US 20250178995A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0178995 A1**

Lee et al. (43) **Pub. Date: Jun. 5, 2025**

(54) **COMPOSITIONS AND METHODS FOR THE PRODUCTION OF GLUCARIC ACID**

(71) Applicant: **Solugen, Inc.**, Houston, TX (US)

(72) Inventors: **Toni M. Lee**, Missouri City, TX (US); **David Weiner**, Houston, TX (US); **Konrad V. Miller**, Sugar Land (TX); **Gaurab Chakrabarti**, Houston, TX (US); **Sean Hunt**, Houston, TX (US)

(73) Assignee: **Solugen, Inc.**, Houston, TX (US)

(21) Appl. No.: **18/965,379**

(22) Filed: **Dec. 2, 2024**

Related U.S. Application Data

(60) Provisional application No. 63/604,831, filed on Nov. 30, 2023.

Publication Classification

(51) **Int. Cl.**

C07C 31/26 (2006.01)

B01J 31/00 (2006.01)

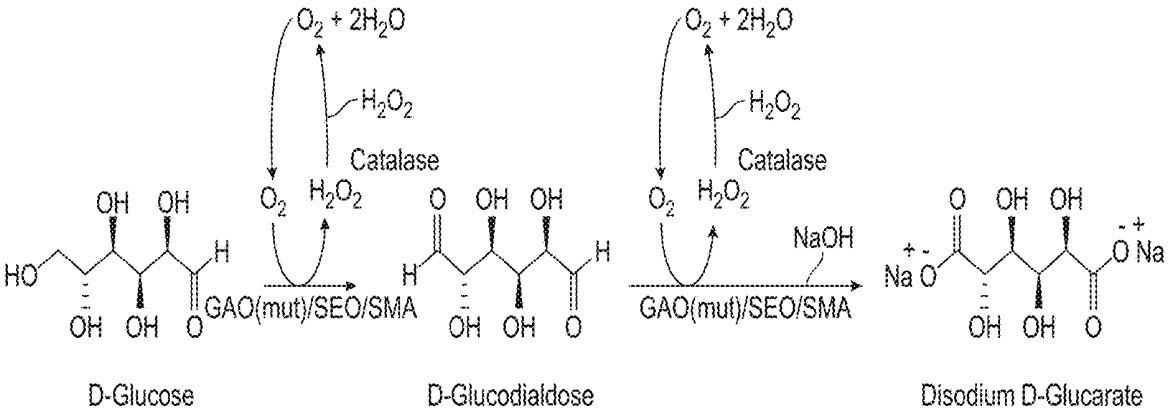
(52) **U.S. Cl.**

CPC ***C07C 31/26*** (2013.01); ***B01J 31/003*** (2013.01)

(57) **ABSTRACT**

A method of preparing glucaric acid comprising contacting D-glucose and oxygen with a first catalyst composition comprising a first copper radical oxidase, a single electron oxidizer, and a small molecule activator under conditions suitable for formation of glucodialdose; and contacting glucodialdose and oxygen with a second catalyst composition comprising a second copper radical oxidase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of a product mixture comprising glucaric acid or salts thereof. A method comprising contacting a sugar with a catalyst composition comprising an oxidoreductase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of one or more oxidized sugar oxidation products comprising glucaric acid wherein the oxidoreductase comprises at least two copper radical oxidases, at least two mutated copper radical oxidases or combinations thereof.

Specification includes a Sequence Listing.



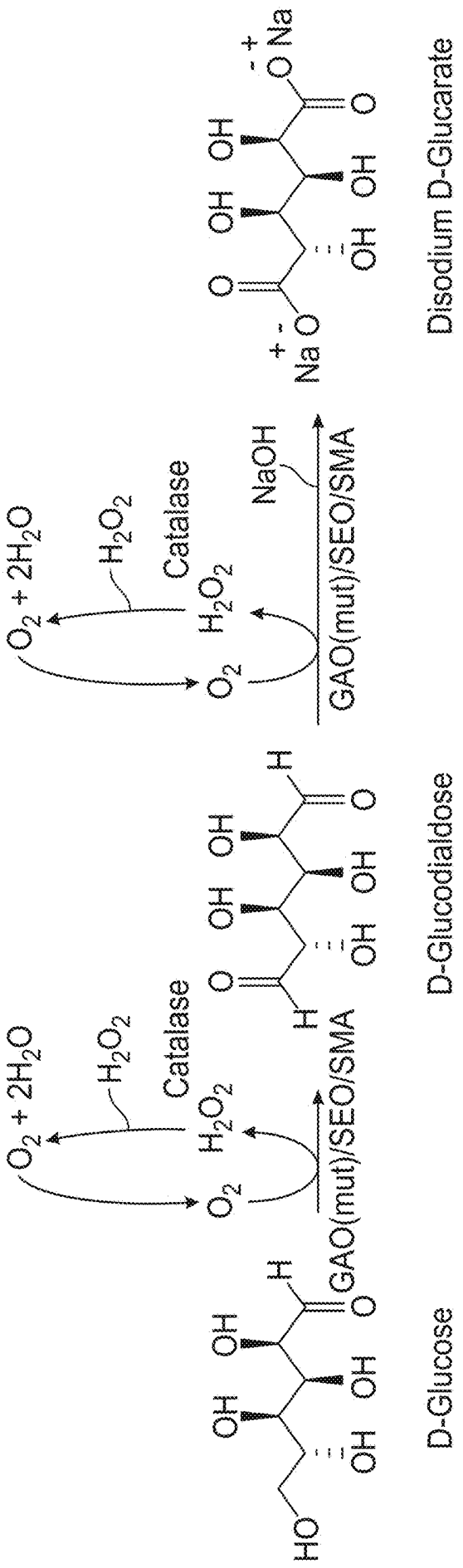

D-Glucose
GAO(mut)/SEO/SMA
D-Glucodialdose
GAO(mut)/SEO/SMA
Disodium D-Glucarate
NaOH
Catalase
H_2O_2
$O_2 + 2H_2O$
O_2
H_2O_2
Catalase
H_2O_2
$O_2 + 2H_2O$
O_2
H_2O_2

FIG. 1

COMPOSITIONS AND METHODS FOR THE PRODUCTION OF GLUCARIC ACID

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application No. 63/604,831 filed on Nov. 30, 3023 and entitled, "COMPOSITIONS AND METHODS FOR THE PRODUCTION OF GLUCARIC ACID," which is hereby incorporated herein by reference in its entirety for all purposes.

REFERENCE TO SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in XML file format and is hereby incorporated by reference in its entirety. Said XML file, created on Nov. 25, 2024, is named "23ENZ008PCT_3416-18001 SEQUENCES11252024.xml" and is 94,208 bytes in size.

TECHNICAL FIELD

[0003] The present disclosure relates generally to the production of high value chemicals. More particularly, the present disclosure relates to the production of glucaric acid from reactants sourced from renewable resources.

BACKGROUND

[0004] Over the last 150 years, synthesis of inexpensive chemicals from fossilized forms of carbon (e.g. oil, coal, natural gas) has dramatically altered society through their broad applications, ranging from cosmetics to plastics. This petroleum-based carbon feedstock generates a small collection of platform chemicals from which highly efficient chemical conversions lead to the manufacture of a large variety of chemical products. However, the current approach to producing these carbon-based chemicals is inherently non-sustainable as feedstocks that required millions of years to form are being depleted. The creation of a truly sustainable chemical industry will only occur when the timescale of the feedstock formation matches the timescale of its utilization to make chemicals.

[0005] For example, the principal methodologies for producing glucaric acid commercially involve 1) nitric acid oxidation or 2) transition metal catalyzed oxidation. The drawbacks of these commercial methods include (i) the generation of a significant amount of hazardous nitrogen oxide (NOx) gas and (ii) the processes are highly exothermic leading to controllability issues. The transition metal catalyzed (e.g., Pt or Pd) method for glucaric acid production uses glucaric acid as an intermediate in the production of bio-based adipic acid.

[0006] Other methods for the production of glucaric acid include a microbial method of producing high-purity glucaric acid using as catalysts myo-inositol-1-phosphate synthase from *S. cerevisiae*, mouse myo-inositol oxygenase, and *P. syringae* uronate dehydrogenase. Microbial methods, however, can suffer from product separation issues leading to high chemical costs, limiting usage as a commodity chemical. An ongoing need exists for novel glucaric acid production methods that lack some of the aforementioned challenges.

SUMMARY

[0007] A method of preparing glucaric acid comprising contacting D-glucose and oxygen with a first catalyst composition comprising a first copper radical oxidase, a single electron oxidizer, and a small molecule activator under conditions suitable for formation of glucodialdose; and contacting glucodialdose and oxygen with a second catalyst composition comprising a second copper radical oxidase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of a product mixture comprising glucaric acid or salts thereof.

[0008] A method comprising contacting a sugar with a catalyst composition comprising an oxidoreductase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of one or more oxidized sugar oxidation products comprising glucaric acid wherein the oxidoreductase comprises at least two copper radical oxidases, at least two mutated copper radical oxidases or combinations thereof.

[0009] Embodiments described herein comprise a combination of features and characteristics intended to address various shortcomings associated with certain prior devices, systems, and methods. The foregoing has outlined rather broadly the features and technical characteristics of the disclosed embodiments in order that the detailed description that follows may be better understood. The various characteristics and features described above, as well as others, will be readily apparent to those skilled in the art upon reading the following detailed description, and by referring to the accompanying drawings. It should be appreciated that the conception and the specific embodiments disclosed may be readily utilized as a basis for modifying or designing other structures for carrying out the same purposes as the disclosed embodiments. It should also be realized that such equivalent constructions do not depart from the spirit and scope of the principles disclosed herein.

BRIEF DESCRIPTION OF DRAWINGS

[0010] For a detailed description of various exemplary embodiments, reference will now be made to the accompanying drawings in which:

[0011] FIG. **1** schematically depicts a reaction of the present disclosure for generation of glucaric acid from D-glucose.

DETAILED DESCRIPTION

[0012] The following discussion is directed to various exemplary embodiments. However, one skilled in the art will understand that the examples disclosed herein have broad application, and that the discussion of any embodiment is meant only to be exemplary of that embodiment, and not intended to suggest that the scope of the disclosure, including the claims, is limited to that embodiment.

[0013] Certain terms are used throughout the following description and claims to refer to particular features or components. As one skilled in the art will appreciate, different persons may refer to the same feature or component by different names. This document does not intend to distinguish between components or features that differ in name but not function. The drawing FIGURES are not necessarily to scale. Certain features and components herein may be shown exaggerated in scale or in somewhat sche-

matic form and some details of conventional elements may not be shown in interest of clarity and conciseness.

[0014] Unless the context dictates the contrary, all ranges set forth herein should be interpreted as being inclusive of their endpoints, and open-ended ranges should be interpreted to include only commercially practical values. Similarly, all lists of values should be considered as inclusive of intermediate values unless the context indicates the contrary.

[0015] Where numerical ranges or limitations are expressly stated, such express ranges or limitations should be understood to include iterative ranges or limitations of like magnitude falling within the expressly stated ranges or limitations (e.g., from about 1 to about 10 includes, 2, 3, 4, etc.; greater than 0.10 includes 0.11, 0.12, 0.13, etc.).

[0016] In the following discussion and in the claims, the terms "including" and "comprising" are used in an open-ended fashion, and thus should be interpreted to mean "including, but not limited to . . . " As used herein, the phrases "consist(s) of" and "consisting of" are used to refer to exclusive components of a composition, meaning only those expressly recited components are included in the composition; whereas the phrases "consist(s) essentially of" and "consisting essentially of" are used to refer to the primary components of a composition, meaning that only small or trace amounts of components other than the expressly recited components (e.g., impurities, byproducts, etc.) may be included in the composition. For example, a composition consisting of X and Y refers to a composition that only includes X and Y, and thus, does not include any other components; and a composition consisting essentially of X and Y refers to a composition that primarily comprises X and Y, but may include small or trace amounts of components other than X and Y. In embodiments described herein, any such small or trace amounts of components other than those expressly recited following the phrase "consist(s) essentially of" or "consisting essentially of" preferably represent less than 5.0 wt % of the composition, more preferably less than 4.0 wt % of the composition, even more preferably less than 3.0 wt % of the composition, and still more preferably less than 1.0 wt % of the composition. Use of broader terms such as comprises, includes, having, etc. should be understood to provide support for narrower terms such as consisting of, consisting essentially of, comprised substantially of, etc. Use of the term "optionally" with respect to any element of a claim is intended to mean that the subject element is required, or alternatively, is not required. Both alternatives are intended to be within the scope of the claim.

[0017] Disclosed herein are chemoenzymatic methods of producing glucaric acid from a renewable resource such as a sugar. Advantageously, the disclosed chemoenzymatic method of glucaric acid production involves the use of benign conditions such as aqueous solvents and low temperatures. In one or more aspects, the disclosed compositions and methods are combined with processes for the formation of hexamethyldiamine which is subsequently utilized in the production of polyamides such as poly [imino (1,6-dioxohexamethylene) iminohexamethylene] (Nylon 66).

[0018] Advantageously, the presently disclosed methods utilize bio-based feedstocks, for example, feedstocks derived from biomass, as reactants resulting in the production of glucaric acid with high selectivity and minimized formation of toxic waste products.

[0019] In one or more aspects, a method of the present disclosure comprises contacting one or more sugars and subsequent reaction intermediates with a plurality of copper radical oxidases (CROs) under conditions suitable for the formation of glucaric acid.

[0020] In one or more aspects, the sugar comprises glucose. CROs are a class of non-flavoprotein alcohol oxidoreductases that employ molecular oxygen as a terminal electron acceptor to generate hydrogen peroxide. CROs have been labeled 'green' small-molecule oxidation catalysts as they lack dependence on an organic cofactor and utilizesmolecular oxygen as a cosubstrate.

[0021] Nonlimiting examples of CROs suitable for use in the present disclosure include glyoxal oxidases (EC 1.1.3.-, GLOX) and galactose 6-oxidases (EC 1.1.3.9, GAO). GLOXes typically function on aldehydes such as methylglyoxal to produce acids. One GLOX from *Phanerochaete chrysosporium* primarily accepts alpha-dicarbonyl and alpha-hydroxycarbonyls. In one or more aspects, the CRO is sourced from *Pycnoporus cinnabarinus*. *Pycnoporus cinnabarinus* was found to express three GLOXes, one of which has been found to function on methylglyoxal while the other two show high catalytic efficiency for glyoxylic acid.

[0022] In an aspect, the CRO is a GAO, alternatively a mutated GAO. GAOs function by oxidizing six-carbon (C6) or similar alcohols of galactose or other sugars to produce aldehydes. A particular example of a CRO is the GAO from *Fusarium graminearum*. Two additional CROs capable of oxidizing aliphatic alcohols are CgrAlcOx and CglAlcOx, which are derived from *Colletotrichum graminicola* and *C. gloeosporioides*, respectively. Another example is aryl-alcohol oxidase (CgrAAO) from *C. graminicola*.

[0023] In one or more aspects, a method of the present disclosure is depicted in FIG. **1**. With reference to FIG. **1**, a sugar (e.g., D-glucose) is reacted with a first CRO (designated CRO1) under conditions suitable for the formation of an intermediate, D-glucodialdose (GDA). In one or more aspects, conditions suitable for the formation of GDA include the presence of a small molecule activator (SMA) and a single electron oxidizer (SEO) such as a peroxidase; both of which facilitate the catalytic activity of the CRO1.

[0024] Without wishing to be limited by theory, the SEO performs a single electron oxidation on the SMA to generate the free radical form ($SMA^{\bullet+}$). The $SMA^{\bullet+}$ can reverse or return a CRO (e.g., GAO) to the active state. Without wishing to be limited by theory, reversing and/or returning a CRO to the active state may comprise oxidation of one or more metals of the CRO at the active site of the enzyme to an active oxidation state.

[0025] In one or more aspects, an SMA suitable for use in this disclosure is a molecule that (i) is capable of stabilizing a free radical, (ii) can serve as a substrate for a SEO in a single electron oxidation reaction, and iii) is capable of restoring the active state of a CRO as evidenced by detectable activity (e.g., formation of product or oxygen consumption).

[0026] Nonlimiting examples of SMAs suitable for use in the present disclosure include L-tryptophan, 2-mercaptobenzothiazole, L-histidine, methylchloroisothiazolinone, o-dianisidine, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), 4-aminoantipyrine, L-tyrosine, (2,2,6,6-tetramethylpiperidin-1-yl)oxyl, chloromethylisothiazolinone, 4-thiazolecarboxylic acid, Sunset yellow FCF, tartrazine, p-benzoquinone, dicoumarol, phthalimide, saccharin,

phthalic anhydride, erythrosine B, 2-aminobenzothiazole, thiabendazole, 2-hydroxybenzothiazole, phenothiazine, 6-aminobenzothiazole, indigo carmine, naphthalimide, 2-aminothiazole, thiazole, 2H-1,4-benzothiazin-3 (4H)-one, 2-oxindole, beta-lapachone, menaquinone, thiamine, 4-methyl-5-thiazoleethanol, Allura Red AC, menadione, p-cresol, Fast green FCF, Brilliant Blue FCF, methylisothiazolinone, caffeine, veratryl alcohol, fluorescein, and combinations thereof. The structures of several of these SMAs are depicted in Table 1.

TABLE 1

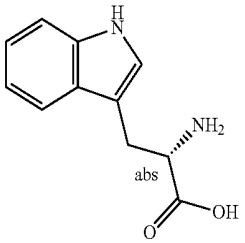
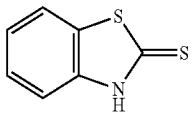
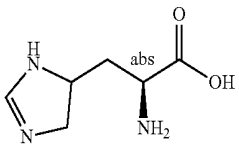
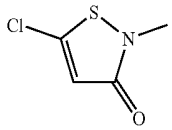
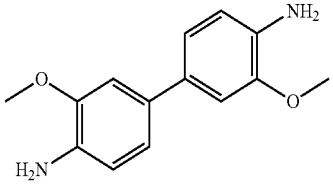
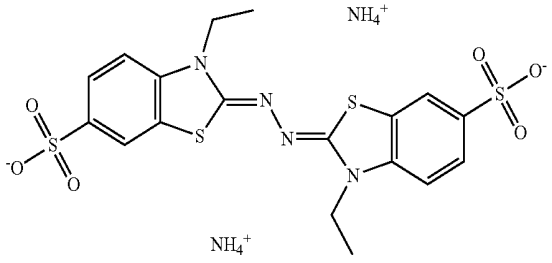
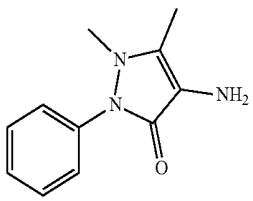
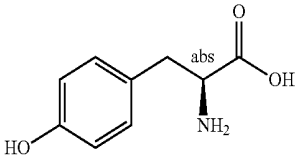
L-tryptophan	
2-mercapto-benzothiazole	
L-histidine	
Methylchloro-isothiazolinone	
o-dianisidine	
ABTS	
4-aminoantipyrine	
L-tyrosine	

TABLE 1-continued

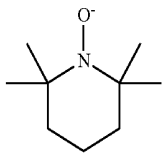
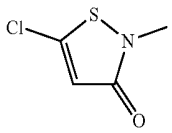
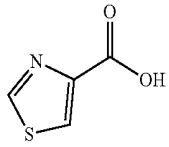
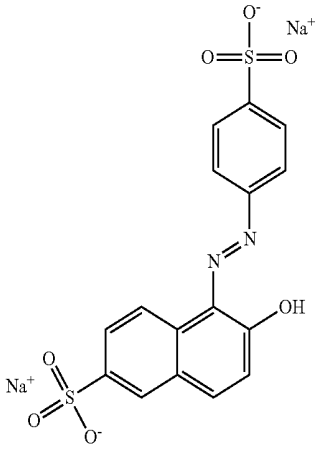
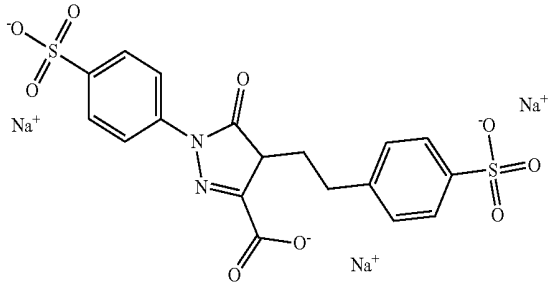
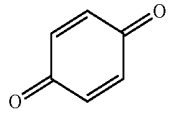
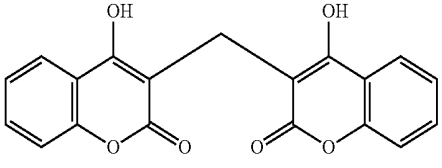
(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl	
chloromethyl-isothiazolinone	
4-thiazolecarboxylic acid	
Sunset yellow FCF	
Tartrazine	
p-benzoquinone	
dicoumarol	

TABLE 1-continued

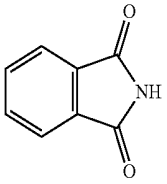
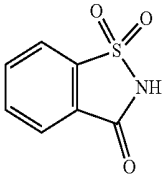
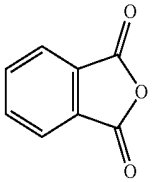
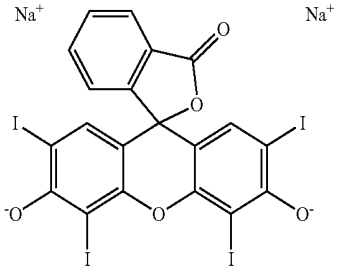
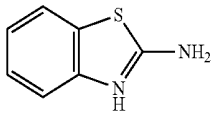
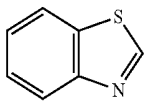
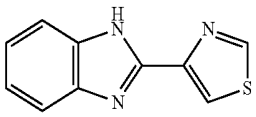
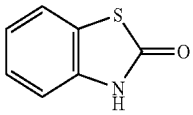
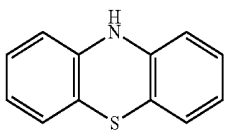
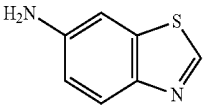
phthalimide	
saccharin	
phthalic anhydride	
Erythrosine B	
2-aminobenzothiazole	
benzothiazole	
thiabendazole	
2-Hydroxybenzothiazole	
Phenothiazine	
6-aminobenzothiazole	

TABLE 1-continued

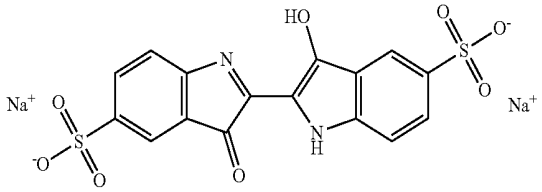
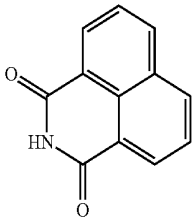
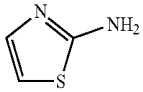
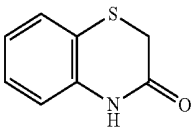
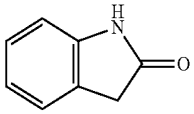
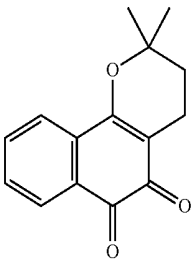
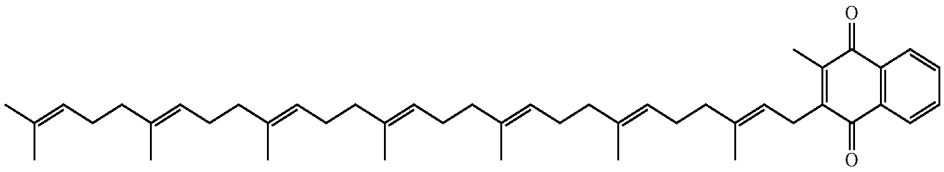
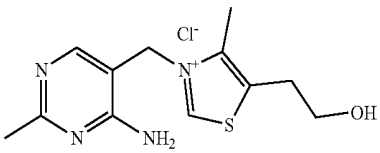
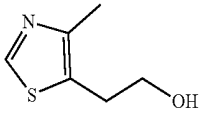
Indigo carmine	
naphthalimide	
2-aminothiazole	
thiazole	
2H-1,4-Benzothiazin-3(4H)-one	
2-oxindole	
beta-lapachone	
menaquinone	
thiamine	
4-Methyl-5-thiazoleethanol	

TABLE 1-continued

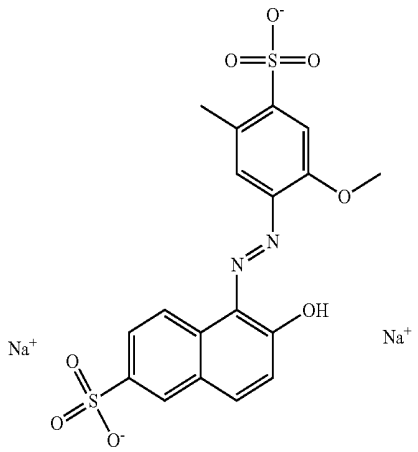
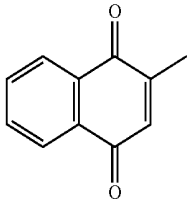
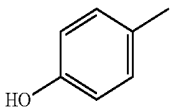
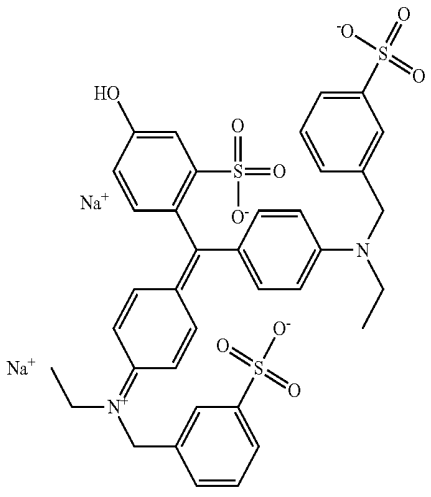
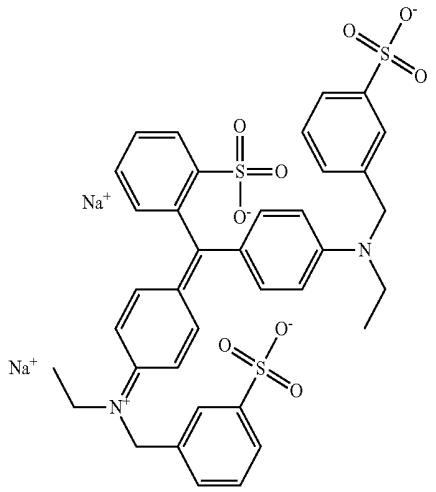
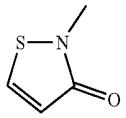
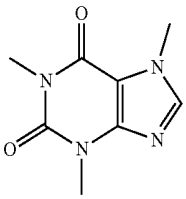
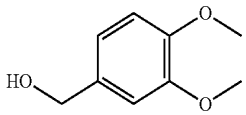
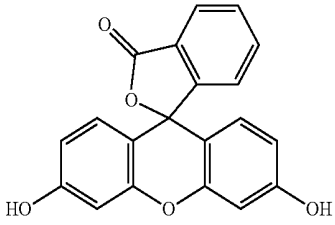
Allura Red AC	
Menadione	
p-cresol	
Fast green FCF	

TABLE 1-continued

Brilliant Blue FCF	
Methylisothiazolinone	
caffeine	
veratryl alcohol	
fluorescein	

[0027] A SEO suitable for use in the present disclosure can be an enzyme or chemical compound (e.g., peroxidase or ferricyanide, catalase, and the like). The SEO can be any molecule that can elicit the radical state of the SMA by performing a single electron oxidation. In one or more aspects, the SEO is an enzyme. For example, the SEO is an enzyme that can use H_2O_2 as the oxidant which is advantageous as H_2O_2 is produced as a coproduct by a CRO.

[0028] In one or more aspects, the SEO is an enzyme such as a laccase, horseradish peroxidase, Dyp-type peroxidase, lactoperoxidase, chloroperoxidase, manganese peroxidase 1, ascorbate peroxidase, dye-decolorizing peroxidase, unspecific peroxygenase, dehaloperoxidase, catalase-peroxidase, lignin peroxidase, soybean seed coat peroxidase, isoforms thereof or combinations thereof.

[0029] In one or more aspects, a catalytic composition comprises a GAO, an SEO and a SMA. The catalytic composition may comprise an GAO in an amount ranging from about 0.01 g/L to about 1 g/L, additionally or alternatively, from about 0.1 g/L to about 1 g/L, additionally or alternatively, from about 0.2 g/L to about 1 g/L, additionally or alternatively, from about 0.4 g/L to about 1 g/L, additionally or alternatively, from about 0.6 g/L to about 1 g/L, additionally or alternatively, from about 0.75 g/L to 1 g/L, additionally or alternatively, about 0.01 g/L, about 0.05 g/L, about 0.1 g/L, about 0.2 g/L, about 0.3 g/L, about 0.4 g/L, about 0.5 g/L, about 0.6 g/L, about 0.7 g/L, about 0.8 g/L, about 0.9 g/L or, additionally or alternatively, about 1 g/L. The catalytic composition may comprise an SEO in an amount ranging from about 1 mg/L to about 250 mg/L, additionally or alternatively, from about 5 mg/L to about 250 mg/L, additionally or alternatively, from about 10 mg/L to about 250 mg/L, additionally or alternatively, from about 25 mg/L to about 250 mg/L, additionally or alternatively, from

about 50 mg/L to find 250 mg/L, additionally or alternatively, from about 75 mg/L to 250 mg/L, additionally or alternatively, from about 100 mg/L to 250 mg/L, additionally or alternatively, from about 150 mg/L to 250 mg/L, additionally or alternatively, about 1 mg/L, about 5 mg/L, about 10 mg/L, about 15 mg/L, about 20 mg/L, about 25 mg/L, about 30 mg/L, about 35 mg/L, about 40 mg/L, about 45 mg/L, about 50 mg/L, about 55 mg/L, about 60 mg/L, about 65 mg/L, about 70 mg/L, about 75 mg/L, about 80 mg/L, about 85 mg/L, about 90 mg/L, about 95 mg/L, about 100 mg/L, about 105 mg/L, about 110 mg/L, about 115 mg/L, about 120 mg/L, about 125 mg/L, about 130 mg/L, about 135 mg/L, about 140 mg/L, about 145 mg/L, about 150 mg/L, about 155 mg/L, about 160 mg/L, about 165 mg/L, about 170 mg/L, about 175 mg/L, about 180 mg/L, about 185 mg/L, about 190 mg/L, about 195 mg/L, about 200 mg/L, about 205 mg/L, about 210 mg/L, about 215 mg/L, about 220 mg/L, about 225 mg/L, about 230 mg/L, about 235 mg/L, about 240 mg/L, about 245 mg/L or, additionally or alternatively, about 250 mg/L. The catalytic composition may comprise an SMA in an amount ranging from about 1 ppm to about 500 ppm, additionally or alternatively, from about 5 ppm to about 500 ppm, additionally or alternatively, from about 10 ppm to about 500 ppm, additionally or alternatively, from about 20 ppm to about 500 ppm, additionally or alternatively, or from about 40 ppm to about 400 ppm, additionally or alternatively, from about 50 ppm to about 350 ppm additionally or alternatively, from about 75 ppm to about 200 ppm additionally or alternatively, about 1 ppm, about 5 ppm, about 10 ppm, about 15 ppm, about 20 ppm, about 25 ppm, about 30 ppm, about 35 ppm, about 40 ppm, about 45 ppm, about 50 ppm, about 55 ppm, about 60 ppm, about 65 ppm, about 70 ppm, about 75 ppm, about 80 ppm, about 85 ppm, about 90 ppm, about 95 ppm, about 100 ppm, about 105 ppm, about 110 ppm, about 115 ppm, about 120 ppm, about 125 ppm, about 130 ppm, about 135 ppm, about 140 ppm, about 145 ppm, about 150 ppm, about 155 ppm, about 160 ppm, about 165 ppm, about 170 ppm, about 175 ppm, about 180 ppm, about 185 ppm, about 190 ppm, about 195 ppm, about 200 ppm, about 205 ppm, about 210 ppm, about 215 ppm, about 220 ppm, about 225 ppm, about 230 ppm, about 235 ppm, about 240 ppm, about 245 ppm, about 250 ppm, about 255 ppm, about 260 ppm, about 265 ppm, about 270 ppm, about 275 ppm, about 280 ppm, about 285 ppm, about 290 ppm, about 295 ppm, about 300 ppm, about 305 ppm, about 310 ppm, about 315 ppm, about 320 ppm, about 325 ppm, about 330 ppm, about 335 ppm, about 340 ppm, about 345 ppm, about 350 ppm, about 355 ppm, about 360 ppm, about 365 ppm, about 370 ppm, about 375 ppm, about 380 ppm, about 385 ppm, about 390 ppm, about 395 ppm, about 400 ppm, about 405 ppm, about 410 ppm, about 415 ppm, about 420 ppm, about 425 ppm, about 430 ppm, about 435 ppm, about 440 ppm, about 445 ppm, about 450 ppm, about 455 ppm, about 460 ppm, about 465 ppm, about 470 ppm, about 475 ppm, about 480 ppm, about 485 ppm, about 490 ppm, about 495 ppm, or, additionally or alternatively, about 500 ppm.

[0030] Conditions suitable for the formation of GDA may include one or more of the following reaction parameters: an amount of sugar (e.g., glucose) of from about 0.1 weight per volume percent (w/v %) to about 60 w/v %, alternatively from about 5 w/v % to about 50 w/v % or alternatively from about 10 w/v % to about 40 w/v %; a temperature ranging from about 1° C. to about 70° C., alternatively from about 5 to about 30° C. or alternatively from about 10° C. to about 25° C.; a suitable buffered media providing a pH ranging from about 5 to about 10, alternatively from about 5.5 to about 9 and a substrate amount ranging from about 6.5 to about 8.5, alternatively from about 6.5 to about 8.5 or alternatively from about 7 to about 8; an oxygen pressure of equal to or less than about 500 psi, alternatively from about 50 psi to about 250 psi or alternatively from about 70 psi to about 150 psi and a reaction time ranging from about 1 hour to about 24 hours or from about 2 hours to about 12 hours or from about 3 hours to about 6 hours.

[0031] GDA may be formed at yields that can range from about 50% to about 99%, additionally or alternatively, from about 50% to about 90%, additionally or alternatively, from about 50% to about 80%, additionally or alternatively, at least about 55%, at least about 58%, at least about 60%, at least about 62%, at least about 64%, at least about 66%, at least about 68%, at least about 70%, at least about 72%, at least about 74%, at least about 76%, at least about 78%, at least about 80%, at least about 82%, at least about 84%, at least about 86%, at least about 88%, at least about 90%, at least about 92%, at least about 94%, at least about 96%, at least about 98%, at least about or at least about 99%. The GDA intermediate formed may be used without further processing in the methods of the present disclosure. In alternative aspects, the GDA intermediate may be subjected to additional processing before (e.g., purification) prior to being utilized in other aspects of the presently disclosed methods. GDA also known as glucodialdose, L-gluco-hexodialdose, and D-gluco-hexodialdose, is a chiral intermediate with multiple industrial applications. The molecular structure corresponds with glucose, which inherently carries an aldehyde at the C1 position, oxidized to the dialdehyde at the C6 position.

[0032] In one or more aspects, a method of the present disclosure further comprises contacting the GDA formed with a second CRO (designated CRO2) under conditions suitable for the formation of glucaric acid. In one or more aspects, conditions suitable for the formation of glucaric acid include the presence of an SEO and/or SMA which may be the same as that used to facilitate the catalytic activity of CRO1. Alternatively, the SEO and/or SMA differ from those used to facilitate the catalytic activity of CRO1. In some aspects, CRO1 and CRO2 are the same enzyme, for example a GAO mutated to accept glucose as a substrate. In the alternative, CRO1 and CRO2 are different. For example, CRO 1 may be a GAO mutated to accept glucose as a substrate while CRO2 may be a GLOX mutated to oxidize GDA at the C1 and C6 positions to glucarate.

[0033] In one or more aspects, GDA is contacted with a catalytic composition comprising a GLOX, an SEO and a SMA. The catalytic composition may comprise a GLOX in an amount ranging from about 0.01 g/L to about 1 g/L, additionally or alternatively, from about 0.1 g/L to about 1 g/L, additionally or alternatively, from about 0.2 g/L to about 1 g/L, additionally or alternatively, from about 0.4 g/L to about 1 g/L, additionally or alternatively, from about 0.6 g/L to about 1 g/L, additionally or alternatively, from about 0.75 g/L to 1 g/L, additionally or alternatively, about 0.01 g/L, about 0.05 g/L, about 0.1 g/L, about 0.2 g/L, about 0.3 g/L, about 0.4 g/L, about 0.5 g/L, about 0.6 g/L, about 0.7 g/L, about 0.8 g/L, about 0.9 g/L or, additionally or alternatively, about 1 g/L. The catalytic composition may comprise an SEO in an amount ranging from about 1 mg/L to about 250

mg/L, additionally or alternatively, from about 5 mg/L to about 250 mg/L, additionally or alternatively, from about 10 mg/L to about 250 mg/L, additionally or alternatively, from about 25 mg/L to about 250 mg/L, additionally or alternatively, from about 50 mg/L to find 250 mg/L, additionally or alternatively, from about 75 mg/L to 250 mg/L, additionally or alternatively, from about 100 mg/L to 250 mg/L, additionally or alternatively, from about 150 mg/L to 250 mg/L, additionally or alternatively, about 1 mg/L, about 5 mg/L, about 10 mg/L, about 15 mg/L, about 20 mg/L, about 25 mg/L, about 30 mg/L, about 35 mg/L, about 40 mg/L, about 45 mg/L, about 50 mg/L, about 55 mg/L, about 60 mg/L, about 65 mg/L, about 70 mg/L, about 75 mg/L, about 80 mg/L, about 85 mg/L, about 90 mg/L, about 95 mg/L, about 100 mg/L, about 105 mg/L, about 110 mg/L, about 115 mg/L, about 120 mg/L, about 125 mg/L, about 130 mg/L, about 135 mg/L, about 140 mg/L, about 145 mg/L, about 150 mg/L, about 155 mg/L, about 160 mg/L, about 165 mg/L, about 170 mg/L, about 175 mg/L, about 180 mg/L, about 185 mg/L, about 190 mg/L, about 195 mg/L, about 200 mg/L, about 205 mg/L, about 210 mg/L, about 215 mg/L, about 220 mg/L, about 225 mg/L, about 230 mg/L, about 235 mg/L, about 240 mg/L, about 245 mg/L or, additionally or alternatively, about 250 mg/L. The catalytic composition may comprise an SMA in an amount ranging from about 1 ppm to about 500 ppm, additionally or alternatively, from about 5 ppm to about 500 ppm, additionally or alternatively, from about 10 ppm to about 500 ppm, additionally or alternatively, from about 20 ppm to about 500 ppm, additionally or alternatively, or from about 40 ppm to about 400 ppm, additionally or alternatively, from about 50 ppm to about 350 ppm additionally or alternatively, from about 75 ppm to about 200 ppm additionally or alternatively, about 1 ppm, about 5 ppm, about 10 ppm, about 15 ppm, about 20 ppm, about 25 ppm, about 30 ppm, about 35 ppm, about 40 ppm, about 45 ppm, about 50 ppm, about 55 ppm, about 60 ppm, about 65 ppm, about 70 ppm, about 75 ppm, about 80 ppm, about 85 ppm, about 90 ppm, about 95 ppm, about 100 ppm, about 105 ppm, about 110 ppm, about 115 ppm, about 120 ppm, about 125 ppm, about 130 ppm, about 135 ppm, about 140 ppm, about 145 ppm, about 150 ppm, about 155 ppm, about 160 ppm, about 165 ppm, about 170 ppm, about 175 ppm, about 180 ppm, about 185 ppm, about 190 ppm, about 195 ppm, about 200 ppm, about 205 ppm, about 210 ppm, about 215 ppm, about 220 ppm, about 225 ppm, about 230 ppm, about 235 ppm, about 240 ppm, about 245 ppm, about 250 ppm, about 255 ppm, about 260 ppm, about 265 ppm, about 270 ppm, about 275 ppm, about 280 ppm, about 285 ppm, about 290 ppm, about 295 ppm, about 300 ppm, about 305 ppm, about 310 ppm, about 315 ppm, about 320 ppm, about 325 ppm, about 330 ppm, about 335 ppm, about 340 ppm, about 345 ppm, about 350 ppm, about 355 ppm, about 360 ppm, about 365 ppm, about 370 ppm, about 375 ppm, about 380 ppm, about 385 ppm, about 390 ppm, about 395 ppm, about 400 ppm, about 405 ppm, about 410 ppm, about 415 ppm, about 420 ppm, about 425 ppm, about 430 ppm, about 435 ppm, about 440 ppm, about 445 ppm, about 450 ppm, about 455 ppm, about 460 ppm, about 465 ppm, about 470 ppm, about 475 ppm, about 480 ppm, about 485 ppm, about 490 ppm, about 495 ppm, or, additionally or alternatively, about 500 ppm.

[0034] In one or more aspects, the reaction disclosed herein is carried out under mild conditions. Conditions suitable for the formation of glucaric acid may include one or more of the following reaction parameters: an amount of GDA of from about 0.1 weight per volume percent (w/v %) to about 60 w/v %, alternatively from about 5 w/v % to about 50 w/v % or alternatively from about 10 w/v % to about 40 w/v %; a temperature ranging from about 5° C. to about 100° C., or from about 10° C. to about 75° C., or from about 15° C. to about 70° C., or from about 25° C. to about 50° C.; a suitable buffered media providing a pH ranging from about 4 to about 9, or from about 5.5 to about 9, or from about 6 to about 8; an oxygen pressure of equal to or less than about 500 psi, alternatively from about 50 psi to about 250 psi or alternatively from about 70 psi to about 150 psi and a reaction time ranging from about 1 hour to about 24 hours or from about 2 hours to about 12 hours or from about 3 hours to about 6 hours.

[0035] Glucaric acid may be formed at yields that can range from about 50% to about 99%, additionally or alternatively, from about 50% to about 90%, additionally or alternatively, from about 50% to about 80%, additionally or alternatively, at least about 55%, at least about 58%, at least about 60%, at least about 62%, at least about 64%, at least about 66%, at least about 68%, at least about 70%, at least about 72%, at least about 74%, at least about 76%, at least about 78%, at least about 80%, at least about 82%, at least about 84%, at least about 86%, at least about 88%, at least about 90%, at least about 92%, at least about 94%, at least about 96%, at least about 98%, at least about or at least about 99%.

[0036] In some aspects, the reaction for formation of glucaric acid further includes a caustic agent such as sodium hydroxide or another base which functions to both control pH and as a counterion for produced glucaric acid or glucarate salts.

[0037] In some aspects, the catalyst compositions disclosed herein further comprise a catalase in order to decompose the hydrogen peroxide produced during the reaction of the CRO. A catalase is used to break down two molecules of hydrogen peroxide (H_2O_2) into one molecule of oxygen (O_2) and two molecules of water (H_2O). In some aspects, oxygen generated by catalase may be captured and used in the oxidation of additional substrate (e.g., glucose or GDA).

[0038] In one or more aspects, the methods of the present disclosure produce high purity glucaric acid or salts thereof, for example, glucaric acid may have a purity of greater than about 70%, or greater than about 90% or greater than about 95%, or from about 70% to about 99%, or from about 72% to about 99%, or from about 75% to about 99%, or from about 78% to about 99%, or from about 80% to about 99%, or from about 82% to about 99%, or from about 84% to about 99%, or from about 86% to about 99%, or from about 88% to about 99%, or from about 90% to about 99%, or from about 92% to about 99%, or from about 94% to about 99%, or from about 96% to about 99%. In an aspect, the product comprises disodium glucarate, monosodium glucarate, potassium glucarate or combinations thereof in crystalline form. In an alternative aspect, the product of the present disclosure comprises a mixture of free acid in combination with salts of glucaric acid (e.g., potassium and/or sodium salt).

[0039] In one or more aspects, CRO1 which catalyzes the formation of glucodialdose can be any CRO or mutant thereof capable of oxidizing the C6 of glucose. In one or more aspects, CRO2 which catalyzes the formation of glucaric acid can be any CRO capable of oxidizing the C1

and C6 aldehydes of glucodialdose. In one or more aspects, a GLOX engineered to oxidize the C1 and C6 of GDA could catalyze this reaction.

[0040] In an aspect, an enzyme of the type disclosed herein is a wild type enzyme, a functional fragment thereof, or a functional variant thereof. "Fragment" as used herein is meant to include any amino acid sequence shorter than the full-length enzyme, but where the fragment maintains a catalytic activity sufficient to meet some user or process goal. Fragments may include a single contiguous sequence identical to a portion of the biocatalyst sequence. Alternatively, the fragment may have or include several different shorter segments where each segment is identical in amino acid sequence to a different portion of the amino acid sequence of the enzyme but linked via amino acids differing in sequence from the enzyme. Herein, a "functional variant" of the enzyme refers to a polypeptide which has at one or more positions of an amino acid insertion, deletion, or substitution, either conservative or non-conservative, and wherein each of these types of changes may occur alone, or in combination with one or more of the others, and/or one or more times in a given sequence but retains catalytic activity.

[0041] In the alternative or in combination with the aforementioned mutations, the enzyme may be mutated to improve the catalytic activity. Mutations may be carried out to enhance the protein or a homolog activity, increase the protein stability in the presence of substrates and products (e.g., hydrogen peroxide) and increase protein yield.

[0042] Herein, reference has been made to "sources" of enzyme. It is to be understood this refers to the biomolecule as expressed by the named organism. It is contemplated the enzyme may be obtained from the organism or a version of said enzyme (wildtype or recombinant) and provided as a suitable construct to an appropriate expression system.

[0043] In an aspect, any enzyme of the type disclosed herein may be cloned into an appropriate expression vector and used to transform cells of an expression system such as *E. coli, Saccharomyces* sp., *Pichia* sp., *Aspergillus* sp., *Trichoderma* sp., or *Myceliophthora* sp. A "vector" is a replicon, such as plasmid, phage, viral construct or cosmid, to which another DNA segment may be attached. Vectors are used to transduce and express a DNA segment in cells. As used herein, the terms "vector" and "construct" may include replicons such as plasmids, phage, viral constructs, cosmids, Bacterial Artificial Chromosomes (BACs), Yeast Artificial Chromosomes (YACs) Human Artificial Chromosomes (HACs) and the like into which one or more gene expression cassettes may be or are ligated. Herein, a cell has been "transformed" by an exogenous or heterologous nucleic acid or vector when such nucleic acid has been introduced inside the cell, for example, as a complex with transfection reagents or packaged in viral particles. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell.

[0044] In an aspect, the gene of an enzyme disclosed herein is provided as a recombinant sequence in a vector where the sequence is operatively linked to one or more control or regulatory sequences. "Operatively linked" expression control sequences refer to a linkage in which the expression control sequence is contiguous with the gene of interest to control the gene of interest, as well as expression control sequences that act in trans or at a distance to control the gene of interest.

[0045] The term "expression control sequence" or "regulatory sequences" are used interchangeably and are used herein refer to polynucleotide sequences which affect the expression of coding sequences to which they are operatively linked. Expression control sequences are sequences that control the transcription, post-transcriptional events, and translation of nucleic acid sequences. Expression control sequences include appropriate transcription initiation, termination, promoter, and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (e.g., ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to include, at a minimum, all components whose presence is essential for expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

[0046] The term "recombinant host cell" ("expression host cell", "expression host system", "expression system" or simply "host cell"), as used herein, is intended to refer to a cell into which a recombinant vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein. A recombinant host cell may be an isolated cell or cell line grown in culture or may be a cell which resides in a living tissue or organism.

[0047] In one or more aspects, a CRO suitable for use in the present disclosure is a CRO comprising any of SEQ ID Nos. 1 through 39 or alternatively from about 85% to about 100% sequence identity with any of SEQ ID Nos. 1 through 39. In one or more aspects, a catalase suitable for use in the present disclosure is a catalase defined by of any of SEQ ID Nos. 40 through 44. In one or more aspects, a peroxidase suitable for use in the present disclosure is a peroxidase defined by of any of SEQ ID Nos. 45 through 61 In one or more aspects, a CRO suitable for use in the present disclosure is a CRO defined by of any of SEQ ID Nos.62 through or alternatively from about 85% to about 100% sequence identity with any of SEQ ID Nos. 62 through 65.

[0048] In one or more aspects, a glucaric acid product of the type disclosed herein may be used in the production of into adipic acid and subsequently adipic adipodinitrile, which can then be hydrogenated to produce hexamethyldiamine.

ADDITIONAL DISCLOSURE

[0049] A first aspect which is a method of preparing glucaric acid comprising contacting D-glucose and oxygen with a first catalyst composition comprising a first copper radical oxidase, a single electron oxidizer, and a small molecule activator under conditions suitable for formation of glucodialdose; and contacting glucodialdose and oxygen with a second catalyst composition comprising a second copper radical oxidase, a single electron oxidizer and a small

molecule activator under conditions suitable for the formation of a product mixture comprising glucaric acid or salts thereof.

[0050] A second aspect which is the method of the first aspect wherein the first copper radical oxidase comprises a galactose oxidase, a galactose oxidase mutant or combinations thereof.

[0051] A third aspect which is the method of any of the first through second aspects wherein the first copper radical oxidase comprises any of SEQ ID Nos. 1 through 39.

[0052] A fourth aspect which is the method of any of the first through third aspects wherein the first copper radical oxidase has from about 85% to about 100% sequence identity with any of SEQ ID Nos. 1 through 39.

[0053] A fifth aspect which is the method of any of the first through fourth aspects wherein the second copper radical oxidase comprises a glyoxal oxidase, a glyoxal oxidase mutant or combinations thereof.

[0054] A sixth aspect which is the method of any of the first through fifth aspects wherein the second copper radical oxidase comprises any of SEQ ID Nos. 62 through 65.

[0055] A seventh aspect which is the method of any of the first through sixth aspects wherein the second copper radical oxidase has from about 85% to about 100% sequence identity with any of SEQ ID Nos. 62 through 65.

[0056] An eighth aspect which is the method of any of the first through seventh aspects wherein the small molecule activator comprises L-tryptophan, 2-mercaptobenzothiazole, L-histidine, methylchloroisothiazolinone, o-dianisidine, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), 4-aminoantipyrine, L-tyrosine, (2,2,6,6-tetramethylpiperidin-1-yl)oxyl, chloromethylisothiazolinone, 4-thiazolecarboxylic acid, Sunset yellow FCF, tartrazine, p-benzoquinone, dicoumarol, phthalimide, saccharin, phthalic anhydride, erythrosine B, 2-aminobenzothiazole, thiabendazole, 2-hydroxybenzothiazole, phenothiazine, 6-aminobenzothiazole, indigo carmine, naphthalimide, 2-aminothiazole, thiazole, 2H-1,4-benzothiazin-3 (4H)-one, 2-oxindole, beta-lapachone, menaquinone, thiamine, 4-methyl-5-thiazoleethanol, Allura Red AC, menadione, p-cresol, Fast green FCF, Brilliant Blue FCF, methylisothiazolinone, caffeine, veratryl alcohol, fluorescein, or combinations thereof.

[0057] A ninth aspect which is the method of any of the first through eighth aspects wherein the small molecule activator is present in an amount ranging from about 1 ppm to about 500 ppm.

[0058] A tenth aspect which is the method of any of the first through ninth aspects wherein the single electron oxidizer comprises laccase, horseradish peroxidase, dyp-type peroxidase, lactoperoxidase, chloroperoxidase, manganese peroxidase 1, ascorbate peroxidase, dye-decolorizing peroxidase, unspecific peroxygenase, dehaloperoxidase, catalase-peroxidase, lignin peroxidase, soybean seed coat peroxidase, isoforms thereof or combinations thereof.

[0059] An eleventh aspect which is the method of any of the first through tenth aspects wherein conditions suitable for formation of glucodialdose, conditions suitable for formation of a product mixture comprising glucaric acid or both comprise an oxygen pressure of equal to or less than about 500 psi.

[0060] A twelfth aspect which is the method of any of the first through eleventh aspects further comprising a catalase.

[0061] A thirteenth aspect which is the method of any of the first through twelfth aspects wherein the catalase is defined by any SEQ ID No. 40 through SEQ ID No. 44.

[0062] A fourteenth aspect which is the method of any of the first through thirteenth aspects further comprising introducing a caustic to the second catalyst composition, the product mixture or both.

[0063] A fifteenth aspect which is the method of any of the first through fourteenth aspects wherein the glucodialdose, glucaric acid are formed at yields ranging from about 50% to about 99%.

[0064] A sixteenth aspect which is the method of any of the first through fifteenth aspects wherein the glucaric acid or salts thereof have a purity of greater than about 70%.

[0065] A seventeenth aspect which is the method of any of the first through sixteenth aspects wherein the product mixture comprising glucaric acid further comprises on or more sugar oxidation products.

[0066] An eighteenth aspect which is the method of the seventeenth aspect wherein the one or more sugar oxidation products comprise aldonic acid, uronic acid, aldaric acid, a gluconic acid oxidation product, a gluconate, gluconic acid, glucuronic acid, glucose oxidation products, galactonic acid, galactaric acid, glutamic acid, a lactone of gluconic acid, a lactone of glucaric acid, a lactone of galactaric acid, a lactone of galactonic acid, glucodialdose, 2-ketoglucose, disaccharides, oxidized disaccharides, n-keto-acids, C2 to C6 diacids, salts thereof or combinations thereof.

[0067] A nineteenth aspect which is a method comprising contacting a sugar with a catalyst composition comprising an oxidoreductase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of one or more oxidized sugar oxidation products comprising glucaric acid wherein the oxidoreductase comprises at least two copper radical oxidases, at least two mutated copper radical oxidases or combinations thereof.

[0068] A twentieth aspect which is the method of the nineteenth aspect wherein the at least two copper radical oxidases comprises any of SEQ ID Nos. 1 through 39, any of SEQ ID Nos. 62 through 65, have 85% to about 100% sequence identity with any of SEQ ID Nos. 1 through 39 or have 85% to about 100% sequence identity with any of SEQ ID Nos. 62 through 65.

EXAMPLES

[0069] The aspects having been generally described, the following example is given as particular aspects of the disclosure and to demonstrate the practice and advantages thereof. It is understood that the example is given by way of illustration and is not intended to limit the specification or the claims in any manner.

Example 1

Generation of GAO Mutant for Producing Glucodialdose from Glucose

[0070] A GAO mutant capable of converting glucose to GDA was engineered. Following directed evolution and rational enzyme engineering, the improved GAO mutant exhibits a specific activity of 35 U mg^{-1} on glucose.

Directed Evolution

[0071] Directed evolution of thirty sites within 10 Å of the catalytic copper was performed on a parent sequence con-

taining the following added mutations: 1) R330, Q406T, W290F to introduce less than 1 U mg^{-1} activity on glucose to GAO, 2) C383S to lower the KM of the enzyme on galactose, and 3) Y405F and Q406E to enhance activity on a D-N-acetyl glucosamine substrate. Other mutations described in Table I were found to have neutral or deleterious effects on glucodialdose-generating activity. The new combination sequence was designated GAO-Mut1. The sequence of GAO-Mut1 contains a "MGHHHHHHSSGHIEGRHM" N-terminal his-tag and linker for expression and purification in *E. coli* and is SEQ ID NO:66.

[0072] While preferred embodiments have been shown and described, modifications thereof can be made by one skilled in the art without departing from the scope or teachings herein. The embodiments described herein are exemplary only and are not limiting. Many variations and modifications of the systems, apparatus, and processes described herein are possible and are within the scope of the disclosure. For example, the relative dimensions of various parts, the materials from which the various parts are made, and other parameters can be varied. Accordingly, the scope of protection is not limited to the embodiments described herein, but is only limited by the claims that follow, the scope of which shall include all equivalents of the subject matter of the claims. Unless expressly stated otherwise, the steps in a method claim may be performed in any order. The recitation of identifiers such as (a), (b), (c) or (1), (2), (3) before steps in a method claim are not intended to and do not specify a particular order to the steps, but rather are used to simplify subsequent reference to such steps.

[0073] Each and every claim is incorporated into the specification as an aspect of the present disclosure. Thus, the claims are a further description and are an addition to the aspects of the present invention. The discussion of a reference herein is not an admission that it is prior art to the presently disclosed subject matter, especially any reference that may have a publication date after the priority date of this application. The disclosures of all patents, patent applications, and publications cited herein are hereby incorporated by reference, to the extent that they provide exemplary, procedural or other details supplementary to those set forth herein.

SEQUENCE LISTING

Sequence total quantity: 66

SEQ ID NO: 1 moltype = AA length = 657
FEATURE Location/Qualifiers
source 1..657
mol_type = protein
organism = synthetic construct

SEQUENCE: 1

```
MGHHHHHHSS GHIEGRHMAS APIGSAIPRN NWAVTCDSAQ SGNECNKAID GNKDTFWHTF  60
YGANGDPKPP HTYTIDMKTT QNVNGLSVLP RQDGNQNGWI GRHEVYLSSD GTNWGSPVAS  120
GSWFADSTTK YSNFETRPAR YVRLVAITEA NGQPWTSIAE INVFQASSYT APQPGLGRWG  180
PTIDLPIVPA AAAIEPTSGR VLMWSSYRND APEGSPGGIT LTSSWDPSTG IVSDRTVTVT  240
KHDMFCPGIS MDGNGQIVVT GGNDAKKTSL YDSSSDSWIP GPDMQVARGY QSSATMSDGR  300
VFTIGGSFSG GVFEKNGEVY SPSSKTWTSL PNAKVNPMLT ADKQGLYKSD NHAWLFGWKK  360
GSVFQAGPST AMNWYYTSGS GDVKSAGKRQ SNRGVAPDAM SGNAVMYDAV KGKILTFGGS  420
PDFEDSDATT NAHIITLGEP GTSPNTVFAS NGLYFARTFH TSVVLPDGST FITGGQRRGI  480
PFEDSTPVFT PEIYVPEQDT FYKQNPNSIV RAYHSISLLL PDGRVFNGGG GLCGDCTTNH  540
FDAQIFTPNY LYDSNGNLAT RPKITRTSTQ SVKVGGRITI STDSSISKAS LIRYGTATHT  600
VNTDQRRIPL TLTNNGGNSY SFQVPSDSGV ALPGYWMLFV MNSAGVPSVA STIRVTQ     657
```

SEQ ID NO: 2 moltype = AA length = 639
FEATURE Location/Qualifiers
source 1..639
mol_type = protein
organism = synthetic construct

SEQUENCE: 2

```
ASAPIGSAIS RNNWAVTCDS AQSGNECNKA IDGNKDTFWH TFYGANGDPK PPHTYTIDMK  60
TTQNVNGLSM LPRQDGNQNG WIGRHEVYLS SDGTNWGSPV ASGSWFADST TKYSNFETRP  120
ARYVRLVAIT EANGQPWTSI AEINVFQASS YTAPQPGLGR WGPTIDLPIV PAAAAIEPTS  180
GRVLMWSSYR NDAFGGSPGG ITLTSSWDPS TGIVSDRTVT VTKHDMFCPG ISMDGNGQIV  240
VTGGNDAKKT SLYDSSSDSW IPGPDMQVAR GYQSSATMSD GRVFTIGGSW SGGVFEKNGE  300
VYSPSSKTWT SLPNAKVNPM LTADKQGLYR SDNHAWLFGW KKGSVFQAGP STAMNWYYTS  360
GSGDVKSAGK RQSNRGVAPD AMCGNAVMYD AVKGKILTFG GSPDYQDSDA TTNAHIITLG  420
EPGTSPNTVF ASNGLYFART FHTSVVLPDG STFITGGQRR GIPFEDSTPV FTPEIYVPEQ  480
DTFYKQNPNS IVRVYHSISL LLPDGRVFNG GGGLCGDCTT NHFDAQIFTP NYLYNSNGNL  540
ATRPKITRTS TQSVKVGGRI TISTDSSISK ASLIRYGTAT HTVNTDQRRI PLTLTNNGGN  600
SYSFQVPSDS GVALPGYWML FVMNSAGVPS VASTIRVTQ                          639
```

SEQ ID NO: 3 moltype = AA length = 680
FEATURE Location/Qualifiers
source 1..680
mol_type = protein
organism = synthetic construct

SEQUENCE: 3

```
MKHFLSLALC FSSINAVAVT VPHKSGGTGS PEGSLQFLSL RASAPIGSAI SRNNWAVTCD  60
SAQSGNECNK AIDGNKDTFW HTFYGANGDP KPPHTYTIDM KTTQNVNGLS MLPRQDGNQN  120
GWIGRHEVYL SSDGTNWGSP VASGSWFADS TTKYSNFETR PARYVRLVAV TEANGQPWTS  180
IAEINVFQAS SYTAPQPGLG RWGPTIDLPI VPAAAAIEPT SGRVLMWSSY RNDAFGGSPG  240
```

-continued

```
GITLTSSWDP STGIVSDRTV TVTKHDMFCP GISMDGNGQI VVTGGNDAKK TSLYDSSSDS  300
WIPGPDMQVA RGYQSSATMS DGRVFTIGGS WSGGVFEKNG EVYSPSSKTW TSLPNAKVNP  360
MLTADKQGLY RSDNHAWLFG WKKGSVFQAG PSTAMNWYYT SGSGDVKSAG KRQSNRGVAP  420
DAMCGNAVMY DAVKGKILTF GGSPDYQDSD ATTNAHIITL GEPGTSPNTV FASNGLYFAR  480
TFHTSVVLPD GSTFITGGQR RGIPFEDSTP VFTPEIYVPE QDTFYKQNPN SIVRVYHSIS  540
LLLPDGRVFN GGGGLCGDCT TNHFDAQIFT PNYLYNSNGN LATRPKITRT STQSVKVGGR  600
ITISTDSSIT KASLIRYGTA THTVNTDQRR IPLTLTNNGG NSYSFQVPSD SGVALPGYWM  660
LFVMNSAGVP SVASTIRVTQ                                              680

SEQ ID NO: 4          moltype = AA  length = 680
FEATURE               Location/Qualifiers
source                1..680
                      mol_type = protein
                      organism = Fusarium graminearum
SEQUENCE: 4
MKHFLSLALC FSSINAVAVT VPHKSGGTGS PEGSLQFLSL RASAPIGSAI SRNNWAVTCD  60
SAQSGNECNK AIDGNKDTFW HTFYGANGDP KPPHTYTIDM KTTQNVNGLS MLPRQDGNQN  120
GWIGRHEVYL SSDGTNWGSP VASGSWFADS TTKYSNFETR PARYVRLVAV TEANGQPWTS  180
IAEINVFQAS SYTAPQPGLG RWGPTIDLPI VPAAAAIEPT SGRVLMWSSY RNDAFGGSPG  240
GITLTSSWDP STGIVSDRTV TVTKHDMFCP GISMDGNGQI VVTGGNDAKK TSLYDSSSDS  300
WIPGPDMQVA RGYQSSATMS DGRVFTIGGS WSGGVFEKNG EVYSPSSKTW TSLPNAKVNP  360
MLTADKQGLY RSDNHAWLFG WKKGSVFQAG PSTAMNWYYT SGSGDVKSAG KRQSNRGVAP  420
DAMCGNAVMY DAVKGKILTF GGSPDYQDSD ATTNAHIITL GEPGTSPNTV FASNGLYFAR  480
TFHTSVVLPD GSTFITGGQR RGIPFEDSTP VFTPEIYVPE QDTFYKQNPN SIVRVYHSIS  540
LLLPDGRVFN GGGGLCGDCT TNHFDAQIFT PNYLYNSNGN LATRPKITRT STQSVKVGGR  600
ITISTDSSIT KASLIRYGTA THTVNTDQRR IPLTLTNNGG NSYSFQVPSD SGVALPGYWM  660
LFVMNSAGVP SVASTIRVTQ                                              680

SEQ ID NO: 5          moltype = AA  length = 559
FEATURE               Location/Qualifiers
source                1..559
                      mol_type = protein
                      organism = Phanerodontia chrysosporium
SEQUENCE: 5
MLSLLAVVSL AAATLAAPAA SDAPGWRFDL KPNLSGIVAL EAIVVNSSLV VIFDRATGDQ  60
PLKINGESTW GALWDLDTST VRPLSVLTDS FCASGALLSN GTMVSMGGTP GGTGGDVAAP  120
PGNQAIRIFE PCASPSGDGC TLFEDPATVH LLEERWYPSS VRIFDGSLMI IGGSHVLTPF  180
YNVDPANSFE FFPSKEQTPR PSAFLERSLP ANLFPRAFAL PDGTVFIVAN NQSIIYDIEK  240
NTETILPDIP NGVRVTNPID GSAILLPLSP PDFIPEVLVC GGSTADTSLP STSLSSQHPA  300
TSQCSRIKLT PEGIKAGWQV EHMLEARMMP ELVHVPNGQI LITNGAGTGF AALSAVADPV  360
GNSNADHPVL TPSLYTPDAP LGKRISNAGM PTTTIPRMYH STVTLTQQGN FFIGGNNPNM  420
NFTPPGTPGI KFPSELRIET LDPPFMFRSR PALLTMPEKL KFGQKVTVPI TIPSDLKASK  480
VQVALMDLGF SSHAFHSSAR LVFMESSISA DRKSLTFTAP PNGRVFPPGP AVVFLTIDDV  540
TSPGERVMMG SGNPPPTLE                                               559

SEQ ID NO: 6          moltype = AA  length = 558
FEATURE               Location/Qualifiers
source                1..558
                      mol_type = protein
                      organism = Pycnoporus cinnabarinus
SEQUENCE: 6
MAPTAFSLVS ALALASLSLA APSAPGWSFD LKKETSGIVA LEAIVVSPTL VVFFDRASDD  60
PLQINNHSAW GALWNLESST VRPLDVLTNS FCASGALLSN GTMASVGGDP DGFVGNPAIR  120
PGNQAIRLFE PCDSPTGDGC TLFEDPATLH LLEKRWYPSS ARIFDGSLII VGGMHEATPF  180
YNTDPALSFE FFPRKEDTPR PSEFLNRSLP ANLFPRVFAL PDGKVFMVAN NQSIIYDIEA  240
KTERILPDVP NNVRVTNPMD GSAILLPLSP PDFVPEVLVC GGSQTDTIDP SLLTSQTPAS  300
SQCSRIRLDE EGIAKGWEVE HMLEGRIMPE LVHLPNGQVL IANGGRTGFA AIASVSEPVG  360
NSNADHPVLV PSLYTPDAPL GRRISNVGLP SSGIPRLYHS SVTLTPQGNF LIAGSNPNNR  420
TTVGPGIKFP SEFRVQTLDP PFMSVERPKI LNMPKKLGFN KSFTVPISVP SSLARPGAKV  480
QISLMDLGFS SHAFHSSARL VFMNGKISQD SKSLTFTTPP NGRVYPPGPA TVFLTIDDVT  540
SEGAWVMMGS GNSPPTLE                                                558

SEQ ID NO: 7          moltype = AA  length = 506
FEATURE               Location/Qualifiers
source                1..506
                      mol_type = protein
                      organism = Colletotrichum graminicola
SEQUENCE: 7
MPTLRSALRN LPAALLALAA ACEAQNVGKW GPMVKFPVVP VAVALVPETG NLLVWSSGWP  60
NRWTTAGNGK TYTSLYNVNT GNISDAIVQN TQHDMFCPGT SLDADGRIIV TGGSSAAKTS  120
VLDFKKGESS PWTPLSNMQI SRGYQSSCTT SEGKIFVIGG SFSGAGTRNG EVYDPKANTW  180
TKLAGCPVKP LVMQRGMFPD SHAWLWSWKN GSVLQAGPSK KMNWYDTKGT GSNTPAGLRG  240
TDEDSMCGVS VMYDAVAGKI FTYGGGKGYT GYDSTSNAHI LTLGEPGQAV QVQKLANGKY  300
NRGFANAVVM PDGKIWVVGG MQKMWLFSDT TPQLTPELFD PATGSFTPTT PHTVPRNYHS  360
TALLMADATI WSGGGGLCGA NCKENHFDGQ FWSPPYLFEA DGVTPAKRPV IQSLSDTAVR  420
AGAPITITMQ DAGAYTFSMI RVSATTHTVN TDQRRIPLDG QDGGDGKSFT VNVPNDYGVA  480
```

-continued

```
IPGYYMLFAM NEAGVPCVAQ FFKVTL                                          506

SEQ ID NO: 8           moltype = AA  length = 497
FEATURE                Location/Qualifiers
source                 1..497
                       mol_type = protein
                       organism = Colletotrichum gloeosporioides
SEQUENCE: 8
EAEAQGLGQW SPLIKFPVVP VSVALLPESG NLLVWSSGWP NRWTTAGNGK TYTSLYNVQT  60
GNVSDAVIQN TQHDMFCPGT SLDAEGRIIV TGGSSAAKTS VLDFKNGESS SWTALSNMQI  120
SRGYQSSCTT SEGKIFVIGG SFSGAGTRNG EIYDTATNKW TKLAGCPVKP LVMQLGMFPD  180
SHAWLWSWKN GSVLQAGPAK QMNWYDTKGT GANTPAGLRG ADQDSMCGVS VMYDAVAGKI  240
FTYGGGKGYT GYQSTSNAHI LTLGEPGQQV QVQKLQNGQY NRGFANAVVM PDGKIWVVGG  300
MKQMALFSDA TPQLTPELFD PATGKFTPTA AHTVPRNYHS TALLMADGTI WSGGGGLCGA  360
GCAANKFDGQ FWSPPYLFEA DGKTPAKRPV IESLSDETVK AGAALTINMQ DEGKYTFSMI  420
RVSATTHTVN TDQRRIPLDG QDGGDGKSFS VNMPSDYGVV IPGYYMMFAM NEAGTPCVAK  480
FFKVSLHHHH HHHHHHV                                                 497

SEQ ID NO: 9           moltype = AA  length = 711
FEATURE                Location/Qualifiers
source                 1..711
                       mol_type = protein
                       organism = Colletotrichum graminicola
SEQUENCE: 9
MVRSCAYKAI AAASLLSQLA SAAITSCPNN ETVWETPIGV KYTLCPGSDY QNGGASLQTV  60
RDIQSSLECA KICDSDARCN RAVYDNVNKA CDVKNSTNPM QWAADDRFET IRLTNDLPEG  120
AFISTCSFNE TSYRVPETNA EYRICPDTDY TGVNAKVVEG VTTIQACAEL CSNTQDCRKS  180
VFDHINNACA IKAAEPATSI FWVQDKQFST IRLPENIDPA VKGKWGDLIR LPVIPVAAYI  240
VPSYPEPSRL LFFSSWSNDA FSGASGMTQF GDYDFATGAI SQRTVTNTHH DMFCPGISQL  300
EDGRILIQGG SDADTVSIYD PATNEFTRGP NMTLARGYQT SCTLSNGKVF TIGGAYSGER  360
VGKNGEVYDP VANAWTYLPG ADFRPMLTND HEGIWREDNH AWLFGWKNGS IFQAGPSKDQ  420
HWYGIQGNGT VAKAATRDDD DAMCGVWVMY DAVAGKIFSA GGSPDYTDSP ATQRAHITTI  480
GEPNTPAEVE RVADMGFPRG FANAVVLPDG QVLVTGGQRM SLVFTNTDGI LVAELFNPET  540
REWKQMAPMA VPRNYHSVSI LLPDATVFSG GGGMCWVQNV GDSTAGCDKT VDHSDGEIFE  600
PPYLFNEDGS RAARPVISAI SADPIKAGAT LTFTVEGVEG QGTAALIRLG SVTHSVNSDQ  660
RRVPLNVTVS GNEYSATLPD DYGILLPGYY YLFVSTPQGT PSIAKTVHVI L           711

SEQ ID NO: 10          moltype = AA  length = 611
FEATURE                Location/Qualifiers
source                 1..611
                       mol_type = protein
                       organism = Streptomyces lividans
SEQUENCE: 10
STEKYHQYKI NQPEYKAANG KWEIIEFPEK YRQNTIHAAL LRTGKVLMVA GSGNNQDNSD  60
DKQYDTRIWD PVKGTIKKVP TPSDLFCTGH TQLANGNLLI AGGTKRYEKL KGDVTKAGGL  120
MVVHNENPDK PITLPAGTKF TGKENGKTFV SKDPVLVPRA EKVFDPATGA FVRNDPGLGR  180
IYVEAQKSGS AYETGTEDNY RVQGLSGADA RNTYGIAQKL ALDKKDFQGI RDAFEFDPVA  240
EKYIKVDPMH EARWYPTLTT LGDGKILSVS GLDDIGQLVP GKNEVYDPKT KAWTYTDKVR  300
QFPTYPALFL MQNGKIFYSG ANAGYGPDDV GRTPGVWDVE TNKFTKVPGM SDANMLETAN  360
TVLLPPAQDE KYMVIGGGGV GESKLSSEKT RIADLKADDP KFVDGPSLEK GTRYPQASIL  420
PDDSVLVSGG SQDYRGRGDS NILQARLYHP DTNEFERVAD PLVGRNYHSG SILLPDGRLM  480
FFGSDSLYAD KANTKPGKFE QRIEIYTPPY LYRDSRPDLS GGPQTIARGG SGTFTSRAAS  540
TVKKVRLIRP SASTHVTDVD QRSIALDFKA DGDKLTVTVP SGKNLVQSGW YMMFVTDGEG  600
TPSKAEWVRV P                                                       611

SEQ ID NO: 11          moltype = AA  length = 646
FEATURE                Location/Qualifiers
source                 1..646
                       mol_type = protein
                       organism = Actinobacteria bacterium
SEQUENCE: 11
MSELGGGRRR ARRLAVGTVV VLALAGMNGP WVYRFSTEKY HEYAINRPEY KAENGKWDII  60
QFPKKYRQDT IHAVLLHTGK VLLIAGSGND QDNFDAKKFD TRIWDPVKGT IKKVPTPADL  120
FCTGHTQLAN GNILIAGGTK KYEKLKGDVK KAGGVMILYN ENPDKPITLP KGTVFTGKQN  180
GKSFVSTDSS LIPRATKVFD KKTGKFLRNN PGYARVVVEA PKEGTTYETG TEDNYTIQGL  240
TGADERNTYG IANKLGMDKR DFEGIKNAYE FDPVAEKYIK VDPMNEARWY PTLTTLSDGK  300
VLSVSGLDDI GQLVPGKNEI FDPKTKTWTY TKTKRQFPTY PALFLMQNGK LFYSGSNAGY  360
GPDDVGRDPG IWDVDTNAFT KLPGLSDPNM LETSATVLLA PAQEEKFMVI GGGGVGESKL  420
STNKTRLIDL KDANPKFVDG PTLEKGTRYA QASVLPDDSV LVSGGSQDYR GRSASDIKQA  480
RIYNPTTNTF KRVADPEVGR DYHSGSILLP DGRVMFFGSN PLYADKANTK PGVFEQRIEI  540
YTPPYLYRDA RPDLSGGPGT IARGASGTFT SKHASSIRKV RLIRPSSSTH VTDVDQRSIA  600
LDYTTSGDKI TVTVPKNRNL VPSGWYMMFV DDDQGTPSKA VWVDVP                 646

SEQ ID NO: 12          moltype = AA  length = 666
FEATURE                Location/Qualifiers
source                 1..666
```

-continued

```
                     mol_type = protein
                     organism = Streptomyces sp.
SEQUENCE: 12
MTRQSGRGAR RSAGPVSRRS GARARRSGLR TRRFAIGTVA VVALAGANGP WLYRFTTERY  60
HHYTINTQSY KEANGHWDFL DIPAEFRVNA VHAALLRTGK VLLVAGSGNN EKNFDAKSFR  120
SILWDPKTNA FKDIHTPKDL FCAGHTQLPG GNLLIAGGTK RYEKLGGDVT KAGGLMIVRN  180
ENPRKPITLP AGTRFTGKES GKTFISQDPL LVEKATKVFD KKTGEYQRNE AGVGRIYVEA  240
EKSGRKYETG TEDNYRISGL STEDTRNVYG IAQKLALDKK DFQGINDAYE FDPVAEKYIP  300
VDPMDKARWY PTLTTLKDGK VIAVSGLDDI GQIDVGKAEI YDPKKKSWAP SGIVRKFATY  360
PALFLLNDGK LFYSGSNSGY GLQSEGRTPG IWDLKTNDFQ EVPGLVDADQ METSATVRLP  420
PAQDERFMVI GGGGVGESDK STPKSRLVDL QQKNPKFTEG ASLDKGTRYP SASLLPDDSL  480
LVAGGAGDYR GRGGSDVLEA RLYDAKSDTY KRVADPAVGR NYHSGSVLLP DGRVMIFGSD  540
PLYMDKINTR PGTFEQRIEI YTPPYLYRGS RPELTAGPKS IKRGGTGMFT TQHASKITSA  600
KLMRPSAVTH VTDTDQRSIA LELEKSADGI TVTVPKNRAL VPSGWYMLFV TDAKGTPSEG  660
TWVEVP                                                             666

SEQ ID NO: 13          moltype = AA  length = 663
FEATURE                Location/Qualifiers
source                 1..663
                       mol_type = protein
                       organism = Streptomyces abyssalis
SEQUENCE: 13
MKYRPSRRTR RTAIGAAVVL VIAGFNGPAL YGIASEKYHD YKINQPEYKA ENGHWKTVQV  60
PEKYRINTIH AALLHTGKVL LVAGSGNDAK QFNAKTFRTV LWDPVKNTFK NIPTPNDLFC  120
AGHTQLPDGK LLVAGGTRRY EKLGGDVKKA GGLMIIHNEN PDRAKTVKAG TRFTGRKNGK  180
TFEAKDAVLV PRAKKKTNPA TGKVEVTPSV ARVYVEAIKR GKKYQTGTED NYRLSGLRGA  240
ERRNIYGIAQ KLSFDKKDFQ GIKDTYEFDP VAERYLKTDF MDEARWYPTL TTLEDGKVLS  300
VSGLDDIGQV VPGKNEIYDP QTKKWEFLPR KRFFPTYPAL HLTDDNKIFY SGANAGYGPA  360
DKGRTPGLWN LETNGFRKIP GMSDPDKLET AMSVPLPPTQ SRKFMVLGGG GVGESKKASE  420
KTRIVDLKDP EPRFQDGPDL YAKARYPSSV ILPNDKVLTT NGSGDYRGRS ESNILKAEIY  480
DPADNTKKPV ADPIVGRNYH SGALLLPDGR VMTFGSDSLF SDKANTKPGK FEQQIDIYTP  540
PYLHTKGERP ELENTGRGEQ TVRLGGTASF NSKDAGSIKK MRLVRPSSFT HVTNVEQSSV  600
ALKFKKSADG VTVKLPEDPS LVPPGWWMVT AVDDKGKPSK SVWLKVPRTA SAQAPTTTPL  660
SPP                                                                663

SEQ ID NO: 14          moltype = AA  length = 652
FEATURE                Location/Qualifiers
source                 1..652
                       mol_type = protein
                       organism = Streptomyces indicus
SEQUENCE: 14
MTVGSTPVKH TRRRARRIAI GAAVVLALAG MNGPALWRFA SDQYHDYKIN RPEYKADNGK  60
WDIVELPDEY KLNTIHAALL HTGKVLLVAG SGNNQKNFDA KKFDTVLWDP VKNTYKKIPT  120
PNDLFCTGHT QLPDGKLLVA GGTKSYEKLK GDVTKAGGVM NVYNENPDEP ITLAKGTKLT  180
GKENGKTFVI TESIVVERAK KRFDKETGKF LGNTPGYARV YVEAEKRGKK YETGTQDNYK  240
IVGLSGADAK NTYGIAQKLA LDKKDFQGID DAYEFDPVAE KYIKVDPMNE ARWYPTLTTL  300
EDGRVLALSG LDDIGQIVPG NRNEVYDPDT KKWTYVKQNR QFPTYPAIFL LPNGKLFYSG  360
ANAGYGPDDV GREPGIWDFD TNGFTKLKGL SDPKLMETAG TVWLPPAQDE KFMVVGGGGV  420
GESEQSSDKT RIIDLSEDAP EFKDGPSFDK GTRYPQSSVL PDDSVLISGG SEDYRGRSGS  480
NILEARMYDA KTGEMRRVAD PLVGRNYHSG SLLLPDGRVV FFGSDSLFAD EANTKPGEFE  540
QRLEIYTPPY LYQGAQPTLT GGPDSIERGE SATFKTQHAS SIKNARLIRP SASTHVTDID  600
QRSIALDVKK GADSVTVTVP ENKNLVQPGW YMLFVTDDQG VPSKAQWVEI PE          652

SEQ ID NO: 15          moltype = AA  length = 645
FEATURE                Location/Qualifiers
source                 1..645
                       mol_type = protein
                       organism = Streptomyces diastatochromogenes
SEQUENCE: 15
MKDQAGRRRA RRLAIGTAVV LALAGMNGPW LYRFGTEQYH QYKINKPEYK AANGHWDIVD  60
FPEEYRQDTI HAALLHTGKV LLVAGSGNNQ DNFDKKKFDT RIWDPVKGTI KKVPTPADLF  120
CTGHTQLANG NLLIAGGTKR YEKLKGDVTK AGGLMVVHNE NPDKPMTLPA GTKFTGKENG  180
KTFVSKDPVL VPRAKKVFDP KTGAFLRNDP GLGRIYVEAQ RSGTQYETGT QDNYRIQGLT  240
GADARNTYGI AQKLALDKKD FQGIRDAYEF DPVAERYIKV DPMNEARWYP TLTTLSDGKI  300
LSVSGLDDIG QLVPGKNEIF DPKTKKWTYT QKIRQFPTYP ALFLMQNGKI FYSGANAGYG  360
PDNVGRVPGI WDVGTNKFTK LPGLSDPNMM ETAGTVLLPP AQDEKFMVIG GGGVGESKLS  420
SKKTRLIDLK AQNPHFVDGP QLEKGTRYPQ SSILPDDTVL VSGGSEDYRG RGASNILQAR  480
MYHPDSNSFT QVADPLVGRN YHSGSILLPD GRVMFFGSDS LYADKANTKP GKFEQRIEIY  540
TPPYLYRDSR PSLSGGPQTI ARGASGTFTS KHASSIKKVR LIRPSASTHV TDVDQRSIAL  600
DFTTSGDKIT VTVPKSRNLV QSGWYMLFVD DDQGTPSKAQ WVKVP                  645

SEQ ID NO: 16          moltype = AA  length = 658
FEATURE                Location/Qualifiers
source                 1..658
                       mol_type = protein
                       organism = Streptomyces carminius
```

-continued

```
SEQUENCE: 16
MTYRPSRRFR RIAIASAVVA ALAGANGPAL YRFGGEAWHD HRINQPGYKA ANGHWRTVDV  60
PERYRVNTIH AALLHTGKVL LIAGSGNDAE QFRAKTFRTV LWDPEKNTFA NIPTPKDLFC  120
AGHTQLPDGR LLVAGGTQRY ETLEGDVEKA GGLMTVHNED PDKPVTLPAG TRFTGRANGK  180
TFASQAPVVV PRAGKTTAPG TGKVTVTVST ARVYVEALKK GRRHQTGTQD QYRVHGLKGD  240
DRRNVYGIAQ KLSFDKKDFQ GIRDSYEFDP VAERYVGVDP MNEARWYPTL TTLADGSVLS  300
VSGLDEIGQV VPGRNEIYDP ETQKWSYLPE ERFFPTYPAL FLAAGGRILY TGSNAGYGPA  360
DEGRTPGLWD LESNEFREVP GLGDPDVLET SMSVLLPPAQ EQRYMVLGGG GVGESRRATA  420
RTGIVDLTEP EPRFTPGPDL YAPVRYPSSV ILPDDTVLTT NGSGGYRGRG GSDVLRAGLY  480
DARTDTARTV ADPLVGRNYH SGALLLPDGR VMTFGSDSLF ADKADTRPGE FQQQIDLYSP  540
PYLFRGKRPE LRGTGEPRTV ELGGKVTFAS EDAARVRTAR LVRPGSFTHV TNVEQRSVAL  600
EFERTADGLT FTLPDDPSLV PPGYYMLNVL DGEGVPSESV WVKVPVSPDA EPVTVAAD    658

SEQ ID NO: 17          moltype = AA  length = 647
FEATURE                Location/Qualifiers
source                 1..647
                       mol_type = protein
                       organism = Streptomyces sp. Ru73
SEQUENCE: 17
MARRRTRFKE RKTWLGIGAV LVLAAFNAPA ATGFAERVYH DYAINRDAYK AENGHWDVLD  60
VPPEFRQNTI HAALLRTGKV LLVAGSGNNV KNFDAKRFTS VLWDPVRNSF KKIHTPNDLF  120
CTGHIHLADG KLLFAGGTKR YEKLKGDVTK AGGLMIVHNE NPNEPVRLPA GTVFTGRRNH  180
RTFVAKDPIR IRPAKKVFDK ETGRFLRTEA SVDRIYVEAR KKGRQYEAGA QDNYRVRGLT  240
RARAHNVYGI ATKLGLDKKD FQGIHDAYEF DPLAEKFTKV QSMHEARWYP TLVRLSDGRV  300
LALSGLDDIG QVVPGKSEIY DPKSRTWAYS GKVRQFPTYP AVFPLSDGTL FYSGSNAGYG  360
PADVGRTPGI WHFRTNTFTK LPGLDDAGLM ETSGTVLLPP AQRQKFMVVG GGGIGESAKS  420
SKKTRLIDLT QPHPSFHDGP SLDEGTRYPI LSVLPDDTVL VTGGSKDYRG PVAENILQAR  480
LYDARSGTMR RVADPQVGRN YHSGAILLPD GRVLTFGGDS LFADKAKTQT GTFEQRLEVY  540
TPPYLYRGAR PALSHGPTTV GRGATAVFRT KDAAAVKTAR LLHPSSTTHV TDVEQRSVAL  600
DFTRTKSGDG IEVTLPKNRN LVPSGWYMLF VTNRQGTPST ARWVKVP                647

SEQ ID NO: 18          moltype = AA  length = 667
FEATURE                Location/Qualifiers
source                 1..667
                       mol_type = protein
                       organism = Streptomyces sp.
SEQUENCE: 18
MAHRPSRRPV RILLACGAVL ALAAANAPAV YDFTSERYHA YEIDQADYKA EKGHWRTVDV  60
PEKYRINTIH AALLNTGKVL LIAGSGNDAK QFDAKTFRTV LWDPKKNSFK EIDTPKDLFC  120
AGHTQLPDGN LLVAGGTQRY ETLKGDVQKA GGLMIVHNEN PDKPKTLPEG TRFTGQENGK  180
TFVSQASVLI PRAKKTVDKK TGKAKVTASS ARVYVEAPEK GEKYQTGTQD QYRVEGLKGD  240
EKQNIYGIAQ KLSFDKKDFQ GIKDSFEFDP VAERYITVDP MAEARWYPTL TTLEDGKVLS  300
VSGLDEIGQV VPGKNEIYDP KSKKWKYLPK SRFFPTYPAL FLTADSRIFF TGSNAGYGPD  360
DQGRVPGLWD LDSNKFSEVP GLSDPDILET SMSVLLPPAQ DQKFLVLGGG GVGESPKSTD  420
RTGIADLKAD KPRFTDGPDL YDKARYPSSV ILPDDTVLTT NGSGDYRGKG ASNVLKASLY  480
DPGTNKFAEA AAPLVGRNYH SGGLLLPDGR VMTFGSDSLF SDKADSKPGK FEQTLEIYTP  540
PYLHKGGERP ELRDSDGARR TVKLGASADF RSKDAASIKK MRLIRPGAFT HVTNVEQRSI  600
AMDFKRTADG VRVTVPHNPS LVPPGWYMLT AVDDRGVPSE SVWLNVPVSK DAPTHKTTAD  660
APAPAPN                                                            667

SEQ ID NO: 19          moltype = AA  length = 654
FEATURE                Location/Qualifiers
source                 1..654
                       mol_type = protein
                       organism = Streptomyces sp.
SEQUENCE: 19
MLPCCQASAA LTARNPCLPA NAHSHCSVSQ APELATLEYA SPPVSTPYHD YTIIKPGYKA  60
ANGHWDLLDV PAGYRINTIH AALLHTGKVL LVAGSGNNRK NFDAKSFRSV LWDPRTGAFK  120
NIPTPDDMFC SGHTQLPDGK LLIAGGTRRY EKLKGDVTKA GGLMIVHNEN PDRPVTLPAG  180
TRFTGRANGR TFVSKDPVLV EKATKVFDRR TGRFLHNDAG LGRIYVEAEK SGARYATGAE  240
DNYRIAGLTG SDTRDVYGIA QKIALDKKDF QGIREAFEFD PVAEKYISVD PMNEARWYPT  300
LTTLKDGRVL ALSGLDEIGQ IVPGKDEIYD PRTKKWRYTG IVRRFPTYPA VFLLGDGKLF  360
YSGSNAGYGP ADVGRDPGVW DLATNSFRKI PGLGDADETE TSATVRLPPA QDERFMVIGG  420
GGVGESDRSS AKSRLVGLKD PAPRFRDGAA LAEGTRYPGV SLLPDDSVLV TGGSSDYRGR  480
GGSNVLQARL YDPGTDRYRR VADPAVGRNY HSGSVLLPDG RVMIFGSDSL YADRANTRPG  540
VFEQRIEIYT PPYLYRSSRP VLTGGPRSLR RGGTGGFTTA PGRAITSAKL IRPSAVTHVT  600
DTDQRSVALG LRRTPGGITV TVPGNRALVP SGWYMLFVTD DRGTPSEGRW VEVP        654

SEQ ID NO: 20          moltype = AA  length = 667
FEATURE                Location/Qualifiers
source                 1..667
                       mol_type = protein
                       organism = Streptomyces mobaraensis
SEQUENCE: 20
MRLRHLTDRR SGTVIGRRTR RTAVGAAVVL VLAGMNGPAI WGYASDQYHE YKINKPEYKA  60
KNGHWDVVDV PDEYRINTIH AALLQTGKVL LVAGSGNNAA NFAAKTFRTV LWDPERNTFK  120
```

-continued

```
NVPTPKDLFC AGHTQLPDGK LLVAGGTQRY EKLEGDVTRA GGLMIVHNEN PDKPMTFPAG  180
TRFTGKQSGK TFESKDAVLV PRAKKEQGKD GKVKVTVSTA RVYVEALEKG QEHATGLTDN  240
YRIEGLTGDD ARNFYGIANK LSFDKKDFQG IKDAFEFDPV AERYVTVDPM NEARWYPTLT  300
TLQDGKVLSV SGLDEIGQVV PGKNEVYDPK TKKWTYLPQE RFFPTYPALF LTDKGKIFYT  360
GSNAGYGPAD KGRDPGVWDL GNNSFIPVPG ISDPDALETS MSVLLPPAQD QRYMVLGGGG  420
VGEDKKSTAR TRIVDLHTER PRFHDGPDLY AKARYPSSVI LPDDTVLTTN GSGDYRGRSA  480
SNVLKAEIYD PKANASHRVA DPLVGRNYHS GALLLPDGRV MTFGSDSLFR DKDNTQPGVF  540
QQQIDLYTPP YLFHKGDRPE IRDTDRRIVK LGDKTTYRIT SAHGVAKARL IRPGSFTHVT  600
NIEQRSIALD LKKEGTDRFT VTLPKDPSLV PPGWYMVIVV DEEGTPSKAV WVKVPAAEKR  660
DDEKKEE                                                            667

SEQ ID NO: 21          moltype = AA  length = 506
FEATURE                Location/Qualifiers
source                 1..506
                       mol_type = protein
                       organism = Colletotrichum tofieldiae
SEQUENCE: 21
MVTVCSVLRC LPAALLAFAA ASEAQNVGQW GPMVKFPVVP VAVALLPETG NMLVWSSGWP  60
NRWTTAGNGK TYTSMYDVKT GKVSDALIQN TQHDMFCPGT SMDADGRIIV TGGSSASKTS  120
VLDFKKGESS SWAPLANMQI SRGYQSSCTT SEGKIFLIGG SFSGAGRRDG EIYNPKANTW  180
TKLPGCPVKP LVMQRGMFPD SHAWLWSWKN GYVLHAGPAK QMNWYDTKGT GANTPAGLRG  240
ADDDSMCGVS VMYDAVAGKI FTYGGGKAYT GVTSSSNAHI LTLGEPGQAV QVQKLQNGKF  300
NRGFANAVVM PDGKIWVVGG MRQMQLFSDN TPQLTPELFD PATGNFTPTT PHAVPRNYHS  360
TALLMADATI WSGGGGLCGA NCKENHFDGQ FWSPPYLFEA DGKTPAKRPV IQNLSETDVK  420
AGAPITITMQ DAGTYTFSML RVSATTHTVN TDQRRIPLDG QDGGDGKDFT VNIPDDYGIA  480
VPGYYMLFAM NEAGVPCVAK FFKVTL                                       506

SEQ ID NO: 22          moltype = AA  length = 680
FEATURE                Location/Qualifiers
source                 1..680
                       mol_type = protein
                       organism = Moelleriella libera RCEF 2490
SEQUENCE: 22
MSRFLLLLAL VAVASSAAVE QQQQQQQQPR AIKILEHYRE SSTFTKLFAA APIGNEIPKN  60
AYTVTCDSYQ PGNECALAID GNNNTFWHTA FSGANLPHQI VVDLGATRNI NGLSALPRQD  120
GNNHGFIAQH EVAVSTDNRN WEVVASGTWY GGDSTTKFAN FETRSIRYIR LKALSEAYGN  180
QWTSVAEVKV YEAKTGPAAF AGTGKWGPTI NFPTIPVAGA VDPLTGQVLI WSSYTYDNYL  240
GSSQDRVFTS IWDPSTGVVT PKLVDNTNHD MFCPGISIDG AGRMVITGGN SAQKTTVYQF  300
GSQTWIPGPD MNTQRGYQAS ATLSDGRVFT IGGCWSGGWF EKNGEVYDPK ASTWTSLPGA  360
LVHPMLTNDA QGIYRADNHA WLFGWKNGSV FQAGPSTAMN WYTTSGNGSV APAGNRTSSR  420
GDAPDAMTGN AVMYDAVNGK ILSFGGSPSY QDSSATTNAH IITIGSPGAP AQSRFASNGL  480
WSPRAFHTSV ILPDGKTFIT GGQTYAVPFS DDNPDLTPEM YDPVADSFVQ QQANTIIRVY  540
HSISLLLPDG RVFNAGGGLC GDCTTNHFDG QIFTPQYLLT RTGQLAARPA INSVSLSGRR  600
LTINTNSAIT SAALMRYGSA THTVNTDQRR IPLKLTTTGT NRYTADAPSD PGILLPGYYM  660
LFVLNSQGVP SVAKTVNFLV                                              680

SEQ ID NO: 23          moltype = AA  length = 685
FEATURE                Location/Qualifiers
source                 1..685
                       mol_type = protein
                       organism = Purpureocillium lilacinum
SEQUENCE: 23
MKLLGTVAAL ALCDAASHTY AVAVKPRPHT AMPNLAARGT EMSLMAAKPI GNAINRAGWK  60
VTCDGEEQGN ECAKAIDGDN NTMWHTAWQN DNPPPPHTIT VDMGSAQTIN GISVLPRQDG  120
SEHGWIARHD VLVSNDGQTW GDPVATGTWY TDATAKYANF EPRSARYVRL VARSEAQGRP  180
WTSIAELNVY RADGPPVPKN GIGKWGLTLD FPVVPVAGVV DPLTGKVVVW SAYENDQYEG  240
SPGGWTLTST WDPATGEVTE RNVTNIGHDM FCPGVSLDAS GRVVVTGGSN AQKTSFYDAA  300
TEAWVPGPDM KTPRGYQASA TCSDGRVFTI GGSWSGGQFE KNGEVWDPKT NSWRALPGAA  360
VKPMLTKDRG GIYRADNHAW LFGWRNGSVF QAGPSTAMNW YYTAGDGRVR SAGQRRAPRG  420
ADPDAMCGNA VMFDATAGKI LTVGGAPHYE DADATTNAHV LTLGDAGAAP KVVFAGNGMA  480
HPRIFANAVV LPDGTVFVTG GQQHAELFKD TTPQLTPELY DPALGAFVEQ APNSVVRVYH  540
SMALLLPDAT VLSGGGGLCG GACDTNHFDA QVFSPRYLFD GEGQPAARPK IRAVASKEVH  600
AGDAIKVTTD GPVKSAALVR YGSATHSVNT DQRRVPLTLR QDGGSYAYSA DLPRDLGVLL  660
PGYWMLFVMN DKGVPSVAAT VKVLL                                        685

SEQ ID NO: 24          moltype = AA  length = 671
FEATURE                Location/Qualifiers
source                 1..671
                       mol_type = protein
                       organism = Penicillium antarcticum
SEQUENCE: 24
MHAKWLSGLL LLGLSQAKDN SQALQTAATT NLQSAAEDPY ATNQAPPIGE KLSRRRWTVT  60
CDNSQTDECT KMLDDNFDTF WQSGNDGSIH EVVVDLSRAE NIQALSIVSR RDTATTGQIF  120
AHQIYVSNDK ENWGNPVAFG TWYNTTDEKL AVFEPKEGQY LRLVSLDGNA ASISELQIYD  180
SDSPKQVSNG GVWGPTIDLP LVAVSGAVLP GTGQVLVWSS WAKDDYLHST SQTLTALWDP  240
KTGNVTQRTV QDTGHDMFCA GTSWDGNNEL LVTGGNNDGQ TSILDPATDT WYPGAAMKLG  300
```

-continued

```
RAYHASATNS EGRVFIIGGS WNGGTNNNRD GEIYDPVAEE FTGLPGARVE PMWTNDANGG  360
YRRDSHGWLF GWKNGTVFQA GPSKAMNWYY TSGQGDQKPA GNRGDADDAM SGNAVMFDAV  420
NGKIISFGGS PSYENTYATK DAHLIEIGEP GTDPKVSTAK NPSGDGMAFA RVFHTSVVLP  480
NGNVFVAGGQ TYATPFNDSG AQFTPEIYDP ATNEFRQGQT NSVVRVYHSI SLLLTDGRVF  540
NGGSGLGVSA PSNHFDAQIY TPDYLLNNDG TLATRPTIDS VDKRTLHPGE KLIVKASIKL  600
TKASLIRYGS TTHTMNTDQR RVVLDSWAAN GETYETTLPS EAGALLPGPW MLFILNDDGV  660
PSIAETIYIE V                                                       671

SEQ ID NO: 25          moltype = AA  length = 673
FEATURE                Location/Qualifiers
source                 1..673
                       mol_type = protein
                       organism = Penicillium flavigenum
SEQUENCE: 25
MRPRWLTGLL LLGLAHADND NSVISDKQVA NQDHSATTDP YATTQAPPRG EKLSRRRWSV  60
TCDSSQSGNE CEKLLDNDLQ SFWLSGAEGN THEVTIDLSR AENVHAISVV SRQDGKTDGQ  120
IAAHQVFVSG DKNDWGEPVA YGTWYEGTGE KLAIFEPKEG QYVRLVSLGS NVASISELNI  180
YDGADPKTVP NGGVWGPTID LPVVAVSGAV IPETNEVLVW SSWAKDDYLH SRGYTLTAVW  240
NMNDNSVTQR KVQETHHDMF CSGMSYDGKG ELLVTGGNND KSTSIFDPAS GNWTEGNTMI  300
ISRGYQASAT LADGRVFIVG GSWNGGTNYD KDGEIYDPNT EEFRFLKNAL VKPMWTDDYN  360
SGYRRDSHGW LFGWKNDTVF QGGPSKNMNW YYTEGDGDQK PAGTRADASD SMSGNAVMFD  420
AVNGKIITFG GSPSYENSYA TTDAYLIEID EPGSQPKVTA AKNPNGEGMA YARTFHTSVV  480
LPDGGVFTAG GQSYGVPFND SNAHLTPELY DPKTNQFNEQ QPNSIVRVYH SISLLLPDGR  540
VFNGGSGLGV SAPTNHFDAQ IYSPHYLFNQ DGSLATRPTI DGVVEKKLRA GDKLTISASI  600
DVKNASLIRY GTTTHTVNTD QRRIALDSWT ANGDSYETTL PGDNGILLPG PWMLFILNDD  660
GVPSVSQTIH IQV                                                     673

SEQ ID NO: 26          moltype = AA  length = 684
FEATURE                Location/Qualifiers
source                 1..684
                       mol_type = protein
                       organism = Fusarium sp. AF-8
SEQUENCE: 26
MWSPILTLCL TGLFWDEAMA IVVPAQNSTK VKVQMEHLSM RATGPLGTAI NRNNWKVTCD  60
SDNPDSDNAC QKAIDGDVNS FWHTAWFEDT SKDPGLPHTL TVDMKTVKNV NGISALPRQD  120
GTTHGWIARH DIFLSTDGKT WGSPVATGTW YADGTEKFSN FETKRARYVR IVAITEAYDG  180
PWTSIAEFNV YKAATYTAPK TGIGMWGPTL DFPIVPVAGA VDPGTGKVLV WSSYYHDTMN  240
GSPGGMTLTS LWDPETGIIT QREVFETNHD MFCPGISMDG TGQIVVTGGN NAARTSTWDP  300
VKNQWVSAPD MKIPRGYQSS ATTSNGKVFT IGGSWGGDRA FKPGEIFDPK SRKWTLLPNA  360
KVEPMLTADA QGIFRSDNHA WLFGWKGETV FQAGPSSAMN WYYTQGKGTV KPAGKRQSSR  420
GVDPDSMCGN AVMFDAAKGK IVTFGGTPNY QDSYATTNAH IITIGAPGTQ ASVAFVSDGM  480
YYPRVFHTSV LLPDGTVFIT GGQEYAIPFE DSTPQLTPEL YIPDSDTFIK QQPNSIVRTY  540
HSMSILLQDA RVFNGGGGLC GDCSTNHFDA QIFTPSYLLT KDGKPAARPK IVSVSATTIK  600
VGGSITVTTG GVVNTASLIR YGTATHTVNT DQRRIPLTLA SAGKNKYTFK VPGDSGIALP  660
GYWMLFVMNS AGVPSVATTI KVTT                                          684

SEQ ID NO: 27          moltype = AA  length = 506
FEATURE                Location/Qualifiers
source                 1..506
                       mol_type = protein
                       organism = Colletotrichum spinosum
SEQUENCE: 27
MVTLCSTLRS LPAALLALAA LSDAQGVGQW GPLVKFPIVP VAVALIPESG NMLVWSSGWP  60
NRWTNAGNGK TFTSLYDVST GKAGDAIVQN TQHDMFCPGT SLDVEGRIIV TGGSSAAKTS  120
VLDFRKGESS SWAPLSNMQI SRGYQSSCTT SEGKIFVIGG SFSGAGTRNG EIYDTASNKW  180
TKLAGCPVKP LVMQRGLFQD SHTWLWSWKN GSVLQAGPAK QMNWYDTKGT GANTPAGLRG  240
ADDDAMCGVS VMFDAVAGKI FTYGGGKAYT GYAASSNAHI LTLGEPGQQV QVQKLANGQY  300
NRGFSNAVVL PDGRIWVVGG MRQMQLFSDN TPQLTPELFD PATGKFTPTA PHAIPRNYHS  360
TALLMADATV WSGGGGLCGA NCKENKFDGQ FWSPPYLFEA DGKTPAQRPV IDSLSDKTVR  420
AGAPLVVTMK DEGKYTFSML RVSATTHTVN TDQRRIPLDG QDGGDGKSFT VNMPGDYGVV  480
IPGYYMMFAM NEAGVPCVAK FFKVAL                                        506

SEQ ID NO: 28          moltype = AA  length = 684
FEATURE                Location/Qualifiers
source                 1..684
                       mol_type = protein
                       organism = Pochonia chlamydosporia 123
SEQUENCE: 28
MELFGAIAFV LCQLSCHVHA VTLKQTSTTA LVHATEQAAL HLLAAAPVGY RIDRANWKVT  60
CDSQEIGYEC GKAIDGDNST VWHTAFRADN PKPPHNITVD MGSIQNVNGL SILPRQDSDS  120
NGWVARHQVL VSADGQRWEH PVATGTWYAD STEKFSNFEP EQARFVRLVA DSEVSGNPWT  180
SVAELNVYKA GTDPAPAKSA SSGKWGPTLD FPIVPVAAAV VPQSGKVLVW SAYENDKFEG  240
SPGGYTLTST WDPATGDVTQ RNVTNIGHDM FCPGISMDSN GQIVVTGGNN AQKTSFYDAA  300
SDSWVPGPDM TVSRGYQSSA TCSDGRIFTI GGSWSGGQFE KNGEIYDPST KTWKMLSLAA  360
VKPMLTADKK GRYRADNHGW LFGWRNGTVF QAGPSTAMNW YFTGGDGGVK SAGSRKSSRG  420
PDPDSMCGNA VMFDATNGKI LTVGGSPWYQ DNDATANAHI ITVRGPGSEP SVSFAGNGMH  480
```

-continued

```
HARIFASSVV LPDGSVFITG GQQHSIPFAD STPQFTPELY DPEKDQFLEQ APNSIVRVYH  540
SLALLLPDAT VFNGGGGLCG GSCATNHFDA QIYSPRYLFD HDGSPAVRPV IKSVSRTSVK  600
PGDSLSFTTD SAVQHASLVR YGTATHTVNT DQRRVPLTLD ASGANTYAVQ LPNDPGIMLP  660
GYWMLFVMNE RGTPSVAASI KAHL                                        684

SEQ ID NO: 29          moltype = AA  length = 684
FEATURE                Location/Qualifiers
source                 1..684
                       mol_type = protein
                       organism = Nectria haematococca
SEQUENCE: 29
MKNLLTLSLC LGSLLDVTDA VVIPETPGPQ PGKFKASPPL GIPIDRKGWT VTCSNQQSGN  60
DCAKAIDGNT ATFWQTNANT KAPVTITVDM KTARNVNGLS QLPLQDNSNN IGNWIERHSV  120
FLSTDGKNWG SPVATGTWWT DKTEKFSNFE PQRARYVRLV AIPDETHDNP RISLAELNIY  180
AASSNSNQEP GLGRWGPTLD FPVVPVASAV EPTSGQVVVW SARAYFEPAY RYDDFQHQNP  240
RGGLTLSAVW DPATGVITQR NITLTHHDMF CPGISMDGDG QIVVTGGNDA KKTSLYDSPS  300
DSWITGPEMN IARGYQSSAT TSDGRVFTIG GSWNGPRGGK NGEIYNPDLK SWTLLPGALV  360
KPMLTADKEG VYRSDNHGWL FGWKKGTVFQ AGPSTAMNWY YTRGADGDVK PAGKRQVGDI  420
VDPDSMCGNS VMYDAVKGKI LTFGGSPNYR FSDSTANAHI ITIGEPGSVA KTAFAGGGQG  480
LHPRIFHTSV VLPDGTVFIT GGQKHSEPFV DSTPQLEPEM YLPASDAFVK QQSNSIVRVY  540
HSISLLLPDG RVFNGGGGLC GTCTTNHFDA QIFTPNYLFD KNGNLATRPR ISSTSTKTAK  600
VGSTITFTTN GPVKQGSLIR YGTATHTVNT DQRRIALTFT NTGTNRYSFK IPNDPGIALP  660
GYWMLFVLNS AGVPSVATTI KVTN                                        684

SEQ ID NO: 30          moltype = AA  length = 679
FEATURE                Location/Qualifiers
source                 1..679
                       mol_type = protein
                       organism = Gibberella subglutinans
SEQUENCE: 30
MKSFYSLALC LGAFFNAATA IDPEEQGQQP GKFAAAPPVG SNPIDRKGWT VKCSSQAPNF  60
PCGRAIDGDK NTFWQTPYGT TNTPPPHTIT IDMKQTQYVS GLQITPRQDG NTRNWIGRHE  120
VYLSSDGTNW GKPVAFGTYW GDKYPWITNF ETHPARYLRF VALSNVNSDY PWIAIADFQV  180
YNALKYNPPA KGLGKWGPTI DFPVIPVAGA VEPVSGKVVI WSAYRYDAFQ GTTPRGGFTL  240
TSIWDPKTNV ISNRNVSNNH HDMFCPGISM DGEGQIVVTG GNDAKKTTIL MPDGNWVPGP  300
DMQIARGYQS SATCSDGRVF TIGGSWSGAR GGKNGEIYDP RAKTWTSLPK CLVGPMLTHD  360
KEGVYKADNH AWLFGWKKGS VFQAGPSTAM NWYYTDRGTQ GNTKAAGTRR KNGRVDPDSM  420
NGNVAMFDAV KGKILTFGGA TSYQQAPATA NAHVLTIDQP GAIAQTALVG NNGAGIHARV  480
FATSTILPDG NVFITGGQSY SNPFTDTNAQ LEPEMFISSS NTFTKQQPNT IPRTYHSMSL  540
LLPDGTVFNG GGGLCGSCKS NHFDAQIFTP QYLLDGNGNL ATRPKITAVS ATTAKVGSTI  600
TVTANSAIKS ASLMRYGTAT HVVNTDQRRI PLALTGAGTN KYSFKIPNDS GIALPGYWML  660
FVINNAGVPS VASTIKVTV                                              679

SEQ ID NO: 31          moltype = AA  length = 681
FEATURE                Location/Qualifiers
source                 1..681
                       mol_type = protein
                       organism = Gibberella subglutinans
SEQUENCE: 31
MKSFWTLAFY LGNANAVAIS QPASKAETPE GSLQFLSLRA SAPIGTAINR DKWKVTCDSQ  60
HQGDECSKAI DGDRNTFWHT NWAAGATNDP KPPHTITIDM GSSQNVNGLS VLPRQDGSDH  120
GWIGRHNVFL STDGKNWGNA VATGTWFADI TEKYSNFETR PARYVRLVAV TEANDQPWTS  180
IAEINVFNAA SYTSPQPGLG LWGPTLDFPI VPVAAAVEPT SGKVLVWSSY KNDAFGGSPG  240
GVTLTSTWDP STGVISQRTV TVTKHDMFCP GISMDGNGQV VVTGGNNAEK TSLYDSSSDS  300
WIPGPDMKVA RGYQFSATLS NGRVFTIGGS WSGGTFEKNG EVYDPSSKTW TSLPGALVKP  360
MLTADQQGIY RSDNHGWLFG WKKGSVFQAG PSTAMNWYYT SGKGNTKSAG KRQSSRGTDP  420
DAMCGNAVMY DAVKGKILTF GGSPSYQDSD ATTNAHIITI GEPGSTPKTV FASNGLYYPR  480
TFHTSVILPD GNVFITGGQQ RGIPFADSTP QLTPELYVPN DDTFYKQQPN SIVRVYHSIS  540
LLLPDGRVFN GGGGLCGDCS TNHFDAQIYT PNNLYDSNGK LATRPKITNV SAKSAKVGGK  600
ITISTDTSIK QASLIRYGTS THTVNTDQRR IPLSLRSTGS GNSYSFQVPS DSGIALPGYW  660
MLFVMNSAGV PSIASTLLVT Q                                           681

SEQ ID NO: 32          moltype = AA  length = 681
FEATURE                Location/Qualifiers
source                 1..681
                       mol_type = protein
                       organism = Gibberella moniliformis
SEQUENCE: 32
MKSFWTLAFY LGGASAVAIS QPAAKSETPA GSLQFLSLRA SAPLGTAINR DKWKVTCDSQ  60
HEGDECSKAI DGDRNTFWHT NWAAGGTNDP KPPHTITIDM GSSQNVNGLS VLPRQDGSNH  120
GWIGRHNVFL STNGKNWGDA VATGTWFADN TEKYSNFETR PARYVRLVAV TEANDQAWTS  180
IAEINVFKAA SYTSPQPGLG RWGPTLDFPI VPVAAAVEPT SGKVLVWSSY KNDEFGGSPG  240
GVTLTSTWDP STGVISQRTV TNTKHDMFCP GISMDGNGQV VVTGGNNAEK TSLYDSSSDS  300
WIPGPDMKVA RGYQSSATLS NGRVFTIGGS WSGGISEKNG EVYDPSSKTW TSLSGALVKP  360
MLTADQQGLY RSDNHGWLFG WKKGSVFQAG PSTAMNWYYT SGNGNTKSAG KRQSSRGTDP  420
DAMCGNAVMY DAVKGKILTF GGSPSYQDSD ATTNAHIITI GEPGSTPKTV FASNGLYYPR  480
```

-continued

```
TFHTSVILPD GNVFITGGQQ RGIPFADSTP QLTPELYVPN DDTFYKQQPN SIVRVYHSVS  540
LLLPDGRVFN GGGGLCGGCT TNHFDAQIYT PNNLYDSNGK LATRPKITKV SAKSVKVGGK  600
ITISTDSSIK QASLIRYGTS THTVNTDQRR IPLSLRSTGS GNSYSFQVPS DSGIALPGYW  660
MLFVMNSAGV PSVASTLLVT Q                                            681

SEQ ID NO: 33          moltype = AA  length = 680
FEATURE                Location/Qualifiers
source                 1..680
                       mol_type = protein
                       organism = Fusarium coffeatum
SEQUENCE: 33
MRHLLTLALC FTSINAVAIT KSHKAAGTGE PEGSFQFLSL RASAPIGTTI SRDKWSVTCD  60
SFHDGEGCDK AIDGDLNTFW HSQWGNGNDP KPPHTYTIDM GSTQNVNGLS VLPRQDGSSN  120
GWIGRHEVFL SSDGKNWGSA VATGTWYADS TTKYSNFETR PARYVRLVAL TEAKGQPWTS  180
IAEINVFKAD SYSAPRAGLG RWGPTLDFPI VPAAAFVEPT SGKVVVFSSY RNDAFGGSPG  240
GITLTSSWDP ANNVIAERTV TVTKHDMFCP GISMDGAGQI VVTGGNDAKK TSLYDSSSDS  300
WIPGPDMQNA RGYQSSATTS DGRVFTIGGS WSGGIFEKNG EIYNPSSKTW TSLSGAKVNP  360
MLTADRQGLY RSDNHAWLFG WKKGIVFQAG PSTAMNWYYT SGNGDVKSAG KRQSNRGVAT  420
DAMCGNAVMY DAVQGKILTF GGAKDYQDTD ATTDAHIITL GEPGTTPKTV YASNGLWYPR  480
TFHTSVVLPD GSTFITGGQV RGIPFEDSTP QLTPELYVPG DDTFYKQQPN SIVRVYHSVS  540
LLLPDGRVFN GGGGLCGDCT TNHFDAQVFT PNYLYDNNGN LVTRPKITNT STKSVKIGGR  600
VTITTDGSIQ KASLVRYGTA THTVNTDQRR IPLTLSNSGR NSYSFTVPND SGVALPGYWM  660
LFVMNSAGVP SVSTTIRITS                                              680

SEQ ID NO: 34          moltype = AA  length = 680
FEATURE                Location/Qualifiers
source                 1..680
                       mol_type = protein
                       organism = Fusarium sporotrichioides
SEQUENCE: 34
MKHLLTLALC FSSINAVAIT NPHKTAGHDH PEGSLQFLSL RASAPIGSAI SRNNWAVTCD  60
SAQSGNECNK AIDGNQDTFW HTFYGANGDP KPPHTYTIDM KSTQNVNGLS MLPRQDGSRN  120
GWIGRHEVYL STDGTNWGSP VAAGSWFADS TTKYSNFETR PARYVRLVAV TEASGQPWTS  180
IAEINVFQAS SYTAPQPGLG RWGPTIDLPI VPAAAAVEPT SGRVLVWSSY RNDAFGGSPG  240
GVTLTSSWDP SSGIVSDRTV TVTKHDMFCP GISMDGNGQI VVTGGNDAKK TSLYDSSSDS  300
WIPGPDMQVA RGYQSSATMS DGRVFTIGGS WSGGIFEKNG EVYSPSSKTW TSLPNAKVNP  360
MLTADKQGVY RSDNHAWLFG WKKGSVFQAG PSTAMNWYYT SGSGDVKSAG KRQSNRGVAP  420
DAMCGNAVMY DAVRGKILTF GGSPDYQDSD ATANAHIITL GEPGSTPNTV FASNGLYFAR  480
TFHTSVVLPD GSTFITGGQR RGIPFEDSTP VFTPEVYVPE QDTFYKQNPN SIVRVYHSIS  540
LLLPDGRVFN GGGGLCGDCT TNHFDAQIFT PNYLYDGNGN LATRPKITRT STQSVKVGGR  600
VTISTDSSIV KASLIRYGTA THTVNTDQRR IPLTLTNNGG NSYSFQVPSD SGIALPGYWM  660
LFVMNSAGVP SVAATIRVTQ                                              680

SEQ ID NO: 35          moltype = AA  length = 682
FEATURE                Location/Qualifiers
source                 1..682
                       mol_type = protein
                       organism = Fusarium longipes
SEQUENCE: 35
MKQLLTLALC FSSINAVAIN PHGHHNNKGT GDHEGSLQFL SLRASAPLGS AIARENWVVT  60
CDSAQPGNEC NKAIDGDTNT FWHTFYGNNG DPKPPHTYTI DMGSNRNVNG LSVLPRQDGN  120
RNGWIGRHEV YLSTDGSNWG SPVAYGSWFA DSTTKYSNFE TRPARYVRLV ALTEASGQPW  180
TSIAEINVYQ AGSYTAPQAG LGRWGPTIDL PIVPAAAAIE PTSGRVLMWS SYRNDAFGGS  240
PGGITLTSSW DPSSGIVSDR TVTVTNHDMF CPGISMDGNG QIVVTGGNDA KKTSLYDSSS  300
DSWIPGPDMQ VARGYQSSAT MSDGRVFTIG GSWSGGIFEK NGEVYSPSSK TWTSLPGAKV  360
NPMLTADKQG LYRSDNHAWL FGWKKGSVFQ AGPSTAMNWY YTSGSGDVKS AGRRQSNRGL  420
DPDAMCGNAV MYDAVAGKIL TFGGSPDYQD SDATTNAHII TLGEPGSTPN TVFASNGLYF  480
PRTFHTSVVL PDGSTFITGG QRRGIPFEDS TPQLTPEIYV PEQDTFYKQN PNSIVRVYHS  540
ISLLLPDGRV FNGGGGLCGD CTTNHFDAQI FTPNYLYNSN GDLATRPKIT RTSAQSVRVG  600
GRITISTDSS IRRASLIRYG TATHTVNTDQ RRIPLTLTNN GGNSYSFQVP SDSGIALPGY  660
WMLFVMNSAG VPSVATTLRV TR                                           682

SEQ ID NO: 36          moltype = AA  length = 645
FEATURE                Location/Qualifiers
source                 1..645
                       mol_type = protein
                       organism = Streptomyces lividans
SEQUENCE: 36
MKDRAGRRRA RRFAIGTAVV VALAGMNGPW LYRFSTEKYH QYKINQPEYK AANGKWEIIE  60
FPEKYRQNTI HAALLRTGKV LMVAGSGNNQ DNSDDKQYDT RIWDPVKGTI KKVPTPSDLF  120
CTGHTQLANG NLLIAGGTKR YEKLKGDVTK AGGLMVVHNE NPDKPITLPA GTKFTGKENG  180
KTFVSKDPVL VPRAEKVFDP ATGAFVRNDP GLGRIYVEAQ KSGSAYETGT EDNYRVQGLS  240
GADARNTYGI AQKLALDKKD FQGIRDAFEF DPVAEKYIKV DPMHEARWYP TLTTLGDGKI  300
LSVSGLDDIG QLVPGKNEVY DPKTKAWTYT DKVRQFPTYP ALFLMQNGKI FYSGANAGYG  360
PDDVGRTPGV WDVETNKFTK VPGMSDANML ETANTVLLPP AQDEKYMVIG GGGVGESKLS  420
SEKTRIADLK ADDPKFVDGP SLEKGTRYPQ ASILPDDSVL VSGGSQDYRG RGDSNILQAR  480
```

-continued

```
LYHPDTNEFE RVADPLVGRN YHSGSILLPD GRLMFFGSDS LYADKANTKP GKFEQRIEIY  540
TPPYLYRDSR PDLSGGPQTI ARGGSGTFTS RAASTVKKVR LIRPSASTHV TDVDQRSIAL  600
DFKADGDKLT VTVPSGKNLV QSGWYMMFVT DGEGTPSKAE WVRVP                  645

SEQ ID NO: 37          moltype = AA  length = 730
FEATURE                Location/Qualifiers
source                 1..730
                       mol_type = protein
                       organism = unidentified
SEQUENCE: 37
MRHFWLLPAV AGIAGAQCPY LSGEMSFTQE QDNAGDTIEV TEQPIDNTLY VNDTGSYMTT  60
DFGTPISDQT SLKAGPRGPT LLEDFIFRQK LQRFDHERVP ERVVHARGAG AYGTFKSYAD  120
WSNVTAADFL SANDKETPMF CRFSTVVGFR GSVDTARDVH GHACRFYTDE GNYDIVGINF  180
APFFIQDAIQ FPDLVHAIKP MPNNEIPQAA TAHTSAWDFF SQQSTALHSA LWLMSGNGIP  240
RSFRHMNGYG VHSFRFVAAN GTSKVVRTPW KSQQGVASLV WDEAQAAAGK NSDYHRQDLY  300
NAMPNGHYPK YELQAQIMDE ADMLRFGFDL LDPTKLVPEE VVPYTPLGMM ELNANPTNYF  360
AEVEQAGFQP GHVVPGIDFT DDPLLQGRLF SYLDTQLTRH GGPNFEQIPV NRPRKPVHNN  420
NRDGFGQQQI PTNNWAYTPN SMSNGYPMQA NQTQGHGFFT APYRYASGHL VRQTSPTFND  480
HWSQPAMFWN SLIPAEQQMV VNAIVFENSK VNSPHVRKNV VNQLNMVNNN LAVRVARGLG  540
LDEPSPNPTY YTSNKTSNVG TFGKPLLSIE GLQVGFLASN SHPESIKQGQ AMAAQFSAAG  600
VDLNIVTEAY ADGVNTTYAL SDAIDFDALI IADGVQSLFA SPALANQMNS TATSTLYPPA  660
RPFQILVDSF RYGKPVAAVG SGSVALKNAG IDSSRSGVYT GSSETTEKIA KEVLEGLYTF  720
RFVDRFALDE                                                         730

SEQ ID NO: 38          moltype = AA  length = 730
FEATURE                Location/Qualifiers
source                 1..730
                       mol_type = protein
                       organism = unidentified
SEQUENCE: 38
MRHFWLLPAV AGIAGAQCPY LSGEMSFTQE QDNAGDTIEV TEQPIDNTLY VNDTGSYMTT  60
DFGTPISDQT SLKAGPRGPT LLEDFIFRQK LQRFDHERVP ERVVHARGAG AYGTFKSYAD  120
WSNVTAADFL SANDKETPMF CRFSTVVGFR GSVDTARDVH GHACRFYTDE GNYDIVGINF  180
APFFIQDAIQ FPDLVHAIKP MPNNEIPQAA TAHTSAWDFF SQQSTALHSA LWLMSGNGIP  240
RSFRHMNGYG VHSFRFVAAN GTSKVVRTPW KSQQGVASLV WDEAQAAAGK NSDYHRQDLY  300
NAMPNGHYPK YELQAQIMDE ADMLRFGFDL LDPTKLVPEE VVPYTPLGMM ELNANPTNYF  360
AEVEQAGFQP GHVVPGIDFT DDPLLQGRLF SYLDTQLTRH GGPNFEQIPV NRPRKPVHNN  420
NRDGFGQQQI PTNNWAYTPN SMSNGYPMQA NQTQGHGFFT APYRYASGHL VRQTSPTFND  480
HWSQPAMFWN SLIPAEQQMV VNAIVFENSK VNSPHVRKNV VNQLNMVNNN LAVRVARGLG  540
LDEPSPNPTY YTSNKTSNVG TFGKPLLSIE GLQVGFLASN SHPESIKQGQ AMAAQFSAAG  600
VDLNIVTEAY ADGVNTTYAL SDAIDFDALI IADGVQSLFA SPALANQMNS TATSTLYPPA  660
RPFQILVDSF RYGKPVAAVG SGSVALKNAG IDSSRSGVYT GSSETTEKIA KEVLEGLYTF  720
RFVDRFALDE                                                         730

SEQ ID NO: 39          moltype = AA  length = 503
FEATURE                Location/Qualifiers
source                 1..503
                       mol_type = protein
                       organism = Micrococcus luteus
SEQUENCE: 39
MEHQKTTPHA TGSTRQNGAP AVSDRQSLTV GSEGPIVLHD THLLETHQHF NRMNIPERRP  60
HAKGSGAFGE FEVTEDVSKY TKALVFQPGT KTETLLRFST VAGELGSPDT WRDVRGFALR  120
FYTEEGNYDL VGNNTPIFFL RDPMKFTHFI RSQKRLPDSG LRDATMQWDF WTNNPESAHQ  180
VTYLMGPRGL PRTWREMNGY GSHTYLWVNA QGEKHWVKYH FISQQGVHNL SNDEATKIAG  240
ENADFHRQDL FESIAKGDHP KWDLYIQAIP YEEGKTYRFN PFDLTKTISQ KDYPRIKVGT  300
LTLNRNPENH FAQIESAAFS PSNTVPGIGL SPDRMLLGRA FAYHDAQLYR VGAHVNQLPV  360
NRPKNAVHNY AFEGQMWYDH TGDRSTYVPN SNGDSWSDET GPVDDGWEAD GTLTREAQAL  420
RADDDDFGQA GTLVREVFSD QERDDFVETV AGALKGVRQD VQARAFEYWK NVDATIGQRI  480
EDEVKRHEGD GIPGVEAGGE ARM                                          503

SEQ ID NO: 40          moltype = AA  length = 527
FEATURE                Location/Qualifiers
source                 1..527
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 40
MADNRDPASD QMKHWKEQRA AQKPDVLTTG GGNPVGDKLN SLTVGPRGPL LVQDVVFTDE  60
MAHFDRERIP ERVVHAKGAG AFGYFEVTHD ITRYSKAKVF EHIGKRTPIA VRFSTVAGES  120
GSADTVRDPR GFAVKFYTED GNWDLVGNNT PIFFIRDALL FPSFIHSQKR NPQTHLKDPD  180
MVWDFWSLRP ESLHQVSFLF SDRGIPDGHR HMNGYGSHTF KLVNANGEAV YCKFHYKTDQ  240
GIKNLSVEDA ARLAHEDPDY GLRDLFNAIA TGNYPSWTLY IQVMTFSEAE IFPFNPFDLT  300
KVWPHGDYPL IPVGKLVLNR NPVNYFAEVE QLAFDPSNMP PGIEPSPDKM LQGRLFAYPD  360
THRHRLGPNY LQIPVNCPYR ARVANYQRDG PMCMMDNQGG APNYYPNSFS APEHQPSALE  420
HRTHFSGDVQ RFNSANDDNV TQVRTFYLKV LNEEQRKRLC ENIAGHLKDA QLFIQKKAVK  480
NFSDVHPEYG SRIQALLDKY NEEKPKNAVH TYVQHGSHLS AREKANL                527
```

-continued

SEQ ID NO: 41 moltype = AA length = 502
FEATURE Location/Qualifiers
source 1..502
mol_type = protein
organism = Corynebacterium glutamicum
SEQUENCE: 41
MRPKLSGNTT RHNGAPVPSE NISATAGPQG PNVLNDIHLI EKLAHFNREN VPERIPHAKG 60
HGAFGELHIT EDVSEYTKAD LFQPGKVTPL AVRFSTVAGE QGSPDTWRDV HGFALRFYTE 120
EGNYDIVGNN TPTFFLRDGM KFPDFIHSQK RLNKNGLRDA DMQWDFWTRA PESAHQVTYL 180
MGDRGTPKTS RHQDGFGSHT FQWINAEGKP VWVKYHFKTR QGWDCFTDAE AAKVAGENAD 240
YQREDLYNAI ENGDFPIWDV KVQIMPFEDA ENYRWNPFDL TKTWSQKDYP LIPVGYFILN 300
RNPRNFFAQI EQLALDPGNI VPGVGLSPDR MLQARIFAYA DQQRYRIGAN YRDLPVNRPI 360
NEVNTYSREG SMQYIFDAEG EPSYSPNRYD KGAGYLDNGT DSSSNHTSYG QADDIYVNPD 420
PHGTDLVRAA YVKHQDDDDF IQPGILYREV LDEGEKERLA DNISNAMQGI SEATEPRVYD 480
YWNNVDENLG ARVKELYLQK KA 502

SEQ ID NO: 42 moltype = AA length = 527
FEATURE Location/Qualifiers
source 1..527
mol_type = protein
organism = Homo sapiens
SEQUENCE: 42
MADSRDPASD QMQHWKEQRA AQKADVLTTG AGNPVGDKLN VITVGPRGPL LVQDVVFTDE 60
MAHFDRERIP ERVVHAKGAG AFGYFEVTHD ITKYSKAKVF EHIGKKTPIA VRFSTVAGES 120
GSADTVRDPR GFAVKFYTED GNWDLVGNNT PIFFIRDPIL FPSFIHSQKR NPQTHLKDPD 180
MVWDFWSLRP ESLHQVSFLF SDRGIPDGHR HMNGYGSHTF KLVNANGEAV YCKFHYKTDQ 240
GIKNLSVEDA ARLSQEDPDY GIRDLFNAIA TGKYPSWTFY IQVMTFNQAE TFPFNPFDLT 300
KVWPHKDYPL IPVGKLVLNR NPVNYFAEVE QIAFDPSNMP PGIEASPDKM LQGRLFAYPD 360
THRHRLGPNY LHIPVNCPYR ARVANYQRDG PMCMQDNQGG APNYYPNSFG APEQQPSALE 420
HSIQYSGEVR RFNTANDDNV TQVRAFYVNV LNEEQRKRLC ENIAGHLKDA QIFIQKKAVK 480
NFTEVHPDYG SHIQALLDKY NAEKPKNAIH TFVQSGSHLA AREKANL 527

SEQ ID NO: 43 moltype = AA length = 353
FEATURE Location/Qualifiers
source 1..353
mol_type = protein
organism = Armoracia rusticana
SEQUENCE: 43
MHFSSSSTLF TCITLIPLVC LILHASLSDA QLTPTFYDNS CPNVSNIVRD TIVNELRSDP 60
RIAASILRLH FHDCFVNGCD ASILLDNTTS FRTEKDAFGN ANSARGFPVI DRMKAAVESA 120
CPRTVSCADL LTIAAQQSVT LAGGPSWRVP LGRRDSLQAF LDLANANLPA PFFTLPQLKD 180
SFRNVGLNRS SDLVALSGGH TFGKNQCRFI MDRLYNFSNT GLPDPTLNTT YLQTLRGLCP 240
LNGNLSALVD FDLRTPTIFD NKYYVNLEEQ KGLIQSDQEL FSSPNATDTI PLVRSFANST 300
QTFFNAFVEA MDRMGNITPL TGTQGQIRLN CRVVNSNSLL HDMVEVVDFV SSM 353

SEQ ID NO: 44 moltype = AA length = 308
FEATURE Location/Qualifiers
source 1..308
mol_type = protein
organism = Armoracia rusticana
SEQUENCE: 44
QLSPSFYDKT CPQVFDIATN TIKTALRSDP RIAASILRLH FHDCFVNGCD ASILLDNTTS 60
FRTEKDAFGN ARSARGFDVI DTMKAAVEKA CPKTVSCADL LAIAAQKSVV LAGGPSWKVP 120
SGRRDSLRGF MDLANDNLPG PSSTLQVLKD KFRNVGLDRP SDLVALSGGH TFGKNQCQFI 180
MDRLYNFSNS GKPDPTLDKS YLSTLRKQCP RNGNLSVLVD FDLRTPTIFD NKYYVNLKEN 240
KGLIQSDQEL FSSPDASDTI PLVRAYADGQ GKFFDAFVEA MIRMGNLSPS TGKQGEIRLN 300
CRVVNSKP 308

SEQ ID NO: 45 moltype = AA length = 308
FEATURE Location/Qualifiers
source 1..308
mol_type = protein
organism = Armoracia rusticana
SEQUENCE: 45
QLSPSFYDKT CPQVFDIATN TIKTALRSDP RIAASILRLH FHDCFVNGCD ASILLDNTTS 60
FRTEKDAFGN ARSARGFDVI DTMKAAVEKA CPKTVSCADL LAIAAQKSVV LAGGPSWKVP 120
SGRRDSLRGF MDLANDNLPG PSSTLQVLKD KFRNVGLDRP SDLVALSGGH TFGKNQCQFI 180
MDRLYNFSNS GKPDPTLDKS YLSTLRKQCP RNGNLSVLVD FDLRTPTIFD NKYYVNLKEN 240
KGLIQSDQEL FSSPDASDTI PLVRAYADGQ GKFFDAFVEA MIRMGNLSPS TGKQGEIRLN 300
CRVVNSKP 308

SEQ ID NO: 46 moltype = AA length = 376
FEATURE Location/Qualifiers
source 1..376
mol_type = protein
organism = Streptomyces lividans

-continued

```
SEQUENCE: 46
ATPLTSLGSE QAMFHGKHQP GITTPMQARG HLVAFDLAAG AGRKEAAALL RRWSDTARRL  60
MAGEPAGSRD TDVARDAGPS SLTVTFGFGH SFFGRTGLEK QRPVALDPLP DFSSDHLDKN  120
RSNGDLWVQI GADDALVAFH ALRAIQRDAG AAARVRWQMN GFNRSPGATA HPMTARNLMG  180
QVDGTRNPKP GEADFDRRIF VPEEPEAGKG GPAWMANGSY VVVRRIRMLL DDWEELSLKA  240
QEDVIGRRKS DGAPLSGGSG ATESTEMDLE KTDGSGELVV PINAHARITR PDQNGGAAMV  300
RRPFSYHDGF DADGVPDAGL LFVCWQADPL RGFVPVQRKL DRGDALSQFI RHEASGLFAV  360
PGGAAEGEYV GQRLLE                                                  376

SEQ ID NO: 47          moltype = AA  length = 712
FEATURE                Location/Qualifiers
source                 1..712
                       mol_type = protein
                       organism = Bos taurus
SEQUENCE: 47
MWVCLQLPVF LASVTLFEVA ASDTIAQAAS TTTISDAVSK VKIQVNKAFL DSRTRLKTTL  60
SSEAPTTQQL SEYFKHAKGR TRTAIRNGQV WEESLKRLRR DTTLTNVTDP SLDLTALSWE  120
VGCGAPVPLV KCDENSPYRT ITGDCNNRRS PALGAANRAL ARWLPAEYED GLALPFGWTQ  180
RKTRNGFRVP LAREVSNKIV GYLDEEGVLD QNRSLLFMQW GQIVDHDLDF APETELGSNE  240
HSKTQCEEYC IQGDNCFPIM FPKNDPKLKT QGKCMPFFRA GFVCPTPPYQ SLAREQINAV  300
TSFLDASLVY GSEPSLASRL RNLSSPLGLM AVNQEAWDHG LAYLPFNNKK PSPCEFINTT  360
ARVPCFLAGD FRASEQILLA TAHTLLLREH NRLARELKKL NPHWNGEKLY QEARKILGAF  420
IQIITFRDYL PIVLGSEMQK WIPPYQGYNN SVDPRISNVF TFAFRFGHME VPSTVSRLDE  480
NYQPWGPEAE LPLHTLFFNT WRIIKDGGID PLVRGLLAKK SKLMNQDKMV TSELRNKLFQ  540
PTHKIHGFDL AAINLQRCRD HGMPGYNSWR GFCGLSQPKT LKGLQTVLKN KILAKKLMDL  600
YKTPDNIDIW IGGNAEPMVE RGRVGPLLAC LLGRQFQQIR DGDRFWWENP GVFTEKQRDS  660
LQKVSFSRLI CDNTHITKVP LHAFQANNYP HDFVDCSTVD KLDLSPWASR EN          712

SEQ ID NO: 48          moltype = AA  length = 372
FEATURE                Location/Qualifiers
source                 1..372
                       mol_type = protein
                       organism = Leptoxyphium fumago
SEQUENCE: 48
MFSKVLPFVG AVAALPHSVR QEPGSGIGYP YDNNTLPYVA PGPTDSRAPC PALNALANHG  60
YIPHDGRAIS RETLQNAFLN HMGIANSVIE LALTNAFVVC EYVTGSDCGD SLVNLTLLAE  120
PHAFEHDHSF SRKDYKQGVA NSNDFIDNRN FDAETFQTSL DVVAGKTHFD YADMNEIRLQ  180
RESLSNELDF PGWFTESKPI QNVESGFIFA LVSDFNLPDN DENPLVRIDW WKYWFTNESF  240
PYHLGWHPPS PAREIEFVTS ASSAVLAASV TSTPSSLPSG AIGPGAEAVP LSFASTMTPF  300
LLATNAPYYA QDPTLGPNDK REAAPAATTS MAVFKNPYLE AIGTQDIKNQ QAYVSSKAAA  360
MASAMAANKA RN                                                      372

SEQ ID NO: 49          moltype = AA  length = 378
FEATURE                Location/Qualifiers
source                 1..378
                       mol_type = protein
                       organism = Phanerodontia chrysosporium
SEQUENCE: 49
MAFKSLIAFV ALAAAVRAAP TAVCPDGTRV SHAACCAFIP LAQDLQETIF QNECGEDAHE  60
VIRLTFHDAI AISRSQGPKA GGGADGSMLL FPTVEPNFSA NNGIDDSVNN LIPFMQKHNT  120
ISAADLVQFA GAVALSNCPG APRLEFLAGR PNKTIAAVDG LIPEPQDSVT KILQRFEDAG  180
GFTPFEVVSL LASHSVARAD KVDQTIDAAP FDSTPFTFDT QVFLEVLLKG VGFPGSANNT  240
GEVASPLPLG SGSDTGEMRL QSDFALAHDP RTACIWQGFV NEQAFMAASF RAAMSKLAVL  300
GHNRNSLIDC SDVVPVPKPA TGQPAMFPAS TGPQDLELSC PSERFPTLTT QPGASQSLIA  360
HCPDGSMSCP GVQFNGPA                                                378

SEQ ID NO: 50          moltype = AA  length = 430
FEATURE                Location/Qualifiers
source                 1..430
                       mol_type = protein
                       organism = Thermobifida fusca DyP
SEQUENCE: 50
MTEPDTERKG SSRRGFLAGL GAAALTGAGI GMAAGEVLRP LLPDSDPAAS PKAEQRLRMA  60
AQRADATAAP QPGISGPAPA FVHVIALDLA EEARKNPDTA RDSAAAALRS WTELAARLHE  120
ESPHDIAEGA ASAGLLPASL MVTVGIGGSL LSAIDAEDRR PDALADLPEF STDDLHPRWC  180
GGDFMLQVGA EDPMVLTAAV EELVAAAADA TAVRWSLRGF RRTAAAARDP DATPRNLMGQ  240
IDGTANPAQD HPLFDRTITA RPADNPAHAW MDGGSYLVVR RIRMLLTEWR KLDVAARERV  300
IGRRLDTGAP LGSRNETDPV VLSARDEEGE PLIPENAHVR LASPENNLGA RMFRRGYSYD  360
QGWRDDGVRD AGLLFMAWQG DPATGFIPVQ RSLADQGDAL NRYIRHEGSA LFAVPAAREG  420
RYLGQDLIEG                                                         430

SEQ ID NO: 51          moltype = AA  length = 403
FEATURE                Location/Qualifiers
source                 1..403
                       mol_type = protein
                       organism = Saccharomonospora viridis
```

-continued

```
SEQUENCE: 51
MKGRRFGRRA FLATGAAALG GAAALGGREL FTGDGGADGR DRPDRGDIGR ATVDFHGERQ  60
AGVATPAQAF ATFVAFDLLD GVDREALIRW MRVWTDDIER LTRGAPALTD TEPELALLPA  120
RLTVTVGFGP GFLAAAGREE LRPSWLAPLP EFPIDRLREE FSGGDLVAQV CADDEVTVAH  180
AVRVLTKQAR SFARPRWVQR GFRNTPGAVP EGATMRNLMG QLDGTRNLRP GPDDRLIWIS  240
DGPEWLRGGT GMVVRRIAMN LDTWDELDRP ARELVIGRRL DNGAPLTGRH EHDEPDLEAV  300
DERGLSVIPM FAHIRRARSD NPDERFLRRS YNYDDPPEPG ELSNSGLVFV TFQADIEAQF  360
TPIQRRLAEL DSLNDWTTPI GSAVFAVPRG CRPGEYLGQP LLE                    403

SEQ ID NO: 52          moltype = AA  length = 299
FEATURE                Location/Qualifiers
source                 1..299
                       mol_type = protein
                       organism = Escherichia coli
SEQUENCE: 52
MSQVQSGILP EHCRAAIWIE ANVKGEVDAL RAASKTFADK LATFEAKFPD AHLGAVVAFG  60
NNTWRALSGG VGAEELKDFP GYGKGLAPTT QFDVLIHILS LRHDVNFSVA QAAMEAFGDC  120
IEVKEEIHGF RWVEERDLSG FVDGTENPAG EETRREVAVI KDGVDAGGSY VFVQRWEHNL  180
KQLNRMSVHD QEMMIGRTKE ANEEIDGDER PETSHLTRVD LKEDGKGLKI VRQSLPYGTA  240
SGTHGLYFCA YCARLHNIEQ QLLSMFGDTD GKRDAMLRFT KPVTGGYYFA PSLDKLMAL   299

SEQ ID NO: 53          moltype = AA  length = 343
FEATURE                Location/Qualifiers
source                 1..343
                       mol_type = protein
                       organism = Cellulomonas bogoriensis
SEQUENCE: 53
MTPPPSFIAL VALDLASTSD RASVERLLRV WTVDIERLTT GRPGLADSEP ELALVPAALT  60
VTVGFGPGLL TAAGLRHRAP AWLHPLPPFG IDRLDPAWCD GDVVLQVCAD DRTTLAHAVR  120
VLTKEAQGLA SVRWVQRGFR RSPGISEPDG TSMRNLMGQV EGTANLDPRT DPDLLWHRDG  180
EPGWLTGGTS MVVRRIAMNL DTWDELSRGA REATIGRTLR TGAPLTGRAE HDEPDLEALD  240
DHGRPVIDLE AHIRRARPTQ REETFLRRAY NYDEAPPPGR ASDSGLLFVT YQRDVDAQFT  300
PVQRRLDAAD LLNEWTFPVG SAVFAVPGGW SAGEYVGQRL LEG                    343

SEQ ID NO: 54          moltype = AA  length = 427
FEATURE                Location/Qualifiers
source                 1..427
                       mol_type = protein
                       organism = Thermobifida cellulosilytica
SEQUENCE: 54
MTGPEPTPAG SSRRGFLAGL GAAALTGAGL GMAAGEAVRP LLPDSDSEQP DPVDAQRLDM  60
ARRRADANAA PQPGISGRAP AFVHVIAFDL AEPARAEPAA AREGAATALR TWAEHAARLH  120
ADGPEGAASA GLLPASLMVT IGIGGSLLEA MDAADRRPDA LADLPEFATD DLRPRWCGGD  180
LMLQVGAEDP MVLAAAVDEL VAATAPTTTV RWSLRGFRRT AAAAQDPDAT PRNLMGQIDG  240
TANPAQDHPL FTRTVTAPPA DDPAHAWMDG GSYLVVRRIR MLLDEWRRLD VPDRERVIGR  300
HLDTGAPLGG EKETDPVVLT ARDADGRLVI PEDAHVRLAN PENNLGARMV RRGYNYDEGW  360
RDDGVRDAGL LFMAWQGNPA TGFVPVQRSL VEQGDALNRY TRHEGSALFA VPAATADRYP  420
GQDLVEG                                                            427

SEQ ID NO: 55          moltype = AA  length = 316
FEATURE                Location/Qualifiers
source                 1..316
                       mol_type = protein
                       organism = Streptomyces coelicolor
SEQUENCE: 55
MGGEVEEPEP QMVLSPLTSA AIFLVVTIDS GGEDTVRDLL SDVASLERAV GFRAQPDGRL  60
SCVTGIGSEA WDRLFSGARP AGLHPFRELD GPVHRAVATP GDLLFHIRAS RLDLCFALAT  120
EIMGRLRGAV TPQDEVHGFK YFDERDMLGF VDGTENPTGA AARRAVLVGA EDPAFAGGSY  180
AVVQKYLHDI DAWEGLSVEA QERVIGRRKM TDVELSDDVK PADSHVALTS VTGPDGSDLE  240
ILRDNMPFGS VGREEFGTYF IGYARTPEVT ETMLERMFLG TASAPHDRIL DFSTAVTGSL  300
FFTPAADFLE DLPARP                                                  316

SEQ ID NO: 56          moltype = AA  length = 138
FEATURE                Location/Qualifiers
source                 1..138
                       mol_type = protein
                       organism = Amphitrite ornata
SEQUENCE: 56
MGFKQDIATI RGDLRTYAQD IFLAFLNKYP DERRYFKNYV GKSDQELKSM AKFGDHTEKV  60
FNLMMEVADR ATDCVPLASD ANTLVQMKQH SSLTTGNFEK LFVALVEYMR ASGQSFDSQS  120
WDRFGKNLVS ALSSAGMK                                                138

SEQ ID NO: 57          moltype = AA  length = 261
FEATURE                Location/Qualifiers
source                 1..261
                       mol_type = protein
```

-continued

organism = Thermothelomyces thermophilus

SEQUENCE: 57
MRASVLPVLI AISPALAGFD TWSPPGPYDV RAPCPMLNTL ANHGFLPHDG KDITREQTEN 60
ALFEALHINK TLASFLFDFA LTTNPKNTST FSLNDLGNHN ILEHDASLSR ADAYFGNVLQ 120
FNQTVFDETK TYWEGDTIDL RMAAKARLGR IKTSQATNPT YSMSELGDAF TYGESAAYVV 180
VLGDKESRTV KRSWVEWFFE HEQLPQHLGW KRPAASFEEE DLNSSMEEIE KYTKELEGSN 240
STSGSQKHRR RLPRRRAHFG F 261

SEQ ID NO: 58 moltype = AA length = 504
FEATURE Location/Qualifiers
source 1..504
mol_type = protein
organism = PLEUROTUS SAPIDUS
SEQUENCE: 58
MTTPAPPLDL NNIQGDILGG LPKKTETYFF FDVTNVDRFK ANMTQFIPHV KTSAGIVKDR 60
EAIKEHKRQK RPGLVPMAAV NVSFSHLGLQ KLGITDDLSD SSFTTGQRKD AEVLGDPGTK 120
NGDTFTPAWE APFLKDIHGV IFVAGDCHAS VHKKLDEIKH IFGVGTSHAS ISEVTHVRGD 180
VRPGQVSAHE HFGFLDGISN PAVDQFDQNP FPGQDSIRPG FILAKENGDS RAAARPDWAK 240
DGSFLTFRYL FQMVPEFDDF LESNPIVLPG LSRKEGSELL GARIVGRWKS GAPIEITPLK 300
DDPKLGADAQ RNNNFDFGDS LVRGDQTKCP FAAHIRKTYP RNDLEGPPLN ADIDNRRIIR 360
RGIQFGPEVT SQEHHDKKTH HGRGLLFVCY SSSIDDGFHF IQQSWANAPN FPVNAVTSAG 420
PIPPLDGVIP GFDAIIGQKV GGGIRQISGT NPNDPTTNIT LPDQDFVIPR GGEYFFSPSI 480
SALKTKFAAG VASSAPQAQA PIST 504

SEQ ID NO: 59 moltype = AA length = 480
FEATURE Location/Qualifiers
source 1..480
mol_type = protein
organism = Irpex lacteus
SEQUENCE: 59
MHVKRARSTP LIGSFPGQPP LPTIAQVQST SAGNDSLPFE NIQGDILVGM KKDKEKFVFF 60
HINNATAFKS VLKTYAPANI TSVATIIGPV ANQPLAFVNL AFSHAGFGAL NVTDDLQDTA 120
FSDGQFKDSP NLGDDTSTWE EAFKGTNVDG VFLIGSNDES ITAQYRDDLN AKFGDAWTIV 180
YDLDSAARPG NEKGHEHFGY LDGISNPTIP GFGTPHPGQA VVDPGIIFTG RSKDPVMNRP 240
SWALDGSFLV FRKLKQLVPE FNKYVLDNAL QNQAGNLTVE EGAELLGSRM FGRWKSGAPI 300
DLSPDFDDPA LGNDIERNNN FNYSHPGSDL ATDQTRCPFT AHIRKTNPRD LEGQGLFGDT 360
FHAIRAGTPY GPEVTDYEAS SNTTTIDRGL AFVEYQSVIG NGFRFQQQAW ANNPRFPFSK 420
GPSIQLGLDP VIGQGSPRET FGLDPRNASE SFTVPQVIIS NGGEYFFSPS ITAIVEKFAA 480

SEQ ID NO: 60 moltype = AA length = 250
FEATURE Location/Qualifiers
source 1..250
mol_type = protein
organism = Nicotiana tabacum
SEQUENCE: 60
MGKCYPTVSE EYLKAVDKCK RKLRGLIAEK NCAPLMLRLA WHSAGTYDVC SKTGGPFGTM 60
RLKAEQGHGA NNGIDIAIRL LEPIKEQFPI LSYGDFYQLA GVVAVEVTGG PDVPFHPGRE 120
DKTEPPVEGR LPDATKGSDH LRDVFVKQMG LSDKDIVALS GGHTLGRCHK ERSGFEGPWT 180
TNPLIFDNSY FTELLSGEKE GLLQLPSDKA LLSDPAFRPL VEKYAADEDA FFADYAEAHL 240
KLSELGFAEA 250

SEQ ID NO: 61 moltype = AA length = 519
FEATURE Location/Qualifiers
source 1..519
mol_type = protein
organism = Trametes versicolor
SEQUENCE: 61
MGLQRFSFFV TLALVARSLA AIGPVASLVV ANAPVSPDGF LRDAIVVNGV VPSPLITGKK 60
GDRFQLNVDD TLTNHSMLKS TSIHWHGFFQ AGTNWADGPA FVNQCPIASG HSFLYDFHVP 120
DQAGTFWYHS HLSTQYCDGL RGPFVVYDPK DPHASRYDVD NESTVITLTD WYHTAARLGP 180
RFPLGADATL INGLGRSAST PTAALAVINV QHGKRYRFRL VSISCDPNYT FSIDGHNLTV 240
IEVDGINSQP LLVDSIQIFA AQRYSFVLNA NQTVGNYWVR ANPNFGTVGF AGGINSAILR 300
YQGAPVAEPT TTQTTSVIPL IETNLHPLAR MPVPGSPTPG GVDKALNLAF NFNGTNFFIN 360
NATFTPPTVP VLLQILSGAQ TAQDLLPAGS VYPLPAHSTI EITLPATALA PGAPHPFHLH 420
GHAFAVVRSA GSTTYNYNDP IFRDVVSTGT PAAGDNVTIR FQTDNPGPWF LHCHIDFHLD 480
AGFAIVFAED VADVKAANPV PKAWSDLCPI YDGLSEANQ 519

SEQ ID NO: 62 moltype = AA length = 559
FEATURE Location/Qualifiers
source 1..559
mol_type = protein
organism = Phanerodontia chrysosporium
SEQUENCE: 62
MLSLLAVVSL AAATLAAPAA SDAPGWRFDL KPNLSGIVAL EAIVVNSSLV VIFDRATGDQ 60
PLKINGESTW GALWDLDTST VRPLSVLTDS FCASGALLSN GTMVSMGGTP GGTGGDVAAP 120
PGNQAIRIFE PCASPSGDGC TLFEDPATVH LLEERWYPSS VRIFDGSLMI IGGSHVLTPF 180

-continued

```
YNVDPANSFE FFPSKEQTPR PSAFLERSLP ANLFPRAFAL PDGTVFIVAN NQSIIYDIEK  240
NTETILPDIP NGVRVTNPID GSAILLPLSP PDFIPEVLVC GGSTADTSLP STSLSSQHPA  300
TSQCSRIKLT PEGIKAGWQV EHMLEARMMP ELVHVPNGQI LITNGAGTGF AALSAVADPV  360
GNSNADHPVL TPSLYTPDAP LGKRISNAGM PTTTIPRMYH STVTLTQQGN FFIGGNNPNM  420
NFTPPGTPGI KFPSELRIET LDPPFMFRSR PALLTMPEKL KFGQKVTVPI TIPSDLKASK  480
VQVALMDLGF SSHAFHSSAR LVFMESSISA DRKSLTFTAP PNGRVFPPGP AVVFLTIDDV  540
TSPGERVMMG SGNPPPTLE                                               559

SEQ ID NO: 63          moltype = AA  length = 928
FEATURE                Location/Qualifiers
source                 1..928
                       mol_type = protein
                       organism = Myceliophthora thermophila
SEQUENCE: 63
MTNAACLSYC ASKGFPYAGT EYSVECFCGT TLASSSAKVA DSECNMPCSG APSEPCGAGS  60
RLSLFHSSAV TGPAANPGVN DFTHLGCYAE GKTGRALTYN PGLPGADMTV AKCTAACRAA  120
NYILAGVEYG GECYCGNTIA NGGAPADSGC SMVCNGNSTE FCGGPDRLNV YSYKNQYEPT  180
ATSTTGAGST SSSSVPSATG LPEGWSYQGC WIDGKQGRIL PYQLPDSQTN SRAACANACA  240
EAGYTVSGTE YAVQCFCGDA IHNGGVETDE ADCSTPCPGA PGEKCGAGDR LSIVSRGPPK  300
IYAPPAPIEK IGDWEYQGCA EDNINDKRTF FWQIFFNDIM TPEMCLDRCA EFGYHAAGLE  360
YGQECYCGDP ANMATHGATF RPESECNVVC AGNSTAICGG LARLTTYFWI GTPFYSWDFP  420
QDWRAGKYEF LVDGVNIPLI THETITGKVS FISKGATGPG NETGAYEFDP ATLEFRELHI  480
KTDVFCAASV TLPDKAGRQL NVGGWAGEAT YGTRLYWPDG APGVPGTHDW QENVNVLHLQ  540
AGRWYPSVLV LTNGSVMVVG GLIGSNDAAT PSIEILPYTG TPPLYMDWLD RTHPNNLYPF  600
LCILPGGGIF VQYWNEARIL DPVTFDTVKT LPDAPGAPND PKGGRTYPLE GTAVLLPQKY  660
PYTDPLGVLI CGGSTEGPGN ALDNCVSIYP EADEPEWQIE RMPSFRVMTC MAPLPDGTYL  720
IANGALHGVA GFGLGVGPNL NALLYDPSKP LGSRITVAAN TTIARMYHSE AITLLDGRVL  780
ISGSNPEDGV NPEEYRVEVF LPPYLLAGKP RPTFTLENRD WAHGQTGIPF TLGSPARNGD  840
ITATLLGSVA STHGNSMGAR TLMPRVSCRG TSCTVDAPPT ANICPPGWYQ FFVLDGGIPA  900
VGVYVRIGGD AGQIGNWPQA PDFSVPGV                                     928

SEQ ID NO: 64          moltype = AA  length = 558
FEATURE                Location/Qualifiers
source                 1..558
                       mol_type = protein
                       organism = Pycnoporus cinnabarinus
SEQUENCE: 64
MAPTAFSLVS ALALASLSLA APSAPGWSFD LKKETSGIVA LEAIVVSPTL VVFFDRASDD  60
PLQINNHSAW GALWNLESST VRPLDVLTNS FCASGALLSN GTMASVGGDP DGFVGNPAIR  120
PGNQAIRLFE PCDSPTGDGC TLFEDPATLH LLEKRWYPSS ARIFDGSLII VGGMHEATPF  180
YNTDPALSFE FFPRKEDTPR PSEFLNRSLP ANLFPRVFAL PDGKVFMVAN NQSIIYDIEA  240
KTERILPDVP NNVRVTNPMD GSAILLPLSP PDFVPEVLVC GGSQTDTIDP SLLTSQTPAS  300
SQCSRIRLDE EGIAKGWEVE HMLEGRIMPE LVHLPNGQVL IANGGRTGFA AIASVSEPVG  360
NSNADHPVLV PSLYTPDAPL GRRISNVGLP SSGIPRLYHS SVTLTPQGNF LIAGSNPNNR  420
TTVGPGIKFP SEFRVQTLDP PFMSVERPKI LNMPKKLGFN KSFTVPISVP SSLARPGAKV  480
QISLMDLGFS SHAFHSSARL VFMNGKISQD SKSLTFTTPP NGRVYPPGPA TVFLTIDDVT  540
SEGAWVMMGS GNSPPTLE                                                558

SEQ ID NO: 65          moltype = AA  length = 559
FEATURE                Location/Qualifiers
source                 1..559
                       mol_type = protein
                       organism = Pycnoporus cinnabarinus
SEQUENCE: 65
MFQTTLHLLF VLVVTGRGLA APSTPTGWQF NLKAERSGIV ALESIVVSPT LVVFFDRATN  60
DPLQINNHSA WGALWNLETS TVRALDVLTN SFCASGALLS NGTMASIGGD PNGFPGNPAI  120
HPGTQAIRLF EPCDSPTGEG CTLFEDPVTL HLLEKRWYPS SVRIFDGSLL IVGGMHEETP  180
FYNTDPALSF EFFPPKESTP RPSEFLNRSL PANLFPRVFA LPDGKVFMVA NNQSIIYDIE  240
ANTERILPDI PNNVRVTNPI DGSAILLPLS PPDFVPEVLV CGGTQTDTID PSLLTSQTPA  300
SSQCSRIRLD EEGIARGWEV EHMLEGRMMP ELVHLPNGQV LIANGARTGF AAIASVSDPV  360
GGSNADHAVL VPSLYTPDAP LGTRISNVGL PSSGIARVYH SSITLTPQGN FLIAGSNPNN  420
NSSVTAGVKF PSEFRVQTLD PPFMFVERPK ILSMPKKLAF GKSFTVPIAV PSTLAHPGAK  480
VQVSLMDLGF SSHAFHSSAR LVFMNAKISQ DGKSLTFTTP PNGRVYPPGP ATIFLTIDDV  540
TSEGAWVMMG SGNPPPTLE                                               559

SEQ ID NO: 66          moltype = AA  length = 656
FEATURE                Location/Qualifiers
source                 1..656
                       mol_type = protein
                       organism = synthetic construct
SEQUENCE: 66
MGHHHHHHSS GHIEGRHMAS APIGSAIPRN NWAVTCDSAQ SGNECNKAID GNKDTFWHTF  60
YGANGDPKPP HTYTIDMKTT QNVNGLSVLP RQDGNQNGWI GRHEVYLSSD GTNWGSPVAS  120
GSWFADSTTK YSNFETRPAR YVRLVAITEA NGQPWTSIAE INVFQASSYT APQPGLGRWG  180
PTIDLPIVPA AAAIEPTSGR VLMWSSYRND AFEGSPGGIT LTSSWDPSTG IVSDRTVTVT  240
KHDMFCPGIS MDGNGQIVVT GGNDAKKTSL YDSSSDSWIP GPDMQVARGY QSSATMSDGR  300
```

-continued

```
VFTIGGSFSG GVFEKNGEVY SPSSKTWTSL PNAKVNPMLT ADKQGLYKSD NHAWLFGWKK  360
GSVFQAGPST AMNWYYTSGS GDVKSAGKRQ SNRGVAPDAM SGNAVMYDAV KGKILTFGGS  420
PDFEDSDATT NAHIITLGEP GTSPNTVFAS NGLYFARTFH TSVVLPDGST FITGGQRRGI  480
PFEDSTPVFT PEIYVPEQDT FYKQNPNSIV RAYHSISLLL PDGRVFNGGG GLCGDCTTNH  540
FDAQIFTPNY LYDSNGNLAT RPKITRTSTQ SVKVGGRITI STDSSISKAS LIRYGTATHT  600
VNTDQRRIPL TLTNNGGNSY SFQVPSDSGV ALPGYWMLFV MNSAGVPSVA STIRVT      656
```

What is claimed is:

1. A method of preparing glucaric acid, comprising:

contacting D-glucose and oxygen with a first catalyst composition comprising a first copper radical oxidase, a single electron oxidizer, and a small molecule activator under conditions suitable for formation of glucodialdose; and

contacting glucodialdose and oxygen with a second catalyst composition comprising a second copper radical oxidase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of a product mixture comprising glucaric acid or salts thereof.

2. The method of claim **1**, wherein the first copper radical oxidase comprises a galactose oxidase, a galactose oxidase mutant or combinations thereof.

3. The method of claim **1**, wherein the first copper radical oxidase comprises any of SEQ ID Nos. 1 through 39.

4. The method of claim **1**, wherein the first copper radical oxidase has from about 85% to about 100% sequence identity with any of SEQ ID Nos. 1 through 39.

5. The method of claim **1**, wherein the second copper radical oxidase comprises a glyoxal oxidase, a glyoxal oxidase mutant or combinations thereof.

6. The method of claim **1**, wherein the second copper radical oxidase comprises any of SEQ ID Nos. 62 through 65.

7. The method of claim **1**, wherein the second copper radical oxidase has from about 85% to about 100% sequence identity with any of SEQ ID Nos. 62 through 65.

8. The method of claim **1**, wherein the small molecule activator comprises L-tryptophan, 2-mercaptobenzothiazole, L-histidine, methylchloroisothiazolinone, o-dianisidine, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), 4-aminoantipyrine, L-tyrosine, (2,2,6,6-tetramethylpiperidin-1-yl)oxyl, chloromethylisothiazolinone, 4-thiazolecarboxylic acid, Sunset yellow FCF, tartrazine, p-benzoquinone, dicoumarol, phthalimide, saccharin, phthalic anhydride, erythrosine B, 2-aminobenzothiazole, thiabendazole, 2-hydroxybenzothiazole, phenothiazine, 6-aminobenzothiazole, indigo carmine, naphthalimide, 2-aminothiazole, thiazole, 2H-1,4-benzothiazin-3 (4H)-one, 2-oxindole, beta-lapachone, menaquinone, thiamine, 4-methyl-5-thiazoleethanol, Allura Red AC, menadione, p-cresol, Fast green FCF, Brilliant Blue FCF, methylisothiazolinone, caffeine, veratryl alcohol, fluorescein, or combinations thereof.

9. The method of claim **1**, wherein the small molecule activator is present in an amount ranging from about 1 ppm to about 500 ppm.

10. The method of claim **1**, wherein the single electron oxidizer comprises laccase, horseradish peroxidase, dyp-type peroxidase, lactoperoxidase, chloroperoxidase, manganese peroxidase 1, ascorbate peroxidase, dye-decolorizing peroxidase, unspecific peroxygenase, dehaloperoxidase, catalase-peroxidase, lignin peroxidase, soybean seed coat peroxidase, isoforms thereof or combinations thereof.

11. The method of claim **1**, wherein conditions suitable for formation of glucodialdose, conditions suitable for formation of a product mixture comprising glucaric acid or both comprise an oxygen pressure of equal to or less than about 500 psi.

12. The method of claim **1**, further comprising a catalase.

13. The method of claim **1**, wherein the catalase is defined by any SEQ ID No. 40 through SEQ ID No. 44.

14. The method of claim **1**, further comprising introducing a caustic to the second catalyst composition, the product mixture or both.

15. The method of claim **1**, wherein the glucodialdose, glucaric acid are formed at yields ranging from about 50% to about 99%.

16. The method of claim **1**, wherein the glucaric acid or salts thereof have a purity of greater than about 70%.

17. The method of claim **1**, wherein the product mixture comprising glucaric acid further comprises on or more sugar oxidation products.

18. The method of claim **17**, wherein the one or more sugar oxidation products comprise aldonic acid, uronic acid, aldaric acid, a gluconic acid oxidation product, a gluconate, gluconic acid, glucuronic acid, glucose oxidation products, galactonic acid, galactaric acid, glutamic acid, a lactone of gluconic acid, a lactone of glucaric acid, a lactone of galactaric acid, a lactone of galactonic acid, glucodialdose, 2-ketoglucose, disaccharides, oxidized disaccharides, n-keto-acids, C2 to C6 diacids, salts thereof or combinations thereof.

19. A method comprising:

contacting a sugar with a catalyst composition comprising an oxidoreductase, a single electron oxidizer and a small molecule activator under conditions suitable for the formation of one or more oxidized sugar oxidation products comprising glucaric acid wherein the oxidoreductase comprises at least two copper radical oxidases, at least two mutated copper radical oxidases or combinations thereof.

20. The method of claim **19**, wherein the at least two copper radical oxidases comprises any of SEQ ID Nos. 1 through 39, any of SEQ ID Nos. 62 through 65, have 85% to about 100% sequence identity with any of SEQ ID Nos. 1 through 39 or have 85% to about 100% sequence identity with any of SEQ ID Nos. 62 through 65.

* * * * *