heterocyclic ring system containing one to two heteroatoms selected from O, S, and N, the fused ring system being unsubstituted or substituted with one to two substituents selected from hydroxy, amino, fluorine, C_{1-4} alkyl, C_{1-4} alkoxy, and trifluoromethyl.

[0015] In some embodiments, the engineered transaminase polypeptide are capable of biocatalytic conversion of compounds of Formula (**II**) to compounds of Formula (**I**) having the indicated stereochemical configuration at the stereogenic center marked with an *; in an enantiomeric excess of at least 70% over the opposite enantiomer.

[0016] In one aspect, there is provided ~~some embodiments, the present disclosure provides~~ an engineered polypeptide having transaminase activity comprising an amino acid sequence having at least 80% sequence identity to reference sequence of SEQ ID NO:2 and (a) an amino acid residue difference as compared to SEQ ID NO:2 selected from X33L, X36C, X41C/F/K/M/N/R, X42G, X48D/E/G/K/T, X51K, X54P, X76S, X122F/Q, X148Q, X152T, X155A/I/K/T/V, X156R, X160P, X215G/H/L, X241R, X270T, X273H, X325M; and X241R, ~~and/~~or (b) a combination of residue differences selected from: X42G, X54P, X152S, and X155T; X42G, X54P, X152S, X155T, and R164P; X42G, X54P, X150F, X152S, and X155T; X42G, X54P, X150F, X152S, X155T, and X267V; X42G, X54P, X150F, X152S, X155L, W156Q, and C215G; X42G, X54P, X150F, X152S, X155T, X215G, and X267V; X33L; X42G, X54P, X117G; X150F, X152S, X155I, X156Q, and C215G; and X41K, X42G, X54P, X150F, X152S, X155K, X156Q, and C215G; X33L, X42G, X54P, X109S, X150F, X152S, X155K, X156Q, and X215H; X33L, X42G, X54P, X150F, X152S, X155I, X156Q, and X215G; X33L, X42G, X54P, X150F, X152S, X155K, X156Q, and X215H; X33L, X42G, X54P, X150F, X152S, X155L, X156Q, and X215H; X33L, X42G, X54P, X150F, X152S, X155L, X156Q, X215H, and X241R; X41F, X42G, X54P, X122Q, X150F, X152T, X155V, X156Q, and X215G; X41F, X42G, X54P, X150F, X152S, X155L, X156Q, X171I, X215G, and X241R; X41F, X42G, X54P, X150F, X152S, X155I, X156Q, V171I, and X215G; X41F, X42G, X54P, X150F, X152S, X155I, X156Q, and X215G; X41F, X42G, X54P, X150F, X152S, X155L, X156Q, X171I, and X215G; X41F, X42G, X54P, X150F, X152S, X155L, X156Q, and X215G; X42G, X48G, X54P, X150F, X152S, X155L, X156Q, and X215H; X42G, X54P, X60V, X150F, X152S, X155L, X156Q, and X215G; X42G, X54P, X68A, X150F, X152S, X155L, X156Q, and X215G; X42G, X54P, X69S, X150F, X152S, X155L, X156Q, and X215G; X42G, X54P, X122Q, X150F, X152S, X155I, X156Q, X215G, and X241R; X42G, X54P, X122Q, X150F, X152S, X155L, X156Q, X171I, X215G, and X241R; X42G, X54P, X122Q, X150F, X152T, X155V, X156Q, X171I, X215G, and X241R; X42G, X54P, X126M, X150F, X152S, X155L, X156Q, and X215G; X42G, X54P, X135I, X136Y, X150F, X152S, X155L, X156Q, X192F, and X215G; X42G, X54P, X136I, X150F, X152S, X155L, X156Q, and X215G; X42G, X54P, X136I, X150F, X152S, X155L, X156Q, X215G, and X224I; X42G, X54P, X136I, X150F, X152S, X155L, X156Y, X215G, X282V, and X284I; X42G, X54P, X136I, X150F, X152S, X155L, X156Y, X215G, and X284P; X42G, X54P, X136Y, X150F, X152S, X155L,

X156Q, X215G, X282V, and X284P; X42G, X54P, X150F, X152S, X155I, X156Q, X171I, X215G, and X241R; X42G, X54P, X150F, X152S, X155L, X156Q, X193M, and X215G; X42G, X54P, X150F, X152S, X155L, X156Q, X215G, X282V, and X284I; X42G, X54P, X150F, X152S, X155L, X156Q, X215G, and X283S; X42G, X54P, X150F, X152S, X155L, X156Q, X215G, and X284I; and X42G, X54P, X150F, X152S, X155L, X156Y, and X215G.

[0017] In some embodiments of the engineered polypeptides having transaminase activity of the present disclosure, the amino acid sequence can further comprise one or more residue differences as compared to SEQ ID NO:2 selected from: X5K, X33L, X36C, X41C/F/K/M/N/R, X42A/G, X44Q, X48D/E/G/K/T, X49T, X51K, X54P, X55L, X76S, X108V, X117G, X122F/Q, X126A, X148Q, X150A/F, X152S/T, X155A/I/K/L/T/V, X156Q/R/S, X160P, X164P, X165N, X182T, X215G/H/L, X218M, X241R, X267V, X270T, X273H, X325M, and X328I.

[0018] In one aspect, there is provided ~~some embodiments, the present disclosure provides~~ an engineered polypeptide having transaminase activity comprising an amino acid sequence having at least 80% sequence identity to reference sequence of SEQ ID NO:2 and (a) an amino acid residue difference as compared to SEQ ID NO:2 selected from G36C, I41C, I41F, I41K, I41M, I41N, I41R, E42G, P48D, P48E, P48G, P48K, P48T, A51K, S54P, M122F, M122Q, Y148Q, C152T, Q155A, Q155I, Q155K, Q155T, Q155V, C215H, C215L, Y273H, L325M, and A241R; or (b) a combination of residue differences selected from: A5K, E42G, S49T, S54P, C152S, Q155T, and W156Q; P33L, I41C, E42G, S54P, S150F, C152S, Q155K, F160P, and C215G; P33L, I41K, E42G, S54P, S150F, C152S, Q155I, F160P, and C215L; P33L, E42G, P48G, S54P, S150F, C152S, Q155T, and C215H; P33L, E42G, S54P, A109S, S150F, C152S, Q155K, W156Q, and C215H; P33L, E42G, S54P, E117G, S150F, C152S, Q155I, W156Q, and C215G; P33L, E42G, S54P, S150F, C152S, Q155I, W156Q, and C215G; P33L, E42G, S54P, S150F, C152S, Q155K, W156Q, and C215H; P33L, E42G, S54P, S150F, C152S, Q155L, W156Q, and C215H; P33L, E42G, S54P, S150F, C152S, Q155L, W156Q, C215H, and A241R; G36C, E42G, P48G, S54P, S150F, C152S, Q155I. and C215H; G36C, E42G, P48K, S54P, S150F, C152S, Q155T, and C215H; G36C, E42G, S54P, S150F, C152S, Q155I, C215H, and A241R; G36C, E42G, S54P, S150F, C152S, Q155K, C215H, and A241R; G36C, E42G, S54P, S150F, C152S, Q155T, and A241R; G36C, E42G, S54P, S150F, C152S, Q155V, and C215H; I41C, E42G, S49T, S54P, S150F, C152S, Q155I, F160P, C215G, and I267V; I41C, E42G, S49T, S54P, S150F, C152S, Q155K, W156Q, C215G and I267V; I41C, E42G, S54P, I108V, S150F, C152S, and Q155K; I41C, E42G, S54P, I108V, S150F, C152S, Q155K, W156Q, C215G, and I267V; I41C, E42G, S54P, I108V, S150F, C152S, Q155T, W156Q, and C215G; I41C, E42G, S54P, E117G, S150F, C152S, Q155K, and F160P; I41C, E42G, S54P, E117G, S150F, C152S, Q155K, and C215L; I41C, E42G, S54P, E117G, S150F, C152S, Q155L, and C215L; I41C, E42G, S54P, S150F, C152S, Q155I, and C215G; I41C, E42G, S54P, S150F, C152S, Q155I, and C215L; I41C, E42G, S54P, S150F, C152S, Q155K, W156Q, C215G, and I267V; I41C, E42G, S54P, S150F, C152S, Q155K, and

196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274, 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 300, 302, 304, and 306. Each of these exemplary polypeptide sequences comprises a different combination of the amino acid differences relative to SEQ ID NO:2 as disclosed herein (See e.g., Tables 2A, 2B, and 2C). In some embodiments, the engineered polypeptide comprises a sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identity to any one of these exemplary sequences, and further comprising a combination of amino acid differences relative to SEQ ID NO:2, as found in any one of these exemplary amino acid sequences. In some embodiments, the engineered polypeptide comprising a combination of amino acid differences relative to SEQ ID NO:2, as found in any one of these exemplary amino acid sequences can further comprise additional amino acid differences as compared to SEQ ID NO:2 selected from: X5K, X33L, X36C, X41C/F/K/M/N/R, X42A/G, X44Q, X48D/E/G/K/T, X49T, X51K, X54P, X55L, X76S, X108V, X117G, X122F/Q, X126A, X148Q, X150A/F, X152S/T, X155A/I/K/L/T/V, X156Q/R/S, X160P, X164P, X165N, X182T, X215G/H/L, X218M, X241R, X267V, X270T, X273H, X325M, and X328I; or other amino acid differences disclosed in the art of engineered transaminase polypeptides (See e.g., amino acid differences disclosed in US Pat. No. 8,293,507 B2, issued Oct. 23, 2012; WO2011/005477A1, published Jan. 13, 2011; WO2012/024104, published Feb. 23, 2012.)

[0021] In some embodiments of the present disclosure, the engineered polypeptide having transaminase activity is immobilized on a solid support, optionally wherein the solid support is selected from a bead or resin comprising polymethacrylate with epoxide functional groups, polymethacrylate with amino epoxide functional groups, styrene/DVB copolymer or polymethacrylate with octadecyl functional groups.

[0021a] In one aspect, there is provided a polynucleotide encoding the engineered polypeptide as described herein.

[0022] In other embodiments ~~aspects~~, the present disclosure provides a polynucleotide encoding the engineered polypeptide having transaminase activity disclosed herein. In some embodiments, the polynucleotide can comprise a nucleotide sequence selected from SEQ ID NO:3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 375, 277, 279, 281, 283, 285, 287, 291, 293, 295, 297, 299, 301, 303, and 305.

[0022a] In one aspect, there is provided an expression vector comprising the polynucleotide as described herein.

[0023] Further, the present disclosure provides expression vectors and host cells comprising a polynucleotide encoding the engineered polypeptide having transaminase activity disclosed herein. Thus, in some embodiments, the present disclosure provides an expression vector comprising the

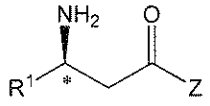
polynucleotide encoding an engineered polypeptide as disclosed herein, and optionally further comprising a control sequence.

[0023a] In one aspect, there is provided a host cell comprising the polynucleotide as described herein or the expression vector as described herein.

[0023b] In other embodiments, the present disclosure provides a host cell comprising a polynucleotide encoding an engineered polypeptide as disclosed herein. In other embodiments, the present disclosure provides a host cell comprising an expression vector, wherein the expression vector comprises a polynucleotide encoding an engineered polypeptide as disclosed herein.

[0023c] In one aspect, there is provided ~~other embodiments, of the present disclosure provides~~ a method of preparing an engineered polypeptide as described ~~disclosed~~ herein, wherein the method comprises culturing a host cell as described herein ~~of~~ under conditions suitable for expression of the polypeptide. In some embodiments, the method of preparing the engineered polypeptide further comprises isolating the polypeptide.

[0024] The present disclosure also provides processes for using the engineered transaminase polypeptides disclosed herein for the preparation of wide range of chiral amine compounds. In one aspect, there is provided ~~some embodiments, the present disclosure provides~~ a use ~~method~~ for preparing a compound of structural Formula **(I)**:



(I)

having the indicated stereochemical configuration at the stereogenic center marked with an *; in an enantiomeric excess of at least 70% over the opposite enantiomer, wherein

Z is OR^2 or NR^2R^3;

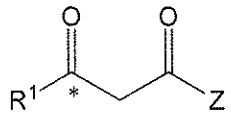
R^1 is C_{1-8} alkyl, aryl, heteroaryl, aryl-C_{1-2} alkyl, heteroaryl-C_{1-2} alkyl, or a 5- to 6-membered heterocyclic ring system optionally containing an additional heteroatom selected from O, S, and N, the heterocyclic ring being unsubstituted or substituted with one to three substituents independently selected from oxo, hydroxy, halogen, C_{1-4} alkoxy, and C_{1-4} alkyl,

wherein alkyl and alkoxy are unsubstituted or substituted with one to five fluorines;;

R^2 and R^3 are each independently hydrogen, C_{1-8} alkyl, aryl, or aryl-C_{1-2} alkyl; or

R^2 and R^3 together with the nitrogen atom to which they are attached form a 4- to 7-membered heterocyclic ring system optionally containing an additional heteroatom selected from O, S, and N, the heterocyclic ring being unsubstituted or substituted with one to three substituents independently selected from oxo, hydroxy, halogen, C_{1-4} alkoxy, and C_{1-4} alkyl, wherein alkyl and alkoxy are unsubstituted or substituted with one to five fluorines; and the heterocyclic ring system being optionally fused with a 5- to 6-membered saturated or aromatic carbocyclic ring system or a 5- to 6-membered saturated or aromatic heterocyclic ring system containing one to two heteroatoms selected from O, S, and N, the fused ring system being unsubstituted or substituted with one to two substituents selected from hydroxy, amino,

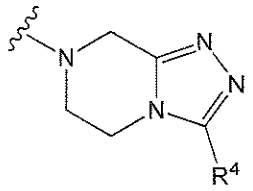
fluorine, C_{1-4} alkyl, C_{1-4} alkoxy, and trifluoromethyl; the process comprising the step of contacting a prochiral ketone of structural Formula (**II**):



(II)

with an engineered polypeptide as disclosed herein in the presence of an amino group donor in a suitable organic solvent under suitable reaction conditions.

[0025] In some embodiments of the process for preparing a compound of structural Formula (**I**), R^1 is benzyl and the phenyl group of benzyl is unsubstituted or substituted one to three substituents selected from the group consisting of fluorine, trifluoromethyl, and trifluoromethoxy. In some embodiments of the process, Z is NR^2R^3, wherein NR^2R^3 is a heterocycle of the structural Formula (**III**):

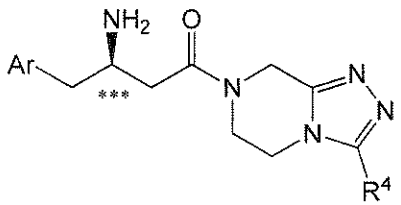


(III)

wherein R^4 is hydrogen or C_{1-4} alkyl which is unsubstituted or substituted with one to five fluorines.

[0026] In some embodiments of the process for preparing a compound of structural Formula (**I**), the compound of Formula (**II**) specifically excludes compound (**2**) and the compound of Formula (**I**) prepared by the method specificall excludes compound (**1**).

[0027] In one aspect, there is provided a ~~some embodiments, the present disclosure provides~~ process for preparing a compound of structural Formula (**Ia**):



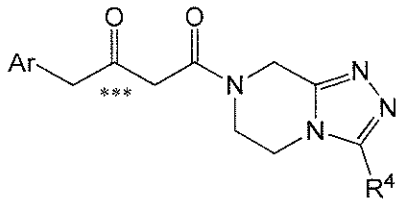
(Ia)

having the (*R*)-configuration at the stereogenic center marked with an ***; in an enantiomeric excess of at least 70% over the enantiomer having the opposite (*S*)-configuration; wherein

Ar is phenyl which is unsubstituted or substituted with one to five substituents independently selected from the group consisting of fluorine, trifluoromethyl, and trifluoromethoxy; and

R^4 is hydrogen or C_{1-4} alkyl unsubstituted or substituted with one to five fluorines; the process comprising the step of:

contacting a prochiral ketone of structural Formula **(IIa)**:

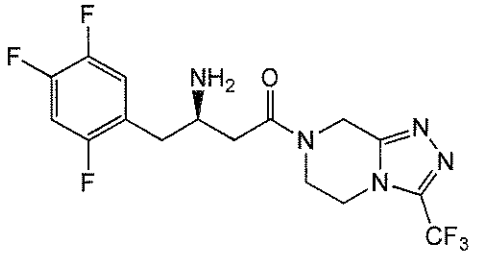


(IIa)

with an engineered polypeptide as disclosed herein in the presence of an amino group donor under suitable reaction conditions. In some embodiments of the process for preparing the compound of Formula (**Ia**), Ar is selected from 2,5-difluorophenyl or 2,4,5-trifluorophenyl, and R^4 is trifluoromethyl.

[0028] In some embodiments of the process for preparing a compound of structural Formula **(Ia)**, the compound of Formula (**IIa**) specifically excludes compound **(2)** and the compound of Formula **(Ia)** prepared by the method specificall excludes compound **(1)**.

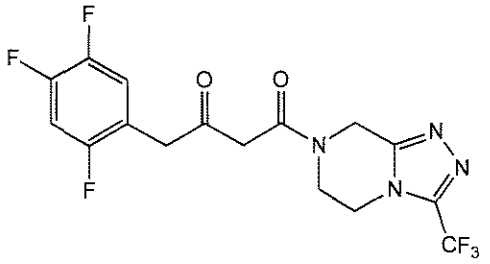
[0029] In one aspect, there is provided ~~some embodiments, the present disclosure provides~~ a process of preparing compound **(1)**



(1)

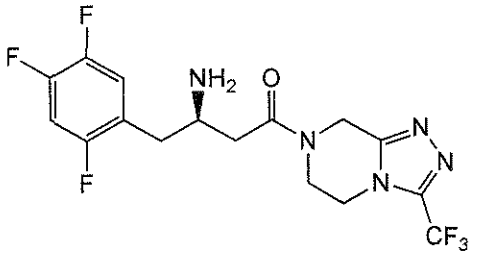
comprising a step of contacting a substrate of compound **(2)**

production of JANUVIA® or related pharmaceutical compounds. In one aspect, there is provided ~~some embodiments the present disclosure also provides~~ a process for the preparation of (2*R*)-4-oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]triazolo[4,3-*a*]pyrazin-7(8*H*)-yl]-1-(2,4,5-trifluorophenyl)butan-2-amine phosphate (1:1) monohydrate, wherein the process comprises a step of converting a substrate compound (**2**)



(2)

to a product compound (**1**)



(1)

by contacting a substrate of compound (**2**) with an engineered polypeptide as disclosed herein, in the presence of an amino group donor under suitable reaction conditions.

[0037] Similarly, the present disclosure provides a process for the preparation of compound (**3**), or a pharmaceutically acceptable salt or acid of compound (**3**), wherein the process comprises a step of converting a substrate compound (**4**), or a substrate of compound (**4**) modified with a protecting group, to a product compound (**3**), by contacting a substrate of compound (**4**), or a substrate of compound (**4**) modified with a protecting group, with an engineered polypeptide as disclosed herein in the presence of an amino group donor under suitable reaction conditions.

[0038] Further guidance on the choice of engineered polypeptides, their preparation, the choice of substrates, and parameters for carrying out the processes are further described in the more detailed description and Examples that follow.

5. DETAILED DESCRIPTION

[0039] 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.

[0040] Similarly, "comprise," "comprises," "comprising" "include," "includes," "including," "have," and "having" are interchangeable and not intended to be limiting.

[0041] It is to be 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."

[0042] It is to be further 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.

[0043] The section headings used herein are for organizational purposes only and not to be construed as limiting the subject matter described.

5.1 Abbreviations

[0044] The abbreviations used for the genetically encoded amino acids are conventional and are as follows:

Amino Acid	Three-Letter Abbreviation	One-Letter Abbreviation
Alanine	Ala	A