(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101511856 B
(45)授权公告日 2016.01.20

(21)申请号 200780033064.X

(22)申请日 2007.09.07

(30)优先权数据
60/843,264 2006.09.08 US

(85)PCT国际申请进入国家阶段日
2009.03.06

(86)PCT国际申请的申请数据
PCT/US2007/019656 2007.09.07

(87)PCT国际申请的公布数据
WO2008/030614 EN 2008.03.13

(73)专利权人 AMBRX公司
地址 美国加利福尼亚州

(72)发明人 田锋 席雅·诺曼 斯蒂芬妮·周

(74)专利代理机构 北京科龙寰宇知识产权代理有限责任公司 11139
代理人 孙皓晨

(51)Int.Cl.
C12N 15/63(2006.01)
C12N 5/10(2006.01)

(56)对比文件
WO 2006/001832 A2,2006.01.05, 第15、31、37、63-67、172段.
US 2004/0265952 A1,2004.12.30,权利要求1、13、20、24-25、136-138.
WO 2004/094593 A2,2004.11.04, 全文.
WO 2006/001832 A2,2006.01.05, 第15、31、37、63-67、172段.

审查员 蔺娜

权利要求书1页 说明书74页
序列表76页 附图4页

(54)发明名称

脊椎动物细胞中抑制因子TRNA的转录

(57)摘要

本发明提供用于产生扩充脊椎动物细胞中遗传编码的氨基酸的数目的翻译组件的组合物和方法。所述组件包括正交tRNA、正交氨酰基tRNA合成酶、tRNA/合成酶的正交对以及非天然氨基酸。本发明还提供具有非天然氨基酸的蛋白质和在脊椎动物细胞中产生具有非天然氨基酸的蛋白质的方法。

CN 101511856 B

1.一种抑制因子 tRNA 转录盒,所述盒由 5’至 3’依序包含:U6 或 H1 的 pol III 启动子;以及嗜热脂肪芽孢杆菌(Bacillus stearothermophilus)或大肠杆菌(Escherichia coli)抑制因子 tRNA。

2.一种哺乳动物细胞系,其包含根据权利要求 1 所述的 tRNA 转录盒。

脊椎动物细胞中抑制因子 TRNA 的转录

技术领域

[0001] 本发明属于脊椎动物细胞中的翻译生物化学领域。本发明涉及在脊椎动物细胞中产生正交 tRNA、正交合成酶和其对的方法以及正交 tRNA、正交合成酶和其对的组合物。本发明还涉及非天然氨基酸的组合物、包括非天然氨基酸的蛋白质和在脊椎动物细胞中产生包括非天然氨基酸的蛋白质的方法。

背景技术

[0002] 从细菌到人类，每种已知生物体的遗传密码都编码相同的二十种常见氨基酸。这相同的二十种天然氨基酸的不同组合形成进行生命的几乎所有的复杂过程（从光合作用到信号转导和免疫反应）的蛋白质。为研究和修饰蛋白质结构和功能，科学家们曾尝试操纵蛋白质的遗传密码和氨基酸序列。然而，很难消除由遗传密码所施加的使蛋白质局限于二十种遗传编码的标准结构单元 (building block) 的限制（其中极个别例外为硒代半胱氨酸 (selenocysteine)（例如参见 A. Bock，(1991)，Molecular Microbiology 5：515-20) 和吡咯赖氨酸 (pyrrolysine)（例如参见 G. Srinivasan 等人，(2002)，Science 296：1459-62))。

[0003] 虽然已取得一些进展来消除这些限制，但这一进展受到限制且合理控制蛋白质结构和功能的能力仍处于初期阶段。举例来说，化学家们已开发出合成和操纵小分子结构的方法和策略（例如参见 E. J. Corey 和 X. M. Cheng，The Logic of ChemicalSynthesis(Wiley-Interscience，New York，1995))。虽然全合成（例如参见 B. Merrifield，(1986)，Science 232：341-7(1986)) 和半合成方法（例如参见 D. Y. Jackson，(1994)，Science 266：243-7；以及 P. E. Dawson 和 S. B. Kent，(2000)，Annual Review ofBiochemistry 69：923-60) 已使得有可能合成肽和小蛋白质，但这些方法限制了超过 10 千道尔顿 (kDa) 的蛋白质的效用。诱变方法尽管有效，但也局限于有限数目的结构变化。在许多情况下，有可能在整个蛋白质中竞争性并入常见氨基酸的闭合结构类似物。例如参见 R. Furter，(1998)，Protein Science 7：419-26；K. Kirshenbaum 等人，(2002)，ChemBioChem 3：235-7；和 V. Doring 等人，(2001)，Science 292：501-4。

[0004] 在尝试扩充操纵蛋白质结构和功能的能力的过程中，开发出使用化学酰化的正交 tRNA 的活体外方法，其使得能在活体外响应无义密码子而选择性并入非天然氨 基酸（例如参见 J. A. Ellman 等人，(1992)，Science，255：197-200)。将具有新颖结构和物理性质的氨基酸选择性并入蛋白质中，从而来研究蛋白质的折叠和稳定性以及生物分子的识别和催化作用。例如参见 D. Mendel 等人，(1995)，Annual Review ofBiophysics and Biomolecular Structure 24：435-462；和 V. W. Cornish 等 人，(1995 年 3 月 31 日)，Angewandte Chemie-International Edition in English 34：621-633。然而，这个过程的化学计量性质严重限制了可生成的蛋白质的量。

[0005] 非天然氨基酸已可微量注射于细胞中。举例来说，通过微量注射化学错酰化的嗜热四膜虫 (Tetrahymena thermophila) tRNA（例如，M. E. Saks 等人，(1996)，Anengineered

Tetrahymena tRNA Glnfor in vivo incorporation of unnatural amino acidsinto proteins by nonsense suppression，J.Biol.Chem.271:23169-23175) 和相关 mRNA 而将非天然氨基酸引入爪蟾 (Xenopus) 卵母细胞的烟碱型乙酰胆碱受体中（例如，M.W.Nowak 等人，(1998)，In vivo incorporation of unnatural amino acids into ionchannels in Xenopus oocyte expression system，Method Enzymol.293:504-529)。这允许通过引入具有独特物理或化学性质的含侧链氨基酸来对卵母细胞中的受体进行详细的生物物理学研究。例如参见 D.A.-Dougherty(2000)，Unnatural amino acids asprobes of protein structure and function，Curr.Opin.Chem.Biol.4:645-652。令人遗憾的是，这种方法局限于细胞中可进行微量注射的蛋白质，且因为相关 tRNA 在活体外化学酰化且不能再酰化，所以蛋白质的产率非常低。

[0006] 为克服这些局限性，向原核生物大肠杆菌 (Escherichia coli，E.coli) 的蛋白质生物合成机构中添加新颖组件（例如，L.Wang 等人，(2001)，Science 292:498-500)，这允许在活体内遗传编码非天然氨基酸。已使用这种方法响应琥珀密码子 TAG 将多种具有新颖化学、物理或生物性质的新颖氨基酸有效且高保真地并入大肠杆菌的蛋白质中，所述新颖氨基酸包括光亲和标记 (photoaffinity label) 和光致异构化氨基酸、酮基氨基酸以及糖基化氨基酸。例如参见 J.W.Chin 等人，(2002)，Journal of theAmerican Chemical Society 124:9026-9027；J.W.Chin 和 P.G.Schultz，(2002)，ChemBioChem 11:1135-1137；J.W.Chin 等人，(2002)，PNAS United States of America99:11020-11024；以及 L.Wang 和 P.G.Schultz，(2002)，Chem.Comm.1-10。然而，原核生物和真核生物的翻译机构并非高度保守的；因此，添加到大肠杆菌中的生物合成机构的组件通常无法用于将非天然氨基酸位点特异性并入脊椎动物细胞中的蛋白质中。举例来说，用于大肠杆菌中的詹氏甲烷球菌 (Methanococcus jannaschii) 酪氨酰基 tRNA 合成酶 /tRNA 对在脊椎动物细胞中并不正交。另外，真核生物而非原 核生物中 tRNA 的转录是通过 RNA 聚合酶 III 进行，且这对脊椎动物细胞中可转录的 tRNA 结构基因的一级序列产生限制。此外，与原核生物细胞相反，脊椎动物细胞中的 tRNA 需要从转录所述 tRNA 的核中输出到细胞质中，以进行翻译。最后，脊椎动物的 80S 核糖体不同于 70S 原核生物核糖体。因此，需要开发出改进的生物合成机构组件，以扩充脊椎动物遗传密码。阅读下列公开内容后将显而易见，本发明满足这些和其它需要。

发明内容

[0007] 本发明对脊椎动物细胞提供翻译组件，所述翻译组件可用于脊椎动物蛋白质生物合成机构中以将非天然氨基酸并入脊椎动物细胞中正在生长的多肽链中，例如正交氨酰基 tRNA 合成酶 (O-RS) 和正交 tRNA(O-tRNA) 的对和其个别组件。

[0008] 本发明的组合物包括包含正交氨酰基 tRNA 合成酶 (O-RS)（例如，源自诸如大肠杆菌、嗜热脂肪芽孢杆菌等非脊椎动物生物体）的脊椎动物细胞（例如，哺乳动物细胞、禽类细胞、鱼类细胞、爬行动物细胞、两栖动物细胞、源自非哺乳动物的细胞等），其中 O-RS 在脊椎动物细胞中优先利用至少一个非天然氨基酸使正交 tRNA(O-tRNA) 氨酰化。在给定的脊椎动物细胞中，可视情况使两个或两个以上 OtRNA 氨酰化。一方面，O-RS 利用非天然氨基酸使 O-tRNA 氨酰化，这与具有（例如）如 SEQ ID NO.:86 或 45 中所述的氨基酸序列的 O-RS

例如至少40%、至少45%、至少50%、至少60%、至少75%、至少80%或者甚至90%或90%以上有效。在一个实施例中，本发明的O-RS利用非天然氨基酸使O-tRNA氨酰化，这比O-RS利用天然氨基酸使O-tRNA氨酰化有效例如至少10倍、至少20倍、至少30倍等。

[0009]　在一个实施例中，O-RS或其部分是由如SEQ ID NO:3-35中任一序列所述的多核苷酸序列或其互补多核苷酸序列编码。在另一实施例中，O-RS包含如SEQ ID NO:36-63和/或86中任一序列所述的氨基酸序列或其保守变异体。在又一实施例中，O-RS包含与天然存在的酪氨酰基氨酰基tRNA合成酶(TyrRS)的氨基酸序列例如至少90%、至少95%、至少98%、至少99%或至少99.5%或99.5%以上一致且包含两个或两个以上来自群组A-E的氨基酸的氨基酸序列。群组A包括在对应于大肠杆菌TyrRS的Tyr37的位置处的缬氨酸、异亮氨酸、亮氨酸、甘氨酸、丝氨酸、丙氨酸或苏氨酸。群组B包括在对应于大肠杆菌TyrRS的Asn126的位置处的天冬氨酸。群组C包括在对应于大肠杆菌TyrRS的Asp182的位置处的苏氨酸、丝氨酸、精氨酸、天冬酰胺或甘氨酸。群组D包括在对应于大肠杆菌TyrRS的Phe183的位置处 的蛋氨酸、丙氨酸、缬氨酸或酪氨酸；且群组E包括在对应于大肠杆菌TyrRS的Leu186的位置处的丝氨酸、蛋氨酸、缬氨酸、半胱氨酸、苏氨酸或丙氨酸。

[0010]　在另一实施例中，与天然氨基酸相比，O-RS具有一种或一种以上针对非天然氨基酸的改进或增强的酶促性质。举例来说，与天然氨基酸相比，针对非天然氨基酸的改进或增强的性质包括例如较高Km、较低Km、较高kcat、较低kcat、较低kcat/km、较高kcat/km等中的任一种。

[0011]　脊椎动物细胞还视情况包括非天然氨基酸。脊椎动物细胞视情况包括正交tRNA(O-tRNA)(例如源自诸如大肠杆菌、嗜热脂肪芽孢杆菌等非脊椎动物生物体)，其中O-tRNA识别选择密码子(selector codon)且优先通过O-RS经非天然氨基酸氨酰化。一方面，O-tRNA以包含如SEQ ID NO:65中所述的多核苷酸序列或在细胞中由如SEQ ID NO:65中所述的多核苷酸序列加工而成的tRNA的效率的例如至少45%、至少50%、至少60%、至少75%、至少80%、至少90%、至少95%或99%或99%以上，介导将非天然氨基酸并入蛋白质中。另一方面，O-tRNA包含SEQ ID NO:65的序列，且O-RS包含选自SEQ ID NO:36-63和/或86中任一序列中所述的氨基酸序列的多肽序列和/或其保守变异体。

[0012]　在另一实施例中，脊椎动物细胞包含包括编码所关注的多肽的多核苷酸的核酸，其中多核苷酸包含由O-tRNA识别的选择密码子。一方面，包含非天然氨基酸的所关注的多肽的产率为从多核苷酸缺乏选择密码子的细胞获得所关注的天然存在的多肽的产率的例如至少2.5%、至少5%、至少10%、至少25%、至少30%、至少40%、50%或50%以上。另一方面，细胞在不存在非天然氨基酸的情况下以一定产率产生所关注多肽，所述产率例如为在存在非天然氨基酸的情况下多肽产率的不到35%、不到30%、不到20%、不到15%、不到10%、不到5%、不到2.5%等。

[0013]　本发明也提供包含正交氨酰基tRNA合成酶(O-RS)、正交tRNA(O-tRNA)、非天然氨基酸和包含编码所关注的多肽的多核苷酸的核酸的脊椎动物细胞。所述多核苷酸包含由O-tRNA识别的选择密码子。另外，O-RS在脊椎动物细胞中优先利用非天然氨基酸使正交tRNA(O-tRNA)氨酰化，且所述细胞在不存在非天然氨基酸的情况下以一定产率产生所关注的多肽，所述产率例如为在存在非天然氨基酸的情况下多肽产率的不到30%、不到20%、不到15%、不到10%、不到5%、不到2.5%等。

[0014] 包括包含正交 tRNA(O-tRNA) 的脊椎动物细胞的组合物也是本发明的特征。O-tRNA 通常介导在活体内将非天然氨基酸并入由包含由 O-tRNA 识别的选择密码 子的多核苷酸编码的蛋白质中。在一个实施例中,O-tRNA 以包含如 SEQ ID NO :65 中所述的多核苷酸序列或在细胞中由如 SEQ ID NO :65 中所述的多核苷酸序列加工而成的 tRNA 的效率的例如至少 45%、至少 50%、至少 60%、至少 75%、至少 80%、至少 90%、至少 95%或甚至 99%或 99%以上,介导将非天然氨基酸并入蛋白质中。在另一实施例中,O-tRNA 包含如 SEQ ID NO :65 中所述的多核苷酸序列或其保守变异体,或由如 SEQ ID NO :65 中所述的多核苷酸序列或其保守变异体加工而成。在又一实施例中,O-tRNA 包含可重复利用的 O-tRNA。

[0015] 在本发明的一方面,O-tRNA 经转录后修饰。本发明也提供编码脊椎动物细胞中的 O-tRNA 的核酸或其互补多核苷酸。在一个实施例中,核酸包含 A 盒 (A box) 和 B 盒 (B box)。

[0016] 本发明的特征也在于产生例如 O-RS 或 O-tRNA/O-RS 对的翻译组件的方法 (和由这些方法产生的翻译组件)。举例来说,本发明提供产生在脊椎动物细胞中优先利用非天然氨基酸使正交 tRNA 氨酰化的正交氨酰基 tRNA 合成酶 (O-RS) 的方法。例如,所述方法包括 (a) 在存在非天然氨基酸的情况下使第一物种的脊椎动物细胞群体经受阳性选择,其中所述脊椎动物细胞各自包含 :i) 氨酰基 tRNA 合成酶 (RS) 的文库的成员,ii) 正交 tRNA(O-tRNA),iii) 编码阳性选择标记的多核苷酸以及 iv) 编码阴性选择标记的多核苷酸 ;其中在阳性选择中存活的细胞包含在存在非天然氨基酸的情况下使正交 tRNA(O-tRNA) 氨酰化的活性 RS。使在阳性选择中存活的细胞在不存在非天然氨基酸的情况下经受阴性选择,以去除利用天然氨基酸使 O-tRNA 氨酰化的活性 RS。这提供优先利用非天然氨基酸使 O-tRNA 氨酰化的 O-RS。

[0017] 在某些实施例中,将编码阳性选择标记的多核苷酸可操作性连接到反应元件,且所述细胞进一步包含多核苷酸,其 a) 编码调节从反应元件进行的转录的转录调节蛋白 (modulator protein)(例如脊椎动物转录调节蛋白等),和 b) 包含至少一个选择密码子。通过经非天然氨基酸氨酰化的 O-tRNA 将非天然氨基酸并入转录调节蛋白中会引起阳性选择标记的转录。在一个实施例中,转录调节蛋白为转录激活蛋白 (例如 GAL4 等),且选择密码子为琥珀终止密码子,举例来说,其中所述琥珀终止密码子位于编码转录激活蛋白的 DNA 结合域的多核苷酸的一部分中,或实质上靠近所述部分。

[0018] 阳性选择标记可为多种分子中的任一种。在一个实施例中,阳性选择标记包含供生长使用的营养增补剂,且所述选择是在缺乏所述营养增补剂的培养基上进行。在另一实施例中,编码阳性选择标记的多核苷酸例如为 ura3、leu2、lys2、lacZ 基因、his3(例如,其中通过提供 3- 氨基三唑 (3-AT) 所检测,his3 基因编码咪唑磷酸甘油酯脱水酶) 等。在又一实施例中,编码阳性选择标记的多核苷酸包含选择密码子。

[0019] 与阳性选择标记一样,阴性选择标记也可为多种分子中的任一种。在某些实施例中,将编码阴性选择标记的多核苷酸可操作性连接到反应元件,其中从反应元件进行的转录是通过转录调节蛋白来介导。通过经天然氨基酸氨酰化的 O-tRNA 将天然氨基酸并入转录调节蛋白中会引起阴性选择标记的转录。在一个实施例中,编码阴性选择标记的多核苷酸例如为 ura3 基因,且阴性选择是在包含 5- 氟乳清酸 (5-FOA) 的培养基上实现。在另一实施例中,用于阴性选择的培养基包含通过阴性选择标记转化为可检测物质的选择或筛选

剂。在本发明的一方面，可检测物质为有毒物质。在一个实施例中，编码阴性选择标记的多核苷酸包含选择密码子。

[0020] 在某些实施例中，阳性选择标记和 / 或阴性选择标记包含发荧光或在存在合适反应物的情况下催化发光反应的多肽。在本发明的一方面，通过荧光激活细胞分类 (fluorescence-activated cell sorting,FACS) 或通过发光性来检测阳性选择标记和 / 或阴性选择标记。在某些实施例中，阳性选择标记和 / 或阴性选择标记包含基于亲和力的筛选标记或转录调节蛋白。在一个实施例中，同一多核苷酸编码阳性选择标记和阴性选择标记。

[0021] 在一个实施例中，编码本发明的阳性选择标记和 / 或阴性选择标记的多核苷酸可包含至少两个选择密码子，所述选择密码子各自或都可包含至少两个不同选择密码子或至少两个相同选择密码子。

[0022] 本发明的方法中也可使用其它水平的选择 / 筛选严格度。在一个实施例中，所述方法可例如包含在步骤 (a)、(b) 或 (a) 和 (b) 中提供不同量的无活性合成酶，其中所述不同量的无活性合成酶提供另一水平的选择或筛选严格度。在一个实施例中，产生 O-RS 的方法的步骤 (a)、(b) 或步骤 (a) 和 (b) 包括改变例如阳性和 / 或阴性选择标记的选择或筛选严格度。所述方法视情况包括使优先利用非天然氨基酸使 O-tRNA 氨酰化的 O-RS 经受另一轮选择，例如另一轮（多轮）阳性选择、另一轮（多轮）阴性选择或另一轮阳性选择和另一轮阴性选择的组合。

[0023] 在一个实施例中，选择 / 筛选包含一种或一种以上选自例如氨基酸渗透性改变、翻译效率改变、翻译保真性改变等的阳性或阴性选择 / 筛选。所述一种或一种以上改变是基于编码用于产生蛋白质的正交 tRNA-tRNA 合成酶对的组件的一个或一个以上多核苷酸的突变。

[0024] RS 的文库（例如突变 RS 的文库）通常包含源自至少一种例如来自非脊椎动物 生物体的氨酰基 tRNA 合成酶 (RS) 的 RS。在一个实施例中，RS 的文库源自无活性 RS，例如，其中无活性 RS 是通过使活性 RS 突变而生成。在另一实施例中，无活性 RS 包含氨基酸结合袋，且包含结合袋的一个或一个以上氨基酸经一个或一个以上不同氨基酸取代，例如所述经取代氨基酸经丙氨酸取代。

[0025] 在某些实施例中，产生 O-RS 的方法进一步包括对编码 RS 的核酸进行随机突变、位点特异性突变、重组、嵌合构筑或其任何组合，从而产生突变 RS 的文库。在某些实施例中，所述方法例如进一步包括 (c) 分离编码 O-RS 的核酸；(d) 由核酸生成一组编码突变 O-RS 的多核苷酸（例如通过随机诱变、位点特异性诱变、嵌合构筑、重组或其任何组合）；和 (e) 重复步骤 (a) 和 / 或 (b)，直到获得优先利用非天然氨基酸使 O-tRNA 氨酰化的突变 O-RS。在本发明的一方面，进行步骤 (c)(e) 至少两次。

[0026] 本发明的特征也在于产生 O-tRNA/O-RS 对的方法。在一个实施例中，如上所述获得 O-RS，且通过使第一物种的脊椎动物细胞群体（其中脊椎动物细胞包含 tRNA 文库的成员）经受阴性选择以去除包含经脊椎动物细胞的内源性氨酰基 tRNA 合成酶 (RS) 氨酰化的 tRNA 文库的成员的细胞来获得 O-tRNA。这提供与第一物种的脊椎动物细胞正交的 tRNA 池。在本发明的一方面，tRNA 的文库包含源自至少一种例如来自非脊椎动物生物体的 tRNA 的 tRNA。在本发明的另一方面，氨酰基 tRNA 合成酶 (RS) 的文库包含源自至少一种例如来自

非脊椎动物生物体的氨酰基 tRNA 合成酶 (RS) 的 RS。在本发明的另一方面,tRNA 的文库包含源自至少一种来自第一种非脊椎动物生物体的 tRNA 的 tRNA。氨酰基 tRNA 合成酶 (RS) 的文库视情况包含源自至少一种来自第二种非脊椎动物生物体的氨酰基 tRNA 合成酶 (RS) 的 RS。在一个实施例中,第一种和第二种非脊椎动物生物体相同。或者,第一种和第二种非脊椎动物生物体可不同。本发明的特征也在于由本发明的方法产生的特异性 O-tRNA/O-RS 对。

[0027] 本发明的另一特征为在一物种中产生翻译组件且在第二物种中引入所选择 / 筛选的翻译组件的方法。举例来说,在第一物种 (例如脊椎动物物种,诸如酵母等) 中产生 O-tRNA/O-RS 对的方法进一步包括在第二物种 (例如哺乳动物、昆虫、真菌、藻类、植物等) 的脊椎动物细胞中引入编码 O-tRNA 的核酸和编码 O-RS 的核酸。第二物种可使用所引入的翻译组件例如在翻译期间将非天然氨基酸在活体内并入正在生长的多肽链中。

[0028] 在另一实例中,产生在脊椎动物细胞中优先利用非天然氨基酸使正交 tRNA 氨酰化的正交氨酰基 tRNA 合成酶 (O-RS) 的方法包括 :(a) 在存在非天然氨基酸的情况下使第一物种 (例如脊椎动物物种,诸如酵母等) 的脊椎动物细胞群体经受阳性选择。第一物种的脊椎动物细胞各自包含 :i) 氨酰基 tRNA 合成酶 (RS) 的文库的成员,ii) 正交 tRNA(O-tRNA),iii) 编码阳性选择标记的多核苷酸,以及 iv) 编码阴性选择标记的多核苷酸。在阳性选择中存活的细胞包含在存在非天然氨基酸的情况下使正交 tRNA(O-tRNA) 氨酰化的活性 RS。使在阳性选择中存活的细胞在不存在非天然氨基酸的情况下经受阴性选择,以去除利用天然氨基酸使 O-tRNA 氨酰化的活性 RS,从而提供优先利用非天然氨基酸使 O-tRNA 氨酰化的 O-RS。将编码 O-tRNA 的核酸和编码 O-RS 的核酸引入第二物种 (例如哺乳动物、昆虫、真菌、藻类、植物等) 的脊椎动物细胞中。当在第二物种中翻译时,可使用这些组件将非天然氨基酸并入第二物种中的所关注的蛋白质或多肽中。在一个实施例中,将 O-tRNA 和 / 或 O-RS 引入第二物种的脊椎动物细胞中。

[0029] 在某些实施例中,通过使第一物种的脊椎动物细胞群体 (其中脊椎动物细胞包含 tRNA 文库的成员) 经受阴性选择以去除包含经脊椎动物细胞的内源性氨酰基 tRNA 合成酶 (RS) 氨酰化的 tRNA 文库的成员的细胞来获得 O-tRNA。这提供与第一物种和第二物种的脊椎动物细胞正交的 tRNA 池。

[0030] 具有至少一个非天然氨基酸的蛋白质 (或所关注的多肽) 也是本发明的特征。在本发明的某些实施例中,具有至少一个非天然氨基酸的蛋白质包括至少一个翻译后修饰。在一个实施例中,所述至少一个翻译后修饰包含通过 [3+2] 环加成使包含第二反应性基团的分子 (例如染料、例如聚乙二醇衍生物的聚合物、光交联剂、细胞毒性化合物、亲和标记、生物素衍生物、树脂、第二蛋白质或多肽、金属螯合剂、辅因子、脂肪酸、碳水化合物、多核苷酸 (例如 DNA、RNA 等) 等) 与至少一个包含第一反应性基团的非天然氨基酸连接。举例来说,第一反应性基团为炔基部分 (例如,在非天然氨基酸对炔丙氧基苯丙氨酸中)(所述基团有时也称为乙炔部分),且第二反应性基团为叠氮基部分。在另一实例中,第一反应性基团为叠氮基部分 (例如,在非天然氨基酸对叠氮基 -L- 苯丙氨酸中),且第二反应性基团为炔基部分。在某些实施例中,本发明的蛋白质包括至少一个包含至少一个翻译后修饰的非天然氨基酸 (例如酮基非天然氨基酸),其中所述至少一个翻译后修饰包含糖部分 (saccharide moiety)。在某些实施例中,翻译后修饰是在脊椎动物细胞中活体内进行。

[0031] 在某些实施例中，蛋白质包括至少一个由脊椎动物细胞在活体内进行的翻译后修饰，其中所述翻译后修饰并非由原核细胞进行。翻译后修饰的实例包括（但不限于）乙酰化、酰化、脂质修饰、棕榈酰化、棕榈酸盐加成、磷酸化、糖脂键修饰等。在一个实施例中，翻译后修饰包含通过 GlcNAc- 天冬酰胺键使寡糖与天冬酰胺连接（例如，其中寡糖包含 $(GlcNAc-Man)_2$-Man-GlcNAc-GlcNAc 等）。在另一实施例中，翻译后修饰包含通过 GalNAc- 丝氨酸、GalNAc- 苏氨酸、GlcNAc- 丝氨酸或 GlcNAc- 苏氨酸键使寡糖（例如 Gal-GalNAc、Gal-GlcNAc 等）与丝氨酸或苏氨酸连接。在某些实施例中，本发明的蛋白质或多肽可包含分泌或定位序列、抗原决定基标签、FLAG 标签、聚组氨酸标签、GST 融合物等。

[0032] 蛋白质通常与任何可利用的蛋白质（例如治疗蛋白、诊断蛋白、工业酶或其部分等）例如至少 60％、至少 70％、至少 75％、至少 80％、至少 90％、至少 95％或甚至至少 99％或 99％以上一致，且其包含一个或一个以上非天然氨基酸。在一个实施例中，本发明的组合物包括所关注的蛋白质或多肽和赋形剂（例如缓冲剂、医药学上可接受的赋形剂等）。

[0033] 所关注的蛋白质或多肽可含有至少一个、至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个或十个或十个以上非天然氨基酸。非天然氨基酸可相同或不同，例如在蛋白质中可存在 1、2、3、4、5、6、7、8、9、10 个或 10 个以上不同位点，其包含 1、2、3、4、5、6、7、8、9、10 个或 10 个以上不同的非天然氨基酸。在某些实施例中，天然存在的蛋白质型式中所存在的至少一个（但少于全部）特定氨基酸是经非天然氨基酸取代。

[0034] 例如，蛋白质（或所关注的多肽）的实例包括（但不限于）细胞因子、生长因子、生长因子受体、干扰素、白细胞介素、发炎性分子、致癌基因产物、肽激素、信号转导分子、类固醇激素受体、促红细胞生成素 (erythropoietin，EPO)、胰岛素、人类生长激素、α-1 抗胰蛋白酶、血管抑素、抗溶血因子、抗体、载脂蛋白 (Apolipoprotein)、载脂蛋白 (Apoprotein)、心房利尿钠因子、心房利尿钠多肽、心房肽、C-X-C 趋化因子、T39765、NAP-2、ENA-78、Gro-a、Gro-b、Gro-c、IP-10、GCP-2、NAP-4、SDF-1、PF4、MIG、降钙素、c-kit 配位体、CC 趋化因子、单核细胞趋化蛋白 -1、单核细胞趋化蛋白 -2、单核细胞趋化蛋白 -3、单核细胞发炎性蛋白 -1α、单核细胞发炎性蛋白 -1β、RANTES、I309、R83915、R91733、HCC1、T58847、D31065、T64262、CD40、CD40 配位体、胶原蛋白、菌落刺激因子 (Colony stimulating factor，CSF)、补体因子 5a、补体抑制剂、补体受体 1、DHFR、上皮中性粒细胞激活肽 -78、GROα/MGSA、GROβ、GROγ、MIP-1α、MIP-1δ、MCP-1、表皮生长因子 (Epidermal Growth Factor，EGF)、上皮中性粒细胞激活肽、剥脱性毒素、因子 IX、因子 VII、因子 VIII、因子 X、成纤维细胞生长因子 (Fibroblast Growth Factor，FGF)、纤维蛋白原、纤维粘连蛋白、G-CSF、GM-CSF、葡糖脑苷脂酶、促性腺激素、刺猬蛋白 (Hedgehog protein)、血红蛋白、肝细胞生长因子 (Hepatocyte Growth Factor，HGF)、水蛭素、人类血清白蛋白、ICAM-1、ICAM-1 受体、LFA-1、LFA-1 受体、胰岛素样生长因子 (Insulin-like Growth Factor，IGF)、IGF-I、IGF-II、干扰素、IFN-α、IFN-β、IFN-γ、IL-1、IL-2、IL-3、IL-4、IL-5、IL-6、IL-7、IL-8、IL-9、IL-10、IL-11、IL-12、角质细胞生长因子 (Keratinocyte Growth Factor，KGF)、乳铁蛋白 (Lactoferrin)、白血病抑制因子、荧光素酶、神经营养因子、中性粒细胞抑制因子 (Neutrophilinhibitory factor，NIF)、抑瘤素 M、成骨蛋白、甲状旁腺激素、PD-ECSF、PDGF、多效营养因子 (Pleiotropin)、蛋白质 A、蛋白质 G、致热外毒素 A、B 或 C、松弛素、肾素、SCF、

可溶性补体受体 I、可溶性 I-CAM 1、可溶性白细胞介素受体、可溶性 TNF 受体、生长调节素、生长抑素、生长激素、链激酶、超抗原、葡萄球菌肠毒素、SEA、SEB、SEC1、SEC2、SEC3、SED、SEE、类固醇激素受体、超氧化物歧化酶 (SOD)、中毒性休克综合症毒素、胸腺素 α 1、组织型纤溶酶原激活剂、肿瘤生长因子 (TGF)、TGF-α、TGF-β、肿瘤坏死因子、肿瘤坏死因子 α、肿瘤坏死因子 β、肿瘤坏死因子受体 (TNFR)、VLA-4 蛋白、VCAM-1 蛋白、血管内皮生长因子 (Vascular Endothelial Growth Factor,VEGEF)、尿激酶、Mos、Ras、Raf、Met、p53、Tat、Fos、Myc、Jun、Myb、Rel、雌激素受体、孕酮受体、睾酮受体、醛固酮受体、LDL 受体、SCF/c-Kit、CD40L/CD40、VLA-4/VCAM-1、ICAM-1/LFA-1、透明质酸 (hyalurin)/CD44、皮质酮、Genebank 或其它可利用数据库中所存在的蛋白质等,和 / 或其部分。在一个实施例中,所关注的多肽包括转录调节蛋白 (例如转录激活蛋白 (诸如 GAL4) 或转录阻遏蛋白等) 或其部分。

[0035] 本发明的脊椎动物细胞提供合成包含大量有用的非天然氨基酸的蛋白质的能力。举例来说,可在细胞提取物、缓冲剂、医药学上可接受的赋形剂等中产生蛋白质浓度为例如至少 10 微克 / 升、至少 50 微克 / 升、至少 75 微克 / 升、至少 100 微克 / 升、至少 200 微克 / 升、至少 250 微克 / 升或至少 500 微克 / 升或 500 微克 / 升以上的包含非天然氨基酸的蛋白质。在某些实施例中,本发明的组合物包括例如至少 10 微克、至少 50 微克、至少 75 微克、至少 100 微克、至少 200 微克、至少 250 微克或至少 500 微克或 500 微克以上的包含非天然氨基酸的蛋白质。

[0036] 在某些实施例中,所关注的蛋白质或多肽 (或其部分) 是由核酸编码。核酸通常包含至少一个选择密码子、至少两个选择密码子、至少三个选择密码子、至少四个选择密码子、至少五个选择密码子、至少六个选择密码子、至少七个选择密码子、至少八个选择密码子、至少九个选择密码子或甚至十个或十个以上选择密码子。

[0037] 本发明还提供在脊椎动物细胞中产生至少一种包含至少一个非天然氨基酸的蛋白质的方法 (以及由这些方法产生的蛋白质)。所述方法例如包括使包含包括至少一个选择密码子且编码蛋白质的核酸的脊椎动物细胞在适当培养基中生长。脊椎动物细胞还包含在细胞中起作用且识别选择密码子的正交 tRNA(O-tRNA) 和优先利用非天然氨基酸使 O-tRNA 氨酰化的正交氨酰基 tRNA 合成酶 (O-RS),且培养基包含非天然氨基酸。在一个实施例中,O-RS 利用非天然氨基酸使 O-tRNA 氨酰化,这与具有 (例如) 如 SEQ ID NO :86 或 45 中所述的氨基酸序列的 O-RS 例如至少 45%、至少 50%、至少 60%、至少 75%、至少 80%、至少 90%、至少 95%或甚至 99%或 99%以上有效。在另一实施例中,O-tRNA 包含 SEQ ID NO :64 或 65 或其互补多核苷酸序列,由 SEQ ID NO :64 或 65 或其互补多核苷酸序列加工而成或由 SEQ ID NO :64 或 65 或其互补多核苷酸序列编码。在又一实施例中,O-RS 包含如 SEQ ID NO :36-63 和 / 或 86 中任一序列所述的氨基酸序列。

[0038] 在一个实施例中,所述方法进一步包括在蛋白质中并入非天然氨基酸,其中所述非天然氨基酸包含第一反应性基团 ;和使所述蛋白质与包含第二反应性基团的分子 (例如染料、例如聚乙二醇衍生物的聚合物、光交联剂、细胞毒性化合物、亲和标记、生物素衍生物、树脂、第二蛋白质或多肽、金属螯合剂、辅因子、脂肪酸、碳水化合物、多核苷酸 (例如 DNA、RNA 等) 等) 接触。第一反应性基团与第二反应性基团反应以通过 [3+2] 环加成使分子与非天然氨基酸连接。在一个实施例中,第一反应性基团为炔基或叠氮基部分,且第二反应性基团为叠氮基或炔基部分。举例来说,第一反应性基团为炔基部分 (例如,在非天然氨

基酸对炔丙氧基苯丙氨酸中),且第二反应性基团为叠氮基部分。在另一实例中,第一反应性基团为叠氮基部分(例如,在非天然氨基酸对叠氮基-L-苯丙氨酸中),且第二反应性基团为炔基部分。

[0039] 在某些实施例中,编码蛋白包含治疗蛋白、诊断蛋白、工业酶或其部分。在一个实施例中,由所述方法产生的蛋白质由非天然氨基酸进一步修饰。举例来说,通过例如亲核-亲电子反应、[3+2]环加成等来修饰非天然氨基酸。在另一实施例中,通过至少一种翻译后修饰(例如N-糖基化、O-糖基化、乙酰化、酰化、脂质修饰、棕榈酰化、棕榈酸盐加成、磷酸化、糖脂键修饰等)在活体内修饰由所述方法产生的蛋白质。

[0040] 也提供产生筛选或选择转录调节蛋白的方法(以及由这些方法产生的筛选或选择转录调节蛋白)。所述方法例如包括选择第一多核苷酸序列,其中所述多核苷酸序列编码核酸结合域;和使第一多核苷酸序列突变以包括至少一个选择密码子。这提供了筛选或选择多核苷酸序列。例如,所述方法还包括选择第二多核苷酸序列,其中所述第二多核苷酸序列编码转录激活域;提供包含可操作性连接到第二多核苷酸序列的筛选或选择多核苷酸序列的构筑体;以及在细胞中引入构筑体、非天然氨基酸、正交tRNA合成酶(O-RS)和正交tRNA(O-tRNA)。利用这些组件,O-RS优先利用非天然氨基酸使O-tRNA氨酰化,且O-tRNA识别选择密码子且响应筛选或选择多核苷酸序列中的选择密码子将非天然氨基酸并入核酸结合域中。这提供了筛选或选择转录调节蛋白。

[0041] 在某些实施例中,本发明的组合物和方法包括脊椎动物细胞。本发明的脊椎动物细胞例如包括哺乳动物细胞、酵母细胞、真菌细胞、植物细胞、昆虫细胞等中的任一个。本发明的翻译组件可源自多种生物体,例如诸如原核生物体(例如大肠杆菌、嗜热脂肪芽孢杆菌等)或古细菌(archaebacterium)等非脊椎动物生物体;或例如脊椎动物生物体。

[0042] 本发明的选择密码子将扩充脊椎动物蛋白生物合成机构的遗传密码子构架。本发明中可使用多种选择密码子中的任一个,包括终止密码子(例如琥珀密码子、赭石密码子(ochre codon)或蛋白石终止密码子(opal stop codon))、无义密码子、稀有密码子、四个(或四个以上)碱基密码子等。

[0043] 可用于本文所述的组合物和方法中的非天然氨基酸的实例包括(但不限于):对乙酰基-L-苯丙氨酸、对碘-L-苯丙氨酸、O-甲基-L-酪氨酸、对炔丙氧基苯丙氨酸、对炔丙基-苯丙氨酸、L-3-(2-萘基)丙氨酸、3-甲基-苯丙氨酸、O-4-烯丙基-L-酪氨酸、4-丙基-L-酪氨酸、三-O-乙酰基-GlcNAcβ-丝氨酸、L-多巴(L-Dopa)、氟化苯丙氨酸、异丙基-L-苯丙氨酸、对叠氮基-L-苯丙氨酸、对酰基-L-苯丙氨酸、对苯甲酰基-L-苯丙氨酸、L-磷酸丝氨酸、膦酰基丝氨酸(phosphonoserine)、膦酰基酪氨酸(phosphonotyrosine)、对溴苯丙氨酸、对氨基-L-苯丙氨酸、异丙基-L-苯丙氨酸、酪氨酸氨基酸的非天然类似物;谷氨酰胺氨基酸的非天然类似物;苯丙氨酸氨基酸的非天然类似物;丝氨酸氨基酸的非天然类似物;苏氨酸氨基酸的非天然类似物;烷基、芳基、酰基、叠氮基、氰基、卤基、肼、酰肼、羟基、烯基、炔基、醚、硫醇、磺酰基、硒基、酯、硫代酸、硼酸酯(borate)、硼酸酯(boronate)、磷酸基(phospho)、膦酸基(phosphono)、膦、杂环、烯酮、亚胺、醛、羟胺、酮基或氨基取代的氨基酸,或其任何组合;具有光活化交联剂的氨基酸;自旋标记的氨基酸;荧光氨基酸;金属结合氨基酸;含金属氨基酸;放射性氨基酸;光笼化和/或光致异构化氨基酸(photocaged and/or photoisomerizable amino acid);含生物素或生物素类似物的氨

基酸；含酮基的氨基酸；包含聚乙二醇或聚醚的氨基酸；重原子取代的氨基酸；化学裂解或光裂解的氨基酸；具有伸长侧链的氨基酸；含毒性基团的氨基酸；糖取代的氨基酸；含碳连接的糖的氨基酸；氧化还原活性氨基酸；含α-羟基的酸；氨基硫代酸(amino thio acid)；α，α-二取代氨基酸；β-氨基酸；除脯氨酸或组氨酸以外的环状氨基酸，除苯丙氨酸、酪氨酸或色氨酸以外的芳香族氨基酸等。

[0044] 本发明还提供多肽(O-RS)和多核苷酸，例如O-tRNA、编码O-RS或其部分(例如合成酶的活性位点)的多核苷酸、用于构筑氨酰基tRNA合成酶突变体的寡核苷酸、编码所关注的蛋白质或多肽的包含一个或一个以上选择密码子的多核苷酸等。举例来说，本发明的多肽包括：包含如SEQ ID NO：36-63和/或86中任一序列所述的氨基酸序列的多肽；包含由如SEQ ID NO：3-35中任一序列所述的多核苷酸序列编码的氨基酸序列的多肽；以及与对包含如SEQ ID NO：36-63和/或86中任一序列所示的氨基酸序列的多肽或包含由如SEQ ID NO：3-35中任一序列所示的多核苷酸序列编码的氨基酸序列的多肽具有特异性的抗体特异性免疫反应的多肽。

[0045] 本发明的多肽中还包括包含与天然存在的酪氨酰基氨酰基tRNA合成酶(TyrRS)的氨基酸序列(例如SEQ ID NO：2)至少90%一致且包含两个或两个以上来自群组A-E(说明如上)的氨基酸的氨基酸序列的多肽。类似地，本发明的多肽还视情况包括包含SEQ ID NO：36-63和/或86中任一序列的至少20个相邻氨基酸以及两个或两个以上如上文在群组A-E中所指示的氨基酸取代的多肽。也包括包含任一上述多肽的保守变异体的氨基酸序列作为本发明的多肽。

[0046] 在一个实施例中，组合物包括本发明的多肽和赋形剂(例如缓冲剂、水、医药学上可接受的赋形剂等)。本发明还提供与本发明的多肽特异性免疫反应的抗体或抗血清。

[0047] 本发明中也提供多核苷酸。本发明的多核苷酸包括那些编码本发明的所关注的蛋白质或多肽的具有一个或一个以上选择密码子的多核苷酸。另外，本发明的多核苷酸例如包括：包含如SEQ ID NO：3-35、64-85中任一序列所述的核苷酸序列的多核苷酸；与其多核苷酸序列互补或编码其多核苷酸序列的多核苷酸；和/或编码包含 如SEQ ID NO：36-63和/或86中任一序列所述的氨基酸序列的多肽的多核苷酸，或其保守变异体。本发明的多核苷酸也包括编码本发明的多肽的多核苷酸。类似地，在核酸的实质上整个长度范围内在高度严格条件下与上文所指示的多核苷酸杂交的核酸为本发明的多核苷酸。

[0048] 本发明的多核苷酸也包括编码多肽的多核苷酸，所述多肽包含与天然存在的酪氨酰基氨酰基tRNA合成酶(TyrRS)的氨基酸序列(例如SEQ ID NO：2)至少90%一致且包含两个或两个以上如上文在群组A-E(说明如上)中所指示的突变的氨基酸序列。本发明的多核苷酸中还包括与上文所指示的多核苷酸至少70%(或至少75%、至少80%、至少85%、至少90%、至少95%、至少98%或至少99%或99%以上)一致的多核苷酸，和/或包含上文所指示的任何多核苷酸的保守变异体的多核苷酸。

[0049] 在某些实施例中，载体(例如质粒、粘粒(cosmid)、噬菌体、病毒等)包含本发明的多核苷酸。在一个实施例中，载体为表达载体。在另一实施例中，表达载体包括可操作性连接到本发明的一个或一个以上多核苷酸的启动子。在另一实施例中，细胞包含包括本发明的多核苷酸的载体。

[0050] 另一方面，本发明提供化合物的组合物和产生这些化合物的方法。举例来说，化合

物例如包括非天然氨基酸（诸如对（炔丙氧基）- 苯丙氨酸、叠氮基染料（诸如化学结构 4 和化学结构 6 中所示）、炔基聚乙二醇（例如，如化学结构 7 中所示），其中 n 为例如 50 与 10,000、75 与 5,000、100 与 2,000、100 与 1,000 之间的整数等。在本发明的实施例中，炔基聚乙二醇具有例如约 5,000 到约 100,000Da、约 20,000 到约 50,000Da、约 20,000 到约 10,000Da（例如，20,000Da）的分子量。

[0051]　还提供包含这些化合物，例如具有蛋白质和细胞的各种组合物。一方面，包括对（炔丙氧基）- 苯丙氨酸非天然氨基酸的组合物进一步包括正交 tRNA。非天然氨基酸可与正交 tRNA 键结（例如共价），例如通过氨基 - 酰基键与正交 tRNA 共价键结，与正交 tRNA 的末端核糖的 3′ OH 或 2′ OH 共价键结等。

[0052]　试剂盒也是本发明的特征。举例来说，提供在细胞中产生包含至少一个非天然氨基酸的蛋白质的试剂盒，其中所述试剂盒包括含有编码 O-tRNA 的多核苷酸序列或 O-tRNA 和编码 O-RS 的多核苷酸或 O-RS 的容器。在一个实施例中，试剂盒进一步包括至少一种非天然氨基酸。在另一实施例中，试剂盒进一步包含用于产生蛋白质的说明材料。

附图说明

[0053]　图 1 表示 tRNA 转录的流程 1。

[0054]　图 2 表示 tRNA 转录的流程 2。

[0055]　图 3 表示 hGH 的琥珀抑制和表达。

[0056]　图 4 表示 hGH 的增加的琥珀抑制和表达。

[0057]　图 5 表示 H1 启动子对 hGH 表达的影响。

[0058]　图 6 表示 tRNA 转录的流程 3。

[0059]　图 7 表示 U6 启动子对 hGH 表达的影响。

[0060]　图 8 表示利用来自嗜热脂肪芽孢杆菌的 tRNA 将亮氨酸并入多肽中的选择密码子的抑制。

具体实施方式

[0061]　在详细描述本发明之前，应理解，本发明不限于当然可变化的特定装置或生物系统。也应理解，本文中所使用的术语仅出于描述特定实施例的目的且不打算限制本发明。如本说明书和随附权利要求书中所使用，除非另外明确说明，否则单数形式“一”和“所述”包括多个指示物。因此，例如提及“一细胞”包括两个或两个以上细胞的组合；提及“细菌”包括细菌的混合物等。

[0062]　除非本文或本说明书的以下剩余部分中另有说明，否则本文中所使用的所有科技术语都具有与本发明所属领域的一般技术人员通常所理解的含义相同的含义。

[0063]　同源：当蛋白质和 / 或蛋白质序列是天然或人工地源自常见的祖蛋白质或蛋白质序列时，其是“同源”的。类似地，当核酸和 / 或核酸序列是天然或人工地源自常见的祖核酸或核酸序列时，其是“同源”的。举例来说，可通过任何可利用的诱变方法来修饰任何天然存在的核酸以使其包括一个或一个以上选择密码子。当这一经诱变核酸表达时，其编码包含一个或一个以上非天然氨基酸的多肽。当然，突变过程可另外改变一个或一个以上标准密码子，从而也改变所得突变蛋白质中的一个或一个以上标准氨基酸。同源性通常是从

两个或两个以上核酸或蛋白质（或其序列）之间的序列相似性推断得出。虽然适用于确定同源性的序列之间的相似性的精确百分比随所讨论的核酸和蛋白质而变化，但通常使用仅25％的序列相似性来确定同源性。也可使用例如30％、40％、50％、60％、70％、80％、90％、95％或99％或99％以上的较高水平的序列相似性来确定同源性。本文中描述用于测定序列相似性百分比的方法（例如使用默认参数的BLASTP和BLASTN），且通常可利用所述方法。

[0064] 正交：如本文中所使用，"正交"是指与细胞或翻译系统的内源性相应分子相比，以降低的效率利用细胞的内源性组件起作用；或无法利用细胞的内源性组件起作用的分子（例如正交tRNA(O-tRNA)和/或正交氨酰基tRNA合成酶(O-RS)）。在tRNA和氨酰基tRNA合成酶的情况下，正交是指与利用内源性tRNA合成酶起作用的内源性tRNA相比，正交tRNA无法利用内源性tRNA合成酶起作用或以降低的效率（例如，低于20％的效率、低于10％的效率、低于5％的效率或低于1％的效率）利用内源性tRNA合成酶起作用；或者与利用内源性tRNA起作用的内源性tRNA合成酶相比，正交氨酰基tRNA合成酶无法利用内源性tRNA起作用或以降低的效率（例如，低于20％的效率、低于10％的效率、低于5％的效率或低于1％的效率）利用内源性tRNA起作用。正交分子在细胞内缺乏功能内源性互补分子。举例来说，当与通过内源性RS使内源性tRNA氨酰化相比时，细胞中的任何内源性RS以降低的效率或甚至为零的效率使所述细胞中的正交tRNA氨酰化。在另一实例中，当与通过内源性RS使内源性tRNA氨酰化相比时，正交RS以降低的效率或甚至为零的效率使所关注的细胞中的任何内源性tRNA氨酰化。可在细胞内引入第二正交分子，从而与第一正交分子一起作用。举例来说，正交tRNA/RS对包括所引入的在细胞中相对于相应的内源性tRNA/RS对以一定效率（例如，50％效率、60％效率、70％效率、75％效率、80％效率、90％效率、95％效率或99％或99％以上效率）共同起作用的互补组件。

[0065] 互补：术语"互补"是指可共同起作用的正交对O-tRNA和O-RS（例如其中O-RS使O-tRNA氨酰化）的组件。

[0066] 优先氨酰化：术语"优先氨酰化"是指与O-RS使天然存在的tRNA或用于生成O-tRNA的起始物质氨酰化相比，O-RS利用非天然氨基酸使O-tRNA氨酰化的效率，例如70％有效、75％有效、85％有效、90％有效、95％有效或99％或99％以上有效。将非天然氨基酸以高保真度并入正在生长的多肽链中，例如对给定选择密码子来说在75％以上的效率下，对给定选择密码子来说在约80％以上的效率下，对给定选择密码子来说在约90％以上的效率下，对给定选择密码子来说在约95％以上的效率下或对给定选择密码子来说在约99％或99％以上的效率下。

[0067] 选择密码子：术语"选择密码子"是指在翻译过程中由O-tRNA识别且不被内源性tRNA识别的密码子。O-tRNA反密码子环识别mRNA上的选择密码子且在这一位点处在多肽中并入其氨基酸，例如非天然氨基酸。例如，选择密码子可包括诸如终止密码子（例如琥珀密码子、赭石密码子和蛋白石密码子）等无义密码子；四个或四个以上碱基的密码子；稀有密码子；源自天然或非天然碱基对的密码子等。

[0068] 抑制因子tRNA：抑制因子tRNA为例如通过提供响应选择密码子在多肽链中并入氨基酸的机构来改变给定翻译系统中信使RNA(mRNA)的阅读的tRNA。举例来说，抑制因子tRNA可通读（例如）终止密码子、四个碱基的密码子、稀有密码子等。

[0069] 可重复利用的tRNA：术语"可重复利用的tRNA"是指在翻译期间经氨酰化且可重

复利用氨基酸(例如非天然氨基酸)再氨酰化以将所述氨基酸(例如所述非天然氨基酸)并入一个或一个以上多肽链中的 tRNA。

[0070] 翻译系统:术语"翻译系统"是指将天然存在的氨基酸并入正在生长的多肽链(蛋白质)中的组件的集合。翻译系统的组件可例如包括核糖体、tRNA、合成酶、mRNA、氨基酸等。本发明的组件(例如 ORS、OtRNA、非天然氨基酸等)可添加到活体外或活体内翻译系统中,所述翻译系统例如脊椎动物细胞,例如酵母细胞、哺乳动物细胞、植物细胞、藻类细胞、真菌细胞、昆虫细胞等。

[0071] 非天然氨基酸:如本文中所使用,术语"非天然氨基酸"是指不为 20 种常见的天然存在的氨基酸、硒代半胱氨酸或吡咯赖氨酸中的一种的任何氨基酸、经修饰氨基酸和/或氨基酸类似物。

[0072] 源自:如本文中所使用,术语"源自"是指从指定分子或生物体分离或使用来自指定分子或生物体的信息制成的组件。

[0073] 无活性 RS:如本文中所使用,术语"无活性 RS"是指已突变以致不能再利用氨基酸使天然的同源 tRNA 氨酰化的合成酶。

[0074] 阳性选择或筛选标记:如本文中所使用,术语"阳性选择或筛选标记"是指当存在(例如,经表达、激活等)时导致鉴别具有阳性选择标记的细胞与不具有阳性选择标记的细胞的标记。

[0075] 阴性选择或筛选标记:如本文中所使用,术语"阴性选择或筛选标记"是指当存在(例如,经表达、激活等)时允许鉴别不具有所需性质的细胞(例如与具有所需性质的细胞相比)的标记。

[0076] 报告基因 (reporter):如本文中所使用,术语"报告基因"是指可用于选择所关注的系统的标靶组件的组件。举例来说,报告基因可包括荧光筛选标记(例如绿色荧光蛋白)、发光标记(例如荧火虫荧光素酶蛋白)、基于亲和力的筛选标记或诸如 his3、ura3、leu2、lys2、lacZ、β-gal/lacZ(β-半乳糖苷酶)、Adh(醇脱氢酶)等的可选择标记基因。

[0077] 脊椎动物:如本文中所使用,术语"脊椎动物"是指属于系统发育域 (phylogenetic domain) 真核域 (Eucarya) 的生物体,诸如动物,例如哺乳动物、爬行动物、鸟类等。

[0078] 非真核生物:如本文中所使用,术语"非真核生物"是指非脊椎动物生物体。举例来说,非脊椎动物生物体可属于真细菌 (Eubacteria)(例如大肠杆菌、嗜热栖热菌 (Thermus thermophilus)、嗜热脂肪芽孢杆菌 (Bacillus stearothermophilus) 等)系统发育域或古细菌 (Archaea)(例如詹氏甲烷球菌、产甲烷杆菌 (Methanobacteriumthermoautotrophicum)、诸如沃氏嗜盐富饶菌 (Haloferax volcanii) 和嗜盐菌属 NRC-1(Halobacterium species NRC-1) 等嗜盐菌、超嗜热古菌 (Archaeoglobus fulgidus)、强烈炽热球菌 (Pyrococcus furiosus)、极端嗜热古菌 (Pyrococcus horikoshii)、嗜热泉生古细菌 (Aeuropyrum pernix) 等)系统发育域。

[0079] 抗体:如本文中所使用,术语"抗体"包括(但不限于)特异性结合并识别分析物(抗原)的实质上由一个或一个以上免疫球蛋白基因编码的多肽或其片段。实例包括多克隆抗体、单克隆抗体、嵌合抗体及单链抗体等。如本文中所使用的术语"抗体"中也包括免疫球蛋白的片段,包括 Fab 片段和由包括噬菌体呈现的表达文库产生的片段。关于抗体结构和术语学,例如参见 Paul,Fundamental Immunology,第 4 版,1999,Raven Press,New York。

[0080] 保守变异体:术语“保守变异体”是指在功能上与得到保守变异体的组件(例如O-tRNA或O-RS)一样表现但序列发生变异的翻译组件,例如保守变异体O-tRNA或保守变异体O-RS。举例来说,虽然O-tRNA和保守变异体O-tRNA不具有相同的序列,但O-RS将利用非天然氨基酸使互补O-tRNA或保守变异体O-tRNA氨酰化。例如,保守变异体的序列可具有一处变异、两处变异、三处变异、四处变异或五处或五处以上变异,只要保守变异体与相应O-tRNA或O-RS互补即可。

[0081] 选择或筛选剂:如本文中所使用,术语“选择或筛选剂”是指当存在时允许从群体中选择/筛选某些组件的试剂。举例来说,选择或筛选剂例如包括(但不限于)营养物、抗生素、光波长、抗体、经表达多核苷酸(例如转录调节蛋白)等。可例如通过浓度、强度等来改变选择剂。

[0082] 可检测物质:如本文中所使用,术语“可检测物质”是指当激活、改变、表达等时允许从群体中选择/筛选某些组件的试剂。举例来说,可检测物质可为化学剂,例如5-氟乳清酸(5-FOA),其在某些情况下(例如表达URA3报告基因时)变得可检测(例如变为杀死表达URA3报告基因的细胞的有毒产物)。

[0083] 突破遗传密码所施加的化学限制,直接在脊椎动物细胞中遗传修饰蛋白质结构的能力将提供探测和操纵细胞过程的强有力的分子工具。本发明提供扩充脊椎动物细胞中遗传编码的氨基酸的数目的翻译组件。这些组件包括tRNA(例如正交tRNA(O-tRNA))、氨酰基tRNA合成酶(例如正交合成酶(O-RS))、O-tRNA/O-RS对以及非天然氨基酸。

[0084] 本发明的O-tRNA通常有效地经表达和加工且在脊椎动物细胞中在翻译中起作用,但不被宿主的氨酰基tRNA合成酶有效氨酰化。本发明的O-tRNA响应选择密码子在mRNA翻译期间将不编码常见的20种氨基酸中的任一种的非天然氨基酸传递到正在生长的多肽链中。

[0085] 本发明的O-RS在脊椎动物细胞中优先利用非天然氨基酸使本发明的O-tRNA氨酰化,但不会使任何细胞质中宿主的tRNA氨酰化。此外,本发明的氨酰基tRNA合成酶的特异性使得能接受非天然氨基酸,同时排除任何内源性氨基酸。包括例示性O-RS的氨基酸序列或其部分的多肽也是本发明的特征。另外,编码翻译组件、O-tRNA、O-RS和其部分的多核苷酸是本发明的特征。

[0086] 本发明还提供利用用于脊椎动物细胞中的非天然氨基酸产生例如O-RS和/或正交对(正交tRNA和正交氨酰基tRNA合成酶)的所需翻译组件的方法(和由这些方法产生的翻译组件)。举例来说,来自大肠杆菌的酪氨酰基tRNA合成酶/$tRNA_{CUA}$对为本发明的O-tRNA/O-RS对。另外,本发明的特征也在于在一种脊椎动物细胞中选择/筛选翻译组件,且当选择/筛选后将这些组件用于另一脊椎动物细胞(未用于选择/筛选的脊椎动物细胞)中的方法。举例来说,可在酵母(例如酿酒酵母(Saccharomyces cerevisiae))中进行产生用于脊椎动物细胞的翻译组件的选择/筛选方法,且然后可将这些所选择的组件用于另一脊椎动物细胞中,例如另一酵母细胞、哺乳动物细胞、昆虫细胞、植物细胞、真菌细胞等。

[0087] 本发明进一步提供在脊椎动物细胞中产生蛋白质的方法,其中所述蛋白质包含非天然氨基酸。所述蛋白质是使用本发明的翻译组件产生。本发明也提供包括非天然氨基酸的蛋白质(和由本发明的方法产生的蛋白质)。所关注的蛋白质或多肽也可包括翻译后修

饰，其例如是通过 [3+2] 环加成或亲核－亲电子反应来添加，并非由原核细胞进行等。在某些实施例中，产生具有非天然氨基酸的转录调节蛋白的方法（和由这些方法产生的蛋白质）也包括在本发明内。包括包含非天然氨基酸的蛋白质的组合物也是本发明的特征。

[0088]　产生具有非天然氨基酸的蛋白质或多肽的试剂盒也是本发明的特征。

[0089]　正交氨酰基 tRNA 合成酶 (O-RS)

[0090]　为了在脊椎动物细胞中在所关注的蛋白质或多肽中特异性并入非天然氨基酸，改变合成酶的底物特异性，以致仅将所需的非天然氨基酸而不是常见的 20 种氨基酸中的任一种装入 tRNA 中。如果正交合成酶是混杂的，那么将产生在标靶位置处具有天然和非天然氨基酸的混合物的突变蛋白。本发明提供对特异性非天然氨基酸具有经修饰的底物特异性的正交氨酰基 tRNA 合成酶的组合物和其产生方法。

[0091]　包括正交氨酰基 tRNA 合成酶 (O-RS) 的脊椎动物细胞是本发明的特征。O-RS 在脊椎动物细胞中优先利用非天然氨基酸使正交 tRNA(O-tRNA) 氨酰化。在某些实施例中，O-RS 利用一个以上非天然氨基酸，例如两个或两个以上、三个或三个以上等非天然氨基酸。因此，本发明的 O-RS 可具有优先利用不同非天然氨基酸使 O-tRNA 氨酰化的能力。这允许通过选择将何种非天然氨基酸或非天然氨基酸的组合与细胞放在一起和／或通过选择不同量的非天然氨基酸与细胞放在一起以供并入来达成另一控制水平。

[0092]　与天然氨基酸相比，本发明的 O-RS 视情况具有一种或一种以上针对非天然氨基酸的改进或增强的酶促性质。与天然存在的氨基酸（例如，20 种已知的常见氨基酸中的一种）相比，针对非天然氨基酸的这些性质包括例如较高 Km、较低 Km、较高 kcat、较低 kcat、较低 kcat/km、较高 kcat/km 等。

[0093]　可视情况通过包括 O-RS 的多肽和／或通过编码 O-RS 或其部分的多核苷酸将 O-RS 提供给脊椎动物细胞。举例来说，O-RS 或其部分是由如 SEQ ID NO :3-35 中任一序列所述的多核苷酸序列或其互补多核苷酸序列编码。在另一实例中，O-RS 包含如 SEQ ID NO :36-63 和／或 86 中任一序列所述的氨基酸序列或其保守变异体。关于例示性 O-RS 分子的序列，例如参见本文中的表 5、表 6 和表 8 以及实例 6。

[0094]　O-RS 也可包含与天然存在的酪氨酰基氨酰基 tRNA 合成酶 (TyrRS) 的氨基酸序列（例如，如 SEQ ID NO :2 中所述）例如至少 90％、至少 95％、至少 98％、至少 99％或甚至至少 99.5％一致且包含来自群组 A-E 的两个或两个以上氨基酸的氨基酸序列。群组 A 包括在对应于大肠杆菌 TyrRS 的 Tyr37 的位置处的缬氨酸、异亮氨酸、亮氨酸、甘氨酸、丝氨酸、丙氨酸或苏氨酸；群组 B 包括在对应于大肠杆菌 TyrRS 的 Asn126 的位置处的天冬氨酸；群组 C 包括在对应于大肠杆菌 TyrRS 的 Asp182 的位置处的苏氨酸、丝氨酸、精氨酸、天冬酰胺或甘氨酸；群组 D 包括在对应于大肠杆菌 TyrRS 的 Phe183 的位置处的蛋氨酸、丙氨酸、缬氨酸或酪氨酸；且群组 E 包括在对应于大肠杆菌 TyrRS 的 Leu186 的位置处的丝氨酸、蛋氨酸、缬氨酸、半胱氨酸、苏氨酸或丙氨酸。也例如参见本文中的表 4、表 6 和表 8。

[0095]　除 O-RS 外，本发明的脊椎动物细胞还可包括其它组件，例如非天然氨基酸。脊椎动物细胞还包括正交 tRNA(O-tRNA)（例如源自诸如大肠杆菌、嗜热脂肪芽孢杆菌等的非脊椎动物生物体），其中 O-tRNA 识别选择密码子且通过 O-RS 优先利用非天然氨基酸使 O-tRNA 氨酰化。细胞中也可存在包含编码所关注的多肽的多核苷酸的核酸，其中多核苷酸包含由 O-tRNA 识别的选择密码子，或一个或一个以上所述选择密码子的组合。

[0096] 一方面,O-tRNA以包含如SEQ ID NO:65中所述的多核苷酸序列或由如SEQID NO:65中所述的多核苷酸序列加工而成的tRNA的效率的例如至少45%、至少50%、至少60%、至少75%、至少80%、至少90%、至少95%或99%或以上,介导将非天然氨基酸并入蛋白质中。另一方面,O-tRNA包含SEQ ID NO:65,且O-RS包含SEQ ID NO:36-63和/或86中任一序列所述的多肽序列和/或其保守变异体。关于例示性O-RS和O-tRNA分子的序列,也例如参见本文中的表5和实例6。

[0097] 在一个实例中,脊椎动物细胞包含正交氨酰基tRNA合成酶(O-RS)、正交tRNA(O-tRNA)、非天然氨基酸和包含编码所关注的多肽的多核苷酸的核酸,所述多核苷酸包含由O-tRNA识别的选择密码子。在脊椎动物细胞中,O-RS优先利用非天然氨基酸使正交tRNA(O-tRNA)氨酰化,且所述细胞在不存在非天然氨基酸的情况下以一定产率产生所关注的多肽,所述产率例如为在存在非天然氨基酸的情况下多肽产率的不到30%、不到20%、不到15%、不到10%、不到5%、不到2.5%等。

[0098] 产生O-RS的方法(其为本发明的特征)视情况包括:由野生型合成酶构架产生突变合成酶的池,且然后根据相对于常见的20种氨基酸,对非天然氨基酸的特异性选择突变RS。为分离这种合成酶,选择方法:(i)灵敏,因为由最初几轮得到的所需合成酶的活性可能较低且群体较小;(ii)"可调",因为希望在各轮选择下改变选择严格度;以及(iii)通用,以使所述方法可用于不同非天然氨基酸。

[0099] 在脊椎动物细胞中产生优先利用非天然氨基酸使正交tRNA氨酰化的正交氨酰基tRNA合成酶(O-RS)的方法通常包括应用阳性选择、接着阴性选择的组合。在阳性选择中,在阳性标记的非必需位置处引入的选择密码子的抑制使得脊椎动物细胞在阳性选择压力下存活。因此,在存在非天然氨基酸的情况下,存活细胞编码在正交抑制因子tRNA中装入非天然氨基酸的活性合成酶。在阴性选择中,阴性标记的非必需位置处所引入的选择密码子的抑制将去除具有天然氨基酸特异性的合成酶。阴性和阳性选择的存活细胞编码只能(或至少优先)利用非天然氨基酸使正交抑制因子tRNA氨酰化(在正交抑制因子tRNA中装入非天然氨基酸)的合成酶。

[0100] 举例来说,所述方法包括:(a)在存在非天然氨基酸的情况下使第一物种的脊椎动物细胞群体经受阳性选择,其中所述脊椎动物细胞各自包含:i)氨酰基tRNA合成酶(RS)文库的成员,ii)正交tRNA(O-tRNA),iii)编码阳性选择标记的多核苷酸以及iv)编码阴性选择标记的多核苷酸;其中在阳性选择中存活的细胞包含在存在非天然氨基酸的情况下使正交tRNA(O-tRNA)氨酰化的活性RS;和(b)在不存在非天然氨基酸的情况下使在阳性选择中存活的细胞经受阴性选择,以去除利用天然氨基酸使O-tRNA氨酰化的活性RS,从而提供利用非天然氨基酸优先使O-tRNA氨酰化的O-RS。

[0101] 阳性选择标记可为多种分子中的任一种。在一个实施例中,阳性选择标记是提供供生长使用的营养增补剂的产品,且所述选择是在缺乏所述营养增补剂的培养基上进行。编码阳性选择标记的多核苷酸的实例例如包括(但不限于)基于补充细胞的氨基酸营养缺陷体的报告基因、his3基因(例如,其中通过提供3-氨基三唑(3-AT)所检测,his3基因编码咪唑磷酸甘油酯脱水酶)、ura3基因、leu2基因、lys2基因、lacZ基因、adh基因等。例如参见G.M.Kishore和D.M.Shah,(1988),Amino acidbiosynthesis inhibitors as herbicides,<u>Annual Review of Biochemistry</u> 57:627-663。在一个实施例中,通过邻硝

基苯基-β-D-吡喃半乳糖苷(ONPG)水解检测 lacZ 的产生。例如参见 I.G.Serebriiskii 和 E.A.Golemis,(2000),Uses of lacZto study gene function:evaluation of beta-galactosidase assays employed in the yeast two-hybrid system,Analytical Biochemistry 285:1-15。其它阳性选择标记例如包括荧光素酶、绿色荧光蛋白(GFP)、YFP、EGFP、RFP、抗生素抗性基因的产物(例如氯霉素乙酰转移酶(CAT))、转录调节蛋白(例如 GAL4)等。编码阳性选择标记的多核苷酸视情况包含选择密码子。

[0102] 可将编码阳性选择标记的多核苷酸可操作性连接到反应元件。也可存在编码调节从反应元件进行的转录的转录调节蛋白且包含至少一个选择密码子的另一多核苷酸。通过经非天然氨基酸氨酰化的 O-tRNA 将非天然氨基酸并入转录调节蛋白中会引起编码阳性选择标记的多核苷酸(例如报告基因)的转录。选择密码子视情况位于编码转录调节蛋白的 DNA 结合域的多核苷酸的一部分中,或实质上靠近所述部分。

[0103] 也可将编码阴性选择标记的多核苷酸可操作性连接到反应元件,其中从反应元件进行的转录是通过转录调节蛋白来介导。例如参见 A.J.DeMaggio 等人,(2000),The yeast split-hybrid system,Method Enzymol.328:128-137;H.M.Shih 等人,(1996),A positive genetic selection for disrupting protein-protein interactions:identificationofCREB mutations that prevent association with the coactivator CBP,Proc.Natl.Acad.Sci.U.S.A.93:13896-13901;M.Vidal 等人,(1996),Genetic characterization of amammalian protein-protein interaction domain by using a yeast reverse two-hybridsystem,[comment],Proc.Natl.Acad.Sci.U.S.A.93:10321-10326 以及 M.Vidal 等人,(1996),Reverse two-hybrid and one-hybrid systems to detect dissociation ofprotein-protein and DNA-protein interactions.[comment]Proc.Natl.Acad.Sci.U.S.A.93:10315-10320。通过经天然氨基酸氨酰化的 O-tRNA 将天然氨基酸并入转录调节蛋白中会引起阴性选择标记的转录。阴性选择标记视情况包含选择密码子。在一个实施例中,本发明的阳性选择标记和/或阴性选择标记可包含至少两个选择密码子,所述选择密码子各自或都可包含至少两个不同选择密码子或至少两个相同选择密码子。

[0104] 转录调节蛋白是结合(直接或间接)核酸序列(例如反应元件)且调节可操作性连接到反应元件的序列的转录的分子。转录调节蛋白可为转录激活蛋白(例如 GAL4、核激素受体、API、CREB、LEF/tcf 家族成员、SMAD、VP16、SP1 等)、转录阻遏蛋白(例如核激素受体、Groucho/tle 家族、Engrailed 家族等)或可视环境而具有这两种活性的蛋白质(例如 LEF/tcf、同源异型盒(homobox)蛋白等)。反应元件通常为由转录调节蛋白识别的核酸序列或与转录调节蛋白协力作用的另一试剂。

[0105] 转录调节蛋白的另一实例为转录激活蛋白,GAL4。例如参见 A.Laughon 等人,(1984),Identification of two proteins encoded by the Saccharomyces cerevisiae GAL4gene,Molecular&Cellular Biology 4:268-275;A.Laughon 和 R.F.Gesteland,(1984),Primary structure of the Saccharomyces cerevisiae GAL4 gene,Molecular&CellularBiology 4:260-267;L.Keegan 等人,(1986),Separation of DNA binding from thetranscription-activating function of a vertebrate regulatory protein,Science231:699-704;以及 M.Ptashne,(1988),How vertebrate

transcriptional activators work, Nature 335:683-689。这一 881 个氨基酸的蛋白质的 N- 末端 147 个氨基酸形成特异性结合 DNA 序列的 DNA 结合域 (DBD)。例如参见 M.Carey 等人,(1989),Anamino-terminal fragment of GAL4 binds DNA as a dimer, J.Mol.Biol.209:423-432；和 E.Giniger 等人,(1985),Specific DNA binding of GAL4,a positive regulatory protein ofyeast,Cell 40:767-774。DBD 通过插入蛋白质序列连接到当与 DNA 结合时可激活转录的 C- 末端 113 个氨基酸的激活域 (AD)。例如参见 J.Ma 和 M.Ptashne,(1987),Deletion analysis of GAL4 defines two transcriptional activating segments,Cell48:847-853；和 J.Ma 和 M.Ptashne,(1987),The carboxy-terminal 30 amino acids ofGAL4 are recognized by GAL80,Cell 50:137-142。通过例如朝向含有 GAL4 的 N- 末端 DBD 和其 C- 末端 AD 的单一多肽的 N- 末端 DBD 放置琥珀密码子,可将通过 O-tRNA/O-RS 对执行的琥珀抑制连接到通过 GAL4 执行的转录激活。GAL4 激活的报告基因可用以利用基因执行阳性和阴性选择。

[0106] 用于阴性选择的培养基可包含通过阴性选择标记转化为可检测物质的选择或筛选剂。在本发明的一方面,可检测物质为有毒物质。编码阴性选择标记的多核苷酸例如可为 ura3 基因。举例来说,可将 URA3 报告基因放于含有 GAL4 DNA 结合位点的启动子的控制下。当例如通过翻译编码具有选择密码子的 GAL4 的多核苷酸产生阴性选择标记时,GAL4 激活 URA3 的转录。阴性选择是在包含 5- 氟乳清酸 (5-FOA) 的培养基上实现,所述 5- 氟乳清酸通过 ura3 基因的基因产物转化为可检测物质(例如,杀死细胞的有毒物质)。例如参见 J.D.Boeke 等人,(1984),A positiveselection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast:5-fluoroorotic acid resistance,Molecular&General Genetics 197:345-346)；M.Vidal 等人,(1996),Genetic characterization of a mammalian protein-protein interactiondomain by using a yeast reverse two-hybrid system.[comment],Proc.Natl.Acad.Sci.U.S.A.93:10321-10326；以及 M.Vidal 等人,(1996),Reverse two-hybrid and one-hybridsystems to detect dissociation of protein-protein and DNA-protein interactions,[comment],Proc.Natl.Acad.Sci.U.S.A.93:10315-10320。

[0107] 与阳性选择标记一样,阴性选择标记也可为多种分子中的任一种。在一个实施例中,阳性选择标记和/或阴性选择标记为发荧光或在存在合适反应物的情况下催化发光反应的多肽。举例来说,阴性选择标记例如包括(但不限于)荧光素酶、绿色荧光蛋白 (GFP)、YFP、EGFP、RFP、抗生素抗性基因的产物(例如氯霉素乙酰转移酶 (CAT))、lacZ 基因的产物、转录调节蛋白等。在本发明的一方面,通过荧光激活细胞分类 (FACS) 或发光性来检测阳性选择标记和/或阴性选择标记。在另一实例中,阳性选择标记和/或阴性选择标记包含基于亲和力的筛选标记。同一多核苷酸可编码阳性选择标记和阴性选择标记。

[0108] 本发明的方法中也可使用其它水平的选择/筛选严格度。选择或筛选严格度可在产生 O-RS 的方法的一个或两个步骤中改变。此可例如包括改变编码阳性和/或阴性选择标记的多核苷酸中的反应元件的量,向一个或两个步骤中添加不同量的无活性 合成酶,改变所使用的选择/筛选剂的量等。也可进行另外数轮阳性和/或阴性选择。

[0109] 选择或筛选也可包含一种或一种以上阳性或阴性选择/筛选,所述选择/筛选例如包括氨基酸渗透性改变、翻译效率改变、翻译保真性改变等。所述一种或一种以上变化通

常是基于包含或编码用于产生蛋白质的正交 tRNA-tRNA 合成酶对的组件的一个或一个以上多核苷酸的突变。

[0110] 也可使用模型富集研究从过量无活性合成酶中快速选择活性合成酶。可进行阳性和/或阴性模型选择研究。举例来说,将包含潜在活性氨酰基 tRNA 合成酶的脊椎动物细胞与不同倍数过量的无活性氨酰基 tRNA 合成酶混合。在非选择性培养基中生长且通过例如 X-GAL 覆盖 (X-GAL overlay) 检定的细胞与在选择培养基(例如,在不存在组氨酸和/或尿嘧啶的情况下)中生长且能够在其中存活且通过例如 X-GAL 检定 (X-GAL assay) 所检定的细胞之间进行比率比较。对于阴性模型选择,将潜在活性氨酰基 tRNA 合成酶与不同倍数过量的无活性氨酰基 tRNA 合成酶混合,且用例如 5-FOA 的阴性选择物质进行选择。

[0111] RS 的文库(例如突变 RS 的文库)通常包含源自至少一种例如来自非脊椎动物生物体的氨酰基 tRNA 合成酶 (RS) 的 RS。在一个实施例中,RS 的文库是源自无活性 RS,例如其中无活性 RS 例如在合成酶的活性位点处、在合成酶的编辑机制位点处、在通过组合合成酶的不同域所得的不同位点处等通过使活性 RS 突变而产生。举例来说,使 RS 的活性位点处的残基突变为例如丙氨酸残基。使用编码丙氨酸突变 RS 的多核苷酸作为将丙氨酸残基诱变到所有 20 种氨基酸中的模板。选择/筛选突变 RS 的文库来产生 O-RS。在另一实施例中,无活性 RS 包含氨基酸结合袋,且包含结合袋的一个或一个以上氨基酸经一个或一个以上不同氨基酸取代。在一个实例中,经取代的氨基酸是经丙氨酸取代。视情况使用编码丙氨酸突变 RS 的多核苷酸作为将丙氨酸残基诱变到所有 20 种氨基酸中的模板,并对其加以筛选/选择。

[0112] 产生 O-RS 的方法可进一步包括通过使用所属领域中已知的各种诱变技术产生 RS 文库。举例来说,可通过位点特异性突变、随机点突变、同源重组、DNA 改组或其它递归诱变方法、嵌合构筑或其任何组合产生突变 RS。举例来说,突变 RS 的文库可由两个或两个以上其它(例如)较小、较少多样性的“子文库”产生。一旦使合成酶经受阳性和阴性选择/筛选策略,然后就可使这些合成酶经受进一步诱变。举例来说,可分离编码 O-RS 的核酸;可由核酸生成一组编码突变 O-RS 的多核苷酸(例如,通过随机诱变、位点特异性诱变、重组或其任何组合);且可重复这些个别步骤或这些步骤的组合,直到获得优先利用非天然氨基酸使 O-tRNA 氨酰化的突 变 O-RS。在本发明的一方面,进行所述步骤至少两次。

[0113] 关于产生 O-RS 的其它细节可见于标题为“Methods and compositions for theproduction of orthogonal tRN A-aminoacyltRNA synthetase pairs” 的 WO 2002/086075 中。也参见 Hamano-Takaku 等人,(2000)A mutant Escherichia coli Tyrosyl-tRNASynthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently thanTyrosine, Journal of Biological Chemistry. 275(51):40324-40328;Kiga 等人,(2002),Anengineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of anunnatural amino acid into proteins in vertebrate translation and its application in awheat germ cell-free system, PNAS 99(15):9715-9723;以及 Francklyn 等人,(2002),Aminoacyl-tRNA synthetases:Versatile players in the changing theater of translation:RNA. 8:1363-1372。

[0114] 正交 tRNA

[0115] 本发明提供包括正交 tRNA(O-tRNA) 的真核细胞。正交 tRNA 介导在活体内将非天

然氨基酸并入由包含由 O-tRNA 识别的选择密码子的多核苷酸编码的蛋白质中。在某些实施例中，本发明的 O-tRNA 以包含如 SEQ ID NO：65 中所述的多核苷酸序列或在细胞中由如 SEQ ID NO：65 中所述的多核苷酸序列加工而成的 tRNA 的效率的例如至少 40％、至少 45％、至少 50％、至少 60％、至少 75％、至少 80％或甚至 90％或 90％以上，介导将非天然氨基酸并入蛋白质中。参见本文中的表 5。

[0116] 本发明的 O-tRNA 的实例为 SEQ ID NO：65。(参见本文中的实例 6 和表 5)。SEQID NO：65 是在细胞中视情况例如使用标准内源性细胞剪接和加工机构加工且经修饰以形成活性 O-tRNA 的剪接前 / 加工前转录物。这些剪接前转录物的群体通常形成细胞中的活性 tRNA 群体。本发明还包括 O-tRNA 和其经加工的细胞产物的保守变异体。举例来说，O-tRNA 的保守变异体包括那些像 SEQ ID NO：65 的 O-tRNA 一样作用且在加工形式中维持 tRNA 的 L 形结构，但不具有相同序列的分子 (且不为野生型 tRNA 分子)。本发明的 O-tRNA 通常为可重复利用的 O-tRNA，这是因为 O-tRNA 可在活体内再氨酰化，从而再次介导响应选择密码子将非天然氨基酸并入由多核苷酸编码的蛋白质中。

[0117] 真核生物而非原核生物中的 tRNA 的转录是通过 RNA 聚合酶 III 进行，这对脊椎动物细胞中可转录的 tRNA 结构基因的一级序列产生限制。另外，在脊椎动物细胞中，tRNA 需要从转录所述 tRNA 的核中输出到细胞质中，以进行翻译。编码本发明的 O-tRNA 或其互补多核苷酸的核酸也是本发明的特征。在本发明的一方面，编 码本发明的 O-tRNA 的核酸包括内部启动子序列，例如 A 盒 (例如 TRGCNNAGY) 和 B 盒 (例如 GGTTCGANTCC，SEQ ID NO：88)。本发明的 O-tRNA 也可经转录后修饰。举例来说，真核生物中 tRNA 基因的转录后修饰包括分别通过 Rnase P 和 3′ - 内切核酸酶去除 5′ - 侧接序列和 3′ - 侧接序列。添加 3′ -CCA 序列也是真核生物中 tRNA 基因的转录后修饰。

[0118] 在一个实施例中，通过使第一物种的脊椎动物细胞的群体经受阴性选择获得 O-tRNA，其中脊椎动物细胞包含 tRNA 文库的成员。阴性选择去除包含经脊椎动物细胞的内源性氨酰基 tRNA 合成酶 (RS) 氨酰化的 tRNA 文库的成员的细胞。这提供与第一物种的脊椎动物细胞正交的 tRNA 池。

[0119] 可替代使用或与上述将非天然氨基酸并入多肽中的其它方法组合使用反式翻译系统。这一系统包括存在于大肠杆菌中的称为 tmRNA 的分子。这一 RNA 分子在结构上与丙氨酰基 tRNA 相关且经丙氨酰合成酶氨酰化。tmRNA 与 tRNA 之间的不同之处在于，反密码子环是由特定的大序列置换。这一序列允许核糖体使用 tmRNA 内编码的开放阅读框架作为模板在已停止的序列上重新开始翻译。在本发明中，可产生优先经正交合成酶氨酰化且加载非天然氨基酸的正交 tmRNA。通过由所述系统转录基因，核糖体停在特定位点处；在这一位点处引入非天然氨基酸，且使用正交 tmRNA 内编码的序列重新开始翻译。

[0120] 关于产生重组正交 tRNA 的其它方法例如可见于标题为“Methods andcompositions for the production of orthogonal tRNA-aminoacyltRNA synthetase pairs”的国际专利申请案 WO 2002/086075 中。也参见 Forster 等人，(2003) Programmingpeptidomimetic synthetases by translating genetic codes designed de novo PNAS 100(11)：6353-6357；和 Feng 等人，(2003)，Expanding tRNA recognition of a tRNAsynthetase by a single amino acid change，PNAS 100(10)：5676-5681。

[0121] 正交 TRNA 和正交氨酰基 tRNA 合成酶对

[0122] 正交对包含例如抑制因子 tRNA、移码 tRNA 等的 O-tRNA 和 O-RS。O-tRNA 未被内源性合成酶酰化且能够在活体内介导将非天然氨基酸并入由包含由 O-tRNA 识别的选择密码子的多核苷酸编码的蛋白质中。在脊椎动物细胞中，O-RS 识别 O-tRNA 且优先利用非天然氨基酸使 O-tRNA 氨酰化。本发明中包括产生正交对的方法和由所述方法产生的正交对，和用于脊椎动物细胞中的正交对的组合物。多种正交 tRNA/ 合成酶对的研发可允许在脊椎动物细胞中使用不同密码子同时合并多个非天然氨基酸。

[0123] 在脊椎动物细胞中，可通过利用无效跨物种氨酰化输入来自不同生物体的对（例如无义抑制因子对）来产生正交 O-tRNA/O-RS。在脊椎动物细胞中有效表达和加工 O-tRNA 和 O-RS，且将 O-tRNA 有效地从核中输出到细胞质中。举例来说，一个这样的对为来自大肠杆菌的酪氨酰基 tRNA 合成酶 /tRNACUA 对（例如参见 H. M. Goodman 等人，(1968)，Nature 217 :1019-24 ;和 D. G. Barker 等人，(1982)，FEBS Letters150 :419-23)。当在酿酒酵母的细胞质中表达大肠杆菌酪氨酰基 tRNA 合成酶和其同源大肠杆菌 $tRNA_{CUA}$时，大肠杆菌酪氨酰基tRNA合成酶有效地使其同源大肠杆菌 $tRNA_{CUA}$氨酰化，但不会使酿酒酵母tRNA氨酰化。例如参见 H. Edwards 和 P. Schimmel，(1990)，Molecular&Cellular Biology 10 :1633-41 ;和 H. Edwards 等人，(1991)，PNAS United States of America 88 :1153-6。另外，大肠杆菌酪氨酰基 $tRNA_{CUA}$是酿酒酵母氨酰基 tRNA 合成酶的不良底物（例如参见 V. Trezeguet 等人，(1991)，Molecular&Cellular Biology 11 :2744-51)，但其在酿酒酵母的蛋白质翻译中有效地起作用。例如参见 H. Edwards 和 P. Schimmel，(1990)Molecular&Cellular Biology10 :1633-41 ;H. Edwards 等人，(1991). PNAS United States of America 88 :1153-6 ;以及 V. Trezeguet 等人，(1991)，Molecular&Cellular Biology 11 :2744-51。此外，大肠杆菌 TyrRS 不具有校正与 tRNA 连接的非天然氨基酸的编辑机构。

[0124] O-tRNA 和 O-RS 可天然存在，或可通过使来自多种生物体的天然存在的 tRNA 和 / 或 RS 突变而产生，所述突变产生 tRNA 文库和 / 或 RS 文库。参见本文中的标题为“来源和宿主”的部分。在各种实施例中，O-tRNA 和 O-RS 是源自至少一种生物体。在另一实施例中，O-tRNA 是源自来自第一种生物体的天然存在或突变的天然存在的 tRNA，且 O-RS 是源自来自第二种生物体的天然存在或突变的天然存在的 RS。在一个实施例中，第一种和第二种非脊椎动物生物体相同。或者，第一种和第二种非脊椎动物生物体可不同。

[0125] 关于产生 O-RS 和 O-tRNA 的方法，参见本文中的标题为“正交氨酰基 tRNA 合成酶”和“O-tRNA”的部分。也参见标题为“Methods and compositions for theproduction of orthogonal tRNA-aminoacyltRNA synthetase pairs”的国际专利申请案 WO 2002/086075。

[0126] 保真性、效率和产率

[0127] 保真性是指将所需分子（例如非天然氨基酸或氨基酸）在所需位置处并入正在生长的多肽中的准确性。本发明的翻译组件响应选择密码子以高保真性将非天然氨基酸并入蛋白质中。举例来说，与在正在生长的多肽链中在所需位置处不合需要地 并入特定天然氨基酸相比，使用本发明的组件将所需非天然氨基酸在所需位置处并入正在生长的多肽链中的效率（例如，响应选择密码子）例如为超过 75％、超过 85％、超过 95％或甚至超过 99％或 99％以上有效。

[0128] 效率也可指与相关对照相比，O-RS 利用非天然氨基酸使 O-tRNA 氨酰化的程度。

本发明的O-RS可通过其效率来定义。在本发明的某些实施例中,比较某种O-RS与另一种O-RS。举例来说,本发明的O-RS利用非天然氨基酸使O-tRNA氨酰化,这与具有(例如)如SEQ ID NO:86或45中所述的氨基酸序列的O-RS或表5中的另一特定RS使O-tRNA氨酰化例如至少40%、至少50%、至少60%、至少75%、至少80%、至少90%、至少95%或甚至99%或99%以上有效。在另一实施例中,本发明的O-RS利用非天然氨基酸使O-tRNA氨酰化,这比O-RS利用天然氨基酸使O-tRNA氨酰化有效至少10倍、至少20倍、至少30倍等。

[0129] 使用本发明的翻译组件,包含非天然氨基酸的所关注的多肽的产率为从多核苷酸缺乏选择密码子的细胞获得所关注的天然存在的多肽的产率的例如至少5%、至少10%、至少20%、至少30%、至少40%、50%或50%以上。另一方面,细胞在不存在非天然氨基酸的情况下以一定产率产生所关注的多肽,所述产率例如为在存在非天然氨基酸的情况下多肽产率的不到30%、不到20%、不到15%、不到10%、不到5%、不到2.5%等。

[0130] 来源和宿主生物体

[0131] 本发明的正交翻译组件通常源自用于脊椎动物细胞或翻译系统中的非脊椎动物生物体。举例来说,正交O-tRNA可源自非脊椎动物生物体,例如真细菌,诸如大肠杆菌、嗜热栖热菌、嗜热脂肪芽孢杆菌等;或古细菌,诸如詹氏甲烷球菌、产甲烷杆菌、诸如沃氏嗜盐富饶菌和嗜盐菌属NRC-1等嗜盐菌、超嗜热古菌、强烈炽热球菌、极端嗜热古菌、嗜热泉生古细菌等,而正交O-RS也可源自非脊椎动物生物体,例如真细菌,诸如大肠杆菌、嗜热栖热菌、嗜热脂肪芽孢杆菌等;或古细菌,诸如詹氏甲烷球菌、产甲烷杆菌、诸如沃氏嗜盐富饶菌和嗜盐菌属NRC-1等嗜盐菌、超嗜热古菌、强烈炽热球菌、极端嗜热古菌、嗜热泉生古细菌等。或者,也可使用脊椎动物来源,例如植物、藻类、原生生物、真菌、酵母、动物(例如哺乳动物、昆虫、节肢动物等)等,例如其中所述组件与所关注的细胞或翻译系统正交,或其中其经修饰(例如突变)而与细胞或翻译系统正交。

[0132] O-tRNA/O-RS对的个别组件可源自相同生物体或不同生物体。在一个实施例中,O-tRNA/O-RS对是来自相同生物体。举例来说,O-tRNA/O-RS对可源自来自大 肠杆菌的酪氨酰基tRNA合成酶/$tRNA_{CUA}$对。或者,O-tRNA/O-RS对的O-tRNA和O-RS视情况来自不同生物体。

[0133] 可选择或筛选正交O-tRNA、O-RS或O-tRNA/O-RS对和/或在脊椎动物细胞将其用于产生具有非天然氨基酸的多肽。脊椎动物细胞可来自多种来源,例如任何脊椎动物(例如哺乳动物、两栖动物、鸟类、爬行动物、鱼类等)等。具有本发明的翻译组件的脊椎动物细胞的组合物也是本发明的特征。

[0134] 本发明也提供在一种物种中的有效筛选,从而任选用于所述物种和/或第二物种中(视情况无其它选择/筛选)。举例来说,在一种物种(例如容易操纵的物种(诸如酵母细胞等))中选择或筛选O-tRNA/O-RS的组件且将其引入例如植物(例如复杂植物,诸如单子叶植物或双子叶植物)、藻类、原生生物、真菌、酵母、动物(例如哺乳动物、昆虫、节肢动物等)等的第二脊椎动物物种中,以用于在第二物种中活体内并入非天然氨基酸。

[0135] 举例来说,可选择酿酒酵母(Saccharomyces cerevisiae,S.cerevisiae)作为第一脊椎动物物种,这是因为其为单细胞物种,具有快速的世代时间和相对明确表征的遗传学。例如参见D.Burke等人,(2000)Methods in Yeast Genetics.Cold SpringHarbor Laboratory Press,Cold Spring Harbor,NY。此外,因为真核生物的

翻译机构高度保守(例如参见(1996)Translational Control.Cold Spring Harbor Laboratory,ColdSpring Harbor,NY;Y.Kwok 和 J.T.Wong,(1980),Evolutionary relationship betweenHalobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNAsynthetases as phylogenetic probes,Canadian Journal of Biochemistry 58:213-218;以及(2001)The Ribosome.Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY),所以可将酿酒酵母中发现的用于并入非天然氨基酸的aaRS基因引入高级脊椎动物生物体中且与同源tRNA合作(例如参见K.Sakamoto等人,(2002)Site-specificincorporation of an unnatural amino acid into proteins in mammalian cells,NucleicAcids Res.30:4692-4699;和C.Kohrer等人,(2001),Import of amber and ochresuppressor tRNA′s into mammalian cells:a general approach to site-specific insertion ofamino acid analogues into proteins,Proc.Natl.Acad.Sci.U.S.A.98:14310-14315)用于并入非天然氨基酸。

[0136] 在一个实例中,在如本文中所述的在第一物种中产生O-tRNA/O-RS的方法进一步包括在第二物种(例如哺乳动物、昆虫、真菌、藻类、植物等)的脊椎动物细胞中引入编码O-tRNA的核酸和编码O-RS的核酸。在另一实例中,产生在脊椎动物 细胞中优先利用非天然氨基酸使正交tRNA氨酰化的正交氨酰基tRNA合成酶(O-RS)的方法包括:(a)在存在非天然氨基酸的情况下使第一物种(例如酵母等)的脊椎动物细胞群体经受阳性选择。脊椎动物细胞各自包含:i)氨酰基tRNA合成酶(RS)文库的成员,ii)正交tRNA(O-tRNA),iii)编码阳性选择标记的多核苷酸,以及iv)编码阴性选择标记的多核苷酸。在阳性选择中存活的细胞包含在存在非天然氨基酸的情况下使正交tRNA(O-tRNA)氨酰化的活性RS。使在阳性选择中存活的细胞在不存在非天然氨基酸的情况下经受阴性选择,以去除利用天然氨基酸使O-tRNA氨酰化的活性RS。这提供优先利用非天然氨基酸使O-tRNA氨酰化的O-RS。将编码O-tRNA的核酸和编码O-RS的核酸(或组件O-tRNA和/或O-RS)引入第二物种(例如哺乳动物、昆虫、真菌、藻类、植物等)的脊椎动物细胞中。通常,通过使第一物种的脊椎动物细胞的群体经受阴性选择来获得O-tRNA,其中脊椎动物细胞包含tRNA文库的成员。阴性选择去除包含经脊椎动物细胞的内源性氨酰基tRNA合成酶(RS)氨酰化的tRNA文库的成员的细胞,此提供与第一物种和第二物种的脊椎动物细胞正交的tRNA池。

[0137] 选择密码子

[0138] 本发明的选择密码子将扩充蛋白质生物合成机构的遗传密码子构架。举例来说,选择密码子例如包括独特的三个碱基密码子、无义密码子(诸如终止密码子,例如琥珀密码子(UAG)、蛋白石密码子(UGA))、非天然密码子、至少四个碱基的密码子、稀有密码子等。可将多个选择密码子引入所需基因中,例如一个或一个以上、两个或两个以上、三个以上等。基因可包括给定选择密码子的多个复本,或可包括多个不同的选择密码子或其任何组合。

[0139] 在一个实施例中,所述方法包括使用作为终止密码子的选择密码子,以在脊椎动物细胞中活体内并入非天然氨基酸。举例来说,产生识别例如UAG的终止密码子的O-tRNA且通过O-RS利用所需的非天然氨基酸使O-tRNA氨酰化。天然存在的宿主氨酰基tRNA合成酶无法识别这一O-tRNA。可使用常规定点诱变在所关注的多肽中的所关注的位点处引入例如TAG的终止密码子。例如参见Sayers,J.R.等人,(1988),5′,3′ Exonuclease in

phosphorothioate-based oligonucleotide-directedmutagenesis. Nucleic Acids Res, 791-802。当活体内组合 O-RS、O-tRNA 以及编码所关注的多肽的核酸时，响应 UAG 密码子并入非天然氨基酸以产生在指定位置处含有非天然氨基酸的多肽。

[0140] 活体内并入非天然氨基酸可在不显著干扰脊椎动物宿主细胞的情况下进行。举例来说，因为 UAG 密码子的抑制效率取决于 O-tRNA（例如琥珀抑制因子 tRNA）与脊椎动物释放因子（例如 eRF）（其与终止密码子结合且起始正在生长的肽从核糖体中的释放）之间的竞争，所以抑制效率可例如通过增加 O-tRNA（例如抑制因子 tRNA）的表达水平来调节。

[0141] 选择密码子还包含延长密码子，例如四个或四个以上碱基的密码子（诸如四个、五个、六个或六个以上碱基的密码子）。四个碱基的密码子的实例例如包括 AGGA、CUAG、UAGA、CCCU 等。五个碱基的密码子的实例例如包括 AGGAC、CCCCU、CCCUC、CUAGA、CUACU、UAGGC 等。本发明的特征包括使用基于移码抑制（frameshift suppression）的延长密码子。四个或四个以上碱基的密码子例如可将一个或一个以上非天然氨基酸插入同一蛋白质中。举例来说，在存在具有反密码子环（例如，具有至少 8-10nt 反密码子环）的突变 O-tRNA（例如特定移码抑制因子 tRNA）的情况下，将四个或四个以上碱基的密码子读作单个氨基酸。在其它实施例中，反密码子环可例如解码至少一个四个碱基的密码子、至少一个五个碱基的密码子或至少一个六个碱基的密码子或更多碱基的密码子。因为存在 256 种可能的四个碱基的密码子，所以可使用四个或四个以上碱基的密码子在同一细胞中编码多个非天然氨基酸。参见 Anderson 等人，(2002)Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology, 9:237-244;Magliery, (2001) Expanding the Genetic Code:Selection of Efficient Suppressors of Four-base Codons and Identification of "Shifty" Four-base Codons with a Library Approach in Escherichia coli, J. Mol. Biol. 307:755-769。

[0142] 举例来说，已经使用活体外生物合成方法利用四个碱基的密码子将非天然氨基酸并入蛋白质中。例如参见 Ma 等人，(1993)Biochemistry. 32:7939;和 Hohsaka 等人，(1999) J. Am. Chem. Soc, 121:34。使用 CGGG 和 AGGU，从而在活体外利用两个化学酰化的移码抑制因子 tRNA 同时将 2-萘基丙氨酸和赖氨酸的 NBD 衍生物并入抗生蛋白链菌素中。例如参见 Hohsaka 等人，(1999)J. Am. Chem. Soc. 121:12194。在活体内研究中，Moore 等人检查具有 NCUA 反密码子的 tRNALeu 衍生物抑制 UAGN 密码子（N 可为 U、A、G 或 C）的能力且发现四联体 UAGA 可通过具有 UCUA 反密码子的 tRNALeu 以 13％到 26％的效率解码，其中在 0 或 -1 框架中极少解码。参见 Moore 等人，(2000)J. Mol. Biol., 298:195。在一个实施例中，可在本发明中使用基于稀有密码子或无义密码子的延长密码子，其可减少其它不合需要的位点处的错义通读和移码抑制。

[0143] 对于给定系统来说，选择密码子也可包括一个天然的三个碱基的密码子，其中内源性系统不使用（或很少使用）天然碱基密码子。举例来说，这包括缺乏识别天然的三个碱基的密码子的 tRNA 的系统和／或三个碱基的密码子为稀有密码子的系统。

[0144] 选择密码子视情况包括非天然碱基对。这些非天然碱基对进一步扩充现有的遗传代码（genetic alphabet）。一个额外的碱基对使三联体密码子的数目从 64 增加到 125。第三碱基对的性质包括稳定且具选择性的碱基配对、通过聚合酶以高保真性有效地酶促并入 DNA 中以及在合成新的非天然碱基对之后有效的持续引物延长。可适于方

法和组合物的非天然碱基对的描述例如包括 Hirao 等人，(2002)An unnaturalbase pair for incorporating amino acid analogues into protein，Nature Biotechnology，20：177-182。其它相关公开文献列于下表中。

[0145] 对于活体内使用来说，非天然核苷是膜可渗透的且可磷酸化以形成相应的三磷酸盐。另外，增加的遗传信息是稳定的且不受细胞酶破坏。Benner 和其它人先前的工作利用不同于典型 Watson-Crick 配对的氢键模式，其中最值得注意的实例是 iso-C：iso-G 配对。例如参见 Switzer 等人，(1989)J.Am.Chem.Soc，111：8322；和 Piccirilli 等人，(1990)Nature，343：33；Kool，(2000)Curr.Opin.Chem.Biol.4：602。这些碱基通常在一定程度上与天然碱基错配且不能酶促复制。Kool 和其同事证明碱基之间的疏水性堆积相互作用(hydrophobic packing interaction)可代替氢键来驱动碱基对的形成。参见 Kool，(2000)Curr.Ooin.Chem.Biol.4：602；和 Guckian 和 Kool，(1998)Aneew.Chem.Int.Ed.Engl.36，2825。在研发满足所有上述要求的非天然碱基对的工作中，Schultz、Romesberg 和其同事系统地合成和研究了一系列非天然疏水性碱基。发现 PICS：PICS 自身配对比天然碱基对稳定且可通过大肠杆菌 DNA 聚合酶 I 的克列诺片段(Klenow fragment)(KF) 有效地并入 DNA 中。例如参见 McMinn 等人，(1999)J.Am.Chem.Soc，121：11586；和 Ogawa 等人，(2000)J.Am.Chem.Soc.122：3274。可通过 KF 以足以用于生物功能的效率和选择性合成 3MN：3MN 自身配对。例如参见 Ogawa 等人，(2000)J.Am.Chem.Soc.122：8803。然而，这两种碱基都充当用于进一步复制的链终止子。突变 DNA 聚合酶近来已得到发展，其可用于复制 PICS 自身配对。另外，也可复制 7AI 自身配对。例如参见 Tae 等人，(2001)J.Am.Chem.Soc.123：7439。也已研发新颖的金属碱基对 Dipic：Py，其在结合 Cu(II) 后形成稳定配对。参见 Meggers 等人，(2000)J.Am.Chem.Soc.122：10714。因为延长密码子和非天然密码子固有地与天然密码子正交，所以本发明的方法可利用这一性质产生正交 tRNA 供其使用。

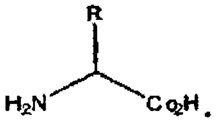
[0146] 也可使用翻译旁路系统(translational bypassing system)在所需多肽中并入非天然氨基酸。在翻译旁路系统中，将大序列并入基因中，但并未将其翻译成蛋白质。所述序列含有充当诱导核糖体跳过所述序列且在插入下游重新开始翻译的提示的结构。

[0147] 非天然氨基酸

[0148] 如本文中所使用，非天然氨基酸是指除硒代半胱氨酸和/或吡咯赖氨酸和下列 20 种遗传编码的 α-氨基酸以外的任何氨基酸、经修饰氨基酸或氨基酸类似物：丙氨酸、精氨酸、天冬酰胺、天冬氨酸、半胱氨酸、谷氨酰胺、谷氨酸、甘氨酸、组氨酸、异亮氨酸、亮氨酸、赖氨酸、蛋氨酸、苯丙氨酸、脯氨酸、丝氨酸、苏氨酸、色氨酸、酪氨酸、缬氨酸。α-氨基酸的一般结构由式 I 说明：

[0149]

I



。

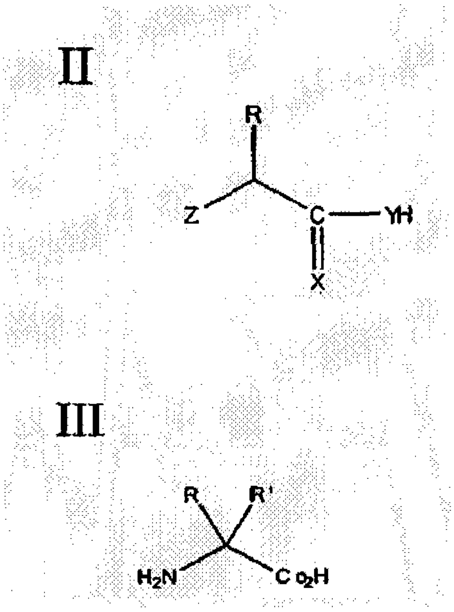
[0150] 非天然氨基酸通常为任何具有式 I 的结构，其中 R 基团是除 20 种天然氨基酸中所用的取代基以外的任何取代基。关于 20 种天然氨基酸的结构，例如参见 L.Stryer，Biochemistry，第 3 版，1988，Freeman and Company，New York。需注意的是，本发明的非天

然氨基酸可为除以上 20 种 α-氨基酸以外的天然存在的化合物。

[0151] 因为本发明的非天然氨基酸不同于天然氨基酸之处通常在于侧链，所以非天然氨基酸以与天然存在的蛋白质中形成酰胺键相同的方式与其它氨基酸（例如天然或非天然）形成酰胺键。然而，非天然氨基酸具有使其区别于天然氨基酸的侧链基团。举例来说，式 I 中的 R 视情况包含烷基、芳基、酰基、酮基、叠氮基、羟基、肼、氰基、卤基、酰肼、烯基、炔基、醚、硫醇、硒基、磺酰基、硼酸酯 (borate)、硼酸酯 (boronate)、磷酸基、膦酸基、膦、杂环、烯酮、亚胺、醛、酯、硫代酸、羟胺、胺等或其任何组合。所关注的其它非天然氨基酸包括（但不限于）包含光活化交联剂的氨基酸、经自旋标记的氨基酸、荧光氨基酸、金属结合氨基酸、含金属氨基酸、放射性氨基酸、具有新颖官能团的氨基酸、与其它分子共价或非共价相互作用的氨基酸、光笼化和/或光致异构化氨基酸、含生物素或生物素类似物的氨基酸、含酮基的氨基酸、包含聚乙二醇或聚醚的氨基酸、重原子取代的氨基酸、化学裂解或光裂解的氨基酸、与天然氨基酸相比具有伸长侧链（例如聚醚或长链烃，例如约 5 个碳以上、约 10 个碳以上等）的氨基酸、含碳连接的糖的氨基酸、氧化还原活性氨基酸、含氨基硫代酸的氨基酸以及含一个或一个以上毒性部分的氨基酸。在一些实施例中，非天然氨基酸具有例如用于使蛋白质与固体支撑物连接的光活化交联剂。在一个实施例中，非天然氨基酸具有连接到氨基酸侧链的糖部分（例如糖基化氨基酸）和/或其它碳水化合物修饰。

[0152] 除含有新颖侧链的非天然氨基酸外，非天然氨基酸还视情况包含经修饰的主链结构，例如，如式 II 和式 III 的结构所说明：

[0153]



，

[0154] 其中 Z 通常包含 OH、NH_2、SH、NH-R′ 或 S-R′；X 和 Y 可相同或不同，通常包含 S 或 O；且 R 和 R′ 视情况相同或不同，通常选自与上文关于具有式 I 的非天然氨基酸所述的 R 基团的组分相同的清单以及氢。举例来说，本发明的非天然氨基酸视情况包含如式 II 和 III 所说明的氨基或羧基的取代。这种类型的非天然氨基酸包括（但不限于）例如具有对应于常见的 20 种天然氨基酸的侧链或非天然侧链的 α-羟基酸、α-硫代酸、α-氨基硫代羧酸酯。另外，α-碳的取代视情况包括 L、D 或 α-α-二取代氨基酸，诸如 D-谷氨酸、D-丙氨酸、D-甲基-O-酪氨酸、氨基丁酸等。其它结构替代物包括环状氨基酸，诸如脯氨

酸类似物以及 3、4、6、7、8 和 9 元环脯氨酸类似物；β 和 γ 氨基酸，诸如经取代 β-丙氨酸和 γ-氨基丁酸。举例来说，多种非天然氨基酸是基于诸如酪氨酸、谷氨酰胺、苯丙氨酸等的天然氨基酸。酪氨酸类似 物包括对位取代的酪氨酸、邻位取代的酪氨酸以及间位取代的酪氨酸，其中经取代酪氨酸例如包含酮基（例如乙酰基）、苯甲酰基、氨基、肼、羟胺、硫醇基、羧基、异丙基、甲基、C_6-C_{20}直链或支链烃、饱和或不饱和烃、O-甲基、聚醚基、硝基、炔基等。另外，也涵盖多取代芳基环。本发明的谷氨酰胺类似物包括（但不限于）α-羟基衍生物、γ 取代衍生物、环状衍生物以及酰胺取代的谷氨酰胺衍生物。例示性苯丙氨酸类似物包括（但不限于）对位取代的苯丙氨酸、邻位取代的苯丙氨酸以及间位取代的苯丙氨酸，其中取代基例如包含羟基、甲氧基、甲基、烯丙基、醛、叠氮基、碘基、溴基、酮基（例如乙酰基）、苯甲酰基、炔基等。非天然氨基酸的特定实例包括（但不限于）对乙酰基-L-苯丙氨酸、对炔丙氧基苯丙氨酸、O-甲基-L-酪氨酸、L-3-(2-萘基）丙氨酸、3-甲基-苯丙氨酸、O-4-烯丙基-L-酪氨酸、4-丙基-L-酪氨酸、三-O-乙酰基-GlcNAcβ-丝氨酸、L-多巴、氟化苯丙氨酸、异丙基-L-苯丙氨酸、对叠氮基-L-苯丙氨酸、对酰基-L-苯丙氨酸、对苯甲酰基-L-苯丙氨酸、L-磷酸丝氨酸、膦酰基丝氨酸、膦酰基酪氨酸、对碘-苯丙氨酸、对溴苯丙氨酸、对氨基-L-苯丙氨酸以及异丙基-L-苯丙氨酸等。例如，多种非天然氨基酸的其它结构是提供于标题为“In vivo incorporation of unnatural amino acids”的 WO 2002/085923 的图 16、17、18、19、26 以及 29 中。关于其它蛋氨酸类似物，也参见 Kiick 等人，(2002)Incorporation ofazides into recombinant proteins for chemoselective modification by theStaudinger ligtation，PNAS 99:19-24 的图 1 结构 2-5。

[0155] 在一个实施例中，提供包括非天然氨基酸（诸如对（炔丙氧基）-苯丙氨酸）的组合物。还提供各种包含对（炔丙氧基）-苯丙氨酸和例如蛋白质和/或细胞的组合物。一方面，包括对（炔丙氧基）-苯丙氨酸非天然氨基酸的组合物进一步包括正交 tRNA。非天然氨基酸可与正交 tRNA 键结（例如共价），例如通过氨基-酰基键与正交 tRNA 共价键结，与正交 tRNA 的末端核糖的 3′ OH 或 2′ OH 共价键结等。

[0156] 可借助于非天然氨基酸并入蛋白质中的化学部分提供蛋白质的多种优点和操纵性。举例来说，酮基官能团的独特反应性允许利用许多含肼或羟胺试剂中的任一个在活体外和活体内选择性修饰蛋白质。举例来说，重原子非天然氨基酸可适用于定相 x 射线结构数据。使用非天然氨基酸位点特异性引入重原子还为选择重原子的位置提供选择性和灵活性。举例来说，光反应性非天然氨基酸（例如，具有二苯甲酮和叠氮芳基（例如叠氮苯）侧链的氨基酸）允许蛋白质的活体内和活体外有效光交联。光反应性非天然氨基酸的实例例如包括（但不限于）对叠氮基-苯丙氨酸和对苯甲酰基-苯丙氨酸。然后，可通过激发光反应性基团（提供时间（和/或空间）控制）使具有光反应性非天然氨基酸的蛋白质随意交联。在一个实例中，非天然氨基酸的甲基可经作为局部结构和动力学的探针（例如，使用核磁共振和振动光谱学）的同位素标记的（例如）甲基取代。举例来说，炔基或叠氮基官能团允许利用分子通过 [3+2] 环加成反应选择性修饰蛋白质。

[0157] 非天然氨基酸的化学合成

[0158] 以上提供的多种非天然氨基酸可例如从 Sigma(USA) 或 Aldrich(Milwaukee，WI，USA) 购得。无法购得的非天然氨基酸是视情况如本文中所提供或如各种公开文献中所提供或使用所属领域技术人员已知的标准方法合成。关于有机合成技术，例

如参见 Fessendon 和 Fessendon，Organic Chemistry(1982，第 2 版，Willard Grant Press，Boston Mass.)；March，Advanced Organic Chemistry(第 3 版，1985，Wiley and Sons，New York)；以及 Carey 和 Sundberg，Advanced Organic Chemistry(第 3 版，A 和 B 部分，1990，Plenum Press，New York)。描述非天然氨基酸的合成的其它公开文献例如包括：名称为“In vivo incorporation of Unnatural Amino Acids”的 WO 2002/085923；Matsoukas 等人，(1995)J. Med. Chem. 38，4660-4669；King，F.E. 和 Kidd，D.A.A.(1949)A New Synthesis of Glutamine and of γ-Dipeptides of Glutamic Acid from PhthylatedIntermediates. J. Chem. Soc. 3315-3319；Friedman，O.M. 和 Chatterrji，R(1959)Synthesis of Derivatives of Glutamine as Model Substrates for Anti-Tumor Agents. J. Am. Chem. Soc. 81，3750-3752；Craig，J.C. 等人，(1988)Absolute Configuration of theEnantiomers of 7-Chloro-4[[4-(diethylamino)-1-methylbutyl]amino]quinoline(Chloroquine). J. Org. Chem. 53，1167-1170；Azoulay，M.，Vilmont，M. 和 Frappier，F.(1991)Glutamine analogues asPotential Antimalarials，. Eur. J. Med. Chem. 26，201-5；Koskinen，A.M.P. 和 Rapoport，H.(1989)Synthesis of 4-Substituted Prolines as Conformationally Constrained Amino AcidAnalogues. J. Org. Chem. 54，1859-1866；Christie，B.D. 和 Rapoport，H.(1985)Synthesisof Optically Pure Pipecolates from L-Asparagine. Application to the Total synthesis of(+)-Apovincamine through Amino Acid Decarbonylation and Iminium Ion Cyclization. J. Org. Chem. 1989：1859-1866；Barton 等人，(1987)Synthesis of Novel α-Amino-Acids andDerivatives Using Radical Chemistry：synthesis of L-and D-α-Amino-Adipic Acids，L-α-aminopimelic Acid and Appropriate Unsaturated Derivatives. Tetrahedron Lett. 43：4297-4308；以及 Subasinghe 等人，(1992)Quisqualic acid analogues：synthesis ofbeta-heterocyclic 2-aminopropanoic acid derivatives and their activity at a novelquisqualate-sensitized site. J. Med. Chem. 35：4602-7。

[0159] 非天然氨基酸的细胞摄取

[0160] 脊椎动物细胞对非天然氨基酸的摄取是在设计和选择例如并入蛋白质中的非天然氨基酸时通常需考虑的一个问题。举例来说，α-氨基酸的高电荷密度提示这些化合物可能无法渗透细胞。借助于一系列基于蛋白质的转运系统将天然氨基酸吸收到脊椎动物细胞中。可进行快速筛选以评定哪些非天然氨基酸(如果有)被细胞吸收。例如，参见例如 2002 年 12 月 22 日申请的代理人案号为 P1001US00 的名称为“ProteinArrays”的申请案以及 Liu，D.R. 和 Schultz，P.G.(1999)Progress toward the evolution ofan organism with an expanded genetic code. PNAS United States 96：4780-4785 中的毒性检定。虽然易于利用各种检定来分析摄取，但设计适于细胞摄取途径的非天然氨基酸的替代方案在于提供活体内产生氨基酸的生物合成途径。

[0161] 非天然氨基酸的生物合成

[0162] 细胞中已存在许多产生氨基酸和其它化合物的生物合成途径。虽然自然界中(例如脊椎动物细胞中)可能不存在特定非天然氨基酸的生物合成方法，但本发明提供这些方法。举例来说，在宿主细胞中通过添加新颖酶或修饰现有的宿主细胞途径而视情况产生非

天然氨基酸的生物合成途径。其它新颖酶视情况为天然存在的酶或人工发展的酶。举例来说,对氨基苯丙氨酸的生物合成(如名称为“In vivo incorporationof unnatural amino acids”的WO 2002/085923中的实例所提供)依赖于添加来自其它生物体的已知酶的组合。可通过用包含这些酶的基因的质粒转化细胞而将所述基因引入脊椎动物细胞中。当在细胞中表达时,这些基因提供合成所需化合物的酶促途径。视情况添加的酶的类型的实例提供于以下实例中。其它酶序列例如可见于Genbank中。也可以相同的方式视情况将人工发展的酶添加到细胞中。以此方式操纵细胞机构和细胞的资源以产生非天然氨基酸。

[0163] 可利用多种方法来产生用于生物合成途径或发展现有途径的新颖酶。举例来说,视情况使用(例如)如由Maxygen, Inc.(可在万维网的www.maxygen.com上获得)研发的递归重组来研发新颖酶和途径。例如参见Stemmer(1994),Rapid evolution of aprotein in vitro by DNA shuffling,Nature370(4):389-391;和Stemmer,(1994),DNAshuffling by random fragmentation and reassembly:In vitro recombination formolecular evolution,Proc.Natl.Acad.Sci.USA.,91:10747-10751。类似地,视情况使用Genencor(可在万维网的genencor.com上获得)研发的DesignPath™进行代谢途径工程化,例如工程化在细胞中产生O-甲基-L-酪氨酸的途径。这种技术使用例如通过功能基因组学鉴别的新颖基因的组合和分子进化和设计在宿主生物体中重新构 筑现有途径。Diversa公司(可在万维网的diversa.com上获得)还提供快速筛选基因文库和基因途径(例如建立新的途径)的技术。

[0164] 本发明包括通过在质粒构筑体中使用约2、3、4、5、6、7、8、9、10、40或多达约80个tRNA复本来增加琥珀抑制和选择密码子识别的方法。在本发明的一个实施例中,这些复本为杂交tRNA且可包括多种如图2和图6中所示的终止序列。本发明还包括其它转染载体(包括诸如λ噬菌体等重组噬菌体)和粘粒DNA表达载体等,且通常包括启动子序列(包括(但不限于)诸如U6和H1等Pol III启动子)。另外,噬菌体允许在转染载体中包括更多tRNA复本(多达约200个),此增加细胞中的总tRNA复本数。

[0165] 通常以足以进行有效蛋白质生物合成(例如天然细胞量)但未达到影响其它氨基酸的浓度或耗尽细胞资源的程度的浓度来产生利用本发明的工程化生物合成途径所产生的非天然氨基酸。以此方式在活体内产生的典型浓度为约10mM到约0.05mM。在用包含用于产生特定途径所需的酶的基因的质粒转化细胞且产生非天然氨基酸后,视情况使用活体内选择以针对核糖体蛋白合成和细胞生长进一步优化非天然氨基酸的产生。

[0166] 具有非天然氨基酸的多肽

[0167] 所关注的具有至少一个非天然氨基酸的蛋白质或多肽是本发明的特征。本发明还包括使用本发明的组合物和方法产生的具有至少一个非天然氨基酸的多肽或蛋白质。赋形剂(例如医药学上可接受的赋形剂)也可与蛋白质一起存在。

[0168] 通过在脊椎动物细胞中产生所关注的具有至少一个非天然氨基酸的蛋白质或多肽,蛋白质或多肽通常将包括脊椎动物转译后修饰。在某些实施例中,蛋白质包括至少一个非天然氨基酸和至少一个由脊椎动物细胞在活体内进行的翻译后修饰,其中所述翻译后修饰并非由原核细胞进行。举例来说,翻译后修饰例如包括乙酰化、酰化、脂质修饰、棕榈酰化、棕榈酸盐加成、磷酸化、糖脂键修饰、糖基化等。一方面,翻译后修饰包括通过GlcNAc-天冬酰胺键使寡糖(例如$(GlcNAc\text{-}Man)_2\text{-}Man\text{-}GlcNAc\text{-}GlcNAc$)与天冬酰胺连

接。也参见表7,表7列出脊椎动物蛋白质的N-连接寡糖的一些实例(也可存在其它未绘示的残基)。另一方面,翻译后修饰包括通过GalNAc-丝氨酸或者GalNAc-苏氨酸键或GlcNAc-丝氨酸或GlcNAc-苏氨酸键使寡糖(例如Gal-GalNAc、Gal-GlcNAc等)与丝氨酸或苏氨酸连接。

[0169] 表7:通过 GlcNAc键连接的寡糖的实例

[0170]

类型	基本结构
高甘露糖	Manα1-6 Manα1-3 → Manα1-6 Manα1-3 → Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn
杂合体	Manα1-6 GlcNAcβ1-2 —— Manα1-3 → Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn
复合物	GlcNAcβ1-2 —— Manα1-6 GlcNAcβ1-2 —— Manα1-3 → Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn
木糖	Manα1-6 Xylβ1-2 → Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn

[0171] 另一方面,翻译后修饰包括前体(例如降钙素前体、降钙素基因相关肽前体、前甲状旁腺激素原(preproparathyroid hormone)、前胰岛素原、胰岛素原、前阿黑皮素原(prepro-opiomelanocortin)、阿黑皮素原等)的蛋白水解加工,组装成多亚单元蛋白质或大分子组装物,翻译到细胞中的另一位点(例如细胞器,诸如内质网、高尔基体、核、溶酶体、过氧化物酶体、线粒体、叶绿体、空泡等,或通过分泌途径)。在某些实施例中,蛋白质包含分泌或定位序列、抗原决定基标签、FLAG标签、聚组氨酸标签、GST融合物等。

[0172] 非天然氨基酸的一个优势在于,其提供可用于添加其它分子的其它化学部分。这些修饰可在脊椎动物细胞中活体内进行或在活体外进行。因此,在某些实施例中,翻译后修饰是通过非天然氨基酸进行。举例来说,翻译后修饰可通过亲核-亲电子反应进行。目前用于选择性修饰蛋白质的大多数反应包括亲核与亲电子反应搭配物之间的共价键形成,例如α-卤基酮与组氨酸或半胱氨酸侧链的反应。这些情况下的选择性是由蛋白质中亲核残基的数目和可及性决定。在本发明的蛋白质中,可在活体外和活体内使用其它更具选择性的反应,诸如非天然酮基氨基酸与酰肼或氨氧基化合物的反应。例如参见Cornish等人,(1996)Am.Chem.Soc,118:8150-8151;Mahal等人,(1997)Science,276:1125-1128;Wang等人,(2001)Science 292:498-500;Chin等人,(2002)Am.Chem.Soc.124:9026-9027;Chin等人,(2002)Proc.Natl.Acad.Sci.99:11020-11024;Wang等人,(2003)Proc.Natl.Acad.Sci.,100:56-61;Zhang等人,(2003)Biochemistry,42:6735-6746;以及Chin等人,(2003)Science,出版中。这允许用许多试剂选择性标记几乎任何蛋白质,所述试剂包括荧光团、交联剂、糖类衍生物以及细胞毒性分子。也参见2003年10月15日申请的名称为"Glycoproteinsynthesis"的专利申请案USSN 10/686,944。例如,通过叠氮基氨基酸

进行的翻译后修饰也可通过斯达汀格连接反应(Staudinger ligation)(例如用三芳基膦试剂)进行。例如参见Kiick等人,(2002)Incorporation of azides into recombinant proteins forchemoselective modification by the Staudinger ligtation,PNAS99:19-24。

[0173]　本发明提供选择性修饰蛋白质的另一种极为有效的方法,其包括响应选择密码子将例如含有叠氮部分或炔基部分的非天然氨基酸遗传并入蛋白质中。然后,可例如通过休斯根(Huisgen)[3+2]环加成反应(例如参见Padwa,A.,ComprehensiveOrganic Synthesis,第4卷,(1991)Trost,B.M.编,Pergamon,Oxford,第1069-1109页;和Huisgen,R.,1,3-Dipolar Cycloaddition Chemistry,(1984)Padwa,A.编,Wiley,NewYork,第1-176页)分别用例如炔基或叠氮衍生物来修饰这些氨基酸侧链。例如参见图16。因为这一方法包括环加成而不是亲核取代,所以可以极高选择性修饰蛋白质。可在室温下在水性条件下通过向反应混合物中添加催化量的Cu(I)盐而以极佳区域选择性(1,4＞1,5)进行这一反应。例如参见Tornoe等人,(2002)Org.Chem.67:3057-3064;和Rostovtsev等人,(2002)Angew.Chem.Int.Ed.41:2596-2599。可使用的另一方法为用四半胱氨酸(tetracysteine)基元对双砷化合物进行配位体交换,例如参见Griffin等人,(1998)Science 281:269-272。

[0174]　可通过非天然编码氨基酸的官能团添加到本发明蛋白质中的分子包括几乎任何具有互补官能团的分子。这些分子包括(但不限于)染料、荧光团、交联剂、糖类衍生物、聚合物(例如聚乙二醇的衍生物)、光交联剂、细胞毒性化合物、亲和标记、生物素衍生物、树脂、珠粒、第二蛋白质或多肽(或更多)、多核苷酸(例如DNA、RNA等)、金属螯合剂、辅因子、脂肪酸、碳水化合物等。另一方面,本发明提供包括这些分子的组合物和产生这些分子(例如聚乙二醇衍生物)的方法,其中n为介于例如50与10,000、75与5,000、100与2,000、100与1,000等之间的整数。在本发明的实施例中,聚乙二醇具有例如约5,000到约100,000Da,约20,000到约30,000、约40,000或约50,000Da、约20,000到约10,000Da等的分子量。

[0175]　也提供包含这些化合物,例如具有蛋白质和细胞的各种组合物。在本发明的一方面,包含叠氮基染料的蛋白质(例如具有化学结构4或化学结构6)进一步包括至少一个非天然氨基酸(例如炔基氨基酸),其中叠氮基染料通过[3+2]环加成与非天然氨基酸连接。

[0176]　本发明的脊椎动物细胞提供合成包含大量有用的非天然氨基酸的蛋白质的能力。一方面,组合物视情况包括例如至少10微克、至少50微克、至少75微克、至少100微克、至少200微克、至少250微克、至少500微克、至少1毫克、至少10毫克或10毫克以上或活体内蛋白质产生方法可实现的量(本文中提供关于重组蛋白质产生和纯化的细节)的包含非天然氨基酸的蛋白质。另一方面,蛋白质视情况以在例如细胞溶解产物、缓冲剂、医药缓冲剂或其它液体悬浮液(例如,体积约为例如约1纳升到约100升)中例如每升至少10微克蛋白质、每升至少50微克蛋白质、每升至少75微克蛋白质、每升至少100微克蛋白质、每升至少200微克蛋白质、每升至少250微克蛋白质、每升至少500微克蛋白质、每升至少1毫克蛋白质或每升至少10毫克或10毫克以上蛋白质的浓度存在于组合物中。在脊椎动物细胞中产生大量(例如,大于通常用其它方法、例如活体外翻译可能获得的量)蛋白质(包括至少一个非天然氨基酸)是本发明的特征。

[0177] 可进行非天然氨基酸的并入，从而来例如调整蛋白质结构和/或功能的改变，例如改变大小、酸度、亲核性、氢键、疏水性、蛋白酶靶位点的可及性、靶向部分（例如，对于蛋白质阵列）等。包括非天然氨基酸的蛋白质可具有增加或甚至完全新颖的催化或物理性质。举例来说，视情况通过在蛋白质中包括非天然氨基酸来改变下列性质：毒性、生物分布(biodistribution)、结构性质、光谱性质、化学和/或光化学性质、催化能力、半衰期（例如血清半衰期）、与其它分子反应（例如共价或非共价）的能力等。包括包含至少一个非天然氨基酸的蛋白质的组合物可用于例如新颖治疗剂、诊断剂、催化酶、工业酶、结合蛋白（例如抗体）以及例如蛋白质结构和功能的研究。例如参见Dougherty,(2000)Unnatural Amino Acids as Probes of ProteinStructure and Function，Current Opinion in Chemical Biology.4：645-652。

[0178] 在本发明的一方面，组合物包括至少一种具有至少一个，例如至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个或至少十个或十个以上非天然氨基酸的蛋白质。非天然氨基酸可相同或不同，例如在蛋白质中可存在1、2、3、4、5、6、7、8、9、10个或10个以上不同位点，其包含1、2、3、4、5、6、7、8、9、10个或10个以上不同的非天然氨基酸。另一方面，组合物包括蛋白质中所存在的至少一个（但少于全部）特定氨基酸经非天然氨基酸取代的蛋白质。对于具有一个以上非天然氨基酸的给定蛋白质，非天然氨基酸可相同或不同（例如，蛋白质可包括两个或两个以上不同类型的非天然氨基酸，或可包括两个相同的非天然氨基酸）。对于具有两个以上非天然氨基酸的给定蛋白质，非天然氨基 酸可相同、不同或为多个相同种类的非天然氨基酸与至少一个不同非天然氨基酸的组合。

[0179] 基本上任何包括非天然氨基酸（和任何相应的编码核酸，例如包括一个或一个以上选择密码子）的蛋白质（或其部分）都可使用本文中的组合物和方法产生。未尝试鉴别成千上万种已知蛋白质，所述蛋白质中任一种都可例如通过调整任何可利用的突变方法以在相关翻译系统中包括一个或一个以上适当选择密码子而经修饰，从而包括一个或一个以上非天然氨基酸。已知蛋白质的常见序列谱系包括GenBankEMBL、DDBJ以及NCBI。其它谱系可通过搜索互联网容易地鉴别。

[0180] 蛋白质通常与任何可利用的蛋白质（例如治疗蛋白、诊断蛋白、工业酶或其部分等）例如至少60％、至少70％、至少75％、至少80％、至少90％、至少95％或至少99％或99％以上一致，且其包含一个或一个以上非天然氨基酸。可经修饰以包含一个或一个以上非天然氨基酸的治疗蛋白、诊断蛋白以及其它蛋白的实例包括（但不限于）例如α-1抗胰蛋白酶、血管抑素、抗溶血因子、抗体（关于抗体的更多细节可见于下文）、载脂蛋白(Apolipoprotein)、载脂蛋白(Apoprotein)、心房利尿钠因子、心房利尿钠多肽、心房肽、C-X-C趋化因子（例如T39765、NAP-2、ENA-78、Gro-a、Gro-b、Gro-c、IP-10、GCP-2、NAP-4、SDF-1、PF4、MIG)、降钙素、CC趋化因子（例如单核细胞趋化蛋白-1、单核细胞趋化蛋白-2、单核细胞趋化蛋白-3、单核细胞发炎性蛋白-1α、单核细胞发炎性蛋白-1β、RANTES、I309、R83915、R91733、HCC1、T58847、D31065、T64262)、CD40配位体、C-kit配位体、胶原蛋白、菌落刺激因子(CSF)、补体因子5a、补体抑制剂、补体受体1、细胞因子（例如上皮中性粒细胞激活肽-78、GROα/MGSA、GROβ、GROγ、MIP-1α、MIP-1δ、MCP-1)、表皮生长因子(EGF)、促红细胞生成素（“EPO”，代表通过并入一个或一个以上非天然氨基酸加以修饰的

优选标靶）、剥脱性毒素 A 和 B、因子 IX、因子 VII、因子 VIII、因子 X、成纤维细胞生长因子 (FGF)、纤维蛋白原、纤维粘连蛋白、G-CSF、GM-CSF、葡糖脑苷脂酶、促性腺激素、生长因子、刺猬蛋白（例如音速 (Sonic)、印度 (Indian)、沙漠 (Desert)）、血红蛋白、肝细胞生长因子 (HGF)、水蛭素、人类血清白蛋白、胰岛素、胰岛素样生长因子 (IGF)、干扰素（例如 IFN-α、IFN-β、IFN-γ）、白细胞介素（例如 IL-1、IL-2、IL-3、IL-4、IL-5、IL-6、IL-7、IL-8、IL-9、IL-10、IL-11、IL-12 等）、角质细胞生长因子 (KGF)、乳铁蛋白、白血病抑制因子、荧光素酶、神经营养因子、中性粒细胞抑制因子 (NIF)、抑瘤素 M、成骨蛋白、甲状旁腺激素、PD-ECSF、PDGF、肽激素（例如人类生长激素）、多效营养因子、蛋 白质 A、蛋白质 G、致热外毒素 A、B 和 C、松弛素、肾素、SCF、可溶性补体受体 I、可溶性 I-CAM 1、可溶性白细胞介素受体 (IL-1、2、3、4、5、6、7、8、9、10、11、12、13、14、15)、可溶性 TNF 受体、生长调节素、生长抑素、生长激素、链激酶、超抗原（即，葡萄球菌肠毒素 (SEA、SEB、SEC1、SEC2、SEC3、SED、SEE)）、超氧化物歧化酶 (SOD)、中毒性休克综合症毒素 (TSST-1)、胸腺素 α1、组织型纤溶酶原激活剂、肿瘤生长因子 β (TGFβ)、肿瘤坏死因子受体 (TNFR)、肿瘤坏死因子 -α (TNFα)、血管内皮生长因子 (VEGEF)、尿激酶等。

[0181] 可使用本文中所述的用于活体内并入非天然氨基酸的组合物和方法产生的一类蛋白质包括转录调节蛋白或其部分。例示性转录调节蛋白包括调节细胞生长、分化、调控等的基因和转录调节蛋白。转录调节蛋白可见于原核生物、病毒以及真核生物（包括真菌、植物、酵母、昆虫以及包括哺乳动物的动物）中，从而提供大量治疗性标靶。应了解表达和转录激活因子通过许多机制调控转录，例如通过与受体结合、刺激信号转导级联、调控转录因子表达、与启动子和增强子结合、与结合启动子和增强子的蛋白质结合、解开 DNA、剪接前体 mRNA、使 RNA 多腺苷酸化以及降解 RNA。举例来说，脊椎动物细胞中的 GAL4 蛋白或其部分的组合物也是本发明的特征。GAL4 蛋白或其部分通常包含至少一个非天然氨基酸。也参见本文中标题为“正交氨酰基 tRNA 合成酶”的部分。

[0182] 一类本发明的蛋白质（例如具有一个或一个以上非天然氨基酸的蛋白质）包括表达激活因子，诸如细胞因子、发炎性分子、生长因子、其受体以及致癌基因产物，例如白细胞介素（例如 IL-1、IL-2、IL-8 等）、干扰素、FGF、IGF-I、IGF-II、FGF、PDGF、TNF、TGF-α、TGF-β、EGF、KGF、SCF/c-Kit、CD40L/CD40、VLA-4/VCAM-1、ICAM-1/LFA-1 以及透明质酸 / CD44；信号转导分子和相应致癌基因产物，例如 Mos、Ras、Raf 以及 Met；和转录激活因子和抑制因子，例如 p53、Tat、Fos、Myc、Jun、Myb、Rel 以及类固醇激素受体（诸如雌激素、孕酮、睾酮、醛固酮、LDL 受体配位体和皮质酮的受体）。

[0183] 本发明还提供具有至少一个非天然氨基酸的酶（例如工业酶）或其部分。酶的实例包括（但不限于）例如酰胺酶、氨基酸消旋酶、酰化酶、脱卤酶、双加氧酶、二芳基丙烷过氧化物酶、表异构酶、环氧化物水解酶、酯酶、异构酶、激酶、葡萄糖异构酶、糖苷酶、糖基转移酶、卤代过氧化物酶 (haloperoxidase)、单加氧酶（例如 p450)、脂肪酶、木质素过氧化物酶、腈水合酶、腈水解酶、蛋白酶、磷酸酶、枯草杆菌蛋白酶、转氨酶以及核酸酶。

[0184] 这些蛋白质中的多者可购得（例如参见 Sigma BioSciences 2002 目录和价格表），且相应蛋白质序列和基因以及通常其许多变异体也为我们所熟知（例如参见 Genbank)。可通过插入本发明的一种或一种以上非天然氨基酸来修饰其中任一个，从而例如改变蛋白质的一种或一种以上所关注的治疗、诊断或酶促性质。治疗上相关的性质的实

例包括血清半衰期、保存半衰期、稳定性、免疫原性、治疗活性、可检测性(例如,通过在非天然氨基酸中包括报告基团(例如标记或标记结合位点))、LD_{50}或其它副作用的减少、通过胃道进入身体的能力(例如口服可用性)等。诊断性质的实例包括保存半衰期、稳定性、诊断活性、可检测性等。相关酶促性质的实例包括保存半衰期、稳定性、酶促活性、生产能力等。

[0185] 也可修饰多种其它蛋白质以使其包括本发明的一个或一个以上非天然氨基酸。举例来说,本发明可包括利用例如以下来源的蛋白质中的非天然氨基酸取代一种或一种以上疫苗蛋白中的一个或一个以上天然氨基酸:感染性真菌,例如曲霉属、念珠菌属;细菌,尤其用作病原菌模型的大肠杆菌以及医学上重要的细菌,诸如葡萄球菌(Staphylococci)(例如金黄色葡萄球菌(aureus))或链球菌(Streptococci)(例如肺炎链球菌(pneumoniae));原生动物,诸如孢子虫纲(sporozoa)(例如疟原虫属(Plasmodia))、根足虫纲(rhizopods)(例如内阿米巴属(Entamoeba)以及鞭毛虫纲(flagellates)(锥虫属(Trypanosoma)、利什曼原虫属(Leishmania)、毛滴虫属(Trichomonas)、贾第鞭毛虫属(Giardia)等);病毒,诸如(+)RNA病毒(实例包括痘病毒(Poxviruses),例如牛痘病毒;微小RNA病毒(Picornaviruses),例如脊髓灰质炎病毒;外衣病毒(Togaviruses),例如风疹病毒;黄病毒(Flaviviruses),例如HCV;以及冠状病毒)、(-)RNA病毒(例如棒状病毒(Rhabdoviruse),例如VSV;副粘病毒(Paramyxoviruse),例如RSV;正粘病毒(Orthomyxoviruse),例如流感病毒;本雅病毒(Bunyaviruse);以及沙状病毒(Arenaviruse))、dsDNA病毒(例如呼肠孤病毒(Reoviruses))、RNA变为DNA的病毒(即,逆转录病毒)(例如HIV和HTLV)以及某些DNA变为RNA的病毒(诸如乙型肝炎病毒)。

[0186] 农业相关的蛋白质也是非天然氨基酸修饰的合适标靶,诸如昆虫抗性蛋白(例如Cry蛋白)、淀粉和脂质产生酶、植物和昆虫毒素、毒素抗性蛋白、霉菌毒素解毒蛋白、植物生长酶(例如1,5-二磷酸核酮糖羧化酶/加氧酶,"RUBISCO")、脂氧合酶(LOX)以及磷酸烯醇丙酮酸(PEP)羧化酶。

[0187] 本发明还提供在脊椎动物细胞中产生至少一种包含至少一个非天然氨基酸的蛋白质的方法(和由这些方法产生的蛋白质)。举例来说,方法包括:使包含包括至少 一个选择密码子且编码蛋白质的核酸的脊椎动物细胞在适当培养基中生长。脊椎动物细胞还包含在细胞中起作用且识别选择密码子的正交tRNA(O-tRNA)和优先利用非天然氨基酸使O-tRNA氨酰化的正交氨酰基tRNA合成酶(O-RS),且培养基包含非天然氨基酸。

[0188] 在一个实施例中,所述方法进一步包括在蛋白质中并入非天然氨基酸,其中非天然氨基酸包含第一反应性基团;和使蛋白质与包含第二反应性基团的分子(例如染料、例如聚乙二醇衍生物等聚合物、光交联剂、细胞毒性化合物、亲和标记、生物素衍生物、树脂、第二蛋白质或多肽、金属螯合剂、辅因子、脂肪酸、碳水化合物、多核苷酸(例如DNA、RNA等)等)接触。第一反应性基团与第二反应性基团反应以通过[3+2]环加成使分子与非天然氨基酸连接。在一个实施例中,第一反应性基团为炔基或叠氮基部分,且第二反应性基团为叠氮基或炔基部分。举例来说,第一反应性基团为炔基部分(例如在非天然氨基酸对炔丙氧基苯丙氨酸中),且第二反应性基团为叠氮基部分。在另一实例中,第一反应性基团为叠氮基部分(例如在非天然氨基酸对叠氮基-L-苯丙氨酸中),且第二反应性基团为炔基部分。

[0189] 在一个实施例中,O-RS以具有(例如)如SEQ ID NO:86或45中所述的氨基酸序

列的O-RS的效率的至少50％利用非天然氨基酸使O-tRNA氨酰化。在另一实施例中，O-tRNA包含SEQ ID NO:65或64或其互补多核苷酸序列，由SEQ ID NO:65或64或其互补多核苷酸序列加工而成或由SEQ ID NO:65或64或其互补多核苷酸序列编码。在又一实施例中，O-RS包含SEQ ID NO:36-63和/或86中任一序列所述的氨基酸。

[0190] 编码蛋白可例如包含治疗蛋白、诊断蛋白、工业酶或其部分。由所述方法产生的蛋白质视情况由非天然氨基酸进一步修饰。举例来说，视情况通过至少一种活体内翻译后修饰来修饰由所述方法产生的蛋白质。

[0191] 还提供产生筛选或选择转录调节蛋白的方法（和由这些方法产生的筛选或选择转录调节蛋白）。举例来说，方法包括选择第一多核苷酸序列，其中多核苷酸序列编码核酸结合域；和使第一多核苷酸序列突变以包括至少一个选择密码子。这提供筛选或选择多核苷酸序列。所述方法还包括选择第二多核苷酸序列，其中第二多核苷酸序列编码转录激活域；提供包含可操作性连接到第二多核苷酸序列的筛选或选择多核苷酸序列的构筑体；以及在细胞中引入构筑体、非天然氨基酸、正交tRNA合成酶(O-RS)以及正交tRNA(O-tRNA)。利用这些组件，O-RS优先利用非天然氨基酸使O-tRNA氨酰化，且O-tRNA识别选择密码子且响应筛选或选择多核苷酸序列中的选择密码子将非天然氨基酸并入核酸结合域中，从而提供筛选或选择转录调节蛋白。

[0192] 在某些实施例中，本发明的方法和/或组合物中的所关注的蛋白质或多肽（或其部分）是由核酸编码。核酸通常包含至少一个选择密码子、至少两个选择密码子、至少三个选择密码子、至少四个选择密码子、至少五个选择密码子、至少六个选择密码子、至少七个选择密码子、至少八个选择密码子、至少九个选择密码子、十个或十个以上选择密码子。

[0193] 可使用所属领域的技术人员所熟知且在本文中在“诱变和其它分子生物学技术”下描述的方法使编码所关注的蛋白质或多肽的基因诱变以例如包括一个或一个以上用于并入非天然氨基酸中的选择密码子。举例来说，使所关注的蛋白质的核酸诱变以包括一个或一个以上选择密码子，从而提供插入一个或一个以上非天然氨基酸。本发明包括例如包括至少一个非天然氨基酸的任何蛋白质的任何此种变异体（例如，突变体）型式。类似地，本发明还包括相应核酸，即，任何编码一个或一个以上非天然氨基酸的具有一个或一个以上选择密码子的核酸。

[0194] 纯化包含非天然氨基酸的重组蛋白

[0195] 根据所属领域技术人员已知且使用的标准程序，可将本发明的蛋白质（例如包含非天然氨基酸的蛋白质、包含非天然氨基酸的蛋白质的抗体等）部分或实质上纯化成均质。因此，可通过所属领域中熟知的许多方法中的任一种来回收和纯化本发明的多肽，所述方法例如包括硫酸铵或乙醇沉淀、酸或碱萃取、柱色谱、亲和柱色谱、阴离子或阳离子交换色谱、磷酸纤维素色谱、疏水性相互作用色谱、羟基磷灰石色谱、凝集素色谱、凝胶电泳等。在制备正确折叠的成熟蛋白质时，必要时可使用蛋白质再折叠步骤。当需要高纯度时，在最后纯化步骤中可使用高效液相色谱(HPLC)、亲和色谱或其它合适方法。在一个实施例中，使用针对非天然氨基酸（或包含非天然氨基酸的蛋白质）制备的抗体作为纯化试剂，从而例如用于包含一个或一个以上非天然氨基酸的蛋白质的基于亲和力的纯化。必要时，在部分纯化或纯化成均质后，视情况将多肽例如用作检定组分、治疗剂或抗体产生的免疫原。

[0196] 除了本文中所说明的其它参考文献外，所属领域中熟知多种纯化/蛋白

质折叠方法，例如包括以下文献中所述的方法：R.Scopes，Protein Purification. Springer-Verlag，N.Y.(1982)；Deutscher，Methods in Enzymology，第182卷：Guide to ProteinPurification. Academic Press，Inc.N.Y.(1990)；Sandana(1997)Bioseparation ofProteins. Academic Press，Inc.；Bollag 等 人，(1996)Protein Methods. 第 2 版，Wiley-Liss，NY；Walker(1996)The Protein Protocols Handbook Humana Press，NJ；Harris 和 Angal(1990)Protein Purification Applications：A Practical Approach IRLPress at Oxford，Oxford，England；Harris 和 Angal Protein Purification Methods：APractical Approach IRL Press at Oxford，Oxford，England；Scopes(1993) ProteinPurification：Principles and Practice，第3版，Springer Verlag，NY；Janson 和 Ryden(1998)Protein Purification：Principles.High Resolution Methods and Applications.第 2版，Wiley-VCH，NY；以及 Walker(1998)Protein Protocols on CD-ROM Humana Press，NJ；以及其中所引用的参考文献。

[0197] 在脊椎动物细胞中产生具有非天然氨基酸的所关注的蛋白质或多肽的一个优势在于，蛋白质或多肽通常将以其天然构象经折叠。然而，在本发明的某些实施例中，所属领域技术人员将认识到在合成、表达和/或纯化后，蛋白质可能具有不同于相关多肽所需的构象的构象。在本发明的一方面，视情况使所表达的蛋白质变性，且然后使其复性。这是例如通过向所关注的蛋白质或多肽中添加伴侣蛋白和/或通过将蛋白质溶解于诸如盐酸胍等离液剂(chaotropic agent)中来实现。

[0198] 一般来说，有时希望使所表达的多肽变性和还原，且然后使多肽再折叠成优选构象。举例来说，可向所关注的翻译产物中添加胍、脲、DTT、DTE 和/或伴侣蛋白。所属领域技术人员熟知使蛋白质还原、变性和复性的方法(参见以上参考文献，以及 Debinski 等人，(1993)L Biol.Chem.，268：14065-14070；Kreitman 和 Pastan(1993) Bioconiug.Chem.，4：581-585；以及 Buchner 等人，(1992)Anal.Biochem.，205：263-270)。举例来说，Debinski 等人描述胍-DTE 中包涵体蛋白的变性和还原。可在例如含有氧化型谷胱甘肽和 L-精氨酸的氧化还原缓冲剂中使蛋白质再折叠。可使再折叠试剂流动或以其它方式移动以与一种或一种以上多肽或其它表达产物接触，或者可使一种或一种以上多肽或其它表达产物流动或以其它方式移动以与再折叠试剂接触。

[0199] 抗体

[0200] 一方面，本发明提供针对本发明的分子，例如合成酶、tRNA 以及包含非天然氨基酸的蛋白质的抗体。本发明的分子的抗体适用作纯化试剂，从而例如用于纯化本发明的分子。另外，抗体可用作指示试剂以指示合成酶、tRNA 或包含非天然氨基酸的蛋白质的存在，从而例如追踪分子的存在或位置(例如活体内或原位)。

[0201] 本发明的抗体可为包含一种或一种以上实质上或部分由免疫球蛋白基因或免疫球蛋白基因片段编码的多肽的蛋白质。所识别的免疫球蛋白基因包括 κ、λ、α、γ、δ、ε 以及 μ 恒定区基因以及多种免疫球蛋白可变区基因。轻链分为 κ 或者 λ。重链分为 γ、μ、α、δ 或 ε，其又分别界定免疫球蛋白的种类 IgG、IgM、IgA、IgD 以及 IgE。典型的免疫球蛋白(例如抗体)结构单元包含四聚物。各四聚物包含两对相同的多肽链，各对具有一个“轻”链(约 25kD)和一个“重”链(约 50-70kD)。各链的 N-末端界定主要负责抗原识别的具有约 100 到 110 个或 110 个以上氨基酸的可变区。术语可变轻链(V_L)和可变

重链(V_H)分别是指这些轻链和重链。

[0202] 抗体是以完整免疫球蛋白或通过由各种肽酶消化产生的多个明确表征的片段形式存在。因此，举例来说，胃蛋白酶消化铰链区中二硫键以下的抗体，从而产生Fab的二聚体F(ab′)$_2$，所述Fab本身为通过二硫键与V_H-C_H1连接的轻链。可在温和条件下还原F(ab′)$_2$以断开铰链区中的二硫键，从而使F(ab′)$_2$二聚体转化为Fab′单体。Fab′单体基本上为具有铰链区的部分的Fab(关于其它抗体片段的更多详细描述，参见Fundamental Immunology，第4版，W.E.Paul编，Raven Press，N.Y.(1999))。虽然各种抗体片段都是根据完整抗体的消化来定义，但技术人员应了解这些Fab′片段等可以化学方法或利用重组DNA方法重新合成。因此，如本文中所使用的术语抗体还视情况包括通过修饰整个抗体所产生或使用重组DNA方法重新合成的抗体片段。抗体包括单链抗体，包括其中可变重链和可变轻链连接在一起(直接或通过肽连接子)形成连续多肽的单链Fv(sFv或scFv)抗体。本发明的抗体例如可为多克隆片段、单克隆片段、嵌合片段、人类化片段、单链片段、Fab片段、由Fab表达文库产生的片段等。

[0203] 一般来说，本发明的抗体在多种分子生物或医药过程中作为一般试剂和治疗性试剂都极具价值。可获得产生多克隆和单克隆抗体的方法，且可将其用于产生本发明的抗体。许多基础文章描述标准抗体产生方法，所述文章例如包括Borrebaeck(编)(1995)Antibody Engineering.第2版 Freeman and Company，NY(Borrebaeck)；McCafferty等人，(1996)Antibody Engineering.A Practical Approach IRL at OxfordPress，Oxford，England(McCafferty)；和Paul(1995)Antibody Engineering ProtocolsHumana Press，Towata，NJ(Paul)；Paul(编)，(1999)Fundamental Immunology.第5版 Raven Press，N.Y.；Coligan(1991)Current Protocols in Immunology Wiley/Greene，NY；Harlow和Lane(1989)Antibodies:A Laboratory Manual Cold Spring Harbor Press，NY；Stites等人(编)Basic and Clinical Immunology(第4版)Lange MedicalPublications，Los Altos，CA，和其中所引用的参考文献；Goding(1986)MonoclonalAntibodies:Principles and Practice(第2版)Academic Press，New York，NY；以及Kohler和Milstein(1975)Nature 256:495-497。

[0204] 已研发多种不依赖于例如将抗原注射到动物中的关于抗体制备的重组技术，且可将其用于本发明的情形中。举例来说，有可能在噬菌体或类似载体中产生和选择重组抗体文库。例如参见Winter等人，(1994)Making Antibodies by Phage DisplayTechnology Annu.Rev.Immunol.12:433-55和其中回顾的所引用的参考文献。也参见Griffiths和Duncan(1998)Strategies for selection of antibodies by phage display CurrOpin Biotechnol 9:102-8；Hoogenboom等人，(1998)Antibody phage displaytechnology and its applications Immunotechnology 4:1-20；Gram等人，(1992)in vitroselection and affinity maturation of antibodies from a naive combinatorialimmunoglobulin library PNAS 89:3576-3580；Huse等人，(1989)Science 246:1275-1281；和Ward等人，(1989)Nature 341:544-546。

[0205] 在一个实施例中，抗体文库可包括经克隆以在丝状噬菌体表面上显示相关的重链和轻链可变域的V基因的谱系(例如从大量淋巴细胞收集或在活体外组装)。通过与抗原结合来选择噬菌体。可溶性抗体是由感染噬菌体的细菌表达且可例如借助于诱变来

改进抗体。例如参见 Balint 和 Larrick(1993)Antibody Engineering byParsimonious Mutagenesis Gene 137 :109-118 ;Stemmer 等人，(1993)Selection of an Active Single Chain Fv Antibody From a Protein Linker Library Prepared by EnzymaticInverse PCR Biotechniques 14(2) :256-65 ;Cramer 等 人，(1996)Construction andevolution ofantibody-phage libraries by DNA shuffling Nature Medicine 2 :100-103 ;以及 Crameri 和 Stemmer(1995)Combinatorial multiple cassette mutagenesis creates allthe permutations of mutant and wildtype cassettes BioTechniques 18 :194-195。

[0206] 也已知且利用克隆和表达重组抗体噬菌体系统的试剂盒，例如来自 Amersham-Pharmacia Biotechnology(Uppsala,Sweden)的"重组噬菌体抗体系统小鼠 ScFv 组件"。也已产生噬菌体抗体文库以通过链改组来制备高亲和力人类抗体(例如参见 Marks 等 人，(1992)By-Passing Immunization :Building High Affinity HumanAntibodies by Chain Shuffling Biotechniques 10 :779-782)。也应认识到，可通过多种商务服务中的任一种(例如 Bethyl Laboratories(Montgomery TX)、Anawa(Switzerland)、Eurogentec(Belgium 和美国的 Philadelphia，PA 等)等)来制备抗体。

[0207] 在某些实施例中，例如当治疗性给予本发明的抗体时，适用于使所述抗体"人类化"。使用人类化抗体倾向于减少针对治疗性抗体的不合需要的免疫反应的发生率(例如当患者为人类时)。上述抗体参考文献描述人类化策略。除人类化抗体外，人 类抗体也是本发明的特征。人类抗体是由特征性人类免疫球蛋白序列组成。人类抗体可使用多种方法产生(关于综述，例如参见 Larrick 等人的美国专利第 5,001,065 号)。通过三源杂交瘤 (trioma) 技术产生人类抗体的一般方法是由 Ostberg 等人，(1983)，Hybridoma 2 :361-367 ;Ostberg，美国专利第 4,634,664 号 ;以及 Engelman 等人，美国专利第 4,634,666 号描述。

[0208] 已知多种使用抗体来纯化和检测蛋白质的方法，且可将其用于检测和纯化如本文中所说明的包含非天然氨基酸的蛋白质。一般来说，抗体为 ELISA、Western 印迹法、免疫化学、亲和色谱法、SPR 以及许多其它方法的适用试剂。以上所说明的参考文献提供关于如何进行 ELISA 检定、Western 印迹、表面等离子体共振 (SPR) 等的细节。

[0209] 在本发明的一方面，本发明的抗体本身包括非天然氨基酸，从而提供具有所关注的性质(例如改进的半衰期、稳定性、毒性等)的抗体。也参见本文中标题为"具有非天然氨基酸的多肽"的部分。抗体占目前临床试验中的所有化合物的几乎 50% (Wittrup,(1999) Phage on display Tibtech 17 :423-424) 且普遍使用抗体作为诊断试剂。因此，利用非天然氨基酸修饰抗体的能力提供修饰这些有价值的试剂的重要工具。

[0210] 举例来说，MAb 在诊断学领域中存在多种应用。检定范围从简单的斑点测试 (spot test) 到较为复杂的方法，诸如来自 DuPont Merck Co. 的用于肿瘤成像的放射性标记 NR-LU-10MAb(Rusch 等人，(1993)NR-LU-10 monoclonal antibody scanning.A helpful new adjunct to computed tomography in evaluating non-small-cell lung cancer. J Thorac Cardiovasc Sure 106 :200-4)。如所说明，MAb 是 ELISA、Western 印迹法、免疫化学、亲和色谱法等的重要试剂。任何这种诊断抗体都可经修饰以包括一个或一个以上非天然氨基酸，从而例如改变 Ab 对标靶的特异性或亲和力，或例如通过在非天然氨基酸中包括可检测标记(例如光谱、荧光、发光等)来改变一种或一种以上可检测性质。

[0211] 一类有价值的抗体试剂为治疗性 Ab。举例来说，抗体可为通过靶向肿瘤细胞以通过抗体依赖性细胞介导的细胞毒性 (ADCC) 或补体介导的细胞溶解 (CML) 破坏来阻滞肿瘤生长的肿瘤特异性 MAb(Ab 的这些一般类型有时称为“魔弹 (magicbullet)”)。一个实例为美罗华 (Rituxan)，一种治疗非何杰金氏淋巴瘤 (Non-Hodgkinslymphoma) 的抗 -CD20MAb(Scott(1998)Rituximab :a new therapeutic monoclonalantibody for non-Hodgkin′ s lymphoma Cancer Pract 6 :195-7)。第二实例涉及干扰肿瘤 生长的关键组分的抗体。赫赛汀是治疗转移性乳癌的抗 HER-2 单克隆抗体且提供具有这种作用机制的抗体的实例 (Baselga 等人，(1998)Recombinant humanizedanti-HER2 antibody(Herceptin)enhances the antitumor activity ofpaclitaxel anddoxorubicin against HER2/neu overexpressing human breast cancer xenografts[公布的勘误表登载在 Cancer Res(1999)59(8) :2020 中]，Cancer Res 58 :2825-31)。第三个实例涉及将细胞毒性化合物 (毒素、放射性核素等) 直接传递到肿瘤或其它所关注的位点的抗体。举例来说，一种应用 Mab 是 CYT-356，一种将辐射直接靶向前列腺肿瘤细胞的 90Y 连接的抗体 (Deb 等人，(1996)Treatment of hormone-refractoryprostate cancer with 90Y-CYT-356 monoclonal antibody Clin Cancer Res 2 :1289-97)。第四种应用是抗体导向的酶前药疗法 (antibody-directed enzyme prodrug therapy)，其中与肿瘤共存的酶激活肿瘤附近的全身给予的前药。举例来说，正研发用于治疗结肠直肠癌的与羧基肽酶 A 连接的抗 Ep-CAM1 抗体 (Wolfe 等人，(1999)Antibody-directed enzyme prodrug therapy with the T268G mutant of humancarboxypeptidase A1 :in vitro and in vivo studies with prodrugs of methotrexate and thethymidylate synthase inhibitors GW1031 and GW1843 Bioconiug Chem 10 :38-48)。其它 Ab(例如拮抗剂) 经设计以特异性抑制正常细胞功能，从而实现治疗益处。一个实例为 Orthoclone OKT3，一种由强生公司 (Johnson and Johnson) 提供的用于减少急性器官移植排斥的抗 CD3MAb(Strate 等人，(1990)Orthoclone OKT3 as first-linetherapy in acute renal allograft rejection Transplant Proc 22 :219-20。另一类抗体产物是激动剂。这些 Mab 经设计以特异性增加正常细胞功能，从而实现治疗益处。举例来说，正研发用于神经病疗法的乙酰胆碱受体的基于 Mab 的激动剂 (Xie 等人，(1997)Direct demonstration of MuSK involvement in acetylcholine receptor clusteringthrough identification of agonist ScFv Nat.Biotechnol.15 :768-71。这些抗体中的任一种都可经修饰以包括一个或一个以上非天然氨基酸，从而增加一种或一种以上治疗性质 (特异性、亲和力、血清半衰期等)。

[0212] 另一类抗体产物提供新颖功能。这一组中的主要抗体是经工程化以模拟酶的催化能力的催化抗体，诸如 Ig 序列 (Wentworth 和 Janda(1998)Catalytic antibodies CurrOpin Chem Biol 2 :138-44)。举例来说，一种引人关注的应用涉及使用用于成瘾疗法的活体内水解可卡因 (cocaine) 的催化抗体 mAb-15A10(Mets 等人，(1998)A catalyticantibody against cocaine prevents cocaine′ s reinforcing and toxic effects in rats ProcNatl Acad Sci U S A 95 :10176-81)。催化抗体也可经修饰以包括一个或一个以上非天 然氨基酸，从而改进所关注的一种或一种以上性质。

[0213] 通过免疫反应性定义多肽

[0214] 因为本发明的多肽提供多种新颖的多肽序列 (例如，在本文中的翻译系统中合成

的蛋白质的情况下包含非天然氨基酸;或例如在本文中的新颖合成酶的情况下,包含标准氨基酸的新颖序列),所以多肽还提供可例如在免疫学检定中识别的新颖结构特征。特异性结合本发明多肽的抗体或抗血清的产生以及这些抗体或抗血清所结合的多肽都是本发明的特征。

[0215] 举例来说,本发明包括合成酶蛋白质,其特异性结合针对包含选自(SEQ ID NO:36-63 和/或 86)中一个或一个以上序列的氨基酸序列的免疫原所产生的抗体或抗血清或与所述抗体或抗血清特异性免疫反应。为消除与其它同源物的交叉反应性,利用诸如野生型大肠杆菌酪氨酰基合成酶(TyrRS)(例如 SEQ ID NO:2)等可利用的对照合成酶同源物来消减抗体或抗血清。

[0216] 在一种典型格式中,免疫检定使用针对一种或一种以上包含一个或一个以上对应于 SEQ ID NO:36-63 和/或 86 中一个或一个以上的序列或其实质性子序列(即,所提供的全长序列的至少约 30%)的多肽培养的多克隆抗血清。源自 SEQ ID NO:36-63 和 86 的潜在多肽免疫原组在下文中总称为"免疫原性多肽"。视情况选择所得抗血清以针对对照合成酶同源物具有低交叉反应性,且在免疫检定中使用多克隆抗血清之前,例如通过用一种或一种以上对照合成酶同源物进行免疫吸附来去除任何这种的交叉反应性。

[0217] 为产生用于免疫检定的抗血清,如本文中所述产生和纯化一种或一种以上免疫原性多肽。举例来说,可在重组细胞中产生重组蛋白。利用免疫原性蛋白以及标准佐剂(诸如弗氏佐剂(Freund′s adjuvant))和标准小鼠免疫方案(关于抗体产生、可用于确定特异性免疫反应性的免疫检定格式和条件的标准说明,例如参见 Harlow 和 Lane(1988) Antibodies,A Laboratory Manual,Cold Spring Harbor Publications,New York。本文中也可见抗体的其它参考文献和论述,且可将其用于此处以产生通过免疫反应性来定义/检测多肽的抗体),使小鼠的近交品系(因为结果因小鼠的实际遗传一致性而更可重复,所以用于这一检定中)免疫。或者,可将一种或一种以上源自本文中所公开的序列的合成或重组多肽与载体蛋白连接并用作免疫原。

[0218] 收集多克隆血清并且在免疫检定,例如,利用一种或一种以上在固体支撑物上固定的免疫原性蛋白的固相免疫检定中,针对免疫原性多肽进行滴定。选择效价为 10^6 或 10^6 以上的多克隆抗血清,汇集并利用对照合成酶多肽进行消减,以产生消减 的经汇集的经滴定多克隆抗血清。

[0219] 在比较免疫检定中,测试消减的经汇集的经滴定多克隆抗血清针对对照同源物的交叉反应性。在这一比较检定中,确定用于消减的经滴定多克隆抗血清的差别结合条件,这使得经滴定多克隆抗血清与免疫原性合成酶的结合的信噪比与经滴定多克隆抗血清与对照合成酶的结合相比高至少约 5-10 倍。也就是说,通过添加诸如白蛋白或脱脂奶粉等非特异性竞争剂和/或通过调节盐条件、温度等来调节结合/洗涤反应的严格度。后续检定中使用这些结合/洗涤条件来确定测试多肽(与免疫原性多肽和/或对照多肽相比的多肽)是否与所汇集的经消减的多克隆抗血清特异性结合。具体说来,在区别结合条件下显示信噪比比对照合成酶同源物高至少 2-5 倍且与免疫原性多肽相比显示至少约 1/2 信噪比的测试多肽与已知合成酶相比与免疫原性多肽共享实质性结构相似性,且因此,所述测试多肽为本发明的多肽。

[0220] 在另一实例中,将竞争性结合格式的免疫检定用于检测测试多肽。举例来说,如所

说明，通过利用对照多肽的免疫吸附从汇集的抗血清混合物中去除交叉反应性抗体。然后，使免疫原性多肽固定于暴露于消减的经汇集的抗血清的固体支撑物上。向检定中添加测试蛋白质以竞争与汇集的经消减抗血清结合。将与固定蛋白质相比，测试蛋白质竞争与所汇集的经消减抗血清结合的能力与添加到所述检定中的免疫原性多肽竞争结合的能力（所述免疫原性多肽有效地与固定的免疫原性多肽竞争与所汇集的抗血清结合）相比较。使用标准计算法计算测试蛋白质的交叉反应性百分比。

[0221] 在平行检定中，视情况与免疫原性多肽竞争与抗血清结合的能力相比较，确定对照蛋白质竞争与汇集的经消减抗血清结合的能力。同样使用标准计算法计算对照多肽的交叉反应性百分比。当测试多肽的交叉反应性百分比与对照多肽相比高至少 5-10 倍时，和/或当测试多肽的结合大致在免疫原性多肽的结合范围内时，认为测试多肽特异性结合汇集的经消减抗血清。

[0222] 一般来说，在如本文中所述的竞争性结合免疫检定中可使用经免疫吸附且汇集的抗血清来比较任何测试多肽与免疫原性和/或对照多肽。为进行这一比较，分别检定多种浓度的免疫原性、测试以及对照多肽且使用标准技术来确定抑制经消减抗血清与例如经固定对照、测试或免疫原性蛋白的结合的 50％所需的各多肽的量。如果竞争检定中结合所需的测试多肽的量比所需的免疫原性多肽的量少两倍，那么只要所述量为对照多肽的量的至少约 5-10 倍高，就认为测试多肽特异性结合针对免疫原性蛋白所产生的抗体。

[0223] 作为特异性的另一测定，视情况利用免疫原性多肽（而不是对照多肽）完全免疫吸附汇集的抗血清，直到可检测到极少所得免疫原性多肽-汇集的经消减抗血清与免疫吸附中所使用的免疫原性多肽的结合或无法检测到结合。然后，测试所述完全免疫吸附的抗血清与测试多肽的反应性。如果几乎未观察到反应性（即，不超过关于完全免疫吸附的抗血清与免疫原性多肽的结合所观察到的信噪比的 2 倍），那么测试多肽与免疫原性蛋白引发的抗血清特异性结合。

[0224] 医药组合物

[0225] 将本发明的多肽或蛋白质（例如合成酶、包含一个或一个以上非天然氨基酸的蛋白质等）视情况例如与合适的医药载剂组合用于治疗用途中。这些组合物例如包含治疗有效量的化合物和医药学上可接受的载剂或赋形剂。所述载剂或赋形剂包括（但不限于）生理盐水、缓冲生理盐水、葡萄糖、水、甘油、乙醇和/或其组合。进行调配以适合给药模式。一般来说，所属领域中熟知给予蛋白质的方法，且可将其用于给予本发明的多肽。

[0226] 根据所属领域中熟知的方法，视情况在疾病的一种或一种以上适当的活体外和/或活体内动物模型中测试包含一种或一种以上本发明的多肽的治疗组合物，从而证实功效、组织代谢并估算剂量。具体说来，最初可通过本文中非天然氨基酸同源物相对于天然氨基酸同源物（例如，比较经修饰以包括一个或一个以上非天然氨基酸的 EPO 与天然氨基酸 EPO）的活性、稳定性或其它合适的量度（即，在相关检定中）来确定剂量。

[0227] 给药是通过通常用于引导分子与血液或组织细胞紧密接触的任何途径来进行。本发明的非天然氨基酸多肽视情况与一种或一种以上医药学上可接受的载剂一起以任何合适的方式给予。可利用在本发明的情况下向患者给予所述多肽的合适方法，且虽然可使用一种以上途径来给予特定组合物，但特定途径通常可提供比另一途径更快捷且更有效的作用或反应。

[0228] 医药学上可接受的载剂是由所给予的特定组合物以及用于给予组合物的特定方法部分地决定。因此,存在多种本发明的医药组合物的合适调配物。

[0229] 多肽组合物可通过许多途径给予,所述途径包括(但不限于)口服、静脉内、腹膜内、肌肉内、经皮、皮下、局部、舌下或直肠方式。非天然氨基酸多肽组合物也可借助于脂质体给予。所属领域技术人员通常已知所述给药途径和适当的调配物。

[0230] 非天然氨基酸多肽也可单独或与其它合适组分组合制成欲经由吸入给予的气溶胶调配物(即,其可“成喷雾状”)。气溶胶调配物可放置在加压的可接受推进剂中,诸如二氯二氟甲烷、丙烷、氮等。

[0231] 适于不经肠给药(诸如通过关节内(关节中)、静脉内、肌肉内、真皮内、腹膜内和皮下途径)的调配物包括可含有抗氧化剂、缓冲剂、抑菌剂以及使调配物与预期接受者的血液等张的溶质的水性和非水性等张无菌注射液和可包括悬浮剂、增溶剂、增稠剂、稳定剂以及防腐剂的水性和非水性无菌悬浮液。可在诸如安瓿和小瓶的单位剂量或多剂量密封容器中提供经包装核酸的调配物。

[0232] 不经肠给药和静脉内给药是优选的给药方法。具体说来,已经用于天然氨基酸同源物治疗剂的给药途径(例如,通常用于 EPO、GCSF、GMCSF、IFN、白细胞介素、抗体和/或任何其它医药学上传递的蛋白质的给药途径)和目前使用的调配物提供包括本发明的非天然氨基酸的蛋白质(例如,目前的治疗蛋白的聚乙二醇化变异体等)的优选给药途径和调配物。

[0233] 在本发明的情况下,视应用而定,向患者给予的剂量足以随时间在患者体内实现有益的治疗反应,或例如抑制病原体感染或其它适当活性。通过特定组合物/调配物的功效和所使用的非天然氨基酸多肽的活性、稳定性或血清半衰期以及患者的状况以及欲治疗的患者的体重或表面积来确定剂量。剂量的大小也通过特定患者中给予特定组合物/调配物所伴随的任何不利副作用的存在、性质以及程度等来确定。

[0234] 在确定治疗或预防疾病(例如癌症、遗传疾病、糖尿病、AIDS 等)时欲给予的组合物/调配物的有效量时,医师评估循环血浆含量、调配物毒性、疾病的进展和/或相关情形下抗非天然氨基酸多肽抗体的产生。

[0235] 向例如 70 公斤患者给予的剂量通常在与针对相关组合物的活性或血清半衰期的改变所调整的目前使用的治疗蛋白的剂量等效的范围内。本发明的组合物/调配物可通过任何已知的常规疗法来补充治疗条件,所述常规疗法包括给予抗体、给予疫苗、给予细胞毒性剂、天然氨基酸多肽、核酸、核苷酸类似物、生物反应调节剂等。

[0236] 就给药来说,本发明的调配物是以由相关调配物的 LD-50 和/或(例如)当应用于患者的质量和总体健康状况时对各种浓度下的非天然氨基酸的任何副作用的观察结果所确定的速率给予。给药可借助于单次剂量或分次给药实现。

[0237] 如果经历调配物输注的患者出现发烧、寒战或肌肉疼痛,那么他/她接受适当剂量的阿司匹林 (aspirin)、布洛芬 (ibuprofen)、对乙酰氨基酚 (acetaminophen) 或其它疼痛/发烧控制药物。在将进行输注之前 30 分钟,对经历输注反应(诸如发烧、肌肉疼痛和寒战)的患者预先给予阿司匹林、对乙酰氨基酚或例如苯海拉明 (diphenhydramine)。将哌替啶 (Meperidine) 用于更为严重的不能快速响应解热剂和抗组胺剂的寒战和肌肉疼痛。治疗视反应的严重程度而减缓或中止。

[0238] 核酸和多肽序列及变异体

[0239] 如上文和下文中所述，本发明提供核酸多核苷酸序列和多肽氨基酸序列（例如O-tRNA和O-RS）以及例如包含所述序列的组合物和方法。本文中公开例如O-tRNA和O-RS的所述序列的实例（参见表5，例如SEQ ID NO.3-65、86，且SEQ ID NO：1和2除外）。然而，所属领域的技术人员应了解，本发明不受本文中所公开的序列（例如实例和表5）限制。所属领域的技术人员应了解，本发明也提供许多具有本文中所述的功能（例如编码O-tRNA或O-RS）的相关和甚至无关的序列。

[0240] 本发明也提供多肽（O-RS）和多核苷酸，例如O-tRNA、编码O-RS或其部分（例如合成酶的活性位点）的多核苷酸、用于构筑氨酰基tRNA合成酶突变体的寡核苷酸等。举例来说，发明的多肽包括：包含如SEQ ID NO：36-63和/或86中任一序列所述的氨基酸序列的多肽；包含由如SEQ ID NO：3-35中任一序列所述的多核苷酸序列编码的氨基酸序列的多肽；以及与对包含如SEQ ID NO：36-63和/或86中任一序列所示的氨基酸序列的多肽或包含由如SEQ ID NO：3-35中任一序列所示的多核苷酸序列编码的氨基酸序列的多肽具有特异性的抗体特异性免疫反应的多肽。

[0241] 本发明的多肽也包括包含与天然存在的酪氨酰基氨酰基tRNA合成酶（TyrRS）的氨基酸序列（例如SEQ ID NO：2）至少90％一致且包含两个或两个以上来自群组A-E的氨基酸的氨基酸序列的多肽。举例来说，群组A包括在对应于大肠杆菌TyrRS的Tyr37的位置处的缬氨酸、异亮氨酸、亮氨酸、甘氨酸、丝氨酸、丙氨酸或苏氨酸；群组B包括在对应于大肠杆菌TyrRS的Asn126的位置处的天冬氨酸；群组C包括在对应于大肠杆菌TyrRS的Asp182的位置处的苏氨酸、丝氨酸、精氨酸、天冬酰胺或甘氨酸；群组D包括在对应于大肠杆菌TyrRS的Phe183的位置处的蛋氨酸、丙氨酸、缬氨酸或酪氨酸；且群组E包括在对应于大肠杆菌TyrRS的Leu186的位置处的丝氨酸、蛋氨酸、缬氨酸、半胱氨酸、苏氨酸或丙氨酸。类似地，本发明的多肽还包括包含SEQ ID NO：36-63和/或86的至少20个相邻氨基酸和两个或两个以上如上文所指示的氨基酸取代的多肽。也包括包含任一上述多肽的保守变异体的氨基酸序列作为本发明的多肽。

[0242] 在一个实施例中，组合物包括本发明的多肽和赋形剂（例如缓冲剂、水、医药学上可接受的赋形剂等）。本发明也提供与本发明的多肽特异性免疫反应的抗体或抗血清。

[0243] 本发明中也提供多核苷酸。本发明的多核苷酸包括编码本发明的所关注的蛋白质或多肽或包括一个或一个以上选择密码子或两者的多核苷酸。举例来说，本发明 的多核苷酸例如包括：包含如SEQ ID NO：3-35、64-85中任一序列所述的核苷酸序列的多核苷酸；与其多核苷酸序列互补或编码其多核苷酸序列的多核苷酸；和/或编码包含如SEQ ID NO：36-63和/或86中任一序列中所述的氨基酸序列的多肽的多核苷酸，或其保守变异体。本发明的多核苷酸也包括编码本发明的多肽的多核苷酸。类似地，在核酸的实质上整个长度范围内在高度严格条件下与上文所指示的多核苷酸杂交的核酸为本发明的多核苷酸。

[0244] 本发明的多核苷酸也包括编码包含与天然存在的酪氨酰基氨酰基tRNA合成酶（TyrRS）的氨基酸序列（例如SEQ ID NO：2）至少90％一致且包含两个或两个以上如以上段落11中在群组A-E中所指示的突变的氨基酸序列的多肽的多核苷酸。本发明的多核苷酸中还包括与以上所指示的多核苷酸至少70％（或至少75％、至少80％、至少85％、至少90％、至少95％、至少98％或至少99％或99％以上）一致的多核苷酸，和/或包含以上所

指示的任何多核苷酸的保守变异体的多核苷酸。

[0245] 在某些实施例中，载体（例如质粒、粘粒、噬菌体、病毒等）包含本发明的多核苷酸。在一个实施例中，载体为表达载体。在另一实施例中，表达载体包括可操作性连接到本发明的一个或一个以上多核苷酸的启动子。在另一实施例中，细胞包含包括本发明的多核苷酸的载体。

[0246] 技术人员应了解，所公开的序列的许多变异体都包括在本发明中。举例来说，产生功能相同的序列的所公开序列的保守变异体包括在本发明中。认为核酸多核苷酸序列的变异体包括在本发明内，其中所述变异体与至少一个所公开的序列杂交。如通过例如标准序列比较技术所确定的本文中所公开的序列的独特子序列也包括在本发明中。

[0247] 保守变异体

[0248] 由于遗传密码的简并性，所以“沉默取代”（即，不会在编码多肽中产生改变的核酸序列中的取代）是编码氨基酸的每个核酸序列的暗含特征。类似地，“保守氨基酸取代”（即，氨基酸序列中一个或一个以上氨基酸经具有极其类似性质的不同氨基酸取代）也易于鉴别为与所公开的构筑体高度相似。各公开序列的所述保守变异体是本发明的特征。

[0249] 特定核酸序列的“保守变异体”是指编码一致或基本上一致的氨基酸序列的核酸，或当核酸不编码氨基酸序列时是指基本上一致的序列。所属领域技术人员应认识到，改变、添加或缺失编码序列中的单一氨基酸或少量氨基酸（通常小于5%，更通常小于4%、2%或1%）的个别取代、缺失或添加为“保守性修饰变异”，其中 所述改变将导致氨基酸的缺失、氨基酸的添加或化学上类似的氨基酸对氨基酸的取代。因此，本发明所列出的多肽序列的“保守变异”包括用具有相同保守性取代基团的保守性选择的氨基酸取代多肽序列的少量、通常少于5%、更通常少于2%或1%的氨基酸。最后，不改变核酸分子的编码活性的序列的添加（诸如，非功能性序列的添加）为基本核酸的保守变异。

[0250] 所属领域中熟知提供功能类似的氨基酸的保守性取代表。下表陈述含有包括彼此互为“保守性取代”的天然氨基酸的例示性基团。

[0251] 保守性取代基团

[0252]

1	丙氨酸 (A)	丝氨酸 (S)	苏氨酸 (T)	
2	天冬氨酸 (D)	谷氨酸 (E)		
3	天冬酰胺 (N)	谷氨酰胺 (Q)		
4	精氨酸 (R)	赖氨酸 (K)		
5	异亮氨酸 (I)	亮氨酸 (L)	蛋氨酸 (M)	缬氨酸 (V)
6	苯丙氨酸 (F)	酪氨酸 (Y)	色氨酸 (W)	

[0253] 核酸杂交

[0254] 可使用比较杂交来鉴别本发明的核酸，包括本发明的核酸的保守变异体，且这种比较杂交方法是区分本发明的核酸的优选方法。另外，在高、超高和极高严格条件下与由SEQ ID NO：3-35、64-85 表示的核酸杂交的标靶核酸是本发明的特征。所述核酸的实例包括与给定核酸序列相比具有一个或一个以上沉默或保守性核酸取代的核酸。

[0255] 当测试核酸以与最佳匹配的互补标靶杂交的至少 1/2 水平与探针杂交时，即，以高达探针与标靶在一定条件下（其中最佳匹配的探针以关于与任何不匹配标靶核酸杂交所观察到的信噪比的至少约 5-10 倍高的信噪比与最佳匹配的互补标靶结合）杂交的信噪比的至少 1/2 信噪比杂交时，认为所述测试核酸与探针核酸特异性杂交。

[0256] 当核酸缔合时,其通常在溶液中"杂交"。核酸因多种明确表征的物理化学力而杂交,所述力诸如氢键、溶剂排阻、碱基堆积等。核酸杂交的详细指南可见于 Tijssen(1993) Laboratory Techniques in Biochemistry and Molecular Biology-Hybridizationwith Nucleic Acid Probes 第 I 部分第 2 章,"Overview of principles of hybridization andthe strategy of nucleic acid probe assays,"(Elsevier,New York) 以及 Ausubel,同上文中。Hames 和 Higgins(1995)Gene Probes 1 IRL Press at Oxford University Press,Oxford,England,(Hames 和 Higgins 1) 以及 Hames 和 Higgins(1995)Gene Probes 2 IRL Press at Oxford University Press, Oxford, England(Hames 和 Higgins 2) 提供关于包括寡核苷酸的 DNA 和 RNA 的合成、标记、检测以及定量的细节。

[0257] 在 Southern 或 Northern 印迹法中在过滤器上具有超过 100 个互补残基的互补核酸的杂交的严格杂交条件的实例为在 42℃下,50%福尔马林 (formalin) 和 1mg 肝素,其中杂交进行整夜。严格洗涤条件的实例为 0.2×SSC 洗涤,在 65℃下,15 分钟 (关于 SSC 缓冲剂的描述,参见 Sambrook,同上文)。通常,在高严格度洗涤后为低严格度洗涤,以去除背景探针信号。例示性低严格度洗涤为 2×SSC,在 40℃下,15 分钟。比特定杂交检定中关于不相关探针所观察到的信噪比高 5 倍 (或更高) 的信噪比通常指示检测到特异性杂交。

[0258] 在诸如 Southern 和 Northern 杂交等核酸杂交实验的情况下,"严格杂交洗涤条件"具有序列依赖性且在不同环境参数下不同。核酸杂交的全面指南可见于 Tijssen(1993),同上文以及 Hames 和 Higgins,1 和 2 中。任何测试核酸的严格杂交和洗涤条件都可凭经验容易地确定。举例来说,在确定高度严格杂交和洗涤条件的过程中,使杂交和洗涤条件逐渐增加 (例如,通过增加温度、降低盐浓度、增加洗涤剂浓度和 / 或增加杂交或洗涤中诸如福尔马林等有机溶剂的浓度),直到符合一组选定的标准。举例来说,使杂交和洗涤条件逐步增加,直到探针与最佳匹配的互补标靶以高达探针与不匹配标靶杂交所观察到的信噪比的至少 5 倍的信噪比结合。

[0259] 选择"极严格"条件以与特定探针的热熔点 (T_m) 等效。T_m是 50%的测试序列与最佳匹配的探针杂交的温度 (在规定的离子强度和 pH 值下)。出于本发明的目的,通常选择"高度严格"杂交和洗涤条件比规定的离子强度和 pH 值下特定序列的 T_m低约 5℃。

[0260] "超高严格度"杂交和洗涤条件是增加杂交和洗涤条件的严格度,直到探针与最佳匹配的互补标靶核酸结合的信噪比高达探针与任何不匹配的标靶核酸杂交所观察到的信噪比的至少 10 倍的条件。认为在所述条件下以最佳匹配的互补标靶核酸的信噪比的至少 1/2 的信噪比与探针杂交的标靶核酸在超高严格条件下与探针结合。

[0261] 类似地,可通过逐步增加相关杂交检定的杂交和 / 或洗涤条件来确定甚至更高的严格水平。举例来说,增加杂交和洗涤条件的严格度,直到探针与最佳匹配的互补标靶核酸结合的信噪比高达探针与任何不匹配的标靶核酸杂交所观察到的信噪比的至少 10 倍、20 倍、50 倍、100 倍或 500 倍或 500 倍以上的条件。认为在所述条件下以最佳匹配的互补标靶核酸的信噪比的至少 1/2 的信噪比与探针杂交的标靶核酸在极高严格条件下与探针结合。

[0262] 如果在严格条件下彼此不杂交的核酸编码的多肽实质上一致,那么所述核酸也实质上一致。例如当使用遗传密码所允许的最大密码子简并性产生核酸复本时,出现这种情形。

[0263] 独特子序列

[0264] 一方面,本发明提供包含选自本文中所公开的 O-tRNA 和 O-RS 的序列的核酸中的独特子序列的核酸。独特子序列与对应于任何已知的 O-tRNA 或 O-RS 核酸序列的核酸相比具有独特性。可使用例如设定为默认参数的 BLAST 进行比对。任何独特的子序列例如都适合用作鉴别本发明核酸的探针。

[0265] 类似地,本发明包括包含选自本文中所公开的 O-RS 的序列的多肽中的独特子序列的多肽。此处,独特子序列与对应于任何已知的多肽序列的多肽相比具有独特性。

[0266] 本发明也提供在严格条件下与编码选自 O-RS 的序列的多肽中的独特子序列的独特编码寡核苷酸杂交的标靶核酸,其中独特子序列与对应于任何对照多肽(例如,例如通过突变得到本发明的合成酶的亲本序列)的多肽相比具有独特性。独特序列是如上文所述确定。

[0267] 序列比较、一致性以及同源性

[0268] 在两个或两个以上核酸或多肽序列的情形中,术语"一致"或"一致性"百分比是指当比较和比对最大对应性时,如使用下文所述的序列比较算法中的一种(或所属领域技术人员可用的其它算法)或通过目测所测量,相同或具有特定的相同氨基酸残基或核苷酸百分比的两个或两个以上序列或子序列。

[0269] 在两个核酸或多肽(例如,编码 O-tRNA 或 O-RS 的 DNA 或 O-RS 的氨基酸序列)的情况下,短语"实质上一致"是指当比较和比对最大对应性时,如使用序列比较算法或通过目测所测量,具有至少约 60%、优选 80%、最优选 90-95%的核苷酸或氨基酸残基一致性的两个或两个以上序列或子序列。在不提及实际祖先的情况下,通常认为所述"实质上一致"的序列是"同源"的。"实质一致性"优选存在于长度为至少约 50 个残基的序列的区内,更优选至少约 100 个残基的区内,且最优选,所述序列在至少约 150 个残基内或在欲比较的两个序列的全长内实质上一致。

[0270] 为进行序列比较和同源性测定,通常使一个序列充当与测试序列相比较的参考序列。当使用序列比较算法时,将测试序列与参考序列输入计算机中,必要时指定子序列坐标,且指定序列算法程序参数。然后,序列比较算法基于指定的程序参数计算测试序列相对于参考序列的序列一致性百分比。

[0271] 可例如通过 Smith 和 Waterman, Adv.AppI.Math.2 :482(1981) 的局部同源算法,通过 Needleman 和 Wunsch, J.Mol.Biol.48 :443(1970) 的同源性比对算法,通过搜索 Pearson 和 Lipman, Proc.Nat' l.Acad.Sci.USA 85 :2444(1988) 的相似方法,通过这些算法 (Wisconsin Genetics Software Package,Genetics Computer Group,575 ScienceDr.,Madison,WI 中的 GAP、BESTFIT、FASTA 以及 TFASTA) 的计算机实施或通过目测检查(通常参见 Ausubel 等人,同下文)来进行用于比较的序列的最佳比对。

[0272] 适于测定序列一致性百分比和序列相似性的算法的一个实例为 BLAST 算法,其描述于 Altschul 等人, J.Mol.Biol.215 :403-410(1990) 中。执行 BLAST 分析的软件可通过美国生物技术信息中心 (National Center for Biotechnology Information, www.ncbi.nlm.nih.gov/) 公开获得。所述算法涉及首先通过鉴别询问序列中长度 W 的短字来鉴别高得分的序列对 (high scoring sequence pair, HSP),当与数据库序列中相同长度的字比对时,所述 HSP 匹配或满足一些正值临界得分 T。T 是指邻近字的临界分值 (neighborhood word score threshold)(Altschul 等人,同上)。这些初始邻近字匹配 (word hit) 充当起

始搜索以发现含有其的较长 HSP 的种子。然后,使字匹配沿各序列的两个方向延伸直到可增加累积比对的分值。对核苷酸序列使用参数 M(一对匹配残基的奖励分值;通常＞ 0) 和 N(错配残基的处罚分值,通常＜ 0) 计算累积分值。对于氨基酸序列,使用得分矩阵计算累积分值。当累积比对分值比其所达到的最大值低数量 X;累积分值因一个或一个以上负得分残基比对的积累而为零或更低值;或到达各序列的末端时,在各个方向上的字匹配延伸停止。BLAST 算法参数 W、T 和 X 将决定比对的灵敏性和速度。BLASTN 程序(对于核苷酸序列)使用字长 (W) 为 11、期望值 (expectation,E) 为 10、截止值为 100、M = 5、N = -4 和两条链的比较作为默认值。对于氨基酸序列来说,BLASTP 程序使用字长 (W) 为 3、期望值 (E) 为 10 和 BLOSUM62 计分矩阵(参见 Henikoff 和 Henikoff(1989) Proc.Natl.Acad.Sci.USA 89:10915) 作为默认值。

[0273] 除计算序列一致性百分比外,BLAST 算法也进行两个序列之间相似性的统计分析(例如参见 Karlin 和 Altschul,Proc.Nat'1.Acad.Sci.USA 90:5873-5787(1993))。由 BLAST 算法提供的一种相似性量度是最小和概率 (smallest sum probability,P(N)),其提供两个核苷酸或氨基酸序列之间偶然会发生匹配的概率的指示。举例来说,如果在测试核酸与参考核酸的比较中,最小和概率小于约 0.1、更优选小于约 0.01 且最优选小于约 0.001,那么认为核酸与参考序列相似。

[0274] 诱变和其它分子生物学技术

[0275] 描述分子生物技术的一般文章包括 Berger 和 Kimmel,Guide to MolecularCloning Techniques.Methods in Enzymology 第 152 卷 Academic Press,Inc.,SanDiego,CA(Berger);Sambrook 等人,Molecular Cloning-A Laboratory Manual(第 2 版)第 1-3 卷.Cold Spring Harbor Laboratory,Cold Spring Harbor,New York,1989("Sambrook")以及 Current Protocols in Molecular Biology.F.M.Ausubel 等人编,Current Protocols,Greene Publishing Associates,Inc. 与 John Wiley&Sons,Inc. 合资 (1999 增刊)("Ausubel")。这些文章描述诱变、载体的使用、质粒和 λ 噬菌体的 DNA 产生、启动子以及许多其它相关主题,所述相关主题涉及例如包括产生包括非天然氨基酸、正交 tRNA、正交合成酶以及其对的蛋白质的选择密码子的基因的产生。

[0276] 本发明中使用各种类型的诱变,从而例如产生 tRNA 文库、产生合成酶文库、在所关注的蛋白质或多肽中插入编码非天然氨基酸的选择密码子。其包括(但不限于)定点诱变、随机点诱变、同源重组、DNA 改组或其它递归诱变方法、嵌合构筑、使用含尿嘧啶的模板的诱变、寡核苷酸定向诱变、硫代磷酸酯修饰 DNA 诱变、使用缺口双螺旋 DNA 的诱变等,或其任何组合。其它合适方法包括点错配修复、使用修复缺陷型宿主品系的诱变、限制选择和限制纯化、缺失诱变、通过总基因合成进行的诱变、双链断裂修复等。举例来说,涉及嵌合构筑体的诱变也包括在本发明内。在一个实施例中,可通过天然存在的分子或已改变或突变的天然存在的分子的已知信息(例如序列、序列比较、物理性质、晶体结构等)指导诱变。

[0277] 上述文章和本文中可见的实例描述这些程序。其它信息可见于下列公开文献和其中所引用的参考文献中:Ling 等人,Approaches to DNA mutagenesis:an overview,Anal Biochem.254(2):157-178(1997);Dale 等人,Oligonucleotide-directed randommutagenesis using the phosphorothioate method,Methods Mol.Biol.57:369-374(1996);Smith,In vitro mutagenesis,Ann.Rev.Genet.19:423-462(1985);

Botstein 和 Shortle, Strategies and applications of in vitro mutagenesis, Science229 :1193-1201(1985) ;Carter, Site-directed mutagenesis, Biochem. J. 237 : 1-7(1986) ;Kunkel, The efficiency of oligonucleotide directed mutagenesis, in Nucleic Acids&Molecular Biology(Eckstein,F. 和 Lilley,D.M.J 编,Springer Verlag, Berlin))(1987) ;Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection,Proc.Natl.Acad.Sci.USA 82 :488-492(1985) ;Kunkel等人,Rapid and efficient site-specificmutagenesis without phenotypic selection,Methods in Enzymol. 154,367-382(1987) ;Bass等人,Mutant Trp repressors with new DNA-binding specificities,Science 242 :240-245(1988) ;Methods in Enzymol. 100 :468-500(1983) ; Methods in Enzymol. 154 :329-350(1987) ;Zoller 和 Smith, Oligonucleotide-directed mutagenesis usingM13-derived vectors :an efficient and general procedure for the production of pointmutations in any DNA fragment, Nucleic Acids Res. 10 : 6487-6500(1982) ;Zoller 和 Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100 :468-500(1983) ; Zoller 和 Smith, Oligonucleotide-directedmutagenesis :a simple method using two oligonucleotide primers and a single-strandedDNA template, Methods in Enzymol. 154 :329-350(1987) ;Taylor 等 人, The use ofphosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl.Acids Res. 13 : 8749-8764(1985) ;Taylor 等 人, The rapid generation ofoligonucleotide-directed mutations at high frequency using phosphorothioate-modifiedDNA, Nucl. Acids Res. 13 :8765-8787(1985) ;Nakamaye 和 Eckstein, Inhibition ofrestriction endonuclease Nci I cleavage by phosphorothioate groups and its applicationto oligonucleotide-directed mutagenesis, Nucl.Acids Res. 14 :9679-9698(1986) ; Sayers 等 人, Y-T Exonucleases in phosphorothioate-based oligonucleotide-directedmutagenesis, Nucl.Acids Res. 16 :791-802(1988) ;Sayers 等 人, Strand specific cleavageof phosphorothioate-containing DNA by reaction with restriction endonucleases in thepresence of ethidium bromide, (1988) Nucl.Acids Res. 16 :803-814 ;Kramer 等 人, Thegapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl.Acids Res. 12 : 9441-9456(1984) ;Kramer 和 Fritz Oligonucleotide-directed constructionof mutations via gapped duplex DNA, Methods in Enzymol. 154 :350-367(1987) ; Kramer 等 人, Improved enzymatic in vitro reactions in the gapped duplex DNA approach tooligonucleotide-directed construction of mutations, Nucl. Acids Res. 16 :7207(1988) ;Fritz 等 人, Oligonucleotide-directed construction of mutations :a gapped duplex DNAprocedure without enzymatic reactions in vitro, Nucl.Acids Res. 16 :6987-6999(1988) ;Kramer 等 人, Point Mismatch Repair, Cell 38 :879-887(1984) ;Carter 等 人, Improvedoligonucleotide site-directed mutagenesis using MIS vectors, Nucl.Acids Res. 13 :4431-4443(1985) ;Carter, Improved oligonucleotide-directed mutagenesis using M13vectors, Methods in

Enzymol. 154 :382-403(1987) ;Eghtedarzadeh 和 Henikoff, Use ofoligonucleotides to generate large deletions, Nucl. Acids Res. 14 :5115(1986) ;Wells 等 人, Importance of hydrogen-bond formation in stabilizing the transition state ofsubtilisin, Phil. Trans. R. Soc. Lond. A 317 :415-423(1986) ;Nambiar 等 人, Total synthesis and cloning of a gene coding for the ribonuclease Sprotein, Science 223 :1299-1301(1984) ;Sakamar 和 Khorana, Total synthesis and expression of a gene for the a-subunit of bovinerod outer segment guanine nucleotide-binding protein(transducin), Nucl. Acids Res. 14 :6361-6372(1988) ;Wells 等 人, Cassette mutagenesis :an efficient method for generationof multiple mutations at defined sites, Gene 34 :315-323(1985) ;Grundstrom 等人, Oligonucleotide-directed mutagenesis by microscale ′ shot-gun ′ gene synthesis, Nucl. Acids Res. 13 :3305-3316(1985) ;Mandecki, Oligonucleotide-directed double-strandbreak repair in plasmids of Escherichia coli :a method for site-specific mutagenesis, Proc. Natl. Acad. Sci. USA. 83 :7177-7181(1986) ;Arnold, Protein engineering forunusual environments, Current Opinion in Biotechnology 4 :450-455(1993) ;Sieber 等 人, Nature Biotechnology, 19 :456-460(2001) ;W. P. C. Stemmer, Nature 370, 389-91(1994) ;以及 I. A. Lorimer, I. Pastan, Nucleic Acids Res. 23, 3067-8(1995)。关于许多上述方法的其它细节可见于 Methods in Enzymology 第 154 卷中,所述文献还描述各种诱变方法对故障查找问题的有效控制。

[0278] 本发明也涉及借助于正交 tRNA/RS 对在活体内并入非天然氨基酸的脊椎动物宿主细胞和生物体。利用本发明的多核苷酸或包括本发明的多核苷酸的构筑体(例如本发明载体,其可例如为克隆载体或表达载体)使宿主细胞遗传工程化(例如转化、转导或转染)。例如,载体可呈质粒、细菌、病毒、裸多核苷酸或结合多核苷酸的形式。通过标准方法将载体引入细胞和/或微生物,所述标准方法包括电穿孔 (From 等人, Proc. Natl. Acad. Sci. USA 82, 5824(1985))、通过病毒载体感染、由具有核酸的小粒子在小珠粒或粒子的基质内或表面上高速弹道渗透 (Klein 等人, Nature 327, 70-73(1987))。

[0279] 可在经改良适用于诸如筛选步骤、激活启动子或选择转化株等活动的常规营养培养基中培养经工程化的宿主细胞。视情况,可将这些细胞培养到转基因有机体中。关于例如细胞分离和培养(例如用于后续的核酸分离)的其它有用参考文献包括 Freshney(1994) Culture of Animal Cells, a Manual of Basic Technique, 第 3 版, Wiley-Liss, New York 和其中所引用的参考文献 ;Payne 等人 (1992) Plant Cell andTissue Culture in Liquid Systems John Wiley&Sons, Inc. New York, NY ;Gamborg 和 Phillips(编)(1995) Plant Cell, Tissue and Organ Culture;Fundamental MethodsSpringer Lab Manual, Springer-Verlag(Berlin Heidelberg New York) 以及 Atlas 和 Parks(编) The Handbook of Microbiological Media(1993) CRC Press, Boca Raton, FL。

[0280] 本发明还涉及具有通过正交 tRNA/RS 对的并入一个或一个以上非天然氨基酸的能力的脊椎动物细胞系。可使用所属领域中已知的细胞培养技术在已经本发明的多核苷酸或包括本发明的多核苷酸的构筑体转化、转导或转染的宿主细胞上建立这些细胞系。将外源核酸引入宿主细胞中的方法已为所属领域中熟知且将随所使用的宿主细胞而变化。技术

包括（但不限于）葡聚糖介导的转染、磷酸钙沉淀、氯化钙处理、聚凝胺(polybrene)介导的转染、原生质体融合、电穿孔、病毒或噬菌体感染、脂质体中多核苷酸的囊封以及定向微注射。

[0281] 可以允许瞬时或稳定并入 DNA 的方式转化或转染细胞。对于重组蛋白的长期、高产率产生来说，优选稳定表达。举例来说，可使稳定表达抗体分子的细胞系进行工程化。除使用含有病毒复制起点的表达载体外，可用通过适当表达控制元件（例如，启动子、增强子、序列、转录终止子、多聚腺苷酸位点等）控制的 DNA 和可选择标记转化宿主细胞。在引入外来 DNA 后，可使经工程化的细胞在富集培养基中生长 1-2 天，且然后转换为选择性培养基。重组质粒中的可选择标记赋予对选择的抗性，并且使细胞能将质粒稳定整合到其染色体组中并生长，从而形成细胞群，又可对其进行克隆并扩充成细胞系。可有利地使用所述方法来工程化表达抗体分子的细胞系。所述经工程化的细胞系尤其适用于筛选和评估与抗体分子直接或间接相互作用的化合物。另外，诸如一些病毒介导的载体转染技术等所属领域技术人员熟知的其它技术可允许瞬时转染细胞。

[0282] 可使用将标靶核酸引入细胞中的数种熟知的方法，其中任一种都可用于本发明中。这些方法包括：接受体细胞与含有 DNA 的细菌原生质体融合、电穿孔、基因枪法(projectile bombardment)和经病毒载体（在下文中进一步论述）感染等。细菌细胞可用于扩增含有本发明的 DNA 构筑体的质粒的数目。使细菌生长到对数生长期且可通过所属领域中已知的多种方法分离细菌中的质粒（例如参见 Sambrook）。另外，可购得试剂盒以从细菌中纯化质粒（例如参见都来自 Pharmacia Biotech 的 EasyPrep™、FlexiPrep™；来自 Stratagene 的 StrataClean™；以及来自 Qiagen 的 QIAprep ™）。然后，进一步操纵经分离和纯化的质粒以产生其它质粒，将其用于转染细胞或并入相关载体中以感染生物体。典型载体含有转录和翻译终止子、转录和翻译起始序列以及可用于调控特定标靶核酸的表达的启动子。载体视情况包含通用表达盒，其含有至少一个独立终止子序列、允许表达盒在真核细胞或原核细胞或两者中复制的序列（例如，穿梭载体(shuttle vector)）以及用于原核系统与脊椎动物系统的选择标记。载体适用于在原核细胞、真核细胞或优选两者中复制和整合。参见 Gillam 和 Smith，Gene 8：81(1979)；Roberts 等人，Nature.328：731(1987)；Schneider,B. 等人，Protein Expr.Purif.6435：10(1995)；Ausubel,Sambrook,Berger（都同上文）。举例来说，ATCC（例如，由 ATCC 出版的 The ATCC Catalogue of Bacteria andBacteriophage(1992)Gherna 等人（编））提供可用于克隆的细菌和噬菌体的目录。用于测序、克隆和分子生物学的其它方面的其它基本程序以及基础理论考虑也见于 Watson 等人，(1992)Recombinant DNA第 2版 Scientific American Books，NY。另外，基本上任何核酸（和几乎任何标记核酸，无论标准或非标准）都可从多种商业来源中的任一种定制或标准定购，这些商业来源诸如 Midland Certified ReagentCompany(Midland，TX，mcrc.com)、The Great American Gene Company(Ramona，CA，可在万维网的 genco.com上获得)、ExpressGen Inc.(Chicago，IL，可在万维网的 expressgen.com上获得)、Operon Technologies Inc.(Alameda，CA) 和许多其它来源。

[0283] 试剂盒

[0284] 试剂盒也是本发明的特征。举例来说，提供在细胞中产生包含至少一个非天然氨基酸的蛋白质的试剂盒，其中所述试剂盒包括含有编码 O-tRNA 的多核苷酸序列和／或

O-tRNA 和 / 或编码 O-RS 的多核苷酸序列和 / 或 O-RS 的容器。在一个实施例中，试剂盒另外包括至少一个非天然氨基酸。在另一实施例中，试剂盒另外包含用于产生蛋白质的说明材料。

[0285] 实例

[0286] 提供以下实例来说明而非限制所主张的本发明。所属领域技术人员将认识到，在不脱离所主张的本发明的范围的情况下，可改变多个不重要的参数。

[0287] 实例 1 :产生在脊椎动物细胞中并入非天然氨基酸的氨酰基 tRNA 合成酶的方法和所述氨酰基 tRNA 合成酶的组合物

[0288] 扩充脊椎动物遗传密码以包括具有新颖物理、化学或生物性质的非天然氨基酸，将提供用于分析和控制这些细胞中的蛋白质功能的有效工具。为此，描述用于分离氨酰基 tRNA 合成酶的一般方法，所述氨酰基 tRNA 合成酶响应酿酒酵母 (Saccharomyces cerevisiae，S. cerevisiae) 中的琥珀密码子以高保真性并入非天然氨基酸。所述方法是建立在通过抑制 GAL4 的 DNA 结合域与转录激活域之间的琥珀密码子来激活 GAL4 反应性报告基因 HIS3、URA3 或 LacZ 的基础上。描述用于阳性选择活性大肠杆菌酪氨酰基 tRNA 合成酶 (EcTyrRS) 变异体的 GAL4 报告基因的优化。还曾利用作为“有毒等位基因 (toxic allele)”添加到培养基中的小分子，利用 URA3 报告基因来进行无活性 EcTyrRS 变异体的阴性选择。重要的是，阳性选 择和阴性选择都可以多种严格度对单个细胞进行。这可有利于从大型突变合成酶文库中分离出多种氨酰基 tRNA 合成酶 (aaRS) 活性。用于分离所需 aaRS 表现型的方法的功效可通过模型选择来证实。

[0289] 实例 2

[0290] 哺乳动物细胞中大肠杆菌 Tyr tRNA 的转录

[0291] 随着基于大肠杆菌 Tyr-RS 的肽文库的酵母基遗传选择的发展，有可能分离对将非天然氨基酸并入真核细胞中的蛋白质中的能力具有选择性的突变合成酶。鉴于酵母与哺乳动物细胞之间的同源性，预想突变合成酶将支持哺乳动物细胞中的非天然氨基酸抑制。然而，Yokoyama 等人为达成此目的所进行的最初尝试碰到了关于在哺乳动物细胞中转录大肠杆菌 Tyr tRNA 的能力的问题。真核 tRNA 的转录是由称为 A 盒和 B 盒的启动子驱动 :这些启动子在 tRNA 序列自身的内部。A 盒和 B 盒序列的其它实例可见于 Geiduschek，(1988)，Transcription By RNA Polymerase III，Ann. Rev. Biochem. 57 :873-914 中，所述文献以引用的方式并入本文中。虽然 Yokoyama 尝试使 tRNA 的 A 盒序列工程化以使其能够活体内转录，但未检测到相应的琥珀抑制产物。

[0292] 在这一实例中，使用两个替代的 pol III 转录流程成功地转录琥珀抑制性大肠杆菌 $tRNA^{Tyr}$。在第一个流程中，使琥珀抑制性大肠杆菌 $tRNA^{Tyr}$突变体与哺乳动物 tRNA 基因 (具有适当的 5′ 和 3′ 侧接序列的人类 $tRNA^{Tyr}$) 的 3′ 端融合。在这两个 tRNA 之间不放置 pol III 终止子。主要的哺乳动物 tRNA 充当启动 pol III 转录的“II 型 polIII 前导启动子 (TIILP)”，所述 pol III 转录持续到通过琥珀抑制性大肠杆菌 $tRNA^{Tyr}$基因。如所得转录物支持通过选择密码子抑制进行非天然氨基酸并入的能力所证明，这一转录物产生哺乳动物细胞中所产生的功能性琥珀抑制性大肠杆菌 $tRNA^{Tyr}$。

[0293] 从具有适当的 5′ 和 3′ 侧接序列的人类 $tRNA^{Tyr}$的转录启动得出，也可转录数个琥珀抑制性大肠杆菌 tRNA 簇。琥珀抑制性大肠杆菌 tRNA 的团簇产生改进的非天然氨基酸

抑制。(图 1)

[0294] 在第二个流程中，琥珀抑制性大肠杆菌的转录是由诸如 U6 或 H1 启动子等 III 型 pol III 转录启动子启动。使用这一流程来评估 5′ 和 3′ 侧接序列以及团簇的琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 的不同 tRNA 盒的效应。(图 2)

[0295] 实验：

[0296] 构筑含有 TIILP 转录盒的质粒

[0297] 借助于 PCR 产生编码单一复本的琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 突变体的 DNA 序列。这一插入物从 5′ 到 3′ 的方向包括 5′ 限制位点 (EcoR I 和 Bgl II)、人类 Tyr tRNA5′ 侧接序列 (CTGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTC(SEQ IDNO:87))、琥珀抑制性大肠杆菌 $tRNA^{Tyr}$、3′ 侧接人类 $tRNA^{Tyr}$ 序列 (GACAAGTGCGGTTTTTTTCTCCAGCTCCCGATGACTTATGGC(SEQ ID NO:88)) 以及 3′ 限制位点 (BamH I 和 Hind III)。使用下列 PCR 引物从已含有 5′ 侧接人类 $tRNA^{Tyr}$ 序列、琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 以及 3′ 侧接人类 $tRNA^{Tyr}$ 序列的模板扩增 DNA 序列：

[0298] FTam97 短序列

[0299] GTACGAATTCCCGAGATCTCTGTGCTGAACCTCAGGGGACGC(SEQ ID NO:89)；

[0300] FTam109

[0301] GATGCAAGCTTGATGGATCCGCCATAAGTCATCGGGAGCTGGAGAAAAAAACCGCACTTGTCTGGTGGGGGAAGGATTCG(SEQ ID NO:90)。消化 (EcoR I 和 Hind III) 所得 PCR 产物，并使其与经相同酶消化的 pUC19 质粒 DNA 连接。

[0302] 消化 (EcoR I 和 Bgl II) 所得质粒并使其与含有人类 $tRNA^{Tyr}$ 基因的插入物 (在 EcoRI 和 Bam H1 处切割) 连接，所述人类 $tRNA^{Tyr}$ 基因含有 5′ 限制位点 (EcoR I 和 Bgl II)、人类 5′ 侧接 $tRNA^{Tyr}$ 序列 (GGATTACGCATGCTCAGTGCAATCTTCGGTTGCCTGGACTAGCGCTCCGGTTTTTCTGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTC(SEQ ID NO:91))、人类 $tRNA^{Tyr}$、人类 3′ 侧接序列 (GACAAGTGCGG(SEQ ID NO:92)) 以及 3′ 限制位点 (BamH I 和 Hind III)。

[0303] 为构筑编码人类 $tRNA^{Tyr}$ 和大肠杆菌 $tRNA^{Tyr}$ 转录盒的两个级联复本的质粒，消化 (EcoR I 和 Bgl II 限制酶) 所得质粒。通过 PCR 使用引物 FT73-out-new(CTTTGTGTAATACTTGTAACGCTGAATTC(SEQ ID NO:93)) 和 FT76-out 反向引物 (ACCATGATTACGCCAAGCTTGAT SEQ ID NO:94)) 来扩增转录盒。消化 (EcoRI 和 BamH I)PCR 产物并使其与切割质粒连接。

[0304] 通过遵循这种策略，构筑多种质粒 (图 1)。

[0305] 构筑驱动大肠杆菌 tRNA 表达质粒的 H1 启动子

[0306] 使用以上所构筑的大肠杆菌 $tRNA^{Tyr}$ 琥珀抑制性突变体作为合成不同 DNA 克隆插入物的模板，然后如流程 2 中所示用不同 5′ 和 3′ 序列侧接所述插入物 (图 2)。用于 PCR 反应的引物是：

[0307] 形式 1：FTam111(SEQ ID NO:95)/FTam113(SEQ ID NO:98)

[0308] 形式 1a：FTam111(SEQ ID NO:95)/FTam13(SEQ ID NO:98)

[0309] 形式 2：FTam102a(SEQ ID NO:97)/FTam113(SEQ ID NO:98)

[0310] 形式 3：FTam111(SEQ ID NO:95)/FTam114(SEQ ID NO:99)

[0311] 形式 4：FTam102a(SEQ ID NO:97)/FTam114(SEQ ID NO:99)

[0312] FTam111：GCATCGGATCCGGTGGGGTTCCCGAGCGGCC(SEQ ID NO:95)

[0313] FTam112 :ACGCCAAGCTTTTCCAAAATGGTGGGGAAGGATTCGAACCTTC(SEQ ID NO :96)

[0314] FTam102a :GCATCGGATCCGTGCTGAACCTCAGGGGACGCCG(SEQ ID NO :97)

[0315] FTam 113 :

[0316] ACGCCAAGCTTTTCCAAAAAATGGTGGGGAAGGATTCGAACCTTC(SEQ IDNO :98)

[0317] FTam 114 :

[0318] ACGCCAAGCTTTTCCAAAAAACCGCACTTGTCTGGTGGGGAAGG(SEQ ID NO :99)。

[0319] 消化 (BamH I 和 Hind III)、纯化 (PCR 纯化试剂盒 :Qiagen) 插入物并使其与经相同酶消化的 pSilenser 载体 (Ambion) 连接 (图 2)。图 2 的完全构筑体的序列如下 :

[0320] 具有 Tx4 终止子的无 5′ 和 3′ 侧接序列的大肠杆菌形式 1 :

[0321] GCATCGGATCCGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCT

[0322] GCCGTCACAGACTTCGAAGGTTCGAATCCTTCCCCCACCATTTTGGAAAAGCTT

[0323] GGCGT(SEQ ID NO :100)

[0324] 具有 Tx6 终止子的无 5′ 和 3′ 侧接序列的大肠杆菌形式 1a :

[0325] GCATCGGATCCGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCT

[0326] GCCGTCACAGACTTCGAAGGTTCGAATCCTTCCCCCACCATTTTTTGGAAAAGC

[0327] TTGGCGT(SEQ ID NO :101)

[0328] 具有 5′ 侧接序列的大肠杆菌形式 2 :

[0329] GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTCGGTG

[0330] GGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCACAGACTTC

[0331] GAAGGTTCGAATCCTTCCCCCACCATTTTTTGGAAAAGCTTGGCGT(SEQ ID NO :102)

[0332] 具有 3′ 侧接序列的大肠杆菌形式 3a :

[0333] GCATCGGATCCGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCT

[0334] GCCGTCACAGACTTCGAAGGTTCGAATCCTTCCCCCACCAGACAAGTGCGGTTT

[0335] TTTGGAAAAGCTTGGCGT(SEQ ID NO :103)

[0336] 具有 3′ 和 5′ 侧接序列的大肠杆菌形式 4 :

[0337] GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTCGGTG

[0338] GGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCACAGACTTC

[0339] GAAGGTTCGAATCCTTCCCCCACCAGACAAGTGCGGTTTTTTGGAAAAGCTTGG

[0340] CGT(SEQ ID NO :104)。

[0341] 另外，使用嗜热脂肪芽孢杆菌 $tRNA^{Tyr}$代替大肠杆菌进行实验，其中如图 6 中的第三个流程所示来转录琥珀抑制性嗜热脂肪芽孢杆菌 $tRNA^{Tyr}$。第三个流程包括使用如上文所讨论的相同技术来产生质粒，且这一实验显示使用启动子 (在这种情况下使用 U6) 会进一步增加嗜热脂肪芽孢杆菌 $tRNA^{Tyr}$的功能性琥珀抑制 (图 7 绘示实验结果)。

[0342] 图 6 的完全构筑体的序列如下 :

[0343] 无 5′ 和 3′ 侧接序列和 3′ CCA 的嗜热脂肪芽孢杆菌形式 1 :

[0344] GCATCGGATCCGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCC

[0345] GCTCCCTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCATTTTTTGGAAAAGCT

[0346] TGGCGT(SEQ ID NO :105)

[0347] 无 5′ 和 3′ 侧接序列和 3′ CCA 的以 4 个 T 终止的嗜热脂肪芽孢杆菌形式 1a :

[0348] GCATCGGATCCGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCC

[0349] GCTCCCTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCATTTTGGAAAAGCTTG

[0350] GCGT(SEQ ID NO:106)

[0351] 具有较短 5′ 侧接序列而无 3′ CCA 和 3’侧接序列的嗜热脂肪芽孢杆菌形式 3a：

[0352] GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTCGGAG

[0353] GGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCCGCTCCCTTTGGGTTCG

[0354] GCGGTTCGAATCCGTCCCCCTCCATTTTTTGGAAAAGCTTGGCGT(SEQ ID NO:107)

[0355] 具有 3′ 侧接序列、终止子且无 3′ CCA 和 5′ 侧接序列的嗜热脂肪芽孢杆菌形式 4：

[0356] GCATCGGATCCGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCC

[0357] GCTCCCTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCAGACAAGTGCGGTTTT

[0358] TTGGAAAAGCTTGGCGT(SEQ ID NO:108)

[0359] 具有较短 5′ 侧接序列和正常 3′ 侧接序列而无 3′ CCA 的嗜热脂肪芽孢杆菌形式 5a：

[0360] GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTCGGAG

[0361] GGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCCGCTCCCTTTGGGTTCG

[0362] GCGGTTCGAATCCGTCCCCCTCCAGACAAGTGCGGTTTTTGAAAAGCTTGGC

[0363] GT(SEQ ID NO:109)。

[0364] H1 启动子序列和 U6 启动子序列如下：

[0365] H1 启动子：

[0366] GAATTCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAA

[0367] ACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACTCG

[0368] GATCC(SEQ ID NO:111)

[0369] U6 启动子：

[0370] CCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAGC

[0371] GTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCC

[0372] CATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAG

[0373] AATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAA

[0374] AGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTA

[0375] TCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGGTTTATATATCTTG

[0376] TGGAAAGGACGCGGGATCC(SEQ ID NO:110)。

[0377] 结果

[0378] 琥珀抑制性大肠杆菌 Tyr tRNA(图 3)

[0379] 在琥珀抑制检定中评估数个 tRNA 构筑体，即三个复本、两个复本以及一个复本的侧接有 5′ 人类 $tRNA^{Tyr}$前导序列的琥珀抑制性大肠杆菌 $tRNA^{Tyr}$。将编码 hGHE88 琥珀突变体 tRNA 和大肠杆菌 $tRNA^{Tyr}$合成酶的质粒共转染到 CHO K1 细胞中。也进行省去合成酶的阴性对照。在转染后 41 小时，检定 hGH 的表达。在所有阴性对照实验中，几乎未检测到 hGH 表达 (hGH 特异性 ELISA)。在存在大肠杆菌 Tyr RS 且不存在人类前导序列 tRNA 的情况下，也几乎未检测到 hGH 表达。仅在存在 5′ 端侧接有人类 $tRNA^{Tyr}$序列的大肠杆菌 $tRNA^{Tyr}$合

成酶的情况下检测到 hGH 表达。在两个复本的琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 的情况下，hGH 表达几乎比单个复本的 hGH 表达高 3 倍。这些结果证明，琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 突变体仅在存在人类前导序列 tRNA 的情况下具有抑制功能。

[0380] 使用利用 TIILP 在活体内转录的琥珀抑制性大肠杆菌 Tyr tRNA，在存在对乙酰基 - 苯丙氨酸 (pAF) 的情况下进行琥珀抑制 (图 4)

[0381] 将编码 hGH E88 琥珀突变体 (特定 tRNA 格式) 和装有 pAF 的大肠杆菌 $tRNA^{Tyr}$ 合成酶突变体的质粒共转染到 CHO K1 细胞中。用含有 1mM pAF 的生长培养基更换转染培养基。在转染后 42 小时，检定 hGH 的表达。一个人类 tRNA 前导序列在下游与琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 的级联复本融合的 tRNA 格式 (h-(EC)2) 产生极有限的 hGH 表达。使用以下两个替代的 tRNA 格式检测到较高 hGH 表达 ：第一个含有与单一复本的琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 融合的 5′ 人类 tRNA 前导序列 (hEC) ；第二个含有刚才描述的整个转录盒的两个复本 (2×(hEC))。对于这两个构筑体，可通过使转染的 tRNA 质粒的量加倍 (即从 1 微克到 2 微克) 来进一步增加琥珀抑制产率。相反，当使转染的编码 pAFRS 的质粒的量加倍时，琥珀抑制产率降低。

[0382] 使用利用 H1 启动子在活体内转录的琥珀抑制性大肠杆菌 Tyr tRNA，在存在 pAF 的情况下进行琥珀抑制 (图 5)

[0383] 将编码 hGH E88 琥珀突变体 (一种由 H1 启动子驱动的大肠杆菌 Tyr 琥珀抑制性 tRNA) 和大肠杆菌 $tRNA^{Tyr}$ 合成酶突变体的质粒共转染到 CHO K1 细胞中。作为阴性对照，使用缺乏 H1 启动子且由在 5′ 和 3′ 侧接人类 Tyr tRNA 侧接序列的单一复本的大肠杆菌 Tyr 琥珀抑制性 tRNA 组成的质粒。用含有 1mM pAF 的生长培养基更换转染培养基，且在转染后 42 小时，通过 ELISA 检定 hGH 的表达。在琥珀抑制性大肠杆菌 $tRNA^{Tyr}$ 的 5′ 端中不存在 H1 启动子的阴性对照实验产生极有限的 hGH 抑制。类似地，利用 tRNA 仅在 5′ 端处侧接紧靠人类 Tyr tRNA 上游的 5′ 序列的 tRNA 转录形式 2(图 2) 检测到 hGH 的极低效价。使用形式 1a、3 和 4 转录盒检测到较高水平的经抑制的 hGH(图 2)。总之，这些结果提示缺乏内部 A 盒的琥珀抑制性大肠杆菌 Tyr tRNA 无法由 CHO 细胞中的 pol III 有效转录 ；此格式 (即形式 1) 的琥珀抑制较差。然而，如在 tRNA 转录构筑体形式 1a、形式 3 和形式 4 中所观察，在 tRNA 的 5′ 端包括诸如 H1 等 pol III 启动子使得能够转录和产生功能性琥珀抑制性 tRNA。构筑体形式 2 产生有限抑制效价的事实提示 5′ 侧接序列对产生功能性琥珀抑制性 tRNA 来说很重要。

[0384] 实例 3

[0385] 利用来自大肠杆菌的野生型 Leu RS 和琥珀抑制性 tRNA，在 CHO K1 细胞中进行琥珀抑制

[0386] 在这一实验中，利用来自大肠杆菌的野生型 Leu RS 和琥珀抑制性 tRNA 来抑制 CHO K1 细胞中的选择密码子 (在这种情况下为琥珀密码子)(GCCCGGATGGTGGAATCGGTaGACACAAGGGATTCTAAATCCCTCGGCGTTCGCGCTgTGCGGGTTCAAGTCCCGCTCCGGGTA(SEQ ID NO ：112))。

[0387] 使用嗜热脂肪芽孢杆菌对作为阳性对照，且图左侧绘示阴性对照。此处使用两组启动子来驱动 tRNA 表达 ：H1 和 U6(2 组)，且在这种情况下，tRNA 不是 tRNA 簇。对于改进的抑制，在转染期间使用两倍的编码 tRNA 的 DNA。

[0388] 图 8 中的结果绘示大肠杆菌 Leu RS/tRNA 对在哺乳动物细胞中正交。新颖 tRNA

转录流程(U6、H1 驱动 tRNA 表达)还对另一也缺乏 A 盒序列的 tRNA 起作用。

[0389] 向具有非天然氨基酸的蛋白质中添加分子。

[0390] 一方面,本发明提供包含与其它取代基分子偶合的非天然氨基酸的蛋白质的方法和相关组合物。

[0391] 应了解,本文中所述的实例和实施例仅出于说明的目的,且将对所属领域的技术人员提示根据其作出各种修改和变化,且所述修改和变化包括在本申请案的精神和权限以及附属权利要求书的范围内。

[0392] 虽然出于清楚和便于理解的目的已较详细地描述上述发明,但通过阅读本公开文献,显而易见所属领域的技术人员可在不脱离本发明的真正范围的情况下进行形式和细节的各种变化。举例来说,本文中所述的所有技术和装置都可以不同组合使用。本申请案中所引用的所有公开文献、专利、专利申请案和/或其它文献出于所有目的都是以引用的方式全部并入本文中,引用的程度就如同已个别地将各个公开文献、专利案、专利申请案和/或其它文献出于所有目的以引用的方式并入一般。

[0393] 表 5

[0394]

SEQ ID NO:	标记	序列

[0395]

SEQ ID NO: 1	大肠杆菌野生型 TyrRS（合成酶）多核苷酸	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTA GCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCGCTCTATTGCGGCTTCGATCCTACCGCTGACAGCTTGCAT TTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGG GCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCG ACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTG TTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCG ATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCGAACAACTATGACT GGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACA CTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCT CAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTG TTGCAGGGTTATGACTTCGCCTGTCTGAACAAACAGTACGGTGTGGTGC TGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGA CCTGACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCG CTGATCACTAAAGCAGATGGCACCAAATTTGGTAAAACTGAAGGCGGC GCAGTCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAG TTCTGGATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCT TCACCTTTATGAGCATTGAAGAGATCAACGCCCTGGAAGAAGAAGATA AAAACAGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAG GTGACTCGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGT ATTACCGAATGCCTGTTCAGCGGTTCTTTGAGTGCGCTGAGTGAAGCGG ACTTCGAACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAAA AGGGCGCAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTT CCCGTGGTCAGGCACGTAAAACTATCGCCTCCAATGCCATCACCATTAA CGGTGAAAAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGATCG TCTGTTTGGTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGT CTGATTTGCTGGAAATAA
SEQ ID NO: 2	大肠杆菌野生型 TyrRS（合成酶）氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALYCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYDFACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 3	pOMe-1 合成酶多核苷酸	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTA GCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGC CCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTT GGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGC CACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGAC CCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTT CAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATT TCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGTT CGGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTC TCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAAC CGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGC AGGGTTATAGTATGGCCTGTTTGAACAAACAGTACGGTGTGGTGCTGCA AATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTG ACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCGCTGA TCACTAAAGCAGATGGCACCAAATTTGGTAAAACTGAAGGCGGCGCAG TCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAGTTCTG GATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCTTCACC TTTATGAGCATTGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAAC AGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAGGTGACT CGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGTATTACC GAATGCCTGTTCAGCGGTTCTTTGAGTGCGCTGAGTGAAGCGGACTTCG AACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAAAAGGGCG CAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTTCCcGTGG TCAGGCACGTAAAACTATCGCCTCCAATGCCATCACCATTAACGGTGAA AAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTG GTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGTCTGATTTGC TGGAAATAA

[0396]

SEQ ID NO: 4	pOMe-2 合成酶多核苷酸	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTA gCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGC CCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTT GGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGC CACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGAC CCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTT CAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATT TCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGTT CAGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTC TCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAAC CGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGC AGGGTTATACGTATGCCTGTCTGAACAAACAGTACGGTGTGGTGCTGCA AATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTG ACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCGCTGA TCACTAAAGCAGATGGCACCAAATTTGGTAAAACTGAAGGCGGCGCAG TCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAGTTCTG GATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCTTCACC TTTATGAGCATTGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAAC AGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAGGTGACT CGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGTATTACC GAATGCCTGTTCAGCGGTTCTTTGAGTGCGCTGAGTGAAGCGGACTTCG AACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAAAAGGGCG CAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTTCCCGTGG TCAGGCACGTAAAACTATCGCCTCCAATGCCATCACCATTAACGGTGAA AAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTG GTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGTCTGATTTGC TGGAAATAA
SEQ ID NO: 5	pOMe-3 合成酶多核苷酸	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTA GCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGC CCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTT GGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGC CACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGAC
		CCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTT CAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATT TCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGTT CGGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTC TCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAAC CGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGC AGGGTTATAGTATGGCCTGTTTGAACAAACAGTACGGTGTGGTGCTGCA AATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTG ACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCGCTGA TCACTAAAGCAGATGGCACCAAATTTGGTAAAACTGAAGGCGGCGCAG TCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAGTTCTG GATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCTTCACC TTTATGAGCATTGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAAC AGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAGGTGACT CGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGTATTACC GAATGCCTGTTCAGCGGTTCTTTGAGTGCGCTGAGTGAAGCGGACTTCG AACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAAAAGGGCG CAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTTCCCGTGG TCAGGCACGTAAAACTATCGCCTCCAATGCCATCACCATTAACGGTGAA AAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTG GTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGTCTGATTTGC TGGAAATAA

[0397]

SEQ ID NO: 6	pOMe-4 合成酶多核苷酸	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAgCGGGGGCTGGTA GCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGC CCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTT GGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGC CACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGAC CCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTT CAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATT TCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGTT CGGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTC TCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAAC CGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGC AGGGTTATAGTATGGCCTGTTTGAACAAACAGTACGGTGTGGTGCTGCA AATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTG ACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCGCTGA TCACTAAAGCAGATGGCACCAAATTTGGTAAAACTGAAGGCGGCGCAG TCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAGTTCTG GATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCTTCACC TTTATGAGCATTGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAAC AGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAGGTGACT CGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGTATTACC GAATGCCTGTTCAGCGGTTCTTTGAGTGCGCTGAGTGAAgCGGACTTCG AACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAAAAGGGCG CAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTTCCCGTGG TCAGGCACGTAAAACTATCGCCTCCAATGCCATCACCATTAACGGTGAA AAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTG GTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGTCTGATTTGC TGGAAATAA
SEQ ID NO: 7	pOMe-5 合成酶多核苷酸	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTA gCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGC CCGATCGCACTCACGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATT TGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGG CCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGA CCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGT TCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGAT TTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGT TCGGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTT
		CTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAA CCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAGCCTGCTG CAGGGTTATACGATGGCCTGTCTGAACAAACAGTACGGTGTGGTGCTGC AAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCT GACCCGTCGTCTGCATCAGAATCAGGTGTTTGGCCTGACCGTTCCGCTG ATCACTAAAGCAGATGGCACCAAATTTGGTAAAACTGAAGGCGGCGCA GTCTGGTTGGATCCGAAGAAAACCAGCCCGTACAAATTCTACCAGTTCT GGATCAACACTGCGGATGCCGACGTTTACCGCTTCCTGAAGTTCTTCAC CTTTATGAGCATTGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAA CAGCGGTAAAGCACCGCGCGCCCAGTATGTACTGGCGGAGCAGGTGAC TCGTCTGGTTCACGGTGAAGAAGGTTTACAGGCGGCAAAACGTATTACC GAATGCCTGTTCAGCGGTTCTTTGAGTGCGCTGAGTGAAGCGGACTTCG AACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAAAAGGGCG CAGACCTGATGCAGGCACTGGTCGATTCTGAACTGCAACCTTCCCGTGG TCAGGCACGTAAAACTATCGCCTCCAATGCCATCACCATTAACGGTGAA AAACAGTCCGATCCTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTG GTCGTTTTACCTTACTGCGTCGCGGTAAAAAGAATTACTGTCTGATTTGC TGGAAATAA

[0398]

SEQ ID NO: 8	pOMe-6（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCAGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGTATGCCTGTCTGAACAAACAGTACG GTGTG
SEQ ID NO: 9	pOMe-7（活性位点）合成酶多核苷酸	CGGGGGCTGGTACCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCAGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGTATGCCTGTCTGAACAAACAGTACG GTGTG
SEQ ID NQ: 10	pOMe-8（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCAGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGTATGCCTGTCTGAACAAACAGTACG GTGTG
SEQ ID NO: 11	pOMe-9（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT
		CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATTCGTATGCCTGTGCGAACAAACAGTACG GTGTG
SEQ ID NO: 12	pOMe-10（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCAGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGTATGCCTGTCTGAACAAACAGTACG GTGTG

[0399]

SEQ ID NO: 13	pOMe-11(活性位点)合成酶多核苷酸	CGGGGGCTGGTACCcCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCCTTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATTCTATTGCCTGTTCGAACAAACAGTACG GTGTG
SEQ ID NO: 14	pOMe-12(活性位点)合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATAGTATTGCCTGTTTGAACAAACAGTACG GTGTG
SEQ ID NO: 15	pOMe-13(活性位点)合成酶多核苷酸	CGGGGGCTGGTACCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATAGTATTGCCTGTTTGAACAAACAGTACG GTGTG
SEQ ID NO: 16	pOMe-14(活性位点)合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCTGGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC
		CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATT GTTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATATGCGTGCCTGTGAGAACAAACAGTACG GTGTG
SEQ ID NO: 17	对乙酰基Phe-1(活性位点)合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCATTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGGTCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATGGTATGGCCTGTGCTAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAATGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG

[0400]

SEQ ID NO: 18	对二苯甲酮-1（活性位点）合成酶多核苷酸	CAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCG ATCGCACTCGGTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGG GGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCA CAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGACCC GAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTTCA GGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATTTC GACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGTTCG GCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTCTC CGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCG TGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGCAG GGTTATGGTTTTGCCTGTTTGAACAAACAGTACGGTGTGGTGCTGCAAA TTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTGAC CCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 19	对二苯甲酮-2（活性位点）合成酶多核苷酸	GCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGGGTGTGGC TTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGTTCCATTGTT ATGCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGT AGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGA GCGTAAGCTGAACACCGAAGAAACTGTTCAGGAGTGGGTGGACAAAAT CCGTAAGCAGGTTGCCCCGTTCCTCGATTTCGACTGTGGAGAAAACTCT GCTATCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTGCTGA CCTTCCTGCGCGATATTGGCAAACACTTCTCCGTTAACCAGATGATCAA CAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTC GTTCACTGAGTTTTCCTACAACCTGCTGCAGGGTTATGGTTATGCCTGTA TGAACAAACAGTACGGTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTG GGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCAGAAT CAGGTG
SEQ ID NO: 20	对叠氮基 Phe-1（活性位点）合成酶多核苷酸	GGGCTGGTAGCCCAGGTGACGGACGNAGAAGCGTTAGCAGAGCGACTG GCGCAAGGCCCGATCGCACTCCTTTGTGGCTTCGATCCTACCGCTGACA GCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAG CAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTG ATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAA GAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCG TTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATAATT ATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGG CAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAAGCA GCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTAC
		AACCTGCTGCAGGGTTATTCTATGGCCTGTGCGAACAAACAGTACGGTG TGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGG TATCGACCTGACCCGTCGTCTGCATCANAATCANGTG
SEQ ID NO: 21	对叠氮基 Phe-2（活性位点）合成酶多核苷酸	TTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGTTTGTGGCTTCG ATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGC CTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGC GGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGT AAGCTGAACACCGAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGT AAGCAGGTTGCCCCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTA TCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTGCTGACCTT CCTGCGCGATATTGGCAAACACTTCTCCGTTAACCAGATGATCAACAAA GAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTC ACTGAGTTTTCCTACAACCTGCTGCAGGGTTATTCTGCGGCCTGTGCGA ACAAACAGTACGGTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGG GTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCAGAATCA GGTG

[0401]

SEQ ID NO: 22	对叠氮基Phe-3（活性位点）合成酶多核苷酸	GACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTC CTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGT TCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTT GCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAA GCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTTCAGGAGTGGGTG GACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATTTCGACTGTGGAG AAAACTCTGCTATCGCGGCCAATAATTATGACTGGTTCGGCAATATGAA TGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTCTCCGTTAACCAG ATGATCAACAAANAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAG GGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGCAGGGTTATTCGGC TGCCTGTGCGAACAAACAGTACGGNGNGGNGCTGCAAATTGGNGGTTC TGACCAGGGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTG CATCAAAATCAGGTG
SEQ ID NO: 23	对叠氮基Phe-4（活性位点）合成酶多核苷酸	GCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGTTTGTGGCT TCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTG TGCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTA GGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAG CGTAAGCTGAACACCGAAGAAACTGTTCAGGAGTGGGTGGACAAAATC CGTAAGCAGGTTGCCCCGTTCCTCGATTTCGACTGTGGAGAAAACTCTG CTATCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTGCTGAC CTTCCTGCGCGATATTGGCAAACACTTCTCCGTTAACCAGATGATCAAC AAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCG TTCACTGAGTTTTCCTACAACCTGCTGCAGGGTTATAGTGCGGCCTGTGT TAACAAACAGTACGGTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGG GGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCAGAATC ANGTG
SEQ ID NO: 24	对叠氮基Phe-5（活性位点）合成酶多核苷酸	GACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTC ATTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGT TCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTT GCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAA GCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTTCAGGAGTGGGTG GACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATTTCGACTGTGGAG AAAACTCTGCTATCGCGGCCAATGATTATGACTGGTTCGGCAATATGAA TGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTCTCCGTTAACCAG ATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAG GGGATTTCGTTCACTGAGTTTTCCTACAACCTGCTGCAGGGTTATAATTT TGCCTGTGTGAACAAACAGTACGGTGTGGTGCTGCAAATTGGTGGTTCT GACCAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGC ATCAGAATCAGGTG
SEQ ID NO: 25	对叠氮基Phe-6（活性位点）合成酶多核苷酸	CGACTGGCGCAAGGCCCGATCGCACTCACGTGTGGCTTCGATCCTACCG CTGACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGC TTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACG GGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAAC
		ACCGAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTT GCCCCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCA ATAATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGA TATTGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTT AAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTT CCTACAATCTGCTGCAGGGTTATTCGGCTGCCTGTCTTAACAAACAGTA CGGTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACT TCTGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG

[0402]

SEQ ID NO: 26	pPR-EcRS-1(炔丙氧基苯丙氨酸合成酶)（活性位点)合成酶多核苷酸	CGGGGGCTGGTANCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGGGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATTCTATGGCCTGTTTGAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGANCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 27	pPR-EcRS-2（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAATCTGCTGCAGGGTTATTCGGCTGCCTGTCTTAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGAACCTGANCCGTCGTCTGCATCAAAATCAAGTG
SEQ ID NO: 28	pPR-EcRS -3（活性位点）合成酶多核苷酸	CGGGGGCTGGTACCCCAAGTGACGGACGAGGAAACGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCTCTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCAGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGATGGCCTGTGTGAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 29	pPR-EcRS-4（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGCGTGCGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATTCTTATGCCTGTCTTAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC
		TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG

[0403]

SEQ ID NO: 30	pPR-EcRS-5（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGCGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGATGGCCTGTTGTAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 31	pPR-EcRS-6（活性位点）合成酶多核苷酸	CGGGGGCTGGTACCCCAAGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCGCTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGTTTGCCTGTATGAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 32	pPR-EcRS-7（活性位点）合成酶多核苷酸	GTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATC GCACTCACGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGC ATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCACAA GCCGGTTGCGCTGGTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAG CTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTTCAGGA GTGGGTGGACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATTTCGAC TGTGGAGAAAACTCTGCTATCGCGGCCAATAATTATGACTGGTTCGGCA ATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACACTTCTCCGT TAACCAGATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGA AGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAATCTGCTGCAGGGT TATTCGGCTGCCTGTCTTAACAAACAGTACGGTGTGGTGCTGCAAATTG GTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCG TCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 33	pPR-EcRS-8（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGTTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATTCGATGGCCTGTACGAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 34	pPR-EcRS-9（活性位点）合成酶多核苷酸	CGGGGGCTGGTANCCCAAGTGACGGACGGGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCAGTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC

[0404]

		CCGTTCCTCGATCTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATAGTTTTGCCTGTCTGAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 35	pPR-EcRS-10（活性位点）合成酶多核苷酸	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCACGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACC GAAGAAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGCGATAT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATACGTTTGCCTGTACTAACAAACAGTACG GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTC TGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
SEQ ID NO: 36	对碘 PheRS-1 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSYACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 37	对碘 PheRS-2 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 38	对碘 PheRS-3 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 39	OMeTyrRS-1 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 40	OMeTyrRS-2 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM

[0405]

		INKEAVKQRLNREDQGISFTEFSYNLLQGYTMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 41	OMeTyrRS-3 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTYACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 42	OMeTyrRS-4 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACSNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 43	OMeTyrRS-5 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 44	OMeTyrRS-6 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYRMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 45	对乙酰基 PheRS-1 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYGMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 46	对苯甲酰基 PheRS-1 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALGCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYGFACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK

[0406]

SEQ ID NO: 47	对苯甲酰基 PheRS-2 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALGCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYGYACMNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 48	对叠氮基 PheRS-1 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 49	对叠氮基 PheRS-2 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 50	对叠氮基 PheRS-3 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 51	对叠氮基 PheRS-4 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACVNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 52	对叠氮基 PheRS-5 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANDYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYNFACVNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 53	对叠氮基 PheRS-6 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACLNKQYGVVLQIGGSDQ

[0407]

		WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 54	pPR-EcRS-1 合成酶氨基酸（aa） *对炔丙氧基苯丙氨酸合成酶*	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALGCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 55	pPR-EcRS-2 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 56	pPR-EcRS-3 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALSCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTMACVNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 57	pPR-EcRS-4 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALACGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSYACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 58	pPR-EcRS-5 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALACGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTMACCNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 59	pPR-EcRS-6 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTFACMNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK

[0408]

SEQ ID NO: 60	pPR-EcRS-7 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSVACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 61	pPR-EcRS-8 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACTNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 62	pPR-EcRS-9 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALSCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSFACLNKQYGVVLQIGGSDQW GNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTSP YKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVLA EQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 63	pPR-EcRS-10 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTFACTNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 64	tRNA/Tyr 多核苷酸	AGCTTCCCGATAAGGGAGCAGGCCAGTAAAAAGCATTACCCCGTGGTG GGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCG ACCTCGAAGGTTCGAATCCTTCCCCCACCACCA
SEQ ID NO: 65	tRNA/Tyr	AGCUUCCCGAUAAGGGAGCAGGCCAGUAAAAAGCAUUACCCCGUGGU GGGGUUCCCGAGCGGCCAAAGGGAGCAGACUCUAAAUCUGCCGUCAU CGACCUCGAAGGUUCGAAUCCUUCCCCCACCACCA
SEQ ID NO: 66	琥珀突变体 L3TAG	5'-ATGAAGTAGCTGTCTTCTATCGAACAAGCATGCG-3'
SEQ ID NO: 67	琥珀突变体 113TAG	5'-CGAACAAGCATGCGATTAGTGCCGACTTAAAAAG-3'
SEQ ID NO: 68	琥珀突变体 T44TAG	5'-CGCTACTCTCCCAAATAGAAAAGGTCTCCGCTG-3'
SEQ ID NO:	琥珀突变体 F68TAG	5'-CTGGAACAGCTATAGCTACTGATTTTTCCTCG-3'

[0409]

69		
SEQ ID NO: 70	琥珀突变体 R110TAG	5'-GCCGTCACAGATTAGTTGGCTTCAGTGGAGACTG-3'
SEQ ID NO: 71	琥珀突变体 V114TAG	5'-GATTGGCTTCATAGGAGACTGATATGCTCTAAC-3'
SEQ ID NO:72	琥珀突变体 T121TAG	5'-GCCTCTATAGTTGAGACAGCATAGAATAATGCG-3'
SEQ ID NO: 73	琥珀突变体 I127TAG	5'-GAGACAGCATAGATAGAGTGCGACATCATCATCGG-3'
SEQ ID NO: 74	琥珀突变体 S131TAG	5'-GAATAAGTGCGACATAGTCATCGGAAGAGAGTAGTAG-3'
SEQ ID NO: 75	琥珀突变体 T145TAG	5'-GGTCAAAGACAGTTGTAGGTATCGATTGACTCGGC-3'
SEQ ID NO: 76	许可位点突变体 T44F	5'-CGCTACTCTCCCCAAATTTAAAAGGTCTCCGCTG-3'
SEQ ID NO: 77	许可位点突变体 T44Y	5'-CGCTACTCTCCCCAAATATAAAAGGTCTCCGCTG-3'
SEQ ID NO: 78	许可位点突变体 T44W	5'-CGCTACTCTCCCCAAATGGAAAAGGTCTCCGCTG-3'
SEQ ID NO: 79	许可位点突变体 T44D	5'-CGCTACTCTCCCCAAAGATAAAAGGTCTCCGCTG-3'
SEQ ID NO: 80	许可位点突变体 T44K	5'-CGCTACTCTCCCCAAAAAAAAAAGGTCTCCGCTG-3'
SEQ ID NO: 81	许可位点突变体 R110F	5'-GCCGTCACAGATTTTTTGGCTTCAGTGGAGACTG-3'
SEQ ID NO: 82	许可位点突变体 R110Y	5'-GCCGTCACAGATTATTTGGCTTCAGTGGAGACTG-3'
SEQ ID NO:	许可位点突变	5'-GCCGTCACAGATTGGTTGGCTTCAGTGGAGACTG-3'

[0410]

83	体 R110W	
SEQ ID NO: 84	许可位点突变体 R110D	5'-GCCGTCACAGATGATTTGGCTTCAGTGGAGACTG-3'
SEQ ID NO: 85	许可位点突变体 R110K	5'-GCCGTCACAGATAAATTGGCTTCAGTGGAGACTG-3'
SEQ ID NO: 86	对乙酰基 PheRS-1 合成酶氨基酸（aa）	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREGQGISFTEFSYNLLQGYGMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
SEQ ID NO: 87	人类 Tyr tRNA5'侧接序列	CTGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTC
SEQ ID NO: 88	3'侧接人类 tRNATyr 序列	GACAAGTGCGGTTTTTTTCTCCAGCTCCCGATGACTTATGGC
SEQ ID NO: 89	FTam97 短引物	GTACGAATTCCCGAGATCTCTGTGCTGAACCTCAGGGGACGC
SEQ ID NO: 90	FTam109	GATGCAAGCTTGATGGATCCGCCATAAGTCATCGGGAGCTGGAGAAAA AAACCGCACTTGTCTGGTGGGGGAAGGATTCG
SEQ ID NO: 91	人类 5'侧接 tRNATyr 序列	GGATTACGCATGCTCAGTGCAATCTTCGGTTGCCTGGACTAGCGCTCCG GTTTTTCTGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTC
SEQ ID NO: 92	人类 3'侧接 tRNATyr 序列	GACAAGTGCGG
SEQ ID NO: 93	FT73-out-new 引物	CTTTGTGTAATACTTGTAACGCTGAATTC
SEQ ID NO: 94	FT76-out 反向引物	ACCATGATTACGCCAAGCTTGAT
SEQ ID NO: 95	FTam 111 引物	GCATCGGATCCGGTGGGGTTCCCGAGCGGCC
SEQ	FTam 112 引物	ACGCCAAGCTTTTCCAAAATGGTGGGGGAAGGATTCGAACCTTC

[0411]

ID NO: 96		
SEQ ID NO: 97	FTam 102a 引物	GCATCGGATCCGTGCTGAACCTCAGGGGACGCCG
SEQ ID NO: 98	FTam 113 引物	ACGCCAAGCTTTTCCAAAAAATGGTGGGGGAAGGATTCGAACCTTC
SEQ ID NO: 99	FTam 114 引物	ACGCCAAGCTTTTCCAAAAAACCGCACTTGTCTGGTGGGGGAAGG
SEQ ID NO: 100	大肠杆菌形式 1 构筑体	GCATCGGATCCGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTA AATCTGCCGTCACAGACTTCGAAGGTTCGAATCCTTCCCCCACCATTTT GGAAAAGCTTGGCGT
SEQ ID NO: 101	大肠杆菌形式 1a 构筑体	GCATCGGATCCGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTA AATCTGCCGTCACAGACTTCGAAGGTTCGAATCCTTCCCCCACCATTTTT TGGAAAAGCTTGGCGT
SEQ ID NO: 102	大肠杆菌形式 2 构筑体	GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGT CGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGT CACAGACTTCGAAGGTTCGAATCCTTCCCCCACCATTTTTGGAAAAGC TTGGCGT
SEQ ID NO: 103	大肠杆菌形式 3 构筑体	GCATCGGATCCGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTA AATCTGCCGTCACAGACTTCGAAGGTTCGAATCCTTCCCCCACCAGACA AGTGCGGTTTTTGGAAAAGCTTGGCGT
SEQ ID NO: 104	大肠杆菌形式 4 构筑体	GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGT CGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGT CACAGACTTCGAAGGTTCGAATCCTTCCCCCACCAGACAAGTGCGGTTT TTGGAAAAGCTTGGCGT
SEQ ID NO: 105	嗜热脂肪芽孢杆菌形式 1	GCATCGGATCCGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTA AATCCGCTCCCTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCATTTTT TGGAAAAGCTTGGCGT
SEQ ID NO: 106	嗜热脂肪芽孢杆菌形式 1a	GCATCGGATCCGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTA AATCCGCTCCCTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCATTTTG GAAAAGCTTGGCGT
SEQ ID NO: 107	嗜热脂肪芽孢杆菌形式 3a	GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGT CGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCCGCTCC CTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCATTTTTGGAAAAGCT TGGCGT
SEQ ID NO: 108	嗜热脂肪芽孢杆菌形式 4	GCATCGGATCCGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTA AATCCGCTCCCTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCAGACA AGTGCGGTTTTTGGAAAAGCTTGGCGT
SEQ ID NO: 109	嗜热脂肪芽孢杆菌形式 5a	GCATCGGATCCGTGCTGAACCTCAGGGGACGCCGACACACGTACACGT CGGAGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCCGCTCC CTTTGGGTTCGGCGGTTCGAATCCGTCCCCCTCCAGACAAGTGCGGTTT TTGGAAAAGCTTGGCGT

[0412]

SEQ ID NO: 110	U6 启动子序列	CCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGA CCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTC CCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGAT AATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACG TGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTAT GTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGA TTTCTTGGGTTTATATATCTTGTGGAAAGGACGCGGGATCC
SEQ ID NO: 111	H1 启动子序列	GAATTCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACG TGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACT CGGATCC
SEQ ID NO: 112	大肠杆菌 Leu 琥珀抑制性 tRNA	GCCCGGATGGTGGAATCGGTaGACACAAGGGATTCTAAATCCCTCGGCG TTCGCGCTgTGCGGGTTCAAGTCCCGCTCCGGGTA

[0413] [a]这些克隆也含有 Asp165Gly 突变。

P095438CN-seq.txt
序列表
<110> 田锋
席雅·诺曼
斯蒂芬妮·周
<120> 脊椎动物细胞中抑制因子 TRNA 的转录
<130>AMBX-0124.00PCT
<150>60/843,264
<151>2006-09-08
<160>112
<170>PatentIn version 3.4
<210>1
<211>1275
<212>DNA
<213>Escherichia coli
<400>1
atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcgctcta ttgcggcttc 120
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 180
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 360
gctatcgcgg cgaacaacta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 420
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gttgcagggt 540
tatgacttcg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720
ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780
atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840
gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900
tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960
aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200
gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260
atttgctgga aataa 1275

<210>2
<211>424
<212>PRT
<213>Escherichia coli
<400>2

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Tyr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Asp Phe Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>3
<211>1275
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>3

atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcgt gtgtggcttc 120
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 180
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 360
gctatcgcgg ccaataatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 420
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540
tatagtatgg cctgtttgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720

ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780
atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840
gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900
tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960
aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200
gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260
atttgctgga aataa 1275

<210>4

<211>1275

<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>4

atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcac ttgtggcttc 120
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 180
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 360
gctatcgcgg ccaataatta tgactggttc agcaatatga atgtgctgac cttcctgcgc 420
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540
tatacgtatg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720
ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780
atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840
gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900
tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960
aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200
gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260
atttgctgga aataa 1275

<210>5
<211>1275
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>5

atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcgt gtgtggcttc 120
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 180
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 360
gctatcgcgg ccaataatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 420
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540
tatagtatgg cctgtttgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720
ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780
atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840
gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900
tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960
aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200
gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260
atttgctgga aataa 1275

<210>6
<211>1275
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>6

atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcgt gtgtggcttc 120
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 180

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 360
gctatcgcgg ccaataatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 420
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540
tatagtatgg cctgtttgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720
ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780
atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840
gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900
tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960
aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200
gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260
atttgctgga aataa 1275

<210>7

<211>1275

<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>7

atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcac gtgtggcttc 120
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 180
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 360
gctatcgcgg ccaataatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 420
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacagcct gctgcagggt 540
tatacgatgg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720
ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780

atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840
gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900
tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960
aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200
gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260
atttgctgga aataa 1275

<210>8
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>8

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacggtgtg 540

<210>9
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>9

cgggggctgg taccccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacggtgtg 540

<210>10
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>10

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacggtgtg 540

<210>11
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>11

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttattcg tatgcctgtg cgaacaaaca gtacggtgtg 540

<210>12
<211>540
<212>DNA
<213> 人工
<220>

<223> 人工合成酶
<400>12
cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacggtgtg 540
<210>13
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>13
cgggggctgg taccccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcctttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttattct attgcctgtt cgaacaaaca gtacggtgtg 540
<210>14
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>14
cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcgtgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatagt attgcctgtt tgaacaaaca gtacggtgtg 540

<210>15
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>15

cgggggctgg taccccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcgtgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatagt attgcctgtt tgaacaaaca gtacggtgtg 540

<210>16
<211>540
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>16

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tctggtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaagg ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaatt gttatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatatg cgtgcctgtg agaacaaaca gtacggtgtg 540

<210>17
<211>624
<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>17

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcatttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaaggtc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatggt atggcctgtg ctaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccaatgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624

<210>18

<211>609

<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>18

caggtgacgg acgaggaagc gttagcagag cgactggcgc aaggcccgat cgcactcggt 60
tgtggcttcg atcctaccgc tgacagcttg catttggggc atcttgttcc attgttatgc 120
ctgaaacgct tccagcaggc gggccacaag ccggttgcgc tggtaggcgg cgcgacgggt 180
ctgattggcg acccgagctt caaagctgcc gagcgtaagc tgaacaccga agaaactgtt 240
caggagtggg tggacaaaat ccgtaagcag gttgccccgt tcctcgattt cgactgtgga 300
gaaaactctg ctatcgcggc caataattat gactggttcg gcaatatgaa tgtgctgacc 360
ttcctgcgcg atattggcaa acacttctcc gttaaccaga tgatcaacaa agaagcggtt 420
aagcagcgtc tcaaccgtga agatcagggg atttcgttca ctgagttttc ctacaacctg 480
ctgcagggtt atggttttgc ctgtttgaac aaacagtacg gtgtggtgct gcaaattggt 540
ggttctgacc agtggggtaa catcacttct ggtatcgacc tgacccgtcg tctgcatcag 600
aatcaggtg 609

<210>19

<211>591

<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>19
gcgttagcag agcgactggc gcaaggcccg atcgcactcg ggtgtggctt cgatcctacc 60
gctgacagct tgcatttggg gcatcttgtt ccattgttat gcctgaaacg cttccagcag 120
gcgggccaca agccggttgc gctggtaggc ggcgcgacgg gtctgattgg cgacccgagc 180
ttcaaagctg ccgagcgtaa gctgaacacc gaagaaactg ttcaggagtg ggtggacaaa 240
atccgtaagc aggttgcccc gttcctcgat ttcgactgtg gagaaaactc tgctatcgcg 300
gccaataatt atgactggtt cggcaatatg aatgtgctga ccttcctgcg cgatattggc 360
aaacacttct ccgttaacca gatgatcaac aaagaagcgg ttaagcagcg tctcaaccgt 420
gaagatcagg ggatttcgtt cactgagttt tcctacaacc tgctgcaggg ttatggttat 480
gcctgtatga acaaacagta cggtgtggtg ctgcaaattg gtggttctga ccagtggggt 540
aacatcactt ctggtatcga cctgacccgt cgtctgcatc agaatcaggt g 591
<210>20
<211>621
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<220>
<221>misc_feature
<222>(26)..(26)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature
<222>(612)..(612)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature
<222>(618)..(618)
<223>n 为 a、c、g 或 t
<400>20
gggctggtag cccaggtgac ggacgnagaa gcgttagcag agcgactggc gcaaggcccg 60
atcgcactcc tttgtggctt cgatcctacc gctgacagct tgcatttggg gcatcttgtt 120
ccattgttat gcctgaaacg cttccagcag gcgggccaca agccggttgc gctggtaggc 180
ggcgcgacgg gtctgattgg cgacccgagc ttcaaagctg ccgagcgtaa gctgaacacc 240
gaagaaactg ttcaggagtg ggtggacaaa atccgtaagc aggttgcccc gttcctcgat 300
ttcgactgtg gagaaaactc tgctatcgcg gccaataatt atgactggtt cggcaatatg 360
aatgtgctga ccttcctgcg cgatattggc aaacacttct ccgttaacca gatgatcaac 420
aaagaagcgg ttaagcagcg tctcaaccgt gaagatcagg ggatttcgtt cactgagttt 480
tcctacaacc tgctgcaggg ttattctatg gcctgtgcga acaaacagta cggtgtggtg 540

ctgcaaattg gtggttctga ccagtggggt aacatcactt ctggtatcga cctgacccgt 600
cgtctgcatc anaatcangt g 621
<210>21
<211>588
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>21
ttagcagagc gactggcgca aggcccgatc gcactcgttt gtggcttcga tcctaccgct 60
gacagcttgc atttggggca tcttgttcca ttgttatgcc tgaaacgctt ccagcaggcg 120
ggccacaagc cggttgcgct ggtaggcggc gcgacgggtc tgattggcga cccgagcttc 180
aaagctgccg agcgtaagct gaacaccgaa gaaactgttc aggagtgggt ggacaaaatc 240
cgtaagcagg ttgccccgtt cctcgatttc gactgtggag aaaactctgc tatcgcggcc 300
aataattatg actggttcgg caatatgaat gtgctgacct tcctgcgcga tattggcaaa 360
cacttctccg ttaaccagat gatcaacaaa gaagcggtta agcagcgtct caaccgtgaa 420
gatcagggga tttcgttcac tgagttttcc tacaacctgc tgcagggtta ttctgcggcc 480
tgtgcgaaca aacagtacgg tgtggtgctg caaattggtg gttctgacca gtggggtaac 540
atcacttctg gtatcgacct gacccgtcgt ctgcatcaga atcaggtg 588
<210>22
<211>600
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<220>
<221>misc_feature
<222>(403)..(403)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature
<222>(513)..(513)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature
<222>(515)..(515)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature

<222>(518)..(518)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature
<222>(531)..(531)
<223>n 为 a、c、g 或 t
<400>22
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcct gtgtggcttc 60
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 120
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 180
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 240
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 300
gctatcgcgg ccaataatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 360
gatattggca aacacttctc cgttaaccag atgatcaaca aanaagcggt taagcagcgt 420
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 480
tattcggctg cctgtgcgaa caaacagtac ggngnggngc tgcaaattgg nggttctgac 540
caggggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca aaatcaggtg 600
<210>23
<211>591
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<220>
<221>misc_feature
<222>(588)..(588)
<223>n 为 a、c、g 或 t
<400>23
gcgttagcag agcgactggc gcaaggcccg atcgcactcg tttgtggctt cgatcctacc 60
gctgacagct tgcatttggg gcatcttgtt ccattgttgt gcctgaaacg cttccagcag 120
gcgggccaca agccggttgc gctggtaggc ggcgcgacgg gtctgattgg cgacccgagc 180
ttcaaagctg ccgagcgtaa gctgaacacc gaagaaactg ttcaggagtg ggtggacaaa 240
atccgtaagc aggttgcccc gttcctcgat ttcgactgtg gagaaaactc tgctatcgcg 300
gccaataatt atgactggtt cggcaatatg aatgtgctga ccttcctgcg cgatattggc 360
aaacacttct ccgttaacca gatgatcaac aaagaagcgg ttaagcagcg tctcaaccgt 420
gaagatcagg ggatttcgtt cactgagttt tcctacaacc tgctgcaggg ttatagtgcg 480
gcctgtgtta acaaacagta cggtgtggtg ctgcaaattg gtggttctga ccagtggggt 540
aacatcactt ctggtatcga cctgacccgt cgtctgcatc agaatcangt g 591
<210>24

<211>600
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>24
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcat ttgtggcttc 60
gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 120
ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 180
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 240
gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 300
gctatcgcgg ccaatgatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 360
gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt 420
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 480
tataattttg cctgtgtgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 540
cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 600
<210>25
<211>579
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>25
cgactggcgc aaggcccgat cgcactcacg tgtggcttcg atcctaccgc tgacagcttg 60
catttggggc atcttgttcc attgttatgc ctgaaacgct tccagcaggc gggccacaag 120
ccggttgcgc tggtaggcgg cgcgacgggt ctgattggcg acccgagctt caaagctgcc 180
gagcgtaagc tgaacaccga agaaactgtt caggagtggg tggacaaaat ccgtaagcag 240
gttgccccgt tcctcgattt cgactgtgga gaaaactctg ctatcgcggc caataattat 300
gactggttcg gcaatatgaa tgtgctgacc ttcctgcgcg atattggcaa acacttctcc 360
gttaaccaga tgatcaacaa agaagcggtt aagcagcgtc tcaaccgtga agatcagggg 420
atttcgttca ctgagttttc ctacaatctg ctgcagggtt attcggctgc ctgtcttaac 480
aaacagtacg gtgtggtgct gcaaattggt ggttctgacc agtggggtaa catcacttct 540
ggtatcgacc tgacccgtcg tctgcatcag aatcaggtg 579
<210>26
<211>624
<212>DNA
<213> 人工
<220>
<223> 人工合成酶

<220>
<221>misc_feature
<222>(13)..(13)
<223>n 为 a、c、g 或 t
<220>
<221>misc_feature
<222>(599)..(599)
<223>n 为 a、c、g 或 t
<400>26
cgggggctgg tancccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcgggtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttattct atggcctgtt tgaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctganc 600
cgtcgtctgc atcagaatca ggtg 624
<210>27
<211>625
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<220>
<221>misc_feature
<222>(600)..(600)
<223>n 为 a、c、g 或 t
<400>27
cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttcctaca atctgctgca gggttattcg gctgcctgtc ttaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgaacctgan 600
ccgtcgtctg catcaaaatc aagtg 625
<210>28
<211>624
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>28
cgggggctgg taccccaagt gacggacgag gaaacgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tctcttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcaggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg atggcctgtg tgaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624
<210>29
<211>624
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>29
cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcgcgtgcgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaagg ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttattct tatgcctgtc ttaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624

<210>30
<211>624
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>30

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcgcgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg atggcctgtt gtaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624

<210>31
<211>624
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>31

cgggggctgg taccccaagt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcgctgag 480
ttttcctaca acctgctgca gggttatacg tttgcctgta tgaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624

<210>32
<211>606
<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>32

gtgacggacg aggaagcgtt agcagagcga ctggcgcaag gcccgatcgc actcacgtgt 60

ggcttcgatc ctaccgctga cagcttgcat ttggggcatc ttgttccatt gttatgcctg 120

aaacgcttcc agcaggcggg ccacaagccg gttgcgctgg taggcggcgc gacgggtctg 180

attggcgacc cgagcttcaa agctgccgag cgtaagctga acaccgaaga aactgttcag 240

gagtgggtgg acaaaatccg taagcaggtt gccccgttcc tcgatttcga ctgtggagaa 300

aactctgcta tcgcggccaa taattatgac tggttcggca atatgaatgt gctgaccttc 360

ctgcgcgata ttggcaaaca cttctccgtt aaccagatga tcaacaaaga agcggttaag 420

cagcgtctca accgtgaaga tcaggggatt tcgttcactg agttttccta caatctgctg 480

cagggttatt cggctgcctg tcttaacaaa cagtacggtg tggtgctgca aattggtggt 540

tctgaccagt ggggtaacat cacttctggt atcgacctga cccgtcgtct gcatcagaat 600

caggtg 606

<210>33

<211>624

<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<400>33

cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60

ccgatcgcac tcgtttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120

gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180

ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttcctaca acctgctgca gggttattcg atggcctgta cgaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210>34

<211>624

<212>DNA

<213> 人工

<220>

<223> 人工合成酶

<220>
<221>misc_feature
<222>(13)..(13)
<223>n 为 a、c、g 或 t
<400>34
cgggggctgg tancccaagt gacggacggg gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcagttgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatctcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatagt tttgcctgtc tgaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624
<210>35
<211>624
<212>DNA
<213> 人工
<220>
<223> 人工合成酶
<400>35
cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60
ccgatcgcac tcacgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt 120
gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180
ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac 240
accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300
gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360
atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420
aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480
ttttcctaca acctgctgca gggttatacg tttgcctgta ctaacaaaca gtacggtgtg 540
gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600
cgtcgtctgc atcagaatca ggtg 624
<210>36
<211>424
<212>PRT
<213> 人工
<220>

<223> 人工合成酶

<400>36

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Tyr Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala

290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420
<210>37
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>37
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp

115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>38
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>38

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420
<210>39
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>39
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu

385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>40
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>40

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Thr Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr

210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu

225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe

245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu

260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu

275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala

290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala

305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser

325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu

340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu

355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile

370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu

385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys

405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys

420

<210>41

<211>424

<212>PRT

<213> 人工

<220>

<223> 人工合成酶

<400>41

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val

1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly

20 25 30

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His

35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala

50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly

65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr

85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu

100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp

115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys

130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg

145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn

165 170 175

Leu Leu Gln Gly Tyr Thr Tyr Ala Cys Leu Asn Lys Gln Tyr Gly Val

180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly

195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr

210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu

225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe

245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu

260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu

275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala

290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala

305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser

325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu

340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>42
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>42

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Ser Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>43
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>43

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala

305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>44
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>44

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys

130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Arg Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>45

<211>424

<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>45
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Gly Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>46
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>46

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Gly Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg AspIle Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Gly Phe Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys

405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys

420

<210>47

<211>424

<212>PRT

<213> 人工

<220>

<223> 人工合成酶

<400>47

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30

Pro Ile Ala Leu Gly Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Gly Tyr Ala Cys Met Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu

225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>48
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>48

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala

50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>49
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>49

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Ala Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>50
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>50

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Ala Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser

325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu

340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu

355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile

370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu

385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys

405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys

420

<210>51

<211>424

<212>PRT

<213> 人工

<220>

<223> 人工合成酶

<400>51

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val

1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly

20 25 30

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His

35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala

50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly

65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr

85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu

100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp

115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys

130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg

145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Ala Ala Cys Val Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>52
<211>424
<212>PRT
<213>人工

<220>

<223> 人工合成酶

<400>52

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30

Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asp Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Asn Phe Ala Cys Val Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>53
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>53

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Ala Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys

420

<210>54

<211>424

<212>PRT

<213> 人工

<220>

<223> 人工合成酶

<400>54

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Gly Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe

245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>55
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>55

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly

65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Ala Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>56

<211>424

<212>PRT

<213> 人工

<220>

<223> 人工合成酶

<400>56

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Ser Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Thr Met Ala Cys Val Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>57
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>57

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30

Pro Ile Ala Leu Ala Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Tyr Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu

340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420
<210>58
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>58
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Ala Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn

165 170 175
Leu Leu Gln Gly Tyr Thr Met Ala Cys Cys Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>59
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶

<400>59

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Phe Ala Cys Met Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>60
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>60

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Val Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>61

<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>61
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Met Ala Cys Thr Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu

260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420
<210>62
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>62
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30
Pro Ile Ala Leu Ser Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr

85 90 95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160
Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175
Leu Leu Gln Gly Tyr Ser Phe Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>63

<211>424

<212>PRT

<213> 人工

<220>

<223> 人工合成酶

<400>63

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Phe Ala Cys Thr Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>64
<211>129
<212>DNA
<213>Escherichia coli
<400>64
agcttcccga taagggagca ggccagtaaa aagcattacc ccgtggtggg gttcccgagc 60
ggccaaaggg agcagactct aaatctgccg tcatcgacct cgaaggttcg aatccttccc 120
ccaccacca 129
<210>65
<211>129
<212>RNA
<213>Escherichia coli
<400>65

agcuucccga uaagggagca ggccaguaaa aagcauuacc ccgugguggg guucccgagc 60
ggccaaaggg agcagacucu aaaucugccg ucaucgaccu cgaagguucg aauccuuccc 120
ccaccacca 129

<210>66
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>66
atgaagtagc tgtcttctat cgaacaagca tgcg 34

<210>67
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>67
cgaacaagca tgcgattagt gccgacttaa aaag 34

<210>68
<211>33
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>68
cgctactctc ccaaatagaa aaggtctccg ctg 33

<210>69
<211>32
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>69
ctggaacagc tatagctact gatttttcct cg 32

<210>70
<211>34
<212>DNA
<213> 人工

<220>
<223> 寡核苷酸引物
<400>70
gccgtcacag attagttggc ttcagtggag actg 34
<210>71
<211>33
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>71
gattggcttc ataggagact gatatgctct aac 33
<210>72
<211>33
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>72
gcctctatag ttgagacagc atagaataat gcg 33
<210>73
<211>35
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>73
gagacagcat agatagagtg cgacatcatc atcgg 35
<210>74
<211>37
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>74
gaataagtgc gacatagtca tcggaagaga gtagtag 37
<210>75
<211>35
<212>DNA

<213> 人工
<220>
<223> 寡核苷酸引物
<400>75
ggtcaaagac agttgtaggt atcgattgac tcggc 35
<210>76
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>76
cgctactctc cccaaattta aaaggtctcc gctg 34
<210>77
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>77
cgctactctc cccaaatata aaaggtctcc gctg 34
<210>78
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>78
cgctactctc cccaaatgga aaaggtctcc gctg 34
<210>79
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>79
cgctactctc cccaaagata aaaggtctcc gctg 34
<210>80
<211>34

<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>80
cgctactctc cccaaaaaaa aaaggtctcc gctg 34
<210>81
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>81
gccgtcacag attttttggc ttcagtggag actg 34
<210>82
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>82
gccgtcacag attatttggc ttcagtggag actg 34
<210>83
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>83
gccgtcacag attggttggc ttcagtggag actg 34
<210>84
<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>84
gccgtcacag atgatttggc ttcagtggag actg 34
<210>85

<211>34
<212>DNA
<213> 人工
<220>
<223> 寡核苷酸引物
<400>85
gccgtcacag ataaattggc ttcagtggag actg 34
<210>86
<211>424
<212>PRT
<213> 人工
<220>
<223> 人工合成酶
<400>86

```
Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1               5                   10                  15
Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
            20                  25                  30
Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
        35                  40                  45
Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
    50                  55                  60
Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65                  70                  75                  80
Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
                85                  90                  95
Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
            100                 105                 110
Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
        115                 120                 125
Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
    130                 135                 140
His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145                 150                 155                 160
Leu Asn Arg Glu Gly Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
                165                 170                 175
Leu Leu Gln Gly Tyr Gly Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
            180                 185                 190
Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
```

195 200 205
Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220
Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240
Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255
Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270
Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285
Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300
Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320
Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335
Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350
Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365
Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380
Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400
Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415
Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>87
<211>40
<212>DNA
<213>Homo sapiens
<400>87
ctgtgctgaa cctcagggga cgccgacaca cgtacacgtc 40

<210>88
<211>42
<212>DNA
<213>Homo sapiens

<400>88
gacaagtgcg gtttttttct ccagctcccg atgacttatg gc 42
<210>89
<211>42
<212>DNA
<213> 人工
<220>
<223> 引物
<400>89
gtacgaattc ccgagatctc tgtgctgaac ctcaggggac gc 42
<210>90
<211>80
<212>DNA
<213> 人工
<220>
<223> 引物
<400>90
gatgcaagct tgatggatcc gccataagtc atcgggagct ggagaaaaaa accgcacttg 60
tctggtgggg gaaggattcg 80
<210>91
<211>95
<212>DNA
<213>Homo sapiens
<400>91
ggattacgca tgctcagtgc aatcttcggt tgcctggact agcgctccgg tttttctgtg 60
ctgaacctca ggggacgccg acacacgtac acgtc 95
<210>92
<211>11
<212>DNA
<213>Homo sapiens
<400>92
gacaagtgcg g 11
<210>93
<211>29
<212>DNA
<213> 人工
<220>
<223> 引物
<400>93

ctttgtgtaa tacttgtaac gctgaattc 29
<210>94
<211>23
<212>DNA
<213> 人工
<220>
<223> 引物
<400>94
accatgatta cgccaagctt gat 23
<210>95
<211>31
<212>DNA
<213> 人工
<220>
<223> 引物
<400>95
gcatcggatc cggtggggtt cccgagcggc c 31
<210>96
<211>44
<212>DNA
<213> 人工
<220>
<223> 引物
<400>96
acgccaagct tttccaaaat ggtgggggaa ggattcgaac cttc 44
<210>97
<211>34
<212>DNA
<213> 人工
<220>
<223> 引物
<400>97
gcatcggatc cgtgctgaac ctcaggggac gccg 34
<210>98
<211>46
<212>DNA
<213> 人工
<220>
<223> 引物

<400>98
acgccaagct tttccaaaaa atggtggggg aaggattcga accttc 46
<210>99
<211>45
<212>DNA
<213> 人工
<220>
<223> 引物
<400>99
acgccaagct tttccaaaaa accgcacttg tctggtgggg gaagg 45
<210>100
<211>112
<212>DNA
<213> 人工
<220>
<223> 引物
<400>100
gcatcggatc cggtggggtt cccgagcggc caaagggagc agactctaaa tctgccgtca 60
cagacttcga aggttcgaat ccttccccca ccattttgga aaagcttggc gt 112
<210>101
<211>114
<212>DNA
<213> 人工
<220>
<223> 引物
<400>101
gcatcggatc cggtggggtt cccgagcggc caaagggagc agactctaaa tctgccgtca 60
cagacttcga aggttcgaat ccttccccca ccatttttg gaaaagcttg gcgt 114
<210>102
<211>152
<212>DNA
<213> 人工
<220>
<223> 引物
<400>102
gcatcggatc cgtgctgaac ctcaggggac gccgacacac gtacacgtcg gtggggttcc 60
cgagcggcca aagggagcag actctaaatc tgccgtcaca gacttcgaag gttcgaatcc 120
ttcccccacc attttttgga aaagcttggc gt 152
<210>103

<211>125
<212>DNA
<213> 人工
<220>
<223> 引物
<400>103
gcatcggatc cggtggggtt cccgagcggc caaagggagc agactctaaa tctgccgtca 60
cagacttcga aggttcgaat ccttccccca ccagacaagt gcggtttttt ggaaaagctt 120
ggcgt 125
<210>104
<211>163
<212>DNA
<213> 人工
<220>
<223> 引物
<400>104
gcatcggatc cgtgctgaac ctcaggggac gccgacacac gtacacgtcg gtggggttcc 60
cgagcggcca aagggagcag actctaaatc tgccgtcaca gacttcgaag gttcgaatcc 120
ttcccccacc agacaagtgc ggttttttgg aaaagcttgg cgt 163
<210>105
<211>114
<212>DNA
<213> 人工
<220>
<223> 质粒构筑体
<400>105
gcatcggatc cggaggggta gcgaagtggc taaacgcggc ggactctaaa tccgctccct 60
ttgggttcgg cggttcgaat ccgtcccccct ccattttttg gaaaagcttg gcgt 114
<210>106
<211>112
<212>DNA
<213> 人工
<220>
<223> 质粒构筑体
<400>106
gcatcggatc cggaggggta gcgaagtggc taaacgcggc ggactctaaa tccgctccct 60
ttgggttcgg cggttcgaat ccgtcccccct ccattttgga aaagcttggc gt 112
<210>107
<211>152

<212>DNA
<213> 人工
<220>
<223> 质粒构筑体
<400>107
gcatcggatc cgtgctgaac ctcaggggac gccgacacac gtacacgtcg gaggggtagc 60
gaagtggcta aacgcggcgg actctaaatc cgctcccttt gggttcggcg gttcgaatcc 120
gtccccctcc attttttgga aaagcttggc gt 152
<210>108
<211>125
<212>DNA
<213> 人工
<220>
<223> 质粒构筑体
<400>108
gcatcggatc cggaggggta gcgaagtggc taaacgcggc ggactctaaa tccgctccct 60
ttgggttcgg cggttcgaat ccgtccccct ccagacaagt gcggtttttt ggaaaagctt 120
ggcgt 125
<210>109
<211>163
<212>DNA
<213> 人工
<220>
<223> 质粒构筑体
<400>109
gcatcggatc cgtgctgaac ctcaggggac gccgacacac gtacacgtcg gaggggtagc 60
gaagtggcta aacgcggcgg actctaaatc cgctcccttt gggttcggcg gttcgaatcc 120
gtccccctcc agacaagtgc ggttttttgg aaaagcttgg cgt 163
<210>110
<211>333
<212>DNA
<213> 人工
<220>
<223> 启动子
<400>110
cccagtggaa agacgcgcag gcaaaacgca ccacgtgacg gagcgtgacc gcgcgccgag 60
cgcgcgccaa ggtcgggcag gaagagggcc tatttcccat gattccttca tatttgcata 120
tacgatacaa ggctgttaga gagataatta gaattaattt gactgtaaac acaaagatat 180
tagtacaaaa tacgtgacgt agaaagtaat aatttcttgg gtagtttgca gttttaaaat 240

tatgttttaa aatggactat catatgctta ccgtaacttg aaagtatttc gatttcttgg 300
gtttatatat cttgtggaaa ggacgcggga tcc 333
<210>111
<211>105
<212>DNA
<213> 人工
<220>
<223> 启动子
<400>111
gaattcatat ttgcatgtcg ctatgtgttc tgggaaatca ccataaacgt gaaatgtctt 60
tggatttggg aatcttataa gttctgtatg agaccactcg gatcc 105
<210>112
<211>84
<212>DNA
<213> 人工
<220>
<223>E.coli Leu 琥珀抑制性 tRNA
<400>112
gcccggatgg tggaatcggt agacacaagg gattctaaat ccctcggcgt tcgcgctgtg 60
cgggttcaag tcccgctccg ggta 84

流程1

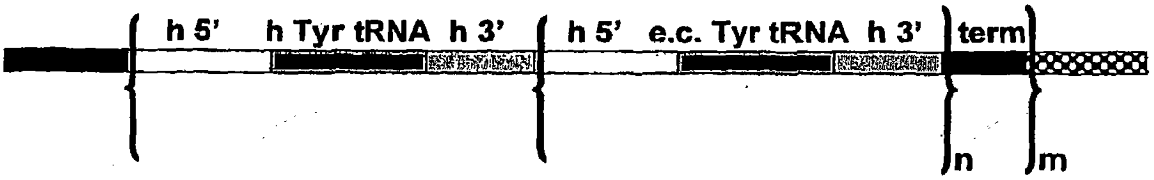

h 5'
h Tyr tRNA
h 3'
h 5'
e.c. Tyr tRNA
h 3'
term
n
m

图 1

流程2

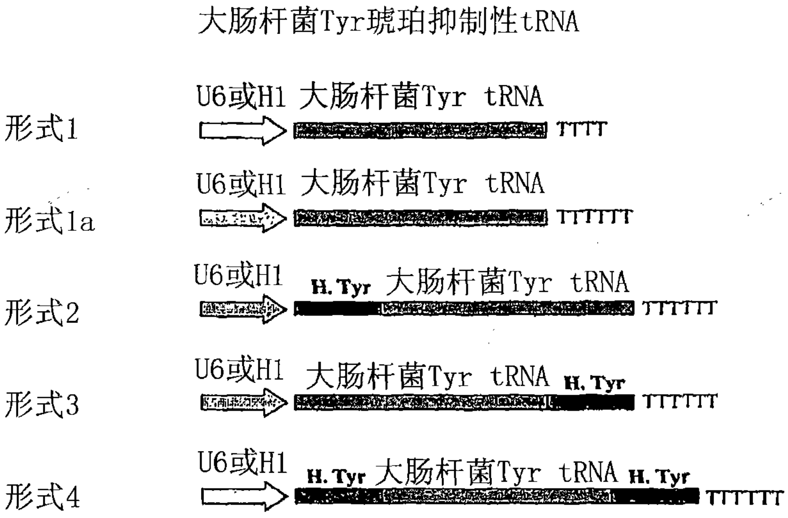

大肠杆菌Tyr琥珀抑制性tRNA
形式1
U6或H1 大肠杆菌Tyr tRNA
TTTT
形式1a
U6或H1 大肠杆菌Tyr tRNA
TTTTTT
形式2
U6或H1 H.Tyr 大肠杆菌Tyr tRNA
TTTTTT
形式3
U6或H1 大肠杆菌Tyr tRNA H.Tyr
TTTTTT
形式4
U6或H1 H.Tyr 大肠杆菌Tyr tRNA H.Tyr
TTTTTT

图 2

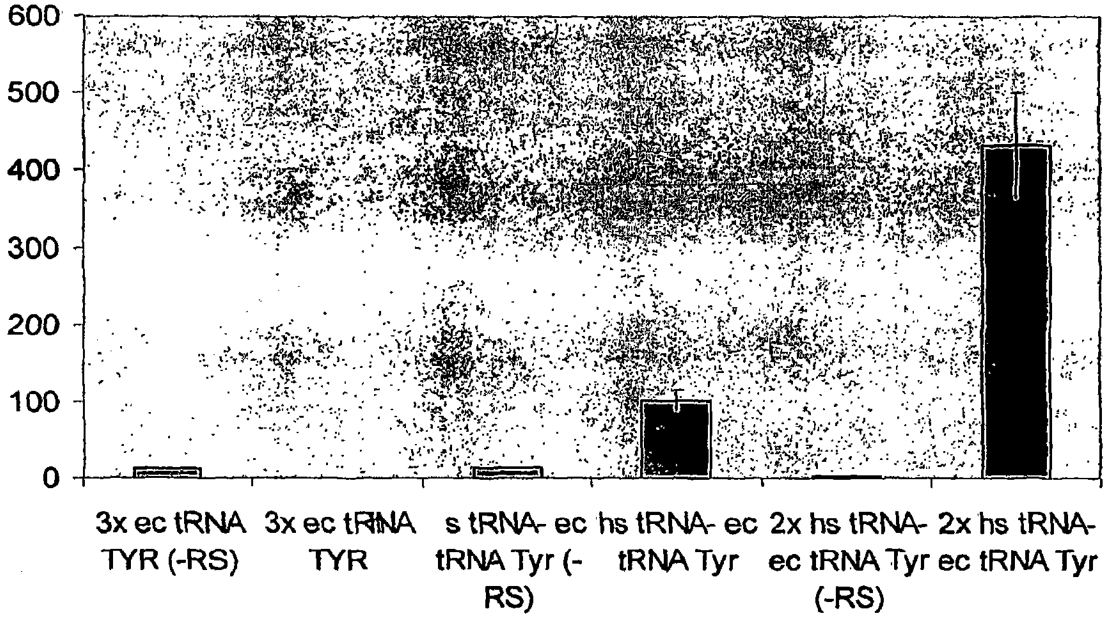

600
500
400
300
200
100
0
3x ec tRNA TYR (-RS)
3x ec tRNA TYR
s tRNA- ec tRNA Tyr (-RS)
hs tRNA- ec tRNA Tyr
2x hs tRNA- ec tRNA Tyr (-RS)
2x hs tRNA- ec tRNA Tyr

图 3

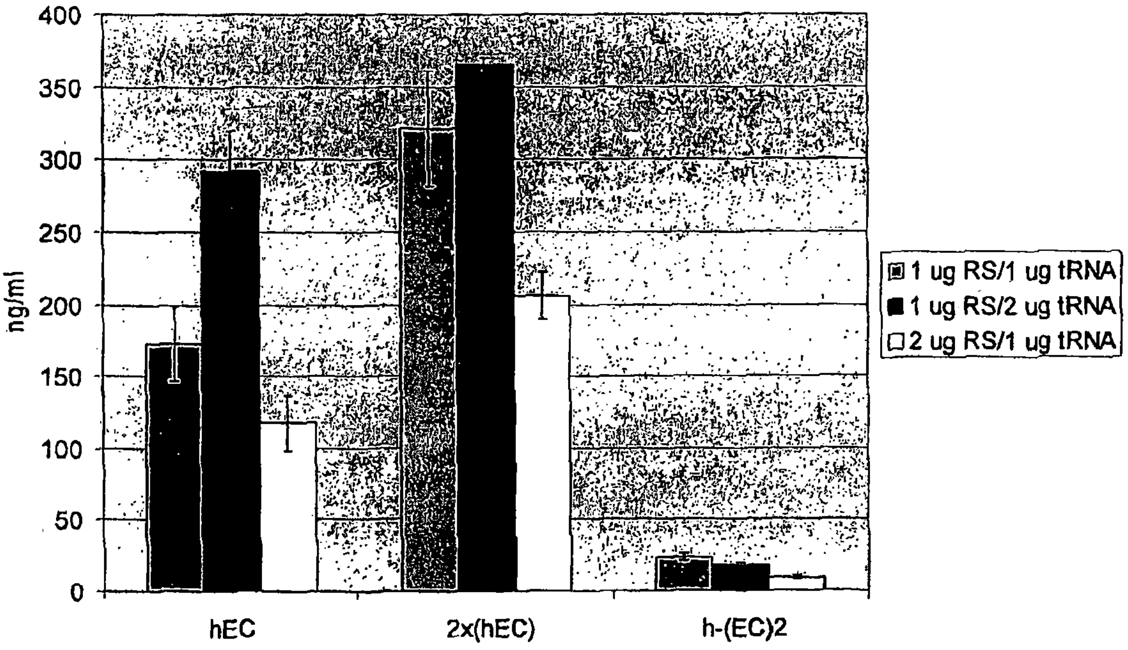

400
350
300
250
200
150
100
50
0
ng/ml
hEC
2x(hEC)
h-(EC)2
1 ug RS/1 ug tRNA
1 ug RS/2 ug tRNA
2 ug RS/1 ug tRNA

图 4

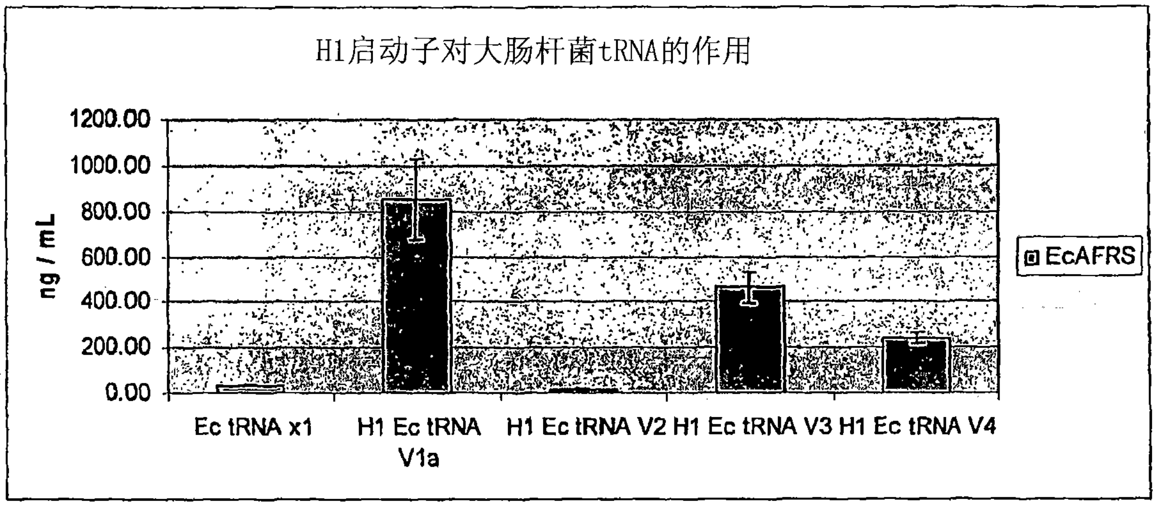

H1启动子对大肠杆菌tRNA的作用
1200.00
1000.00
800.00
600.00
400.00
200.00
0.00
ng / mL
EcAFRS
Ec tRNA x1
H1 Ec tRNA V1a
H1 Ec tRNA V2
H1 Ec tRNA V3
H1 Ec tRNA V4

图 5

流程3:

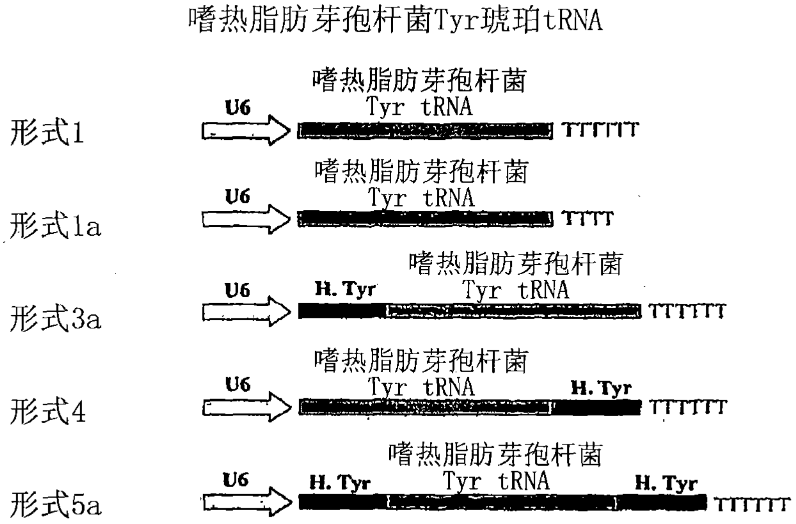
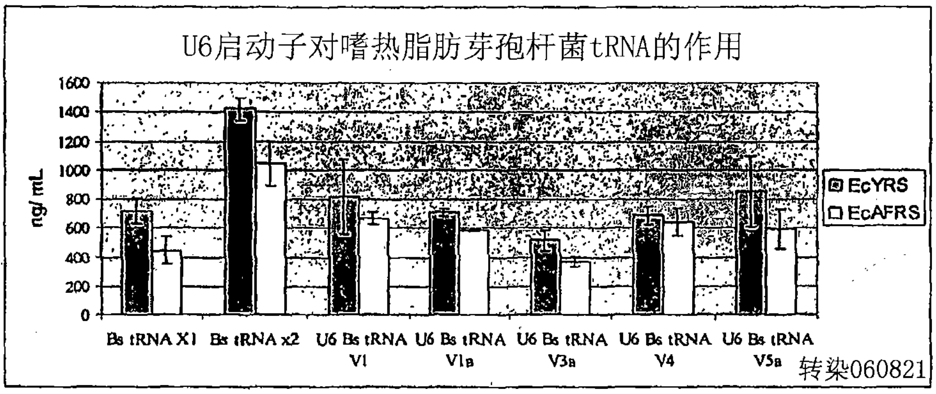
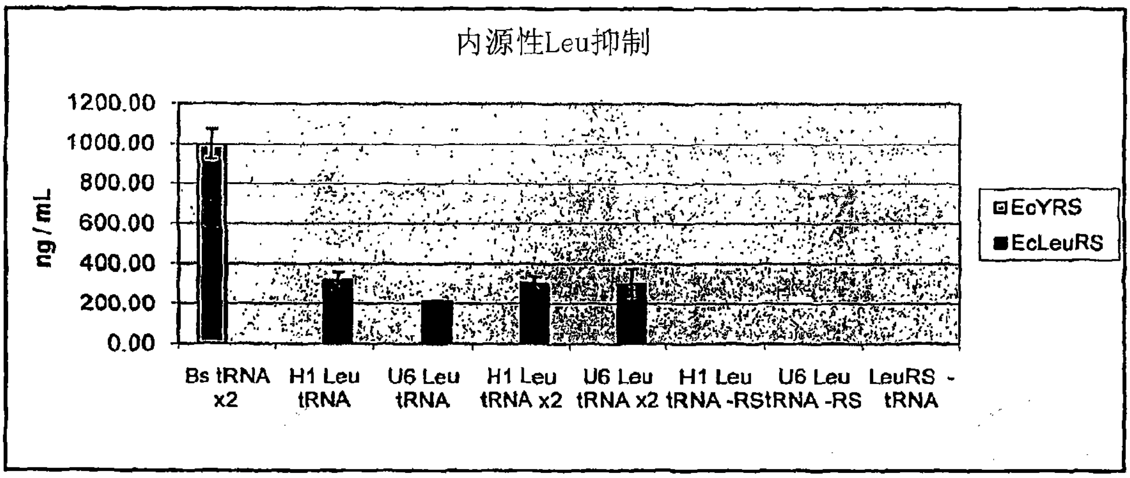

嗜热脂肪芽孢杆菌Tyr琥珀tRNA
形式1
U6
嗜热脂肪芽孢杆菌
Tyr tRNA
TTTTTT
形式1a
U6
嗜热脂肪芽孢杆菌
Tyr tRNA
TTTT
形式3a
U6
H. Tyr
嗜热脂肪芽孢杆菌
Tyr tRNA
TTTTTT
形式4
U6
嗜热脂肪芽孢杆菌
Tyr tRNA
H. Tyr
TTTTTT
形式5a
U6
H. Tyr
嗜热脂肪芽孢杆菌
Tyr tRNA
H. Tyr
TTTTTT

图 6



• Lipofectamine 2000：转染后40小时
• 1ug E88am hGH / 1ug Bs tRNA/ 1ug RS / 1ug lacZ
• 由新制的1 M pAF制成的1 mM pAF CHO培养基

图 7



Lipofectamine 2000：转染后42小时

1ug E88am hGH / 1ug tRNA / 1ug RS / 1ug lacZ

约60%的转染效率

图 8