



(12) 发明专利申请

(10) 申请公布号 CN 102888387 A

(43) 申请公布日 2013. 01. 23

(21) 申请号 201110205760. 2

(22) 申请日 2011. 07. 21

(71) 申请人 中国科学院生物物理研究所
地址 100101 北京市朝阳区大屯路 15 号

(72) 发明人 王江云 张维 江欢欢

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 王旭

(51) Int. Cl.

C12N 9/10 (2006. 01)

C12N 15/54 (2006. 01)

C12N 1/21 (2006. 01)

C12N 15/63 (2006. 01)

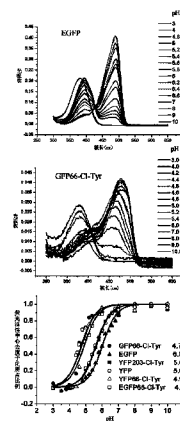
权利要求书 1 页 说明书 10 页
序列表 12 页 附图 7 页

(54) 发明名称

3- 氯代酪氨酸翻译系统及其应用

(57) 摘要

本发明涉及氨酰基-tRNA 合成酶突变体, 其含有的氨基酸序列选自 SEQ ID NO :2、3、4 所示氨基酸和它们的保守性变体构成的组。本发明提供利用正交 tRNA、正交氨酰基-tRNA 合成酶和它们的配对将 3- 氯代酪氨酸掺入目标蛋白质的 3- 氯代酪氨酸翻译系统, 和利用所述翻译系统在目标蛋白质中掺入 3- 氯代酪氨酸的方法。所述 3- 氯代酪氨酸翻译系统包含:(i) 3- 氯代酪氨酸;(ii) 正交氨酰基-tRNA 合成酶;(iii) 正交 tRNA, 其中所述正交氨酰基-tRNA 合成酶用所述 3- 氯代酪氨酸优先氨酰化所述正交 tRNA; 和 (iv) 编码目标蛋白质的核酸, 其中所述核酸含有所述正交 tRNA 特异性识别的至少一个选择密码子。



1. 一种 3- 氯代酪氨酸翻译系统,所述系统包含:
 - (i) 3- 氯代酪氨酸;
 - (ii) 正交氨酰基 -tRNA 合成酶;
 - (iii) 正交 tRNA,其包含 SEQ ID NO:1 所示的多核苷酸序列;其中所述正交氨酰基 -tRNA 合成酶用所述 3- 氯代酪氨酸优先氨酰化所述正交 tRNA ;和
 - (iv) 编码目标蛋白质的核酸,其中所述核酸含有所述正交 tRNA 特异性识别的至少一个选择密码子。
2. 如权利要求 1 所述的翻译系统,其特征在于,所述正交氨酰基 -tRNA 合成酶含有的氨基酸序列选自 SEQ ID NO:2、3、4 所示氨基酸和它们的保守性变体构成的组。
3. 如权利要求 1 所述的翻译系统,其特征在于,所述正交 tRNA 是琥珀抑制型 tRNA,并且所述选择密码子是琥珀密码子。
4. 如权利要求 1 所述的翻译系统,其还包含编码正交氨酰基 -tRNA 合成酶的核苷酸序列。
5. 一种宿主细胞,其包含所述正交 tRNA 序列和编码所述正交氨酰基 -tRNA 合成酶的核苷酸序列。
6. 如权利要求 5 所述的宿主细胞,其中所述宿主细胞是真细菌细胞,优选大肠杆菌细胞。
7. 一种产生在至少一个所选位置掺入 3- 氯代酪氨酸的突变蛋白质的方法,所述方法包括下述步骤:
 - (a) 提供权利要求 1 所述的 3- 氯代酪氨酸翻译系统,该系统包含:
 - (i) 3- 氯代酪氨酸;
 - (ii) 正交氨酰基 -tRNA 合成酶;
 - (iii) 正交 tRNA,其包含 SEQ ID NO:1 所示的多核苷酸序列;其中所述正交氨酰基 -tRNA 合成酶用所述 3- 氯代酪氨酸优先氨酰化所述正交 tRNA ;和
 - (iv) 编码所述目标蛋白质的核酸,其中所述核酸在所选的位置包含所述正交 tRNA 特异性识别的至少一个选择密码子;和
 - (b) 将编码所述目标蛋白质的核酸转化到权利要求 5 所述的宿主细胞中,在所述蛋白质的翻译期间,3- 氯代酪氨酸氨酰化的正交 tRNA 对所述选择密码子起反应而将培养基中的 3- 氯代酪氨酸掺入所述目标蛋白质的所述所选位置,从而产生在所选位置含 3- 氯代酪氨酸的所述目标蛋白质。
8. 如权利要求 7 所述的方法,其中所述正交氨酰基 -tRNA 合成酶含有的氨基酸序列选自 SEQ ID NO:2、3、4 所示氨基酸和它们的保守性变体构成的组。
9. 如权利要求 7 所述的方法,其中所述正交 tRNA 是琥珀抑制型 tRNA,并且所述选择密码子是琥珀密码子。
10. 一种氨酰基 tRNA 合成酶突变体,其含有的氨基酸序列选自 SEQ ID NO:2、3、4 所示氨基酸和它们的保守性变体构成的组。

3- 氯代酪氨酸翻译系统及其应用

技术领域

[0001] 本发明属于生物化学领域。具体地,本发明提供氨酰基-tRNA 合成酶突变体,其含有的氨基酸序列选自 SEQ ID NO :2、3、4 所示氨基酸和它们的保守性变体构成的组。本发明还涉及一种 3- 氯代酪氨酸 (3-Cl-Tyr) 翻译系统。更具体地,本发明涉及利用正交 tRNA、正交氨酰基-tRNA 合成酶和它们的配对将 3- 氯代酪氨酸掺入目标蛋白质的 3- 氯代酪氨酸翻译系统,和利用所述翻译系统在目标蛋白质中掺入 3- 氯代酪氨酸的方法。本发明还涉及用这种翻译系统和这种方法产生的含有 3- 氯代酪氨酸的突变蛋白质。

背景技术

[0002] 生物体氧化应激状态产生的活性物质一方面可以起到防御和免疫作用,另一方面会对自身蛋白造成氧化损伤,其中蛋白酪氨酸残基氯代是一种常见的氧化修饰,目前已知这种氧化修饰与帕金森氏病,哮喘,动脉粥样硬化,急性心肌梗死等疾病密切相关 (Choi DK, Pennathur S, Przedborski S, et al. J Neurosci 2005,25(28) :6594-6600 ;Aldridge RE, Chan T, Kettle AJ, et al. Free Radic Biol Med 2002,33(6) :847-856.)。载脂蛋白 ApoA1 是高密度脂蛋白主要成分,介导胆固醇从血管壁到血液反向传输。ApoA1 192 位酪氨酸是髓过氧化物酶氯代修饰的主要位点,在动脉粥样硬化患者中发现有高水平的 ApoA1 氯代 (Heinecke JW. Am J Cardiol 2003,91(3A) :12A-16A.)。目前研究 ApoA1 氯代与功能主要采用位点突变和氧化剂氧化的方法,除了 192 位酪氨酸可以被氯化,其它氨基酸和位点也可以被氧化。从而使目前 ApoA1 氧化与功能损伤机理仍然存在争议 (Shao B, Bergt C, Heinecke JW, et al. J Biol Chem 2005,280(7) :5983-5993 ;Peng DQ, Wu Z, Brubaker G, Smith JD, et al. J Biol Chem 2005,280(40) :33775-33784.)。

[0003] 另一方面,生物体内有一些酸性的环境,如溶酶体 (lysosome),以及与溶酶体相关的自噬体 (autophagosome) 及抗原呈递过程中的吞噬酶体 (phagolysosome),它们内部 pH 低于 6,在 4-5 左右。目前用来标记这类酸性细胞器主要使用 EGFP-LC3 载体 (Ni HM, Bockus A, Ding WX, et al. Autophagy,7(2) :188-204.)。但是当 EGFP 进入酸性细胞器内部,会引起 488nm 激发荧光减弱。目前通过引入红色荧光蛋白 mRFP 来标记自噬体 (Kimura S, Noda T, Yoshimori T. Autophagy 2007,3(5) :452-460.)。由于 mRFP pKa 为 4.5,在 600nm 处发红光,在自噬体中不影响其荧光亮度。

[0004] 虽然目前有红色荧光蛋白 dsRed pKa 4.3 (Baird GS, Zacharias DA, Tsien RY. Proc Natl Acad Sci U S A 2000,97(22) :11984-11989.),青色荧光蛋白 ECFP pKa 4.7 (Patterson G, Day RN, Piston D. J Cell Sci 2001,114(Pt5) :837-838.),但是目前使用的绿色荧光蛋白 EGFP pKa 6.0,黄色荧光蛋白 YFP pKa 5.6 均不太适合标记溶酶体等酸性细胞器,而研究细胞内蛋白与蛋白相互作用,以及多种蛋白的动态情况,往往需要多种不同颜色的荧光蛋白来标记。

[0005] 对比 Tyr, Cl-Tyr 的酚环羟基更容易解离去质子化,具有更低的 pKa 值。在绿色荧光蛋白 GFP 中包含有 Ser-Tyr-Gly 交联形成的荧光活性中心 (Reid B G, Flynn

GC. Biochemistry 1997, 36 (22) :6786-6791.), 我们设想在 GFP 荧光活性中心用 Cl-Tyr 取代 Tyr 会降低荧光活性中心 pKa, 使其在酸性条件下, 荧光活性中心主要以去质子化形式存在。由于中性状态 GFP 荧光活性中心激发峰波长为 397nm, 在紫外区域, 这个波段的光会给细胞带来光损伤。而去质子化状态 GFP 荧光活性中心激发峰波长在 470nm, 不会给细胞带来损伤。同时我们也可对单分子成像中应用最突出的一种荧光蛋白 mEOS2 荧光活性中心用 3-Cl-Tyr 取代 Tyr 来研究其 pKa 及光学性质 (McKinney SA, Murphy CS, Looger LL. A, et al. Nat Methods. 2009, 6 (2) :131-133.)。

[0006] 为了研究这些蛋白质的结构和功能, 本领域需要能将非天然 3- 氯代酪氨酸掺入蛋白质的新方案。现已开发了在原核和真核生物中将各种非天然氨基酸体内位点特异性地掺入蛋白质的通用方法。这些方法依赖于正交蛋白质翻译组分, 所述组分识别合适的选择密码子 (selector codon) 从而能在体内多肽翻译期间将所需的非天然氨基酸插入限定位置。这些方法利用识别选择密码子的正交 tRNA (O-tRNA), 而相应的特异性正交氨酰基-tRNA 合成酶 (O-RS) 用非天然氨基酸加载该 O-tRNA。这些组分不与宿主生物体内的任何内源性 tRNA、氨酰基-tRNA 合成酶 (RS)、氨基酸或密码子交叉反应 (即, 它必须是正交的)。利用这种正交 tRNA-RS 配对可能遗传编码大量结构各异的非天然氨基酸。

[0007] 本领域普遍知道利用适合于制备含一个或多个非天然氨基酸的蛋白质的正交翻译系统, 例如产生正交翻译系统的通用方法。例如, 参见国际公布号 WO 2002/086075, 其名为“METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL-tRNA SYNTHETASE PAIRS”; WO 2002/085923, 其名为“IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS”; WO 2004/094593, 其名为“EXPANDING THE EUKARYOTIC GENETIC CODE”。掺入非天然氨基酸的正交翻译系统及它们的产生和使用方法的其他讨论还可参见 Wang 和 Schultz, Chem. Commun. (Camb) 1 :1-11 (2002) ; Wang 和 Schultz, Angewandte Chemie Int. Ed. 44 (1) :34-66 (2005) ; Xie 和 Schultz, Methods 36 (3) :227-238 (2005) ; Xie 和 Schultz, Curr. Opinion in Chemical Biology 9 (6) :548-554 (2005) ; Wang 等, Annu. Rev. Biophys. Biomol. Struct. 35 :225-249 (2006)。

发明内容

[0008] 1、技术问题

[0009] 本发明提供氨酰基-tRNA 合成酶突变体, 其含有的氨基酸序列选自由 SEQ ID NO : 2、3、4 所示氨基酸和它们的保守性变体构成的组。本发明涉及利用正交 tRNA、正交氨酰基-tRNA 合成酶和它们的配对将 3- 氯代酪氨酸掺入目标蛋白质的 3- 氯代酪氨酸翻译系统, 和利用所述翻译系统在目标蛋白质中掺入 3- 氯代酪氨酸的方法。本发明还涉及用这种翻译系统和这种方法产生的含有 3- 氯代酪氨酸的突变蛋白质。

[0010] 因此, 本发明的目的在于提供利用正交 tRNA、正交氨酰基-tRNA 合成酶和它们的配对将 3- 氯代酪氨酸掺入蛋白质的 3- 氯代酪氨酸翻译系统, 并且提供该翻译系统在目标蛋白质中掺入 3- 氯代酪氨酸的方法。

[0011] 在本发明的优选方面中, 本发明人利用这种方法将 3- 氯代酪氨酸分别掺入肌红蛋白 (myoglobin)、ApoA1、mEOS2 和 GFP 及其系列突变体中来研究这些蛋白的性质。

[0012] 2、技术方案

[0013] 本发明提供在体内（例如在宿主细胞内）对选择密码子（selector codon）如琥珀终止密码子（TAG）起反应而将非天然氨基酸 3-氯代酪氨酸掺入延伸中的多肽链的 3-氯代酪氨酸翻译系统。所述 3-氯代酪氨酸翻译系统包含不与宿主细胞翻译机制相互作用的正交-tRNA (O-tRNA) 和正交氨酰基-tRNA 合成酶 (O-RS) 配对。即，宿主细胞内源性氨酰基-tRNA 合成酶不会用氨基酸（天然的或非天然的）加载 O-tRNA。类似地，本发明提供的 O-RS 不以显著水平或者某些情况下不以可检测水平地用氨基酸（天然的或非天然的）加载内源性 tRNA。利用所述翻译系统能够产生含有在翻译过程中掺入 3-氯代酪氨酸的大量蛋白质。

[0014] 在一些方面中，本发明提供 3-氯代酪氨酸翻译系统。所述翻译系统包含：(a) 非天然氨基酸，即 3-氯代酪氨酸，(b) 正交氨酰-tRNA 合成酶 (O-RS)，和 (c) 正交 tRNA (O-tRNA)，其包含 SEQ ID NO:1 所示的多核苷酸序列，其中所述正交氨酰-tRNA 合成酶用所述非天然氨基酸（即 3-氯代酪氨酸），优先氨酰化所述 O-tRNA。

[0015] 优选地，本发明的 3-氯代酪氨酸翻译系统还包含编码目标蛋白质的核酸，其中所述核酸含有由正交 tRNA (O-tRNA) 特异性识别的至少一个选择密码子，优选地为琥珀密码子。更优选地，本发明的 3-氯代酪氨酸翻译系统还包含编码正交氨酰基-tRNA 合成酶的核苷酸序列。

[0016] 所述系统中所用的正交氨酰基-tRNA 合成酶 (O-RS) 即为本发明人发现的氨酰基 tRNA 合成酶突变体，其含有的氨基酸序列选自 SEQ ID NO:2、3、4 所示氨基酸和它们的保守性变体构成的组。

[0017] 在本发明的优选方面中，本发明提供一种 3-氯代酪氨酸翻译系统，所述系统包含：

[0018] (i) 3-氯代酪氨酸；

[0019] (ii) 正交氨酰基-tRNA 合成酶；

[0020] (iii) 正交 tRNA，其包含 SEQ ID NO:1 所示的多核苷酸序列；其中所述正交氨酰基-tRNA 合成酶用所述 3-氯代酪氨酸优先氨酰化所述正交 tRNA；和

[0021] (iv) 编码目标蛋白质的核酸，其中所述核酸含有所述正交 tRNA 特异性识别的至少一个选择密码子。

[0022] 优选地，所述 3-氯代酪氨酸翻译系统还包含编码正交氨酰基-tRNA 合成酶的核苷酸序列。

[0023] 该翻译系统中的各种组分可以衍生自各种物种来源，例如，该翻译系统中的各组分衍生自詹氏甲烷球菌 (*Methanococcus jannaschii*)。例如，正交 tRNA (O-tRNA) 为古菌来源的反密码子突变为与琥珀密码互补的酪氨酸 tRNA。在一些实施方式中，O-tRNA 是琥珀抑制型 tRNA。在一些实施方式中，O-tRNA 包含 SEQ ID NO:1 所示的多核苷酸序列，优选地，O-tRNA 的序列如 SEQ ID NO:1 所示。在一个实施方式中，用于该系统的正交氨酰基-tRNA 合成酶 (O-RS) 可以包含 SEQ ID NO:2、3 或 4 所示的氨基酸序列及该序列的保守变体。

[0024] 在一些方面中，本发明的 3-氯代酪氨酸翻译系统还包含编码目标蛋白质的核酸，其中所述核酸具有由正交 tRNA (O-tRNA) 特异性识别的至少一个选择密码子。在优选方面中，所述正交 tRNA 是琥珀抑制型 tRNA，并且所述选择密码子是琥珀密码子。

[0025] 在一些方面中，本发明提供包含正交 tRNA 序列和编码正交氨酰基-tRNA 合成酶的

核苷酸序列的宿主细胞。所用的宿主细胞不作具体限定,只要 O-RS 和 O-tRNA 在它们的宿主细胞环境中保留它们的正交性即可。例如,所述宿主细胞可以是真细菌细胞,如大肠杆菌。

[0026] 本发明还提供产生在至少一个所选位置掺入 3- 氯代酪氨酸的突变蛋白质的方法。所述方法利用上述 3- 氯代酪氨酸翻译系统。所述方法通常始于提供含有以下组分的 3- 氯代酪氨酸翻译系统的步骤:(i) 非天然氨基酸,即 3- 氯代酪氨酸;(ii) 正交氨酰基-tRNA 合成酶(O-RS);(iii) 正交 tRNA(O-tRNA),其包含 SEQ ID NO:1 所示的多核苷酸序列,其中所述 O-RS 用所述非天然氨基酸(即 3- 氯代酪氨酸)优先氨酰化所述 O-tRNA;和(iv) 编码目标蛋白质的核酸,其中所述核酸含有 O-tRNA 特异性识别的至少一个选择密码子(任选地为琥珀密码子);然后将编码所述目标蛋白质的核酸转化到包含正交 tRNA 序列和编码正交氨酰基-tRNA 合成酶的核苷酸序列的宿主细胞中,在所述蛋白质的翻译过程中,3- 氯代酪氨酸氨酰化的 O-tRNA 对所述选择密码子起反应而将培养基中的 3- 氯代酪氨酸掺入所述目标蛋白质的所选位置,从而产生在所选位置含有 3- 氯代酪氨酸的蛋白质。

[0027] 在所述方法的一些实施方式中,提供翻译系统的步骤包括通过定点诱变使野生型氨酰基-tRNA 合成酶的氨基酸结合口袋发生突变,选择用所述非天然氨基酸(即 3- 氯代酪氨酸)优先氨酰化所述 O-tRNA 的氨酰基-tRNA 合成酶突变体(即,本发明所用的正交氨酰基-tRNA 合成酶)。所述选择步骤包括定点诱变后从得到的氨酰基-tRNA 合成酶分子库进行所述 O-RS 的正选择和负选择(参见下述实施例 1)。在一些实施方式中,提供翻译系统的步骤还包括提供 O-tRNA 的序列,O-tRNA 为古菌来源的反密码子突变为与琥珀密码互补的酪氨酸 tRNA,例如,所述 O-tRNA 是琥珀抑制型 tRNA,或者 O-tRNA 包含 SEQ ID NO:1 所示的多核苷酸序列。在这些方法中,提供翻译系统的步骤还包括提供含有所述翻译系统所用的琥珀选择密码子的编码目标蛋白质的核酸。

[0028] 还可在宿主细胞内实施产生含有 3- 氯代酪氨酸的突变蛋白质的方法。在这些情况中,提供的宿主细胞包含本发明的 3- 氯代酪氨酸翻译系统(即,包含编码 O-RS 的核苷酸序列、O-tRNA 序列和含有至少一个选择密码子的编码目标蛋白质的核酸),而在适宜的培养条件下(例如,在培养基中添加 3- 氯代酪氨酸等)培养该宿主细胞可导致在所述目标蛋白质中掺入 3- 氯代酪氨酸。在一些实施方式中,提供步骤包括提供真细菌宿主细胞(例如,大肠杆菌)。

[0029] 本发明还提供改变荧光蛋白 pKa 及光学性质的方法,所述方法利用上述 3- 氯代酪氨酸翻译系统。这些方法通常始于提供含有以下组分的 3- 氯代酪氨酸翻译系统的步骤:(i) 3- 氯代酪氨酸;(ii) 正交氨酰基-tRNA 合成酶(O-RS);(iii) 正交 tRNA(O-tRNA),其包含 SEQ ID NO:1 所示的多核苷酸序列,其中所述 O-RS 用所述 3- 氯代酪氨酸优先氨酰化所述 O-tRNA;和(iv) 编码所述荧光蛋白质的核酸,其中所述核酸含有所述 O-tRNA 特异性识别的至少一个选择密码子(任选地为琥珀密码子);然后在所述蛋白质的翻译过程中,3- 氯代酪氨酸氨酰化的 O-tRNA 对所述选择密码子起反应而将培养基中的所述 3- 氯代酪氨酸掺入所述荧光蛋白质的所选位置,使其 pKa 降低,同时改变光学性质。

[0030] 3、有益效果

[0031] 通常研究蛋白酪氨酸氯化损伤影响蛋白功能,在体外研究主要采用次氯酸等氧化剂进行体外氧化,这样的缺点是对酪氨酸位点没有特异性,很多酪氨酸位点都可以氧化,同时又会引起假阳性,因为其它氨基酸被氧化同样会引起功能损伤,此外这种方法比较难以

应用于体内。另外是采用酪氨酸位点突变为苯丙氨酸或其它氨基酸,但是这种方法难以消除氨基酸本身突变而引起的功能变化。我们通过基因密码扩展方法,在目标蛋白任意位点通过引入 TAG 琥珀密码子,可以特异性引入 3- 氯代酪氨酸,从而可以位点特异性研究某个位点酪氨酸氯化引起的蛋白功能损伤。

[0032] 我们通过这种方法,在 ApoA1-192 位成功引入 3- 氯代酪氨酸,可以进一步结晶 ApoA1 氯代型,与野生型结构比较,从而可以从结构上解释酪氨酸氯代与功能关系。另外在细胞中,可以研究氯代 ApoA1 介导的胆固醇反向运输,提供酪氨酸氯代与胆固醇反向运输障碍的新方法和直接证据。

[0033] 由于氯原子的影响,Cl-Tyr 的 pKa 为 8.3,低于 Tyr 的 pKa(为 10.5),一些荧光蛋白荧光活性中心包含有酪氨酸交联形成的共轭体系,因此具有质子化和去质子化两种存在形式。据此我们推断在荧光活性中心用 3-Cl-Tyr 取代 Tyr 会降低 pKa,适合标记酸性细胞器。

[0034] 因此,我们系统的研究了 3-Cl-Tyr 突变的绿色荧光蛋白突变体,发现在绿色荧光蛋白活性中心氯代能够明显的降低绿色荧光蛋白的 pKa,其中 GFP66-Cl-Tyr 降低最明显,由 pKa 高于 10.8 降低到 pKa 4.7。推断其降低 pKa 主要有两个方面的原因,一方面,从结构上来看,GFP66-Cl-Tyr 跟 EGFP(S65T) 有相似的结构效果,荧光活性中心增加一个 Cl 离子,阻碍了 GFP 的 66 位 Tyr,水分子, Glu222 之间的质子传递体系,从而以 66 位 Tyr 主要以去质子化形式存在;另一方面,3-Cl-Tyr 比 Tyr 有更低的 pKa,从而也降低了荧光活性中心的 pKa。而 Cl 原子的空间位阻效应应该是主要原因。

[0035] 在 EGFP66-Cl-tyr,以及 YFP66-Cl-Tyr 突变体中均比相应野生型 pKa 降低 pH 1 左右,因为 EGFP 和 YFP 在结构上有利于荧光活性中心以去质子化形式存在,推断其影响 pKa 主要为 3-Cl-Tyr 比 Tyr 有更低的 pKa。

[0036] 从荧光量子产率来看,荧光活性中心氯代突变体均比野生型量子产率有所降低。但是 GFP66-Cl-Tyr 荧光量子产率为 0.57,与 EGFP 荧光量子产率 0.60 相当,但是相比 EGFP 具有更低的 pKa,EGFP pKa 为 6.0,而 GFP66-Cl-Tyr pKa 为 4.7。

[0037] 光激活绿红转化荧光蛋白 mEOS2 由于在 400nm 强光照射情况下,会由发绿色荧光转化为发红色荧光,是研究单分子成像的理想荧光分子。其荧光活性中心由 His62-Tyr63-Gly64 共轭交联形成。绿色状态的 mEOS2 pKa 为 5.6,我们通过用 3-Cl-Tyr 取代 Tyr 使其 pKa 降低到 4.2,使其可以在溶酶体和吞噬酶体中保持去质子化状态,在 502nm 光照情况下发绿光。

[0038] 我们根据氨基酸的 pKa 性质,理性的设计出了 pKa 降低的绿色荧光蛋白 GFP,黄色荧光蛋白 YFP,以及光转化荧光蛋白 mEOS2。

[0039] 传统的进化酶或蛋白的方式是通过在酶活中心或附近做定点突变库,或者易错 PCR 来引入随机突变。这往往需要很大的库容量以及合适的筛选策略。而这种根据氨基酸性质定点引入突变的理性设计提高了目标的靶向性。另外可以根据进化目标需要引入相应功能的非天然氨基酸,为酶和蛋白的定向进化提供了新的思路和解决办法。

[0040] 总之,我们进化出了可以位点特异性插入 3-Cl-Tyr 的氨酰基-tRNA 合成酶,为体内外研究蛋白氯代氧化损伤提供了新的方法。我们通过荧光活性中心氯代,得到了具有更低 pKa,以及荧光波谱稍有红移的突变体。尤其是 GFP66-Cl-Tyr,具有跟 EGFP 相似的荧光

量子产率,和更低的 pKa。另外 mEOS2 氯代使其 pKa 由 5.6 降低到 4.2 是标记溶酶体、吞噬酶体(内部 pH4-5)等酸性细胞器的合适的荧光蛋白。最后由于这种理性设计进化荧光蛋白方法,突破了传统定向进化中突变库容量和筛选策略的限制,引入了带有新性质的非天然氨基酸,为酶和蛋白定向进化提供了新的思路。

附图说明

- [0041] 从下面结合附图的详细描述中,本发明的上述特征和优点将更明显,其中:
- [0042] 图 1 是 3-氯代酪氨酸结构式;
- [0043] 图 2 是正交 tRNA 及特异性识别 3-Cl-Tyr 的氨酰基-tRNA 合成酶的序列;
- [0044] 图 3 是 3-Cl-Tyr-肌红蛋白的 SDS-PAGE 电泳图;
- [0045] 图 4 是 3-Cl-Tyr-肌红蛋白的质谱图;
- [0046] 图 5 是 3-Cl-Tyr-ApoA1 的质谱图:上图是 3-Cl-Tyr-ApoA1 的 HPLC 图谱,下图是 3-Cl-Tyr-ApoA1 的核质比图谱;
- [0047] 图 6 是绿色荧光蛋白不同突变体的 pKa:上图是 EGFP 的 pKa 滴定曲线,488nm 处波峰曲线由上到下依次指 pH 为 10.0、9.0、8.0、7.0、6.6、6.4、6.2、6.0、5.8、5.6、5.4、5.2、5.0、4.8、4.0、3.0;中图是 GFP-66-Cl-Tyr 的 pKa 滴定曲线,484nm 处波峰曲线由上到下依次指 pH 为 10.0、9.0、8.0、7.0、6.0、5.4、5.2、5.0、4.8、4.6、4.5、4.4、4.2、4.0、3.0;下图是 GFP 不同突变体的 pKa 滴定曲线;
- [0048] 图 7 是绿色荧光蛋白不同突变体的荧光性质:上图是 GFP 不同突变体的吸收光谱,下图是 GFP 不同突变体的荧光光谱;
- [0049] 图 8 是 mEOS2-63-Cl-Tyr 的荧光光谱;
- [0050] 图 9 是 mEOS2-63-Cl-Tyr 的 pKa 滴定,其中 502nm 处波峰曲线由上到下依次指 pH 为 10.0、9.0、8.5、8.0、7.5、6.8、6.6、6.4、6.2、6.0、5.8、5.6、5.4、5.2、5.0、4.8、4.6、4.4、4.2、4.0、3.8、3.6、3.6、3.6、3.4、3.2、3.0) 及 Henderson-Hasselbalch 方程曲线(方框内的曲线图)。

具体实施方式

[0051] 以下通过实施例来进一步阐明本发明。但是应该理解,所述实施例只是举例说明的目的,并不意欲限制本发明的范围和精神。

[0052] 本领域技术人员应该理解,除非特别说明,下述实施例中所用的化学试剂均为可通过商业途径购得的分析纯级别的试剂。

[0053] 实施例 1:进化 3-Cl-Tyr 特异性氨酰基-tRNA 合成酶

[0054] 为了在基因中位点特异性插入 3-Cl-Tyr,需要在所用的 *E. coli* 宿主细胞中引入氨酰基-tRNA 合成酶/tRNA 正交对,这个正交对来源于詹氏甲烷球菌 (*Methanococcus jannaschii*) 琥珀抑制酪氨酰 tRNA ($Mj\text{tRNA}_{\text{CUA}}^{\text{Tyr}}$)/酪氨酰 tRNA 合成酶 ($Mj\text{TyrRS}$,野生型,其氨基酸序列为 SEQ ID NO:12) 对。 $Mj\text{TyrRS}$ 突变库构建在卡纳霉素抗性 pBK 质粒(购自美国 scripps 研究所 Peter G. Schultz 实验室)中,位于该质粒上 *E. coli* 谷氨酰胺合成酶的启动子和终止子之间。所使用的合成酶突变库为 pBk-lib-jw1 库,该突变库的构建方法为:在 $Mj\text{TyrRS}$ 基因上挑选 6 个位点 (Tyr32,Leu65,Phe108,Gln109,Asp158,和 Leu162) 引入 NNK

突变 ($N = A+T+C+G$; $K = T+G$), 另外6个位点 (Ile63, Ala67, His70, Tyr114, Ile159, Val164) 或随机突变为 Gly 或保持不变 (参见 Xie, J.; Liu, W. S.; Schultz, P. G. *Angew. Chem., Int. Ed.* 2007, 46, 9239-9242; Wang, JY.; Zhang W.; Song WJ; et al. *J. Am. Chem. Soc.* 2010, 132, 14812-14818)。

[0055] 通过正负筛选来进化特异性识别 3-Cl-Tyr 的氨酰基-tRNA 合成酶。正筛选质粒包含 MjtRNA_{CUA}^{Tyr}, TAG 突变的氯霉素乙酰转移酶基因, 启动表达绿色荧光蛋白的琥珀突变的 T7RNA 聚合酶, 四环素抗性基因。负筛选质粒包含 MjtRNA_{CUA}^{Tyr}, 在阿拉伯糖操纵子下的琥珀突变芽孢杆菌 RNA 酶基因, 以及氨基青霉素抗性基因。进行 3 轮正负筛选: 包含有正筛选质粒的 *E. coli* DH10B 细胞作为正筛选寄主细胞。细胞电转 pbk-lib-jw1 库, SOC 培养基 (2% (W/V) 胰蛋白胨, 0.5% (W/V) 酵母粉, 0.05% (W/V) NaCl, 2.5mM KCl, 10mM MgCl₂, 20mM 葡萄糖) 在 37°C 培养 1 小时。之后换用极限培养基 (GMML 极限培养基的配方: M9 盐 / 甘油: 764g Na₂HPO₄·7H₂O 或者 30g Na₂HPO₄, 15g KH₂PO₄, 2.5g NaCl, 5g NH₄Cl, 50ml 甘油, 高压灭菌, pH 7.0; 1M MgSO₄: 高压灭菌; 50mM CaCl₂: 高压灭菌; 25mM FeCl₂: 过滤灭菌; 0.3M 亮氨酸: 溶于 0.3M NaOH 中, 过滤灭菌; 1L 液体 GMML 培养基: 200ml M9 盐 / 甘油, 2ml MgSO₄, 2ml CaCl₂, 2ml FeCl₂, 1ml 亮氨酸) 洗两次, 铺板固体极限培养基 (在液体 GMML 培养基中加入 500ml 3% 琼脂粉, 1mM 3-Cl-Tyr, 50mg/L 卡那霉素, 60mg/L 氯霉素, 15mg/L 四环素), 37°C 培养 60 小时。收取细胞, 提取质粒 DNA, 电泳分离, 胶回收。然后, 将经过正筛选的 pbk-lib-jw1 转化到包含负筛选质粒的 DH10B 感受态细胞中。SOC 培养基中恢复 1 小时。之后涂板包含 0.2% 阿拉伯糖 (购自 sigma 公司) 的 LB 固体培养基 (每升培养基含 10g 胰蛋白胨, 5g 酵母粉, 10g NaCl)。37°C 培养 8-12 小时。共重复 3 轮。

[0056] 最后一轮正筛选挑 384 个克隆, 分别点板在含有 1mM 3-Cl-Tyr、氯霉素 60, 80, 100, 120mg/L 的 GMML 固体培养基上, 及不包含 3-Cl-Tyr、但包含氯霉素 0, 20, 40, 60mg/L 的 GMML 固体培养基。挑选在 1mM 3-Cl-Tyr 120mg/L 氯霉素的培养基上生长, 而在 0mM 3-Cl-Tyr 40mg/L 氯霉素培养基中不生长的克隆进行进一步验证。挑出 3 个克隆, 测序结果如图 2 所述。其中克隆 1 的 3-氯代酪氨酸插入效率最高, 测序表明, 克隆 1 所包含的氨酰基-tRNA 合成酶突变体 1 的氨基酸序列为 SEQ ID NO: 2 所示, 其中的突变位点为 H70A, D158S, I159S。

[0057] 另外两种氨酰基-tRNA 合成酶突变体 2 和 3 的氨基酸序列分别为 SEQ ID NO: 3 和 4 所示。氨酰基-tRNA 合成酶突变体 1-3 的相应的核苷酸序列分别为 SEQ ID NO: 5-7。

[0058] 实施例 2: 表达 3-Cl-Tyr-肌红蛋白及质谱鉴定

[0059] 将正交 tRNA (SEQ ID NO: 1) 和筛选出来的氨酰基-tRNA 合成酶突变体 1 (SEQ ID NO: 5) 分别构建到 pEVOL 载体 (购自美国 scripps 研究所 Peter G. Schultz 实验室) 上, 然后共转化到包含有 pbad-肌红蛋白 (4TAG) (该质粒购自美国 scripps 研究所 Peter G. Schultz 实验室) (其中肌红蛋白 (4TAG) 的核苷酸序列为 SEQ ID NO: 8) 的 DH10B 细胞 (购自全式金公司) 中。挑取单个克隆在 37°C 培养到 OD₆₀₀ 约等于 0.5 时, 向 LB 培养基中加入 1mM 3-Cl-Tyr (购自上海吉尔生化公司), 及 0.2% 阿拉伯糖 (购自 sigma 公司) 培养细胞, 对照不加入 3-Cl-Tyr。6-8 小时之后, 收菌, Ni-NTA 纯化蛋白, 并用 SDS-PAGE 电泳分析 (图 3)。

[0060] 我们发现, 只有在存在 3-Cl-Tyr 的培养基中才能纯化出全长的肌红蛋白, 这说明

氨酰基-tRNA 合成酶突变体可以特异性的识别 3-C1-Tyr。在 LB 培养基中 3-C1-Tyr 肌红蛋白的产率为 2-5mg/L。为了检测 3-C1-Tyr 仅仅插入到肌红蛋白的 4 位琥珀突变位点,我们对 3-C1-Tyr-肌红蛋白进行了 ESI-TOF 质谱检测,检测结果分子量为 18465Da(图 4),与计算的分子量 18465.5Da 吻合。

[0061] 实施例 3:表达插入 3-C1-Tyr 的 ApoA1 及其质谱鉴定

[0062] 将 AopA1(核苷酸序列为 SEQ ID NO:9)构建在 pet24a 载体(购自 novagen 公司)上,同时将正交 tRNA(SEQ ID NO:1)和筛选出来的氨酰基-tRNA 合成酶突变体 1(SEQ ID NO:5)分别构建到 pEVOL 载体(购自美国 scripps 研究所 Peter G.Schultz 实验室)上。通过重叠 PCR 方法,在 ApoA1 192 位引入 TAG。共转化 pEVOL-tRNA, pEVOL-3-C1-TyrRS 及 pet24a-ApoA1-192TAG 到 BL21(DE3)细胞(购自全式金公司)中,挑单克隆到 2YT 培养基(每升培养基含 16g 胰蛋白胨,5g 酵母粉,5gNaCl),加相应的抗生素(卡纳霉素 50ug/ml 四环素 10ug/ml)。待细胞生长到 OD₆₀₀ 约等于 1.0 左右时,加入 1mM 3-C1-Tyr, IPTG 0.5mM,阿拉伯糖 0.2%,37℃继续培养 6-8h。收菌,Ni-NTA 纯化蛋白。通过 ESI 质谱鉴定,我们发现纯化出的 3-C1-Tyr-ApoA1 分子量为 25156,对比野生型分子量 25119,相差 37,与一个 C1 分子量相当(图 5)。

[0063] 实施例 4:表达插入 3-C1-Tyr 的荧光蛋白并测定其 pKa 及检测光性质

[0064] 将 GFP 系列突变体(GFP 的核苷酸序列见 SEQ ID NO:10,EGFP,YFP 均通过 GFP 基因改造获得)均分别构建在 pet24a 载体(购自 novagen 公司)上,同时将正交 tRNA(SEQ ID NO:1)和筛选出来的氨酰基-tRNA 合成酶突变体 1(SEQ ID NO:5)分别构建到 pEVOL 载体(购自美国 scripps 研究所 Peter G.Schultz 实验室)上。Pet24a-EGFP 通过重叠 PCR 方法,在 GFP 基因上引入 F64L,S65T。pet24a-YFP 通过重叠 PCR 方法,在 GFP 基因引入 S65G,S72A,K79R,T203Y。用同样方法,在 GFP,EGFP,YFP 的 66 位引入 TAG,在 GFP,YFP 的 203 位引入 TAG。共转化 pEVOL-tRNA, pEVOL-3-C1-TyrRS 及相应的荧光蛋白载体如 pet24a-GFP66TAG, pet24a-GFP203TAG, pet24a-EGFP66TAG, pet24a-YFP66TAG, pet24a-YFP203TAG, pe24at-mEOS2-63TAG 到 BL21(DE3)细胞(购自全式金公司)中,表达纯化条件同 ApoA1-192-C1-Tyr。

[0065] 将上述各种荧光蛋白溶在不同 pH 值范围的缓冲液中:(1)不同比例的 0.1M 柠檬酸-0.2M 的 Na₂HPO₄ 缓冲液中,pH 值 3.0-7.0;(2)pH7.5-10.0 为 50mM Tris 缓冲液;(3)pH 值 10.5,10.8 为 0.1M Na₂CO₃-NaHCO₃。紫外分光光度计扫描光吸收,从 240nm 到 700nm。

[0066] 从 pH3.0-10.8,测定荧光蛋白去质子化状态吸收峰,YFP 514nm, YFP203-C1-Tyr 514nm, YFP66-C1-Tyr 519nm, EGFP 488nm, EGFP66-C1-Tyr498nm, GFP66-C1-Tyr 484nm, mEOS2-63-C1-Tyr 502nm。去质子化状态最高吸收峰值对应去质子化比例为 1,其它 pH 值光吸收峰与最高吸收峰比例为相应去质子化比例。测定值做图,pH 值为 x 轴,去质子化比例为 y 轴,根据 Henderson-Hasselbalch 公式:

$$[0067] \quad pH = pK_a + \log \frac{[A^-]}{[HA]} = pK_a - \log \frac{[HA]}{[A^-]} \quad (1)$$

$$[0068] \quad \text{推出:} \log \frac{[HA]}{[A^-]} = pK_a - pH \quad (2)$$

[0069] 推出：
$$\frac{[HA]}{[A^-]} = 10^{pK_a - pH} \quad (3)$$

[0070] 而
$$y = \frac{[A^-]}{[A^-] + [HA]} = \frac{1}{\frac{[HA]}{[A^-]} + 1}$$

[0071]
$$= \frac{1}{10^{pK_a - pH} + 1} = \frac{1}{10^{pK_a - x} + 1} \quad (4)$$

[0072] 测定的 pH 值 (x), 去质子化比例 (y) 根据公式 4 拟合曲线, 计算 pKa。

[0073] 其中 GFP, GFP203-Cl-Tyr, GFP203-Tyr, 滴定 pH3.0-10.8, 质子化形式荧光活性中心光吸收没有明显的降低。故其 pKa 高于我们测定的 pH 范围, 在 pH 10.8 以上。其它绿色荧光蛋白突变体的 pKa 如图 6 所示, EGFP66-Cl-Tyr pKa 为 pH4.7, EGFP pKa 为 pH6.0, YFP66-Cl-Tyr pKa 为 pH4.9, YFP pKa 5.6, GFP66-Cl-Tyr pKa 为 pH4.7。荧光活性中心氯代能够明显降低绿色荧光蛋白的 pKa, 其中 EGFP66-Cl-Tyr 和 YFP66-Cl-Tyr 均比野生型降低 pKa pH 值 1 左右。而 GFP66-Cl-Tyr 降低 pKa 最明显, 由原来的 pH 10.8 以上, 降低到 pH 4.7。

[0074] Hi-tachi F4500 荧光分光光度计 (日本, 日立公司) 检测, 发射波谱测定, EFP66-Cl-Tyr, EFP, GFP66-Cl-Tyr 固定 450nm 激发, 扫描 450-700 发射谱。GFP203Y, GFP203-Cl-Tyr, GFP 固定 375nm 激发, 扫描 380-700nm 发射波谱。YFP, YFP203-Cl-Tyr, YFP66-Cl-Tyr 固定 470nm 激发, 扫描 470-700nm 发射波谱。

[0075] 激发波谱测定, EGFP66-Cl-Tyr, EGFP, GFP66-Cl-Tyr 固定 545 光发射, 扫描 200-540 光激发波谱。GFP203Y, YFP66-Cl-Tyr, GFP 固定 565nm 光发射, 扫描 200-560nm 光激发波谱。YFP, YFP203-Cl-Tyr 固定 560nm 光发射, 扫描 200-555nm 光激发波谱。

[0076] GFP66-Cl-Tyr, EGFP, EGFP66-Cl-Tyr, GFP 溶在 pH7.050mM Tris 缓冲液中, 488nm 光吸收为 0.1, 514nm 光激发, 扫描 495-700nm 发射波谱。激发 495-620nm 发射波谱, 以已经发表的 EGFP 荧光量子产率 0.60 为标准。YFP, YFP66-Cl-Tyr, YFP203-Cl-Tyr, GFP203-Cl-Tyr, GFP203Y 溶在 pH7.050mM Tris 缓冲液中, 稀释 514nm 光吸收为 0.1, 514nm 光激发, 扫描 520-700nm 光发射波谱。激发 520-620nm 发射波谱, 根据已经发表的 YFP 量子产率 0.60 为标准。

[0077] 在 pH7.0 情况下, 从最大光吸收来看, 荧光活性中心氯代突变体均比野生型光吸收向红光方向偏移约 5-10nm。其中 GFP 最大光吸收为 396nm, 是质子化荧光活性中心的光吸收, 而 GFP66-Cl-Tyr 最大光吸收是 488nm, 是去质子化荧光活性中心光吸收。

[0078] 从摩尔消光系数来看, 荧光活性中心氯代对比野生型荧光蛋白摩尔消光系数均有增加。其中 EGFP66-Cl-Tyr 消光系数增加 41%。而在 203 位 Cl-Tyr 取代 Tyr 没有怎么影响绿色荧光蛋白光吸收和摩尔消光系数。

[0079] 我们进一步在 pH7.0 情况下, 测定了绿色荧光蛋白不同突变体的荧光性质 (图 7)。

[0080] 从荧光激发波谱和荧光发射波谱来看, 荧光活性中心氯代的突变体均比野生型最大激发波长及最大发射波长向红光方向偏移, 其中 GFP66-Cl-Tyr 最大激发波长为 492nm, 对比 GFP 的最大发射波长 397nm, 荧光活性中心主要为去质子化形式。

[0081] 而不同类型绿色荧光蛋白 203 位 Cl-Tyr 取代 Tyr 没有怎么影响绿色荧光蛋白的

荧光发射波谱及荧光激发波谱。从不同绿色荧光蛋白突变体量子产率来看,荧光活性中心氯代比野生型蛋白荧光量子产率有所降低。但是 GFP203-Cl-Tyr 突变体比 GFP203-Tyr 在 405nm 左右光激发荧光量子产率有所提高,由 0.23 增高到 0.39。

[0082] 实施例 5 :3-Cl-Tyr 插入 mEOS2 荧光蛋白并检测其光学性质

[0083] 通过与实施例 2 或 3 相同的方法,将 mEOS2 基因(其核苷酸序列为 SEQ ID NO :11)构建在 pet24a 载体(购自 novagen 公司)上,同时将正交 tRNA(SEQ ID NO :1)和筛选出来的氨酰基-tRNA 合成酶突变体 1(SEQ ID NO :5)分别构建到 pEVOL 载体(购自美国 scripps 研究所 Peter G. Schultz 实验室)上。通过重叠 PCR 方法,在 mEOS2 63 位引入 TAG。共转化 pEVOL-tRNA, pEVOL-3-Cl-TyrRS 及 pet24a-mEOS2-63TAG 到 BL21 (DE3) 细胞(购自全式金公司)中,在蛋白翻译过程中,通过本发明的翻译系统将 mEOS2 Tyr63 替换为 3-Cl-Tyr,并测量 mEOS2 发射光谱和激发光谱。

[0084] 对于 mEOS2 的绿色荧光状态,其吸收光谱最大吸收峰为 506nm,发射光谱峰值为 519nm。对于 mEOS2-63-Cl-Tyr 的绿色荧光状态,我们测得吸收峰为 502nm,发射峰为 518nm。可见吸收峰向短波长移动 4nm,而发射峰几乎没变(图 8)。

[0085] 对于绿色荧光状态的 3-Cl-Tyr-mEos2,随着 pH 的降低,502nm 吸收峰逐渐降低,同时在 381nm 处出现吸收峰并逐渐升高(图 9)。在 pH3.0 时,502nm 吸收峰几乎完全消失。绿色荧光状态的 mEos2 随着 pH 降低会在 395nm 出现吸收峰。与野生型类似,3-Cl-Tyr-mEos2 的 381nm 吸收峰对应着发色团的质子化状态的光吸收,而 502nm 吸收峰对应着去质子化发色团的光吸收。发色团的质子化与去质子化发生在酪氨酸的酚羟基上。3-Cl-Tyr 使得质子化状态的吸收峰向短波方向移动了 14nm。吸收光谱随 pH 变化的过程中,在 426nm 处存在一个等消光点,其吸光值不随 pH 的变化而变化。3-Cl-Tyr 的替代不仅改变了荧光发色团在质子化以及去质子化状态的吸收峰,而且改变了发色团的 pKa。mEos2 的 $pK_a = 5.6$ 。3-Cl-TyrmEos2 在 502nm 的吸收峰随 pH 的变化可以用 Henderson-Hasselbalch 方程拟合,得到 $pK_a = 4.2$ 。

[0086] 应该理解,尽管参考其示例性的实施方案,已经对本发明进行具体地显示和描述,但是本领域的普通技术人员应该理解,在不背离由后附的权利要求所定义的本发明的精神和范围的前提下,可以在其中进行各种形式和细节的变化,可以进行各种实施方案的任意组合。

[0001]

序列表

- <110> 中国科学院生物物理研究所
- <120> 3-氯代酪氨酸翻译系统及其应用
- <130> IB111460
- <160> 12
- <170> PatentIn version 3.1
- <210> 1
- <211> 77
- <212> DNA
- <213> 人工合成的
- <400> 1
 tggtcggcg ggcggattt gaaccagcgc catgcgatt tagagtccgc cgttctgccc 60
 tgctgaacta ccgcccg 77
- <210> 2
- <211> 312
- <212> PRT
- <213> 人工的
- <400> 2
- Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser
 1 5 10 15
- Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Tyr
 20 25 30
- Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln
 35 40 45
- Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile
 50 55 60
- Leu Leu Ala Asp Leu Ala Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
 65 70 75 80
- Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met
 85 90 95

[0002]

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Gln Leu Asp Lys
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Ser Ser His
 145 150 155 160

Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser
 195 200 205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala
 210 215 220

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro
 225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys
 245 250 255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu
 260 265 270

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys
 275 280 285

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys
 290 295 300

Arg Leu His His His His His His
 305 310

<210> 3
 <211> 312
 <212> PRT
 <213> 人工的

<400> 3

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Glu
 20 25 30

[0003]

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

<210> 4

<211> 312

[0004]

<212> PRT
 <213> 人工的

 <400> 4
 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser
 1 5 10 15
 Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Glu
 20 25 30
 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln
 35 40 45
 Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Gly
 50 55 60
 Ile Leu Ala Asp Leu Thr Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
 65 70 75 80
 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met
 85 90 95
 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Val Leu Asp Lys
 100 105 110
 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys
 115 120 125
 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro
 130 135 140
 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Thr Ile His
 145 150 155 160
 Tyr Thr Gly Gly Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile
 165 170 175
 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His
 180 185 190
 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser
 195 200 205
 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala
 210 215 220
 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro
 225 230 235 240
 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys
 245 250 255
 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu
 260 265 270

[0005]

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys
275 280 285

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys
290 295 300

Arg Leu His His His His His His
305 310

<210> 5

<211> 921

<212> PRT

<213> 人工的

<400> 5

Ala Thr Gly Gly Ala Cys Gly Ala Ala Thr Thr Thr Gly Ala Ala Ala
1 5 10 15

Thr Gly Ala Thr Ala Ala Ala Gly Ala Gly Ala Ala Ala Cys Ala Cys
20 25 30

Ala Thr Cys Thr Gly Ala Ala Ala Thr Thr Ala Thr Cys Ala Gly Cys
35 40 45

Gly Ala Gly Gly Ala Ala Gly Ala Gly Thr Thr Ala Ala Gly Ala Gly
50 55 60

Ala Gly Gly Thr Thr Thr Thr Ala Ala Ala Ala Ala Ala Ala Gly Ala
65 70 75 80

Thr Gly Ala Ala Ala Ala Ala Thr Cys Thr Gly Cys Thr Thr Ala Thr
85 90 95

Ala Thr Ala Gly Gly Thr Thr Thr Thr Gly Ala Ala Cys Cys Ala Ala
100 105 110

Gly Thr Gly Gly Thr Ala Ala Ala Ala Thr Ala Cys Ala Thr Thr Thr
115 120 125

Ala Gly Gly Gly Cys Ala Thr Thr Ala Thr Cys Thr Cys Cys Ala Ala
130 135 140

Ala Thr Ala Ala Ala Ala Ala Ala Gly Ala Thr Gly Ala Thr Thr Gly
145 150 155 160

Ala Thr Thr Thr Ala Cys Ala Ala Ala Ala Thr Gly Cys Thr Gly Gly
165 170 175

Ala Thr Thr Thr Gly Ala Thr Ala Thr Ala Ala Cys Thr Ala Thr Ala
180 185 190

Cys Thr Thr Thr Thr Gly Gly Cys Thr Gly Ala Thr Thr Thr Ala Gly
195 200 205

[0006]

Cys Thr Gly Cys Cys Thr Ala Thr Thr Thr Ala Ala Ala Cys Cys Ala
 210 215 220

Gly Ala Ala Ala Gly Gly Ala Gly Ala Gly Thr Thr Gly Gly Ala Thr
 225 230 235 240

Gly Ala Gly Ala Thr Thr Ala Gly Ala Ala Ala Ala Ala Thr Ala Gly
 245 250 255

Gly Ala Gly Ala Thr Thr Ala Thr Ala Ala Cys Ala Ala Ala Ala Ala
 260 265 270

Ala Gly Thr Thr Thr Thr Thr Thr Gly Ala Ala Gly Cys Ala Ala Thr Gly
 275 280 285

Gly Gly Gly Thr Thr Ala Ala Ala Gly Gly Cys Ala Ala Ala Ala Thr
 290 295 300

Ala Thr Gly Thr Thr Thr Thr Ala Thr Gly Gly Ala Ala Gly Thr Gly Ala
 305 310 315 320

Ala Thr Thr Cys Cys Ala Gly Cys Thr Thr Gly Ala Thr Ala Ala Gly
 325 330 335

Gly Ala Thr Thr Ala Thr Ala Cys Ala Cys Thr Gly Ala Ala Thr Gly
 340 345 350

Thr Cys Thr Ala Thr Ala Gly Ala Thr Thr Gly Gly Cys Thr Thr Thr
 355 360 365

Ala Ala Ala Ala Ala Cys Thr Ala Cys Cys Thr Thr Ala Ala Ala Ala
 370 375 380

Ala Gly Ala Gly Cys Ala Ala Gly Ala Ala Gly Gly Ala Gly Thr Ala
 385 390 395 400

Thr Gly Gly Ala Ala Cys Thr Thr Ala Thr Ala Gly Cys Ala Ala Gly
 405 410 415

Ala Gly Ala Gly Gly Ala Thr Gly Ala Ala Ala Ala Thr Cys Cys Ala
 420 425 430

Ala Ala Gly Gly Thr Thr Gly Cys Thr Gly Ala Ala Gly Thr Thr Ala
 435 440 445

Thr Cys Thr Ala Thr Cys Cys Ala Ala Thr Ala Ala Thr Gly Cys Ala
 450 455 460

Gly Gly Thr Thr Ala Ala Thr Thr Cys Thr Ala Gly Thr Cys Ala Thr
 465 470 475 480

Thr Ala Thr Cys Thr Thr Gly Gly Cys Gly Thr Thr Gly Ala Thr Gly
 485 490 495

Thr Thr Gly Cys Ala Gly Thr Thr Gly Gly Ala Gly Gly Gly Ala Thr
 500 505 510

[0007]

Gly Gly Ala Gly Cys Ala Gly Ala Gly Ala Ala Ala Ala Ala Thr Ala
 515 520 525

Cys Ala Cys Ala Thr Gly Thr Thr Ala Gly Cys Ala Ala Gly Gly Gly
 530 535 540

Ala Gly Cys Thr Thr Thr Thr Ala Cys Cys Ala Ala Ala Ala Ala Ala
 545 550 555 560

Gly Gly Thr Thr Gly Thr Thr Thr Gly Thr Ala Thr Thr Cys Ala Cys
 565 570 575

Ala Ala Cys Cys Cys Thr Gly Thr Cys Thr Thr Ala Ala Cys Gly Gly
 580 585 590

Gly Thr Thr Thr Gly Gly Ala Thr Gly Gly Ala Gly Ala Ala Gly Gly
 595 600 605

Ala Ala Ala Gly Ala Thr Gly Ala Gly Thr Thr Cys Thr Thr Cys Ala
 610 615 620

Ala Ala Ala Gly Gly Gly Ala Ala Thr Thr Thr Thr Ala Thr Ala Gly
 625 630 635 640

Cys Thr Gly Thr Thr Gly Ala Thr Gly Ala Cys Thr Cys Thr Cys Cys
 645 650 655

Ala Gly Ala Ala Gly Ala Gly Ala Thr Thr Ala Gly Gly Gly Cys Thr
 660 665 670

Ala Ala Gly Ala Thr Ala Ala Ala Gly Ala Ala Ala Gly Cys Ala Thr
 675 680 685

Ala Cys Thr Gly Cys Cys Cys Ala Gly Cys Thr Gly Gly Ala Gly Thr
 690 695 700

Thr Gly Thr Thr Gly Ala Ala Gly Gly Ala Ala Ala Thr Cys Cys Ala
 705 710 715 720

Ala Thr Ala Ala Thr Gly Gly Ala Gly Ala Thr Ala Gly Cys Thr Ala
 725 730 735

Ala Ala Thr Ala Cys Thr Thr Cys Cys Thr Thr Gly Ala Ala Thr Ala
 740 745 750

Thr Cys Cys Thr Thr Thr Ala Ala Cys Cys Ala Thr Ala Ala Ala Ala
 755 760 765

Ala Gly Gly Cys Cys Ala Gly Ala Ala Ala Ala Ala Thr Thr Thr Gly
 770 775 780

Gly Thr Gly Gly Ala Gly Ala Thr Thr Thr Gly Ala Cys Ala Gly Thr
 785 790 795 800

Thr Ala Ala Thr Ala Gly Cys Thr Ala Thr Gly Ala Gly Gly Ala Gly
 805 810 815

[0008]

Thr Thr Ala Gly Ala Gly Ala Gly Thr Thr Thr Ala Thr Thr Thr Ala
 820 825 830

Ala Ala Ala Ala Thr Ala Ala Gly Gly Ala Ala Thr Thr Gly Cys Ala
 835 840 845

Thr Cys Cys Ala Ala Thr Gly Gly Ala Thr Thr Thr Ala Ala Ala Ala
 850 855 860

Ala Ala Thr Gly Cys Thr Gly Thr Ala Gly Cys Thr Gly Ala Ala Gly
 865 870 875 880

Ala Ala Cys Thr Thr Ala Thr Ala Ala Ala Gly Ala Thr Thr Thr Thr
 885 890 895

Ala Gly Ala Gly Cys Cys Ala Ala Thr Thr Ala Gly Ala Ala Ala Gly
 900 905 910

Ala Gly Ala Thr Thr Ala Thr Ala Ala
 915 920

- <210> 6
- <211> 921
- <212> DNA
- <213> 人工的

<400> 6
 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
 agagaggttt taaaaaaga tgaaaaatct gctgagatag gtttgaacc aagtgttaa 120
 atacatttag ggcattatct ccaataaaa aagatgattg atttcaaaa tgctggattt 180
 gatataatta taactttggc tgatttaact gcctatttaa accagaaagg agagtggat 240
 gagattagaa aaataggaga ttatacaaaa aaagtttttg aagcaatggg gtaaaggca 300
 aaatatgttt atggaagtga aagtggctct gataaggata gcacactgaa tgtctataga 360
 ttggctttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
 gatgaaaatc caaaggtgc tgaagttatc tatccaataa tgcaggtaa tcagattcat 480
 tatcttggcg ttgatgttc agttggaggg atggagcaga gaaaaataca catgttagca 540
 agggagcttt taccaaaaaa gttgtttgt attcacaacc ctgtcttaac gggtttgat 600
 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660
 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720
 ataatggaga tagctaaata cttccttgaa tatectttaa ccataaaaag gccagaaaaa 780
 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
 gaattgcatc caatggattt aaaaaatget gtagctgaag aacttataaa gattttagag 900
 ccaattagaa agagattata a 921

- <210> 7
- <211> 921

[0009]

<212> DNA	
<213> 人工的	
<400> 7	
atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta	60
agagagggtt taaaaaaga tgaaaaatct gctgagatag gttttgaacc aagtggtaaa	120
atacatttag ggcattatct ccaataaaa aagatgattg atttacaana tgctggattt	180
gatataggta taattttggc tgatttaact gcctatntaa accagaaaag agagtggat	240
gagattagaa aaataggaga ttatacaaaa aaagtntttg aagcaatggg gttaaaggea	300
aaatatgttt atggaagtga attccgtctt gataaggatt atacactgaa tgcctataga	360
ttggctntaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag	420
gatgaaaate caaaggttgc tgaagttatc tatccaataa tgcaggttaa tactattcat	480
tatactggcg gtgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca	540
agggagcttt taccaaaaa ggttgtttgt attcacaacc ctgtctaac gggtttgat	600
ggagaaggaa agatgagttc ttcaaaaggg aatnttatag ctgttgatga ctctccagaa	660
gagallagg ctaagalaaa gaaagcalac tgcctcagct gagttgttga aggaaatcca	720
ataatggaga tagctaaata cttccttgaa tatectntaa ccataaaaag gccagaaaa	780
tttgggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag	840
gaattgcatc caatggattt aaaaaatgct gtagctgaag aactntataa gattnttagag	900
ccaattagaa agagattata a	921

<210> 8

<211> 492

<212> DNA

<213> 人工的

<400> 8

atggttctgt ctgaaggatga atggcagctg gttctgcatg tttgggctaa agttgaagct	60
gacgtcgtg gtcatgttca ggacatcttg attcactgt tcaaatctca tccgaaaact	120
ctgaaaaat tcgatcgttt caaacatctg aaaactgaag ctgaaatgaa agcttctgaa	180
gatctgaaaa aacatgggtg taccgtnta actgcctag gtgctatcct taagaaaaa	240
gggatcatg aagctgagct caaacctctt gcacaategc atgctactaa acataagatc	300
ccgatcaaat acctggaatt catctctgaa gcgatcatcc atgttctgca ttctagacat	360
ccaggtgact tcggtgctga cgctcagggt gctatgaaca aagctctga gctgttccgt	420
aaagatatcg ctgctaagta caaagaactg ggttaccagg gtggctcggg acatcatcac	480
catcaccatt ga	492

<210> 9

<211> 654

<212> DNA

[0010]

<213> 人工的

<400> 9

```

atgccggagc agagtccatg ggatcgcgtg aaggacctgg ccaactgtgta cgtggatgtg      60
ctcaaagaca gcggcagaga ctatgtgtct cagtttgaag gatccgcctt gggcaaaaaa      120
ttgaacctta agctgctgga caactgggac agcgtgacgt ccaccttcag caagctgcgc      180
gaacagctcg gccctgtgac ccaggaattc tgggataacc tggaaaagga gacagagggc      240
ctgcccagg agatgagcaa ggatctggag gaggtgaagg ccaaggtgca gccgtacctg      300
gacgacttcc agaagaagtg gcaggaggag atggagctct accgccagaa ggtggagccc      360
ctgcgcgcag agctgcagga gggcgcgcgc cagaagctgc acgagctgca agagaagctg      420
agcccactgg gcgaggagat gcgcgacccg gcgcgcgccc atgtcgacgc gctgcgcaeg      480
catctggcgc cgtacagcga cgagctgcgc cageccttag cggcgcgcct tgaggctctc      540
aaggagaacg gcggggcccc cctggccgag taccacgcca aggccaccga gcatctgagc      600
acgctcagcg agaaggccaa gccggcgtc gagcaccacc accaccacca ctga      654

```

<210> 10

<211> 759

<212> DNA

<213> 人工的

<400> 10

```

atgagtaaag gagaagaact tttcactgga gttgtcccaa ttcttgttga attagatggt      60
gatgttaatg gccacaaatt ttctgtcagt ggagagggtg aaggtgatgc aacatacgga      120
aaacttacc ttaaatttat ttgactact ggaaaactac ctgttccatg gccaacactt      180
gtcactactt tctcttatgg tgttcaatgc ttttccggtt atccgatca catgaaacgg      240
catgactttt tcaagagtgc catgcccgaa ggttatgtac aggaacgcac tatactttc      300
aaagatgacg ggaactacaa gacgcgtgct gaagtcaagt ttgaagtgta tacccttgtt      360
aatcgtatcg agttaaagg tattgatttt aaagaagatg gaaacattct cggacacaaa      420
ctcgaataca actataactc acacaatgta tacatcacgg cagacaaaca aaagaatgga      480
atcaaageta acttcaaat tcgcacaaac attgaagatg gatccgttca actagcagac      540
cattataac aaaatactcc aattggcgat ggccctgtcc tttaccaga caaccattac      600
ctgtcgacac aatctgccct ttcgaaagat cccaacgaaa agegtgacca catggtcctt      660
cttgagtttg taactgctgc tgggattaca catggcatgg atgagctcgt cccgcgcggc      720
agctagggtg gtctcgagca ccaccaccac caccactga      759

```

<210> 11

<211> 678

<212> DNA

<213> 人工的

[0011]

<400> 11
 atgagtgcga ttaagccaga catgaagatc aaactccgta tggaaggcaa cgtaaaccggg 60
 caccactttg tgatcgacgg agatggtaca ggcaagcctt ttgagggaaa acagagtatg 120
 gatcttgaag tcaaagaggg eggacctctg ccttttgect ttgatatcct gaccactgca 180
 ttccattacg gcaacagggt attcgccaaa tateccagaca acatacaaga ctattttaag 240
 cagtcgtttc ctaaggggta ttctgtggaa cgaagettga ctttgaaga cgggggcatt 300
 tgcattgcca gaaacgacat aacaatggaa ggggacactt ctataataa agttcgattt 360
 tatggtacca actttccgc caatggtcca gttatgcaga agaagacgct gaaatgggag 420
 ccctccactg agaaaatgta tgtgcgtgat ggagtctga cgggtgatat tcatatggct 480
 ttgttgcttg aaggaaatgc ccattaccga tgtgacttca gaactactta caaagctaag 540
 gagaaggggt tcaagttacc aggctaccac tttgtggacc actgcattga gattttaagc 600
 catgacaaag attacaaca ggттаactg tatgagcatg ctgttgetca tcttgattg 660
 cctgacaatg ccagacga 678

<210> 12

<211> 312

<212> PRT

<213> 来源于詹氏甲烷球菌(Methanococcus jannaschii)

<400> 12

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser
1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Tyr
20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln
35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile
50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met
85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Gln Leu Asp Lys
100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys
115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro
130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Asp Ile His
145 150 155 160

[0012]

Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser
 195 200 205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala
 210 215 220

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro
 225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys
 245 250 255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu
 260 265 270

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys
 275 280 285

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys
 290 295 300

Arg Leu His His His His His His
 305 310

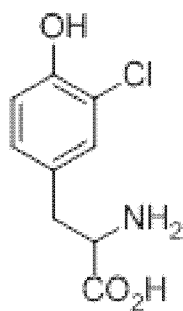


图 1

<p>氨基-tRNA 合成酶 突变体 3</p>	<p>VVCIHNPVLTGLDGEKMSSSKGNFIAVDDSPPEIRAKIKKAYCPAGVVEGNPIMEIAKYFLE YPLTIKRPEKFGGDLTVNSYEELLESLEFKNKELHPMDLKNVAEELIKILEPIRKRLHHHHHH</p> <p>核苷酸序列 (SEQ ID NO: 7): ATGGACGAATTTGAAATGATAAAGAGAAAACACATCTGAAATTATCAGCGAGGAAGAGTTA AGAGAGGTTTTAAAAAAGATGAAAAATCTGCTGAGATAGGTTTTGAAACCAAGTGGTAA AATACATTTAGGGCATTATCTCCAATAAAAAAGATGATTGATTTACAAAATGCTGGATTT GATATAGGTATAATTTGGCTGATTAACTGCCATTAAACCAGAAAAGGAGAGTTGGATG AGATTAGAAAAATAGGAGATTATAACAAAAAGTTTTTGAAGCAATGGGGTTAAAGGCA AAATATGTTTATGGAAGTGAATCCGCTTGATAAGGATTATACACTGAATGTCTATAGATT GGCTTTAAAAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACCTTATAGCAAGAGAGG ATGAAAATCCAAAGGTTGCTGAAGTTATCTCCAATAATGCAGGTTAATACATTATAT ACTGGCGGTGATGTTGCAGTTGGAGGGATGGAGCAGAGAAAATAACACATGTTAGCAAG GGAGCTTTTACCAAAAAAGTTGTTTGTATTACAACCTGTTTAAACGGTTTGGATGG AGAAGGAAAGATGAGTTCTTCAAAAAGGAATTTATAGCTGTTGATGACTCTCCAGAAGA GATTAGGGCTAAGATAAAGAAAGCATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAAT AATGGAGATAGCTAAATACTTCTTGAATATCCTTAAACCATAAAAAAGGCCAGAAAAATTT GGTGGAGATTGACAGTTAATAGCTATGAGGAGTTAGAGAGTTTATTAAAAAATGAA TTGCATCCAATGGATTTAAAAAATGCTGTAGCTGAAGAAGTTATAAAGATTTTAGACCAA TTAGAAAAGAGATTATAA</p>
<p>肌红蛋白</p>	<p>SEQ ID NO: 8 Atgttctgtctgaagtgatggcagctggttctgcatgtttggcctaaagttgaagctgacgtcgtggtcattggtcaggacatctgattcagct gttcaaatctatccggaaactctggaaaaatcctgctgttcaaacatctgaaaactgaagctgaaatgaaagctctgaaatctgaaaaaacatg gtgtaccgtgtaactgcctaggtgctatccttaagaaaaaggcctcatgaagctgagctcaaacctgctcaaacatcgtactactaaaca taagatcccgatcaaatcctggaattcctctgaaagcgtatccatgttctgattctagacatccaggctgactcgtgctgagctcagggctg ctatgaacaaagctctcagctgttccgtaagatatcgtcctaagtacaaagaactgggtaccagggtgctcgggacatcatccatcacc attga</p>
<p>ApoA1</p>	<p>SEQ ID NO: 9 ATGCCGGAGCAGAGTCCATGGGATCGCGTGAAGGACCTGGCCACTGTGTACGTGGATGT GCTCAAAGACAGCGGCAGAGACTATGTCTCAGTTTGAAGGATCCGCCCTGGGCAAA AATTGAACCTTAAAGTCTGGACAACCTGGGACAGCGTGACGTTCCACTTACGCAAGCTG CGGAACAGCTCGGCCCTGTGACCCAGGAATTTCTGGGATAACCTGGAAAAGGAGACAG AGGGCCTGCGCCAGGAGATGAGCAAGGATCTGGAGGAGGTGAAAGGCCAAGGTGCAAGC GTACCTGGACGACTTCCAGAAGAAGTGGCAGGAGGATGCGAGACTTACCGCCAGAAG GTGGAGCCGCTGCGCGCAGAGCTGCAGGAGGGCGCGCCAGCTGCACGAGCTGC AAGAGAAGCTGAGCCACTGGGCGAGGAGATGCGCGACCCGCGCGCGCCCATGTCGA CGCGTGCACGACATCTGCGCCGTACAGCGACGAGCTGCGCCAGCGCTTAGCGGCGC GCCTTGAGGCTCTCAAGGAGAACGGCGGGGCCCGCTGGCCGAGTACCACGCCAAGGC CACCAGCATCTGAGCACGCTCAGCGAGAAGGCCAAGCCGGCTCGAGCACCACCAC CACCACACTGA</p>
<p>GFP</p>	<p>SEQ ID NO: 10 ATGAGTAAAGGAGAAGAAGACTTTTCACTGGAGTTGTCCCAATTCTTGTGAATTAGATGGT GATGTTAATGGGCACAAATTTCTGTCAAGTGGAGAGGGTGAAGGTGATGCAACATACGGA AACTTACCCTTAAATTTATTTGCACTACTGGAAAACCTGTTCCATGGCCAACTTTG TCACTACTTTCTTATGGTGTTCATGCTTTTCCCGTATCCGGATCACATGAAACGGCAT GACTTTTCAAGAGTGCCATGCCGGAAGGTTATGTACAGGAACGCACTATATCTTTCAA GATGACGGAACTACAAGACCGCTGTAAGTCAAGTTGAAGGTGATACCCCTTTGTTAA TCGTATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAACATTCTCGGACACAACT CGAATACAATACTACACAATGTATACATCACGGCAGACAAAACAAAAGAATGGAA TCAAAGCTAACTTCAAATTCGCCACAACATTGAAGATGGATCCGTTCAACTAGCAGACC ATTATCAACAATACTCCAATTTGGCGATGGCCCTGTCTTTTACAGACAACCTTACCT GTCGACACAATCTGCCCTTTTCAAAGATCCCAACGAAAAGCGTGACCACATGGTCTTCT TTGAGTTTGTAAGTCTGCTGGGATTACACATGGCATGGATGAGCTCGTCCCGCGCGGA GCTAGGGTGGTCTCGAGCACCACCACCACCACACTGA</p>
<p>mEOS2</p>	<p>SEQ ID NO: 11 ATGAGTGCATTAAGCCAGACATGAAGATCAAACCTCCGTATGGAAGGCAACGTAACCG GGCACCACCTTTGTGATCGACGGAGATGGTACAGGCAAGCCCTTTGAGGGAAAACAGAG TATGGATCTTGAAGTCAAAGAGGGCGGACCTCTGCCCTTTGCTTTGATATCCTGACCAC TGCATTCCATTACGGCAACAGGGTATTTCGCCAAATATCCAGACAACATACAAGACTATT TTAAGCAGTCGTTTCTAAGGGGATTCGTGGGAACGAAGCTTACTTTTGAAGACGGG GGCATTGCTGATGCCAGAAACGACATAACAATGGAAGGGGACACTTCTATAATAAAGT TCGATTTTATGGTACCAACTTTCCCGCAATGGTCCAGTTATGCAGAAGAAGACGCTGA AATGGGAGCCCTCCACTGAGAAAATGTATGTGCGTGATGGATGGATGCTGACGGGTGATATT CATATGGCTTTGTTGCTTGAAGGAAATGCCATTACCGATGTGACTTTCAGAACTACTTAC AAAGCTAAGGAGAAGGGTGTCAAGTTACCAGGCTACCCTTTGTGGACCCTGACATTGA GATTTTAAAGCATGACAAAGATTACAACAAGGTTAAGCTGTATGAGCATGCTGTTGCTC ATTCTGGATTGCCTGACAAATGCCAGACGA</p>
<p>野生型酪氨酰 tRNA 合 成酶 (MjTyrRS), 来源于詹氏甲烷球菌</p>	<p>SEQ ID NO: 12 MDEFEMIKRNTSEIIEEELREVLKCKDEK SAYIGFEPGSKIHLGHYLQIKKMIDLQNAFGDIIIL LADLHAYLNQKQELDEIRKIGDYNKKVFEAMGLKAKYVYGSEFQLDKDYTLNVYRLALKT TLKRARRSMELIAREDENPKVAEVIYPIMQVNDIHYLGVDVAVGGMEQRKIHMLARELLPK KVVCIHNPVLTGLDGEKMSSSKGNFIAVDDSPPEIRAKIKKAYCPAGVVEGNPIMEIAKYFL EYPLTIKRPEKFGGDLTVNSYEELLESLEFKNK ELHPMDLKNVAEELIKILEPIRKRLHHHHHHH</p>

图 2

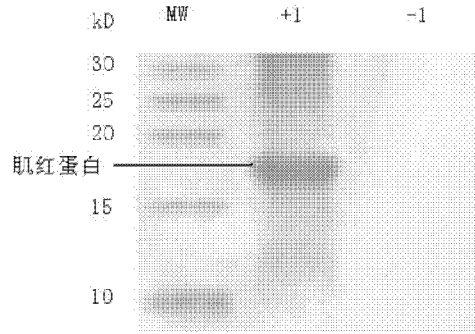


图 3

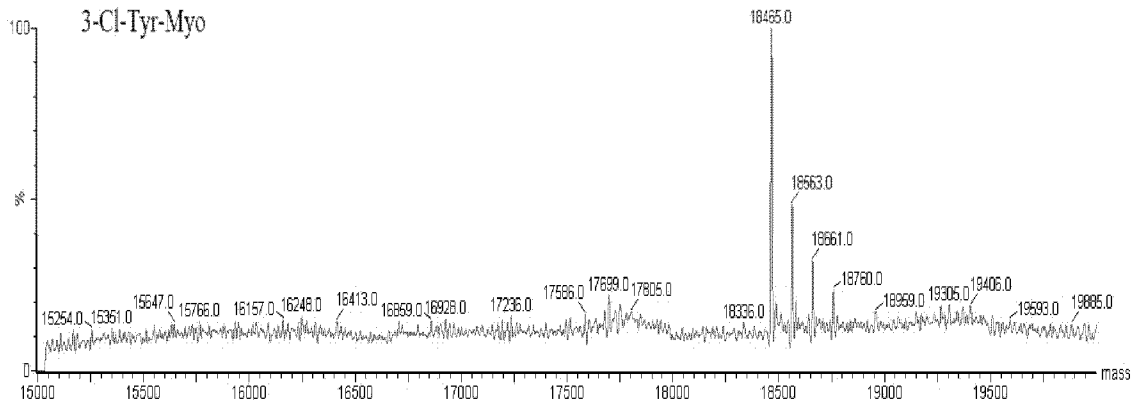


图 4

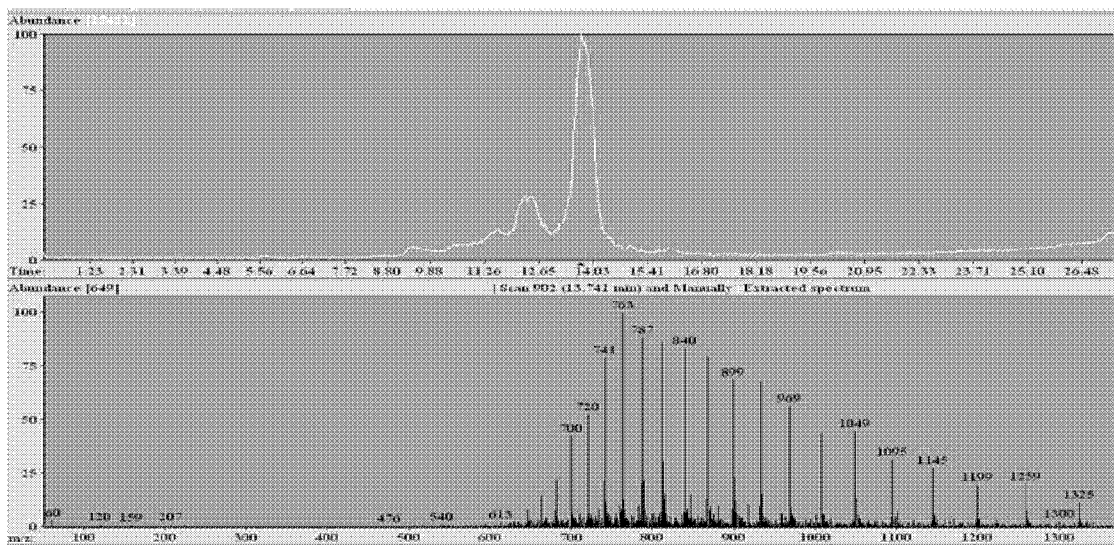


图 5

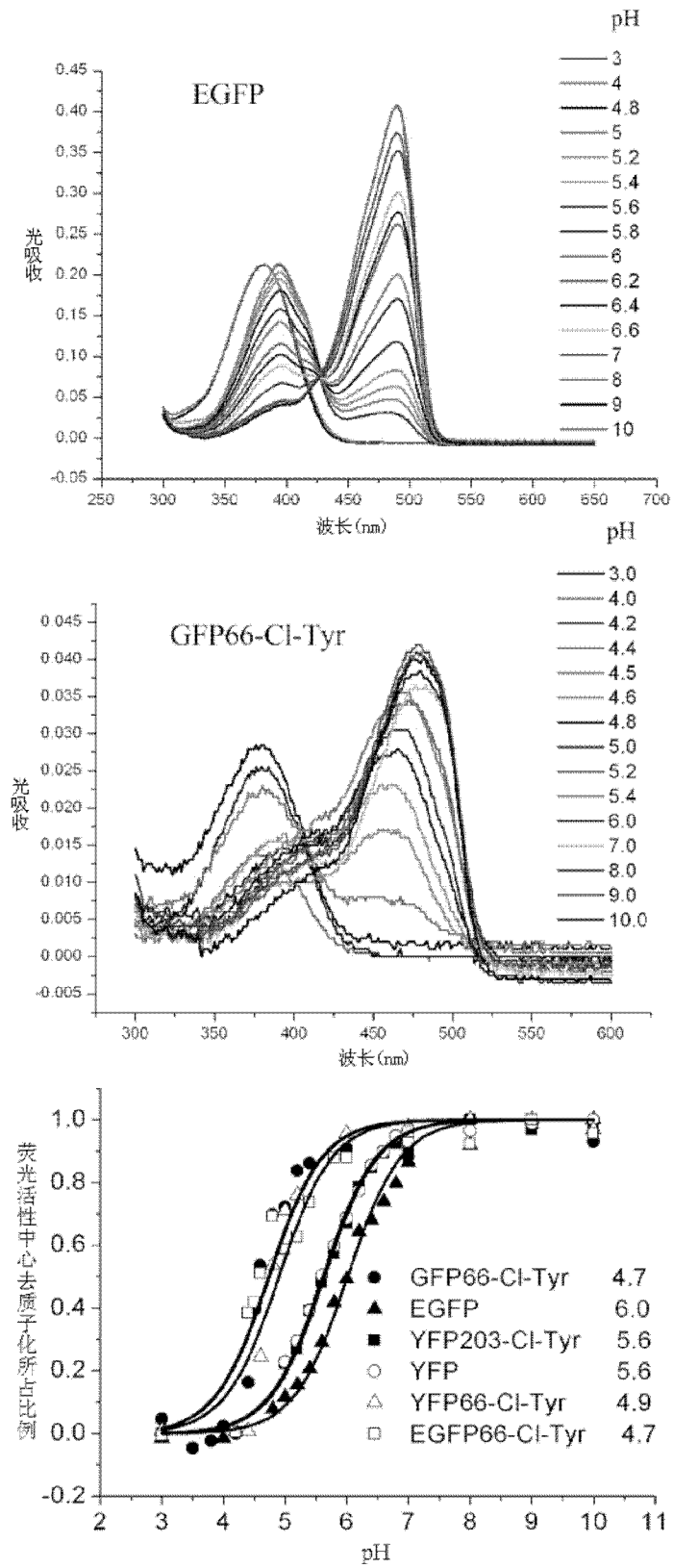


图 6

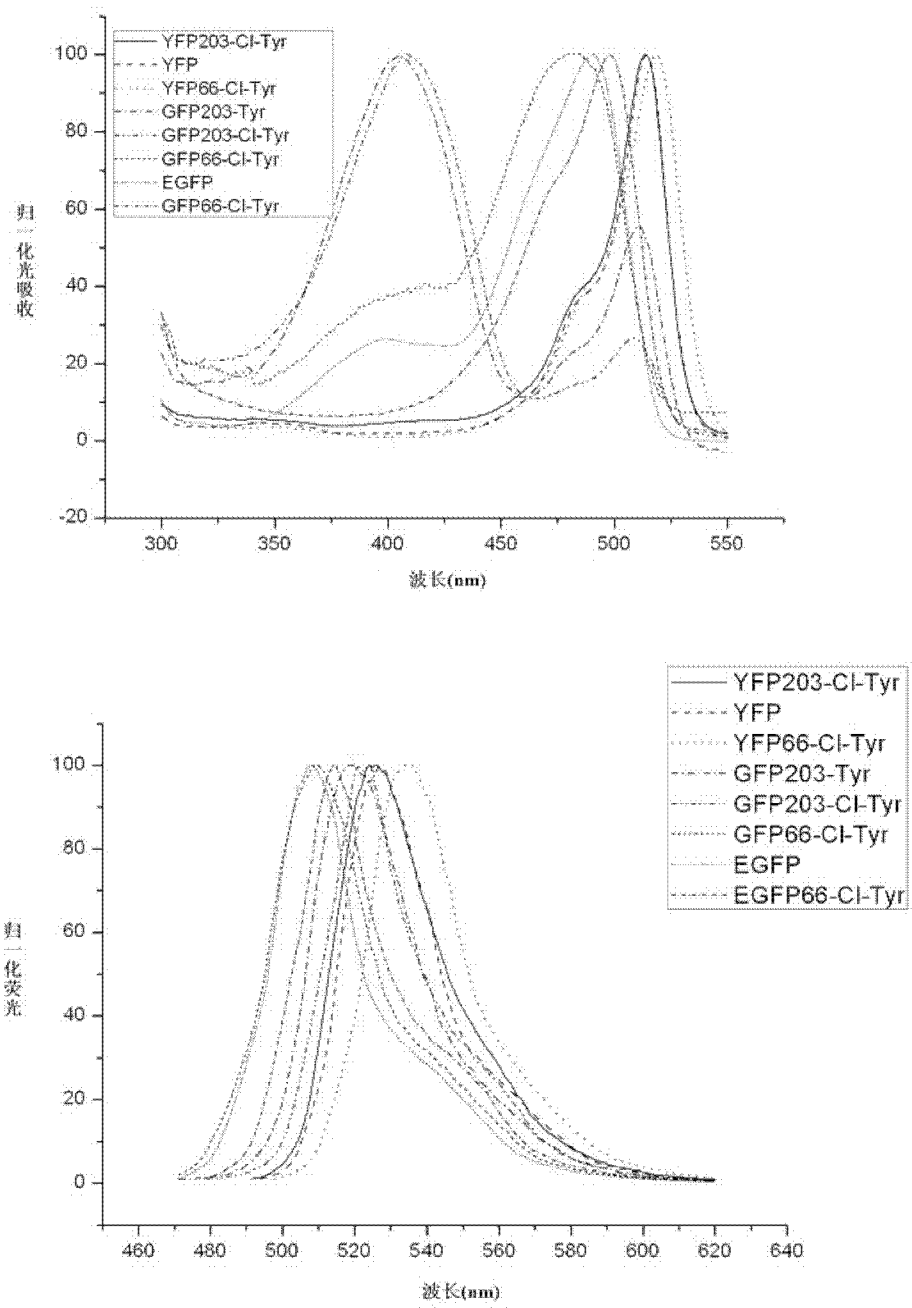


图 7

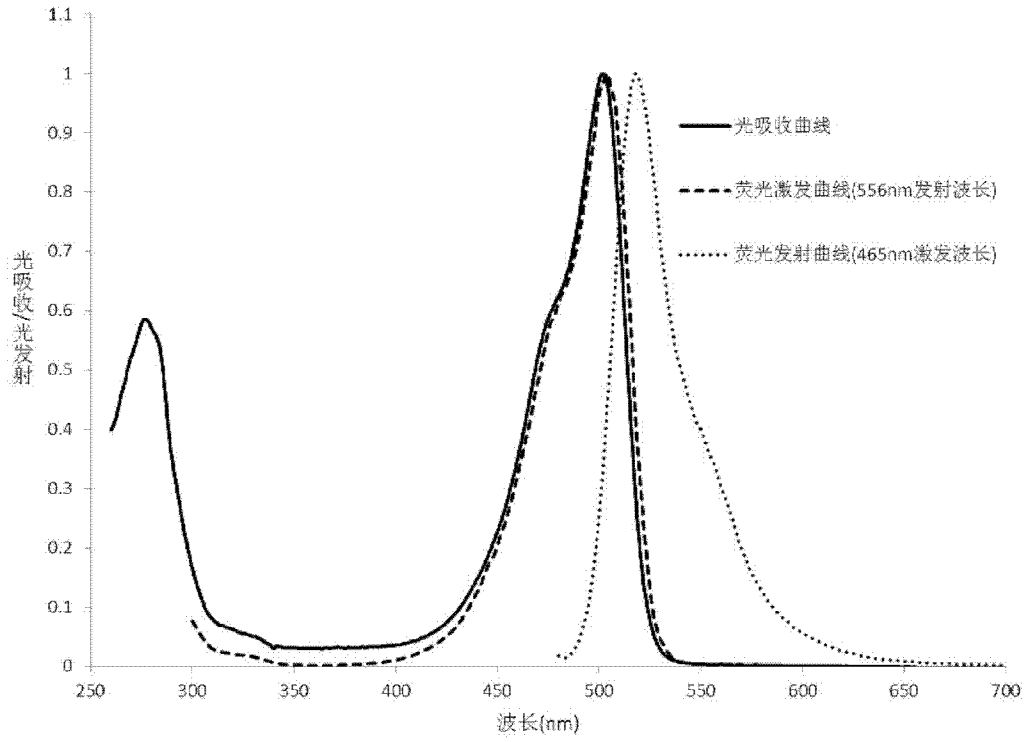


图 8

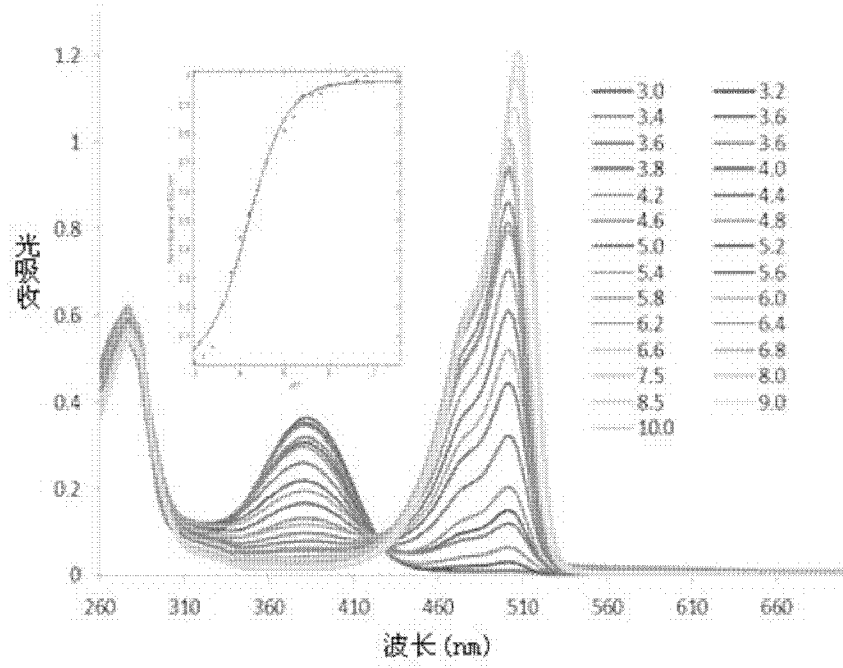


图 9