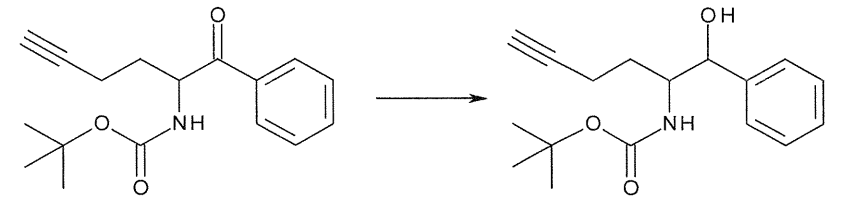
[0008a] In one aspect, there is provided an engineered ketoreductase polypeptide capable of converting substrate compound (2), tert-butyl(1-oxo-1-phenylhex-5-yn-2-yl)carbamate, to product compound (1), tert-butyl(1-hydroxy-1-phenylhex-5-yn-2-yl)carbamate,



(2) (1)

the ketoreductase polypeptide comprising an amino acid sequence having at least 80% sequence identity to SEQ ID NO:4 and the features X40R and X206L or F, with the proviso that when X206 is F, the sequence further comprises one or more residue differences as compared to SEQ ID NO:4 selected from: X7S; X17M; X17Q; X17R; X23V; X27L; X29G; X60I; X64V; X71P; X87L; X94A; X94P; X94S; X95M; X96L; X96Y; X105G; X113I; X122A; X127R; X131S; X144V; X145L; X147I; X147L; X147Q; X150Y; X152G; X153G; X157C; X173L; X195A; X195G; X196M; X198S; X199H; X208R; X216R; X221S; X243S; X245I; X249F; X249G; and X249Y.

[0009] In one ~~aspect~~embodiment, the engineered ketoreductase polypeptide comprises an amino acid sequence having at least 80% sequence identity to a reference sequence selected from SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 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, 96, 98, 102, 104, 106, 108, 110, 114, 116, 122, 124, 126, 128, 130, 132, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 168, 170 and 172 and one or more residue differences as compared to SEQ ID NO:2 or 4 selected from X7S; X17M; X17Q; X17R; X23V; X27L; X29G; X40R; X60I; X64V; X71P; X87L; X94A; X94P; X94S; X95M; X96L; X96Y; X105G; X113I; X122A; X127R; X131S; X144V; X145L; X147I; X147L; X147Q; X150Y; X152G; X153G; X157C; X173L; X195A; X195G; X196M; X198S; X199H; X206F; X206L; X208R; X216R; X221S; X243S; X245I; X249F; X249G; and X249Y. As provided herein, in some embodiments, the disclosed amino acid differences can be used singly or in various combinations to generate the engineered ketoreductase polypeptides having the improved enzyme properties.

[0010] In some embodiments, the ketoreductase polypeptide comprises an amino acid sequence having at least 80% sequence identity to SEQ ID NO:4 and the features X40R and X206L or F, with the proviso that when X206 is F, the sequence further comprises one or more residue differences as compared to SEQ ID NO:4 selected from: X7S; X17M; X17Q; X17R; X23V; X27L; X29G; X60I; X64V; X71P; X87L; X94A; X94P; X94S; X95M; X96L; X96Y; X105G; X113I; X122A; X127R; X131S; X144V; X145L; X147I; X147L; X147Q; X150Y; X152G; X153G; X157C; X173L; X195A; X195G; X196M; X198S; X199H; X208R; X216R; X221S; X243S; X245I; X249F; X249G; and X249Y.

[0011] In some embodiments, the engineered ketoreductase comprises an amino acid sequence having the features X40R and X206F or L and at least one or more residue differences as compared to SEQ ID NO:4 selected from: X17Q/R/M; X64V; X94P; X144V; X147Q/I/L; X157C; and X196M.

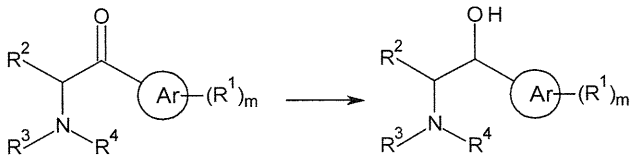
[0012] In some embodiments, the engineered ketoreductase polypeptide with improved properties in the conversion of substrate compound (**2**) to product compound (**1**) comprises an amino acid sequence selected from SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 96, 98, 102, 104, 106, 108, 110, 114, 116, 122, 124, 126, 128, 130, 132, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 168, 170 and 172, and in particular an amino acid sequence selected from SEQ ID NO: 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 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, 154, 156, 158, 160, 162, 164, 166, 168, 170 and 172.

[0013] In some embodiments, the engineered ketoreductase polypeptide is capable of forming product compounds (**1a**) and (**1c**)

[0019] In another ~~aspect~~embodiment, the present disclosure provides polynucleotides encoding the engineered ketoreductases, as well as expression vectors comprising the polynucleotides, and host cells capable of expressing the polynucleotides encoding the polypeptides. In one aspect, there is provided a polynucleotide encoding an engineered ketoreductase polypeptide as defined herein. In another aspect, there is provided an expression vector comprising the polynucleotide as defined herein. In another aspect, there is provided a host cell comprising the polynucleotide as defined herein. In some embodiments, the present disclosure also provides methods of manufacturing the engineered ketoreductase polypeptides. In one aspect, there is provided a method for preparing an engineered ketoreductase polypeptide as defined herein, comprising culturing the host cell as defined herein under culture conditions suitable for expression of the engineered polypeptide. Exemplary polynucleotide sequences encoding the engineered ketoreductases include the sequences of SEQ ID NO: 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 and 171, in particular the sequences of SEQ ID NO: 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 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 and 171.

[0020] In another ~~aspect~~embodiment, the engineered ketoreductase polypeptides can be used in a process for the preparation of hydroxyl compounds of formula (**I**) useful for synthesis of drug compounds. Accordingly, in some embodiments, the engineered ketoreductases can be used in the conversion of substrate compounds of formula (**II**) to product compounds of formula (**I**), as shown in **Scheme 2**

Scheme 2



(II) **(I)**

wherein

Ar is a 5 to 7-membered carbocyclic or heterocyclic ring;

each occurrence of R^1 is independently selected from halo, $-COOR^a$, $-C(O)R^b$, $-OR^c$, $-SO_2$,$-SR^d$, $-S(O)R^e$, $-NR^fR^g$, $-C(O)NR^hR^i$,$-NO_2$, -CN, alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl and heteroarylalkyl, wherein R^a, R^b, R^c, R^d, R^e, R^f, R^g, R^h and R^i are independently selected from H and an optionally substituted alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl;

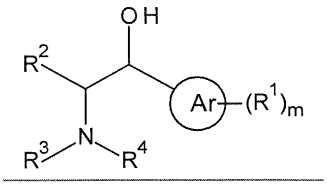
R^2 is selected from an optionally substituted alkyl, alkenyl, alkynyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, cycloalkylalkenyl, cycloalkylalkynyl, heterocycloalkyl, heterocycloalkylalkyl, heterocycloalkylalkenyl, heterocycloalkylalkynyl, aryl, arylalkyl, arylalkenyl, arylalkynyl, heteroaryl, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, alkylthio, alkylsulfinyl, and arylsulfinyl;

R^3 and R^4 are, independently of the other, selected from H, an optionally substituted alkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, heteroarylalkyl, alkyloxycarbonyl, arylalkyloxycarbonyl, heteroarylalkyloxycarbonyl, and a protecting

group, or one of R^3 and R^4 forms an optionally substituted 5 to 7-membered heterocyclic ring with R^2, and

m is 0 to 10.

[0020a] In one aspect, there is provided a process for preparing the product compound of formula (I)



(I)

wherein

Ar is a 5 to 7-membered carbocyclic or heterocyclic ring;

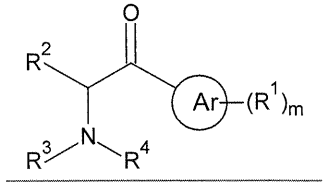
each occurrence of R^1 is independently selected from halo, $-COOR^a$, $-C(O)R^b$, $-OR^c$, $-SO_2$,$-SR^d$, $-S(O)R^e$, $-NR^fR^g$, $-C(O)NR^hR^i$,$-NO_2$, -CN, alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl and heteroarylalkyl, wherein R^a, R^b, R^c, R^d, R^e, R^f, R^g, R^h and R^i are independently selected from H and an optionally substituted alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl;

R^2 is selected from an optionally substituted alkyl, alkenyl, alkynyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, cycloalkylalkenyl, cycloalkylalkynyl, heterocycloalkyl, heterocycloalkylalkyl, heterocycloalkylalkenyl, heterocycloalkylalkynyl, aryl, arylalkyl, arylalkenyl, arylalkynyl, heteroaryl, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, alkylthio, alkylsulfinyl, and arylsulfinyl;

R^3 and R^4 are, independently of the other, selected from H, an optionally substituted alkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, heteroarylalkyl, alkyloxycarbonyl, arylalkyloxycarbonyl, heteroarylalkyloxycarbonyl, and a protecting group, or one of R^3 and R^4 forms an optionally substituted 5 to 7-membered heterocyclic ring with R^2, and

m is 0 to 10;

comprising contacting the substrate compound of formula (II),



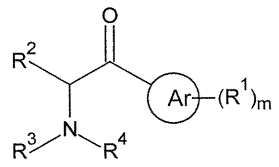
(II)

wherein

R^1, R^2, R^3, R^4, and *m* are as defined above,

with an engineered ketoreductase polypeptide as defined herein in presence of cofactor NADPH or NADH under suitable reaction conditions.

[0021] Accordingly, a process for preparing the compound of formula (**I**) can comprise contacting the compound of formula (**II**)



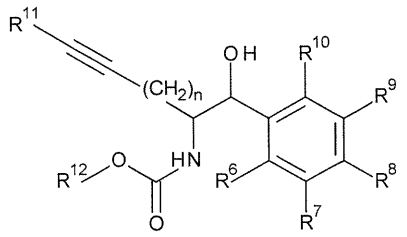
(II)

wherein

R^1, R^2, R^3, R^4, and *m* are as defined above;

with an engineered ketoreductase polypeptide described herein in presence of cofactor NADPH or NADH under suitable reaction conditions.

[0022] In some embodiments of the process, the compound of formula (**I**) comprises the compound of formula (**Ib**)



(Ib)

wherein

R^6, R^7, R^8, R^9, and R^{10}, are each independently selected from H, halo, -$COOR^a$, -$C(O)R^b$, -OR^c, -SO_2,-SR^d, -$S(O)R^e$, -NR^fR^g, -$C(O)NR^hR^i$,-NO_2, -CN, alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl and heteroarylalkyl, wherein R^a, R^b, R^c, R^d, R^e, R^f, R^g, R^h and R^i are independently selected from H and an optionally substituted alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl;

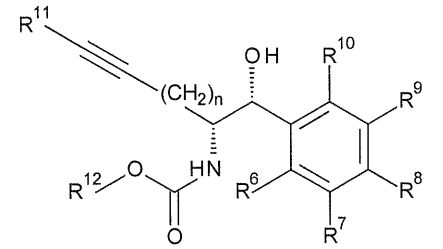
n is 1 to 4,

R^{11} is selected from H and an optionally substituted alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl; and

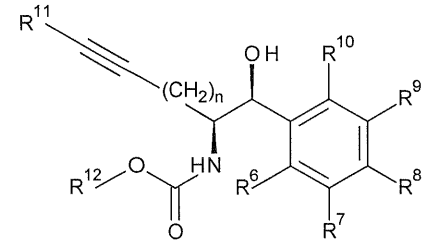
R^{12} is selected from H, an optionally substituted alkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, heteroarylalkyl and a protecting group.

compound of formula (**Ib1**) in diastereomeric excess of compounds of formula (**Ib3**) comprises contacting the substrate compound of formula (**IIb**) with an engineered ketoreductase having diastereoselectivity for compound of formula (**Ib1**) over compound of formula (**Ib3**).

[0026a] In one aspect, there is provided a process for preparing product compound of formula (**Ib1**) in diastereomeric excess of compound of formula (**Ib3**),



(Ib1)



(Ib3)

wherein,

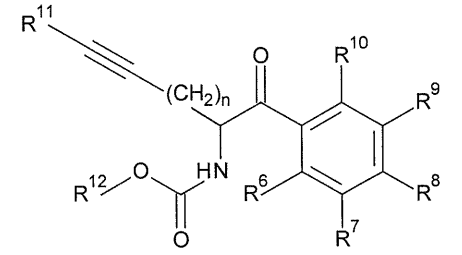
R^6, R^7, R^8, R^9, and R^{10} are each independently selected from H, halo, $-COOR^a$, $-C(O)R^b$, $-OR^c$, $-SO_2$, $-SR^d$, $-S(O)R^e$, $-NR^fR^g$, $-C(O)NR^hR^i$, $-NO_2$, -CN, alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl and heteroarylalkyl, wherein R^a, R^b, R^c, R^d, R^e, R^f, R^g, R^h and R^i are independently selected from H and an optionally substituted alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, heterocycloalkylalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl;

n is 1 to 4;

R^{11} is selected from H or an optionally substituted alkyl, heteroalkyl, cycloalkyl, cycloalkylalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl; and

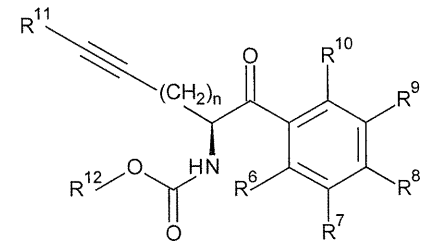
R^{12} is selected from H, an optionally substituted alkyl, cycloalkyl, heterocycloalkyl, aryl, heteroaryl, and a protecting group;

the process comprising contacting the substrate compound of formula (**IIb**)

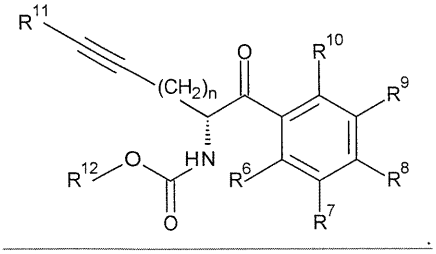


(IIb)

wherein R^6, R^7, R^8, R^9, R^{10}, R^{11}, R^{12} and *n* are as defined above, with an engineered ketoreductase as defined herein in presence of NADPH or NADH under suitable reaction conditions for equilibration of substrate compounds of formula (**IIb1**) and (**IIb2**)



(IIb1)



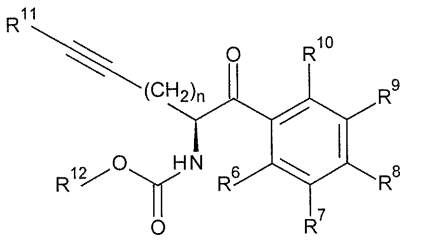
(IIb2)

[0027] Exemplary engineered ketoreductase polypeptides having diastereoselectivity for compounds of formula (**Ib1**) and (**Ib3**) over compounds of formula (**Ib2**) and (**Ib4**) or diastereoselectivity for compound of formula (**Ib1**) over compound of formula (**Ib3**) can comprise an amino acid sequence selected from SEQ ID NO:18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 56, 58, 60, 62, 64, 66, 68, 70, 72, 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, 154, 156, 158, 160, 162, 164, 166, 168, 170 and 172.

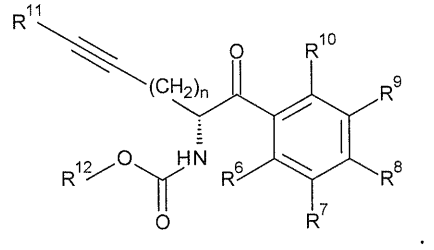
[0028] In some embodiments, the product compound of formula (**Ib**) comprises the compounds of formula (**Ib2**) and (**Ib4**), and the process forms the product compounds of formula (**Ib2**) and (**Ib4**) in diastereomeric excess of compounds of formula (**Ib1**) and (**Ib3**). In some embodiments, the process for preparing product compounds of formula (**Ib2**) and (**Ib4**) in diastereomeric excess of compounds of formula (**Ib1**) and (**Ib3**) comprises contacting the substrate compound of formula (**IIb**) with an engineered ketoreductase having diastereoselectivity for compounds of formula (**Ib2**) and (**Ib4**) over compounds of formula (**Ib1**) and (**Ib3**).

[0029] Exemplary engineered ketoreductase polypeptides having diastereoselectivity for compounds of formula (**Ib2**) and (**Ib4**) over compounds of formula (**Ib1**) and (**Ib3**) can comprise an amino acid sequence selected from SEQ ID NO: 50 and 74.

[0030] In some embodiments, the processes for preparing compounds of formula (**Ib1**) and (**Ib3**) in diastereomeric excess of compounds of formula (**Ib2**) and (**Ib4**), the compound of formula (**Ib1**) in diastereomeric excess of compound of formula (**Ib3**), and the compounds of formula (**Ib2**) and (**Ib4**) in diastereomeric excess of the compounds of formula (**Ib1**) and (**Ib3**) can be carried out under suitable reaction conditions that promote equilibration between the substrate compounds of formula (**IIb1**) and (**IIb2**)



(IIb1)



(IIb2)

[0031] The equilibration between substrate compounds **(IIb1)** and **(IIb2)** during the conversion to product can - by dynamic kinetic resolution - increase the yield of product compound beyond the initial concentration of compounds of formula **(IIb1)** and **(IIb2)** present in the reaction solution.