



(12) 发明专利申请

(10) 申请公布号 CN 113166738 A

(43) 申请公布日 2021.07.23

(21) 申请号 201980069492.0

(22) 申请日 2019.10.24

(30) 优先权数据

62/749,855 2018.10.24 US

(85) PCT国际申请进入国家阶段日

2021.04.21

(86) PCT国际申请的申请数据

PCT/JP2019/041751 2019.10.24

(87) PCT国际申请的公布数据

W02020/085441 JA 2020.04.30

(71) 申请人 摩大力斯医疗株式会社

地址 日本东京都

(72) 发明人 秦园博

(74) 专利代理机构 中原信达知识产权代理有限  
责任公司 11219

代理人 鲁雯雯 金龙河

(51) Int.Cl.

C12N 9/16 (2006.01)

C12N 15/09 (2006.01)

C12N 15/55 (2006.01)

权利要求书2页 说明书13页

序列表29页 附图2页

(54) 发明名称

经改造的Cas9蛋白及其用途

(57) 摘要

本发明涉及一种蛋白质,其由下述序列构成且具有与向导RNA结合的能力,所述序列包含如下氨基酸序列:在序列号2所示的氨基酸序列中,具有在第721~755位之间连续的缺失区域,与该缺失区域分别相邻的氨基酸通过由3~10个氨基酸残基构成的接头连接。该蛋白质尽管存在缺失区域、与全长的dSaCas9相比缩小,但是与向导RNA结合的能力以及DNA结合亲和性得到了维持。小型化的dSaCas9蛋白的使用使更多的基因向载体中的搭载成为可能。

1. 一种蛋白质,其由下述序列构成且具有与向导RNA结合的能力,所述序列包含如下氨基酸序列:

在序列号2所示的氨基酸序列中,具有在第721~755位之间连续的缺失区域,与该缺失区域分别相邻的氨基酸通过由3~10个氨基酸残基构成的接头连接。

2. 根据权利要求1所述的蛋白质,其中,接头是由甘氨酸(G)及丝氨酸(S)构成的长度为5~9个氨基酸的接头。

3. 根据权利要求2所述的蛋白质,其中,接头从以下接头中选择:

-SGGGS-

-GSGGGS-

-SGSGSGSG-

-SGSGSGSGS-

4. 根据权利要求1~3中任一项所述的蛋白质,其中,该缺失区域为第721~745位的区域。

5. 根据权利要求4所述的蛋白质,其由序列号4表示。

6. 根据权利要求1~3中任一项所述的蛋白质,其中,该缺失区域为第721~755位的区域。

7. 根据权利要求6所述的蛋白质,其由序列号6表示。

8. 根据权利要求1~7中任一项所述的蛋白质,其中,进一步地,第45位和/或第163位的谷氨酸(E)被置换为其它氨基酸。

9. 根据权利要求8所述的蛋白质,其中,其它氨基酸为碱性氨基酸。

10. 根据权利要求9所述的蛋白质,其中,碱性氨基酸为赖氨酸(K)。

11. 根据权利要求1~10中任一项所述的蛋白质,其中,在序列号2的除实施了变异和/或缺失的位置以外的部位具有80%以上的一致性。

12. 根据权利要求1~10中任一项所述的蛋白质,其中,在序列号2的除实施了变异和/或缺失的位置以外的部位置换、缺失、插入和/或添加了1个~几个氨基酸。

13. 根据权利要求1~12中任一项所述的蛋白质,其连接有转录调控因子蛋白或结构域。

14. 根据权利要求13所述的蛋白质,其中,转录调控因子为转录活化因子。

15. 根据权利要求13所述的蛋白质,其中,转录调控因子为转录沉默子或转录抑制因子。

16. 一种核酸,其编码权利要求1~15中任一项所述的蛋白质。

17. 一种蛋白质-RNA复合物,其具备权利要求1~16中任一项所述的蛋白质和包含下述多核苷酸的向导RNA,所述多核苷酸由与靶双链多核苷酸中的PAM(前间隔序列邻近基序)序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

18. 一种用于对靶双链多核苷酸进行位点特异性改造的方法,

其具备:

将靶双链多核苷酸、蛋白质和向导RNA混合并进行温育的步骤;以及

所述蛋白质在位于PAM序列的上游的结合位点处对所述靶双链多核苷酸进行改造的步骤,

所述蛋白质为权利要求1~15中任一项所述的蛋白质，

所述向导RNA包含下述多核苷酸，所述多核苷酸由与所述靶双链多核苷酸中的所述PAM序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

19. 一种使细胞的靶基因的表达增加的方法，其包括：使权利要求14所述的蛋白质和针对所述靶基因的一种或两种以上向导RNA在所述细胞内表达。

20. 一种使细胞的靶基因的表达减少的方法，其包括：使权利要求15所述的蛋白质和针对所述靶基因的一种或两种以上向导RNA在所述细胞内表达。

21. 根据权利要求19或20所述的方法，其中，细胞为真核细胞。

22. 根据权利要求19或20所述的方法，其中，细胞为酵母细胞、植物细胞或动物细胞。

## 经改造的Cas9蛋白及其用途

### 技术领域

[0001] 本发明涉及维持了与向导RNA结合的能力且小型化的经改造的Cas9蛋白及其用途。

### 背景技术

[0002] 已知成簇的规律性间隔的短回文重复序列(Clustered Regularly Interspaced Short Palindromic Repeats:CRISPR)与Cas(CRISPR相关)基因一起构成细菌及古细菌中提供对入侵外源核酸的获得性抗性的适应性免疫系统。CRISPR大多起源于噬菌体或质粒DNA,由在其间插入了尺寸相似的被成为间隔子的独特的可变DNA序列的、24~48bp的短的保守重复序列构成。另外,在重复序列及间隔子序列附近存在编码Cas蛋白家族的基因群。

[0003] CRISPR/Cas系统中,外源性DNA被Cas蛋白家族切成30bp左右的片段并被插入到CRISPR中。作为Cas蛋白家族之一的Cas1及Cas2蛋白识别外源性DNA的被成为前间隔序列邻近基序(proto-spacer adjacent motif,PAM)的碱基序列,切取其上游并插入到宿主的CRISPR序列中,其形成细菌的免疫记忆。转录包含免疫记忆的CRISPR序列而生成的RNA(称为pre-crRNA。)与部分互补的RNA(trans-activating crRNA:tracrRNA)配对,整合到作为Cas蛋白家族之一的Cas9蛋白中。整合到Cas9中的pre-crRNA及tracrRNA被RNaseIII切断,成为包含外源序列(向导序列)的小的RNA片段(CRISPR-RNAs:crRNAs),形成Cas9-crRNA-tracrRNA复合物。Cas9-crRNA-tracrRNA复合物结合到与crRNA互补的外源入侵DNA上,作为切断DNA的酶(nuclease)的Cas9蛋白将外源入侵DNA切断,由此抑制和消除从外部入侵的DNA的功能。

[0004] 近年来,正在积极开发将CRISPR/Cas系统应用于基因组编辑的技术。使crRNA和tracrRNA融合而以tracrRNA-crRNA嵌合体(以下称为向导RNA(guide RNA:gRNA。))形式表达,并有效利用。由此引进核酸酶(RNA-guided nuclease:RGN)并在目标位点切断基因组DNA。

[0005] 另一方面,通过在作为基因组编辑系统之一的CRISPR/Cas9中的Cas9蛋白的核酸酶已失活的变异体(nuclease-null,dCas9)中融合VP64、VP160等转录活化因子或KRAB等转录抑制因子等转录调控因子,从而能够形成能调节靶基因的表达水平的系统。例如,为了提高基因活化的效率而与将2个转录活性因子连接而成的活化因子(VP64-Rta)、将3个转录活性因子连接而成的活化因子(VP64-p65-Rta,VPR)融合,该融合而得的dCas9蛋白(dCas9-VPR;dCas9融合蛋白)在不切断DNA的情况下强烈活化靶基因的表达。

[0006] 对于Cas9蛋白,以PAM特异性的缓和、核酸酶活性的改造(活化/失活)、小型化为目的,已研制并报道了各种变异体(专利文献1~3)。

[0007] 现有技术文献

[0008] 专利文献

[0009] 专利文献1:W02016/141224A1

[0010] 专利文献2:W02017/010543A1

[0011] 专利文献3:W02018/074979A1

## 发明内容

[0012] 发明所要解决的问题

[0013] 在体内的dCas9融合蛋白的表达中,需要表达载体。基因治疗中,从安全性和效率高角度来看,腺相关病毒载体(AAV)逐渐成为主流,但是AAV的可搭载的尺寸为4.4kb左右,而dCas9蛋白已占据4kb左右,搭载于AAV中时,融合蛋白的构成极其受限。

[0014] 因此,本发明人的目的在于,提供具有与全长蛋白质实质上同等的DNA结合亲和性且进一步小型化的dCas9蛋白的变异体。

[0015] 用于解决问题的方法

[0016] 本发明人着眼于作为Cas9蛋白的源自金黄色葡萄球菌(*S.aureus*)的Cas9(本说明书中也称为SaCas9)的无核酸酶(nuclease-null)变异体(dSaCas9),为了解决上述问题进行了深入研究。结果发现了即使发生缺失也较少影响与向导RNA结合的能力的特定区域,进而通过将规定位置的氨基酸置换为特定氨基酸,成功地制造出DNA结合亲和性得到了维持、甚至增强且小型化的dSaCas9蛋白,从而完成了本发明。

[0017] 将缺失及置换一并称为变异。

[0018] 本说明书中,有时将导入变异前的dSaCas9蛋白称为野生型dSaCas9(蛋白)、将导入变异后的dSaCas9蛋白称为dSaCas9变异体(蛋白)。

[0019] 即,本发明如下所述。

[0020] [1]一种蛋白质,其由下述序列构成且具有与向导RNA结合的能力,所述序列包含如下氨基酸序列:

[0021] 在序列号2所示的氨基酸序列中,具有在第721~755位之间连续的缺失区域,

[0022] 与该缺失区域分别相邻的氨基酸通过由3~10个氨基酸残基构成的接头连接。

[0023] [2]根据上述[1]所述的蛋白质,其中,接头是由甘氨酸(G)及丝氨酸(S)构成的、长度为5~9个氨基酸的接头。

[0024] [3]根据上述[2]所述的蛋白质,其中,接头从以下接头中选择:

[0025] -SGGGS-

[0026] -GGSGGS-

[0027] -SGSGSGSG-

[0028] -SGSGSGSGS-

[0029] [4]根据上述[1]~[3]中任一项所述的蛋白质,其中,该缺失区域为第721~745位的区域。

[0030] [5]根据上述[4]所述的蛋白质,其由序列号4表示。

[0031] [6]根据上述[1]~[3]中任一项所述的蛋白质,其中,该缺失区域为第721~755位的区域。

[0032] [7]根据上述[6]所述的蛋白质,其由序列号6表示。

[0033] [8]根据上述[1]~[7]中任一项所述的蛋白质,其中,进一步地,第45位和/或第163位的谷氨酸(E)被置换为其它氨基酸。

[0034] [9]根据上述[8]所述的蛋白质,其中,其它氨基酸为碱性氨基酸。

- [0035] [10]根据上述[9]所述的蛋白质,其中,碱性氨基酸为赖氨酸(K)。
- [0036] [11]根据上述[1]~[10]中任一项所述的蛋白质,其中,在序列号2的除实施了变异和/或缺失的位置以外的部位具有80%以上的一致性。
- [0037] [12]根据上述[1]~[10]中任一项所述的蛋白质,其中,在序列号2的除实施了变异和/或缺失的位置以外的部位置换、缺失、插入和/或添加了1个~几个氨基酸。
- [0038] [13]根据上述[1]~[12]中任一项所述的蛋白质,其连接有转录调控因子蛋白或结构域。
- [0039] [14]根据上述[13]所述的蛋白质,其中,转录调控因子为转录活化因子。
- [0040] [15]根据上述[13]所述的蛋白质,其中,转录调控因子为转录沉默子或转录抑制因子。
- [0041] [16]一种核酸,其编码上述[1]~[15]中任一项所述的蛋白质。
- [0042] [17]一种蛋白质-RNA复合物,其具备上述[1]~[16]中任一项所述的蛋白质和包含下述多核苷酸的向导RNA,所述多核苷酸由与靶双链多核苷酸中的PAM(Proto-spacer Adjacent Motif,前间隔序列邻近基序)序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。
- [0043] [18]一种用于对靶双链多核苷酸进行位点特异性改造的方法,
- [0044] 其具备:
- [0045] 将靶双链多核苷酸、蛋白质和向导RNA混合并进行温育的步骤;以及
- [0046] 上述蛋白质在位于PAM序列的上游的结合位点处对上述靶双链多核苷酸进行改造的步骤,
- [0047] 上述蛋白质为上述[1]~[15]中任一项所述的蛋白质,
- [0048] 上述向导RNA包含下述多核苷酸,所述多核苷酸由与上述靶双链多核苷酸中的上述PAM序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。
- [0049] [19]一种使细胞的靶基因的表达增加的方法,其包括:使上述[14]所述的蛋白质和针对上述靶基因的一种或两种以上向导RNA在上述细胞内表达。
- [0050] [20]一种使细胞的靶基因的表达减少的方法,其包括:使上述[15]所述的蛋白质和针对上述靶基因的一种或两种以上向导RNA在上述细胞内表达。
- [0051] [21]根据上述[19]或[20]所述的方法,其中,细胞为真核细胞。
- [0052] [22]根据上述[19]或[20]所述的方法,其中,细胞为酵母细胞、植物细胞或动物细胞。
- [0053] 发明效果
- [0054] 根据本发明,能够得到具有向导RNA结合能力、并且进一步小型化的dSaCas9蛋白。利用该小型化的dSaCas9蛋白,能够使更多的基因搭载到容量方面有制限的表达载体中。

## 附图说明

- [0055] 图1为示意性示出野生型dSaCas9(FL)及dSaCas9变异体(T10~T11)的结构图。T10中的缺失区域(第721~745位)及T11中的缺失区域(721~755位)的两端通过接头肽(SGGGS)连接。

[0056] 图2为示出利用使用A2(序列号12)作为crRNA的向导RNA、和在野生型dSaCas9(FL)或dSaCas9变异体(T10、T11)上连接有KRAB基因的分子来抑制MyD88基因表达的结果的图。

[0057] 图3为示出利用使用A2(序列号12)或A3(序列号13)作为crRNA的向导RNA、和在野生型dSaCas9(FL)或dSaCas9变异体(T10)上连接有KRAB基因的分子来抑制MyD88基因表达的结果的图。

### 具体实施方式

[0058] 以下,对本发明进行说明。只要没有特别声明,则本说明书中使用的术语具有该领域中通常使用的含义。

[0059] <dSaCas9变异体>

[0060] 本发明的dSaCas9变异体为具有与向导RNA结合的能力且进一步小型化的dSaCas9蛋白。若使用小型化的dSaCas9变异体,则能够使更多的基因搭载到载体中。

[0061] 本说明书中,“向导RNA”是指模仿tracrRNA-crRNA的发夹结构的RNA,在其5'末端区域包含由与靶双链多核苷酸中的PAM序列的上游1个碱基至优选20个碱基以上且24个碱基以下、更优选22个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成的多核苷酸。此外,也可以包含1个以上的下述多核苷酸,所述多核苷酸由包含与靶双链多核苷酸不互补的碱基序列、按照以一点为轴对称地形成互补的序列的方式排列、并能够形成发夹结构的碱基序列构成。

[0062] 向导RNA具有与本发明的dSaCas9变异体结合、并且将该蛋白质引导至靶DNA的功能。向导RNA在其5'末端具有与靶DNA互补的序列,介由该互补的序列与靶DNA结合而将本发明的dSaCas9变异体引导至靶DNA。dSaCas9变异体不具有DNA内切核酸酶,因此与靶DNA进行结合但不会将其切断。

[0063] 向导RNA可以基于靶DNA的序列信息来设计、制备。具体而言,可列举实施例中所使用的序列。

[0064] 本说明书中,“多肽”、“肽”及“蛋白质”是指氨基酸残基的聚合物,可互换使用。另外,是指一个或两个以上氨基酸为天然存在的对应氨基酸的化学类似物或修饰衍生物的氨基酸聚合物。

[0065] 本说明书中,“碱性氨基酸”是指赖氨酸、精氨酸、组氨酸等在分子内除具有一个氨基以外还具有显示碱性的残基的氨基酸。

[0066] 本说明书中,“序列”是指任意长度的核苷酸序列,为脱氧核糖核苷酸或核糖核苷酸,为线状、环状或支链状,为单链或双链。

[0067] 本说明书中,“PAM序列”是指存在于靶双链多核苷酸中、能够被Cas9蛋白识别的序列,PAM序列的长度、碱基序列根据细菌种类而不同。

[0068] 需要说明的是,本说明书中,“N”是指选自腺嘌呤、胞嘧啶、胸腺嘧啶及鸟嘌呤组成的组中的任意的一种碱基,“A”是指腺嘌呤,“G”是指鸟嘌呤,“C”是指胞嘧啶,“T”是指胸腺嘧啶,“R”是指具有嘌呤骨架的碱基(腺嘌呤或鸟嘌呤),“Y”是指具有嘧啶骨架的碱基(胞嘧啶或胸腺嘧啶)。

[0069] 本说明书中,“多核苷酸”是指为线状或环状构象、为单链或双链中的任一形态的脱氧核糖核苷酸或核糖核苷酸聚合物,对于聚合物的长度,不进行限定性解释。另外,包括

天然核苷酸的公知的类似物以及碱基部分、糖部分及磷酸部分中的至少一个部分经修饰的核苷酸(例如硫代磷酸酯骨架)。通常,特定核苷酸的类似物具有相同的碱基配对特异性,例如A的类似物与T进行碱基配对。

[0070] 本发明提供一种蛋白质(方式1),其由下述序列构成且具有与向导RNA结合的能力,所述序列包含如下氨基酸序列:在序列号2所示的氨基酸序列中,具有在第721~755位之间连续的缺失区域,与该缺失区域分别相邻的氨基酸通过由3~10个氨基酸残基构成的接头连接。

[0071] 序列号2为dSaCas9蛋白的全长氨基酸序列。dSaCas9蛋白由REC叶(第41~425位的残基)及NUC叶(第1~40位的残基及第435~1053位的残基)这2个叶构成。2个叶通过富含精氨酸的桥状螺旋(BH:第41~73位的残基)和接头环(第426~434位的残基)而连接。NUC叶由RuvC结构域(第1~40位、第435~480位及第650~774位的残基)、HNH结构域(第520~628位的残基)、WED结构域(第788~909位的残基)及PI结构域(第910~1053位的残基)构成。PI结构域分为拓扑异构酶同源(TOP0)结构域和C末端结构域(CTD)。RuvC结构域由3个被隔开的基序(RuvC-I~III)构成,与HNH结构域及PI结构域相关联。HNH结构域介由L1(第481~519位的残基)接头及L2(第629~649位的残基)接头分别与RuvC-II及RuvC-III连接。WED结构域及RucV结构域通过“phosphate lock(磷酸酯锁)”环(第775~787位的残基)连接。

[0072] 本发明的一个实施方式中,在序列号2所示的氨基酸序列的第721~755位之间连续存在的缺失区域为第721~745位的区域(方式1-1)。

[0073] 作为方式1-1的蛋白质及编码其的基因,可分别列举由序列号4所示的氨基酸序列构成的蛋白质、及由序列号3所示的碱基序列构成的基因。该蛋白质相当于后述的dSaCas9变体T10。

[0074] 本发明的一个实施方式中,在序列号2所示的氨基酸序列的第721~755位之间连续存在的缺失区域为第721~755位的区域(方式1-2)。

[0075] 作为方式1-2的蛋白质及编码其的基因,可分别列举由序列号6所示的氨基酸序列构成的蛋白质、及由序列号5所示的碱基序列构成的基因。该蛋白质相当于后述的dSaCas9变体T11。

[0076] 本发明的另一实施方式中,本发明提供一种蛋白质(方式2),其在上述方式1、1-1及1-2各自的变异的基础上还在第45位和/或第163位具有变异、并且具有与向导RNA结合的能力。

[0077] 第45位和/或第163位的变异具体而言是谷氨酸置换为碱性氨基酸,优选置换为赖氨酸、精氨酸或组氨酸,更优选置换为赖氨酸。

[0078] 作为在序列号2所示的氨基酸序列中任选形成“在第721~755位之间连续的缺失区域”时的方法及“将第45位和/或第163位的谷氨酸置换为其它氨基酸”时的方法,可列举如下方法:对于编码规定的氨基酸序列的DNA实施惯用的定点诱变,然后通过常规方法使该DNA表达。在此,作为定点诱变法,可列举例如利用琥珀突变的方法(缺口双链体法、Nucleic Acids Res., 12, 9441-9456 (1984))、利用使用变异导入用引物的PCR的方法等。另外,可使用Q5定点诱变试剂盒(NEB)按照手册简便地实施。

[0079] 本发明的另一实施方式中,本发明提供与上述方式1、1-1、1-2及2的蛋白质在功能上同等的蛋白质(方式3)。为了与上述方式1、1-1、1-2及2的蛋白质在功能上同等,在序列号



2所示的氨基酸序列中,在上述方式1、1-1、1-2及2中除实施了变异的位置以外的部位具有80%以上的序列一致性、且具有与向导RNA结合的能力。在由于变异而使氨基酸存在增减的情况下,该“除实施了变异的位置以外的部位”可理解为“除与实施了变异的位置相当的位置以外的部位”。作为所述一致性,优选为80%以上,更优选为85%以上,进一步优选为90%以上,特别优选为95%以上,最优选为99%以上。氨基酸序列一致性可通过本身公知的方法来确定。例如,氨基酸序列一致性(%)可以使用该领域中惯用的程序(例如BLAST、FASTA等)在初始设定下确定。另外,另外,在另一侧面中,一致性(%)可以使用该领域中公知的任意算法、例如Needleman等(1970)(J.Mol.Biol.48:444-453)、Myers及Miller(CABIOS,1988,4:11-17)的算法等来确定。Needleman等的算法已整合到GCG软件包(可由www.gcg.com获得)的GAP程序中,一致性(%)例如可通过使用BLOSUM 62matrix或PAM250 matrix;以及gap weight:16、14、12、10、8、6或4,及length weight:1、2、3、4、5或6中的任一者来确定。另外,Myers及Miller的算法已整合到作为GCG序列比对软件包的一部分的ALIGN程序中。为了比较氨基酸序列而利用ALIGN程序时,例如可以使用PAM120 weight residue table、gap length penalty 12、gap penalty 4。

[0080] 作为与上述方式1、1-1、1-2及2的蛋白质在功能上同等的蛋白质,提供在序列号2所示的氨基酸序列中在上述方式1、1-1、1-2及2中除实施了变异的位置以外的部位置换、缺失、插入和/或添加1个~几个氨基酸、并且具有与向导RNA结合的能力的蛋白质(方式3-1)。在由于变异而使氨基酸存在增减的情况下,该“除实施了变异的位置以外的部位”可以理解为“除与实施了变异的位置相当的位置以外的部位”。

[0081] 作为人工进行“氨基酸的置换、缺失、插入和/或添加”时的方法,可列举例如以下方法:对于编码规定的氨基酸序列的DNA实施惯用的定点诱变,然后通过常规方法使该DNA表达。在此,作为定点诱变法,可列举例如利用琥珀突变的方法(缺口双链体法、Nucleic Acids Res.,12,9441-9456(1984))、利用使用变异导入用引物的PCR的方法等。另外,可以使用Q5定点诱变试剂盒(NEB)按照手册简便地实施。

[0082] 关于上述所改造的氨基酸的数量,为至少1个残基,具体为1个或几个或其以上。另外,在上述置换、缺失、插入或添加中,特别优选氨基酸的置换。该置换更优选置换为在疏水性、电荷、pK、立体结构上的特征等方面具有类似性质的氨基酸。作为这样的置换,可列举例如i)甘氨酸、丙氨酸;ii)缬氨酸、异亮氨酸、亮氨酸;iii)天冬氨酸、谷氨酸、天冬酰胺、谷氨酰胺;iv)丝氨酸、苏氨酸;v)赖氨酸、精氨酸;vi)苯丙氨酸、酪氨酸中的组内置换。

[0083] 本发明的dSaCas9变异体由于缺失变异而呈dSaCas9变异体蛋白被切断的状态,该缺失区域的两端通过接头连接。即,本发明的dSaCas9变异体中,位于与缺失区域分别相邻的位置的氨基酸通过由3~10个氨基酸残基构成的接头连接。通过所述连接,本申请发明的dSaCas9变异体具有连续的氨基酸序列。

[0084] 接头只要能够连接被切断的蛋白质的两端、且不影响其功能就没有特别限定,优选能够采取根据其它蛋白质而自由地改变自身形状并进行结合的天然变性结构的基团,更优选由甘氨酸(G)及丝氨酸(S)构成的长度为5~9个氨基酸的肽残基。具体可列举以下的残基。

[0085] -SGGGS-(序列号7)

[0086] -GGSGGS-(序列号8)

[0087] -SGSGSGSG- (序列号9)

[0088] -SGSGSGSGS- (序列号10)

[0089] dSaCas9变异体T10及T11中,该接头使用的是-SGGGS-。

[0090] 接头向各变异体中的导入也可以通过如下方法来实施:对于编码规定的氨基酸序列的DNA实施惯用的定点诱变并插入编码接头的碱基序列,然后通过常规方法使该DNA表达。作为定点诱变法,可列举与上述相同的方法。

[0091] 本实施方式中的dSaCas9变异体例如可通过如下方法来制作。首先,使用包含编码上述本发明的dSaCas9变异体的核酸的载体转化宿主。接着,对该宿主进行培养而表达上述蛋白质。培养基的组成、培养的温度、时间、诱导物质的添加等条件可以按照公知的方法由本领域技术人员来决定,以使转化体能够生长、高效地产生上述蛋白质。另外,例如在表达载体中整合有作为选择标志物的抗生素抗性基因的情况下,可以向培养基中加入抗生素来选择转化体。接着,通过适宜地利用本身公知的方法来纯化宿主所表达的上述蛋白质,可得到本发明的dSaCas9变异体。

[0092] 作为宿主,没有特别限定,可列举动物细胞、植物细胞、昆虫细胞或大肠杆菌、枯草杆菌、酵母等微生物。优选动物细胞。

[0093] <dSaCas9变异体-向导RNA复合物>

[0094] 一个实施方式中,本发明提供如下蛋白质-RNA复合物,其具备:上述的<dSaCas9变异体>中所示的蛋白质和包含下述多核苷酸的向导RNA,所述多核苷酸由与靶双链多核苷酸中的PAM(前间隔序列邻近基序)序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

[0095] 上述蛋白质及上述向导RNA可通过在体外及体内在温和的条件下混合而形成蛋白质-RNA复合物。温和的条件表示蛋白质不发生分解或变性的程度的温度及pH,温度优选为4℃以上且40℃以下,pH优选为4以上且10以下。

[0096] 另外,将上述蛋白质及上述向导RNA混合并进行温育的时间优选为0.5小时以上且1小时以下。由上述蛋白质及上述向导RNA形成的复合物稳定,即使在室温下静置数小时也可以保持稳定性。

[0097] <CRISPR-Cas载体系统>

[0098] 一个实施方式中,本发明提供如下CRISPR-Cas载体系统,其具备第一载体和第二载体,所述第一载体包含编码上述的<dSaCas9变异体>中所示的蛋白质的基因;所述第二载体含有包含下述多核苷酸的向导RNA,所述多核苷酸由与靶双链多核苷酸中的PAM序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

[0099] 另一实施方式中,本发明提供如下CRISPR-Cas载体系统,其具有在同一载体中具有编码上述的<dSaCas9变异体>中所示的蛋白质的基因和包含下述多核苷酸的向导RNA,所述多核苷酸由与靶双链多核苷酸中的PAM序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

[0100] 向导RNA可以适宜地设计成:在其5'末端区域包含由与靶双链多核苷酸中的PAM序列的上游1个碱基至优选20个碱基以上且24个碱基以下、更优选22个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成的多核苷酸。此外,也可以包含1个以上的下述多核苷酸,所述多核苷酸由包含与靶双链多核苷酸不互补的碱基序列、按照以一点为轴对称地形

成互补的序列的方式排列、并能够形成发夹结构的碱基序列构成。

[0101] 本实施方式的载体优选为表达载体。作为表达载体,可使用例如pBR322、pBR325、pUC12、pUC13等大肠杆菌来源的质粒;pUB110、pTP5、pC194等枯草杆菌来源的质粒;pSH19、pSH15等酵母来源的质粒; $\lambda$ 噬菌体等噬菌体;腺病毒、腺相关病毒、慢病毒、牛痘病毒、杆状病毒、巨细胞病毒等病毒;及对这些进行改造而成的载体等,在考虑体内的基因表达的活化的情况下,优选病毒载体、特别是腺相关病毒。

[0102] 上述表达载体中,作为上述dSaCas9变异体蛋白及上述向导RNA表达用启动子,没有特别限定,可以使用例如EF1 $\alpha$ 启动子、SR $\alpha$ 启动子、SV40启动子、LTR启动子、CMV(巨细胞病毒)启动子、HSV-tk启动子等用于在动物细胞中表达的启动子、花椰菜花叶病毒(CaMV)的35S启动子、REF(rubber elongation factor,橡胶延长因子)启动子等用于在植物细胞中表达的启动子、多角体启动子、p10启动子等用于在昆虫细胞中表达的启动子等。这些启动子可以根据上述dSaCas9变异体蛋白及上述向导RNA、或表达上述Cas9蛋白及上述向导RNA的细胞的种类来适宜选择。

[0103] 上述表达载体可以还具有多克隆位点、增强子、剪接信号、多聚腺苷酸化信号、选择标志物(抗药性)及其启动子、复制起点等。

[0104] <用于对靶双链多核苷酸进行位点特异性改造的方法>

[0105] [第一实施方式]

[0106] 一个实施方式中,本发明提供用于对靶双链多核苷酸进行位点特异性改造的方法,所述方法具备:

[0107] 将靶双链多核苷酸、蛋白质和向导RNA混合并进行温育的步骤;和上述蛋白质在位于PAM序列的上游的结合位点处对上述靶双链多核苷酸进行改造的步骤,

[0108] 上述靶双链多核苷酸具有PAM序列,

[0109] 上述蛋白质为上述的<dSaCas9变异体>中所示的蛋白质,

[0110] 上述向导RNA包含下述多核苷酸,所述多核苷酸由与上述靶双链多核苷酸中的上述PAM序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

[0111] 本实施方式中,靶双链多核苷酸只要具有PAM序列即可,没有特别限定。

[0112] 本实施方式中,关于蛋白质及向导RNA,如上述的<dSaCas9变异体>中所示。

[0113] 以下对用于对靶双链多核苷酸进行位点特异性改造的方法进行详细说明。

[0114] 首先,将上述蛋白质及上述向导RNA在温和的条件下混合并进行温育。温和的条件如上所述。进行温育的时间优选为0.5小时以上且1小时以下。由上述蛋白质及上述向导RNA形成的复合物稳定,在室温下静置数小时也能够保持稳定。

[0115] 然后,上述蛋白质及上述向导RNA在上述靶双链多核苷酸上形成复合物。上述蛋白质识别PAM序列并在位于PAM序列的上游的结合位点处与上述靶双链多核苷酸结合。接着,可以得到在通过上述向导RNA与上述双链多核苷酸的互补性结合而确定的区域中实施了目标改造的靶双链多核苷酸。

[0116] 本说明书中,“改造”是指靶双链多核苷酸结构上或功能上发生变化。可列举例如通过添加功能性蛋白质、碱基序列而进行的靶双链多核苷酸的结构上或功能上的变化。通过该改造,能够使靶双链多核苷酸的功能发生改变、缺失、增强或抑制而添加新功能。

[0117] 本发明的dSaCas9变异体不具有内切核酸酶活性,因此该蛋白质能够在位于PAM序列的上游的结合位点处与上述靶双链多核苷酸结合,但止步于此,不会将其切断。因此,例如预先在该蛋白质上融合荧光蛋白(例如GFP)等标记蛋白时,可以介由dSaCas9变异体蛋白-向导RNA使标记蛋白与靶双链多核苷酸结合。通过适宜地选择与dSaCas9变异体结合的物质,能够给靶双链多核苷酸带来多样的功能。

[0118] 进一步地,可以在dSaCas9变异体蛋白的N末端或C末端连接转录调控因子蛋白或结构域。作为转录调控因子或其结构域,可列举转录活化因子或其结构域(例如VP64、VP160、NF- $\kappa$ B p65)及转录沉默子或其结构域(例如异染色质蛋白1(HP1))或转录抑制因子或其结构域(例如KRUPPEL相关盒(KRAB)、ERF阻抑结构域(ERD)、mSin3A相互作用结构域(SID))。

[0119] 也可以连接修饰DNA的甲基化状态的酶(例如DNA甲基转移酶(DNMT)、TET)、修饰组蛋白亚基的酶(例如组蛋白乙酰基转移酶(HAT)、组蛋白去乙酰化酶(HDAC)、组蛋白甲基转移酶、组蛋白去甲基化酶)。

[0120] [第二实施方式]

[0121] 本实施方式中,在温育步骤之前可以还具备下述表达步骤:使用上述CRISPR-Cas载体系统表达上述的<dSaCas9变异体>中所示的蛋白质和向导RNA。

[0122] 本实施方式的表达步骤中,首先使用上述CRISPR-Cas载体系统表达dSaCas9变异体蛋白及向导RNA。作为表达的具体方法,使用包含编码dSaCas9变异体蛋白的基因的表达载体及包含向导RNA的表达载体各载体(或同时包含编码dSaCas9变异体蛋白的基因及向导RNA的表达载体)转化宿主。接着,对该宿主进行培养,表达dSaCas9变异体蛋白及向导RNA。培养基的组成、培养的温度、时间、诱导物质的添加等条件可以按照公知的方法由本领域技术人员决定,以使转化体能够生长、高效地产生融合蛋白。另外,例如在表达载体中整合有作为选择标志物的抗生素抗性基因时,可以向培养基中加入抗生素来选择转化体。接着,通过利用适宜的方法纯化宿主所表达的dSaCas9变异体蛋白及向导RNA,可得到dSaCas9变异体蛋白及向导RNA。

[0123] <用于在细胞内对靶双链多核苷酸进行位点特异性改造的方法>

[0124] 一个实施方式中,本发明提供在细胞内对靶双链多核苷酸进行位点特异性改造的方法,所述方法具备:

[0125] 将上述CRISPR-Cas载体系统导入到细胞中而表达上述<dSaCas9变异体>中所示的蛋白质和向导RNA的表达步骤;

[0126] 上述蛋白质在位于PAM序列的上游的结合位点处与上述靶双链多核苷酸结合的步骤;和

[0127] 得到在通过上述向导RNA与上述靶双链多核苷酸的互补性结合而确定的区域中进行了改造的上述靶双链多核苷酸的步骤,

[0128] 上述向导RNA包含如下多核苷酸,所述多核苷酸由与上述靶双链多核苷酸中的上述PAM序列的上游1个碱基至上游20个碱基以上且24个碱基以下的碱基序列互补的碱基序列构成。

[0129] 本实施方式的表达步骤中,首先使用上述CRISPR-Cas载体系统在细胞内表达dSaCas9变异体蛋白及向导RNA。

[0130] 作为成为本实施方式的方法的适用对象的细胞所来源的生物,可列举例如原核生物、酵母、动物、植物、昆虫等。作为上述动物,没有特别限定,可列举例如人、猴、狗、猫、兔、猪、牛、小鼠、大鼠等,但不限于这些。另外,成为细胞的来源的动物的种类也可以根据期望的靶双链多核苷酸的种类、目的等来任意选择。

[0131] 作为成为本实施方式的方法的适用对象的动物来源的细胞,可列举例如生殖细胞(精子、卵子等)、构成生物体的体细胞、干细胞、祖细胞、从生物体分离出的癌细胞、从生物体分离且获得了永生能力而在体外稳定维持的细胞(细胞系)、从生物体分离且人工进行了基因改造的细胞、从生物体分离且人工替换了细胞核的细胞等,但不限于这些。

[0132] 作为构成生物体的体细胞,可列举例如从皮肤、肾脏、脾脏、肾上腺、肝脏、肺、卵巢、胰腺、子宫、胃、结肠、小肠、大肠、膀胱、前列腺、睾丸、胸腺、肌肉、结缔组织、骨、软骨、血管组织、血液、心脏、眼、脑、神经组织等任意组织采集的细胞等,但不限于这些。作为体细胞,更具体而言,可列举例如成纤维细胞、骨髓细胞、免疫细胞(例如B淋巴细胞、T淋巴细胞、嗜中性粒细胞、巨噬细胞、单核细胞等)、红细胞、血小板、骨细胞、骨髓细胞、周皮细胞、树突细胞、角质形成细胞、脂肪细胞、间充质细胞、上皮细胞、表皮细胞、内皮细胞、血管内皮细胞、淋巴管内皮细胞、肝细胞、胰岛细胞(例如 $\alpha$ 细胞、 $\beta$ 细胞、 $\delta$ 细胞、 $\epsilon$ 细胞、PP细胞等)、软骨细胞、卵丘细胞、胶质细胞、神经细胞(神经元)、少突胶质细胞、小胶质细胞、星形胶质细胞、心肌细胞、食道细胞、肌肉细胞(例如平滑肌细胞、骨骼肌细胞等)、黑色素细胞、单核细胞等,但不限于这些。

[0133] 干细胞是兼具自我复制能力和分化为其它多系统的细胞的能力的细胞。作为干细胞,可列举例如胚胎干细胞(ES细胞)、胚胎肿瘤细胞、胚胎生殖干细胞、诱导多能干细胞(iPS细胞)、神经干细胞、造血干细胞、间充质干细胞、肝干细胞、胰干细胞、肌干细胞、生殖干细胞、肠干细胞、癌干细胞、毛囊干细胞等,但不限于这些。

[0134] 癌细胞是指由体细胞派生并获得了无限增殖能力的细胞。作为成为癌细胞的来源的癌症,可列举例如乳腺癌(例如浸润性导管癌、非浸润性导管癌、炎性乳腺癌等)、前列腺癌(例如激素依赖性前列腺癌、激素非依赖性前列腺癌等)、胰腺癌(例如胰管癌等)、胃癌(例如乳头状腺癌、粘液性腺癌、腺鳞状细胞癌等)、肺癌(例如非小细胞肺癌、小细胞肺癌、恶性间皮瘤等)、结肠癌(例如消化道间质瘤等)、直肠癌(例如消化道间质瘤等)、大肠癌(例如家族性大肠癌、遗传性非息肉性大肠癌、消化道间质瘤等)、小肠癌(例如非霍奇金淋巴瘤、消化道间质瘤等)、食道癌、十二指肠癌、舌癌、咽癌(例如上咽癌、中咽癌、下咽癌等)、头颈癌、唾液腺癌、脑瘤(例如松果体星形细胞瘤、毛细胞星形细胞瘤、弥漫性星形细胞瘤、退行性星形细胞瘤等)、神经鞘瘤、肝癌(例如原发性肝癌、肝外胆管癌等)、肾癌(例如肾细胞癌、肾盂和输尿管移行上皮癌等)、胆囊癌、胆管癌、胰腺癌、子宫内膜癌、宫颈癌、卵巢癌(例如上皮性卵巢癌、性腺外生殖细胞瘤、卵巢性生殖细胞瘤、卵巢低恶性度肿瘤等)、膀胱癌、尿道癌、皮肤癌(例如眼内(眼)黑色素瘤、默克细胞癌等)、血管瘤、恶性淋巴瘤(例如网状细胞肉瘤、淋巴肉瘤、霍奇金病等)、黑色素瘤(恶性黑色素瘤)、甲状腺癌(例如甲状腺髓样癌等)、甲状旁腺癌、鼻癌、鼻窦癌、骨癌(例如骨肉瘤、尤文肉瘤、子宫肉瘤、软组织肉瘤等)、转移性髓母细胞瘤、血管纤维瘤、隆突性皮肤纤维肉瘤、视网膜肉瘤、阴茎癌、睾丸肿瘤、小儿实体癌(例如威尔姆氏瘤、小儿肾肿瘤等)、卡波西肉瘤、起因于AIDS的卡波西肉瘤、上颌窦肿瘤、纤维组织细胞瘤、平滑肌肉瘤、横纹肌肉瘤、慢性骨髓增殖性疾病、白血病(例如急性

髓系白血病、急性淋巴母细胞性白血病等)等,但不限于这些。

[0135] 细胞系是通过生物体外的人工操作而获得了无限增殖能力的细胞。作为细胞系,可列举例如HCT116、Huh7、HEK293(人胎肾细胞)、HeLa(人宫颈癌细胞系)、HepG2(人肝癌细胞系)、UT7/TPO(人白血病细胞系)、CHO(中华仓鼠卵巢细胞系)、MDCK、MDBK、BHK、C-33A、HT-29、AE-1、3D9、Ns0/1、Jurkat、NIH3T3、PC12、S2、Sf9、Sf21、High Five、Vero等,但不限于这些。

[0136] 作为将CRISPR-Cas载体系统导入到细胞中的方法,可通过适合于所用的活细胞的方法来进行,可列举电穿孔法、热休克法、磷酸钙法、脂质体转染法、DEAE葡聚糖法、显微注射法、基因枪法、使用病毒的方法以及使用FuGENE(注册商标)6转染试剂(罗氏公司制)、Lipofectamine 2000试剂(Invitrogen公司制)、Lipofectamine LTX试剂(Invitrogen公司制)、Lipofectamine 3000试剂(Invitrogen公司制)等市售的转染试剂的方法等。

[0137] 接下来的改造步骤与上述的<用于对靶双链核苷酸进行位点特异性改造的方法>的[第一实施方式]中所示的方法相同。

[0138] 通过本实施方式的靶双链多核苷酸的修饰,可得到靶双链多核苷酸被改造的细胞。

[0139] 以下示出实施例来更详细地说明本发明,但是这些并非对本发明的范围进行限定。

[0140] 实施例

[0141] 实施例1:dSaCas9变异体的DNA结合亲和性评价

[0142] (方法)

[0143] 1.克隆

[0144] 使用NEB Q5定点诱变试剂盒,在dSaCas9基因内形成规定的缺失区域,导入编码接头的基因,进一步融合作为转录调控因子的KRAB基因,由此制作各种dSaCas9(图1)。对于这些dSaCas9,使用MYD88基因调查其表达抑制活性。将全部dSaCas9的基因构建物整合到pX601载体中(F. Ann Ran et al., Nature 2015;520(7546); pp.186-191)。各载体表达以下的dSaCas9构建物(FL、T10、T11)及sgRNA。

[0145] [dSaCas9变异体]

[0146] FL:野生型dSaCas9(序列号1、序列号2)

[0147] T10:在野生型dSaCas9中具有缺失区域(第721~745位)的缺失变异体(序列号3、序列号4)

[0148] T11:在野生型dSaCas9中具有缺失区域(第721~755位)的缺失变异体(序列号5、序列号6)

[0149] T10及T11中的各缺失区域的两端通过接头肽连接。

[0150] dSaCas9构建物以dSaCas9-KRAB-P2A-sfGFP或dSaCas9-VR-P2A-sfGFP的融合蛋白形式进行表达。

[0151] [crRNA序列]

[0152] C1:ACGGAGGCUAAGCGUCGCAA(序列号11)

[0153] A2:GGAGCCACAGUUCUCCACGG(序列号12)

[0154] A3:CUCUACCCUUGAGGUCUCGAG(序列号13)

[0155] 上述crRNA序列与以下的tracrRNA序列融合形成sgRNA并从载体表达。

[0156] [tracrRNA序列]

[0157] GUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCACGUCAACUUGUUGGCGAGAUUUUUUU (序列号14)

[0158] 2. 细胞转染

[0159] 在转染前24小时,将HEK293FT细胞以每孔中75000个细胞的密度播种于24孔板,用添加有10%FBS、2mM的新鲜的L-谷氨酰胺、1mM的丙酮酸钠及非必需氨基酸的DMEM培养基进行培养。使用1.5 $\mu$ l的Lipofectamine 2000 (Life technologies),按照操作说明书使各质粒(500ng)转染细胞。为了进行基因表达分析,在转染后48~72小时回收细胞,溶解于RLT缓冲液(Qiagen),使用RNeasy试剂盒(Qiagen)提取总RNA。

[0160] 3. 基因表达分析

[0161] 为了进行Taqman分析,以20 $\mu$ l的容量使用TaqMan<sup>TM</sup>高容量RNA-to-cDNA试剂盒(Applied Biosystems)由1.5 $\mu$ g的总RNA制备cDNA。将制备的cDNA稀释至20倍,每个Taqman反应使用6.33 $\mu$ l。针对MYD88基因及HPRT基因的Taqman引物及探针由Applied Biosystems获得。在Roche LightCycler 96或LightCycler 480中,使用Taqman基因表达预混物(gene expression master mix)(ThermoFisher)进行Taqman反应,使用LightCycler 96 analysis软件进行分析。

[0162] Taqman探针产物IDs:

[0163] MYD88:Hs01573837\_g1 (FAM)

[0164] HPRT:Hs99999909\_m1 (FAM, VIC)

[0165] Taqman QPCR条件:

[0166] 步骤1:95 $^{\circ}$ C 10分钟

[0167] 步骤2:95 $^{\circ}$ C 15秒

[0168] 步骤3:60 $^{\circ}$ C 30秒

[0169] 重复步骤2和3:40次

[0170] (结果)

[0171] 利用表达pX601-dSaCas9 (FL、T10或T11) -KRAB-P2A-sfGFP及sgRNA-A2的质粒载体转染HEK293FT细胞。使用利用表达sgRNA-C1来代替表达sgRNA-A2的载体进行转染而得的细胞作为阴性对照。3天后回收所转染的细胞,单离出RNA。通过Taqman分析来分析MYD88基因的表达,利用HPRT基因的表达水平来标准化。将MYD88基因的相对表达水平的结果示于图2。

[0172] 本发明的dSaCas9变异体的MYD88基因表达水平与对照相比降低至与野生型dSaCas9匹敌的程度。该结果表明,本发明的dSaCas9变异体尽管存在缺失区域、与全长的dSaCas9相比缩小,但是与向导RNA结合的能力以及DNA结合亲和性得到了维持。

[0173] 上述结果中,特别是对于DNA结合亲和性高的T10,即使改变向导RNA也确认到其结合亲和性。利用表达pX601-dSaCas9 (FL或T10) -KRAB-P2A-sfGFP及sgRNA-A2或sgRNA-A3的质粒载体转染HEK293FT细胞。使用利用不表达dSaCas9变异体的载体转染而得的细胞作为阴性对照。3天后回收所转染的细胞,单离出RNA。利用Taqman分析来分析MYD88基因的表达,利用HPRT基因的表达水平进行标准化。将MYD88基因的相对表达水平的结果示于图3。在使用sgRNA-A3的情况下也确认到与sgRNA-A2同样的效果。

[0174] 本发明的dSaCas9变异体中的MYD88基因表达水平与对照相比降低至与野生型dSaCas9匹敌的程度。该结果表明,本发明的dSaCas9变异体尽管存在缺失区域、与全长的dSaCas9相比缩小,但是与向导RNA结合的能力以及DNA结合亲和性得到了维持。

[0175] 产业上的可利用性

[0176] 根据本发明,能够得到保持了DNA结合亲和性且小型化的dSaCas9蛋白。小型化的dSaCas9蛋白的使用使更多的基因向载体中的搭载成为可能,因此能够提供多样的基因组编辑技术。

[0177] 本申请以在美国提出的美国临时专利申请第62/749855(申请日2018年10月24日)为基础,其内容全部包含在本说明书中。



## 序列表

<110> 摩大力斯医疗株式会社  
 <120> 经改造的Cas9蛋白及其用途  
 <130> 092961  
 <150> US62/749855  
 <151> 2018-10-24  
 <160> 14  
 <170> PatentIn version 3.5  
 <210> 1  
 <211> 3162  
 <212> DNA  
 <213> 金黄色葡萄球菌 (Staphylococcus aureus)  
 <220>  
 <221> CDS  
 <222> (1) .. (3162)  
 <400> 1

```

atg aag cgg aac tac atc ctg ggc ctg gcc atc ggc atc acc agc gtg 48
Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
1           5           10           15
ggc tac ggc atc atc gac tac gag aca cgg gac gtg atc gat gcc ggc 96
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
           20           25           30
gtg cgg ctg ttc aaa gag gcc aac gtg gaa aac aac gag ggc agg cgg 144
Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
           35           40           45
agc aag aga ggc gcc aga agg ctg aag cgg cgg agg cgg cat aga atc 192
Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
           50           55           60
cag aga gtg aag aag ctg ctg ttc gac tac aac ctg ctg acc gac cac 240
Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
65           70           75           80
agc gag ctg agc ggc atc aac ccc tac gag gcc aga gtg aag ggc ctg 288
Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
           85           90           95
agc cag aag ctg agc gag gaa gag ttc tct gcc gcc ctg ctg cac ctg 336
Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
           100          105          110
gcc aag aga aga ggc gtg cac aac gtg aac gag gtg gaa gag gac acc 384
  
```

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr  
 115 120 125  
 ggc aac gag ctg tcc acc aaa gag cag atc agc cgg aac agc aag gcc 432  
 Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala  
 130 135 140  
 ctg gaa gag aaa tac gtg gcc gaa ctg cag ctg gaa cgg ctg aag aaa 480  
 Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys  
 145 150 155 160  
 gac ggc gaa gtg cgg ggc agc atc aac aga ttc aag acc agc gac tac 528  
 Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr  
 165 170 175  
 gtg aaa gaa gcc aaa cag ctg ctg aag gtg cag aag gcc tac cac cag 576  
 Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln  
 180 185 190  
 ctg gac cag agc ttc atc gac acc tac atc gac ctg ctg gaa acc cgg 624  
 Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg  
 195 200 205  
 cgg acc tac tat gag gga cct ggc gag ggc agc ccc ttc ggc tgg aag 672  
 Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys  
 210 215 220  
 gac atc aaa gaa tgg tac gag atg ctg atg ggc cac tgc acc tac ttc 720  
 Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe  
 225 230 235 240  
 ccc gag gaa ctg cgg agc gtg aag tac gcc tac aac gcc gac ctg tac 768  
 Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr  
 245 250 255  
 aac gcc ctg aac gac ctg aac aat ctc gtg atc acc agg gac gag aac 816  
 Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn  
 260 265 270  
 gag aag ctg gaa tat tac gag aag ttc cag atc atc gag aac gtg ttc 864  
 Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe  
 275 280 285  
 aag cag aag aag aag ccc acc ctg aag cag atc gcc aaa gaa atc ctc 912  
 Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu  
 290 295 300  
 gtg aac gaa gag gat att aag ggc tac aga gtg acc agc acc ggc aag 960  
 Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys  
 305 310 315 320  
 ccc gag ttc acc aac ctg aag gtg tac cac gac atc aag gac att acc 1008



Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu  
 530 535 540  
 gat ctg ctg aac aac ccc ttc aac tat gag gtg gac cac atc atc ccc 1680  
 Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro  
 545 550 555 560  
 aga agc gtg tcc ttc gac aac agc ttc aac aac aag gtg ctc gtg aag 1728  
 Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys  
 565 570 575  
 cag gaa gaa gcc agc aag aag ggc aac cgg acc cca ttc cag tac ctg 1776  
 Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu  
 580 585 590  
 agc agc agc gac agc aag atc agc tac gaa acc ttc aag aag cac atc 1824  
 Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile  
 595 600 605  
 ctg aat ctg gcc aag ggc aag ggc aga atc agc aag acc aag aaa gag 1872  
 Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu  
 610 615 620  
 tat ctg ctg gaa gaa cgg gac atc aac agg ttc tcc gtg cag aaa gac 1920  
 Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp  
 625 630 635 640  
 ttc atc aac cgg aac ctg gtg gat acc aga tac gcc acc aga ggc ctg 1968  
 Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu  
 645 650 655  
 atg aac ctg ctg cgg agc tac ttc aga gtg aac aac ctg gac gtg aaa 2016  
 Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys  
 660 665 670  
 gtg aag tcc atc aat ggc ggc ttc acc agc ttt ctg cgg cgg aag tgg 2064  
 Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp  
 675 680 685  
 aag ttt aag aaa gag cgg aac aag ggg tac aag cac cac gcc gag gac 2112  
 Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp  
 690 695 700  
 gcc ctg atc att gcc aac gcc gat ttc atc ttc aaa gag tgg aag aaa 2160  
 Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys  
 705 710 715 720  
 ctg gac aag gcc aaa aaa gtg atg gaa aac cag atg ttc gag gaa aag 2208  
 Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys  
 725 730 735  
 cag gcc gag agc atg ccc gag atc gaa acc gag cag gag tac aaa gag 2256

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu  
 740 745 750  
 atc ttc atc acc ccc cac cag atc aag cac att aag gac ttc aag gac 2304  
 Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp  
 755 760 765  
 tac aag tac agc cac cgg gtg gac aag aag cct aat aga gag ctg att 2352  
 Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile  
 770 775 780  
 aac gac acc ctg tac tcc acc cgg aag gac gac aag ggc aac acc ctg 2400  
 Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu  
 785 790 795 800  
 atc gtg aac aat ctg aac ggc ctg tac gac aag gac aat gac aag ctg 2448  
 Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu  
 805 810 815  
 aaa aag ctg atc aac aag agc ccc gaa aag ctg ctg atg tac cac cac 2496  
 Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His  
 820 825 830  
 gac ccc cag acc tac cag aaa ctg aag ctg att atg gaa cag tac ggc 2544  
 Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly  
 835 840 845  
 gac gag aag aat ccc ctg tac aag tac tac gag gaa acc ggg aac tac 2592  
 Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr  
 850 855 860  
 ctg acc aag tac tcc aaa aag gac aac ggc ccc gtg atc aag aag att 2640  
 Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile  
 865 870 875 880  
 aag tat tac ggc aac aaa ctg aac gcc cat ctg gac atc acc gac gac 2688  
 Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp  
 885 890 895  
 tac ccc aac agc aga aac aag gtc gtg aag ctg tcc ctg aag ccc tac 2736  
 Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr  
 900 905 910  
 aga ttc gac gtg tac ctg gac aat ggc gtg tac aag ttc gtg acc gtg 2784  
 Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val  
 915 920 925  
 aag aat ctg gat gtg atc aaa aaa gaa aac tac tac gaa gtg aat agc 2832  
 Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser  
 930 935 940  
 aag tgc tat gag gaa gct aag aag ctg aag aag atc agc aac cag gcc 2880

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala  
 945 950 955 960  
 gag ttt atc gcc tcc ttc tac aac aac gat ctg atc aag atc aac ggc 2928  
 Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly  
 965 970 975  
 gag ctg tat aga gtg atc ggc gtg aac aac gac ctg ctg aac cgg atc 2976  
 Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile  
 980 985 990  
 gaa gtg aac atg atc gac atc acc tac cgc gag tac ctg gaa aac atg 3024  
 Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met  
 995 1000 1005  
 aac gac aag agg ccc ccc agg atc att aag aca atc gcc tcc aag 3069  
 Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys  
 1010 1015 1020  
 acc cag agc att aag aag tac agc aca gac att ctg ggc aac ctg 3114  
 Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu  
 1025 1030 1035  
 tat gaa gtg aaa tct aag aag cac cct cag atc atc aaa aag ggc 3159  
 Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly  
 1040 1045 1050  
 taa 3162  
 <210> 2  
 <211> 1053  
 <212> PRT  
 <213> 金黄色葡萄球菌 (Staphylococcus aureus)  
 <400> 2  
 Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val  
 1 5 10 15  
 Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly  
 20 25 30  
 Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg  
 35 40 45  
 Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile  
 50 55 60  
 Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His  
 65 70 75 80  
 Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu  
 85 90 95  
 Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu

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

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



	725		730		735
Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu					
	740		745		750
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp					
	755		760		765
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile					
	770		775		780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu					
785		790		795	800
Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu					
	805		810		815
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His					
	820		825		830
Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly					
	835		840		845
Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr					
	850		855		860
Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile					
865		870		875	880
Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp					
	885		890		895
Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr					
	900		905		910
Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val					
	915		920		925
Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser					
	930		935		940
Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala					
945		950		955	960
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly					
	965		970		975
Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile					
	980		985		990
Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met					
	995		1000		1005
Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys					
	1010		1015		1020
Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu					
	1025		1030		1035

Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly  
 1040 1045 1050  
 <210> 3  
 <211> 3102  
 <212> DNA  
 <213> 金黄色葡萄球菌 (Staphylococcus aureus)  
 <220>  
 <221> CDS  
 <222> (1) .. (3102)  
 <400> 3  
 atg aag cgg aac tac atc ctg ggc ctg gcc atc ggc atc acc agc gtg 48  
 Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val  
 1 5 10 15  
 ggc tac ggc atc atc gac tac gag aca cgg gac gtg atc gat gcc ggc 96  
 Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly  
 20 25 30  
 gtg cgg ctg ttc aaa gag gcc aac gtg gaa aac aac gag ggc agg cgg 144  
 Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg  
 35 40 45  
 agc aag aga ggc gcc aga agg ctg aag cgg cgg agg cgg cat aga atc 192  
 Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile  
 50 55 60  
 cag aga gtg aag aag ctg ctg ttc gac tac aac ctg ctg acc gac cac 240  
 Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His  
 65 70 75 80  
 agc gag ctg agc ggc atc aac ccc tac gag gcc aga gtg aag ggc ctg 288  
 Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu  
 85 90 95  
 agc cag aag ctg agc gag gaa gag ttc tct gcc gcc ctg ctg cac ctg 336  
 Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu  
 100 105 110  
 gcc aag aga aga ggc gtg cac aac gtg aac gag gtg gaa gag gac acc 384  
 Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr  
 115 120 125  
 ggc aac gag ctg tcc acc aaa gag cag atc agc cgg aac agc aag gcc 432  
 Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala  
 130 135 140  
 ctg gaa gag aaa tac gtg gcc gaa ctg cag ctg gaa cgg ctg aag aaa 480  
 Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys

145	150	155	160
gac ggc gaa gtg cgg ggc agc atc aac aga ttc aag acc agc gac tac	528		
Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr			
	165	170	175
gtg aaa gaa gcc aaa cag ctg ctg aag gtg cag aag gcc tac cac cag	576		
Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln			
	180	185	190
ctg gac cag agc ttc atc gac acc tac atc gac ctg ctg gaa acc cgg	624		
Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg			
	195	200	205
cgg acc tac tat gag gga cct ggc gag ggc agc ccc ttc ggc tgg aag	672		
Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys			
	210	215	220
gac atc aaa gaa tgg tac gag atg ctg atg ggc cac tgc acc tac ttc	720		
Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe			
	225	230	235
ccc gag gaa ctg cgg agc gtg aag tac gcc tac aac gcc gac ctg tac	768		
Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr			
	245	250	255
aac gcc ctg aac gac ctg aac aat ctc gtg atc acc agg gac gag aac	816		
Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn			
	260	265	270
gag aag ctg gaa tat tac gag aag ttc cag atc atc gag aac gtg ttc	864		
Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe			
	275	280	285
aag cag aag aag aag ccc acc ctg aag cag atc gcc aaa gaa atc ctc	912		
Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu			
	290	295	300
gtg aac gaa gag gat att aag ggc tac aga gtg acc agc acc ggc aag	960		
Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys			
	305	310	315
ccc gag ttc acc aac ctg aag gtg tac cac gac atc aag gac att acc	1008		
Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr			
	325	330	335
gcc cgg aaa gag att att gag aac gcc gag ctg ctg gat cag att gcc	1056		
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala			
	340	345	350
aag atc ctg acc atc tac cag agc agc gag gac atc cag gaa gaa ctg	1104		
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu			

355	360	365	
acc aat ctg aac tcc gag ctg acc cag gaa gag atc gag cag atc tct			1152
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser			
370	375	380	
aat ctg aag ggc tat acc ggc acc cac aac ctg agc ctg aag gcc atc			1200
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile			
385	390	395	400
aac ctg atc ctg gac gag ctg tgg cac acc aac gac aac cag atc gct			1248
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala			
	405	410	415
atc ttc aac cgg ctg aag ctg gtg ccc aag aag gtg gac ctg tcc cag			1296
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln			
	420	425	430
cag aaa gag atc ccc acc acc ctg gtg gac gac ttc atc ctg agc ccc			1344
Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro			
	435	440	445
gtc gtg aag aga agc ttc atc cag agc atc aaa gtg atc aac gcc atc			1392
Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile			
	450	455	460
atc aag aag tac ggc ctg ccc aac gac atc att atc gag ctg gcc cgc			1440
Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg			
465	470	475	480
gag aag aac tcc aag gac gcc cag aaa atg atc aac gag atg cag aag			1488
Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys			
	485	490	495
cgg aac cgg cag acc aac gag cgg atc gag gaa atc atc cgg acc acc			1536
Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr			
	500	505	510
ggc aaa gag aac gcc aag tac ctg atc gag aag atc aag ctg cac gac			1584
Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp			
	515	520	525
atg cag gaa ggc aag tgc ctg tac agc ctg gaa gcc atc cct ctg gaa			1632
Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu			
	530	535	540
gat ctg ctg aac aac ccc ttc aac tat gag gtg gac cac atc atc ccc			1680
Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro			
545	550	555	560
aga agc gtg tcc ttc gac aac agc ttc aac aac aag gtg ctc gtg aag			1728
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys			

565	570	575	
cag gaa gaa gcc agc aag aag ggc aac cgg acc cca ttc cag tac ctg			1776
Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu			
580	585	590	
agc agc agc gac agc aag atc agc tac gaa acc ttc aag aag cac atc			1824
Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile			
595	600	605	
ctg aat ctg gcc aag ggc aag ggc aga atc agc aag acc aag aaa gag			1872
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu			
610	615	620	
tat ctg ctg gaa gaa cgg gac atc aac agg ttc tcc gtg cag aaa gac			1920
Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp			
625	630	635	640
ttc atc aac cgg aac ctg gtg gat acc aga tac gcc acc aga ggc ctg			1968
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu			
645	650	655	
atg aac ctg ctg cgg agc tac ttc aga gtg aac aac ctg gac gtg aaa			2016
Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys			
660	665	670	
gtg aag tcc atc aat ggc ggc ttc acc agc ttt ctg cgg cgg aag tgg			2064
Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp			
675	680	685	
aag ttt aag aaa gag cgg aac aag ggg tac aag cac cac gcc gag gac			2112
Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp			
690	695	700	
gcc ctg atc att gcc aac gcc gat ttc atc ttc aaa gag tgg aag aaa			2160
Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys			
705	710	715	720
tcc ggc ggc ggt tcg acc gag cag gag tac aaa gag atc ttc atc acc			2208
Ser Gly Gly Gly Ser Thr Glu Gln Glu Tyr Lys Glu Ile Phe Ile Thr			
725	730	735	
ccc cac cag atc aag cac att aag gac ttc aag gac tac aag tac agc			2256
Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr Lys Tyr Ser			
740	745	750	
cac cgg gtg gac aag aag cct aat aga gag ctg att aac gac acc ctg			2304
His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn Asp Thr Leu			
755	760	765	
tac tcc acc cgg aag gac gac aag ggc aac acc ctg atc gtg aac aat			2352
Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile Val Asn Asn			

770	775	780	
ctg aac ggc ctg tac gac aag gac aat gac aag ctg aaa aag ctg atc			2400
Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys Lys Leu Ile			
785	790	795	800
aac aag agc ccc gaa aag ctg ctg atg tac cac cac gac ccc cag acc			2448
Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His Asp Pro Gln Thr			
805	810	815	
tac cag aaa ctg aag ctg att atg gaa cag tac ggc gac gag aag aat			2496
Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp Glu Lys Asn			
820	825	830	
ccc ctg tac aag tac tac gag gaa acc ggg aac tac ctg acc aag tac			2544
Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu Thr Lys Tyr			
835	840	845	
tcc aaa aag gac aac ggc ccc gtg atc aag aag att aag tat tac ggc			2592
Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile Lys Tyr Tyr Gly			
850	855	860	
aac aaa ctg aac gcc cat ctg gac atc acc gac gac tac ccc aac agc			2640
Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp Tyr Pro Asn Ser			
865	870	875	880
aga aac aag gtc gtg aag ctg tcc ctg aag ccc tac aga ttc gac gtg			2688
Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr Arg Phe Asp Val			
885	890	895	
tac ctg gac aat ggc gtg tac aag ttc gtg acc gtg aag aat ctg gat			2736
Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val Lys Asn Leu Asp			
900	905	910	
gtg atc aaa aaa gaa aac tac tac gaa gtg aat agc aag tgc tat gag			2784
Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys Cys Tyr Glu			
915	920	925	
gaa gct aag aag ctg aag aag atc agc aac cag gcc gag ttt atc gcc			2832
Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu Phe Ile Ala			
930	935	940	
tcc ttc tac aac aac gat ctg atc aag atc aac ggc gag ctg tat aga			2880
Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu Leu Tyr Arg			
945	950	955	960
gtg atc ggc gtg aac aac gac ctg ctg aac cgg atc gaa gtg aac atg			2928
Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu Val Asn Met			
965	970	975	
atc gac atc acc tac cgc gag tac ctg gaa aac atg aac gac aag agg			2976
Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn Asp Lys Arg			

	980	985	990	
ccc	ccc	agg	atc	att
aag	aca	atc	gcc	tcc
aag	acc	cag	agc	att
aag				3024
Pro	Pro	Arg	Ile	Ile
Lys	Thr	Ile	Ala	Ser
Lys	Thr	Gln	Ser	Ile
Lys				
	995	1000	1005	
aag	tac	agc	aca	gac
att	ctg	ggc	aac	ctg
tat	gaa	gtg	aaa	tct
				3069
Lys	Tyr	Ser	Thr	Asp
Ile	Leu	Gly	Asn	Leu
Tyr	Glu	Val	Lys	Ser
	1010	1015	1020	
aag	aag	cac	cct	cag
atc	atc	aaa	aag	ggc
taa				3102
Lys	Lys	His	Pro	Gln
Ile	Ile	Lys	Lys	Gly
	1025	1030		
<210>	4			
<211>	1033			
<212>	PRT			
<213>	金黄色葡萄球菌 (Staphylococcus aureus)			
<400>	4			
Met	Lys	Arg	Asn	Tyr
Ile	Leu	Gly	Leu	Ala
Ile	Gly	Ile	Thr	Ser
Val				
	1	5	10	15
Gly	Tyr	Gly	Ile	Ile
Asp	Tyr	Glu	Thr	Arg
Asp	Val	Ile	Asp	Ala
Gly				
	20	25	30	
Val	Arg	Leu	Phe	Lys
Glu	Ala	Asn	Val	Glu
Asn	Asn	Glu	Gly	Arg
Arg				
	35	40	45	
Ser	Lys	Arg	Gly	Ala
Arg	Arg	Leu	Lys	Arg
Arg	Arg	Arg	Arg	His
Arg	His	Arg	Ile	
	50	55	60	
Gln	Arg	Val	Lys	Lys
Leu	Leu	Phe	Asp	Tyr
Asn	Leu	Leu	Thr	Asp
His				
	65	70	75	80
Ser	Glu	Leu	Ser	Gly
Ile	Asn	Pro	Tyr	Glu
Ala	Arg	Val	Lys	Gly
Leu				
	85	90	95	
Ser	Gln	Lys	Leu	Ser
Glu	Glu	Glu	Phe	Ser
Ala	Ala	Leu	Leu	His
Leu				
	100	105	110	
Ala	Lys	Arg	Arg	Gly
Val	His	Asn	Val	Asn
Glu	Val	Glu	Glu	Asp
Thr				
	115	120	125	
Gly	Asn	Glu	Leu	Ser
Thr	Lys	Glu	Gln	Ile
Ser	Arg	Asn	Ser	Lys
Ala				
	130	135	140	
Leu	Glu	Glu	Lys	Tyr
Val	Ala	Glu	Leu	Gln
Leu	Glu	Arg	Leu	Lys
Lys				
	145	150	155	160
Asp	Gly	Glu	Val	Arg
Gly	Ser	Ile	Asn	Arg
Phe	Lys	Thr	Ser	Asp
Tyr				
	165	170	175	
Val	Lys	Glu	Ala	Lys
Gln	Leu	Leu	Lys	Val
Gln	Lys	Ala	Tyr	His
Gln				
	180	185	190	

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



	500		505		510										
Gly	Lys	Glu	Asn	Ala	Lys	Tyr	Leu	Ile	Glu	Lys	Ile	Lys	Leu	His	Asp
	515		520		525										
Met	Gln	Glu	Gly	Lys	Cys	Leu	Tyr	Ser	Leu	Glu	Ala	Ile	Pro	Leu	Glu
	530		535		540										
Asp	Leu	Leu	Asn	Asn	Pro	Phe	Asn	Tyr	Glu	Val	Asp	His	Ile	Ile	Pro
545			550		555				560						
Arg	Ser	Val	Ser	Phe	Asp	Asn	Ser	Phe	Asn	Asn	Lys	Val	Leu	Val	Lys
			565		570				575						
Gln	Glu	Glu	Ala	Ser	Lys	Lys	Gly	Asn	Arg	Thr	Pro	Phe	Gln	Tyr	Leu
			580		585				590						
Ser	Ser	Ser	Asp	Ser	Lys	Ile	Ser	Tyr	Glu	Thr	Phe	Lys	Lys	His	Ile
	595		600		605										
Leu	Asn	Leu	Ala	Lys	Gly	Lys	Gly	Arg	Ile	Ser	Lys	Thr	Lys	Lys	Glu
	610		615		620										
Tyr	Leu	Leu	Glu	Glu	Arg	Asp	Ile	Asn	Arg	Phe	Ser	Val	Gln	Lys	Asp
625			630		635				640						
Phe	Ile	Asn	Arg	Asn	Leu	Val	Asp	Thr	Arg	Tyr	Ala	Thr	Arg	Gly	Leu
			645		650				655						
Met	Asn	Leu	Leu	Arg	Ser	Tyr	Phe	Arg	Val	Asn	Asn	Leu	Asp	Val	Lys
	660		665		670										
Val	Lys	Ser	Ile	Asn	Gly	Gly	Phe	Thr	Ser	Phe	Leu	Arg	Arg	Lys	Trp
	675		680		685										
Lys	Phe	Lys	Lys	Glu	Arg	Asn	Lys	Gly	Tyr	Lys	His	His	Ala	Glu	Asp
	690		695		700										
Ala	Leu	Ile	Ile	Ala	Asn	Ala	Asp	Phe	Ile	Phe	Lys	Glu	Trp	Lys	Lys
705			710		715				720						
Ser	Gly	Gly	Gly	Ser	Thr	Glu	Gln	Glu	Tyr	Lys	Glu	Ile	Phe	Ile	Thr
			725		730				735						
Pro	His	Gln	Ile	Lys	His	Ile	Lys	Asp	Phe	Lys	Asp	Tyr	Lys	Tyr	Ser
	740		745		750										
His	Arg	Val	Asp	Lys	Lys	Pro	Asn	Arg	Glu	Leu	Ile	Asn	Asp	Thr	Leu
	755		760		765										
Tyr	Ser	Thr	Arg	Lys	Asp	Asp	Lys	Gly	Asn	Thr	Leu	Ile	Val	Asn	Asn
	770		775		780										
Leu	Asn	Gly	Leu	Tyr	Asp	Lys	Asp	Asn	Asp	Lys	Leu	Lys	Lys	Leu	Ile
785			790		795				800						
Asn	Lys	Ser	Pro	Glu	Lys	Leu	Leu	Met	Tyr	His	His	Asp	Pro	Gln	Thr
			805		810				815						



ggc tac ggc atc atc gac tac gag aca cgg gac gtg atc gat gcc ggc 96  
 Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly  
                   20                                  25                                  30  
 gtg cgg ctg ttc aaa gag gcc aac gtg gaa aac aac gag ggc agg cgg 144  
 Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg  
                   35                                  40                                  45  
 agc aag aga ggc gcc aga agg ctg aag cgg cgg agg cgg cat aga atc 192  
 Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile  
                   50                                  55                                  60  
 cag aga gtg aag aag ctg ctg ttc gac tac aac ctg ctg acc gac cac 240  
 Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His  
 65                                  70                                  75                                  80  
 agc gag ctg agc ggc atc aac ccc tac gag gcc aga gtg aag ggc ctg 288  
 Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu  
                                   85                                  90                                  95  
 agc cag aag ctg agc gag gaa gag ttc tct gcc gcc ctg ctg cac ctg 336  
 Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu  
                                   100                                  105                                  110  
 gcc aag aga aga ggc gtg cac aac gtg aac gag gtg gaa gag gac acc 384  
 Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr  
                   115                                  120                                  125  
 ggc aac gag ctg tcc acc aaa gag cag atc agc cgg aac agc aag gcc 432  
 Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala  
                   130                                  135                                  140  
 ctg gaa gag aaa tac gtg gcc gaa ctg cag ctg gaa cgg ctg aag aaa 480  
 Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys  
 145                                  150                                  155                                  160  
 gac ggc gaa gtg cgg ggc agc atc aac aga ttc aag acc agc gac tac 528  
 Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr  
                                   165                                  170                                  175  
 gtg aaa gaa gcc aaa cag ctg ctg aag gtg cag aag gcc tac cac cag 576  
 Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln  
                   180                                  185                                  190  
 ctg gac cag agc ttc atc gac acc tac atc gac ctg ctg gaa acc cgg 624  
 Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg  
                   195                                  200                                  205  
 cgg acc tac tat gag gga cct ggc gag ggc agc ccc ttc ggc tgg aag 672  
 Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys  
                   210                                  215                                  220

```

gac atc aaa gaa tgg tac gag atg ctg atg ggc cac tgc acc tac ttc 720
Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
225                230                235                240
ccc gag gaa ctg cgg agc gtg aag tac gcc tac aac gcc gac ctg tac 768
Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
                245                250                255
aac gcc ctg aac gac ctg aac aat ctc gtg atc acc agg gac gag aac 816
Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
                260                265                270
gag aag ctg gaa tat tac gag aag ttc cag atc atc gag aac gtg ttc 864
Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
                275                280                285
aag cag aag aag aag ccc acc ctg aag cag atc gcc aaa gaa atc ctc 912
Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
                290                295                300
gtg aac gaa gag gat att aag ggc tac aga gtg acc agc acc ggc aag 960
Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305                310                315                320
ccc gag ttc acc aac ctg aag gtg tac cac gac atc aag gac att acc 1008
Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
                325                330                335
gcc cgg aaa gag att att gag aac gcc gag ctg ctg gat cag att gcc 1056
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
                340                345                350
aag atc ctg acc atc tac cag agc agc gag gac atc cag gaa gaa ctg 1104
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
                355                360                365
acc aat ctg aac tcc gag ctg acc cag gaa gag atc gag cag atc tct 1152
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
                370                375                380
aat ctg aag ggc tat acc ggc acc cac aac ctg agc ctg aag gcc atc 1200
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
385                390                395                400
aac ctg atc ctg gac gag ctg tgg cac acc aac gac aac cag atc gct 1248
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
                405                410                415
atc ttc aac cgg ctg aag ctg gtg ccc aag aag gtg gac ctg tcc cag 1296
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
                420                425                430

```

cag aaa gag atc ccc acc acc ctg gtg gac gac ttc atc ctg agc ccc 1344  
 Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro  
 435 440 445  
 gtc gtg aag aga agc ttc atc cag agc atc aaa gtg atc aac gcc atc 1392  
 Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile  
 450 455 460  
 atc aag aag tac ggc ctg ccc aac gac atc att atc gag ctg gcc cgc 1440  
 Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg  
 465 470 475 480  
 gag aag aac tcc aag gac gcc cag aaa atg atc aac gag atg cag aag 1488  
 Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys  
 485 490 495  
 cgg aac cgg cag acc aac gag cgg atc gag gaa atc atc cgg acc acc 1536  
 Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr  
 500 505 510  
 ggc aaa gag aac gcc aag tac ctg atc gag aag atc aag ctg cac gac 1584  
 Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp  
 515 520 525  
 atg cag gaa ggc aag tgc ctg tac agc ctg gaa gcc atc cct ctg gaa 1632  
 Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu  
 530 535 540  
 gat ctg ctg aac aac ccc ttc aac tat gag gtg gac cac atc atc ccc 1680  
 Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro  
 545 550 555 560  
 aga agc gtg tcc ttc gac aac agc ttc aac aac aag gtg ctc gtg aag 1728  
 Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys  
 565 570 575  
 cag gaa gaa gcc agc aag aag ggc aac cgg acc cca ttc cag tac ctg 1776  
 Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu  
 580 585 590  
 agc agc agc gac agc aag atc agc tac gaa acc ttc aag aag cac atc 1824  
 Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile  
 595 600 605  
 ctg aat ctg gcc aag ggc aag ggc aga atc agc aag acc aag aaa gag 1872  
 Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu  
 610 615 620  
 tat ctg ctg gaa gaa cgg gac atc aac agg ttc tcc gtg cag aaa gac 1920  
 Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp  
 625 630 635 640

ttc atc aac cgg aac ctg gtg gat acc aga tac gcc acc aga ggc ctg 1968  
 Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu  
                         645                                650                                655

atg aac ctg ctg cgg agc tac ttc aga gtg aac aac ctg gac gtg aaa 2016  
 Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys  
                         660                                665                                670

gtg aag tcc atc aat ggc ggc ttc acc agc ttt ctg cgg cgg aag tgg 2064  
 Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp  
                         675                                680                                685

aag ttt aag aaa gag cgg aac aag ggg tac aag cac cac gcc gag gac 2112  
 Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp  
                         690                                695                                700

gcc ctg atc att gcc aac gcc gat ttc atc ttc aaa gag tgg aag aaa 2160  
 Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys  
 705                                710                                715                                720

tcc ggc ggc ggt tcg acc ccc cac cag atc aag cac att aag gac ttc 2208  
 Ser Gly Gly Gly Ser Thr Pro His Gln Ile Lys His Ile Lys Asp Phe  
                                 725                                730                                735

aag gac tac aag tac agc cac cgg gtg gac aag aag cct aat aga gag 2256  
 Lys Asp Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu  
                         740                                745                                750

ctg att aac gac acc ctg tac tcc acc cgg aag gac gac aag ggc aac 2304  
 Leu Ile Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn  
                         755                                760                                765

acc ctg atc gtg aac aat ctg aac ggc ctg tac gac aag gac aat gac 2352  
 Thr Leu Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp  
                         770                                775                                780

aag ctg aaa aag ctg atc aac aag agc ccc gaa aag ctg ctg atg tac 2400  
 Lys Leu Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr  
 785                                790                                795                                800

cac cac gac ccc cag acc tac cag aaa ctg aag ctg att atg gaa cag 2448  
 His His Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln  
                                 805                                810                                815

tac ggc gac gag aag aat ccc ctg tac aag tac tac gag gaa acc ggg 2496  
 Tyr Gly Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly  
                         820                                825                                830

aac tac ctg acc aag tac tcc aaa aag gac aac ggc ccc gtg atc aag 2544  
 Asn Tyr Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys  
                         835                                840                                845

aag att aag tat tac ggc aac aaa ctg aac gcc cat ctg gac atc acc 2592  
Lys Ile Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr  
850 855 860  
gac gac tac ccc aac agc aga aac aag gtc gtg aag ctg tcc ctg aag 2640  
Asp Asp Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys  
865 870 875 880  
ccc tac aga ttc gac gtg tac ctg gac aat ggc gtg tac aag ttc gtg 2688  
Pro Tyr Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val  
885 890 895  
acc gtg aag aat ctg gat gtg atc aaa aaa gaa aac tac tac gaa gtg 2736  
Thr Val Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val  
900 905 910  
aat agc aag tgc tat gag gaa gct aag aag ctg aag aag atc agc aac 2784  
Asn Ser Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn  
915 920 925  
cag gcc gag ttt atc gcc tcc ttc tac aac aac gat ctg atc aag atc 2832  
Gln Ala Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile  
930 935 940  
aac ggc gag ctg tat aga gtg atc ggc gtg aac aac gac ctg ctg aac 2880  
Asn Gly Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn  
945 950 955 960  
cgg atc gaa gtg aac atg atc gac atc acc tac cgc gag tac ctg gaa 2928  
Arg Ile Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu  
965 970 975  
aac atg aac gac aag agg ccc ccc agg atc att aag aca atc gcc tcc 2976  
Asn Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser  
980 985 990  
aag acc cag agc att aag aag tac agc aca gac att ctg ggc aac ctg 3024  
Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu  
995 1000 1005  
tat gaa gtg aaa tct aag aag cac cct cag atc atc aaa aag ggc 3069  
Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly  
1010 1015 1020  
taa 3072  
<210> 6  
<211> 1023  
<212> PRT  
<213> 金黄色葡萄球菌 (Staphylococcus aureus)  
<400> 6

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



305	310	315	320
Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr			
	325	330	335
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala			
	340	345	350
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu			
	355	360	365
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser			
	370	375	380
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile			
385	390	395	400
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala			
	405	410	415
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln			
	420	425	430
Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro			
	435	440	445
Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile			
	450	455	460
Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg			
465	470	475	480
Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys			
	485	490	495
Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr			
	500	505	510
Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp			
	515	520	525
Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu			
	530	535	540
Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro			
545	550	555	560
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys			
	565	570	575
Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu			
	580	585	590
Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile			
	595	600	605
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu			
	610	615	620



930	935	940
Asn Gly Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn		
945	950	955
Arg Ile Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu		
	965	970
Asn Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser		
	980	985
Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu		
	995	1000
Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly		
1010	1015	1020
<210> 7		
<211> 5		
<212> PRT		
<213> 人工序列		
<220>		
<223> 接头		
<400> 7		
Ser Gly Gly Gly Ser		
1	5	
<210> 8		
<211> 6		
<212> PRT		
<213> 人工序列		
<220>		
<223> 接头		
<400> 8		
Gly Gly Ser Gly Gly Ser		
1	5	
<210> 9		
<211> 8		
<212> PRT		
<213> 人工序列		
<220>		
<223> 接头		
<400> 9		
Ser Gly Ser Gly Ser Gly Ser Gly		
1	5	
<210> 10		



---

guuuuaguac ucuggaaaca gaaucuacua aaacaaggca aaugccgug uuuaucacgu 60  
caacuuguug gcgagauuuu uu 83



图1

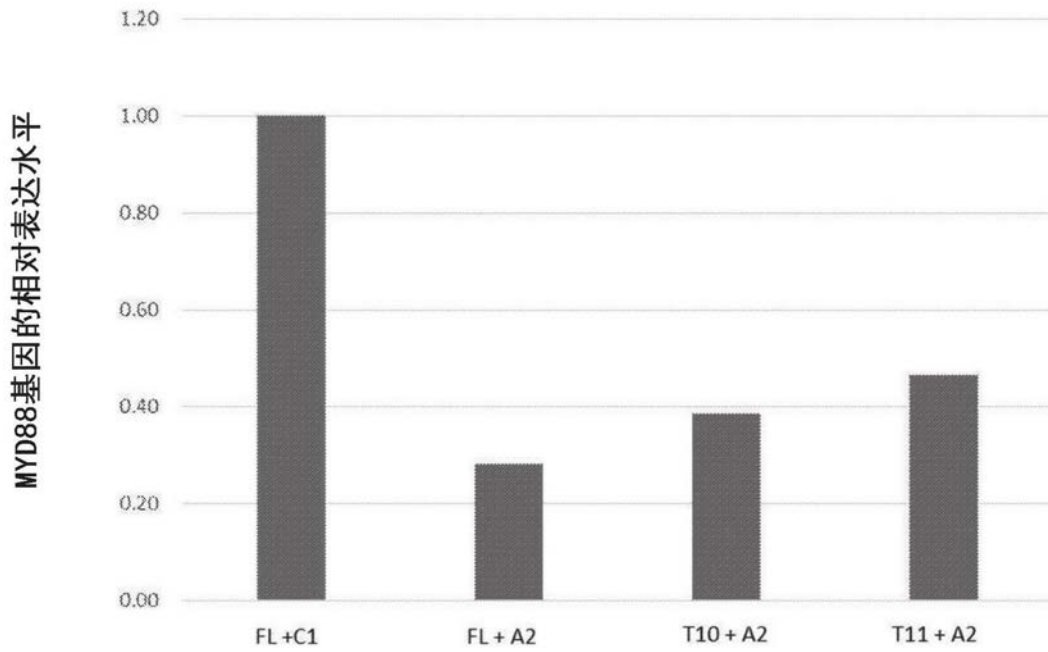


图2

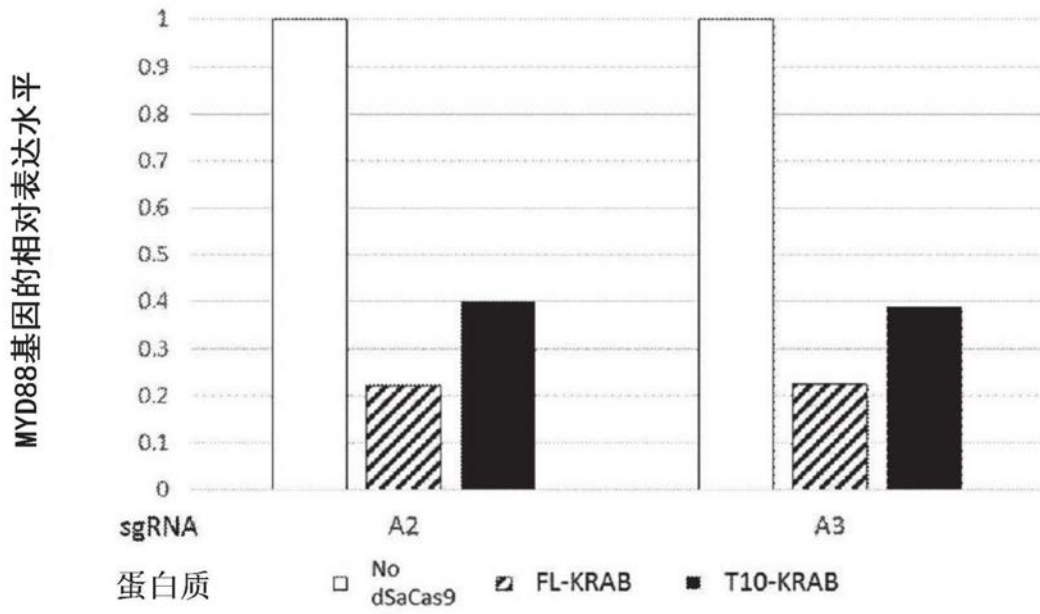


图3