(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104328079 A
(43)申请公布日 2015.02.04

(21)申请号 201410591520.4

(22)申请日 2011.01.13

(30)优先权数据
1000591.6 2010.01.14 GB

(62)分案原申请数据
201180005978.1 2011.01.13

(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.
C12N 1/21(2006.01)
C12P 21/00(2006.01)
C12R 1/19(2006.01)

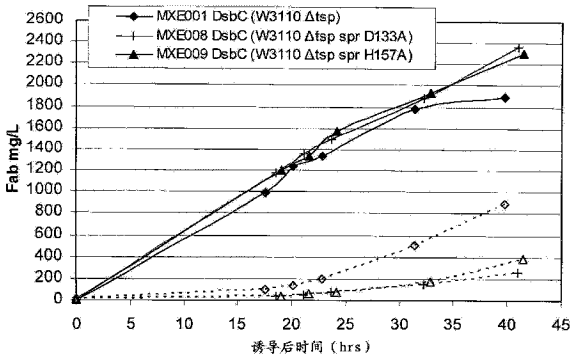
权利要求书2页 说明书38页
序列表27页 附图10页

(54)发明名称

表达重组DSBC并具有降低的TSP活性的细菌宿主菌株

(57)摘要

本发明提供一种重组革兰氏阴性细菌细胞，细胞特征在于：a. 包含编码DsbC的重组多核苷酸；以及b. 与野生型细胞相比其具有降低的Tsp蛋白质活性。



CN 104328079 A

1. 重组革兰氏阴性细菌细胞，特征在于所述细胞：

a. 包含含有编码 DsbC 的重组多核苷酸的表达载体；

b. 具有突变的 Tsp 基因，其编码与野生型非突变 Tsp 相比，具有 50％或更低蛋白酶活性的 Tsp 蛋白质；并且

c. 具有除了突变的 Tsp 基因外，与野生型大肠杆菌 W3110 菌株同基因的基因组。

2. 根据权利要求 1 的细胞，其中所述细胞进一步包含编码突变 spr 蛋白质的突变 spr 基因，所述突变 spr 蛋白质具有选自 H157A、N31Y、R62C、I70T、Q73R、C94A、S95F、V98E、Q99P、R100G、L108S、Y115F、D133A、V135D、V135G、L136P、G140C、R144C、H145A 和 G147C 的一个或多个突变。

3. 根据权利要求 2 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。

4. 根据权利要求 3 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。

5. 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 C94A、D133A、H145A 和 H157A 的突变。

6. 根据权利要求 1 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 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。

7. 根据权利要求 1 的细胞，其中所述细胞包含的突变 Tsp 基因编码与野生型非突变 Tsp 相比，具有 50％或更低的蛋白酶活性的 Tsp 蛋白质或为敲除突变 Tsp 基因。

8. 根据权利要求 6 的细胞，其中所述细胞包含的突变 Tsp 基因编码与野生型蛋白质 Tsp 相比，具有 50％或更低的蛋白酶活性的 Tsp 蛋白质或为敲除突变 Tsp 基因。

9. 根据权利要求 7 或权利要求 9 的细胞，其中所述细胞具有敲除突变基因，其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和基因终止密码子上游的一个或多个终止密码子。

10. 根据权利要求 7 或权利要求 9 的细胞，其中所述敲除突变 Tsp 基因包含由对基因起始密码子的错义突变和任选地一个或多个其它点突变产生的限制性标记物位点。

11. 根据权利要求 7 或权利要求 9 的细胞，其中所述敲除突变 Tsp 基因包含 SEQ ID NO:3。

12. 根据权利要求 1 或 6 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。

13. 根据权利要求 1 的细胞，其中所述细胞包含含有编码 DsbC 的重组多核苷酸以及编码目的蛋白质的多核苷酸序列的载体。

14. 根据权利要求 13 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。

15. 根据权利要求 14 的细胞，其中所述目的蛋白质为抗体或其抗原结合片段。

16. 根据权利要求 15 的细胞，其中所述抗体或其抗原结合片段对 TNF 具有特异性。

17. 用于产生目的蛋白质的方法，其包括：

a. 在能够有效表达目的重组蛋白质及编码 DsbC 的重组多核苷酸的条件下，在培养基中培养根据权利要求 1-16 中任何一项限定的重组革兰氏阴性细菌细胞；以及

b. 从重组革兰氏阴性细菌细胞的细胞周质和 / 或培养基中回收目的蛋白质。

18. 根据权利要求 17 的方法，其中编码目的蛋白质的多核苷酸序列与编码 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。

19. 根据权利要求 17 或 18 的方法，其中所述方法进一步包括将目的蛋白质与 DsbC 分离。

表达重组 DSBC 并具有降低的 TSP 活性的细菌宿主菌株

[0001] 本申请是申请日为 2011 年 1 月 13 日、申请号为 201180005978.1、发明名称为“表达重组 DSBC 并具有降低的 TSP 活性的细菌宿主菌株”的发明专利申请的分案申请。

[0002] 本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。

[0003] 发明背景

[0004] 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。

[0005] 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，包括难以生产蛋白酶敏感型蛋白质。蛋白酶在大肠杆菌细胞周质和细胞质中转换老旧、损坏或折叠错误的蛋白质上起到重要作用。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。

[0006] 已经鉴定出大量的细菌蛋白酶。大肠杆菌中的蛋白酶包括蛋白酶 III(ptr)、DegP、OmpT、Tsp、prlC、ptrA、ptrB、pepA-T、tsh、espc、eatA、clpP 和 lon。

[0007] Tsp(也称为 Prc) 为一种 60kDa 的细胞周质蛋白酶。第一个已知的 Tsp 底物为青霉素结合蛋白质 -3(PBP3)(Determination of the cleavage site involved in C-terminal processing of penicillin-binding protein 3of Escherichia coli ; Nagasawa H,Sakagami Y,Suzuki A,Suzuki H,Hara H,Hirota Y.J Bacteriol.1989Nov ; 171(11):5890-3 和 Cloning,mapping and characterization of the Escherichia coli Tsp gene which is involved in C-terminal processing of penicillin-binding protein 3 ;Hara H,Yamamoto Y,Higashitani A,Suzuki H,Nishimura Y.J Bacteriol.1991Aug ;173(15):4799-813)，但后来发现 Tsp 还能够裂解噬菌体尾巴蛋白质，因此将其改名为尾特异性蛋白酶 (Tail Specific Protease，Tsp)(Silber 等人，Proc. Natl.Acad.Sci.USA,89:295-299(1992))。Silber 等 人 (Deletion of the prc(tsp) gene provides evidence for additional tail-specific proteolytic activity in Escherichia coli K-12 ;Silber,K.R.,Sauer,R.T. ;Mol Gen Genet 1994242:237-240) 描述了一种 prc 删除菌株 (KS1000)，其中的突变是通过将包含 Kanr 标记物的片段取代 prc 基因的片段而形成的。

[0008] Tsp(prc) 活性降低对降低目的蛋白的蛋白水解是可取的。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 等人分离了含有基因外抑制子 (spr) 突变的耐热性回复突变株 (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)。

[0009] 包含有突变 spr 基因的蛋白酶缺陷菌株已有描述。Chen et al(Chen C,Snedecor B,Nishihara JC,Joly JC,McFarland N,Andersen DC,Battersby JE,Champion

KM.Biotechnol Bioeng.2004Mar 5;85(5):463-74)描述了携带prc(Tsp)及另一种蛋白酶DegP的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sprW174R突变的载体上而产生的(重组抗体片段在大肠埃希式杆菌的细胞周质中高水平积累需要三重突变体(ΔDegPΔprc sprW174R)宿主菌株。发现ΔDegP、Δprc和sprW174R突变的组合提供最高水平的抗体轻链、抗体重链及F(ab')2-LZ。EP1341899公开了在编码蛋白酶DegP和Prc的染色体DegP和prc分别有缺陷的大肠杆菌菌株,此菌株还具有突变spr基因,其编码的蛋白质抑制了包含prc突变体的菌株表现出的生长表型。

[0010] 蛋白质二硫键异构酶为在蛋白质折叠时,催化其中半胱氨酸残基见二硫键形成和断裂的酶。已知其与催化二硫键形成的蛋白质共表达而改善宿主细胞中的蛋白质表达。WO98/56930公开了一种在细菌中产生含二硫键的异源多肽的方法,其中原核二硫键异构酶如DsbC或DsbG与真核多肽共表达。US6673569公开了包含编码DsbA、DsbB、DsbC和DsbD中每种的多肽的人工操纵子,用于产生外源蛋白质。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。

[0011] 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 Jounral vol 13,no 8,p2007-2012,1994)。

[0012] 出乎意料地发现在革兰氏阴性细菌细胞中由重组DsbC过表达DsbC改善了缺乏蛋白酶Tsp的细胞的细胞裂解表型。因此,本发明提供了新菌株,其具有用于产生目的蛋白质的优势特性。

[0013] 发明概述

[0014] 本发明提供了重组革兰氏阴性细菌细胞,细胞的特征在于:

[0015] a)包含编码DsbC的重组多核苷酸;以及

[0016] b)与野生型细胞相比,具有降低的Tsp蛋白质活性。

[0017] 在一个实施方式中,细胞包含野生型spr基因。在此实施方式中,优选地,除了降低Tsp蛋白质活性所需的修饰之外,细胞基因组与野生型细菌细胞是同基因的。

[0018] 在一个进一步的实施方式中,根据本发明的细胞与野生型细胞相比具有降低的Tsp蛋白质活性,并包含编码DsbC的重组多核苷酸以及突变spr基因。在此实施方式中,优选地,除了突变spr基因以及相比野生型细胞降低Tsp蛋白质活性所需的修饰之外,细胞基因组与野生型细菌细胞是同基因的。

[0019] 具有上述特定遗传修饰特定组合的革兰氏阴性细菌细胞显示出优势的生长以及蛋白质生产表型。

[0020] 本发明还提供了用于生产目的蛋白质的方法,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蛋白质。所述方法包括在能够有效表达目的蛋白质及编码DsbC的

重组多核苷酸的条件下，在培养基中培养如上文限定的重组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 / 或培养基中回收目的蛋白质。

[0021] 附图简述

[0022] 图 1 显示了表达抗 -TNF Fab' 的菌株 MXE001、MXE008 以及表达抗 -TNFα Fab' 和重组 DsbC 的菌株 MXE001 和 MXE008 的生长概况。

[0023] 图 2 显示了表达抗 -TNF Fab' 和重组 DsbC 的大肠杆菌菌株 MXE001、MXE008 和 MXE009 的细胞周质（阴影图标）及上清（空的非阴影图标）中抗 -TNF Fab' 产量。

[0024] 图 3 显示了表达抗 -TNF Fab' 的大肠杆菌菌株 MXE001 和 MXE008 以及表达抗 -TNF Fab' 和重组 DsbC 的大肠杆菌菌株 MXE001 和 MXE008 的细胞周质（阴影图标）及上清（空的非阴影图标）中抗 -TNF Fab' 产量。

[0025] 图 4 显示表达 Fab A 和 Fab B 的大肠杆菌菌株 W3110 以及表达 Fab A 和重组 DsbC 或 Fab B 和重组 DsbC 的大肠杆菌菌株 MXE008 的细胞周质的抗 -TNF Fab' 产量。

[0026] 图 5a 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。

[0027] 图 5b 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。

[0028] 图 5c 显示了野生型 DegP 及突变 DegP 蛋白质和基因序列的区域。

[0029] 图 6 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。

[0030] 图 7 显示了表达抗 -TNF Fab' 及重组 DsbC 的大肠杆菌菌株 MXE008 的 5L 和 200L 发酵物的生长概况比较。

[0031] 图 8 显示了表达抗 -TNF Fab' 及重组 DsbC 的大肠杆菌菌株 MXE008 哦 5L 和 200L 发酵物的 Fab' 滴度比较。

[0032] 图 9 显示了表达抗 -TNF Fab' 及重组 DsbC 的大肠杆菌菌株 MXE008 和 MXE009 的发酵物的生长概况的比较。

[0033] 图 10 显示了表达抗 -TNF Fab' 及重组 DsbC 的大肠杆菌菌株 MXE008 和 MXE009 的发酵物的 Fab' 滴度的比较。

[0034] 序列简述

[0035] 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。

[0036] SEQ ID NO:2 是野生型 Tsp 蛋白质的氨基酸序列。

[0037] 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。

[0038] SEQ ID NO:4 是野生型蛋白酶 III 基因的 DNA 序列。

[0039] SEQ ID NO:5 是野生型蛋白酶 III 蛋白质的氨基酸序列。

[0040] SEQ ID NO:6 是突变的敲除蛋白酶 III 基因的 DNA 序列

[0041] SEQ ID NO:7 是野生型 DegP 基因的 DNA 序列。

[0042] SEQ ID NO:8 是野生型 DegP 蛋白质的氨基酸序列。

[0043] SEQ ID NO:9 是突变 DegP 基因的 DNA 序列。

[0044] SEQ ID NO:10 是突变 DegP 蛋白质的氨基酸序列。

[0045] SEQ ID NO:11 是抗 -TNF 抗体轻链可变区的氨基酸序列。

[0046] SEQ ID NO:12 是抗 -TNF 抗体重链可变区的氨基酸序列。

[0047] SEQ ID NO:13 是抗 -TNF 抗体轻链的氨基酸序列。

[0048] SEQ ID NO:14 是抗 -TNF 抗体重链的氨基酸序列。

[0049] 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。

[0050] 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。

[0051] 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。

[0052] 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。

[0053] 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。

[0054] 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。

[0055] 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。

[0056] SEQ ID NO:22 是野生型的 spr 基因序列，不包括信号序列。

[0057] SEQ ID NO:23 是突变 OmpT 序列的核苷酸序列，包括 D210A 和 H212A 突变。

[0058] SEQ ID NO:24 是突变 OmpT 序列的氨基酸序列，包括 D210A 和 H212A 突变。

[0059] SEQ ID NO:25 是突变敲除 OmpT 序列的核苷酸序列。

[0060] SEQ ID NO:26 是带 his 标签的 DsbC 的核苷酸序列。

[0061] SEQ ID NO:27 是带 his 标签的 DsbC 的氨基酸序列。

[0062] SEQ ID NO:28 显示了 hTNF40 的 CDRH1 氨基酸序列。

[0063] 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。

[0064] SEQ ID NO:30 显示了 hTNF40 的 CDRH3 氨基酸序列。

[0065] SEQ ID NO:31 显示了 hTNF40 的 CDRL1 氨基酸序列。

[0066] SEQ ID NO:32 显示了 hTNF40 的 CDRL2 氨基酸序列。

[0067] SEQ ID NO:33 显示了 hTNF40 的 CDRL3 氨基酸序列。

[0068] SEQ ID NO:34 显示了 hTNF40 的 CDRH2 氨基酸序列。

[0069] SEQ ID NO:35 显示了 OmpA 寡核苷酸转接子的序列。

[0070] 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。

[0071] 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。

[0072] 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。

[0073] 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。

[0074] 本发明优选实施方式详述

[0075] 现在将以更详细的方式描述本发明。

[0076] 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或更少的氨基酸。

[0077] 除非上下文另有说明，术语“多核苷酸”包括基因、DNA、cDNA、RNA、mRNA等。

[0078] 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”。

[0079] 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，其中所述细胞未修饰以降低Tsp蛋白质活性或携带重组DsbC基因和任选的突变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。

[0080] 词句“细胞”、“细胞系”、“细胞培养物”及“菌株”可交换使用。

[0081] 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裂解。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℃或更高）下，细胞在低渗培养基中死亡。

[0082] 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与野生型细胞相比降低Tsp蛋白质活性所需的修饰和任选的突变spr基因外具有几乎相同或相同的核算序列。在此实施方式中，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，与野生型细胞相比，本发明的细胞的基因组除了与野生型细胞相比降低Tsp蛋白质活性所需的修饰和任选的突变spr基因外具有基本上相同或相同的基因组序列。在一个实施方式中，考虑到任何可以天然发生的突变，根据本发明的细胞与野生型细胞相比除了与野生型细胞相比降低Tsp蛋白质活性所需的修饰和任选的突变spr基因外有几乎相同的基因组序列，包含任何可以发生的天然发生的突变。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与野生型细胞相比降低Tsp蛋白质活性所需的修饰（任选地为突变spr基因）外可以具有完全相同的基因组序列。

[0083] 编码DsbC的重组多核苷酸可存在于转化入细胞和/或整合入宿主细胞基因组的合适的表达载体上。在编码DsbC的重组多核苷酸插入到宿主基因组的实施方式中，由于具有插入的编码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上，因此对宿主细胞基因组造成破坏最小。

[0084] 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于转化入细胞和/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，本发明的细胞也与野生型细胞不同。优选地，编码目的蛋白质的多核苷酸是在细胞内表达载体上，因此对宿主细胞基因组造成破坏最小。

[0085] 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实例为

W3110，如 W3110K-12 菌株。

[0086] 本发明的发明者提供了适合于表达目的蛋白质的重组革兰氏阴性细菌细胞，其与野生型细胞相比具有降低的 Tsp 蛋白质活性并包含编码 DsbC 的重组多核苷酸。

[0087] 根据本发明的细胞包括编码 DsbC 的重组多核苷酸。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，和 / 或具有编码 his 标签的序列。用于本发明的修饰 DsbC 核苷酸序列的一个实例显示于 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。

[0088] 我们已经发现，在与野生型细胞相比具有降低的 Tsp 蛋白质活性的细菌细胞中表达编码 DsbC 的重组多核苷酸这样的特定组合和在在优选的实施方式中再组合突变 spr 基因，提供了用于表达目的蛋白质的改良的宿主。上述遗传突变的特定组合提供了新菌株，其相比于携带敲除突变 Tsp 基因的细胞具有改善的细胞健康和生长表型。携带突变 Tsp 基因的细胞可以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，尤其是在高的细胞密度下。因此，包含突变 Tsp 基因的细胞表型为易于裂解，尤其在高的细胞密度下。然而，本发明细胞中 DsbC 的表达抑制了 Tsp 降低的表型，因此细胞显示出降低的裂解情况。

[0089] 根据本发明的细胞相比于携带敲除突变 Tsp 基因的细胞显示出提高的蛋白质生产产量。提高的蛋白质产量可以为细胞中蛋白质生产速率和 / 或蛋白质生产的持续时间。所述提高的蛋白质产量可以为细胞周质蛋白质产量和 / 或上清蛋白质产量。在一个实施方式中，由于从细胞的渗漏减少，相比于携带突变 Tsp 基因的细胞，本发明的细胞显示出提高的细胞周质蛋白质产量。所述重组细菌细胞可以能够以更快的速率产生目的蛋白质，因此，相比于包含突变 Tsp 基因的细胞，相同量的目的蛋白质可在更短的时间内产生。所述产生目的蛋白质的更快速率可以在细胞生长的起始阶段内表现尤其显著，例如在诱导蛋白质表达后的前 5、10、20 或 30 小时内。

[0090] 根据本发明的细胞优选地在细胞周质和 / 或培养基中表达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。

[0091] 与野生型细胞相比，本发明提供的细胞具有降低的蛋白酶活性，其可降低目的重组蛋白质的蛋白质裂解，尤其对于对 Tsp 蛋白酶具有蛋白质裂解敏感型的目的蛋白质。因此，与野生型细菌细胞相比，由本发明提供的细胞提供了完整蛋白质优选目的蛋白质的更高产量以及蛋白质优选目的蛋白质的的更低产量或优选地不产生蛋白质裂解片段。

[0092] 本领域技术人员能够使用本领域熟知的方法容易地测试候选的细胞克隆以确定其是否具有想要的目的蛋白产量，所述方法包括发酵方法、ELISA 以及蛋白质 G HPLC。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):1133-48,2001 Sep;and 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。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。

[0093] 在一个实施方式中,根据本发明的细胞还表达如下的一种或多种蛋白质:

[0094] •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

[0095] •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

[0096] •一种或多种能够促进二硫键形成的蛋白质,如DsbA、DsbB、DsbD、DsbG。

[0097] 可将一种或多种以上蛋白质整合入细胞基因组和/或插入表达载体。

[0098] 在一个实施方式中,根据本发明的细胞包含编码以下一种或多种其他蛋白质的重组多核苷酸:

[0099] •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

[0100] •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

[0101] •一种或多种能够促进二硫键形成的蛋白质,如DsbA、DsbB、DsbD、DsbG。

[0102] 在本发明的实施方式中,细胞还包含突变spr基因。Spr蛋白质为大肠杆菌膜结合细胞周质蛋白酶。

[0103] spr蛋白质的野生型氨基酸序列显示于SEQ ID NO:21(包含N-端信号序列)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蛋白质序列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)。

[0104] 在根据本发明的细胞包含突变spr基因的实施方式中,突变spr基因优选地为细胞的染色体spr基因。突变spr基因编码能够抑制包含突变Tsp基因的细胞表型的spr蛋白质。携带突变Tsp基因的细胞可以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,尤其是在高的细胞密度下。因此,包含突变Tsp基因的细胞表型为易于裂解,尤其在高的细胞密度下。携带突变Tsp基因的细胞还显示出在低渗条件下的热敏感型生长。然而,本发明细胞携带的spr突变在引入具有降低的Tsp活性的细胞中抑制了Tsp降低的表型,因此细胞显示出降低的裂解情况,尤其在高的细胞密度下。细胞生长表型可由本领域技术人员在摇瓶培养或高细胞密度发酵技术中容易地进行测量。对细胞裂解表型的抑制可见于改善的生长速率和/或重组蛋白质生产尤其是细胞周质中,由携带spr突变体

并具有降低 Tsp 活性的细胞相比于携带 Tsp 突变体及野生型 spr 的细胞所表现出的。

[0105] 根据本发明的细胞优选地包含编码 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，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。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，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 蛋白质选自 N31、R62、I70、Q73、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144 和 G147 的在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。在此实施方式中，spr 蛋白质优选地不具有任何其他突变。

[0106] 本发明的发明者已鉴定出能够抑制包含突变 Tsp 基因的细胞的生长表型的 spr 突变。

[0107] 出乎意料地，本发明者发现相比相比于包含突变 Tsp 基因的细胞，携带重组 DsbC 基因、新的突变 spr 基因并与野生型细胞具有降低的 Tsp 蛋白质活性的细胞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 突变以及具有降低的 Tsp 蛋白质活性的细胞显示出降低的细胞裂解表型。

[0108] 对一个或多个以上 spr 氨基酸的突变可为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和 / 或具有不同化学特性的氨基酸。

[0109] 在一个实施方式中，突变是对 C94、H145 和 H157 构成的催化三分子中的一个、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)。

[0110] 因此，突变 spr 基因可包含：

[0111] • 对 C94 的突变；或

[0112] • 对 H145 的突变；或

[0113] • 对 H157 的突变；或

[0114] • 对 C94 和 H145 的突变；或

[0115] • 对 C94 和 H157 的突变；或

[0116] • 对 H145 和 H157 的突变；或

[0117] • 对 C94、H145 和 H157 的突变。

[0118] 在此实施方式中，spr 蛋白质优选地不具有任何其他突变。

[0119] 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 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 Tsp 基因的细胞的表型的 spr 蛋白质。

[0120] 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。

[0121] 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 ：

[0122] • 对 N31 的突变 ；或

[0123] • 对 R62 的突变 ；或

[0124] • 对 I70 的突变 ；或

[0125] • 对 Q73 的突变 ；或

[0126] • 对 S95 的突变 ；或

[0127] • 对 V98 的突变 ；或

[0128] • 对 Q99 的突变 ；或

[0129] • 对 R100 的突变 ；或

[0130] • 对 L108 的突变 ；或

[0131] • 对 Y115 的突变 ；或

[0132] • 对 D133 的突变 ；或

[0133] • 对 V135 的突变 ；或

[0134] • 对 L136 的突变 ；或

[0135] • 对 G140 的突变 ；或

[0136] • 对 R144 的突变 ；或

[0137] • 对 G147 的突变。

[0138] 在一个实施方式中，突变 spr 蛋白质包含氨基酸的多重突变 ：

[0139] • S95 和 Y115 ；

[0140] • N31、Q73、R100 和 G140 ；

[0141] • Q73、R100 和 G140 ；

[0142] • R100 和 G140 ；

[0143] • Q73 和 G140 ；或

[0144] • Q73 和 R100 ；或

[0145] • R62、Q99 和 R144 ；或

[0146] • Q99 和 R144。

[0147] 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 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。

[0148] 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突变。

[0149] 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。

[0150] 在一个进一步的实施方式中，spr 蛋白质具有多重突变，选自 :

[0151] • S95F 和 Y115F

[0152] • N31Y、Q73R、R100G 和 G140C ;

[0153] • Q73R、R100G 和 G140C ;

[0154] • R100G 和 G140C ;

[0155] • Q73R 和 G140C ;

[0156] • Q73R 和 R100G ;

[0157] • R62C、Q99P 和 R144C ;或

[0158] • Q99P 和 R144C。

[0159] 优选地，突变 spr 基因编码的 spr 蛋白质具有选自 C94A、D133A、H145A 和 H157A 的突变。

[0160] 在一个进一步的实施方式中，突变 spr 基因编码的 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。

[0161] 根据本发明的细胞与野生型细胞相比具有降低的 Tsp 蛋白质活性。词句“与野生型细胞相比具有降低的 Tsp 蛋白质活性”指的是与野生型细胞的 Tsp 活性相比，所述细胞的 Tsp 活性有所降低。细胞可通过任何合适的方式进行修饰以降低 Tsp 活性。

[0162] 在一个实施方式中，降低的 Tsp 活性来自对编码 Tsp 的内源多核苷酸和 / 或相关的调控表达序列进行的修饰。所述修饰可减少或停止 Tsp 基因转录及翻译或可提供表达的 Tsp 蛋白质，其与野生型 Tsp 蛋白质相比具有降低的蛋白酶活性。

[0163] 在一个实施方式中，修饰了相关的调控表达序列以降低 Tsp 表达。例如，可对 Tsp 基因的启动子进行突变以阻止基因的表达。

[0164] 在一个优选的实施方式中，根据本发明的细胞携带编码具有降低蛋白酶活性的 Tsp 蛋白质突变 Tsp 基因或敲除突变 Tsp 基因。优选地，突变的为染色体 Tsp 基因。

[0165] 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 为一种细胞周质蛋白酶，能够作用于青霉素结合蛋白质 -3(PBP3) 及噬菌体尾巴蛋白质。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。

[0166] 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的蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因或者敲除突变 Tsp 基因。

[0167] 词句“编码具有降低的蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因”在本发明上下文中指的是相比于野生型非突变 Tsp 基因，突变 Tsp 基因不具有完全的蛋白酶活性。

[0168] 优选地，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。更

优选地，突变 Tsp 基因编码不具有蛋白酶活性的 Tsp 蛋白质。在此实施方式中，细胞没有染色体 Tsp 缺陷，即，Tsp 基因序列未经删除或突变以致阻止任何形式的 Tsp 蛋白质的表达。

[0169] 任何合适的突变可引入 Tsp 基因中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由本领域技术人员容易地通过本领域任何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.28864－28868,1995Kenneth C.Keiler和Robert T.Sauer)的方法，其中测试了 Tsp 的蛋白酶活性。

[0170] 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、G376E433 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。

[0171] 因此，突变的 Tsp 基因可包含：

[0172] • 对 S430 的突变；或

[0173] • 对 D441 的突变；或

[0174] • 对 K455 的突变；或

[0175] • 对 S430 和 D441 的突变；或

[0176] • 对 S430 和 K455 的突变；或

[0177] • 对 D441 和 K455 的突变；或

[0178] • 对 S430、D441 和 K455 的突变。

[0179] 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：

[0180] • S430A 或 S430C；和／或

[0181] • D441A；和／或

[0182] • K455A 或 K455H 或 K455R。

[0183] 优选地，Tsp 基因包含点突变 S430A 或 S430C。

[0184] 词句“敲除突变 Tsp 基因”在本发明上下文中指的是 Tsp 基因包含一个或多个突变，阻止由野生型基因编码的的 Tsp 蛋白质表达，以提供细胞的 Tsp 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 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。

[0185] 在一个优选的实施方式中，Tsp 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。在一个实施方式中，Tsp 基因包含对基因起始密码子和 / 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 5b 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。在优选的实施方式中，T 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 5b 所示。在优选的实施方式中，Tsp 还可突变为插入 Ase I 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 5b 所示。

[0186] 在一个优选的实施方式中，敲除突变 Tsp 基因具有 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 序列中所做的突变显示于图 5b。在一个实施方式中，突变 Tsp 基因具有 SEQ ID NO:3 的 7-2048 号核苷酸的 DNA 序列。

[0187] 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还包含突变 DegP 基因和 / 或突变 ptr 基因和 / 或突变 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。

[0188] 在一个实施方式中，本发明提供重组革兰氏阴性细菌细胞，其包含

[0189] a. 编码 DsbC 的重组多核苷酸；

[0190] b. 编码具有降低的蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因或敲除突变 Tsp 基因；

[0191] 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，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；以及

[0192] d. 任选地，突变 spr 基因。

[0193] 优选地，在此实施方式中，除了以上突变 b、c、d 之外，所述细胞的基因组与野生型细菌细胞是同基因的。

[0194] 在一个实施方式中，本发明提供重组革兰氏阴性细菌细胞，其包含

[0195] a. 编码 DsbC 的重组多核苷酸；

[0196] b. 编码具有降低的蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因或敲除突变 Tsp 基因；

[0197] 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；以及

[0198] d. 任选地，突变 spr 基因。

[0199] 优选地，在此实施方式中，除了以上突变 b、c、d 之外，所述细胞的基因组与野生型细菌细胞是同基因的。

[0200] 在一个实施方式中本发明提供的细胞包含

[0201] a. 编码 DsbC 的重组多核苷酸；

[0202] b. 编码具有降低的蛋白酶活性的 Tsp 蛋白质的突变 Tsp 基因或敲除突变 Tsp 基因；

[0203] 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；

[0204] 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 蛋白质或为敲除突变 ptr 基因；

[0205] e. 任选地包含突变 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；以及

[0206] f. 任选地，突变 spr 基因。

[0207] 优选地，在此实施方式中，除了以上突变 b、c、d、e 和 f 之外，所述细胞的基因组与野生型细菌细胞是同基因的。

[0208] 在本发明实施方式中，细胞包含突变 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。

[0209] 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 et al Microbiology 2008, 154, 3649-3658)。

[0210] 在细胞包含 DegP 突变的这些实施方式中，细胞中的 DegP 突变提供了突变 DegP 基因，编码了具有分子伴侣活性但不具有完全蛋白酶活性的 DegP 蛋白质。

[0211] 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。

[0212] 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 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而阻止任何形式的 DegP 蛋白质的表达。

[0213] 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) 以及描述于 The proteolytic activity of the HtrA(DegP) protein from Escherichia coli at low temperatures, Skorko-Glonek J et al Microbiology 2008, 154, 3649-3658 中的方法。

[0214] 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, 244:19-61Rawlings 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、两个或三个产生的错义突变。

[0215] 因此，突变 DegP 基因可包含：

[0216] • 对 His105 的突变；或

[0217] • 对 Asp135 的突变；或

[0218] • 对 Ser210 的突变；或

[0219] • 对 His105 和 Asp135 的突变；或

[0220] • 对 His105 和 Ser210 的突变；或

[0221] • 对 Asp135 和 Ser210 的突变；或

[0222] • 对 His105、Asp135 和 Ser210 的突变。

[0223] 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 5c 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)。

[0224] 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)。

[0225] 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 5c 所示。

[0226] 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 5c。

[0227] 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。

[0228] 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。

[0229] 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℃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℃或更高的温度下生长，包括大约 30℃ - 大约 37℃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。

[0230] 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。

[0231] 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。

[0232] 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。

[0233] 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。

[0234] 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。

[0235] 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。

[0236] 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。

[0237] 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。

[0238] 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 5a。

[0239] 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。

[0240] 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，会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，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取代基因序列的一段而实现。例如，可插入 Ase I 限制性位点，其包含终止密码子 TAA。

[0241] 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 5a 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 基因序列中所做的突变显示于图 5a.

[0242] 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，一种或多种根据本发明的细胞相比于通过将外源 DNA 插入基因编码序列而敲除蛋白酶基因的细胞，其可以显示出改善的生长特征和 / 或蛋白质表达。

[0243] 在实施方式中，根据本发明的细胞包含突变 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。

[0244] 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，除非另有说明。

[0245] 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蛋白质的蛋白酶活性的50％或更少、40％或更少、30％或更少、20％或更少、10％或更少或者5％或更少。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任何合适的突变而阻止任何形式的OmpT蛋白质的表达。

[0246] 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)以及Dekker等人(Substrate Specitificity of the Integral Membrane Protease OmpT Determined by Spatially Addressed Peptide Libraries,Biochemistry 2001,40,1694-1701)的方法。

[0247] 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为活性降低～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)提出其的一个新型催化位点，其活性中心包含一个Asp83-Asp85对以及一个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。

[0248] 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、H212、G216、K217、R218&E250中一个或多个残基的错义突变。

[0249] 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、H101A E111A、E136A、E193A、D206A、D208A、D210A、H212A、H212N、H212Q、G216K、K217G、K217T、R218L&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。

[0250] 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敲除。

[0251] 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 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。

[0252] 合适的突变敲除 OmpT 序列显示于 SEQ ID NO:24。

[0253] 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。

[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。

[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敲除突变 ptr 基因，如染色体 ptr 缺陷。

[0256] 包括敲除突变在内的许多遗传工程改造突变涉及抗生素抗性标记物的用途，这样能够选择和鉴定成功突变的细胞。然而，如上文所讨论的，使用抗生素抗性标记物有大量的不利之处。

[0257] 在本发明的一个进一步的实施方式中克服了上文所述使用抗生素抗性标记物的劣势，其中突变 Tsp 基因，任选的突变 spr、任选的突变 DegP 基因和 / 或突变 ptr 基因和 / 或突变 OmpT 基因可经突变为包含一个或多个标记物位点。所述限制性位点通过遗传工程改造进入基因，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筛选和鉴定。经修饰携带一个或多个所述突变蛋白酶基因的细胞可通过对来自细胞裂解物的基因组 DNA 进行 PCR 而分析，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。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，然后可通过琼脂糖凝胶电泳分析孵育之前和孵育之后的扩增 DNA。以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。

[0258] 在细胞携带具有 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子细胞的基因组 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分析以证实基因组 DNA 中有突变 ptr 基因存在。

[0259] 在细胞包含具有 SEQ ID NO:3 的 DNA 序列或 SEQ ID NO:3 的 7-2048 号核苷酸的敲除突变 Tsp 基因的实施方式中，SEQ ID NO:15 和 SEQ ID NO:16 中所示的寡核苷酸引物序列可用于扩增来自转化子细胞的基因组 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分析以证实基因组 DNA 中有突变 Tsp 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进行突变和 / 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，阻止基因编码序列转录完全。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 5a 所示。例如，在一个实施方式中，其中通过引入 Ase I 位点在 Tsp 基因 21 号密码子处引入终止密码子，这也创造出限制性位点，如图 5b 所示。

[0263] 限制性标记物位点可通过对起始密码子的突变和任选地一个或多个其他点突变而插入。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。这一点尤其有利，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 5a 所示。例如，在一个实施方式中，其中 Tsp 基因的起始密码子 (3 号密码子) 由 ATG 变为 TCG，如图 1b 所示，进一步的对 2 号密码子的点突变 AAC 至 AAT 以及对 3 号密码子的突变 ATG 至 TCG 制造出 EcoR I 限制性标记物位点，如图 5b 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，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 基因包括 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 DegP 基因或 spr 基因中，标记物限制性位点可使用沉默密码子的变化而引入。例如，Ase I 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 5c 所示的 DegP 序列。

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

[0267] 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。任何大肠杆菌菌株可用于本发明，但优选地所使用的为野生型 W3110 菌株，如 K-12W3110。

[0268] 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，尤其是用于治疗的蛋白质，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。

[0269] 在一个优选的实施方式中，细胞只携带引入编码 DsbC 的重组多核苷酸的最小突变；降低 Tsp 蛋白质活性所需的修饰，以及任选地突变 spr 基因。对细胞基因组只进行最小的突变以引入重组多核苷酸和突变。细胞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的其他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可显示有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进一步破坏的细胞，本发明提供的细胞还更适合用于生产治疗性蛋白质。

[0270] 因此，本发明还提供了革兰氏阴性细菌细胞，其具有除了降低 Tsp 蛋白质活性所需的重组修饰以及任选的突变 spr 基因之外与野生型细菌细胞是同基因的基因组。所述细胞进一步携带编码 DsbC 的重组多核苷酸以及任选地携带编码目的蛋白质的多核苷酸序列。

[0271] 在一个优选的实施方式中，除了与野生型细胞相比降低 Tsp 蛋白质活性所需的修饰及任选的突变 spr 基因外，所述细胞与野生型大肠杆菌细胞如 W3110 菌株是同基因的。所述细胞进一步包含编码 DsbC 的重组多核苷酸并任选地包含编码目的蛋白质的多核苷酸序列。

[0272] 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。优选地，所述多核苷酸为细胞内的表达载体，这样对宿主细胞基因组造成的破坏最小。

[0273] “目的蛋白质”在本发明的上下文中意欲指代要进行表达的多肽，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。

[0274] 如本文所使用的，“重组多肽”指的是使用重组 DNA 技术构建的蛋白质。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。

[0275] 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。

[0276] 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。

[0277] 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、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。

[0278] 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。

[0279] 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。

[0280] 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 具有蛋白水解敏感性。

[0281] 优选地,蛋白质为真核多肽。

[0282] 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 (llama)、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。

[0283] 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’)2、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。

[0284] 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。

[0285] 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。

[0286] 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。

[0287] 本发明还提供了重组革兰氏阴性细菌细胞，其包含编码 DsbC 的重组多核苷酸、突变 Tsp 基因（其中突变 Tsp 基因编码的 Tsp 蛋白质具有降低的蛋白酶活性或为敲除突变 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。在一个优选的实施方式中，所述细胞进一步包含编码突变 spr 蛋白质的突变 spr 基因。

[0288] 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）。

[0289] 在一个实施方式中，抗体具有针对人 TNFα 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。

[0290] 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 .

[0291] 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。

[0292] 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。

[0293] 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。

[0294] 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。

[0295] 优选地，抗体分子具有针对人 TNF（之前称为 TNFα）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。

[0296] 抗 -TNF 抗体优选地为 Fab 或 Fab' 片段 .

[0297] 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。

[0298] 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。

[0299] 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 WO05/003171 或 WO05/003170（其内容并入本文作为参考）所描述的。WO05/003171 或 WO05/003170 也描述了合适的效应子分子。

[0300] 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-α Fab'，如 WO01/094585 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）。

[0301] 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。

[0302] 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 WO04/035792（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。

[0303] 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。

[0304] 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 蛋白质要尽量避开。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。

[0305] 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。

[0306] 在本发明的实施方式中，其中染色体 Tsp 基因发生突变，并且任选地染色体 spr 基因、DegP 基因、ptr 基因和 OmpT 基因中一种或多种为突变的，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的合适的技术。可采用合适的载体，其能使自身通过同源重组整合入宿主染色体。在细胞包含两种或三种上述突变基因的实施方式中，突变基因可在相同或不同载体上引入革兰氏阴性细菌。

[0307] 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 pSC101replicon, 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 pKO3 质粒 (Link 等人，1997, Journal of Bacteriology, 179, 6228-6237)。

[0308] 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。编码 DsbC 的重组多核苷酸可使用合适的载体如 pKO3 质粒整合入细胞基因组。

[0309] 备选地或另外地，编码 DsbC 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，其携带编码 DsbC 的重组多核苷酸，其一般包含编码 DsbC 的蛋白质编码序列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。

[0310] 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 P31。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。

[0311] 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。

[0312] 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。

[0313] 在细胞还表达一种或多种其他如下蛋白质的实施方式中，

[0314] • 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

[0315] • 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

[0316] • 一种或多种能够促进二硫键形成的蛋白质，如 DsbA、DsbB、DsbD、DsbG。

[0317] 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。

[0318] 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。

[0319] 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。

[0320] 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 ：

[0321] • 质粒，如 pBR322 或 pACYC184 和 / 或

[0322] • 病毒载体如细菌噬菌体

[0323] • 转座遗传元件如转座子

[0324] 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、启动子以和及转录终止子，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。

[0325] 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。

[0326] 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。

[0327] 所使用的一种或多种启动子可为可诱导启动子。

[0328] 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。

[0329] 用于细菌系统的表达单位通常还包含 Shine-Dalgamo(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。

[0330] 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 P33，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。

[0331] 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。

[0332] 终止子可以是宿主细胞内源或外源的。合适的终止子为 rrnB。

[0333] 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,Sept1996,p 512-538。

[0334] 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。

[0335] 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。

[0336] 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。

[0337] 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。

[0338] 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。

[0339] 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。

[0340] 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 6。

[0341] 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。

[0342] 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、表达盒和 / 或包含其中所采用的组分的宿主细胞。

[0343] 根据本发明的第三方面，提供了用于产生目的重组蛋白质的方法，包括如上文本文的第一或第二个方面所描述的，将所述目的重组蛋白质在重组革兰氏阴性细菌细胞中表达。用于产生目的重组蛋白质的方法包括 ：

[0344] a. 在有效表达重组目的蛋白质以及编码 DsbC 的重组多核苷酸的条件下，于培养基中培养如上文所限定的革兰氏阴性细菌细胞 ；以及

[0345] b. 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重组目的蛋白质。

[0346] 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。

[0347] 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的方式并入宿主细胞。如上文所讨论的，一般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包含编码 DsbC 的多核苷酸的表达盒。

[0348] 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。

[0349] 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。

[0350] 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的人工设置条件。

[0351] 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、卡那霉素或氨苄青霉素基因。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。

[0352] 在一个实施方式中，根据本发明的方法进一步包括在培养基中培养转化子细胞以表达目的蛋白质的步骤。

[0353] 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和 / 或 DsbC。合适的可诱导表达系统以及组成型启动子在本领域熟知。

[0354] 在其中编码 DsbC 的多核苷酸以及编码目的蛋白质的多核苷酸处于相同或分别的可诱导启动子控制的实施方式中，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。

[0355] 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。

[0356] 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。

[0357] 可从菌株中回收多肽，包括从细胞质、细胞周质及 / 或上清中。

[0358] 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。

[0359] 在一个实施方式中，所述方法进一步包括将重组目的蛋白质与 DsbC 分离。

[0360] 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。

[0361] 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。

[0362] 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。

[0363] 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

[0364] 细胞系

[0365] 对于所有实验，使用大肠杆菌细胞系 W3110 作为亲本野生型细胞系。

[0366] 产生携带以下突变的细胞系：

[0367] a. 突变 Tsp 基因；

[0368] b. 突变 Tsp 基因并携带重组 DsbC；

[0369] c. 突变 Tsp 基因以及突变 spr 基因；

[0370] d. 突变 Tsp 基因及突变 spr 基因并携带重组 DsbC。

[0371] 实施例 1——产生携带突变 Tsp 基因的细胞 MXE001(ΔTsp)

[0372] 通过如下方法产生 MXE001 菌株：

[0373] 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 pKO3 质粒内。pKO3 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，pKO3 系统将所有选择性标记物从宿主基因组移除。

[0374] 构建了以下质粒。

[0375] pMXE191：包含如 SEQ ID NO:3 所示的敲除突变 Tsp 基因，其包含 EcoR I 和 Ase I 限制性标记物。

[0376] 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。

[0377] 第 1 天将 40μl 的大肠杆菌细胞与 (10pg)1μl 的 pKO3DNA 混合于冰冷的 BioRad 0.2cm 电穿孔小杯，然后以 2500V、25μF、200Ω 进行电穿孔。立即加入 1000μl 的 2xPY，细胞在培养箱中 30℃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 100μl 的等分试样铺于 30℃和 43℃下预热的、含有 20μg/ml 氯霉素的 2xPY 琼脂平板。将平板在 30℃和 43℃下孵育过夜。

[0378] 第 2 天在 30℃下生长的菌落数量给出电穿孔效率的估算值，而在 43℃下存活生长的菌落代表潜在的整合事件。挑取 43℃平板上的单一菌落并重悬于 10ml 的 2xPY 中。将 100μl 的悬液铺于含有 5% (w/v) 蔗糖并预热到 30℃的 2xPY 琼脂平板上以产生单菌落。将平板在 30℃孵育过夜。

[0379] 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 20μg/ml 氯霉素或 5% (w/v) 蔗糖的 2xPY 琼脂平板。将平板在 30℃下孵育过夜。

[0380] 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℃孵育过夜产生单一菌落。

[0381] 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。

[0382] 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：

[0383] 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：

[0384] 6284 Tsp 3'5'-GCATCATAATTTTCTTTTTACCTC-3'(SEQ ID NO:15)

[0385] 6283 Tsp 5'5'-GGGAAATGAACCTGAGCAAAACGC-3'(SEQ ID NO:16)

[0386] 通过将单克隆细胞在 20ul 1x PCR 缓冲液中于 95℃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。

[0387] 使用 Tsp 寡核苷酸对扩增各菌株。

[0388] 使用标准 PCR 步骤扩增 DNA。

[0389]

5ul	缓冲液 x10 （Roche）
1ul	dNTP 混合物 （Roche, 10mM mix）
1.5ul	5'寡核苷酸（5 pmol）
1.5ul	3'寡核苷酸（5 pmol）
2ul	细胞裂解物

[0390]

0.5ul	Taq DNA 聚合酶（Roche 5U/ul）
38.5ul	H2O

[0391] PCR 循环。

[0392]

94 ℃	1 分钟
94 ℃	1 分钟）
55 ℃	1 分钟）重复进行 30 个循环
72 ℃	1 分钟）
72 ℃	10 分钟

[0393] 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℃下孵育 2 小时。

[0394] 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)。

[0395] 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。

[0396] 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。

[0397] 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。

[0398] W3110PCR 扩增的 DNA 无法由 Ase I 限制性酶进行消化。

[0399] 实施例 2 产生携带突变 spr 基因的细胞系以及携带突变 Tsp 基因和突变 spr 基因的细胞系

[0400] 使用互补测定法产生并选择 spr 突变。

[0401] 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(ΔTsp)。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、25μF、200Ω 的条件下进行电转化。电转化后,加入 1ml 的 SOC(Invitrogen)(预热至 37℃)并对细胞轻柔摇动,使之在 37℃下复苏 1 小时。

[0402] 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℃孵育。将形成菌落的细胞在 43℃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。

[0403] 使用此方法分离出与 ΔTsp 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:

[0404] 1.V98E

[0405] 2.D133A

[0406] 3.V135D

[0407] 4.V135G

[0408] 5.G147C

[0409] 6.S95F 和 Y115F

[0410] 7.I70T

[0411] 8.N31T、Q73R、R100G、G140C

[0412] 9.R62C、Q99P、R144C

[0413] 10.L108S

[0414] 11.L136P

[0415] 使用以上鉴定的 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)1 rph1(ATCC 号 27325)) 产生携带 spr 突变菌株或使用来自实施例 1 的 MXE001 菌株产生 ΔTsp/ 突变 SPR 组合突变菌株。

[0416] 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。

[0417] 表 1

[0418]

大肠杆菌细胞突变体菌株	基因型	Spr 载体
MXE001	ΔTsp	-
MXE008	ΔTsp、spr D133A	pMXE339、pK03 spr D133A(-SalI)
MXE009	ΔTsp、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	ΔTsp、spr V135D	pMXE340、pK03 spr V135D(-SalI)
MXE015	ΔTsp、spr V98E	pMXE342、pK03 spr V98E(-SalI)
MXE016	ΔTsp、spr C94A	pMXE343、pK03 spr C94A(-SalI)
MXE017	ΔTsp、spr H145A	pMXE344、pK03 spr H145A(-SalI)
MXE018	ΔTsp、spr V135G	pMXE341、pK03 spr V135G(-SalI)

[0419] 突变 SPR 整合盒以 Sal I、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。

[0420] 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, 4617-4622and 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。

[0421] 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。

[0422] pMXE336、pKO3spr S95F(-SalI)

[0423] pMXE337、pKO3spr Y115F(-SalI)

[0424] pMXE338、pKO3spr G147C(-SalI)

[0425] pMXE339、pKO3spr D133A(-SalI)

[0426] pMXE340、pKO3spr V135D(-SalI)

[0427] pMXE341、pKO3spr V135G(-SalI)

[0428] pMXE342、pKO3spr V98E(-SalI)

[0429] pMXE343、pKO3spr C94A(-SalI)

[0430] pMXE344、pKO3spr H145A(-SalI)

[0431] pMXE345、pKO3spr H157A(-SalI)

[0432] pMXE346、pKO3spr W174R(-SalI)

[0433] 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 ΔTsp/ 突变 SPR 组合突变菌株,如表 1 所示。

[0434] 第 1 天将 40μl 的电感受态大肠杆菌细胞或MXE001细胞与(10pg)1μl 的pKO3DNA混合于冰冷的 BioRad 0.2cm 电穿孔小杯,然后以 2500V、25μF、200Ω 进行电穿孔。立即加入 1000μl 的 2xPY,细胞在培养箱中 30℃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 100μl 的等分试样铺于 30℃和 43℃下预热的含有 20μg/ml 氯霉素的 2xPY 琼脂平板。将平板在 30℃和 43℃下孵育过夜。

[0435] 第 2 天在 30℃下生长的菌落数量给出电穿孔效率的估算值,而在 43℃下存活生长的菌落代表潜在的整合事件。挑取 43℃平板上的单一菌落并重悬于 10ml 的 2xPY。将 100μl 的悬液铺于含有 5%(w/v) 蔗糖并预热到 30℃的 2xPY 琼脂平板上以产生单克隆 P51。将平板在 30℃孵育过夜。

[0436] 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 20μg/ml 氯霉素或 5%(w/v) 蔗糖的 2xPY 琼脂平板。将平板在 30℃下孵育过夜。

[0437] 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℃孵育过夜产生单一菌落。

[0438] 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。

[0439] 实施例 3- 产生用于 Fab’和 DsbC 表达的质粒

[0440] 构建了包含抗 -TNF Fab’的重链和轻链序列以及编码 DsbC 序列的质粒。

[0441] WO01/94585 所描述的抗 -TNFα 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)而下游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 (IGS)2。

[0442] 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(pTTO 40.4CDP870IGS2),其为 CDP870Fab'(抗 -TNF Fab')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 6 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。

[0443] Fab 轻链、重链基因及 DcbC 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。

[0444] 实施例 4- 抗 -TNF Fab' 与 DsbC 在所述突变体细胞系中的表达

[0445] 抗 -TNF Fab' 与 DsbC 的表达

[0446] 以实施例 3 中产生的质粒转化实施例 1 中提供的 MXE001 菌株以及实施例 2 提供的突变体菌株 MXE008 和 MXE009。

[0447] 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。

[0448] 抗 -TNF Fab' 的表达

[0449] 如实施例 3 所描述的,用质粒 pMXE117(pTTO CDP870 或 40.4IGS2) 转化 MXE001、MXE008 和 MXE009 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 40.4IGS2) 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多顺反子信息。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 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。

[0450] 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。

[0451] 实施例 5- 高密度发酵中的突变大肠杆菌的生长状况以及突变大肠杆菌中抗 -TNF Fab' 的表达情况

[0452] 在发酵实验中测试实施例 4 中产生的以下菌株，比较其生长状况及抗 -TNFα Fab' 的表达：

[0453] 表达抗 -TNF Fab' 的 MXE001

[0454] 表达抗 -TNF Fab' 的 MXE008

[0455] 表达抗 TNF Fab' 的 MXE009

[0456] 表达抗 TNF Fab' 和 DsbC 的 MXE001

[0457] 表达抗 TNF Fab' 和 DsbC 的 MXE008

[0458] 表达抗 TNF Fab' 和 DsbC 的 MXE009

[0459] 生长培养基

[0460] 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 $NaH_2PO_4.H_2O$ 及 112g/l 的甘油。

[0461] 接种。接种培养物生长于相同的培养基，其中补充了 10μg/ml 的四环素。培养物在 30℃下搅拌孵育约 22 小时。

[0462] 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_{600}。在生长期将温度维持在 30℃，到诱导前降低至 25℃。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_4OH 和 10％ (v/v) 浓度的 H_2SO_4 自动滴定使培养物的 pH 控制在 7.0。通过添加 10％ (v/v) Struktol J673 溶液 (Schill 和 Seilacher) 控制泡沫产生。

[0463] 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 400OD_{600} 时，添加镁盐及 $NaH_2PO_4.H_2O$。在诱导期之前和之内进一步添加 $NaH_2PO_4.H_2O$ 以保证磷酸盐过量。当发酵开始存在的甘油耗尽后（大约 750OD_{600}），连续进料 80％ (w/w) 的甘油。在发酵中所述相同时间点以 170μM 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。

[0464] 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。

[0465] 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℃）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℃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℃）用于进一步分析。

[0466] 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℃，经 0.2um 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，然后注入 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。

[0467] 图 1 显示了表达抗 -TNF Fab' 的 MXE001 和 MXE008 菌株的生长概况以及表达抗 -TNFα Fab' 和重组 DsbC 的菌株 MXE001 和 MXE008 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 的菌株显示有改善的生长。还可看出菌株中 spr 突变

的存在改善了细胞生长。

[0468] 图 2 显示了来自表达抗 -TNF Fab' 和重组 DsbC 的大肠杆菌菌株 MXE001、MXE008 和 MXE009 的细胞周质（阴影标记）及上清（空心非阴影标记）的 Fab 产量。可从此图中看出，所有三种菌株都产生了高产量的抗 -TNF α Fab'，其中 MXE008 和 MXE009 菌株在 40 小时内产生了超过 2.2g/L 的细胞周质抗 -TNF α Fab'。还可看出相比于 MXE001，携带突变 spr 基因的 MXE008 和 MXE009 菌株显示了降低的细胞裂解，这一点可由上清的抗 -TNF α Fab' 量（空心标记）较少看出。

[0469] 图 3 显示了来自表达抗 -TNF α Fab' 的大肠杆菌菌株 MXE001 和 MXE008 以及表达抗 -TNF α Fab' 和重组 DsbC 的大肠杆菌菌株 MXE001 和 MXE008 的细胞周质（阴影标记）及上清（空心非阴影标记）的 Fab 产量。可从此图中看出，表达重组 DsbC 的菌株都产生了高产量的抗 -TNF α Fab'，其中 MXE008 菌株在 40 小时内产生了 2.4g/L 的细胞周质抗 -TNF α Fab'。还可看出相比于 MXE001 菌株，携带突变 spr 基因的 MXE008 菌株显示了降低的细胞裂解，这一点可由上清的抗 -TNF α Fab' 量（空心标记）较少看出。

[0470] 实施例 6 - 高密度发酵中突变的大肠杆菌菌株的生长状况以及突变的大肠杆菌菌株中 Fab A 和 Fab B 的表达情况

[0471] 对本发明细胞中蛋白质生产产量的作用还进一步在两种 Fab 蛋白质上进行——Fab A 具有对抗原 A 的特异性而 Fab B 具有对抗原 B 的特异性，使用的方法与实施例 5 中所描述的相同，并将其与表达 FabA 和 Fab B 的 W3110 菌株相比较。

[0472] 图 4 显示了来自表达 Fab A 和 Fab B 的大肠杆菌菌株 W3110 和来自表达 Fab A 及重组 DsbC 或表达 Fab B 及重组 DsbC 的大肠杆菌菌株 MXE008 的细胞周质中总的 Fab 产量。由此图可以看出，相比于 W3110，表达重组 DsbC 的菌株在 MXE008 中产生了显著高的 Fab A 和 Fab B 的产量。

[0473] 实施例 7 - 大规模发酵中表达抗 -TNF Fab' 和 DsbC 的大肠杆菌菌株生长状况

[0474] 在发酵实验中测试如实施例 4 中产生的以下菌株，比较菌株的生长状况及抗 -TNF α Fab' 的表达：

[0475] 表达抗 -TNF Fab' 和 DsbC 的 MXE008

[0476] 发酵按如下方法进行：

[0477] 表达抗 -TNF Fab' 和 DsbC 的 MXE008 细胞起始生长于摇瓶培养系统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 5L 或 250L 的生产发酵罐，使用类似的化学配方培养基，终体积分别为大约 3.3L 或 230L。培养物起始时为分批式模式生长，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℃提取细胞中的 Fab'。

[0478] 发酵物的生长概况通过测量 600nm 处培养物光密度而确定。

[0479] Fab' 的滴度通过蛋白质 G HPLC 而确定，如上文实施例 5 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 5 描述的方法测量上清及细胞周质 Fab'。图 8 显示了细胞周质 Fab' 滴度 .

[0480] 图 7 显示了表达抗 -TNF Fab’和重组 DsbC 的大肠杆菌菌株 MXE008 的 5L 和 200L 发酵物的生长概况比较。

[0481] 图 8 显示了表达抗 -TNF Fab’和重组 DsbC 的大肠杆菌菌株 MXE008 的 5L 和 200L 发酵物中的 Fab’滴度比较。

[0482] 实施例 8- 大规模发酵中表达抗 -TNF Fab’和 DsbC 的大肠杆菌菌株生长状况

[0483] 在发酵实验中测试如实施例 4 中产生的以下菌株，比较菌株的生长状况及抗 -TNFα Fab’的表达 ：

[0484] 表达抗 -TNF Fab’和 DsbC 的 MXE008

[0485] 表达抗 -TNF Fab’和 DsbC 的 MXE009

[0486] 发酵如上文实施例 7 中描述的方法进行，使用的是 20L 的生产发酵罐。

[0487] 通过测量 600nm 处培养物的光学密度确定发酵物的生长概况。

[0488] 通过如上文实施例 7 中描述的蛋白质 G HPLC 确定 Fab’滴度。

[0489] 图 9 显示了表达抗 -TNF Fab’和重组 DsbC 的大肠杆菌菌株 MXE008 和 MXE009 的发酵物的生长概况比较。

[0490] 图 10 显示了表达抗 -TNF Fab’和重组 DsbC 的大肠杆菌菌株 MXE008 和 MXE009 的发酵物中 Fab’滴度的比较。

[0491] 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。

[0492] 优选的实施方式 ：

[0493] 1. 重组革兰氏阴性细菌细胞，特征在于所述细胞 ：

[0494] a. 包含编码 DsbC 的重组多核苷酸 ；并且

[0495] b. 与野生型细胞相比，具有降低的 Tsp 蛋白质活性。

[0496] 2. 根据项目 1 的细胞，其中所述细胞进一步包含编码突变 spr 蛋白质的突变 spr 基因。

[0497] 3. 根据项目 2 的细胞，其中所述突变 spr 基因编码的 spr 蛋白质在选自 H157、N31、R62、I70、Q73、C94、S95、V98、Q99、R100、L108、Y115、D133、V135、L136、G140、R144、H145 和 G147 的一个或多个氨基酸处具有突变。

[0498] 4. 根据项目 3 的细胞，其中所述突变 spr 基因编码的 spr 蛋白质具有选自 H157A、N31Y、R62C、I70T、Q73R、C94A、S95F、V98E、Q99P、R100G、L108S、Y115F、D133A、V135D、V135G、L136P、G140C、R144C、H145A 和 G147C 的一个或多个突变。

[0499] 5. 根据项目 4 的细胞，其中所述 spr 蛋白质的一个或多个选自 S95F、V98E、Y115F、D133A、V135D、V135G 和 G147C。

[0500] 6. 根据项目 5 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。

[0501] 7. 根据项目 4 的细胞，其中所述突变 spr 基因编码的 spr 蛋白质具有选自 C94A、D133A、H145A 和 H157A 的突变。

[0502] 8. 根据任何前述项目的细胞，其中所述细胞进一步包含一种或多种以下突变基因 ：

[0503] 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；

[0504] 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 或其为敲除突变 ptr 基因；以及

[0505] 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。

[0506] 9. 根据任何前述项目的细胞，其中所述细胞包含的突变 Tsp 基因编码具有降低的蛋白酶活性的 Tsp 蛋白质或为敲除突变 Tsp 基因。

[0507] 10. 根据项目 9 的细胞，其中除了突变 Tsp 基因及任选的突变 spr 基因外，所述细胞的基因组与野生型细菌细胞是同基因的。

[0508] 11. 根据项目 9 或项目 10 的细胞，其中所述细胞具有敲除突变基因，其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和基因终止密码子上游的一个或多个终止密码子。

[0509] 12. 根据项目 11 的细胞，其中所述敲除突变 Tsp 基因包含由对基因起始密码子的错义突变和任选地一个或多个其它点突变产生的限制性标记物位点。

[0510] 13. 根据项目 12 的细胞，其中所述敲除突变 Tsp 基因包含 SEQ ID NO:3。

[0511] 14. 根据任何前述项目的细胞，其中所述细胞为大肠杆菌。

[0512] 15. 根据任何前述项目的细胞，其中所述细胞包含编码目的蛋白质的多核苷酸序列。

[0513] 16. 据项目 15 的细胞，其中所述细胞包含含有编码 DsbC 的重组多核苷酸以及编码目的蛋白质的多核苷酸序列的载体。

[0514] 17. 据项目 16 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。

[0515] 18. 根据项目 15-17 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。

[0516] 19. 根据项目 18 的细胞，其中所述抗体或其抗原结合片段对 TNF 具有特异性。

[0517] 20. 用于产生目的蛋白质的方法，其包括：

[0518] a. 在能够有效表达目的重组蛋白质及编码 DsbC 的重组多核苷酸的条件下，在培养基中培养根据项目 1-19 中任何一项限定的重组革兰氏阴性细菌细胞；以及

[0519] b. 从重组革兰氏阴性细菌细胞的细胞周质和 / 或培养基中回收目的蛋白质。

[0520] 21. 根据项目 20 的方法，其中编码目的蛋白质的多核苷酸序列与编码 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。

[0521] 22. 根据项目 20 或 21 的方法，其中所述方法进一步包括将目的蛋白质与 DsbC 分离。

[0001]

序列表

<110> UCB PHARMA S.A.

<120> 表达重组DSBC并具有降低的TSP活性的细菌宿主菌株

<130> G0109-WO

<150> GB1000591.6
<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 tttttaggct taccgcgtta gctggcctgc ttgcaatagc aggccagacc 60

ttcgctgtag aagatatcac gcgtgctgat caaattccgg tattaaagga agagacgcag 120

catgcgacgg taagtgagcg cgtaacgtcg cgcttcaccc gttctcatta tcgccagttc 180

gacctcgatc aggcatttte ggccaaaatc tttgaccgct acctgaatct gctcgattac 240

agccacaacg tgctgctggc aagcgatgtt gaacagttcg cgaaaaagaa aaccgagtta 300

ggcgatgaac tgcgttcagg caaactcgac gttttctacg atctctacaa tctggcgcaa 360

aagcgccgtt ttgagcgtta ccagtacgct ttgtcggtac tggaaaagcc gatggatttc 420

accggcaacg acacttataa ccttgaccgc agcaaagcgc cctggccgaa aaacgaggct 480

gagttgaacg cgctgtggga cagtaaagtc aaattcgacg agttaagcct gaagctgaca 540

ggaaaaacgg ataaagaaat tcgtgaaacc ctgactcgcc gctacaaatt tgccattcgt 600

cgtctggcgc aaaccaacag cgaagatgtt ttctcgctgg caatgacggc gtttgcgcgt 660

gaaatcgacc cgcataccaa ctatctttcc ccgcgtaata ccgaacagtt caacactgaa 720

atgagtttgt cgctggaagg tattggcgca gtgctgcaaa tggatgatga ctacaccgtt 780

atcaattcga tggtggcagg tggtccggca gcgaagagta aagctatcag cgttggtgac 840

aaaattgtcg gtgttggtca aacaggcaag ccgatggttg acgtgattgg ctggcgtctt 900

gatgatgtgg ttgccttaat taaagggccg aagggcagta aagttcgtct ggaaatttta 960

cctgctggta aagggaccaa gacccgtact gtaacgttga cccgtgaacg tattcgtctc 1020

gaagaccgcg cggttaaaat gtcggtgaag accgtcggta aagagaaagt cggcgtgctg 1080

gatattccgg gcttctatgt gggtttgaca gacgatgtca aagtgcaact gcagaaactg 1140

gaaaaacaga atgtcagcag cgtcatcatc gacctgcgta gcaatggcgg tggggcgtta 1200

actgaagccg tatcgctctc cggtctgttt attcctgcgg gtcccattgt tcaggtccgc 1260

gataacaacg gcaaggttcg tgaagatagc gataccgacg gacaggtttt ctataaaggc 1320

ccgctggtgg tgctggttga ccgcttcagt gcttcggctt cagaaatctt tgccgcggca 1380

atgcaggatt acggtcgtgc gctggttgtg ggtgaaccga cgtttggtaa aggcaccgtt 1440

cagcaatacc gttcattgaa ccgtatttac gatcagatgt tacgtcctga atggccagcg 1500

ctgggttctg tgcagtacac gatccagaaa ttctatcgcg ttaacggcgg cagtacgcaa 1560

cgtaaaggcg taacgccaga catcatcatg ccgacgggta atgaagaaac ggaaacgggt 1620

gagaaattcg aagataacgc gctgccgtgg gatagcattg atgccgcgac ttatgtgaaa 1680

[0002]

tcaggagatt taacggcctt tgaaccggag ctgctgaagg aacataatgc gcgtatcgcg 1740

aaagatcctg agttccagaa catcatgaag gatatcgcgc gcttcaacgc tatgaaggac 1800

aagcgcaata tcgtttctct gaattacgct gtgcgtgaga aagagaataa tgaagatgat 1860

gcgacgcgtc tggcgcgttt gaacgaacgc tttaaacgcg aaggtaaacc ggagttgaag 1920

aaactggatg atctaccgaa agattaccag gagccggatc cttatctgga tgagacggtg 1980

aatatcgcac tcgatctggc gaagcttgaa aaagccagac ccgcggaaca acccgctccc 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 ttttaggctt accgcgttag ctggcctgct tgcaatagca ggccagacat 60

taattgtaga agatatcacg cgtgctgatc aaattccggt attaaaggaa gagacgcagc 120

atgcgacggt aagtgagcgc gtaacgtcgc gcttcacccg ttctcattat cgccagttcg 180

acctcgatca ggcattttcg gccaaaatct ttgaccgcta cctgaatctg ctcgattaca 240

gccacaacgt gctgctggca agcgatgttg aacagttcgc gaaaaagaaa accgagttag 300

gcgatgaact gcgttcaggc aaactcgacg ttttctacga tctctacaat ctggcgcaaa 360

agcgccgttt tgagcgttac cagtacgctt tgtcggtact ggaaaagccg atggatttca 420

ccggcaacga cacttataac cttgaccgca gcaaagcgcc ctggccgaaa aacgaggctg 480

agttgaacgc gctgtgggac agtaaagtca aattcgacga gttaagcctg aagctgacag 540

gaaaaacgga taaagaaatt cgtgaaaccc tgactcgccg ctacaaattt gccattcgtc 600

gtctggcgca aaccaacagc gaagatgttt tctcgctggc aatgacggcg tttgcgcgtg 660

aaatcgaccc gcataccaac tatctttccc cgcgtaatac cgaacagttc aacactgaaa 720

tgagtttgtc gctggaaggt attggcgcag tgctgcaaat ggatgatgac tacaccgtta 780

tcaattcgat ggtggcaggt ggtccggcag cgaagagtaa agctatcagc gttggtgaca 840

[0005]

aaattgtcgg tgttggtcaa acaggcaagc cgatggttga cgtgattggc tggcgtcttg 900

atgatgtggt tgccttaatt aaagggccga agggcagtaa agttcgtctg gaaattttac 960

ctgctggtaa agggaccaag acccgtactg taacgttgac ccgtgaacgt attcgtctcg 1020

aagaccgcgc ggttaaaatg tcggtgaaga ccgtcggtaa agagaaagtc ggcgtgctgg 1080

atattccggg cttctatgtg ggtttgacag acgatgtcaa agtgcaactg cagaaactgg 1140

aaaaacagaa tgtcagcagc gtcatcatcg acctgcgtag caatggcggt ggggcgttaa 1200

ctgaagccgt atcgctctcc ggtctgttta ttcctgcggg tcccattgtt caggtccgcg 1260

ataacaacgg caaggttcgt gaagatagcg ataccgacgg acaggttttc tataaaggcc 1320

cgctggtggt gctggttgac cgcttcagtg cttcggcttc agaaatcttt gccgcggcaa 1380

tgcaggatta cggtcgtgcg ctggttgtgg gtgaaccgac gtttggtaaa ggcaccgttc 1440

agcaataccg ttcattgaac cgtatttacg atcagatgtt acgtcctgaa tggccagcgc 1500

tgggttctgt gcagtacacg atccagaaat tctatcgcgt taacggcggc agtacgcaac 1560

gtaaaggcgt aacgccagac atcatcatgc cgacgggtaa tgaagaaacg gaaacgggtg 1620

agaaattcga agataacgcg ctgccgtggg atagcattga tgccgcgact tatgtgaaat 1680

caggagattt aacggccttt gaaccggagc tgctgaagga acataatgcg cgtatcgcga 1740

aagatcctga gttccagaac atcatgaagg atatcgcgcg cttcaacgct atgaaggaca 1800

agcgcaatat cgtttctctg aattacgctg tgcgtgagaa agagaataat gaagatgatg 1860

cgacgcgtct ggcgcgtttg aacgaacgct ttaaacgcga aggtaaaccg gagttgaaga 1920

aactggatga tctaccgaaa gattaccagg agccggatcc ttatctggat gagacggtga 1980

atatcgcact cgatctggcg aagcttgaaa aagccagacc cgcggaacaa cccgctcccg 2040

tcaagtaa 2048

<210> 4
<211> 2889
<212> DNA
<213> 大肠杆菌

<400> 4
atgccccgca gcacctggtt caaagcatta ttgttgttag ttgcctttg ggcaccctta 60

agtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120

aaccgccagt atcaggctat acgtctggat aacggtatgg tggtcttgct ggtttctgat 180

ccgcaggcag ttaaatcgct ctcggcgctg gtggtgcccg ttgggtcgct ggaagatccc 240

gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag 300

tacccgcagg ctgacagtct ggccgaatat ctcaaaatgc acggcggtag tcacaatgcc 360

agcactgcgc cgtatcgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt 420

gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa 480

cgtgagcgta atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgcgc 540

atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt 600

ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat 660

ttccacgaga agtactattc cgccaatttg atgaaggcgg ttatttacag taataaaccg 720

ctgccggagt tggcaaaaat ggcggcggac acctttggtc gcgtgccgaa caaagagagc 780

[0006]

aaaaaaccgg aaatcaccgt gccggtagtc accgacgcgc aaaagggcat tatcattcat 840

tacgtcctg cgctgccgcg taaagtgttg cgcgttgagt ttcgcatcga taacaactca 900

gcgaagttcc gtagtaaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960

ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020

gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa 1080

ggcctggcta atcgcgatca ggttgtggcg gcaattttta gctatctcaa tctgttacgt 1140

gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200

cgttatccgt cgatcacccg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260

cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320

gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380

ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgccgtatca ggtcgataaa 1440

atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500

ccagagctta acccttatat tcctgatgat ttctcgctga ttaagtcaga gaagaaatac 1560

gaccatccag agctgattgt tgatgagtcg aatctgcgcg tggtgtatgc gccaagccgt 1620

tattttgcca gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680

gacagcgccc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagg gctggcgctt 1740

gatcagttaa gcaaccaggc gtcggttggt ggcataagtt tttccaccaa cgctaacaac 1800

ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctgtt ccaggcattg 1860

ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcaggc gaagtcctgg 1920

tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgccc 1980

gcgcagatgc tctcgcaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040

tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100

gagtttatgg ttatcggcaa catgaccgag gcccaggcaa caacgctggc acgcgatgtg 2160

caaaaacagt tgggcgctga tggttcagag tggtgtcgaa acaaagatgt agtggtcgat 2220

aaaaaacaat ccgtcatctt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280

gtatttgtac cgactggcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340

cagatcgtac agccgtggtt ctacaatcag ttgcgtaccg aagaacaatt gggctatgcc 2400

gtgtttgcgt ttccaatgag cgtggggcgt cagtggggca tgggcttcct tttgcaaagc 2460

aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520

gcaaaattgc gagcgatgaa gccagatgag tttgcgcaaa tccagcaggc ggtaattacc 2580

cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640

gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700

acgccgcaaa aacttgctga tttcttccat caggcggtgg tcgagccgca aggcatggct 2760

attctgtcgc agatttccgg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820

tggaaagtgt gggagaacgt cagcgcgttg cagcaaacaa tgcccctgat gagtgaaaag 2880

aatgagtga 2889

<210> 5
<211> 962
<212> PRT

[0007]

<213> 大肠杆菌

<400> 5

```
Met Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
1               5                  10                  15

Trp Ala Pro Leu Ser Gln Ala Glu Thr Gly Trp Gln Pro Ile Gln Glu
            20                  25                  30

Thr Ile Arg Lys Ser Asp Lys Asp Asn Arg Gln Tyr Gln Ala Ile Arg
        35                  40                  45

Leu Asp Asn Gly Met Val Val Leu Leu Val Ser Asp Pro Gln Ala Val
    50                  55                  60

Lys Ser Leu Ser Ala Leu Val Val Pro Val Gly Ser Leu Glu Asp Pro
65                  70                  75                  80

Glu Ala Tyr Gln Gly Leu Ala His Tyr Leu Glu His Met Ser Leu Met
                85                  90                  95

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

Met His Gly Gly Ser His Asn Ala Ser Thr Ala Pro Tyr Arg Thr Ala
        115                 120                 125

Phe Tyr Leu Glu Val Glu Asn Asp Ala Leu Pro Gly Ala Val Asp Arg
    130                 135                 140

Leu Ala Asp Ala Ile Ala Glu Pro Leu Leu Asp Lys Lys Tyr Ala Glu
145                 150                 155                 160

Arg Glu Arg Asn Ala Val Asn Ala Glu Leu Thr Met Ala Arg Thr Arg
                165                 170                 175

Asp Gly Met Arg Met Ala Gln Val Ser Ala Glu Thr Ile Asn Pro Ala
            180                 185                 190

His Pro Gly Ser Lys Phe Ser Gly Gly Asn Leu Glu Thr Leu Ser Asp
        195                 200                 205

Lys Pro Gly Asn Pro Val Gln Gln Ala Leu Lys Asp Phe His Glu Lys
    210                 215                 220

Tyr Tyr Ser Ala Asn Leu Met Lys Ala Val Ile Tyr Ser Asn Lys Pro
225                 230                 235                 240

Leu Pro Glu Leu Ala Lys Met Ala Ala Asp Thr Phe Gly Arg Val Pro
                245                 250                 255

Asn Lys Glu Ser Lys Lys Pro Glu Ile Thr Val Pro Val Val Thr Asp
            260                 265                 270

Ala 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 gcacctggtt caaagcatta ttgttgttag ttgcctttg ggcacattaa 60

tgtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120

aaccgccagt atcaggctat acgtctggat aacggtatgg tggtcttgct ggtttctgat 180

ccgcaggcag ttaaatcgct ctcggcgctg gtggtgcccg ttgggtcgct ggaagatccc 240

gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag 300

tacccgcagg ctgacagtct ggccgaatat ctcaaaatgc acggcggtag tcacaatgcc 360

agcactgcgc cgtatcgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt 420

gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa 480

cgtgagcgta atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgcgc 540

atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt 600

ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat 660

ttccacgaga agtactattc cgccaatttg atgaaggcgg ttatttacag taataaaccg 720

ctgccggagt tggcaaaaat ggcggcggac acctttggtc gcgtgccgaa caaagagagc 780

aaaaaaccgg aaatcaccgt gccggtagtc accgacgcgc aaaagggcat tatcattcat 840

tacgtccctg cgctgccgcg taaagtgttg cgcgttgagt ttcgcatcga taacaactca 900

gcgaagttcc gtagtaaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960

ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020

gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa 1080

ggcctggcta atcgcgatca ggttgtggcg gcaattttta gctatctcaa tctgttacgt 1140

gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200

cgttatccgt cgatcacccg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260

cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320

gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380

ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgccgtatca ggtcgataaa 1440

atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500

[0011]

ccagagctta acccttatat tcctgatgat ttctcgctga ttaagtcaga gaagaaatac 1560

gaccatccag agctgattgt tgatgagtcg aatctgcgcg tggtgtatgc gccaagccgt 1620

tattttgcca gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680

gacagcgccc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagg gctggcgctt 1740

gatcagttaa gcaaccaggc gtcggttggt ggcataagtt tttccaccaa cgctaacaac 1800

ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctgtt ccaggcattg 1860

ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcaggc gaagtcctgg 1920

tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgccc 1980

gcgcagatgc tctcgcaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040

tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100

gagtttatgg ttatcggcaa catgaccgag gcccaggcaa caacgctggc acgcgatgtg 2160

caaaaacagt tgggcgctga tggttcagag tggtgtcgaa acaaagatgt agtggtcgat 2220

aaaaaacaat ccgtcatctt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280

gtatttgtac cgactggcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340

cagatcgtac agccgtggtt ctacaatcag ttgcgtaccg aagaacaatt gggctatgcc 2400

gtgtttgcgt ttccaatgag cgtggggcgt cagtggggca tgggcttcct tttgcaaagc 2460

aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520

gcaaaattgc gagcgatgaa gccagatgag tttgcgcaaa tccagcaggc ggtaattacc 2580

cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640

gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700

acgccgcaaa aacttgctga tttcttccat caggcggtgg tcgagccgca aggcatggct 2760

attctgtcgc agatttccgg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820

tggaaagtgt gggagaacgt cagcgcgttg cagcaaacaa tgcccctgat gagtgaaaag 2880

aatgagtgat gtcgccgaga cactagatcc tttgc 2915

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

<400> 7
atgaaaaaaa ccacattagc actgagtgca ctggctctga gtttaggttt ggcgttatct 60

ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaagc 120

cttgcaccga tgctcgaaaa ggtgatgcct tcagtggtca gcattaacgt agaaggtagc 180

acaaccgtta atacgccgcg tatgccgcgt aatttccagc agttcttcgg tgatgattct 240

ccgttctgcc aggaaggttc tccgttccag agctctccgt tctgccaggg tggccagggc 300

ggtaatggtg gcggccagca acagaaattc atggcgctgg gttccggcgt catcattgat 360

gccgataaag gctatgtcgt caccaacaac cacgttgttg ataacgcgac ggtcattaaa 420

gttcaactga gcgatggccg taagttcgac gcgaagatgg ttggcaaaga tccgcgctct 480

gatatcgcgc tgatccaaat ccagaacccg aaaaacctga ccgcaattaa gatggcggat 540

tctgatgcac tgcgcgtggg tgattacacc gtagcgattg gtaacccgtt tggtctgggc 600

gagacggtaa cttccgggat tgtctctgcg ctggggcgta gcggcctgaa tgccgaaaac 660

[0012]

tacgaaaact tcatccagac cgatgcagcg atcaaccgtg gtaactccgg tggtgcgctg 720

gttaacctga acggcgaact gatcggtatc aacaccgcga tcctcgcacc ggacggcggc 780

aacatcggta tcggttttgc tatcccgagt aacatggtga aaaacctgac ctcgcagatg 840

gtggaatacg gccaggtgaa acgcggtgag ctgggtatta tggggactga gctgaactcc 900

gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960

cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctcactgaac 1020

ggtaagccga tcagcagctt tgccgcactg cgtgctcagg tgggtactat gccggtaggc 1080

agcaaactga ccctgggctt actgcgcgac ggtaagcagg ttaacgtgaa cctggaactg 1140

cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgct 1200

gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcact 1260

ccggctgcgc agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcaggca 1320

gtgaaaaaca tcgctgaact gcgtaaagtt ctcgacagca aaccgtctgt gctggcactc 1380

aacattcagc gcggcgacag 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 actgagtgca ctggctctga gtttaggttt ggcgttatct 60

ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaagc 120

cttgcaccga tgctcgaaaa ggtgatgcct tcagtggtca gcattaacgt agaaggtagc 180

acaaccgtta atacgccgcg tatgccgcgt aatttccagc agttcttcgg tgatgattct 240

ccgttctgcc aggaaggttc tccgttccag agctctccgt tctgccaggg tggccagggc 300

ggtaatggtg gcggccagca acagaaattc atggcgctgg gttccggcgt catcattgat 360

gccgataaag gctatgtcgt caccaacaac cacgttgttg ataacgcgac ggtcattaaa 420

gttcaactga gcgatggccg taagttcgac gcgaagatgg ttggcaaaga tccgcgctct 480

gatatcgcgc tgatccaaat ccagaacccg aaaaacctga ccgcaattaa gatggcggat 540

tctgatgcac tgcgcgtggg tgattacacc gtagcgattg gtaacccgtt tggtctgggc 600

gagacggtaa cttccgggat tgtctctgcg ctggggcgta gcggcctgaa tgccgaaaac 660

tacgaaaact tcatccagac cgatgcagcg attaatcgtg gtaacgccgg tggtgcgctg 720

gttaacctga acggcgaact gatcggtatc aacaccgcga tcctcgcacc ggacggcggc 780

aacatcggta tcggttttgc tatcccgagt aacatggtga aaaacctgac ctcgcagatg 840

gtggaatacg gccaggtgaa acgcggtgag ctgggtatta tggggactga gctgaactcc 900

gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960

cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctcactgaac 1020

ggtaagccga tcagcagctt tgccgcactg cgtgctcagg tgggtactat gccggtaggc 1080

agcaaactga ccctgggctt actgcgcgac ggtaagcagg ttaacgtgaa cctggaactg 1140

cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgct 1200

gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcact 1260

ccggctgcgc agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcaggca 1320

gtgaaaaaca tcgctgaact gcgtaaagtt ctcgacagca aaccgtctgt gctggcactc 1380

aacattcagc 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-gL1

<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
gcatcataat 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
gtgccaggag 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 ttttcgccgg aacg 24

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

<220>
<223> 寡核苷酸引物

<400> 20
cgcatggtac gtgccacgat 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 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OmpT序列

<400> 23
atgcgggcga aacttctggg aatagtcctg acaaccccta ttgcgatcag ctcttttgct 60

tctaccgaga ctttatcgtt tactcctgac aacataaatg cggacattag tcttggaact 120

ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180

caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240

atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300

atggtcgatc aggactggat ggattccagt aaccccggaa cctggacgga tgaaagtaga 360

caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420

ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480

tttacagcca gaggtggttc ctatatctac agttctgagg agggattcag agatgatatc 540

ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac 600

[0022]

attggcttga ctggaagtta tcgttatgaa gattttgaac tcggtggcac atttaaatac 660

agcggctggg tggaatcatc tgataacgct gaagcttatg acccgggaaa aagaatcact 720

tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780

gtcacaccta acgcaaaagt ttatgttgaa ggcgcatgga atcgggttac gaataaaaaa 840

ggtaatactt 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

Asn 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]

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 aacttctggg aatagtcctg acaaccccta ttgcgatcag ctcttttgct 60

tctaccgaga ctttatcgtt tactcctgac aacataaatg cggacattag tcttggaact 120

ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180

caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240

atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300

atggtcgatc aggactggat ggattccagt aaccccggaa cctggacgga tgaaagtaga 360

caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420

ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480

tttacagcca gaggtggttc ctatatctac agttctgagg agggattcag agatgatatc 540

ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac 600

attggcttga ctggaagtta tcgttatgaa gattttgaac tcggtggcac atttaaatac 660

agcggctggg tggaatcatc tgataacgat gaacactatg acccgggaaa aagaatcact 720

tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780

gtcacaccta acgcaaaagt ttatgttgaa ggcgcatgga atcgggttac gaataaaaaa 840

ggtaatactt cactttatga tcacaataat aacacttcag actacagcaa aaatggagca 900

ggtatagaaa actataactt catcactact gctggtctta agtacacatt ttaa 954

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

<220>
<223> 突变DsbC序列

[0024]

<400> 26
atgaagaaag gttttatgtt gtttactttg ttagcggcgt tttcaggctt tgctcaggct 60

gatgacgcgg caattcaaca aacgttagcc aaaatgggca tcaaaagcag cgatattcag 120

cccgcgcctg tagctggcat gaagacagtt ctgactaaca gcggcgtgtt gtacatcacc 180

gatgatggta aacatatcat tcaggggcca atgtatgacg ttagtggcac ggctccggtc 240

aatgtcacca ataagatgct gttaaagcag ttgaatgcgc ttgaaaaaga gatgatcgtt 300

tataaagcgc cgcaggaaaa acacgtcatc accgtgttta ctgatattac ctgtggttac 360

tgccacaaac tgcatgagca aatggcagac tacaacgcgc tggggatcac cgtgcgttat 420

cttgctttcc cgcgccaggg gctggacagc gatgcagaga aagaaatgaa agctatctgg 480

tgtgcgaaag ataaaaacaa agcgtttgat gatgtgatgg caggtaaaag cgtcgcacca 540

gccagttgcg acgtggatat tgccgaccat tacgcacttg gcgtccagct tggcgttagc 600

ggtactccgg cagttgtgct gagcaatggc acacttgttc cgggttacca gccgccgaaa 660

gagatgaaag aatttctcga cgaacaccaa aaaatgacca gcggtaaaca ccatcaccat 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 gctatcgcaa 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 gctttaatag 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

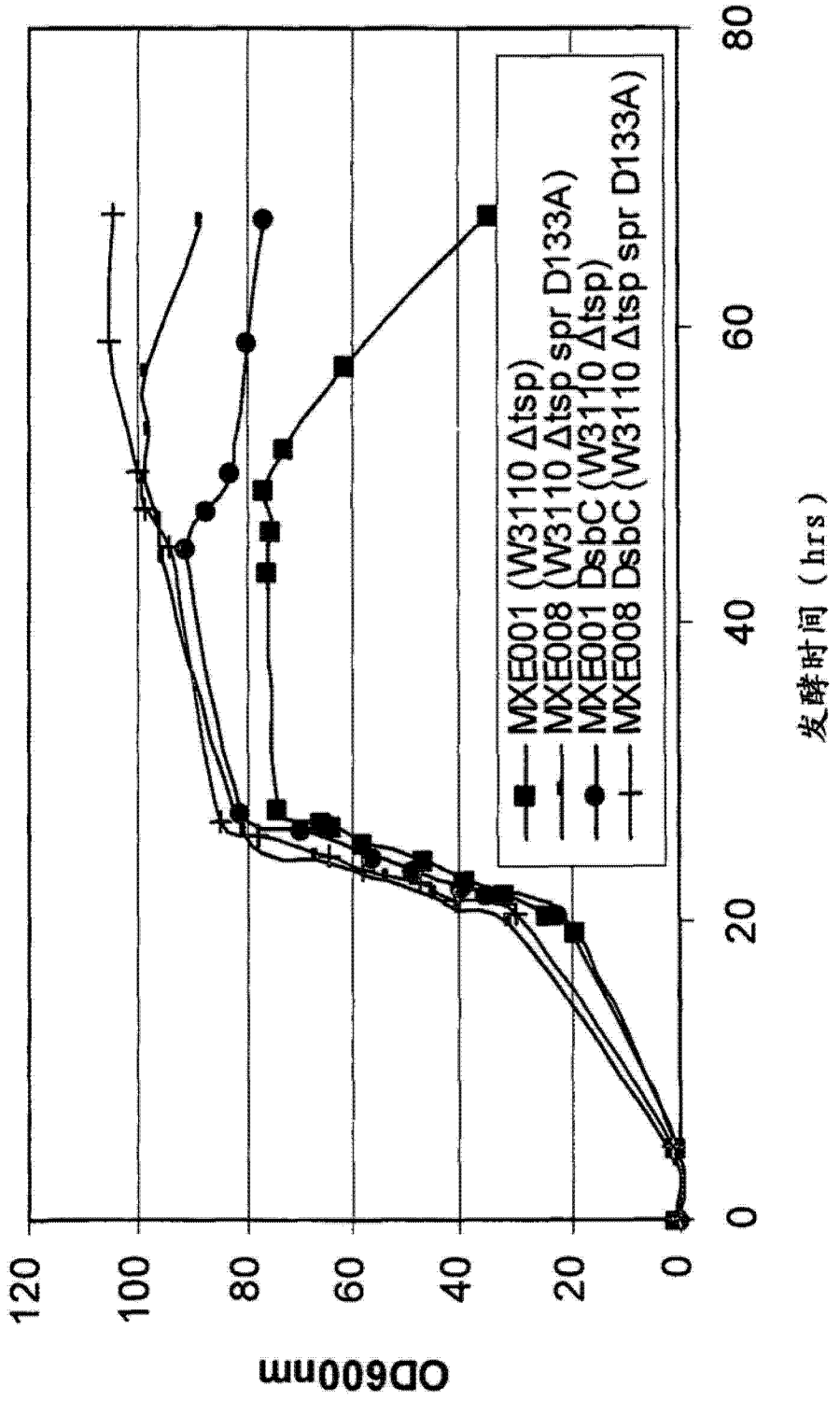

OD600nm
0
20
40
60
80
100
120
发酵时间（hrs）
0
20
40
60
80
MXE001 (W3110 Δtsp)
MXE008 (W3110 Δtsp spr D133A)
MXE001 DsbC (W3110 Δtsp)
MXE008 DsbC (W3110 Δtsp spr D133A)

图 1

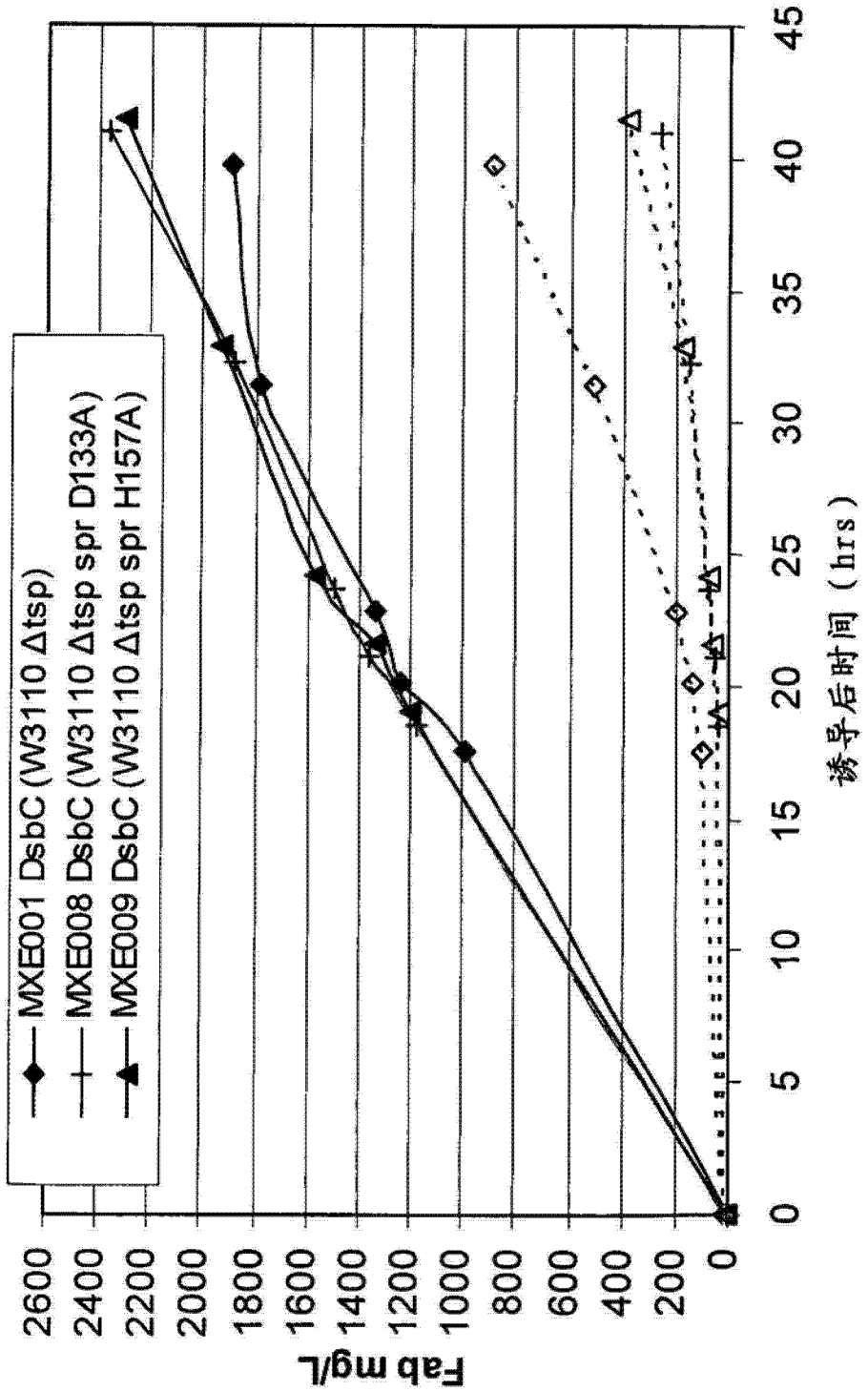

MXE001 DsbC (W3110 Δtsp)
MXE008 DsbC (W3110 Δtsp spr D133A)
MXE009 DsbC (W3110 Δtsp spr H157A)
Fab mg/L
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
5
10
15
20
25
30
35
40
45
诱导后时间（hrs）

图 2

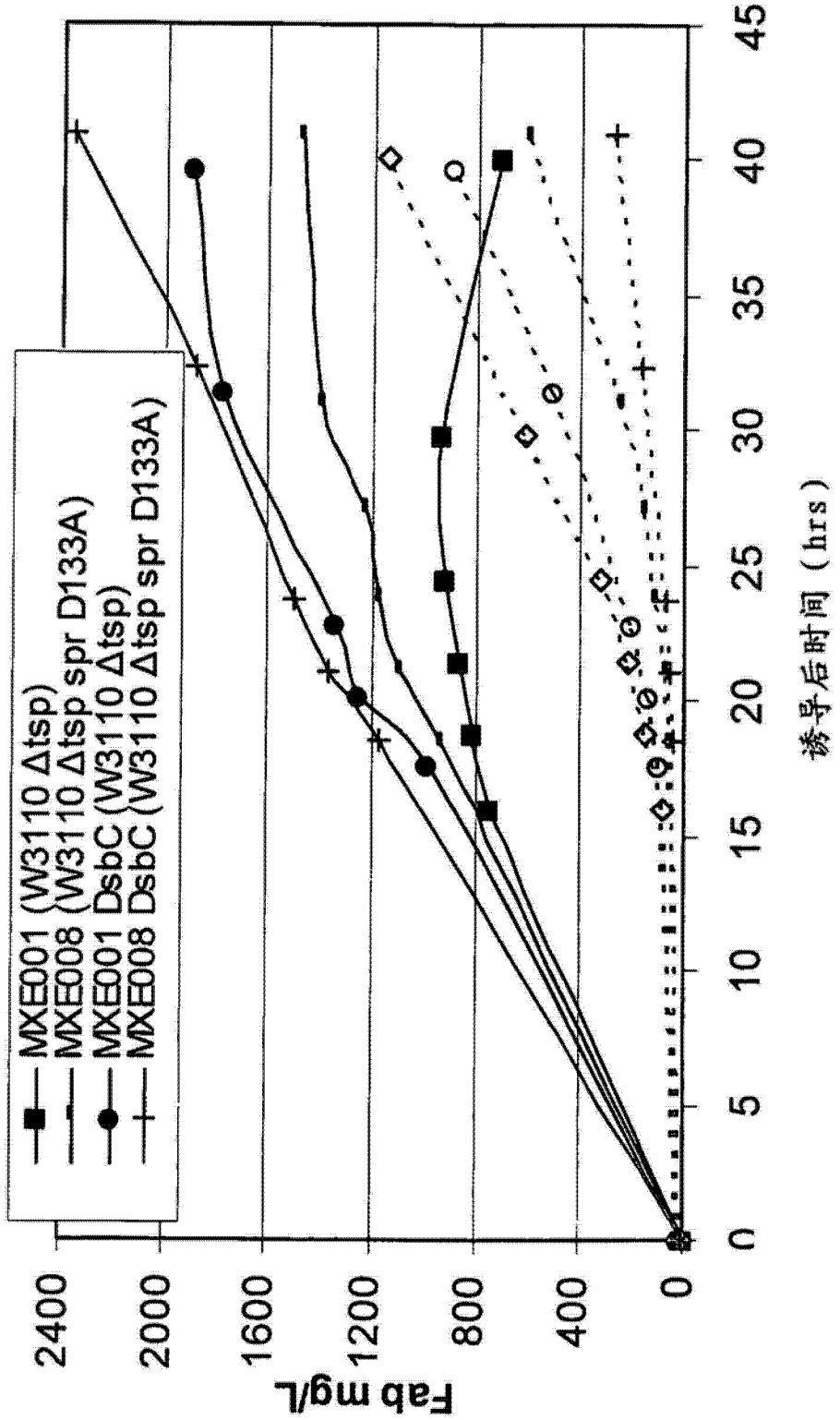

MXE001 (W3110 Δtsp)
MXE008 (W3110 Δtsp spr D133A)
MXE001 DsbC (W3110 Δtsp)
MXE008 DsbC (W3110 Δtsp spr D133A)
Fab mg/L
2400
2000
1600
1200
800
400
0
0
5
10
15
20
25
30
35
40
45
诱导后时间（hrs）

图 3

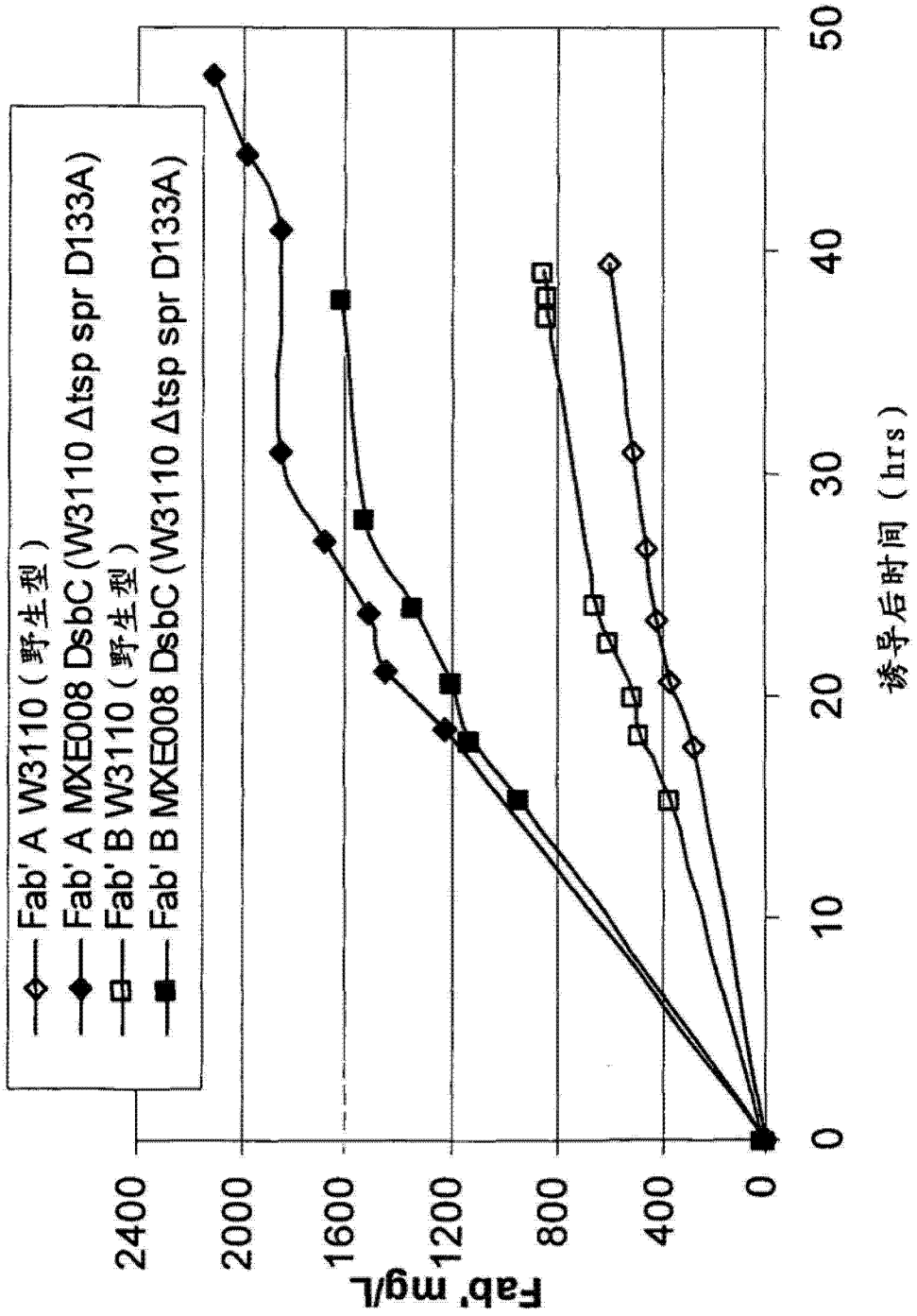

Fab' A W3110 (野生型)
Fab' A MXE008 DsbC (W3110 Δtsp spr D133A)
Fab' B W3110 (野生型)
Fab' B MXE008 DsbC (W3110 Δtsp spr D133A)
Fab' mg/L
2400
2000
1600
1200
800
400
0
0
10
20
30
40
50
诱导后时间 (hrs)

图 4

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 5a

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 5b

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
949   GAT GCA GCG ATT AAT CGT GGT AAC GCC GGT GGT
```

图 5c

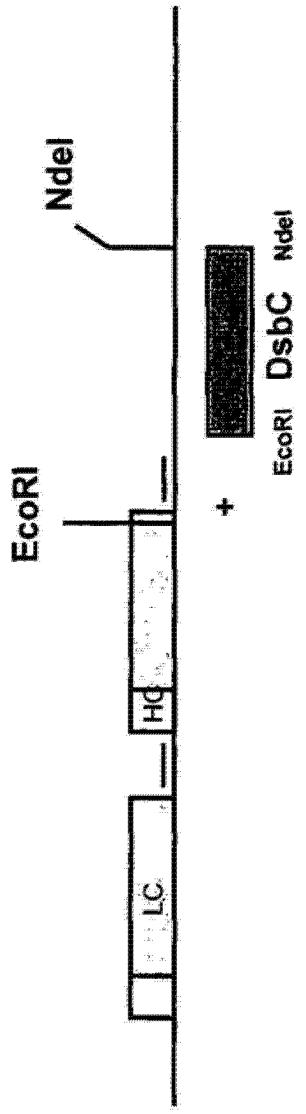


图 6

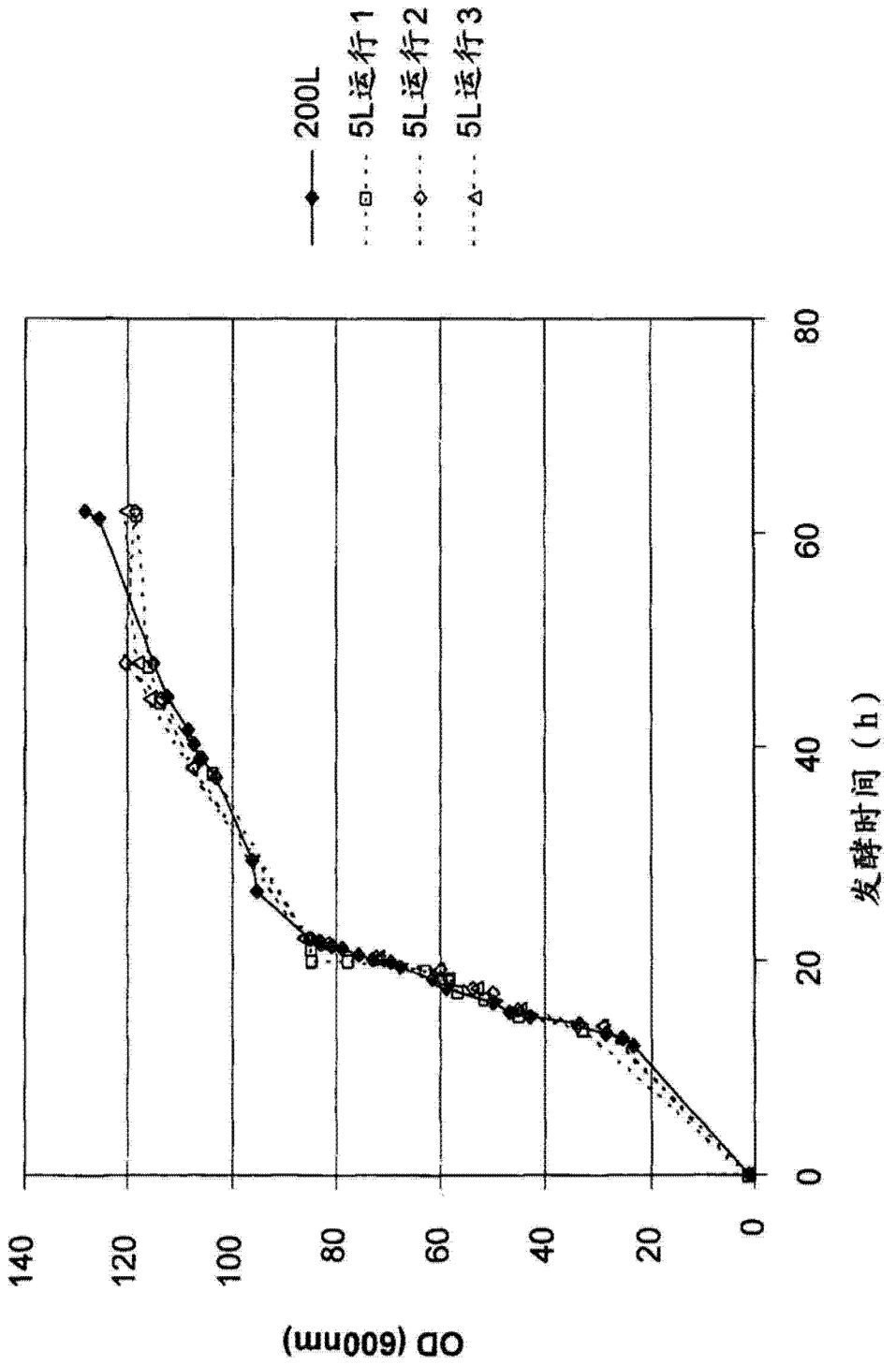

200L
5L运行1
5L运行2
5L运行3
OD (600nm)
140
120
100
80
60
40
20
0
0
20
40
60
80
发酵时间（h）

图 7

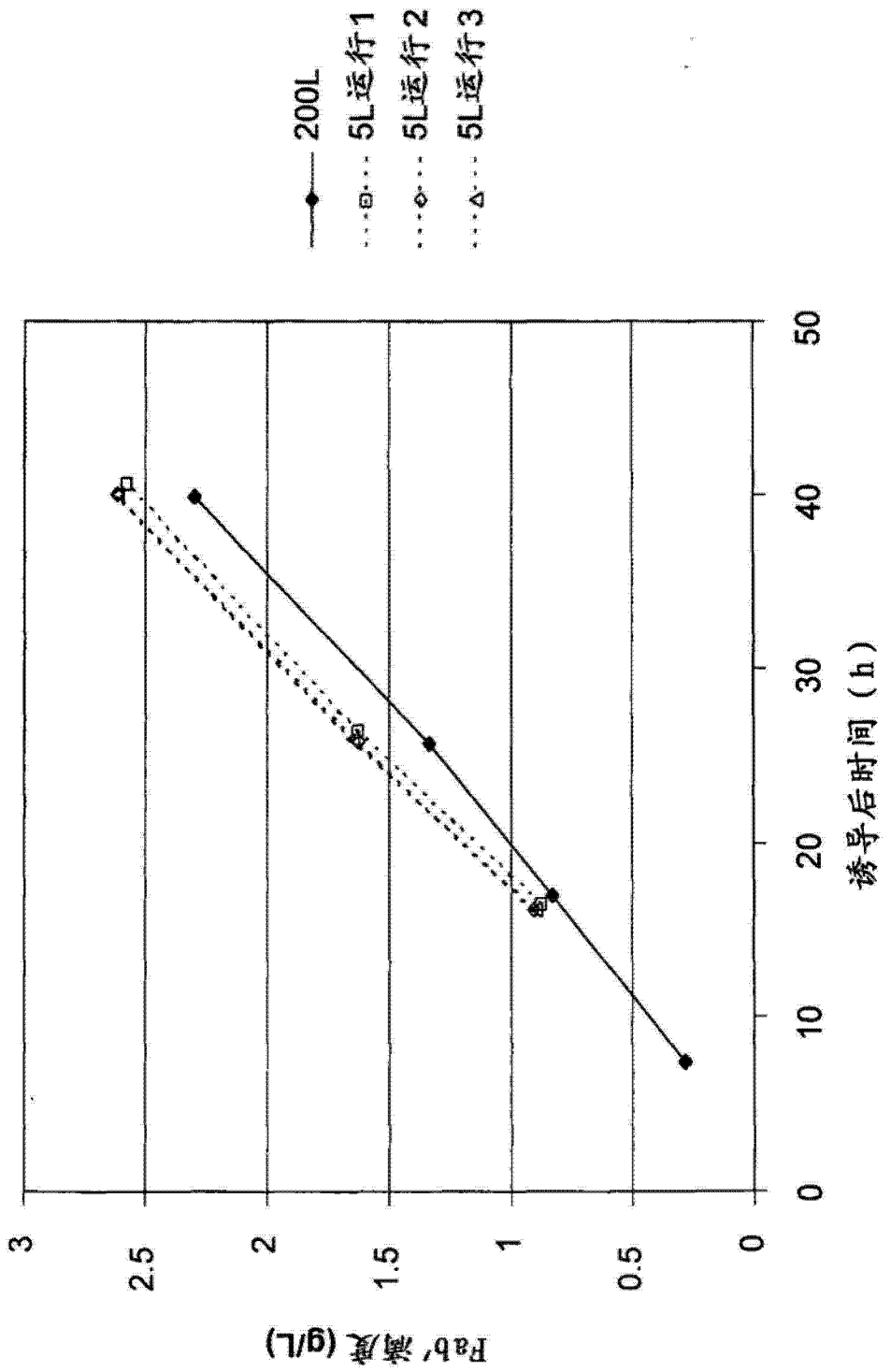

200L
5L运行1
5L运行2
5L运行3
Fab'滴度 (g/L)
诱导后时间 (h)
0
0.5
1
1.5
2
2.5
3
0
10
20
30
40
50

图 8

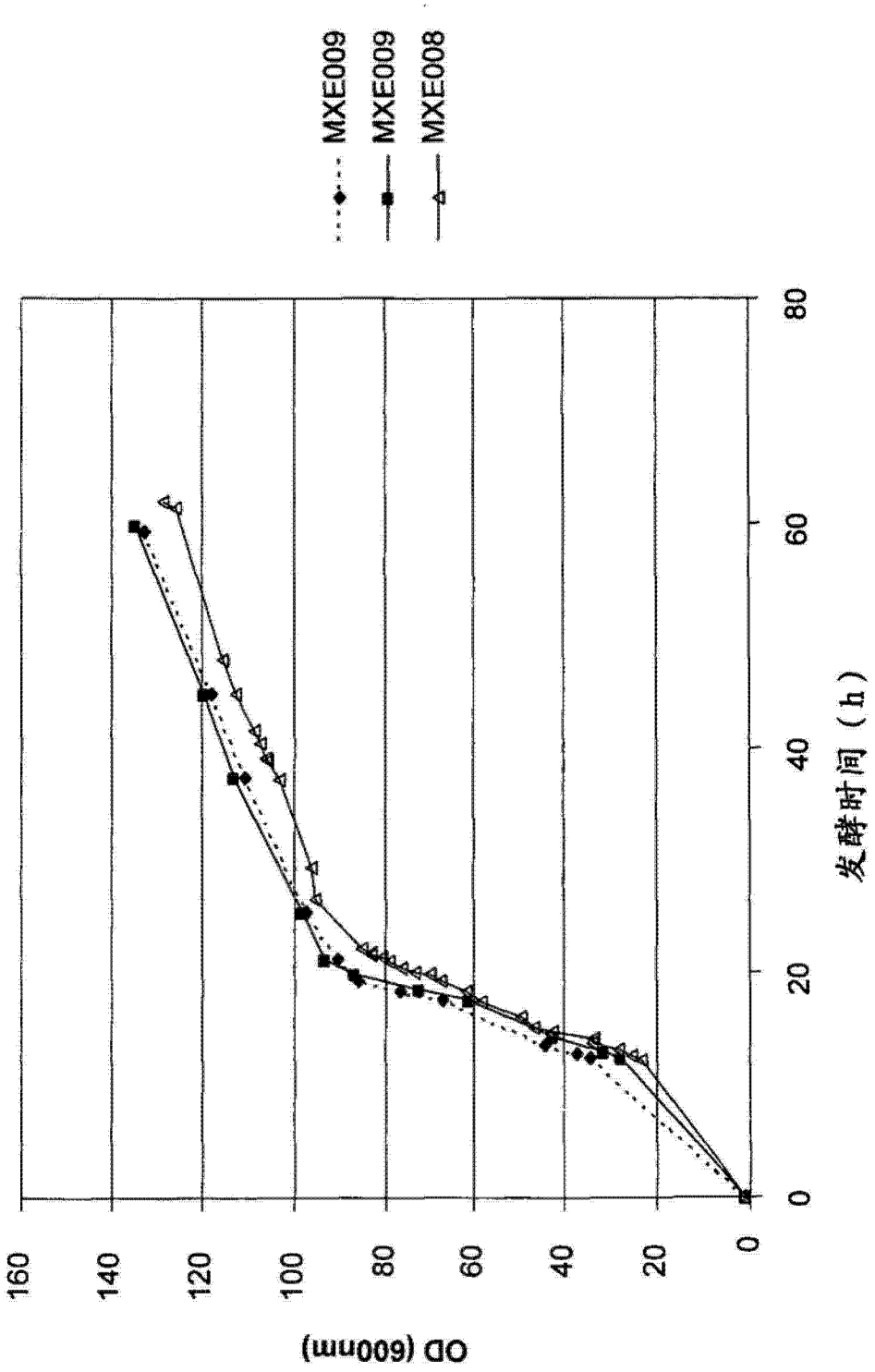

MXE009
MXE009
MXE008
OD (600nm)
0
20
40
60
80
100
120
140
160
发酵时间（h）
0
20
40
60
80

图 9

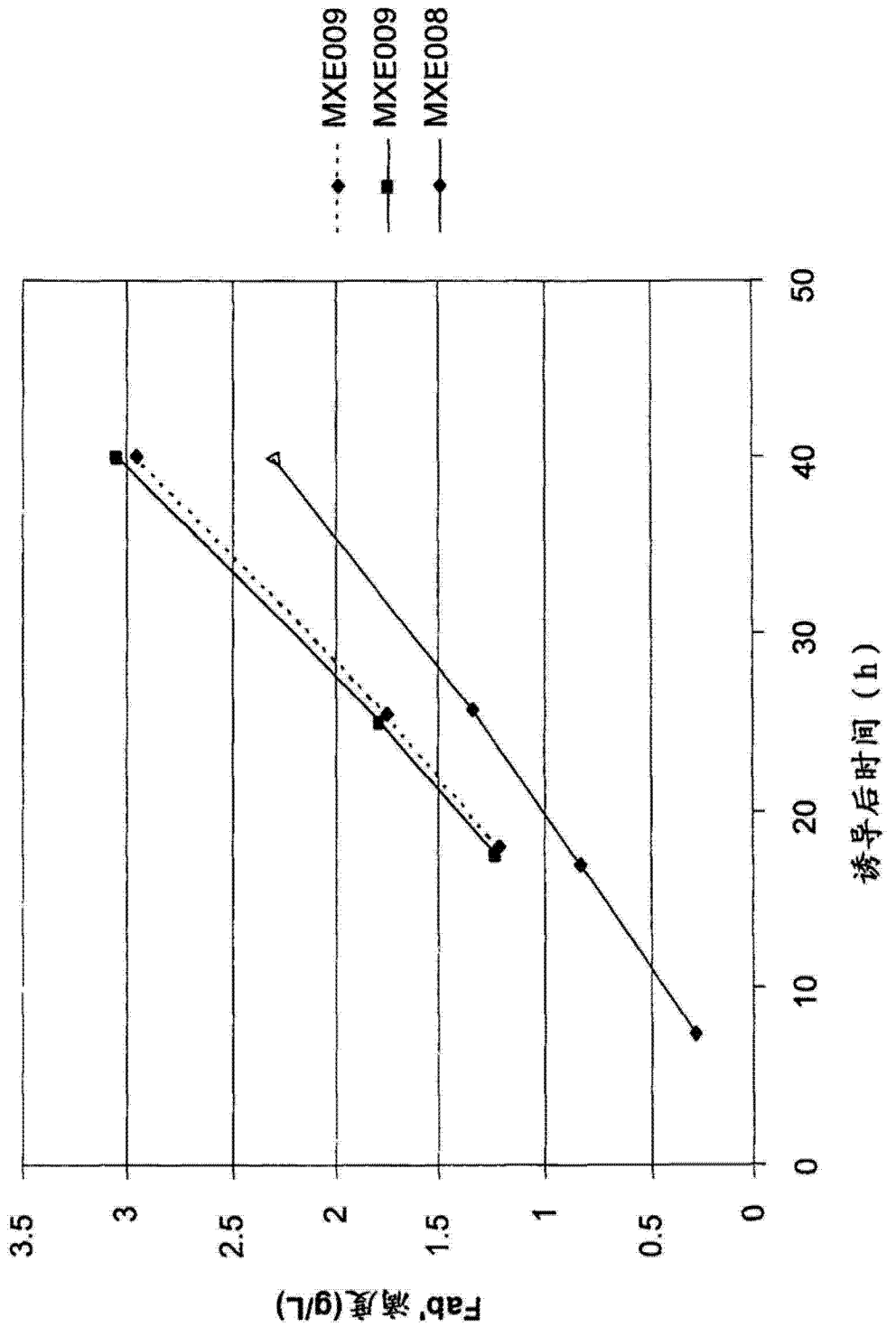

MXE009
MXE009
MXE008
Fab' 滴度(g/L)
3.5
3
2.5
2
1.5
1
0.5
0
0
10
20
30
40
50
诱导后时间（h）

图 10