



# (12) 发明专利申请

(10) 申请公布号 CN 102712694 A

(43) 申请公布日 2012. 10. 03

(21) 申请号 201180005972. 4

C12N 15/70(2006. 01)

(22) 申请日 2011. 01. 13

C12P 21/02(2006. 01)

(30) 优先权数据

1000590. 8 2010. 01. 14 GB

(85) PCT申请进入国家阶段日

2012. 07. 13

(86) PCT申请的申请数据

PCT/EP2011/050415 2011. 01. 13

(87) PCT申请的公布数据

WO2011/086138 EN 2011. 07. 21

(71) 申请人 UCB 医药有限公司

地址 比利时布鲁塞尔

(72) 发明人 M·埃里斯 D·P·哈姆费雷斯

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 袁泉

(51) Int. Cl.

C07K 16/24(2006. 01)

C12N 9/48(2006. 01)

权利要求书 2 页 说明书 37 页

序列表 27 页 附图 15 页

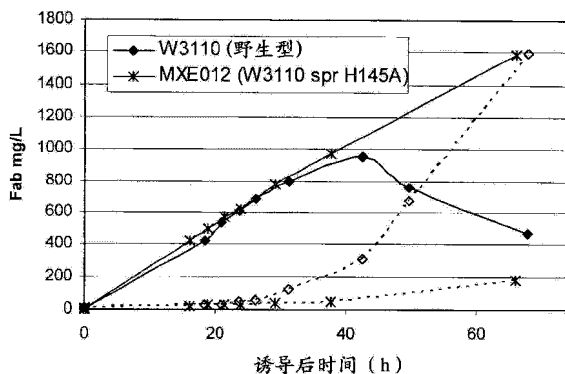
(54) 发明名称

包含突变 SPR 基因以及野生型 TSP 基因的细菌宿主菌株

(57) 摘要

本发明提供了重组革兰氏阴性细菌细胞, 其包含编码突变 SPR 蛋白质的突变 SPR 基因, 且其中的细胞包含非重组野生型染色体 Tsp 基因。

W3110 (野生型)对比 MXE012 (W3110 spr H145A)



1. 重组革兰氏阴性细胞,其包含编码突变 spr 蛋白质的突变 spr 基因,并且其中细胞包含非重组野生型染色体 Tsp 基因。
2. 根据权利要求 1 的细胞,其中所述突变 spr 基因编码的 spr 蛋白质在选自 H145、N31、R62、I70、Q73、C94、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144、G147、H157 和 W174 的一个或多个氨基酸处具有突变。
3. 根据权利要求 2 的细胞,其中所述突变 spr 基因编码的 spr 蛋白质具有选自 H145A、N31Y、R62C、I70T、Q73R、C94A、S95F、V98E、Q99P、R100G、L108S、Y115F、D133A、V135D、V135G、L136P、G140C、R144C、G147C、H157A 和 W174R 的一个或多个突变。
4. 根据权利要求 3 的细胞,其中所述 spr 蛋白质的一个或多个突变选自 S95F、V98E、Y115F、D133A、V135D、V135G 和 G147C。
5. 根据权利要求 4 的细胞,其中所述突变 spr 基因编码的 spr 蛋白质具有突变 S95F 和 Y115F。
6. 根据权利要求 3 的细胞,其中所述 spr 蛋白质突变为 H145A。
7. 根据任何前述权利要求的细胞,其中除了突变 spr 基因外,所述细胞与野生型细菌细胞是同基因的。
8. 根据任何前述权利要求的细胞,其中所述细胞进一步包含编码 DsbC 的重组多核苷酸。
9. 根据权利要求 1-8 中任何一项的细胞,其中所述细胞进一步包含一种或多种以下突变基因:
  - a. 编码具有分子伴侣活性和降低的蛋白酶活性的 DegP 蛋白质的突变 DegP 基因;
  - b. 突变 ptr 基因,其中所述突变 ptr 基因编码具有降低的蛋白酶活性的蛋白酶 III 蛋白质或其为敲除突变 ptr 基因;以及
  - c. 突变 OmpT 基因,其中所述突变 OmpT 基因编码具有降低的蛋白酶活性的 OmpT 蛋白质或其为敲除突变 OmpT 基因。
10. 根据任何前述权利要求的细胞,其中所述细胞为大肠杆菌。
11. 根据任何前述权利要求的细胞,其中所述细胞包含编码目的蛋白质的多核苷酸序列。
12. 根据权利要求 11 的细胞,其中所述细胞包含的载体含有编码 DsbC 的重组多核苷酸以及编码目的蛋白质的多核苷酸序列。
13. 根据权利要求 12 的细胞,其中所述载体包含控制编码 DsbC 的重组多核苷酸以及编码目的蛋白质的多核苷酸序列的表达的启动子。
14. 根据权利要求 11-13 中任何一项的细胞,其中所述目的蛋白质为抗体或其抗原结合片段。
15. 根据权利要求 14 的细胞,其中所述抗体或其抗原结合片段对 TNF 具有特异性。
16. 重组革兰氏阴性细菌细胞,其包含编码突变 spr 蛋白质的突变 spr 基因、野生型 Tsp 基因以及编码对 TNF 具有特异性的抗体或其抗原结合片段的多核苷酸序列。
17. 根据权利要求 16 的细胞,其中所述细胞包含编码 DsbC 的重组多核苷酸。
18. 用于生产目的重组蛋白质的方法,包括于培养基中在能够有效表达目的重组蛋白质的条件下,培养权利要求 1-17 中任何一项限定的重组革兰氏阴性细菌细胞,以及从所述

重组革兰氏阴性细菌细胞的细胞周质和 / 或培养基中回收目的重组蛋白质。

19. 根据权利要求 18 的方法,其中所述目的重组蛋白质回收自细胞周质和 / 或上清。

20. 根据权利要求 18 或 19 的方法,其中所述细胞包含编码 DsbC 的重组多核苷酸,且细胞培养于能够有效表达编码 DsbC 的重组多核苷酸的条件。

21. 根据权利要求 20 的方法,其中编码目的蛋白质的多核苷酸序列以及编码 DsbC 的重组多核苷酸的表达是通过向培养基中添加诱导子而进行诱导的。

22. 根据权利要求 20 或权利要求 21 的方法,其中所述方法进一步包括将目的重组蛋白质与 DsbC 相分离。

## 包含突变 SPR 基因以及野生型 TSP 基因的细菌宿主菌株

[0001] 本发明涉及重组细菌宿主菌株,特别是大肠杆菌(E. Coli)。本发明还涉及在此细胞中产生目的蛋白质的方法。

[0002] 发明背景

[0003] 通常将细菌细胞,如大肠杆菌,用于生产重组蛋白质。使用细菌细胞如大肠杆菌生产重组蛋白质有很多优势,具体而言是因为细菌细胞作为宿主细胞可以使得基因通过质粒插入而具有多样的属性。大肠杆菌已经用于很多重组蛋白质的生产,包括人胰岛素。

[0004] 尽管使用细菌细胞产生重组蛋白质有很多优势,但是仍然有显著的局限性,包括细菌细胞易于在目的重组蛋白质的表达过程中裂解。这种裂解表型可见于野生型细菌细胞,也可见于遗传工程改造的细胞,如细菌蛋白酶缺陷的细胞。蛋白酶在转换大肠杆菌细胞周质和细胞质中老旧、损坏或折叠错误的细胞上起到重要作用。细菌蛋白酶起降解目的重组蛋白质的作用,因此常常显著降低活性蛋白质的产量。因此,需要降低蛋白酶活性以降低目的蛋白质的蛋白质水解。然而,缺乏蛋白酶如 Tsp (也称为 Prc) 的细菌菌株也表现有细胞裂解。

[0005] Tsp (也称为 Prc) 为一种 60kDa 的细胞周质蛋白酶。Tsp (prc) 活性降低对降低目的蛋白质的蛋白质水解是可取的。然而,发现缺少蛋白酶 prc 的细胞在低渗透压下显示出热敏感生长概况。Hara 等人分离了 Tsp 缺陷菌株,其为耐热性回复突变株,含有基因外抑制子(spr)突变(Hara 等人, Microbial Drug Resistance, 2:63-72(1996))。Spr 为 18kDa 的膜结合细胞周质蛋白酶, spr 的底物为 Tsp 以及外膜中参与细胞分裂中细胞壁水解的肽聚糖。spr 基因标记为 UniProtKB/Swiss-Prot P0AFV4 (SPR\_ECOLI)。携带有突变 spr 基因的蛋白酶缺陷细菌菌株已在 Chen 等人(Chen C, Snedecor B, Nishihara JC, Joly JC, McFarland N, Andersen DC, Batters by JE, Champion KM. Biotechnol Bioeng. 2004 Mar 5; 85(5):463-74) 中有描述,其中描述了携带 prc (Tsp) 及另一种蛋白酶 DegP 不同突变组合的大肠杆菌菌株的构建,是通过扩增基因上游及下游区域并将其一起连接于包含选择性标记物及 spr W174R 突变体的载体上产生的。

[0006] 出乎意料地发现携带突变 spr 基因及野生型 Tsp 基因的革兰氏阴性细菌细胞为具有降低的裂解水平的细胞。因此,本发明人提供了新菌株,其具有用于产生目的蛋白质的优势特性。

[0007] 出乎意料地,根据本发明的细胞显示出优势的生长及蛋白质生产表型,因为已知 spr 和 Tsp 为相互抑制子,因此,可以预测如果让其中之一成为主导则细胞可以显示出生长不良的表型,如变得渗漏或更倾向发生细胞裂解。然而,与野生型细胞相比以及包含敲除突变 Tsp 基因的细胞,本发明的细胞的细胞裂解表型有显著的降低。

[0008] 发明概述

[0009] 本发明提供了重组的革兰氏阴性细菌细胞,其包含编码突变 spr 蛋白质的突变 spr 基因且其中细胞包含非重组野生型染色体 Tsp 基因。

[0010] 在一个实施方式中,除了突变 spr 基因外,根据本发明的细胞的基因组与野生型细菌细胞基因组是同基因的(isogenic)。

[0011] 本发明提供的细胞显示出优势的生长以及蛋白质生产表型。

[0012] 本发明还提供了用于生产目的重组蛋白质的方法,包括在如上文限定的重组革兰氏阴性细菌细胞中表达目的重组蛋白质。

[0013] 附图简述

[0014] 图 1 显示了相比于野生型 W3110 和 MXE001, MXE012 和 MXE017 的生长概况。

[0015] 图 2 显示了相比于野生型 W3110 和 MXE001, 在 MXE012 和 MXE017 中抗 -TNF  $\alpha$  Fab' 的表达情况。

[0016] 图 3 显示了在抗 -TNF  $\alpha$  Fab' 生产发酵期间, W3110 和 MXE012 的生长概况。

[0017] 图 4 显示了在抗 -TNF  $\alpha$  Fab' 生产发酵中 W3110 和 MXE012 (W3110spr H119A) 的细胞周质抗 -TNF  $\alpha$  Fab 的积累情况(实线及实心图标)以及培养基 Fab' 的积累情况(虚线及空心图标)。

[0018] 图 5 显示了表达抗 -TNF  $\alpha$  Fab' 的菌株 W3110 和 MXE012 以及表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 W3110 和 MXE012 的生长概况。

[0019] 图 6 显示了表达抗 -TNF  $\alpha$  Fab' 菌株 W3110 和 MXE012 以及表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 W3110 和 MXE012 的细胞周质(阴影图标)以及上清(空心非阴影图标)中的抗 -TNF  $\alpha$  Fab' 产量。

[0020] 图 7 显示了菌株 W3110、MXE001、MXE008 和 MXE012 的 dsDNA 测定的结果。

[0021] 图 8 显示了菌株 W3110、MXE001、MXE008 和 MXE012 的蛋白质测定的结果。

[0022] 图 9a 显示了野生型 ptr (蛋白酶 III) 和敲除突变的 ptr (蛋白酶 III) 的蛋白质及基因序列的 5' 末端。

[0023] 图 9b 显示了野生型 Tsp 和敲除突变的 Tsp 的蛋白质及基因序列的 5' 末端。

[0024] 图 9c 显示了野生型 DegP 及突变 DegP 蛋白质及基因序列的区域。

[0025] 图 10 显示了用于产生根据本发明实施方式的细胞的载体的构建。

[0026] 图 11 显示了表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 MXE012 的 200L 发酵物的生长概况。

[0027] 图 12 显示了表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 MXE012 的 200L 发酵物的抗 -TNF  $\alpha$  Fab' 滴度。

[0028] 图 13 显示了表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 MXE012 的 200L 发酵物的活力。

[0029] 图 14 显示了表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 MXE012 的 3000L 发酵物的生长概况。

[0030] 图 15 显示了显示了表达抗 -TNF  $\alpha$  Fab' 及重组 DsbC 的菌株 MXE012 的 3000L 发酵物的抗 -TNF  $\alpha$  Fab' 滴度。

[0031] 序列简述

[0032] SEQ ID NO:1 是野生型 Tsp 基因的 DNA 序列,包括起始密码子上游的 6 个核苷酸 ATGAAC。

[0033] SEQ ID NO:2 是野生型 Tsp 蛋白质的氨基酸序列。

[0034] SEQ ID NO:3 是突变的敲除 Tsp 基因的 DNA 序列,包括起始密码子上游的 6 个核苷酸 ATGAAT。

- [0035] SEQ ID NO:4 是野生型蛋白酶 III 基因的 DNA 序列。
- [0036] SEQ ID NO:5 是野生型蛋白酶 III 蛋白质的氨基酸序列。
- [0037] SEQ ID NO:6 是突变的敲除蛋白酶 III 基因的 DNA 序列
- [0038] SEQ ID NO:7 是野生型 DegP 基因的 DNA 序列。
- [0039] SEQ ID NO:8 是野生型 DegP 蛋白质的氨基酸序列。
- [0040] SEQ ID NO:9 是突变 DegP 基因的 DNA 序列。
- [0041] SEQ ID NO:10 是突变 DegP 蛋白质的氨基酸序列。
- [0042] SEQ ID NO:11 是抗 -TNF 抗体轻链可变区的氨基酸序列。
- [0043] SEQ ID NO:12 是抗 -TNF 抗体重链可变区的氨基酸序列。
- [0044] SEQ ID NO:13 是抗 -TNF 抗体轻链的氨基酸序列。
- [0045] SEQ ID NO:14 是抗 -TNF 抗体重链的氨基酸序列。
- [0046] SEQ ID NO:15 是用于突变 Tsp 基因区域的 3' 寡核苷酸引物的序列, 包含 Ase I 限制性位点。
- [0047] SEQ ID NO:16 是用于突变 Tsp 基因区域的 5' 寡核苷酸引物的序列, 包含 Ase I 限制性位点。
- [0048] SEQ ID NO:17 是用于突变蛋白酶 III 基因区域的 3' 寡核苷酸引物的序列, 包含 Ase I 限制性位点。
- [0049] SEQ ID NO:18 是用于突变蛋白酶 III 基因区域的 5' 寡核苷酸引物的序列, 包含 Ase I 限制性位点。
- [0050] SEQ ID NO:19 是用于突变 DegP 基因区域的 5' 寡核苷酸引物的序列, 包含 Ase I 限制性位点。
- [0051] SEQ ID NO:20 是用于突变 DegP 基因区域的 3' 寡核苷酸引物的序列, 包含 Ase I 限制性位点。
- [0052] SEQ ID NO:21 是野生型 spr 基因的序列, 包括信号序列, 即, 前 26 个氨基酸残基。
- [0053] SEQ ID NO:22 是非突变 spr 基因序列, 不包括信号序列。
- [0054] SEQ ID NO:23 是突变 OmpT 序列的核苷酸序列, 包括 D210A 和 H212A 突变。
- [0055] SEQ ID NO:24 是突变 OmpT 序列的氨基酸序列, 包括 D210A 和 H212A 突变。
- [0056] SEQ ID NO:25 是突变敲除 OmpT 序列的核苷酸序列。
- [0057] SEQ ID NO:26 是带 his 标签的 DsbC 的核苷酸序列。
- [0058] SEQ ID NO:27 是带 his 标签的 DsbC 的氨基酸序列。
- [0059] SEQ ID NO:28 显示了 hTNF40 的 CDRH1 氨基酸序列。
- [0060] SEQ ID NO:29 显示了 hTNF40 的 CDRH2 的氨基酸序列, 其为杂合 CDR, 其中 C-端六个氨基酸来自人亚群 3 种系抗体 (human subgroup3) 的 H2CDR 序列, 且由此杂合产生的序列中氨基酸变化由下划线标出, 如下 :WINTYIGEPI YADSVKG。
- [0061] SEQ ID NO:30 显示了 hTNF40 的 CDRH3 氨基酸序列。
- [0062] SEQ ID NO:31 显示了 hTNF40 的 CDRL1 氨基酸序列。
- [0063] SEQ ID NO:32 显示了 hTNF40 的 CDRL2 氨基酸序列。
- [0064] SEQ ID NO:33 显示了 hTNF40 的 CDRL3 氨基酸序列。
- [0065] SEQ ID NO:34 显示了 hTNF40 的 CDRH2 氨基酸序列。

[0066] SEQ ID NO:35 显示了 OmpA 寡核苷酸转接子的序列。

[0067] SEQ ID NO:36 显示了编码用于大肠杆菌 Fab 表达的基因间序列 1 (IGS1) 的寡核苷酸盒。

[0068] SEQ ID NO:37 显示了编码用于大肠杆菌 Fab 表达的基因间序列 2 (IGS2) 的寡核苷酸盒。

[0069] SEQ ID NO:38 显示了编码用于大肠杆菌 Fab 表达的基因间序列 3 (IGS3) 的寡核苷酸盒。

[0070] SEQ ID NO:39 显示了编码用于大肠杆菌 Fab 表达的基因间序列 4 (IGS4) 的寡核苷酸盒。

[0071] 本发明优选实施方式详述

[0072] 本发明提供了适合于表达目的蛋白质的重组的革兰氏阴性细菌细胞,其包含突变 spr 基因及非重组野生型染色体 Tsp 基因。

[0073] 出乎意料地发现与野生型细胞相比或包含突变 Tsp 基因的细胞,携带突变 spr 及非重组野生型染色体 Tsp 的细胞显示出细胞生长的改善,并显示出降低的细胞裂解表型。

[0074] 进一步地,在一个实施方式中,与野生型细胞相比或包含突变 Tsp 基因的细胞,携带突变 spr 及非重组野生型染色体 Tsp 的细胞显示出目的重组蛋白质产量的提高。提高的蛋白质产量可以是细胞周质蛋白质产量及 / 或上清蛋白质产量。在一个实施方式中,由于细胞渗漏减少,本发明的细胞与野生型细胞相比显示出提高的细胞周质蛋白质产量。由于细胞裂解减少,重组细菌细胞能够延长地表达目的重组蛋白质。

[0075] 根据本发明的细胞优选地在细胞周质及 / 或培养基中表达的目的蛋白质的最大产量为大约 1.0g/L、1.5g/L、1.8g/L、2.0g/L、2.4g/L、2.5g/L、3.0g/L、3.5g/L 或 4.0g/L。

[0076] 与已知的遗传工程改造菌株,如蛋白酶缺陷细菌菌株(先前产生此种菌株用以表达重组蛋白质)相关的缺点包括使用了参与细胞新陈代谢及 DNA 复制的基因突变,例如,大肠杆菌菌株中的 phoA、fhuA、lac、rec、gal、ara、arg、thi 和 pro。这些突变可以对宿主细胞具有很多有害作用,包括对细胞生长、稳定性、重组蛋白质表达产量及毒性的作用。具有一种或多种这些基因组突变的菌株,特别是具有大量这些突变的菌株可以丧失健康,这使得细菌生长速率降低到不适合工业化蛋白质生产的水平。进一步地,任何以上基因组突变可以以不可预测的有害方式顺式或反式影响其他基因并由此改变菌株表型、健康以及蛋白质概况。进一步地,使用严重突变的细胞一般不适合于生产用于商业化的重组蛋白质,尤其是用于治疗蛋白质,因为这些菌株一般具有缺陷的代谢途径,因此在基本或化学限定的培养基中生长不良或根本不生长。

[0077] 在本发明一个优选的实施方式中,细胞只携带能够引入 spr 突变体的基因组最小突变。在此实施方式中,细菌细胞的基因组与野生型细菌细胞基因组的差异只有对 spr 基因的一个或两个突变,且不携带任何其他可以对细胞生长和 / 或表达目的蛋白质能力有不利作用的突变。因此,相比于包含对基因组序列有进一步遗传工程改造突变的细胞,根据本发明的一种或多种重组宿主细胞展示了改进的蛋白质表达和 / 或改进的生长特征。相比于包含对细胞基因组的其他破坏的细胞,本发明提供的细胞还更适合于用于产生治疗用蛋白质。

[0078] 本发明还提供了重组革兰氏阴性细菌细胞,其含有编码突变 spr 蛋白质的突变

spr 基因,其中除了突变 spr 基因外细胞的基因组与野生型细菌细胞基因组是同基因的。在本发明的这个方面,所述细胞携带野生型 Tsp 基因。野生型染色体 Tsp 基因优选地为非重组染色体 Tsp 基因。

[0079] 技术人员可容易地使用本领域熟知的方法,包括发酵方法、ELISA 及蛋白质 G hplc,测试候选的细胞克隆以确定其是否具有目的蛋白质的所需要产量。合适的发酵方法描述于 Humphreys D P 等人 (1997)。Formation of dimeric Fabs in E.coli:effect of hinge size and isotype,presence of interchain disulphide bond,Fab' expression levels,tail piece sequences and growth conditions. J. IMMUNOL. METH. 209:193-202;Backlund E.Reeks D.Markland K.Weir N.Bowering L.Larsson G.Fedbatch design for periplasmic product retention in Escherichia coli, Journal Article. Research Support, Non-U. S. Gov't Journal of Biotechnology. 1 35(4):358-65, 2008 Jul 31;Champion KM.Nishihara JC. Joly JC.Arnott D. Similarity of the Escherichia coli proteome upon completion of different biopharmaceutical fermentation processes. [Journal Article]Proteomics. 1(9):1133-48, 2001Sep; 以及 Horn U.Strittmatter W.Krebber A.Knupfer U.Kujau M.Wenderoth R.Muller K.Matzku S.Pluckthun A.Riesenberg D.High volumetric yields of functional dimeric miniantibodies in Escherichia coli,using an optimized expression vector and high-cell-density fermentation under non-limited growth conditions, Journal Article. Research Support, Non-U. S. Gov't Applied Microbiology & Biotechnology. 46(5-6):524-32, 1996Dec。本领域技术人员还可容易地使用本领域熟知方法,如蛋白质 G HPLC、圆二色谱、NMR, X-光晶体成像及表位亲和力测量方法,测试分泌蛋白质以确定蛋白质是否正确折叠。

[0080] 在本发明一个优选的实施方式中,细胞进一步包含编码 DsbC 的重组多核苷酸。

[0081] 现在将以更详细的方式描述本发明。

[0082] 除非上下文另有说明,术语“蛋白质”和“多肽”在本文交换使用。“肽”意欲指代 10 个或更少的氨基酸。

[0083] 除非上下文另有说明,术语“多核苷酸”包括基因、DNA、cDNA、RNA、mRNA 等。

[0084] 如本文所使用的,术语“包含”在本说明书的上下文中应解释为“包括”。

[0085] 在本发明的上下文中,非突变的细胞或对照细胞指的是与本发明重组革兰氏阴性细胞相同类型的细胞,其中所述细胞未被修饰为携带突变 spr 基因。例如,非突变细胞可为野生型细胞,并可与本发明细胞在进行修饰引入任何所述突变之前衍生自相同的宿主细胞群体。

[0086] 词句“细胞”、“细胞系”、“细胞培养物”及“菌株”可交换使用。

[0087] 词句“包含突变 Tsp 基因的细胞的表型”在本发明上下文中指的是带有突变 Tsp 基因的细胞所表现出的表型。一般地,包含突变 Tsp 基因的细胞会裂解,尤其是在高细胞密度下。这些细胞的裂解引起任何重组蛋白质渗漏入上清。携带突变 Tsp 基因的细胞还可以显示出在低渗透压下的热敏型生长。例如,细胞显示不出生长速率或生长速率降低,或在高温(如 40°C 或更高)下,细胞在低渗培养基中死亡。

[0088] 术语“同基因的”在本发明的上下文中指的是与所述细胞来源的野生型细胞相比,



本发明细胞的基因组除了突变 SPR 基因外具有几乎相同或相同的基因组序列。在此实施方式中,根据本发明的细胞基因组不包含其他非天然发生的或遗传工程改造的突变。在一个实施方式中,考虑到任何可以发生的天然发生的突变,与野生型细胞相比,根据本发明的细胞除了突变 spr 基因外有几乎相同的基因组序列。在一个实施方式中,与野生型细胞相比,根据本发明的细胞除了突变 spr 基因外可以具有完全相同的基因组序列。

[0089] 在其中细胞包含编码 DsbC 的重组多核苷酸的本发明实施方式中,编码 DsbC 的重组多核苷酸可存在于转化入细胞及 / 或整合入宿主细胞基因组的合适的表达载体上。在其中编码 DsbC 的重组多核苷酸插入宿主基因组的实施方式中,由于插入的编码 DsbC 的重组多核苷酸,本发明的细胞也与野生型细胞不同。优选地,编码 DsbC 的重组多核苷酸是在细胞内表达载体中,因此对宿主细胞基因组造成破坏最小。

[0090] 术语“野生型”在本发明的上下文中指的是可以在自然中发生或可分离自环境而不携带任何遗传工程改造突变的革兰氏阴性细胞菌株。大肠杆菌野生型菌株的一个实例为 W3110,如 W3110K-12 菌株。

[0091] 任何合适的革兰氏阴性细菌都可用作生产本发明重组细胞的亲本细胞。合适的革兰氏阴性细菌包括鼠伤寒沙门氏菌(*Salmonella typhimurium*)、荧光假单胞菌(*Pseudomonas fluorescens*)、胡萝卜软腐欧文氏菌(*Erwinia carotovora*)、志贺杆菌(*Shigella*)、肺炎克雷伯菌(*Klebsiella pneumoniae*)、嗜肺军团菌(*Legionella pneumophila*)、铜绿假单胞菌(*Pseudomonas aeruginosa*)、鲍氏不动杆菌(*Acinetobacter baumannii*)及大肠杆菌。优选地,亲本细胞为大肠杆菌。大肠杆菌的任何合适菌株可用于本发明,但优选地所使用的为野生型 W3110 菌株,如 K-12W3110。

[0092] 在一个优选的实施方式中,除了突变 SPR 基因外细胞与野生型大肠杆菌细胞如 W3110 是同基因的。

[0093] 在一个实施方式中,本发明的细胞包含编码目的蛋白质的多核苷酸。在此实施方式中,编码目的蛋白质的多核苷酸可包含在转化入细胞和 / 或整合入宿主细胞基因组的合适的表达载体内。在编码目的蛋白质的多核苷酸插入宿主基因组的实施方式中,由于插入的编码目的蛋白质的多核苷酸,本发明的细胞与野生型细胞也不同。优选地,多核苷酸是在细胞内表达载体中,因此对宿主细胞基因组造成破坏最小。

[0094] 根据本发明的细胞携带野生型 Tsp 基因。在本发明的一个方面,细胞携带野生型非重组染色体 Tsp 基因。野生型非重组染色体 Tsp 基因指的是染色体 Tsp 基因,其不是使用重组 DNA 技术构建、产生或插入染色体的。

[0095] 如本文所使用的,“Tsp 基因”指的是编码蛋白酶 Tsp (也称为 Prc) 的基因,Tsp 是一种能够作用于青霉素结合蛋白质-3 (PBP3) 及噬菌体尾巴蛋白质的细胞周质蛋白酶。野生型 Tsp 基因的序列显示于 SEQ ID NO:1,野生型 Tsp 蛋白质序列显示于 SEQ ID NO:2。

[0096] Spr 蛋白质为 18kDa 的膜结合细胞周质蛋白酶,spr 的底物为 Tsp 及外膜中参与细胞分裂期间细胞壁水解的肽聚糖。

[0097] Spr 蛋白质的野生型氨基酸序列显示于 SEQ ID NO:21 (包含在 N-端的信号序列) 及 SEQ ID NO:22 (不含 26 个氨基酸的信号序列) (根据 UniProt 登记号 P0AFV4)。本发明中 Spr 蛋白质的氨基酸编号包括信号序列。因此,spr 蛋白质的 1 号氨基酸为显示于 SEQ ID NO:21 中的第一个氨基酸(Met)。

[0098] 突变 spr 基因优选地为细胞的染色体 spr 基因。

[0099] 突变 spr 基因编码能够抑制包含突变 Tsp 基因的细胞的表型的 spr 蛋白质。携带突变 Tsp 基因的细胞可具有良好的细胞生长速率,但是这些细胞的一个限制在于其易于裂解,尤其在高细胞密度下。因此,包含突变 Tsp 基因的细胞的表型为易于裂解,尤其在高的细胞密度下。携带突变 Tsp 基因的细胞还显示了低渗透压下的热敏型生长方式。然而,本发明的细胞所携带的 spr 突变在引入携带突变 Tsp 基因的细胞中时抑制了突变 Tsp 表型,因此,(尤其在高细胞密度下)细胞显示出裂解减少。本领域技术人员可容易地测量摇瓶或高细胞密度发酵技术中的细胞的生长表型。相比于携带 Tsp 突变体及野生型 spr 的细胞,从提高的生长速率及 / 或重组蛋白质生产(尤其是细胞周质中)上可见携带 spr 突变体及 Tsp 突变体的细胞展示出对细胞裂解表型的抑制。

[0100] 可对 spr 基因进行任何合适的突变或多种突变产生能够抑制包含突变 Tsp 基因的细胞的表型的 spr 蛋白质。该活性可由本领域技术人员通过产生携带突变 spr 基因及突变 Tsp 基因的细胞并将其表型与只携带突变 Tsp 基因的细胞相比较而进行测试。对 Tsp 基因进行的合适突变在下文有详细描述。除非另有说明,提到突变 Tsp 基因或编码 Tsp 的突变 Tsp 基因指的是编码具有降低蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因或者敲除突变的 Tsp 基因。

[0101] 词句“编码具有降低蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因”指的是相比于野生型非突变 Tsp 基因, Tsp 基因不具有完全的蛋白酶活性。突变 Tsp 基因编码的 Tsp 蛋白质具有野生型非突变 Tsp 蛋白质的 50% 或更低、40% 或更低、30% 或更低、20% 或更低、10% 或更低、或者 5% 或更低的蛋白酶活性。突变的 Tsp 基因可以编码不具有蛋白酶活性的 Tsp 蛋白质。细胞没有染色体 Tsp 缺陷,即, Tsp 基因序列未经删除或突变以阻止任何形式的 Tsp 蛋白质表达。

[0102] 任何合适的突变可引入 Tsp 基因以产生具有降低的蛋白酶活性的蛋白质。自革兰氏阴性细菌表达的 Tsp 蛋白质的蛋白酶活性可容易地由本领域技术人员通过任何本领域合适的方法进行测试,如描述于 Keiler 等人(Identification of Active Site Residues of the Tsp Protease\* THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 270, No. 48, Issue of December 1, pp. 28864 - 28868, 1995 Kenneth C. Keiler 和 Robert T. Sauer)的方法,其中测试了 Tsp 的蛋白酶活性。

[0103] Keiler 等人(见上文)已经报道 Tsp 的活性位点包含残基 S430、D441 和 K455,且残基 G375、G376、E433 和 T452 对 Tsp 的结构维持很重要。Keiler 等人(见上文)报道其发现突变的 Tsp 基因 S430A、D441A、K455A、K455H、K455R、G375A、G376A、E433A 和 T452A 不可检测到蛋白酶活性。进一步报道突变的 Tsp 基因 S430C 显示出野生型活性的大约 5-10%。因此,产生具有降低的蛋白酶活性的蛋白质的 Tsp 突变可包含突变,如对残基 S430、D441、K455、G375、G376、E433 和 T452 中的一种或多种进行错义突变。优选地,产生具有降低的蛋白酶活性的蛋白质的 Tsp 突变可包含突变,如对残基 S430、D441 和 K455 中的一种或多种进行错义突变。

[0104] 因此,突变的 Tsp 基因可包含:

[0105] • 对 S430 的突变; 或

[0106] • 对 D441 的突变; 或

[0107] • 对 K455 的突变；或

[0108] • 对 S430 和 D441 的突变；或

[0109] • 对 S430 和 K455 的突变；或

[0110] • 对 D441 和 K455 的突变；或

[0111] • 对 S430、D441 和 K455 的突变。

[0112] S430、D441、K455、G375、G376、E433 和 T452 中的一个或多个可突变为任何合适的氨基酸，产生蛋白质具有降低的蛋白酶活性。合适的突变的实例为 S430A、S430C、D441A、K455A、K455H、K455R、G375A、G376A、E433A 和 T452A。突变 Tsp 基因可包含对活性位点残基的一个、两个或三个突变，例如基因可包含：

[0113] • S430A 或 S430C；和 / 或

[0114] • D441A；和 / 或

[0115] • K455A 或 K455H 或 K455R。

[0116] 优选地，Tsp 基因包含点突变 S430A 或 S430C。

[0117] 词句“敲除突变 Tsp 基因”指的是所述基因包含一个或多个突变由此导致此基因编码的蛋白质不表达，造成细胞中由所述敲除突变基因编码的蛋白质缺陷。敲除基因可部分或全部地转录，但不翻译为编码的蛋白质。敲除突变的 Tsp 基因可以任何合适的方式，例如通过删除、插入、点突变、错义、无义及移码突变中一种或多种方式进行突变，使得蛋白质不表达。例如，基因可通过将外源 DNA 序列如抗生素抗性标记物插入基因编码序列而被敲除。

[0118] 突变的 Tsp 基因可包含基因起始密码子和 / 或位于基因起始密码子下游、基因终止密码子上游的一个或多个终止密码子的突变，由此阻止 Tsp 蛋白质表达。对起始密码子的突变可为起始密码子核苷酸中一个、两个或所有三个核苷酸的错义突变。备选地或另外地，起始密码子可通过插入或删除移码突变进行突变。Tsp 基因在编码序列的 5' 末端包含两个 ATG 密码子，其中两个 ATG 密码子之一或全部可通过错义突变而进行突变。Tsp 基因可在第二个 ATG 密码子（3 号密码子）处突变为 TCG，如图 9b 所示。Tsp 基因可备选地或另外地包含位于基因起始密码子下游、基因终止密码子上游的一个或多个终止密码子。优选地，敲除突变 Tsp 基因包含起始密码子的错义突变以及一个或多个插入的终止密码子。Tsp 基因可突变以删除第五个密码子中的“T”而形成移码，在 11 和 16 号密码子处产生终止密码子，如图 9b 所示。Tsp 还可突变为插入 AseI 限制性位点而在 21 号密码子处产生第三个框内终止密码子，如图 9b 所示。

[0119] 敲除突变的 Tsp 基因可具有 SEQ ID NO:3 的 DNA 序列，其包括起始密码子上游的 6 个核苷酸 ATGAAT。SEQ ID NO:3 的敲除突变 Tsp 序列中所作出的突变显示于图 9b。在一个实施方式中，突变的 Tsp 基因具有 SEQ ID NO:3 的 7-2048 号核苷酸的 DNA 序列。

[0120] 因此，一旦鉴定出携带有合适的突变 Tsp 基因的细胞，即可鉴定产生能够抑制包含突变 Tsp 基因的细胞的表型的 spr 蛋白质的合适 spr 基因突变。

[0121] 根据本发明一个优选的实施方式的细胞包含编码 spr 蛋白质的突变 spr 基因，其在选自 N31、R62、I70、Q73、C94、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144、H145、G147、H157 和 W174 的一个或多个氨基酸处具有突变，更优选地在选自 C94、S95、V98、Y115、D133、V135、H145、G147、H157 和 W174 的一个或多个氨基酸处具有突变。在此实

施方式中, spr 蛋白质优选地不具有任何其他突变。优选地, 突变 spr 基因编码的 spr 蛋白质在选自 N31、R62、I70、Q73、C94、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144、H145、G147 和 H157 的一个或多个氨基酸处具有突变, 更优选地, 在选自 C94、S95、V98、Y115、D133、V135、H145、G147 和 H157 的一个或多个氨基酸处具有突变。在此实施方式中, spr 蛋白质优选地不具有任何其他突变。优选地, 突变 SPR 基因编码的 spr 蛋白质在选自 N31、R62、I70、Q73、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144 和 G147 的一个或多个氨基酸处具有突变, 更优选地, 在选自 S95、V98、Y115、D133、V135 和 G147 的一个或多个氨基酸处具有突变。在此实施方式中, spr 蛋白质优选地不具有任何其他突变。

[0122] 在本发明的一个方面, 提供了革兰氏阴性细菌细胞, 其包含突变 spr 基因, 编码的 spr 蛋白质在选自 C94、S95、V98、Y115、D133、V135、H145、G147 和 H157 的一个或多个氨基酸处具有突变, 优选地, 在选自 S95、V98、Y115、D133、V135 和 G147 的一个或多个氨基酸处具有突变, 其中的细胞包含野生型 Tsp 基因。在此实施方式中, spr 蛋白质优选地不具有任何其他突变。

[0123] 野生型染色体 Tsp 基因优选地为非重组染色体基因。优选地, 细胞还包含编码 DsbC 的重组多核苷酸。

[0124] 对一个或多个以上氨基酸的突变可以是对编码所述氨基酸的一个、两个或三个核苷酸的任何合适的错义突变。突变将氨基酸残基变成任何合适的氨基酸, 产生能够抑制包含突变 Tsp 基因的细胞的表型的突变 SPR 蛋白质。错义突变可将氨基酸变为相比于野生型氨基酸不同大小及 / 或具有不同化学特性的氨基酸。

[0125] 在一个实施方式中, 突变是对 C94、H145 和 H157 构成的催化单分子中的一个、两个或三个氨基酸残基进行的突变 (Solution NMR Structure of the NlpC/P60 Domain of Lipoprotein Spr from Escherichia coli Structural Evidence for a Novel Cysteine Peptidase Catalytic Triad, Biochemistry, 2008, 47, 9715-9717)。

[0126] 因此, 突变 SPR 基因可包含:

- [0127] • 对 C94 的突变; 或
- [0128] • 对 H145 的突变; 或
- [0129] • 对 H157 的突变; 或
- [0130] • 对 C94 和 H145 的突变; 或
- [0131] • 对 C94 和 H157 的突变; 或
- [0132] • 对 H145 和 H157 的突变; 或
- [0133] • 对 C94、H145 和 H157 的突变。

[0134] 在此实施方式中, spr 蛋白质优选地不具有任何其他突变。

[0135] C94、H145 和 H157 中的一个、两个或三个可突变为任何合适的氨基酸, 产生能够抑制包含突变的 Tsp 基因的细胞的表型的 spr 蛋白质。例如, C94、H145 和 H157 中的一个、两个或三个可突变为小氨基酸如 Gly 或 Ala。因此, spr 蛋白质可具有 C94A、H145A 和 H157A 突变中的一个、两个或三个。优选地, spr 基因包含错义突变 H145A, 已经发现其产生能够抑制包含突变的 Tsp 基因的细胞的表型的 spr 蛋白质。

[0136] 对本文的置换突变体的标记由字母 + 数字 + 字母构成。第一个字母标记野生型蛋

白质中的氨基酸。数字指的是进行氨基酸置换的位置,第二个字母标记用于置换野生型氨基酸的氨基酸。

[0137] 在一个优选的实施方式中,突变 spr 蛋白质包含在选自 N31、R62、I70、Q73、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144 和 G147 的一个或多个氨基酸处的突变,优选地为选自 S95、V98、Y115、D133、V135 和 G147 的一个或多个氨基酸处的突变。在此实施方式中, spr 蛋白质优选地不具有任何其他突变。因此,突变 SPR 基因可包含:

[0138] • 对 N31 的突变 ;或

[0139] • 对 R62 的突变 ;或

[0140] • 对 I70 的突变 ;或

[0141] • 对 Q73 的突变 ;或

[0142] • 对 S95 的突变 ;或

[0143] • 对 V98 的突变 ;或

[0144] • 对 Q99 的突变 ;或

[0145] • 对 R100 的突变 ;或

[0146] • 对 L108 的突变 ;或

[0147] • 对 Y115 的突变 ;或

[0148] • 对 D133 的突变 ;或

[0149] • 对 V135 的突变 ;或

[0150] • 对 L136 的突变 ;或

[0151] • 对 G140 的突变 ;或

[0152] • 对 R144 的突变 ;或

[0153] • 对 G147 的突变。

[0154] 在一个实施方式中,突变 SPR 蛋白质包含多重氨基酸突变:

[0155] • S95 和 Y115 ;或

[0156] • N31、Q73、R100 和 G140 ;或

[0157] • Q73、R100 和 G140 ;或

[0158] • R100 和 G140 ;或

[0159] • Q73 和 G140 ;或

[0160] • Q73 和 R100 ;或

[0161] • R62、Q99 和 R144 ;或

[0162] • Q99 和 R144。

[0163] 氨基酸 N31、R62、I70、Q73、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144 和 G147 中的一个或多个可突变为任何合适的氨基酸,产生能够抑制包含突变的 Tsp 基因的细胞的表型的 spr 蛋白质。例如,N31、R62、I70、Q73、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140 和 R144 中的一个或多个可突变为小氨基酸如 Gly 或 Ala。

[0164] 在一个优选的实施方式中, spr 蛋白质包含以下突变中的一个或多个 :N31Y、R62C、I70T、Q73R、S95F、V98E、Q99P、R100G、L108S、Y115F、D133A、V135D 或 V135G、L136P、G140C、R144C 和 G147C。优选地, spr 蛋白质包含以下突变中的一个或多个 :S95F、V98E、Y115F、D133A、V135D 或 V135G 和 G147C。在此实施方式中, spr 蛋白质优选地不含有任何其他

他突变。

[0165] 在一个实施方式中, spr 蛋白质具有选自 N31Y、R62C、I70T、Q73R、S95F、V98E、Q99P、R100G、L108S、Y115F、D133A、V135D 或 V135G、L136P、G140C、R144C 和 G147C 的一个突变。在此实施方式中, spr 蛋白质优选地不具有任何其他突变。

[0166] 在一个进一步的实施方式中, spr 蛋白质具有多重突变, 选自:

[0167] • S95F 和 Y115F;

[0168] • N31Y、Q73R、R100G 和 G140C;

[0169] • Q73R、R100G 和 G140C;

[0170] • R100G 和 G140C;

[0171] • Q73R 和 G140C;

[0172] • Q73R 和 R100G;

[0173] • R62C、Q99P 和 R144C; 或

[0174] • Q99P 和 R144C。

[0175] 在一个实施方式中, spr 蛋白质具有突变 W174R。在一个备选的实施方式中, spr 蛋白质不具有突变 W174R。

[0176] 在一个优选的实施方式中, 根据本发明的细胞包含突变 SPR 基因以及编码 DsbC 的重组多核苷酸。

[0177] 如本文所使用的, “重组多肽”指的是使用重组 DNA 技术构建或产生的蛋白质。编码 DsbC 的多核苷酸序列可与细菌细胞中发现的编码 DsbC 的内源序列同一。备选地, 编码 DsbC 的重组多核苷酸序列为野生型 DsbC 序列的突变形式, 例如, 其有移除的限制性位点, 如 EcoRI 位点, 及 / 或具有编码 his 标签的序列。用于本发明的经修饰 DsbC 核苷酸序列的实例显示于 SEQ ID NO:26, 其编码带 his 标签的 DsbC 氨基酸序列, 显示于 SEQ ID NO:27。

[0178] 在本发明的一个方面, 提供了革兰氏阴性细菌细胞, 其包含编码突变 spr 蛋白质的突变 spr 基因以及编码 DsbC 的重组的多核苷酸, 并且其中细胞包含野生型 Tsp 基因。野生型 Tsp 基因优选地为非重组染色体 Tsp 基因。

[0179] DsbC 是一种发现于大肠杆菌细胞周质中的原核蛋白质, 其催化大肠杆菌中二硫键的形成。DsbC 的氨基酸序列长度为 236 (包括信号肽), 分子量为 25.6KDa (UniProt No. P0AEG6)。DsbC 于 1994 年首次发现 (Missiakas 等人 The Escherichia coli dsbC(xprA) gene encodes a periplasmic protein involved in disulfide bond formation, The EMBO Journal vol 13, no 8, p2013-2020, 1994 以及 Shevchik 等人 Characterization of DsbC, a periplasmic protein of Erwinia chrysanthemi and Escherichia coli with disulfide isomerase activity, The EMBO Journal vol 13, no 8, p2007-2012, 1994)。

[0180] 已知其与催化二硫键形成的蛋白质共表达以改善宿主细胞中的蛋白质表达。W098/56930 公开了用于在细菌细胞中生产异源的含二硫键的多肽的方法, 其中原核的二硫键异构酶如 DsbC 或 DsbG 与真核多肽共表达。US6673569 公开了一种人工操纵子, 其包含编码 DsbA、DsbB、DsbC 和 DsbD 各个的多核苷酸, 用于生产外源蛋白质。EP0786009 公开了用于在细菌中生产异源多肽的过程, 其中在诱导编码异源多肽的核酸表达之前诱导编码 DsbA 或 DsbC 的核酸表达。

[0181] 我们已经发现在包含突变 SPR 基因及野生型 Tsp 基因的细菌细胞中表达编码 DsbC

的重组多核苷酸这一特定组合提供了用于表达目的蛋白质的改良的宿主。出乎意料地发现与野生型细胞相比或包含突变 Tsp 基因的细胞,所述新菌株显示出增加的细胞生长速率及增加的细胞存活时间。具体地,相比于携带突变 Tsp 基因的细胞,携带重组 DsbC 基因、spr 突变及野生型 Tsp 的细胞显示出细胞裂解减少的表型。

[0182] 在一个实施方式中,根据本发明的细胞还进一步表达如下的一种或多种蛋白质:

[0183] • 一种或多种能够促进细胞折叠的蛋白质,如 FkpA、Skp、SurA、PPIa 和 PPIi;和 / 或

[0184] • 一种或多种能够促进蛋白质分泌或易位的蛋白质,如 SecY、SecE、SecG、SecYEG、SecA、SecB、FtsY 和 Lep;和 / 或

[0185] • 一种或多种能够促进二硫键形成的蛋白质,如 DsbA、DsbB、DsbD、DsbG。

[0186] 可将一种或多种以上蛋白质整合入细胞基因组且 / 或插入表达载体。

[0187] 在一个实施方式中,根据本发明的细胞不含有编码以下一种或多种其他蛋白质的重组多核苷酸:

[0188] • 一种或多种能够促进细胞折叠的蛋白质,如 FkpA、Skp、SurA、PPIa 和 PPIi;和 / 或

[0189] • 一种或多种能够促进蛋白质分泌或易位的蛋白质,如 SecY、SecE、SecG、SecYEG、SecA、SecB、FtsY 和 Lep;和 / 或

[0190] • 一种或多种能够促进二硫键形成的蛋白质,如 DsbA、DsbB、DsbD、DsbG。

[0191] 在本发明一个优选的实施方式中,重组革兰氏阴性细菌细胞进一步包含突变 DegP 基因和 / 或突变 ptr 基因和 / 或包含突变 OmpT 基因,所述突变 DegP 基因编码具有分子伴侣活性和降低的蛋白酶活性的 DegP 蛋白质,其中突变 ptr 基因编码具有降低的蛋白酶活性的蛋白酶 III 蛋白质或为敲除突变 ptr 基因,其中突变 OmpT 基因编码具有降低的蛋白酶活性的 OmpT 蛋白质或为敲除突变 OmpT 基因。

[0192] 在一个实施方式中,本发明提供重组的革兰氏阴性细菌细胞,其包含:

[0193] a. 突变 spr 基因;

[0194] b. 野生型非重组染色体 Tsp 基因;以及

[0195] c. 编码具有分子伴侣活性和降低的蛋白酶活性的 DegP 蛋白质的突变的 DegP 基因,和 / 或突变 OmpT 基因,其中突变 OmpT 基因编码具有降低的蛋白酶活性的 OmpT 蛋白质或为敲除突变 OmpT 基因。

[0196] 优选地,在此实施方式中,除了以上突变之外所述细胞与野生型细菌细胞是同基因的。

[0197] 在一个实施方式中,本发明提供了重组革兰氏阴性细菌细胞,包含:

[0198] a. 突变 SPR 基因;

[0199] b. 野生型非重组染色体 Tsp 基因;以及

[0200] c. 突变 ptr 基因和 / 或突变 OmpT 基因,其中突变 ptr 基因编码具有降低的蛋白酶活性的蛋白酶 III 蛋白质或为突变 ptr 基因,其中突变 OmpT 基因编码具有降低的蛋白酶活性的 OmpT 蛋白质或为敲除突变 OmpT 基因。

[0201] 优选地,在此实施方式中,除了以上突变之外所述细胞与野生型细菌细胞是同基因的。

[0202] 在一个实施方式中,本发明提供了细胞,包含:

[0203] a. 突变 SPR 基因;

[0204] b. 野生型非重组染色体 Tsp 基因;

[0205] c. 编码具有分子伴侣活性和降低的蛋白酶活性的 DegP 蛋白质的突变 DegP 基因;

[0206] d. 突变 ptr 基因,其中突变 ptr 基因编码具有降低的蛋白酶活性的蛋白酶 III 蛋白质或为突变 ptr 基因;以及

[0207] e. 任选地,突变 OmpT 基因,其中突变 OmpT 基因编码具有降低的蛋白酶活性的 OmpT 蛋白质或为敲除突变 OmpT 基因。

[0208] 优选地,在此实施方式中,除了以上突变之外所述细胞与野生型细菌细胞是同基因的。

[0209] 在本发明一个实施方式中,细胞携带突变 DegP 基因。如本文所使用的,“DegP”指的是编码 DegP 蛋白质(也称为 HtrA)的基因,其具有作为分子伴侣和蛋白酶双重功能(Families of serine peptidases;Rawlings ND,Barrett AJ.Methods Enzymol. 1994;244:19-61)。非突变 DegP 基因的序列显示于 SEQ ID NO:7,非突变的 DegP 氨基酸序列显示于 SEQ ID NO:8。

[0210] 在低温下,DegP 起分子伴侣的作用,在高温下,DegP 偏向于起蛋白酶的作用(A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein.Cell,Volume 97,Issue 3,Pages 339 - 347.Spiess C,Beil A,Ehrmann M 以及 The proteolytic activity of the HtrA(DegP)protein from Escherichia coli at low temperatures,Skorko-Glonek J 等人 Microbiology 2008, 154, 3649-3658)。

[0211] 在这些细胞包含 DegP 突变的实施方式中,细胞中的 DegP 突变提供了突变 DegP 基因,其编码具有分子伴侣活性但不具有完全蛋白酶活性的 DegP 蛋白质。

[0212] 词句“具有分子伴侣活性”在本发明的上下文中指的是与野生型非突变 DegP 蛋白质相比,突变 DegP 蛋白质具有相同或几乎相同的分子伴侣活性。优选地,突变 DegP 基因编码的 DegP 蛋白质具有野生型非突变 DegP 蛋白质的 50% 或更高、60% 或更高、70% 或更高、80% 或更高、90% 或更高或者 95% 的分子伴侣活性。更优选地,与野生型 DegP 相比,突变 DegP 基因编码的 DegP 蛋白质具有相同的分子伴侣活性。

[0213] 词句“具有降低的蛋白酶活性”在本发明的上下文中指的是与野生型非突变 DegP 蛋白质相比,突变 DegP 蛋白质不具有完全的蛋白酶活性。优选地,突变 DegP 基因编码的 DegP 蛋白质具有野生型非突变 DegP 蛋白质 50% 或更低、40% 或更低、30% 或更低、20% 或更低、10% 或更低或者 5% 或更低的蛋白酶活性。更优选地,突变 DegP 基因编码的 DegP 蛋白质不具有蛋白酶活性。所述细胞不是染色体 DegP 缺陷的,即,DegP 基因序列未经删除或突变以阻止任何形式的 DegP 蛋白质的表达。

[0214] 任何合适的突变可引入 DegP 基因以产生具有分子伴侣活性和降低的蛋白酶活性的蛋白质。表达自革兰氏阴性细菌的 DegP 蛋白质的蛋白酶和分子伴侣活性可由本领域技术人员通过任何合适的方法进行测试,如描述于 Spiess 等人的方法,其中对 DegP 的蛋白酶和分子伴侣活性的测试在 DegP 的天然底物 MalS 上进行(A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein.



Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M); 以及在 The proteolytic activity of the HtrA (DegP) protein from Escherichia coli at low temperatures, Skorko-Glonek J 等人 Microbiology 2008, 154, 3649-3658 中描述的方法。

[0215] DegP 为丝氨酸蛋白酶, 其活性中心由 His105、Asp135 和 Ser210 的氨基酸残基组成的催化三分子构成 (Families of serine peptidases, Methods Enzymol., 1994, 244:19-61 Rawlings N and Barrett A)。产生具有分子伴侣活性和降低的蛋白酶活性的 DegP 突变可包含的突变如对 His105、Asp135 和 Ser210 中一个、两个或三个的错义突变。

[0216] 因此, 突变 DegP 基因可包含:

[0217] • 对 His105 的突变; 或

[0218] • 对 Asp135 的突变; 或

[0219] • 对 Ser210 的突变; 或

[0220] • 对 His105 和 Asp135 的突变; 或

[0221] • 对 His105 和 Ser210 的突变; 或

[0222] • 对 Asp135 和 Ser210 的突变; 或

[0223] • 对 His105、Asp135 和 Ser210 的突变。

[0224] His105, Asp135 和 Ser210 中的一个、两个或三个可突变为任何合适的氨基酸, 产生具有分子伴侣活性和降低的蛋白酶活性的蛋白质。例如, His105, Asp135 和 Ser210 中的一个、两个或三个可突变为小氨基酸如 Gly 或 Ala。其他合适的突变是将 His105, Asp135 和 Ser210 中的一个、两个或三个变为具有相反特性的氨基酸, 如 Asp135 突变为 Lys 或 Arg, 极性的 His105 突变为非极性氨基酸如 Gly、Ala、Val 或 Leu, 而小分子亲水性 Ser210 突变为大分子或疏水性残基如 Val, Leu, Phe 或 Tyr。优选地, DegP 基因包含点突变 S210A, 如图 9c 中显示, 发现其产生了具有分子伴侣活性但不具有蛋白酶活性的蛋白质 (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M)。

[0225] DegP 具有两个 PDZ 结构域, PDZ1 (260-358 号残基) 和 PDZ2 (359-448 号残基), 其调节蛋白质-蛋白质相互作用 (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M)。在本发明的一个实施方式中, 将 DegP 基因突变以删除 PDZ1 结构域和 / 或 PDZ2 结构域。PDZ1 和 PDZ2 的删除导致相比于野生型 DegP 蛋白质, DegP 蛋白质的蛋白酶活性完全丧失, 且分子伴侣活性降低, 而删除 PDZ1 或 PDZ2 之一导致相比于野生型 DegP 蛋白质, 具有 5% 的蛋白酶活性和类似的分子伴侣活性 (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, Volume 97, Issue 3, Pages 339 - 347. Spiess C, Beil A, Ehrmann M)。

[0226] 突变 DegP 基因还可包含沉默的非天然发生的限制性位点, 如 AseI, 以助于鉴定和筛选方法, 例如, 如图 9c 所示。

[0227] SEQ ID NO:9 中提供了突变 DegP 基因的优选序列, 其包含点突变 S210A 及 Ase I

限制性标记物位点,所编码的蛋白质序列显示于 SEQ ID NO:10。SEQ ID NO:9 的突变 DegP 序列中所作出的突变显示于图 9c。

[0228] 在本发明的实施方式中,其中细胞包含突变 DegP 基因,其编码的 DegP 蛋白质具有分子伴侣活性和降低的蛋白酶活性,相对于其中 DegP 基因突变为敲除 DegP 而阻止 DegP 表达(如染色体缺陷 DegP)的突变细胞,本发明提供的一种或多种细胞可以提供产自所述细胞正确折叠蛋白质的提高的产量。在包含敲除突变 DegP 基因而阻止 DegP 表达的细胞中, DegP 的分子伴侣活性完全丧失,而根据本发明的细胞中, DegP 的分子伴侣活性得到保留,同时却丧失了全部的蛋白酶活性。在这些实施方式中,根据本发明的一种或多种细胞具有降低的蛋白酶活性而阻止蛋白质的蛋白水解作用,同时却保留了分子伴侣活性以使蛋白质在宿主细胞中得以正确的折叠和运输。

[0229] 本领域技术人员能够使用本领域已知方法容易地测试分泌的蛋白质以确定蛋白质是否正确折叠,所述方法如蛋白质 G HPLC、圆二色谱、NMR、X-光晶体成像和表位亲和测量方法。

[0230] 在这些方法中,相比于携带突变敲除 DegP 基因而阻止 DegP 表达的细胞,根据本发明的一种或多种细胞具有改善的细胞生长。不意欲受到任何理论的约束,由于保持分子伴侣活性的 DegP 蛋白酶增加细胞处理所有需要分子伴侣活性的蛋白质的能力从而显示出细胞生长改善。因此,相比于携带 DegP 敲除突变的细胞,在本发明的一种或多种细胞中,产生对细胞的生长和繁殖必需的正确折叠蛋白质有所增加,由此改善调控生长的细胞途径。进一步地,已知的 DegP 蛋白酶缺陷菌株一般为温度敏感型,一般在超过大约 28°C 的温度下不生长。然而,根据本发明的细胞是非温度敏感型,可在 28°C 或更高的温度下生长,包括大约 30°C - 大约 37°C 的温度,这些温度一般用于细菌生产蛋白质的工业规模生产。

[0231] 在本发明一个实施方式中,细胞携带突变 ptr 基因。如本文所使用的,“ptr 基因”指的是编码蛋白酶 III 的基因,其为降解高分子量蛋白质的蛋白酶。非突变 ptr 基因的序列显示于 SEQ ID NO:4,非突变蛋白酶 III 蛋白质的序列显示于 SEQ ID NO:5。

[0232] 除非另有说明。提到突变 ptr 基因或编码蛋白酶 III 的突变 ptr 基因时,指的是编码具有降低的蛋白酶活性的蛋白酶 III 的突变 ptr 基因或敲除突变 ptr 基因。

[0233] 词句“编码具有降低的蛋白酶活性的蛋白酶 III 的突变 ptr 基因”在本发明的上下文中指的是相比于野生型非突变 ptr 基因,突变 ptr 基因不具有完全的蛋白酶活性。

[0234] 优选地,突变 ptr 基因编码的蛋白酶 III 具有野生型非突变蛋白酶 III 蛋白质的 50% 或更低、40% 或更低、30% 或更低、20% 或更低、10% 或更低或者 5% 或更低的蛋白酶活性。更优选地,突变 ptr 基因编码的蛋白酶 III 蛋白质不具有蛋白酶活性。在此实施方式中,细胞在染色体 ptr 上没有缺陷,即 ptr 基因序列未经删除或突变而阻止任何形式蛋白酶 III 蛋白质的表达。

[0235] 任何合适的突变可引入 ptr 基因以产生具有降低的蛋白酶活性的蛋白酶 III。表达自革兰氏阴性细菌的蛋白酶 III 蛋白质的蛋白酶活性可容易地由本领域技术人员通过任何本领域合适的方法进行测试。

[0236] 词句“敲除突变 ptr 基因”在本发明的上下文中指的是所述基因包含一个或多个突变,由此导致此基因编码的蛋白质不表达,造成细胞中由所述敲除突变基因编码的蛋白质缺陷。敲除基因可部分或全部地转录,但不翻译为编码的蛋白质。敲除突变 ptr 基因可

以任何合适的方式,例如通过删除、插入、点突变、错义、无义及移码突变中一种或多种方式进行突变,使得蛋白质不表达。例如,基因可通过将外源 DNA 序列如抗生素抗性标记物插入基因编码序列而被敲除。

[0237] 在一个优选的实施方式中,基因不是通过将外源 DNA 序列如抗生素抗性标记物插入基因编码序列而突变的。优选地,蛋白酶 III 基因包含对基因起始密码子的突变和 / 或形成位于基因起始密码子下游、基因终止密码子上游的一个或多个终止密码子的突变,由此阻止蛋白酶 III 蛋白质的表达。

[0238] 对靶标敲除基因起始密码子进行突变引起起始密码子功能的丧失,并由此确保靶标基因在编码序列的起始不含有合适的起始密码子。对起始密码子的突变可为起始密码子核苷酸中一个、两个或三个核苷酸的错义突变。备选地或另外地,起始密码子可通过插入或删除移码突变进行突变。

[0239] 在一个优选的实施方式中,ptr 基因突变使得 ATG 起始密码子变为 ATT,如图 9a。

[0240] 敲除突变 ptr 基因可备选地或另外地包含位于基因起始密码子下游、基因终止密码子上游的一个或多个终止密码子。优选地,敲除突变 ptr 基因包含对起始密码子的错义突变以及一个或多个插入的终止密码子。

[0241] 所述一个或多个插入的终止密码子优选地为框内终止密码子。然而,所述一个或多个插入的终止密码子可备选地或另外地为框外密码子。当框外起始密码子通过插入或删除移码突变转变为框内起始密码子时,需要一个或多个框外终止密码子以终止翻译。可通过任何合适的突变引入所述一个或多个终止密码子,所述突变包括无义点突变和移码突变。优选地,可通过移码突变和 / 或插入突变引入所述一个或多个终止密码子,优选地通过以包含终止密码子的序列置换基因序列的一段而实现。例如,可插入 AseI 限制性位点,其包含终止密码子 TAA。

[0242] 在一个优选的实施方式中,将 ptr 基因突变以通过插入 Ase I 限制性位点而插入框内终止密码子,如图 9a 所示。在一个优选的实施方式中,敲除突变 ptr 基因具有 SEQ ID NO:6 的 DNA 序列。在 SEQ ID NO:6 所示的敲除突变 ptr 基因序列中所作出的突变显示于图 9a。

[0243] 上文描述的敲除突变具有优势,因为其对靶标敲除基因位点上游或下游的染色体 DNA 造成最小的破坏或无破坏,且不需要插入和保留外源 DNA,如抗生素抗性标记物,这些可以影响细胞用于表达目的蛋白质,尤其是治疗性蛋白质的合适度。因此,相比于通过将外源 DNA 插入基因编码序列而敲除蛋白酶基因的细胞,根据本发明的一种或多种细胞可以显示出改善的生长特征和 / 或蛋白质表达。

[0244] 在一个实施方式中,根据本发明的细胞携带突变 OmpT 基因。如本文所使用的,“OmpT 基因”指的是编码蛋白酶 OmpT (外膜蛋白酶 T) 的基因,其为外膜蛋白酶。野生型非突变 OmpT 基因的序列为 SWISS-PROT P09169。

[0245] 除非另有说明,提到突变 OmpT 基因或编码 OmpT 的突变 OmpT 基因时,指的是编码具有降低的蛋白酶活性的 OmpT 蛋白质的突变 OmpT 基因或敲除突变 OmpT 基因,

[0246] 词句“编码具有降低的蛋白酶活性的 OmpT 蛋白质的突变 OmpT 基因”在本发明的上下文中指的是相比于野生型非突变 OmpT 基因,突变 OmpT 基因不具有完全的蛋白酶活性。突变 OmpT 基因可以编码的 OmpT 蛋白质具有野生型非突变 OmpT 蛋白质的 50% 或更低、40%

或更低、30% 或更低、20% 或更低、10% 或更低或者 5% 或更低的蛋白酶活性。突变 OmpT 基因可以编码不具有蛋白酶活性的 OmpT 蛋白质。在此实施方式中,细胞在染色体 OmpT 上没有缺陷,即 OmpT 基因序列未经删除或突变而阻止任何形式的 OmpT 蛋白质的表达。

[0247] 任何合适的突变可引入 OmpT 基因以产生具有降低的蛋白酶活性的蛋白质。表达自革兰氏阴性细菌的 OmpT 蛋白质的蛋白酶活性可容易地由本领域技术人员通过任何本领域合适的方法进行测试,如描述于 Kramer 等人(Identification of essential acidic residues of outer membrane protease OmpT supports a novel active site,FEBS Letters 505(2001)426-430)和 Dekker 等人(Substrate Specificity of the Integral Membrane Protease OmpT Determined by Spatially Addressed Peptide Libraries,Biochemistry 2001,40,1694-1701)的方法。

[0248] Kramer 等人(Identification of active site serine and histidine residues in Escherichia coli outer membrane protease OmpT FEBS Letters 2000468,220-224)已经报道了 OmpT,公开了由丙氨酸置换丝氨酸、组氨酸和酸性残基导致活性降低:对于 Glu27、Asp97、Asp208 或 His101 活性降低~10 倍,对于 Ser99 活性降低~500 倍,对于 Asp83、Asp85、Asp210 或 His212 活性降低~10000 倍。Vandeputte-Rutten 等人(Crystal Structure of the Outer Membrane Protease OmpT from Escherichia coli suggests a novel catalytic site,The EMBO Journal2001,Vol 20No 185033-5039)提出具有活性中心,包含 Asp83-Asp85 对以及 His212-Asp210 对。进一步地,Kramer 等人(Lipopolysaccharide regions involved in the activation of Escherichia coli outer membrane protease OmpT, Eur. J. Biochem. FEBS 2002, 269, 1746-1752)公开了 L4 环中的 D208A、D210A、H212A、H212N、H212Q、G216K/K217G、K217T 和 R218L 突变,这些都导致酶活性部分或几乎全部丧失。

[0249] 因此,产生具有降低的蛋白酶活性的 OmpT 突变可以包含的突变如对 E27、D43、D83、D85、D97、S99、H101E111、E136、E193、D206、D208、D210、H212G216、K217、R218 和 E250 中一个或多个残基的错义突变。

[0250] E27、D43、D83、D85、D97、S99、H101E111、E136、E193、D206、D208、D210、H212G216、K217、R218 和 E250 中一个或多个可突变为任何合适的氨基酸以产生具有降低的蛋白酶活性的蛋白质。例如,E27、D43、D83、D85、D97、S99、H101E111、E136、E193、D206、D208、D210、H212G216、K217、R218 和 E250 中的一个或多个可突变为丙氨酸。合适的突变的实例为 E27A、D43A、D83A、D85A、D97A、S99A、H101AE111A、E136A、E193A、D206A、D208A、D210A、H212A、H212N、H212Q、G216K、K217G、K217T、R218L 和 E250A。在一个实施方式中,突变 OmpT 基因包含 D210A 和 H212A 突变。包含 D210A 和 H212A 突变的合适的突变 OmpT 序列显示于 SEQ ID NO:23。

[0251] 词句“敲除突变 OmpT 基因”在本发明的上下文中指的是所述基因包含一个或多个突变由此导致此基因编码的蛋白质不表达,造成细胞中由所述敲除突变基因编码的蛋白质缺陷。敲除基因可部分或全部地转录,但不翻译为编码的蛋白质。敲除突变 OmpT 基因可以任何合适的方式,例如通过删除、插入、点突变、错义、无义及移码突变中一种或多种方式进行突变,使得蛋白质不表达。例如,基因可通过将外源 DNA 序列如抗生素抗性标记物插入基因编码序列而被敲除。

[0252] 在一个实施方式中,OmpT 基因包含对基因起始密码子的突变和 / 或位于基因起始

密码子下游、基因终止密码子上游的一个或多个终止密码子的突变,由此阻止 OmpT 蛋白质的表达。对起始密码子的突变可为起始密码子核苷酸中一个、两个或三个核苷酸的错义突变。备选地或另外地,起始密码子可通过插入或删除移码突变进行突变。

[0253] 合适的突变敲除 OmpT 序列显示于 SEQ ID NO:24。

[0254] 在一个实施方式中,根据本发明的革兰氏阴性细菌细胞不携带有敲除突变 OmpT 基因,如染色体 ompT 缺陷。

[0255] 在一个实施方式中,根据本发明的革兰氏阴性细菌细胞不携带有敲除突变 DegP 基因,如染色体 degP 缺陷。在一个实施方式中,根据本发明的革兰氏阴性细菌细胞不携带突变 DegP 基因。

[0256] 在一个实施方式中,根据本发明的革兰氏阴性细菌细胞不携带敲除突变 ptr 基因,如染色体 ptr 缺陷。

[0257] 包括敲除突变在内的许多遗传工程改造突变涉及使用抗生素抗性标记物,这样能够选择和鉴定成功突变的细胞。然而,使用抗生素抗性标记物有大量的不利之处。

[0258] 在本发明的一个实施方式中,突变基因可包含一个或多个标记物位点。因此, spr 基因及 / 或编码具有分子伴侣活性而不具有蛋白酶活性的 DegP 蛋白质的突变 DegP 基因及 / 或突变 ptr 基因及 / 或突变 OmpT 基因可突变为包含一个或多个限制性标记物位点。限制性位点通过遗传工程改造而进入基因,其为非天然发生的。限制性标记物位点是有利的,因为其能够使得对包含所需的染色体突变的正确修饰的细胞进行选择和鉴定。经修饰携带一个或多个所述突变基因的细胞可通过对来自细胞裂解物的基因组 DNA 进行 PCR 而进行分析,其中所用的寡核苷酸对设计为扩增包含非天然发生的限制性标记物位点的基因组 DNA 区域。随后将扩增的 DNA 与能够在非天然发生限制性标记物位点消化 DNA 的合适的限制性酶孵育之前和之后通过琼脂糖凝胶电泳进行。用限制性酶孵育后产生的 DNA 片段证实了细胞已经成功地修饰为携带所述一个或多个突变的基因。

[0259] 在细胞包含具有 SEQ ID NO:6 的 DNA 序列的敲除突变 ptr 基因的实施方式中,SEQ ID NO:17 和 SEQ ID NO:18 中所示的寡核苷酸引物序列可用于扩增来自转化细胞的基因组 DNA 中包含非天然发生的 AseI 限制性位点的 DNA。扩增的基因组 DNA 然后可与 Ase I 限制性酶一同孵育并通过凝胶电泳进行分析以证实基因组 DNA 中有突变 ptr 基因存在。

[0260] 在细胞包含具有 SEQ ID NO:9 的 DNA 序列的敲除突变 DegP 基因的实施方式中,SEQ ID NO:19 和 SEQ ID NO:20 中所示的寡核苷酸引物序列可用于扩增来自转化细胞的基因组 DNA 中包含非天然发生的 Ase I 限制性位点的 DNA。扩增的基因组 DNA 然后可与 Ase I 限制性酶一同孵育并通过凝胶电泳进行分析以证实基因组 DNA 中有突变 DegP 基因存在。

[0261] 一个或多个限制性位点可通过任何合适的突变而引入,包括删除、插入、点突变、错义、无义和移码突变中的一种或多种。限制性位点可通过对起始密码子的突变及 / 或引入一个或多个终止密码子的突变而引入,如上文所描述的。此实施方式是有利的,因为限制性标记物位点是所引入的敲除突变的直接且唯一的标记物。

[0262] 可插入包含框内终止密码子的限制性标记物位点,如 Ase I 限制性位点。这一点尤其有利,因为插入的限制性位点同时起到限制性标记物位点以及终止密码子的作用,以阻止基因编码序列的完全转录。例如,在通过引入 Ase I 位点向 ptr 基因引入终止密码子的实施方式中,这还制造出限制性位点,如图 9a 所示。

[0263] 可通过突变起始密码子和任选地一个或多个其他点突变,将限制性标记物位点插入。在此实施方式中,限制性标记物位点优选地为 EcoR I 限制性位点。这一点有其有利,因为对起始密码子的突变还制造出限制性标记物位点。例如,在一个实施方式中,ptr 基因的起始密码子变为 ATT,这制造出 EcoR I 标记物位点,如图 9a 所示。

[0264] 在细胞携带突变 OmpT 基因的本发明的实施方式中,所述一个或多个限制性位点可通过任何合适的突变引入,所述突变包括删除、插入、点突变、错义、无义和移码突变中的一种或多种。例如,在 OmpT 基因包含 D210A 和 H212A 突变的实施方式中,这些突变引入的沉默的 HindIII 限制性位点,其可用作选择性标记物。

[0265] 在突变 SPR 基因和突变 DegP 基因中,标记物限制性位点可使用沉默密码子改变而引入。例如,AseI 位点可用作沉默限制性标记物位点,其中 TAA 终止密码子在框外,如图 9c 所示的 DegP 序列。

[0266] 在本发明的实施方式中,其中 ptr 基因突变为编码具有降低的蛋白酶活性的蛋白酶 III,可使用沉默密码子改变引入一个或多个标记物限制性位点。

[0267] 根据本发明的重组革兰氏阴性细菌细胞可通过任何合适的方式产生。

[0268] 本领域技术人员了解可用于将染色体基因序列用突变的基因序列置换而引入 spr 突变体基因的合适的技术。可采用合适的载体,其能通过同源重组整合入宿主染色体。

[0269] 合适的基因置换的方法描述于,例如,Hamilton 等人(New Method for Generating Deletions and Gene Replacements in Escherichia coli,Hamilton C.M. 等人,Journal of Bacteriology Sept. 1989,Vol. 171, No. 9p 4617-4622)、Skorupski 等人(Positive selection vectors for allelic exchange,Skorupski K 和 Taylor R. K.,Gene, 1996, 169, 47-52)、Kiel 等人(A general method for the construction of Escherichia coli mutants by homologous recombination and plasmid segregation,Kiel J. A. K. W. 等人 Mol Gen Genet 1987, 207:294-301)、Blomfield 等人(Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature sensitive pSC101 replicon,Blomfield I. C. 等人, Molecular Microbiology 1991, 5(6), 1447-1457)以及 Ried 等人(An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in Gram-negative bacteria by marker exchange-eviction mutagenesis,Ried J. L. and Collmer A.,Gene 57(1987) 239-246)。能使同源重组/置换发生的合适质粒为 pK03 质粒(Link 等人,1997, Journal of Bacteriology, 179, 6228-6237)。

[0270] 在细胞包含编码 DsbC 的重组多核苷酸的实施方式中,本领域技术人员了解可用于插入编码 DsbC 的重组多核苷酸的合适技术。可使用合适的载体如 pK03 质粒,将编码 DsbC 的重组多核苷酸整合入细胞基因组。

[0271] 在细胞包含编码目的蛋白质的重组多核苷酸的实施方式中,本领域技术人员也了解可用于插入编码目的蛋白质的重组多核苷酸的合适技术。可使用合适的载体如 pK03 质粒,将编码目的蛋白质的重组多核苷酸整合入细胞基因组。

[0272] 备选地或另外地,编码 DsbC 的重组多核苷酸和/或编码目的蛋白质的重组多核苷酸可以不整合在重组表达盒中。在一个实施方式中,本发明采用了表达盒,以携带编码 DsbC 和/或目的蛋白质的多核苷酸以及一种或多种调控表达序列。所述一种或多种调控表达序

列可包含启动子。所述一种或多种调控表达序列还可包含 3' 不翻译区如终止序列。合适的启动子在下文详细讨论。

[0273] 在一个实施方式中,本发明采用表达盒以携带编码目的蛋白质的多核苷酸和 / 或编码 DsbC 的重组多核苷酸。表达盒一般包含一种或多种调控表达序列、编码一种或多种目的蛋白质的一种或多种编码序列及 / 或编码 DsbC 的编码序列。所述一种或多种调控表达序列可包括启动子。所述一种或多种调控表达序列还可包含 3' 不翻译区如终止序列。合适的启动子在下文详细讨论。

[0274] 在一个实施方式中,根据本发明的细胞包含一种或多种载体,如质粒。载体优选地包含一种或多种如上文限定的表达盒。在一个实施方式中,编码目的蛋白质的多核苷酸序列和编码 DsbC 的多核苷酸插入到一个载体中。备选地,编码目的蛋白质的多核苷酸序列和编码 DsbC 的多核苷酸插入到分别的载体中。

[0275] 在目的蛋白质为包含重链和轻链的抗体的实施方式中,可以用两种载体转染细胞系,第一种载体编码了轻链多肽,第二种载体编码重链多肽。备选地,可使用单一载体,此载体包括编码轻链和重链多肽的序列。备选地,编码抗体的多核苷酸序列和编码 DsbC 的多核苷酸插入到一种载体。优选地,此载体包含编码抗体的轻链和重链多肽的序列。

[0276] 在细胞还表达一种或多种其他如下蛋白质的实施方式中,

[0277] • 一种或多种能够促进细胞折叠的蛋白质,如 FkpA、Skp、SurA、PPiA 和 PPiD; 和 / 或

[0278] • 一种或多种能够促进蛋白质分泌或易位的蛋白质,如 SecY、SecE、SecG、SecYEG、SecA、SecB、FtsY 和 Lep; 和 / 或

[0279] • 一种或多种能够促进二硫键形成的蛋白质,如 DsbA、DsbB、DsbD、DsbG。

[0280] 所述一种或多种其他蛋白质可表达自一种或多种与编码 DsbC 的多核苷酸和 / 或编码目的蛋白质的多核苷酸序列插入相同载体的多核苷酸。备选地,所述一种或多种多核苷酸可插入分别的载体。

[0281] 用于本发明的载体可通过将如上文限定的一种或多种表达盒插入合适的载体而产生。备选地,用于指导多核苷酸序列表达的调控表达序列可包含于载体中,因此,要完整地构建载体只需要多核苷酸的编码区域。

[0282] 合适地,编码 DsbC 的多核苷酸和 / 或编码目的蛋白质的多核苷酸插入可复制载体,一般为自主复制载体,以在细胞的合适的启动子控制下在细胞内进行表达。本领域已知许多载体可用于此目的,选择合适的载体取决于核酸大小以及特定的细胞类型。

[0283] 可用来将根据本发明的多核苷酸转化入宿主细胞的载体的实例包括:

[0284] • 质粒,如 pBR322 或 pACYC184 和 / 或

[0285] • 病毒载体如细菌噬菌体

[0286] • 转座遗传元件如转座子

[0287] 这些载体通常包含 DNA 复制的质粒起点、抗生素选择性标记物、启动子以及及转录终止子,由多克隆位点分隔开(表达盒)以及编码核糖体结合位点的 DNA 序列。

[0288] 本发明采用的启动子可直接与相关的多核苷酸相连,或备选地可位于恰当的位置,例如位于载体中,使得当相关多核苷酸插入时,相关的启动子可对其起作用。在一个实施方式中,启动子位于其作用的多核苷酸编码部分之前,例如,相关的启动子在多核苷酸各

个编码部分之前。如本文所使用的，“在……之前”意为启动子位于相对编码多核苷酸部分的 5 引物端。

[0289] 启动子可以为宿主细胞内源或外源的。合适的启动子包括 Lac、tac、trp、PhoA、Ipp、Arab、Tet 和 T7。

[0290] 所使用的一种或多种启动子可为可诱导启动子。

[0291] 在其中的编码 DsbC 的多核苷酸及编码目的蛋白质的多核苷酸插入一个载体的实施方式中，编码 DsbC 和目的蛋白质的核苷酸序列可处于单一的启动子或分别的启动子的控制之下。在其中编码 DsbC 和目的蛋白质的核苷酸序列处于分别的启动子控制下时，启动子可为独立的可诱导启动子。

[0292] 用于细菌系统的表达单位通常还包含 Shine-Dalgarno (S. D.) 序列，其可操作地连接于编码目的多肽的 DNA。启动子可通过限制性酶消化从细菌来源的 DNA 移除并插入包含所需的 DNA 的载体。

[0293] 在其中的多核苷酸序列包含用于两种或更多目的蛋白质（例如抗体轻链和抗体重链）的两种或更多编码序列的本发明的实施方式中，多核苷酸序列可包含一个或多个内核糖体插入位点 (IRES) 序列，其使得能够在 mRNA 的中间起始翻译。IRES 序列可位于编码的多核苷酸序列之间，以增强 mRNA 分别翻译而产生多种编码的多肽序列。

[0294] 表达载体优选地还包含用于产生抗体或其抗原结合片段的双顺反子信息，如 WO 03/048208 或 WO2007/039714（其内容并入本文作为参考）中所描述。优选地，上游的顺反子包含编码抗体轻链的 DNA，下游顺反子包括编码对应的重链的 DNA，双顺反子基因间序列 (IGS) 优选地包含选自 IGS1 (SEQ ID NO:36)、IGS2 (SEQ ID NO:37)、IGS3 (SEQ ID NO:38) 和 IGS4 (SEQ ID NO:39) 的序列。

[0295] 终止子可以是宿主细胞内源或外源的。合适的终止子为 rrnB。

[0296] 进一步的合适的转录调控子包括启动子和终止子，蛋白质靶向方法可见于“Strategies for Achieving High-Level Expression of Genes in Escherichia coli” Savvas C. Makrides, Microbiological Reviews, Sept 1996, p 512-538。

[0297] 插入表达载体的 DsbC 多核苷酸优选地包含编码 DsbC 信号序列的核酸以及 DsbC 编码序列。所述载体优选地包含能够使得载体在一种或多种所选的宿主细胞中复制的核酸序列，优选地，其能够不依赖宿主染色体而复制。对于多种细菌这样的序列已经熟知。

[0298] 在一个实施方式中，DsbC 及 / 或目的蛋白质在 N- 端及 / 或 C- 端包含组氨酸标签。

[0299] 抗体分子可分泌自细胞或通过合适的信号序列进入细胞周质。备选地，抗体分子可在细胞的细胞质内积累。优选地，抗体分子进入细胞周质。

[0300] 编码目的蛋白质的多核苷酸可与另一种多肽融合表达，另一种多肽优选地为信号序列或其他在成熟多肽的 N- 端具有特异性剪切位点的多肽。所选择的异源信号序列应当为可以被宿主细胞识别并加工的肽。对于不能识别并加工原生或真核多肽信号序列的原核宿主细胞，则用原核信号序列置换信号序列。合适的信号序列包括 OmpA、PhoA、LamB、PelB、DsbA 和 DsbC。

[0301] 构建包含一种或多种上文列出的组分的合适的载体采用了标准的连接技术。对分离的质粒或 DNA 片段进行剪切、加尾并连接为产生所需质粒要求的形式。

[0302] 在本发明一个优选的实施方式中，本发明提供了多顺反子载体，其包含编码 DsbC



的多核苷酸序列以及编码目的蛋白质的多核苷酸序列。多顺反子载体可通过有利的克隆方法产生,其使多核苷酸序列以重复顺序地克隆进入载体。所述方法利用了一对限制性位点的可兼容末端,如 AseI 和 NdeI 限制性位点的“AT”末端。包含编码序列并具有可兼容粘性末端的多核苷酸如 AseI-NdeI 片段,可克隆入载体中的限制性位点,如 NdeI。多核苷酸序列的插入破坏了 5' 限制性位点但创造出一个新的 3' 限制性位点,如 NdeI,然后可插入包含可兼容粘性末端的其他多核苷酸序列。此过程然后可进行重复以插入其他序列。插入载体的每段多核苷酸序列包含终止密码子 3' 方向的非编码序列,这段序列可包含 Ssp I 位点用以进行选择、Shine-Dalgarno 核糖体结合序列、富含 A 的间隔子和编码起始密码子的 NdeI 位点。

[0303] 构建包含编码抗体轻链(LC)、抗体重链(HC)、DsbC 多核苷酸序列以及进一步的多核苷酸序列的载体示意图显示于图 10。

[0304] 成功突变的菌株可使用本领域熟知的方法进行鉴定,包括菌落 PCR DNA 测序以及菌落 PCR 限制性酶切作图。

[0305] 在其中细胞包含两种或更多种突变基因的实施方式中,突变的蛋白酶可在相同或不同载体上引入革兰氏阴性细菌。

[0306] 在一个实施方式中,根据本发明的革兰氏阴性细菌细胞不携带敲除突变 OmpT 基因,如染色体 ompT 缺陷。

[0307] 根据本发明的细胞可以进一步包含编码目的蛋白质的多核苷酸序列。编码目的蛋白质的多核苷酸序列可以为外源或内源的。编码目的蛋白质的多核苷酸序列可整合入宿主染色体或可非整合的存在于载体中,通常为质粒。

[0308] 在一个实施方式中,根据本发明的细胞表达目的蛋白质。“目的蛋白质”在本发明的上下文中意欲指代用于表达的多肽,通常为重组多肽。然而,目的蛋白质可以为表达自宿主细胞中内源基因的内源蛋白质。

[0309] 如本文所使用的,“重组多肽”指的是使用重组 DNA 技术构建或产生的蛋白质。目的蛋白质可以为表达自外源载体的与内源蛋白质同一的外源序列或其突变形式(例如,其具有减弱的生物活性)或其片段。备选地,目的蛋白质可为并非宿主细胞正常表达的异源蛋白质。

[0310] 目的蛋白质可以为任何合适的蛋白质,包括治疗性、预防性或诊断性蛋白质。

[0311] 在一个实施方式中,目的蛋白质可用于治疗包括炎性疾病及障碍、免疫疾病及障碍、纤维化障碍以及癌症在内的疾病或障碍。

[0312] 术语“炎性疾病”或“障碍”以及“免疫疾病或障碍”包括类风湿关节炎、牛皮癣关节炎、斯蒂尔病(still's disease)、Muckle Wells 病、牛皮癣、克罗恩氏病、溃疡性结肠炎、SLE(系统性红斑狼疮)、哮喘、过敏性关节炎、特应性皮炎、多发性硬化、脉管炎、I 型糖尿病、器官移植以及移植物抗宿主病。

[0313] 术语“纤维化障碍”包括特发性肺纤维化(IPF)、系统性硬化(或硬皮病)、肾纤维化、糖尿病肾病、IgA 肾病、高血压、晚期肾病、腹膜纤维化(持续不卧床腹膜透析)、肝硬化、老年性黄斑退化症(ARMD)、视网膜病、心反应性纤维化、瘢痕、瘢痕瘤(keloid)、烧伤、皮肤溃疡、血管成形术、冠状动脉架搭桥术、关节成形术以及白内障手术。

[0314] 术语“癌”包括产生自上皮的恶性新生生长,其见于皮肤,或更通常地见于身体器

官内膜(lining),例如:乳腺、卵巢、前列腺、肾、胰腺、胃、膀胱或肠。癌倾向于浸润进入临近阻止病扩散(转移)到远处器官,例如:至骨、肝脏、肺或脑。

[0315] 所述蛋白质可以为蛋白水解敏感型多肽,即,易于裂解、对裂解易感,或在原生状态或在分泌过程中被一种或多种革兰氏阴性细菌如大肠杆菌、蛋白酶裂解。在一个实施方式中,目的蛋白质对选自DegP、蛋白酶III和Tsp的蛋白酶具有蛋白水解敏感性。在一个实施方式中,目的蛋白质对蛋白酶Tsp具有蛋白水解敏感性。在一个实施方式中,目的蛋白质对蛋白酶DegP和蛋白酶III具有蛋白水解敏感性。在一个实施方式中,目的蛋白质对蛋白酶DegP和Tsp具有蛋白质水解敏感性。在一个实施方式中,目的蛋白质对蛋白酶Tsp和蛋白酶III具有蛋白质水解敏感性。在一个实施方式中,目的蛋白质对蛋白酶DegP、蛋白酶III和Tsp具有蛋白质水解敏感性。

[0316] 优选地,蛋白质为真核多肽。

[0317] 由根据本发明的细胞表达的目的蛋白质可以是例如为免疫原,包含两种异源蛋白质的融合蛋白质或抗体。作为目的蛋白质的抗体包括单克隆、多价、多特异性、人源化、全长人抗体或嵌合抗体。所述抗体可来自任何物种,但其优选地衍生自单克隆抗体、人抗体或人源化片段。抗体可衍生自免疫球蛋白分子的任何种型(例如IgG、IgE、IgM、IgD或IgA)或亚型,可获自任何物种包括例如小鼠、大鼠、鲨鱼、兔、猪、仓鼠、骆驼、美洲驼(1lama)、山羊或人。抗体片段的各个部分可获自超过一种物种,例如,抗体片段可以是嵌合的。在一个实施例中,恒定区来自一个物种而可变区来自另一物种。

[0318] 抗体可以为具有全长重链和轻链或其片段,例如VH、VL、VHH、Fab、修饰的Fab、Fab'、F(ab')<sub>2</sub>、Fv、scFv片段、Fab-Fv的完整的抗体分子,或双重特异性抗体如Fab-dAb,描述于PCT/GB2008/003331。

[0319] 抗体可特异于任何靶标抗原。抗体可以为细胞相关的蛋白质,如细胞表面蛋白质,其存在于细胞上如细菌细胞、酵母细胞、T细胞、内皮细胞或肿瘤细胞;或其可为可溶性蛋白质。目的抗原也可为任何医疗相关蛋白质如在疾病或感染中上调的那些蛋白质,例如受体及/或其对应配体。细胞表面蛋白的具体实例包括粘附分子,例如整合素如β1整合素(如VLA-4)、E-选择素、P-选择素或L-选择素、CD2、CD3、CD4、CD5、CD7、CD8、CD11a、CD11b、CD18、CD19、CD20、CD23、CD25、CD33、CD38、CD40、CD40L、CD45、CDW52、CD69、CD134(OX40)、ICOS、BCMP7、CD137、CD27L、CDCP1、CSF1或CSF1-受体、DPCR1、DPCR1、dudulin2、FLJ20584、FLJ40787、HEK2、KIAA0634、KIAA0659、KIAA1246、KIAA1455、LTBP2、LTK、MAL2、MRP2、nectin-样蛋白2、NKCC1、PTK7、RAIG1、TCAM1、SC6、BCMP101、BCMP84、BCMP11、DTD、胚癌抗原(CEA)、人乳脂球蛋白(HMFG1和2)、I型MHC及II型MHC抗原、KDR和VEGF,恰当的时候还包括其受体。

[0320] 可溶性抗原包括白介素如IL-1、IL-2、IL-3、IL-4、IL-5、IL-6、IL-8、IL-12、IL-13、IL-14、IL-16或IL-17,如IL17A和/或IL17F,病毒抗原如呼吸道合胞体病毒或巨细胞病毒抗原,免疫球蛋白如IgE,干扰素如α-干扰素、β-干扰素或γ-干扰素,肿瘤坏死因子TNF(之前称为肿瘤坏死因子-α),肿瘤坏死因子-β,集落刺激因子如G-CSF或GM-CSF,以及血小板衍生生长因子如PDGF-α和PDGF-β,恰当的时候还包括其受体。其他抗原包括细菌细胞表面抗原,细菌毒素,病毒如流感病毒、EBV、HepA、B和C,生化武器试剂,放射性核素和重金属,以及蛇和蜘蛛毒液及毒素。

[0321] 在一个实施方式中,抗体可用于功能性地改变目的抗原的活性。例如,抗体可直接或间接地中和、拮抗或抵消所述抗原的活性。

[0322] 在本发明的一个方面,提供了重组的革兰氏阴性细菌细胞,其包含编码突变 SPR 蛋白质的突变 SPR 基因,野生型 Tsp 基因以及编码特异于 TNF 的抗体或其抗原结合片段的多核苷酸徐良。野生型染色体 Tsp 基因优选地为非重组染色体 Tsp 基因。优选地,细胞进一步包含编码 DsbC 的重组多核苷酸。

[0323] 在一个优选的实施方式中,由根据本发明的细胞表达的目的蛋白质为抗-TNF 抗体,更优选地为抗-TNF Fab',如 WO01/094585 中所描述的(其内容并入本文作为参考)。

[0324] 在一个实施方式中,抗体具有针对人 TNF  $\alpha$  的特异性,其包含的重链中可变结构域包含的 CDR 具有 SEQ ID NO:28 所示的 CDRH1 序列,SEQ ID NO:29 或 SEQ ID NO:34 所示的 CDRH2 序列或 SEQ ID NO:30 所示的 CDRH3 序列。

[0325] 在一个实施方式中,抗体包含的轻链中的可变结构域包含的 CDR 具有 SEQ ID NO:31 所示的 CDRL1 序列,SEQ ID NO:32 所示的 CDRL2 序列或 SEQ ID NO:33 所示的 CDRL3 序列。

[0326] 上文提到的 SEQ IDS NOS:28 及 30-34 给出的 CDR 衍生自小鼠单克隆抗体 hTNF40。然而,SEQ ID NO:29 由杂合 CDR 构成。所述杂合 CDR 包括来自小鼠单克隆抗体 hTNF40 的重链 CDR2 的部分(SEQ ID NO:34)以及来自人 3 组种系区域 V 序列的重链 CDR2 的部分。

[0327] 在一个实施方式中,抗体包含重链和轻链,其重链中可变结构域包含的 CDR 具有 SEQ ID NO:28 所示的 CDRH1 序列,SEQ ID NO:29 或 SEQ ID NO:34 所示的 CDRH2 序列或 SEQ ID NO:30 所示的 CDRH3 序列;其轻链中可变结构域包含的 CDR 具有 SEQ ID NO:31 所示的 CDRL1 序列,SEQ ID NO:32 所示的 CDRL2 序列或 SEQ ID NO:33 所示的 CDRL3 序列。

[0328] 在一个实施方式中,抗体包含 SEQ ID NO:28 所示的 CDRH1,SEQ ID NO:29 或 SEQ ID NO:34 所示的 CDRH2,SEQ ID NO:30 所示的 CDRH3,SEQ ID NO:31 所示的 CDRL1,SEQ ID NO:32 所示的 CDRL2 以及 SEQ ID NO:33 所示的 CDRL3。优选地,抗体包含 SEQ ID NO:29 所示的 CDRH2。

[0329] 抗-TNF 抗体优选地为植入 CDR 的抗体分子。在一个优选的实施方式中,可变结构域包含人受体框架区及非人供体 CDR。

[0330] 优选地,抗体分子具有针对人 TNF(之前称为 TNF  $\alpha$ ) 的特异性的抗体分子,其中轻链包含 SEQ ID NO:11 的轻链可变区,其中重链包含 SEQ ID NO:12 的重链可变区。

[0331] 抗-TNF 抗体优选地为 Fab 或 Fab' 片段。

[0332] 优选地,具有针对人 TNF 的特异性的抗体分子是 Fab',并且具有的轻链序列包含或由 SEQ ID NO:13 组成,重链序列包含或由 SEQ ID NO:14 组成。

[0333] 在表达后,抗体片段可进行进一步加工,例如通过缀合至另外的实体,如效应子分子。

[0334] 例如,如本文所使用的术语效应子分子包括抗肿瘤试剂、药物、毒素(如细菌或植物来源的酶活性毒素或其片段,例如蓖麻毒素及其片段)、生物学活性蛋白质例如酶、其他的抗体或抗体片段、合成或天然发生的多聚体、核酸及其片段(如 DNA、RNA 及其片段)、放射性核素(尤其是放射性碘素、放射性同位素)、螯合金属、纳米颗粒、及报告子基团如荧光化合物或可由 NMR 或 ESR 光谱检测到的化合物。效应子分子可由任何合适的方法连接于

抗体或其片段,例如抗体片段可经过修饰连接于至少一种效应子分子,如 W005/003171 或 W005/003170(其内容并入本文作为参考)所描述的。W005/003171 或 W005/003170 也描述了合适的效应子分子。

[0335] 在一个实施方式中,如果有需要,抗体或其片段如 Fab 是 PEG 化的,以产生具有所需(例如与完全抗体类似)特性的产物。例如,抗体可以为 PEG 化的抗-TNF- $\alpha$  Fab',如 W001/094585 中所描述的,优选地其在重链 C 末端的一个半胱氨酸残基上连接有赖氨酰-顺丁烯二酰亚胺-衍生基团,其中赖氨酸残基的两个氨基基团中每个都与具有大约 20,000Da 的分子量的甲氧基聚(乙二醇)残基相连,这样甲氧基聚(乙二醇)残基的总平均分子量为大约 40,000Da,更优选地,所述赖氨酰-顺丁烯二酰亚胺-衍生基团为 [1-[[[2-[[3-(2,5-二氧-1-吡咯烷基)-1-氧丙基]氨基]乙基]氨基]-羧基]-1,5-戊烷二基]二(亚胺羰基)。

[0336] 细胞还可包含编码一种或多种其他目的蛋白质的其他多核苷酸序列。

[0337] 在一个实施方式中,选择一种或多种在野生型时已知与目的重组蛋白质在纯化中共纯化的大肠杆菌宿主蛋白质进行遗传修饰,如描述于 Humphreys 等人“Engineering of Escherichia coli to improve the purification of periplasmic Fab' fragments:changing the pI of the chromosomally encoded PhoS/PstS protein”,Protein Expression and Purification 37(2004)109-118 和 W004/035792(其内容并入本文作为参考)。使用这些修饰的宿主蛋白质改善了目的蛋白质(尤其是抗体)的纯化过程,是通过改变所选的大肠杆菌蛋白质的物理特性使之不再与重组抗体共纯化而在大肠杆菌中生产。优选地,受到改变的大肠杆菌蛋白质选自一种或多种磷酸结合蛋白质(PhoS/PstS)、二肽结合蛋白质(DppA)、麦芽糖结合蛋白质(MBP)和硫氧还蛋白质。

[0338] 在一个实施方式中,污染性宿主蛋白质的物理特性通过向 C 端或 N 端添加氨基酸标签得到改变。在一个优选的实施方式中,改变的物理属性为等电点,氨基酸标签为连接于 C 末端的多天冬氨酸标签。在一个实施方式中,通过添加所示标签得到改变的大肠杆菌蛋白质为二肽结合蛋白质(DppA)、麦芽糖结合蛋白质(MBP)、硫氧还蛋白和磷酸结合蛋白质(PhoS/PstS)。在一个特异性实施方式中,大肠杆菌磷酸结合蛋白质(PhoS/PstS)的 pI 通过添加多天冬氨酸标签(polyD)由 7.2 减少至 5.1,其 C 端包含了 6 个天冬氨酸的标签。

[0339] 还优选地是对污染性大肠杆菌蛋白质的特定残基进行修饰以改变其物理特性,不管是单独修饰还是与 N 或 C 末端添加标签相结合。这些改变可包括插入或删除以改变蛋白质大小,或氨基酸置换以改变 pI 或疏水性。在一个实施方式中,这些残基位于蛋白质表面。在一个优选的实施方式中,改变了 PhoS 蛋白质表面的残基以降低蛋白质的 pI。优选地,所提到的残基对于磷酸结合(Bass, US5, 304, 472)很重要,因此为了维持功能性 PhoS 蛋白质要尽量避免。优选地,靶向的是那些伸出蛋白质表面很远或在碱性残基大基团之内或附近的赖氨酸残基。在一个实施方式中,PhoS 蛋白质具有连接于 C 末端的六聚多天冬氨酸标签,而分子相反的末端上帝表面残基用于置换。优选地,所选的赖氨酸残基置换为谷氨酸或天冬氨酸以形成比将中性残基变为酸性残基更大的 pI 电势变化。本文对于置换突变体的标记由字母+数字+字母构成。第一个字母标记野生型蛋白质中的氨基酸。数字指的是做出氨基酸置换的氨基酸位置,第二个字母标记用于置换野生型氨基酸的氨基酸。在本发明优选的 PhoS 突变中,275、107、109、110、262、265、266、309、313 号赖氨酸残基(K)以单一或组

合突变的形式置换为谷氨酸(E)或谷氨酰胺(Q),另外,318号赖氨酸(K)可以单一或组合突变的形式置换为天冬氨酸(D)。优选地,单一突变为K262E、K265E和K266E。优选地,组合突变为K265/266E和K110/265/266E。更优选地,所有突变都与连接于C端的多天冬氨酸(polyD)标签相组合,优选地还与K318D置换相组合。在一个优选的实施方式中,这些突变导致pI降低至少2个单位。优选地,本发明的突变将PhoS的pI从7.2降低至大约4-大约5.5。在本发明的一个实施方式中,使用突变polyD K318D、polyD K265/266E和polyD K110/265/266E分别将大肠杆菌PhoS蛋白质的pI从7.2降低至大约4.9、大约4.8以及大约4.5。

[0340] 编码目的蛋白质的多核苷酸可以以与另一种多肽融合的形式表达,另一种多肽优选地为信号序列或其他在成熟多肽的N端有特异性剪切位点的多肽。所选的异源信号序列应当为可由宿主细胞识别并加工的序列。对于不能识别和加工原生或真核多肽信号序列的原核宿主细胞,信号序列由原核信号序列置换。合适的信号序列包括OmpA、PhoA、LamB、PelB、DsbA和DsbC。

[0341] 只要相关方面能适用于这些实施方式,则本文描述的本发明的实施方式关于所述多核苷酸方面对本发明备选实施方式同等适用,例如载体、表达盒及/或包含其中所采用的组分的宿主细胞。

[0342] 本发明还提供了用于生产目的重组蛋白质的方法,包括在有效表达目的重组蛋白质的条件下,于培养基中培养如上文所描述的革兰氏阴性细菌细胞,以及从重组革兰氏阴性细菌细胞的细胞周质和/或培养基中回收所述目的重组蛋白质。在一个实施方式中,其中细胞包含编码DsbC的重组多核苷酸,细胞是在有效表达编码DsbC的重组多核苷酸的条件下进行培养的。

[0343] 本发明的方法中优选采用的革兰氏阴性细菌细胞及目的蛋白质已在上文详细描述。

[0344] 当编码目的蛋白质的多核苷酸是外源时,多核苷酸可使用任何本领域已知的合适方式并入宿主细胞。编码DsbC的多核苷酸序列也可使用任何本领域已知的合适的方式并入宿主细胞。一般地,多核苷酸作为转化入细胞的表达载体的一部分而并入。因此,在一方面,根据本发明的细胞包含含有编码目的蛋白质的多核苷酸的表达盒以及含有编码DsbC的多核苷酸的表达盒。

[0345] 编码目的蛋白质的多核苷酸序列以及编码DsbC的多核苷酸序列可使用标准技术转化入细胞,如使用氯化铷、PEG或电穿孔方法。

[0346] 根据本发明的方法还可采用选择系统以促进对成功转化有编码目的蛋白质的多核苷酸的稳定细胞进行选择。选择系统一般采用编码选择性标记物的多核苷酸共转化。在一个实施方式中,每种转化入细胞的多核苷酸进一步包含编码一种或多种选择性标记物的多核苷酸序列。因此,转化编码目的蛋白质的多核苷酸和任选的编码DsbC的多核苷酸以及编码标记物的一种或多种多核苷酸可同时发生,然后采用选择系统以选择产生所需的蛋白质的细胞。

[0347] 能够表达所述一种或多种标记物的细胞能够在特定人工设置的条件下存活/生长/扩增,例如在添加毒素或抗生素条件下,因为其中并入的多肽/基因或选择系统的多肽组分提供了某些特性(例如,抗生素抗性)。不能表达所述一种或多种标记物的那些细胞则

不能在这些人工设置的条件下存活 / 生长 / 扩增。可根据需要选择更严格或不那么严格的人工设置条件。

[0348] 本发明可采用任何合适的选择系统。一般地,选择系统可基于包括在载体中的一种或多种提供对已知抗生素抗性的基因,例如四环素、氯霉素、卡那霉素或氨基青霉素基因。选择在相关抗生素存在下生长的细胞,因为其表达提供对抗生素抗性的基因以及所需要的蛋白质。

[0349] 在本发明中可使用可诱导表达系统或组成型启动子以表达目的蛋白质及 / 或 DsbC。在一个实施方式中,通过向培养基中添加诱导子,诱导编码目的蛋白质的多核苷酸序列以及编码 DsbC 的重组多核苷酸的表达。合适的可诱导表达系统以及组成型启动子在本领域熟知。

[0350] 可使用任何适合的培养基培养转化的细胞。可为具体的选择系统采用恰当的培养基,例如培养基可以包含抗生素以使只有已经成功转化的细胞能够在培养基中生长。

[0351] 根据需要,可对获自培养基的细胞进行进一步的选择和 / 或纯化。根据需要,此方法可进一步包括一个或多个提取及纯化目的蛋白质的步骤。

[0352] 可从菌株中回收多肽,包括从细胞质、细胞周质及 / 或上清中。

[0353] 用于纯化蛋白质的具体方法取决于蛋白质类型。合适的方法包括在免疫亲和或离子交换柱上的分级;乙醇沉淀;反相 HPLC; 疏水相互作用色谱;硅色谱;离子交换树脂色谱如 S-SEPHAROSE 和 DEAE; 层析聚焦;硫酸铵沉淀;以及凝胶过滤。

[0354] 在一个实施方式中,所述方法进一步包括将目的重组蛋白质与 DsbC 分离。

[0355] 可通过常规抗体纯化步骤,合适地将抗体与培养基和 / 或细胞质提取物和 / 或细胞周质提取物分离,例如,通过蛋白质 A- 琼脂糖凝胶、蛋白质 G 色谱,蛋白质 L 色谱,嗜硫的、混合形式的树脂,His- 标签,FLAG 标签,羟磷灰石色谱,凝胶电泳,透析,亲和色谱,硫酸铵、乙醇或 PEG 分馏 / 沉淀,离子交换膜,扩张的床吸附色谱 (EBA) 或模拟移动床色谱。

[0356] 所述方法还可以包括进一步的步骤,测量目的蛋白质表达量以及选择具有高表达水平的目的蛋白质的细胞。

[0357] 此方法还可包括一个或多个其他下游加工步骤,如对目的蛋白质如抗体或抗体片段进行 PEG 化。

[0358] 本文所描述的一个或多个方法步骤可在合适的容器如生物反应器中组合进行。

## 实施例

[0359] 实施例 1——产生细胞菌株 MXE001 ( $\Delta$  Tsp)

[0360] 通过如下方法产生 MXE001 菌株:

[0361] 将 Tsp 盒以 Sal I、Not I 限制性片段的形式移至进行了类似的限制性酶切的 pK03 质粒内。pK03 质粒使用 pSC101 的温度敏感型突变体的复制起点 (RepA) 以及氯霉素标记物以加强和选择染色体整合事件。编码果聚糖蔗糖酶的 sacB 基因对生长于蔗糖上的大肠杆菌是致死的,因此(与氯霉素标记物及 pSC101 起点一同)用于加强并选择去整合以及质粒消除。此方法理论在之前有所描述 (Hamilton 等人, 1989, Journal of Bacteriology, 171, 4617-4622 以及 Blomfield 等人, 1991, Molecular Microbiology, 5, 1447-1457)。除了插入的基因之外, pK03 系统将所有选择性标记物从宿

主基因组移除。

[0362] 构建了以下质粒。

[0363] pMXE191:包含如 SEQ ID NO:3 所示的敲除突变 Tsp 基因,其包含 EcoRI 和 AseI 限制性标记物。

[0364] 然后将质粒转化入电感受态大肠杆菌 W3110 细胞,其使用 Methods in Molecular Biology, vol. 47, Nickoloff, J. A. (ed.), Humana Press, Totowa, NJ, 105 (1995) 中 Miller, E. M. 和 Nickoloff, J. A., “Escherichia coli electrotransformation” 的方法进行制备。

[0365] 第 1 天将 40  $\mu$  l 的大肠杆菌细胞与 (10pg) 1  $\mu$  l 的 pK03DNA 混合于冰冷的 BioRad 0.2cm 电穿孔小杯,然后以 2500V、25  $\mu$  F、200  $\Omega$  进行电穿孔。立即加入 1000  $\mu$  l 的 2xPY,细胞在培养箱中 30°C 下、250rpm 摇 1 小时进行复苏。将细胞进行 1/10 的系列稀释在 2xPY 中,然后将 100  $\mu$  l 的等分试样铺于 30°C 和 43°C 下预热的、含有 20  $\mu$  g/ml 氯霉素的 2xPY 琼脂平板。将平板在 30°C 和 43°C 下孵育过夜。

[0366] 第 2 天在 30°C 下生长的菌落数量给出电穿孔效率的估算值,而在 43°C 下存活生长的菌落代表潜在的整合事件。挑取 43°C 平板上的单一菌落并重悬于 10ml 的 2xPY 中。将 100  $\mu$  l 的悬液铺于含有 5% (w/v) 蔗糖并预热到 30°C 的 2xPY 琼脂平板上以产生单菌落。将平板在 30°C 下孵育过夜。

[0367] 第 3 天此时的菌落代表了潜在的自发去整合及质粒消除的事件。如果去整合和消除事件在生长早期发生,则所述克隆实体的大部分是纯系的 (clonal)。挑取单一菌落并将其复制物铺于含有 20  $\mu$  g/ml 氯霉素或 5% (w/v) 蔗糖的 2xPY 琼脂平板。将平板在 30°C 下孵育过夜。

[0368] 第 4 天在蔗糖上生长且在氯霉素中死亡的菌落代表了潜在的染色体置换质粒消除事件。挑取这些菌落并通过使用突变特异性寡核苷酸的 PCR 进行筛选。将产生具有正确大小阳性 PCR 条带的菌落划线至含有 5% (w/v) 蔗糖的 2xPY 琼脂平板上并将平板在 30°C 孵育过夜产生单一菌落。

[0369] 第 5 天使用 PCR 阳性、氯霉素敏感型以及蔗糖抵抗型大肠杆菌的单菌落制作甘油贮备株,制作化学感受态细胞并用作 PCR 模板用于使用 5' 和 3' 两侧的寡核苷酸的 PCR 反应,以产生 PCR 产物用于使用 Taq 聚合酶的直接的 DNA 测序。

[0370] 通过对包含非天然发生 Ase I 限制性位点的 Tsp 基因区域 (如图 1a、1b、1c 所示) 使用寡核苷酸引物进行 PCR 扩增,以测试细胞菌株 MXE001 以确认基因组 DNA 是否成功修饰为携带突变 Tsp 基因。随后将扩增的 DNA 区域与限制性酶 Ase I 孵育,然后通过琼脂糖凝胶电泳分析孵育之前和孵育之后 DNA 所扩增的区域以确认非天然发生的 Ase I 限制性位点是否存在于突变的基因中。此方法按如下步骤进行:

[0371] 使用 PCR 及以下寡核苷酸对从 MXE001 和 W3110 的大肠杆菌细胞裂解物的制备的基因组 DNA 进行扩增:

[0372] 6284Tsp3' 5' -GCATCATAATTTTCTTTTACCTC-3' (SEQ ID NO:15)

[0373] 6283Tsp5' 5' -GGGAAATGAACCTGAGCAAAACGC-3' (SEQ ID NO:16)

[0374] 通过将单克隆细胞在 20ul 1x PCR 缓冲液中于 95°C 下加热 10 分钟制备裂解物。使混合物冷却至室温,然后在 13,200rpm 下离心 10 分钟。移除上清并贴上“细胞裂解物”的

标签。

[0375] 使用 Tsp 寡核苷酸对扩增各菌株。

[0376] 使用标准 PCR 步骤扩增 DNA。

[0377]

<b>5ul</b>	<b>缓冲液 x10 (Roche)</b>
<b>1ul</b>	<b>dNTP 混合物 (Roche, 10mM mix)</b>
<b>1.5ul</b>	<b>5'寡核苷酸 (5 pmol)</b>
<b>1.5ul</b>	<b>3'寡核苷酸 (5 pmol)</b>
<b>2ul</b>	<b>细胞裂解物</b>
<b>0.5ul</b>	<b>Taq DNA 聚合酶 (Roche 5U/ul)</b>
<b>38.5ul</b>	<b>H2O</b>

[0378] PCR 循环。

[0379]

<b>94 °C</b>	<b>1 分钟</b>
<b>94 °C</b>	<b>1 分钟)</b>
<b>55 °C</b>	<b>1 分钟) 重复进行 30 个循环</b>
<b>72 °C</b>	<b>1 分钟)</b>
<b>72 °C</b>	<b>10 分钟</b>

[0380] 反应完成时,将 25ul 移至新离心管用 Ase I 进行消化。向 25ul 的 PCR 反应物中加入 19ul 的 H2O、5ul 缓冲液 3 (NEB)、1ul 的 Ase I (NEB),混合,在 37°C 下孵育 2 小时。

[0381] 向余下的 PCR 反应物中加入 5ul 上样缓冲液(x6),取出 20ul 上样于 0.8%TAE 200ml 琼脂糖凝胶(Invitrogen)加溴化乙锭(5ul 的 10mg/ml 储备液),并于 100 伏电泳 1 小时。最后一个泳道上样 10ul 的分子量标记物(Perfect DNA marker 0.1-12Kb, Novagen)。

[0382] 在 Ase I 消化完成时,加入 10ul 上样缓冲液(x6),取 20ul 上样于 0.8%TAE 琼脂糖凝胶(Invitrogen)加溴化乙锭(5ul 的 10mg/ml 储备液)并于 100 伏下电泳 1 小时。最后一个泳道加入了 10ul 的分子量标记物(Perfect DNA marker 0.1-12Kb, Novagen)。两块胶都使用 UV 透照仪进行观察。

[0383] 扩增的基因组片段显示 Tsp 的正确条带大小,2.8kb。以 Ase I 消化后确认了在 Tsp 缺陷菌株 MXE001 中存在有引入的 Ase I 限制性位点,而在 W3110 对照中没有。

[0384] MXE001: 使用 Tsp 引物组扩增的基因组 DNA,并将扩增所产生的 DNA 用 Ase I 消化,产生 2.2 和 0.6Kb 的条带。

[0385] W3110PCR 扩增的 DNA 无法由 Ase I 限制性酶进行消化。

[0386] 实施例 2 - 产生 spr 突变体

[0387] 使用互补测定法产生并选择 spr 突变。

[0388] 使用 Clontech® 随机诱变多样性 PCR 试剂盒将 spr 基因突变,该试剂盒在每 1000bp 引入 1-2 个突变。将突变 SPR PCR DNA 克隆入表达 CDP870Fab' 及 spr 突变体



的可诱导表达载体 [pTTO CDP870]。然后将此连接产物电转化至使用 Miller, E. M. 和 Nickoloff, J. A., “Escherichia coli electrotransformation,” 在 *Methods in Molecular Biology*, vol. 47, Nickoloff, J. A. (ed.), Humana Press, Totowa, NJ, 105 (1995) 中的方法制备的大肠杆菌菌株 MXE001 ( $\Delta$ Tsp)。使用以下的操作流程: 将 40ul 的电感受态 MXE001、2.5ul 的连接产物 (100pg 的 DNA) 加入 0.2cm 的电穿孔小杯, 使用 BioRad Genepulser Xcell 在 2500V、25  $\mu$ F、200  $\Omega$  的条件下进行电转化。电转化后, 加入 1ml 的 SOC (Invitrogen) (预热至 37°C) 并对细胞轻柔摇动, 使之在 37°C 下复苏 1 小时。

[0389] 将细胞铺于低渗琼脂糖平板 (5g/L 酵母提取物、2.5g/L 蛋白胨、15g/L 琼脂 (皆为 Difco) 并于 40°C 孵育。将形成菌落的细胞在 43°C 下重新铺于 HLB 以确认 MXE001 菌株重新获得了在高温低渗条件下生长的能力。从所选的克隆制备质粒 DNA 并对其测序以鉴定 spr 突变。

[0390] 使用此方法分离出与  $\Delta$ Tsp 表型互补的 spr 蛋白质的八个单一突变、一个双突变以及两个多重突变, 如下:

[0391] 1. V98E

[0392] 2. D133A

[0393] 3. V135D

[0394] 4. V135G

[0395] 5. G147C

[0396] 6. S95F 和 Y115F

[0397] 7. I70T

[0398] 8. N31T、Q73R、R100G、G140C

[0399] 9. R62C、Q99P、R144C

[0400] 10. L108S

[0401] 11. L136P

[0402] 实施例 3 产生携带 spr 突变的突变体大肠杆菌细胞菌株

[0403] 使用实施例 2 中鉴定的 spr 的 1-5 号个体突变以及三个催化三分子突变 (C94A, H145A, H157A) 和 W174R 产生新菌株, 使用野生型 W3110 大肠杆菌菌株 (基因型: F-LAM-IN (rrnD, rrnE) lrph1 (ATCC 号 27325)) 产生携带野生型非重组染色体 Tsp 基因的 spr 突变菌株或使用来自实施例 1 的 MXE001 ( $\Delta$ Tsp) 菌株产生  $\Delta$ Tsp/ 突变 SPR 组合突变菌株。

[0404] 以下突变体大肠杆菌细胞菌株是使用基因置换载体系统产生的, 使用的是 pK03 同源重组 / 置换质粒 (Link 等人, 1997, *Journal of Bacteriology*, 179, 6228-6237), 如实施例 1 中用于产生 MXE001 的方法。

[0405] 表 1.

[0406]

突变体大肠杆菌细胞菌株	基因型	Spr 载体
MXE001	$\Delta$ Tsp	-

MXE008	$\Delta$ Tsp, spr D133A	pMXE339, pK03 spr D133A(-SalI)
MXE009	$\Delta$ Tsp, spr H157A	pMXE345, pK03 spr H157A(-SalI)
MXE010	spr G147C	pMXE338, pK03 spr G147C(-SalI)
MXE011	spr C94A	pMXE343, pK03 spr C94A(-SalI)
MXE012	spr H145A	pMXE344, pK03 spr H145A(-SalI)
MXE013	spr W174R	pMXE346, pK03 spr W174R(-SalI)
MXE014	$\Delta$ Tsp, spr V135D	pMXE340, pK03 spr V135D(-SalI)
MXE015	$\Delta$ Tsp, spr V98E	pMXE342, pK03 spr V98E(-SalI)
MXE016	$\Delta$ Tsp, spr C94A	pMXE343, pK03 spr C94A(-SalI)
MXE017	$\Delta$ Tsp, spr H145A	pMXE344, pK03 spr H145A(-SalI)
MXE018	$\Delta$ Tsp, spr V135G	pMXE341, pK03 spr V135G(-SalI)

[0407]

[0408] 突变 SPR 整合盒以 SalI、Not I 限制性片段形式移至进行了类似限制性酶切的 pK03 质粒。

[0409] 所述质粒使用 pSC101 的温度敏感型突变体的复制起点 (RepA) 以及氯霉素标记物以加强并选择染色体整合时间。编码果聚糖蔗糖酶的 sacB 基因对生长于蔗糖上的大肠杆菌是致死的, 因此(与氯霉素标记物及 pSC101 起点一同) 用于加强和选择去整合以及质粒消除事件。此方法理论在之前有所描述 (Hamilton 等人, 1989, Journal of Bacteriology, 171, 4617-4622 and Blomfield 等人, 1991, Molecular Microbiology, 5, 1447-1457)。除了插入的基因之外, pK03 系统将所有选择性标记物从宿主基因组移除。

[0410] 构建了以下 pK03 载体, 其包含突变 SPR 基因, 包括 spr 序列内移除了用于克隆鉴定的 SalI 限制性位点的沉默突变。

[0411] pMXE336, pK03 spr S95F(-SalI)

[0412] pMXE337, pK03 spr Y115F(-SalI)

[0413] pMXE338, pK03 spr G147C(-SalI)

[0414] pMXE339, pK03 spr D133A(-SalI)

[0415] pMXE340, pK03 spr V135D(-SalI)

[0416] pMXE341, pK03 spr V135G(-SalI)

[0417] pMXE342, pK03 spr V98E(-SalI)

[0418] pMXE343, pK03 spr C94A(-SalI)

[0419] pMXE344、pK03 spr H145A(-SalI)

[0420] pMXE345、pK03 spr H157A(-SalI)

[0421] pMXE346、pK03 spr W174R(-SalI)

[0422] 然后使用 *Methods in Molecular Biology*, vol. 47, Nickoloff, J. A. (ed.), Humana Press, Totowa, NJ, 105 (1995) 中 Miller, E. M. 和 Nickoloff, J. A., “*Escherichia coli* electrotransformation” 的方法将这些质粒转化入化学感受态大肠杆菌 W3110 细胞中, 或转化入来自实施例 1 的 MXE001 菌株以产生  $\Delta T_{sp}$ / 突变 SPR 组合突变菌株, 如表 1 所示。

[0423] 第 1 天将 40  $\mu$  l 的电感受态大肠杆菌细胞或 MXE001 细胞与 (10pg) 1  $\mu$  l 的 pK03DNA 混合于冰冷的 BioRad 0.2cm 电穿孔小杯, 然后以 2500V、25  $\mu$  F、200  $\Omega$  进行电穿孔。立即加入 1000  $\mu$  l 的 2xPY, 细胞在培养箱中 30°C 下、250rpm 摇 1 小时进行复苏。将细胞进行 1/10 的系列稀释在 2xPY 中, 然后将 100  $\mu$  l 的等分试样铺于 30°C 和 43°C 下预热的含有 20  $\mu$  g/ml 氯霉素的 2xPY 琼脂平板。将平板在 30°C 和 43°C 下孵育过夜。

[0424] 第 2 天在 30°C 下生长的菌落数量给出电穿孔效率的估算值, 而在 43°C 下存活生长的菌落代表潜在的整合事件。挑取 43°C 平板上的单一菌落并重悬于 10ml 的 2xPY。将 100  $\mu$  l 的悬液铺于含有 5% (w/v) 蔗糖并预热到 30°C 的 2xPY 琼脂平板上以产生单克隆。将平板在 30°C 孵育过夜。

[0425] 第 3 天此时的菌落代表了潜在的自发去整合及质粒消除的事件。如果去整合和消除事件在生长早期发生, 则所述克隆实体的大部分是纯系的。挑取单一菌落并将其复制物铺于含有 20  $\mu$  g/ml 氯霉素或 5% (w/v) 蔗糖的 2xPY 琼脂平板。将平板在 30°C 下孵育过夜。

[0426] 第 4 天在蔗糖上生长且在氯霉素中死亡的菌落代表了潜在的染色体置换和质粒消除事件。挑取这些菌落并以进行 PCR 筛选且进行限制性消化检测 SalI 位点的丧失情况。将产生具有正确大小阳性 PCR 条带并抵抗 SalI 消化的菌落划线至含有 5% (w/v) 蔗糖的 2xPY 琼脂平板上并将平板在 30°C 孵育过夜产生单一菌落。

[0427] 第 5 天使用 PCR 阳性、氯霉素敏感型以及蔗糖抵抗型大肠杆菌的单菌落制作甘油贮藏株, 制作化学感受态细胞并用作 PCR 模板用于使用 5' 和 3' 两侧的寡核苷酸的 PCR 反应, 以产生 PCR 产物用于使用 Taq 聚合酶的直接的 DNA 测序, 来验证正确的突变。

[0428] 实施例 4- 产生用于 Fab' 和 DsbC 共表达的质粒

[0429] 构建了包含抗 -TNF Fab' (具有如 SEQ ID NO:13 所示轻链以及如 SEQ ID NO:14 所示的重链的抗 -TNF Fab') 的重链和轻链序列以及编码 DsbC 序列的质粒。

[0430] W001/94585 所描述的抗 -TNF  $\alpha$  Fab' 片段 (称为 CDP870) 创造了双顺反子信息。上游顺反子编码抗体轻链 (SEQ ID NO:13) 而下流的顺反子编码了抗体重链 (SEQ ID NO:14)。编码 OmpA 信号肽的 DNA 序列融合于编码轻链和重链的 DNA 各自的 5' 末端以使之有效地分泌入细胞周质。所使用的为如 SEQ ID NO:37 所示的基因间序列 (IGS2)。

[0431] 使用见于 Sambrook 等人 1989, *Molecular cloning: a laboratory manual*. CSHL press, N. Y 的常规限制性克隆方法理论构建了质粒 pDPH358 (pTT0 40.4CDP870IGS2), 其为 CDP870Fab' (抗 -TNFFab') 和 DsbC (细胞周质多肽) 的表达载体。质粒 pDPH358 包含以下特征: 强启动子 tac 及 lac 操纵子序列。如图 10 所示, 该质粒包含位于 Fab' 重链编码区之后的唯一的 EcoRI 限制性位点, 随后为非编码序列, 接着为唯一的 NdeI 限制性位点。DsbC 基因是使用 W3110 原始染色体 DNA 作为模板进行 PCR 克隆产生的, 这样 PCR 产物编码了

5' EcoRI 位点,随后为强的核糖体结合序列、再随后为原生起始密码子、信号序列以及 DsbC 成熟序列,以 C 端 His 标签终止,最后为非编码的 NdeI 位点。对此 EcoRI-NdeI PCR 片段进行限制性酶切并连接入表达载体使得所有三种多肽——Fab' 轻链、Fab' 重链及 DsbC 在单个多顺反子 mRNA 上编码。

[0432] Fab 轻链、重链基因及 DsbC 基因作为单个多顺反子信息转录。编码来自大肠杆菌 OmpA 蛋白质的信号肽(其引导多肽易位进入大肠杆菌细胞周质)的 DNA 基因与轻链和重链基因序列各自的 5' 末端融合。使用双重转录终止子 rrnB t1t2 终止转录。lacIq 基因编码组成型表达的 Lac I 抑制子蛋白质。这样,tac 启动子的转录一直受到抑制直到加入异乳糖或 IPTG 来诱导去抑制。所使用的复制起点为 p15A,这维持了低的拷贝数量。质粒包含四环素抗性基因以用于抗生素筛选。

[0433] 实施例 5 - 抗-TNF Fab' 或抗-TNF Fab' 与 DsbC 在大肠杆菌菌株中的表达

[0434] 抗-TNF Fab' 与 DsbC 的表达

[0435] 用实施例 4 中产生的质粒转化野生型 W3110 细胞系,实施例 1 中提供的 MXE001 菌株及实施例 3 中提供的突变体菌株 MXE012 (H145A spr 突变体菌株)。

[0436] 使用 Chung C.T 等人 Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989) 的方法对菌株进行转化。

[0437] 抗-TNF Fab' 的表达

[0438] 如实施例 4 所描述的,用质粒 pMXE117(pTTO CDP870 或 40.4IGS17) 转化野生型 W3110 细胞系、实施例 3 中提供的 spr 突变体菌株 MXE008、MXE012、MXE017 和 MXE012(H145A spr 突变体菌株)以及实施例 1 中提供的 MXE001 菌株,该质粒为 CDP870Fab' (一种具有如 SEQ ID NO:13 所示的轻链序列及如 SEQ ID NO:14 所示的重链序列的抗-TNF Fab') 表达载体,是使用见于 Sambrook 等人 1989, Molecular cloning: a laboratory manual. CSHL press, N.Y 的常规限制性克隆方法理论构建的。质粒 pMXE117(pTTO CDP870 或 40.4IGS17) 包含以下特征:强启动子 tac 及 lac 操纵子序列。Fab 轻链和重链基因转录为单个双顺反子信息。编码来自大肠杆菌 OmpA 蛋白质的信号肽(其引导多肽易位进入大肠杆菌细胞周质)的 DNA 基因与轻链和重链基因序列各自的 5' 末端融合。使用双重转录终止子 rrnB t1t2 终止转录。lacIq 基因编码组成型表达的 Lac I 抑制子蛋白质。这样,tac 启动子的转录一直受到抑制直到加入异乳糖或 IPTG 而诱导去抑制。所使用的复制起点为 p15A,这维持了低的拷贝数量。质粒包含四环素抗性基因以用于抗生素筛选。

[0439] 使用 Chung C.T 等人 Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989) 中的方法对菌株进行转化。

[0440] 实施例 6 - 用摇瓶培养在突变的大肠杆菌菌株中表达抗-TNF Fab'

[0441] 在摇瓶实验测试以下如实施例 5 中产生的表达抗-TNF Fab' 的菌株以比较 Fab' 的生长和表达:W3110、MXE001、MXE012 和 MXE017。

[0442] 所使用的摇瓶实验操作流程如下进行:

[0443] 5ml 摇瓶实验

[0444] 挑取单菌落接入 5ml 含有 10ug/ml 四环素的 LB 中并在 30°C 下以 250rpm 摇动生长过夜。

[0445] 用过夜培养物接种至 100ml 含有四环素的培养基中至 0.1 OD<sub>600</sub>。(即,对于 4 的

OD, 100/4x01=100ml 中加入 2.5mls)

[0446] 使用此主培养物为每个所需的时间点设置 3x5ml 的培养管。取 1 个参考培养物作为测量 OD 的样品。

[0447] 在 30°C 下以 250rpm 摇动培养物, 起初通过观察检测生长概况, 然后取参考培养物样品获得 0.5 OD<sub>600</sub> 的培养物 (通常大约 2 小时)。一旦培养物达到超过 0.5OD 时, 将 IPTG 加入各培养管至浓度为 200uM (0.04M 取 25ul)。

[0448] 在所需的时间点例如 1 小时、2 小时、诱导后取出培养管并置于冰上。

[0449] 在以 13, 200rpm 离心 5 分钟后, 用 200ul 的细胞周质提取缓冲液 (100mM Tris, Cl/10mM EDTA pH 7.4) 重悬细胞沉淀。将细胞周质提取物在 30°C 下以 250rpm 摇动过夜。第二天, 将提取物在 13, 200rpm 下离心 10 分钟, 弃去上清并保存于 -20°C, 标记为“细胞周质提取物”。弃去用过的细胞沉淀。

[0450] ELISA 定量。

[0451] 在 4°C 下用 PBS 中 2 μ gml<sup>-1</sup> 的 AB141 (兔抗人 CH1, UCB) 将 96 孔 ELISA 平板包被过夜。在以 300ul 的样品 / 结合物缓冲液 (PBS、BSA 0.2% (w/v)、Tween 20 0.1% (v/v)) 洗涤 3 次后, 在平板上用 100 μ l 的样品 / 结合物缓冲液对样品和标准品进行 1/2 的系列稀释, 将平板于室温下以 250r. p. m 摇动 1 小时。以 300ul 的洗涤缓冲液 (PBS, Tween 20 0.1% (v/v)) 洗涤 3 次后, 以样品 / 结合物缓冲液进行 1/1000 的稀释, 加入 100 μ l 的展示抗体 6062 (缀合的兔抗人 κ HRP, The Binding Site, Birmingham, U. K.)。然后将平板于室温下以 250r. p. m 摇动 1 小时。以 300ul 的洗涤缓冲液洗涤 3 次后, 加入 100 μ l 的 TMB 底物 (TMB 溶液 (Calbiochem) : dH<sub>2</sub>O 的 50:50 混合物) 并使用自动平板读数仪记录 A<sub>630</sub>。细胞周质提取物中 Fab' 的浓度通过与纯化的合适同种型 Fab' 标准品比较而计算。

[0452] 图 1 显示相比于野生型 W3110 和 MXE001, MXE012 和 MXE017 的生长得到改善。

[0453] 图 2 显示相比于野生型 W3110 和 MXE001, MXE012 和 MXE017 的 Fab' 的表达得到改善。

[0454] 实施例 7 - 使用高密度发酵培养表达抗 -TNF Fab' 或抗 -TNF Fab' 及 DsbC 的大肠杆菌菌株

[0455] 在发酵实验中测试如实施例 5 中所产生的下列菌株以比较其生长和抗 -TNF α Fab' 的表达:

[0456] 实施例 5 中产生的表达抗 -TNF Fab' 的菌株:

[0457] W3100

[0458] MXE012 (H145A spr 突变体菌株)

[0459] 实施例 5 中产生的表达抗 -TNF Fab' 及 DsbC 的菌株:

[0460] W3110

[0461] MXE012 (H145A spr 突变体菌株)

[0462] 生长培养基

[0463] 发酵生长培养基基于 SM6E 培养基 (描述于 Humphreys 等人, 2002, Protein Expression and Purification, 26, 309-320), 添加了 3.86g/l NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O 及 112g/l 的甘油。

[0464] 接种。接种培养物生长于相同的培养基, 其中补充了 10 μ g/ml 的四环素。培养物

在 30°C 下搅拌孵育约 22 小时。

[0465] 发酵。用接种培养物接种入发酵罐(总体积 2.5 升)达到 0.3-0.5OD<sub>600</sub>。在生长期将温度维持在 30°C,到诱导前降低至 25°C。通过不断变化搅拌和气流,使溶解氧浓度控制在超过 30% 的空气饱和度。通过以 15%(v/v)NH<sub>4</sub>OH 和 10%(v/v) 浓度的 H<sub>2</sub>SO<sub>4</sub> 自动滴定使培养物的 pH 控制在 7.0。通过添加 10%(v/v)Strukto1 J673 溶液 (Schill 和 Seilacher) 控制泡沫产生。

[0466] 在发酵不同阶段添加不同的添加物。当生物质浓度达到大约 400D<sub>600</sub> 时,添加镁盐及 NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O。在诱导期之前和之内进一步添加 NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O 以保证磷酸盐过量。当发酵开始存在的甘油耗尽后(大约 75 OD<sub>600</sub>),连续进料 80%(w/w) 的甘油。在发酵中所述相同时间点以 170 μM 进行 IPTG 进料。IPTG 进料的开始被认为是诱导的开始。一般地,发酵以甘油进料速率(范围为 0.5-2.5ml/h)进行 64-120 小时。

[0467] 生物质浓度和生长速率的测量。生物质浓度是通过测量 600nm 处培养物的光学密度而确定的。

[0468] 细胞周质提取。通过离心从培养物样品中收集细胞。保留上清级分(于 -20° C)用于进一步分析。细胞沉淀级分以原始培养物体积重悬于提取缓冲液(100mM Tris-HCl, 10mM EDTA;pH 7.4)。在 60° C 下孵育约 16 小时后,离心提取物使之澄清,保留上清部分(于 -20° C)用于进一步分析。

[0469] Fab' 定量。细胞周质提取物及培养物上清中的 Fab' 浓度通过 Fab' 装联 ELISA (其在 Humphreys 等人,2002,Protein Expression and Purification, 26, 309-320 中有所描述)以及使用蛋白质 G-hplc 进行测定。将分析物(大约为中性 pH,30°C,经 0.2 μm 过滤)以 2ml/min 加样于 HiTrap 蛋白质 -G HP 1ml 柱子 (GE-Healthcare 或同等产品),用 20mM 磷酸、50mM NaCl pH 7.4 洗涤柱子,然后注 50mM Glycine/HCl pH 2.7 洗脱 Fab'。同在 Agilent 1100 或 1200HPLC 系统上的 A280 而测量洗脱的 Fab',并通过对照已知浓度的纯化 Fab' 蛋白质的标准曲线进行定量。

[0470] 图 3 显示了运行时间延长的发酵过程中,表达抗 -TNF Fab' 的 W3110 和 MXE012 的生长概况。数据表明在生物质积累过程中 spr 菌株相对于野生型在起始生长速率上有少量增加,而在最后 ~20 小时的发酵中, spr 突变体菌株 MXE012 相对于野生型菌株 W3110 其存活时间增加。

[0471] 图 4 显示了运行时间延长的发酵过程中,表达抗 -TNF Fab' 的 W3110 和 MXE012 (W3110spr H145A) 细胞周质 Fab' 积累(实线和实心标记)以及培养基 Fab' 积累(虚线和空心标记)。数据显示两种菌株起始的细胞周质 Fab' 积累速率非常相似,但后来的发酵中相比于 MXE012,野生型 W3110 细胞有细胞周质 Fab' 的渗漏。

[0472] 图 5 显示了表达抗 -TNF α Fab' 的菌株 W3110 和 MXE012 以及表达抗 -TNF α Fab' 及重组 DsbC 的菌株 W3110 和 MXE012 的生长概况。可以看出,相比于对应的不表达重组 DsbC 的菌株,表达 DsbC 的菌株显示出生长有所改善。还可看出菌株中 spr 突变的存在改善了细胞生长。

[0473] 图 6 显示了来自表达抗 -TNF α Fab' 的大肠杆菌菌株 W3110 和 MXE012 以及来自表达抗 -TNF α Fab' 和重组 DsbC 的大肠杆菌菌株 W3110 和 MXE012 的细胞周质(阴影标记)及上清(空心非阴影标记)的 Fab 总产量。可从此图中看出,表达重组 DsbC 的菌株产生了高

产量的抗-TNF $\alpha$  Fab', 其中菌株 MXE012 在大约 92 小时内产生了超过 3.0g/L。还可看出相比于 W3110 菌株, 携带突变 SPR 基因的 MXE012 菌株展示了降低的细胞裂解, 这一点可由上清(空心标记)的抗-TNF $\alpha$  Fab' 量较少看出。

#### [0474] 实施例 8- 确定菌株中 DNA 渗漏情况及总蛋白量

[0475] dsDNA 测定:

[0476] 使用 Quant-IT Picogreen dsDNA 测定试剂盒(Invitrogen, Ref:P11496)确定 W3110, MXE001, MXE008 和 MXE012 菌株渗漏入上清的双链 DNA。将所提供的 DNA 标准品稀释到 1-1000ng/mL 的范围, 制作标准曲线。将样品稀释于 TE 缓冲液, 以使荧光读数处于此方法的线性范围之内(500-1000 倍)。在 96-孔板中, 将 100  $\mu$ L 的稀释的样品或标准品与 100  $\mu$ L 的 Picogreen 反应剂混合, 将平板在室温下避光孵育 5 分钟。使用 485nm 激发滤片以及 535nm 的发射滤片对荧光数进行 0.1s 的测量。结果示于图 7。

[0477] 蛋白质测定:

[0478] 使用 Coomassie Plus Bradford 测定试剂盒(Pierce, Ref:23236)测定 W3110、MXE001、MXE008 和 MXE012 菌株中总蛋白质的浓度。通过将牛血清白蛋白标准品稀释至 25-1000  $\mu$ g/mL 的范围制作标准曲线。将样品稀释于水中以使光学密度处于此方法的线性范围(5-10 倍), 将 33  $\mu$ L 的样品或标准品与 1mL 考马斯反应剂混合。室温下孵育 10 分钟后, 在分光光度计上读出 OD<sub>595nm</sub> 值, 使用考马斯反应剂作为空白对照。总蛋白质的浓度根据标准曲线进行计算。结果显示于图 8。

#### [0479] 实施例 9 表达抗-TNF Fab' 及 DsbC 的大肠杆菌菌株在使用大规模发酵中的生长概况

[0480] 在发酵实验中测试如实施例 5 中产生的以下菌株以比较菌株的生长和活力以及抗-TNF $\alpha$  Fab' 的表达:

[0481] 实施例 5 中产生的表达抗-TNF Fab' 和 DsbC 的 MXE012(spr H145A 突变体)

[0482] 发酵按如下方法进行:

[0483] 一开始, 表达抗-TNF Fab' 和 DsbC 的 MXE012 细胞生长于摇瓶培养中的酵母提取物及蛋白胨组成的复合培养基。然后将细胞转移至种子期的发酵罐, 使用了化学配方培养基。细胞于非营养的限制性条件下生长, 直到规定的转移点。然后将细胞转移至 250L 的生产发酵罐, 使用类似的化学配方培养基, 终体积为大约 230L。培养物起始时以分批模式生长, 直到碳源耗尽。在碳源耗尽后, 以指数增加的速率添加碳源限制性进料。在添加规定量的碳源后, 降低进料溶液添加的速率并添加 IPTG 以诱导 Fab' 表达。然后继续进行发酵, Fab' 在细胞周质中积累。在诱导后的规定时期, 通过离心收获培养物并通过将收获的细胞重悬于 Tris 和 EDTA 缓冲液并加热至 59°C 提取细胞中的 Fab'。

[0484] 发酵物的生长概况通过测量 600nm 处培养物光密度而确定。

[0485] Fab' 的滴度通过蛋白质 G HPLC 而确定, 如上文实施例 7 所描述, 所不同的是在细胞周质提取中使用的是新鲜细胞, 加入细胞培养物的为 1mL 的提取缓冲液。按照实施例 7 描述的方法测量上清及细胞周质 Fab'。图 12 显示了 Fab' 滴度。

[0486] 细胞培养物的活力通过流式细胞术测量, 使用荧光激活细胞分选术进行。

[0487] 图 11 显示了表达抗-TNF $\alpha$  Fab' 和重组 DsbC 的菌株 MXE012 的 200L 发酵物的生长概况。

[0488] 图 12 显示了表达抗 -TNF  $\alpha$  Fab' 和重组 DsbC 的菌株 MXE012 的 200L 发酵物的细胞周质抗 -TNF  $\alpha$  Fab' 滴度。

[0489] 图 13 显示了表达抗 -TNF  $\alpha$  Fab' 和重组 DsbC 的菌株 MXE012 的 200L 发酵物的活力。

[0490] 实施例 10 表达抗 -TNF Fab' 和 DsbC 的大肠杆菌菌株在使用大规模发酵中的生长概况

[0491] 在发酵实验中测试如实施例 5 中产生的以下菌株以比较菌株的生长以及抗 -TNF  $\alpha$  Fab' 的表达：

[0492] 实施例 5 中产生的表达抗 -TNF Fab' 和 DsbC 的 MXE012

[0493] 发酵如实施例 9 中描述的方法进行, 其中使用 3000L 的生产发酵罐, 包含终体积为大约 2650L。

[0494] 发酵物的生长概况通过测量 600nm 处培养物光密度而确定。

[0495] Fab' 的滴度通过蛋白质 G HPLC 而确定, 如上文实施例 9 所描述。

[0496] 图 14 显示了表达抗 -TNF  $\alpha$  Fab' 和重组 DsbC 的菌株 MXE012 的 3000L 发酵物的生长概况。

[0497] 图 15 显示了表达抗 -TNF  $\alpha$  Fab' 和重组 DsbC 的菌株 MXE012 的 3000L 发酵物的细胞周质中的抗 -TNF  $\alpha$  Fab' 滴度。

[0498] 对于优选的实施方式本发明具体地对其进行了显示和描述, 但是本领域技术人员应当理解的是可作出形式及细节上的多种变化而不偏离附随的权利要求所限定的本发明的范围。



[0001]

序列表

<110> UCB PHARMA S. A.  
 <120> 包含突变SPR基因以及野生型TSP基因的细菌宿主菌株  
 <130> G0105-W0  
 <150> GB1000590.8  
 <151> 2010-01-14  
 <160> 39  
 <170> PatentIn version 3.5  
 <210> 1  
 <211> 2049  
 <212> DNA  
 <213> 大肠杆菌

<400> 1  
 atgaacatgt ttttaggct taccgcgtta gctggcctgc ttgcaatage aggccagacc 60  
 ttegetgtag aagatacac gcgtgetgat caaattccgg tattaaagga agagacgcag 120  
 catgcgacgg taagtgagcg cgtaacgtcg cgettcacce gttctcatta tcgccagttc 180  
 gaectcgatc aggcattttc ggccaaaatc ttgaccgct acctgaatct gctcgattac 240  
 agccacaacg tgcctcggc aagegatggt gaacagttcg cgaaaaagaa aaccgagtta 300  
 ggcgatgaac tgcgttcagg caaactcgac gttttctacg atctctacaa tctggcgcaa 360  
 aagegcggtt ttgagcgta ccagtaeget ttgtcggtae tggaaaagcc gatggatttc 420  
 accggcaacg acacttataa cctgaccgcg agcaaagegc cctgcccga aaacgagget 480  
 gagtgaacg cgcctgggga cagtaaagtc aaattcgacg agitaagcct gaagctgaca 540  
 ggaaaaacgg ataaagaaat tcgtgaaacc ctgactcgcc gctacaaatt tgcattcgt 600  
 cgtctggcgc aaaccaacag cgaagatggt ttctcctgg caatgaecgc gtttgcgcgt 660  
 gaaatcgacc cgcatacaca ctatctttcc ccgcgtaata ccgaacagtt caaacctgaa 720  
 atgagtttgt cgcctggaagg tattggcgca gtgctgcaaa tggatgatga ctacacogtt 780  
 atcaattcga tggtagcagg tggcccgca gcgaagagta aagctatcag cgttggtagc 840  
 aaaattgtcg gtgttggtca aacaggcaag ccgatggttg acgtgattgg ctggcgtctt 900  
 gatgatgigg ttgccttaat taaaggcgcg aagggcagta aagttcgtct gaaaatttta 960  
 cctgctggtg aagggaacca gaccctact gtaacgttga cccgtgaacg tattegtctc 1020  
 gaagaccgcg cggitaaaa ttcgggtgaag accgtcggta aagagaaagt cggcgtgctg 1080  
 gatattccgg gcctctatgt gggtttgaca gacgatgca aagtgaact gcagaaactg 1140  
 gaaaaacaga atgtcagcag cgtcatcacc gacctgctga gcaatggcgg tggggcggtt 1200  
 aatgaagcgg tategtctct cggctctggt attcctcggg gtcccattgt tcaggtccgc 1260  
 gataacaacg gcaaggttcg tgaagatage gataccgacg gacaggtttt ctataaagcc 1320  
 ccgctgggtg tgcctgttga ccgcttcagt gcttcgctt cagaaatctt tgcccgcgca 1380  
 atgcaggatt accgtcgtgc gctggttctg ggtgaaccga cgtttggtaa aggcaccggt 1440  
 cagcaatacc gttcattgaa ccgtatttac gatcagatgt tacgtctctga atggccagcg 1500  
 ctgggttctg tgcagtagac gatccagaaa ttctatcgcg ttaacggcgg cagtacgcaa 1560  
 cgtaaaggcg taacgccaga catcatcatg ccgacgggta atgaagaaac gaaaaccgggt 1620  
 gagaaattcg aagataacgc gctgcctggt gatagcattg atgcccgac ttatgtgaaa 1680

[0002]



tcaggagatt taacggcctt tgaaccggag ctgctgaagg aacataatgc gcgiatecgc 1740  
 aaagaicctg agttccagaa caicatgaag gatatecgcg cttcaacgc tatgaaggac 1800  
 aagcgaata tcgtttctct gaattacgct gtgcgtgaga aagagaataa tgaagatgat 1860  
 gcgacgcgtc tggcgcgttt gaacgaacgc tttaaacgcg aaggtaaacc ggagtigaag 1920  
 aaactggatg atctaccgaa agattaccag gagccggate cttatctgga tgagacggtg 1980  
 aatategcac tcgatctggc gaagcttgaa aaagccagac ccgcggaaca acccgcetecc 2040  
 gtcaagtaa 2049

<210> 2  
 <211> 682  
 <212> PRT  
 <213> 大肠杆菌

<400> 2

Met Asn Met Phe Phe Arg Leu Thr Ala Leu Ala Gly Leu Leu Ala Ile  
 1 5 10 15

Ala Gly Gln Thr Phe Ala Val Glu Asp Ile Thr Arg Ala Asp Gln Ile  
 20 25 30

Pro Val Leu Lys Glu Glu Thr Gln His Ala Thr Val Ser Glu Arg Val  
 35 40 45

Thr Ser Arg Phe Thr Arg Ser His Tyr Arg Gln Phe Asp Leu Asp Gln  
 50 55 60

Ala Phe Ser Ala Lys Ile Phe Asp Arg Tyr Leu Asn Leu Leu Asp Tyr  
 65 70 75 80

Ser His Asn Val Leu Leu Ala Ser Asp Val Glu Gln Phe Ala Lys Lys  
 85 90 95

Lys Thr Glu Leu Gly Asp Glu Leu Arg Ser Gly Lys Leu Asp Val Phe  
 100 105 110

Tyr Asp Leu Tyr Asn Leu Ala Gln Lys Arg Arg Phe Glu Arg Tyr Gln  
 115 120 125

Tyr Ala Leu Ser Val Leu Glu Lys Pro Met Asp Phe Thr Gly Asn Asp  
 130 135 140

Thr Tyr Asn Leu Asp Arg Ser Lys Ala Pro Trp Pro Lys Asn Glu Ala  
 145 150 155 160

Glu Leu Asn Ala Leu Trp Asp Ser Lys Val Lys Phe Asp Glu Leu Ser  
 165 170 175

Leu Lys Leu Thr Gly Lys Thr Asp Lys Glu Ile Arg Glu Thr Leu Thr  
 180 185 190

Arg Arg Tyr Lys Phe Ala Ile Arg Arg Leu Ala Gln Thr Asn Ser Glu  
 195 200 205

Asp Val Phe Ser Leu Ala Met Thr Ala Phe Ala Arg Glu Ile Asp Pro

[0003]

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

[0004]

515 520 525  
 Ile Met Pro Thr Gly Asn Glu Glu Thr Glu Thr Gly Glu Lys Phe Glu  
 530 535 540  
 Asp Asn Ala Leu Pro Trp Asp Ser Ile Asp Ala Ala Thr Tyr Val Lys  
 545 550 555 560  
 Ser Gly Asp Leu Thr Ala Phe Glu Pro Glu Leu Leu Lys Glu His Asn  
 565 570 575  
 Ala Arg Ile Ala Lys Asp Pro Glu Phe Gln Asn Ile Met Lys Asp Ile  
 580 585 590  
 Ala Arg Phe Asn Ala Met Lys Asp Lys Arg Asn Ile Val Ser Leu Asn  
 595 600 605  
 Tyr Ala Val Arg Glu Lys Glu Asn Asn Glu Asp Asp Ala Thr Arg Leu  
 610 615 620  
 Ala Arg Leu Asn Glu Arg Phe Lys Arg Glu Gly Lys Pro Glu Leu Lys  
 625 630 635 640  
 Lys Leu Asp Asp Leu Pro Lys Asp Tyr Gln Glu Pro Asp Pro Tyr Leu  
 645 650 655  
 Asp Glu Thr Val Asn Ile Ala Leu Asp Leu Ala Lys Leu Glu Lys Ala  
 660 665 670  
 Arg Pro Ala Glu Gln Pro Ala Pro Val Lys  
 675 680

<210> 3  
 <211> 2048  
 <212> DNA  
 <213> 大肠杆菌

<400> 3  
 atgaattcgt ttttaggett accgcgttag ctggcctgct tgcaatagca gcccagacat 60  
 taattgtaga agatatacag cgtgctgata aaattccggt attaaaggaa gagacgcage 120  
 atgacacggt aagtgagegc gtaacgtcgc gttcaccgc ttctcattat ccccagttcg 180  
 acctcgatca ggcattttcg gccaaaatct ttgaccgcta cctgaatctg ctcgattaca 240  
 gccacaacgt gctgctggca agcgaigtg aacagttcgc gaaaaagaaa accgagttag 300  
 gcgatgaact gcgttcagge aaactcgacg tttctacga tctctacaat ctggcgcaaa 360  
 agcgcgcttt tgagcgttac cagtaecgctt tgteggctact ggaaaagccg atggatttca 420  
 cggcaacga cacttataac cttgaccgca gcaaagcgc ctggccgaaa aacgagcctg 480  
 agttgaacgc gctgtggac agtaaagta aattcgacga gtttaagcctg aagctgacag 540  
 gaaaaacgga taaagaaatt cgtgaaaccc tgactcgcgc ctacaaattt gccattcgtc 600  
 gtctggcgca aaccaacagc gaagatgttt tctcgtgga aatgacggcg tttgcgcgtg 660  
 aaatcgacce gcataccaac tatcttcccg cgcgtaatac egaacagttc aacctgaaa 720  
 tgagtttgcg gctggaaggt attggcgacg tgcgcaaat ggatgatgac tacaccgtta 780  
 tcaattcgat ggtggcaggt ggtccggcag cgaagagtaa agctatcagc gttggtgaca 840

[0005]

aaattgtcgg tgttggicaa acaggcaagc cgatggitga cgtgattggc tggcgtcttg 900  
 atgatgtggt tgccttaatt aaagggccga agggcagtaa agttcgtctg gaaattttac 960  
 ctgctggtaa agggaccaag accectactg taacgttgac ccgtgaacgt attcgtctcg 1020  
 aagaccgcgc ggttaaaatg tcggtgaaga ccgtcggtaa agagaaagtc ggcgtgctgg 1080  
 atattccggg ctctatgtg ggtttgacag acgatgicaa agtgcaactg cagaaactgg 1140  
 aaaaacagaa tgtcagcagc gtcacatcg accitgctag caatggcggg ggggcgttaa 1200  
 ctgaagccgt atcgtctcc ggctgttta ttctcgggg tccattgtt caggctccgg 1260  
 ataacaacgg caaggttcgt gaagatagcg ataccgacgg acaggttttc tataaaggcc 1320  
 cgctgggtgt gctggitgac cgttcagtg ctccgcttc agaaatcttt gccgcggcaa 1380  
 tgcaggatfa cggctgtgcg ctggttgtgg gfgaacggac gtttggtaaa ggcaccgttc 1440  
 agcaataacg ttcattgaac cgtatttac atcagatgtt acgtcctgaa tggccagcgc 1500  
 tgggtctctg gcagtacacg atccagaaat tctatcgcgt taacggcggc agtacgcaac 1560  
 gtaaaggcgt aacgccagac atcatcatgc cgaacggtaa tgaagaaacg gaaacgggtg 1620  
 agaaattcga agataacggc ctgccgtggg atagcattga tgccgcgact tatgtgaaat 1680  
 caggagattt aacggccttt gaaccggagc tgcctgaagga acataatgcg cgtatcgcga 1740  
 aagatcctga gtccagAAC atcatgaagg atatcgcgcg ctccaacgt atgaaggaca 1800  
 agecgaatat cgtttctctg aattacgctg tgcgtgagaa agagaataat gaagatgatg 1860  
 cgacgcgtct ggcgcgtttg aacgaacgt ttaaaccgca aggtaaaccg gagttgaaga 1920  
 aactggatga tetaccgaaa gattaccagg agccggatcc ttatctggat gagacggatg 1980  
 atategcact cgatctggcg aagcttgaaa aagccagacc cgcggaacaa cccgtcccg 2040  
 tcaagtaa 2048

<210> 1  
 <211> 2839  
 <212> DNA  
 <213> 大肠杆菌

<400> 4  
 atgcaccgca gcacctggtt caaagcatta ttgtgttag ttgccctitg ggcaccetta 60  
 agtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120  
 aacgccagc atcaggctat acgtctggat aacggtatgg ttggtcttct ggtttctgat 180  
 ccgcaggcag ttaaactcgt ctccggcctg gtggtgcccg ttgggtcgtt ggaagatccc 240  
 gaggcgtacc aggggctgge acattacctt gaacatatga gtctgatggg gtcgaaaaag 300  
 taccgcagc ctgacagctt ggcgcaatat ctcaaaatgc accgctgtag tcacaatgcc 360  
 agcactgcgc cgtatcgcac ggctttctat ctggaagtig agaacgacgc ctctcctggt 420  
 gcggtagacc gcctggccga tctatttctt gaacctttgc tcgacaagaa atatgccgaa 480  
 cgtgagcgtg atcgggtgaa cgtgaatta accatggcgc gtacgcgiga cgggatgcgc 540  
 atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt 600  
 ggtaacctcg aaactttaag cgacaaacct ggtaaacggc tgcagcaggc cctgaaagat 660  
 ttccacgaga agtactatct cgcgaatttg atgaaggcgg ttatttacag taataaacgc 720  
 ctgcggagtg tggcaaaaat ggcggcggac acctttggtc gcgtgccgaa caaagagagc 780

[0006]

aaaaaaccgg aaatcacctg gccgtagtc accgacgcgc aaaagggcat tatcattcat 840  
 tacgtccctg cgetgccgcg taaagtgttg cgcgttgagt ttcgcatoga taacaactca 900  
 gcgaagltcc gtagtaaac cgaigaattg aflacciate tgattggcaa tcgcagecca 960  
 ggtacaattt ctgactggct gcaaaagcag ggattagtig agggcattag cgccaaotcc 1020  
 gatcctatcg tcaacggcaa cagcggcgta ttagegatet ctgcgtcttt aaccgataaa 1080  
 ggccctggeta atcgcgatca ggittgtggcg gcaattitia gctatctcaa tetgttaegt 1140  
 gaaaanggea ttgataaaca ataactcgat gaactggcga atgtgctgga fatcgacttc 1200  
 cgttatccgt cgatcacccg tgatatggat taegtogaat ggetggcaga taccatgatt 1260  
 cgcgttccgt ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320  
 gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380  
 ccgaagagc cgcaacaaca aacggcttac ttgtcgatg cgcctatca ggtcgataaa 1440  
 atcagcgcac aaactttcgc cgaactggcag aaaaaagccg ccgacattgc gctctctttg 1500  
 ccagagetta acccttatat tctgatgat ttctcgtga ttaagtcaga gaagaaatac 1560  
 gaccatccag agctgattgt tgatgagtcg aatctgcgcg tgggtatgc gccaaagcgt 1620  
 tattttgcca gegageccaa agctgatgtc agcctgattt tgegtaatcc gaaagecatg 1680  
 gacagcgcgc gcaatcaggt gatgtttcgc ctcaatgatt atctcgcagg gctggcgtt 1740  
 gatcagttaa gcaaccagge gtcggttggg ggcataagtt ttccacca cgetaacaac 1800  
 ggccctatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860  
 ctcgaggsgt actttagcta tacegetacg gaagatcagc ttgagcagge gaagtcctgg 1920  
 tataaccaga tgatggatc cgcagaaaag ggtaaagcgt ttgagcagge gattatgccc 1980  
 gcgcagatgc tctcgaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040  
 tccattacgt tgaagaggt gctggcctat cgcgagcct taaaatcagg ggetcgacca 2100  
 gaggttatgg ttatcgcaa catgacegag gcccagcga caacgtgge acgcgatgtg 2160  
 caaaaacagt tgggcctga tggttcagag tgggtcga acaaagatg agtggctgat 2220  
 aaaaaacaat ccgtcaictt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280  
 gtatttgta cgaactggcta cgaigaatac accagctcag cctatagctc tetgttgggg 2340  
 cagatcgtac agcctgggt ctacaatcag ttgcgtaccg aagaacaatt gggetatgcc 2400  
 gtgtttcgt ttccaatgag cgtggggcgt cagtggggca tgggcttctt ttgcaaaage 2460  
 aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520  
 gcaaaattgc gagegatgaa gccagatgag ttgcgcaaa tccagcagge ggtaattacc 2580  
 cagatgetgc aggcacegca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640  
 gatcggcgca atatgcctt cgattecgt gataaaatcg tggcccagat aaaactgctg 2700  
 acgccgcaaa aacttctga ttcttccat cagcgggtgg tcgagccgca agcctatgct 2760  
 attctgtcgc agatttcgg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820  
 tggaaagtgt gggagaacgt cagcggctg cagcaaaaa tgcccctgat gagtgaanaag 2880  
 aatgagtga 2889

<210> 5  
 <211> 962  
 <212> PRI

[0007]

&lt;213&gt; 大肠杆菌

&lt;400&gt; 5

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

[0008]

Val Leu Arg Val Glu Phe Arg Ile Asp Asn Asn Ser Ala Lys Phe Arg  
 290 295 300

Ser Lys Thr Asp Glu Leu Ile Thr Tyr Leu Ile Gly Asn Arg Ser Pro  
 305 310 315 320

Gly Thr Leu Ser Asp Trp Leu Gln Lys Gln Gly Leu Val Glu Gly Ile  
 325 330 335

Ser Ala Asn Ser Asp Pro Ile Val Asn Gly Asn Ser Gly Val Leu Ala  
 340 345 350

Ile Ser Ala Ser Leu Thr Asp Lys Gly Leu Ala Asn Arg Asp Gln Val  
 355 360 365

Val Ala Ala Ile Phe Ser Tyr Leu Asn Leu Leu Arg Glu Lys Gly Ile  
 370 375 380

Asp Lys Gln Tyr Phe Asp Glu Leu Ala Asn Val Leu Asp Ile Asp Phe  
 385 390 395 400

Arg Tyr Pro Ser Ile Thr Arg Asp Met Asp Tyr Val Glu Trp Leu Ala  
 405 410 415

Asp Thr Met Ile Arg Val Pro Val Glu His Thr Leu Asp Ala Val Asn  
 420 425 430

Ile Ala Asp Arg Tyr Asp Ala Lys Ala Val Lys Glu Arg Leu Ala Met  
 435 440 445

Met Thr Pro Gln Asn Ala Arg Ile Trp Tyr Ile Ser Pro Lys Glu Pro  
 450 455 460

His Asn Lys Thr Ala Tyr Phe Val Asp Ala Pro Tyr Gln Val Asp Lys  
 465 470 475 480

Ile Ser Ala Gln Thr Phe Ala Asp Trp Gln Lys Lys Ala Ala Asp Ile  
 485 490 495

Ala Leu Ser Leu Pro Glu Leu Asn Pro Tyr Ile Pro Asp Asp Phe Ser  
 500 505 510

Leu Ile Lys Ser Glu Lys Lys Tyr Asp His Pro Glu Leu Ile Val Asp  
 515 520 525

Glu Ser Asn Leu Arg Val Val Tyr Ala Pro Ser Arg Tyr Phe Ala Ser  
 530 535 540

Glu Pro Lys Ala Asp Val Ser Leu Ile Leu Arg Asn Pro Lys Ala Met  
 545 550 555 560

Asp Ser Ala Arg Asn Gln Val Met Phe Ala Leu Asn Asp Tyr Leu Ala  
 565 570 575

Gly Leu Ala Leu Asp Gln Leu Ser Asn Gln Ala Ser Val Gly Gly Ile  
 580 585 590

[0009]



Ser Phe Ser Thr Asn Ala Asn Asn Gly Leu Met Val Asn Ala Asn Gly  
 595 600 605

Tyr Thr Gln Arg Leu Pro Gln Leu Phe Gln Ala Leu Leu Glu Gly Tyr  
 610 615 620

Phe Ser Tyr Thr Ala Thr Glu Asp Gln Leu Glu Gln Ala Lys Ser Trp  
 625 630 635 640

Tyr Asn Gln Met Met Asp Ser Ala Glu Lys Gly Lys Ala Phe Glu Gln  
 645 650 655

Ala Ile Met Pro Ala Gln Met Leu Ser Gln Val Pro Tyr Phe Ser Arg  
 660 665 670

Asp Glu Arg Arg Lys Ile Leu Pro Ser Ile Thr Leu Lys Glu Val Leu  
 675 680 685

Ala Tyr Arg Asp Ala Leu Lys Ser Gly Ala Arg Pro Glu Phe Met Val  
 690 695 700

Ile Gly Asn Met Thr Glu Ala Gln Ala Thr Thr Leu Ala Arg Asp Val  
 705 710 715 720

Gln Lys Gln Leu Gly Ala Asp Gly Ser Glu Trp Cys Arg Asn Lys Asp  
 725 730 735

Val Val Val Asp Lys Lys Gln Ser Val Ile Phe Glu Lys Ala Gly Asn  
 740 745 750

Ser Thr Asp Ser Ala Leu Ala Ala Val Phe Val Pro Thr Gly Tyr Asp  
 755 760 765

Glu Tyr Thr Ser Ser Ala Tyr Ser Ser Leu Leu Gly Gln Ile Val Gln  
 770 775 780

Pro Trp Phe Tyr Asn Gln Leu Arg Thr Glu Glu Gln Leu Gly Tyr Ala  
 785 790 795 800

Val Phe Ala Phe Pro Met Ser Val Gly Arg Gln Trp Gly Met Gly Phe  
 805 810 815

Leu Leu Gln Ser Asn Asp Lys Gln Pro Ser Phe Leu Trp Glu Arg Tyr  
 820 825 830

Lys Ala Phe Phe Pro Thr Ala Glu Ala Lys Leu Arg Ala Met Lys Pro  
 835 840 845

Asp Glu Phe Ala Gln Ile Gln Gln Ala Val Ile Thr Gln Met Leu Gln  
 850 855 860

Ala Pro Gln Thr Leu Gly Glu Glu Ala Ser Lys Leu Ser Lys Asp Phe  
 865 870 875 880

Asp Arg Gly Asn Met Arg Phe Asp Ser Arg Asp Lys Ile Val Ala Gln  
 885 890 895

[0010]

Ile Lys Leu Leu Thr Pro Gln Lys Leu Ala Asp Phe Phe His Gln Ala  
 900 905 910

Val Val Glu Pro Gln Gly Met Ala Ile Leu Ser Gln Ile Ser Gly Ser  
 915 920 925

Gln Asn Gly Lys Ala Glu Tyr Val His Pro Glu Gly Trp Lys Val Trp  
 930 935 940

Glu Asn Val Ser Ala Leu Gln Gln Thr Met Pro Leu Met Ser Glu Lys  
 945 950 955 960

Asn Glu

- <210> 6
- <211> 2915
- <212> DNA
- <213> 大肠杆菌

<400> 6  
 attccccgca gcaccfggtt caaagcatta ttgttgtag ttgcccttg ggcacattaa 60  
 tgtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120  
 aaccgccagt atcaggctat acgtctggat aacggatgg tggcttctgt gtttctgat 180  
 ccgcagcag ttaateget ctoggecctg gtggtgccg ttgggtcgt ggaagatccc 240  
 gaggegtacc aggggctggc acattacct gaacatatga gtctgatggg gtcgaaaaag 300  
 taccgcaagg ctgacagtet ggccgaatat ctcaaatgc acggcgtag tcacaatgcc 360  
 agcactgccc cgtatcgcac ggtttctat ctggaagttg agaacgacgc cttgectggt 420  
 gcggtagccc gcttgccga tgetattget gaaccttgc tcgacaagaa atatgccgaa 480  
 cgtgagccta atgcggtgaa cgtgaatta accatggcgc gtacgcgta cgggatgcgc 540  
 atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gtttctggt 600  
 ggtaacctcg aaactttaag cgacaaacct ggtaatcogg tgcagcaggc gctgaaagat 660  
 ttccacgaga agtactatc cgcgaatttg atgaaggcgg ttatttacag taataaacgc 720  
 ctgcccgagt tggcaaaaat ggcggcggac accittggtc gcgtgccgaa caaagagagc 780  
 aaaaaacagg aaatcacctg gcggtagtc accgacgcgc aaaagggcac tatcattcat 840  
 tacgtccctg cgtgcccgcg taaagtgttg cgcgttagt ttccatcga taacaactca 900  
 gcgaattcc gtagtaaac cgaatgaattg attacctatc tgattggcaa tcgcagccca 960  
 ggtacacttt ctgactgget gcaaaagcag ggattagttg agggcattag cgcacaactcc 1020  
 gatccatcgc tcaacggcaa cagcggcgta ttagegatct ctgcgtcttt aaccgataaa 1080  
 ggccctggta atcgcgatac gttgtggcg gcaatttita gctatctcaa tctgttacgt 1140  
 gaaaaagcca ttgataaaca atacttgat gaactggcga atgtgtgga tategacttc 1200  
 cgttatccgt cgaacacccg tgatatggat lacgtcgaat ggetggcaga taccatgatt 1260  
 cgcgttccgt ttgagcatac gctggatgca gtaaatattg ccgatcgta cgaatgctaa 1320  
 gcagtaaagg aacgtctggc gatgatgac ccgcagaatg ccggtatctg gtatcfcagc 1380  
 ccgaaagagc cgcacaacaa aacggcttac ttgtcgatg ccgcgatca ggtcgataaa 1440  
 atcagcgcac aaactttcgc cgaactggcag aaaaaagccg ccgacattgc gctctctttg 1500

[0011]



ccagagctta acccttatat tctgatgat ttctcgtga ttaagtcaga gaagaaatac 1560  
 gaccatccag agctgattgt tgaigagteg aatctgecgg tgggtgatge gccaaagcgt 1620  
 tattttgcca gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680  
 gacagcgcce gcaatecagg gatgtttgcg ctcaatgatt atctcgcagg gctggcgctt 1740  
 gatcagttaa gcaaccagge gtcggttggc ggcataagtt ttccaccaaa cgctaacaac 1800  
 ggccttatgg ttaatgetaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860  
 ctcgaggggt actttageta taccgtacg gaagatcage ttgagcagge gaagtcctgg 1920  
 tataaccaga tgatggattc cgcagaaaag ggtaaaagct ttgagcagge gattatgccc 1980  
 gcgcagatgc tctcgaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040  
 tccattactg tgaagaggt gctggcctat cgcgacgctt taaaatcagg ggctcgacca 2100  
 gagtttatgg ttatcgcaa catgaccgag gcccaggcaa caacgctggc acgcgatgtg 2160  
 caaaaacagt tgggcgctga tggttcagag tgggtctgaa acaaagatgt agtggctgat 2220  
 aaaaaacaat cctcatctt tgaaaaaagc ggtaacagca ccgactccgc actggcagcg 2280  
 gtattttgac cgaactgcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340  
 cagatcgtac agcctgtggt ctacaatcag ttgcgtaccg aagaacaatt ggctatgccc 2400  
 gtgtttgctg ttccaatgag cgtggggcgt cagtggggca tgggcttctt ttgcaaagc 2460  
 aatgataaac agccttcatt ctgtgggag cgttacaagg cgtttttccc aaccgcagag 2520  
 gcaaaattgc gagcgaigaa gccagatgag ttgcgcaaaa tccagcagge ggtaattacc 2580  
 cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640  
 gatcggcgca atatgctctt cgattcgcgt gataaaatcg tgcccagat aaaactgctg 2700  
 acgcccgaaa aacttctga ttcttccat cagcgggtgg tcgagccgca aggcattgct 2760  
 attctgtcgc agatttcggg cagccagaac gggaaagccg aatatgtaca cctgaagge 2820  
 tggaaagtgt gggagaacgt cagcgcgttg cagcaaaaaa tgcctctgat gattgaaaag 2880  
 aatgagtgat gtcgcccaga cactagatcc ttgccc 2915

<210> 7  
 <211> 1425  
 <212> DNA  
 <213> 大肠杆菌

<400> 7  
 atgaaaaaaaa ccacattage actgagtcca ctggetctga gtttaggttt ggcgttatct 60  
 ccgctctctg caacggcgge tgagacttct tcagcaacga cagcccagca gatgccaagc 120  
 ctgcaccga tgcctgaaaa ggtgatgcct tcagtggcca gcattaacgt agaaggttagc 180  
 acaacegtta ataegccgcg tatgccgcgt aatttcagc agttctctgg tgatgattct 240  
 cegtctgccc aggaaggctc tccgttccag agctctccgt tctgcccagg tgcccagggc 300  
 ggtaatggtg gcgcccagca acagaaatc atggcgttgg gtcccggcgt catcattgat 360  
 gccgataaag gctatgtctt caccacaac cagtttctt ataacgcgac ggtcattaaa 420  
 gttcaactga gcgatggccc taagttcgac gcgaagatg ttggcaaga tccgcgctct 480  
 gatctcgcgc tgatccaaat ccagaaccgc aaaaacctga ccgcaattaa gatggcggat 540  
 tctgatgcac tgcgcgiggg tgattacacc gtagcgattg gtaaccgctt tggctctggc 600  
 gagacggtaa ctccgggat tctctctgcg ctggggcgta gcggcctgaa tgcgcaaaac 660

[0012]

tacgaaaact tcateccagac cgatgcagcg atcaaccgtg gtaactcogg tggigcgtg 720  
 gttaacctga acggcgaaact gatcggiatc aacaccgcga tcctcgcacc ggacggcgcg 780  
 aacatcggta tcggttttgc tatcccgagt aacatggtga aaaacctgac ctccgagatg 840  
 gtggaatacg gccaggtgaa acgcgggtgag ctgggtatta tggggactga getgaactcc 900  
 gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960  
 cetaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgateac ctcaactgaac 1020  
 ggtaagccga tcagcagctt tgcgcactg cgtgctcagg tgggtactat gccggtaggc 1080  
 agcaaaactga ccttgggctt actgcgcgac ggtaagcagg ttaacctgaa cctggaactg 1140  
 cagcagagea gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgt 1200  
 gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacctgaa aacgggcaact 1260  
 ccggtcgcgc agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcagcga 1320  
 gtgaaaaaca tegetgaact gcgtaaagtt ctgacagca aacctctgt getggcacte 1380  
 aacattcagc gggcgacag caccatctac ctgttaatgc agtaa 1425

<210> 8  
 <211> 474  
 <212> PRT  
 <213> 大肠杆菌  
 <400> 8

Met Lys Lys Thr Thr Leu Ala Leu Ser Ala Leu Ala Leu Ser Leu Gly  
 1 5 10 15

Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala  
 20 25 30

Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val  
 35 40 45

Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn  
 50 55 60

Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser  
 65 70 75 80

Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln  
 85 90 95

Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala  
 100 105 110

Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr  
 115 120 125

Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser  
 130 135 140

Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser  
 145 150 155 160

Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile

[0013]

165 170 175  
 Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Ala  
 180 185 190  
 Ile Gly Asn Pro Phe Gly Leu Gly Glu Thr Val Thr Ser Gly Ile Val  
 195 200 205  
 Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe  
 210 215 220  
 Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ser Gly Gly Ala Leu  
 225 230 235 240  
 Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala  
 245 250 255  
 Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met  
 260 265 270  
 Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg  
 275 280 285  
 Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys  
 290 295 300  
 Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu  
 305 310 315 320  
 Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile  
 325 330 335  
 Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala  
 340 345 350  
 Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu  
 355 360 365  
 Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser  
 370 375 380  
 Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala  
 385 390 395 400  
 Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val  
 405 410 415  
 Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val  
 420 425 430  
 Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg  
 435 440 445  
 Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg  
 450 455 460  
 Gly Asp Ser Thr Ile Tyr Leu Leu Met Gln

[0014]

465 470

<210> 9  
 <211> 1425  
 <212> DNA  
 <213> 大肠杆菌

<400> 9  
 atgaaaaaaa ccacattagc acfgagtgea ctggctctga gtttaggttt ggcgttaet 60  
 ccgctctctg caacggcgcg tgagacttct tcagcaacga cagcccagca gatgecaagc 120  
 cttgcaecga tgcctgaaaa ggigatgect tcagtgtgca gcattaacgt agaaggtagc 180  
 acaaccgtta atacgcgcg gatgccgcgt aatttccagc agttcttcgg tgatgatctt 240  
 ccgttctgce aggaaggttc tccgttccag agctctccgt tctgccaggg tggccagggc 300  
 ggtaaiggtg gcggccagca acagaaattc atggcctggg gttccggcgt catcattgat 360  
 gccgataaag gctatgtcgt caccaacaac cagttgttg ataacgcgac ggtcattaaa 420  
 gttcaactga gcgatggccg taagttcgac gcgaagatgg ttggcaanga tccgcctctt 480  
 gatategcgc tgatecaaat ccagaaccgc aaaaacctga ccgaattaa gatggcggat 540  
 tctgatgcac tgcgcgtggg tgattacacc gtagcgattg gtaaccctgt ttgtctgggc 600  
 gagacggtaa ctccgggat tctctctgag ctggggcgta gcggcctgaa tgcgaaaaac 660  
 tacgaaaact teatccagac cgatgcagcg attaatctgt gtaacgcggg tggctgcctg 720  
 gttaacctga acggcgaact gatcggatc aacaccgcga tctctgcacc ggacggcggc 780  
 aacatcggtg tgggttttgc tctcccgagt aacatggtga aaaacctgac ctccagatg 840  
 gtggaatagc gccaggtgaa acgcggtgag ctgggtatta tggggaactga gctgaacctc 900  
 gaactggcga aagcgaatga agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960  
 cetaattctt ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac etcactgaac 1020  
 ggtaagccga teagcagett tgccgcactg cgtgctcagg tggtaactat gccggtagcc 1080  
 agcaaaatga cctggggett actgcgcgac ggtaagcagg ttaacgtgaa cctggaaactg 1140  
 cagcagagca gccagaatca gttgatctcc agctccatct tcaacggcat tgaaggcct 1200  
 gagatagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcaact 1260  
 ccgctgcgcg agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcaggea 1320  
 gtaaaaaaca tcctgaaact gctaaagtt ctgcacagca aaccgtctgt gctggcaact 1380  
 caattcagc gcggcgacag caccatctac ctgttaatgc agtaa 1425

<210> 10  
 <211> 474  
 <212> PRT  
 <213> 大肠杆菌

<400> 10

Met Lys Lys Thr Thr Leu Ala Leu Ser Ala Leu Ala Leu Ser Leu Gly  
 1 5 10 15

Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala  
 20 25 30

Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val  
 35 40 45

[0015]

Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn  
50 55 60

Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser  
65 70 75 80

Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln  
85 90 95

Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala  
100 105 110

Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr  
115 120 125

Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser  
130 135 140

Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser  
145 150 155 160

Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile  
165 170 175

Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Ala  
180 185 190

Ile Gly Asn Pro Phe Gly Leu Gly Glu Thr Val Thr Ser Gly Ile Val  
195 200 205

Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe  
210 215 220

Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ala Gly Gly Ala Leu  
225 230 235 240

Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala  
245 250 255

Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met  
260 265 270

Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg  
275 280 285

Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys  
290 295 300

Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu  
305 310 315 320

Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile  
325 330 335

Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala  
340 345 350

[0016]

Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu  
 355 360 365

Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser  
 370 375 380

Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala  
 385 390 395 400

Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val  
 405 410 415

Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val  
 420 425 430

Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg  
 435 440 445

Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg  
 450 455 460

Gly Asp Ser Thr Ile Tyr Leu Leu Met Gln  
 465 470

<210> 11  
 <211> 107  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> hTNF40-gLi

<400> 11

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val Gly Thr Asn  
 20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Ala Leu Ile  
 35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Tyr Arg Phe Ser Gly  
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
 65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Leu  
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys  
 100 105

<210> 12  
 <211> 118  
 <212> PRT  
 <213> 人工序列

[0017]



<220>  
 <223> gh3h TNF40.4  
 <400> 12  
 Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly  
 1 5 10 15  
 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Val Phe Thr Asp Tyr  
 20 25 30  
 Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met  
 35 40 45  
 Gly Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val  
 50 55 60  
 Lys Gly Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr  
 65 70 75 80  
 Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys  
 85 90 95  
 Ala Arg Gly Tyr Arg Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr  
 100 105 110  
 Leu Val Thr Val Ser Ser  
 115  
 <210> 13  
 <211> 214  
 <212> PRT  
 <213> 人工序列  
 <220>  
 <223> 移植的轻链  
 <400> 13  
 Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
 1 5 10 15  
 Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val Gly Thr Asn  
 20 25 30  
 Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Ala Leu Ile  
 35 40 45  
 Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Tyr Arg Phe Ser Gly  
 50 55 60  
 Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
 65 70 75 80  
 Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Leu  
 85 90 95  
 Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala  
 100 105 110  
 Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly

[0018]

115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala  
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln  
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser  
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr  
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser  
195 200 205

Phe Asn Arg Gly Glu Cys  
210

<210> 14  
<211> 229  
<212> PRT  
<213> 人工序列

<220>  
<223> 移植的重链

<400> 14

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly  
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Val Phe Thr Asp Tyr  
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met  
35 40 45

Gly Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val  
50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr  
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys  
85 90 95

Ala Arg Gly Tyr Arg Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr  
100 105 110

Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro  
115 120 125

Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly  
130 135 140

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn  
145 150 155 160

[0019]

Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln  
 165 170 175

Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser  
 180 185 190

Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser  
 195 200 205

Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr  
 210 215 220

His Thr Cys Ala Ala  
 225

<210> 15  
 <211> 24  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 寡核苷酸引物

<400> 15  
 gcaacataat tttcttttta cctc 24

<210> 16  
 <211> 24  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 寡核苷酸引物

<400> 16  
 gggaaatgaa cctgagcaaa acgc 24

<210> 17  
 <211> 24  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 寡核苷酸引物

<400> 17  
 gtgccagcag atgcagcagc ttgc 24

<210> 18  
 <211> 21  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 寡核苷酸引物

<400> 18  
 tttgcagcca gtcagaaagt g 21

<210> 19  
 <211> 24  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 寡核苷酸引物

[0020]

<400> 19  
ctgcctgcga ttttcgccg aacg 24

<210> 20  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 寡核苷酸引物

<400> 20  
cgcatggtac gtgccagat atcc 24

<210> 21  
<211> 188  
<212> PRT  
<213> 大肠杆菌

<400> 21  
Met Val Lys Ser Gln Pro Ile Leu Arg Tyr Ile Leu Arg Gly Ile Pro  
1 5 10 15

Ala Ile Ala Val Ala Val Leu Leu Ser Ala Cys Ser Ala Asn Asn Thr  
20 25 30

Ala Lys Asn Met His Pro Glu Thr Arg Ala Val Gly Ser Glu Thr Ser  
35 40 45

Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu Asn Leu Val Arg Asn Val  
50 55 60

Asp Val Lys Ser Arg Ile Met Asp Gln Tyr Ala Asp Trp Lys Gly Val  
65 70 75 80

Arg Tyr Arg Leu Gly Ser Thr Lys Lys Gly Ile Asp Cys Ser Gly  
85 90 95

Phe Val Gln Arg Thr Phe Arg Glu Gln Phe Gly Leu Glu Leu Pro Arg  
100 105 110

Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys Ser Val Ser Arg Ser Asn  
115 120 125

Leu Arg Thr Gly Asp Leu Val Leu Phe Arg Ala Gly Ser Thr Gly Arg  
130 135 140

His Val Gly Ile Tyr Ile Gly Asn Asn Gln Phe Val His Ala Ser Thr  
145 150 155 160

Ser Ser Gly Val Ile Ile Ser Ser Met Asn Glu Pro Tyr Trp Lys Lys  
165 170 175

Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser Arg Ser  
180 185

<210> 22  
<211> 162  
<212> PRT  
<213> 大肠杆菌

[0021]

<400> 22

Cys Ser Ala Asn Asn Thr Ala Lys Asn Met His Pro Glu Thr Arg Ala  
1 5 10 15

Val Gly Ser Glu Thr Ser Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu  
20 25 30

Asn Leu Val Arg Asn Val Asp Val Lys Ser Arg Ile Met Asp Gln Tyr  
35 40 45

Ala Asp Trp Lys Gly Val Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys  
50 55 60

Gly Ile Asp Cys Ser Gly Phe Val Gln Arg Thr Phe Arg Glu Gln Phe  
65 70 75 80

Gly Leu Glu Leu Pro Arg Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys  
85 90 95

Ser Val Ser Arg Ser Asn Leu Arg Thr Gly Asp Leu Val Leu Phe Arg  
100 105 110

Ala Gly Ser Thr Gly Arg His Val Gly Ile Tyr Ile Gly Asn Asn Gln  
115 120 125

Phe Val His Ala Ser Thr Ser Ser Gly Val Ile Ile Ser Ser Met Asn  
130 135 140

Glu Pro Tyr Trp Lys Lys Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser  
145 150 155 160

Arg Ser

<210> 23

<211> 951

<212> DNA

<213> 人工序列

<220>

<223> 突变Omp1序列

<400> 23

```

atgcggcgga aacttctggg aatagtcctg acaacccta ttgcgatcag ctcttttgc 60
tetaccgaga ctttatcggt tactectgac aacataaatg cggacattag tcttggaact 120
ctgagcggaa aaacaaaaga gegtgtttat ctagecgaag aaggaggcgc aaaagtcagt 180
caactcgact ggaaattcaa taacgetgca attattaag gtgcaattaa ttgggatttg 240
atgceceaga tatctatcgg ggetgetggc tggacaactc tggcagcgc aggtggcaat 300
atggctgac aggactggat ggattccagt aaccocegaa cctggacgga tgaaagtaga 360
cacctgata cacaactcaa ttatgcaac gaatttgatc tgaatatcaa aggetggetc 420
ctcaacgaac ecaattaccg cctgggactc atggcggat atcaggaaag cegttatagc 480
tttacagcca gaggtgggtc ctatctctac agttctgagg agggatteag agatgatata 540
ggctctctcc cgaatggaga aagagcaate ggcfacaaac aegtttttaa aatgcctac 600
    
```

[0022]

attggcctga ctggaagtta tcgttaigaa gattttgaac tcgggtggcac atttaaatac 660  
 agcgctgggg tggaateatc tgataaogct gaagcttatg acccgggaaa aagaatcact 720  
 tatecgagta aggtcaaaaga ccaaaaattac tattctgttg cagtcaatgc aggttattac 780  
 gtcacaccta acgcaaaaagt ttafgttgaa ggcgcgatgga atcgggttac gaataaaaaa 840  
 ggtaataactt cactttatga tcacaataat aacacttcag actacagcaa aaatggagca 900  
 ggtatagaaa actataactt catcactact gctggtctta agtacacatt t 951

<210> 24  
 <211> 317  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 突变OmpT序列

<400> 24

Met Arg Ala Lys Leu Leu Gly Ile Val Leu Thr Thr Pro Ile Ala Ile  
 1 5 10 15

Ser Ser Phe Ala Ser Thr Glu Thr Leu Ser Phe Thr Pro Asp Asn Ile  
 20 25 30

Asn Ala Asp Ile Ser Leu Gly Thr Leu Ser Gly Lys Thr Lys Glu Arg  
 35 40 45

Val Tyr Leu Ala Glu Glu Gly Gly Arg Lys Val Ser Gln Leu Asp Trp  
 50 55 60

Lys Phe Asn Asn Ala Ala Ile Ile Lys Gly Ala Ile Asn Trp Asp Leu  
 65 70 75 80

Met Pro Gln Ile Ser Ile Gly Ala Ala Gly Trp Thr Thr Leu Gly Ser  
 85 90 95

Arg Gly Gly Asn Met Val Asp Gln Asp Trp Met Asp Ser Ser Asn Pro  
 100 105 110

Gly Thr Trp Thr Asp Glu Ser Arg His Pro Asp Thr Gln Leu Asn Tyr  
 115 120 125

Ala Asn Glu Phe Asp Leu Asn Ile Lys Gly Trp Leu Leu Asn Glu Pro  
 130 135 140

Asp Tyr Arg Leu Gly Leu Met Ala Gly Tyr Gln Glu Ser Arg Tyr Ser  
 145 150 155 160

Phe Thr Ala Arg Gly Gly Ser Tyr Ile Tyr Ser Ser Glu Glu Gly Phe  
 165 170 175

Arg Asp Asp Ile Gly Ser Phe Pro Asn Gly Glu Arg Ala Ile Gly Tyr  
 180 185 190

Lys Gln Arg Phe Lys Met Pro Tyr Ile Gly Leu Thr Gly Ser Tyr Arg  
 195 200 205

Tyr Glu Asp Phe Glu Leu Gly Gly Thr Phe Lys Tyr Ser Gly Trp Val

[0023]

www.patviewer.com

210 215 220

Glu Ser Ser Asp Asn Ala Glu Ala Tyr Asp Pro Gly Lys Arg Ile Thr  
 225 230 235 240

Tyr Arg Ser Lys Val Lys Asp Gln Asn Tyr Tyr Ser Val Ala Val Asn  
 245 250 255

Ala Gly Tyr Tyr Val Thr Pro Asn Ala Lys Val Tyr Val Glu Gly Ala  
 260 265 270

Trp Asn Arg Val Thr Asn Lys Lys Gly Asn Thr Ser Leu Tyr Asp His  
 275 280 285

Asn Asn Asn Thr Ser Asp Tyr Ser Lys Asn Gly Ala Gly Ile Glu Asn  
 290 295 300

Tyr Asn Phe Ile Thr Thr Ala Gly Leu Lys Tyr Thr Phe  
 305 310 315

<210> 25  
 <211> 954  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 突变OmpT序列

<400> 25  
 attcgggcga aactctfggg aatagtctcg acaaceceta ttgcgacag ctcttttgcct 60  
 tetaccgaga attatcggt tactctcgac aacataaatg eggacattag tcttggaact 120  
 ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggagccg aaaagtcagt 180  
 caactcgact ggaattcaa taacgtgca attattaag gtgcaattaa ttgggatttg 240  
 atgccccaga tctctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300  
 atggtcgate agactggat ggatecagt aaceccggaa cctggacgga tgaagttaga 360  
 caccctgata cacaactcaa ttatgccaac gaatttgate tgaatataca aggcctggctc 420  
 ctcaacgaac ccaattaccg cctgggacte atggcggat atcaggaaag ccgttatagc 480  
 ttacagcca gagctggitc ctatatctac agttctgagg agggattcag agatgatatac 540  
 ggctcttccc cgaatggaga aagagcaate ggctacaaac aacgttttaa aatgcccctac 600  
 attggettga ctggaagtta tcgttatgaa gattttgaac tcggtggeac atttaaatac 660  
 agcgctggg tggaaatcfc tgataacgat gaacaatag accegggaaa aagaateact 720  
 tatcgagta aggtcaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780  
 gtcacaceta acgcaaaagt ttatgttgaa ggcgcagga ategggttac gaataaaaaa 840  
 ggtaataact cactttatga tcacaataat aacacttcag actacagcaa aaatggagca 900  
 ggtatagaaa actataactt cataactact gctggcttta agtacacatt ttaa 954

<210> 26  
 <211> 729  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 突变DsbC序列

[0024]

<400> 26  
 atgaagaag gttttatggt gtttactttg tttagcggegt tttcaggett tgctcagget 60  
 gatgacgegg caattcaaca aacgttagcc aaaatgggca tcaaaageag cgatattcag 120  
 cccgcgctg tagotggcat gaagacagtt ctgactaaca gcggcgtggt gtacatcacc 180  
 gatgatggta aacataatcat tcaggggcca atgtatgacg ttagtggcac ggcctcggtc 240  
 aatgtcacea ataagatget gttaaagcag ttgaatgcgc ttgaaaaaga gatgateggt 300  
 tataaaagcgc cgcaggaaaa acacgtcacc accgtgitta ctgatattac ctgtggttac 360  
 tgccacaaac tgcatgacea aatggcagac tacaacgcgc tggggatcac cgtgctgtat 420  
 ctgtctttcc cgcgccaggg gcctggacagc gatgcagaga aagaaatgaa agctatctgg 480  
 tgtgcgaaag ataaaaaca agegtttgat gatgtgatgg caggtaaaag cgtcgcacca 540  
 gccagttgcg acgtggatat tgccgacat tacgcaettg gcgtccagct tggcgttagc 600  
 ggtactccgg cagttgtget gagcaatgga acactgttc cgggttaccg gccgccgaaa 660  
 gagatgaaag aatttctega cgaacaccaa aaaatgacca gcggtaaaaca ccacaccat 720  
 caccactaa 729

<210> 27  
 <211> 242  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 突变DsbC序列

<400> 27

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

[0025]



145 150 155 160

Cys Ala Lys Asp Lys Asn Lys Ala Phe Asp Asp Val Met Ala Gly Lys  
165 170 175

Ser Val Ala Pro Ala Ser Cys Asp Val Asp Ile Ala Asp His Tyr Ala  
180 185 190

Leu Gly Val Gln Leu Gly Val Ser Gly Thr Pro Ala Val Val Leu Ser  
195 200 205

Asn Gly Thr Leu Val Pro Gly Tyr Gln Pro Pro Lys Glu Met Lys Glu  
210 215 220

Phe Leu Asp Glu His Gln Lys Met Thr Ser Gly Lys His His His His  
225 230 235 240

His His

<210> 28  
<211> 5  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述:hTNF40 CDRH1

<400> 28

Asp Tyr Gly Met Asn  
1 5

<210> 29  
<211> 17  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述:hTNF40人杂合CDRH2

<400> 29

Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val Lys  
1 5 10 15

Gly

<210> 30  
<211> 9  
<212> PRT  
<213> 人工序列

<220>  
<223> 人工序列的描述:hTNF40 CDRH3

<400> 30

Gly Tyr Arg Ser Tyr Ala Met Asp Tyr  
1 5

<210> 31  
<211> 11  
<212> PRT

[0026]

<213> 人工序列

<220>

<223> 人工序列的描述:hTNF40 CDRL1

<400> 31

Lys Ala Ser Gln Asn Val Gly Thr Asn Val Ala  
1 5 10

<210> 32

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:hTNF40 CDRL2

<400> 32

Ser Ala Ser Phe Leu Tyr Ser  
1 5

<210> 33

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:hTNF40 CDRL3

<400> 33

Gln Gln Tyr Asn Ile Tyr Pro Leu Thr  
1 5

<210> 34

<211> 17

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:hTNF40 CDRH2

<400> 34

Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Val Asp Asp Phe Lys  
1 5 10 15

Gly

<210> 35

<211> 84

<212> DNA

<213> 人工序列

<220>

<223> OmpA寡核苷酸适体

<400> 35

tcgagttcta gataacgagg cgtaaaaaat gaaaaagaca gctatcgeaa ttgcagtggc 60

cttggctctg acgtacgagt cagg 84

<210> 36

<211> 67

<212> DNA

<213> 人工序列

[0027]

<220>		
<223>	IGS盒-1	
<400>	36	
	gagctcacca gtaacaaaaa gttttaatag aggagagtgt taatgaagaa gactgctata	60
	gcaattg	67
<210>	37	
<211>	69	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	IGS盒-2	
<400>	37	
	gagctcacca gtaacaaaaa gttttaatag aggggagtgt taaaatgaag aagactgcta	60
	tagcaattg	69
<210>	38	
<211>	81	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	IGS盒-3	
<400>	38	
	gagctcacca gtaacaaaaa gttttaatag aggagagtgt tgaggaggaa aaaaaaatga	60
	agaaaactgc tatagcaatt g	81
<210>	39	
<211>	81	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	IGS盒-4	
<400>	39	
	gagctcacca gtaacaaaaa gttttaatag aggagagtgt tgacgaggat tatataatga	60
	agaaaactgc tatagcaatt g	81

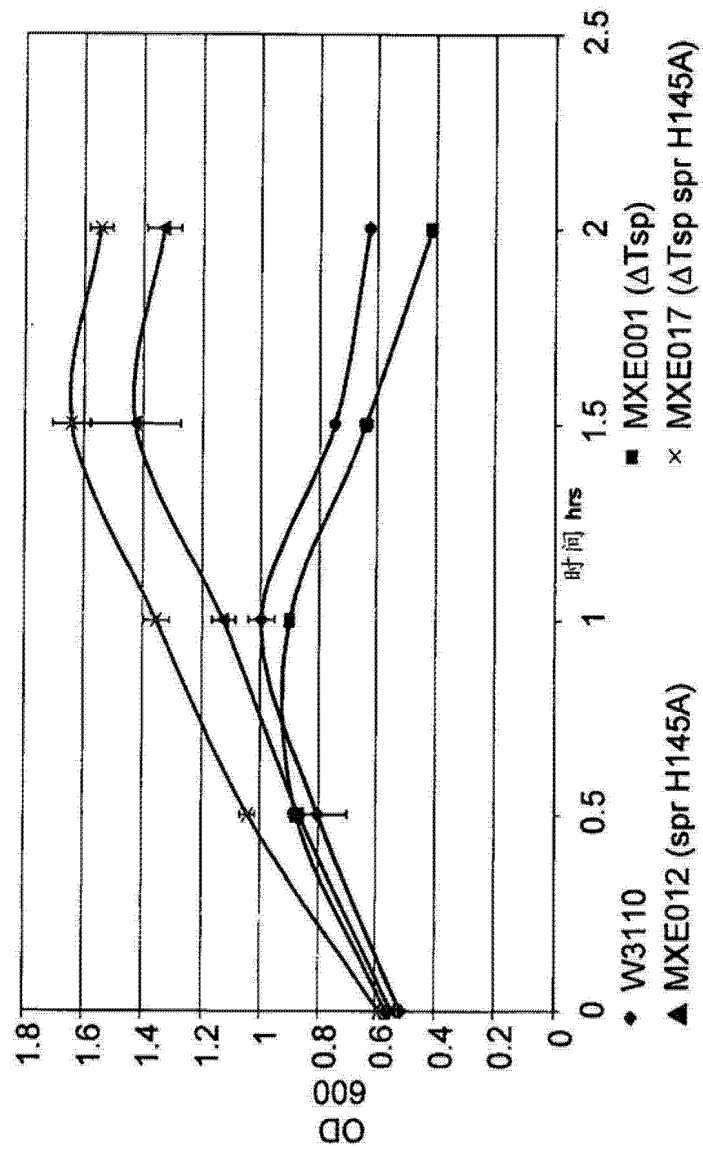


图 1

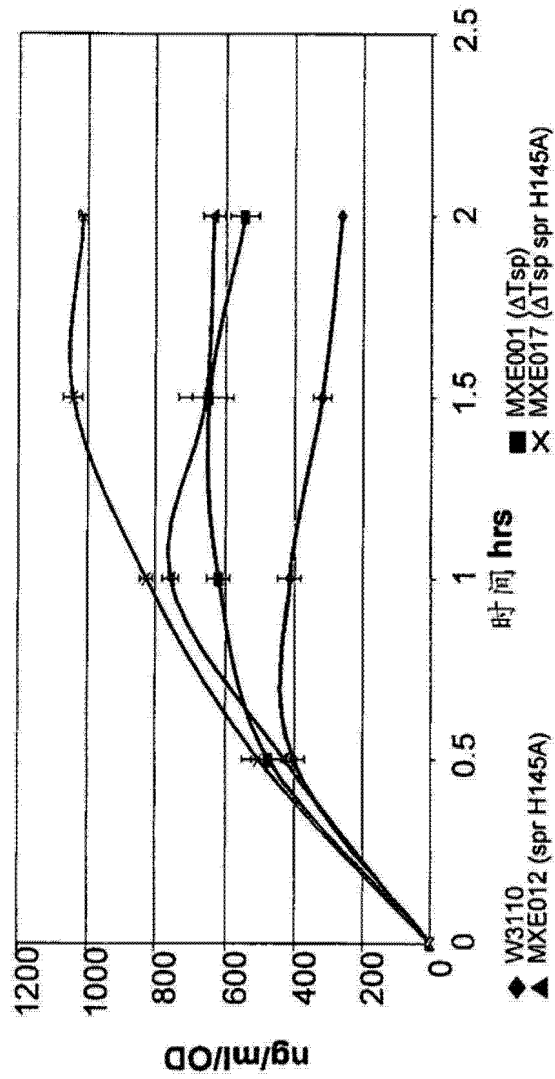


图 2

www.patviewer.com

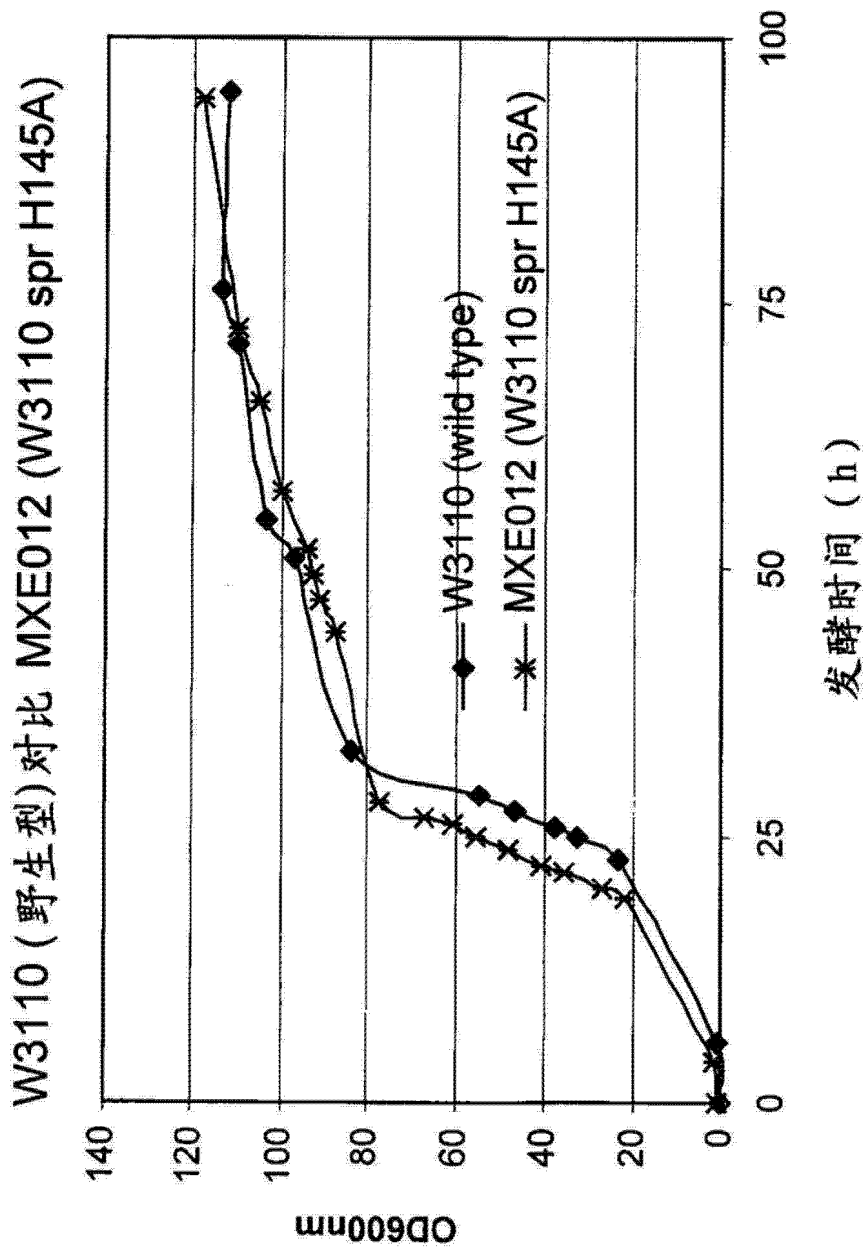


图 3

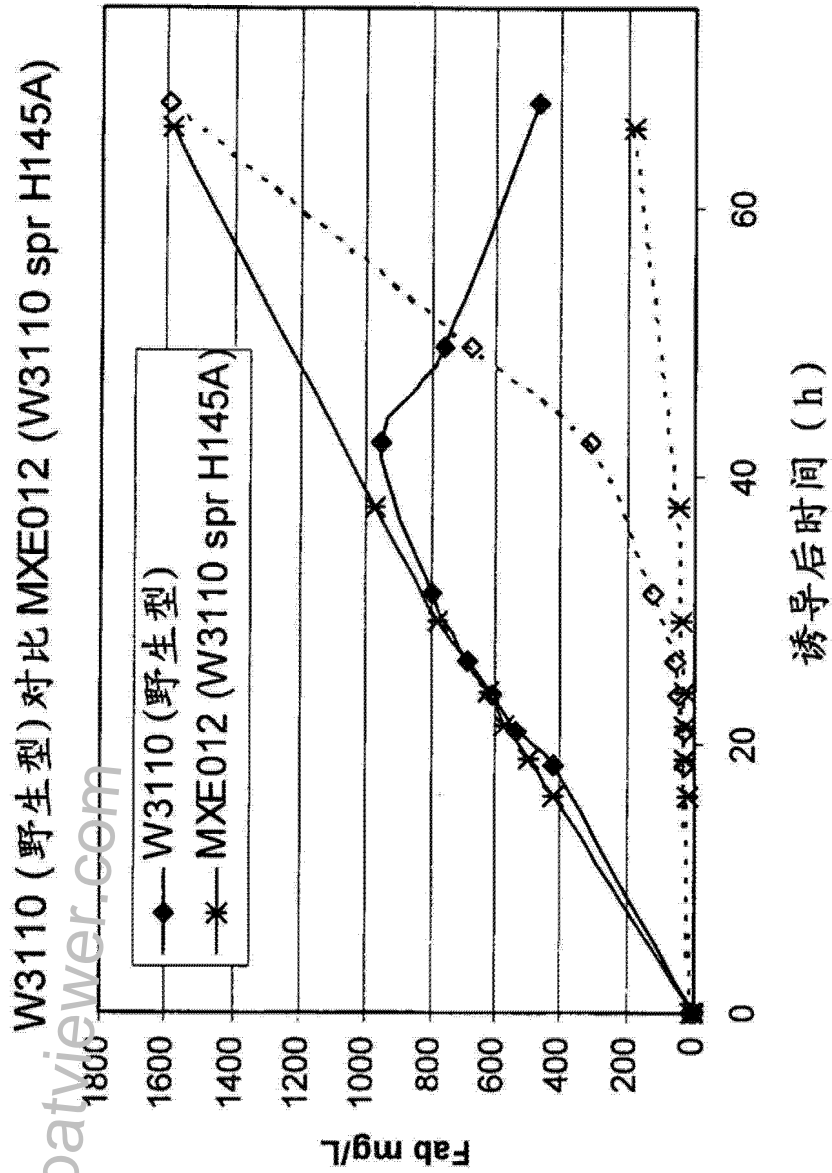


图 4

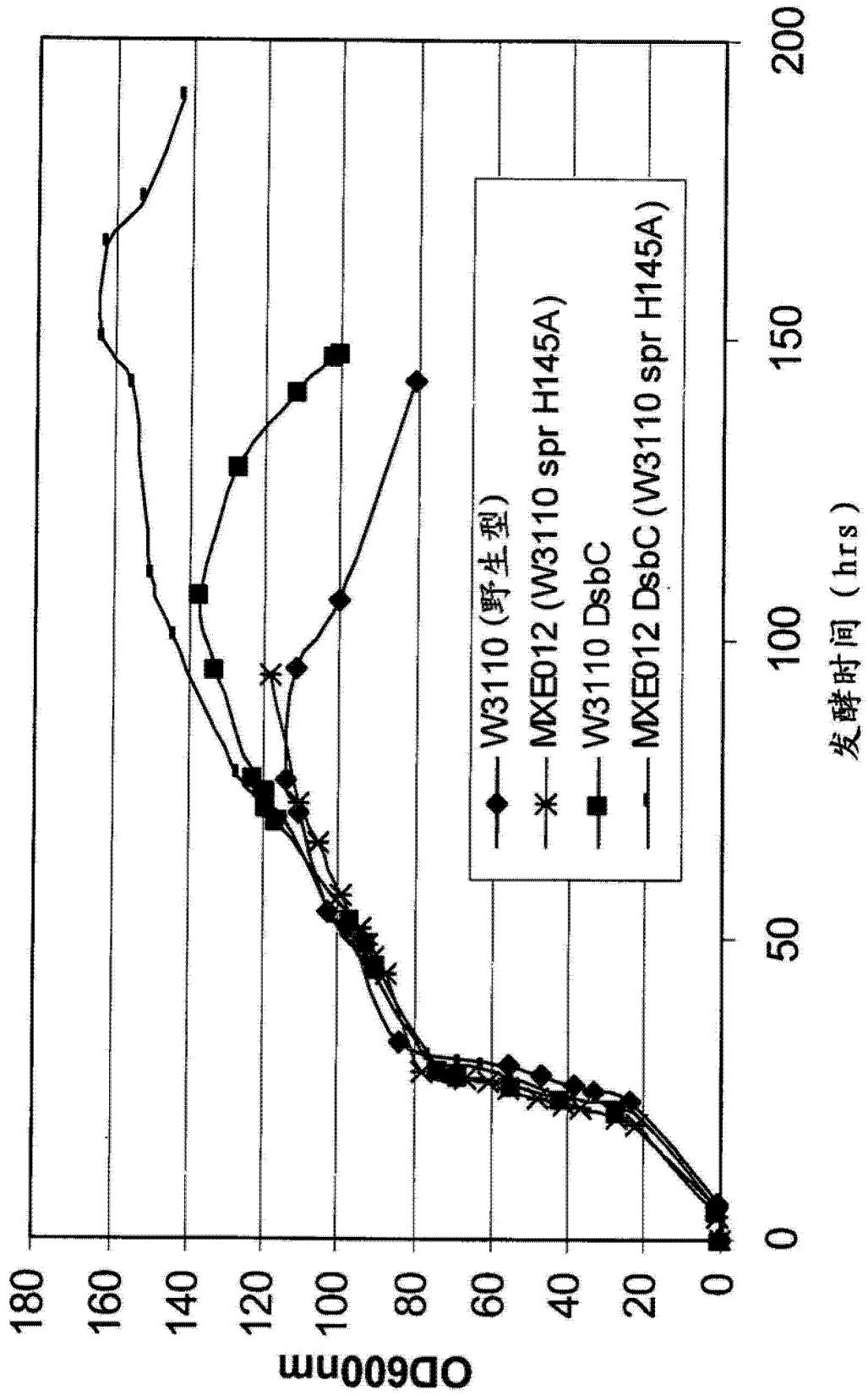


图 5

patviewer.com



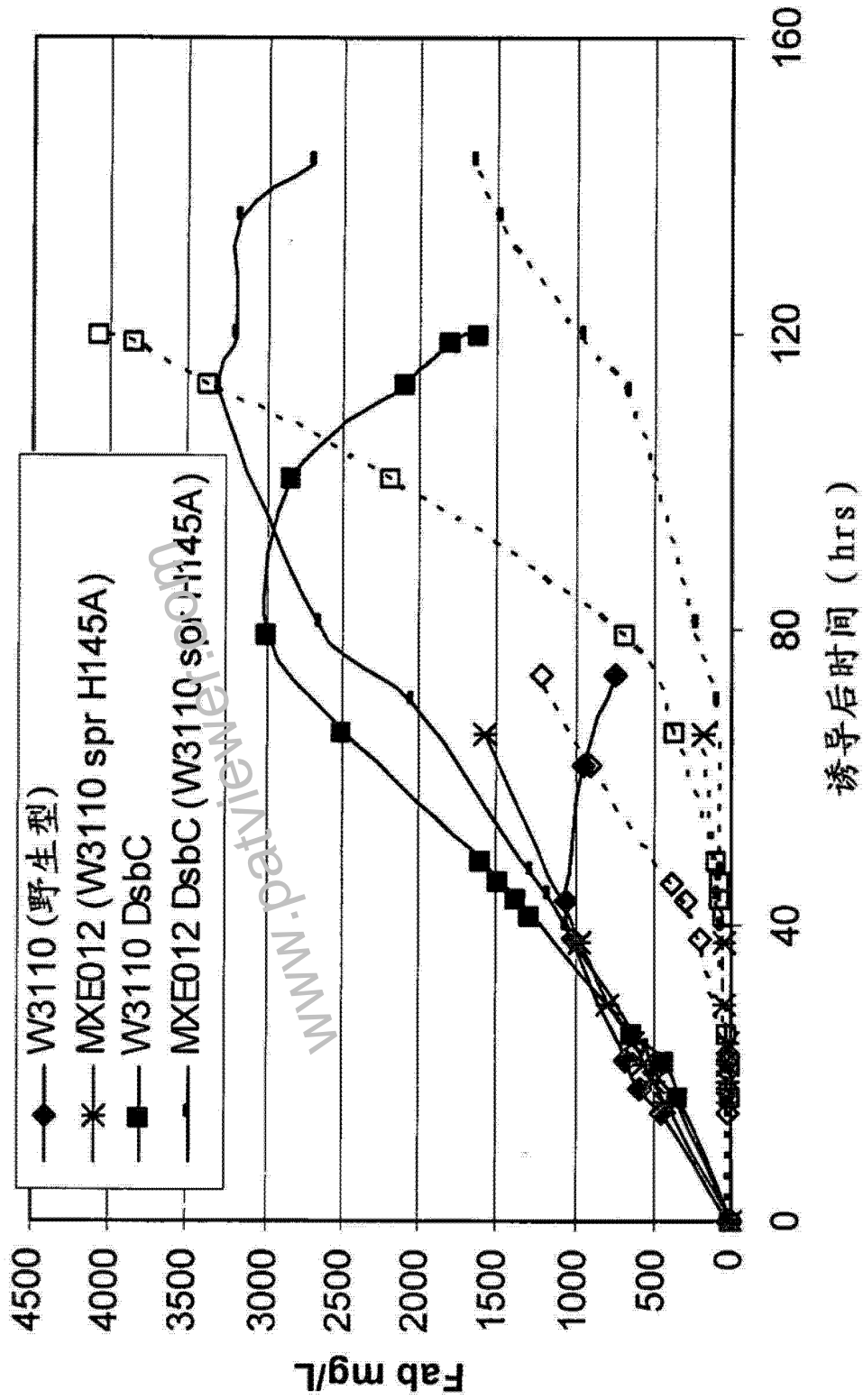


图 6

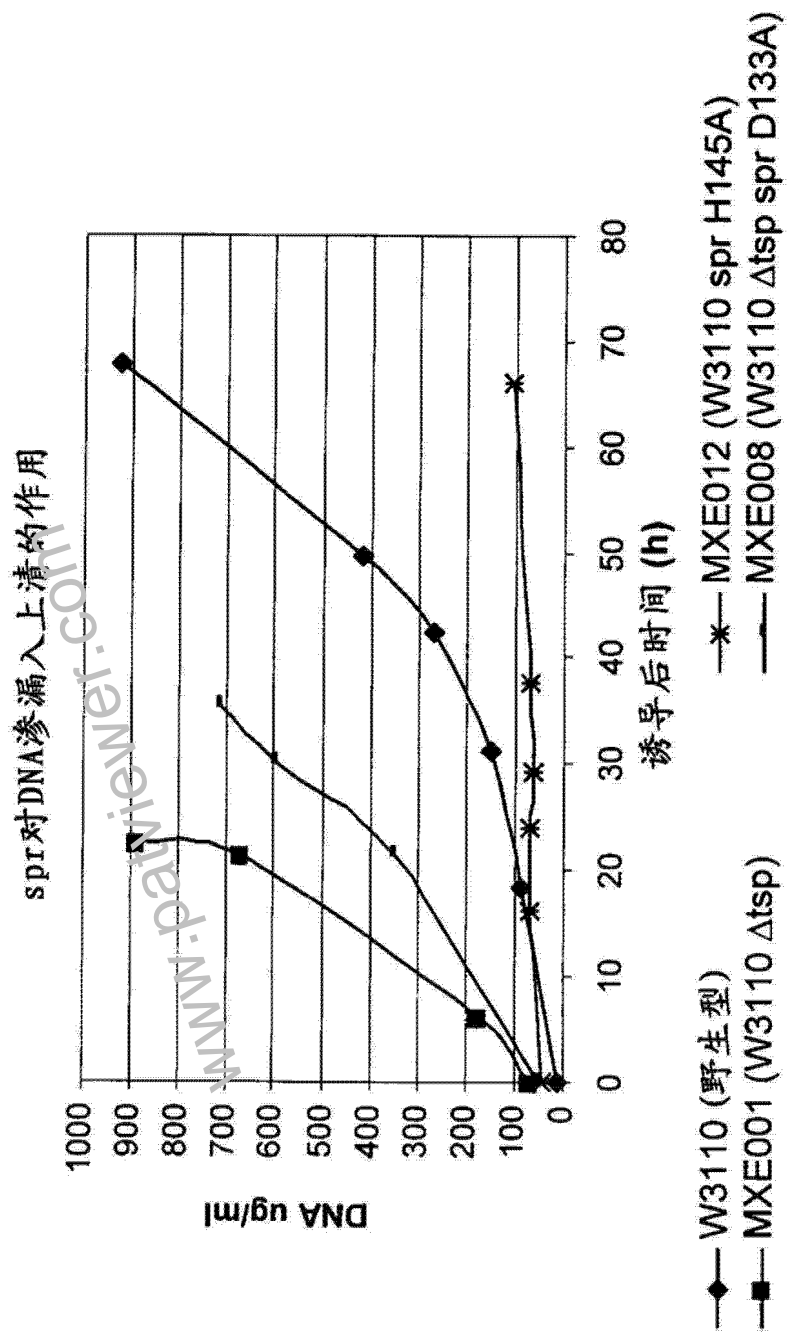


图 7

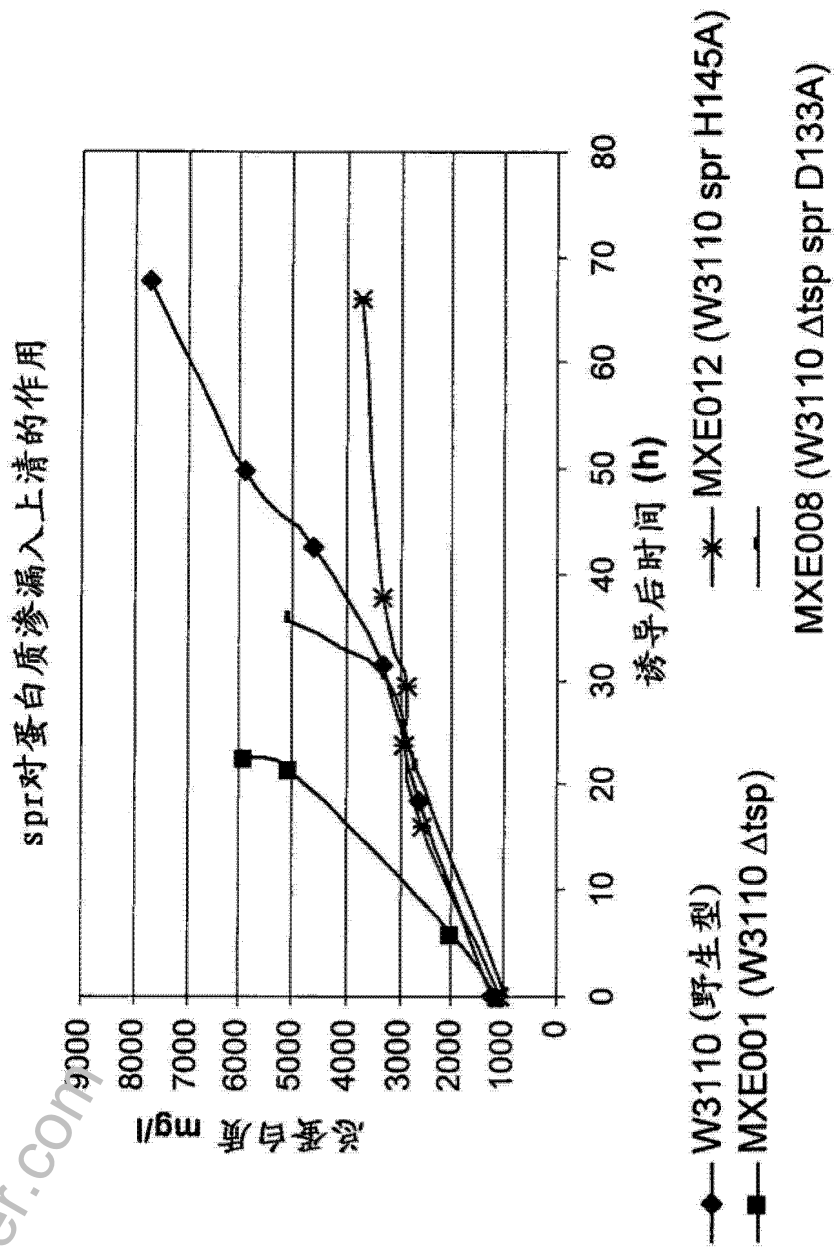


图 8

www.patviewer.com

## 野生型 ptr (蛋白酶 III) 5'.

```

* M P R S T W F K A L L L L V
TGA ATG CCC CGC AGC ACC TGG TTC AAA GCA TTA TTG TTG TTA GTT

A L W A P L S
GCC CTT TGG GCA CCC TTA AGT

```

## 突变 Δ ptr (蛋白酶 III) 5'.

```

EcoR I
~~~~~
* I P R S T W F K A L L L L V
TGA ATT CCC CGC AGC ACC TGG TTC AAA GCA TTA TTG TTG TTA GTT

Ase I
~~~~~
A L W A H * C
GCC CTT TGG GCA CAT TAA TGT

```

图 9a

## 野生型 Tsp 5'.

```

M N M F F R L T A L A G L L A
ATG AAC ATG TTT TTT AGG CTT ACC GCG TTA GCT GGC CTG CTT GCA

I A G Q T F A
ATA GCA GGC CAG ACC TTC GCT

```

## 突变 Δ Tsp 5'.

```

EcoR I
~~~~~
M N S F L G L P R * L A C L Q
ATG AAT TCG TTT TTA GGC TTA CCG CGT TAG CTG GCC TGC TTG CAA

Ase I
~~~~~
* Q A R H * L
TAG CAG GCC AGA CAT TAA TTG

```

图 9b

**野生型 DegP**

202 D A A I N R G N S G G  
 949 GAT GCA GCG ATC AAC CGT GGT AAC TCC GGT GGT

**突变 DegP S210A**

*Ase I*

202 D A A I N R G N A G G

图 9c

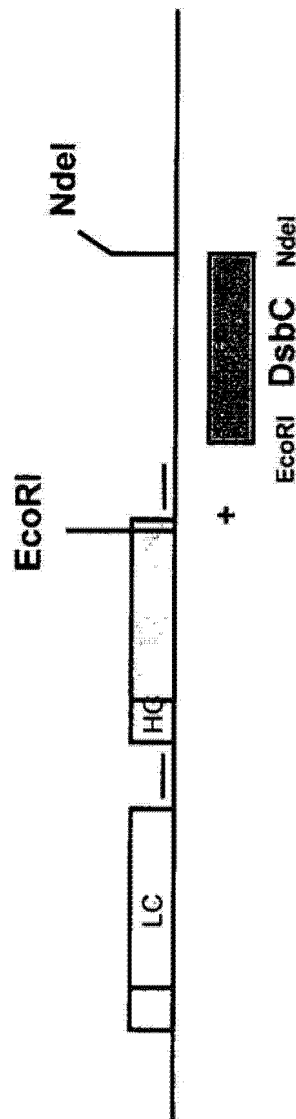


图 10

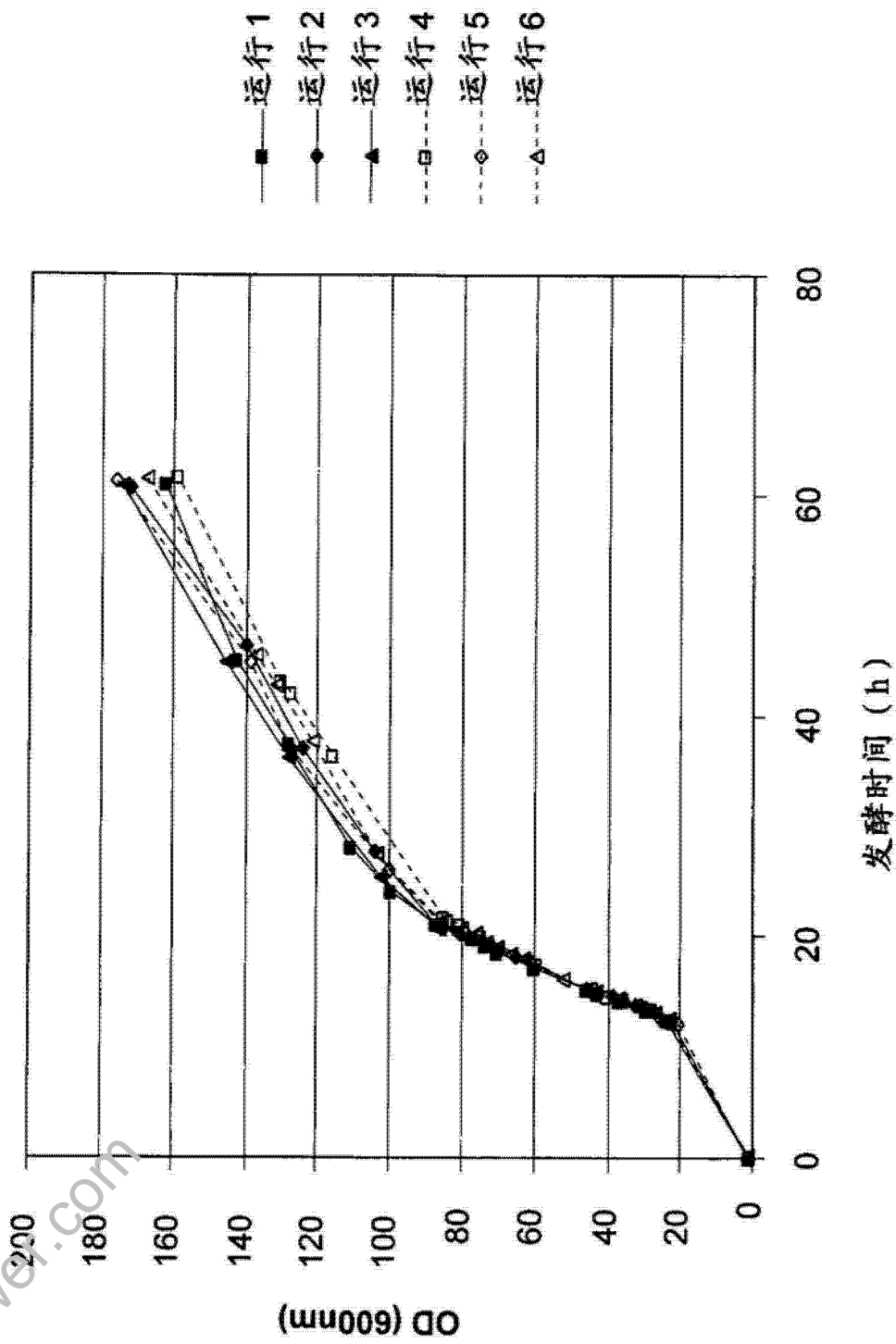


图 11

www.patviewer.com

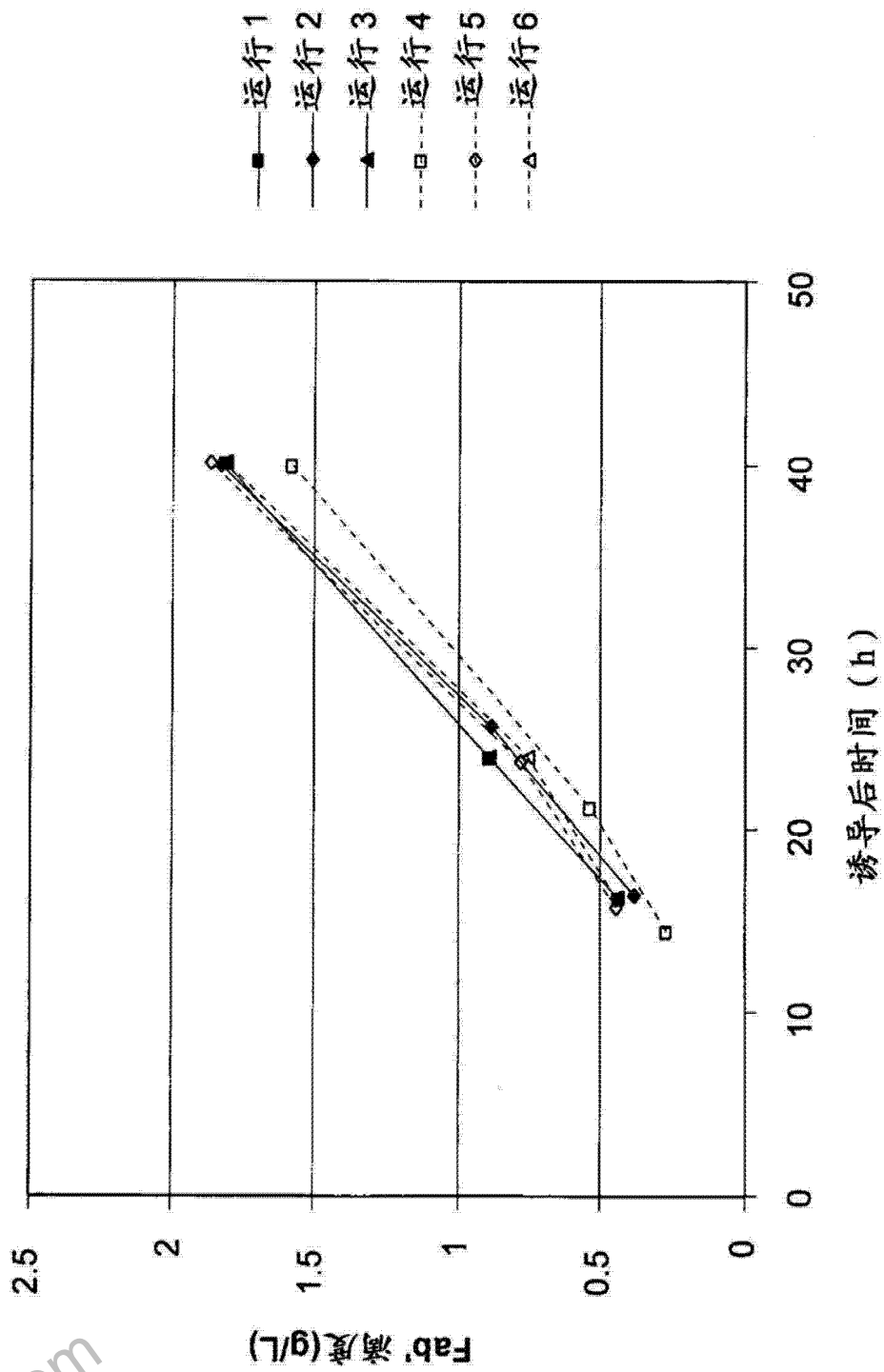


图 12

www.patviewer.com

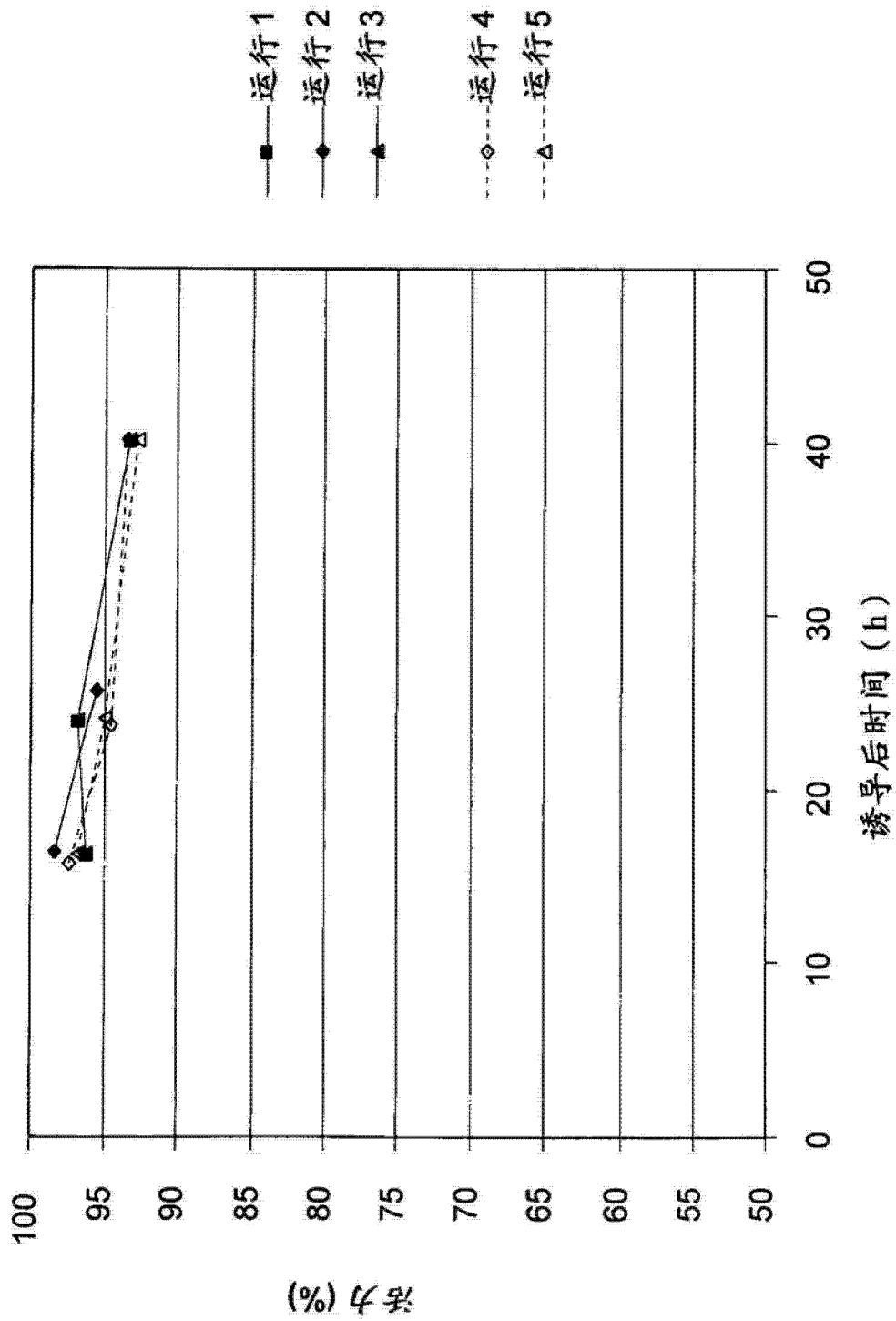


图 13



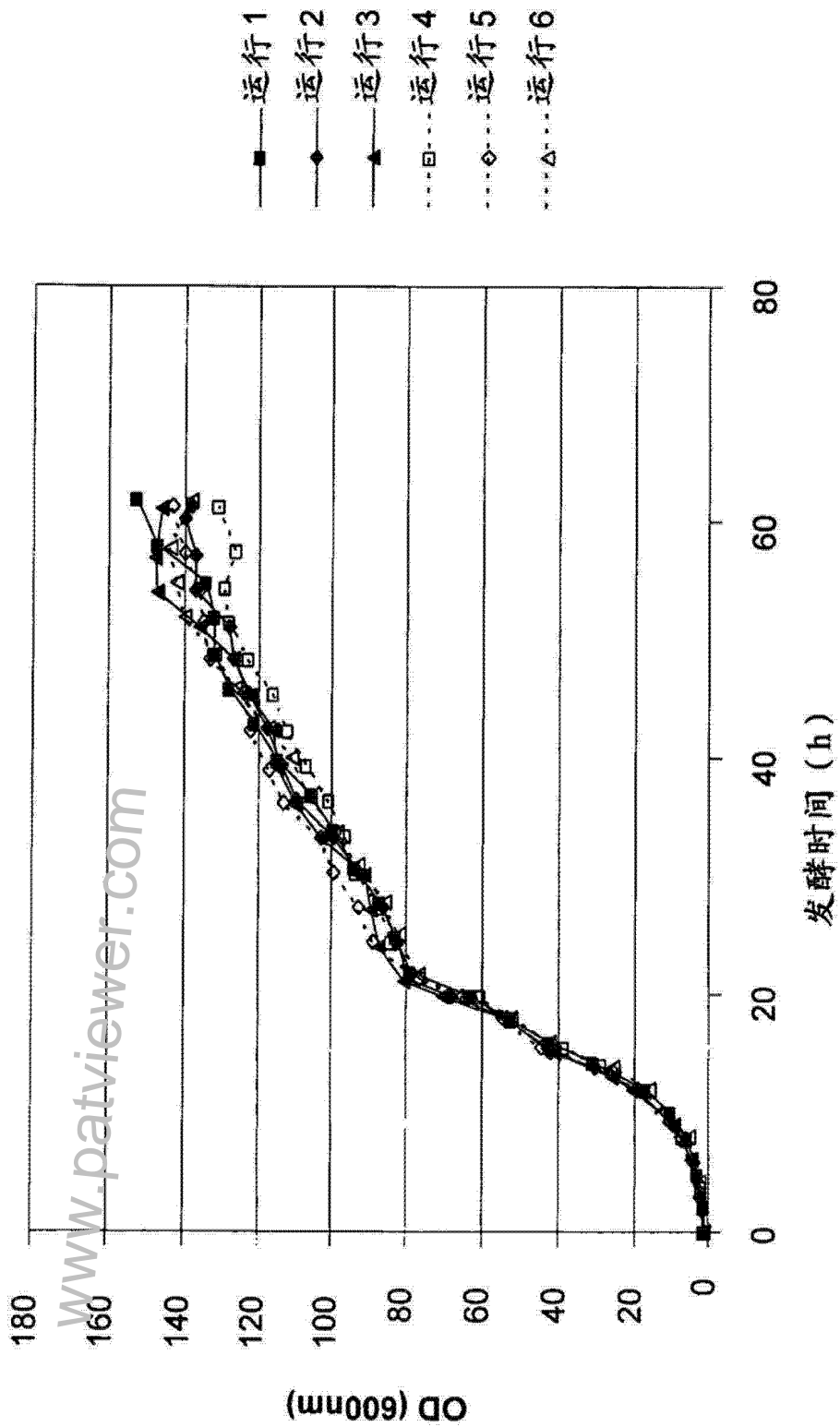


图 14

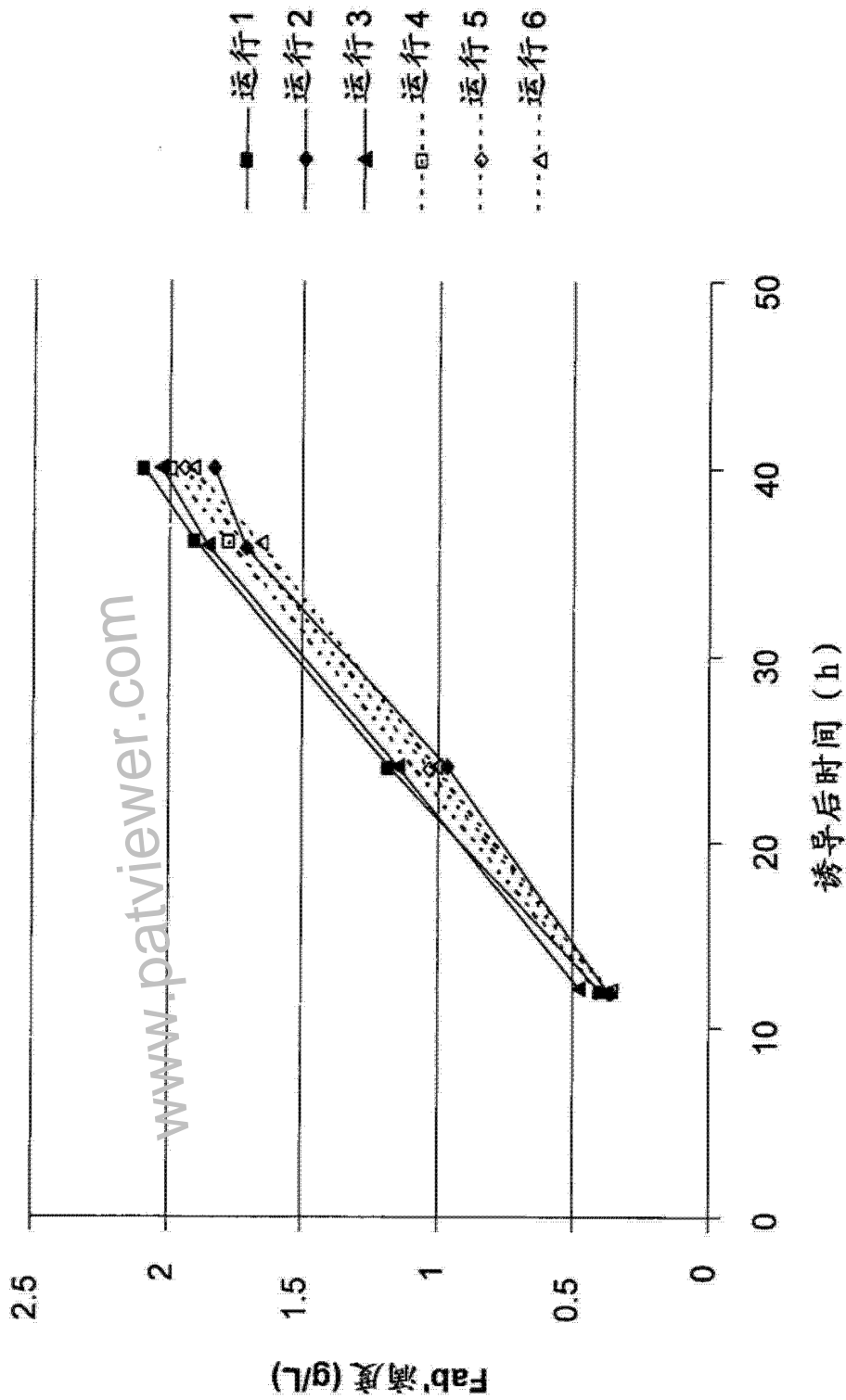


图 15