



(12) 发明专利申请

(10) 申请公布号 CN 104105794 A

(43) 申请公布日 2014. 10. 15

(21) 申请号 201280068693. 7

(51) Int. Cl.

(22) 申请日 2012. 12. 21

C12P 7/42 (2006. 01)

C12N 9/00 (2006. 01)

(30) 优先权数据

10-2012-0009300 2012. 01. 31 KR

(85) PCT国际申请进入国家阶段日

2014. 07. 31

(86) PCT国际申请的申请数据

PCT/KR2012/011295 2012. 12. 21

(87) PCT国际申请的公布数据

W02013/115484 KO 2013. 08. 08

(71) 申请人 尹喆湖

地址 韩国大田广域市

(72) 发明人 尹喆湖 柳相勋 姜馨植

(74) 专利代理机构 北京三友知识产权代理有限公司

11127

代理人 庞东成 丁香兰

权利要求书3页 说明书10页

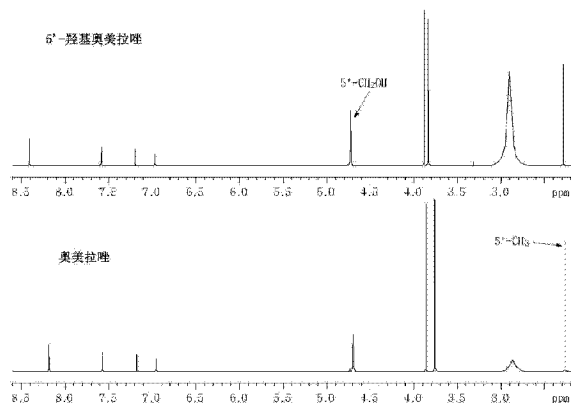
序列表9页 附图6页

(54) 发明名称

使用细菌细胞色素 P450 从奥美拉唑制备代谢物的新方法和用于其的组合物

(57) 摘要

本发明涉及一种使用细菌细胞色素 P450 从奥美拉唑制备代谢物的新方法和用于该新方法的组合物,更具体而言,涉及用于从奥美拉唑制备 5'-羟基产物的组合物和试剂盒及其制备方法,所述组合物和试剂盒含有细菌细胞色素 P450BM3(CYP102A1) 或其突变形式。所述组合物、试剂盒和方法能够从奥美拉唑经济高效地大规模制备 5'-羟基产物,因此会对使用奥美拉唑的代谢物的新型药物的开发作出显著贡献。



1. 选自由氨基酸序列 SEQ ID NO:16 所表示的野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶,

其中,所述突变的 CYP102A1 具有野生型 CYP102A1 被修饰后的序列,所述修饰为选自由以下替换组成的组的至少一种:第 48 位氨基酸精氨酸(R)被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 52 位氨基酸酪氨酸(Y)被选自由丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 65 位氨基酸谷氨酸(E)被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换;第 75 位氨基酸丙氨酸(A)被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换;第 82 位氨基酸苯丙氨酸(F)被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的一种氨基酸替换;第 87 位氨基酸亮氨酸(L)被选自由丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 88 位氨基酸苯丙氨酸(F)被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的一种氨基酸替换;第 144 位氨基酸谷氨酸(E)被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换;第 189 位氨基酸亮氨酸(L)被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换;和第 268 位氨基酸谷氨酸(E)被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换。

2. 如权利要求 1 所述的酶,其中,所述突变的 CYP102A1 具有野生型 CYP102A1 被修饰后的序列,所述修饰为选自由以下替换组成的组的至少一种:第 48 位氨基酸精氨酸(R)被亮氨酸(L)替换;第 52 位氨基酸酪氨酸(Y)被苯丙氨酸(F)替换;第 65 位氨基酸谷氨酸(E)被甘氨酸(G)替换;第 75 位氨基酸丙氨酸(A)被甘氨酸(G)替换;第 82 位氨基酸苯丙氨酸(F)被异亮氨酸(I)替换;第 87 位氨基酸亮氨酸(L)被异亮氨酸(I)替换;第 88 位氨基酸苯丙氨酸(F)被缬氨酸(V)替换;第 144 位氨基酸谷氨酸(E)被甘氨酸(G)替换;第 189 位氨基酸亮氨酸(L)被谷氨酰胺(Q)替换;和第 268 位氨基酸谷氨酸(E)被缬氨酸(V)替换。

3. 如权利要求 2 所述的酶,其中,在所述突变的 CYP102A1 中,野生型 CYP102A1 氨基酸中的替换位置和替换氨基酸是选自由以下组成的组的至少一种:F88A、R48L/Y52F、A75G/F88V/L189Q、R48L/L87I/L189Q、R48L/F88V/L189Q、R48L/F88V/L189Q/E268V、R48L/L87I/L189Q/E268V、48L/L87I/F88V/L189Q、R48L/F88V/E144G/L189Q/E268V、48L/E65G/F88V/E144G/L189Q/E268V、R48L/F82I/F88V/E144G/L189Q/E268V 和 R48L/E65G/F82I/F88V/E144G/L189Q/E268V。

4. 一种用于从奥美拉唑制备 5'-羟基产物的组合物,所述组合物含有选自由氨基酸序列 SEQ ID NO:16 所表示的野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶,

其中,所述突变的 CYP102A1 具有野生型 CYP102A1 被修饰后的序列,所述修饰为选自由以下替换组成的组的至少一种:第 48 位氨基酸精氨酸(R)被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 52 位氨基酸酪氨酸(Y)被选自由丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 65 位氨基酸谷氨酸(E)被选自由甘氨酸、丝氨酸、苏氨

酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换；第 75 位氨基酸丙氨酸 (A) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换；第 82 位氨基酸苯丙氨酸 (F) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的一种氨基酸替换；第 87 位氨基酸亮氨酸 (L) 被选自由丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换；第 88 位氨基酸苯丙氨酸 (F) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的一种氨基酸替换；第 144 位氨基酸谷氨酸 (E) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换；第 189 位氨基酸亮氨酸 (L) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换；和第 268 位氨基酸谷氨酸 (E) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换。

5. 如权利要求 4 所述的组合物,其中,所述突变的 CYP102A1 具有野生型 CYP102A1 被修饰后的序列,所述修饰为选自由以下替换组成的组的至少一种:第 48 位氨基酸精氨酸 (R) 被亮氨酸 (L) 替换;第 52 位氨基酸酪氨酸 (Y) 被苯丙氨酸 (F) 替换;第 65 位氨基酸谷氨酸 (E) 被甘氨酸 (G) 替换;第 75 位氨基酸丙氨酸 (A) 被甘氨酸 (G) 替换;第 82 位氨基酸苯丙氨酸 (F) 被异亮氨酸 (I) 替换;第 87 位氨基酸亮氨酸 (L) 被异亮氨酸 (I) 替换;第 88 位氨基酸苯丙氨酸 (F) 被缬氨酸 (V) 替换;第 144 位氨基酸谷氨酸 (E) 被甘氨酸 (G) 替换;第 189 位氨基酸亮氨酸 (L) 被谷氨酰胺 (Q) 替换;和第 268 位氨基酸谷氨酸 (E) 被缬氨酸 (V) 替换。

6. 如权利要求 5 所述的组合物,其中,在所述突变的 CYP102A1 中,野生型 CYP102A1 氨基酸中的替换位置和替换氨基酸是选自由以下组成的组的至少一种:F88A、R48L/Y52F、A75G/F88V/L189Q、R48L/L87I/L189Q、R48L/F88V/L189Q、R48L/F88V/L189Q/E268V、R48L/L87I/L189Q/E268V、48L/L87I/F88V/L189Q、R48L/F88V/E144G/L189Q/E268V、48L/E65G/F88V/E144G/L189Q/E268V、R48L/F82I/F88V/E144G/L189Q/E268V 和 R48L/E65G/F82I/F88V/E144G/L189Q/E268V。

7. 如权利要求 4 所述的组合物,其中,所述奥美拉唑是含有作为对映异构体的 S- 奥美拉唑或 R- 奥美拉唑的外消旋体,或是比例为 50:50 的 S- 奥美拉唑和 R- 奥美拉唑的对映异构体。

8. 一种从奥美拉唑制备 5'-羟基产物的方法,所述方法包括使奥美拉唑与选自由氨基酸序列 SEQ ID NO:16 所表示的野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶反应,

其中,所述突变的 CYP102A1 具有野生型 CYP102A1 被修饰后的序列,所述修饰为选自由以下替换组成的组的至少一种:第 48 位氨基酸精氨酸 (R) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 52 位氨基酸酪氨酸 (Y) 被选自由丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换;第 65 位氨基酸谷氨酸 (E) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换;第 75 位氨基酸丙氨酸 (A) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的

组的一种氨基酸替换；第 82 位氨基酸苯丙氨酸 (F) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的一种氨基酸替换；第 87 位氨基酸亮氨酸 (L) 被选自由丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换；第 88 位氨基酸苯丙氨酸 (F) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的一种氨基酸替换；第 144 位氨基酸谷氨酸 (E) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换；第 189 位氨基酸亮氨酸 (L) 被选自由甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的一种氨基酸替换；和第 268 位氨基酸谷氨酸 (E) 被选自由丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的一种氨基酸替换。

9. 如权利要求 8 所述的方法,其中,所述突变的 CYP102A1 具有野生型 CYP102A1 被修饰后的序列,所述修饰为选自由以下替换组成的组的至少一种:第 48 位氨基酸精氨酸 (R) 被亮氨酸 (L) 替换;第 52 位氨基酸酪氨酸 (Y) 被苯丙氨酸 (F) 替换;第 65 位氨基酸谷氨酸 (E) 被甘氨酸 (G) 替换;第 75 位氨基酸丙氨酸 (A) 被甘氨酸 (G) 替换;第 82 位氨基酸苯丙氨酸 (F) 被异亮氨酸 (I) 替换;第 87 位氨基酸亮氨酸 (L) 被异亮氨酸 (I) 替换;第 88 位氨基酸苯丙氨酸 (F) 被缬氨酸 (V) 替换;第 144 位氨基酸谷氨酸 (E) 被甘氨酸 (G) 替换;第 189 位氨基酸亮氨酸 (L) 被谷氨酰胺 (Q) 替换;和第 268 位氨基酸谷氨酸 (E) 被缬氨酸 (V) 替换。

10. 如权利要求 9 所述的方法,其中,在所述突变的 CYP102A1 中,野生型 CYP102A1 氨基酸中的替换位置和替换氨基酸是选自由以下组成的组的至少一种:F88A、R48L/Y52F、A75G/F88V/L189Q、R48L/L87I/L189Q、R48L/F88V/L189Q、R48L/F88V/L189Q/E268V、R48L/L87I/L189Q/E268V、R48L/L87I/F88V/L189Q、R48L/F88V/E144G/L189Q/E268V、R48L/E65G/F88V/E144G/L189Q/E268V、R48L/F82I/F88V/E144G/L189Q/E268V 和 R48L/E65G/F82I/F88V/E144G/L189Q/E268V。

11. 如权利要求 8 所述的方法,所述方法还包括:添加产生 NADPH 的体系。

12. 一种用于从奥美拉唑制备 5'-羟基产物的试剂盒,所述试剂盒包含:产生 NADPH 的体系,和选自由氨基酸序列 SEQ ID NO:16 所表示的野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶,

其中,在所述突变的 CYP102A1 中,野生型 CYP102A1 氨基酸中的替换位置和替换氨基酸是选自由以下组成的组的至少一种:F88A、R48L/Y52F、A75G/F88V/L189Q、R48L/L87I/L189Q、R48L/F88V/L189Q、R48L/F88V/L189Q/E268V、R48L/L87I/L189Q/E268V、R48L/L87I/F88V/L189Q、R48L/F88V/E144G/L189Q/E268V、R48L/E65G/F88V/E144G/L189Q/E268V、R48L/F82I/F88V/E144G/L189Q/E268V 和 R48L/E65G/F82I/F88V/E144G/L189Q/E268V。

13. 如权利要求 12 所述的试剂盒,其中,所述产生 NADPH 的体系含有 6-磷酸葡萄糖、NADP<sup>+</sup> 和酵母 6-磷酸葡萄糖脱氢酶。

## 使用细菌细胞色素 P450 从奥美拉唑制备代谢物的新方法和用于其的组合物

### 技术领域

[0001] 本发明涉及一种使用细菌细胞色素 P450 从奥美拉唑制备代谢物的新方法和用于该新方法的组合物,更具体而言,涉及用于从奥美拉唑制备 5'-羟基产物的组合物和试剂盒及其制备方法,所述组合物和试剂盒含有细菌细胞色素 P450BM3 (CYP102A1) 或其突变形式。

### 背景技术

[0002] 奥美拉唑是质子泵抑制剂,已知其是消化不良、胃溃疡、胃食管反流病和咽喉反流病的治疗剂。奥美拉唑是外消旋体,含有比例为 50:50 的 S 和 R 对映异构体。在以上酸性条件下,这两种对映异构体都转化成非手性化合物,并且与 H<sup>+</sup>/K<sup>+</sup>ATP 酶的半胱氨酸基团反应,从而抑制胃壁细胞中的胃酸产生。奥美拉唑和对映异构体被存在于人肝脏中的细胞色素 P450CYP2C19 和 CYP3A4 代谢,其主要代谢物包括 5'-O-脱甲基奥美拉唑、5'-和 3'-羟基奥美拉唑以及奥美拉唑砒(参见 Renberg 等, Drug Metab Dispos 17:69-76, 1989; Andersson 等, Clin Pharmacokinet 40:411-426, 2001; Li 等, J Pharmacol Exp Ther 315:777-787, 2005)。据报道, R 对映异构体通常被 CYP2C19 代谢为 5'-O-脱甲基奥美拉唑、5'-羟基奥美拉唑, S 对映异构体通常被 CYP3A4 代谢为奥美拉唑砒、3'-羟基奥美拉唑。

[0003] 细胞色素 P450 (P450 或 CYP) 酶是个大家族,该家族由多种酶组成,这些酶充当从古细菌至细菌、真菌、植物、动物和人的整个自然界中的多种多样的氧化反应的催化剂。由于催化功能的多样性以及其底物范围广, P450 可大规模地用作制备精细化工品(包括药物供应等)时的生物催化剂(参见 Guengerich, Nat Rev Drug Discov 1:359-366, 2002; Urlacher 等, Trends Biotechnol 24:324-330, 2006; Yun CH 等, Trends Biotechnol 25:289-298, 2007; Lamb 等, Curr Opin Biotechnol 18:504-512, 2007)。但是,虽然哺乳动物的细胞色素 P450 酶在各种上述生物技术领域中具有潜在的可用性,但是 P450 的稳定性、催化活性和可获得性较低,因此不适合作为生物催化剂。

[0004] 当在开发药物期间人利用 P450 将前药转化为生物“活性代谢物”时(参见 Johnson 等, Breast Cancer Res. Treat 85:151-159, 2004),需要大量的纯代谢物来研究该药物的功效、毒性和药代动力学等。此外,当代谢物自身具有生物活性时,体内直接施用所述代谢物具有巨大裨益,因此大规模生产所述代谢物是重要的。

[0005] 当将奥美拉唑施用至人体中时,由于奥美拉唑被 CYP2C19 和 CYP3A4 代谢,代谢物产生的速率会根据酶的表达程度而有所不同。另外,会出现药物与被所述酶代谢的其他药物相互作用的问题。因此,当将奥美拉唑代谢物直接用作药物时,可以避免药物相互作用问题。

[0006] 但是,由于在化学合成纯代谢物时存在各种问题,为了以替代代谢物化学合成的方式产生药物代谢物或药物候选物,使用了 P450。已经报道了使用表达自大肠杆菌(参见 Yun 等, Curr Drug Metab 7:411-429, 2006) 或昆虫细胞(参见 Rushmore 等, Metab

Eng2:115-125, 2000 ;Vail 等, J Ind Microbiol Biotechnol32:67-74, 2005) 的人 P450 来生产所述代谢物。但是, 由于稳定性有限和反应速率低等原因, 这些体系具有诸如成本昂贵和产生率低等问题 (参见 Guengerich 等, Crit Rev Toxicol26:551-583, 1996)。因此, 已经提出了使用具有所需催化活性的工程化细菌 P450 酶作为用于产生人体中的代谢物的替代途径的方法 (参见 Yun CH 等, Trends Biotechnol25:289-298, 2007)。

[0007] 同时, 源自巨大芽孢杆菌 (*Bacillus megaterium*) 的 P450BM3(CYP102A1) 的血红素结构域具有单加氧酶活性, 其与 CYP4A(脂肪酸羟化酶) 家族的哺乳动物成员极其相似。天然地, 其由单个的多肽构成, 其中, 具有哺乳动物样二黄素还原酶功能的 CYP102A1 还原酶结构域融合至 P450 血红素结构域的 C 末端。两种酶活性的融合使得融合性 CYP102A1 成为期望的哺乳动物模型, 特别是人 P450 酶的期望模型。据报道, 通过逻辑性设计或定向进化而遗传改造得到的突变 CYP102A1 可氧化人 P450 的若干种底物, 从而产生具有更高活性的代谢物 (参见 Kim 等, Drug Metab Dispos36:2166-2170, 2008 ;Kim 等, Drug Metab Dispos37:932-936, 2009 ;Kim 等, J Mol Catal B:Enzym63:179-187, 2010 ;Otey 等, Biotechnol Bioeng93:494-499, 2006 ;Yun CH 等, Trends Biotechnol25:289-298, 2007)。

[0008] 基于上述描述, 已提出可以开发突变的 CYP102A1 作为用于药物的检测和合成的生物催化剂。最近已报道, 数种选出的突变形式可以允许 CYP102A1 酶产生人体中的代谢物作为药物 (参见 Kim 等, Drug Metab Dispos36:2166-2170, 2008 ;Kim 等, Drug Metab Dispos37:932-936, 2009), 但是, 尚未报道过以生物学手段从奥美拉唑产生人体中的代谢物的方法。

## 发明内容

[0009] 本发明的目的在于提供一种在通过氧化奥美拉唑而产生 5'-羟基产物的选择性转化反应中能够更稳定更有效地发挥催化功能的酶。

[0010] 另外, 本发明的另一目的在于提供一种用于从奥美拉唑制备 5'-羟基产物的组合物, 所述组合物含有所述酶。

[0011] 此外, 本发明的另一目的在于提供一种从奥美拉唑制备 5'-羟基产物的方法, 所述方法包括使所述酶与奥美拉唑反应。

[0012] 另外, 本发明的另一目的在于提供一种用于从奥美拉唑制备 5'-羟基产物的试剂盒, 所述试剂盒含有所述酶和产生 NADPH 的体系。

[0013] [技术方案]

[0014] 在一个总体方面, 本发明提供选自野生型 CYP102A1 和突变的 CYP102A1 组成的至少一种酶。

[0015] 所述酶可以在通过氧化奥美拉唑而产生 5'-羟基产物的选择性转化反应中更稳定更有效地发挥催化功能。

[0016] 在另一总体方面, 本发明提供一种使用野生型 CYP102A1 和突变的 CYP102A1 从奥美拉唑选择性地大规模制备人体中的代谢物、特别是 5'-羟基产物的方法, 以及用于该方法的组合物和试剂盒, 其中, CYP102A1 是细菌 P450 酶。

[0017] 本发明的野生型 CYP102A1 和突变的 CYP102A1 可以用作以奥美拉唑作为底物的氧

化反应中的催化剂,已知奥美拉唑是人 P450 的底物;特别而言,使用人 CYP2C19 作为催化剂时产生的奥美拉唑代谢物包括两种代谢物;而使用本发明的细菌 CYP102A1 或其突变形式作为催化剂时,可以选择性地产生 5'-羟基产物。

[0018] 在本发明的优选示例性实施方式中,本发明人确认了当细菌野生型 CYP102A1 和其定点突变形式在大肠杆菌中大规模表达(参见表 1 和 2)且使奥美拉唑与产生 NADPH 的体系反应时,通过 HPLC(参见图 3、4 和 5)和 LC-MS 谱(参见图 6)确认奥美拉唑转化成了人体中的代谢物。经确认,在使用人 CYP2C19 时,奥美拉唑被氧化产生两种主要代谢物,即 3'-羟基奥美拉唑和 5'-羟基奥美拉唑;但在使用细菌野生型 CYP102A1 及其突变形式时,选择性地产生了一种主要产物,即 5'-羟基奥美拉唑。

[0019] 在产物的产生方面,野生型 CYP102A1、17 种突变形式和 6 种突变嵌合形式具有范围很宽的分子催化活性(转换数)(参见图 5)。经确认,对于在总分子催化活性方面显示出高活性的突变 #10,在以 1mM 浓度(A)的奥美拉唑反应 2~4 小时(B)的反应中显示出了最高活性;同时,对于野生型 CYP102A1 酶,几乎不显示对于奥美拉唑的活性(参见图 7)。

[0020] 基于上述检验结果,在另一总体方面,本发明提供一种用于从奥美拉唑制备 5'-羟基产物的组合物,所述组合物含有选自野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶,

[0021] 其中,所述突变的 CYP102A1 具有由氨基酸序列 SEQ ID NO:16 表示的野生型 CYP102A1 被修饰后的序列,所述修饰为选自以下替换组成的组的至少一种:第 48 位氨基酸精氨酸(R)被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换;第 52 位氨基酸酪氨酸(Y)被选自丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换;第 65 位氨基酸谷氨酸(E)被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;第 75 位氨基酸丙氨酸(A)被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;第 82 位氨基酸苯丙氨酸(F)被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的氨基酸替换;第 87 位氨基酸亮氨酸(L)被选自丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换;第 88 位氨基酸苯丙氨酸(F)被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的氨基酸替换;第 144 位氨基酸谷氨酸(E)被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;第 189 位氨基酸亮氨酸(L)被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;和第 268 位氨基酸谷氨酸(E)被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换。

[0022] 所述奥美拉唑可以是含有作为对映异构体的 S- 或 R- 奥美拉唑的外消旋体,或是比例为 50:50 的 S- 和 R- 奥美拉唑的对映异构体,但是本发明不限于此。

[0023] 在另一总体方面,本发明提供一种从奥美拉唑制备 5'-羟基产物的方法,所述方法包括使奥美拉唑与选自野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶反应。

[0024] 在所述制备 5'-羟基产物的方法中,所述突变的 CYP102A1 可以优选具有由氨基酸

序列 SEQ ID NO:16 表示的野生型 CYP102A1 被修饰后的序列,所述修饰选自以下替换组成的组的至少一种:第 48 位氨基酸精氨酸 (R) 被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换;第 52 位氨基酸酪氨酸 (Y) 被选自丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换;第 65 位氨基酸谷氨酸 (E) 被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;第 75 位氨基酸丙氨酸 (A) 被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;第 82 位氨基酸苯丙氨酸 (F) 被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的氨基酸替换;第 87 位氨基酸亮氨酸 (L) 被选自丙氨酸、缬氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换;第 88 位氨基酸苯丙氨酸 (F) 被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸和色氨酸组成的组的氨基酸替换;第 144 位氨基酸谷氨酸 (E) 被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;第 189 位氨基酸亮氨酸 (L) 被选自甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺组成的组的氨基酸替换;和第 268 位氨基酸谷氨酸 (E) 被选自丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸组成的组的氨基酸替换。

[0025] 在本发明中,突变的 CYP102A1 可以通过本领域已知的任何已知突变方法来制造,例如缺失突变法(参见 Kowalski D. 等, *J. Biochem.*, 15, 4457)、PCT 法、Kunkel 法、定点突变法、DNA 改组、交错延伸法 (StEP) 和易错 PCR 等。

[0026] 在本发明的突变的 CYP102A1 中,由 SEQ ID NO:16 表示的野生型 CYP102A1 蛋白的氨基酸序列通过天然或人工的替换、缺失、添加和 / 或插入而被修饰。优选的是,按照以下分类,所换上的氨基酸可以具有与待替换的氨基酸相似的性质。例如,丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、甲硫氨酸、苯丙氨酸和色氨酸都被分类成非极性氨基酸,并且彼此具有相似的性质。不带电氨基酸的实例可以包括甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺等,酸性氨基酸的实例可以包括天冬氨酸和谷氨酸,碱性氨基酸的实例可以包括赖氨酸、精氨酸和组氨酸。

[0027] 本发明的突变的 CYP102A1 包括含有与由 SEQ ID NO:16 表示的 CYP102A1 蛋白序列具有至少 50% 同一性、优选至少 75% 同一性、更优选至少 90% 同一性的氨基酸序列的多肽。

[0028] 野生型 CYP102A1 的期望突变形式可以包括对 SEQ ID NO:16 所表示的野生型 CYP102A1 进行的选自以下替换组成的组的至少一种替换:第 48 位氨基酸精氨酸 (R) 被亮氨酸 (L) 替换;第 52 位氨基酸酪氨酸 (Y) 被苯丙氨酸 (F) 替换;第 65 位氨基酸谷氨酸 (E) 被甘氨酸 (G) 替换;第 75 位氨基酸丙氨酸 (A) 被甘氨酸 (G) 替换;第 82 位氨基酸苯丙氨酸 (F) 被异亮氨酸 (I) 替换;第 87 位氨基酸亮氨酸 (L) 被异亮氨酸 (I) 替换;第 88 位氨基酸苯丙氨酸 (F) 被缬氨酸 (V) 替换;第 144 位氨基酸谷氨酸 (E) 被甘氨酸 (G) 替换;第 189 位氨基酸亮氨酸 (L) 被谷氨酰胺 (Q) 替换;和第 268 位氨基酸谷氨酸 (E) 被缬氨酸 (V) 替换。

[0029] 在最优选的突变 CYP102A1 中,由 SEQ ID NO:16 表示的野生型 CYP102A1 氨基酸中的替换位置和替换氨基酸可以选自以下组成的组:F88A、R48L/Y52F、A75G/F88V/L189Q、



R48L/L87I/L189Q、R48L/F88V/L189Q、R48L/F88V/L189Q/E268V、R48L/L87I/L189Q/E268V、R48L/L87I/F88V/L189Q、R48L/F88V/E144G/L189Q/E268V、R48L/E65G/F88V/E144G/L189Q/E268V、R48L/F82I/F88V/E144G/L189Q/E268V 和 R48L/E65G/F82I/F88V/E144G/L189Q/E268V。

[0030] 本发明的蛋白质可以通过本领域公知的方法产生,例如使用基因工程技术的肽合成法(Merrifield, J. Am. Chem. Soc., 85:2149-2154, 1963 文献)、固相技术或利用合适的肽酶切割本发明的蛋白的方法。本发明的蛋白质可以作为天然蛋白产生,或者可以通过培养经转化而具有编码 CYP102A1 或其突变形式的 DNA 的细胞、并回收所转化的细胞的重组方法产生。本发明的蛋白质可以通过下述过程产生:将编码本发明的蛋白质的核酸分子插入合适的表达载体中,培养通过将该载体递送至合适的细胞而产生的转化体,并对转化体所表达的蛋白质进行纯化。

[0031] 载体可以具有例如质粒、粘粒、病毒颗粒或噬菌体形式。克隆或表达载体中的 DNA 的宿主细胞的实例可以包括原核细胞、酵母和高级真核细胞。本领域技术人员可以对诸如培养基、温度和 pH 等培养条件进行合适地选择而无需过多的实验。通常,可以参考 Mammalian Cell Biotechnology:a Practical Approach, M. Butler 编。(IRL Press, 1991) 而利用使细胞培养的生产率最大化的技术。

[0032] 表达和克隆载体通常可以含有与编码 CYP102A1 或其突变形式的核酸序列可操作地连接的诱导 mRNA 合成的启动子。被宿主细胞识别的各种启动子是已知的。适用于原核细胞宿主的启动子的实例包括  $\beta$ -内酰胺酶和乳糖启动子系统、碱性磷酸酶、色氨酸启动子系统和杂合启动子,例如 tac 启动子。此外,用于细菌系统的启动子可以含有与编码 SISP-1 的 DNA 可操作地连接的 Shine-Dalgarno (S. D.) 序列。适用于酵母宿主的启动子序列的实例可以包括 3-磷酸甘油酸酯激酶或其他糖解酶。

[0033] 从奥美拉唑产生 5'-羟基产物的方法可以进一步包括:添加产生 NADPH 的体系。

[0034] 在另一总体方面,本发明提供一种用于从奥美拉唑制备 5'-羟基产物的试剂盒,所述试剂盒包含:产生 NADPH 的体系,和选自由野生型 CYP102A1 和突变的 CYP102A1 组成的组的至少一种酶。

[0035] 在突变的 CYP102A1 中,由 SEQ ID NO:16 表示的野生型 CYP102A1 氨基酸中的替换位置和替换氨基酸可以优选是选自由以下组成的组的至少一种:F88A、R48L/Y52F、A75G/F88V/L189Q、R48L/L87I/L189Q、R48L/F88V/L189Q、R48L/F88V/L189Q/E268V、R48L/L87I/L189Q/E268V、R48L/L87I/F88V/L189Q、R48L/F88V/E144G/L189Q/E268V、R48L/E65G/F88V/E144G/L189Q/E268V、R48L/F82I/F88V/E144G/L189Q/E268V 和 R48L/E65G/F82I/F88V/E144G/L189Q/E268V,但是本发明不限于此。此外,所述试剂盒还可以含有进行所述反应所需的试剂。

[0036] 产生 NADPH 的体系可以含有 6-磷酸葡萄糖、 $\text{NADP}^+$  和酵母 6-磷酸葡萄糖脱氢酶,但是本发明不限于此。

[0037] CYP102A1 及其突变形式是在通过氧化奥美拉唑而产生 5'-羟基产物的选择性转化反应中能够稳定有效地发挥催化功能的细菌酶(奥美拉唑已知是人 P450 的底物),因此可有效地用于以生物学手段从奥美拉唑产生人体中的代谢物。

[0038] [有益效果]

[0039] 本发明的细菌野生型 CYP102A1 及其突变形式可以在将奥美拉唑转化成 5'-羟基产物的转化反应中更稳定更有效地发挥催化功能,从而能够环保地、选择性地大规模制备 5'-羟基产物。本发明的产生 5'-羟基产物的组合物、试剂盒和方法可以包括所述细菌野生型 CYP102A1 及其突变形式,从而能够经济高效地从奥美拉唑大规模制备 5'-羟基产物,因此显著地有助于开发使用奥美拉唑代谢物的新型药物。

#### 附图说明

[0040] 以下将结合附图对优选实施方式进行了描述,由此会使本发明的上述和其他目的、特征和优点变得显而易见,附图中:

[0041] 图 1 显示本发明的野生型 CYP102A1 (SEQ ID NO:16) 的氨基酸序列。

[0042] 图 2 显示本发明的野生型 CYP102A1 (SEQ ID NO:17) 的碱基序列。

[0043] 图 3 显示由本发明的野生型 CYP102A1 及其突变形式产生的奥美拉唑代谢物的 HPLC 色谱(在 302nm 测得的 UV 吸光度)(峰:通过由人 CYP2C19 产生的代谢物的峰的保留时间来确认;箭头:指示底物和作为主要产物的 5'-羟基产物)。

[0044] 图 4 显示由本发明的野生型 CYP102A1 的突变形式(#10)产生的奥美拉唑代谢物衍生物的 HPLC 色谱(A:外消旋体;B:R 对映异构体;C:S 对映异构体)。

[0045] 图 5 显示了本发明的野生型 CYP102A1 及其突变形式产生奥美拉唑氧化物的产生速率。

[0046] 图 6 显示由人 CYP2C19 和本发明的突变 CYP102A1 (#10)产生的奥美拉唑和其代谢物的 LC-MS 洗脱图(A:突变 CYP102A1#10;B:人 CYP2C19)。

[0047] 图 7 显示了本发明的突变 CYP102A1 (#10)产生 5'-羟基产物的总分子催化活性随奥美拉唑的浓度(A)和处理时间(B)的变化。

[0048] 图 8 显示通过核磁共振(NMR)图谱观察到的由本发明的突变 CYP102A1 (#10)产生的奥美拉唑代谢物的结构。

#### 具体实施方式

[0049] 以下将参照附图来详细描述本发明的实施方式。

[0050] 但是,该详细描述是为了帮助具体理解本发明,本发明的保护范围不限于以下实施例。

[0051] <实施例 1>通过定点诱变来构建突变的 P450BM3

[0052] 按照 Kim 等(参见 Drug Metab Dispos 第 35 卷,第 2166-2170 页,2008)所述的方法产生 CYP102A1 的 17 种定点突变形式。用于引入 BamHI/SacI 的识别位点的引物和用于进行诱变的 PCR 引物示于以下表 1 中。对应于氨基酸替换的密码子以斜体加下划线表示。PCR 引物购自 Genotech Company (Daejeon, Korea)。使用设计用来促进表达载体 pCWori(获自 Dr. F. W. Dahlquist, University of California, Santa Barbara, CA)或 pSE420 (Invitrogen) 克隆的引物,通过 PCR 法从 pCWBM3 扩增编码突变的 CYP102A1 的基因。

[0053] 使用下表 1 所示的 14 种设计的引物组来实施寡核苷酸组装。将扩增的基因克隆至 PCWBM3BamHI/SacI 载体的 BamHI/SacI 识别位点。使用经质粒转化的大肠杆菌 DH5  $\alpha$  F'-IQ (Invitrogen) 来表达 CYP102A1 突变蛋白。诱变后,通过 Genotech

Company (Daejeon) 的 DNA 测序来确认是否出现了所需的突变。

[0054] [表 1]

[0055] 用于突变的引物

[0056]

名称	序列
BamHI 正向 (SEQ ID NO: 1)	5' -AGC <del>GGA</del> TCC ATG ACA ATT AAA GAA ATG CCT C-3'
SacI 反向 (SEQ ID NO: 2)	5' -ATC GAG CTC GTA GTT TGT AT-3'
R47L (SEQ ID NO: 3)	5' -GCG CCT GGT <del>CAG</del> GTA ACG CG-3'
Y51F (SEQ ID NO: 4)	5' -GTA ACG CGC <del>TTC</del> TTA TCA AGT-3'
E64G (SEQ ID NO: 5)	5' -GCA TGC GAT <del>GAC</del> TCA CGC TTT-3'
A74G (SEQ ID NO: 6)	5' -TA AGT CAA <del>GAC</del> CTT AAA TTT GTA CG-3'
F81I (SEQ ID NO: 7)	5' -GTA CGT GAT <del>ATT</del> GCA GGA GAC-3'
L86I (SEQ ID NO: 8)	5' -GGA GAC GGG <del>ATT</del> TTT ACA AGC T-3'
F87A (SEQ ID NO: 9)	5' -GAC GGG TTA <del>GAG</del> ACA AGC TGG-3'
F87V (SEQ ID NO: 10)	5' -GAC GGG TTA <del>GAG</del> ACA AGC TGG-3'
E143G (SEQ ID NO: 11)	5' -GAA GTA CCG <del>GAC</del> GAC ATG ACA-3'
L188Q (SEQ ID NO: 12)	5' -ATG AAC AAG <del>CAG</del> CAG CGA GCA A-3'
A264G (SEQ ID NO: 13)	5' -TTC TTA ATT <del>GAG</del> GGA CAC GTG-3'
E267V (SEQ ID NO: 14)	5' -T GCG GGA CAC <del>GAG</del> ACA ACA AGT-3'
L86I/F87V (SEQ ID NO: 15)	5' -GGA GAC GGG <del>ATT</del> <del>GAG</del> ACA AGC TG-3'

[0057] < 实施例 2> 野生型 CYP102A1 (pCWBM3) 及其突变形式的表达和纯化

[0058] 用含有野生型 CYP102A1 和突变 CYP102A1 的基因的质粒转化大肠杆菌 DH5  $\alpha$  F' -IQ (参见 Kim 等, 2008b)。从一个菌落取适当的量接种到 5ml 添加有氨苄青霉素 (100  $\mu$ g/ml) 的 Luria-Bertani 培养基中, 然后在 37°C 培养, 将培养物接种到 250ml 添加有氨苄青霉素 (100  $\mu$ g/ml) 的 Terrific Broth 培养基中, 并在 37°C 以 250rpm 振荡的同时培养至 OD600 为 0.8, 向其中添加异丙基- $\beta$ -D- 硫代半乳糖苷至终浓度为 0.5mM, 由此诱导基因表达。向其中添加  $\delta$ -氨基乙酰丙酸 (0.1mM)。在诱导表达后, 在 30°C 下再另外培养 36 小时, 离心 (15 分钟, 5000g, 4°C), 收获细胞。用 TES 缓冲液 (100mM Tris-HCl, pH7.6, 500mM 蔗糖, 0.5mM EDTA) 将细胞沉淀块再悬浮, 超声 (超声发生器; Misonix, Inc., Farmingdale, NY) 裂解细胞。在 100,000g、90 分钟和 4°C 的条件下离心细胞裂解物, 收集可溶性细胞溶质组分以测量活性。将细胞溶质组分透析至 50mM 磷酸钾缓冲液 (pH7.4) 中, 并在 -80°C 贮存, 使用制备后 1 个月内的组分进行实验。通过 CO 差光谱来确定 CYP102A1 的浓度, 其中  $\epsilon$  为 91mM/cm。在野生型 CYP102A1 和突变的 CYP102A1 中, 都大体上获得了 300 ~ 700nM 的 P450。野生型 CYP102A1 及其突变形式的表达程度为 1.0 ~ 2.0nmol P450/mg 细胞基质蛋白 (cell substrate protein)。在所产生的突变形式中, 选出针对人体中数种底物都具有高催化活性的突变形式, 并且在下表 2 中示出了各突变形式中的氨基酸替换区域。

[0059] [表 2]

[0060]

简称	BM3 野生型及突变形式	参考文献
WT	BM3 野生型	
突变 #1	F87A	Carmichael 等, 2001
突变 #2	A264G	Carmichael 等, 2001
突变 #3	F87A/A264G	Carmichael 等, 2001
突变 #4	R47L/Y51F	Carmichael 等, 2001
突变 #5	R47L/Y51F/A264G	Carmichael 等, 2001
突变 #6	R47L/Y51F/F87A	Carmichael 等, 2001
突变 #7	R47L/Y51F/F87A/A264G	Carmichael 等, 2001
突变 #8	A74G/F87V/L188Q	Li 等, 2001
突变 #9	R47L/L86V/L188Q	Kim 等, 2008a
突变 #10	R47L/F87V/L188Q	van Vugt-Lussenburg 等, 2007
突变 #11	R47L/F87V/L188Q/E267V	van Vugt-Lussenburg 等, 2007
突变 #12	R47L/L86V/L188Q/E267V	Kim 等, 2008
突变 #13	R47L/L86V/F87V/L188Q	van Vugt-Lussenburg 等, 2007
突变 #14	R47L/F87V/E143G/L188Q/E267V	Kim 等, 2008a
突变 #15	R47L/E64G/F87V/E143G/L188Q/E267V	Kim 等, 2008a
突变 #16	R47L/F81V/F87V/E143G/L188Q/E267V	Kim 等, 2008a
突变 #17	R47L/E64G/F81V/F87V/E143G/L188Q/E267V	van Vugt-Lussenburg 等, 2007

[0061] < 实施例 3> 野生型 CYP102A1 或其突变形式对奥美拉唑的氧化

[0062] 对野生型 CYP102A1 或其突变形式是否能够氧化奥美拉唑进行了确认。将 50pmol CYP102A1 和 100  $\mu$  M 底物放入 0.25ml 100mM 的磷酸钾缓冲液 (pH7.4) 中, 并进行典型的稳态反应。为了引发反应, 添加产生 NADPH 的体系 (终浓度: 10mM  $\beta$ -磷酸葡萄糖, 0.5mM NADP<sup>+</sup> 和每毫升 1IU 的酵母  $\beta$ -磷酸葡萄糖脱氢酶)。用 DMSO 制备 20mM 的奥美拉唑溶液, 并用酶反应液将其稀释以使有机溶剂的终浓度为 1% (v/v) 以下。为了测量人 CYP2C19 的活性, 使用 50pmol P450、100pmol NADPH-P450 还原酶 (CPR)、100pmol 细胞色素 b5 和 45  $\mu$  M L- $\alpha$ -二月桂酰基-sn-甘油-3-磷酸胆碱 (DLPC) 来代替 50pmol CYP102A1。在 37°C 下使反应液反应 30 分钟, 使用在冷态用两倍的冰制备的二氯甲烷来终止反应。

[0063] (1) HPLC 分析

[0064] 将反应混合物离心以除去上清液, 在氮气下蒸发掉其溶剂 (参见 Vickers

等, 1990), 并用 HPLC 分析所获得的混合物 (参见 Piver 等, 2004)。将样品 (30  $\mu$ l) 注入 Gemini C18 柱 (4.6mm $\times$ 150mm, 5  $\mu$ m, Phenomenex, Torrance, CA) 中。使用 30% 乙腈作为流动相。流动相以 1ml/ 分钟的速率流动, 并且用 302nm 的 UV 测量洗脱液。为了调查 CYP102A1 (P450BM3) 是否能够氧化奥美拉唑, 将底物的浓度固定至 100  $\mu$ M, 并且测量使用野生型 CYP102A1 及其突变形式时奥美拉唑的氧化程度。

[0065] 结果, 如图 3 的 HPLC 色谱所确认的, 可以确认所产生的代谢物的峰的保留时间与标准 5'-羟基奥美拉唑的峰的保留时间精确相同。

#### [0066] (2) LC-MS 分析和 NMR 分析

[0067] 为了鉴定由突变的 CYP102A1 产生的奥美拉唑代谢物, 通过比较奥美拉唑和代谢物的 LC 图谱与片段图来进行 LC-MS 分析。使突变的 CYP102A1 和人 CYP2C19 在 100  $\mu$ M 奥美拉唑和产生 NADPH 的体系的存在下于 37 $^{\circ}$ C 反应 30 分钟。添加两倍量的经冰冷冷却的  $\text{CH}_2\text{Cl}_2$  来终止反应。离心后, 移去并弃去上清液, 在氮气的存在下使有机溶剂层干燥。用 100  $\mu$ l 流动相以漩涡混合重建反应物, 超声处理 20 秒。将适量 (5  $\mu$ l) 的所制备的溶液注入 LC 柱。

[0068] 使用安装有 LC-MS 软件的 Shimadzu LCMS-2010EV 系统 (Shimadzu, Kyoto, Japan) 以电喷雾电离 (正) 模式进行 LC-MS 分析。在 Shim-pack VP-ODS 柱 (250mm $\times$ 2.0mm i. d.; Shimadzuco., Japan) 中, 使用 30% 乙腈作为流动相。以 0.1ml/ 分钟的流动速率分离流动相。为了确认代谢物, 以电喷雾电离 (正) 模式记录质谱。接口和检测器电压分别为 4.4kV 和 1.5kV。将雾化气体流速设为 1.5ml/ 分钟, 接口、曲形脱溶剂管 (CDL) 和热封闭温度分别为 250 $^{\circ}$ C、250 $^{\circ}$ C 和 200 $^{\circ}$ C。研究了由突变的 CYP102A1#10 和人 CYP2C19 产生的代谢物的总离子流 (TIC) 图谱。

[0069] 结果, 如图 6 所示, 反应样品的质谱显示了在 6.200 分钟的峰 (5'-羟基奥美拉唑) 和 15.267 分钟的峰 (奥美拉唑), 并且当计算由突变的 CYP102A1#10 产生的 5'-羟基产物和奥美拉唑的质谱中的  $[\text{M}+\text{H}]^+$  时, 观察到的值分别为 362 和 346。

[0070] 此外, 对反应混合物的 LC-MS 分析确认了突变的 CYP102A1 产生了 5'-羟基奥美拉唑。经确认, CYP102A1 代谢物的保留时间和片段图与人 CYP2C19 产生的真正代谢物的保留时间和片段图精确相同。

[0071] 利用 NMR 分析法对由细菌 CYP102A1 的突变 #10 产生的代谢物的结构进行分析, 所获得的结果是, 如图 8 所示, 可以确认所产生的产物不是 3'-羟基奥美拉唑而是 5'-羟基奥美拉唑。

#### [0072] (3) 确定转换数

[0073] 确定了野生型 CYP102A1 及其突变形式产生奥美拉唑氧化物的产生速率。使用 100  $\mu$ M 奥美拉唑, 添加产生 NADPH 的体系以引发反应, 在 37 $^{\circ}$ C 进行反应 30 分钟以确定转换数。如上所述通过 HPLC 确定奥美拉唑的产生速率。

[0074] 从图 5 的结果可以确认, 野生型 CYP102A1、其 17 种突变形式以及 6 种突变嵌合形式的转换数的变化范围很大。另外, 研究了突变的 CYP102A1 的总转换数 (TTN) (mol 产物 / mol 催化剂)。为了测量突变的 CYP102A1 的 TTN, 使用 0.1mM ~ 2mM 的奥美拉唑, 并以 30 分钟至 5 小时的间隔进行反应。通过 HPLC 确定奥美拉唑代谢物的产生速率。

[0075] 结果, 如图 7 所示, 可以确认: 当用 1mM 奥美拉唑进行 2 ~ 4 小时的反应时, 在 TTN 方面显示出高活性的 CYP102A1 突变 #10 具有最高的活性; 同时, 野生型 CYP102A1 酶对于

奥美拉唑几乎没有活性。尚未报道过通过化学合成来产生奥美拉唑代谢物。这意味着使用 CYP102A1 酶产生奥美拉唑代谢物是代谢物化学合成的替代途径。

[0076] 从以上结果可以确认,利用细菌 CYP102A1 酶来催化与人 CYP2C19 相同的反应,产生了作为人代谢物的 5'-OH 产物。可以确认,作为人 P450 底物的奥美拉唑的氧化受到野生型 CYP102A1 及其突变形式的催化,并且产生了作为主要代谢物的羟基产物,即 5'-OH 产物,并且,利用 HPLC 和 LC-MS 将该产物与人 CYP2C19 所产生的产物进行比较,确认了所产生的代谢物的产生。

[0077] 从上述结果可以确认,突变的 CYP102A1 能够有效地从奥美拉唑产生人体中的代谢物,其中所述代谢物可以在药物开发中用于评价药物的功效、毒性和药代动力学等,并且可以用于产生人体中的代谢物衍生物,其将是药物开发中的先导化合物。

[0001]

## 核苷酸和氨基酸序列表

FCII4KR2821(提交)

- <110> 尹喆湖  
 <120> 使用细菌细胞色素 P450 从奥美拉唑制备代谢物的新方法和用于其的组合物  
 <130> 2012-OPA-5610/PCT  
  
 <150> KR 10-2012-0009300  
 <151> 2012-01-31  
  
 <160> 17  
  
 <170> KopatentIn 2.0  
  
 <210> 1  
 <211> 31  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> BamHI 正向引物  
 <400> 1  
 agcggatcca tgacaattaa agaaatgect c 31  
  
 <210> 2  
 <211> 20  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> SacI 反向引物  
 <400> 2  
 atcgagctcg tagtttgtat 20  
  
 <210> 3  
 <211> 20  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> R47L 引物  
 <400> 3  
 gcgcctggtc tggtaacgcg 20  
  
 <210> 4

[0002]

<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	Y51F 引物	
<400>	4	
	gtaacgcgct tcttatcaag t	21
<210>	5	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	E64G 引物	
<400>	5	
	gcattgatg gctcacgett t	21
<210>	6	
<211>	25	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	A74G 引物	
<400>	6	
	taagtcaagg ccttaaattt gtacg	25
<210>	7	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	F81I 引物	
<400>	7	
	gtacgtgata ttgcaggaga c	21
<210>	8	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	L86I 引物	

[0003]



<400>	8		
	ggagacggga tttttacaag c		21
<210>	9		
<211>	21		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	F87A 引物		
<400>	9		
	gacgggtag cgacaagetg g		21
<210>	10		
<211>	21		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	F87V 引物		
<400>	10		
	gacgggtag tgacaagetg g		21
<210>	11		
<211>	21		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	E143G 引物		
<400>	11		
	gaagtaccgg gcgacatgac a		21
<210>	12		
<211>	22		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	L188Q 引物		
<400>	12		
	atgaacaagc agcagcgagc aa		22
<210>	13		

[0004]

<211> 21  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> A264G 引物  
 <400> 13  
 ttcttaattg ggggacacgt g 21

<210> 14  
 <211> 22  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> E267V 引物  
 <400> 14  
 tgcgggacac gtgacaaca gt 22

<210> 15  
 <211> 23  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> L86I/F87V 引物  
 <400> 15  
 ggagacggga ttgtgacaag ctg 23

<210> 16  
 <211> 1049  
 <212> PRT  
 <213> 未知  
 <220>  
 <223> CYP102A1  
 <400> 16  
 Met Thr Ile Lys Glu Met Pro Gln Pro Lys Thr Phe Gly Glu Leu Lys  
 1 5 10 15  
 Asn Leu Pro Leu Leu Asn Thr Asp Lys Pro Val Gln Ala Leu Met Lys  
 20 25 30  
 Ile Ala Asp Glu Leu Gly Glu Ile Phe Lys Phe Glu Ala Pro Gly Leu  
 35 40 45  
 Val Thr Arg Tyr Leu Ser Ser Gln Arg Leu Ile Lys Glu Ala Cys Asp  
 50 55 60

[0005]

Gly	Ser	Arg	Phe	Asp	Lys	Asn	Leu	Ser	Gln	Ala	Leu	Lys	Phe	Val	Arg	65	70	75	80
Asp	Ile	Ala	Gly	Asp	Gly	Leu	Val	Thr	Ser	Trp	Thr	His	Glu	Lys	Asn	85	90	95	
Trp	Lys	Lys	Ala	His	Asn	Ile	Leu	Leu	Pro	Ser	Phe	Ser	Gln	Gln	Ala	100	105	110	
Met	Lys	Gly	Tyr	His	Ala	Met	Met	Val	Asp	Ile	Ala	Val	Gln	Leu	Val	115	120	125	
Gln	Lys	Trp	Glu	Arg	Leu	Asn	Ala	Asp	Glu	His	Ile	Glu	Val	Pro	Gly	130	135	140	
Asp	Met	Thr	Arg	Leu	Thr	Leu	Asp	Thr	Ile	Gly	Leu	Cys	Gly	Phe	Asn	145	150	155	160
Tyr	Arg	Phe	Asn	Ser	Phe	Tyr	Arg	Asp	Gln	Pro	His	Pro	Phe	Ile	Thr	165	170	175	
Ser	Met	Val	Arg	Ala	Leu	Asp	Glu	Ala	Met	Asn	Lys	Gln	Gln	Arg	Ala	180	185	190	
Asn	Pro	Asp	Asp	Pro	Ala	Tyr	Asp	Glu	Asn	Lys	Arg	Gln	Phe	Gln	Glu	195	200	205	
Asp	Ile	Lys	Val	Met	Asn	Asp	Leu	Val	Asp	Lys	Ile	Ile	Ala	Asp	Arg	210	215	220	
Lys	Ala	Ser	Gly	Glu	Gln	Ser	Asp	Asp	Leu	Leu	Thr	His	Met	Leu	Asn	225	230	235	240
Gly	Lys	Asp	Pro	Glu	Thr	Gly	Glu	Pro	Leu	Asp	Asp	Glu	Asn	Ile	Arg	245	250	255	
Tyr	Gln	Ile	Ile	Thr	Phe	Leu	Ile	Ala	Gly	His	Val	Thr	Thr	Ser	Gly	260	265	270	
Leu	Leu	Ser	Phe	Ala	Leu	Tyr	Phe	Leu	Val	Lys	Asn	Pro	His	Val	Leu	275	280	285	
Gln	Lys	Ala	Ala	Glu	Glu	Ala	Ala	Arg	Val	Leu	Val	Asp	Pro	Val	Pro	290	295	300	
Ser	Tyr	Lys	Gln	Val	Lys	Gln	Leu	Lys	Tyr	Val	Gly	Met	Val	Leu	Asn	305	310	315	320
Glu	Ala	Leu	Arg	Leu	Trp	Pro	Thr	Ala	Pro	Ala	Phe	Ser	Leu	Tyr	Ala	325	330	335	
Lys	Glu	Asp	Thr	Val	Leu	Gly	Gly	Glu	Tyr	Pro	Leu	Glu	Lys	Gly	Asp	340	345	350	
Glu	Leu	Met	Val	Leu	Ile	Pro	Gln	Leu	His	Arg	Asp	Lys	Thr	Ile	Trp	355	360	365	
Gly	Asp	Asp	Val	Glu	Glu	Phe	Arg	Pro	Glu	Arg	Phe	Glu	Asn	Pro	Ser	370	375	380	
Ala	Ile	Pro	Gln	His	Ala	Phe	Lys	Pro	Phe	Gly	Asn	Gly	Gln	Arg	Ala	385	390	395	400
Cys	Ile	Gly	Gln	Gln	Phe	Ala	Leu	His	Glu	Ala	Thr	Leu	Val	Leu	Gly	405	410	415	

[0006]

Met Met Leu Lys His Phe Asp Phe Glu Asp His Thr Asn Tyr Glu Leu  
 420 425 430  
 Asp Ile Lys Glu Thr Leu Thr Leu Lys Pro Glu Gly Phe Val Val Lys  
 435 440 445  
 Ala Lys Ser Lys Lys Ile Pro Leu Gly Gly Ile Pro Ser Pro Ser Thr  
 450 455 460  
 Glu Gln Ser Ala Lys Lys Val Arg Lys Lys Val Glu Asn Ala His Asn  
 465 470 475 480  
 Thr Pro Leu Leu Val Leu Tyr Gly Ser Asn Met Gly Thr Ala Glu Gly  
 485 490 495  
 Thr Ala Arg Asp Leu Ala Asp Ile Ala Met Ser Lys Gly Phe Ala Pro  
 500 505 510  
 Gln Val Ala Thr Leu Asp Ser His Ala Gly Asn Leu Pro Arg Glu Gly  
 515 520 525  
 Ala Val Leu Ile Val Thr Ala Ser Tyr Asn Gly His Pro Pro Asp Asn  
 530 535 540  
 Ala Lys Gln Phe Val Asp Trp Leu Asp Gln Ala Ser Ala Asp Asp Val  
 545 550 555 560  
 Lys Gly Val Arg Tyr Ser Val Phe Gly Cys Gly Asp Lys Asn Trp Ala  
 565 570 575  
 Thr Thr Tyr Gln Lys Val Pro Ala Phe Ile Asp Glu Thr Leu Ala Ala  
 580 585 590  
 Lys Gly Ala Glu Asn Ile Ala Asp Arg Gly Glu Ala Asp Ala Ser Asp  
 595 600 605  
 Asp Phe Glu Gly Thr Tyr Glu Glu Trp Arg Glu His Met Trp Ser Asp  
 610 615 620  
 Val Ala Ala Tyr Phe Asn Leu Asp Ile Glu Asn Ser Glu Asp Asn Lys  
 625 630 635 640  
 Ser Thr Leu Ser Leu Gln Phe Val Asp Ser Ala Ala Asp Met Pro Leu  
 645 650 655  
 Ala Lys Met His Gly Ala Phe Ser Ala Asn Val Val Ala Ser Lys Glu  
 660 665 670  
 Leu Gln Gln Leu Gly Ser Glu Arg Ser Thr Arg His Leu Glu Ile Ala  
 675 680 685  
 Leu Pro Lys Glu Ala Ser Tyr Gln Glu Gly Asp His Leu Gly Val Ile  
 690 695 700  
 Pro Arg Asn Tyr Glu Gly Ile Val Asn Arg Val Thr Ala Arg Phe Gly  
 705 710 715 720  
 Leu Asp Ala Ser Gln Gln Ile Arg Leu Glu Ala Glu Glu Glu Lys Leu  
 725 730 735  
 Ala His Leu Pro Leu Gly Lys Thr Val Ser Val Glu Glu Leu Leu Gln  
 740 745 750  
 Tyr Val Glu Leu Gln Asp Pro Val Thr Arg Thr Gln Leu Arg Ala Met  
 755 760 765

[0007]

Ala Ala Lys Thr Val Cys Pro Pro His Lys Val Glu Leu Glu Ala Leu  
770 775 780  
Leu Glu Lys Gln Ala Tyr Lys Glu Gln Val Leu Ala Lys Arg Leu Thr  
785 790 795 800  
Met Leu Glu Leu Leu Glu Lys Tyr Pro Ala Cys Glu Met Glu Phe Ser  
805 810 815  
Glu Phe Ile Ala Leu Leu Pro Ser Ile Ser Pro Arg Tyr Tyr Ser Ile  
820 825 830  
Ser Ser Ser Pro His Val Asp Glu Lys Gln Ala Ser Ile Thr Val Ser  
835 840 845  
Val Val Ser Gly Glu Ala Trp Ser Gly Tyr Gly Glu Tyr Lys Gly Ile  
850 855 860  
Ala Ser Asn Tyr Leu Ala Asn Leu Gln Glu Gly Asp Thr Ile Thr Cys  
865 870 875 880  
Phe Val Ser Thr Pro Gln Ser Gly Phe Thr Leu Pro Lys Asp Ser Glu  
885 890 895  
Thr Pro Leu Ile Met Val Gly Pro Gly Thr Gly Val Ala Pro Phe Arg  
900 905 910  
Gly Phe Val Gln Ala Arg Lys Gln Leu Lys Glu Gln Gly Gln Ser Leu  
915 920 925  
Gly Glu Ala His Leu Tyr Phe Gly Cys Arg Ser Pro His Glu Asp Tyr  
930 935 940  
Leu Tyr Gln Glu Glu Leu Glu Asn Ala Gln Asn Glu Gly Ile Ile Thr  
945 950 955 960  
Leu His Thr Ala Phe Ser Arg Val Pro Asn Gln Pro Lys Thr Tyr Val  
965 970 975  
Gln His Val Met Glu Arg Asp Gly Lys Lys Leu Ile Glu Leu Leu Asp  
980 985 990  
Gln Gly Ala His Phe Tyr Ile Cys Gly Asp Gly Ser Gln Met Ala Pro  
995 1000 1005  
Asp Val Glu Ala Thr Leu Met Lys Ser Tyr Ala Asp Val Tyr Glu Val  
1010 1015 1020  
Ser Glu Ala Asp Ala Arg Leu Trp Leu Gln Gln Leu Glu Glu Lys Gly  
1025 1030 1035 1040  
Arg Tyr Ala Lys Asp Val Trp Ala Gly  
1045

<210> 17  
<211> 3150  
<212> DNA  
<213> 未知  
<220>  
<223> CYP102A1

[0008]

<400>	17					
atgacaallta	aagaaatgcc	tcagccaaaa	acgtttggag	agctlaaaaa	tttaccglla	60
ttaaacacag	ataaacccgtt	tcaagctttg	atgaaaattg	eggatgaatt	aggagaaatc	120
tttaaattcg	aggcgcctgg	tcgtgtaacg	cgctacttat	caagtcagcg	tctaattaaa	180
gaagcatgcg	atgaaatcag	ctttgataaa	aacttaagtc	aagcgcitaa	atftgtacgt	240
gattttgcag	gagacgggtt	atftacaagc	tggacgcatt	aaaaaaattg	gaaaaaagcg	300
cataatafct	tacttccaag	cttcagtcag	caggcaatga	aaggetatca	tgcgatgatg	360
gtcgatatcg	ccgtgcaget	tgttcaaaaag	tgggagcgtc	taaattgcaga	tgagcatatt	420
gaagtaccgg	aagacatgac	acgtttaaac	cttgatacaa	ttggtctttg	cggctttaac	480
tategcittta	acagctttta	ccgagatcag	cctcatccat	ttattacaag	tatggtccgt	540
gcaactggatg	aagcaatgaa	caagctgcag	cgagcaaatc	cagacgacce	agcttatgat	600
gaaaacaagc	gccagtttca	agaagatatc	aaggtgatga	acgacctagt	agataaaatt	660
aitgcagatc	gcaaagcaag	cggatgaaca	agcgatgatt	tattaacgca	tatgctaaac	720
ggaaaagatc	cagaaacggg	tgagccgctt	gatgacgaga	acattcgcta	tcaaattatt	780
acattcttaa	ttgcgggaca	cgaacaaca	agtgtctttt	tatcatttgc	gctgtatttc	840
ttagtgaaaa	atccacatgt	attacaaaa	gcagcagaag	aagcagcacg	agttctagta	900
gactctgttc	caagctacaa	acaagtcaaa	cagcttaaat	atgtoggeat	ggtcttaaac	960
gaagcgcctc	gcttatggcc	aactgctcct	gcgttttccc	tatatgcaaa	agaagatacg	1020
gtgcttggag	gagaatatec	tttagaaaa	ggcgacgaac	taatggttct	gattcctcag	1080
cttcaccgtg	ataaaacaat	ttggggagac	gatgttggag	agttccgtec	agagcgtttt	1140
gaaaatccaa	gtgcgattcc	gcagcatgcg	tttaaacctg	ttggaacggg	tcagcgtgcg	1200
tgtatcggtc	agcaagttcg	tcttcatgaa	gcaacgcctg	tacttgggat	gatgetaaaa	1260
cactttgaat	ttgaagatca	tacaaactac	gagctcgata	ttaaagaaac	tttaacglla	1320
aaacctgaag	gctttgtggt	aaaagcaaaa	tcgaaaaaaa	ttccgcttgg	cggtatctct	1380
tcacctagca	ctgaacagtc	tgcataaaaa	gtacgcaaaa	aggcagaaaa	cgtctcataat	1440
acgccgctgc	ttgigtata	cgtttcaaat	atgggaacag	ctgaaggaac	ggcgcgtgat	1500
ttagcagata	ttgcaatgag	caaaggattt	gcaccgcagg	tcgcaacgct	tgattccaac	1560
gcgggaaatc	ttccgcgcga	aggagctgta	tttaattgtaa	cggcgtctta	taacggtcat	1620
ccgcctgata	acgcaaagea	atftgtcgac	tgtttagacc	aagcgtctgc	tgatgaagta	1680
aaaggcgttc	gctactccgt	atftggatgc	ggcgataaaa	actgggctac	tacgtatcaa	1740
aaagtgcctg	cttttatcga	lgaaacgctt	gccgctaaag	gggcagaaaa	calcctgac	1800
cgcggtgaag	cagatgcaag	cgacgacttt	gaaggccat	atgaagaatg	gcgtgaacat	1860
atgtggatg	acgtagcagc	ctactttaac	ctcgacattg	aaaacagtga	agataataaa	1920
tctactcttt	cacttcaatt	tgtcgacagc	gccgcggata	tgccgcttgc	gaaaatgcac	1980
ggtgcgtttt	caagcaagct	cgtagcaagc	aaagaacttc	aacagccagg	cagtgcacga	2040
agcacgcgac	atcttgaaat	tgaacttcca	aaagaagctt	cttatcaaga	aggagatcat	2100
ttagggtgta	ttctcgcgaa	ctatgaagga	atagtaaac	gtgtaacagc	aaggttcgge	2160
ctagatgcat	cacagcaaat	ccgtctggaa	gcagaagaag	aaaaattagc	tcatttgcga	2220
ctcgtcaaaa	cagtatccgt	agaagagctt	ctgcaatacg	tggagcttca	agatectggt	2280
acgcgcacgc	agcttcgcgc	aatggtgctt	aaaacggtct	gcccgccgca	taaagttagag	2340
cttgaagcct	tgcctgaaaa	gcaagcctac	aaagaacaag	tgctggcaaa	acgtttaaca	2400
atgcttgaac	tgcctgaaaa	atacccggcg	tgtgaaatga	aattcagcga	atftatcgcc	2460
cttctgccc	gcatacgecc	gcctatttac	tcgatttctt	catcacctcg	tgtcgatgaa	2520
aaacaagcaa	gcatcacggt	cagcgttgct	tcaggagaag	cgtggagcgg	atatggagaa	2580
tataaaggaa	ttgcgtcgaa	ctatcttggc	gagctgcaag	aaggagatac	gattacgtgc	2640
tttatttcca	cacegcagtc	agaatttacg	ctgccaaaa	acctgaaac	gccgcttate	2700
atggtcggac	cgggaacagg	cgctcgcgcc	tttagagctt	ttgtgcaggc	gcgcaaacag	2760
ctaaaagaac	aaggacagtc	acttggagaa	gcacatttat	acttcggctg	ccgttcaact	2820
catgaagact	atctgtatec	agaagagctt	gaaaacgccc	aaagcgaagg	catcatttacg	2880
cttcaataccg	cttttctctg	catgcaaat	cagccgaaaa	catacgttca	gcacgtaatg	2940

[0009]

---

gaacaagacg gcaagaaatt gattgaactt cttgatcaag gagegcactt ctatatttgc	3000
ggagacggaa gccaaatggc acctgccgtt gaagcaacgc ttatgaaaag ctatctgac	3060
gttcaccaag tgagtgaagc agacgctcgc ttatggctgc agcagctaga agaaaaagge	3120
cgatacgcaa aagacgtgtg ggctgggtaa	3150

1	MTIKEMPQPKTFGELKNLPLLNTRDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIK
61	EACDESRFDKNSQALKFVRDFAGDGLFTSWTHEKNWKKAHNILLPSFSQQAMKGYHAMM
121	VDIAVQLVQKWERLNADEHIEVPEDMTRLTLDITGLCGFNRFNSFYRDQPHPFITSMVR
181	ALDEAMNKLQRANFDDPAYDENKRQFQEDIKVMNDLVDKIADRKASGEQSDDLTHMLN
241	GKDPETGEPLDDENIRYQIITFLIAGHETTSGLLSFALYFLVKNPHVLQKAAEEAARVLV
301	DPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLEKGDLMVLI PQ
361	LHRDKTIWGDDEVEFRPERFENPSAIPQHAFKPFNGQRACIGQQFALHEATLVLGMLK
421	HFDFEDHTNYELDIKETLTLKPEGFVVKAKSKKIPGGIPSPSTEQSAKKVRKKAENAHN
481	TPLLVLGYSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGH
541	PPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIAD
601	RGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMH
661	GAFSTNVVASKELQQPGSARSTRHLEIELPKESYQEGDHLGVI PRNYEGIVNRVTAREG
721	LDASQQIRLEAEEERLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAMAAKTVCPPHKVE
781	LEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKFSFIALLPSIRPRYYISSSPRVDE
841	KQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDITTCFISTPQSEFTLPRDPETPLI
901	MVPGTGVAPFRGFVQARKQLKEQGSLGEAHLVFGCRSPHEDYLYQELENAQSEGITIT
961	LHTAPSRMPNQPKTYVQHVMEDGKKLIELLDQGAHFYICGDGSMAPAVEATLMKSYAD
1021	VHQVSEADARLWLQQLLEEKGRYAKDVWAG-

图 1



5' -ATGACAATTAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGTTATTA  
AACACAGATAAAACCGGTTCAAGCTTTGATGAAAATTTGCGGATGAATTAGGAGAAATCTTTAAA  
TTCGAGGCGCCTGGTTCGTGTAACCGCCTACTTATCAAGTCAGCGTCTAATTAAGAAGCATGC  
GATGAATCACGCTTTTGATAAAAACTTAAGTCAAGCGCTTAAATTTGTACGTGATTTTGCAGGA  
GACGGGTTATTTACAAGCTGGACGCATGAAAAAAATTTGGAAAAAGCGCATAATATCTTACTT  
CCAAGCTTCAGTCAGCAGGCAATGAAAAGGCTATCATGCGATGATGGTTCGATATCGCCGTCAG  
CTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTACCGGAAGACATGACA  
CGTTTAAACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTTTTACCGA  
GATCAGCCTCATCCATTTATTACAAGTATGGTCCGTGCACCTGGATGAAGCAATGAACAAGCTG  
CAGCGAGCAAAATCCAGACGACCAGCTTATGATGAAAAACAAGCGCCAGTTTCAAGAAGATATC  
AAGGTGATGAACGACCTAGTAGATAAAAATTATTGCAGATCGCAAAGCAAGCGGTGAACAAAGC  
GATGATTTATTAACGCATATGCTAAACGGAAAAAGATCCAGAAACGGGTGAGCCGCTTGATGAC  
GAGAACATTCGCTATCAAATTAATTACATTCTTAATTGCGGGACACGAAACAACAAGTGGTCTT  
TTATCATTTGCGCTGTATTTCTTAGTGAAAAATCCACATGTATTACAAAAAGCAGCAGAAGAA  
GCAGCACGAGTTCTAGTAGATCCTGTTCCAAGCTACAAAACAAGTCAAACAGCTTAAATATGTC  
GGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAACCTGCTCCTGCGTTTTCCCTATATGCA  
AAAGAAGATACGGTGGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAACTAATGGTTCTG  
ATTCTCAGCTTCACCGTGATAAAACAATTTGGGGAGACGATGTGGAAGAGTTCCGTCACAG  
CGTTTTGAAAATCCAAGTGGGATTCGGCAGCATGCGTTTTAAACCGTTTGGAAACGGTCAGCGT  
GCGTGTATCGGTGAGCAGTTGCTCTTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAA  
CACTTTGACTTTGAAGATCATACAAACCTACGAGCTCGATATTAAGAAACTTTAACGTTAAAA  
CCTGAAGGCTTTGTGGTAAAAAGCAAAATCGAAAAAAATTCGCTTGGCGGTATTCCTTCACCT  
AGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCGAGAAAACGCTCATAATACGCCGCTG  
CTTGTGCTATACGGTTCAAATATGGGAACAGCTGAAGGAACGGCGCGTGAATTTAGCAGATATT  
GCAATGAGCAAAGGATTTGCACCCGAGGTCGCAACGCTTGATTCACACGCCGGAATCTTCCG  
CGCGAAGGAGCTGTATTAATTTGTAACGGCGTCTTATAACGGTCATCCGCTGATAAAGCAAG  
CAATTTGTGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTGCGTACTCCGTA  
TTGGATCGGGGATAAAAACTGGGCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAA  
ACGCTTTGCGCTAAAAGGGGAGAAAAACATGCTGACCGCGGTGAAGCAGATGCAAGCGACGAC  
TTTGAAGGCACATATGAAGAATGGCGTGAACATATGTGGAGTGAAGTAGCAGCCTACTTTAAC  
CTCGACATTGAAAACAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTGACAGCGCC  
GCGGATATGCCGCTTGGCAAAAATGCACGGTGGTTTTTCAACGAACGTCGTAGCAAGCAAAGAA  
CTTCAACAGCCAGGCAGTGCACGAAGCAGCGGACATCTTGAAATTGAACCTTCAAAAAGAAGCT  
TCTTATCAAGAAGGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGT  
GTAACAGCAAGGTTCCGGCTAGATGCATCAGCAAAATCCGTCCTGGAAGCAGAAAGAAAGAAA  
TTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTGTGCAATACGTGGAGCTT  
CAAGATCCTGTTACGCGCACGCAGCTTCCGCGCAATGGCTGCTAAAACGGTCTGCCCGCCGAT  
AAAGTAGAGCTTGAAGCCTTGGTTGAAAAGCAAGCCTACAAAAGAACAAGTGTGGCAAAACGT  
TTAACAATGCTTGAACCTGCTTGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTATC  
GCCCTTCTGCCAAGCATAACGCCCGCGCTATTACTCGATTTCTTCATCACCTCGTGTGATGAA  
AAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGAAGCGTGGAGCGGATATGGAGAATAT  
AAAGGAATTCGCTGCAACTATCTTCCGAGCTGCAAGAAGGAGATACGATTACGTGCTTTATT  
TCCACACCCGAGTCAGAAATTTACGCTGCCAAAAGACCCCTGAAAACGCCGCTTATCATGGTCGGA  
CCGGGAACAGGCGTCCGCGCGTTTAGAGGCTTTGTGCAGGCGCGCAACAGCTAAAAGAACA  
GGACAGTCACTTGGAGAAGCACATTTATACTTCCGCTGCGGTTACCTCATGAAGACTATCTG  
TATCAAGAAGAGCTTGAACCGCCAAAAGCGAAGGCATCATTACGCTTCATACCGCTTTTCT  
CGCATGCCAAATCAGCCGAAAACATACGTTACGCACGTAATGGAACAAGACGGCAAGAAATTG  
ATTGAACCTTCTGATCAAGGAGCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCT  
GCCGTTGAAGCAACGCTTATGAAAAGCTATGCTGACGTTACCAAGTGAGTGAAGCAGACGCT  
CGCTTATGGCTGCAGCAGCTAGAAGAAAAAGCGGATACGCAAAAAGACGTTGTTGGGCTGGGTAA-3'

图 2

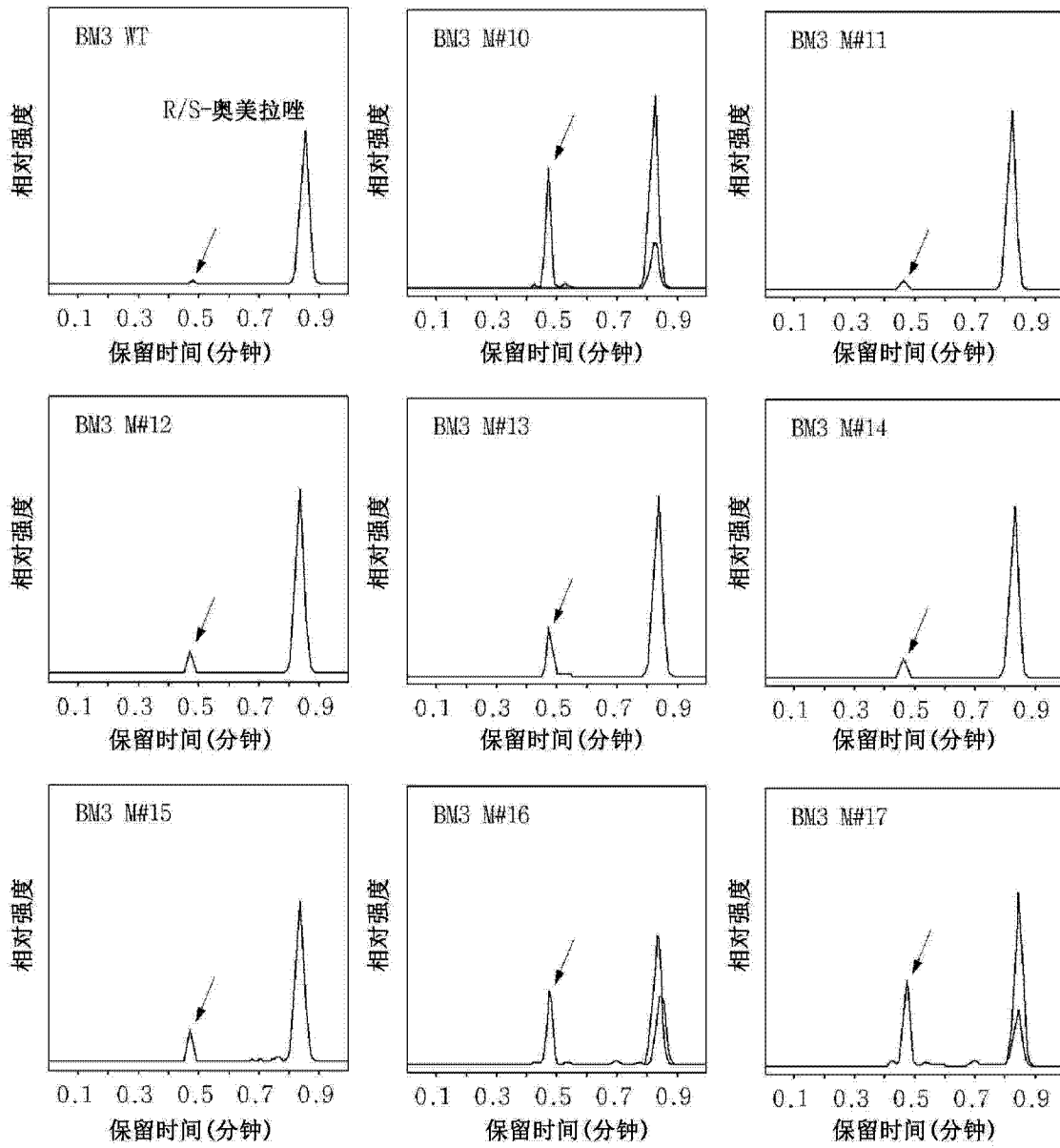


图 3

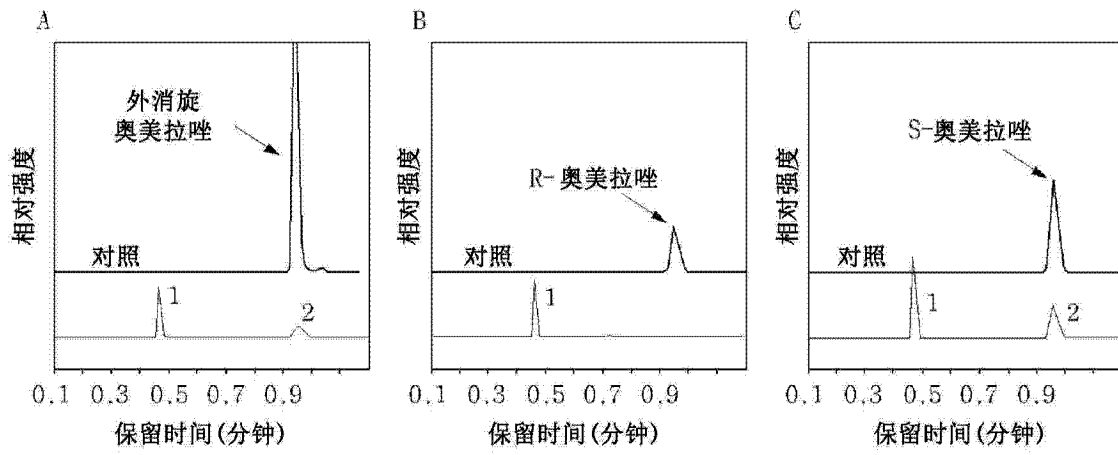


图 4

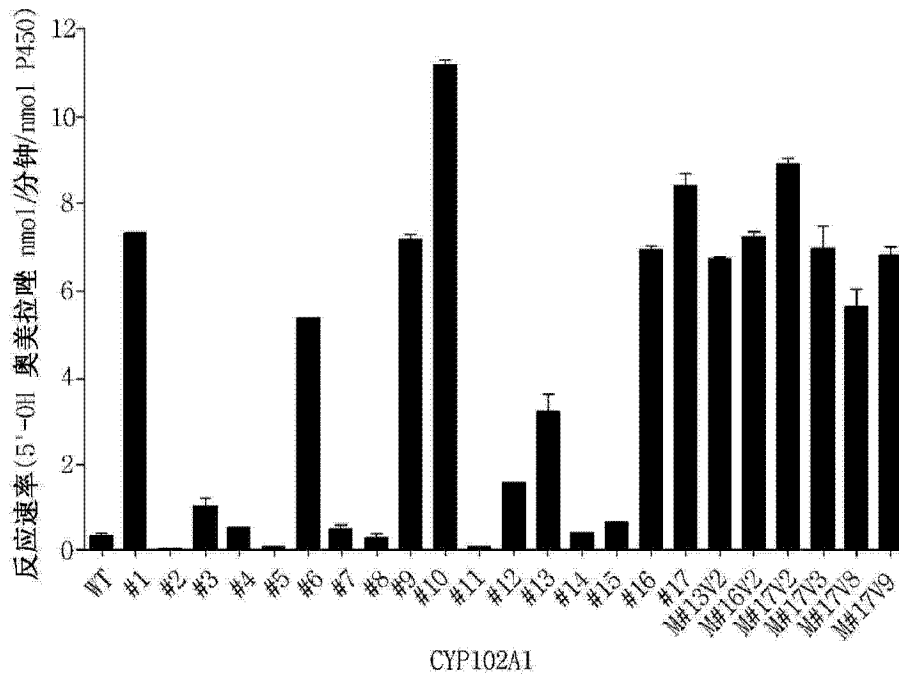


图 5

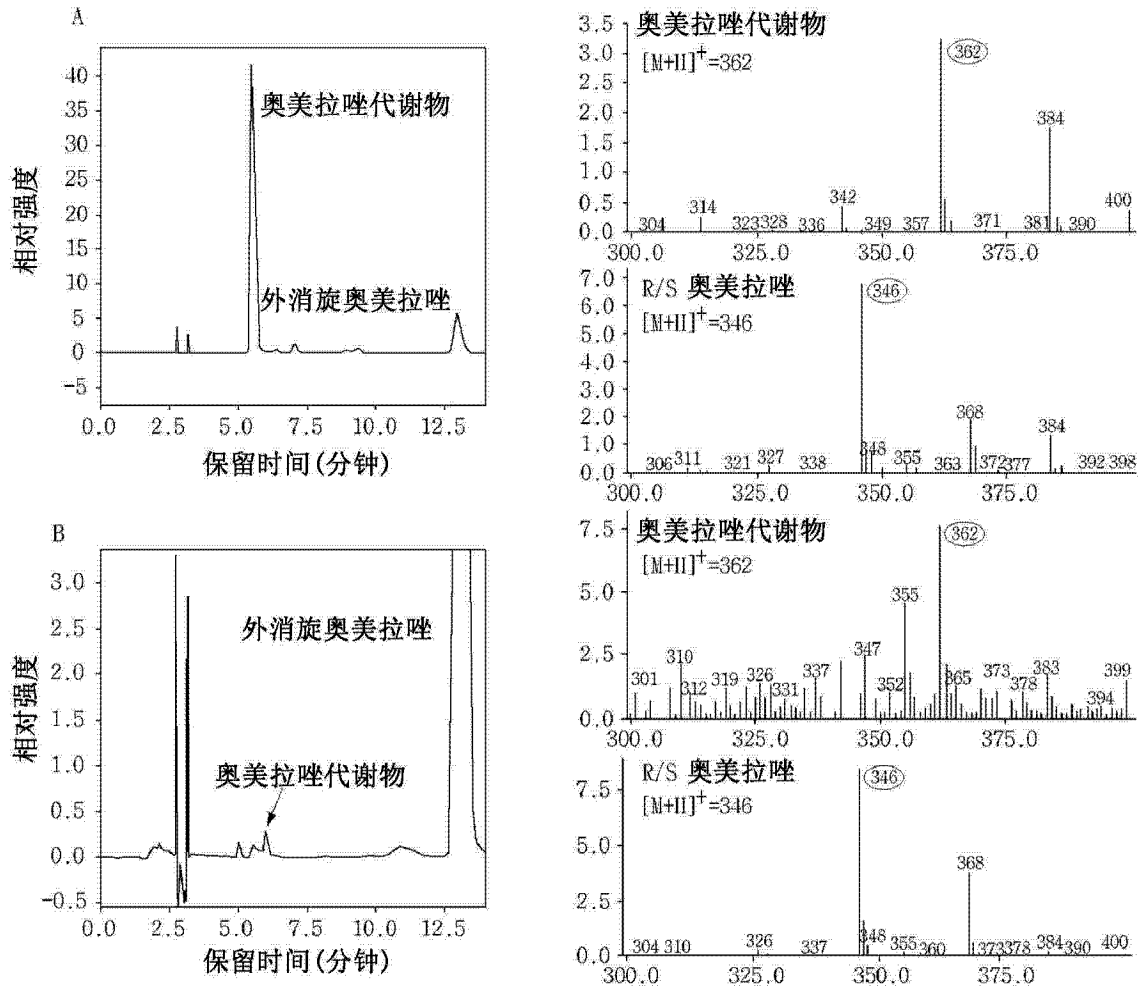


图 6

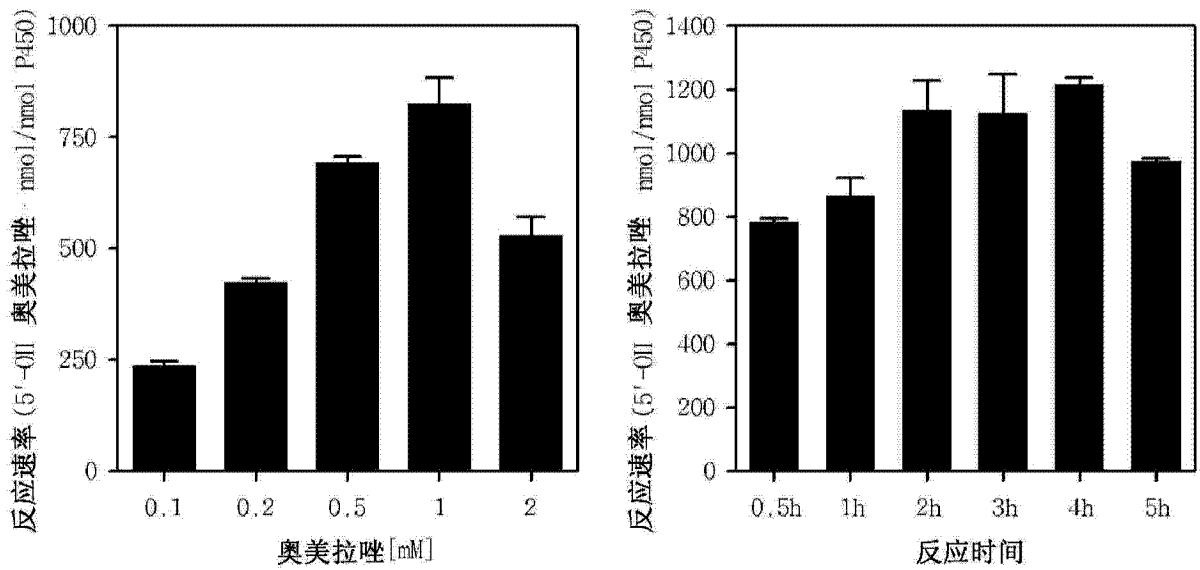


图 7

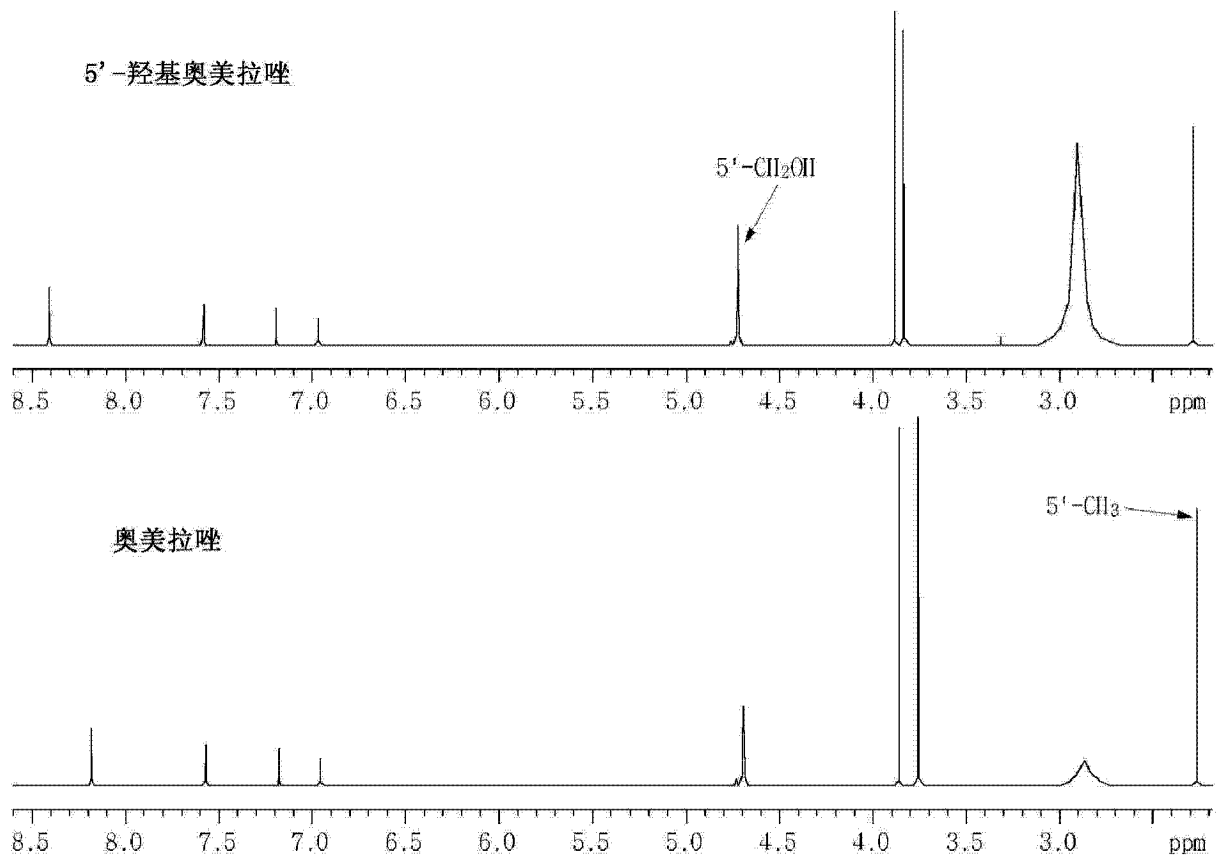


图 8