-indol-3-yl)propan-2-amine comprises an amino acid sequence having one or more residue differences as compared to SEQ ID NO:4 selected from: X26R; X70A; X86D; X88A/L; X132F; X163L; X315G; X395P; X398L; and X419S.

In some embodiments, the engineered transaminase polypeptide with improved properties in the conversion of Compound (IA) to Compound (HA) has an amino acid sequence comprising a sequence selected from SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154.

In some embodiments, the engineered transaminase capable of converting Compound (IA) to Compound (IIA) under suitable reaction conditions, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to one of SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154, and the amino acid residue differences as compared to SEQ ID NO:4 present in any one of SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112,
114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154, as provided in Tables 2A and 2B.

In some embodiments, the engineered polypeptides can be in various forms, for example, such as an isolated preparation, as a substantially purified enzyme, whole cells transformed with gene(s) encoding the enzyme, and/or as cell extracts and/or lysates of such cells. The enzymes can be lyophilized, spray-dried, precipitated or be in the form of a crude paste, as further discussed below.

In some embodiments, the transaminase polypeptides can be bound on a physical substrate. The transaminase polypeptide can be bound non-covalently or covalently on a physical substrate such that they retain their improved activity, stereoselectivity, product tolerance, stability, and/or other improved properties relative to the reference polypeptide of SEQ ID NO: 4. In such embodiments, the immobilized polypeptides can facilitate the biocatalytic conversion of the suitable ketone substrates, e.g., Compound (IA) or structural analogs thereof, to the corresponding amine product, e.g., Compound (IIA) or corresponding structural analogs, and after the reaction is complete are easily retained (e.g., by retaining beads on which polypeptide is immobilized) and then reused or recycled in subsequent reactions. Such immobilized enzyme processes allow for further efficiency and cost reduction. Methods of enzyme immobilization are well-known in the art. Various methods for conjugation to substrates, e.g., membranes, beads, glass, etc. are described in, among others, Hermanson, G.T., Bioconjugate Techniques, Second Edition, Academic Press; (2008), and Bioconjugation Protocols: Strategies and Methods, In Methods in Molecular Biology, C M. Niemeyer ed., Humana Press (2004); the disclosures of which are incorporated herein by reference.

Polynucleotides Encoding Engineered Transaminases, Expression Vectors and Host Cells

In another aspect, the present disclosure provides polynucleotides encoding the engineered transaminase polypeptides described herein. The polynucleotides may be operatively linked to one or more heterologous regulatory sequences that control gene expression to create a recombinant polynucleotide capable of expressing the polypeptide. Expression constructs containing a heterologous polynucleotide encoding the engineered transaminase can be introduced into appropriate host cells to express the corresponding transaminase polypeptide.
As will be apparent to the skilled artisan, availability of a protein sequence and the knowledge of the codons corresponding to the various amino acids provide a description of all the polynucleotides capable of encoding the subject polypeptides. The degeneracy of the genetic code, where the same amino acids are encoded by alternative or synonymous codons, allows an extremely large number of nucleic acids to be made, all of which encode the improved transaminase enzymes. Thus, having knowledge of a particular amino acid sequence, those skilled in the art could make any number of different nucleic acids by simply modifying the sequence of one or more codons in a way which does not change the amino acid sequence of the protein. In this regard, the present disclosure specifically contemplates each and every possible variation of polynucleotides that could be made encoding the polypeptides described herein by selecting combinations based on the possible codon choices, and all such variations are to be considered specifically disclosed for any polypeptide described herein, including the amino acid sequences presented in Tables 2A and 2B, and disclosed in the sequence listing incorporated by reference herein as SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154.

In various embodiments, the codons are preferably selected to fit the host cell in which the protein is being produced. For example, preferred codons used in bacteria are used to express the gene in bacteria; preferred codons used in yeast are used for expression in yeast; and preferred codons used in mammals are used for expression in mammalian cells. In some embodiments, all codons need not be replaced to optimize the codon usage of the transaminases since the natural sequence will comprise preferred codons and because use of preferred codons may not be required for all amino acid residues. Consequently, codon optimized polynucleotides encoding the transaminase enzymes may contain preferred codons at about 40%, 50%, 60%. 70%, 80%, or greater than 90% of codon positions of the full length coding region.

An isolated polynucleotide encoding an improved transaminase polypeptide may be manipulated in a variety of ways to provide for expression of the polypeptide. In some embodiments, the polynucleotides encoding the polypeptides can be provided as expression vectors where one or more control sequences is present to regulate the

73

In the embodiments herein, the improved polypeptides and corresponding polynucleotides can be obtained using methods used by those skilled in the art. The parental polynucleotide sequence encoding the wild-type polypeptide of *Vibrio fluvialis* is described in Shin *et al.*, 2003, Appl. Microbiol. Biotechnol. 61(5-6):463-471, and methods of generating engineered transaminase polypeptides with improved stability and substrate recognition properties are disclosed in patent application publications WO201 0081 053 and US201 00209981, incorporated herein by reference.

Where the sequence of the engineered polypeptide is known, the polynucleotides encoding the enzyme can be prepared by standard solid-phase methods, according to known synthetic methods. In some embodiments, fragments of up to about 100 bases can be individually synthesized, then joined (e.g., by enzymatic or chemical litigation methods, or polymerase mediated methods) to form any desired continuous sequence. For example, polynucleotides and oligonucleotides of the invention can be prepared by chemical synthesis using, e.g., the classical phosphoramidite method described by Beaucage *et al.*, 1981, Tet Lett 22:1859-69, or the method described by Matthes *et al.*, 1984, EMBO J. 3:801-05, e.g., as it is typically practiced in automated synthetic methods. According to the phosphoramidite method, oligonucleotides are synthesized, e.g., in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors. In addition, essentially any nucleic acid can be obtained from any of a variety of commercial sources.

Accordingly, in some embodiments, a method for preparing the engineered transaminases polypeptide can comprise: (a) synthesizing a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136,
138, 140, 142, 144, 146, 148, 150, 152, and 154 and having one or more residue differences as compared to SEQ ID NO:4 at residue positions selected from: X14; X26; X31; X33; X41; X47; X57; X70; X86; X88; X107; X132; X148; X163; X168; X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X423; X448; and X451, wherein the residue differences at residue positions X31; X57; X86; X163; X168; X314; X324; X398; and X417 are selected from: X31S; X57Y; X86D; X163I; X163L; X163R; X163V; X168S; X314N; X324H; X398L; X398V; X398W; and X417M.; and (b) expressing the transaminase polypeptide encoded by the polynucleotide.

The expressed engineered transaminase can be measured for the desired improved property, e.g., activity, enantioselectivity, stability, and product tolerance, in the conversion of compound (IA) to compound (HA) by any of the assay conditions described herein.

In some embodiments, any of the engineered transaminase enzymes expressed in a host cell can be recovered from the cells and or the culture medium using any one or more of the well known techniques for protein purification, including, among others, lysozyme treatment, sonication, filtration, salting-out, ultra-centrifugation, and chromatography. Suitable solutions for lysing and the high efficiency extraction of proteins from bacteria, such as E. coli, are commercially available under the trade name CelLytic™ from Sigma-Aldrich of St. Louis MO.

Chromatographic techniques for isolation of the transaminase polypeptide include, among others, reverse phase chromatography high performance liquid chromatography, ion exchange chromatography, gel electrophoresis, and affinity chromatography. Conditions for purifying a particular enzyme will depend, in part, on factors such as net charge, hydrophobicity, hydrophilicity, molecular weight, molecular shape, etc., and will be apparent to those having skill in the art.

In some embodiments, affinity techniques may be used to isolate the improved transaminase enzymes. For affinity chromatography purification, any antibody which specifically binds the transaminase polypeptide may be used. For the production of antibodies, various host animals, including but not limited to rabbits, mice, rats, etc., may be immunized by injection with a transaminase polypeptide, or a fragment thereof. The transaminase polypeptide or fragment may be attached to a suitable carrier, such as BSA, by means of a side chain functional group or linkers attached to a side chain functional group. Various adjuvants may be used to increase the immunological
response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, poiyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacilli Calmette Guerin) and Corynebacterium parvum.

Methods of Using the Engineered Transaminase Enzymes

In another aspect, the transaminases described herein can be used in a process for carrying out transaminase reactions in which an amino group from an amino donor is transferred to an amino acceptor, e.g., ketone substrate, to produce an amine. Use of a prochiral ketone acceptor can result in the production of a chiral amine in enantiomeric excess. Generally, the process for performing the transamination reaction comprises contacting or incubating an amino donor and an amino acceptor with an engineered transaminase polypeptide of the disclosure under reaction conditions suitable for converting the amino acceptor to an amine.

In some embodiments, the transaminases can be used in the conversion of substrate compound of formula (I) to product compound of formula (II). Accordingly, in some embodiments, a process for preparing the compound of formula (II) in enantiomeric excess comprises contacting the compound of formula (I) in the presence of an amino donor under suitable reaction conditions with an engineered transaminase polypeptide described herein. In some embodiments of the process, the compound of formula (II) can be formed in at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater enantiomeric excess.

For the foregoing processes, any of the engineered transaminases described herein can be used. By way of example and without limitation, in some embodiments, the process can use an engineered transaminases polypeptide comprising comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identity to a reference sequence selected from SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154, and one or more residue differences as compared to SEQ ID NO:4 at residue positions selected from: X14; X26; X31; X33; X41; X47; X57; X70; X86; X88; X107; X132; X148; X163; X168;
X173; X203; X250; X284; X314; X315; X324; X346; X395; X398; X400; X417; X419; X423; X448; and X451, wherein the residue differences at residue positions X31; X57; X86; X163; X168; X314; X324; X398; and X417 are selected from: X31S; X57Y; X86D; X163L; X163R; X163V; X168S; X314N; X324H; X398L; X398V; X398W; and X417M. In some embodiments, the reference sequence is selected from SEQ ID NO: 4, 8, 14, 16, 132, 134, and 146. In some embodiments, the reference sequence is SEQ ID NO:4. In some embodiments, the reference sequence is SEQ ID NO:8. In some embodiments, the reference sequence is SEQ ID NO:134. In some embodiments, the reference sequence is SEQ ID NO:146.

In some embodiments, exemplary transaminases capable of carrying out the processes herein can be a polypeptide comprising an amino acid sequence selected from SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, and 154. Guidance on the choice and use of the engineered transaminases is provided in the descriptions herein, for example Table 2 and the Examples.

In the embodiments herein and illustrated in the Examples, various ranges of suitable reaction conditions that can be used in the processes, including but not limited, to ranges of amino donor, pH, temperature, buffer, solvent system, substrate loading, polypeptide loading, cofactor loading, pressure, and reaction time. Further suitable reaction conditions for carrying out the process for biocatalytic conversion of substrate compounds to product compounds using an engineered transaminase polypeptide described herein can be readily optimized in view of the guidance provided herein by routine experimentation that includes, but is not limited to, contacting the engineered transaminase polypeptide and substrate compound under experimental reaction conditions of concentration, pH, temperature, solvent conditions, and detecting the product compound.

In the embodiments described herein, the transaminase polypeptide uses an amino donor to form the product compounds. In some embodiments, the amino donor in the reaction condition can be selected from isopropylamine (also referred to herein as “IPM”), putrescine, L-lysine, ct-phenethylamine, D-alanine, L-alanine, or D,L-alanine, or D,L-ornithine. In some embodiments, the amino donor is selected from the group
consisting of IPM, putrescine, L-lysine, D- or L-alanine. In some embodiments, the amino donor is IPM. In some embodiments, the suitable reaction conditions comprise the amino donor, in particular IPM, present at a concentration of at least about 0.1 to about 3.0 M, 0.2 to about 2.5 M, about 0.5 to about 2 M or about 1 to about 2 M. In some embodiments, the amino donor is present at a concentration of about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1, 1.5, 2, 2.5 or 3 M. Higher concentrations of amino donor, e.g., IPM, can be used to shift the equilibrium towards amine product formation.

Suitable reaction conditions using the engineered transaminase polypeptides also typically comprise a cofactor. Cofactors useful in the processes using the engineered transaminase enzymes include, but are not limited to, pyridoxal-5'-phosphate (also known as pyridoxal-phosphate, PLP, P5P), pyridoxine (PN), pyridoxal (PL), pyridoxamine (PM), and their phosphorylated counterparts; pyridoxine phosphate (PNP), and pyridoxamine phosphate (PMP). In some embodiments, the cofactor PLP is present naturally in the cell extract and does not need to be supplemented. In some embodiments of the methods, the suitable reaction conditions comprise cofactor added to the enzyme reaction mixture, for example, when using partially purified, or purified transaminase enzyme. In some embodiments, the suitable reaction conditions can comprise the presence of a cofactor selected from PLP, PN, PL, PM, PNP, and PMP, at a concentration of about 0.1 g/L to about 10 g/L, about 0.2 g/L to about 5 g/L, about 0.5 g/L to about 2.5 g/L. In some embodiments, the reaction conditions comprise a PLP concentration of about 0.1 g/L or less, 0.2 g/L or less, 0.5 g/L or less, 1 g/L or less, 2.5 g/L or less, 5 g/L or less, or 10 g/L or less. In some embodiments, the cofactor can be added either at the beginning of the reaction and/or additional cofactor is added during the reaction.

Substrate compound in the reaction mixtures can be varied, taking into consideration, for example, the desired amount of product compound, the effect of substrate concentration on enzyme activity, stability of enzyme under reaction conditions, and the percent conversion of substrate to product. In some embodiments, the suitable reaction conditions comprise a substrate compound loading of at least about 0.5 to about 200 g/L, 1 to about 200 g/L, 5 to about 150 g/L, about 10 to about 100 g/L, 20 to about 100 g/L or about 50 to about 100 g/L. In some embodiments, the suitable reaction conditions comprise a substrate compound loading of at least about 0.5 g/L, at least about 1 g/L, at least about 5 g/L, at least about 10 g/L, at least about 15 g/L, at least about 20 g/L, at
least about 30 g/L, at least about 50 g/L, at least about 75 g/L, at least about 100 g/L, at least about 150 g/L or at least about 200 g/L, or even greater.

The improved activity and/or stereoselectivity of the engineered transaminase polypeptides disclosed herein provides for processes wherein higher percentage conversion can be achieved with lower concentrations of the engineered polypeptide. In some embodiments of the process, the suitable reaction conditions comprise an engineered polypeptide concentration of about 0.01 to about 50 g/L; about 0.05 to about 50 g/L; about 0.1 to about 40 g/L; about 1 to about 40 g/L; about 2 to about 40 g/L; about 5 to about 40 g/L; about 5 to about 30 g/L; about 0.1 to about 10 g/L; about 0.5 to about 10 g/L; about 1 to about 10 g/L; about 0.1 to about 5 g/L; about 0.5 to about 5 g/L; or about 0.1 to about 2 g/L. In some embodiments, the transaminase polypeptide is concentration at about 0.01, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, or 50 g/L.

During the course of the transamination reactions, the pH of the reaction mixture may change. The pH of the reaction mixture may be maintained at a desired pH or within a desired pH range. This may be done by the addition of an acid or a base, before and/or during the course of the reaction. Alternatively, the pH may be controlled by using a buffer. Accordingly, in some embodiments, the reaction condition comprises a buffer. Suitable buffers to maintain desired pH ranges are known in the art and include, by way of example and not limitation, borate, carbonate, phosphate, triethanolamine buffer, and the like. In some embodiments, the buffer is borate. In some embodiments of the process, the suitable reaction conditions comprise a buffer solution of triethanolamine, where the triethanolamine concentration is from about 0.01 to about 0.4 M, 0.05 to about 0.4 M, 0.1 to about 0.3 M, or about 0.1 to about 0.2 M. In some embodiments, the reaction condition comprises a triethanolamine concentration of about 0.01, 0.02, 0.03, 0.04, 0.05, 0.07, 0.1, 0.12, 0.14, 0.16, 0.18, 0.2, 0.3, or 0.4 M. In some embodiments, the reaction conditions comprise water as a suitable solvent with no buffer present.

In the embodiments of the process, the reaction conditions can comprise a suitable pH. The desired pH or desired pH range can be maintained by use of an acid or base, an appropriate buffer, or a combination of buffering and acid or base addition. The pH of the reaction mixture can be controlled before and/or during the course of the reaction. In some embodiments, the suitable reaction conditions comprise a solution pH comprise a pH from about 6 to about 12, pH from about 6 to about 10, pH from about 6 to about 8,
pH from about 7 to about 10, pH from about 7 to about 9, or pH from about 7 to about 8. In some embodiments, the reaction conditions comprise a solution pH of about 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 11.5 or 12.

In the embodiments of the processes herein, a suitable temperature can be used for the reaction conditions, for example, taking into consideration the increase in reaction rate at higher temperatures, and the activity of the enzyme during the reaction time period. For example, the engineered polypeptides of the present disclosure have increased stability relative to naturally occurring transaminase polypeptide e.g., the wild type polypeptide of SEQ ID NO: 2, which allow the engineered polypeptides to be used at higher temperatures for increased conversion rates and improved substrate solubility characteristics for the reaction. Accordingly, in some embodiments, the suitable reaction conditions comprise a temperature of about 10°C to about 70°C, about 10°C to about 65°C, about 15°C to about 60°C, about 20°C to about 60°C, about 20°C to about 55°C, about 30°C to about 55°C, or about 40°C to about 50°C. In some embodiments, the suitable reaction conditions comprise a temperature of about 10°C, 15°C, 20°C, 25°C, 30°C, 35°C, 40°C, 45°C, 50X, 55°C, 60°C, 65°C, or 70°C. In some embodiments, the temperature during the enzymatic reaction can be maintained at a temperature throughout the course of the reaction. In some embodiments, the temperature during the enzymatic reaction can be adjusted over a temperature profile during the course of the reaction.

The processes of the disclosure are generally carried out in a solvent. Suitable solvents include water, aqueous buffer solutions, organic solvents, polymeric solvents, and/or co-solvent systems, which generally comprise aqueous solvents, organic solvents and/or polymeric solvents. The aqueous solvent (water or aqueous co-solvent system) may be pH-buffered or unbuffered. In some embodiments, the processes using the engineered transaminase polypeptides are generally carried out in an aqueous co-solvent system comprising an organic solvent (e.g., ethanol, isopropanol (IPA), dimethyl sulfoxide (DMSO), ethyl acetate, butyl acetate, 1-octanol, heptane, octane, methyl t butyl ether (MTBE), toluene, and the like), ionic or polar solvents (e.g., 1 ethyl 4 methylimidazolium tetrafluoroborate, 1 butyl 3 methylimidazolium tetrafluoroborate, 1 butyl 3 methylimidazolium hexafluorophosphate, glycerol, polyethylene glycol, and the like). In some embodiments, the co-solvent can be a polar solvent, such as a polyl, dimethylsulfoxide, DMSO, or lower alcohol. The non-aqueous co-solvent component of
an aqueous co-solvent system may be miscible with the aqueous component, providing
a single liquid phase, or may be partly miscible or immiscible with the aqueous
component, providing two liquid phases. Exemplary aqueous co-solvent systems can
comprise water and one or more co-solvents selected from an organic solvent, polar
solvent, and polyol solvent. In general, the co-solvent component of an aqueous co-
solvent system is chosen such that it does not adversely inactivate the transaminase
enzyme under the reaction conditions. Appropriate co-solvent systems can be readily
identified by measuring the enzymatic activity of the specified engineered transaminase
enzyme with a defined substrate of interest in the candidate solvent system, utilizing an
enzyme activity assay, such as those described herein.

In some embodiments of the process, the suitable reaction conditions comprise an
aqueous co-solvent, where the co-solvent comprises a polyol solvent, particularly
glycols. Examples of suitable polyol solvents include, by way of example and not
limitation, polyethylene glycol, polyethylene glycol methyl ether, diethylene glycol
dimethyl ether, triethylene glycol dimethyl ether, and polypropylene glycol. In some
embodiments, the aqueous co-solvent comprises polyethylene glycol, which is available
in different molecular weights. Particularly useful are lower molecular weight glycols,
such as PEG200 to PEG600. Accordingly, in some embodiments, the aqueous co-
solvent comprises PEG200 of about 1% to about 40% v/v; about 1% to about 40% v/v;
about 2% to about 40% v/v; about 5% to about 40% v/v; 2% to about 30% v/v; 5% to
about 30% v/v; 1 to about 20% v/v; about 2% to about 20% v/v; about 5% to about 20%
v/v; about 1% to about 10% v/v; about 2% to about 10% v/v. In some embodiments, the
suitable reaction conditions comprise an aqueous co-solvent comprising PEG200 at
about 1%, 2%, 5%, 10%, 15%, 20%; 25%; 30%; 35%; 35% or about 40% v/v.

in some embodiments of the process, the suitable reaction conditions comprise an
aqueous co-solvent, where the co-solvent comprises DMSO at about 1% to about 80%
(v/v), about 1 to about 70% (v/v), about 2% to about 60% (v/v), about 5% to about 40%
(v/v), 10% to about 40% (v/v), 10% to about 30% (v/v), or about 10% to about 20%
(v/v). In some embodiments of the process, the suitable reaction conditions comprise an
aqueous co-solvent comprising DMSO at least about 1%, 5%, 10%, 15%, 20%, 25%,
30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, or 80% (v/v).

The quantities of reactants used in the transamination reaction will generally vary
depending on the quantities of product desired, and concomitantly the amount of
transaminase substrate employed. Those having ordinary skill in the art will readily understand how to vary these quantities to tailor them to the desired level of productivity and scale of production.

In some embodiments, the order of addition of reactants is not critical. The reactants may be added together at the same time to a solvent (e.g., monophasic solvent, biphasic aqueous co-solvent system, and the like), or alternatively, some of the reactants may be added separately, and some together at different time points. For example, the cofactor, transaminase, and transaminase substrate may be added first to the solvent.

The solid reactants (e.g., enzyme, salts, etc.) may be provided to the reaction in a variety of different forms, including powder (e.g., lyophilized, spray dried, and the like), solution, emulsion, suspension, and the like. The reactants can be readily lyophilized or spray dried using methods and equipment that are known to those having ordinary skill in the art. For example, the protein solution can be frozen at -80°C in small aliquots, then added to a pre-chilled lyophilization chamber, followed by the application of a vacuum.

For improved mixing efficiency when an aqueous co-solvent system is used, the transaminase, and cofactor may be added and mixed into the aqueous phase first. The organic phase may then be added and mixed in, followed by addition of the transaminase substrate. Alternatively, the transaminase substrate may be premixed in the organic phase, prior to addition to the aqueous phase.

The transamination reaction is generally allowed to proceed until further conversion of ketone substrate to amine product does not change significantly with reaction time, e.g., less than 10% of substrate being converted, or less than 5% of substrate being converted). In some embodiments, the reaction is allowed to proceed until there is complete or near complete conversion of substrate ketone to product amine.

Transformation of substrate to product can be monitored using known methods by detecting substrate and/or product. Suitable methods include gas chromatography, HPLC, and the like. Conversion yields of the chiral amine product generated in the reaction mixture are generally greater than about 50%, may also be greater than about 60%, may also be greater than about 70%, may also be greater than about 80%, may also be greater than 90%, and are often greater than about 97%. In some embodiments, the methods for preparing compounds of formula (II) using an engineered transaminase polypeptide under suitable reaction conditions results in at least about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater conversion of ketone substrate, e.g.
compound of formula (I), to the amine product compound, e.g., compound of formula (II) in about 48h or less, in about 36 h or less, in about 24 h or less, or even less time.

In some embodiments of the process, the suitable reaction conditions comprise a substrate loading of substrate compound of at least about 20 g/L, 30 g/L, 40 g/L, 50 g/L, 60 g/L, 70 g/L, 100 g/L, or more, and wherein the method results in at least about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater conversion of substrate compound to product compound in about 48h or less, in about 36 h or less, or in about 24 h or less.

The engineered transaminase polypeptides of the present disclosure when used in the process under suitable reaction conditions result in an enantiomeric excess of the chiral amine in at least 97%, 98, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9%. e.e.

In a further embodiment of the methods for converting substrate compound to amine product compound using the engineered transaminase polypeptides, the suitable reaction conditions comprise an initial substrate loading to the reaction solution which is then contacted by the polypeptide. This reaction solution is the further supplemented with additional substrate of compound as a continuous addition over time at a rate of at least about 1 g/L/h, at least about 2 g/L/h, at least about 4 g/L/h, at least about 6 g/L/h, or higher. Thus, according to these suitable reaction conditions polypeptide is added to a solution having an initial substrate loading of at least about 20 g/L, 30 g/L, or 40 g/L. This addition of polypeptide is then followed by continuous addition of further substrate to the solution at a rate of about 2 g/L/h, 4 g/L/h, or 6 g/L/h until a much higher final substrate loading of at least about 30 g/L, 40 g/L, 50 g/L, 60 g/L, 70 g/L, 100 g/L, 150 g/L, 200 g/L or more, is reached. Accordingly, in some embodiments of the method, the suitable reaction conditions comprise addition of the polypeptide to a solution having an initial substrate loading of at least about 20 g/L, 30 g/L, or 40 g/L followed by addition of further substrate to the solution at a rate of about 2 g/L/h, 4 g/L/h, or 6 g/L/h until a final substrate loading of at least about 30 g/L, 40 g/L, 50 g/L, 60 g/L, 70 g/L, 100 g/L or more, is reached. This substrate supplementation reaction condition allows for higher substrate loadings to be achieved while maintaining high rates of conversion of ketone substrate to amine product of at least about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater. In some embodiments of this method, the further substrate added is in a solution comprising isopropylamine or isopropylamine acetate at a concentration
of at least about 0.5 M, at least about 1.0 M, at least about 2.5 M, at least about 5.0 M, 
at least about 7.5 M, at least about 10.0 M.

In some embodiments of the processes, the transamination reaction using an 
engineered transaminase polypeptide can comprise the following suitable reaction 
conditions: (a) substrate loading at about 5 g/L to 200 g/L; (b) about 0.1 to 50 g/L of 
transaminases polypeptide; (c) about 0.1 to 4 M of isopropylamine (IPM); (d) about 0.1 
to 10 g/L of pyridoxal phosphate (PLP) cofactor; (e) pH of about 6 to 9; and (f) 
temperature of about 30 to 60°C.

In some embodiments of the processes, the transamination reaction using an 
engineered transaminase polypeptide can comprise the following suitable reaction 
conditions: (a) substrate loading at about 5 to about 20 g/L; (b) about 0.05 to 2 g/L of 
transaminases polypeptide; (c) about 1 to 10% v/v of PEG200; (d) about 1 to 2 M of 
isopropylamine (IPM); (e) about 0.1 to 1 g/L of pyridoxal phosphate (PLP) cofactor; (f) 
about 0.1 to about 0.5 M of triethanolamine (TEA); (g) pH of about 6 to 8; and (h) 
temperature of about 45 to 55°C.

In some embodiments of the processes, the transamination reaction using an 
engineered transaminase polypeptide can comprise the following suitable reaction 
conditions: (a) substrate loading of about 25 to about 100 g/L; (b) about 0.5 to 10 g/L of 
transaminases polypeptide; (c) about 1 to 10% v/v of PEG200; (d) about 1 to 2 M of 
isopropylamine (IPM); (e) about 0.1 to 1 g/L of pyridoxal phosphate (PLP) cofactor; (f) 
about 0.1 to about 0.5 M of triethanolamine; (g) pH of about 6 to 8; and (h) temperature 
of about 45 to 55°C.

In some embodiments, additional reaction components or additional techniques carried 
out to supplement the reaction conditions. These can include taking measures to 
stabilize or prevent inactivation of the enzyme, reduce product inhibition, shift reaction 
equilibrium to product amine, formation.

Accordingly, in some embodiments of the process for preparing an amine, such as a 
chiral amine, additional quantities of the amino acceptor can be added (up to saturation) 
and/or the amino acceptor (ketone) formed can be continuously removed from the 
reaction mixture. For example, a solvent bridge or a two phase co-solvent system can be 
used to move the amine product to an extraction solution, and thereby reduce inhibition

In some embodiments of the processes, the suitable reaction conditions comprise the presence of the reduced cofactor, nicotinamide adenine dinucleotide (NADH), which can act to limit the inactivation of the transaminase enzyme (see e.g., van Ophem et al., 1998, Biochemistry 37(9):2879-88). In such embodiments where NADH is present, a cofactor regeneration system, such as glucose dehydrogenase (GDH) and glucose or formate dehydrogenase and formate can be used to regenerate the NADH in the reaction medium.

In some embodiments, the process can further comprise removal of the carbonyl by-product formed from the amino group donor when the amino group is transferred to the amino group acceptor. Such removal in situ can reduce the rate of the reverse reaction such that the forward reaction dominates and more substrate is then converted to product. Removal of the carbonyl by-product can be carried in a number of ways. Where the amino group donor is an amino acid, such as alanine, the carbonyl by product, a keto acid, can be removed by reaction with a peroxide (see, e.g., US 2008/0213845, incorporated herein by reference). Peroxides which can be used include, among others, hydrogen peroxide; peroxyacids (peracids) such as peracetic acid (CH₃C0₂H), trifluoroperacetic acid and metachloroperoxybenzoic acid; organic peroxides such as t-butyl peroxide ((CH₃)₂C0₂H), or other selective oxidants such as tetrapropylammonium perruthenate, MnO₂, KMnO₄, ruthenium tetroxide and related compounds. Alternatively, pyruvate removal can be achieved via its reduction to lactate by employing lactate dehydrogenase to shift equilibrium to the product amine (see, e.g., Koszelewski et al., 2008, Adv. Syn. Catal. 350: 2761-2766). Pyruvate removal can also be achieved via its decarboxylation by employing pyruvate decarboxylase (see, e.g., Hohne et al., 2008, Chem BioChem 9: 363-365) or acetolactate synthase (see, Yun and Kim, supra).

Alternatively, in embodiments where an amino acid is used as amino group donor, the keto acid carbonyl by-product can be recycled back to the amino acid by reaction with ammonia and NADH using an appropriate dehydrogenase enzyme, e.g., amino acid dehydrogenase, in presence of an amine donor, such as ammonia, thereby replenishing the amino group donor.

In some embodiments, where the choice of the amino donor results in a carbonyl by-product that has a vapor pressure higher than water (e.g., a low boiling co-product such
as a volatile organic carbonyl compound), the carbonyl by-product can be removed by sparging the reaction solution with a non-reactive gas or by applying a vacuum to lower the reaction pressure and removing the carbonyl by-product present in the gas phase. A non-reactive gas is any gas that does not react with the reaction components. Various non-reactive gases include nitrogen and noble gases (e.g., inert gases). In some embodiments, the non-reactive gas is nitrogen gas. In some embodiments, the amino donor used in the process is isopropylamine (IPM), which forms the carbonyl by-product acetone upon transfer of the amino group to the amino group acceptor. The acetone can be removed by sparging with nitrogen gas or applying a vacuum to the reaction solution and removing the acetone from the gas phase by an acetone trap, such as a condenser or other cold trap. Alternatively, the acetone can be removed by reduction to isopropanol using a transaminase.

In some embodiments of the processes above where the carbonyl by-product is removed, the corresponding amino group donor can be added during the transamination reaction to replenish the amino group donor and/or maintain the pH of the reaction. Replenishing the amino group donor also shifts the equilibrium towards product formation, thereby increasing the conversion of substrate to product. Thus, in some embodiments wherein the amino group donor is isopropylamine and the acetone product is removed in situ, isopropylamine can be added to the solution to replenish the amino group donor lost during the acetone removal and to maintain the pH of the reaction (e.g., at about 8.5).

In further embodiments, any of the above described process for the conversion of substrate compound to product compound can further comprise one or more steps selected from: extraction; isolation; purification; and crystallization of product compound. Methods, techniques, and protocols for extracting, isolating, purifying, and/or crystallizing the product amine from biocatalytic reaction mixtures produced by the above disclosed methods are known to the ordinary artisan and/or accessed through routine experimentation. Additionally, illustrative methods are provided in the Examples below.

Section J: Examples

The following Examples serve to illustrate the invention without limiting the scope thereof, while they on the other hand represent preferred embodiments of the reaction steps, intermediates and/or the process of the present invention.
Abbreviations:

δ  chemical shift
µl  microlitre
µm  micrometre
aq  aqueous
(+)CSA  (+) Camphorsulphonic Acid
ESTP  Ethyl Acetate
IPA  Isopropyl Acetate
TRMA  Treithyl Amine
Ac  acetyl
AcOH  acetic acid
br  broad
brm  broad multiplet
br. m.  broad multiplet
br. mult.  broad multiplet
br. s.  broad signal
br. sign.  broad signal
cat.  catalytic amount
compl. m  complex multiplet
compl. mult.  complex multiplet
copl. m.  complex multiplet
cpl. m.  complex multiplet
CDCl₃  deuterated chloroform
CH₂Cl₂  dichloromethane
CH₂O  formaldehyde
CO₂  carbon dioxide
d  doublet
dd  doublet of doublet
de  diastereomeric excess
dr  diastereomeric ratio
DCM  dichloromethane
DEA  diethyl amine
DMSO  dimethylsulfoxide
DMSO-d₆  dimethylsulfoxide deuterated
ee enantiomeric excess
equiv equivalent
ES electrospray
ES+ positive electrospray ionisation
ESI electrospray ionisation
Et ethyl
Et₂O diethyl ether
EtOAc ethyl acetate
EtOH ethanol
FTIR fourier transform infrared
g gram(s)
GC gass chromatography
h hour(s)
HCl hydrochloric acid
HCl(aq) hydrogen chloride aqueous solution
HNMR proton nuclear magnetic resonance
¹HNMR proton nuclear magnetic resonance
HPLC high performance liquid chromatography
H₃PO₄ phosphoric acid
HRMS high resolution mass spectroscopy
H₂SO₄ sulfuric acid
Hz hertz
iPr isopropyl
iPr₂NEt N-ethyldiisopropylamine
iPrOAc isopropyl acetate
iPrOH isopropanol
IR infrared
J coupling constant
K₂CO₃ potassium carbonate
L litre
LC-MS liquid chromatography-mass spectrometry
LCMS liquid chromatography-mass spectrometry
LRMS low resolution mass spectroscopy
m multiplet
Example 1: Synthesis of Ketone Intermediate

6-Chloro-5-fluoro-3-hydroxy-3-(2-oxo-propyl)-1,3-dihydro-indol-2-one

To the suspension of 6-Chloro-5-fluoro-isatin (25 g, 125 mmol) in acetone (620 ml), Et₂NH (0.9 g, 12.5 mmol) and K₂CO₃ (1.7 g, 12.5 mmol) were added. The reaction mixture was heated to reflux for 2 hours. After cooled down to the r.t., the reaction mixture was filtrated and the solvent was removed under reduced pressure. The crude
product was re-dissolved in MeOH (300 ml) at 50 °C and water (300 ml) was added. With addition of water, the solid was precipitate. MeOH (250 ml) was removed under reduce pressure at 50 °C; The mixture was cooled down to 0-5 °C and stirred for 30mins; The solid was filtered and dried to produce 6-Chloro-5-fluoro-3-hydroxy-3-(2-oxo-propyl)-1,3-dihydro-indol-2-one (27g).  

\[ ^1H \text{ NMR (DMSO-}d_6\text{): 2.00(3H, s, CH}_3\text{), 3.09(1H, d, CH, J =16Hz), 3.39(1H, d, CH, J =16Hz), 6.16(1H, s, OH); 6.88(1H, d, CH, J =8Hz), 7.36(1H, d, CH, J =12Hz), 10.37(1H, s, NH). MS (ESI) m/z 258.0 (M+H)\] 

HPLC method : Column: Agilent Extend-C18; 150x3.0mm; 3.5 μm. Mobile Phase A (0.1 % H3PO4) in water; Mobile Phase B (Acetonitrile). Gradient: 0 min (95 % A); 0.5 min (95% A), 14 min (5% A), 19.5 min (5% A),. Flow rate: 0.5 ml min-1 . Wavelength: 210 nm. Temperature: 40 °C. Retention time: 7.52 min.

**6-Chloro-5-fluoro-3-hydroxy-3-(2-methyl-[1,3]dioxolan-2-ylmethyl)-1,3-dihydro-indol-2-one**

![Chemical structure](image)

To the suspension of 6-chloro-5-fluoro-3-hydroxy-3-(2-oxopropyl)indolin-2-one (25g, 97 mmol) in ethylene glycol (250 mL), triethyl orthoformate (43g, 291mmol) and toluene-4-sulfonic acid monohydrate (0.9g, 4.8mmol) were added. The reaction mixture was heated at 40-50 °C for 2 hours. MeTHF (500 mL) and 10% NaCl aqueous solution (500 mL) were successively added. The organic layer was separated and washed with water twice. After the solvent was evaporated under vacuum, the white solid was washed with TBME (400 mL) and dried to produce 6-Chloro-5-fluoro-3-hydroxy-3-(2-methyl-[1,3]dioxolan-2-ylmethyl)-1,3-dihydro-indol-2-one (27.8 g).  

\[ ^1H \text{ NMR (DMSO-}d_6\text{): 1.00(3H, s, CH}_3\text{), 2.37(2H, d, CH}_2\text{), 3.13(1H, q, CHH), 3.58(1H, q, CHH), 3.59(1H, q, CHH), 3.68(1H, q, CHH), 6.00(1H, s, OH); 6.85(1H, d, CH, J =8Hz), 7.31(1H, d, CH, J =12Hz), 10.24(1H, s, NH). MS (ESI) m/z 302.0 (M+H)\] 

HPLC method :Column: Agilent Extend-C18; 150x3.0mm; 3.5 pm. Mobile Phase A (0.1 % H3PO4) in water; Mobile Phase B (Acetonitrile). Gradient: 0 min (95 % A); 0.5 min (95% A), 14 min (5% A), 19.5
min (5% A). Flow rate: 0.5 ml min⁻¹. Wavelength: 210 nm. Temperature: 40 °C. Retention time: 7.94 min.

1-(6-Chloro-5-fluoro-1H-indol-3-yl)-propan-2-one

To the solution of 6-chloro-5-fluoro-3-hydroxy-3-((2-methyl-1,3-dioxolan-2-yl)methyl)indolin-2-one (27.8g, 92mmol) in THF (250 mL), Red-Al (79.8g, 276mmol) in THF (100 mL) was added slowly at 60 °C under N₂ atmosphere. After the addition, the mixture was stirring at r.t. for 3 hours. 20 % HCl (400 mL) was added and the mixture was stirred for 20mins. Ethyl acetate (500 mL) was added and stirred for 10mins. The organic layer was separated and washed with water (100ml x 2); After treated with Na₂SO₄ (50 g)and active carbon (5 g) at r.t., the mixture was filtered and the filtration was concentrated. The crude product was recrystallized with MeOH (80 mL)and water (60 mL) to produce 1-(6-Chloro-5-fluoro-1H-indol-3-yl)-propan-2-one (15.5 g).

1H NMR (DMSO-d₆): 2.21 (3H, s, CH₃), 3.78(2H, s, CH₂), 7.13(1H, d, CH, J = 4Hz); 7.23(1H, d, CH, J =8Hz), 7.33(1H, d, CH, J =8Hz), 8.23(1H, s, NH).

MS (ESI) m/z 226.0 (M+H)⁺

HPLC method :Column: Agilent Extend-C18; 150x3.0mm; 3.5 μm. Mobile Phase A (0.1 % H₃PO₄) in water; Mobile Phase B (Acetonitrile). Gradient: 0 min (95 % A); 0.5 min (95% A), 14 min (5% A), 19.5 min (5% A). Flow rate: 0.5 ml min⁻¹. Wavelength: 210 nm. Temperature: 40 °C Retention time: 10.75 min.

Example 2: Second Alternative Synthesis of Ketone Starting Material

1-(6-Chloro-5-fluoro-1H-indol-3-yl)-propan-2-one
Under \( \text{N}_2 \), \( \text{Pt/C} \) (10 g, 5% Pt on Carbon) was charged to the suspension of 6-Chloro-5-fluoro-3-((Z)-2-nitro-propenyl)-1H-indole (100 g, 0.39 mol) in ethyl acetate (500 mL). The reaction mixture vacuumed and purged with \( \text{H}_2 \) three times and stirred at r.t. for overnight. The catalyst was filtered off and the organic solvent was removed under reduced pressure to give brown oil. The oil was re-dissolved in EtOH (500 mL) and sodium bisulfite (81.7 g, 0.79 mol) in \( \text{H}_2\text{O} \) (1) was added. The suspension was heated up to 80 °C for overnight. The precipitate was filtered off and ethanol was removed under pressure. The residue was diluted with water (500 mL) and extracted with ethyl acetate (300 mL X2). The combined organic layer was dried over \( \text{Na}_2\text{SO}_4 \) and condensed on the rota vapor to give brown oil (53 g, 60% yield).

**Example 3: Second Alternative Synthesis of Ketone Intermediate**

1-(6-Chloro-5-fluoro-1H-indol-3-yl)-propan-2-one

Raney/Ni (0.75g) and acetic acid (1.1 mL, 19.6 mmol) were charged to the suspension of 6-Chloro-5-fluoro-3-((Z)-2-nitro-propenyl)-1H-indole (5g, 19.6 mmol) in the mixture of MeOH and \( \text{H}_2\text{O} \) (2:1, 50 mL) at r.t. The reaction mixture vacuumed and purged with \( \text{H}_2 \) three times and heated at 50 °C for overnight. \( \text{MeOH} \) (50 mL) was added to the reaction mixture and the catalyst was filtered off over celite. The solvent was concentrated to 50 mL and \( \text{H}_2\text{O} \) (300 ml) was added. With the addition of \( \text{H}_2\text{O} \), off-white solid was produced. The solid was filtered and washed by water. After dried in the oven, 2.53 g off-white solid was isolated.
Example 4: Synthesis, Optimization, and Screening Engineered Transaminase Polypeptides

Gene synthesis and optimization: The polynucleotide sequence encoding the reported wild-type omega transaminase polypeptide from *Vibrio fluvialis* of SEQ ID NO: 2 was codon optimized and synthesized as the gene of SEQ ID NO: 1. The synthetic gene of SEQ ID NO: 1 was cloned into a pCK1 10900 vector system (see e.g., US Patent Application Publication 20060195947, which is hereby incorporated by reference herein) and subsequently expressed in *E. coli* W3110. The *E. coli* W3110 expresses the transaminase polypeptides as an intracellular protein under the control of the lac promoter. The polynucleotide (SEQ ID NO:3) encoding the engineered transaminase polypeptide of SEQ ID NO: 4 was obtained by directed evolution of the codon-optimized gene SEQ ID NO:1. The polypeptide of SEQ ID NO:4 has 24 amino acid residue differences (A9T; G18A; D21H; V31M; N45H; F86Y; A133R; R146L; W147K; V153S; K163F; V177L; R211K; P233T; A235P; P244T; M294V; P318D; A323T; S324G; A383V; T391A; C424A; F427Y) relative to SEQ ID NO:2. This synthetic gene SEQ ID NO: 3 (encoding the polypeptide of SEQ ID NO: 4) was used as the starting backbone for further optimization using standard methods of directed evolution via iterative variant library generation by gene synthesis followed by screening and sequencing of the hits to generate genes encoding engineered transaminases capable of converting compound (IA) to compound (MA) with improved enzyme properties relative to the polypeptides SEQ ID NOs: 2 and 4. The resulting engineered transaminase polypeptide sequences and specific mutations and relative activities are listed in Table 2A.

Example 5: Production of Engineered Transaminases

The engineered transaminase polypeptides were produced in *E. coli* W3110 as an intracellular protein expressed under the control of the lac promoter. The polypeptide accumulates primarily as a soluble cytosolic active enzyme. A shake-flask procedure is used to generate engineered polypeptide powders that can be used in activity assays or biocatalytic process disclosed herein.

High-throughput growth & expression. Cells were picked and grown overnight in LB media containing 1% glucose and 30 µg/mL chloramphenicol (CAM), 30°C, 200 rpm, 85% humidity. 20 µL of overnight growth were transferred to a deep well plate containing 380 µL 2xYT growth media containing 30 µg/mL CAM, 1 mM IPTG, 1 mM MgSO4, and incubated for ~18 h at 30°C, 200 rpm, 85% humidity. Subculture TB media was made up
of TB media (380 uL/well), 30 ug/mL CAM, 1mM MgS04, and 1mM IPTG. Cell cultures were centrifuged at 4000 rpm, 4°C for 10 min., and the media discarded. Cell pellets were resuspended in 200 or 400 µL. Lysis Buffer (0.1 M triethanolamine (TEA) buffer, pH 9.0, containing 1 mM MgS0₄, 400 µg/mL PMBS and 500 µg/mL Lysozyme), as described below.

Production of shake flask powders (SFP): A shake-flask procedure was used to generate engineered transaminase polypeptide powders used in secondary screening assays or in the biocatalytic process disclosed herein. Shake flask powder (SFP) includes approximately 30% total protein and accordingly provide a more purified preparation of an engineered enzyme as compared to the cell lysate used in HTP assays. A single microbial colony of E. coli containing a plasmid encoding an engineered transaminase of interest is inoculated into 50 mL Luria Bertani broth containing 30 µg/ml chloramphenicol and 1% glucose. Cells are grown overnight (at least 16 hours) in an incubator at 30°C with shaking at 250 rpm. The culture is diluted into 250 mL Terrific Broth (12 g/L bacto-tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 65 mM potassium phosphate, pH 7.0, 1 mM MgS0₄) containing 30 µg/ml chloramphenicol, in a 1 liter flask to an optical density at 600 nm (OD600) of 0.2 and allowed to grow at 30°C. Expression of the transaminase gene is induced by addition of isopropyl-β-D-thiogalactoside ("IPTG") to a final concentration of 1 mM when the OD600 of the culture is 0.6 to 0.8 and incubation is then continued overnight (at least 16 hours). Cells are harvested by centrifugation (5000 rpm, 15 min, 4°C) and the supernatant discarded. The cell pellet is resuspended with an equal volume of cold (4°C) 100 mM triethanolamine (chloride) buffer, pH 7.0 (optionally including 2 mM MgS0₄), and harvested by centrifugation as above. The washed cells are resuspended in two volumes of the cold triethanolamine (chloride) buffer and passed through a French Press twice at 12,000 psi while maintained at 4°C. Cell debris is removed by centrifugation (9000 rpm, 45 minutes, 4°C). The clear lysate supernatant was collected and stored at -20°C. Lyophilization of frozen clear lysate provides a dry shake-flask powder of crude transaminase polypeptide. Alternatively, the cell pellet (before or after washing) can be stored at 4°C or -80°C.

Production of downstream process (DSP) powders: DSP powders contains approximately 80% total protein and accordingly provide a more purified preparation of the engineered transaminase enzyme as compared to the cell lysate used in the
high throughput assay. Larger-scale (-100 - 120 g) fermentation of the engineered transaminase for production of DSP powders can be carried out as a short batch followed by a fed batch process according to standard bioprocess methods. Briefly, transaminase expression is induced by addition of IPTG to a final concentration of 1 mM. Following fermentation, the cells are harvested and resuspended in 100 mM Triethanolamine-H$_2$SO$_4$ buffer, then mechanically disrupted by homogenization. The cell debris and nucleic acid are flocculated with polyethylenimine (PEI) and the suspension clarified by centrifugation. The resulting clear supernatant is concentrated using a tangential cross-flow ultrafiltration membrane to remove salts and water. The concentrated and partially purified enzyme concentrate can then be dried in a lyophilizer and packaged (e.g., in polyethylene containers).

**Example 6: Analytical Procedures**

**HPLC Analysis of HTP Reactions:** An aliquot of the quenched reaction was subject to HPLC analysis under the following conditions.

<table>
<thead>
<tr>
<th>Column</th>
<th>Mightysil RP-18 GP Aqua 150x4.6 mm, 5 μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>30°C</td>
</tr>
<tr>
<td>Mobile Phase</td>
<td>Isocratic, 60% 20 mM NH$_4$Ac (pH 6.6)/40% acetonitrile</td>
</tr>
<tr>
<td>Flow Rate</td>
<td>2.5 mL/min</td>
</tr>
<tr>
<td>Detection</td>
<td>254 nm</td>
</tr>
<tr>
<td>Retention Times</td>
<td>Amine Product compound (IIA): 0.96 min; Substrate Impurity: 1.8 min; Ketone Substrate compound (IA): 2.7 min</td>
</tr>
</tbody>
</table>

Conversion of compound (IA) to compound (IIA) was determined from the resulting chromatograms as follows:

\[
\text{Conversion (\%)} = \frac{\text{Product Area}}{\text{Product Area} + \text{Substrate Area} \times 0.73} \times 100\%
\]

**HPLC Analysis of 5 mL and 100 mL scale reactions:** Aliquots of the quenched reaction was subject to HPLC analysis under the following conditions.

<table>
<thead>
<tr>
<th>Column</th>
<th>Mightysil RP-18 GP Aqua 250x4.6 mm, 5 μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>30°C</td>
</tr>
<tr>
<td>Mobile Phase</td>
<td>Isocratic: 60% 20 mM NH$_4$Ac (pH 6.6)/40% acetonitrile</td>
</tr>
<tr>
<td>Flow Rate</td>
<td>2.0 mL/min</td>
</tr>
<tr>
<td>Detection</td>
<td>254 nm</td>
</tr>
</tbody>
</table>
Retention Times

<table>
<thead>
<tr>
<th></th>
<th>Amine Product compound (IIA): 2.1 min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ketone Substrate compound (IA): 7.6 min</td>
</tr>
</tbody>
</table>

Conversion of compound (IA) to compound (IIA) was determined from the chromatograms as follows:

Conversion (%) = Product Area / (Product Area + Substrate Area x 0.73) x 100%

Determination of product chiral purity (%ee): The chiral purity or enantiomeric excess of compound (IIA) was assessed by HPLC using the following conditions.

<table>
<thead>
<tr>
<th>Column</th>
<th>Astec Chirobiotic TAG column</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>15°C</td>
</tr>
<tr>
<td>Mobile Phase</td>
<td>Methanol/Acetic acid/Triethylamine (100/0.2/0.1)</td>
</tr>
<tr>
<td>Flow Rate</td>
<td>1.0 mL/min</td>
</tr>
<tr>
<td>Detection Wavelength</td>
<td>225 nm</td>
</tr>
<tr>
<td>Retention Times</td>
<td>Ketone substrate: 3.6 min</td>
</tr>
<tr>
<td></td>
<td>R-product: 17.9 min;</td>
</tr>
<tr>
<td></td>
<td>S-product: 18.9 min</td>
</tr>
</tbody>
</table>

Determination of product purity: The purity of product was determined by HPLC using the following conditions.

<table>
<thead>
<tr>
<th>Column</th>
<th>MightySil Rp-18 GP aqua 250 x 4.6 mm, 5 μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>30 °C</td>
</tr>
<tr>
<td>Mobile Phase</td>
<td>Gradient. A: 20 mM NH₄Ac (pH 6.6); B: acetonitrile</td>
</tr>
<tr>
<td>Time</td>
<td>Composition</td>
</tr>
<tr>
<td>0 min</td>
<td>5% B</td>
</tr>
<tr>
<td>1.5 min</td>
<td>5% B</td>
</tr>
<tr>
<td>35 min</td>
<td>70% B</td>
</tr>
<tr>
<td>40 min</td>
<td>70% B</td>
</tr>
<tr>
<td>45 min</td>
<td>5% B</td>
</tr>
<tr>
<td>60 min</td>
<td>5% B</td>
</tr>
<tr>
<td>Flow Rate</td>
<td>1.5 mL/min</td>
</tr>
<tr>
<td>Detection Wavelength</td>
<td>254 nm</td>
</tr>
<tr>
<td>Retention Times</td>
<td>Major impurity: 22.2 min; amine product: 17.5 min; ketone substrate: 25.4 min</td>
</tr>
</tbody>
</table>

Example 7: High Throughput (HTP) Screening of Transaminases for Conversion of Compound (IA) to Compound (IIA)
HTP Screening Assays: High-throughput screening used to guide primary selection of variants was carried out in 96-well plates using cell lysates under assay conditions of 10 g/L compound (IA); 1 mM pyridoxal phosphate (PLP); 2 M isopropylamine (IPM), pH 7.0; 0.1 M triethanolamine (TEA), pH 7.0; 5% v/v (or 10%) PEG 200; 2 g/L lysozyme; 200 mg/mL polymyxin B sulfate; 100 mM TEA, pH 7.0). The mixture was then stirred and heated to 50°C.

Cells were grown in 96-well plates as described above and lysates prepared by dispensing 200 uL (for Round 1) or 400 uL (for Round 2) of Lysis Buffer (1 mg/mL lysozyme, 0.5 mg/mL polymyxin B sulfate, 1 mM PLP, 0.1 M triethanolamine (TEA), pH 7.0) into each well. Plates were sealed, shaken for 2 h, and then centrifuged for 20 min at 4,000 rpm at 4°C to pellet the cell debris.

A 10 uL of stock substrate solution (200 g/L compound (IIA) dissolved in PEG200) was added to each well of a 96-well plate followed by 180 uL of a stock solution of isopropylamine (IPM)/pyridoxal phosphate (PLP) (2.2 M IPM and 1.06 mM PLP in 100 mM TEA, pH 7.0. For assessing refractoriness to product compound (IIA) inhibition, compound (IIA) was added to the reaction mixture to a final 14 g/L for Round 1 and 16 g/L Round 2 assays. Reactions were initiated by adding 10 uL of cell lysate/well. Plates were sealed and incubated with shaking at 50 or 55°C for 24 h. Reactions were quenched with 600 uL of acetonitrile and samples examined by HPLC as described in Example 5.

Example 8: Process for Conversion of Compound (IA) to Compound (IIA) in 5 mL

Scale

A 5 ml scale reactions were carried out as follows. To a 20 mL glass vial with equipped a cross-shaped magnetic stirring bar was added 0.75 mL (or 0.5 mL in case of 10 % v/v PEG 200 concentration) of 100 mM TEA buffer (pH 7.0). 2 mL of 0.5% 5 M IPM-HCl stock solution was added to the vial followed by 1 mL of 5 mM PLP stock solution. The mixture was stirred at 500 rpm (magnetic stirring). The pH of the mixture was then adjusted to 7.0 using 1 M NaOH solution. 125 mg (or 250 mg for 50 g/L or 500 mg for 100 g/L concentration) substrate (solid) was then added to the vial. 0.25 mL (or 0.5 mL for 10% v/v) PEG 200 was then added to the mixture. Final concentrations of components were: 25 g/L (or 50 or 100 g/L) of compound (IA); 1 mM PLP; 2 M IPM; 5% v/v (or 10%) PEG 200; 2 g/L TA enzyme; 100 mM TEA, pH 7.0). The mixture was then stirred and heated to 50°C.
Reactions were initiated by adding 1 mL of the enzyme stock solution (10 g/L). 20 uL samples were taken at different time points and diluted with 750 uL methanol and analyzed by HPLC. After 24 h, the reaction mixtures were quenched with 5 mL acetonitrile and samples analyzed by HPLC to get the final %conversion.

Example 9: Process for Conversion of Compound (IA) to Compound (IIA) in 100 mL Scale

A 100 mL scale reaction was carried out as follows. To a 250 mL round bottom flask with an anchor-shaped stirring blade was added 15 mL of 100 mM TEA buffer (pH 7). 40 mL of the 5 M IPM'HCl stock solution was added to the round bottom flask followed by 20 mL of the 5 mM PLP stock solution. The mixture was stirred at 100 rpm (overhead stirring) and the pH adjusted to 7.0 using 10 M NaOH. Solid substrate Compound (IA) was then added to the round bottom flask over ~5 min with stirring. PEG 200 (5 mL) was then added and the mixture heated to 50°C. 20 mL of the enzyme stock solution (10 g/L) was then added to start the reaction. The reaction was stirred at 200 rpm (overhead stirring). 20 uL samples were taken at different time points and diluted with 750 uL methanol and analyzed by HPLC. In some cases, where work up was not pursued, after 24 h, the reaction mixtures were quenched with 100 mL acetonitrile and samples analyzed by HPLC to get the final %conversion (Figure 9). Concentration of different components in the reaction: ketone: 25 g/L (or 50 or 100 g/L); PLP: 1 mM; IPM: 2 M; PEG200: 5 % v/v; TA enzyme:2 g/L; buffer: 100 mM TEA, pH 7.0. Conversion of substrate to product was analyzed by HPLC as described in Example 3.

Product Workup: After 24 h reaction under the conditions described above (up to point 11) the reaction mixture was cooled to room temperature. The mixture was filtered through a standard G4 sintered glass funnel with a piece of filter paper (Whatman 1, pore size 11 μm). The round bottom flask was rinsed with 20 mL deionized water which was then filtered through the same funnel. The filter cake was washed twice with 20 mL deionized water. The pH of the filtrate was adjusted from 6.8 to 3 using a 5 M HCl solution. The filtrate was then transferred into a separatory funnel and extracted with 100 mL MTBE. The biphasic mixture was allowed to separate. The MTBE layer containing unreacted substrate and impurities was discarded. The aqueous layer was transferred into a beaker and 100 mL MTBE was added. The pH of the aqueous layer was then adjusted from 3 to 10 using 10 M NaOH solution. The mixture was transferred into a separatory funnel and the phases allowed to separate.
The aqueous layer was then extracted with 100 ml MTBE three times (nb: repeated until product was not present in aqueous layer). The MTBE phases from the three extractions were combined and evaporated to dryness using a rotary evaporator. The crude product was further dried under vacuum for 48 h.

**Example 10:** Process for Conversion of Compound (IA) to Compound (IIA) in 30g Scale

\[
\begin{align*}
\text{F} & | \text{Cl} & | \text{H} & | \text{H} \\
\text{Cl} & | \text{H} & | \text{F} & | \text{Cl} \\
\end{align*}
\]

152.48g /so-propylamine hydrochloride and 0.204g pyridoxal phosphate monohydrate were dissolved in 495ml water while stirring. To this yellow clear solution a solution of 30.0g ketone in 85ml poly ethylene glycol (average mol weight 200) within 15 minutes. Upon addition the ketone precipitates as fine particles which are evenly distributed in the reaction media. To the suspension 180ml triethanolamine buffer (0.1 mol/l, pH 7) were added and the pH was adjusted to 7 by addition of aqueous sodium hydroxide solution (1 mol/l). The reaction mixture is heated to 50°C and a solution of 1.62g transaminase SEQ ID NO: 134 dissolved in 162ml triethanolamine buffer (0.1 mol/l, pH 7) is added. The reaction mixture is continuously kept at pH 7 by addition of 1 mol/l aqueous sodium hydroxide solution. The reaction mixture is stirred 24h at 50°C and a stream of Nitrogen is blown over the surface of the reaction mixture to strip off formed acetone. The reaction mixture is then cooled to 25°C and filtered over a bed of cellulose flock. The pH of the filtrate is adjusted to $\approx 1$ by addition of concentrated sulfuric acid. The acidified filtrated is extracted with 250 ml /so-Propyl acetate. The layers are separated and the pH of the aqueous phase is adjusted to $\approx 10$ by addition of concentrated aqueous sodium hydroxide.
solution. The basified aqueous phase is extracted with /so-propyl acetate. The layers are seperated and the organic phase is washed with 100 ml water. The organic phase is concentrated by distillation to 2/3 of its origin volume. In a second reactor 33.98g (+)-camphor sulfoinic acid is dissolved in 225 ml /so-propyl acetate upon refluxing and the concentrated organic phase is added within 10 minutes. After complete addition the formed thin suspension is cooled to 0°C within 2 hours and kept at 0°C for 15 hours. The precipitated amine-(+) camphor sulfonate salt is filtered, washed with 70 ml /so-propyl acetate and dried at 40°C in vacum yielding 51.57g of colourless crystals (84.5% yield t.q.)

Analytical Data

IR:

IR (cm⁻¹)=3296, 3061, 2962, 2635, 2531, 2078, 1741, 1625, 1577, 1518, 1461, 1415, 1392, 1375, 1324, 1302, 1280, 1256, 1226, 1170, 1126, 1096, 1041, 988, 966, 937, 868, 834, 814, 790, 766, 746, 719, 669, 615.

LC-MS (ESI +):

Ammonium ion: m/z =227 ([M+H]), 268 ([M+H+CH₃CN]), 453 ([2M+H]).

Camphorsulfonate ion: m/z =250 ([M+NH₄]), 482 ([2M+NH4]).

LC-MS (ESI -):

Camphorsulfonate ion: m/z=231 ([M-H]), 463 ([2M-H]).

¹H-NMR (DMSO-d₆, 400 MHz):

11.22 (br. s., 1H), 7.75 (br. s., 3H), 7.59 (d, J = 10.3 Hz, 1H), 7.54 (d, J = 6.5 Hz, 1H), 7.36 (d, J = 2.3 Hz, 1H), 3.37 - 3.50 (m, 1H), 2.98 (dd, J = 14.3, 5.8 Hz, 1H), 2.91 (d, J = 14.8 Hz, 1H), 2.31 (dd, J = 14.3, 8.0 Hz, 1H), 2.63 - 2.74 (m, 1H), 2.41 (d, J = 14.6 Hz, 1H), 2.24 (dt, J = 18.3, 3.8 Hz, 1H), 1.94 (t, J = 4.4 Hz, 1H), 1.86 (dt, J = 7.4, 3.6 Hz, 1H), 1.80 (d, J = 18.1 Hz, 1H), 1.23 - 1.35 (m, 2H), 1.15 (d, J = 6.3 Hz, 3H), 1.05 (s, 3H), 0.74 (s, 3H)

Free Amine (obtained by evaporating the iso-Propylacetate layer after extraction of the basified aqueous layer):

¹H NMR (400MHz, DMSO-d₆):
11.04 (br. s., 1H), 7.50 (d, J = 10.5 Hz, 1H), 7.48 (d, J = 6.5 Hz, 1H), 7.25 (s, 1H), 3.03 (sxt, J = 6.3 Hz, 1H), 2.61 (dd, J = 14.3, 6.5 Hz, 1H), 2.57 (dd, J = 14.1, 6.5 Hz, 1H), 1.36 (br. s., 2H), 0.96 (d, J = 6.3 Hz, 3H)

Example 1: Process for Conversion of Compound (HA) to Compound (IVB)

13.62 g 5-chloroisatin is suspended in 35 ml /iso-propanol and 2.3 g triethyl amine is added. The suspension is heated to reflux and a solution of 34.42 g amine-(+)-camphor sulfonate salt dissolved in 300 ml /iso-propanol is added within 50 minutes. The reaction mixture is stirred at reflux for 17 hours. The reaction mixture is cooled to 75°C and 17.4 g (+)-camphorsulfonic acid are added to the reaction mixture. Approximately 300 ml /iso-propanol are removed by vacuum distillation. Distilled off /iso-propanol is replaced by iso-propyl acetate and vacuum distillation is continued. This is distillation is repeated a second time. To the distillation residue 19 ml ethanol and 265 ml ethyl acetate is added and the mixture is heated to reflux. The mixture is cooled in ramps to 0°C for 24 hours. The beige to off white crystals are filtered off, washed with 3 portions (each 25 ml) precooled (0°C) ethylacetate and dried in vacuum yielding 40.3 g beige to off white crystals. (86.3% yield t.q.)

IR:
\[\text{IR:} \quad \nu (c m^{-1}) = 3229, 3115, 3078, 3052, 2971, 2890, 2841, 2772, 2722, 2675, 2605, 2434, 1741, 1718, 1621, 1606, 1483, 1460, 1408, 1391, 1372, 1336, 1307, 1277, 1267, 1238, 1202, 1184, 1162, 1149, 1128, 1067, 1036, 987, 973, 939, 919, 896, 871, 857, 843, 785, 771, 756, 717, 690, 678, 613.\]

LC-MS (ESI +):
Ammonium ion: m/z =390 ([M+H]), 431 ([M+H+CH\textsubscript{3}CN])
Camphorsulfonate ion: m/z =250 ([M+NH₄⁺]), 482 ([2M+NH₄⁺])

**LC-MS (ESI -):**

Camphorsulfonate ion: m/z=231 ([M-H]), 403 ([2M-H]), 622.54 (m/z=463([2M-H])

**1H NMR (DMSO-d₆, 600 MHz):**

11.49 (s, 1H), 11.23 (s, 1H), 10.29 - 10.83 (m, 1H), 9.78 - 10.31 (m, 1H), 7.55 - 7.60 (m, 2H), 7.52 (s, 1H), 7.40 (d, J = 6.2 Hz, 1H), 7.16 (d, J = 8.8 Hz, 1H), 4.52 - 4.63 (m, 1H).

3.20 (dd, J = 16.3, 4.2 Hz, 1H), 2.96 (dd, J = 16.1, 11.3 Hz, 1H), 2.90 (d, J = 15.0 Hz, 1H), 2.56 - 2.63 (m, 1H), 2.39 (d, J = 14.6 Hz, 1H), 2.21 (dt, J = 18.0, 3.8 Hz, 1H), 1.89 - 1.93 (m, 1H), 1.81 (ddd, J = 15.3, 7.8, 3.7 Hz, 1H), 1.76 (d, J = 18.3 Hz, 1H), 1.53 (d, J = 6.6 Hz, 3H), 1.20 - 1.33 (m, 2H), 0.98 (s, 3H), 0.70 (s, 3H)

**Example 12: Process for Preparing a Compound of formula (IVA) ½ Hydrate**

\[
\text{F} \quad \begin{array}{c}
\text{N} \\
\text{H} \\
\text{N} \\
\text{H} \\
\text{Cl} \\
\text{Cl} \\
\text{Cl} \\
\text{SO₃} \\
\text{H} \\
\end{array}
\quad \text{1) } \text{Na₂CO₃ (aq)} \\
\text{EtOH/H₂O} \quad \text{2) crystallization} \\
\text{F} \quad \begin{array}{c}
\text{N} \\
\text{H} \\
\text{N} \\
\text{H} \\
\text{Cl} \\
\text{Cl} \\
\text{Cl} \\
\text{SO₃} \\
\text{H} \\
\end{array}
\]

In a 750ml reactor with impeller stirrer 50g of compound (IVB) salt were dissolved in 300ml Ethanol (ALABD) and 100 ml deionised Water (WEM). The clear, yellowish solution was heated to 58°C internal temperature. To the solution 85 g of a 10% aqueous sodium carbonate solution was added within 10 minutes. The clear solution was particle filtered into a second reaction vessel. Vessel and particle filter were each rinsed with 25 ml of a mixture of ethanohwater (3:1 v/v) in the second reaction vessel. The combined particle filtered solution is heated to 58°C internal temperature and 200ml water (WEM) were added dropwise within 15 minutes. Towards the end of the addition the solution gets turbid. The mixture is stirred for 10 minutes at 58°C internal temperature and is then cooled slowly to room temperature within 4 hours 30 minutes forming a thick, well stirable white suspension. To the suspension 200 ml water are added and the mixture is stirred for additional 15 hours 20 minutes at room temperature. The suspension is filtered and the filter cake is washed twice with 25 ml portions of a mixture of ethanohwater 9:1 (v/v). The colourless crystals are dried at 60°C in vacuum yielding 26.23g (=91.2% yield).
$^1\text{H NMR (400 MHz, DMSO-d}_6\text{)}$

10.70 (s, 1H), 10.52 (s, 1H), 7.44 (d, J = 10.0 Hz, 1H), 7.33 (dd, J = 8.4, 2.1 Hz, 1H), 7.26 (d, J = 6.5 Hz, 1H), 7.05 (d, J = 2.3 Hz, 1H), 6.93 (d, J = 8.3 Hz, 1H), 3.83 - 4.00 (m, 1H), 3.13 (d, J = 6.0 Hz, 1H), 2.77 (dd, J = 15.1, 3.8 Hz, 1H), 2.38 (dd, J = 15.1, 10.5 Hz, 1H), 1.17 (d, J = 6.3 Hz, 3H).
Claims:

1. A process for preparing a compound of formula (II), or a salt or solvate or hydrate thereof,

\[ \text{Formula (II)} \]

comprising converting a compound of formula (I) to compound of formula (II), or a salt, solvate or hydrate thereof,

\[ \text{Formula (I)} \]

wherein:

- the dashed line is a bond or absent;
- A is selected from C=O and C=NH; or when the dashed line is a double bond A-
- \( R^8 \) is:

\[ R^a \text{ is } C_{1-6} \text{ alkyl: } \]
R₁ is H, -CH₃, OMe
R₄ and R₇ are each, independently, H or -Cl;
R₅ is H, -OH, -CH₃, -OCH₃, -F, -Cl, -CF₃ or -CN;
R₆ is H, -OH, -OCH₃, -F or -Cl;
R₈ is H, -CH₃, -CH₂CH₃, -CH₂OH, -CO₂H, -CO₂CH₃, -CO₂CH₂CH₃ or -CF₃; and
n is 1 or 2.

2. The process of claim 1, wherein R₅ and R₆ are fluoro when: R₈ is -CH₃, and n is 1.

3. The process of claim 1, wherein R₅ and R₆ are fluoro and chloro, when: R₈ is -CH₃, and n is 1.

4. The process of claim 1, wherein R₅ and R₆ are hydrogen when: R₈ is -CH₃ and n is 1.

5. The process of claim 1, wherein R₅ is fluoro when: n is 1, and R₆ is hydrogen.
6. The process of claim 1, wherein the compound of formula (II) is of formula (IIA), or a salt or solvate or hydrate thereof,

![Chemical structure of (IIA)](#)

7. The process of claim 1, wherein the compound of formula (II) is converted from the compound of formula (I) under a condition selected from enzymatic transamination, chemical asymmetric catalysis, asymmetric reduction and chiral resolution, or a combination of two or more conditions.

8. The process of claim 1, wherein A is C=0.

9. The process according to claim 1, wherein the compound of formula (I) is a compound of formula (IA) or a salt or hydrate or solvate thereof,

![Chemical structure of (IA)](#)

10. A process for preparing a compound of formula (IIA) or a salt or hydrate or solvate thereof,

![Chemical structure of (IIA)](#)
comprising enzymatically transaminating a compound of formula (IA) or a salt or solvate or hydrate thereof,

\[(\text{IA}).\]

\[\text{to provide the compound of formula (HA).}\]

11. A process according to any one of the preceding claims wherein the enzyme is SEQ ID NO: 134.

12. A process for preparing a compound of formula (IV), or a salt or hydrate or solvate thereof,

\[(\text{IV}).\]

comprising,

reacting a compound of formula (III), or a salt or hydrate or solvate thereof,

\[(\text{III}).\]

with a compound of formula (II), or a salt or hydrate or solvate thereof,
wherein the compound of formula (II), or a salt or hydrate or solvate thereof, is prepared from a compound of formula (I), or a salt or hydrate or solvate thereof,

wherein:

the dashed line is a bond or absent;

A is selected from C=0 and C=NH; or when the dashed line is a double bond A-

R^8 is:

R^8 is C_{1-6} alkyl;

R^1 is (C, C_6) alkyl, optionally substituted with an amino, (C_{1-6}) alkyl amino, (C_{1-6}) alkyl di-alkyl amino or (d-C_6) alkyl C(0)NH (C_{1-6}) alkyl;

R^4 and R^7 are each, independently, H or halo;

R^5 and R^6 are each, independently hydrogen, halo, hydroxyl, (C_{1-6}) alkyl, trihalo (C_{1-6}) alkyl, cyano or (C_{1-6}) alkoxy;
R₈ is (C₁-C₆) alkyl, optionally substituted with a hydroxyl:
n is 1 or 2;

R¹ is hydrogen or (C₁-C₆) alkyl; and

R⁴, R⁵, R⁶ and R⁷, are each, independently hydrogen, halo, hydroxy, amino, alkylamino, dialkylamino, (C₁-C₃) alkyl, and (C₁-C₆) alkoxy.

13. The process of claim 12, wherein R⁵ and R⁶ are fluoro when: R₈ is -CH₃, and n is 1.

14. The process of claim 12, wherein R⁵ and R⁶ are fluoro and chloro when: R₈ is -CH₃, and n is 1.

15. The process of claim 12, wherein R⁵ and R⁶ are hydrogen when: R₈ is -CH₃ and n is 1.

16. The process of claim 12, wherein R⁵ is fluoro when: n is 1, and R₆ is hydrogen.

17. The process of claim 12, wherein R¹ is H, -CH₃,
R is H, -CH₃, -CH₂CH₃, -CH₂OH, -CO₂H, -C₀₂CH₃, -C₀₂CH₂CH₃ or CF₃;
and n is 1 or 2.

18. The process of claim 12, wherein the compound of formula (II) is of formula (IIA), or a salt or solvate or hydrate thereof,

![Formula II](image)

(IIA).

19. The process of claim 12, wherein R¹ is a halo.

20. The process of claim 12, wherein R² is a chloro, and R³, R⁴, R⁵ and R⁶ are each hydrogen.

21. The process of claim 12, wherein the compound of formula (III) is of formula (IIIA), or a salt hydrate or solvate thereof.

![Formula III](image)

(IIIA).

22. A process for preparing a compound of formula (IVA), or a salt or solvate or hydrate thereof.
comprising
reacting a compound of formula (IIIA) or a salt or solvate or hydrate thereof,

with a compound of formula (II) or a salt or solvate or hydrate thereof,

to provide the compound of formula (IVA), or a salt or solvate or hydrate thereof,

wherein the compound of formula (HA) is prepared from the compound of formula (IA), or a salt or hydrate or solvate thereof,

23. The process according to any one of claims 12-22, wherein the enzyme is SEQ IS NO: 134.
24. The process of claim 22, wherein the compound of formula (HA) is converted to a salt of formula (IIB):

![Chemical structure](image)

(IIB)

prior to reaction with the compound of formula (MIA).

25. The process of claim 22, wherein the reaction is carried out under basic conditions.

26. The process of claim 22, wherein the reaction is carried out in the presence of triethyl amine.

27. The process according to claim 22, wherein the compound of formula (IVA) is isolated as a salt of formula (IVB):

![Chemical structure](image)

(IVB).

28. The process of claim 22, wherein the salt of formula (IVB) is converted to the compound of formula (IVA) in free base.

29. The process of claim 28, wherein the conversion is carried out with sodium carbonate.
30. The process of claim 22, wherein the compound of formula (IVA) is a hydrate.

31. The process of claim 22, wherein the compound of formula (IVA) is a $\frac{1}{2}$ hydrate.

32. The process of claim 22, wherein the compound of formula (IVA) is milled after isolation.

33. A compound of formula (IC):

![Chemical Structure](image)

wherein:

- $R^1$ is ($C_1$-$C_6$) alkyl, optionally substituted with an amino, ($C_1$-$C_6$) alkyl amino, ($C_1$-$C_6$) alkyl di-alkyl amino or ($C_1$-$C_6$) alkyl C(0)NH ($C_1$-$C_6$) alkyl;
- $R^5$ and $R^6$ are each, independently hydrogen, halo, hydroxy, ($C_1$-$C_6$) alkyl, trihalo ($C_1$) alkyl, cyano or ($C_1$-$C_6$) alkoxy;
- $R^8$ is ($C_1$-$C_6$) alkyl, optionally substituted with a hydroxyl:
  - $n$ is 1 or 2;
- or a pharmaceutically acceptable salt or hydrate or solvate thereof.

34. The compound of claim 33, wherein $R^5$ and $R^6$ are fluoro when: $R^8$ is -CH$_3$, and $n$ is 1.

35. The compound of claim 33, wherein $R^5$ and $R^6$ are fluoro and chloro, when: $R^8$ is -CH$_3$, and $n$ is 1.
36. The compound of claim 33, wherein $R^5$ and $R^9$ are hydrogen when: $R^8$ is -CH and $n$ is 1.

37. The compound of claim 33, wherein $R^5$ is fluoro when: $n$ is 1, and $R^8$ is hydrogen.

38. The compound of claim 33, wherein $R^1$ is H, -CH₃.

39. The compound of claim 33, wherein the compound is of formula (IA): 

or a pharmaceutically acceptable salt or hydrate or solvate thereof.
A compound selected from:

\[
\begin{align*}
&\text{or} \\
&\text{or a pharmaceutically acceptable salt, solvate or hydrate thereof.}
\end{align*}
\]
### A. CLASSIFICATION OF SUBJECT MATTER

INV. C07D405/06  C07D471/20  C07D209/12  C07D209/14  C07D209/34
C12P17/10

### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07D  C12P

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 , BIOSIS, CHEM ABS Data, EMBASE, WPI Data

### C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>wo 2009/132921 AI (NOVARTIS AG [CH]; ANG SHI HUA [SG]; KRASTEL PHI LI PP [CH]; LEONG SEH Y0) 5 November 2009 (2009-11-05) cited in the application</td>
<td>1-9 , 12-22 , 24-40</td>
</tr>
<tr>
<td>Y</td>
<td>Schemes A and C; examples 1,2,7-31,35</td>
<td>10,11,23</td>
</tr>
<tr>
<td>X</td>
<td>GB 974 895 A (PARKE DAVIS &amp; CO) 11 November 1964 (1964-11-11)</td>
<td>1-9 , 33-40</td>
</tr>
<tr>
<td>Y</td>
<td>Page 1, line 29 - line 35; Page 1, line 61 - page 2, line 13; page 1, line 29 - line 35; claims 1-7; example lb)</td>
<td>10-32</td>
</tr>
</tbody>
</table>

-/-/-

* 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) in which it 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

"A" document member of the same patent family

Date of the actual completion of the international search

31 May 2013

Date of mailing of the international search report

17/06/2013

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Fax (+31-70) 340-3016

Authorized officer

Kirsch, Cecile
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 3 282 959 A (ROBERT VANE JOHN ET AL) 1 November 1966 (1966-11-01)</td>
<td>1-9, .34,</td>
</tr>
<tr>
<td></td>
<td>Column 2, line 14 - line 24, line 45 - line 65; Column 6, line 64 - col umn 7,</td>
<td>35, .37,</td>
</tr>
<tr>
<td></td>
<td>line 6; examples 4-8</td>
<td>39, .40</td>
</tr>
<tr>
<td></td>
<td>Scheme 1; Table 1; compounds la, lc, ld</td>
<td>33-40</td>
</tr>
<tr>
<td>X</td>
<td>IWAO TABUSHI ET AL: &quot;A- (modified B6)-B-[omega.-ami no(ethyl ami no)] -beta. -c cyclodextrin as an artificial B6 enzyme for chiral amine transfer reaction&quot;, JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 107, no. 19, 1 September 1985 (1985-09-01), pages 5545-5546, XP055064105. ISSN: 0002-7863, DOI: 10.1021/ja00305a048</td>
<td>1-9, .34, .35, .37,</td>
</tr>
<tr>
<td></td>
<td>Equations 1; compounds 5b</td>
<td>10-32</td>
</tr>
<tr>
<td>X</td>
<td>US 4 073 795 A (BATCH0 ANDREW DAVID ET AL) 14 February 1978 (1978-02-14)</td>
<td>1-7</td>
</tr>
<tr>
<td>A</td>
<td>Column 3 - Column 6; Column 13 - Column 14; claims 1-11; examples 34,39; compounds XV, XVI, XIX, XX</td>
<td>8-40</td>
</tr>
<tr>
<td>A</td>
<td>page 5, line 7 - line 25; claims 7-8; example X</td>
<td>8-40</td>
</tr>
<tr>
<td>X</td>
<td>EP 0 135 846 A2 (GENETICS INST [US]) 3 April 1985 (1985-04-03)</td>
<td>1-11, .34, .35, .37,</td>
</tr>
<tr>
<td>Y</td>
<td>claims 1-10; example 7</td>
<td>12-32, .40</td>
</tr>
</tbody>
</table>

Form PCT/ISA/210 (continuation of second sheet) (April 2008)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>Scheme 1; table 1; compound 2e</td>
<td>12-32, 40</td>
</tr>
<tr>
<td>X</td>
<td>EP 0 120 208 A2 (DEGUSSA [DE]); BIOTECHNOLOG FÖRSCHUNG GMBH [DE]</td>
<td>1-9, 33-39</td>
</tr>
<tr>
<td>Y</td>
<td>claim 7; example 9; tables 4-5</td>
<td>10-32, 40</td>
</tr>
<tr>
<td>Y</td>
<td>table 1</td>
<td>10-32, 40</td>
</tr>
<tr>
<td>Y</td>
<td>AT 507 050 AI (ANGEWANDE BIOKATALYSE KOMPETE [AT]; UNIV GRAZ [AT]) 15 January 2010 (2010-01-15) claims; examples</td>
<td>1-40</td>
</tr>
<tr>
<td>Y</td>
<td>wo 2012/007548 A1 (DSM IP ASSETS BV [NL]; SCHUERMANN MARTIN [NL]; PEETERS WUHAND PETER H) 19 January 2012 (2012-01-19) claims; examples</td>
<td>1-40</td>
</tr>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>wo 2009132921</td>
<td>05-11-2009</td>
<td>AR 071222 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2009242279 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2719365 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 102015711 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CO 6300940 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EA 201001689 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EC SP10016646 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 2285808 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2011518852 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 20100135320 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NZ 588217 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PE 17872009 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TW 200948357 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2009275560 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2011275613 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wo 2009132921 Al</td>
</tr>
<tr>
<td>GB 974895</td>
<td>11-11-1964</td>
<td>NONE</td>
</tr>
<tr>
<td>US 3282959</td>
<td>01-11-1966</td>
<td>GB 974893 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 3282959 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BE 855904 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 27279671 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FR 2378759 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FR 2378760 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GB 1569341 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GB 1569342 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IT 1143671 B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP S521588865 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NL 7706906 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 40739795 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 44140697 A</td>
</tr>
<tr>
<td>EP 0036741</td>
<td>30-09-1981</td>
<td>CA 1182462 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 3163355 Di</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0036741 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE 51022 B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP S56167695 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 4331818 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 4397787 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 4440936 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0135846 A2</td>
</tr>
<tr>
<td>EP 0120208</td>
<td>03-10-1984</td>
<td>CA 1209933 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 3307095 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 142184 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0120208 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP S59198972 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 4590161 A</td>
</tr>
<tr>
<td>wo 2010081053</td>
<td>15-07-2010</td>
<td>CN 102341494 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 2385983 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG 172891 Al</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2010209981 Al</td>
</tr>
</tbody>
</table>

Form PCT/ISA/210 (patent family annex) (April 2005)
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO 2010081053 A2</td>
<td>15-07-2010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT 507050 A1</td>
<td>15-01-2010</td>
<td>NONE</td>
<td></td>
</tr>
<tr>
<td>WO 2012007548 A1</td>
<td>19-01-2012</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>