



(12) 发明专利

(10) 授权公告号 CN 102245771 B

(45) 授权公告日 2013. 04. 10

(21) 申请号 200980150330. 6

W0 2007/046825 A2, 2007. 04. 26, 全文.

(22) 申请日 2009. 12. 14

Regis CEBE et al. Size of the ligand complex between the N-terminal domain of the gene III coat protein and the non-infectious phage strongly influences the usefulness of in vitro selective infective phage technology. 《Biochem. J. 》. 2000, 第 352 卷 841-849.

(30) 优先权数据

61/122, 910 2008. 12. 16 US

(85) PCT 申请进入国家阶段日

2011. 06. 15

(86) PCT 申请的申请数据

PCT/EP2009/067066 2009. 12. 14

审查员 孙婷婷

(87) PCT 申请的公布数据

W02010/069913 EN 2010. 06. 24

(73) 专利权人 诺瓦提斯公司

地址 瑞士巴塞尔

(72) 发明人 A·洛

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 黄革生 魏永良

(51) Int. Cl.

C12N 15/10 (2006. 01)

(56) 对比文件

W0 99/36569 A1, 1999. 07. 22, 全文.

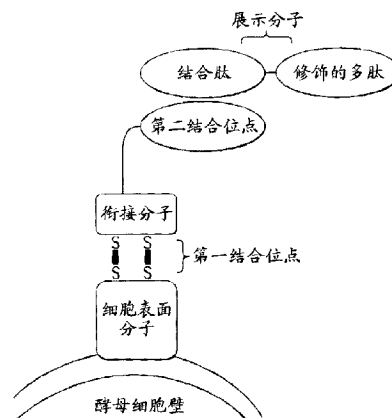
权利要求书 2 页 说明书 19 页
序列表 14 页 附图 12 页

(54) 发明名称

酵母展示系统

(57) 摘要

本发明涉及蛋白质展示文库和文库筛选领域。在优选的实施方案中, 本发明提供了包含细胞表面分子、衔接分子和展示分子的用于展示的三个组分系统。



1. 酵母宿主细胞文库, 其中每个宿主细胞包含
 - (a) 连接到细胞表面的细胞表面分子,
 - (b) 包含第一结合位点和第二结合位点的衔接分子, 和
 - (c) 包含经修饰多肽的展示分子;

其中, 所述第一结合位点特异地与细胞表面分子结合并且不能与展示分子结合, 并且所述第二结合位点特异地与展示分子结合并且不能与细胞表面分子结合, 其中所述衔接分子不是所述经修饰多肽的组分, 并且其中所述每一宿主细胞包含不同的经修饰多肽。

2. 权利要求 1 所述的酵母宿主细胞文库, 其包含多个展示分子。

3. 权利要求 1 或权利要求 2 所述的酵母宿主细胞文库, 其中所述宿主细胞表面分子与所述第一结合位点共价连接。

4. 权利要求 1 所述的酵母宿主细胞文库, 其中所述宿主细胞表面分子包含为 Aga1p 的第一凝集素, 并且其中所述第一结合位点包含为 Aga2p 的第二凝集素。

5. 权利要求 1 所述的酵母宿主细胞文库, 其中所述宿主细胞表面分子通过磷脂酰肌醇聚糖锚与细胞膜连接。

6. 权利要求 1 所述的酵母宿主细胞文库, 其中所述第二结合位点与展示分子共价连接。

7. 权利要求 1 所述的酵母宿主细胞文库, 其中所述第二结合位点包含 InaD 的 PDZ 结构域。

8. 权利要求 7 的酵母宿主细胞文库, 其中所述展示分子包含羧基端 NorpA 配体。

9. 权利要求 1 所述的酵母宿主细胞文库, 其中所述展示分子包含 InaD 的 PDZ 结构域。

10. 权利要求 9 的酵母宿主细胞文库, 其中所述第二结合位点包含羧基端 NorpA 配体。

11. 权利要求 1 所述的酵母宿主细胞文库, 其中所述经修饰多肽选自支架蛋白、信号转导蛋白质、抗体、免疫球蛋白、免疫粘附素、受体、配体、癌蛋白质、转录因子、酶和纤连蛋白多肽。

12. 权利要求 11 所述的酵母宿主细胞文库, 其中所述纤连蛋白多肽包含 F10 多肽。

13. 权利要求 1 所述的酵母宿主细胞文库, 其中所述展示分子包含分泌信号肽。

14. 权利要求 13 所述的酵母宿主细胞文库, 其中所述分泌信号肽包含 MF α /HSA 杂合前导肽或 MF α 前导肽。

15. 权利要求 1 所述的酵母宿主细胞文库, 其中所述展示分子的表达在第一诱导型启动子的控制下, 所述第一诱导型启动子选自 AOX1 启动子、Cup 1 启动子和 Gal 1 启动子。

16. 权利要求 1 所述的酵母宿主细胞文库, 其中所述衔接分子的表达在第二诱导型启动子的控制下, 所述第二诱导型启动子选自 AOX1 启动子、Cup 1 启动子和 Gal 1 启动子。

17. 用于展示经修饰多肽的方法, 包括:

(a) 提供包含与细胞的表面连接的细胞表面分子和编码包含经修饰多肽的展示多肽的第一核酸的酵母宿主细胞,

(b) 在其中第一结合位点与细胞表面分子结合的条件下, 将酵母宿主细胞与包含第一结合位点和第二结合位点的衔接分子接触, 和

(c) 在其中酵母宿主细胞将展示多肽输出到酵母宿主细胞外的条件下, 在其中第二结合位点与展示多肽结合的条件下, 温育酵母宿主细胞,

其中所述第一结合位点与细胞表面分子特异结合并且不能与展示分子结合,并且所述第二结合位点与展示多肽特异结合并且不能与细胞表面分子结合,并且其中所述衔接分子不是所述经修饰多肽的组分。

18. 产生酵母宿主细胞展示文库的方法,包括:

将各自编码包含经修饰多肽的展示多肽的第一核酸的展示文库导入到多个酵母宿主细胞中,其中至少两种导入的第一核酸编码不同的经修饰多肽,其中每一酵母宿主细胞包含编码细胞表面多肽的第二核酸和编码包含第一结合位点和第二结合位点的衔接分子的第三核酸,其中第一结合位点与细胞表面分子结合但是不与展示多肽结合,并且第二结合位点与展示多肽结合但是不与细胞表面分子结合,并且其中衔接分子不是经修饰多肽的组分。

酵母展示系统

[0001] 领域

[0002] 本发明涉及蛋白质展示文库和文库筛选领域。更具体地,本发明涉及产生用于在细胞表面上展示的蛋白质。

[0003] 背景

[0004] 可以从序列数据预测蛋白质结合结构域,然而重新设计具有改进的或者改变的结合亲和力的蛋白质常常需要测试许多重新设计的蛋白质的变体。目前,用于获得具有需要的结合亲和力的蛋白质的最好方法是产生并筛选包括此类变体的蛋白质文库,所述变体包括合理地重新设计的蛋白质、随机改变的蛋白质或者它们的组合。已经成功地构建了许多类型的蛋白质,例如免疫球蛋白和支架蛋白和受体或者受体配体的文库,并用于结合亲和力的筛选。

[0005] 筛选文库的方法很多,但最常用的方法之一是噬菌体展示方法,该方法包括将蛋白质文库与丝状噬菌体的外壳蛋白融合(例如,Huse 等人, ' 89 ;Clackson 等人, ' 91 ; Marks 等人, ' 92)。最常见的是与称为基因 III 蛋白质(pIII)的次要外壳蛋白融合,所述次要外壳蛋白在噬菌体的顶端存在 3 至 5 个拷贝。然后将融合的文库展示于病毒颗粒的表面上。然后可以筛选针对固定的靶蛋白质的噬菌体展示文库。然而,该方法的一个主要的缺点是不能总是鉴定出与文库非常高亲和力结合的靶蛋白质,因为洗脱结合的噬菌体所需的条件通常会使噬菌体颗粒变性,以致于不能鉴定目标蛋白质。噬菌体展示文库的另一个缺点是需要将靶蛋白质固定于固体表面上,这对确定靶蛋白质与噬菌体展示蛋白质的实际亲和力造成了困难。此外,一些目标蛋白质需要翻译后修饰,例如糖基化、甲基化或者二硫键,而这在噬菌体颗粒中表达时不能实现。

[0006] 用于筛选蛋白质文库的备选方法是在细菌细胞的表面上展示文库。该方法解决了与噬菌体展示相关的许多缺点,但也具有其固有问题。细菌展示的一个问题是,细菌荚膜可能对展示在细菌表面上的蛋白质造成位阻。此外,细菌不含有正确折叠真核蛋白质的机器,因此并不是总可以在细菌内表达目标蛋白质。与在噬菌体中的问题类似,细菌也不能为真核蛋白质提供翻译后修饰,如二硫键。

[0007] Wittrup 等人(美国专利号 6,699,658 和 6,696,251)已经研发了用于酵母细胞展示文库的方法。其是一种双组分系统,其中第一个组分涉及表达酵母接合吸附蛋白的一个亚基凝集素,其与酵母细胞壁锚定。第二个组分涉及表达与凝集素蛋白质的第二个亚基融合的蛋白质文库,其与第一个凝集素亚基形成高亲和性二硫键。然后将与凝集素融合的蛋白质文库展示于细胞的表面上。然后可以筛选文库。该方法允许真核蛋白质的正确折叠和翻译后修饰。

[0008] Rakestraw 等人(PCT/US2008/003978)已经研发了用于在酵母细胞的表面上展示蛋白质文库的三组分系统。第一个组分涉及表达与生物素结合肽融合的蛋白质文库,第二个组分涉及修饰酵母细胞壁以表达生物素,和第三个组分涉及将抗生物素蛋白与表达在细胞表面上的生物素结合。然后将融合蛋白质文库生物素化并从酵母细胞分泌,并且与酵母细胞表面上的抗生物素蛋白结合,从而,在酵母细胞的表面上展示蛋白质文库。该系统的一

个可能的缺点是,抗生物素蛋白非特异性结合所有生物素。另一个可能的缺点是抗生物素蛋白含有 4 个结合位点,其可能引起空间位阻,因此阻碍生物素化的蛋白质文库与细胞表面结合的抗生物素蛋白结合。类似地,抗生物素蛋白分子可能在结合生物素化的酵母细胞壁之前结合生物素化蛋白质文库,因此阻碍了抗生物素蛋白与酵母细胞壁的结合。此外,该方法还包含不得不在酵母细胞内生物素化蛋白质文库的额外复杂性。该必需的额外步骤需要进一步对酵母细胞进行修饰。已经公认,对生物系统进行了越多的修饰,该系统就越不可能按照所设计来进行运转。此外,由于抗生物素蛋白/链霉抗生物素蛋白是多价的,所以必需注意不要交联生物素化细胞。最后,生物素/链霉抗生物素蛋白和生物素/抗生物素蛋白已经用于许多市售的标记试剂盒。此类试剂盒将很难在该生物素/抗生物素蛋白展示系统中使用。

[0009] 现有技术缺少简单、有效的能特异结合分泌蛋白质文库的系统,所述系统使用衔接分子通过不同的结合部分结合蛋白质文库和真核细胞的表面。

[0010] 发明概述

[0011] 本发明通过提供本说明书所公开的方法和组合物满足了该需求。一个方面包括具有与细胞表面连接的细胞表面分子的宿主细胞、包含第一结合位点和第二结合位点的衔接分子和包含经修饰多肽的展示分子,其中第一结合位点特异地与细胞表面分子结合并且不与展示分子结合,第二结合位点特异地与展示分子结合且不与细胞表面分子结合,并且衔接分子不是经修饰多肽的组分。在某些实施方案中,宿主细胞具有多个展示分子。在其他可以与前述实施方案组合的实施方案中,宿主细胞表面分子可以与第一结合位点共价地连接。在其他可以与前述实施方案组合的实施方案中,宿主细胞表面分子可以通过二硫键与第一结合位点共价地连接。在其他可以与任一前述实施方案组合的实施方案中,宿主细胞表面分子可以包括可以是 Aga1p 的第一凝集素。在其他可以与任一前述实施方案组合的实施方案中,宿主细胞表面分子可以通过 GPI 锚与细胞膜连接。在其他可以与任一前述实施方案组合的实施方案中,其中第一结合位点包含可以是 Aga2p 的第二凝集素。在其他可以与任一前述实施方案组合的实施方案中,第二结合位点可以与展示分子共价地连接。在其他可以与任一前述实施方案组合的实施方案中,第二结合位点可以通过二硫键与展示分子共价地连接。在其他可以与任一前述实施方案组合的实施方案中,第二结合位点包括 PDZ 结构域,其可以是具有氨基酸序列 SEQ ID NO:8 的 InaD 的 PDZ 结构域。在其他可以与任一前述实施方案组合的实施方案中,展示分子包括在羧基端具有氨基酸序列 SEQ ID NO:9 的 NorpA 配体。在其他可以与任一前述实施方案组合的实施方案中,除其中的第二结合位点包括 PDZ 结构域或者其中展示分子包括 NorpA 配体外,展示分子包括 PDZ 结构域,其可以是具有氨基酸序列 SEQ ID NO:8 的 InaD 的 PDZ 结构域。在其他可以与任一前述实施方案组合的实施方案中,除其中第二结合位点包括 PDZ 结构域或者其中的展示分子包括 NorpA 配体外,第二结合位点还包括在羧基端具有氨基酸序列 SEQ ID NO:9 的 NorpA 配体。在其他可以与任一前述实施方案组合的实施方案中,其中经修饰多肽可以是支架蛋白、信号转导蛋白质、抗体、免疫球蛋白、免疫粘附素、受体、配体、癌蛋白质、转录因子或酶。在其他可以与任一前述实施方案组合的实施方案中,展示分子可以是包括 F10 多肽的纤连蛋白多肽。在其他可以与任一前述实施方案组合的实施方案中,展示分子包括可以是 MF α 的分泌信号序列、葡糖淀粉酶、Aga2 分泌信号序列、Flo 1p 分泌信号序列、转化酶分泌信号序列或者酸性磷酸酶分泌信号序列的分泌信号肽。在其他可以与任一前述实施方案组合的实施方案中,分泌

信号肽可以是 MF α /HSA 杂合前导肽。在其他可以与任一前述实施方案组合的实施方案中,展示分子的表达在可以是 AOX1 启动子、Cup1 启动子或者 Gal 启动子的第一诱导型启动子的控制下。在其他可以与任一前述实施方案组合的实施方案中,衔接分子的表达在可以是 AOX1 启动子、Cup1 启动子或者 Gal 启动子的第二诱导型启动子的控制下。在其他可以与任一前述实施方案组合的实施方案中,宿主细胞可以是巴斯德毕赤酵母(*Pichia pastoris*)或者酿酒酵母(*Saccharomyces cerevisiae*)的酵母细胞。

[0012] 根据前述方面和其任一和所有实施方案,另外的方面包括包含至少两种宿主细胞的宿主细胞的文库,其中每一宿主细胞包括不同的经修饰多肽。

[0013] 再一方面包括用于展示修饰的多肽的方法,包括 (a) 提供包含与细胞的表面连接的细胞表面分子的宿主细胞和编码包含经修饰多肽的展示多肽的第一核酸, (b) 在第一结合位点与细胞表面分子结合的情况下,将宿主细胞与包含第一结合位点和第二结合位点的衔接分子接触,并然后 (c) 在宿主细胞将展示多肽输出到宿主细胞之外的条件下,在第二结合位点与展示多肽结合的条件下,温育宿主细胞,其中第一结合位点特异地与细胞表面分子结合并且不能与展示分子结合,第二结合位点特异地与展示多肽结合并且不能与细胞表面分子结合,并且衔接分子不是经修饰多肽的组分。在其他实施方案中,宿主细胞可以展示至少 102、至少 103、至少 104 或者至少 105 的经修饰多肽。在其他可以与前述实施方案组合的实施方案中,宿主细胞表面分子可以与第一结合位点共价地连接。在其他可以与前述实施方案组合的实施方案中,宿主细胞表面分子可以通过二硫键与第一结合位点共价地连接。在其他可以与任一前述实施方案组合的实施方案中,宿主细胞表面分子可以包括可以是 Aga1p 的第一凝集素。在其他可以与任一前述实施方案组合的实施方案中,宿主细胞表面分子可以通过 GPI 锚与细胞膜连接。在其他可以与任一前述实施方案组合的实施方案中,其中第一结合位点包含可以是 Aga2p 的第二凝集素。在其他可以与任一前述实施方案组合的实施方案中,第二结合位点可以与展示分子共价地连接。在其他可以与任一前述实施方案组合的实施方案中,第二结合位点可以通过二硫键与展示分子共价地连接。在其他可以与任一前述实施方案组合的实施方案中,第二结合位点包括具有氨基酸序列 SEQ ID NO:8 的 InaD 的 PDZ 结构域的 PDZ 结构域。在其他可以与任一前述实施方案组合的实施方案中,展示分子包括在羧基端具有氨基酸序列 SEQ ID NO:9 的 NorpA 配体。在其他可以与任一前述实施方案组合的实施方案中,除其中第二结合位点包含 PDZ 结构域或者其中的展示分子包含 NorpA 配体外,展示分子包含 PDZ 结构域,其可以是具有氨基酸序列 SEQ ID NO:8 的 InaD 的 PDZ 结构域。在其他可以与任一前述实施方案组合的实施方案中,除其中的第二结合位点包括 PDZ 结构域或者其中的展示分子包括 NorpA 配体外,第二结合位点还包括在羧基端具有氨基酸序列 SEQ ID NO:9 的 NorpA 配体。在其他可以与任一前述实施方案组合的实施方案中,其中经修饰多肽可以是支架蛋白、信号转导蛋白质、抗体、免疫球蛋白、免疫粘附素、受体、配体、癌蛋白质、转录因子或酶。在其他可以与任一前述实施方案组合的实施方案中,展示分子可以是包括 F10 多肽的纤连蛋白多肽。在其他可以与任一前述实施方案组合的实施方案中,展示分子包括可以是 MF α 分泌信号序列、葡糖淀粉酶、Aga2 分泌信号序列、Flo 1p 分泌信号序列、转化酶分泌信号序列或者酸性磷酸酶分泌信号序列的分泌信号肽。在其他可以与任一前述实施方案组合的实施方案中,分泌信号肽可以是 MF α /HSA 杂合前导肽。在其他可以与任一前述实施方案组合的实施方案中,展示分子的表达在可以是

AOX1 启动子、Cup1 启动子或者 Gal 启动子的第一诱导型启动子的控制下。在其他可以与任一前述实施方案组合的实施方案中,衔接分子的表达在可以是 AOX1 启动子、Cup1 启动子或者 Gal 启动子的第二诱导型启动子的控制下。在其他可以与任一前述实施方案组合的实施方案中,宿主细胞可以是巴斯德毕赤酵母或者酿酒酵母的酵母细胞。

[0014] 仍在另一方面包括用于产生宿主细胞展示文库的方法,其包括将分别编码包含经修饰多肽的展示多肽的第一核酸的展示文库导入到多个宿主细胞中,其中至少两种导入的第一核酸编码不同的经修饰多肽,其中每一宿主细胞包含编码细胞表面多肽的第二核酸和编码包含第一结合位点和第二结合位点的衔接分子的第三核酸,其中第一结合位点与细胞表面分子而不与展示多肽结合,第二结合位点与展示多肽而不与细胞表面分子结合,并且衔接分子不是经修饰多肽的组分。该方面可以与前述方面的任一实施方案组合。

[0015] 以上所述的是本发明的非限制性实施例。其他方面和实施方案可见于本说明书中。

[0016] 附图简述

[0017] 图 1 显示三组分系统的示意图,该系统包括细胞表面分子、包括第一和第二结合位点的衔接分子,以及具有针对第二结合位点的结合配偶体(在该实施方案中是结合多肽)和作为待展示分子的经修饰多肽的展示分子。在该实施方案中,宿主细胞是酵母细胞。

[0018] 图 2 显示用于酵母表达作为细胞表面分子的 Aga1p 的 pPIC3.5AGA1 载体。

[0019] 图 3 显示用于酵母表达作为衔接分子的 Aga2p-InaD 融合多肽的 pPIC6A AGA2-InaD 载体。

[0020] 图 4 显示用于酵母表达作为展示分子的 MF α 1Hsa-Fn10-NorpA 融合多肽的 pPICHOLI-1MF α 1Hsa-Fn10-NorpA 载体,其中 NorpA 是衔接分子(即,InaD)的第二结合位点的结合配偶体并且纤连蛋白 F10 结构域是经修饰多肽。

[0021] 图 5 显示用于酵母表达作为展示分子的 MF α 1Hsa-Fn10-NorpA 融合多肽的 pPICHOLI-C MF α 1Hsa-Fn10-NorpA 载体,其中 NorpA 是衔接分子(即,InaD)的第二结合位点的结合配偶体并且纤连蛋白 F10 结构域是经修饰多肽。

[0022] 图 6 显示用于酵母表达作为衔接分子的 Aga2p-InaD 融合多肽的 pYDNBC1Aga2-InaD 载体。

[0023] 图 7 显示用于酵母表达作为展示分子的 MF α 1HSA-Fn10-NorpA 融合多肽的 pYS HSA MF α 1Fn10-NorpA 载体,其中 NorpA 是衔接分子(即,InaD)的第二结合位点的结合配偶体并且纤连蛋白 F10 结构域是经修饰多肽。

[0024] 图 8 显示用于酵母表达作为展示分子的 MF α 1HSA--NorpA 融合多肽的 pYS MF α 1HSA NorpA 载体,其中 NorpA 是衔接分子(即,InaD)的第二结合位点的结合配偶体且 HSA 是展示多肽。

[0025] 图 9A 是显示表达纤连蛋白(pYS6/CT HSA MF α 1Fn10NorpA)或者 HSA (pYS6/CT MF α 1-HSA-NorpA)的酵母细胞的蛋白质表面表达的 FACS 分析图,其中将酵母细胞用抗 myc 抗体和 APC 标记的二级抗小鼠抗体染色。图 9A 是未染色的酵母细胞对照;图 9B 是未诱导的表达纤连蛋白的酵母细胞;图 9C 是诱导的表达纤连蛋白的酵母细胞,显示出在曲线中的偏移;图 9D 是未诱导的表达 HSA 的酵母细胞;和图 9E 是诱导的表达 HSA 的酵母细胞,显示出在曲线中的偏移。

[0026] 图 10A-E 是表达纤连蛋白(pYS6/CT HSA MF α 1Fn10NorpA) 或者 HSA (pYS6/CT HSA-NorpA)的 FMAT 分析的图像,其中将酵母细胞用抗 myc 小鼠单克隆抗体和 APC 标记的二级抗小鼠抗体染色并进行 FMAT 共焦荧光分析。图 10A 是用未染色酵母细胞的对照;图 10B 是未诱导的表达纤连蛋白的酵母细胞;图 10C 是显示为白色斑点的诱导的表达纤连蛋白的酵母细胞;图 10D 是未诱导的表达 HSA 的酵母细胞;和图 10E 是显示为白色斑点的诱导的表达 HSA 的酵母细胞。

[0027] 图 11 显示用于酵母表达作为展示分子的 MF α 1scFv 溶菌酶 NorpA 融合多肽的 pYS MF α 1scFv 溶菌酶 NorpA 载体,其中 NorpA 是衔接分子(即,InaD)的第二结合位点的结合配偶体且溶菌酶是经修饰多肽。

[0028] 图 12 显示其中将 pYD NBC1Aga2-NorpA 载体用于酵母表达 NBC1Aga2-NorpA 的反向系统。

[0029] 图 13 显示其中将 pYS6/CT*MF α 1-InaD-Fn10 载体用于酵母表达 MF α 1-InaD-Fn10 的反向系统。

[0030] 发明详述

[0031] 定义

[0032] 如本文中所用,术语“宿主细胞”指经修饰以表达细胞表面分子、衔接分子和展示分子的真核细胞。此外,应当理解,宿主细胞在展示分子与宿主细胞表面上的衔接分子结合前分泌或者排出展示分子。

[0033] 如本文中所用,术语“细胞表面分子”指指向宿主细胞的胞外表面的肽、多肽、结合结构域、配体、脂类或者糖类。细胞表面分子可以通过共价结合或非共价结合锚定于细胞表面。细胞表面分子可包括磷脂、糖类或者蛋白质,通过其可以与宿主细胞的表面连接。细胞表面分子可以是与细胞表面的磷脂、糖类或者多肽结合或者缀合的多肽。例如,多肽,例如 α -凝集素、 α -凝集素和絮片蛋白可以使用磷脂酰肌醇聚糖(GPI)与宿主细胞的表面锚定或者连接。细胞表面分子也可以是具有定位于宿主细胞的表面的、可以与衔接分子的第一结合位点结合的结合结构域的跨膜蛋白质。

[0034] 如本文中所用,术语“衔接分子”指具有两个不同的结合位点的肽、多肽、结合结构域、配体、脂类或糖类或者前述的组合。衔接分子具有特异结合细胞表面分子的结合位点和特异结合展示分子的第二个不同的结合位点。并不限制本发明,多肽的两个结合位点可以是分别对融合在一起的不同分子具有其自己的结合亲和性的多肽结构域。例如,多肽可以是与 PDZ 结构域融合的 α -凝集素亚基如 Aga2p,或者其可以是与 PDZ 结构域融合的絮片蛋白如 Flol1。

[0035] 如本文中所用,术语“第一结合位点”指特异地识别并结合至少一部分细胞表面分子的衔接分子的区域。例如,第一结合位点可以包含与细胞表面分子特异结合的肽、多肽、结合结构域、配体、脂类或糖类或者它们的组合,所述细胞表面分子可以包括但不限于,肽、结合结构域、配体、蛋白质、脂蛋白、脂类或糖类。更具体地,但并不限制本发明,第一结合位点可以指通过二硫键与 α -凝集素的 Aga1p 亚基特异结合的 α -凝集素的 Aga2p 亚基。总而言之,可以将任意的两个分子结合配偶体用于第一结合位点和细胞表面分子的对应的部分。实例包括 Ni^{2+} 离子和多组氨酸标签、糖残基和伴刀豆球蛋白 A、对氨基苯基- β -D-硫代半乳糖苷和 β -半乳糖苷酶(Germino 等人, roc. Natl. Acad. Sci. USA 80 :6848 (1983),

谷胱甘肽和谷胱甘肽-S 转移酶 (Smith, D. B 和 Johnson, K. S. *Gene* 67 :31 (1988)); 葡萄糖菌 A 蛋白和 IgG (Uhlen, M. 等人 *Gene* 23 :369 (1983))、钙调蛋白镍结合蛋白 (CNBP) 和钙调蛋白琼脂糖 (Stofko-Hahn, R. E. 等人 *FEBS Lett.* 302 (3):274-278); 链霉抗生素蛋白或抗生物素蛋白和生物素 (Takashige, S. 和 Dale, G. L., *Proc. Natl. Acad. Sci. USA.* 85 :1647-1651 (1988)); 来自大肠杆菌 *malE* 基因的淀粉酶和麦芽糖结合蛋白结构域 (Bach, H. 等人, *J. Mol. Biol.* 312 :79-93 (2001))、任意表位和其对应的抗体 (参见, Kolodziej, P. A. 和 Young, R. A., *Methods Enzymol.* 194 :508-519 (1991)), 例如, FLAG (TM) 八肽或者抗洋地黄毒苷抗体和洋地黄毒苷)。

[0036] 如本文中所用, 术语“第二结合位点”指特异识别并结合展示分子的衔接分子的区域。例如, 第二结合位点可以包含与包含展示分子的肽、配体、蛋白质、脂蛋白、脂类或糖类特异结合的肽、多肽、结合结构域、配体、脂类或糖类。更具体地, 但并非限制, 第二结合位点可以指与 NorpA 配体特异结合的 PMZ 结构域。也可以将适合于第一结合位点的任一结合对用于第二结合位点, 只要它们不识别相同的配偶体。

[0037] 如本文中所用, 术语“展示分子”指可以通过结合宿主细胞表面的衔接分子定位于宿主细胞表面的分子。展示分子一般包含待展示分子 (或分子文库) 和与衔接分子的第二结合位点特异结合的结合配偶体。在某些实例中, 待展示分子和结合配偶体可以是相同的分子。通过示例的方式, 展示分子可以包含肽、多肽、结合结构域、配体、脂类或糖类或者它们的组合。应当理解, 展示分子在宿主细胞内表达或者以其他方式产生, 并分泌或者排出于细胞之外以展示在所述细胞的表面上。展示分子可以包含不同分子的文库, 其可以筛选与靶的结合或者用于改进或改变的活性。在某些实施方案中, 文库可以包含经修饰多肽。展示分子也可以包含可以标记的标签或多肽以检测展示分子与细胞表面的结合, 或者分选展示所述分子的宿主细胞。

[0038] 如本文中所用, 术语“经修饰多肽”指任一目标多肽, 其融合到与包含衔接分子的第二结合位点的肽、配体、蛋白质、脂蛋白、脂类或糖类特异结合的肽、多肽、结合结构域、配体、脂类或糖类, 并在宿主细胞的表面上展示 (并因此是展示分子的组分)。经修饰多肽的非限制性示例是支架蛋白、信号转导蛋白质、抗体、免疫球蛋白、免疫粘附素、受体、配体、癌蛋白质、转录因子和酶。

[0039] 如本文中所用, 术语“多个展示分子”指在宿主细胞的表面上展示的展示分子的至少两个拷贝。在某些实例中, 每一独特的展示分子通过不同的宿主细胞展示。

[0040] 如本文中所用, 术语“经修饰多肽的组分”指修饰肽的任意片段的任一天然存在的结合配偶体。非限制性实例包括与免疫球蛋白重链结合的免疫球蛋白轻链、结合生物素分子的抗生物素蛋白、二聚化 α -血红蛋白的两个亚基、与肌球蛋白轻链结合的肌球蛋白重链、二聚化的血型糖蛋白 A 的两个单体或者相互结合的任一天然存在的二聚体蛋白质的两个单体。

[0041] 如本文中所用, 术语“宿主细胞的文库”指多个宿主细胞, 其中每一宿主细胞包含在细胞的表面上展示的不同经修饰多肽。

[0042] 如本文中所用, 术语“不同的经修饰多肽”指至少两种经修饰多肽的氨基酸序列, 其中每一氨基酸序列包含氨基酸替换、插入或者缺失, 其将在宿主细胞表面上展示的一种经修饰多肽与在第二宿主细胞的表面上展示的另一种经修饰多肽相区别。

[0043] 如本文中所用,术语“Fn10”指人类纤连蛋白的第 10 型 III 结构域。

[0044] 展示分子

[0045] 展示分子可以用于展示在宿主细胞中表达或者以其他方式产生的任一分子。此类分子的非限制性示例如下。

[0046] 抗体支架

[0047] 展示分子可以是免疫球蛋白。最初研发了产生免疫球蛋白文库的方法以通过噬菌体展示免疫球蛋白,但是后来已经研发了其他展示方法。对于这些备选展示方法,可以采用所有产生文库的方法以允许使用本文中公开的方法展示。产生抗体文库的示例方法可以见于 Ladner 等人的美国专利号 5,223,409 ;5,403,484 和 5,571,698 ;Dower 等人的美国专利号 5,427,908 和 5,580,717 ;McCafferty 等人的美国专利号 5,969,108 和 6,172,197 和 Griffiths 等人的美国专利号 5,885,793 ;6,521,404 ;6,544,731 ;6,555,313 ;6,582,915 和 6,593,081。

[0048] 非抗体支架

[0049] 使用本文中公开的方法可以展示的公知的非免疫球蛋白构架或支架包括但不限于,纤连蛋白(Compound Therapeutics 公司,Waltham, MA)、锚蛋白(Molecular Partners AG, Zurich, 瑞士)、结构域抗体(Domantis 有限公司(Cambridge, MA) 和 Ablynx nv (Zwijnaarde, 比利时))、脂笼蛋白(Anticalin)(Pieris Proteolab AG, Freising, 德国)、小型模块免疫药物(small modular immuno-pharmaceuticals) (Trubion Pharmaceuticals 公司, Seattle, WA)、maxybody (Avidia 公司(Mountain View, CA))、A 蛋白(Affibody AG, 瑞典)和 affilin (γ 晶体蛋白或泛素) (Scil Proteins GmbH, Halle, 德国)、蛋白质表位模拟物(protein epitope mimetics) (Polyphor 有限公司, Allschwil, 瑞士)。

[0050] (i) 纤连蛋白

[0051] adnectin 支架基于纤连蛋白 III 型结构域(例如,纤连蛋白 III 型的第 10 个模块(10Fn3 结构域))。纤连蛋白 III 型结构域具有分布在两个 β 折叠之间的 7 或者 8 条 β 链,它们相对于彼此包装以形成蛋白质的核心,并进一步含有将 β 链相互连接的环(类似于 CDR) 并且是溶剂暴露的。在 β 折叠夹层的每一边缘存在至少 3 个此类环,其中该边缘是垂直于 β 链的方向的蛋白质的边界。

[0052] 这些基于纤连蛋白的支架并不是免疫球蛋白,尽管整体折叠与最小的功能性抗体片段重链的可变区密切相关,所述重链的可变区包含在骆驼和骆马 IgG 中的整个抗原识别单位。由于这种结构,非免疫球蛋白纤连蛋白分子模拟抗原结合特性,它们的性质和亲和力与抗体的相似。这些支架可以用于体外环随机化和改组策略,其与抗体在体内的亲和成熟过程类似。这些基于纤连蛋白的分子可以用作为支架,其中分子的环区域可以使用标准的克隆技术用本公开的 CDR 取代。

[0053] (ii) 锚蛋白 - 分子配偶体

[0054] 该技术基于使用具有锚蛋白衍生的重复模块作为支架的蛋白质,用于产生可以用于与不同靶结合的可变区。锚蛋白重复模块是由两个反平行的 α 螺旋和 β 转角组成的 33 个氨基酸多肽。通过使用核糖体展示最优化了可变区的结合。

[0055] (iii) Maxybodies/Avimers-Avidia

[0056] Avimer 来源于天然含有 A 结构域的蛋白质例如 LRP-1。这些结构域天然地用于

蛋白质-蛋白质相互作用,并且在人类中超过 250 种蛋白质结构上基于 A 结构域。Avimer 由许多不同的“A 结构域”单体(2-10)通过氨基酸连接体连接组成。使用描述于,例如 20040175756 ;20050053973 ;20050048512 和 20060008844 中的方法可以产生能与靶抗原结合的 Avimer。

[0057] (vi) A 蛋白-Affibody

[0058] Affibody®亲和配体是由基于 A 蛋白的一个 IgG 结合结构域的支架的 3 螺旋束组成的小的、简单的蛋白质。A 蛋白是来自细菌金黄色葡萄球菌的表面蛋白质。该支架结构域由 58 个氨基酸组成,其中 13 个氨基酸随机化以产生具有大量配体变体的 Affibody®文库(参见例如,US 5,831,012)。Affibody®分子模拟抗体,与分子量为 150kDa 的抗体比较,它们具有 6kDa 的分子量。尽管其小尺寸,但是 Affibody®分子的结合位点与抗体的结合位点类似。

[0059] (v) Anticalins - Pieris

[0060] Anticalins®是由 Pieris ProteoLab AG 公司研发的产品。它们来源于脂笼蛋白,一组广泛分布的小而且结实的蛋白质,它们通常涉及化学敏感或不溶性化合物的生理转运或贮藏。几种天然的脂笼蛋白存在于人组织或体液中。

[0061] 蛋白质结构表明为免疫球蛋白,在刚性构架的顶部具有高变环。然而,与抗体或它们的重组片段相比,脂笼蛋白由具有 160 至 180 个氨基酸残基的单一多肽链组成,仅比单一的免疫球蛋白结构域稍大。

[0062] 组成结合袋的 4 个环的集合显示出明显的结构可塑性并容忍多种侧链。因此,可以以合适的方法重新塑造结合位点,以高亲和性和特异性识别指定的不同形状的靶分子。

[0063] 通过诱变 4 个环的集合,已经将脂笼蛋白家族的一种蛋白质,大菜粉蝶(Pieris Brassicae)的后胆色素结合蛋白质(BBP)用于研发 anticalins。描述“anticalins”的专利申请的一个实例是 PCT WO 199916873。

[0064] (vi) Affilin - Scil Proteins

[0065] Affilin™分子是设计用于对蛋白质和小分子具有特异亲和性的小的非免疫球蛋白蛋白质。新的 Affilin™分子可以非常快的选自两个文库,每一所述文库都基于不同的人类来源的支架蛋白。

[0066] Affilin™分子与免疫球蛋白蛋白质不显示出任何的结构同源性。Scil Proteins 使用两种 Affilin™支架,其中一个为 γ 晶体蛋白——一种人类结构性眼晶状体蛋白质,另一个是“泛素”超家族蛋白质。两种人类支架都非常小,显示出高温稳定性并几乎抗 pH 变化和变性剂。这种高稳定性主要是由于该蛋白质扩展的 β 折叠结构。 γ 晶体蛋白衍生的蛋白质的实例描述于 W0200104144 中,并且“泛素样”蛋白质的实例描述于 W02004106368 中。

[0067] (vii) 蛋白质表位模拟物(PEM)

[0068] PEM 是模拟蛋白质的 β 发夹二级结构的中等大小的、环状、肽样分子(MW 1-2kDa),主要的二级结构涉及蛋白质-蛋白质相互作用。

[0069] 非支架

[0070] 除用于重新产生具有特定亲和性的分子的支架外,在本文中公开的方法可以用于展示任一其他的可以在宿主细胞中表达或以其他方式产生的生物分子。可以将此类生物分子,特别是可以由多核苷酸编码的便于宿主细胞表达的多肽的文库,用于筛选改进的目标

特性例如受体和配体之间改进的结合性,其中受体或配体是展示分子的一部分,或者改进的酶活性,其中酶是展示分子的一部分。

[0071] 表达系统

[0072] 表达载体可以用于在宿主细胞中表达一种或多种细胞表面分子、衔接分子和展示分子。用于真核宿主细胞的表达载体一般包括 (i) 控制转录起始的真核 DNA 元件,例如启动子, (ii) 控制转录物加工的真核 DNA 元件,例如转录终止 / 聚腺苷酸化信号序列,和 (iii) 任选地,如果该载体是独立复制的(例如,非整合载体),在真核宿主细胞中控制复制的真核 DNA 元件。为了便于构建此类表达载体,载体可以任选地包括 (iv) 编码细菌复制起点和抗生素抗性标记物的真核 DNA 元件,以在细菌宿主细胞中操作载体时提供表达载体的生长和选择。用于真菌、酵母和哺乳动物细胞宿主的合适的真核表达载体是本领域已知的,并描述于,例如,Powels 等人(*Cloning Vectors: A Laboratory Manual*, Elsevier, N. Y., 1985)中。

[0073] 具体的目标酵母宿主细胞包括酿酒酵母、巴斯德毕赤酵母和甲醇毕赤酵母(*Pichia methanolica*)。这些载体包括基于 Yip 的载体例如 Yip5、YRp 载体例如 YRp17、YEp 载体例如 YEp13 和 YCp 载体例如 YCp19。存在许多用于在酵母中表达重组蛋白质的载体。YEp 载体的其他实例包括 YEp24、YEp51 和 YEp52,其是用于将基因构建体导入到酿酒酵母中的克隆和表达载体(参见,例如 Broach 等人(1983)在 *Experimental Manipulation of Gene Expression*, M. Inouye 编辑 Academic Press, 第 83 页中)。这些载体也是穿梭载体,因为由于 pBR322 起点的存在,它们可以在大肠杆菌中复制,并且由于酵母 2 微米质粒的复制决定子,它们可以在酿酒酵母中复制。

[0074] 在酵母中发挥功能的合适的启动子包括金属硫蛋白、3-磷酸甘油酸激酶(Hitzeman 等人, *J. Biol. Chem.* 255, 2073 (1980) 或者其他糖酵解酶(Hess 等人, *J. Adv. Enzyme Req.* 7, 149 (1968); 和 Holland 等人 *Biochemistry* 17, 4900 (1978)), 例如烯醇酶、甘油醛-3-磷酸脱氢酶、己糖激酶、丙酮酸脱羧酶、磷酸果糖激酶、葡萄糖-6-磷酸异构酶、3-磷酸甘油酸变位酶、丙酮酸激酶、磷酸丙糖异构酶、磷酸-葡萄糖异构酶和葡糖激酶的启动子。在酵母表达中使用的合适的载体和启动子进一步描述于 R. Hitzeman 等人, EP073, 657 中。其他在酵母中表达的合适的启动子包括来自 GAL1 (半乳糖)、PGK (磷酸甘油酸激酶)、ADH (醇脱氢酶)、AOX1 (醇氧化酶)、HIS4 (组氨酸脱氢酶)等的启动子。许多酵母克隆载体是容易得到的并且可以按上述讨论进行修饰。其他具有通过生长条件控制转录的额外优点的启动子是醇脱氢酶 2、异细胞色素 C、酸性磷酸酶、与氮代谢相关的降解酶,和前述金属硫蛋白和甘油醛-3-磷酸脱氢酶,以及负责麦芽糖和半乳糖利用的酶的启动子区域。最后,仅在两种单倍体接合型之一中起作用的启动子可能适合于某些情况。在这些单倍体特异的启动子中,信息素启动子 MFa1 和 MF α 1 是特别重要的。

[0075] 可以通过使用酵母蛋白质的任一可用的分泌信号序列来增加包括衔接分子(如果在宿主细胞中产生)和展示分子的组分从酵母宿主细胞中分泌。一个实例是酵母接合信息素的前体 α 因子的前导序列,其也用于指导酵母中异源蛋白质的分泌(参见,例如, Valenzuela, P. 编辑, 第 269-280 页, Butterworths, London ; Brake, A. J. (1990) *Meth. Enzymol.* 185, 408-441)。除 17 个残基的氨基端信号肽外, α 因子前导序列还包括含有 72 个残基的亲水性前区,并具有 3 个 N-连接的糖基化位点。前区在 ER 和 Golgi 中是广泛糖基化的,并且在晚高尔基区室中被 Kex2 内肽酶剪切。认为氨基端前区的存在允许一些异源

蛋白质通过 ER 中的质量控制并到达周质。

[0076] 另一实例是来自酵母转化酶的前导序列(MLLQAFLFLLAGFAAKISADAHKS)(SEQ ID NO:1)。已经表明,当进入内质网时,该前导序列将从新生的异源肽中剪切。负责剪切前序列的酶 Kex2 定位于高尔基外侧区中。另外的实例是酵母酸性磷酸酶的信号序列,其可以用于指导在本文中公开的组分的分泌。

[0077] 用外源 DNA 转化酿酒酵母细胞并从其产生重组多肽的方法由,例如, Kawasaki, 美国专利号 4,599,311、Kawasaki 等人,美国专利号 4,931,373、Brake, 美国专利号 4,870,008、Welch 等人,美国专利号 5,037,743 和 Murray 等人,美国专利号 4,845,075 公开。通过选择标记,通常是药物抗性或者在不存在特定营养物(例如,亮氨酸)中生长的能力确定的表型来选择转化的细胞。在酿酒酵母中使用的优选的载体系统是由 Kawasaki 等人(美国专利号 4,931,373)公开的 POT1 载体系统,其允许通过在含葡萄糖的培养基中生长来选择转化的细胞。

[0078] 用于其他酵母的转化系统是本领域已知的,所述酵母包括多形汉逊酵母(*Hansenula polymorpha*)、粟酒裂殖酵母(*Schizosaccharomyces pombe*)、乳酸克鲁维酵母(*Kluyveromyces lactis*)、脆壁克鲁维酵母(*Kluyveromyces fragilis*)、玉蜀黍黑粉菌(*Ustilago maydis*)、巴斯德毕赤酵母(*Pichia pastoris*)、甲醇毕赤酵母(*Pichia methanolica*)、季也蒙毕赤酵母(*Pichia guilliermondii*)和麦牙糖假丝酵母(*Candida maltosa*)。参见,例如 Gleeson 等人, *J. Gen. Microbiol.* 132 :3459 (1986) 和 Cregg, 美国专利号 4,882,279。根据 McKnight 等人,美国专利号 4,935,349 的方法可以使用曲霉属(*Aspergillus*)细胞。转化产黄顶孢霉(*Acremonium chrysogenum*)的方法由 Sumino 等人,美国专利号 5,162,228 公开。转化脉孢菌属(*Neurospora*)的方法由 Laambowitz, 美国专利号 4,486,533 公开。

[0079] 例如,甲醇毕赤酵母作为用于产生重组蛋白质的宿主的用途由 Raymond, 美国专利号 5,716,808、Raymond, 美国专利号 5,736,383、Raymond 等人, *Yeast* 14 :11-23 (1998) 和在国际公开号 WO 97/17450、WO 97/17451、WO 98/02536 和 WO 98/02565 中公开。通常将在转化甲醇毕赤酵母中使用的 DNA 分子制备为双链、环状质粒,其优选地在转化前线性化。对于在甲醇毕赤酵母中的多肽产生,优选地,质粒中的启动子和终止子是甲醇毕赤酵母基因,例如甲醇毕赤酵母醇利用基因(AUG1 或者 AUG2)的启动子和终止子。其他有用的启动子包括二羟丙酮合酶(DHAS)、甲酸脱氢酶(FMD)和过氧化氢酶(CAT)基因的启动子。为了便于 DNA 整合到宿主染色体中,优选地是质粒的整个表达区段的两端侧翼都具有宿主 DNA 序列。对于希望最小化甲醇使用的大规模、工业化方法,优选地使用将两个甲醇利用基因(AUG1 和 AUG2)都缺失了的宿主细胞。对于产生分泌蛋白质,优选地是缺失了液泡蛋白酶基因(PEP4 和 PRB1)的宿主细胞。使用电穿孔法以便于将含有编码目标多肽的 DNA 的质粒导入到甲醇毕赤酵母细胞中。可以通过使用指数衰减的、具有 2.5 至 4.5kV/cm,优选地约 3.75kV/cm 场强的脉动电场和 1 至 40 毫秒,最优选地约 20 毫秒时间常数(τ)的电穿孔法转化甲醇毕赤酵母细胞。

[0080] 对于哺乳动物宿主细胞的用途,哺乳动物表达载体也是本领域熟知的并也可以使用。合适的哺乳动物宿主细胞的实例包括非洲绿猴肾细胞(Vero ;ATCC CRL 1587)、人胚肾细胞(293-HEK ;ATCC CRL 1573)、幼仓鼠肾细胞(BHK-21, BHK-570 ;ATCC CRL 8544, ATCC

CRL 10314)、犬肾细胞(MDCK ;ATCC CCL 34)、中国仓鼠卵巢细胞(CHO-K1 ;ATCC CCL61 ;CHO DG44 (Chasin 等人, Som. Cell. Molec. Genet. 12 :555, 1986))、大鼠垂体细胞(GH1 ;ATCC CCL82)、HeLa S3 细胞(ATCC CCL2. 2)、大鼠肝癌细胞(H4-II-E ;ATCC CRL 1548)、SV40 转化的猴肾细胞(COS-1 ;ATCC CRL 1650) 和鼠胚胎细胞(NIH-3T3 ;ATCC CRL1658)。

实施例

[0081] 以下提供了本中所公开的系统、组合物和方法的非限制性实施例。可以通过使用酵母 InaD 蛋白质的 PDZ 结构域和酵母 NorpA 蛋白质的羧基端 5 个氨基酸在酵母细胞的表面上展示蛋白质。该三组分蛋白质展示系统由以下组成:表达具有在其氨基端融合了分泌信号并在羧基端融合了 NorpA 配体的待展示蛋白质的载体;表达能与酵母细胞壁蛋白质特异结合并与 InaD 的 PDZ 结构域融合的、与 NorpA 配体特异结合的衔接蛋白质的第二载体;和表达与衔接蛋白质特异结合的酵母细胞壁蛋白质的第三载体。该系统适合于在酿酒酵母和巴斯德毕赤酵母中使用。

[0082] 实施例 1:在巴斯德毕赤酵母中使用 InaD/NorpA 相互作用的酵母展示

[0083] 通过在 Fn10 的氨基端融合杂合的分泌序列(MF α /HSA) 或者酵母前导序列(MF α 1) 并在其羧基端融合 NorpA 配体,研发了展示纤连蛋白 III 型结构域(Fn10) 的巴斯德毕赤酵母的蛋白质展示系统。一旦表达,Fn10 从细胞中分泌,并且 NorpA 配体通过二硫键与 InaD 的 PDZ 结构域特异结合。将 InaD 与 Aga2p 蛋白质的羧基端融合。Aga2p-InaD 融合蛋白质用作为衔接蛋白质,并且将氨基端 Aga2p 与固定于细胞表面上的 Aga1p 结合。Aga2p 通过二硫键与 Aga1p 特异地结合。

[0084] 将由 Fn10-NorpA 融合蛋白质、Aga2p-InaD 融合蛋白质和 Aga1p 细胞表面蛋白质组成的三组分系统克隆到 pPIC 表达载体中,处于诱导型启动子控制下。使用的诱导型启动子是由甲醇诱导的 AOX1 启动子。因此,当将甲醇添加到转化了载体的酵母细胞时,这 3 种蛋白质将得到表达。Aga1p 表达在细胞的表面上。Aga2p-InaD 定位于细胞表面,其中 Aga2p-InaD 的氨基端区与 Aga1p 结合。Fn10-NorpA 定位于分泌途径,从细胞中分泌,并通过羧基端 NorpA 配体结合 InaD (图 1)。也可以变换该系统,使得将 InaD 与 Aga1 融合并将 NorpA 与 Aga2p 融合(参见图 12 和 13)。

[0085] 将 c-myc 表位标记融合在 NorpA 配体和 Fn10 的羧基端之间。通过使用 c-myc 抗体, c-myc 表位允许检测展示的纤连蛋白。通过荧光激活细胞分选术(FACS) 检测与细胞表面上 c-myc 表位结合的荧光标记的 c-myc 抗体。

[0086] 菌株和培养基

[0087] 将大肠杆菌 Top10 菌株(Invitrogen Carlsbad, CA) 用作为用于重组 DNA 操作的宿主菌株。将巴斯德毕赤酵母 GS115 菌株(Invitrogen Co., Carlsbad, CA) 用于融合蛋白质 AGA2-InaD 和 HSA/MF α 1-Fn10-NorpA 的产生。将大肠杆菌培养于含有 100 μ g/mL 氨苄青霉素的 LB 培养基(1% 胰蛋白胨、0.5% 酵母提取物和 0.5% 氯化钠) 中。将巴斯德毕赤酵母培养于 BMGY 培养基(1% 酵母提取物、2% 蛋白胨、100mM 磷酸钾缓冲液(pH6.0)、1.34% 酵母氮碱基、 4×10^{-5} % 生物素和 1% 甘油) 和 BMMY 培养基(1% 酵母提取物、2% 蛋白胨、100mM 磷酸钾缓冲液(pH 6.0)、1.34% 酵母氮碱基、 4×10^{-5} % 生物素和 0.5-2.0% 甲醇) 中。

[0088] 表达质粒的构建

[0089] 通过 Geneart 合成对应于 AGA1 的基因,并亚克隆到 pPIC3.5(Invitrogen)中。得到的载体命名为 pPIC3.5-AGA1 (图 2)。通过 Geneart (德国)合成 AGA2-InaD 锚定基因,并使用 BstI 和 EcoRI 限制位点亚克隆到表达载体 pPIC6a (Invitrogen) 中。得到的载体命名为 pPIC6-AGA2-InaD (图 4)。纤连蛋白构建体由 MF α 1/HSA 杂合前导区,接着是在羧基端与 NorpA 配体序列融合的纤连蛋白组成。通过 Geneart (德国)合成全基因并亚克隆到 pPICHOLI-1 (Mobictec) 中。得到的载体命名为 pPICHOLI-1MF α 1Hsa-Fn10-NorpA (图 6)。

[0090] 酵母转化

[0091] 根据由供应商指定的方案制备电感受态巴斯德毕赤酵母 GS115 (Invitrogen) 菌株并用 SalI 消化的 pPIC3.5-AGA1、pPIC6-AGA2-InaD 和 pPICHOLI-1MF α 1Hsa-Fn10-NorpA 共转化。

[0092] 培养条件

[0093] 在 30°C 将酵母转化体预培养于含有 100 μ g/ml Zeocin 和 200 μ g/ml 杀稻瘟素的 BMGY 培养基中 16 小时,并用于接种在 1L 挡板瓶中的 200ml 的 BMGY 培养基(含有 100 μ g/ml Zeocin 和 200 μ g/ml 杀稻瘟素)以产生 0.1 的初始 OD₆₀₀ 值。培养 24 小时后,将培养物在 1000g 离心 10 分钟,并重悬于含有 0.5%、1.0% 或 2.0% 甲醇的 BMMY 培养基(+100 μ g/ml Zeocin 和 200 μ g/ml 杀稻瘟素)中。为了维持融合蛋白质的诱导,每 24 小时添加一次 100% 甲醇到培养物中以达到上述终浓度。使用 FACS 和抗 myc 抗体进行酵母表面上展示的纤连蛋白的分析。

[0094] 实施例 2:在巴斯德毕赤酵母的表面上分泌或者展示纤连蛋白的开关系统

[0095] 上述展示系统的一种变体使得可以在分泌还是展示来自巴斯德毕赤酵母的蛋白质之间选择。为了实现这一点,将由 MF α 1/HSA 杂合前导区接着在羧基端与 NorpA 配体序列融合的纤连蛋白组成的纤连蛋白构建体克隆到 pPICHOLI-C 而不是 pPICHOLI-1 中。得到的载体命名为 pPICHOLI-CMF α 1Hsa-Fn10-NorpA (图 8)。两种载体之间关键的不同是启动子,pPICHOLI-1 中的启动子是由甲醇诱导的 AOX1 启动子,并且在 pPICHOLI-C 中的启动子是由铜诱导的 Cup1 启动子。为了在巴斯德毕赤酵母的表面上展示纤连蛋白,用甲醇诱导 AGA1 和 AGA2-InaD,而用铜诱导 pPICHOLI-C。这允许捕获通过紧密的 InaD/NorpA 相互作用介导的在酵母的表面的分泌纤连蛋白。对于没有在酵母的表面上展示蛋白质的纤连蛋白,用铜诱导就足够了。在没有诱导 AGA1 和 AGA2-InaD (由甲醇驱动)的情况下,针对 NorpA 的结合配偶体(AGA1/AGA2-InaD)将不会在酵母的表面上存在,并因此纤连蛋白被分泌。

[0096] 实施例 3:在酿酒酵母中使用 InaD/NorpA 相互作用的酵母展示

[0097] 该实施例描述了用其他酵母菌株例如酿酒酵母使用 InaD/NorpA 系统。

[0098] 菌株和培养基

[0099] 将大肠杆菌 Top10 菌株(Invitrogen, Carlsbad, CA)用作为用于重组 DNA 操作的宿主菌株。将酿酒酵母菌株 EBY100 (Invitrogen Co., Carlsbad, CA)用于产生融合蛋白质 AGA2-InaD 和 MS α 1/HSA-Fn10-NorpA 或者 pYS6CT*MF α 1-HSA-NorpA。将大肠杆菌培养于含有 100 μ g/mL 氨苄青霉素或 100 μ g/mL 杀稻瘟素的 LB 培养基(1% 胰蛋白胨、0.5% 酵母提取物和 0.5% 氯化钠)中培养。将 EBY100 培养于 CM 培养基 -URA 中。

[0100] 表达质粒的构建

[0101] 通过 Geneart (德国)合成 InaD 锚定基因,并使用 HindIII 和 EcoRI 限制位点与

AGA2 锚定蛋白质符合读框地克隆到表达载体 pYD_NBC1(pYD1 的衍生物, Invitrogen) 中。得到的载体命名为 pYD_NBC1AGA2-InaD (图 10)。纤连蛋白构建体由 MF α 1/HSA 杂合前导区接着是在羧基端与 NorpA 配体融合的纤连蛋白组成。通过 Geneart (德国) 合成全基因并亚克隆到 pYS6CT (Invitrogen) 中, 其中将复制起点用 CEN6/ARS4 区取代。得到的载体命名为 pYS6CT_HSA_MF α 1_Fn10_NorpA (图 12)。

[0102] 质粒分离自大肠杆菌并确认序列。然后将纯化的质粒共转化到 EBY100 中并接种在含有 CM-TRP、+200 μ g/ml 杀稻瘟素的选择性培养基中。转化菌落在 2 天内出现, 并使用抗 myc 抗体(ccccc) 通过 FACS 分析测试纤连蛋白的展示。

[0103] 实施例 4: 在巴斯德毕赤酵母中使用 Flo1-InaD/NorpA 的酵母展示

[0104] 该实施例描述了在毕赤酵母中使用 InaD/NorpA 的备选表达系统 Flo1 的用途。

[0105] 菌株和培养基

[0106] 将大肠杆菌 Top10 菌株(Invitrogen Carlsbad, CA) 用作为用于重组 DNA 操作的宿主菌株。将巴斯德毕赤酵母 GS115 菌株(Invitrogen Co., Carlsbad, CA) 用于融合蛋白质 Flo1-InaD 和 HSA/MF α 1-Fn10-NorpA 的产生。将大肠杆菌培养于含有 100 μ g/mL 氨苄青霉素的 LB 培养基(1% 胰蛋白胨、0.5% 酵母提取物和 0.5% 氯化钠) 中。将巴斯德毕赤酵母培养于 BMGY 培养基(1% 酵母提取物、2% 蛋白胨、100mM 磷酸钾缓冲液(pH6.0)、1.34% 酵母氮碱基、4 $\times 10^{-5}$ % 生物素和 1% 甘油) 和 BMMY 培养基(1% 酵母提取物、2% 蛋白胨、100mM 磷酸钾缓冲液(pH 6.0)、1.34% 酵母氮碱基、4 $\times 10^{-5}$ % 生物素和 0.5-2.0% 甲醇) 中。

[0107] 表达质粒的构建

[0108] 通过 Geneart (德国) 合成在羧基端与 InaD 的 PDZ 结构域融合的 Flo1 基因, 并使用 5' EcoR1 位点和 3' Not1 位点亚克隆到 pPIC3.5 (Invitrogen) 中。得到的载体命名为 pPIC3.5-Flo1-InaD。融合蛋白质的表达通过甲醇诱导的启动子 AOX1 驱动。纤连蛋白构建体由 MF α 1/HSA 杂合前导区, 接着在羧基端与 NorpA 配体序列融合的纤连蛋白组成。通过 Geneart (德国) 合成全基因并亚克隆到 pPICHOLI-1 (Mobictec) 中。得到的载体命名为 pPICHOLI-1MF α 1Hsa-Fn10-NorpA (图 6)。纤连蛋白构建体的表达通过甲醇诱导的启动子 AOX1 驱动。

[0109] 酵母转化

[0110] 根据由供应商指定的方案制备电感受态巴斯德毕赤酵母 GS115 (Invitrogen) 菌株并用 SalI 消化的 pPIC3.5-Flo1-InaD 和 pPICHOLI-1MF α 1Hsa-Fn10-NorpA 共转化。

[0111] 培养条件

[0112] 在 30 $^{\circ}$ C 将酵母转化体预培养于含有 100 μ g/ml Zeocin 和 200 μ g/ml 杀稻瘟素的 BMGY 培养基中 16 小时, 并用于接种在 1L 挡板瓶中的 200ml 的 BMGY 培养基(含有 100 μ g/ml Zeocin 和 200 μ g/ml 杀稻瘟素) 以产生 0.1 的初始 OD₆₀₀ 值。培养 24 小时后, 将培养物在 1000g 离心 10 分钟, 并重悬于含有 0.5%、1.0% 或 2.0% 甲醇的 BMMY 培养基(+100 μ g/ml Zeocin 和 200 μ g/ml 杀稻瘟素) 中。为了维持融合蛋白质的诱导, 每 24 小时添加一次 100% 甲醇到培养物中以达到上述终浓度。使用 FACS 和抗 myc 抗体进行酵母表面上展示的纤连蛋白的分析。

[0113] 实施例 5: 筛选纤连蛋白文库

[0114] 纤连蛋白文库展示

[0115] 通过本领域熟知的方法,包括在美国专利号 6,673,901 中公开的方法产生纤连蛋白文库。其他方法,例如使用易错聚合酶链式反应、使用随机引发技术或者使用计算机技术都是本领域熟知的,并且也可以使用。设计的纤连蛋白文库具有合适的限制酶切割位点,以将文库克隆到酵母表达载体中。

[0116] 如实施例 1 中所述,将纤连蛋白文库展示在多个巴斯德毕赤酵母细胞上。修饰纤连蛋白文库以含有与氨基端融合的 MF α /HSA 杂合前导序列和与羧基端融合的 NorpA 配体序列。然后,将修饰的纤连蛋白文库克隆到 pPICHOLI-1 载体中。与上面的实施例一样,纤连蛋白文库的表达在 AOX1 启动子的控制下。将巴斯德毕赤酵母用表达纤连蛋白文库的 pPICHOLI-1 载体和表达 Aga1p 和 Aga2p-InaD 的载体转化。通过添加甲醇到细胞中诱导组分的表达,并将纤连蛋白文库展示于多个巴斯德毕赤酵母细胞上。

[0117] 筛选展示文库

[0118] 使用本领域已知的许多方法之一,针对目标靶蛋白质筛选酵母展示纤连蛋白文库。例如,将靶蛋白质在允许靶蛋白质与文库的任一成员特异结合的条件下与酵母展示纤连蛋白文库接触。此时,所有结合的靶蛋白质都固定在酵母细胞的表面。将所有未结合的靶蛋白质洗掉。通过本领域熟知的方法,例如特异针对靶蛋白质的荧光标记抗体荧光标记结合的靶蛋白质。此时,标记的靶蛋白质固定在酵母细胞的表面,然后使用流式细胞术,即 FACS 检测。结合标记的靶蛋白质的酵母细胞将发荧光,并且可以从那些不结合靶蛋白质的酵母细胞中分选出来。将已经结合靶蛋白质的分选的酵母细胞克隆扩增,并确定结合靶蛋白质的克隆细胞系或者含有纤连蛋白文库成员的细胞系。

[0119] 实施例 6:筛选蛋白质文库

[0120] 蛋白质文库展示

[0121] 通过本领域熟知的方法,例如使用易错聚合酶链式反应、使用随机引物技术或者使用计算机技术产生蛋白质文库。设计的蛋白质文库具有合适的限制酶切割位点,以将文库克隆到酵母表达载体中。

[0122] 如上面的实施例 1 所述,将克隆的蛋白质文库展示在多个巴斯德毕赤酵母细胞上。修饰蛋白质文库以含有与氨基端融合的 MF α /HSA 杂合前导序列和与羧基端融合的 NorpA 配体序列。然后,将修饰的蛋白质文库克隆到 pPICHOLI-1 载体中。与上面的实施例,蛋白质文库的表达在 AOX1 启动子的控制下。将巴斯德毕赤酵母用表达蛋白质文库的 pPICHOLI-1 载体和表达 Aga1p 和 Aga2p-InaD 的载体转化。通过添加甲醇到细胞中诱导组分的表达,并将蛋白质文库展示于多个巴斯德毕赤酵母细胞上。

[0123] 筛选展示文库

[0124] 使用本领域已知的许多方法之一,针对目标靶蛋白质筛选酵母展示蛋白质文库。例如,将靶蛋白质在允许靶蛋白质与文库的任一成员特异结合的条件下与酵母展示蛋白质文库接触。此时,所有结合的靶蛋白质都固定在酵母细胞的表面。将所有未结合的靶蛋白质洗掉。通过本领域熟知的方法,例如特异针对靶蛋白质的荧光标记抗体荧光标记结合的靶蛋白质。此时,标记的靶蛋白质固定在酵母细胞的表面上,然后使用流式细胞术,即 FACS 检测。结合标记的靶蛋白质的酵母细胞会发荧光,并且可以从那些不结合靶蛋白质的酵母细胞中分选出来。将分选的已经结合靶蛋白质的酵母细胞克隆扩增,并确定结合靶蛋白质的克隆细胞系或者含有蛋白质文库的成员的细胞系。

[0125] 实施例 7:筛选纤连蛋白或者 HSA 文库

[0126] 纤连蛋白或者 HSA 文库展示

[0127] 通过本领域熟知的方法,例如使用易错聚合酶链式反应、使用随机引物技术或者使用计算机技术产生纤连蛋白或 HSA 文库。设计的文库具有合适的限制酶切割位点,以将文库克隆到酵母表达载体中(参见图 8 和 SEQ ID NO:10)。

[0128] 如上面的实施例 1 所述,将克隆的纤连蛋白文库或 HSA 文库展示在多个巴斯德毕赤酵母细胞上。修饰纤连蛋白文库以含有与氨基端融合的 MF α /HSA 杂合前导序列和与羧基端融合的 NorpA 配体序列(pYS HSA_MF α 1 Fn10 NorpA)。修饰 HSA 文库以含有 MF α CT(羧基端)。其是用于构建的初始 Invitrogen 载体的剩余物。如果查找载体图(例如图 13),你会发现羧基端 v5 和插入片段的羧基端的 6 $\times$ 组氨酸序列。我们已经在其前面放置了终止子,并且其不会在最终的展示蛋白质前导序列(pYS6/CT HSA-NorpA)中翻译。然后,将修饰的纤连蛋白文库或 HSA 文库克隆到 pYS 载体中。纤连蛋白文库或 HSA 文库的表达在 T7 启动子的控制下。将巴斯德毕赤酵母细胞用表达纤连蛋白文库或 HSA 文库的 pYS 载体和表达 Aga1p 和 Aga2p-InaD 的载体转化。通过添加甲醇到细胞中诱导组分的表达,并将纤连蛋白文库或 HSA 文库展示于多个巴斯德毕赤酵母细胞上。

[0129] (a) 蛋白质表面表达的 FACS 分析

[0130] 将表达纤连蛋白(pYS HSA_MF α 1 Fn10 NorpA)或者 HSA(pYS6/CTHSA-NorpA)的酵母细胞用抗 myc 抗体,然后用 APC 标记的二级抗小鼠抗体染色,并然后进行 FACS 分析。分析的结果显示于图 9A-E 中。具体地,图 9A 是显示出未染色酵母细胞的对照样品;图 9B 是表达纤连蛋白的未诱导的酵母细胞的样品;图 9C 是表达纤连蛋白的诱导的酵母细胞的样品,与未诱导的细胞比较,显示出细胞的偏移;图 9D 是表达 HSA 的未诱导酵母细胞的样品;和图 9E 是表达 HSA 的诱导的酵母细胞的样品,与未诱导的细胞比较,再次显示出细胞的偏移。结果清楚地表明,酵母展示系统能表达纤连蛋白分子和蛋白质,例如 HSA。

[0131] (b) 荧光微体积测定技术(FMAT、Perkin Elmer)分析表达纤连蛋白或 HSA 的酵母

[0132] 也通过 FMAT 通过用抗 myc 抗体和 APC 标记的二级抗小鼠抗体分析表达纤连蛋白(质粒)或 HSA(质粒)的酵母细胞。然后,将样品进行 FMAT 共焦荧光显微镜分析,并显示于图 10A-E 中。那些表达纤连蛋白或 HSA 的菌落相对于黑色背景呈现出白点。具体地,图 10A 是显示未染色的酵母细胞的对照样品,并呈现出完全的黑色;图 10B 是表达纤连蛋白的未诱导样品。未诱导酵母细胞不产生纤连蛋白并且不会被检测到(图像呈现黑色);图 10C 是表达纤连蛋白的诱导的酵母细胞的样品。在该实例中,诱导导致具有 myc 标记的纤连蛋白表达并可以使用抗 myc 抗体检测。随后用 APC 二级抗小鼠抗体和 FMAT 共焦荧光显微镜检测,产生了检测到的可见白色菌落。图 10D 是表达 HSA 的未诱导酵母细胞的样品,如前所述,未诱导的酵母细胞不产生纤连蛋白并且图像呈现黑色;和图 10E 是表达 HSA 的诱导的酵母细胞的样品,与未诱导细胞比较,再次显示出小的白色菌落。这些结果进一步证实,酵母展示系统能表达纤连蛋白分子和蛋白质,例如 HSA。

[0133] 实施例 8:筛选单链 Fv 文库

[0134] 单链 Fv 文库展示

[0135] 通过本领域熟知的方法,例如使用易错聚合酶链式反应、使用随机引物技术或者使用计算机技术产生单链 Fv 文库。设计的文库具有合适的限制酶切割位点,以将文库克隆

到酵母表达载体中(参见图 11 和 SEQ ID NO:11)。

[0136] 如上面的实施例 1 所述,将克隆的 scFv 溶菌酶文库展示在多个巴斯德毕赤酵母细胞上。修饰 scFv 文库以含有与氨基端融合的 MF α 前导序列和与羧基端融合的 NorpA 配体序列(pYS6/CT*MF α 1-scFv 溶菌酶-NorpA)。然后,将修饰的 scFv 溶菌酶文库克隆到 pYS 载体中。scFv 溶菌酶的表达在 T7 启动子的控制下。将巴斯德毕赤酵母用表达 scFv 溶菌酶文库的 pYS 载体和表达 Aga1p 和 Aga2p-InaD 的载体转化。通过添加甲醇到细胞中诱导组分的表达,并将纤连蛋白文库或 HSA 文库展示在多个巴斯德毕赤酵母细胞上。

[0137] (a) 蛋白质表面表达的 FACS 分析

[0138] 用抗 myc 抗体然后 APC 标记的二级抗小鼠抗体染色表达 scFv 溶菌酶的酵母细胞,并然后进行 FACS 分析。

[0139] (b) FMAT 分析表达纤连蛋白或 HSA 的酵母

[0140] 通过 FMAT 通过用抗 myc 抗体和 APC 标记的二级抗小鼠抗体染色分析表达 scFv 溶菌酶的酵母细胞。然后,将样品进行 FMAT 共焦荧光显微镜分析。

[0141] 实施例 9:筛选具有反向系统的蛋白质文库

[0142] 蛋白质文库展示

[0143] 在该实施例中,将本文中所述的酵母展示系统反向,以使 NorpA 与 Aga2 融合并使 InaD 与 Aga1 融合。通过本领域熟知的方法,例如使用易错聚合酶链式反应、使用随机引物技术或者使用计算机技术产生蛋白质文库。设计的文库具有合适的限制酶切割位点,以将文库克隆到酵母表达载体中(参见图 12 和 13 以及 SEQ ID NO:12 和 13)。

[0144] 通过在 InaD 的氨基端融合前导序列(MF α 1)并在其羧基端融合 Fn10,研发了展示纤连蛋白 III 型结构域(Fn10)的巴斯德毕赤酵母的蛋白质展示系统。一旦表达,Fn10 从细胞中分泌并且 InaD 的 PDZ 结构域通过二硫键与 NorpA 配体结合。将 NorpA 与 Aga2p 蛋白质的羧基端融合。Aga2p-NorpA 融合蛋白质用作为衔接蛋白质,并且氨基端 Aga2p 与固定于细胞表面上的 Aga1p 结合。Aga2p 通过二硫键与 Aga1p 特异地结合。

[0145] 将由 Aga2p-NorpA 融合蛋白质、Aga1-InaD 融合蛋白质和 Aga1p 细胞表面蛋白质组成的三组分系统分别克隆到 pPD 和 pYS 表达载体中,处于 Gal 诱导型启动子控制下。

[0146] 使用的诱导型启动子是由半乳糖诱导的 Gal1 启动子。因此,当将半乳糖添加到用载体转化了的酵母细胞中时,这 3 种蛋白质将得到表达。Aga1p 表达在细胞的表面上。Aga2p-Norp 定位于细胞表面,其中 Aga2p-Norp 的氨基端区与 Aga1p 结合。从 Fn10-InaD 定位于分泌途径,从细胞中分泌,并通过羧基端 InaD 配体结合 NorpA。

[0147] 序列

[0148] pPIC3.5AGA1 (956bp-3136bp,直接的) 242aa

[0149]

MTLSFAHFTY LFTILLGLTN IALASDPETI LVTITKTND NGVVTPTVSP
 ALVSTSTIVQ AGTTTLYTTW CPLTVSTSSA AEISPSISYA TTLSRFSTLT
 LSTEVCSSHEA CPSSSTLFTT TLSVTSKFTS YICPTCHTTA ISSLSEVGTT
 TVVSSSAIEP SSASIIISPVT STLSSTTSSN PTTTSLSSTS TSPSSTSTSP
 SSTSTSSSST STSSSSTSTS SSSTSTSPSS TSTSSSLTST SSSSTSTSQS
 STSTSSSSTS TSPSSTSTSS SSTSTSPSSK STSASSTSTS SYSTSTSPSL
 TSSSPTLAST SPSSTISISST FTDSTSSLGS SIASSSTSVS LYSPSTPVYS
 VPSTSSNVAT PSMTSSTVET TVSSQSSSEY ITKSSISTTI PSFSMSTYFT
 TVSGVTTMYT TWCPSSESE TSTLTSMHET VPTDATVCTH ESCMPSQTTS
 LITSSIKMST KNVATSVSTS TVESSYACST CAETSHSYSS VQTASSSSVT
 QOTTSTKSWV SSMTTSDDEF NKHATCKYHV TSSGTSTIST SVSEATSTSS
 IDSESQEQSS HLLSTSVLSS SLSATLSSD STILLFSSVS SLSEVQSPVT
 TLOISSTSEI LQPTSSTAIA TISASTSSL SATSISTPSTS VESTIESSSL
 TPTVSSIFLS SSSAPSSLQT SVTTTEVSTT SISIQYQTSS MVTISQYMGs
 GSQTRLPLGK LVFAIMAVAC NVIFS (SEQ ID NO: 2)

[0150] pPIC6 A AGA2-InaD (941bp-1648bp,直接的) 78aa

[0151]

MQLLRCSIF SVIASVLAQE LTTICEQIPS PTLESTPYSL STPTILANGK
 AMQGVFEYYK SVTFVSNCGS HPSTTSKGGP INTQYVFKLL QASGGGGSGG
 GSGGGGSAS MTGGQQMGRE NLYFQGVPGS SVVSRAGELI HMTLDKTKG
 KSFGICIVRG EVKDSPTTKT TGIFIKGIVF DSPARLCGRL KVGDRILSLN
 GKDVNRNTEQ AVIDLIKEAD FKIELEIQTDF DK (SEQ ID NO: 3)

[0152] pPICHOLI-1_Mf α 1Hsa-Fn10-NorpA (884bp-1441bp,直接的) 62aa

[0153]

MKWVSFISLL FLFSAYSRS LDKRENLYFQ GGSVSDVPRD LEVVAATPTS
 LLISWDAPAV TVRYRITYG ETGGNSPVQE FTVPGSKSTA TISGLKPGVD
 YTITVYAVTG RGDSPASSKP ISINYRTEFE NLYFQGGGGG GEQKLISEED
 LHHHHHPST PPTPSPSTPP TPSPSYKTQG KTEFCA (SEQ ID NO: 4)

[0154] pPICHOLI-C Mf α 1Hsa-Fn10-NorpA (691bp-1248bp,直接的) 62aa

[0155]

MKWVSFISLL FLFSAYSRS LDKRENLYFQ GGSVSDVPRD LEVVAATPTS
 LLISWDAPAV TVRYRITYG ETGGNSPVQE FTVPGSKSTA TISGLKPGVD
 YTITVYAVTG RGDSPASSKP ISINYRTEFE NLYFQGGGGG GEQKLISEED
 LHHHHHPST PPTPSPSTPP TPSPSYKTQG KTEFCA (SEQ ID NO: 5)

[0156] pYD_NBC1_Aga2-InaD (534bp-1235bp,直接的) 78aa

[0157]

MQLLRCSIF SVIASVLAQE LTTICEQIPS PTLESTPYSL STPTILANGK
 AMQGVFEYK SVTFVSNCGS HPSTTSKGGP INTQYVFKLL QASGGGGSGG
 GSGGGGSAS MTGGQOMGRE NLYFQGVPGS SVVSRAGELI HMVTLDKTGK
 KSPGICIVRG EVKDSPNTKT TGIFIKGIVP DSPAHLGRL KVGDRILSLN
 GKDVNRNSTEQ AVIDLIKEAD FKIELEIQTf DK (SEQ ID NO: 6)

[0158] pYS6/CT_HSA_MF α 1_Fn10_NorpA (513bp-1080bp, 直接的) 63aa

[0159]

MKWVSFISLL FLESSAYSRS LDKRENLYFQ GGSVSDVPRD LEVVAATPTS
 LLISWDAPAV TVRYRITYG ETGGNSPVQE FTVPGSKSTA TISGLKPGVD
 YTITVYAVTG RGDSPASSKP ISINYRTEFE NLYFQGGGGG GEQKLISEED
 LHHHHHPST PPTPSPSTPP TPSPSYKTQG KTEFCA (SEQ ID NO: 7)

[0160] InaD PDZ 结构域氨基酸序列(InaD aa 11-107)

[0161]

AGELINMVTI DKTGKKSFGI CIVRGEVKDS PNTKTTGIFI KGIVPDSPAH
 LCGRLKVGDR ILSLNGKDVR NSTEQAVIDL IKEADFKIEL EIQTfDK (SEQ ID
 NO: 8)

[0162] 包括 EFCA 基序的 NorpA 羧基端 11 个氨基酸

[0163] YKTQGKTEFC A (SEQ ID NO :9)

[0164] pYS6/CT*MF α HSA-NorpA

[0165] MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPFSNST

[0166] NNGLLFINTTIIASIAAKEEGVSLEKREAEASDAHKSEVAHRFKDLGEENFKALVLI AF

[0167] AQYLQQCPFEDHVKL VNEVTEFAKTCVADESAENCDSLHTLFGDKLCTVATLRETYGE

[0168] MADCCAKQEPERNECFLQHKDDNP NLPRLVRPEVDVMCTAFHDNEETFLKKYLYE IARR

[0169] HPYFYAPELLFFAKRYKAAFTGCCAADKAACLLPKLDEL RDEGKASSAKQRLKCASLQ

[0170] KFGERAFAKAWAVARLSQRFPKAEFAEVSKLVTDLT KVHTECCHGDLLECADDRADLAKY

[0171] ICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLP SLAADFVESKDVCKNYAEA

[0172] KDVFLGMFLYEYARRHPDYSVVL LRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLV

[0173] EEPQNLIKQNCLEFELGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCKH

[0174] PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVDETYVP

[0175] KEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKC

[0176] CKADDKETCFAEEGKKLVAASQAALGLGSENLYFQGGGGGEQKLISEEDLHHHHHHHHH

[0177] PSTPPTPSPSTPPTPSPSYKTQGKTEFCA (SEQ ID NO :10).

[0178] pYS6/CT*MF α 1-scFv 溶菌酶 -NorpA

[0179] MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPFSNS

[0180] TNGLLFINTTIIASIAAKEEGVSLEKREAEASQVKLQQSGAELVKPGASVKLSCTASG

[0181] FNIKDTYMHVVKRPEQGLEWIGRIDPANGNTKYDPKFQ GKATITADTSSNTAYLQLSS

[0182] LTSED TAVYYCARWDWYFDVWGQGTTVT VSSGGGGSGGGSGGGSDIELTQSPSSMYT

[0183] SLGERVTITCKASQDINSYLRWFQKPKGSPKTLIYYATSLADGVPSRFSGSGSGQDYS

[0184] LTISSLESDDTTTYYCLQHGESPYTFGGG TKLEIKRAAAEQKLISEEDLNGSENLYFQG

- [0185] SGGGGEQKLISEEDLHHHHHHHPSTPPTPSPSTPPTPSPSYKTQGKTEFCA (SEQ ID NO :11)
- [0186] pYD_NBC1_Aga2-NorpA
- [0187] MQLLRCSIFSIVIASVLAQELTTICEQIPSPITLESTPYSLSSTTTILANGKAMQGVFEYY
- [0188] KSVTFVSNCGSHPTTSTKGSPINTQYVFKLLQASGGGSGGGGSYKTQGKTEFCA
- [0189] (SEQ ID NO :12)
- [0190] pYS6/CT*MF α 1-InaD-Fn10 (507bp-1487bp, 直接的) 109aa
- [0191] MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNST
- [0192] NNGLLFINTTIAASIAAKEEGVSLEKREAEAAASAGELIHMVTLDKTGKKSFGICIVRGEV
- [0193] KDSPNTKTTGIFIKGIVPDSPAHL CGRLKVGDRILSLNGKDVRNSTEQAVIDLKEADF
- [0194] KIELEIQTFDKSGGGGEQKLISEEDLHHHHHHHPSTPPTPSPSTPPTPSPENLYFQGVSD
- [0195] VPRDLEVVAATPTSLLISWDAPAVTVRYRITYGETGGNSPVQEFTVPGSKSTATISGL
- [0196] KPGVDYTITVYAVTGRGDSPASSKPISINYRT (SEQ ID NO :13)

[0001]

序列表

<110> A·洛

<120> 酵母展示系统

<130> 53011 WO PCT

<140> PCT/EP2009/067066

<141> 2009-12-14

<150> 60/122910

<151> 2008-12-16

<160> 13

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> 酿酒酵母

<400> 1

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys

1 5 10 15

Ile Ser Ala Asp Ala His Lys Ser

20

<210> 2

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> pPIC3.5 AGA1

[0002]

```

<400> 2
Met Thr Leu Ser Phe Ala His Phe Thr Tyr Leu Phe Thr Ile Leu Leu
 1           5           10           15
Gly Leu Thr Asn Ile Ala Leu Ala Ser Asp Pro Glu Thr Ile Leu Val
          20           25           30
Thr Ile Thr Lys Thr Asn Asp Ala Asn Gly Val Val Thr Thr Thr Val
        35           40           45
Ser Pro Ala Leu Val Ser Thr Ser Thr Ile Val Gln Ala Gly Thr Thr
        50           55           60
Thr Leu Tyr Thr Thr Trp Cys Pro Leu Thr Val Ser Thr Ser Ser Ala
65           70           75           80
Ala Glu Ile Ser Pro Ser Ile Ser Tyr Ala Thr Thr Leu Ser Arg Phe
          85           90           95
Ser Thr Leu Thr Leu Ser Thr Glu Val Cys Ser His Glu Ala Cys Pro
        100          105          110
Ser Ser Ser Thr Leu Pro Thr Thr Thr Leu Ser Val Thr Ser Lys Phe
        115          120          125
Thr Ser Tyr Ile Cys Pro Thr Cys His Thr Thr Ala Ile Ser Ser Leu
130          135          140
Ser Glu Val Gly Thr Thr Thr Val Val Ser Ser Ser Ala Ile Glu Pro
145          150          155          160
Ser Ser Ala Ser Ile Ile Ser Pro Val Thr Ser Thr Leu Ser Ser Thr
          165          170          175
Thr Ser Ser Asn Pro Thr Thr Thr Ser Leu Ser Ser Thr Ser Thr Ser
          180          185          190
Pro Ser Ser Thr Ser Thr Ser Pro Ser Ser Thr Ser Thr Ser Ser Ser
          195          200          205
Ser Thr Ser Thr Ser Ser Ser Ser Thr Ser Thr Ser Ser Ser Ser Thr
          210          215          220
Ser Thr Ser Pro Ser Ser Thr Ser Thr Ser Ser Ser Leu Thr Ser Thr
225          230          235          240
Ser Ser Ser Ser Thr Ser Thr Ser Gln Ser Ser Thr Ser Thr Ser Ser
          245          250          255
Ser Ser Thr Ser Thr Ser Pro Ser Ser Thr Ser Thr Ser Ser Ser Ser
          260          265          270
Thr Ser Thr Ser Pro Ser Ser Lys Ser Thr Ser Ala Ser Ser Thr Ser
          275          280          285
Thr Ser Ser Tyr Ser Thr Ser Thr Ser Pro Ser Leu Thr Ser Ser Ser
          290          295          300
Pro Thr Leu Ala Ser Thr Ser Pro Ser Ser Thr Ser Ile Ser Ser Thr
305          310          315          320
Phe Thr Asp Ser Thr Ser Ser Leu Gly Ser Ser Ile Ala Ser Ser Ser
          325          330          335
Thr Ser Val Ser Leu Tyr Ser Pro Ser Thr Pro Val Tyr Ser Val Pro

```

[0003]

	340		345		350
Ser Thr Ser Ser Asn Val Ala Thr Pro Ser Met Thr Ser Ser Thr Val					
	355		360		365
Glu Thr Thr Val Ser Ser Gln Ser Ser Ser Glu Tyr Ile Thr Lys Ser					
	370		375		380
Ser Ile Ser Thr Thr Ile Pro Ser Phe Ser Met Ser Thr Tyr Phe Thr					
385		390		395	400
Thr Val Ser Gly Val Thr Thr Met Tyr Thr Thr Trp Cys Pro Tyr Ser					
	405		410		415
Ser Glu Ser Glu Thr Ser Thr Leu Thr Ser Met His Glu Thr Val Thr					
	420		425		430
Thr Asp Ala Thr Val Cys Thr His Glu Ser Cys Met Pro Ser Gln Thr					
	435		440		445
Thr Ser Leu Ile Thr Ser Ser Ile Lys Met Ser Thr Lys Asn Val Ala					
450		455		460	
Thr Ser Val Ser Thr Ser Thr Val Glu Ser Ser Tyr Ala Cys Ser Thr					
465		470		475	480
Cys Ala Glu Thr Ser His Ser Tyr Ser Ser Val Gln Thr Ala Ser Ser					
	485		490		495
Ser Ser Val Thr Gln Gln Thr Thr Ser Thr Lys Ser Trp Val Ser Ser					
	500		505		510
Met Thr Thr Ser Asp Glu Asp Phe Asn Lys His Ala Thr Gly Lys Tyr					
	515		520		525
His Val Thr Ser Ser Gly Thr Ser Thr Ile Ser Thr Ser Val Ser Glu					
530		535		540	
Ala Thr Ser Thr Ser Ser Ile Asp Ser Glu Ser Gln Glu Gln Ser Ser					
545		550		555	560
His Leu Leu Ser Thr Ser Val Leu Ser Ser Ser Ser Leu Ser Ala Thr					
	565		570		575
Leu Ser Ser Asp Ser Thr Ile Leu Leu Phe Ser Ser Val Ser Ser Leu					
	580		585		590
Ser Val Glu Gln Ser Pro Val Thr Thr Leu Gln Ile Ser Ser Thr Ser					
	595		600		605
Glu Ile Leu Gln Pro Thr Ser Ser Thr Ala Ile Ala Thr Ile Ser Ala					
610		615		620	
Ser Thr Ser Ser Leu Ser Ala Thr Ser Ile Ser Thr Pro Ser Thr Ser					
625		630		635	640
Val Glu Ser Thr Ile Glu Ser Ser Ser Leu Thr Pro Thr Val Ser Ser					
	645		650		655
Ile Phe Leu Ser Ser Ser Ser Ala Pro Ser Ser Leu Gln Thr Ser Val					
	660		665		670
Thr Thr Thr Glu Val Ser Thr Thr Ser Ile Ser Ile Gln Tyr Gln Thr					
	675		680		685
Ser Ser Met Val Thr Ile Ser Gln Tyr Met Gly Ser Gly Ser Gln Thr					

[0004]

690	695	700
Arg Leu Pro Leu Gly Lys Leu Val Phe Ala Ile Met Ala Val Ala Cys		
705	710	715
Asn Val Ile Phe Ser		720
	725	
<210> 3		
<211> 232		
<212> PRT		
<213> 人工序列		
<220>		
<223> pPIC6 A AGA2-InaD		
<400> 3		
Met Gln Leu Leu Arg Cys Phe Ser Ile Phe Ser Val Ile Ala Ser Val		
1	5	10
Leu Ala Gln Glu Leu Thr Thr Ile Cys Glu Gln Ile Pro Ser Pro Thr		
	20	25
Leu Glu Ser Thr Pro Tyr Ser Leu Ser Thr Thr Thr Ile Leu Ala Asn		30
	35	40
Gly Lys Ala Met Gln Gly Val Phe Glu Tyr Tyr Lys Ser Val Thr Phe		45
	50	55
Val Ser Asn Cys Gly Ser His Pro Ser Thr Thr Ser Lys Gly Ser Pro		60
65	70	75
Ile Asn Thr Gln Tyr Val Phe Lys Leu Leu Gln Ala Ser Gly Gly Gly		80
	85	90
Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Ser Met Thr		95
	100	105
Gly Gly Gln Gln Met Gly Arg Glu Asn Leu Tyr Phe Gln Gly Val Pro		110
	115	120
Gly Ser Ser Val Val Ser Arg Ala Gly Glu Leu Ile His Met Val Thr		125
	130	135
Leu Asp Lys Thr Gly Lys Lys Ser Phe Gly Ile Cys Ile Val Arg Gly		140
145	150	155
Glu Val Lys Asp Ser Pro Asn Thr Lys Thr Thr Gly Ile Phe Ile Lys		160
	165	170
Gly Ile Val Pro Asp Ser Pro Ala His Leu Cys Gly Arg Leu Lys Val		175
	180	185
Gly Asp Arg Ile Leu Ser Leu Asn Gly Lys Asp Val Arg Asn Ser Thr		190
	195	200
Glu Gln Ala Val Ile Asp Leu Ile Lys Glu Ala Asp Phe Lys Ile Glu		205
210	215	220

[0005]

Leu Glu Ile Gln Thr Phe Asp Lys
225 230

<210> 4
<211> 186
<212> PRT
<213> 人工序列

<220>
<223> pPICHOLI-1_MFa1pha1Hsa-Fn10-Norpa

<400> 4
Met Lys Trp Val Ser Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15
Tyr Ser Arg Ser Leu Asp Lys Arg Glu Asn Leu Tyr Phe Gln Gly Gly
20 25 30
Ser Val Ser Asp Val Pro Arg Asp Leu Glu Val Val Ala Ala Thr Pro
35 40 45
Thr Ser Leu Leu Ile Ser Trp Asp Ala Pro Ala Val Thr Val Arg Tyr
50 55 60
Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val Gln Glu
65 70 75 80
Phe Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys
85 90 95
Pro Gly Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly
100 105 110
Asp Ser Pro Ala Ser Ser Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu
115 120 125
Phe Glu Asn Leu Tyr Phe Gln Gly Ser Gly Gly Gly Gly Glu Gln Lys
130 135 140
Leu Ile Ser Glu Glu Asp Leu His His His His His His Pro Ser Thr
145 150 155 160
Pro Pro Thr Pro Ser Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Tyr
165 170 175
Lys Thr Gln Gly Lys Thr Glu Phe Cys Ala
180 185

<210> 5
<211> 186
<212> PRT
<213> 人工序列

[0006]

<220>

<223> pPICHOLI-C Mfalpha1Hsa-Fn10-NorpA

<400> 5

```

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

```

<210> 6

<211> 232

<212> PRT

<213> 人工序列

<220>

<223> pYD_NBC1_Aga2-InaD

<400> 6

```

Met Gln Leu Leu Arg Cys Phe Ser Ile Phe Ser Val Ile Ala Ser Val
 1           5           10           15
Leu Ala Gln Glu Leu Thr Thr Ile Cys Glu Gln Ile Pro Ser Pro Thr
      20           25           30
Leu Glu Ser Thr Pro Tyr Ser Leu Ser Thr Thr Thr Ile Leu Ala Asn

```

[0007]

```

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

```

<210> 7

<211> 186

<212> PRT

<213> 人工序列

<220>

<223> pYS6/CT_HSA_MFalpha1_Fn10_Norpa

<400> 7

```

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

```

[0008]

Tyr	Arg	Ile	Thr	Tyr	Gly	Glu	Thr	Gly	Gly	Asn	Ser	Pro	Val	Gln	Glu
65					70					75					80
Phe	Thr	Val	Pro	Gly	Ser	Lys	Ser	Thr	Ala	Thr	Ile	Ser	Gly	Leu	Lys
				85					90					95	
Pro	Gly	Val	Asp	Tyr	Thr	Ile	Thr	Val	Tyr	Ala	Val	Thr	Gly	Arg	Gly
				100				105					110		
Asp	Ser	Pro	Ala	Ser	Ser	Lys	Pro	Ile	Ser	Ile	Asn	Tyr	Arg	Thr	Glu
				115				120					125		
Phe	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Ser	Gly	Gly	Gly	Gly	Glu	Gln	Lys
				130				135					140		
Leu	Ile	Ser	Glu	Glu	Asp	Leu	His	His	His	His	His	His	Pro	Ser	Thr
145						150					155				160
Pro	Pro	Thr	Pro	Ser	Pro	Ser	Thr	Pro	Pro	Thr	Pro	Ser	Pro	Ser	Tyr
						165				170					175
Lys	Thr	Gln	Gly	Lys	Thr	Glu	Phe	Cys	Ala						
				180					185						

<210> 8

<211> 97

<212> PRT

<213> 未知

<220>

<223> InaD PDZ 结构域氨基酸序列

<400> 8

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

<210> 9

[0009]

<211> 11
 <212> PRT
 <213> 未知

<220>
 <223> NorpA C-末端

<400> 9
 Tyr Lys Thr Gln Gly Lys Thr Glu Phe Cys Ala
 1 5 10

<210> 10
 <211> 737
 <212> PRT
 <213> 人工序列

<220>
 <223> pYS6/CT* MFalpha HSA-NorpA

<400> 10
 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15
 Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
 20 25 30
 Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45
 Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60
 Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80
 Ser Leu Glu Lys Arg Glu Ala Glu Ala Ala Ser Asp Ala His Lys Ser
 85 90 95
 Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala
 100 105 110
 Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu
 115 120 125
 Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys
 130 135 140
 Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu
 145 150 155 160
 Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly
 165 170 175
 Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys

[0010]

	180		185		190										
Phe	Leu	Gln	His	Lys	Asp	Asp	Asn	Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg
	195						200				205				
Pro	Glu	Val	Asp	Val	Met	Cys	Thr	Ala	Phe	His	Asp	Asn	Glu	Glu	Thr
	210						215				220				
Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu	Ile	Ala	Arg	Arg	His	Pro	Tyr	Phe
225					230				235				240		
Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala	Lys	Arg	Tyr	Lys	Ala	Ala	Phe
			245					250					255		
Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	Cys	Leu	Leu	Pro	Lys
	260						265				270				
Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	Ser	Ala	Lys	Gln	Arg
	275						280				285				
Leu	Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	Arg	Ala	Phe	Lys	Ala
290					295						300				
Trp	Ala	Val	Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	Lys	Ala	Glu	Phe	Ala
305					310				315				320		
Glu	Val	Ser	Lys	Leu	Val	Thr	Asp	Leu	Thr	Lys	Val	His	Thr	Glu	Cys
			325						330				335		
Cys	His	Gly	Asp	Leu	Leu	Glu	Cys	Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala
	340						345				350				
Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser	Ile	Ser	Ser	Lys	Leu	Lys	Glu
	355						360				365				
Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His	Cys	Ile	Ala	Glu	Val
370					375						380				
Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala	Ala	Asp	Phe
385					390				395				400		
Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala	Lys	Asp	Val
			405						410				415		
Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr
	420						425					430			
Ser	Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu
	435						440				445				
Glu	Lys	Cys	Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr	Ala	Lys	Val
450					455						460				
Phe	Asp	Glu	Phe	Lys	Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys
465					470				475				480		
Gln	Asn	Cys	Glu	Leu	Phe	Glu	Gln	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn
			485						490				495		
Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys	Val	Pro	Gln	Val	Ser	Thr	Pro
	500						505				510				
Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys	Val	Gly	Ser	Lys	Cys
	515						520				525				
Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu	Asp	Tyr	Leu

[0011]

530	535	540
Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val		
545	550	555
Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg		
565	570	575
Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu		
580	585	590
Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser		
595	600	605
Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val		
610	615	620
Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp		
625	630	635
Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu		
645	650	655
Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala		
660	665	670
Ala Leu Gly Leu Gly Ser Glu Asn Leu Tyr Phe Gln Gly Ser Gly Gly		
675	680	685
Gly Gly Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu His His His His		
690	695	700
His His His His Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Thr Pro		
705	710	715
Pro Thr Pro Ser Pro Ser Tyr Lys Thr Gln Gly Lys Thr Glu Phe Cys		
725	730	735
Ala		

<210> 11

<211> 405

<212> PRT

<213> 人工序列

<220>

<223> pYS6/CT* MFalpha1-scFv lysozyme-Norpa

<400> 11

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

[0012]

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu		
50	55	60
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val		
65	70	75 80
Ser Leu Glu Lys Arg Glu Ala Glu Ala Ala Ser Gln Val Lys Leu Gln		
85	90	95
Gln Ser Gly Ala Glu Leu Val Lys Pro Gly Ala Ser Val Lys Leu Ser		
100	105	110
Cys Thr Ala Ser Gly Phe Asn Ile Lys Asp Thr Tyr Met His Trp Val		
115	120	125
Lys Gln Arg Pro Glu Gln Gly Leu Glu Trp Ile Gly Arg Ile Asp Pro		
130	135	140
Ala Asn Gly Asn Thr Lys Tyr Asp Pro Lys Phe Gln Gly Lys Ala Thr		
145	150	155 160
Ile Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Leu Gln Leu Ser Ser		
165	170	175
Leu Thr Ser Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Trp Asp Trp		
180	185	190
Tyr Phe Asp Val Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser Gly		
195	200	205
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile		
210	215	220
Glu Leu Thr Gln Ser Pro Ser Ser Met Tyr Thr Ser Leu Gly Glu Arg		
225	230	235 240
Val Thr Ile Thr Cys Lys Ala Ser Gln Asp Ile Asn Ser Tyr Leu Arg		
245	250	255
Trp Phe Gln Gln Lys Pro Gly Lys Ser Pro Lys Thr Leu Ile Tyr Tyr		
260	265	270
Ala Thr Ser Leu Ala Asp Gly Val Pro Ser Arg Phe Ser Gly Ser Gly		
275	280	285
Ser Gly Gln Asp Tyr Ser Leu Thr Ile Ser Ser Leu Glu Ser Asp Asp		
290	295	300
Thr Thr Thr Tyr Tyr Cys Leu Gln His Gly Glu Ser Pro Tyr Thr Phe		
305	310	315 320
Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Ala Ala Glu Gln Lys		
325	330	335
Leu Ile Ser Glu Glu Asp Leu Asn Gly Ser Glu Asn Leu Tyr Phe Gln		
340	345	350
Gly Ser Gly Gly Gly Gly Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu		
355	360	365
His His His His His His His His Pro Ser Thr Pro Pro Thr Pro Ser		
370	375	380
Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Tyr Lys Thr Gln Gly Lys		
385	390	395 400

[0013]

Thr Glu Phe Cys Ala
405

<210> 12

<211> 114

<212> PRT

<213> 人工序列

<220>

<223> pYD_NBC1_Aga2-NorpA

<400> 12

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

<210> 13

<211> 327

<212> PRT

<213> 人工序列

<220>

<223> pYS6/CT* MFalpha1-InaD-Fn10

<400> 13

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

[0014]

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

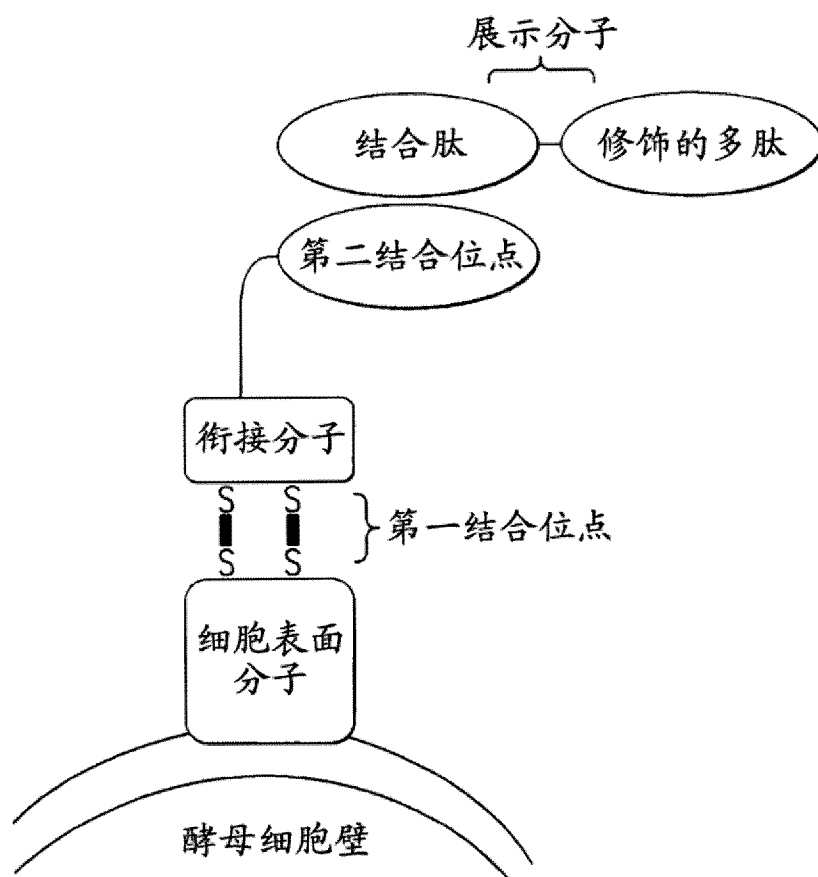


图 1

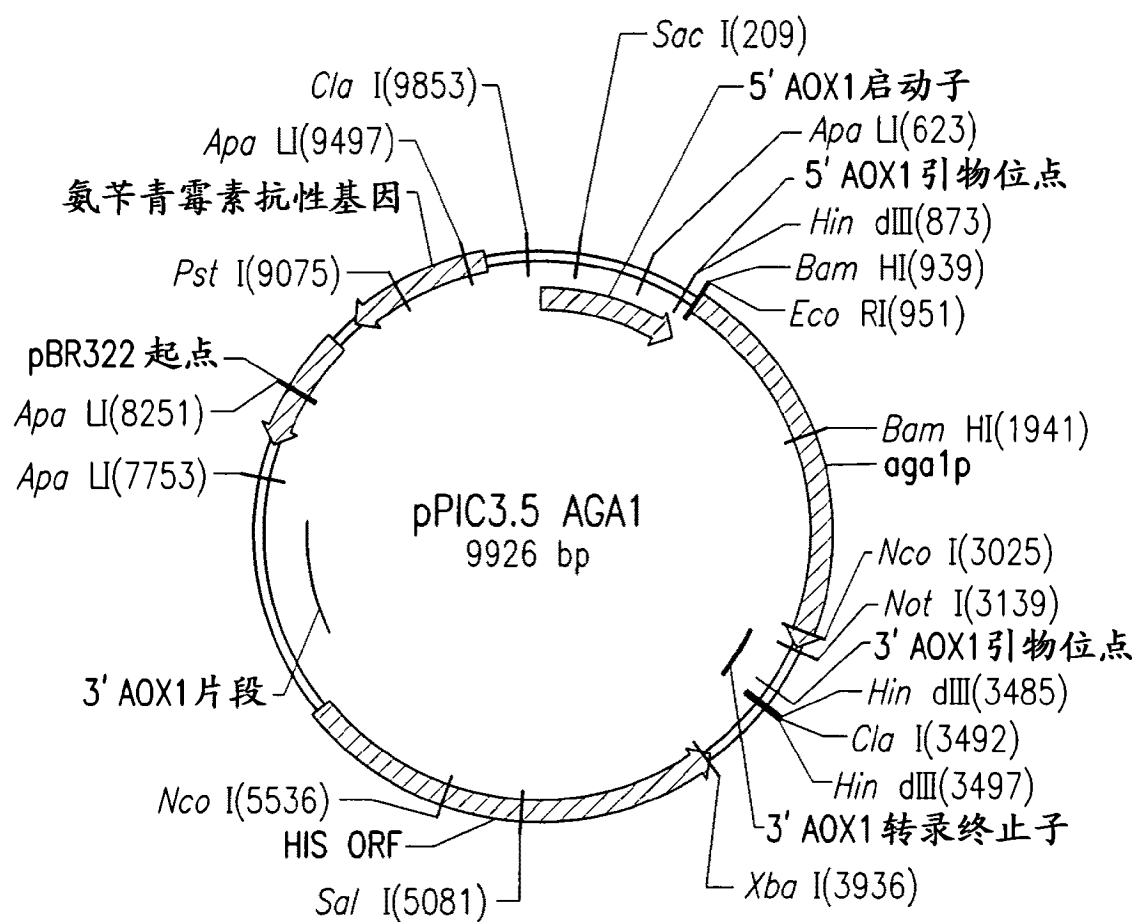


图 2

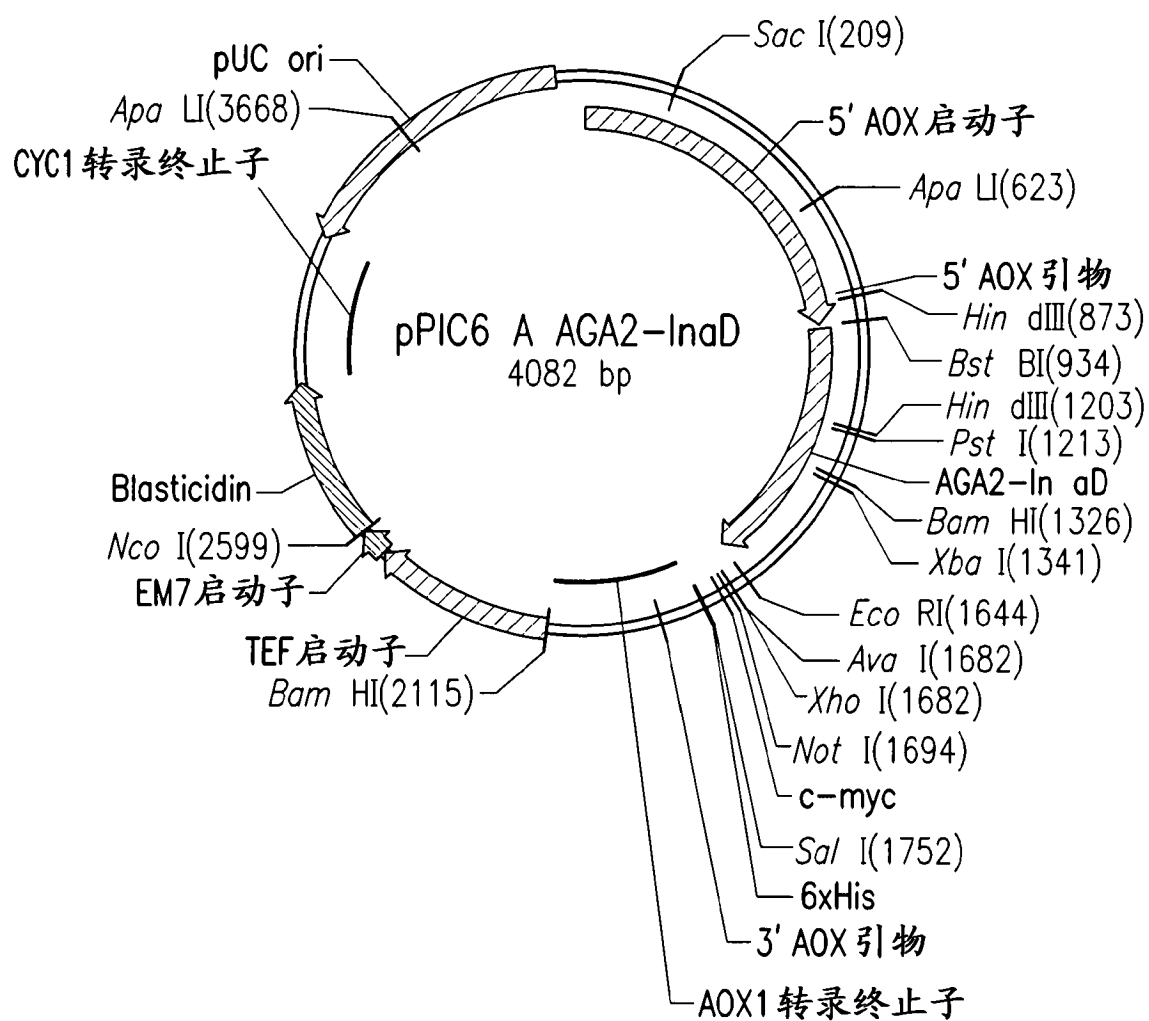


图 3

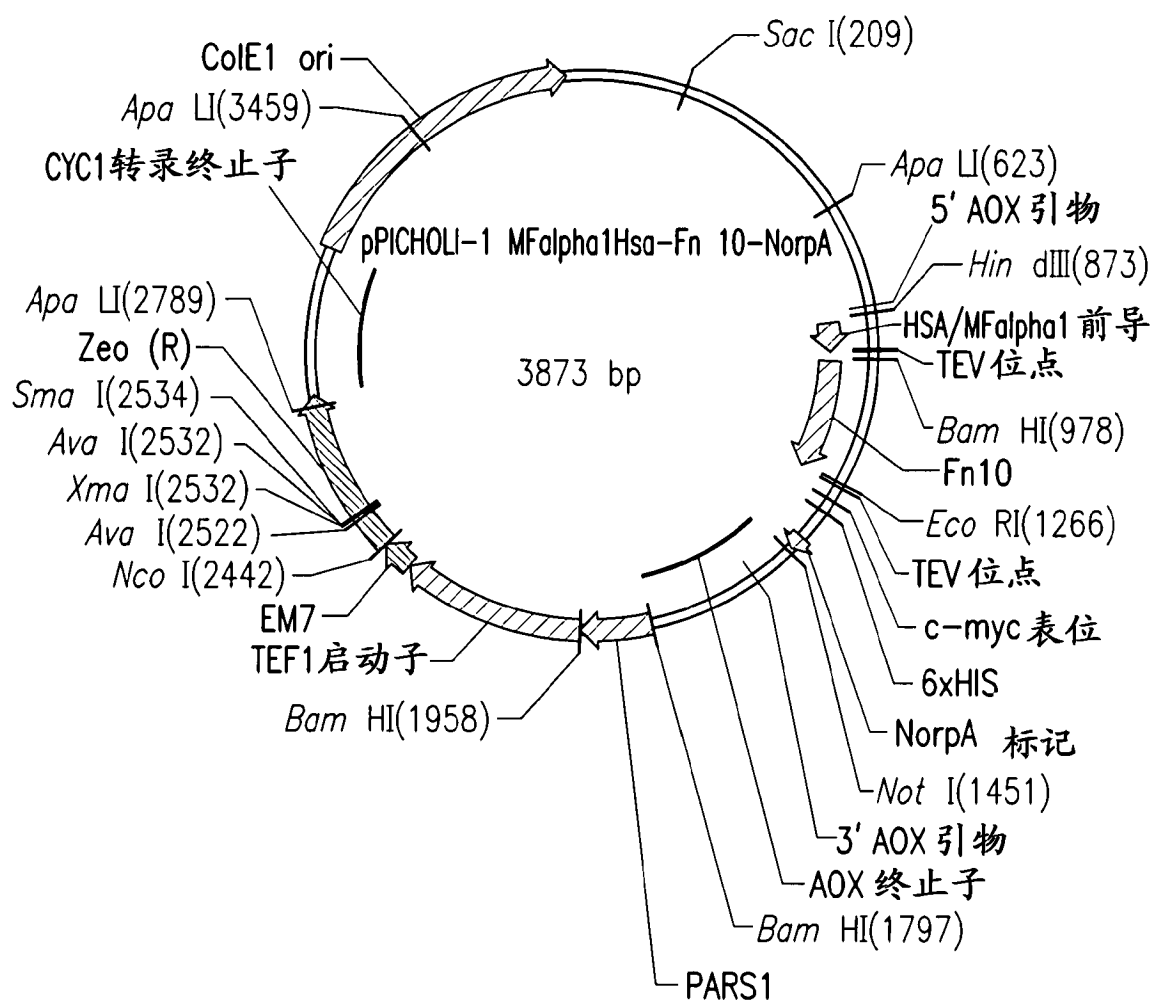


图 4

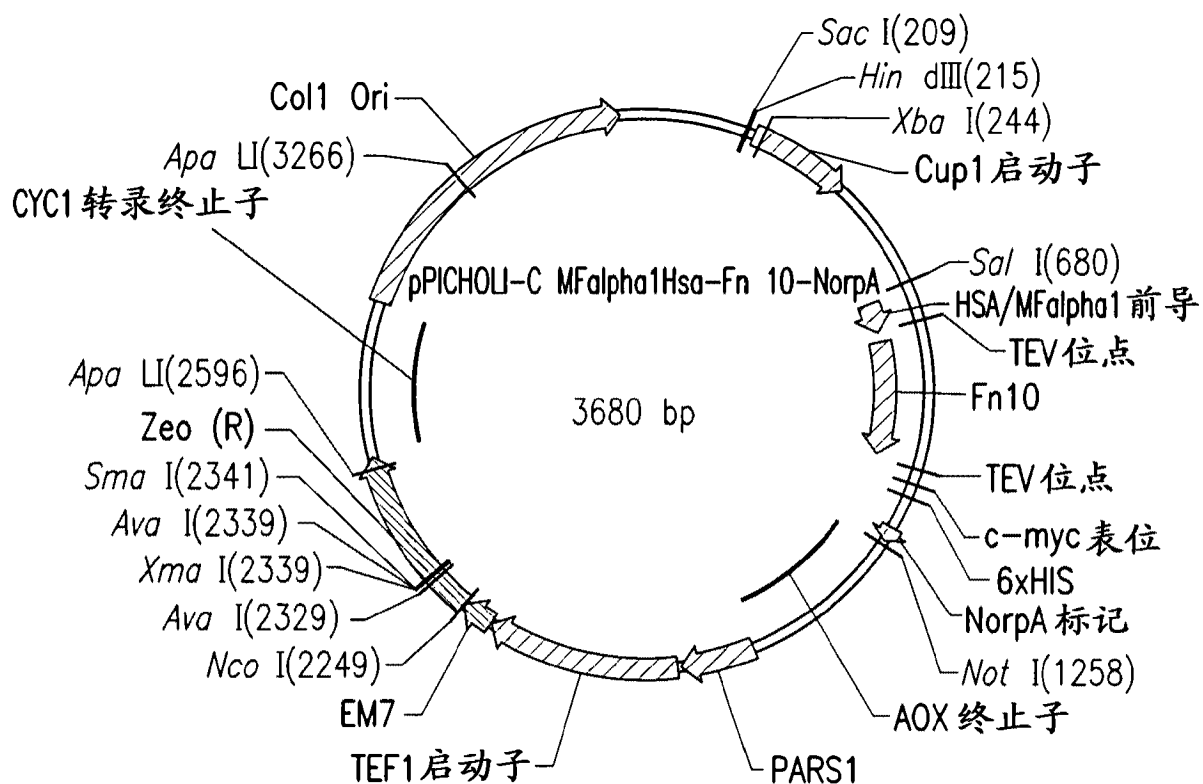


图 5

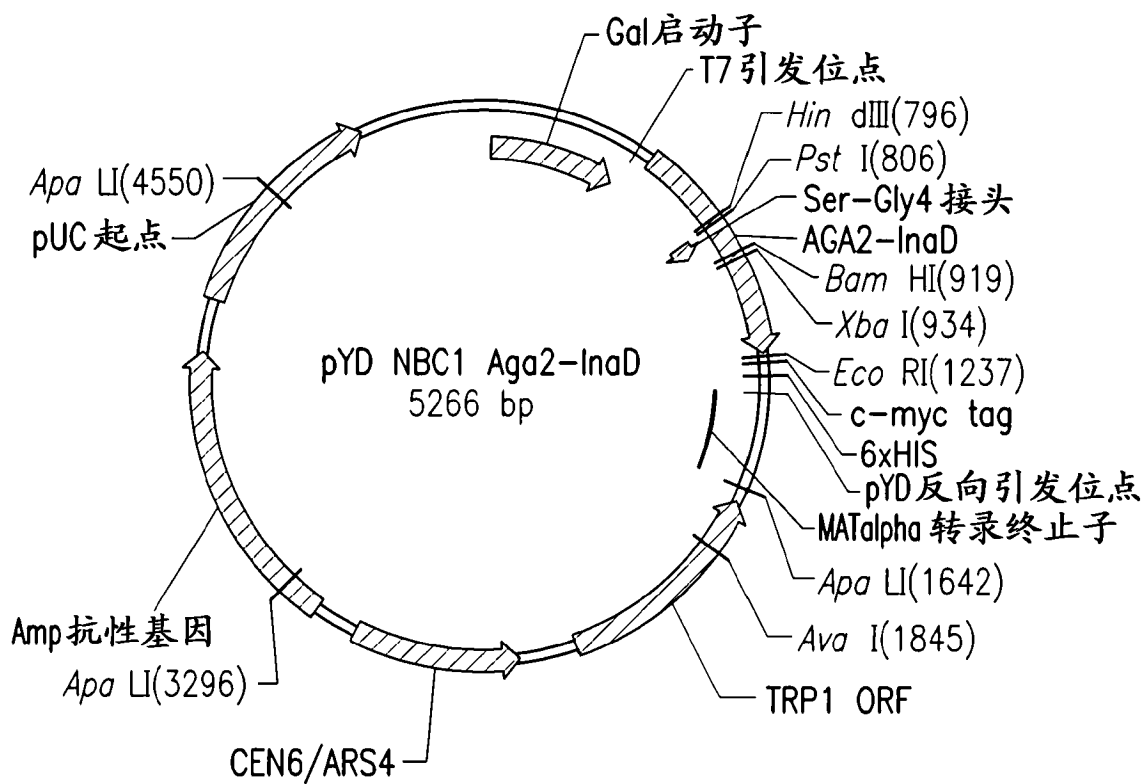


图 6

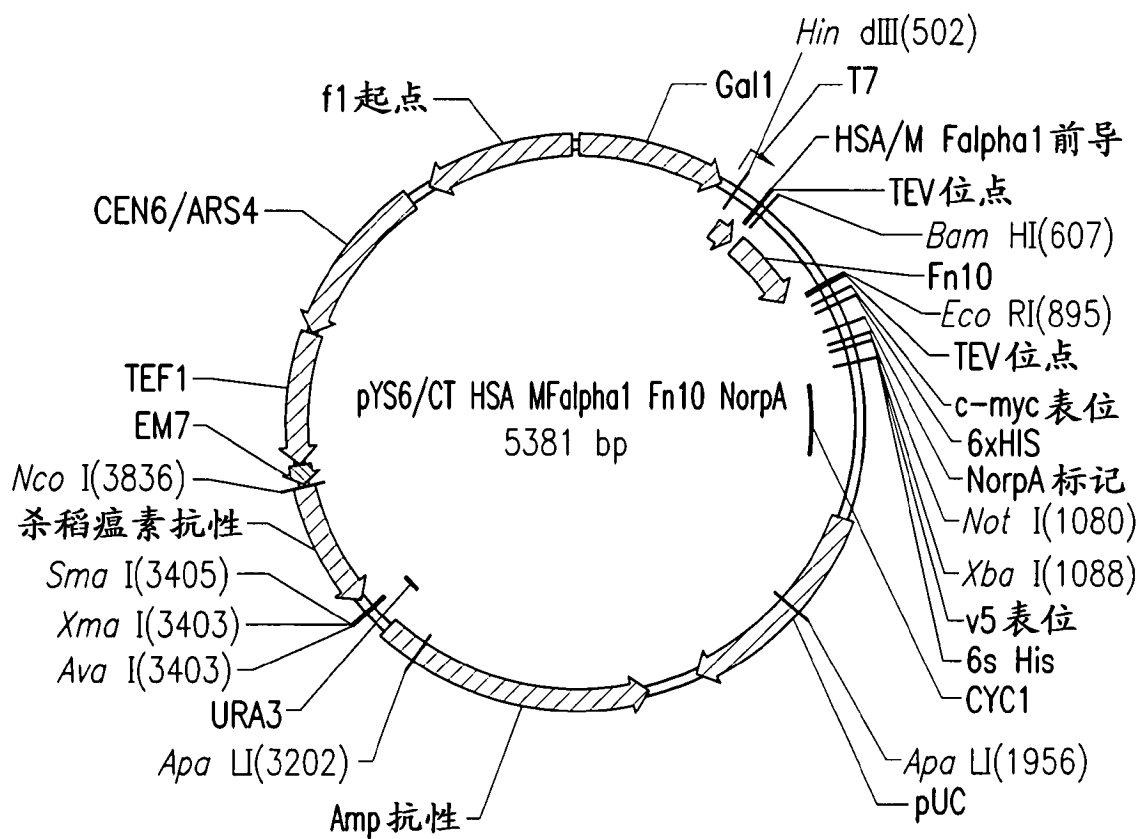


图 7

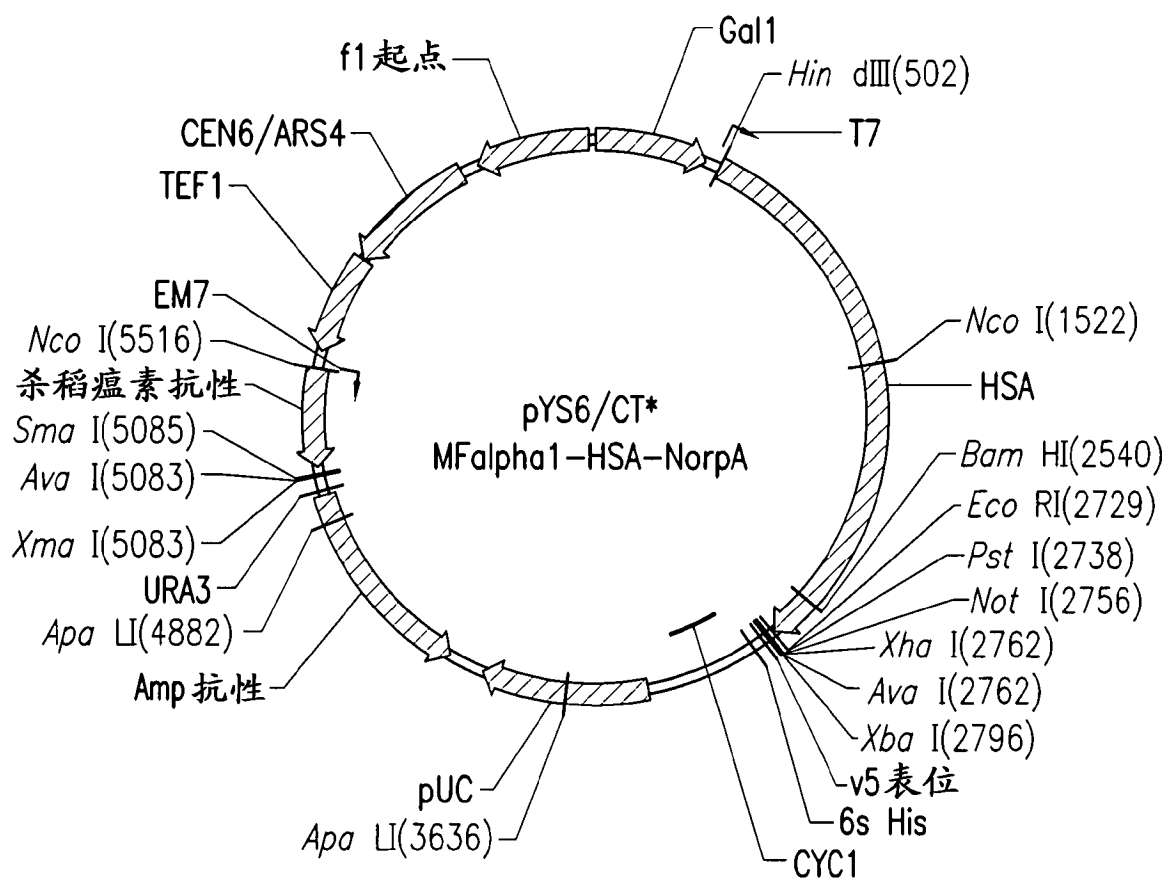


图 8

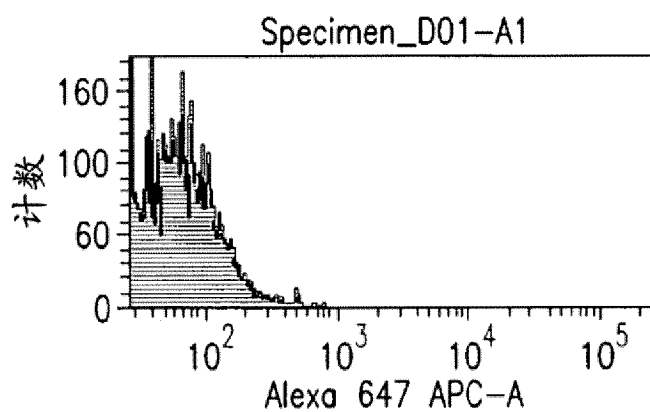


图 9A

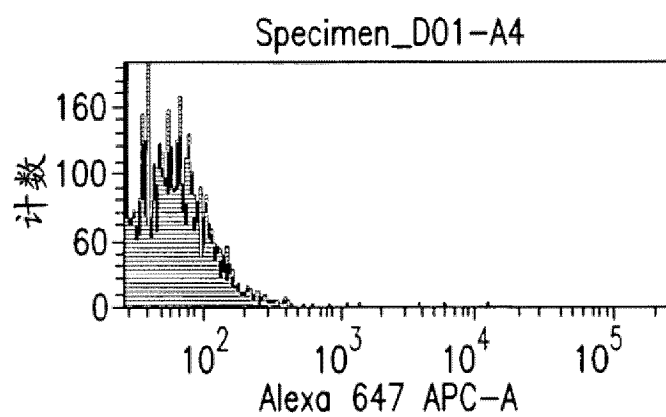


图 9B

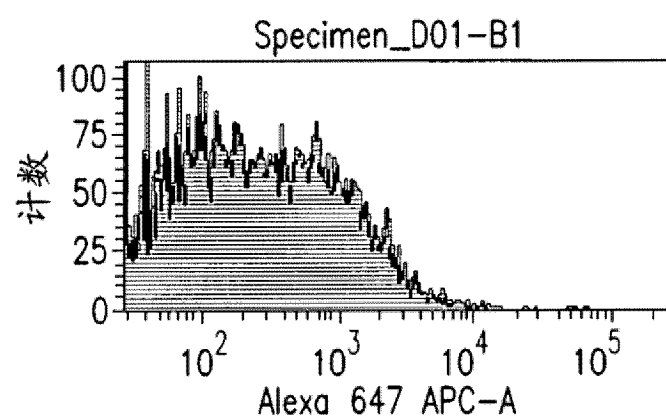


图 9C

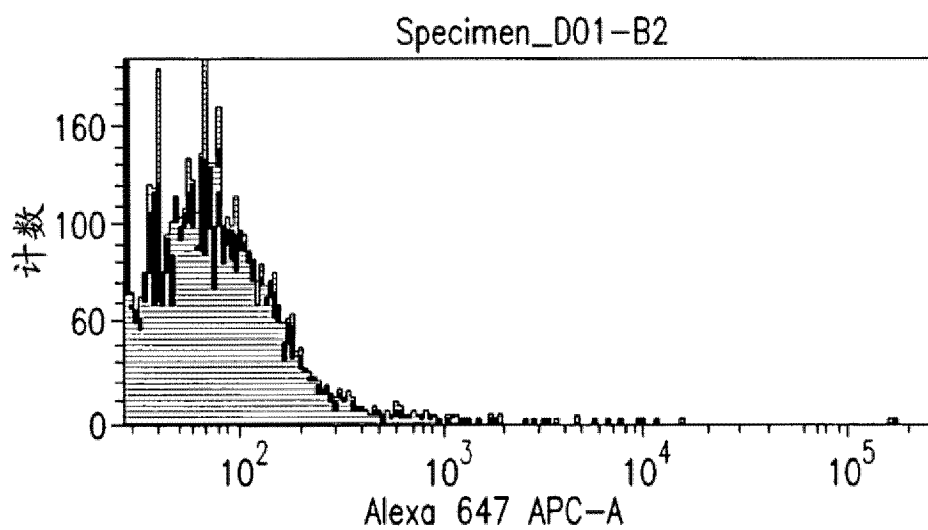


图 9D

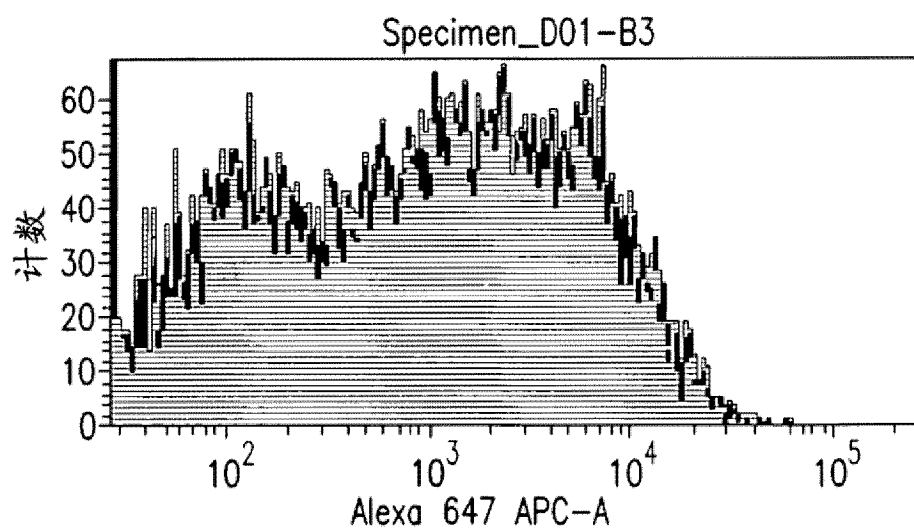


图 9E



图 10A

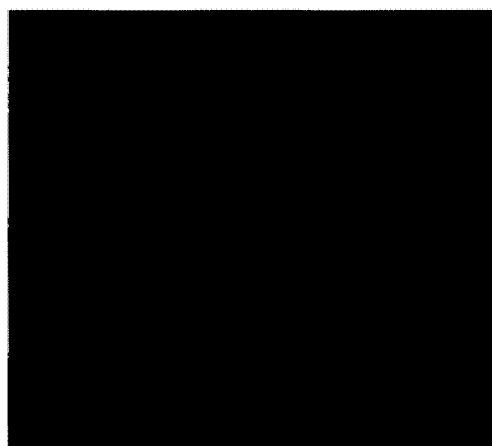


图 10B

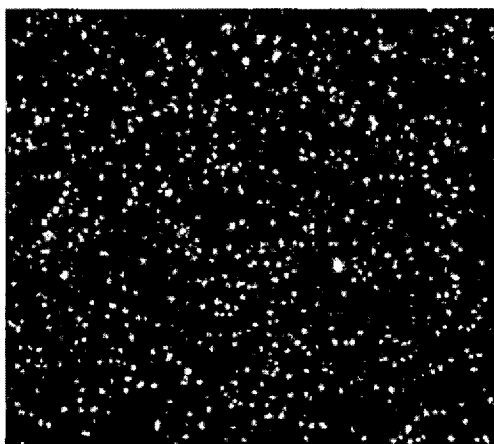


图 10C

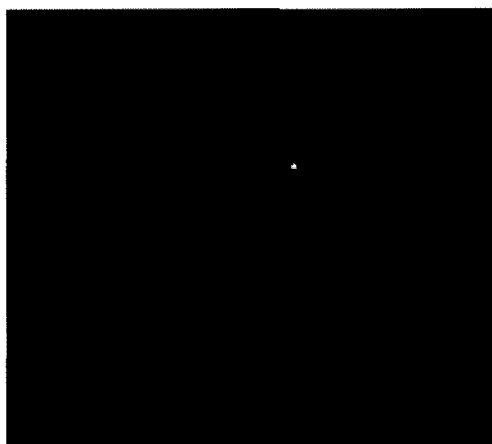


图 10D

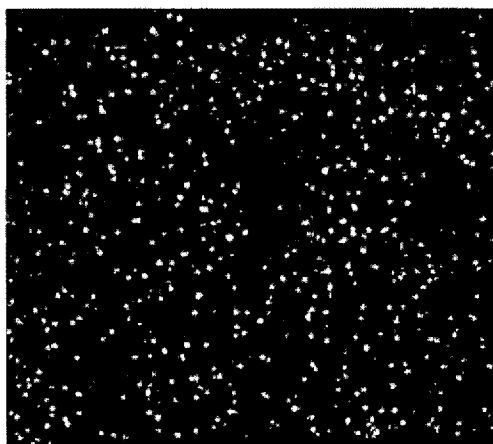


图 10E

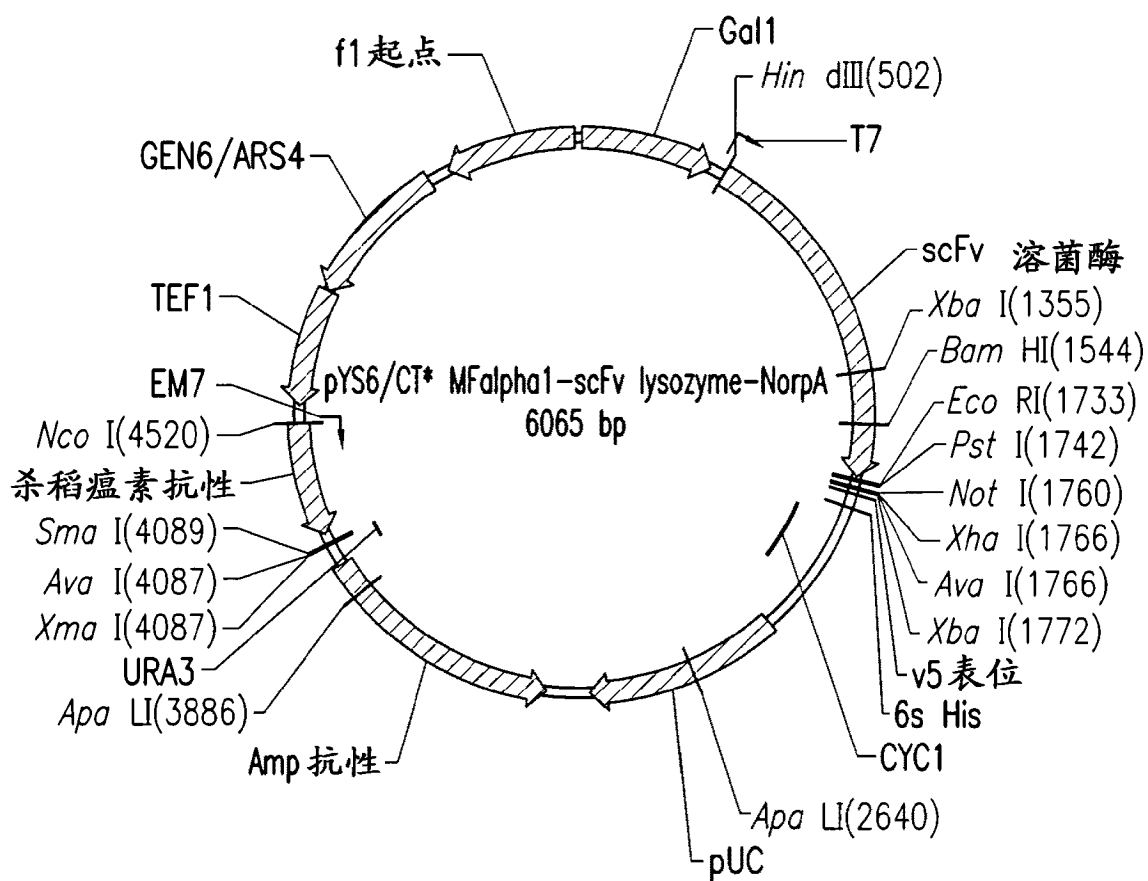


图 11

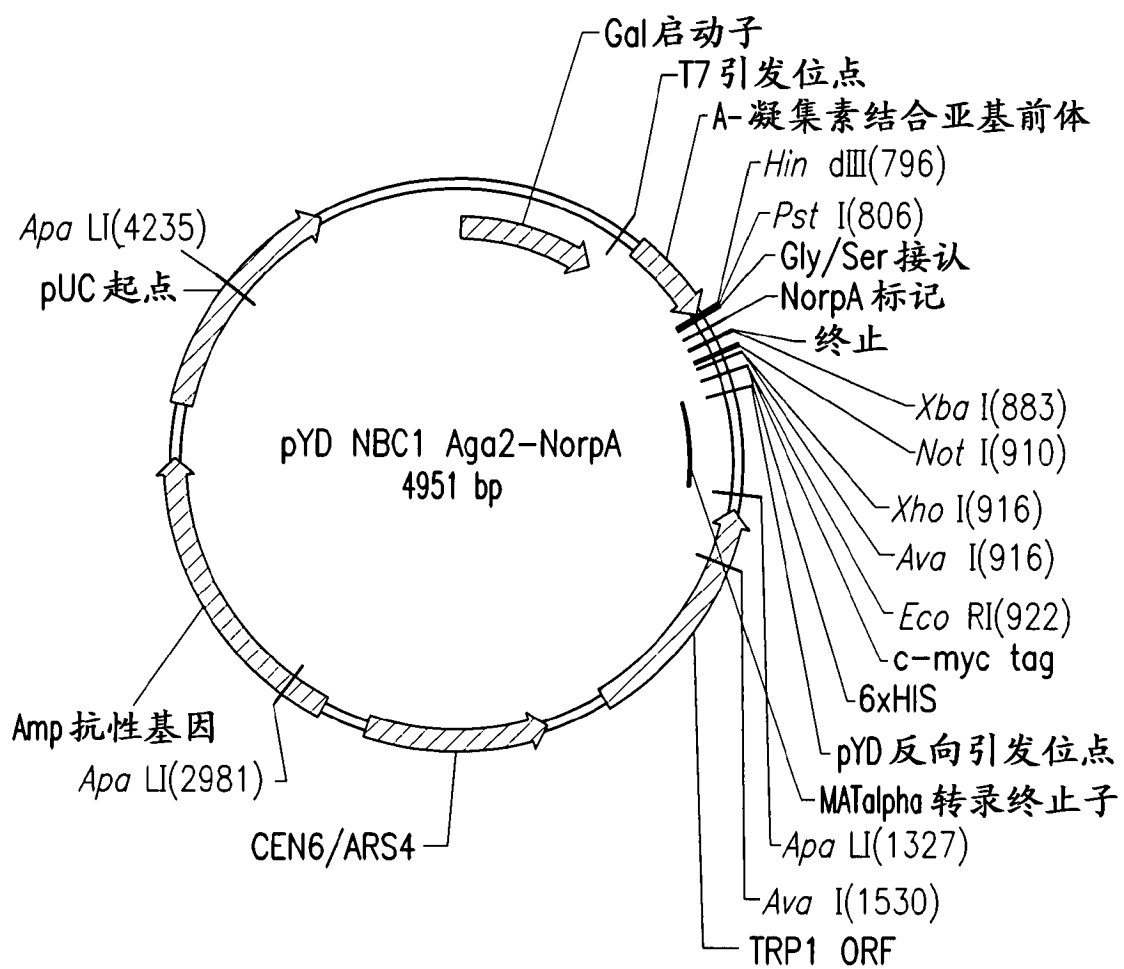


图 12

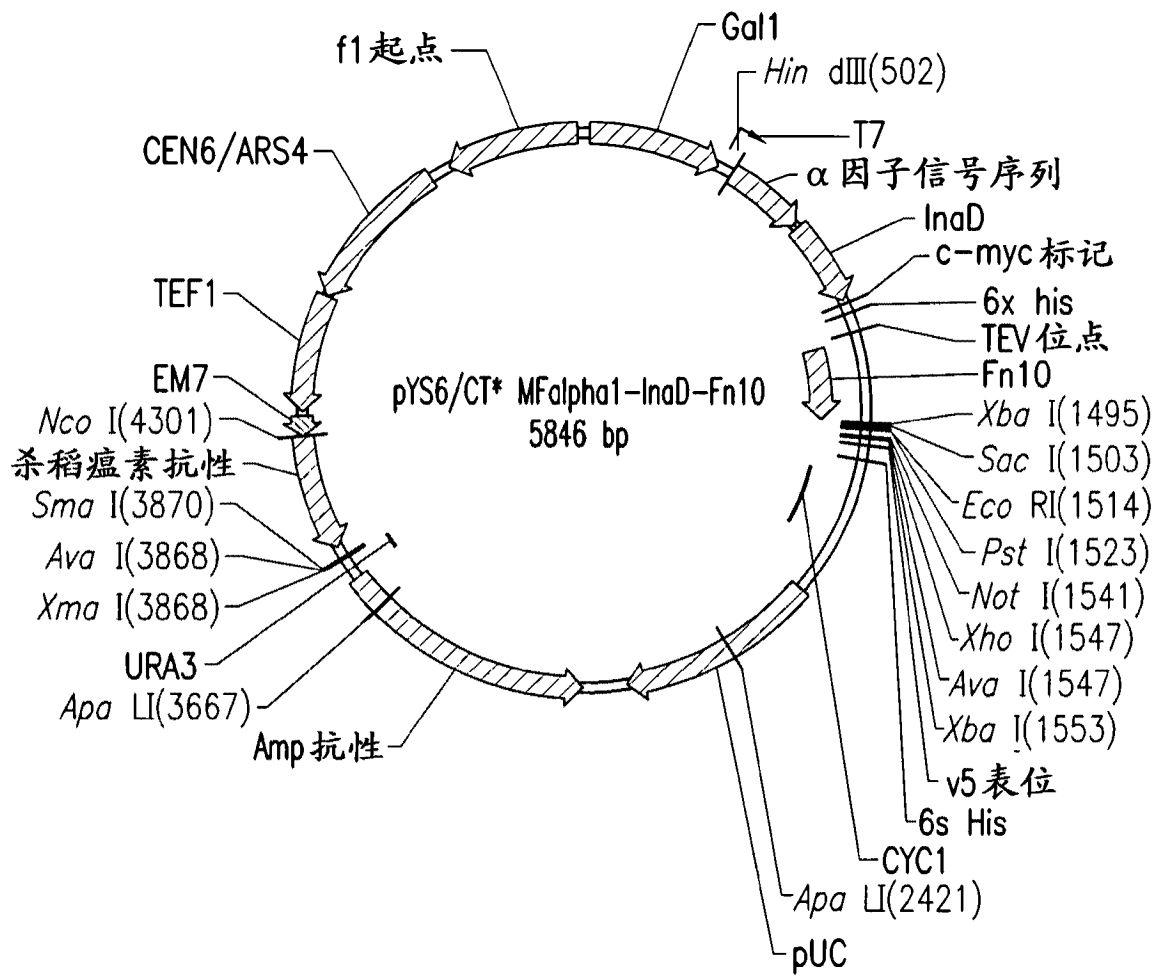


图 13