(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116157144 A
(43)申请公布日 2023.05.23

(21)申请号 202180055113.X
(22)申请日 2021.07.15
(30)优先权数据
63/052,175 2020.07.15 US
(85)PCT国际申请进入国家阶段日
2023.03.07
(86)PCT国际申请的申请数据
PCT/US2021/041809 2021.07.15
(87)PCT国际申请的公布数据
WO2022/015969 EN 2022.01.20
(71)申请人 生命编辑制药股份有限公司
地址 美国北卡罗来纳
(72)发明人 T·D·博文 A·B·克拉维利
T·D·埃里驰
(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
专利代理师 谢丽菲
(51)Int.Cl.
A61K 38/43(2006.01)

权利要求书8页 说明书59页
序列表267页

(54)发明名称

尿嘧啶稳定蛋白及其活性片段和变体以及使用方法

(57)摘要

提供了包含用于核酸的靶向编辑的尿嘧啶稳定多肽的组合物和方法。组合物包含尿嘧啶稳定多肽。还提供了融合蛋白,其包含i)DNA结合多肽;ii)脱氨酶;和iii)尿嘧啶稳定多肽(USP)。融合蛋白包含与脱氨酶融合并进一步与USP融合的RNA引导的核酸酶,其任选地与引导RNA形成复合物。组合物还包含编码USP或融合蛋白的核酸分子。还提供了包含编码USP或融合蛋白的核酸分子的载体和宿主细胞。

CN 116157144 A

1.一种分离的多肽,其包含:

a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

其中所述多肽具有尿嘧啶稳定活性并且其中所述多肽还包含异源氨基酸序列。

2.权利要求1的分离的多肽,其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

3.一种药物组合物,其包含非天然存在的药学上可接受的承载体和多肽,所述多肽包含:

a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

其中所述多肽具有尿嘧啶稳定活性。

4.一种药物组合物,其包含非天然存在的药学上可接受的承载体和包含编码多肽的多核苷酸的核酸分子,所述多肽包含:

a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

其中所述多肽具有尿嘧啶稳定活性。

5.权利要求3或4的药物组合物,其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

6.权利要求3-5中任一项的药物组合物,其还包含氟嘧啶。

7.一种包含编码多肽的多核苷酸的核酸分子,所述多肽包含:

a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

其中所述多肽具有尿嘧啶稳定活性;和

其中所述核酸分子还包含可操作地连接至所述多核苷酸的异源启动子。

8.权利要求7的核酸分子,其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

9.一种包含氟嘧啶和多肽的组合物,所述多肽包含:

a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

其中所述多肽具有尿嘧啶稳定活性。

10.一种包含氟嘧啶和编码多肽的核酸分子的组合物,所述多肽包含:

a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

其中所述多肽具有尿嘧啶稳定活性。

11.权利要求9或10的组合物,其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

12.一种融合蛋白,其包含:(i)DNA结合多肽;(ii)脱氨酶;和(iii)至少一种尿嘧啶稳定多肽(USP),其与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

13.权利要求12的融合蛋白,其中所述USP具有SEQ ID NO:33-39中的任一个的序列。

14.权利要求12或13的融合蛋白,其中所述脱氨酶是胞苷脱氨酶。

15.权利要求14的融合蛋白,其中所述胞苷脱氨酶是活化诱导的胞苷脱氨酶(AID)或脱氨酶的载脂蛋白B mRNA编辑复合物(APOBEC)家族的成员。

16.权利要求15的融合蛋白,其中所述胞苷脱氨酶包含与SEQ IDNO:47、48和76-94中的任一个具有至少80%序列同一性的氨基酸序列。

17.权利要求12-16中任一项的融合蛋白,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

18.权利要求12-16中任一项的融合蛋白,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

19.权利要求18的融合蛋白,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

20.权利要求19的融合蛋白,其中所述RGN是II型CRISPR-Cas多肽。

21.权利要求19的融合蛋白,其中所述RGN是V型CRISPR-Cas多肽。

22.权利要求19的融合蛋白,其中所述RGN包含与SEQ ID NO:40和95-142中的任一个具有至少80%序列同一性的氨基酸序列。

23.权利要求19-22中任一项的融合蛋白,其中所述RGN是RGN切口酶。

24.权利要求19-23中任一项的融合蛋白,其中所述融合蛋白包含RGN、胞苷脱氨酶和USP。

25.权利要求24的融合蛋白,其中所述RGN与SEQ ID NO:40、41和95-142中的任一个具有至少80%的序列同一性,所述胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个具有至少80%的序列同一性,并且所述USP与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

26.权利要求12-25中任一项的融合蛋白,其中所述融合蛋白还包含至少一个核定位信号(NLS)。

27.一种包含编码融合蛋白的多核苷酸的核酸分子,所述融合蛋白包含:(i)DNA结合多肽;(ii)脱氨酶;和(iii)至少一种尿嘧啶稳定多肽(USP),其中USP由以下核苷酸序列编码:

a)与SEQ ID NO:17-32中的任一个具有至少80%序列同一性的核苷酸序列,

b)在SEQ ID NO:17-32中的任一个示出的核苷酸序列,

c)编码与SEQ ID NO:1-16至少80%相同的氨基酸序列并且进一步具有SEQ ID NO:33-39中的任一个的序列的核苷酸序列,

d)编码与SEQ ID NO:1-16中的任一个示出的氨基酸序列至少80%相同的氨基酸序列的核苷酸序列,或

e)编码SEQ ID NO:1-16中的任一个示出的氨基酸序列的核苷酸序列。

28.权利要求27的核酸分子,其中所述脱氨酶是胞苷脱氨酶。

29.权利要求28的核酸分子,其中所述胞苷脱氨酶是活化诱导的胞苷脱氨酶(AID)或脱氨酶的载脂蛋白B mRNA编辑复合物(APOBEC)家族的成员。

30.权利要求29的核酸分子,其中所述胞苷脱氨酶包含与SEQ IDNO:47、48和76-94中的任一个具有至少80%序列同一性的氨基酸序列。

31.权利要求27-30中任一项的核酸分子,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

32.权利要求27-30中任一项的核酸分子,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

33.权利要求32的核酸分子,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

34.权利要求33的核酸分子,其中所述RGN是II型CRISPR-Cas多肽。

35.权利要求33的核酸分子,其中所述RGN是V型CRISPR-Cas多肽。

36.权利要求33的核酸分子,其中所述RGN包含与SEQ ID NO:40和95-142中的任一个具有至少80%序列同一性的氨基酸序列。

37.权利要求33-36中任一项的核酸分子,其中所述RGN是RGN切口酶。

38.权利要求33-36中任一项的核酸分子,其中所述融合蛋白包含RGN、胞苷脱氨酶和USP。

39.权利要求38的核酸分子,其中所述RGN与SEQ ID NO:40、41和95-142中的任一个具有至少80%的序列同一性,所述胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个具有至少80%的序列同一性,并且所述USP与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

40.权利要求27-39中任一项的核酸分子,其中编码融合蛋白的多核苷酸在其5'端可操作地连接至异源启动子。

41.权利要求27-39中任一项的核酸分子,其中编码融合蛋白的多核苷酸在其3'端可操作地连接至异源终止子。

42.权利要求29-41中任一项的核酸分子,其中所述融合蛋白包含一个或多个核定位信号。

43.权利要求27-42中任一项的核酸分子,其中所述融合蛋白针对在真核细胞中的表达进行了密码子优化。

44.权利要求27-43中任一项的核酸分子,其中所述融合蛋白针对在原核细胞中的表达进行了密码子优化。

45.一种载体,其包含权利要求27-44中任一项的核酸分子。

46.一种包含权利要求33-39中任一项的核酸分子的载体,其还包含编码能够与靶序列杂交的引导RNA(gRNA)的至少一个核苷酸序列。

47.权利要求46的载体,其中所述gRNA是单引导RNA。

48.权利要求46的载体,其中所述gRNA是双引导RNA。

49.一种细胞,其包含权利要求12-26中任一项的融合蛋白。

50.一种细胞,其包含权利要求18-25中任一项的融合蛋白,其中所述细胞还包含引导RNA。

51.一种细胞,其包含权利要求27-44中任一项的核酸分子。

52.一种细胞,其包含权利要求45至48中任一项的载体。

53.权利要求49-52中任一项的细胞,其中所述细胞是原核细胞。

54.权利要求49-52中任一项的细胞,其中所述细胞是真核细胞。

55.权利要求54的细胞,其中所述细胞是昆虫、鸟类或哺乳动物细胞。

56.权利要求54的细胞,其中所述细胞是植物细胞或真菌细胞。

57.一种药物组合物,其包含药学上可接受的承载体和权利要求7、8、27-44中任一项的核酸分子、权利要求9-11中任一项的组合物、权利要求12-26中任一项的融合蛋白、权利要求45-48任一项的载体或权利要求49-55任一项的细胞。

58.一种制备融合蛋白的方法,包括在表达融合蛋白的条件下培养权利要求49-56中任一项的细胞。

59.一种制备融合蛋白的方法,包括将权利要求27-44中任一项的核酸分子或权利要求45-48中任一项的载体引入细胞中并在表达融合蛋白的条件下培养细胞。

60.权利要求58或59的方法,还包括纯化所述融合蛋白。

61.一种制备RGN融合核糖核蛋白复合物的方法,包括将权利要求33-39中任一项的核酸分子和包含编码引导RNA的表达盒的核酸分子或权利要求46-48中任一项的载体引入细胞中,以及在表达融合蛋白和gRNA并形成RGN融合核糖核蛋白复合物的条件下培养细胞。

62.权利要求61的方法,还包括纯化所述RGN融合核糖核蛋白复合物。

63.一种用于修饰包含靶DNA序列的靶DNA分子的系统,所述系统包括:

a)包含RNA引导的核酸酶多肽(RGN)、胞苷脱氨酶和至少一种尿嘧啶稳定多肽(USP)的融合蛋白,其中USP与SEQ ID NO:1-16中的任一个至少80%相同,或编码所述融合蛋白的核苷酸序列;和

b)能够与所述靶DNA序列杂交的一种或多种引导RNA或编码所述一种或多种引导RNA(gRNA)的一种或多种核苷酸序列;

其中编码一种或多种引导RNA和编码融合蛋白的所述核苷酸序列各自可操作地连接至与所述核苷酸序列异源的启动子;

和

其中所述一种或多种引导RNA能够与融合蛋白形成复合物以引导所述融合蛋白结合所述靶DNA序列并修饰靶DNA分子。

64.权利要求63的系统,其中所述靶DNA序列位于邻近被RGN识别的前间区序列邻近基序(PAM)的位置。

65.权利要求63或64的系统,其中所述靶DNA分子在细胞内。

66.权利要求65的系统,其中所述细胞是真核细胞。

67.权利要求66的系统,其中所述真核细胞是植物细胞。

68.权利要求66的系统,其中所述真核细胞是哺乳动物细胞。

69.权利要求66的系统,其中所述真核细胞是昆虫细胞。

70.权利要求65的系统,其中所述细胞是原核细胞。

71.权利要求63-70中任一项的系统,其中所述融合蛋白的RGN是II型CRISPR-Cas多肽。

72.权利要求63-70中任一项的系统,其中所述融合蛋白的RGN是V型CRISPR-Cas多肽。

73.权利要求63-70中任一项的系统,其中所述融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少80%相同。

74.权利要求63-73中任一项的系统,其中所述融合蛋白的RGN是RGN切口酶。

75.权利要求63-74中任一项的系统,其中所述胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少80%相同。

76.权利要求63-75中任一项的系统,其中所述USP包含SEQ IDNO:33-39中的任一个的序列。

77.权利要求63-76中任一项的系统,其中所述RGN与SEQ IDNO:40、41和95-142中的任一个具有至少80%的序列同一性,所述胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个具有至少80%的序列同一性,并且所述USP与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

78.权利要求63-77中任一项的系统,其中所述融合蛋白包含一个或多个核定位信号。

79.权利要求63-78中任一项的系统,其中所述融合蛋白针对在真核细胞中的表达进行了密码子优化。

80.权利要求63-79中任一项的系统,其中编码一种或多种引导RNA的核苷酸序列和编码融合蛋白的核苷酸序列位于一个载体上。

81.一种修饰包含靶DNA序列的靶DNA分子的方法,所述方法包括将根据权利要求63-80中任一项的系统递送至所述靶DNA分子或包含所述靶DNA分子的细胞。

82.权利要求81的方法,其中所述修饰的靶DNA分子包含靶DNA分子内至少一个核苷酸的C>T突变。

83.权利要求81的方法,其中所述修饰的靶DNA分子包含靶DNA序列内至少一个核苷酸的C>T突变。

84.一种修饰包含靶序列的靶DNA分子的方法,包括:

a)通过在适合形成RGN-脱氨酶-USP核糖核苷酸复合物的条件下组合以下i)和ii)在体外组装RGN-脱氨酶-USP核糖核苷酸复合物:

i)能够与靶DNA序列杂交的一种或多种引导RNA;和

ii)包含RNA引导的核酸酶多肽(RGN)、胞苷脱氨酶和至少一种尿嘧啶稳定多肽(USP)的融合蛋白,其中USP与SEQ ID NO:1-16中的任一个至少80%相同;和

b)使所述靶DNA分子或包含所述靶DNA分子的细胞与体外组装的RGN-脱氨酶-USP核糖核苷酸复合物接触;

其中所述一种或多种引导RNA与靶DNA序列杂交,从而指导所述融合蛋白结合所述靶DNA序列并且发生靶DNA分子的修饰。

85.权利要求84的方法,其中所述修饰的靶DNA分子包含靶DNA分子内至少一个核苷酸的C>T突变。

86.权利要求84的方法,其中所述修饰的靶DNA分子包含靶DNA序列内至少一个核苷酸的C>T突变。

87.权利要求84-86中任一项的方法,其中所述融合蛋白的RGN是II型CRISPR-Cas多肽。

88.权利要求84-86中任一项的方法,其中所述融合蛋白的RGN是V型CRISPR-Cas多肽。

89.权利要求84-86中任一项的方法,其中所述融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少80%相同。

90.权利要求84-89中任一项的方法,其中所述融合蛋白的RGN是RGN切口酶。

91.权利要求84-90中任一项的方法,其中所述胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少80%相同。

92.权利要求84-91中任一项的方法,其中所述USP包含SEQ IDNO:33-39中的任一个的序列。

93.权利要求84-92中任一项的方法,其中所述RGN与SEQ IDNO:40、41和95-142中的任一个具有至少80%的序列同一性,所述胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个具有至少80%的序列同一性,并且所述USP与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

94.权利要求84-93中任一项的方法,其中所述融合蛋白包含一个或多个核定位信号。

95.权利要求84-94中任一项的方法,其中所述融合蛋白针对在真核细胞中的表达进行了密码子优化。

96.权利要求84-95中任一项的方法,其中所述靶DNA序列位于邻近前间区序列邻近基序(PAM)的位置。

97.权利要求84-96中任一项的方法,其中所述靶DNA分子在细胞内。

98.权利要求97的方法,其中所述细胞是真核细胞。

99.权利要求98的方法,其中所述真核细胞是植物细胞。

100.权利要求98的方法,其中所述真核细胞是哺乳动物细胞。

101.权利要求98的方法,其中所述真核细胞是昆虫细胞。

102.权利要求97的方法,其中所述细胞是原核细胞。

103.权利要求97-102中任一项的方法,还包括选择包含所述修饰的DNA分子的细胞。

104.一种细胞,其包含根据权利要求103的方法修饰的靶DNA序列。

105.权利要求104的细胞,其中所述细胞是真核细胞。

106.权利要求105的细胞,其中所述真核细胞是植物细胞。

107.一种包含权利要求106的细胞的植物。

108.一种包含权利要求106的细胞的种子。

109.权利要求105的细胞,其中所述真核细胞是哺乳动物细胞。

110.权利要求105的细胞,其中所述真核细胞是昆虫细胞。

111.权利要求104的细胞,其中所述细胞是原核细胞。

112.一种产生具有遗传性疾病的致病突变的校正的遗传修饰的细胞的方法,所述方法包括向细胞中引入:

a)包含RNA引导的核酸酶多肽(RGN)、胞苷脱氨酶和至少一种尿嘧啶稳定多肽(USP)的融合蛋白,其中USP与SEQ ID NO:1-16中的任一个至少80%相同,或编码所述融合蛋白的多核苷酸,其中编码融合蛋白的所述多核苷酸可操作地连接至启动子以使融合蛋白能够在细胞中表达;和

b)能够与靶DNA序列杂交的一种或多种引导RNA(gRNA),或编码所述gRNA的多核苷酸,其中编码gRNA的所述多核苷酸可操作地连接至启动子以使gRNA能够在细胞中表达;

由此融合蛋白和gRNA靶向致病突变的基因组位置并修饰基因组序列以去除致病突变。

113.权利要求112的方法,其中所述融合蛋白的RGN是切口酶。

114.权利要求112或113的方法,其中所述USP包含SEQ ID NO:33-39中的任一个的序

列。

115.权利要求112-114中任一项的方法,其中基因组修饰包括在所述靶DNA序列内引入至少一个核苷酸的C＞T突变。

116.权利要求112-115中任一项的方法,其中所述细胞是动物细胞。

117.权利要求116的方法,其中所述动物细胞是哺乳动物细胞。

118.权利要求117的方法,其中所述细胞源自狗、猫、小鼠、大鼠、兔、马、绵羊、山羊、牛、猪或人。

119.权利要求112-118中任一项的方法,其中致病突变的校正包括引入终止密码子。

120.一种组合物,其包含:

a)融合蛋白,其包含:(i)DNA结合多肽;和(ii)脱氨酶;或编码所述融合蛋白的核酸分子;和

b)尿嘧啶稳定多肽(USP),其与SEQ ID NO:1-16中的任一个具有至少80%序列同一性;或编码所述USP的核酸分子。

121.权利要求120的组合物,其中所述融合蛋白还包含与SEQ IDNO:1-16中的任一个具有至少80%序列同一性的尿嘧啶稳定多肽(USP)。

122.权利要求120或121的组合物,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

123.权利要求120或121的组合物,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

124.权利要求123的组合物,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

125.权利要求124的组合物,其中所述RGN是RGN切口酶。

126.一种包含编码融合蛋白的核酸分子和编码尿嘧啶稳定多肽(USP)的核酸分子的载体,其中所述融合蛋白包含DNA结合多肽和脱氨酶,并且其中所述USP与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

127.权利要求126的载体,其中所述融合蛋白还包含与SEQ IDNO:1-16中的任一个具有至少80%序列同一性的尿嘧啶稳定多肽(USP)。

128.权利要求126或127的载体,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

129.权利要求126或127的载体,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

130.权利要求129的载体,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

131.权利要求130的载体,其中所述RGN是RGN切口酶。

132.一种细胞,其包含权利要求126-131中任一项的载体。

133.一种细胞,其包含:

a)融合蛋白,其包含:(i)DNA结合多肽;和(ii)脱氨酶;或编码所述融合蛋白的核酸分子;和

b)尿嘧啶稳定多肽(USP),其与SEQ ID NO:1-16中的任一个具有至少80%序列同一性;或编码所述USP的核酸分子。

134.权利要求133的细胞,其中所述融合蛋白还包含与SEQ IDNO:1-16中的任一个具有

至少80％序列同一性的尿嘧啶稳定多肽(USP)。

135.权利要求134的细胞,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

136.权利要求134的细胞,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

137.权利要求136的细胞,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

138.权利要求137的细胞,其中所述RGN是RGN切口酶。

尿嘧啶稳定蛋白及其活性片段和变体以及使用方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年7月15日提交的美国临时申请号63/052,175的优先权,其全部内容通过引用并入本文。

[0003] 关于序列表的声明

[0004] 与本申请相关的序列表以ASCII格式代替纸质副本提供,并在此通过引用并入说明书中。名为L103438_1220WO_0106_5_SL.txt的ASCII副本大小为658,586字节,创建于2021年7月14日,并通过EFS-Web以电子方式提交。

技术领域

[0005] 本发明涉及分子生物学和基因编辑领域。

背景技术

[0006] 靶向基因组编辑或修饰正迅速成为基础及应用研究的重要工具。最初的方法涉及诸如大范围核酸酶(meganuclease)、锌指融合蛋白或TALEN之类的工程化核酸酶,需要生成具有特异于每一种特定靶序列的工程化、可程序化、序列特异性的DNA结合结构域的嵌合核酸酶。RNA引导的核酸酶(RNG)(诸如CRISPR-Cas细菌系统的成簇规则间隔短回文重复序列(CRISPR)相关(Cas)蛋白)通过将核酸酶与引导RNA(其与特定靶序列特异性地杂交)复合而允许靶向特定序列。相较于为每个靶序列生成嵌合核酸酶,制造靶特异性的引导RNA的成本更低且更有效。这种RNA引导的核酸酶可用于通过序列特异性的、双链断裂(其经由易错的非同源末端连接(NHEJ)被修复)的引入来编辑基因组,以在特定的基因组位置引入突变。

[0007] 此外,RGN可用于靶向DNA编辑方法。靶向编辑(例如靶向切割)核酸序列以允许在基因组DNA中引入特定修饰使得能够进行高度细微的方法来研究基因功能和基因表达。这种靶向编辑也可用于靶向人的遗传性疾病或在农作物基因组中引入农艺学上有益的突变。基因组编辑工具的发展为基于基因编辑的哺乳动物疗法和农业生物技术提供了新方法。

发明内容

[0008] 提供了用于修饰靶DNA分子的组合物和方法。该组合物可用于修饰感兴趣的靶DNA分子。提供的组合物包含尿嘧啶稳定多肽。还提供了包含DNA结合多肽、脱氨酶多肽和尿嘧啶稳定多肽的融合蛋白。提供的组合物还包含编码尿嘧啶稳定多肽或融合蛋白的核酸分子,以及包含该核酸分子的载体和宿主细胞。本文所公开的方法涉及结合感兴趣的靶DNA分子内的感兴趣的靶序列并修饰感兴趣的靶DNA分子。

[0009] 发明详述

[0010] 受益于前述描述中呈现的教导内容,这些发明所属领域的技术人员将想到本文所示的本发明的许多修改和其他实施方案。因此,应当理解,本发明不限于所公开的特定实施方案,并且修改和其他实施方案旨在包括在所附权利要求的范围内。尽管本文使用了特定术语,但它们仅在一般和描述性意义上使用,并且不是出于限制目的。

[0011] I.概述

[0012] 本公开内容提供了尿嘧啶稳定多肽(USP)(其稳定DNA分子中的尿嘧啶残基)以及编码其的核酸分子。

[0013] Komor等人于2016年使用可操作地连接至修饰的RNA引导核酸酶(SpCas9)的胞嘧啶脱氨酶(rAPOBEC1)开发了靶向核碱基编辑(也称为碱基编辑)(Nature 533:420-424)。在Komor等人描述的系统中,引导RNA将rAPOBEC1-Cas9融合蛋白引导至靶DNA序列,其中rAPOBEC1将靶胞嘧啶(C)脱氨为尿嘧啶(U),尿嘧啶(U)具有胸腺嘧啶(T)的碱基配对特性。使用该系统,可以将靶向的C>T突变引入DNA分子中。

[0014] 在体内使用rAPOBEC1-Cas9融合进行碱基编辑的一个主要缺点是细胞尿嘧啶DNA糖基化酶(UDG)识别U:G异源双链DNA并催化从DNA中去除尿嘧啶以留下无碱基位点,从而启动碱基切除修复,其中U:G对恢复为C:G对是最常见的结果,尽管也观察到插入缺失(插入或缺失)形成。通过将尿嘧啶DNA糖基化酶抑制剂(UGI)并入到rAPOBEC1-Cas9融合蛋白上,尿嘧啶保持存在足够长的时间以发生复制并引入所需的C>T突变。

[0015] 本发明发现,通过稳定由脱氨基的胞嘧啶产生的尿嘧啶,可以防止无碱基位点的产生,并且更有可能引入所需的C>T突变。这是通过鉴定尿嘧啶稳定蛋白(也称为尿嘧啶稳定多肽或USP)来实现的。

[0016] 在一些实施方案中,USP作为融合蛋白的一部分提供,所述融合蛋白包含DNA结合多肽、脱氨酶多肽和尿嘧啶稳定多肽。在一些实施方案中,DNA结合多肽是或源自大范围核酸酶、锌指融合蛋白或TALEN。在一些实施方案中,DNA结合多肽是RNA引导的核酸酶,例如Cas9多肽,其结合引导RNA(也称为gRNA),其进而通过链杂交结合靶核酸序列。在其他实施方案中,单独提供USP。

[0017] 在一些实施方案中,脱氨酶多肽可以是可以使核碱基(例如胞苷)脱氨的脱氨酶结构域。脱氨酶对核碱基的脱氨可导致相应残基处的点突变,这在本文中被称为“核酸编辑”或“碱基编辑”。因此,包含RNA引导的核酸酶(RGN)多肽和脱氨酶多肽的融合蛋白可用于核酸序列的靶向编辑。

[0018] 此类融合蛋白可用于DNA的体外靶向编辑,例如用于突变细胞的产生。这些突变细胞可以在植物或动物中。此类融合蛋白也可用于引入靶向突变,例如,用于离体校正哺乳动物细胞中(例如,从受试者获得的随后再引入相同或另一受试者的细胞中)的遗传缺陷;以及用于体内引入靶向突变,例如,校正哺乳动物受试者中疾病相关基因中的遗传缺陷或在哺乳动物受试者中的疾病相关基因中引入失活突变。此类融合蛋白还可用于在植物细胞中引入靶向突变,例如,用于引入有益或农艺学上有价值的性状或等位基因。

[0019] 术语“蛋白质”、“肽”和“多肽”在本文中可互换使用,是指通过肽(酰胺)键连接在一起的氨基酸残基的聚合物。这些术语是指任何大小、结构或功能的蛋白质、肽或多肽。通常,蛋白质、肽或多肽的长度为至少三个氨基酸。蛋白质、肽或多肽可以指个体蛋白质或蛋白质的集合。蛋白质、肽或多肽中的一个或多个氨基酸可以被修饰,例如,通过添加化学实体如糖基、羟基、磷酸基、法呢基、异法呢基、脂肪酸基团、用于缀合的接头、功能化或其他修饰等。蛋白质、肽或多肽也可以是单个分子或可以是多分子复合物。蛋白质、肽或多肽可能只是天然存在的蛋白质或肽的片段。蛋白质、肽或多肽可以是天然存在的、重组的或合成的,或其任何组合。

[0020] 本文提供的任何蛋白质均可通过本领域已知的任何方法产生。例如，本文提供的蛋白质可以通过重组蛋白表达和纯化产生，这尤其适用于包含肽接头的融合蛋白。重组蛋白表达和纯化的方法是众所周知的，并且包括Green和Sambrook,Molecular Cloning:A Laboratory Manual(4th ed.,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.(2012))描述的那些，其全部内容通过引用并入本文。

[0021] II.尿嘧啶稳定蛋白

[0022] 本文公开了新型尿嘧啶稳定多肽(USP)，如SEQ ID NO:1-16所示。本文所述的USP可用于需要稳定DNA分子中的尿嘧啶的应用。

[0023] 如本文所用，术语“尿嘧啶稳定蛋白”、“尿嘧啶稳定多肽”和“USP”是指具有尿嘧啶稳定活性的多肽。如本文所用，术语“尿嘧啶稳定活性”是指与不存在分子(例如，尿嘧啶稳定多肽)时由脱氨酶引起的突变率相比，分子(例如，多肽)增加由脱氨酶引起的核酸分子中至少一个胞苷、脱氧胞苷或胞嘧啶向胸苷、脱氧胸苷或胸腺嘧啶的突变率的能力。不受理论或作用机制的束缚，据信，本文公开的USP可以通过维持经由胞苷、脱氧胞苷或胞嘧啶碱基的脱氨基作用生成的单链DNA中尿嘧啶存在持续足够的时间以允许复制发生和引入所需的C>T突变来发挥作用。尿嘧啶稳定活性可通过抑制尿嘧啶DNA糖基化酶、碱基切除修复途径或错配修复机制而发生。

[0024] 在一些实施方案中，与不存在USP时脱氨酶引起的突变率相比，本文公开的保留尿嘧啶稳定活性的USP或其活性变体或片段将由脱氨酶引起的核酸分子中至少一个胞苷、脱氧胞苷或胞嘧啶向胸苷、脱氧胸苷或胸腺嘧啶的突变率增加至少10％、至少15％、至少20％、至少25％、至少30％、至少35％、至少40％、至少45％、至少50％、至少55％、至少60％、至少70％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％、至少100％、至少150％、至少200％或更多。相反，与不存在USP时由脱氨酶引起的突变率相比，本文公开的USP或其活性变体或片段将由脱氨酶引起的核酸分子中至少一个胞苷、脱氧胞苷或胞嘧啶向除胸苷、脱氧胸苷或胸腺嘧啶以外的任何核碱基(即鸟苷、脱氧鸟苷、鸟嘌呤、腺苷、脱氧腺苷、腺嘌呤)的突变率减少至少10％、至少15％、至少20％、至少25％、至少30％、至少35％、至少40％、至少45％、至少50％、至少55％、至少60％、至少70％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％，或更多。胞苷、脱氧胞苷或胞嘧啶向另一种核碱基的突变率的增加或减少可以通过比较在存在或不存在USP的情况下特定脱氨酶对特定核碱基的突变率来测量。在其中脱氨酶通过与DNA结合多肽融合而靶向至核酸分子的特定区域的那些实施方案中，胞苷、脱氧胞苷或胞嘧啶在与DNA结合多肽结合的靶序列内或附近的突变率可以使用本领域已知的任何方法测量，包括聚合酶链式反应(PCR)、限制性片段长度多态性(RFLP)或DNA测序。

[0025] 本文公开的保留尿嘧啶稳定活性的新型USP或其活性变体或片段可作为脱氨酶-DNA结合多肽融合体的一部分引入细胞，和/或可与DNA结合多肽-脱氨酶融合体或与DNA结合多肽-脱氨酶-USP融合体共表达，以提高在靶DNA分子中引入所需C>T突变的效率。本文公开的保留尿嘧啶稳定活性的USP具有SEQ ID NO:1-16中的任一个的氨基酸序列或其变体或片段。在一些实施方案中，USP具有与SEQ ID NO:1-16中的任一个的氨基酸序列具有至少50％、至少55％、至少60％、至少65％、至少70％、至少75％、至少80％、至少81％、至少82％、至少83％、至少84％、至少85％、至少86％、至少87％、至少88％、至少89％、至少90％、至少

91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列。在具体实施方案中，USP包含与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80％序列同一性的氨基酸序列。在其他实施方案中，USP包含与SEQ ID NO:3或16具有至少81％序列同一性的氨基酸序列。在另外其他实施方案中，USP包含与SEQ ID NO:6具有至少82％序列同一性的氨基酸序列。

[0026] III.融合蛋白

[0027] 本公开内容的一些方面提供融合蛋白，其包含DNA结合多肽和脱氨酶多肽，并且在一些实施方案中，包含USP多肽。这种融合蛋白可用于DNA的体外、离体或体内靶向编辑。

[0028] 本文所用的术语“融合蛋白”是指包含来自至少两种不同蛋白质的蛋白质结构域的杂合多肽。融合蛋白可包含不同的结构域，例如DNA结合结构域和脱氨酶。在一些实施方案中，融合蛋白与核酸例如RNA复合或缔合。

[0029] 脱氨酶多肽包含可以使核碱基(例如胞苷)脱氨的脱氨酶结构域。脱氨酶对核碱基的脱氨作用可导致相应残基处的点突变，这在本文中被称为“核酸编辑”或“碱基编辑”。因此，包含RGN多肽变体或结构域和脱氨酶结构域的融合蛋白可用于核酸序列的靶向编辑。在一些实施方案中，脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个至少80％、至少81％、至少82％、至少83％、至少84％、至少85％、至少86％、至少87％、至少88％、至少89％、至少90％、至少91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％至少98％、至少99％或更多相同的氨基酸序列。在一些实施方案中，脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个至少80％相同的氨基酸序列。在一些实施方案中，脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个至少85％相同的氨基酸序列。在一些实施方案中，脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个至少90％相同的氨基酸序列。在一些实施方案中，脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个至少95％相同的氨基酸序列。在其他实施方案中，脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个至少99％相同的氨基酸序列。在一些特定的实施方案中，脱氨酶包含如SEQ ID NO:47、48和76-94中的任一个示出的氨基酸序列。

[0030] 本文公开的融合蛋白包含DNA结合多肽。如本文所用，术语“DNA结合多肽”是指能够结合DNA的任何多肽。在某些实施方案中，本文公开的融合蛋白的DNA结合多肽部分与双链DNA结合。在具体实施方案中，DNA结合多肽以序列特异性方式结合DNA。如本文所用，术语“序列特异性”或“序列特异性方式”是指与特定核苷酸序列的选择性相互作用。

[0031] 当两个多核苷酸序列在严格条件下彼此杂交时，可以认为该两个多核苷酸序列基本上互补。类似地，如果DNA结合多肽在严格条件下与其靶序列结合，则认为该DNA结合多肽以序列特异性方式结合特定靶序列。“严格条件”或“严格杂交条件”旨在指两个多核苷酸序列(或多肽结合其特定靶序列)将彼此结合至高于与其他序列结合的可检测程度(例如，至少是背景的2倍)的条件。严格条件是序列依赖性的，并且在不同环境下将会不同。通常情况下，严格条件将是这样的条件，其中：在pH 7.0至8.3下，盐类浓度小于约1.5M Na离子，通常情况下约0.01至1.0M Na离子浓度(或其它盐类)，并且针对短序列(例如，10至50个核苷酸)，温度为至少约30℃，而针对长序列(例如，大于50个核苷酸)，温度为至少约60℃。通过添加诸如甲酰胺的去稳定剂也可以达到严格条件。示例性的低严格条件包括在37℃下30至35％甲酰胺、1M NaCl、1％SDS(十二烷基硫酸钠)的缓冲溶液中进行杂交和在50至55℃下在

1X至2X SSC(20X SSC＝3.0M NaCl/0.3M柠檬酸三钠)中洗涤。示例性的中等严格条件包括在37℃下于40至45％甲酰胺、1.0M NaCl、1％SDS中杂交和在55至60℃下在0.5X至1X SSC中洗涤。示例性的高严格条件包括在37℃下于50％甲酰胺、1M NaCl、1％SDS中杂交和在60至65℃下在0.1X SSC中洗涤。任选地，洗涤缓冲液可包括约0.1％至约1％的SDS。杂交持续时间一般小于约24小时，通常约4至约12小时。洗涤时间的持续时间至少是足以达到平衡的时间长度。

[0032] Tm是50％的互补靶序列与完全匹配的序列杂交的温度(在规定的离子强度和pH下)。对于DNA-DNA杂交体，Tm可以由Meinkoth和Wahl(1984)Anal.Biochem.138:267-284中的等式:Tm＝81.5℃+16.6(log M)+0.41(％GC)-0.61(％form)-500/L大致估计;其中M是单价阳离子的摩尔浓度，％GC是鸟苷和胞嘧啶核苷酸在DNA中的百分比，％form是杂交溶液中甲酰胺的百分比，以及L是碱基对中的杂交体的长度。一般而言，严格条件被选择为比在规定的离子强度和pH下的特定序列及其补体的热熔点(Tm)低约5℃。然而，极度严格条件可利用在比热熔点(Tm)低1、2、3或4℃的温度下的杂交和/或洗涤;中等严格条件可利用在比热熔点(Tm)低6、7、8、9或10℃的温度下的杂交和/或洗涤;低严格条件可利用在比热熔点(Tm)低11、12、13、14、15或20℃的温度下的杂交和/或洗涤。使用该等式、杂交和洗涤组合物以及要求的Tm，本领域技术人员将明白固有地描述了杂交和/或洗涤溶液的严格性的变化。核酸杂交的广泛指南可在Tijssen(1993)Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes,Part I,Chapter 2(Elsevier,New York);和Ausubel等人,eds.(1995)Current Protocols in Molecular Biology,Chapter 2(Greene Publishing and Wiley-Interscience,New York)中得到。参见Sambrook等人(1989)Molecular Cloning:A Laboratory Manual(2d ed.,Cold Spring Harbor Laboratory Press,Plainview,New York)。

[0033] 在某些实施方案中，序列特异性DNA结合多肽是RNA引导的DNA结合多肽(RGDBP)。如本文所用，术语“RNA引导的DNA结合多肽”和“RGDBP”是指能够通过相关RNA分子与靶DNA序列的杂交而结合DNA的多肽。

[0034] 在一些实施方案中，融合蛋白的DNA结合多肽是核酸酶，例如序列特异性核酸酶。如本文所用，术语“核酸酶”是指催化核酸分子中核苷酸之间磷酸二酯键断裂的酶。在一些实施方案中，DNA结合多肽是核酸内切酶，其能够裂解核酸分子内核苷酸之间的磷酸二酯键，而在其他实施方案中，DNA结合多肽是能够在核酸分子的任一末端(5'或3')切割核苷酸的核酸外切酶。在一些实施方案中，序列特异性核酸酶选自大范围核酸酶、锌指核酸酶、TAL-效应物DNA结合结构域-核酸酶融合蛋白(TALEN)和RNA引导的核酸酶(RGN)或其在其中核酸酶活性已被降低或抑制的变体。

[0035] 如本文所用，术语“大范围核酸酶”或“归巢核酸内切酶”是指结合双链DNA内12至40bp长度的识别位点的核酸内切酶。大范围核酸酶的非限制性实例是属于LAGLIDADG家族的那些，其包含保守氨基酸基序LAGLIDADG(SEQ ID NO:75)。术语“大范围核酸酶”可以指二聚体或单链大范围核酸酶。

[0036] 如本文所用，术语“锌指核酸酶”或“ZFN”是指包含锌指DNA结合结构域和核酸酶结构域的嵌合蛋白。

[0037] 如本文所用，术语“TAL-效应物DNA结合结构域-核酸酶融合蛋白”或“TALEN”是指

包含TAL效应物DNA结合结构域和核酸酶结构域的嵌合蛋白。

[0038] 如本文所用,术语“RNA引导的核酸酶”或“RGN”是指具有核酸酶活性的RNA引导的DNA结合多肽。RGN被认为是“RNA引导的”,因为引导RNA与RNA引导的核酸酶形成复合物以指导RNA引导的核酸酶结合靶序列和在一些实施方案中在靶序列处引入单链或双链断裂。

[0039] 可用于本公开内容的组合物和方法的RGN的非限制性实例包括在公开号WO2020/139783、WO2019/236566、WO2021/030344、WO/2021/138247和分别于2021年4月23日和2021年5月11日提交的申请号PCT/US2021/028843和PCT/US2021/031794中公开的那些,其各自通过引用整体并入本文。在一些实施方案中,本文公开的融合蛋白包含RGN,其包含与SEQ ID NO:40和95-142中的任一个至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更多相同的氨基酸序列。在一些实施方案中,本文公开的融合蛋白包含具有与SEQ ID NO:40和95-142中的任一个至少80%相同的氨基酸序列的RGN。在一些实施方案中,本文公开的融合蛋白包含具有与SEQ ID NO:40和95-142中的任一个至少85%相同的氨基酸序列的RGN。在一些实施方案中,本文公开的融合蛋白包含具有与SEQ ID NO:40和95-142中的任一个至少90%相同的氨基酸序列的RGN。在一些实施方案中,本文公开的融合蛋白包含具有与SEQ ID NO:40和95-142中的任一个至少95%相同的氨基酸序列的RGN。在其他实施方案中,本文公开的融合蛋白包含具有与SEQ ID NO:40和95-142中的任一个至少99%相同的氨基酸序列的RGN。在一些具体的实施方案中,本文公开的融合蛋白包含具有如SEQ ID NO:40和95-142中的任一个示出的氨基酸序列的RGN。

[0040] 根据本发明,已突变为核酸酶失活的或“死”的RGN蛋白(例如dCas9)在本文中被称为RNA引导的DNA结合多肽。一种示例性的合适的核酸酶失活的Cas9结构域是D10A/H840A Cas9结构域突变体(参见,例如,Qi等人,Cell.2013;152(5):1173-83,其全部内容通过引用并入本文)。另外,可以确定其他已知RNA引导的核酸酶(RGN)的合适的核酸酶失活的Cas9结构域(例如,美国专利公开号2019/0367949中公开的RGN APG08290.1的核酸酶失活变体,其全部内容通过引用并入本文)。

[0041] 术语“RGN多肽”涵盖仅切割靶核苷酸序列的单条链的RGN多肽,其在本文中被称为切口酶。此类RGN具有单一功能性核酸酶结构域。RGN切口酶可以是天然存在的切口酶或可以是已在额外的核酸酶结构域内发生突变以使得这些突变结构域的核酸酶活性降低或消除从而成为切口酶的天然切割双链核酸分子的两条链的RGN蛋白。在一些实施方案中,融合蛋白的切口酶RGN包含D10A突变(例如nAPG07433.1(SEQ ID NO:41)),这使得RGN能够仅切割核酸双链体的非碱基编辑的靶标链(包含PAM并与gRNA碱基配对的链)。在一些实施方案中,融合蛋白的切口酶RGN在SEQ ID NO:40和95-142中的任一个中包含D10A突变或其等效突变。在一些实施方案中,融合蛋白的切口酶RGN包含H840A突变,这使得RGN能够仅切割核酸双链体的碱基编辑的非靶标链(不包含PAM且未与gRNA碱基配对的链)。包含H840A突变或等效突变的切口酶RGN具有失活的HNH结构域。包含D10A突变或等效突变的切口酶RGN具有失活的RuvC结构域。脱氨酶作用于非靶标链。包含D10A突变或等效突变的切口酶具有无活性的RuvC核酸酶结构域并且不能切割DNA的非靶标链,即需要进行碱基编辑的链。

[0042] 其他额外的示例性合适的核酸酶失活Cas9结构域包括但不限于D10A/D839A/

H840A和D10A/D839A/H840A/N863A突变结构域(参见,例如Mali等人,Nature Biotechnology.2013;31(9):833-838,其全部内容通过引用并入本文)。基于本公开内容和本领域的知识,被突变为切口酶的另外的合适的RGN蛋白对于本领域技术人员来说将是明显的(例如PCT公开号WO2019/236566中公开的RGN)并且在本公开内容的范围内。

[0043] 在一些实施方案中,保留切口酶活性的RGN切口酶包含与nAPG07433.1(SEQ ID NO:41)至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%相同的氨基酸序列。

[0044] 本领域已知的用于将突变引入氨基酸序列的任何方法(例如PCR介导的诱变和定点诱变)可用于产生切口酶或核酸酶死亡的RGN。参见,例如,美国公开号2014/0068797和美国专利号9,790,490;其各自通过引用整体并入本文。RNA引导的核酸酶(RGN)允许对基因组内的单个位点进行靶向操作,并且在基因靶向背景下可用于治疗和研究应用。在包括哺乳动物在内的多种生物体中,RNA引导的核酸酶已被用于通过刺激非同源末端连接或同源重组来进行基因组工程化。RGN包括CRISPR-Cas蛋白(其是作为成簇规则间隔短回文重复序列(CRISPR)RNA-引导核酸酶系统的一部分被引导RNA(gRNA)导向到靶序列的RNA引导的核酸酶)或其活性变体或片段。

[0045] 本公开内容的一些方面提供包含RNA引导的DNA结合多肽、脱氨酶多肽和USP的融合蛋白。在一些实施方案中,RNA引导的DNA结合多肽是RNA引导的核酸酶。在进一步的实施方案中,RNA引导的核酸酶是天然存在的CRISPR-Cas蛋白或其活性变体或片段。CRISPR-Cas系统分类为I类或II类系统。II类系统包含单一效应核酸酶,并且包括II、V和VI型。每个类别又细分为类型(I、II、III、IV、V、VI型),其中一些类型进一步分为亚型(例如,II-A型、II-B型、II-C型、V-A型、V-B型)。

[0046] 在某些实施方案中,CRISPR-Cas蛋白是天然存在的II型CRISPR-Cas蛋白或其活性变体或片段。如本文所用,术语“II型CRISPR-Cas蛋白”、“II型CRISPR-Cas效应蛋白”或“Cas9”是指需要反式激活RNA(tracrRNA)并且包含两个核酸酶结构域(RuvC和HNH)(其各自负责切割双链DNA分子的单条链)的CRISPR-Cas效应蛋白。

[0047] 在其他实施方案中,CRISPR-Cas蛋白是天然存在的V型CRISPR-Cas蛋白或其活性变体或片段。如本文所用,术语“V型CRISPR-Cas蛋白”、“V型CRISPR-Cas效应蛋白”或“Cas12”是指切割dsDNA并包含单个RuvC核酸酶结构域或分裂的RuvC核酸酶结构域并且缺少HNH结构域的CRISPR-Cas效应蛋白(Zetsche等人2015,Cell doi:10.1016/j.cell.2015.09.038;Shmakov等人2017,Nat Rev Microbiol doi:10.1038/nrmicro.2016.184;Yan等人2018,Science doi:10.1126/science.aav7271;Harrington等人2018,Science doi:10.1126/science.aav4294)。需要注意的是,Cas12a也被称为Cpf1,并且不需要tracrRNA,尽管其他V型CRISPR-Cas蛋白(例如Cas12b)确实需要tracrRNA。大多数V型效应子还可以靶向ssDNA(单链DNA),通常不需要PAM(Zetsche等人2015;Yan等人2018;Harrington等人2018)。术语“V型CRISPR-Cas蛋白”涵盖包含分裂的RuvC核酸酶结构域的独特RGN,例如2019年12月30日提交的美国临时申请号62/955,014中公开的那些,其内容通过引用整体并入本文。

[0048] 在其他实施方案中,CRISPR-Cas蛋白是天然存在的VI型CRISPR-Cas蛋白或其活性变体或片段。如本文所用,术语“VI型CRISPR-Cas蛋白”、“VI型CRISPR-Cas效应蛋白”或

“Cas13”是指不需要tracrRNA且包含两个切割RNA的HEPN结构域的CRISPR-Cas效应蛋白。术语“引导RNA”是指与靶核苷酸序列具有足够互补性以与靶序列杂交并指导相关RGN与靶核苷酸序列的序列特异性结合的核苷酸序列。对于CRISPR-Cas RGN,相应的引导RNA是一个或多个RNA分子(通常是一个或两个),它们可以结合RGN并指导RGN结合特定的靶核苷酸序列,并且在其中RGN具有切口酶或核酸酶活性的那些情况下还切割靶核苷酸序列。引导RNA包含CRISPR RNA(crRNA),并且在一些实施方案中,包含反式激活CRISPR RNA(tracrRNA)。

[0049] CRISPR RNA包含间隔序列和CRISPR重复序列。“间隔序列”是直接与感兴趣的靶核苷酸序列杂交的核苷酸序列。间隔序列被工程化为与感兴趣的靶序列完全或部分互补。在各种实施方案中,间隔序列可包含约8个核苷酸至约30个核苷酸或更多。例如,间隔序列的长度可以是约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30或更多个核苷酸。在一些实施方案中,间隔序列的长度为约10至约26个核苷酸,或长度为约12至约30个核苷酸。在特定的实施方案中,间隔序列的长度为约30个核苷酸。在一些实施方案中,当使用合适的比对算法进行最佳比对时,间隔序列与其相应靶序列之间的互补程度为约或超过约50%、约60%、约70%、约75%、约80%、约81%、约82%、约83%、约84%、约85%、约86%、约87%、约88%、约89%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或更高。在特定的实施方案中,间隔序列不含二级结构,其可以使用本领域已知的任何合适的多核苷酸折叠算法预测,包括但不限于mFold(参见,例如,Zuker和Stiegler(1981)Nucleic Acids Res.9:133-148)和RNAfold(参见,例如,Gruber等人(2008)Cell106(1):23-24)。

[0050] CRISPR RNA重复序列包含本身或与杂交的tracrRNA协同形成被RGN分子识别的结构的核苷酸序列。在各种实施方案中,CRISPR RNA重复序列可包含约8个核苷酸至约30个核苷酸或更多。例如,CRISPR重复序列的长度可以是约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30或更多个核苷酸。在一些实施方案中,当使用合适的比对算法进行最佳比对时,CRISPR重复序列与其对应的tracrRNA序列之间的互补程度为约或超过约50%、约60%、约70%、约75%、约80%、约81%、约82%、约83%、约84%、约85%、约86%、约87%、约88%、约89%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或更高。

[0051] 在一些实施方案中,引导RNA还包含tracrRNA分子。反式激活CRISPR RNA或tracrRNA分子包含一个核苷酸序列,该核苷酸序列包含一个区域,该区域具有足够的互补性以与crRNA的CRISPR重复序列杂交,该区域在本文中被称为抗重复序列区域。在一些实施方案中,tracrRNA分子进一步包含具有二级结构(例如茎环)的区域或在与其对应的crRNA杂交时形成二级结构。在特定的实施方案中,与CRISPR重复序列完全或部分互补的tracrRNA区域位于分子的5’端并且tracrRNA的3’端包含二级结构。该二级结构区域通常包含几个发夹结构,包括与抗重复序列相邻的连接发夹。tracrRNA的3’端通常有末端发夹,其结构和数量可能会有所不同,但通常包含富含GC的Rho独立转录终止子发夹,随后是在3’端的一串U。参见,例如,Briner等人(2014)Molecular Cell 56:333-339、Briner和Barrangou(2016)Cold Spring Harb Protoc;doi:10.1101/pdb.top090902和美国公开号2017/0275648,其各自通过引用整体并入本文。

[0052] 在各种实施方案中,与CRISPR重复序列完全或部分互补的tracrRNA的抗重复序列

区域包含约6个核苷酸至约30个核苷酸或更多。例如，tracrRNA抗重复序列和CRISPR重复序列之间的碱基配对区域的长度可以是约6、约7、约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30或更多个核苷酸。在特定的实施方案中，与CRISPR重复序列完全或部分互补的tracrRNA的抗重复序列区域的长度为约10个核苷酸。在一些实施方案中，当使用合适的比对算法进行最佳比对时，CRISPR重复序列与其对应的tracrRNA抗重复序列之间的互补程度为约或大于约50％、约60％、约70％、约75％、约80％、约81％、约82％、约83％、约84％、约85％、约86％、约87％、约88％、约89％、约90％、约91％、约92％、约93％、约94％、约95％、约96％、约97％、约98％、约99％或更高。

[0053] 在各种实施方案中，整个tracrRNA可包含约60个核苷酸至多于约210个核苷酸。例如，tracrRNA的长度可以是约60、约65、约70、约75、约80、约85、约90、约95、约100、约105、约110、约115、约120、约125、约130、约135、约140、约150、约160、约170、约180、约190、约200、约210或更多个核苷酸。在具体实施方案中，tracrRNA的长度为约100至约201个核苷酸，包括约95、约96、约97、约98、约99、约100、约105、约106、约107、约108、约109和约100个核苷酸的长度。

[0054] 引导RNA与RNA引导的核酸酶形成复合物以指导RNA引导的核酸酶结合靶序列并在靶序列处引入单链或双链断裂。在靶序列被切割后，断裂可以被修复，使得靶序列的DNA序列在修复过程中被修饰。本文提供了使用RNA引导的核酸酶的突变变体的方法，这些突变变体是无活性的核酸酶或切口酶，它们与脱氨酶连接以修饰宿主细胞DNA中的靶序列。其中核酸酶活性被灭活或显著降低的RNA引导的核酸酶的突变变体可被称为RNA引导的DNA结合多肽，因为该多肽能够结合但不一定切割靶序列。仅能够切割双链核酸分子的单条链的RNA引导的核酸酶在本文中被称为切口酶。

[0055] 靶核苷酸序列被RGN结合并与与RGN缔合的引导RNA杂交。如果多肽具有核酸酶活性(包括作为切口酶的活性)，则靶序列随后可以被RGN切割。

[0056] 引导RNA可以是单引导RNA或双引导RNA系统。单引导RNA包括单一RNA分子上的crRNA和任选tracrRNA，而双引导RNA系统包括存在于两个不同RNA分子上的crRNA和tracrRNA，crRNA和tracrRNA通过crRNA的CRISPR重复序列的至少一部分和tracrRNA的可与crRNA的CRISPR重复序列完全或部分互补的至少一部分而彼此杂交。在其中引导RNA是单引导RNA的那些实施方案的一些实施方案中，crRNA和任选tracrRNA被接头核苷酸序列分开。

[0057] 一般来说，接头核苷酸序列是不包含互补碱基的核苷酸序列，以避免在该接头核苷酸序列的核苷酸内形成二级结构或形成包含该接头核苷酸序列的核苷酸的二级结构。在一些实施方案中，crRNA与tracrRNA之间的接头核苷酸序列的长度是至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12个或更多个核苷酸。在具体实施方案中，单引导RNA的接头核苷酸序列的长度是至少4个核苷酸。

[0058] 在某些实施方案中，引导RNA可作为RNA分子而被引入靶细胞、细胞器或胚胎中。引导RNA可以在体外被转录或被化学合成。在其他实施方案中，将编码引导RNA的核苷酸序列引入细胞、细胞器或胚胎中。在这些实施方案的一些实施方案中，编码引导RNA的核苷酸序列可操作地连接至启动子(例如，RNA聚合酶III启动子)。启动子可以是天然启动子或与引导RNA编码的核苷酸序列异源。

[0059] 在各种实施方案中，如本文所述，引导RNA可作为核糖核蛋白复合物而被引入靶细胞、细胞器或胚胎中，其中引导RNA与RNA引导的核酸酶多肽结合。

[0060] 引导RNA通过该引导RNA与靶核苷酸序列的杂交将缔合的RNA引导的核酸酶导向至感兴趣的特定靶核苷酸序列。靶核苷酸序列可包括DNA、RNA或两者的组合，并且可以是单链或双链的。靶核苷酸序列可以是基因组DNA(即，染色体DNA)、质粒DNA或RNA分子(例如，信使RNA、核糖体RNA、转移RNA、微RNA、小干扰RNA)。靶核苷酸序列可在体外或在细胞中被RNA引导的核酸酶结合(并且在一些实施方案中，被切割)。被RGN靶向的染色体序列可以是细胞核、质体或线粒体染色体序列。在一些实施方案中，靶核苷酸序列在靶基因组中是独特的。

[0061] 在一些实施方案中，靶核苷酸序列与前间区序列邻近基序(protospacer adjacent motif，PAM)相邻。PAM一般在距靶核苷酸序列约1至约10个核苷酸内，包括距靶核苷酸序列约1、约2、约3、约4、约5、约6、约7、约8、约9或约10个核苷酸。PAM可以在靶序列的5’或3’。在一些实施方案中，PAM在靶序列的3’。一般而言，PAM是约2-6个核苷酸的共有序列，但在具体实施方案中，PAM的长度可以是1、2、3、4、5、6、7、8、9或更多个核苷酸。

[0062] 在识别其对应的PAM序列时，RGN可在特定切割位点切割靶核苷酸序列。如本文所使用的，切割位点由靶核苷酸序列内的两个特定核苷酸组成，该两个特定核苷酸之间的核苷酸序列被RGN切割。切割位点可以包括在5’或3’方向上距离PAM的第1和第2、第2和第3、第3和第4、第4和第5、第5和第6、第7和第8或第8和第9个核苷酸。由于RGN可切割靶核苷酸序列从而导致交错的末端，所以在一些实施方案中，基于多核苷酸的正(+)链上的两个核苷酸距离PAM的距离和多核苷酸的负(-)链上的两个核苷酸距离PAM的距离来定义切割位点。

[0063] RGN可用于将融合多肽、多核苷酸或小分子有效载荷递送至特定基因组位置。在一些实施方案中，无活性的核酸酶或切口酶RGN可操作地连接至脱氨酶以及本发明的USP。

[0064] 如本文所用，术语“脱氨酶”或“脱氨酶多肽”是指催化脱氨反应的多肽。脱氨酶可以是天然存在的脱氨酶或其活性片段或变体。在一些实施方案中，脱氨酶是胞苷脱氨酶，其催化胞苷或脱氧胞苷分别水解脱氨为尿嘧啶或脱氧尿嘧啶。胞苷脱氨酶可作用于DNA或RNA，通常作用于单链核酸分子。在优选的实施方案中，在靶标链上具有切口酶活性的RGN在靶标链上切出切口，而互补的非靶标链被脱氨酶修饰。细胞DNA修复机制可以使用修饰的非靶标链作为模板修复带切口的靶标链，从而在DNA中引入突变。

[0065] 在一些实施方案中，胞苷脱氨酶是载脂蛋白B mRNA编辑复合物(APOBEC)家族脱氨酶。在这些实施方案的一些中，脱氨酶是APOBEC1家族脱氨酶。在一些实施方案中，胞苷脱氨酶是活化诱导的胞苷脱氨酶(AID)。在一些实施方案中，脱氨酶是ACF1/ASE脱氨酶。在某些实施方案中，脱氨酶是腺苷脱氨酶。在这些实施方案的一些中，脱氨酶是ADAT家族脱氨酶。基于本公开内容，其他合适的脱氨酶或结构域对本领域技术人员而言将是明显的。

[0066] 一种示例性合适类型的脱氨酶多肽是胞苷脱氨酶，例如APOBEC家族的脱氨酶。胞嘧啶脱氨酶的载脂蛋白B mRNA编辑复合物(APOBEC)家族包含11种蛋白质，它们用于以受控和有益的方式启动诱变(Conticello等人，2008.Genome Biology，9(6):229)。一个家族成员(活化诱导的胞苷脱氨酶(AID))通过以转录依赖性链偏向方式将ssDNA中的胞嘧啶转化为尿嘧啶来引起抗体的成熟(Reynaud等人，2003.Nature Immunology，4(7):631-638)。载脂蛋白B编辑复合物3(APOBEC3)酶通过逆转录病毒ssDNA中胞嘧啶的脱氨基作用为人类细胞提供针对某种HIV-1毒株的保护(Bhagwat等人，2004.DNA Repair(Amst)，3(1):85-9)。这

些蛋白质都需要Zn^{2+}配位基序(HisX-Glu-X_{23-26}-Pro-Cys-X_{2-4}-Cys)和结合的水分子以实现催化活性。Glu残基用于将水分子活化至氢氧化锌以用于在脱氨反应中的亲核攻击。每个家族成员优先在其自己特定的“热点”处脱氨基,范围从hAID的WRC(其中W是A或T,R是A或G)到hAPOBEC3F的TTC(Navaratnam等人,2006.Intl J Hematol 83(3):195-200)。APOBEC3G的催化结构域的最新晶体结构揭示了一个二级结构,该二级结构由侧接六个α螺旋的五链β折叠核心组成,据信其在整个家族中是保守的(Holden等人,2008.Nature 456(7218):121-124)。活性中心环已显示负责ssDNA结合和确定“热点”身份(Chelico等人,2009.J Biol Chem 284(41):27761-27765)。这些酶的过表达与基因组不稳定性和癌症有关,因此突出了序列特异性靶向的重要性(Pham等人,2005.Biochem 44(8):2703-2715)。在一些实施方案中,脱氨酶多肽可以是可使胞苷脱氨基以产生尿嘧啶的脱氨酶多肽。脱氨酶对核碱基的脱氨作用可导致相应残基的点突变,从而修饰DNA分子。这种修饰行为在本文中也被称为核酸编辑或碱基编辑。因此,包含Cas9变体或结构域、脱氨酶结构域和USP结构域的融合蛋白可用于核酸序列的靶向编辑。

[0067] 在一些实施方案中,本发明的与脱氨酶和USP融合的核酸酶失活的RGN或切口酶RGN可以被靶向至核酸分子(即,靶核酸分子)的特定位置(其在一些实施方案中是特定的基因组位点),以改变所需序列的表达。在一些实施方案中,融合蛋白与靶序列的结合导致核苷酸碱基的脱氨基,导致从一种核苷酸碱基转化为另一种。在一些实施方案中,该融合蛋白与靶序列的结合导致与靶序列相邻的核苷酸碱基的脱氨基。使用本文公开的组合物和方法脱氨基和突变的与靶序列相邻的核苷酸碱基可以是距离靶核酸分子内的靶序列(由gRNA结合)的5’或3’端的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个碱基对。本公开内容的一些方面提供了融合蛋白,其包含(i)核酸酶失活的或切口酶RGN多肽;(ii)脱氨酶多肽;和(iii)尿嘧啶稳定多肽。

[0068] 本公开内容提供了各种构型的融合蛋白。在一些实施方案中,脱氨酶多肽融合至RGN多肽的N端。在一些实施方案中,脱氨酶多肽融合至RGN多肽的C端。

[0069] 在一些实施方案中,USP结构域、脱氨酶结构域和RNA引导的DNA结合多肽通过接头彼此融合。在融合蛋白的三个功能结构域之间可以使用各种接头长度和灵活性(例如,范围从$(GGGGS)_n$和$(G)_n$形式的非常灵活的接头到$(EAAAK)_n$和$(XP)_n$形式的更刚性的接头,以便针对特定应用实现脱氨酶活性的最佳长度。如本文所用,术语“接头”指连接两个分子或部分(例如核酸酶的结合结构域和切割结构域)的化学基团或分子。在一些实施方案中,接头连接RNA引导的核酸酶和脱氨酶。在一些实施方案中,接头连接dCas9和脱氨酶。通常,接头位于两个基团、分子或其他部分之间或两侧,并通过共价键连接至每个基团、分子或其他部分,从而连接两者。在一些实施方案中,接头是氨基酸或多个氨基酸(例如,肽或蛋白质)。在一些实施方案中,接头是有机分子、基团、聚合物或化学部分。在一些实施方案中,接头长度为5-100个氨基酸,例如长度为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、30-35、35-40、40-45、45-50、50-60、60-70、70-80、80-90、90-100、100-150或150-200个氨基酸。也考虑更长或更短的接头。

[0070] 在一些实施方案中,接头包含$(GGGGS)_n$、$(G)_n$、$(EAAAK)_n$或$(XP)_n$基序,或任何这些的组合,其中n独立地是1至30的整数。在一些实施方案中,n独立地是1、2、3、4、5、6、7、8、9、

10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30，或者，如果存在多于一个接头或多于一个接头基序，其任何组合。其他合适的接头基序和接头构型对本领域技术人员而言将是明显的。在一些实施方案中，合适的接头基序和构型包括Chen等人，2013(Adv Drug Deliv Rev.65(10):1357-69，其全部内容通过引用并入本文)中描述的那些。其他合适的接头序列对于本领域技术人员来说将是明显的。

[0071] 在一些实施方案中，本文提供的示例性融合蛋白的一般架构包含结构：[NH_2]-[脱氨酶]-[RGN多肽]-[USP]-[COOH]；[NH_2]-[USP]-[脱氨酶]-[RGN多肽]-[COOH]；[NH_2]-[USP]-[RGN多肽]-[脱氨酶]-[COOH]；[NH_2]-[RGN多肽]-[脱氨酶]-[USP]-[COOH]；[NH_2]-[RGN多肽]-[USP多肽]-[脱氨酶多肽]-[COOH]；或[NH_2]-[脱氨酶多肽]-[USP多肽]-[RGN多肽]-[COOH]，其中NH_2在融合蛋白的N端，COOH在融合蛋白的C端。

[0072] 本公开内容的一些方面提供脱氨酶-RGN-USP融合蛋白、脱氨酶-核酸酶失活的RGN-USP融合蛋白和脱氨酶-切口酶RGN-USP融合蛋白，其与不包含USP结构域的类似融合蛋白相比，具有提高的C→T核碱基编辑效率。

[0073] 在一些实施方案中，融合蛋白包含以下结构：[NH_2]-[脱氨酶]-[核酸酶失活的RGN]-[USP]-[COOH]；[NH_2]-[脱氨酶多肽]-[USP]-[核酸酶失活的RGN]-[COOH]；[NH_2]-[USP]-[脱氨酶]-[核酸酶失活的RGN]-[COOH]；[NH_2]-[USP]-[核酸酶失活的RGN]-[脱氨酶]-[COOH]；[NH_2]-[核酸酶失活的RGN]-[脱氨酶]-[USP]-[COOH]；或[NH_2]-[核酸酶失活的RGN]-[USP]-[脱氨酶]-[COOH]。应当理解，“核酸酶失活的RGN”代表已被突变为核酸酶失活的任何RGN，包括任何CRISPR-Cas蛋白。还应当理解，“USP”代表一种或多种USP多肽。

[0074] 在其他实施方案中，融合蛋白包含以下结构：[NH_2]-[脱氨酶]-[RGN切口酶]-[USP]-[COOH]；[NH_2]-[脱氨酶]-[USP]-[RGN切口酶]-[COOH]；[NH_2]-[USP]-[脱氨酶]-[RGN切口酶]-[COOH]；[NH_2]-[USP]-[RGN切口酶]-[脱氨酶]-[COOH]；[NH_2]-[RGN切口酶]-[脱氨酶]-[USP]-[COOH]；或[NH_2]-[RGN切口酶]-[USP]-[脱氨酶]-[COOH]。应当理解，“RGN切口酶”代表已被突变为具有作为切口酶的活性的任何RGN，包括任何CRISPR-Cas蛋白。还应当理解，“USP”代表一种或多种USP多肽。

[0075] 在一些实施方案中，融合蛋白包含与SEQ ID NO:47、48和76-94中的任一个具有至少80％序列同一性的胞苷脱氨酶、与SEQ ID NO:40、41、95-142中的任一个具有至少80％序列同一性的RGN(或其切口酶)以及与SEQ ID NO:1-16中的任一个具有至少80％序列同一性的USP。

[0076] 在一些实施方案中，融合蛋白包含与SEQ ID NO:47、48和76-94中的任一个具有至少85％序列同一性的胞苷脱氨酶、与SEQ ID NO:40、41和95-142中的任一个具有至少85％序列同一性的RGN(或其切口酶)以及与SEQ ID NO:1-16中的任一个具有至少85％序列同一性的USP。

[0077] 在一些实施方案中，融合蛋白包含与SEQ ID NO:47、48和76-94中的任一个具有至少90％序列同一性的胞苷脱氨酶、与SEQ ID NO:40、41和95-142中的任一个具有至少90％序列同一性的RGN(或其切口酶)以及与SEQ ID NO:1-16中的任一个具有至少90％序列同一性的USP。

[0078] 在一些实施方案中，融合蛋白包含与SEQ ID NO:47、48和76-94中的任一个具有至少95％序列同一性的胞苷脱氨酶、与SEQ ID NO:40、41和95-142中的任一个具有至少95％

序列同一性的RGN(或其切口酶)以及与SEQ ID NO:1-16中的任一个具有至少95％序列同一性的USP。

[0079] 在一些实施方案中,融合蛋白包含具有SEQ ID NO:47、48和76-94中的任一个所示氨基酸序列的胞苷脱氨酶、具有SEQ ID NO:40、41和95-142中的任一个所示氨基酸序列的RGN(或其切口酶)以及具有SEQ ID NO:1-16中的任一个所示氨基酸序列的USP。

[0080] 在一些实施方案中,上述通用架构中使用的“-”表示存在任选的接头序列。在一些实施方案中,本文提供的融合蛋白不包含接头序列。在一些实施方案中,存在至少一个任选的接头序列。

[0081] 可能存在的其他示例性特征是定位序列,例如核定位序列、细胞质定位序列、输出序列例如核输出序列或其他定位序列,以及可用于溶解、纯化或检测融合蛋白的序列标签。本文提供的合适的定位信号序列和蛋白质标签序列包括但不限于生物素羧化酶载体蛋白(BCCP)标签、myc-标签、钙调蛋白-标签、FLAG-标签、血凝素(HA)-标签,多组氨酸标签(也称为组氨酸标签或His-标签)、麦芽糖结合蛋白(MBP)-标签、nus-标签、谷胱甘肽-S-转移酶(GST)-标签、绿色荧光蛋白(GFP)-标签、硫氧还蛋白-标签、S-标签、Softag(例如,Softag 1、Softag 3)、strep标签、生物素连接酶标签、FlAsH标签、V5标签和SBP标签。其他合适的序列对本领域技术人员来说是明显的。

[0082] 在某些实施方案中,本文公开的融合蛋白包含促进融合蛋白的细胞摄取的至少一个细胞穿透结构域。细胞穿透结构域在本领域中是已知的并且通常包含带正电的氨基酸残基(即聚阳离子细胞穿透结构域)、交替的极性氨基酸残基和非极性氨基酸残基(即两亲性细胞穿透结构域)或疏水性氨基酸残基(即疏水性细胞穿透结构域)(参见,例如,Milletti F.(2012)Drug Discov Today 17:850-860)的区段。细胞穿透结构域的一个非限制性例子是来自人类免疫缺陷病毒1的反式激活转录激活因子(TAT)。

[0083] 在一些实施方案中,本文提供的USP或融合蛋白还包含核定位序列(NLS)。核定位信号、质体定位信号、线粒体定位信号、双靶向定位信号和/或细胞穿透结构域可位于融合蛋白的氨基末端(N端)、羧基末端(C端)或者在融合蛋白的内部位置。

[0084] 在一些实施方案中,NLS融合到融合蛋白或USP的N端。在一些实施方案中,NLS融合到融合蛋白或USP的C端。在一些实施方案中,NLS融合到融合蛋白的USP的N端。在一些实施方案中,NLS融合到融合蛋白的USP的C端。在一些实施方案中,NLS融合到融合蛋白的RGN多肽的N端。在一些实施方案中,NLS融合到融合蛋白的RGN多肽的C端。在一些实施方案中,NLS融合到融合蛋白的脱氨酶多肽的N端。在一些实施方案中,NLS融合到融合蛋白的脱氨酶多肽的C端。在一些实施方案中,NLS通过一个或多个接头融合到融合蛋白或UPS。在一些实施方案中,NLS在没有接头的情况下与融合蛋白或UPS融合。在一些实施方案中,NLS包含本文提供或提及的任一NLS序列的氨基酸序列。在一些实施方案中,NLS包含如SEQ ID NO:42或SEQ ID NO:45示出的氨基酸序列。

[0085] 在一些实施方案中,本文提供的融合蛋白包含尿嘧啶稳定蛋白的全长序列,例如,SEQ ID NO:1-16中的任一个。然而,在其他实施方案中,本文提供的融合蛋白不包含USP的全长序列,而仅包含其片段。例如,在一些实施方案中,本文提供的融合蛋白进一步包含RNA引导的DNA结合结构域、脱氨酶结构域和USP的活性片段。

[0086] 在一些实施方案中,本发明的融合蛋白包含RGN、脱氨酶和USP,其中USP具有与SEQ

ID NO:1-16中的任一个至少50％、至少55％、至少60％、至少65％、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％，或100％相同的氨基酸序列。此类融合蛋白的实例在本文的实施例部分中有所描述。

[0087] 在一些实施方案中，融合蛋白包含一种USP多肽。在一些实施方案中，融合蛋白包含直接或通过接头可操作地连接的至少两种USP多肽。在一些实施方案中，融合蛋白包含一种USP多肽，并且第二种USP多肽与融合蛋白共表达。

[0088] 本发明的另一个实施方案是包含融合蛋白和引导RNA(作为单引导RNA或作为双引导RNA(统称为gRNA))的核糖核蛋白复合物。

[0089] IV.编码尿嘧啶稳定多肽、融合蛋白和/或gRNA的核苷酸

[0090] 本公开内容提供了编码本文公开的尿嘧啶稳定多肽的多核苷酸(SEQ ID NO:17-32)。本公开内容还提供了编码融合蛋白的多核苷酸，所述融合蛋白包含脱氨酶和DNA结合多肽，例如大范围核酸酶、锌指融合蛋白或TALEN。本公开内容还提供了编码融合蛋白的多核苷酸，所述融合蛋白包含USP、脱氨酶结构域和RNA引导的DNA结合多肽。这种RNA引导的DNA结合多肽可以是RGN或RGN变体。蛋白质变体可以是核酸酶失活的或切口酶。RGN可以是CRISPR-Cas蛋白或其活性变体或片段。SEQ ID NO:40和41分别是RGN和切口酶RGN变体的非限制性实例。CRISPR-Cas核酸酶的实例是本领域众所周知的，并且类似的对应突变可以产生突变变体，其也是切口酶或是核酸酶失活的。

[0091] 本发明的一个实施方案提供了编码包含本文所述的RGN、脱氨酶和USP的融合蛋白的多核苷酸(SEQ ID NO:1-16，或其变体)。在一些实施方案中，第二多核苷酸编码RGN靶向感兴趣的核苷酸序列所需的引导RNA。在其他实施方案中，引导RNA和融合蛋白由相同的多核苷酸编码。

[0092] 术语“多核苷酸”的使用不旨在将本公开内容局限于包括DNA的多核苷酸。本领域普通技术人员将认可多核苷酸可包括核糖核苷酸(RNA)以及核糖核苷酸与脱氧核糖核苷酸的组合。这种脱氧核糖核苷酸和核糖核苷酸既包含天然存在的分子又包含合成类似物。本文公开的多核苷酸还涵盖所有形式的序列，包含但不限于单链形式、双链形式、茎环结构等。

[0093] 本发明的一个实施方案是包含与SEQ ID NO:17-32中的任一个至少50％、至少55％、至少60％、至少65％、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或100％相同的序列的核酸分子，其中所述核酸分子编码具有尿嘧啶稳定活性的USP。核酸分子还可以包含异源启动子或终止子。核酸分子可以编码融合蛋白，其中编码的USP可操作地连接至DNA结合多肽和/或脱氨酶。在一些实施方案中，核酸分子编码融合蛋白，其中编码的USP可操作地连接至RGN和/或脱氨酶。

[0094] 包含编码本发明的USP的多核苷酸的核酸分子可以针对在感兴趣的生物体中的表达进行密码子优化。“密码子优化”的编码序列是使其密码子使用频率设计成模拟特定宿主细胞的优选密码子使用频率或转录条件的多核苷酸编码序列。由于核酸水平上的一个或多个密码子的改变使得特定宿主细胞或生物体中的表达被增强而翻译的氨基酸序列未改变。核酸分子可以全部或部分地被密码子优化。密码子表和提供多种生物体的偏好信息的其他参考文献在本领域中是可获得的(有关植物优选密码子使用的讨论，参见，例如，Campbell和Gowri(1990)Plant Physiol.92:1-11)。本领域中用于合成植物优选基因的方法是可获

得的。参见，例如，美国专利号5,380,831和5,436,391以及Murray等人(1989)Nucleic Acids Res.17:477-498，它们通过引用并入本文。

[0095] 编码本文中描述的USP、融合蛋白和/或gRNA的多核苷酸可在表达盒中提供，以用于体外表达或在感兴趣的细胞、细胞器、胚胎或生物体中表达。该盒将包含允许多核苷酸表达的可操作地连接至编码USP和/或包含USP、RNA引导的DNA结合多肽和脱氨酶的融合蛋白和/或本文提供的gRNA的多核苷酸的5'和3'调节序列。该盒可额外含有将被共转化至生物体中的至少一种额外基因或遗传元件。如果包含额外基因或元件，则成分被可操作地连接。术语"可操作地连接"旨在意指两个或更多个元件之间的功能性连接。例如，启动子与感兴趣的编码区域(例如，编码USP、脱氨酶、RNA引导的DNA结合多肽和/或gRNA的区域)之间的可操作地连接是允许感兴趣的编码区域表达的功能性连接。可操作地连接的元件可以是连续的或非连续的。当用于指两个蛋白质编码区域的连接时，可操作地连接意指编码区域在相同的阅读框中。可选择地，可在多个表达盒上提供(多个)额外基因或元件。例如，编码本文公开的尿嘧啶稳定多肽(单独地或作为融合蛋白的组分)的核苷酸序列可存在于一个表达盒上，而编码gRNA的核苷酸序列可在分开的表达盒上。另一个实例可以具有在第一表达盒上的单独的编码本文公开的USP的核苷酸序列，第二表达盒编码包含USP的融合蛋白，以及在第三表达盒上的编码gRNA的核苷酸序列。这种表达盒提供有多个限制性位点和/或重组位点，以使多核苷酸的插入受调节区域的转录调节。也可以存在包含选择性标记基因的表达盒。

[0096] 表达盒在转录的5'-3'方向上将包含在感兴趣的生物体中起作用的转录(以及在一些实施方案中，翻译)起始区域(即，启动子)、本发明的USP编码多核苷酸、以及转录(以及在一些实施方案中，翻译)终止区(即，终止区域)。本发明的启动子能够在宿主细胞中指导或驱动编码序列的表达。调节区域(例如，启动子、转录调节区域以及翻译终止区域)对于宿主细胞或彼此可以是内源的或异源的。如本文所使用的关于序列的"异源"是源自外来物种的序列，或者，如果来自相同物种，则是通过蓄意的人为干预从其在组合物和/或基因组位点中的天然形式基本上被修饰的序列。如本文中所使用的嵌合基因包括与转录起始区域可操作地连接的编码序列，该转录起始区与编码序列是异源的。

[0097] 合宜的终止区域可从根癌农杆菌(A.tumefaciens)的Ti质粒获得，例如，章鱼碱合成酶和胭脂碱合成酶终止区域。还参见Guerineau等人(1991)Mol.Gen.Genet.262:141-144;Proudfoot(1991)Cell64:671-674;Sanfacon等人(1991)Genes Dev.5:141-149;Mogen等人(1990)Plant Cell 2:1261-1272;Munroe等人(1990)Gene 91:151-158;Ballas等人(1989)Nucleic Acids Res.17:7891-7903;和Joshi等人(1987)Nucleic Acids Res.15:9627-9639。

[0098] 额外调节信号包含但不限于转录起始初始位点、操纵子、活化子、增强子、其他调节元件、核糖体结合位点、起始密码子、终止信号等。参见，例如，美国专利号5,039,523和4,853,331;EPO 0480762A2;Sambrook等人(1992)Molecular Cloning:A Laboratory Manual,Maniatis等人编辑(Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.)，后称"Sambrook 11";Davis等人编辑(1980)Advanced Bacterial Genetics(Cold Spring Harbor Laboratory Press),Cold Spring Harbor,N.Y.,以及其中引用的参考文献。

[0099] 在制备表达盒时,可以操纵各种DNA片段,以便在适宜的方向上和恰当的情况下于适宜阅读框中提供DNA序列。为此,可以使用衔接子或接头来连接DNA片段,或者可涉及其他操纵来提供合宜的限制位点、除去多余的DNA、除去限制位点等。为此目的,可涉及体外诱变、引物修复、限制性酶切、退火、重新取代,例如转换和颠换。

[0100] 许多启动子可用于本发明的实施。可基于期望结果来选择启动子。核酸可以与组成型、诱导型、生长阶段特异性、细胞类型特异性、组织优选、组织特异性的启动子或其他启动子组合以用于在感兴趣的生物体中的表达。参见,例如,WO 99/43838中和美国专利号8,575,425;7,790,846;8,147,856;8,586832;7,772,369;7,534,939;6,072,050;5,659,026;5,608,149;5,608,144;5,604,121;5,569,597;5,466,785;5,399,680;5,268,463;5,608,142;以及6,177,611中所列的启动子;它们通过引用并入本文。

[0101] 对于在植物中的表达,组成型启动子还包括CaMV 35S启动子(Odell等人(1985)Nature 313:810-812);水稻肌动蛋白(McElroy等人(1990)Plant Cell 2:163-171);泛素(Christensen等人(1989)Plant Mol.Biol.12:619-632和Christensen等人(1992)Plant Mol.Biol.18:675-689);pEMU(Last等人(1991)Theor.Appl.Genet.81:581-588);以及MAS(Velten等人(1984)EMBO J.3:2723-2730)。

[0102] 诱导型启动子的实例是可通过缺氧或冷应激诱导的Adh1启动子、可通过热应激诱导的Hsp70启动子、可由光诱导的PPDK启动子和磷酸烯醇丙酮酸羧化酶(pepcarboxylase)启动子。同样有用的是化学诱导的启动子,诸如,安全剂诱导的In2-2启动子(美国专利号5,364,780)、生长素诱导的和绒毡层特异性的但在愈伤组织中仍有活性的Axig1启动子(PCT US01/22169)、类固醇反应性启动子(参见,例如,Schena等人(1991)Proc.Natl.Acad.Sci.USA 88:10421-10425和McNellis等人(1998)Plant J.14(2):247-257中的雌激素诱导的ERE启动子和糖皮质激素诱导型启动子)、以及四环素诱导型和四环素抑制型启动子(参见,例如,Gatz等人(1991)Mol.Gen.Genet.227:229-237以及美国专利号5,814,618和5,789,156),它们通过引用并入本文。

[0103] 组织特异性或组织优选的启动子可用于靶向特定组织内的表达构建体的表达。在某些实施方案中,组织特异性或组织优选的启动子在植物组织中是有活性的。在植物中受发育控制的启动子的实例包括优选地在诸如叶、根、果实、种子或花的某些组织中起始转录的启动子。“组织特异性的”启动子是仅在某些组织中起始转录的启动子。与基因的组成型表达不同,组织特异性表达是基因调节的几种相互作用水平的结果。因此,来自同源或密切相关的植物物种的启动子可优选地用于在特定组织中实现转基因的有效且可靠的表达。在一些实施方案中,表达包括组织优选的启动子。“组织优选”的启动子是优选地但不一定完全地或仅在某些组织中起始转录的启动子。

[0104] 在一些实施方案中,编码本文所述的USP的核酸分子包括细胞类型特异性启动子。“细胞类型特异性”启动子是主要在一个或多个器官的某些细胞类型中驱动表达的启动子。例如,其中在植物中具有功能的细胞类型特异性启动子可具有主要活性的植物细胞的一些实例包括BETL细胞、根、叶中的维管细胞、柄细胞以及茎细胞(stem cell)。核酸分子还可包含细胞类型优选的启动子。“细胞类型优选”的启动子是主要驱动一个或多个器官中的大部分但不一定全部或仅在某些细胞类型中的表达的启动子。例如,其中在植物中具有功能的细胞类型优选的启动子可具有优先活性的植物细胞的一些实例包括BETL细胞、根、叶中的

维管细胞、柄细胞以及茎细胞。

[0105] 编码USP、融合蛋白和/或gRNA的核酸序列可与噬菌体RNA聚合酶识别的启动子序列可操作地连接以例如用于体外mRNA合成。在此类实施方案中，可纯化体外转录的RNA以用于本文所述的方法。例如，启动子序列可以是T7、T3或SP6启动子序列，或T7、T3或SP6启动子序列的变体。在此类实施方案中，可纯化表达蛋白和/或RNA以用于本文所述的基因组修饰方法。

[0106] 在某些实施方案中，编码USP、融合蛋白和/或gRNA的多核苷酸还可与多腺苷酸化信号(例如，SV40多聚A信号和在植物中起作用的其他信号)和/或至少一个转录终止序列连接。另外，编码脱氨酶或融合蛋白的序列还可与编码至少一个核定位信号、至少一个细胞穿透结构域和/或能够将蛋白质运输到特定亚细胞位置的至少一个信号肽的序列连接，如本文中别处所述的。

[0107] 编码USP、融合蛋白和/或gRNA的多核苷酸可存在于一个载体或多个载体中。“载体”是指用于将核酸转移、递送或引入宿主细胞内的多核苷酸组合物。合适的载体包括质粒载体、噬菌粒、黏粒、人工/微型染色体、转座子以及病毒载体(例如，慢病毒载体、腺相关病毒载体、杆状病毒载体)。载体可包括额外的表达控制序列(例如，增强子序列、Kozak序列、多腺苷酸化序列、转录终止序列)、选择性标记序列(例如，抗生素抗性基因)、复制起点等。额外信息可在"Current Protocols in Molecular Biology"Ausubel等人，John Wiley& Sons，New York，2003或"Molecular Cloning:A Laboratory Manual"Sambrook&Russell，Cold Spring Harbor Press，Cold Spring Harbor，N.Y.，第3版，2001中找到。

[0108] 载体还可包括用于选择转化细胞的选择性标记基因。选择性标记基因用于转化的细胞或组织的选择。标记基因包括编码抗生素抗性的基因，诸如，编码新霉素磷酸转移酶II(NEO)和潮霉素磷酸转移酶(HPT)的基因，以及赋予针对诸如草铵膦、溴苯腈、咪唑啉酮以及2,4-二氯苯氧乙酸酯(2,4-D)的除草性化合物的抗性的基因。

[0109] 在一些实施方案中，包含编码包含RNA引导的DNA结合多肽例如RGN的融合蛋白的序列的表达盒或载体可以进一步包含编码gRNA的序列。编码gRNA的(多个)序列可与至少一个转录控制序列可操作地连接以用于gRNA在感兴趣的生物体或宿主细胞中的表达。例如，编码gRNA的多核苷酸可与由RNA聚合酶III(Pol III)识别的启动子序列可操作地连接。合适的Pol III启动子的实例包括但不限于哺乳动物U6、U3、H1和7SL RNA启动子以及水稻U6和U3启动子。

[0110] 如所指示的，包含编码USP、融合蛋白和/或gRNA的核苷酸序列的表达构建体可用于转化感兴趣的生物体。用于转化的方法涉及将核苷酸构建体引入感兴趣的生物体。“引入”旨在以使得构建体进入宿主细胞内部的方式将核苷酸构建体引入宿主细胞。本发明的方法不需要用于将核苷酸构建体引入宿主生物体的特定方法，而是仅需要该核苷酸构建体进入宿主生物体的至少一个细胞的内部。宿主细胞可以是真核细胞或原核细胞。在具体实施方案中，真核宿主细胞是植物细胞、哺乳动物细胞或昆虫细胞。用于将核苷酸构建体引入植物以及其他宿主细胞内的方法是本领域中已知的，其包括但不限于稳定转化方法、瞬时转化方法以及病毒介导的方法。

[0111] 这些方法导致转化的生物体，诸如植物，包括整株植物以及植物器官(例如，叶、茎、根等)、种子、植物细胞、繁殖体、胚胎及其后代。植物细胞可以是分化的或未分化的(例

如,愈伤组织、悬浮培养细胞、原生质体、叶细胞、根细胞、韧皮部细胞、花粉)。

[0112] “转基因生物体”或“转化的生物体”或“稳定转化的”生物体或细胞或组织是指已并入或整合了编码本发明的脱氨酶的多核苷酸的生物体。应认识到其他外源或内源核酸序列或DNA片段也可被并入到宿主细胞内。农杆菌介导和基因枪介导的转化仍然是用于植物细胞转化的两种主要采用的方法。然而,宿主细胞的转化可通过感染、转染、显微注射、电穿孔、显微投影(microprojection)、基因枪或粒子轰击、电穿孔、二氧化硅/碳纤维、超声波介导的、PEG介导的、磷酸钙共沉淀、聚阳离子DMSO技术、DEAE葡聚糖程序以及病毒介导的、脂质体介导的等进行。编码脱氨酶、融合蛋白和/或gRNA的多核苷酸的病毒介导的引入包括逆转录病毒、慢病毒、腺病毒以及腺相关病毒介导的引入和表达,以及花椰菜花叶病毒属(Caulimoviruses)(例如花椰菜花叶病毒)、双生病毒(例如豆金黄色花叶病毒或玉米条纹病毒)以及RNA植物病毒(例如烟草花叶病毒)的使用。

[0113] 转化方案以及用于将多肽或多核苷酸序列引入植物内的方案可根据被靶向用于转化的宿主细胞的类型(例如,单子叶植物或双子叶植物细胞)而不同。用于转化的方法是本领域中已知的,并且包括美国专利号8,575,425、7,692,068、8,802,934、7,541,517中阐述的那些方法,这些专利中的每一篇都通过引用并入本文。还参见Rakoczy-Trojanowska,M.(2002)Cell Mol Biol Lett.7:849-858;Jones等人(2005)Plant Methods 1:5;Rivera等人(2012)Physics of Life Reviews 9:308-345;Bartlett等人(2008)Plant Methods 4:1-12;Bates,G.W.(1999)Methods in Molecular Biology 111:359-366;Binns和Thomashow(1988)Annual Reviews in Microbiology 42:575-606;Christou,P.(1992)The Plant Journal 2:275-281;Christou,P.(1995)Euphytica 85:13-27;Tzfira等人(2004)TRENDS in Genetics 20:375-383;Yao等人(2006)Journal of Experimental Botany 57:3737-3746;Zupan和Zambryski(1995)Plant Physiology 107:1041-1047;Jones等人(2005)Plant Methods 1:5。

[0114] 转化可导致核酸稳定或瞬时掺入细胞内。“稳定转化”旨在指引入宿主细胞内的核苷酸构建体整合到宿主细胞的基因组内,并且能够被其后代遗传。“瞬时转化”旨在指多核苷酸被引入宿主细胞内并且不整合到宿主细胞的基因组内。

[0115] 用于叶绿体转化的方法是本领域中已知的。参见,例如,Svab等人(1990)Proc.Natl.Acad.Sci.USA 87:8526-8530;Svab和Maliga(1993)Proc.Natl.Acad.Sci.USA 90:913-917;Svab和Maliga(1993)EMBO J.12:601-606。该方法依赖于含有选择性标记的DNA的粒子枪递送和通过同源重组将该DNA靶向至质体基因组。另外,质体转化可通过核编码的和质体指导的RNA聚合酶的组织优选的表达来反式活化(transactivation)沉默的质体携带的转基因来实现。在McBride等人(1994)Proc.Natl.Acad.Sci.USA 91:7301-7305中报导了这种系统。

[0116] 已经被转化的细胞可以按照传统方法生长成转基因生物体,诸如植物。参见,例如McCormick等人(1986)Plant Cell Reports 5:81-84。然后,可以使这些植物生长,并用相同转化品系或不同品系授粉,并且鉴别出具有期望表型特征的组成型表达的所得杂交体。可生长两代或更多代,以确保稳定维持并遗传期望表型特征的表达,然后,收获种子以确保已经实现期望表型特征的表达。以此方式,本发明提供具有稳定地掺入其基因组内的本发明的核苷酸构建体(例如,本发明的表达盒)的转化种子(也被称为“转基因种子”)。

[0117] 可选择地，可以将已经被转化的细胞引入生物体内。这些细胞可以源自该生物体，其中细胞以离体方法被转化。

[0118] 本文提供的序列可用于转化任何植物物种，包含但不限于：单子叶植物和双子叶植物。感兴趣的植物的实例包含但不限于：玉米(玉蜀黍)、高粱、小麦、向日葵、西红柿、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、甜菜、甘蔗、烟草、大麦及油菜、芸苔属物种、苜蓿、黑麦、小米、红花、花生、甘薯、木薯、咖啡、椰子、菠萝、柑橘树、可可、茶、香蕉、鳄梨、无花果、番石榴、芒果、橄榄、木瓜、腰果、澳洲胡桃、杏仁、燕麦、蔬菜、观赏植物以及针叶树。

[0119] 蔬菜包括但不限于西红柿、莴苣、绿豆、利马豆、豌豆、以及诸如黄瓜、哈密瓜和香瓜的Curcumins属的成员。观赏植物包括但不限于杜鹃花、绣球花、木槿、玫瑰、郁金香、水仙、矮牵牛、康乃馨、猩猩木以及菊花。优选地，本发明的植物是农作物(例如，玉蜀黍、高粱、小麦、向日葵、西红柿、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、甜菜、甘蔗、烟草、大麦、油菜等)。

[0120] 如本文所使用的，术语“植物”包括植物细胞、植物原生质体、可以从其再生植物的植物细胞组织培养物、植物愈伤组织、植物块、以及在植物或植物的诸如胚胎、花粉、胚珠、种子、叶子、花、枝、果实、仁、穗、穗轴、壳、茎、根、根尖、花药等的部分中是完整的植物细胞。谷物意指出于种植或繁殖物种之外的目的而由商业种植者生产的成熟种子。再生植物的后代、变体以及突变体也包括在本发明的范围内，前提条件是这些部分包括引入的多核苷酸。还提供了保留本文中公开的序列的经加工的植物产品或副产物，包括例如豆粕。

[0121] 编码USP、融合蛋白和/或gRNA的多核苷酸可用于转化任何真核物种，包括但不限于动物(例如，哺乳动物、昆虫、鱼类、鸟类以及爬行动物)、真菌、变形虫、藻类以及酵母。编码USP、融合蛋白和/或gRNA的多核苷酸也可用于转化任何原核物种，包括但不限于古生菌和细菌(例如，芽孢杆菌属的几个种(Bacillus spp.)、克雷伯氏菌属的几个种(Klebsiella spp.)、链霉菌属的几个种(Streptomyces spp.)、根瘤菌属的几个种(Rhizobium spp.)、埃希氏菌属的几个种(Escherichia spp.)、假单胞菌属的几个种(Pseudomonas spp.)、沙门氏菌属的几个种(Salmonella spp.)、志贺氏杆菌属的几个种(Shigella spp.)、弧菌属的几个种(Vibrio spp.)、耶尔森菌属的几个种(Yersinia spp.)、支原体菌属的几个种(Mycoplasma spp.)、农杆菌属的几个种(Agrobacterium spp.)和乳杆菌属的几个种(Lactobacillus spp.)。

[0122] 传统的基于病毒和非病毒的基因转移方法可用于将核酸引入哺乳动物细胞或靶组织中。这些方法可用于将编码本发明的融合蛋白和任选gRNA的核酸施用给培养中的细胞或宿主生物体中的细胞。非病毒载体递送系统包括DNA质粒、RNA(例如，本文描述的载体的转录物)、裸核酸以及与诸如脂质体的递送载剂复合的核酸。病毒载体递送系统包括DNA和RNA病毒，其在递送至细胞后具有附加型或整合的基因组。非限制性实例包括利用花椰菜花叶病毒属(Caulimoviruses)(例如花椰菜花叶病毒)、双生病毒(例如豆金黄色花叶病毒或玉米条纹病毒)以及RNA植物病毒(例如烟草花叶病毒)的载体。有关基因治疗程序的综述，参见Anderson,Science 256:808-813(1992);Nabel&Feigner,TIBTECH 11:211-217(1993);Mitani&Caskey,TIBTECH 11:162-166(1993);Dillon,TIBTECH 11:167-175(1993);Miller,Nature 357:455-460(1992);Van Brunt,Biotechnology 6(10):1149-1154(1988);Vigne,Restorative Neurology and Neuroscience 8:35-36(1995);Kremer&

Perricaudet,British Medical Bulletin 51(1):31-44(1995);Haddada等人,于Current Topics in Microbiology and Immunology,Doerfler和Bohm(eds)(1995);和Yu等人,Gene Therapy 1:13-26(1994)。

[0123] 核酸的非病毒递送方法包括脂质体转染、农杆菌介导的转化、核转染、显微注射、基因枪、病毒微体、脂质体、免疫脂质体、聚阳离子或脂质:核酸缀合物、裸DNA、人工病毒粒子以及DNA的药剂增强摄取。例如美国专利号5,049,386、4,946,787以及4,897,355中描述了脂质体转染,并且脂质体转染试剂是市售的(例如,Transfectam™和Lipofectin™)。适合用于多核苷酸的有效受体识别脂质体转染的阳离子和中性脂质包括Feigner的WO 91/17424;WO 91/16024中的那些阳离子和中性脂质。递送可以是递送至细胞(例如,体外或离体施用)或靶组织(例如,体内施用)。脂质:核酸复合物包括靶向的脂质体诸如免疫脂质复合物的制备是本领域技术人员众所周知的(参见,例如Crystal,Science 270:404-410(1995);Blaese等人,Cancer Gene Ther.2:291-297(1995);Behr等人,Bioconjugate Chem.5:382-389(1994);Remy等人,Bioconjugate Chem.5:647-654(1994);Gao等人,Gene Therapy 2:710-722(1995);Ahmad等人,Cancer Res.52:4817-4820(1992);美国专利号4,186,183、4,217,344、4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028以及4,946,787)。

[0124] 用于核酸的递送的基于RNA或DNA病毒的系统的使用利用了高度进化的过程以用于将病毒靶向至体内的特定细胞并将病毒载荷运输至细胞核。病毒载体可以直接施用于患者(体内),或者该病毒载体可以用于体外处理细胞,并且可任选地将修饰后的细胞施用于患者(离体)。传统的基于病毒的系统可包括用于基因转移的逆转录病毒、慢病毒、腺病毒、腺相关以及单纯疱疹病毒载体。使用逆转录病毒、慢病毒以及腺相关病毒基因转移方法,在宿主基因组中的整合是可能的,常常导致插入的转基因的长期表达。另外,在许多不同细胞类型和靶组织中已经观察到高转导功效。

[0125] 逆转录病毒的向性可通过掺入外来包膜蛋白而被改变,从而扩展靶细胞的潜在靶群。慢病毒载体是能够转导或感染非分裂细胞并且通常情况下生成高病毒效价的逆转录病毒载体。因此,逆转录病毒基因转移系统的选择将取决于靶组织。逆转录病毒载体由顺式作用长末端重复序列构成,具有达到6-10kb外源序列的包装能力。最小顺式作用LTR足够用于载体的复制和包装,然后,将其用于将治疗基因整合到靶细胞中,以提供永久转基因表达。广泛使用的逆转录病毒载体包括基于鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猿猴免疫缺陷病毒(SIV)、人免疫缺陷病毒(HIV)及其组合的载体。(参见,例如,Buchscher等人,J.Virol.66:2731-2739(1992);Johann等人,J.Virol.66:1635-1640(1992);Sommnerfelt等人,Virol.176:58-59(1990);Wilson等人,J.Virol.63:2374-2378(1989);Miller等人,J.Virol.65:2220-2224(1991);PCT/US94/05700)。

[0126] 在瞬时表达优选的应用中,可以使用基于腺病毒的系统。基于腺病毒的载体能够在许多细胞类型中具有非常高的转导效率,并且不需要细胞分裂。使用这样的载体,已经获得了高效价和表达水平。此载体可在相对简单的系统中大量生产。腺相关病毒("AAV")载体也可用于转导具有靶核酸的细胞,例如在核酸和肽的体外生成中,并且用于体内和离体基因治疗程序(参见,例如,West等人,Virology160:38-47(1987);美国专利号4,797,368;WO93/24641;Katin,Human Gene Therapy 5:793-801(1994);Muzyczka,

J.Clin.Invest.94:1351(1994))。重组AAV载体的构建描述于许多出版物中,包括美国专利号5,173,414;Tratschin等人,Mol.Cell.Biol.5:3251-3260(1985);Tratschin等人,Mol.Cell.Biol.4:2072-2081(1984);Hermonat&Muzyczka,PNAS 81:6466-6470(1984);和Samulski等人,J.Virol.63:03822-3828(1989)。包装细胞通常情况下用于形成能够感染宿主细胞的病毒颗粒。这种细胞包括包装腺病毒的293细胞和包装逆转录病毒的ψJ2细胞或PA317细胞。

[0127] 用于基因治疗的病毒载体通常通过生成将核酸载体包装于病毒颗粒内的细胞系而被产生。载体通常情况下含有用于包装且随后整合至宿主内所需的最小病毒序列,其他病毒序列被用于要表达的(多个)多核苷酸的表达盒取代。遗失的病毒功能通常情况下由包装细胞系反式提供。例如,基因治疗中使用的AAV载体通常情况下仅具有来自用于包装及整合至宿主基因组内所需的AAV基因组的ITR序列。病毒DNA被包装在细胞系中,该细胞系含有编码其他AAV基因(即rep和cap)的辅助质粒,但缺少ITR序列。

[0128] 也可以用腺病毒作为辅助者来感染细胞系。辅助病毒促进AAV载体的复制和来自辅助质粒的AAV基因的表达。因缺少ITR序列,未以显著量包装辅助质粒。腺病毒的污染可通过例如热处理来降低,腺病毒比AAV对热处理更敏感。将核酸递送至细胞的额外方法是本领域技术人员已知的。例如,参见US20030087817,其通过引用并入本文。

[0129] 在一些实施方案中,用本文描述的一个或多个载体瞬时或非瞬时转染宿主细胞。在一些实施方案中,细胞在其天然存在于受试者中时被转染。在一些实施方案中,被转染的细胞取自受试者。在一些实施方案中,细胞衍生自取自受试者的细胞,诸如细胞系。在一些实施方案中,细胞或细胞系是原核的。在其他实施方案中,细胞或细胞系是真核的。在进一步的实施方案中,细胞或细胞系来源于昆虫、鸟类、植物或真菌物种。在一些实施方案中,细胞或细胞系可以是哺乳动物,例如人、猴、小鼠、牛、猪、山羊、仓鼠、大鼠、猫或狗。用于组织培养的多种细胞系是本领域中已知的。细胞系的实例包括但不限于C8161、CCRF-CEM、MOLT、mIMCD-3、NHDF、HeLaS3、Huh1、Huh4、Huh7、HUVEC、HASMC、HEKn、HEKa、MiaPaCell、Panel、PC-3、TF1、CTLL-2、CIR、Rat6、CVI、RPTE、A10、T24、182、A375、ARH-77、Calu1、SW480、SW620、SKOV3、SK-UT、CaCo2、P388D1、SEM-K2、WEHI-231、HB56、TIB55、lurkat、145.01、LRMB、Bcl-1、BC-3、IC21、DLD2、Raw264.7、NRK、NRK-52E、MRC5、MEF、Hep G2、HeLa B、HeLa T4.COS、COS-1、COS-6、COS-M6A、BS-C-1猴肾上皮细胞、BALB/3T3小鼠胚胎纤维母细胞、3T3 Swiss、3T3-L1、132-d5人胎儿纤维母细胞;10.1小鼠纤维母细胞、293-T、3T3、721、9L、A2780、A2780ADR、A2780cis、A172、A20、A253、A431、A-549、ALC、B16、B35、BCP-I细胞、BEAS-2B、bEnd.3、BHK-21、BR 293、BxPC3、C3H-10T1/2、C6/36、Cal-27、CHO、CHO-7、CHO-IR、CHO-K1、CHO-K2、CHO-T、CHO Dhfr-/-、COR-L23、COR-L23/CPR、COR-L23501O、CORL23/R23、COS-7、COV-434、CML T1、CMT、CT26、D17、DH82、DU145、DuCaP、EL4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、HEK-293、HeLa、Hepalclc7、HL-60、HMEC、HT-29、lurkat、lY细胞、K562细胞、Ku812、KCL22、KG1、KY01、LNCap、Ma-Mel1-48、MC-38、MCF-7、MCF-10A、MDA-MB-231、MDA-MB-468、MDA-MB-435、MDCKII、MDCKII、MOR/0.2R、MONO-MAC 6、MTD-1A、MyEnd、NCI-H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、NALM-1、NW-145、OPCN/OPCT细胞系、Peer、PNT-1A/PNT 2、RenCa、RIN-5F、RMA/RMAS、Saos-2细胞、Sf-9、SkBr3、T2、T-47D、T84、THP1细胞系、U373、U87、U937、VCaP、Vero细胞、WM39、WT-49、X63、YAC-1、YAR及其转基因品

种。细胞系可从本领域技术人员已知的多种来源获得(参见,例如美国典型培养物保藏中心(American Type Culture Collection)(ATCC)(Manassas,Va.))。

[0130] 在一些实施方案中,用本文描述的一个或多个载体转染的细胞被用于建立包含一个或多个载体衍生序列的新细胞系。在一些实施方案中,用本发明的融合蛋白和任选gRNA或用本发明的核糖核蛋白复合物瞬时转染并通过融合蛋白或核糖核蛋白复合物的活性修饰的细胞被用于建立包括含有修饰但缺少任何其他外源序列的细胞的新细胞系。在一些实施方案中,用本文描述的一个或多个载体瞬时或非瞬时转染的细胞或从此类细胞衍生的细胞系用于评估一个或多个测试化合物。

[0131] 在一些实施方案中,本文描述的一个或多个载体用于生成非人的转基因动物或转基因植物。在一些实施方案中,转基因动物是昆虫。在进一步的实施方案中,昆虫是害虫,例如蚊子或蜱。在其他实施方案中,昆虫是植物害虫,例如玉米根虫或秋粘虫。在一些实施方案中,转基因动物是鸟类,例如鸡、火鸡、鹅或鸭。在一些实施方案中,转基因动物是哺乳动物,例如人、小鼠、大鼠、仓鼠、猴子、猿、兔、猪、牛、马、山羊、绵羊、猫或狗。

[0132] V.多肽及多核苷酸的变体和片段

[0133] 本公开内容提供了天然存在的(即,野生型)尿嘧啶稳定多肽的活性变体和片段(其氨基酸序列如SEQ ID NO:1-16所示)以及编码它们的多核苷酸。

[0134] 虽然与感兴趣的多核苷酸或多肽相比较可以改变变体或片段的活性,但变体和片段应保留感兴趣的多核苷酸或多肽的功能性。例如,当与感兴趣的多核苷酸或多肽相比较时,变体或片段可具有增加的活性、降低的活性、不同的活性谱或任何其他活性改变。

[0135] 天然存在的USP的片段和变体(如本文公开的那些)将保留活性,使得如果它们是进一步包含脱氨酶或其片段和/或DNA结合多肽或其片段的融合蛋白的一部分,所述融合蛋白将与不包含USP结构域的类似融合蛋白相比表现出增加的C→T核碱基编辑效率。

[0136] 术语“片段”是指本发明的多核苷酸或多肽序列的一部分。“片段”或“生物活性部分”包含多核苷酸,该多核苷酸包含足以保持生物活性(即对核酸的脱氨酶活性)的数量的连续核苷酸。“片段”或“生物活性部分”包含多肽,该多肽包含足以保持生物活性的数量的连续氨基酸残基。USP的片段包含由于使用代替的下游开始位点而比全长序列短的那些。USP的生物活性部分可以是多肽,其包含例如SEQ ID NO:1-16中的任一个的10、20、30、40、50、60、70、80、90、100、110或更多个连续氨基酸残基,或其变体。此类生物活性部分可通过重组技术制备并评估活性。

[0137] 一般来说,“变体”旨在意指基本上相似的序列。对于多核苷酸,变体包括在天然多核苷酸内的一个或多个内位点处的一个或多个核苷酸的缺失和/或添加和/或天然多核苷酸中一个或多个位点处的一个或多个核苷酸的取代。如本文所使用的,“天然”或“野生型”多核苷酸或多肽分别包括天然存在的核苷酸序列或氨基酸序列。对于多核苷酸,保守变体包括因为遗传密码的简并性而编码感兴趣的基因的天然氨基酸序列的那些序列。与例如利用下面阐述的聚合酶链式反应(PCR)和杂交技术一样,可使用众所周知的分子生物学技术鉴定诸如这些等位基因变体的天然存在的等位基因变体。变体多核苷酸还包含合成衍生的多核苷酸,诸如通过使用定点诱变产生的但其仍然编码感兴趣的多肽或多核苷酸的多核苷酸。一般而言,如通过本文其他地方描述的序列比对程序和参数所确定的,本文公开的特定多核苷酸的变体与该特定多核苷酸具有至少约40%、45%、50%、55%、60%、65%、70%、

75％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更高的序列同一性。

[0138] 本文公开的特定多核苷酸(即,参考多核苷酸)的变体也可以通过比较由变体多核苷酸所编码的多肽与该参考多核苷酸所编码的多肽之间的序列同一性百分比来评价。可使用本文其他地方描述的序列比对程序和参数来计算任何两个多肽之间的序列同一性百分比。当通过比较本文公开的任何给定的多核苷酸对编码的两个多肽共有的序列同一性百分比来评价本文公开的任何给定的多核苷酸对时,这两个编码的多肽之间的序列同一性百分比为至少约40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更高的序列同一性。

[0139] 在具体实施方案中,本文公开的多核苷酸编码USP,其包含与SEQ ID NO:1-16中的任一个的氨基酸序列具有至少40％、45％、50％、55％、60％、65％、70％、75％、80％、81％、82％、83％、84％、85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更高的同一性的氨基酸序列。

[0140] 本发明的尿嘧啶稳定多肽的生物活性变体可以相差少至约1-15个氨基酸残基、少至约1-10个(诸如,约6-10个)、少至5个、少至4个、少至3个、少至2个或少至1个氨基酸残基。在具体实施方案中,多肽可包括N端或C端截短,该截短可包括从该多肽的N端或C端缺失至少5、10、15、20、25、30、35、40、45、50、55、60个或更多个氨基酸。

[0141] 已经认识到的是,可对本文提供的USP进行修饰,从而建立变体蛋白质和多核苷酸。人为设计的变化可通过位点指导的诱变技术的应用来引入。可选择地,在结构上和/或功能上与本文公开的序列有关的天然的、尚不知晓的或尚未鉴别的多核苷酸和/或多肽也可被认为落入本发明范围内。可以在不改变尿嘧啶稳定多肽功能的非保守区域中进行保守式氨基酸取代。可选择地,可以进行改进尿嘧啶稳定多肽活性的修饰。

[0142] 变体多核苷酸和蛋白质还涵盖从诸如DNA改组的诱变和引起重组程序衍生的序列和蛋白质。利用这种程序,操纵本文公开的一个或多个不同的USP(例如,SEQ ID NO:1-16),以建立具有期望特性的新USP。以这种方式,重组多核苷酸文库是从包含序列区域的相关序列多核苷酸群产生的,该序列区域具有基本的序列同一性并且可以在体外或体内同源重组。例如,使用这种方法,编码感兴趣的结构域的序列基序可在本文提供的USP序列与其他随后鉴定的USP基因之间改组,以获得编码具有改善的感兴趣的性质(诸如在酶的情况下增加的K_m)的蛋白质的新基因。这种DNA改组的策略在本领域中是已知的。参见,例如Stemmer(1994)Proc.Natl.Acad.Sci.USA 91:10747-10751;Stemmer(1994)Nature 370:389-391;Crameri等人(1997)Nature Biotech.15:436-438;Moore等人(1997)J.Mol.Biol.272:336-347;Zhang等人(1997)Proc.Natl.Acad.Sci.USA 94:4504-4509;Crameri等人(1998)Nature 391:288-291;以及美国专利号5,605,793和5,837,458。“改组的”核酸是通过改组程序(诸如本文阐述的任何改组程序)生成的核酸。改组的核酸是通过例如以人工方式和任选地递归方式(实际或虚拟地)重组两种或更多种核酸(或字符串)生成的。一般而言,在改组过程中使用一个或多个筛选步骤来识别感兴趣的核酸;此筛选步骤可在任何重组步骤之前或之后进行。在一些(但不是全部)改组实施方案中,期望在选择之前进行多轮重组,以增加要筛选的库的多样性。重组和选择的整个过程可任选递归地重复。取决于上下文,改组可以指重组和选择的整个过程,或者,可选择地,可仅指整个过程的重组部分。

[0143] 如本文在两个多核苷酸或多肽序列的背景中所使用的“序列同一性”或“同一性”涉及当在指定的比较窗口上就最大对应性进行对齐时为相同的两个序列中的残基。当关于蛋白质使用序列同一性百分比时，应认识到不相同的残基位置常常因保守氨基酸取代而不同，其中氨基酸残基被取代成具有类似化学性质(例如，电荷或疏水性)的其他氨基酸残基，因此不改变分子的功能性质。当序列在保守取代上不同时，可以上调序列同一性百分比以校正取代的保守性质。因这种保守取代而不同的序列被称为具有“序列相似性”或“相似性”。用于进行这种调节的手段是本领域技术人员众所周知的。通常情况下，这涉及将保守取代作为部分错配而非完全错配进行评分，从而增加序列同一性百分比。由此，例如，当对相同氨基酸给予1分，而对非保守取代给予零分时，对保守取代给予零与1之间的得分。例如如在程序PC/GENE(Intelligenetics,Mountain View,California)中所实施的计算保守取代的得分。

[0144] 如本文所使用的，“序列同一性百分比”是指通过在比较窗口上比较两个最佳对齐序列所确定的值，其中多核苷酸序列在比较窗口中的部分可包括相较于参考序列(其不包括添加或缺失)的添加或缺失(即，缺口)以用于该两个序列的最佳对齐。通过确定在两个序列中出现相同的核酸碱基或氨基酸残基的位置的数目以得到匹配位置的数目、将匹配位置的数目除以比较窗口中的位置总数并将该结果乘以100以得到序列同一性百分比来计算该百分比。

[0145] 除非另外说明，否则本文提供的序列同一性/相似性值是指使用GAP版本10采用以下参数获得的值：核苷酸序列的％同一性和％相似性使用GAP权重50和长度权重3，以及nwsgapdna.cmp评分矩阵；氨基酸序列的％同一性和％相似性使用GAP权重8和长度权重2，以及BLOSUM62评分矩阵；或其任何等同程序。“等同程序”是指对于所考虑的任何两个序列，当与GAP版本10产生的对应对齐相比较时，产生具有相同的核苷酸或氨基酸残基匹配和相同的序列同一性百分比的对齐的任何序列比较程序。

[0146] 当为了进行相似性评分而使用限定的氨基酸取代矩阵(例如，BLOSUM62)、缺口存在罚分以及缺口延伸罚分对齐两个序列，以达到该对序列可能的最高分数时，该两个序列被“最佳对齐”。氨基酸取代矩阵及其在定量两个序列之间的相似性中的用途在本领域中是已知的，并且被描述于例如Dayhoff等人(1978)"A model of evolutionary change in proteins."In"Atlas of Protein Sequence and Structure,"Vol.5,Suppl.3(ed.M.O.Dayhoff),pp.345-352.Natl.Biomed.Res.Found.,Washington,D.C.和Henikoff等人(1992)Proc.Natl.Acad.Sci.USA 89:10915-10919。BLOSUM62矩阵常常用作序列对齐操作流程中的默认评分取代矩阵。缺口存在罚分是针对单一氨基酸缺口在其中一个对齐序列中的引入而施加的，而缺口延伸罚分是针对被插入已经打开的缺口内的每一个额外空的氨基酸位置而施加的。对齐是由在对齐开始和结束处的每一个序列的氨基酸位置限定的，并且可任选地通过一个缺口或多个缺口在一个或两个序列中的插入来限定，以便达到最高可能分数。尽管可以手动完成最佳对齐和评分，但是通过使用计算机实施的对齐算法(例如Altschul等人(1997)Nucleic Acids Res.25:3389-3402中描述的gapped BLAST 2.0)促进该过程，并且所述算法在美国国家生物技术信息中心网站(www.ncbi.nlm.nih.gov)向大众开放。可以使用例如可通过www.ncbi.nlm.nih.gov获得的和Altschul等人(1997)Nucleic Acids Res.25:3389-3402中描述的PSI-BLAST来准备包含多重对齐的最佳对齐。

[0147] 关于与参考序列最佳对齐的氨基酸序列，氨基酸残基“对应于”参考序列中在该对齐中与该残基配对的位置。“位置”由数字标示，该数字基于其相对于N端的位置依序识别该参考序列中的每一个氨基酸。归因于在确定最佳对齐时必须考虑的缺失、插入、截短、融合等，一般来说，通过简单地从N端计数所确定的测试序列中的氨基酸残基数目不一定与其在该参考序列中的对应位置的数目相同。例如，在对齐的测试序列中存在缺失的情况中，将不存在与缺失位点处的参考序列中的位置对应的氨基酸。当在对齐的参考序列中存在插入时，该插入将不对应于该参考序列中的任何氨基酸位置。在截短或融合的情况中，参考序列或对齐序列中可存在不对应于相应序列中的任何氨基酸的氨基酸区段。

[0148] VI.抗体

[0149] 还涵盖了针对本发明的USP、融合蛋白或包含USP的核糖核蛋白(包括具有如SEQ ID NO:1-16示出的氨基酸序列或其活性变体或片段的那些)的抗体。生成抗体的方法是在本领域中众所周知的(参见，例如，Harlow和Lane(1988)Antibodies:A Laboratory Manual,Cold Spring Harbor Laboratory,Cold Spring Harbor,N.Y.;以及美国专利号4,196,265)。这些抗体可用于试剂盒中以用于本文所述的USP或融合蛋白或包含USP的核糖核蛋白的检测和分离。因此，本公开内容提供了包含特异性结合本文描述的多肽或核糖核蛋白(包括例如包含与SEQ ID NO:1-16中的任一个具有至少85％同一性的序列的多肽)的抗体的试剂盒。

[0150] VII.用于结合感兴趣的靶序列的系统和核糖核蛋白复合物及其制备方法

[0151] 本公开内容提供了一种靶向核酸序列并修饰靶核酸序列的系统。在一些实施方案中，RNA引导的DNA结合多肽例如RGN和gRNA负责将核糖核蛋白复合物靶向至感兴趣的核酸序列;脱氨酶多肽负责从C>U修饰靶核酸序列;尿嘧啶稳定多肽允许尿嘧啶在DNA分子中持续存在，使得发生所需的DNA修复，从而引入C>T突变。引导RNA与感兴趣的靶序列杂交，并且还与RNA引导的DNA结合多肽形成复合物，从而指导RNA引导的DNA结合多肽结合靶序列。RNA引导的DNA结合多肽是3-结构域融合蛋白的一个结构域;第二个结构域是脱氨酶，第三个结构域是本文所述的USP。在一些实施方案中，RNA引导的DNA结合多肽是RGN，例如Cas9。RNA引导的DNA结合多肽的其他实例包括RGN，例如美国专利申请公开号2019/0367949(通过引用整体并入本文)中描述的那些。在一些实施方案中，RNA引导的DNA结合多肽是II型CRISPR-Cas多肽或其活性变体或片段。在一些实施方案中，RNA引导的DNA结合多肽是V型CRISPR-Cas多肽或其活性变体或片段。在其他实施方案中，RNA引导的DNA结合多肽是VI型CRISPR-Cas多肽。在其他实施方案中，融合蛋白的DNA结合结构域不需要RNA引导物，例如锌指核酸酶、TALEN或大范围核酸酶多肽。在这些实施方案中的一些中，每一个的核酸酶活性已被灭活。在进一步的实施方案中，RNA引导的DNA结合多肽包含RGN的氨基酸序列，例如与APG07433.1(SEQ ID NO:40)具有至少80％序列同一性的氨基酸序列或其活性变体或片段，例如切口酶APG07433.1(SEQ ID NO:41)。在进一步的实施方案中，RNA引导的DNA结合多肽包含RGN的氨基酸序列，例如与APG07433.1(SEQ ID NO:40)具有至少85％序列同一性的氨基酸序列或其活性变体或片段，例如切口酶APG07433.1(SEQ ID NO:41)。在进一步的实施方案中，RNA引导的DNA结合多肽包含RGN的氨基酸序列，例如与APG07433.1(SEQ ID NO:40)具有至少90％序列同一性的氨基酸序列或其活性变体或片段，例如切口酶APG07433.1(SEQ ID NO:41)。在进一步的实施方案中，RNA引导的DNA结合多肽包含RGN的氨基酸序列，例如与

APG07433.1(SEQ ID NO:40)具有至少95%序列同一性的氨基酸序列或其活性变体或片段，例如切口酶APG07433.1(SEQ ID NO:41)。在进一步的实施方案中，RNA引导的DNA结合多肽包含RGN的氨基酸序列，例如APG07433.1(SEQ ID NO:40)或其活性变体或片段，例如切口酶APG07433.1(SEQ ID NO:41)。

[0152] 本文提供的用于结合感兴趣的靶序列的系统可以是核糖核蛋白复合物，其是结合至少一种蛋白质的至少一种RNA分子。本文提供的核糖核蛋白复合物包含至少一种引导RNA作为RNA成分和包含脱氨酶、本发明的USP和RNA引导的DNA结合多肽的融合蛋白作为蛋白质成分。可以从已经用编码融合蛋白和引导RNA的多核苷酸转化并在允许表达融合蛋白和引导RNA的条件下培养的细胞或生物体纯化核糖核蛋白复合物。因此，提供了制备USP、融合蛋白或融合蛋白核糖核蛋白复合物的方法。此类方法包括在表达USP或融合蛋白(以及在一些实施方案中，引导RNA)的条件下培养包含编码USP、融合蛋白的核苷酸序列以及在一些实施方案中编码引导RNA的核苷酸序列的细胞。然后可以从培养细胞的裂解物中纯化USP、融合蛋白或融合核糖核蛋白。

[0153] 用于从生物样品的裂解物中纯化USP、融合蛋白或融合核糖核蛋白复合物的方法是本领域中已知的(例如，尺寸排阻层析法和/或亲和层析法、2D-PAGE、HPLC、反相层析法、免疫沉淀法)。在特定方法中，USP或融合蛋白是重组生成的并且包含纯化标签以帮助其纯化，纯化标签包括但不限于：谷胱甘肽-S-转移酶(GST)、几丁质结合蛋白(CBP)、麦芽糖结合蛋白、硫氧还蛋白(TRX)、聚(NANP)、串联亲和纯化(TAP)标签、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、HA、nus、Softag 1、Softag 3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、6xHis、生物素羧基载体蛋白(BCCP)以及钙调蛋白。通常，使用免疫沉淀或本领域已知的其他类似方法纯化标记的USP、融合蛋白或融合核糖核蛋白复合物。

[0154] “分离的”或“纯化的”多肽或其生物活性部分基本上或实质上不含在其天然存在的环境中发现的、正常情况下与该多肽相伴或相互作用的成分。因此，分离的或纯化的多肽当通过重组技术被生成时基本上不含其他细胞物质或培养基，或当被化学合成时基本上不含化学前体物或其他化学物质。基本上不含细胞物质的蛋白质包含具有少于约30%、20%、10%、5%或1%(以干重计)的污染蛋白质的蛋白质制剂。当重组产生本发明的蛋白质或其生物活性部分时，最佳地培养基呈现少于约30%、20%、10%、5%或1%(以干重计)的化学前体物或非感兴趣的蛋白质的化学物质。

[0155] VIII.修饰靶序列的方法

[0156] 本公开内容提供了用于修饰感兴趣的靶核酸分子(例如，靶DNA分子)的方法。所述方法包括将包含至少一种引导RNA或编码其的多核苷酸以及包含本发明的USP、脱氨酶和RNA引导的DNA结合多肽的至少一种融合蛋白或编码其的多核苷酸的系统递送至靶序列或包含靶序列的细胞、细胞器或胚胎。在这些实施方案的一些中，融合蛋白包含SEQ ID NO:1-16的氨基酸序列中的任一个或其活性变体或片段。

[0157] 在一些实施方案中，所述方法包括使DNA分子与以下接触：(a)包含USP、脱氨酶和RNA引导的DNA结合多肽例如核酸酶失活的或切口酶Cas9结构域的融合蛋白；和(b)将(a)的融合蛋白靶向至DNA分子的靶核苷酸序列的gRNA；其中所述DNA分子与融合蛋白及gRNA以有效量并且在适于核苷酸碱基的脱氨基的条件下接触。在一些实施方案中，靶DNA分子包含与疾病或病症相关的序列，并且其中核苷酸碱基的脱氨基导致与疾病或病症无关的序列。在

一些实施方案中，疾病或病症影响动物。在进一步的实施方案中，疾病或病症影响哺乳动物，例如人、牛、马、狗、猫、山羊、绵羊、猪、猴子、大鼠、小鼠或仓鼠。在一些实施方案中，靶DNA序列驻留于农作物的等位基因中，其中感兴趣的性状的特定等位基因导致具有较低农艺价值的植物。核苷酸碱基的脱氨基导致改善了性状并增加了植物的农艺价值的等位基因。

[0158] 在一些实施方案中，所需突变包括与疾病或病症相关的T→C点突变，并且其中突变体C碱基的脱氨基导致与疾病或病症不相关的序列。在一些实施方案中，脱氨基校正与疾病或病症相关的序列中的点突变。

[0159] 在一些实施方案中，与疾病或病症相关的序列编码蛋白质，并且其中脱氨基将终止密码子引入与疾病或病症相关的序列中，导致编码蛋白质的截短。在一些实施方案中，该接触在易患有、患有或诊断患有该疾病或病症的受试者中体内进行。在一些实施方案中，该疾病或病症是与基因组中的点突变或单碱基突变相关的疾病。在一些实施方案中，该疾病是遗传性疾病、癌症、代谢疾病或溶酶体贮积病。

[0160] 因此，本文公开的组合物和方法可用于治疗与突变的序列相关(即，该序列是疾病或病症的原因或与疾病或病症相关的症状的原因)的疾病或病症以治疗疾病或病症或减轻与疾病或病症相关的症状。如本文所用，术语“治疗”或“医治”是指将本文公开的包含USP或融合蛋白的药物组合物施用至患有疾病或病症的受试者。治疗可以通过预防与疾病或病症相关的症状在对疾病或病症易感(例如遗传易感的)的受试者中的发作而是预防性的。治疗的理想效果包括但不限于预防疾病的发生或复发、减轻症状、减少疾病的任何直接或间接病理后果、降低疾病进展的速度、改善或缓解疾病状态和缓解或改善的预后。

[0161] 药物组合物是用于预防疾病或病症、降低疾病或病症的强度、治愈或以其他方式治疗疾病或病症的组合物，其包含活性成分(即USP或融合蛋白或编码其的核酸分子)和药学上可接受的承载体。药学上可接受的承载体是指一种或多种相容的固体或液体填充剂、稀释剂或包封物质，其适合施用于人或其他脊椎动物。在一些实施方案中，药物组合物包含非天然存在的药学上可接受的承载体。

[0162] 用于本文公开的方法的药物组合物可以与合适的承载体、赋形剂和提供合适的转移、递送、耐受性等的其他试剂一起配制。许多合适的制剂是本领域技术人员已知的。非限制性实例包括无菌稀释剂，例如注射用水、盐水溶液、固定油、聚乙二醇、甘油、丙二醇或其他合成溶剂；抗细菌剂例如苯甲醇或对羟基苯甲酸甲酯；抗氧化剂例如抗坏血酸或亚硫酸氢钠；螯合剂例如乙二胺四乙酸；缓冲液，例如醋酸盐、柠檬酸盐或磷酸盐，以及用于调节张力的试剂，例如氯化钠或葡萄糖。静脉内施用的特定承载体是生理盐水或磷酸盐缓冲盐水(PBS)。可以将用于口服或肠胃外使用的药物组合物制备成适合于活性成分剂量的单位剂量的剂型。这种单位剂量的剂型包括例如片剂、丸剂、胶囊剂、注射剂(安瓿剂)、栓剂等。这些组合物还可以含有助剂，包括防腐剂、润湿剂、乳化剂和分散剂。可通过各种抗细菌剂和抗真菌剂(例如对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸等)确保防止微生物的作用。还可能需要包括等渗剂，例如糖、氯化钠等。可通过使用延迟吸收剂例如单硬脂酸铝和明胶来实现可注射药物形式的延长的吸收。

[0163] 包含USP或融合蛋白或编码其的核酸分子或包含它们的细胞的药物组合物可以通过任何途径施用于受试者，例如口服、含服、肠胃外、局部、通过吸入或吹入(即通过口或通

过鼻)或经直肠。施用可以进行例如一次、多次和/或在一个或多个延长的时期内进行。

[0164] 本发明的药物组合物的有效量是有效实现其目的(例如,预防由特定序列引起的病症或疾病或从由特定序列引起的病症或疾病中恢复(包括部分恢复)或预防或减缓由特定序列引起的病症或疾病)的任何量。通常以mg/kg表示的有效量可以由本领域技术人员在临床前和临床试验期间通过常规方法确定。

[0165] 在其中方法包括递送编码引导RNA和/或融合蛋白的多核苷酸的那些实施方案中,随后可以在表达引导RNA和/或融合蛋白的条件下培养细胞或胚胎。在各种实施方案中,所述方法包括使靶序列与包含gRNA和融合蛋白(其包含本发明的USP、脱氨酶和RNA引导的DNA结合多肽)的核糖核蛋白复合物接触。在某些实施方案中,所述方法包括将本发明的核糖核蛋白复合物引入包含靶序列的细胞、细胞器或胚胎中。本发明的核糖核蛋白复合物可以是已经从生物样品中纯化、重组产生并随后纯化或如本文所述体外组装的核糖核蛋白复合物。在其中与靶序列或细胞器或胚胎接触的核糖核蛋白复合物已在体外组装的那些实施方案中,所述方法可进一步包括在与靶序列、细胞、细胞器或胚胎接触之前体外组装复合物。

[0166] 可以使用本领域已知的任何方法(包括但不限于电穿孔)将本发明的纯化或体外组装的核糖核蛋白复合物引入细胞、细胞器或胚胎。可选择地,可以使用任何本领域已知的任何方法(例如,电穿孔)将包含本发明的USP、脱氨酶和RNA引导的DNA结合多肽的融合蛋白以及编码或包含引导RNA的多核苷酸引入细胞、细胞器或胚胎中。

[0167] 在递送至或接触靶序列或包含靶序列的细胞、细胞器或胚胎后,引导RNA指导融合蛋白以序列特异性方式结合至靶序列。随后可以通过融合蛋白的脱氨酶结构域和USP结构域对靶序列进行修饰。在一些实施方案中,该融合蛋白与靶序列的结合导致与靶序列相邻的核苷酸的修饰。与靶序列相邻的核苷酸碱基可以是距靶序列的5'或3'端的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个碱基对。与包含单独的脱氨酶和RNA引导的DNA结合多肽的融合蛋白相比,包含本发明的USP、脱氨酶和RNA引导的DNA结合多肽的融合蛋白可以更有效地引入靶向C>T突变。

[0168] 测量融合蛋白与靶序列结合的方法是本领域已知的并且包括染色质免疫沉淀测定、凝胶迁移率变动测定、DNA下拉测定、报告因子测定、微孔板捕获和检测测定。类似地,测量靶序列的切割或修饰的方法在本领域中是已知的并且包括体外或体内切割测定,其中,在有或没有适当的标记(例如,放射性同位素、荧光物质)与靶序列的附接以促进降解产物的检测的情况下使用PCR、测序或凝胶电泳确认切割。可选择地,可以使用切口触发的指数扩增反应(NTEXPAR)测定(参见,例如,Zhang等人(2016)Chem.Sci.7:4951-4957)。可以使用Surveyor测定评估体内切割(Guschin等人(2010)Methods Mol Biol649:247-256)。

[0169] 在一些实施方案中,所述方法涉及与多于一种引导RNA复合的作为融合蛋白的一部分的RNA结合的DNA引导的结构域的使用。多于一种引导RNA可以靶向单个基因的不同区域或可以靶向多个基因。这种多重靶向使融合蛋白的脱氨酶结构域能够修饰核酸,从而在感兴趣的靶核酸分子(例如,基因组)中引入多重突变。融合蛋白的USP结构域提高了引入所需突变的效率。

[0170] 在其中该方法涉及使用RNA引导的核酸酶(RGN)例如切口酶RGN(即,仅能够切割双链多核苷酸的单条链,例如nAPG07433.1(SEQ ID NO:41))的那些实施方案中,该方法可包

括引入两种不同的RGN或RGN变体，其靶向相同或重叠的靶序列并切割多核苷酸的不同链。例如，可以将仅切割双链多核苷酸的正(+)链的RGN切口酶与仅切割双链多核苷酸的负(-)链的第二RGN切口酶一起引入。可选择地，可以提供两种不同的融合蛋白，其中每种融合蛋白包含具有不同PAM识别序列的不同RGN，使得可以靶向更大多样性的核苷酸序列以用于突变。

[0171] 本领域的普通技术人员将理解，任何本文公开的方法都可以用于靶向单个靶序列或多个靶序列。因此，方法包括使用融合蛋白，该融合蛋白包含与可以靶向单个基因和/或多个基因内的多个不同的序列的多个不同的引导RNA组合的单个RNA引导的DNA结合多肽。然后融合蛋白的脱氨酶结构域会在每个靶序列处引入突变。融合蛋白的USP结构域提高了引入所需突变的效率。本文还包括其中引入与多种不同的RNA引导的DNA结合多肽组合的多种不同的引导RNA的方法。此类RNA引导的DNA结合多肽可以是多种RGN或RGN变体。这些引导RNA和引导RNA/融合蛋白系统可以靶向单个基因和/或多个基因内的多个不同序列。

[0172] IX.包含多核苷酸遗传修饰的细胞

[0173] 一些实施方案提供了使用本文提供的融合蛋白的方法。在一些实施方案中，融合蛋白用于通过使靶核碱基例如C残基脱氨基将点突变引入靶核酸分子中。在一些实施方案中，靶核碱基的脱氨基导致遗传缺陷的校正，例如导致基因产物功能丧失的点突变的校正。在一些实施方案中，遗传缺陷与疾病或病症(例如溶酶体贮积症或代谢疾病，例如I型糖尿病)相关。在一些实施方案中，本文提供的方法用于将失活点突变引入编码与疾病或病症相关的基因产物的基因或等位基因中。例如，在一些实施方案中，本文提供了使用融合蛋白将失活点突变引入致癌基因的方法(例如，在增殖性疾病的治疗中)。在一些实施方案中，失活突变可在编码序列中产生过早的终止密码子，这导致截短基因产物(例如缺乏全长蛋白质的功能的截短蛋白质)的表达。在一些实施方案中，本文提供的方法的目的是通过基因组编辑恢复功能障碍基因的功能。本文提供的融合蛋白可以例如通过校正人类细胞培养物中的疾病相关突变在体外针对基于基因编辑的人类治疗进行验证。技术人员将理解本文提供的融合蛋白例如包含RNA引导的DNA结合结构域、脱氨酶结构域和本发明的USP的融合蛋白可用于校正任何单点C>T突变。将突变C至U的脱氨基导致突变的校正。

[0174] 在一些实施方案中，包含RNA引导的DNA结合结构域、脱氨酶结构域和本发明的USP的融合蛋白可用于在感兴趣的靶基因或基因的靶区域中产生突变。在一些实施方案中，本发明的融合蛋白可用于感兴趣的靶基因或靶基因区的域的饱和诱变，然后进行高通量正向遗传筛选以鉴定新突变和/或表型。在其他实施方案中，本文所述的融合蛋白可用于在靶基因组位置产生突变，其可能包含或可能不包含编码DNA序列。通过上述靶向诱变产生的细胞系的文库也可用于基因功能或基因表达的研究。

[0175] 本发明的融合蛋白还可用于通过无义突变的靶向插入有效地产生敲除(KO)系，包括KO系的整个文库。包含RNA引导的DNA结合结构域、脱氨酶结构域和本发明的USP的融合蛋白如果靶向编码DNA株系则可以将3个密码子(CAA、CAG和CGA)转化为终止密码子(TAG、TAA或TGA)，并且如果靶向非编码DNA株系则可以将TGG转化为终止密码子。Billon等人(2017，Mol Cell 67:1068-1079;通过引用并入本文)提供了可用于生成终止密码子的八种真核生物物种中超过340万个引导RNA的数据库。此外，Kuscu等人(2017,Nature Methods 14(7):710-714;通过引用并入本文)在人类基因组中鉴定出约260,000个独特的i-stop sgRNA，它

们可以靶向近17,000个基因以引入早期终止密码子。在一些实施方案中,KO系是真核细胞。在其他实施方案中,KO系是原核细胞。在一些实施方案中,使用本发明的融合蛋白产生的KO系是人细胞系。在其他实施方案中,KO细胞系是哺乳动物细胞系,例如小鼠、大鼠、猴子、猫、狗、牛、猪、羊或马细胞系。在其他实施方案中,KO系是禽类细胞。在其他实施方案中,KO系是昆虫细胞。在其他实施方案中,KO系是微生物细胞。在其他实施方案中,KO系是植物细胞。在进一步的实施方案中,KO系是拟南芥、大豆、玉蜀黍、棉花、番茄、马铃薯或豆类细胞。在进一步的实施方案中,细胞系是植物种子。

[0176] 在一些实施方案中,本文提供的融合蛋白可用于治疗性基因组编辑。例如,包含RNA引导的DNA结合结构域、脱氨酶结构域和本发明的USP的融合蛋白可用于产生PCSK9(前蛋白转化酶枯草杆菌蛋白酶/kexin 9型)的靶向无义突变。PCSK9参与脂蛋白稳态,并且阻断PCSK9的药物可以降低血液中的低密度脂蛋白颗粒(LDL)浓度。个体中PCSK9中天然发生的无义变体导致显著降低的血液胆固醇水平并降低患冠心病的风险(Cohen等人(2006)N Engl J Med 354:1264-1272)。Chadwick等人(2017,Arteriscler Thromb Vasc Biol 37:1741-1747;通过引用并入本文)发现他们可以成功地将靶向的C>T突变引入PCSK9基因,从而降低小鼠中的PCSK9蛋白水平和血浆胆固醇水平。在一些实施方案中,可以对本发明的融合蛋白采取类似的方法。此外,本领域技术人员将理解,本文公开的融合蛋白可用于校正其他点突变和与其他癌症和与癌症以外的疾病包括其他增殖性疾病相关的突变。

[0177] 本文提供包含感兴趣的靶核酸分子的细胞和生物体,所述靶核酸分子已使用由融合蛋白介导的过程修饰,任选地用本文所述的gRNA。在这些实施方案的一些中,融合蛋白包含含有SEQ ID NO:1-16中的任一个的氨基酸序列或其活性变体或片段的USP。在一些实施方案中,融合蛋白包含USP,其包含与SEQ ID NO:1-16中的任一个至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%相同的氨基酸序列。在一些实施方案中,融合蛋白还包含脱氨酶和RNA引导的DNA结合多肽。在进一步的实施方案中,融合蛋白包含脱氨酶和RGN或其变体,例如与APG07433.1(SEQ ID NO:40)或其切口酶变体nAPG07433.1(SEQ ID NO:41)具有至少80%序列同一性的氨基酸序列。在进一步的实施方案中,融合蛋白包含脱氨酶和RGN或其变体,例如与APG07433.1(SEQ ID NO:40)或其切口酶变体nAPG07433.1(SEQ ID NO:41)具有至少85%序列同一性的氨基酸序列。在进一步的实施方案中,融合蛋白包含脱氨酶和RGN或其变体,例如与APG07433.1(SEQ ID NO:40)或其切口酶变体nAPG07433.1(SEQ ID NO:41)具有至少90%序列同一性的氨基酸序列。在进一步的实施方案中,融合蛋白包含脱氨酶和RGN或其变体,例如与APG07433.1(SEQ ID NO:40)或其切口酶变体nAPG07433.1(SEQ ID NO:41)具有至少95%序列同一性的氨基酸序列。在进一步的实施方案中,融合蛋白包含脱氨酶和RGN或其变体,例如APG07433.1(SEQ ID NO:40)或其切口酶变体nAPG07433.1(SEQ ID NO:41)。在其他实施方案中,融合蛋白包含脱氨酶和Cas9或其变体,例如dCas9或切口酶Cas9。在一些实施方案中,融合蛋白包含II型CRISPR-Cas多肽的核酸酶失活或切口酶变体。在其他实施方案中,融合蛋白包含V型CRISPR-Cas多肽的核酸酶失活或切口酶变体。在其他实施方案中,融合蛋白包含VI型CRISPR-Cas多肽的核酸酶失活或切口酶变体。

[0178] 修饰的细胞可以是真核细胞(例如哺乳动物、植物、昆虫、鸟类细胞)或原核细胞。还提供了包含已经通过利用本文所述的融合蛋白的方法进行了修饰的至少一个核苷酸序

列的细胞器和胚胎。对于修饰的核苷酸序列，遗传修饰的细胞、生物体、细胞器和胚胎可以是杂合的或纯合的。

[0179] 通过融合蛋白的脱氨酶结构域引入的(多个)突变可导致改变的表达(上调或下调)、失活或改变的蛋白质产物或整合序列的表达。在突变导致基因失活或非功能性蛋白质产物表达的那些情况下，遗传修饰的细胞、生物体、细胞器或胚胎被称为是“敲除”的。敲除表型可以是缺失突变(即至少一个核苷酸的缺失)、插入突变(即至少一个核苷酸的插入)或无义突变(即至少一个核苷酸的取代以引入终止密码子)的结果。

[0180] 在其他实施方案中，由融合蛋白的脱氨酶结构域引入的突变导致产生变体蛋白产物。表达的变体蛋白质产物可以具有至少一个氨基酸取代和/或至少一个氨基酸的添加或缺失。与野生型蛋白质相比，变体蛋白质产物可以表现出修饰的特征或活性，包括但不限于改变的酶活性或底物特异性。

[0181] 在其他实施方案中，由融合蛋白的脱氨酶结构域引入的突变可导致改变的蛋白质表达模式。作为非限制性实例，控制蛋白质产物表达的调节区中的突变可导致蛋白质产物的过表达或下调或改变的组织或时间表达模式。

[0182] 本公开内容的一些方面提供包含融合蛋白的试剂盒，所述融合蛋白包含RNA引导的DNA结合多肽，例如RGN多肽，例如核酸酶失活的Cas9结构域，和本发明的脱氨酶，以及任选地位于Cas9结构域和脱氨酶之间的接头。此外，在一些实施方案中，试剂盒包含合适的试剂、缓冲液和/或使用融合蛋白(例如，用于体外或体内DNA或RNA编辑)的说明书。在一些实施方案中，试剂盒包含关于设计和使用合适的gRNA用于核酸序列的靶向编辑的说明书。

[0183] X:USP的其他应用

[0184] 本文描述的USP还具有超越基因组碱基编辑的效用。通常，USP可用于需要稳定DNA分子中的尿嘧啶核碱基的应用。通过天然过程或人为操作，尿嘧啶可能因由活性氧、电离辐射和/或烷化剂引起的DNA损伤而被引入基因组DNA。对DNA修复机制(例如碱基切除修复(BER)途径)的研究或测量DNA修复能力的研究可以使用本发明的USP来抑制尿嘧啶的修复。

[0185] 此外，本文所述的USP可用于治疗各种癌症。例如，包括5-氟尿嘧啶(5-FU)及其脱氧核糖核苷代谢物5-氟脱氧尿苷(5-FdU)在内的氟嘧啶类药物已广泛用于治疗各种实体瘤，包括结直肠癌。5-FdU通过抑制胸苷酸合酶发挥活性，从而将尿嘧啶和5-FU掺入细胞基因组中。如上所述，碱基修复酶例如UDG识别基因组DNA中的尿嘧啶核碱基并将其去除。Yan等人(2016;Oncotarget 7(37):59299-59313)发现UDG耗尽的细胞在5-FdU处理后被阻滞并表现出持续的DNA损伤，表明在去除尿嘧啶和5-FU中的UDG作用在肿瘤的5-FdU治疗的有效性中发挥作用。与氟嘧啶组合的本发明的USP的递送可以增强这种肿瘤治疗的有效性。因此，提供了包含本文公开的USP和氟嘧啶的药物组合物，以及用有效量的此类药物组合物治疗癌症的方法。氟嘧啶是一类抗癌抗代谢物，包括卡培他滨、卡莫氟、多西氟尿苷、氟尿嘧啶、5-氟脱氧尿苷和替加氟。

[0186] 冠词“一个”和“一种”在本文中用于指代冠词的一个或多于一个/一种或多于一种(即，至少一个/至少一种)语法对象。举例来说，“一个/一种多肽”是指一个或多个/一种或多种多肽。

[0187] 说明书中提及的所有出版物和专利申请都表明了本公开内容所属领域的技术人员的水平。所有出版物和专利申请均以引用方式并入本文，其程度如同每个单独的出版物

或专利申请都具体且单独地指明以引用方式并入。

[0188] 尽管为了清楚理解的目的,已经通过说明和示例的方式对前述发明进行了一些详细描述,但是显然可以在所附权利要求的范围内实施某些改变和修改。

[0189] 非限制性实施方案包括:

[0190] 1.一种分离的多肽,其包含:

[0191] a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

[0192] b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

[0193] c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

[0194] 其中所述多肽具有尿嘧啶稳定活性并且其中所述多肽还包含异源氨基酸序列。

[0195] 2.实施方案1的分离的多肽,其中所述氨基酸序列与SEQ ID NO:1-16中的任一个具有至少85%的序列同一性。

[0196] 3.实施方案1的分离的多肽,其中所述氨基酸序列与SEQ ID NO:1-16中的任一个具有至少90%的序列同一性。

[0197] 4.实施方案1的分离的多肽,其中所述氨基酸序列与SEQ ID NO:1-16中的任一个具有至少95%的序列同一性。

[0198] 5.实施方案1的分离的多肽,其中所述氨基酸序列与SEQ ID NO:1-16中的任一个具有100%的序列同一性。

[0199] 6.实施方案1的分离的多肽,其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

[0200] 7.一种药物组合物,包含非天然存在的药学上可接受的承载体和多肽,所述多肽包含:

[0201] a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

[0202] b)与SEQ ID NO:3或16具有至少81%序列同一性的氨基酸序列;或

[0203] c)与SEQ ID NO:6具有至少82%序列同一性的氨基酸序列;

[0204] 其中所述多肽具有尿嘧啶稳定活性。

[0205] 8.实施方案7的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的氨基酸序列。

[0206] 9.实施方案7的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的氨基酸序列。

[0207] 10.实施方案7的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的氨基酸序列。

[0208] 11.实施方案7的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有100%序列同一性的氨基酸序列。

[0209] 12.一种药物组合物,其包含非天然存在的药学上可接受的承载体和包含编码多肽的多核苷酸的核酸分子,所述多肽包含:

[0210] a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80%序列同一性的氨基酸序列;

[0211] b)与SEQ ID NO:3或16具有至少81％序列同一性的氨基酸序列;或

[0212] c)与SEQ ID NO:6具有至少82％序列同一性的氨基酸序列;

[0213] 其中所述多肽具有尿嘧啶稳定活性。

[0214] 13.实施方案12的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少85％序列同一性的氨基酸序列。

[0215] 14.实施方案12的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少90％序列同一性的氨基酸序列。

[0216] 15.实施方案12的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少95％序列同一性的氨基酸序列。

[0217] 16.实施方案12的药物组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有100％序列同一性的氨基酸序列。

[0218] 17.实施方案7或12的药物组合物,其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

[0219] 18.实施方案7-17中任一项的药物组合物,其还包含氟嘧啶。

[0220] 19.一种包含编码多肽的多核苷酸的核酸分子,所述多肽包含:

[0221] a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80％序列同一性的氨基酸序列;

[0222] b)与SEQ ID NO:3或16具有至少81％序列同一性的氨基酸序列;或

[0223] c)与SEQ ID NO:6具有至少82％序列同一性的氨基酸序列;

[0224] 其中所述多肽具有尿嘧啶稳定活性;和

[0225] 其中所述核酸分子还包含可操作地连接至所述多核苷酸的异源启动子。

[0226] 20.实施方案19的核酸分子,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少85％序列同一性的氨基酸序列。

[0227] 21.实施方案19的核酸分子,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少90％序列同一性的氨基酸序列。

[0228] 22.实施方案19的核酸分子,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少95％序列同一性的氨基酸序列。

[0229] 23.实施方案19的核酸分子,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有100％序列同一性的氨基酸序列。

[0230] 24.实施方案19的核酸分子,其中多肽具有SEQ ID NO:33-39中的任一个的序列。

[0231] 25.一种包含氟嘧啶和多肽的组合物,所述多肽包含:

[0232] a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80％序列同一性的氨基酸序列;

[0233] b)与SEQ ID NO:3或16具有至少81％序列同一性的氨基酸序列;或

[0234] c)与SEQ ID NO:6具有至少82％序列同一性的氨基酸序列;

[0235] 其中所述多肽具有尿嘧啶稳定活性。

[0236] 26.实施方案25的组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少85％序列同一性的氨基酸序列。

[0237] 27.实施方案25的组合物,其中所述多肽包含与SEQ ID NO:1-16中的任一个具有

至少90％序列同一性的氨基酸序列。

[0238] 28.实施方案25的组合物，其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少95％序列同一性的氨基酸序列。

[0239] 29.实施方案25的组合物，其中所述多肽包含与SEQ ID NO:1-16中的任一个具有100％序列同一性的氨基酸序列。

[0240] 30.一种包含氟嘧啶和编码多肽的核酸分子的组合物，所述多肽包含：

[0241] a)与SEQ ID NO:1、2、4、5和7-15中的任一个具有至少80％序列同一性的氨基酸序列；

[0242] b)与SEQ ID NO:3或16具有至少81％序列同一性的氨基酸序列；或

[0243] c)与SEQ ID NO:6具有至少82％序列同一性的氨基酸序列；

[0244] 其中所述多肽具有尿嘧啶稳定活性。

[0245] 31.实施方案30的组合物，其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少85％序列同一性的氨基酸序列。

[0246] 32.实施方案30的组合物，其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少90％序列同一性的氨基酸序列。

[0247] 33.实施方案30的组合物，其中所述多肽包含与SEQ ID NO:1-16中的任一个具有至少95％序列同一性的氨基酸序列。

[0248] 34.实施方案30的组合物，其中所述多肽包含与SEQ ID NO:1-16中的任一个具有100％序列同一性的氨基酸序列。

[0249] 35.实施方案25或30的组合物，其中所述多肽具有SEQ ID NO:33-39中的任一个的序列。

[0250] 36.一种融合蛋白，其包含：(i)DNA结合多肽；(ii)脱氨酶；和(iii)至少一种尿嘧啶稳定多肽(USP)，其与SEQ ID NO:1-16中的任一个具有至少80％的序列同一性。

[0251] 37.实施方案36的融合蛋白，其中至少一种USP与SEQ ID NO:1-16中的任一个具有至少85％的序列同一性。

[0252] 38.实施方案36的融合蛋白，其中至少一种USP与SEQ ID NO:1-16中的任一个具有至少90％的序列同一性。

[0253] 39.实施方案36的融合蛋白，其中至少一种USP与SEQ ID NO:1-16中的任一个具有至少95％的序列同一性。

[0254] 40.实施方案36的融合蛋白，其中至少一种USP与SEQ ID NO:1-16中的任一个具有100％的序列同一性。

[0255] 41.实施方案36的融合蛋白，其中USP具有SEQ ID NO:33-39中的任一个的序列。

[0256] 42.实施方案36或41的融合蛋白，其中脱氨酶是胞苷脱氨酶。

[0257] 43.实施方案42的融合蛋白，其中胞苷脱氨酶是活化诱导的胞苷脱氨酶(AID)或脱氨酶的载脂蛋白B mRNA编辑复合物(APOBEC)家族的成员。

[0258] 44.实施方案43的融合蛋白，其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少80％序列同一性的氨基酸序列。

[0259] 45.实施方案43的融合蛋白，其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少85％序列同一性的氨基酸序列。

[0260] 46.实施方案43的融合蛋白,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少90%序列同一性的氨基酸序列。

[0261] 47.实施方案43的融合蛋白,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少95%序列同一性的氨基酸序列。

[0262] 48.实施方案43的融合蛋白,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有100%序列同一性的氨基酸序列。

[0263] 49.实施方案36-44中任一项的融合蛋白,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

[0264] 50.实施方案36-44中任一项的融合蛋白,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

[0265] 51.实施方案50的融合蛋白,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

[0266] 52.实施方案51的融合蛋白,其中RGN是II型CRISPR-Cas多肽。

[0267] 53.实施方案51的融合蛋白,其中RGN是V型CRISPR-Cas多肽。

[0268] 54.实施方案51的融合蛋白,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少80%序列同一性的氨基酸序列。

[0269] 55.实施方案51的融合蛋白,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少85%序列同一性的氨基酸序列。

[0270] 56.实施方案51的融合蛋白,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少90%序列同一性的氨基酸序列。

[0271] 57.实施方案51的融合蛋白,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少95%序列同一性的氨基酸序列。

[0272] 58.实施方案51的融合蛋白,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有100%序列同一性的氨基酸序列。

[0273] 59.实施方案51-58中任一项的融合蛋白,其中RGN是RGN切口酶。

[0274] 60.实施方案51-59中任一项的融合蛋白,其中所述融合蛋白包含RGN、胞苷脱氨酶和USP。

[0275] 61.实施方案60的融合蛋白,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少80%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少80%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少80%序列同一性的USP。

[0276] 62.实施方案60的融合蛋白,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少85%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少85%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的USP。

[0277] 63.实施方案60的融合蛋白,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少90%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少90%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的USP。

[0278] 64.实施方案60的融合蛋白,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少95%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少95%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的USP。

[0279] 65.实施方案60的融合蛋白,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有100%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有100%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有100%序列同一性的USP。

[0280] 66.实施方案36-66中任一项的融合蛋白,其中所述融合蛋白还包含至少一个核定位信号(NLS)。

[0281] 67.一种包含编码融合蛋白的多核苷酸的核酸分子,所述融合蛋白包含:(i)DNA结合多肽;(ii)脱氨酶;和(iii)至少一种尿嘧啶稳定多肽(USP),其中USP由以下核苷酸序列编码:

[0282] a)与SEQ ID NO:17-32中的任一个具有至少80%序列同一性的核苷酸序列,

[0283] b)在SEQ ID NO:17-32中的任一个示出的核苷酸序列,

[0284] c)编码与SEQ ID NO:1-16至少80%相同的氨基酸序列并且进一步具有SEQ ID NO:33-39中的任一个的序列的核苷酸序列,

[0285] d)编码与SEQ ID NO:1-16中的任一个示出的氨基酸序列至少80%相同的氨基酸序列的核苷酸序列,或

[0286] e)编码SEQ ID NO:1-16中的任一个示出的氨基酸序列的核苷酸序列。

[0287] 68.实施方案67的核酸分子,其中USP由以下核苷酸序列编码:

[0288] a)与SEQ ID NO:17-32中的任一个具有至少85%序列同一性的核苷酸序列,

[0289] b)在SEQ ID NO:17-32中的任一个示出的核苷酸序列,

[0290] c)编码与SEQ ID NO:1-16至少85%相同的氨基酸序列并且进一步具有SEQ ID NO:33-39中的任一个的序列的核苷酸序列,

[0291] d)编码与SEQ ID NO:1-16中的任一个示出的氨基酸序列至少85%相同的氨基酸序列的核苷酸序列,或

[0292] e)编码SEQ ID NO:1-16中的任一个示出的氨基酸序列的核苷酸序列。

[0293] 69.实施方案67的核酸分子,其中USP由以下核苷酸序列编码:

[0294] a)与SEQ ID NO:17-32中的任一个具有至少90%序列同一性的核苷酸序列,

[0295] b)在SEQ ID NO:17-32中的任一个示出的核苷酸序列,

[0296] c)编码与SEQ ID NO:1-16至少90%相同的氨基酸序列并且进一步具有SEQ ID NO:33-39中的任一个的序列的核苷酸序列,

[0297] d)编码与SEQ ID NO:1-16中的任一个示出的氨基酸序列至少90%相同的氨基酸序列的核苷酸序列,或

[0298] e)编码SEQ ID NO:1-16中的任一个示出的氨基酸序列的核苷酸序列。

[0299] 70.实施方案67的核酸分子,其中USP由以下核苷酸序列编码:

[0300] a)与SEQ ID NO:17-32中的任一个具有至少95%序列同一性的核苷酸序列,

[0301] b)在SEQ ID NO:17-32中的任一个示出的核苷酸序列,

[0302] c)编码与SEQ ID NO:1-16至少95%相同的氨基酸序列并且进一步具有SEQ ID

NO:33-39中的任一个的序列的核苷酸序列，

[0303] d)编码与SEQ ID NO:1-16中的任一个示出的氨基酸序列至少95％相同的氨基酸序列的核苷酸序列,或

[0304] e)编码SEQ ID NO:1-16中的任一个示出的氨基酸序列的核苷酸序列。

[0305] 71.实施方案67的核酸分子,其中USP由以下核苷酸序列编码：

[0306] a)在SEQ ID NO:17-32中的任一个示出的核苷酸序列，

[0307] b)编码与SEQ ID NO:1-16 100％相同的氨基酸序列并且进一步具有SEQ ID NO:33-39中的任一个的序列的核苷酸序列,或

[0308] c)编码SEQ ID NO:1-16中的任一个示出的氨基酸序列的核苷酸序列。

[0309] 72.实施方案67的核酸分子,其中脱氨酶是胞苷脱氨酶。

[0310] 73.实施方案72的核酸分子,其中胞苷脱氨酶是活化诱导的胞苷脱氨酶(AID)或脱氨酶的载脂蛋白B mRNA编辑复合物(APOBEC)家族的成员。

[0311] 74.实施方案73的核酸分子,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少80％序列同一性的氨基酸序列。

[0312] 75.实施方案73的核酸分子,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少85％序列同一性的氨基酸序列。

[0313] 76.实施方案73的核酸分子,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少90％序列同一性的氨基酸序列。

[0314] 77.实施方案73的核酸分子,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有至少95％序列同一性的氨基酸序列。

[0315] 78.实施方案73的核酸分子,其中胞苷脱氨酶包含与SEQ ID NO:47、48和76-94中的任一个具有100％序列同一性的氨基酸序列。

[0316] 79.实施方案67-78中任一项的核酸分子,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

[0317] 80.实施方案67-78中任一项的核酸分子,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

[0318] 81.实施方案80的核酸分子,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

[0319] 82.实施方案81的核酸分子,其中RGN是II型CRISPR-Cas多肽。

[0320] 83.实施方案81的核酸分子,其中RGN是V型CRISPR-Cas多肽。

[0321] 84.实施方案81的核酸分子,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少80％序列同一性的氨基酸序列。

[0322] 85.实施方案82的核酸分子,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少85％序列同一性的氨基酸序列。

[0323] 86.实施方案81的核酸分子,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少90％序列同一性的氨基酸序列。

[0324] 87.实施方案81的核酸分子,其中RGN包含与SEQ ID NO:40和95-142中的任一个具有至少95％序列同一性的氨基酸序列。

[0325] 88.实施方案81的核酸分子,其中RGN包含与SEQ ID NO:40和95-142中的任一个具

有100%序列同一性的氨基酸序列。

[0326] 89.实施方案81-88中任一项的核酸分子,其中RGN是RGN切口酶。

[0327] 90.实施方案81-89中任一项的核酸分子,其中融合蛋白包含RGN、胞苷脱氨酶和USP。

[0328] 91.实施方案90的核酸分子,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少80%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少80%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少80%序列同一性的USP。

[0329] 92.实施方案90的核酸分子,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少85%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少85%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的USP。

[0330] 93.实施方案90的核酸分子,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少90%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少90%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的USP。

[0331] 94.实施方案90的核酸分子,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少95%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少95%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的USP。

[0332] 95.实施方案90的核酸分子,其中融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有100%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有100%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有100%序列同一性的USP。

[0333] 96.实施方案67-95中任一项的核酸分子,其中编码融合蛋白的多核苷酸在其5'端可操作地连接至异源启动子。

[0334] 97.实施方案67-95中任一项的核酸分子,其中编码融合蛋白的多核苷酸在其3'端可操作地连接至异源终止子。

[0335] 98.实施方案67-97中任一项的核酸分子,其中所述融合蛋白包含一个或多个核定位信号。

[0336] 99.实施方案67-98中任一项的核酸分子,其中所述融合蛋白针对在真核细胞中的表达进行了密码子优化。

[0337] 100.实施方案67-99中任一项的核酸分子,其中所述融合蛋白针对在原核细胞中的表达进行了密码子优化。

[0338] 101.一种包含实施方案67-100中任一项的核酸分子的载体。

[0339] 102.一种包含实施方案81-95中任一项的核酸分子的载体,其还包含编码能够与靶序列杂交的引导RNA(gRNA)的至少一个核苷酸序列。

[0340] 103.实施方案102的载体,其中gRNA是单引导RNA。

[0341] 104.实施方案102的载体,其中gRNA是双引导RNA。

[0342] 105.一种包含实施方案36-66中任一项的融合蛋白的细胞。

[0343] 106.一种包含实施方案51-66中任一项的融合蛋白的细胞,其中所述细胞还包含引导RNA。

[0344] 107.一种包含实施方案67-100中任一项的核酸分子的细胞。

[0345] 108.一种包含实施方案101至104中任一项的载体的细胞。

[0346] 109.实施方案105-108中任一项的细胞,其中所述细胞是原核细胞。

[0347] 110.实施方案105-108中任一项的细胞,其中所述细胞是真核细胞。

[0348] 111.实施方案110的细胞,其中所述细胞是昆虫、鸟类或哺乳动物细胞。

[0349] 112.实施方案110的细胞,其中所述细胞是植物细胞或真菌细胞。

[0350] 113.一种药物组合物,其包含药学上可接受的承载体和实施方案19-24和67-100中任一项的核酸分子、实施方案25-35中任一项的组合物、实施方案36-66中任一项的融合蛋白、实施方案101-104中任一项的载体或实施方案105-111中任一项的细胞。

[0351] 114.一种制备融合蛋白的方法,包括在表达融合蛋白的条件下培养实施方案105-112中任一项的细胞。

[0352] 115.一种制备融合蛋白的方法,包括将实施方案67-100中任一项的核酸分子或实施方案101-104中任一项的载体引入细胞中,并在表达融合蛋白的条件下培养细胞.

[0353] 116.实施方案114或115的方法,还包括纯化所述融合蛋白。

[0354] 117.一种用于制备RGN融合核糖核蛋白复合物的方法,包括将实施方案81-95中任一项的核酸分子和包含编码引导RNA的表达盒的核酸分子或实施方案102-104中任一项的载体引入细胞,以及在表达融合蛋白和gRNA并形成RGN融合核糖核蛋白复合物的条件下培养细胞。

[0355] 118.实施方案117的方法,还包括纯化所述RGN融合核糖核蛋白复合物。

[0356] 119.一种用于修饰包含靶DNA序列的靶DNA分子的系统,所述系统包括:

[0357] a)包含RNA引导的核酸酶多肽(RGN)、胞苷脱氨酶和至少一种尿嘧啶稳定多肽(USP)的融合蛋白,其中USP与SEQ ID NO:1-16中的任一个至少80%相同,或编码所述融合蛋白的核苷酸序列;和

[0358] b)能够与所述靶DNA序列杂交的一种或多种引导RNA或编码所述一种或多种引导RNA(gRNA)的一种或多种核苷酸序列;

[0359] 其中编码一种或多种引导RNA和编码融合蛋白的所述核苷酸序列各自可操作地连接至与所述核苷酸序列异源的启动子;

[0360] 和

[0361] 其中所述一种或多种引导RNA能够与融合蛋白形成复合物以引导所述融合蛋白结合所述靶DNA序列并修饰靶DNA分子。

[0362] 120.实施方案119的系统,其中USP与SEQ ID NO:1-16中的任一个至少85%相同。

[0363] 121.实施方案119的系统,其中USP与SEQ ID NO:1-16中的任一个至少90%相同。

[0364] 122.实施方案119的系统,其中USP与SEQ ID NO:1-16中的任一个至少95%相同。

[0365] 123.实施方案119的系统,其中USP与SEQ ID NO:1-16中的任一个100%相同。

[0366] 124.实施方案119的系统,其中USP包含SEQ ID NO:33-39中的任一个示出的序列。

[0367] 125.实施方案119-124中任一项的系统,其中靶DNA序列位于邻近由RGN识别的原前间区序列邻近基序(PAM)的位置。

[0368] 126.实施方案119-125中任一项的系统,其中靶DNA分子在细胞内。

[0369] 127.实施方案126的系统,其中细胞是真核细胞。

[0370] 128.实施方案127的系统,其中真核细胞是植物细胞。

[0371] 129.实施方案127的系统,其中真核细胞是哺乳动物细胞。

[0372] 130.实施方案127的系统,其中真核细胞是昆虫细胞。

[0373] 131.实施方案126的系统,其中所述细胞是原核细胞。

[0374] 132.实施方案119-131中任一项的系统,其中融合蛋白的RGN是II型CRISPR-Cas多肽。

[0375] 133.实施方案119-131中任一项的系统,其中融合蛋白的RGN是V型CRISPR-Cas多肽。

[0376] 134.实施方案119-131中任一项的系统,其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少80%相同。

[0377] 135.实施方案119-131中任一项的系统,其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少85%相同。

[0378] 136.实施方案119-131中任一项的系统,其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个90%相同。

[0379] 137.实施方案119-131中任一项的系统,其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个95%相同。

[0380] 138.实施方案119-131中任一项的系统,其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个100%相同。

[0381] 139.实施方案119-138中任一项的系统,其中融合蛋白的RGN是RGN切口酶。

[0382] 140.实施方案119-139中任一项的系统,其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少80%相同。

[0383] 141.实施方案119-139中任一项的系统,其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少85%相同。

[0384] 142.实施方案119-139中任一项的系统,其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少90%相同。

[0385] 143.实施方案119-139中任一项的系统,其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少95%相同。

[0386] 144.实施方案119-139中任一项的系统,其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个100%相同。

[0387] 145.实施方案119-144中任一项的系统,其中融合蛋白包含与SEQ ID NO:40、41、95和142中的任一个具有至少80%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少80%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少80%序列同一性的USP。

[0388] 146.实施方案119-144中任一项的系统,其中融合蛋白包含与SEQ ID NO:40、41、95和142中的任一个具有至少85%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少85%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的USP。

[0389] 147.实施方案119-144中任一项的系统,其中融合蛋白包含与SEQ ID NO:40、41、95和142中的任一个具有至少90%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少90%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的USP。

[0390] 148.实施方案119-144中任一项的系统,其中融合蛋白包含与SEQ ID NO:40、41、95和142中的任一个具有至少95%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有至少95%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的USP。

[0391] 149.实施方案119-144中任一项的系统,其中融合蛋白包含与SEQ ID NO:40、41、95和142中的任一个具有100%序列同一性的RGN,与SEQ ID NO:47、48和76-94中的任一个具有100%序列同一性的胞苷脱氨酶,以及与SEQ ID NO:1-16中的任一个具有100%序列同一性的USP。

[0392] 150.实施方案119-149中任一项的系统,其中所述融合蛋白包含一个或多个核定位信号。

[0393] 151.实施方案119-150中任一项的系统,其中所述融合蛋白针对在真核细胞中的表达进行了密码子优化。

[0394] 152.实施方案119-151中任一项的系统,其中编码一种或多种引导RNA的核苷酸序列和编码融合蛋白的核苷酸序列位于一个载体上。

[0395] 153.一种修饰包含靶DNA序列的靶DNA分子的方法,所述方法包括将实施方案119-152中任一项的系统递送至所述靶DNA分子或包含靶DNA分子的细胞。

[0396] 154.实施方案153的方法,其中所述修饰的靶DNA分子包含靶DNA分子内至少一个核苷酸的C>T突变。

[0397] 155.实施方案153的方法,其中所述修饰的靶DNA分子包含靶DNA序列内至少一个核苷酸的C>T突变。

[0398] 156.一种修饰包含靶序列的靶DNA分子的方法,包括:

[0399] a)通过在适合形成RGN-脱氨酶-USP核糖核苷酸复合物的条件下组合以下i)和ii)在体外组装RGN-脱氨酶-USP核糖核苷酸复合物:

[0400] i)能够与靶DNA序列杂交的一种或多种引导RNA;和

[0401] ii)包含RNA引导的核酸酶多肽(RGN)、胞苷脱氨酶和至少一种尿嘧啶稳定多肽(USP)的融合蛋白,其中USP与SEQ ID NO:1-16中的任一个至少80%相同;和

[0402] b)使所述靶DNA分子或包含所述靶DNA分子的细胞与体外组装的RGN-脱氨酶-USP核糖核苷酸复合物接触;

[0403] 其中所述一种或多种引导RNA与靶DNA序列杂交,从而指导所述融合蛋白结合所述靶DNA序列并且发生靶DNA分子的修饰。

[0404] 157.实施方案156的方法,其中USP与SEQ ID NO:1-16中的任一个至少85%相同。

[0405] 158.实施方案156的方法,其中USP与SEQ ID NO:1-16中的任一个至少90%相同。

[0406] 159.实施方案156的方法,其中USP与SEQ ID NO:1-16中的任一个至少95%相同。

[0407] 160.实施方案156的方法,其中USP与SEQ ID NO:1-16中的任一个100%相同。

[0408] 161.实施方案156的方法,其中USP包含SEQ ID NO:33-39中的任一个的序列。

[0409] 162.实施方案156-161中任一项的方法，其中所述修饰的靶DNA分子包含靶DNA分子内至少一个核苷酸的C＞T突变。

[0410] 163.实施方案156-161中任一项的方法，其中所述修饰的靶DNA分子包含靶DNA序列内至少一个核苷酸的C＞T突变。

[0411] 164.实施方案156-163中任一项的方法，其中融合蛋白的RGN是II型CRISPR-Cas多肽。

[0412] 165.实施方案156-163中任一项的方法，其中融合蛋白的RGN是V型CRISPR-Cas多肽。

[0413] 166.实施方案156-163中任一项的方法，其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少80％相同。

[0414] 167.实施方案156-163中任一项的方法，其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少85％相同。

[0415] 168.实施方案156-163中任一项的方法，其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少90％相同。

[0416] 169.实施方案156-163中任一项的方法，其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个至少95％相同。

[0417] 实施方案156-163中任一项的方法，其中融合蛋白的RGN与SEQ ID NO:40和95-142中的任一个100％相同。

[0418] 171.实施方案156-170中任一项的方法，其中融合蛋白的RGN是RGN切口酶。

[0419] 172.实施方案156-171中任一项的方法，其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少80％相同。

[0420] 173.实施方案156-171中任一项的方法，其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少85％相同。

[0421] 174.实施方案156-171中任一项的方法，其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少90％相同。

[0422] 175.实施方案156-171中任一项的方法，其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个至少95％相同。

[0423] 176.实施方案156-171中任一项的方法，其中胞苷脱氨酶与SEQ ID NO:47、48和76-94中的任一个100％相同。

[0424] 177.实施方案156-176中任一项的方法，其中所述融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少80％序列同一性的RGN，与SEQ ID NO:47、48和76-94中的任一个具有至少80％序列同一性的胞苷脱氨酶，以及与SEQ ID NO:1-16中的任一个具有至少80％序列同一性的USP。

[0425] 178.实施方案156-176中任一项的方法，其中所述融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少85％序列同一性的RGN，与SEQ ID NO:47、48和76-94中的任一个具有至少85％序列同一性的胞苷脱氨酶，以及与SEQ ID NO:1-16中的任一个具有至少85％序列同一性的USP。

[0426] 179.实施方案156-176中任一项的方法，其中所述融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少90％序列同一性的RGN，与SEQ ID NO:47、48和76-94中的

任一个具有至少90％序列同一性的胞苷脱氨酶，以及与SEQ ID NO:1-16中的任一个具有至少90％序列同一性的USP。

[0427] 180.实施方案156-176中任一项的方法，其中所述融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有至少95％序列同一性的RGN，与SEQ ID NO:47、48和76-94中的任一个具有至少95％序列同一性的胞苷脱氨酶，以及与SEQ ID NO:1-16中的任一个具有至少95％序列同一性的USP。

[0428] 181.实施方案156-176中任一项的方法，其中所述融合蛋白包含与SEQ ID NO:40、41和95-142中的任一个具有100％序列同一性的RGN，与SEQ ID NO:47、48和76-94中的任一个具有100％序列同一性的胞苷脱氨酶，以及与SEQ ID NO:1-16中的任一个具有100％序列同一性的USP。

[0429] 182.实施方案156-181中任一项的方法，其中所述融合蛋白包含一个或多个核定位信号。

[0430] 183.实施方案156-182中任一项的方法，其中所述融合蛋白针对在真核细胞中的表达进行了密码子优化。

[0431] 184.实施方案156-183中任一项的方法，其中所述靶DNA序列位于邻近前间区序列邻近基序(PAM)的位置。

[0432] 185.实施方案156-184中任一项的方法，其中靶DNA分子在细胞内。

[0433] 186.实施方案185的方法，其中所述细胞是真核细胞。

[0434] 187.实施方案186的方法，其中真核细胞是植物细胞。

[0435] 188.实施方案186的方法，其中真核细胞是哺乳动物细胞。

[0436] 189.实施方案186的方法，其中真核细胞是昆虫细胞。

[0437] 190.实施方案185的方法，其中所述细胞是原核细胞。

[0438] 191.实施方案185-190中任一项的方法，还包括选择包含所述修饰的DNA分子的细胞。

[0439] 192.一种包含根据实施方案191的方法的修饰的靶DNA序列的细胞。

[0440] 193.实施方案192的细胞，其中所述细胞是真核细胞。

[0441] 194.实施方案193的细胞，其中真核细胞是植物细胞。

[0442] 195.一种包含实施方案194的细胞的植物。

[0443] 196.一种包含实施方案194的细胞的种子。

[0444] 197.实施方案193的细胞，其中真核细胞是哺乳动物细胞。

[0445] 198.实施方案193的细胞，其中真核细胞是昆虫细胞。

[0446] 199.实施方案192的细胞，其中所述细胞是原核细胞。

[0447] 200.一种产生具有遗传性疾病的致病突变的校正的遗传修饰的细胞的方法，所述方法包括向细胞中引入：

[0448] a)包含RNA引导的核酸酶多肽(RGN)、胞苷脱氨酶和至少一种尿嘧啶稳定多肽(USP)的融合蛋白，其中USP与SEQ ID NO:1-16中的任一个至少80％相同，或编码所述融合蛋白的多核苷酸，其中编码融合蛋白的所述多核苷酸可操作地连接至启动子以使融合蛋白能够在细胞中表达；和

[0449] b)能够与靶DNA序列杂交的一种或多种引导RNA(gRNA)，或编码所述gRNA的多核苷

酸,其中编码gRNA的所述多核苷酸可操作地连接至启动子以使gRNA能够在细胞中表达;

[0450] 由此融合蛋白和gRNA靶向致病突变的基因组位置并修饰基因组序列以去除致病突变。

[0451] 201.实施方案200的方法,其中USP与SEQ ID NO:1-16中的任一个至少85%相同。

[0452] 202.实施方案200的方法,其中USP与SEQ ID NO:1-16中的任一个至少90%相同。

[0453] 203.实施方案200的方法,其中USP与SEQ ID NO:1-16中的任一个至少95%相同。

[0454] 204.实施方案200的方法,其中USP与SEQ ID NO:1-16中的任一个100%相同。

[0455] 205.实施方案200的方法,其中融合蛋白的所述RGN是切口酶。

[0456] 206.实施方案200或205的方法,其中USP包含SEQ ID NO:33-39中的任一个的序列。

[0457] 207.实施方案200-206中任一项的方法,其中基因组修饰包括在靶DNA序列内引入至少一个核苷酸的C>T突变。

[0458] 208.实施方案200-207中任一项的方法,其中细胞是动物细胞。

[0459] 209.实施方案208的方法,其中动物细胞是哺乳动物细胞。

[0460] 210.实施方案209的方法,其中所述细胞来源于狗、猫、小鼠、大鼠、兔、马、绵羊、山羊、牛、猪或人。

[0461] 211.实施方案200-210中任一项的方法,其中致病突变的校正包括引入终止密码子。

[0462] 212.一种组合物,其包含:

[0463] a)融合蛋白,其包含:(i)DNA结合多肽;和(ii)脱氨酶;或编码所述融合蛋白的核酸分子;和

[0464] b)尿嘧啶稳定多肽(USP),其与SEQ ID NO:1-16中的任一个具有至少80%序列同一性;或编码所述USP的核酸分子。

[0465] 213.实施方案212的组合物,其中USP与SEQ ID NO:1-16中的任一个具有至少85%的序列同一性。

[0466] 214.实施方案212的组合物,其中USP与SEQ ID NO:1-16中的任一个具有至少90%的序列同一性。

[0467] 215.实施方案212的组合物,其中USP与SEQ ID NO:1-16中的任一个具有至少95%的序列同一性。

[0468] 216.实施方案212的组合物,其中USP与SEQ ID NO:1-16中的任一个具有100%的序列同一性。

[0469] 217.实施方案212的组合物,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少80%序列同一性的尿嘧啶稳定多肽(USP)。

[0470] 218.实施方案212的组合物,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的尿嘧啶稳定多肽(USP)。

[0471] 219.实施方案212的组合物,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的尿嘧啶稳定多肽(USP)。

[0472] 220.实施方案212的组合物,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的尿嘧啶稳定多肽(USP)。

[0473] 221.实施方案212的组合物,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有100%序列同一性的尿嘧啶稳定多肽(USP)。

[0474] 222.实施方案212的组合物,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

[0475] 223.实施方案212的组合物,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

[0476] 224.实施方案223的组合物,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

[0477] 225.实施方案224的组合物,其中RGN是RGN切口酶。

[0478] 226.一种包含编码融合蛋白的核酸分子和编码尿嘧啶稳定多肽(USP)的核酸分子的载体,其中所述融合蛋白包含DNA结合多肽和脱氨酶,并且其中所述USP与SEQ ID NO:1-16中的任一个具有至少80%的序列同一性。

[0479] 227.实施方案226的载体,所述USP与SEQ ID NO:1-16中的任一个具有至少85%的序列同一性。

[0480] 228.实施方案226的载体,所述USP与SEQ ID NO:1-16中的任一个具有至少90%的序列同一性。

[0481] 229.实施方案226的载体,所述USP与SEQ ID NO:1-16中的任一个具有至少95%的序列同一性。

[0482] 230.实施方案226的载体,所述USP与SEQ ID NO:1-16中的任一个具有100%的序列同一性。

[0483] 231.实施方案226的载体,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少80%序列同一性的尿嘧啶稳定多肽(USP)。

[0484] 232.实施方案226的载体,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的尿嘧啶稳定多肽(USP)。

[0485] 233.实施方案226的载体,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的尿嘧啶稳定多肽(USP)。

[0486] 234.实施方案226的载体,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的尿嘧啶稳定多肽(USP)。

[0487] 235.实施方案226的载体,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有100%序列同一性的尿嘧啶稳定多肽(USP)。

[0488] 236.实施方案226的载体,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

[0489] 237.实施方案226的载体,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

[0490] 238.实施方案237的载体,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

[0491] 239.实施方案238的载体,其中RGN是RGN切口酶。

[0492] 240.一种包含实施方案226-239中任一项的载体的细胞。

[0493] 241.一种细胞,其包含:

[0494] a)融合蛋白,其包含:(i)DNA结合多肽;和(ii)脱氨酶;或编码所述融合蛋白的核酸分子;和

[0495] b)尿嘧啶稳定多肽(USP),其与SEQ ID NO:1-16中的任一个具有至少80%序列同一性;或编码所述USP的核酸分子。

[0496] 242.实施方案241的细胞,其中USP与SEQ ID NO:1-16中的任一个具有至少85%的序列同一性。

[0497] 243.实施方案241的细胞,其中USP与SEQ ID NO:1-16中的任一个具有至少90%的序列同一性。

[0498] 244.实施方案241的细胞,其中USP与SEQ ID NO:1-16中的任一个具有至少95%的序列同一性。

[0499] 245.实施方案241的细胞,其中USP与SEQ ID NO:1-16中的任一个具有100%序列同一性。

[0500] 246.实施方案241的细胞,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少80%序列同一性的尿嘧啶稳定多肽(USP)。

[0501] 247.实施方案241的细胞,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少85%序列同一性的尿嘧啶稳定多肽(USP)。

[0502] 248.实施方案241的细胞,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少90%序列同一性的尿嘧啶稳定多肽(USP)。

[0503] 249.实施方案241的细胞,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有至少95%序列同一性的尿嘧啶稳定多肽(USP)。

[0504] 250.实施方案241的细胞,其中所述融合蛋白还包含与SEQ ID NO:1-16中的任一个具有100%序列同一性的尿嘧啶稳定多肽(USP)。

[0505] 251.实施方案241的细胞,其中所述DNA结合多肽是大范围核酸酶、锌指融合蛋白或TALEN。

[0506] 252.实施方案241的细胞,其中所述DNA结合多肽是RNA引导的DNA结合多肽。

[0507] 253.实施方案252的细胞,其中所述RNA引导的DNA结合多肽是RNA引导的核酸酶多肽(RGN)。

[0508] 254.实施方案253的细胞,其中RGN是RGN切口酶。

[0509] 以下实施例以说明而非限制的方式提供。

实施例

[0510] 实施例1:尿嘧啶稳定多肽

[0511] 本发明的尿嘧啶稳定多肽(USP)的氨基酸序列作为SEQ ID NO:1-16提供,如表1所示。公开的所有USP均来自葡萄球菌属的几个种,并且长度范围为112至116个氨基酸。本文描述的USP的多肽具有16.4%至20.7%的负电荷,预期pI为3.86-4.55,APG05963除外,其预期pI为5.26。

[0512] USP APG06351、APG03399、APG04638、APG09242、APG02463、APG04080、APG01791、APG04001和APG03327共享独特的共有C端序列“KEGGNDHE”(SEQ ID NO:33)。USP APG05198和APG05756共享独特的C端序列“EKENYNNE”(SEQ ID NO:34)。APG05963具有独特的C端序列“EKEKHKNE”(SEQ ID NO:35);APG06702具有独特的C端序列“DKGDDNHD”(SEQ ID NO:36);APG05316具有独特的C端序列“QKGGQ”(SEQ ID NO:37);APG09230具有独特的C端序列

“KGENKYE”(SEQ ID NO:38);和APG04100具有独特的C端序列“KQGENNHE”(SEQ ID NO:39)。

[0513] 表1:尿嘧啶稳定多肽

[0514]

USP ID	SEQ ID NO
APG03399	1
APG06702	2
APG05198	3
APG04638	4
APG06351	5
APG05963	6
APG09230	7
APG04100	8
APG04001	9
APG02463	10

[0515]

APG09242	11
APG04080	12
APG05316	13
APG03327	14
APG01791	15
APG05756	16

[0516] 实施例2:USP融合蛋白在哺乳动物细胞中表现出增加的碱基编辑活性

[0517] 鉴定了预计使RGN APG07433.1(SEQ ID NO:40;PCT公开WO2019/236566,通过引用并入本文)的RuvC结构域失活的残基,并将RGN修饰为切口酶变体(nAPG07433.1;SEQ ID NO:41)。产生了包含胞苷脱氨酶的融合蛋白,即APG09980(SEQ ID NO:47;参见PCT/US2019/068079,通过引用并入本文)或APG07386CTD(SEQ ID NO:48;参见PCT/US2019/068079)。在表1中的USP中,选择了三种以用于测定在融合蛋白中的活性,即APG03399、APG06702和APG05198。针对表达进行了密码子优化的脱氨酶、USP和nRGN核苷酸序列被合成为具有N端核定位标签的融合蛋白,并克隆到pTwist CMV(Twist Biosciences)表达质粒中。缺少本发明的USP的融合蛋白包含(从氨基末端开始)SV40 NLS(SEQ ID NO:42),其在C端可操作地连接至3X FLAG标签(SEQ ID NO:43),在C端可操作地连接至脱氨酶,在C端可操作地连接至肽接头(SEQ ID NO:44),在C端可操作地连接至nRGN(例如,nAPG07433.1,其是SEQ ID NO:41),最终在C端可操作地连接至核质蛋白NLS(SEQ ID NO:45)。包含本发明的USP的融合蛋

白包含(从氨基末端开始)SV40 NLS(SEQ ID NO:42),其在C端可操作地连接至3X FLAG标签(SEQ ID NO:43),在C端可操作地连接至脱氨酶,在C端可操作地连接至肽接头(SEQ ID NO:44),在C端可操作地连接至nRGN(例如,SEQ ID NO:41),在C端可操作地连接至第二接头序列(SEQ ID NO:46),在C端可操作地连接至本发明的USP,最终在C端可操作地连接至核质蛋白NLS(SEQ ID NO:45)。表2显示了经产生并测试活性的融合蛋白。如上所述,所有融合蛋白都包含至少一个NLS和3X FLAG标签。

[0518] 表2:针对C>T编辑进行测定的融合蛋白

[0519]

融合蛋白	SEQ ID NO
APG09980-nAPG07433.1	49
APG09980-nAPG07433.1-APG03399	50
APG09980-nAPG07433.1-APG06702	51
APG09980-nAPG07433.1-APG05198	52
APG07386-CTD-nAPG07433.1	53
APG07386-CTD-nAPG07433.1-APG03399	54
APG07386-CTD-nAPG07433.1-APG06702	55
APG07386-CTD-nAPG07433.1-APG05198	56

[0520] 还产生了包含编码sgRNA的表达盒的表达质粒。表3显示了人类基因组靶序列和用于将融合蛋白引导至基因组靶标的sgRNA序列。还指出了每个靶序列的基因组位点。

[0521] 表3:引导RNA序列

[0522]

sgRNA ID	靶序列	sgRNA 序列	基因组位点
SGN000143	57	63	69
SGN000169	58	64	70
SGN000173	59	65	71
SGN000929	60	66	72
SGN000930	61	67	73
SGN001101	62	68	74

[0523] 使用Lipofectamine 2000试剂(Life Technologies)将包含含有表2所示的融合蛋白的编码序列的表达盒的500ng质粒和包含编码表3所示的sgRNA的表达盒的500ng质粒共转染到24孔板中75-90%汇合度的HEK293FT细胞中。然后将细胞在37℃下孵育72小时。孵

育后，按照制造商的方案使用NucleoSpin 96Tissue(Macherey-Nagel)提取基因组DNA。将靶基因组位点侧翼的基因组区域进行PCR扩增，并按照制造商的方案使用ZR-96DNA Clean and Concentrator(Zymo Research)纯化产物。然后将纯化的PCR产物发送到Illumina MiSeq(2x250)上进行下一代测序。针对插入缺失形成或特定胞嘧啶突变分析结果至+30个核苷酸，其中表3中描述的靶序列3'末端的最后一个核苷酸是+1，并且其中+30核苷酸是SEQ ID NO:57-62所示靶序列中的+1核苷酸的5'或上游的29个核苷酸(SEQ ID NO:57-62的靶序列分别在SEQ ID NO:69-74所示基因组位点序列中表示为小写字母文本)。

[0524] 表4至15显示了针对表2的融合蛋白和表3的引导RNA的每种组合的胞苷碱基编辑。融合蛋白通过它们的SEQ ID NO来识别。表4-15中胞苷(C)的编号如上一段所述，其中靶序列3'端的最后一个核苷酸为+1，并且编号在靶序列和相应的基因组位点序列中从3'到5'方向进行。有趣的是，当比较脱氨酶-nRGN融合蛋白的活性与相应脱氨酶-nRGN加USP的活性时，使用相同的引导RNA，胞苷向所需胸苷的转化率更高，向腺苷或鸟苷的转化率更低。

[0525] 表4:使用脱氨酶APG09980和引导SGN000143的C>N编辑率

[0526]

融合蛋白(SEQ ID NO)		C10	C15	C20	C21	C22	C25
49	A	2.8	0.1	0	0.2	0.6	0.1
	G	1.1	0.1	0.1	0.1	0.2	0
	T	11	0.4	0.2	0.3	2	0.7
50	A	0	0	0	0	0	0
	G	0	0	0	0	0	0
	T	23.5	1.4	1.7	2	4.6	0.7
51	A	0.1	0	0	0	0	0
	G	0	0	0	0	0	0

[0527]

融合蛋白(SEQ ID NO)		C10	C15	C20	C21	C22	C25
	T	16.3	1.2	1.5	1.6	3.1	0.6
52	A	0.2	0	0	0	0	0
	G	0	0	0	0	0	0
	T	21.4	1.6	1.1	1.4	2.5	0.7

[0528] 表4的结果显示，在具有USP的样品中，C10位置处的C>T编辑率增加。

[0529] 表5:使用脱氨酶APG07386和引导SGN000143的C>N编辑率

[0530]

融合蛋白 (SEQ ID NO.)		C10	C15	C20	C21	C22	C25
53	A	0.1	0.4	0	0.5	1	0
	G	0	0.8	0	0	0.1	0
	T	0.5	4	0.3	0.7	3.8	0
54	A	0	0	0	0	0.1	0
	G	0	0	0	0	0	0
	T	1	4.5	1	1.7	3.1	0.1
55	A	0	0.1	0	0	0.1	0
	G	0	0	0	0	0	0
	T	1.6	9.7	2.5	3.2	5.8	0.2
56	A	0	0.1	0	0	0	0
	G	0	0	0	0	0	0
	T	1.9	10.2	2.6	3.1	7	0.2

[0531] 表5的结果显示,在具有USP的样品中,C15位置处的C>T编辑率增加。

[0532] 表6:使用脱氨酶APG09980和引导SGN000169的C>N编辑率

[0533]

融合蛋白 (SEQ ID NO.)		C9	C13	C15	C18	C20	C23

[0534]

融合蛋白(SEQ ID NO.)		C9	C13	C15	C18	C20	C23
49	A	0.5	0.8	3.6	0.3	1.3	1.3
	G	1.5	1.8	8.3	0	0.2	0.1
	T	1.5	23.9	16.3	0.3	8.6	4.9
50	A	0	0	0.6	0	0.1	0.2
	G	0.2	0	2.5	0	0	0
	T	7	38.1	35.8	0.5	16.6	13.4
51	A	0.2	0.3	1.3	0.2	0.3	0.4
	G	0.6	0.6	2.1	0.1	0.1	0
	T	9.2	42.4	40.8	0.8	20.4	16.6
52	A	0.8	1	2.4	0.7	0.6	1.5
	G	0.5	1.7	3.7	0.5	0.5	0.2
	T	8.5	37.8	35.7	1.2	17.1	14.9

[0535] 表6的结果显示,C13和C15位置处的C>T形成率在添加USP的情况下增加。

[0536] 表7:使用脱氨酶APG07386和引导SGN000169的C>N编辑率

[0537]

融合蛋白(SEQ ID NO.)		C9	C13	C15	C18	C20	C23
53	A	0.8	0.7	2	0	0.5	0.3
	G	5.4	0.8	10.2	0.1	0.1	0
	T	3.5	5.1	11.2	0.3	2.6	2.4
54	A	0.1	0.5	0.8	0	0.3	0.2
	G	0.7	0.2	2.5	0.2	0.1	0
	T	8.5	15.9	24.8	1.5	6.4	7.1
55	A	0.3	0.5	1.3	0.4	0.2	0.4
	G	2	0.5	3.5	0.1	0.2	0.1

[0538]

融合蛋白(SEQ ID NO.)		C9	C13	C15	C18	C20	C23
	T	14.3	22.7	37.1	2	9.2	7.9
56	A	0.1	0.1	0.7	0	0.2	0
	G	0.7	0	2.9	0	0	0
	T	14.7	24	36.4	1.3	8.4	8.9

[0539] 表7的结果显示,C13和C15位置处的C>T形成率在添加USP的情况下增加。

[0540] 表8:使用脱氨酶APG09980和引导SGN000930的C>N编辑率

[0541]

融合蛋白(SEQ ID NO.)		C17	C19	C22
49	A	0.7	0.4	1.4
	G	34.5	0.4	0
	T	2.6	2.4	2.2
50	A	0.3	0.1	0.3
	G	8.6	0.3	0
	T	33	2.8	3.2
51	A	0.2	0.1	0.2
	G	5.1	0.1	0.1
	T	35.7	3.6	3.3
52	A	0.4	0.3	0.9
	G	11.2	0	0
	T	23.7	2.5	2.5

[0542] 表8的结果显示,在具有有利于C>T变化的USP的所有样品中,C17位置处的C>G编辑率降低。

[0543] 表9:使用脱氨酶APG07386和引导SGN000930的C>N编辑率

[0544]

融合蛋白(SEQ ID NO.)		C17	C19	C22

[0545]

融合蛋白 (SEQ ID NO.)		C17	C19	C22
53	A	0.2	0	0.4
	G	14.7	0	0.4
	T	1.1	0.2	2.2
54	A	0	0	0
	G	1.4	0	0
	T	9.1	0.4	3.7
55	A	0.4	0.1	0.3
	G	2.6	0.2	0.2
	T	11.8	0.9	4.5
56	A	0.1	0.2	0
	G	2.7	0	0.1
	T	17.4	0.8	6.7

[0546] 表9的结果显示,在具有有利于C>T变化的USP的所有样品中,C17位置处的C>G编辑减少。

[0547] 表10:使用脱氨酶APG09980和引导SGN00173的C>N编辑率

[0548]

融合蛋白 (SEQ ID NO.)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
49	A	0	0	0.1	0.3	3.1	6.9	0.3	0	3.4	0.1
	G	0	0.1	0	0.2	0.3	1	0.1	0.4	3.4	0.1
	T	0.1	0	0	2.5	11.1	12.3	1.7	0.2	12.1	1.7
50	A	0	0.1	0	0.1	0.9	2.9	0	0.4	1.5	0
	G	0.1	0.1	0	0	0.5	0.8	0.4	0.3	1.9	0.2
	T	0.1	0.1	0.3	14	39	39.3	28.7	13	31.1	12
51	A	0	0	0	0	1.2	1.6	0	0	0.9	0
	G	0	0	0	0	0	1.4	0.1	0	2.2	0.1
	T	0	0	0.3	10.4	33.5	33.4	23.8	10.9	26.8	10.3

[0549]

融合蛋白(SEQ ID NO.)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
52	A	0	0	0	0	1.8	1.2	0.3	0	1.5	0
	G	0	0	0	0	0	0.8	0	0.5	2.7	0
	T	0	0	0.6	10.5	33.1	36.8	26.7	13.2	28.6	9.6

[0550] 表10的结果显示，在C4、C7、C8、C10、C11、C17和C20位置处的C>T形成率随着USP的添加而增加。

[0551] 表11：使用脱氨酶APG07386和引导SGN00173的C>N编辑率

[0552]

融合蛋白(SEQ ID NO.)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
53	A	0	0	0	0.1	0.4	1.1	0	2.5	2.8	0
	G	0	0	0	0	0.1	0.3	0	0.5	2	0
	T	0	0	0	0.5	1.4	2.5	0	7	11.3	0.7
54	A	0	0	0	0	0	0.1	0	0.7	0.8	0
	G	0	0	0	0	0	0	0	0	0.6	0
	T	0	0	0	1.6	6.1	8.6	2.2	13.3	23	1.8
55	A	0	0	0	0	0.2	0	0	1	0.9	0
	G	0	0	0	0	0	0.3	0.1	0.1	0.5	0
	T	0	0	0.2	2.2	9.1	11.2	3.2	13	21.3	2
56	A	0	0	0	0	0	0	0	0.9	0.8	0
	G	0	0	0	0	0	0.1	0	0.2	0.5	0
	T	0	0	0.1	2.7	12.3	15	4.3	18.3	27.1	1.9

[0553] 表11的结果显示，在C7、C8、C11和C17位置处的C>T形成率随着USP的添加而增加。

[0554] 表12：使用脱氨酶APG09980和引导SGN000929的C>N编辑率

[0555]

融合蛋白(SEQ ID NO.)		C6	C9	C13	C23

[0556]

融合蛋白 (SEQ ID NO.)		C6	C9	C13	C23
49	A	0.2	1.1	0.2	5.2
	G	0.2	1.5	1.4	20.1
	T	1.2	5.2	1	7.4
50	A	0	0.1	0	1.3
	G	0	0.1	0	3.3
	T	2.2	9.4	3.2	30.8
51	A	0	0.2	0	0.8
	G	0	0.1	0	1.6
	T	1.5	9.2	3.5	34
52	A	0.5	0.6	0	1.8
	G	0.1	0.5	0	6.5
	T	1.6	8.6	2.7	22.3

[0557] 表12的结果显示,C23位置处的C>T形成率随着USP的添加而增加。

[0558] 表13:使用脱氨酶APG07386和引导SGN000929的C>N编辑率

[0559]

融合蛋白 (SEQ ID NO.)		C6	C9	C13	C23
53	A	1.5	0	0	1.5
	G	0.6	0	0.1	5.2
	T	5	0.3	0.1	2.1
54	A	0	0	0	0.1
	G	0	0	0	0.1
	T	4.2	0.3	0	4.6
55	A	0.2	0	0.1	0
	G	0	0	0	1.4
	T	10.4	0.9	0.3	8.2
56	A	0.1	0	0	0.2

[0560]

融合蛋白 (SEQ ID NO.)		C6	C9	C13	C23
	G	0.2	0	0	0.5
	T	10.8	0.6	0.4	8.5

[0561] 表13的结果显示，在C6和C23位置处的C>T形成率随着USP的添加而增加。

[0562] 表14：使用脱氨酶APG09980和引导SGN001101的C>N编辑率

[0563]

融合蛋白 (SEQ ID NO.)		C2	C6	C7	C18
49	A	0.1	1.8	0.4	3.3
	G	0	2.8	0.1	18.5
	T	0.1	6.5	1.8	5.3
50	A	0.2	0.3	0.2	0.6
	G	0.1	0.1	0.2	2.1
	T	0.1	9.1	0.7	25.4
51	A	0.2	0.1	0.1	0.3
	G	0.1	0.1	0	2.4
	T	0.1	7.9	0.7	21.8
52	A	0.2	0.9	0.2	1.2
	G	0	0.8	0	4.8
	T	0.2	7.7	1.3	19.4

[0564] 表14的结果显示，C18位置处的C>T形成率随着USP的添加而增加。

[0565] 表15：使用脱氨酶APG07386和引导SGN001101的C>N编辑率

[0566]

融合蛋白 (SEQ ID NO.)		C2	C6	C7	C18
53	A	0.1	0	4	2.2
	G	0	0	2.1	11

[0567]

融合蛋白 (SEQ ID NO.)		C2	C6	C7	C18
	T	0	0	10.4	3.9
54	A	0.1	0	0.1	0
	G	0	0	0.1	0.6
	T	0	1.7	8.5	6.7
55	A	0.2	0.1	0.6	0.3
	G	0	0	0.5	1.4
	T	0	1.7	10.7	9.2
56	A	0.1	0	0.3	0.2
	G	0	0	0.4	1.4
	T	0	1.9	12.1	9.8

[0568] 表15的结果显示,C18位置处的C>T形成率随着USP的添加而增加。在C18位置处的C>G转化率随着USP的添加而降低。

[0569] 表16和17显示了测试的每种融合蛋白/引导物组合的插入缺失形成率。融合蛋白由SEQ ID NO表示。数据表明包含本文所述的USP的融合蛋白降低了测试的所有靶基因组位置处的插入缺失形成率。

[0570] 表16:APG09980和USP的插入和缺失率

[0571]

sgRNA ID	SEQ ID NO: 49	SEQ ID NO: 50	SEQ ID NO: 51	SEQ ID NO: 52
SGN000143	2.78	0	0	0.62
SGN000169	18.02	1.62	3.89	9.1
SGN000173	27.46	7.86	7.17	7.48
SGN000929	3.49	0.2	0.65	2.36
SGN000930	6.03	0.81	0.65	1.65
SGN001101	2.5	1.66	0.92	1.86

[0572] 表17:APG07386-CTD和USP的插入和缺失率

[0573]

sgRNA ID	SEQ ID NO: 53	SEQ ID NO: 54	SEQ ID NO: 55	SEQ ID NO: 56

[0574]

SGN000143	**0.13**	**0.05**	**0**	**0.06**
SGN000169	**5.6**	**1.38**	**3.11**	**0.32**
SGN000173	**13.8**	**0.99**	**3.49**	**2.57**
SGN000929	**0.5**	**0**	**0.3**	**0.16**
SGN000930	**1.7**	**0**	**0.23**	**0.31**
SGN001101	**1.78**	**0**	**0.48**	**0.39**

[0575] 实施例3:测试不同的递送形式

[0576] 为了确定碱基编辑器是否能够以不同的形式递送,使用原代T细胞测试了mRNA递送。纯化的CD3+T细胞或PBMC被解冻,使用CD3/CD28珠子(ThermoFisher)激活3天,然后使用Lonza4D-Nucleofector X单元和Nucleocuvette条进行核转染。P3原代细胞试剂盒用于mRNA和RNP递送。使用EO-115和EH-115程序分别转染细胞以进行mRNA和RNP递送。细胞在核转染后在含有IL-2、IL-7和IL-15(Miltenyi Biotec)的CTS OpTimizer T细胞扩增培养基(ThermoFisher)中培养4天,然后使用Nucleospin Tissue基因组DNA分离试剂盒(Machery Nagel)收获。

[0577] 编辑位点周围的扩增子通过PCR生成,并使用Illumina Nexterra平台使用2x250bp配对末端测序进行NGS测序。估计的碱基编辑率是通过计算每个样品的总体取代率来确定的。下表18和19显示了每个测试的引导物的平均值和样品数。

[0578] APG09980-nAPG07433.1-APG03399和APG05840-nAPG07433.1-APG03399在通过mRNA递送时显示出作为多个靶标的高碱基编辑率。尽管具有高取代率,但由于在碱基编辑构建体中并入了USP2,插入缺失形成率非常低。

[0579] 表18:APG09980-nAPG07433.1-APG03399的平均碱基编辑率

[0580]

融合构建体	**基因名称**	**SGN**	**平均取代%**	**平均插入缺失%**	**N**
APG09980-nAPG07433.1-APG03399	**基因 1**	**SGN000754**	**23.32917436**	**1.128931**	**6**

[0581]

融合构建体	基因名称	SGN	平均取代%	平均插入缺失%	N
APG09980-nAPG07433.1-APG03399	基因 1	SGN000755	59.37254849	7.1037823	4
APG09980-nAPG07433.1-APG03399	基因 2	SGN001061	13.60100568	0.4214674	3
APG09980-nAPG07433.1-APG03399	基因 2	SGN001062	26.9304354	4.3225871	4
APG09980-nAPG07433.1-APG03399	基因 2	SGN001063	75.27761104	0.8163273	4
APG09980-nAPG07433.1-APG03399	基因 2	SGN001064	72.94658862	1.0468487	3

[0582] 表19:APG05840-nAPG07433.1-APG03399的平均碱基编辑率

[0583]

融合构建体	基因名称	SGN	平均取代%	平均插入缺失%	N
APG05840-nAPG07433.1-APG03399	基因 1	SGN000754	57.7775198	5.14624384	4
APG05840-nAPG07433.1-APG03399	基因 1	SGN000755	68.352455	4.98538891	3
APG05840-nAPG07433.1-APG03399	基因 2	SGN001061	14.6830209	0	2
APG05840-nAPG07433.1-APG03399	基因 2	SGN001062	39.7312597	2.9887885	4
APG05840-nAPG07433.1-APG03399	基因 2	SGN001063	70.4564399	0.25727852	4
APG05840-nAPG07433.1-APG03399	基因 2	SGN001064	53.2112842	1.98008536	3

序列表

<110> 生命编辑制药股份有限公司

<120> 尿嘧啶稳定蛋白及其活性片段和变体以及使用方法

<130> L103438 1220WO (0106.5)

<140>

<141>

<150> 63/052,175

<151> 2020-07-15

<160> 142

<170> PatentIn version 3.5

<210> 1

<211> 116

<212> PRT

<213> 葡萄球菌属的几个种

<223> APG03399

<400> 1

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

<210> 2

<211> 116

<212> PRT

<213> 葡萄球菌属的几个种

<223> APG06702

<400> 2

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

<210> 3
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG05198
<400> 3

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

<210> 4

<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG04638
<400> 4

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

<210> 5
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG06351
<400> 5

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

Thr Gln Asn Ile Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Gly
100 105 110
Gln Asp His Glu
115

<210> 6
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG05963
<400> 6

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

<210> 7
<211> 115
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG09230
<400> 7

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

50 55 60
Leu Ser Ile Tyr Pro Phe Leu Met Tyr Tyr Gln Asn Gln Leu Ile Ser
65 70 75 80
Ile Gly Tyr Leu Asp His Tyr Asp Met Asp Phe Leu Tyr Leu Asn Asp
85 90 95
Thr Gln Thr Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gly Glu Asn
100 105 110
Lys Tyr Glu
115

<210> 8
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG04100
<400> 8
Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15
Gln Leu Ser Asn Glu Glu Ser Trp Glu Cys Glu Val Leu Asp Glu Val
20 25 30
Ala Glu Asn Ile Leu Pro Pro Arg Phe Val Asp Gly Ser Pro Leu Thr
35 40 45
His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asn Glu Leu His Asp
50 55 60
Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80
Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95
Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Glu
100 105 110
Asn Asn His Glu
115

<210> 9
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG04001
<400> 9
Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

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

<210> 10
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG02463
<400> 10

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

<210> 11
<211> 116
<212> PRT

<213> 葡萄球菌属的几个种
<223> APG09242
<400> 11

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

<210> 12
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG04080
<400> 12

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

Asn Asn His Glu
115

<210> 13
<211> 112
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG05316
<400> 13

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

<210> 14
<211> 113
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG03327
<400> 14

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

85 90 95
Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Gln Lys Gly Arg
100 105 110
Glu

<210> 15
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG01791
<400> 15

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

<210> 16
<211> 116
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG05756
<400> 16

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

50 55 60

His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile
65 70 75 80

Gly Tyr Ile Asp Glu Ile Asn Asp Met Asp Phe Leu Tyr Leu His Asn
85 90 95

Thr Val Met Pro Leu Leu Asp Gln Arg His Leu Leu Glu Lys Glu Asn
100 105 110

Tyr Asn Asn Glu
115

<210> 17
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG03399哺乳动物密码子优化序列
<400> 17

atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg 120
ttcatcaaca acagcctgct ggcccacctg accctgaaca cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacgacca gctgatcagc 240
atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgaagcag ggcgagaaca accacgag 348

<210> 18
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG06702哺乳动物密码子优化序列
<400> 18

atgaccatcg agaacgagct gaaggagtac atcaccgacc tgttcgagct gcccaaggac 60
gagcagtggc tgtgcgaggc cgtggacgag gtggccgaca acatcctgcc cgacgagtac 120
gtgaccctgg gccccctgag gaacaagacc ctggagacct tcacctacta cagcaacacc 180
ctgcacgaga ggaacatcta ccccttcatc ctgtactacc agaacaggct gatcgccatc 240
ggctacatcg acgagaacta cgagatggac ttcatgtacc tgcacaacac cgccatcacc 300
ctgatcgacc agaggtacct gatcgacaag ggcgacgaca accacgac 348

<210> 19
<211> 348
<212> DNA
<213> 人工序列

<220>
<223> APG05198哺乳动物密码子优化序列
<400> 19
atgaccctgg agaagcagct gaaggagtac atcaccaacc tgttcaacct gcccaaggac 60
gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac 120
atcaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacacc 180
ctgcacgaga ggcacatcta ccccttcatc ctgtactacc agaagcagct gatcgccatc 240
ggctacatcg acgagaccaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc 300
ctgctggacc agaggcacct gctggagaag gagaactaca acaacgag 348
<210> 20
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG04638哺乳动物密码子优化序列
<400> 20
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg 120
ttcatcaacg acagcccccc ggcccacctg accctggaga cctacaccta ctacaacaac 180
gagctgcacg acctgagcat ctacccccttc ctgatgtacg ccaacaacca gctgatcagc 240
gtgggctacc tgaaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgagggag ggcggcaccc accacgag 348
<210> 21
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG06351哺乳动物密码子优化序列
<400> 21
atgaagacca tcacccagga gctgaagcag tacatcaccc acctgttcca gctgagcaac 60
aacgaggcct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg 120
ttcatcaacg acagcccccc ggtgcacctg accctggaga cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccccttc ctgatgtacg ccaacaacca gctgatcagc 240
gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac ccagaacatc 300
atcatcgacg agaggcacct gctgaagcag ggcggccagg accacgag 348
<210> 22
<211> 348
<212> DNA
<213> 人工序列

<220>
<223> APG05963哺乳动物密码子优化序列
<400> 22
atgaccctgg agcagcagct gaagcactac atcaccaacc tgttcaagct gcccaaggac 60
gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac 120
gtgaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacacc 180
ctgcacgaga ggcacatcta ccccttcatc ctgtactacc agaagcagct gatcgccatc 240
ggctacatcg acgagaccaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc 300
ctgctggacc agaggcacct gctggagaag gagaagcaca agaacgag 348
<210> 23
<211> 345
<212> DNA
<213> 人工序列
<220>
<223> APG09230哺乳动物密码子优化序列
<400> 23
atgaagaccc agatcaacga gctgaagagc tacatcacca acctgttcca gctgaccagc 60
caggagaact gggagtgcga ggccctgaac gagagcgccg acaacatcct gcccgagagg 120
ttcctgaacg acagcccccc gagcaacagg atcctggaga cctacaccta ctacaacaac 180
gagctgcacg agctgagcat ctacccccttc ctgatgtact accagaacca gctgatcagc 240
atcggctacc tggaccacta cgacatggac ttcctgtacc tgaacgacac ccagaccacc 300
atcatcgacg agaggcacct gctgaagggc gagaacaagt acgag 345
<210> 24
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG04100哺乳动物密码子优化序列
<400> 24
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
gaggagagct gggagtgcga ggtgctggac gaggtggccg agaacatcct gccccccagg 120
ttcgtggacg gcagcccccc gacccacctg accctggaga cctacaccta ctacaacaac 180
gagctgcacg acctgagcat ctacccccttc ctgatgtacg ccaacaacca gctgatcagc 240
gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgaagcag ggcgagaaca accacgag 348
<210> 25
<211> 348
<212> DNA
<213> 人工序列

<220>
<223> APG04001哺乳动物密码子优化序列
<400> 25
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccaccagg 120
ttcgtggacc acacccccct ggcccacctg accctggaca cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccctTc ctgatgtacg ccaacaacca gctgatcagc 240
atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacatc 300
atcatcgacg agaggcacct gctgagggag ggcggcaacg acagggag 348
<210> 26
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG02463哺乳动物密码子优化序列
<400> 26
atgcagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg 120
ttcatcaacg gcagcccccT ggcccacctg accctggaga cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccctTc ctgatgtacg ccaacaacca gctgatcagc 240
atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgaaggag ggcggcaaca accacgag 348
<210> 27
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG09242哺乳动物密码子优化序列
<400> 27
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gacaccgccg agaacatcct gcccgagtgc 120
ttcatcaacg acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccctTc ctgatgtacg ccaacaacca gctgatcagc 240
gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgaaggag ggcggcaacg accacgag 348
<210> 28
<211> 348
<212> DNA
<213> 人工序列

<220>
<223> APG04080哺乳动物密码子优化序列
<400> 28
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg ccaacatcct gcccgagagg 120
ttcatcaaca acagcccccc ggcccacctg accctggaga cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccctcc ctgatgtacg ccaacaacca gctgatcagc 240
gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgaaggag ggcggcaaca accacgag 348
<210> 29
<211> 336
<212> DNA
<213> 人工序列
<220>
<223> APG05316哺乳动物密码子优化序列
<400> 29
atgaccctga gcaagcagct gaagacctac atcaccgaga ggttcaagct gaactaccag 60
gagacctggg cctgcgagac catcgacgcc gtggccgagg acgtgctgcc cgagaagtac 120
atcaagaaca gccccctgga gcacaagatc ctgaacacct tcacctacta caacgacgag 180
ctgcacgaga tcagcatcta ccccttcctg tgctacctgg acaaggagct ggtggccatc 240
ggctacctgg acaacttcga cctggacttc atcttcctga acgacaccca ccaggtgatc 300
atcgacgaga ggtacctgct gcagaagggc ggccag 336
<210> 30
<211> 339
<212> DNA
<213> 人工序列
<220>
<223> APG03327哺乳动物密码子优化序列
<400> 30
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gacgccgccg agaacatcct gcccaccagg 120
ttcgtggacc acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccctcc ctgatctacg ccaacaacca gctgatcagc 240
atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgcagaag ggcagggag 339
<210> 31
<211> 348
<212> DNA
<213> 人工序列

<220>
<223> APG01791哺乳动物密码子优化序列
<400> 31
atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccaccagg 120
ttcgtgggcc acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccctttc ctgatgtacg ccaacgacca gctgatcagc 240
atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
ctgatcgacg agaggcacct gctgaagcag ggcggcaacg accacgag 348
<210> 32
<211> 348
<212> DNA
<213> 人工序列
<220>
<223> APG05756哺乳动物密码子优化序列
<400> 32
atgaccctgg agcagcagct gaagcactac atcaccaacc tgttcagcct gccccaggac 60
gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac 120
gtgaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacaag 180
ctgcacgaga ggcacatcta ccccttcatc ctgtactacc agaagcagct gatcgccatc 240
ggctacatcg acgagatcaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc 300
ctgctggacc agaggcacct gctggagaag gagaactaca acaacgag 348
<210> 33
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 共享的共有C端序列
<400> 33
Lys Glu Gly Gly Asn Asp His Glu
1 5
<210> 34
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> APG05756和APG05198独特C末端序列
<400> 34
Glu Lys Glu Asn Tyr Asn Asn Glu

1 5

<210> 35

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> APG05963独特C末端序列

<400> 35

Glu Lys Glu Lys His Lys Asn Glu

1 5

<210> 36

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> APG06702独特C末端序列

<400> 36

Asp Lys Gly Asp Asp Asn His Asp

1 5

<210> 37

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> APG05316独特C末端序列

<400> 37

Gln Lys Gly Gly Gln

1 5

<210> 38

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> APG09230独特C末端序列

<400> 38

Lys Gly Glu Asn Lys Tyr Glu

1 5

<210> 39

<211> 8

<212> PRT
<213> 人工序列
<220>
<223> APG04100独特C末端序列
<400> 39
Lys Gln Gly Glu Asn Asn His Glu
1 5
<210> 40
<211> 1071
<212> PRT
<213> 人工序列
<220>
<223> RGN APG07433.1
<400> 40
Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15
Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Arg
20 25 30
Tyr Glu Lys Val Arg Ile Val Asp Gln Gly Val Arg Met Phe Asp Arg
35 40 45
Ala Glu Met Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60
Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80
Asn Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu
85 90 95
Leu Asp Ser Leu Tyr Pro Leu Ser Lys Lys Ser Met Asp Ile Trp Gly
100 105 110
Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Phe Glu Trp Ala
115 120 125
Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140
Lys Ser Glu Leu Lys Asp Thr Glu Thr Gly Lys Val Leu Ser Ser Ile
145 150 155 160
Gln Leu Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175
Trp Met Lys Asp Pro Asp Phe Ser Lys Tyr Asp Arg Lys Arg Asn Ser
180 185 190
Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Glu Leu Glu Lys Glu

195 200 205

Ile Val Thr Leu Phe Ala Ala Gln Arg Arg Phe Gln Ser Pro Tyr Ala
210 215 220

Ser Lys Asp Leu Gln Glu Thr Tyr Leu Gln Ile Trp Thr His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
245 250 255

Leu Leu Lys Gly Lys Glu Arg Arg Ile Pro Lys Ala Thr Tyr Thr Phe
260 265 270

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asp Phe Gln Pro Phe Thr Lys Glu Gln Arg Glu Ile Ile Leu Asn Asn
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val
355 360 365

Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln
435 440 445

Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr
450 455 460

Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val
500 505 510

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

820 825 830

Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly

835 840 845

Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp

850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr

865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys

885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly

900 905 910

Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val

915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val

930 935 940

Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val

945 950 955 960

Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser

965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys

980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu

995 1000 1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg

1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr

1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val

1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg

1055 1060 1065

Asn Phe His

1070

<210> 41

<211> 1071

<212> PRT

<213> 人工序列

<220>

<223> nRGN APG07433.1

<400> 41

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

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

610 615 620
Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640
Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
645 650 655
Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
660 665 670
Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685
Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
690 695 700
Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
705 710 715 720
Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735
Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750
Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
755 760 765
Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
770 775 780
Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
785 790 795 800
Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
805 810 815
Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
820 825 830
Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910
Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val
915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
1055 1060 1065
Asn Phe His
1070

<210> 42
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> SV40 NLS
<400> 42
Pro Lys Lys Lys Arg Lys Val
1 5
<210> 43
<211> 22
<212> PRT
<213> 人工序列
<220>
<223> 3X FLAG tag
<400> 43
Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr
1 5 10 15
Lys Asp Asp Asp Asp Lys

20

<210> 44
<211> 32
<212> PRT
<213> 人工序列
<220>
<223> 肽接头1
<400> 44

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Ser Glu Thr Pro Gly Thr
1 5 10 15
Ser Glu Ser Ala Thr Pro Glu Ser Ser Gly Gly Ser Ser Gly Gly Ser
20 25 30

<210> 45
<211> 16
<212> PRT
<213> 人工序列
<220>
<223> 核质蛋白NLS
<400> 45

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 46
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 肽接头2
<400> 46

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
1 5 10

<210> 47
<211> 230
<212> PRT
<213> 葡萄球菌属的几个种
<223> APG09980_胞苷脱氨酶
<400> 47

Met Ala Ala Gly Pro Ala Pro Glu Ala Arg Ser Leu Met Asp Glu Gln
1 5 10 15
Thr Phe Leu Asp Asn Phe Asn Asn Leu Lys Tyr Pro Arg Lys Thr Tyr

20 25 30
Leu Cys Tyr Glu Val Glu Leu Leu Val Gly Glu Asn His Ile Pro Leu
35 40 45
Asp Asp Tyr Lys Gly Phe Val His Asn Glu Gly Phe Asp Met Gly Leu
50 55 60
Glu Arg Cys His Ala Glu Leu Ile Phe Leu Glu Arg Met Ala Ser Trp
65 70 75 80
Asn Leu Asp Thr Glu Leu Arg Tyr Arg Ile Thr Val Phe Ile Ser Trp
85 90 95
Ser Pro Cys Pro Glu Cys Ala Asp Glu Leu Val Lys Phe Leu Arg Glu
100 105 110
Asn Arg His Val Asn Leu Arg Ile Phe Ala Ala Arg Ile Tyr Asp Trp
115 120 125
Tyr Gln Gly Tyr Glu Ala Gly Leu Arg Ala Leu Lys Ala Ala Gly Ala
130 135 140
Glu Val Ala Met Met Thr Leu His Glu Phe Glu Tyr Cys Trp Asn Asn
145 150 155 160
Phe Val Asp His Gln Gln Asp Glu Asp Thr Pro Phe Pro Pro Trp Asp
165 170 175
Asn Leu Val Ala Arg Ser Glu Glu Leu Ser Gln Arg Leu Glu Gly Ile
180 185 190
Leu Gln Pro Ser Val Leu Val Phe Cys Trp Pro Ser Gln Val Ser Val
195 200 205
Thr Ala Ala His Ser Asp Ile Met Ser Gln Ala Ser Arg Ala Trp Glu
210 215 220
Lys Arg Arg Asp Pro Pro
225 230

<210> 48
<211> 182
<212> PRT
<213> 人工序列
<220>
<223> APG07386-CTD_胞苷脱氨酶
<400> 48

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

35 40 45

Asn Gln Val Tyr Pro Glu Ser Leu Cys His Ala Glu Arg Cys Phe Leu
50 55 60

Ser Trp Phe Cys Glu Asp Ile Leu Ser Pro Asn Thr Asp Tyr Arg Val
65 70 75 80

Thr Trp Tyr Thr Ser Trp Ser Pro Cys Leu Asp Cys Ala Gly Glu Val
85 90 95

Ala Glu Phe Leu Ala Arg His Ser Asn Val Lys Leu Ala Ile Phe Ala
100 105 110

Ala Arg Leu Tyr Tyr Phe Trp Asp Pro His Tyr Gln Gln Gly Leu Arg
115 120 125

Ser Leu Ser Glu Lys Gly Ala Ser Val Glu Ile Met Gly Tyr Lys Asp
130 135 140

Phe Lys Tyr Cys Trp Glu Asn Phe Val Tyr Asn Gly Asp Glu Pro Phe
145 150 155 160

Lys Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu Phe Leu Asp Ser Lys
165 170 175

Leu Gln Glu Ile Leu Gln
180

<210> 49
<211> 4158
<212> DNA
<213> 人工序列
<220>
<223> APG09980_nAPG07433.1
<400> 49

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc 120
agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga 180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac 240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc 300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg 360
atcaccgtgt tcatctcttg gtccccatgt cctgagtgcg ccgacgagct ggttaagttc 420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat 480
cagggctacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg 540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat 600
acacctttcc ccccctggga caacctggtc gccagatccg aggaactgag ccagagactg 660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc 720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc 780

cccagcggcg gatctagcgg cggctctagc ggcagtgaaa cccctggtac aagcgagtct 840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga 900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac 960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac 1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc 1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg 1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt acccccctgag caaaaagagc 1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg 1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtgaa 1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg 1380
agcttgtata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat 1440
aggaagagaa atagccccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag 1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat 1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc 1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag 1680
gctacataca ccttccaata cttctctgct ctggaccagg tgaatcggac cagactggga 1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag 1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacggaag 1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa 1920
ctgaagaaga tcgagaagaa gccttttatc aatctgaagg ccttctacga gatcaacaag 1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc 2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc 2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg 2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg 2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt 2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc 2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga 2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg 2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc 2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg 2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc 2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc 2700
tacagccagt ccttcatcga ctcctaccac aacaaggtcc tggtgtacag cgacgaaaac 2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag 2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt 2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac 2940
accagatacg cctctacctt cctgaagaac ttcatcgagc agaacctgca gttcaaggaa 3000
gccgaggaca accccagaaa aagacgggtg caaaccgtga acggcgttat caccgcccac 3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg 3120

gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgac cgagtactat 3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc 3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag 3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga 3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag 3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga 3480
gatttcaaga tggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga 3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc 3600
aaaaagggca caggtaacct gatcaagcgg gtgaaggtgg aaggacaggc caagagcttt 3660
gtgagggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt 3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag 3780
ctgcccaaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac 3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caagggagat 3900
gaggatcggt tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag 3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggtccggaa ggaaactcga 4080
agaaacttcc acagcggcgg aagcaaaaga cctgccgcta caaagaaggc cggccaggcc 4140
aagaaaaaga agtaatga 4158

<210> 50
<211> 4530
<212> DNA
<213> 人工序列
<220>
<223> APG09980_nAPG07433.1_APG03399
<400> 50

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc 120
agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga 180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac 240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc 300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg 360
atcaccgtgt tcatctcttg gtccccatgt cctgagtgcg ccgacgagct ggttaagttc 420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat 480
cagggctacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg 540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat 600
acacctttcc ccccctggga caacctggtc gccagatccg aggaactgag ccagagactg 660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc 720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc 780
cccagcggcg gatctagcgg cggctctagc ggcagtgaaa cccctggtac aagcgagtct 840

gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga 900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac 960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac 1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc 1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg 1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt acccccctgag caaaaagagc 1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg 1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtgaa 1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg 1380
agcttgtata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat 1440
aggaagagaa atagccccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag 1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat 1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc 1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag 1680
gctacataca ccttccaata cttctctgct ctggaccagg tgaatcggac cagactggga 1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag 1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacggaag 1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa 1920
ctgaagaaga tcgagaagaa gccttttatc aatctgaagg ccttctacga gatcaacaag 1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc 2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc 2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg 2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg 2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt 2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc 2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga 2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg 2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc 2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg 2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc 2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc 2700
tacagccagt ccttcatcga ctcctaccac aacaaggtcc tggtgtacag cgacgaaaac 2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag 2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt 2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac 2940
accagatacg cctctacctt cctgaagaac ttcatcgagc agaacctgca gttcaaggaa 3000
gccgaggaca accccagaaa aagacgggtg caaaccgtga acggcgttat caccgcccac 3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg 3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgac cgagtactat 3180

cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc 3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag 3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga 3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag 3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga 3480
gatttcaaga tggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga 3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc 3600
aaaaagggca caggtaacct gatcaagcgg gtgaaggtgg aaggacaggc caagagcttt 3660
gtgagggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt 3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag 3780
ctgcccaaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac 3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caagggagat 3900
gaggatcggt tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag 3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggtccggaa ggaaactcga 4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaaaac caccacacag 4140
gagctgaagc agtacatcac cagactgttc caactgagca acaacgagac atgggagtgc 4200
gaggccctgg aagaggccgc cgagaacatc ctgcctgaga gattcatcaa caacagcctg 4260
ctggcccacc tgacactgaa cacctacacc tactacaacg atgagctgca cgagctttct 4320
atctacccct tcctgatgta cgccaacgac caactgatca gcatcggcta cctggaccac 4380
ttcgacatgg acttccttta tctgaccgac accaagaaca ccatcatcga cgagcggcac 4440
ctgctgaaac agggcgagaa caaccacgag agcggcggaa gcaaaagacc tgccgctaca 4500
aagaaggccg gccaggccaa gaaaaagaag 4530

<210> 51
<211> 4530
<212> DNA
<213> 人工序列
<220>
<223> APG09980_nAPG07433.1_APG06702
<400> 51

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc 120
agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga 180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac 240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc 300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg 360
atcaccgtgt tcatctcttg gtccccatgt cctgagtgcg ccgacgagct ggttaagttc 420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat 480
cagggctacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg 540

accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat 600
acacctttcc ccccctggga caacctggtc gccagatccg aggaactgag ccagagactg 660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc 720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc 780
cccagcggcg gatctagcgg cggctctagc ggcagtgaaa cccctggtac aagcgagtct 840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga 900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac 960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac 1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc 1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg 1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt acccccctgag caaaaagagc 1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg 1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtgaa 1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg 1380
agcttgtata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat 1440
aggaagagaa atagccccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag 1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat 1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc 1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag 1680
gctacataca ccttccaata cttctctgct ctggaccagg tgaatcggac cagactggga 1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag 1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacggaag 1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa 1920
ctgaagaaga tcgagaagaa gccttttatc aatctgaagg ccttctacga gatcaacaag 1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc 2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc 2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg 2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg 2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt 2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc 2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga 2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg 2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc 2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg 2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc 2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc 2700
tacagccagt ccttcatcga ctcctaccac aacaaggtcc tggtgtacag cgacgaaaac 2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag 2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt 2880

ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac 2940
accagatacg cctctacctt cctgaagaac ttcatcgagc agaacctgca gttcaaggaa 3000
gccgaggaca accccagaaa aagacgggtg caaaccgtga acggcgttat caccgcccac 3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg 3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgac cgagtactat 3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc 3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag 3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga 3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag 3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga 3480
gatttcaaga tggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga 3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc 3600
aaaaagggca caggtaacct gatcaagcgg gtgaaggtgg aaggacaggc caagagcttt 3660
gtgagggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt 3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag 3780
ctgcccaaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac 3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caagggagat 3900
gaggatcggt tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag 3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggtccggaa ggaaactcga 4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaccat cgagaacgag 4140
ctgaaggaat acatcactga tctgttcgag ctgcccaagg acgagcagtg gctgtgcgag 4200
gccgtggacg aggtggccga taacatcctg cctgatgagt acgtgacctt gggccctctg 4260
agaaacaaga ccctggaaac cttcacctac tacagcaaca cactgcacga gcggaacatc 4320
taccccttca tcctgtacta ccagaacaga ctgatcgcca taggatatat cgacgaaaac 4380
tacgagatgg acttcatgta cctgcacaac accgccatca ccctgatcga ccagagatac 4440
ctgattgata agggcgacga caaccacgac agcggcggaa gcaaaagacc tgccgctaca 4500
aagaaggccg gccaggccaa gaaaaagaag 4530

<210> 52
<211> 4530
<212> DNA
<213> 人工序列
<220>
<223> APG09980_nAPG07433.1_APG05198
<400> 52

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc 120
agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga 180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac 240

gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc 300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg 360
atcaccgtgt tcatctcttg gtccccatgt cctgagtgcg ccgacgagct ggttaagttc 420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat 480
cagggctacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg 540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat 600
acacctttcc ccccctggga caacctggtc gccagatccg aggaactgag ccagagactg 660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc 720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc 780
cccagcggcg gatctagcgg cggctctagc ggcagtgaaa cccctggtac aagcgagtct 840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga 900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac 960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac 1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc 1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg 1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt acccccctgag caaaaagagc 1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg 1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtgaa 1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg 1380
agcttgtata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat 1440
aggaagagaa atagccccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag 1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat 1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc 1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag 1680
gctacataca ccttccaata cttctctgct ctggaccagg tgaatcggac cagactggga 1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag 1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacggaag 1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa 1920
ctgaagaaga tcgagaagaa gccttttatc aatctgaagg ccttctacga gatcaacaag 1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc 2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc 2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg 2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg 2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt 2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc 2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga 2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg 2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc 2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg 2580

aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc 2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc 2700
tacagccagt ccttcatcga ctcctaccac aacaaggtcc tggtgtacag cgacgaaaac 2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag 2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt 2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac 2940
accagatacg cctctacctt cctgaagaac ttcatcgagc agaacctgca gttcaaggaa 3000
gccgaggaca accccagaaa aagacgggtg caaaccgtga acggcgttat caccgcccac 3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg 3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgac cgagtactat 3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc 3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag 3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga 3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag 3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga 3480
gatttcaaga tggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga 3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc 3600
aaaaagggca caggtaacct gatcaagcgg gtgaaggtgg aaggacaggc caagagcttt 3660
gtgagggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt 3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag 3780
ctgcccaaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac 3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caagggagat 3900
gaggatcggt tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag 3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggtccggaa ggaaactcga 4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaccct ggaaaagcaa 4140
ctgaaggaat acatcaccaa cctgttcaac ctgcccaagg acgagaagtg ggagtgcgag 4200
agcatcgagg aagtggccga caacatcctg cctgatcagt acatcagact gggccctctg 4260
accaacaaga tcctgcatac atacacctac tacagcgaca ccctgcacga gcggcacatc 4320
taccccttca tcctttatta ccagaaacaa ctgatcgcca tcggctacat cgacgagaca 4380
aacgacatgg acttcctgta cctgcacaac accgtgatgc ctctgctgga ccagagacac 4440
ctgctggaaa aagagaacta caacaacgag agcggcggaa gcaaaagacc tgccgctaca 4500
aagaaggccg gccaggccaa gaaaaagaag 4530

<210> 53
<211> 4014
<212> DNA
<213> 人工序列
<220>
<223> APG07386-CTD_nAPG07433.1

〈400〉 53

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc 120
tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtccccatgt 360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaacgt gaagctggct 420
atctttgccg ctagactgta ctatttctgg gaccctcact accagcaggg cctgagaagc 480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac 600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc 660
tctagcggca gtgaaacccc tggtacaagc gagtctgcca cacctgagag cagcggcgga 720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840
gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900
gccagcttag ctgaacccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020
gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggctc 1080
gacggcctgg acagactcct caatcatttc gagtgggcca gactgctgat ccacctggct 1140
cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag 1200
gtgctgagca gcatccaact gaacgagaaa cggctgagct tgtatagaac cgtgggcgag 1260
atgtggatga aggaccccga cttctctaaa tacgatagga agagaaatag ccccaacgaa 1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgaccct gttcgccgcc 1380
cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc 1440
tggacccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc 1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc 1560
tctgctctgg accaggtgaa tcggaccaga ctgggacctg atttccagcc cttcaccaag 1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa 1680
accatccccg aggtgaccta ctatgacata cggaagtggc tggaattgga cgagacaatt 1740
cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct 1800
tttatcaatc tgaaggcctt ctacgagatc aacaaggtgg tggccaacta cagcgaaaga 1860
accaacgaga ccttctccac cctggactac gacggcatcg gctacgccct gaccgtgtac 1920
aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc 1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg 2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa 2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc 2160
ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggccct gacccaggcc 2220
agaaaagtgg tgaacgccat aattagaaga cacggatctc ctcactccgt gcacatcgag 2280

ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat 2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac 2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc 2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac 2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc 2580
taccacaaca aggtcctggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct 2640
tatacctact tcctggaaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc 2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga 2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgcctc taccttcctg 2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga 2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag 2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca 3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc 3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac 3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc 3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag 3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg 3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa 3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc 3420
aagaaggcct tcgagacccc tctgtacaag cccagcaaaa agggcacagg taacctgatc 3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtgaa cggcggagtg 3540
gcccaaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg 3600
gtccccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc 3660
tctaagggct atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg 3720
tacccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc 3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag 3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg 3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttccacag cggcggaagc 3960
aaaagacctg ccgctacaaa gaaggccggc aggccaaga aaaagaagta atga 4014

<210> 54
<211> 4386
<212> DNA
<213> 人工序列
<220>
<223> APG07386-CTD_nAPG07433.1_APG03399
<400> 54

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc 120
tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180

acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtccccatgt 360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaacgt gaagctggct 420
atctttgccg ctagactgta ctatttctgg gaccctcact accagcaggg cctgagaagc 480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac 600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc 660
tctagcggca gtgaaacccc tggtacaagc gagtctgcca cacctgagag cagcggcgga 720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840
gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900
gccagcttag ctgaacccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020
gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggctc 1080
gacggcctgg acagactcct caatcatttc gagtgggcca gactgctgat ccacctggct 1140
cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag 1200
gtgctgagca gcatccaact gaacgagaaa cggctgagct tgtatagaac cgtgggcgag 1260
atgtggatga aggaccccga cttctctaaa tacgatagga agagaaatag ccccaacgaa 1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgaccct gttcgccgcc 1380
cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc 1440
tggacccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc 1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc 1560
tctgctctgg accaggtgaa tcggaccaga ctgggacctg atttccagcc cttcaccaag 1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa 1680
accatccccg aggtgaccta ctatgacata cggaagtggc tggaattgga cgagacaatt 1740
cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct 1800
tttatcaatc tgaaggcctt ctacgagatc aacaaggtgg tggccaacta cagcgaaaga 1860
accaacgaga ccttctccac cctggactac gacggcatcg gctacgccct gaccgtgtac 1920
aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc 1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg 2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa 2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc 2160
ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggccct gacccaggcc 2220
agaaaagtgg tgaacgccat aattagaaga cacggatctc ctcactccgt gcacatcgag 2280
ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat 2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac 2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc 2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac 2520

ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc 2580
taccacaaca aggtcctggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct 2640
tatacctact tcctggaaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc 2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga 2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgcctc taccttcctg 2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga 2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag 2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca 3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc 3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac 3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc 3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag 3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg 3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa 3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc 3420
aagaaggcct tcgagacccc tctgtacaag cccagcaaaa agggcacagg taacctgatc 3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtgaa cggcggagtg 3540
gcccaaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg 3600
gtccccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc 3660
tctaagggct atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg 3720
taccccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc 3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag 3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg 3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttccacag tggaggcagc 3960
ggcggatctg gaggcagcat gaaaaccacc acacaggagc tgaagcagta catcaccaga 4020
ctgttccaac tgagcaacaa cgagacatgg gagtgcgagg ccctggaaga ggccgccgag 4080
aacatcctgc ctgagagatt catcaacaac agcctgctgg cccacctgac actgaacacc 4140
tacacctact acaacgatga gctgcacgag ctttctatct acccccttcct gatgtacgcc 4200
aacgaccaac tgatcagcat cggctacctg gaccacttcg acatggactt cctttatctg 4260
accgacacca agaacaccat catcgacgag cggcacctgc tgaaacaggg cgagaacaac 4320
cacgagagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa 4380
aagaag 4386

<210> 55
<211> 4386
<212> DNA
<213> 人工序列
<220>
<223> APG07386-CTD_nAPG07433.1_APG06702
<400> 55

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc 120
tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtccccatgt 360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaacgt gaagctggct 420
atctttgccg ctagactgta ctatttctgg gaccctcact accagcaggg cctgagaagc 480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac 600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc 660
tctagcggca gtgaaacccc tggtacaagc gagtctgcca cacctgagag cagcggcgga 720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840
gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900
gccagcttag ctgaacccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020
gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggctc 1080
gacggcctgg acagactcct caatcatttc gagtgggcca gactgctgat ccacctggct 1140
cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag 1200
gtgctgagca gcatccaact gaacgagaaa cggctgagct tgtatagaac cgtgggcgag 1260
atgtggatga aggaccccga cttctctaaa tacgatagga agagaaatag ccccaacgaa 1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgaccct gttcgccgcc 1380
cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc 1440
tggacccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc 1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc 1560
tctgctctgg accaggtgaa tcggaccaga ctgggacctg atttccagcc cttcaccaag 1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa 1680
accatccccg aggtgaccta ctatgacata cggaagtggc tggaattgga cgagacaatt 1740
cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct 1800
tttatcaatc tgaaggcctt ctacgagatc aacaaggtgg tggccaacta cagcgaaaga 1860
accaacgaga ccttctccac cctggactac gacggcatcg gctacgccct gaccgtgtac 1920
aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc 1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg 2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa 2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc 2160
ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggccct gacccaggcc 2220
agaaaagtgg tgaacgccat aattagaaga cacggatctc ctcactccgt gcacatcgag 2280
ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat 2340

gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac 2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc 2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac 2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc 2580
taccacaaca aggtcctggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct 2640
tatacctact tcctggaaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc 2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga 2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgcctc taccttcctg 2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga 2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag 2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca 3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc 3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac 3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc 3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag 3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg 3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa 3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc 3420
aagaaggcct tcgagacccc tctgtacaag cccagcaaaa agggcacagg taacctgatc 3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtgaa cggcggagtg 3540
gcccaaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg 3600
gtccccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc 3660
tctaagggct atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg 3720
tacccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc 3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag 3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg 3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttccacag tggaggcagc 3960
ggcggatctg gaggcagcat gaccatcgag aacgagctga aggaatacat cactgatctg 4020
ttcgagctgc ccaaggacga gcagtggctg tgcgaggccg tggacgaggt ggccgataac 4080
atcctgcctg atgagtacgt gaccttgggc cctctgagaa acaagaccct ggaaaccttc 4140
acctactaca gcaacacact gcacgagcgg aacatctacc ccttcatcct gtactaccag 4200
aacagactga tcgccatagg atatatcgac gaaaactacg agatggactt catgtacctg 4260
cacaacaccg ccatcaccct gatcgaccag agatacctga ttgataaggg cgacgacaac 4320
cacgacagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa 4380
aagaag 4386

<210> 56
<211> 4386
<212> DNA
<213> 人工序列

<220>

<223> APG07386-CTD_nAPG07433.1_APG05198

<400> 56

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc 120
tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtccccatgt 360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaacgt gaagctggct 420
atctttgccg ctagactgta ctatttctgg gaccctcact accagcaggg cctgagaagc 480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac 600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc 660
tctagcggca gtgaaacccc tggtacaagc gagtctgcca cacctgagag cagcggcgga 720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840
gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900
gccagcttag ctgaacccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020
gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggctc 1080
gacggcctgg acagactcct caatcatttc gagtgggcca gactgctgat ccacctggct 1140
cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag 1200
gtgctgagca gcatccaact gaacgagaaa cggctgagct tgtatagaac cgtgggcgag 1260
atgtggatga aggaccccga cttctctaaa tacgatagga agagaaatag ccccaacgaa 1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgaccct gttcgccgcc 1380
cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc 1440
tggacccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc 1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc 1560
tctgctctgg accaggtgaa tcggaccaga ctgggacctg atttccagcc cttcaccaag 1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa 1680
accatccccg aggtgaccta ctatgacata cggaagtggc tggaattgga cgagacaatt 1740
cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct 1800
tttatcaatc tgaaggcctt ctacgagatc aacaaggtgg tggccaacta cagcgaaaga 1860
accaacgaga ccttctccac cctggactac gacggcatcg gctacgccct gaccgtgtac 1920
aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc 1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg 2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa 2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc 2160

ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggccct gacccaggcc 2220
agaaaagtgg tgaacgccat aattagaaga cacggatctc ctcactccgt gcacatcgag 2280
ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat 2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac 2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc 2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac 2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc 2580
taccacaaca aggtcctggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct 2640
tatacctact tcctggaaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc 2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga 2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgcctc taccttcctg 2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga 2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag 2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca 3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc 3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac 3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc 3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag 3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg 3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa 3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc 3420
aagaaggcct tcgagacccc tctgtacaag cccagcaaaa agggcacagg taacctgatc 3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtgaa cggcggagtg 3540
gcccaaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg 3600
gtccccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc 3660
tctaagggct atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg 3720
tacccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc 3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag 3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg 3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttccacag tggaggcagc 3960
ggcggatctg gaggcagcat gaccctggaa aagcaactga aggaatacat caccaacctg 4020
ttcaacctgc ccaaggacga gaagtgggag tgcgagagca tcgaggaagt ggccgacaac 4080
atcctgcctg atcagtacat cagactgggc cctctgacca acaagatcct gcatacatac 4140
acctactaca gcgacaccct gcacgagcgg cacatctacc ccttcatcct ttattaccag 4200
aaacaactga tcgccatcgg ctacatcgac gagacaaacg acatggactt cctgtacctg 4260
cacaacaccg tgatgcctct gctggaccag agacacctgc tggaaaaaga gaactacaac 4320
aacgagagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa 4380
aagaag 4386

〈210〉 57

<211> 20
<212> DNA
<213> 人工序列
<220>
<223> SGN000143靶序列
<400> 57
catggcagta cattagagca 20
<210> 58
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> SGN000169靶序列
<400> 58
cacatctcga gcaagacgtt 20
<210> 59
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> SGN000173靶序列
<400> 59
cttctatagc ctccttcccc 20
<210> 60
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> SGN000929靶序列
<400> 60
atctggaggg aacttacagc atatg 25
<210> 61
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> SGN000930靶序列
<400> 61
gaacaactca aatggaaatg aatat 25

<210> 62
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> SGN001101靶序列
<400> 62
atattttctt taatggtgcc aggca 25
<210> 63
<211> 130
<212> RNA
<213> 人工序列
<220>
<223> SGN000143 sgRNA序列
<400> 63
cauggcagua cauuagagca gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60
uaugauaagg guuucgaccc guggcgucgg ggaucgccug cccauugaaa ugggcuucuc 120
cccauuuauu 130
<210> 64
<211> 130
<212> RNA
<213> 人工序列
<220>
<223> SGN000169 sgRNA序列
<400> 64
cacaucucga gcaagacguu gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60
uaugauaagg guuucgaccc guggcgucgg ggaucgccug cccauugaaa ugggcuucuc 120
cccauuuauu 130
<210> 65
<211> 130
<212> RNA
<213> 人工序列
<220>
<223> SGN000173 sgRNA序列
<400> 65
cuucuauagc cuccuucccc gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60
uaugauaagg guuucgaccc guggcgucgg ggaucgccug cccauugaaa ugggcuucuc 120
cccauuuauu 130
<210> 66

<211> 135
<212> RNA
<213> 人工序列
<220>
<223> SGN000929 sgRNA序列
<400> 66
aucuggaggg aacuuacagc auauggucau aguuccauga aagccaaaag uggcuuugau 60
guuucuauga uaaggguuuc ggcccguggc gucggggauc gccugcccau uccgaugggc 120
uucucccccau uuauu 135
<210> 67
<211> 135
<212> RNA
<213> 人工序列
<220>
<223> SGN000930 sgRNA序列
<400> 67
gaacaacuca aauggaaaug aauaugucau aguuccauga aagccaaaag uggcuuugau 60
guuucuauga uaaggguuuc ggcccguggc gucggggauc gccugcccau uccgaugggc 120
uucucccccau uuauu 135
<210> 68
<211> 130
<212> RNA
<213> 人工序列
<220>
<223> SGN001101 sgRNA序列
<400> 68
auauuuucuu uaauggugcc aggcagucau aguuccauua aagccaaaag uggcuuugau 60
guuucuauga uaaggguuuc gacccguggc gucggggauc gccugcccau ugaaaugggc 120
uucucccccau 130
<210> 69
<211> 200
<212> DNA
<213> 人工序列
<220>
<223> SGN000143基因组位点
<400> 69
gaagtcagca atcttcagct ctcccttgag ccctaagagc agattttctg gctttatgtc 60
tctgtgaatc accttcttcc catggcagta cattagagca tctgccaact cctccatgat 120
ctgggagggg caacgacagg caatcagaca aggatggacc tccagctaca agcagcacac 180

cctcagcctc gagcccccta 200

<210> 70

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> SGN000169基因组位点

<400> 70

tccagcaggt cagcaaagaa tttatagccc cccttgagca cacagagggc tacaatgtga 60

tggcctccca tctccttcat cacatctcga gcaagacgtt cagtcctaca gaaataaaat 120

caggaattta atagaaagtt tcatacatta aactttataa caaacacctc ttagtcatta 180

aacttccaca ccaacctggg 200

<210> 71

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> SGN000173基因组位点

<400> 71

gtataactgg ccagccttta atacattgct gcttagagtc aaagcatgta cttagagttg 60

gtatgattta tctttttggt cttctatagc ctccttcccc atccccatca gtcttaatca 120

gtcttgttac gttatgacta atctttgggg atttgtgcag aatgttattt tagataagca 180

aaaacgagca aaatagggga 200

<210> 72

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> SGN000929基因组位点

<400> 72

gactctccca ccccagacct cctagttctt gggcaggtga acaactacag cctttccatg 60

cctccttgtg gagagatctg gagggaactt acagcatatg agcacctgct agggagagcc 120

cctggccctg gcaggaccga tacctcggct cctggacacc atttctattt actcaatata 180

cactgtttgt acaaagtaca 200

<210> 73

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> SGN000930基因组位点
<400> 73
tttatcatca caatacctcg ggtccagatg ggcccttag tgagttccgc taccacagga 60
aggatgactc aataagaaca actcaaatgg aaatgaatat gggcccaggg tgggtgtgag 120
gagtggcagg ggattcttga gagacaggac cattgcccac agcctatgtg agtactgttg 180
cttgtcttga aagaaggcaa 200
<210> 74
<211> 200
<212> DNA
<213> 人工序列
<220>
<223> SGN001101基因组位点
<400> 74
tttcttacct cttctagttg gcatgctttg atgacgcttc tgtatctata ttcatcatag 60
gaaacaccaa agatgatatt ttctttaatg gtgccaggca taatccagga aaactgagaa 120
cagaatgaaa ttcttccact gtgcttaatt ttaccctctg aaggctccag ttctcccata 180
atcaccatta gaagtgaagt 200
<210> 75
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 大范围核酸酶家族基序
<400> 75
Leu Ala Gly Leu Ile Asp Ala Asp Gly
1 5
<210> 76
<211> 182
<212> PRT
<213> 人工序列
<220>
<223> 截短的APG09980_胞苷脱氨酶
<400> 76
Met Asp Glu Gln Thr Phe Leu Asp Asn Phe Asn Asn Leu Lys Tyr Pro
1 5 10 15
Arg Lys Thr Tyr Leu Cys Tyr Glu Val Glu Leu Leu Val Gly Glu Asn
20 25 30
His Ile Pro Leu Asp Asp Tyr Lys Gly Phe Val His Asn Glu Gly Phe
35 40 45

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

<210> 77
<211> 201
<212> PRT
<213> 人工序列
<220>
<223> APG05840_胞苷脱氨酶
<400> 77

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

Ala Glu Phe Leu His Glu Asn Pro His Val Asn Leu Arg Ile Phe Ala
115 120 125
Ala Arg Ile Tyr Asp Tyr Leu Pro Arg Tyr Glu Glu Gly Leu Gln Met
130 135 140
Leu Gln Asn Ala Gly Ala Gln Val Ser Ile Met Thr Ser Glu Glu Phe
145 150 155 160
Gly His Cys Trp Asp Thr Phe Val Asp Arg Gln Gly His Pro Phe Gln
165 170 175
Pro Trp Glu Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu
180 185 190
Gln Ala Ile Leu Gln Asn Gln Gly Asn
195 200

<210> 78
<211> 185
<212> PRT
<213> 人工序列
<220>
<223> 截短的APG05840_胞苷脱氨酶
<400> 78

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

His Pro Phe Gln Pro Trp Glu Gly Leu Asp Glu His Ser Gln Ala Leu
165 170 175
Ser Gly Arg Leu Gln Ala Ile Leu Gln
180 185

<210> 79
<211> 192
<212> PRT
<213> 人工序列
<220>
<223> APG00868_胞苷脱氨酶
<400> 79

Met Glu Pro Trp Arg Pro Ser Pro Arg Asn Pro Met Asp Arg Ile Asp
1 5 10 15
Pro Asn Thr Phe Arg Phe His Phe Pro Asn Leu Leu Tyr Ala Ser Gly
20 25 30
Arg Lys Leu Cys Tyr Leu Cys Phe Gln Val Glu Thr Gly Asp Tyr Phe
35 40 45
Ser Cys Asp Asp Ser Asp Arg Gly Val Phe Arg Asn Lys Val His Pro
50 55 60
Trp Ala Arg Cys His Ala Glu Gln Cys Phe Leu Ser Trp Phe Arg Asp
65 70 75 80
Gln Tyr Pro Cys Arg Asp Glu Tyr Tyr Asn Val Thr Trp Phe Leu Ser
85 90 95
Trp Ser Pro Cys Pro Thr Cys Ala Glu Glu Val Val Glu Phe Leu Glu
100 105 110
Glu Tyr Arg Asn Leu Thr Leu Ser Ile Phe Thr Ser Arg Leu Tyr Tyr
115 120 125
Phe Tyr His Pro Asn Tyr Gln Gln Gly Leu Arg Lys Leu Trp Asp Ala
130 135 140
Gly Val Gln Leu Asp Ile Met Ser Cys Asp Asp Phe Glu His Cys Trp
145 150 155 160
Asp Asn Phe Val Asp His Lys Gly Met Arg Phe Gln Arg Arg Asn Leu
165 170 175
Leu Lys Asp Tyr Asp Phe Leu Ala Ala Glu Leu Gln Glu Ile Leu Arg
180 185 190

<210> 80
<211> 178
<212> PRT
<213> 人工序列

<220>
<223> 截短的APG00868_胞苷脱氨酶
<400> 80

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

<210> 81
<211> 152
<212> PRT
<213> 人工序列
<220>
<223> APG30125_胞苷脱氨酶
<400> 81

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

Gly Asn Ser Arg Tyr Ala Phe Gly Asn Ser Gln Met Gln Val Glu Met
50 55 60
Pro Pro Asn Ile Tyr Ala His Asn Trp Val Lys Leu His Gly Glu Met
65 70 75 80
Thr Ala Leu Val Ala Ala Ile Arg Arg Ile Glu Arg Val Asp Ser Thr
85 90 95
Ser Ala Val Leu Pro Ile Thr Ser Ala Tyr Ile Glu Leu Arg Pro Cys
100 105 110
Glu Ala Ser Cys Leu Pro Ala Leu His Asn Met Leu Pro Asp Asn Ile
115 120 125
Thr Val Tyr Phe Ser Phe Leu His Pro Thr Gln Val Asp Gln Trp Lys
130 135 140
Gln Ser Ala Arg Ala Leu Cys Ala
145 150

<210> 82
<211> 152
<212> PRT
<213> 人工序列
<220>
<223> APG30126_胞苷脱氨酶
<400> 82

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

Ala Ala Ala Lys His Leu Cys Ala
145 150

<210> 83
<211> 148
<212> PRT
<213> 人工序列
<220>
<223> APG30127_胞苷脱氨酶
<400> 83

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

<210> 84
<211> 147
<212> PRT
<213> 人工序列
<220>
<223> APG30128_胞苷脱氨酶
<400> 84

Met Ile Lys Thr Glu Tyr Thr Leu Arg Ser Gly Tyr Pro Ile Val Arg
1 5 10 15
Arg Thr Leu Glu Asn Lys Lys Asn Leu Val Lys Gln Pro Gly Phe Gly

20 25 30
Pro Glu Ser Cys Cys Ala Val Val Glu Tyr Arg Leu Arg Gly Asn Ile
35 40 45
Arg Tyr Ala Phe Gly Asn Ser Arg Met Gln Val Ser Val Pro Pro Gly
50 55 60
Ile Tyr Thr Asn Asn Trp Val Arg Leu His Gly Glu Met Ala Ala Leu
65 70 75 80
Val Ala Ala Ile Glu Arg Ile Glu Arg Phe Ser Ser Asp Asp Val Ile
85 90 95
Pro Ile Thr Ala Ala Tyr Ile Glu Leu Arg Pro Cys Glu Ala Asn Cys
100 105 110
Met Gln Ala Leu His Asn Ile Leu Pro Glu Asn Ala Asn Val Tyr Tyr
115 120 125
Ser Phe Asp His Pro Thr Gln Val Glu Glu Trp Lys Leu Arg Ala His
130 135 140
Glu Leu Cys
145

<210> 85
<211> 147
<212> PRT
<213> 人工序列
<220>
<223> APG30129_胞苷脱氨酶
<400> 85
Met Ile Lys Thr Glu Tyr Ala Leu Asn Ser Gly Tyr Pro Ile Val Arg
1 5 10 15
Arg Thr Leu Glu Asp Lys Lys Lys Leu Val Lys Gln Pro Gly Phe Gly
20 25 30
Pro Glu Ser Cys Cys Ala Val Val Glu Tyr Arg Leu Arg Gly Asn Ile
35 40 45
Arg Tyr Ala Phe Gly Asn Ser Gln Met His Val Ser Met Pro Pro Gly
50 55 60
Ile Tyr Thr His Asn Trp Val Lys Leu His Gly Glu Met Ala Ala Leu
65 70 75 80
Val Ala Ala Ile Asn Arg Ile Glu Arg Phe Ser Thr Asp Asp Val Ile
85 90 95
Pro Ile Thr Gly Ala Tyr Ile Glu Leu Arg Pro Cys Glu Ala Asn Cys
100 105 110
Leu Gln Ala Leu His Asn Ile Leu Pro Glu Asp Ala Asn Val Tyr Tyr

115 120 125
Ser Phe Asp His Pro Thr Gln Leu Asp Glu Trp Lys Leu Arg Ala His
130 135 140
Glu Leu Cys
145
<210> 86
<211> 147
<212> PRT
<213> 人工序列
<220>
<223> APG30130_胞苷脱氨酶
<400> 86
Met Ile Gln Leu Val Gln Ala Glu Tyr Ser Ile Lys Ser Gly Tyr Pro
1 5 10 15
Ile Val Arg Arg Thr Leu Glu Asp Lys Lys Lys Leu Ile Glu Lys Pro
20 25 30
Gly Phe Gly Pro Glu Ser Cys Cys Ala Thr Ile Glu Tyr Gln Leu Arg
35 40 45
Gly Ser Thr Arg Tyr Ala Phe Gly Asn Ser Gln Met Lys Met Glu Met
50 55 60
Pro Pro Asp Ile Tyr Thr His Asn Trp Val Lys Leu His Ala Glu Met
65 70 75 80
Ala Ala Leu Val Ala Ala Ile Arg Arg Ile Glu Arg Phe Asp Ala Asp
85 90 95
Lys Glu Gln Val Pro Ile Thr Asn Val Tyr Ile Glu Leu Arg Pro Cys
100 105 110
Glu Ala Asn Cys Met Gln Ala Leu Gln Asn Ile Leu Pro Asp Gly Thr
115 120 125
Thr Val Tyr Tyr Ser Phe Leu His Pro Thr Glu Val Glu Glu Trp Lys
130 135 140
Arg Ser Ala
145
<210> 87
<211> 235
<212> PRT
<213> 人工序列
<220>
<223> APG05241_胞苷脱氨酶
<400> 87

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

<210> 88
<211> 262
<212> PRT
<213> 人工序列
<220>
<223> APG07280_胞苷脱氨酶
<400> 88

Met His His Ser Ala Arg Leu Pro Pro Asn Cys Ile Val Ser Arg Tyr
1 5 10 15

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

<210> 89
<211> 211
<212> PRT
<213> 人工序列
<220>
<223> APG09260_胞苷脱氨酶
<400> 89

Met Glu Asp Phe His Cys Val Cys Val Tyr Cys Pro Leu Gln Val Ser
1 5 10 15
Val Met Ser Asp Tyr Gln Ala Ala Ser Ser Ser Pro Asp Cys Pro Gln
20 25 30
Cys Pro Gln His Ile Leu Thr Gly Ala Glu Ala Ala Val Ser Tyr Ser
35 40 45
Gln Phe Cys Glu Thr Phe Gly Phe His Met Gly Pro Ser Gly Ala Arg
50 55 60
Ala Leu Leu Leu Phe Tyr Glu Leu Trp Gly Pro Ser Gly Thr Leu Val
65 70 75 80
Gln Arg Gly Gln Ala Ser Asn Leu Leu Glu Val Cys Glu Glu Val Leu
85 90 95
Tyr Ser Asn Tyr Met Pro Cys Gln Glu Cys Ser Gln Thr Leu Ile Ser
100 105 110
Phe Leu Leu Arg Tyr Pro Trp Val Arg Leu Asp Leu Leu Phe Ser Gln
115 120 125
Leu Tyr His Thr Ala Pro Ser Gln Thr His Ser Leu Asp Asn Gln Thr
130 135 140
Gly Leu Arg Ser Leu Ala Val Leu Thr Leu Ser Pro Asn Ser Gly Ala
145 150 155 160
Ala Trp Gly His Leu Leu Arg Cys Phe Val Arg Asp Val Pro Pro Ser
165 170 175
Ala Leu Gln Leu Pro Leu Leu Pro Glu Arg Val Glu Ala Asp Arg Val
180 185 190
Asn Ala Ile His Ile Ser Ala Thr Thr Gly Ile Gly Pro Ala Phe Trp
195 200 205
Thr Ser His
210

<210> 90
<211> 234
<212> PRT
<213> 人工序列
<220>
<223> APG08360_胞苷脱氨酶
<400> 90

Met Asp Pro Gln Arg Leu Arg Gln Trp Pro Gly Pro Gly Pro Ala Ser
1 5 10 15
Arg Gly Gly Tyr Gly Gln Arg Pro Arg Ile Arg Asn Pro Glu Glu Trp
20 25 30

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

<210> 91
<211> 395
<212> PRT
<213> 人工序列
<220>
<223> APG07386_胞苷脱氨酶
<400> 91

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

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

355 360 365

Gly Asp Glu Pro Phe Lys Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu
370 375 380

Phe Leu Asp Ser Lys Leu Gln Glu Ile Leu Gln
385 390 395

<210> 92

<211> 213

<212> PRT

<213> 人工序列

<220>

<223> APG07386-NTD_胞苷脱氨酶

<400> 92

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Gly Met Asp Arg His Ala
1 5 10 15

Phe Asn Tyr Asn Phe Glu Asn Glu Pro Ile Leu Tyr Gly Arg Ser Tyr
20 25 30

Thr Trp Leu Cys Tyr Glu Val Lys Ile Arg Lys Asp Pro Ser Lys Leu
35 40 45

Pro Trp Asp Thr Gly Val Phe Arg Gly Gln Val Arg Pro Lys Leu Gln
50 55 60

Ser Asn Arg Arg Tyr Glu Leu Ser Asn Trp Glu Cys Arg Lys His Val
65 70 75 80

Tyr Phe Gln Pro Gln Tyr His Ala Glu Met Cys Phe Leu Ser Trp Phe
85 90 95

Cys Gly Asn Gln Leu Pro Ala His Lys Arg Phe Gln Ile Thr Trp Phe
100 105 110

Val Ser Trp Thr Pro Cys Pro Asp Cys Val Ala Lys Val Thr Glu Phe
115 120 125

Leu Ala Glu His Pro Asn Val Thr Leu Thr Ile Ser Val Ala Arg Leu
130 135 140

Tyr Tyr Tyr Arg Gly Lys Asp Trp Arg Arg Ala Leu Cys Arg Leu His
145 150 155 160

Gln Ala Gly Ala Arg Val Lys Ile Met Asp Tyr Glu Glu Phe Ala Tyr
165 170 175

Cys Trp Glu Asn Phe Val Tyr Asn Glu Gly Gln Ser Phe Met Pro Trp
180 185 190

Asp Lys Phe Asp Asp Asn Tyr Ala Phe Leu His His Lys Leu Lys Glu
195 200 205

Ile Leu Arg Asn Pro

210

<210> 93
<211> 267
<212> PRT
<213> 人工序列
<220>
<223> APG09688_胞苷脱氨酶
<400> 93

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

245 250 255

Thr Gly Leu Leu Gln Pro Ser Val Pro Trp Arg
260 265

<210> 94
<211> 373
<212> PRT
<213> 人工序列
<220>
<223> APG02316_胞苷脱氨酶
<400> 94

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

225 230 235 240
Asp Asp Ile Phe Asn Ile Asn Gln Ser Leu Ala Leu Gln Asp Ser Ile
245 250 255
Asn Ile Ala Leu Ala Thr Phe Asp Asn Lys Glu Phe Lys Ala Ile Ser
260 265 270
Gly Lys Asn Asp Ile Glu Asp Tyr Cys Lys Val Ile Pro Ile Glu Lys
275 280 285
Ser Arg Phe Lys Ser Arg Lys Val Asn Lys Leu Gly Glu Leu Gly Ile
290 295 300
Gly Tyr Asn Arg Ile Asn Asp Ser Glu Lys Lys Ile Leu Glu Glu Ile
305 310 315 320
His Glu Lys Ile Lys Gln Lys Ile Leu Lys Asp Arg Gly Asn Leu Thr
325 330 335
Leu Tyr Thr Lys Trp Glu Pro Cys Pro Ser Cys Tyr Phe Val Ile Ser
340 345 350
Gln Phe Cys Glu Lys Tyr Pro Asn Ile Lys Val Glu Val Lys Tyr Asn
355 360 365
Lys Lys Tyr Gly Glu
370

<210> 95
<211> 1060
<212> PRT
<213> 芽孢杆菌属物种
<223> APG00969_RGN
<400> 95

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

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

420 425 430

Asp Leu His Asn Cys Asn Glu Ile Pro Val Asn Met Ile Asn Asp Leu

435 440 445

Ile Val Ser Pro Val Val Arg Arg Ser Leu Thr Gln Ser Ile Glu Met

450 455 460

Ile Asn Gln Ile Ile Lys Asp Tyr Gly His Pro Arg Glu Ile Val Ile

465 470 475 480

Glu Leu Ala Arg Glu Lys Asn Ser Glu Glu Lys Lys Asn Phe Ile Lys

485 490 495

Ser Leu Asn Glu Lys Asn Lys Gln Ile Asn Asp Glu Val Ile Glu Lys

500 505 510

Leu Asn Ala Ser Asn His Arg Asp Asn Lys Gly Met Phe Asn Lys Val

515 520 525

Lys Leu Trp Ile Leu Gln Asp Gly His Cys Leu Tyr Ser Leu Lys Pro

530 535 540

Ile Arg Leu Glu Asp Leu Leu Asn Asn Pro Asn His Tyr Glu Ile Asp

545 550 555 560

His Ile Ile Pro Lys Ser Val Ser Phe Asp Asp Ser Met Ser Asn Lys

565 570 575

Val Leu Val Tyr Gln Ile Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro

580 585 590

Tyr Gln Tyr Leu Thr Ser Ala Asp Lys Thr Ile Thr Tyr Glu Lys Phe

595 600 605

Lys Ala Asn Ile Thr Gln Leu Ala Lys Ser Asn His Lys Ile Ser Lys

610 615 620

Lys Lys Leu Asp Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe His

625 630 635 640

Ile Lys Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala

645 650 655

Thr Arg Ser Leu Ile Asn Leu Leu Lys Tyr Tyr Phe Ser Glu Lys Asp

660 665 670

Ile Asn Val Lys Val Lys Ser Ile Asn Gly Ser Phe Thr Asp Tyr Leu

675 680 685

Arg Lys Leu Trp Asn Phe Pro Lys Asp Arg Glu Phe Tyr His Lys His

690 695 700

His Ala Glu Asp Ala Leu Ile Ile Ala Met Ala Asn Lys Ile Phe Thr

705 710 715 720

Thr Arg Lys Ile Phe Lys Glu Gln Asn Ser Val Phe Ser Asp Glu Gln

725 730 735

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

1040 1045 1050

Gln Thr Phe Arg Gly Asp Leu

1055 1060

<210> 96

<211> 1090

<212> PRT

<213> 根瘤菌属物种

<223> APG03128_RGN

<400> 96

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

Phe Arg Val Met Phe Tyr Gln Arg Pro Leu Lys Lys Pro Leu Val Gly
245 250 255
Arg Cys Ser Phe Asn Pro Ala Glu Ala Arg Leu Ser Arg Ala His Pro
260 265 270
Leu Phe Gln Glu Phe Arg Leu Tyr Lys Glu Val Asn Asp Leu Glu Val
275 280 285
Val Leu Pro Asp Gln Ser His Arg Lys Leu Thr Leu Asp Glu Arg Asn
290 295 300
Ala Leu Val Ala Lys Leu Lys Ser Ser Arg Lys Ala Ser Phe Ser Val
305 310 315 320
Leu Arg Arg Thr Leu Lys Leu Thr Pro Asp Leu Ala Phe Asn Lys Glu
325 330 335
Ser Glu Ala Arg Lys Asp Leu Leu Gly Asp Glu Ile Asn Ser Ala Leu
340 345 350
Ala Asp Ala Lys Met Phe Gly Val Arg Trp Gly Gly Phe Pro Arg Ala
355 360 365
Arg Gln Trp Glu Ile Ile Thr His Leu Lys Glu Glu Glu Asn Pro Ala
370 375 380
Arg Leu Ser Asp Trp Leu Lys Ser Glu Phe Gly Leu Asp Asp Glu Arg
385 390 395 400
Val Val Ala Ile Ala Asn Ile Ala Leu Pro Glu Gly Tyr Gly Arg Leu
405 410 415
Gly Glu Thr Ala Leu Ala Ser Met Leu Glu Glu Met Lys Thr Ala Val
420 425 430
Ile Pro Glu Ser Glu Ala Ala Lys Arg Cys Gly Tyr Asp His Ser Asn
435 440 445
Leu Ala Lys Glu Gln Asp Glu Gly Leu Asp Ile Leu Pro Glu Tyr Gln
450 455 460
Glu Ile Leu Glu Arg His Ile Pro Pro Gly Thr Gly Asp Pro Asp Asp
465 470 475 480
Ile Tyr Asp Ile Arg Lys Gly Arg Ile Thr Asn Pro Thr Val His Ile
485 490 495
Gly Leu Asn Gln Leu Arg Arg Val Val Asn Ala Leu Ile Lys Arg His
500 505 510
Gly Lys Pro Arg His Ile Val Val Glu Leu Ala Arg Asp Leu Gln Leu
515 520 525
Ser Glu Lys Gln Lys Ala Asp Val Asn Arg Arg Ile Ala Lys Asn Thr
530 535 540
Arg Glu Ala Glu Gly Arg Ser Gln Lys Leu Ile Glu Met Gly Gln Leu

545 550 555 560
Asp Thr Gly Tyr Asn Arg Leu Leu Leu Lys Leu Trp Glu Glu Leu Asn
565 570 575
Gln Asp Lys Pro Glu Asp Arg Val Cys Ile Tyr Ser Gly Lys Pro Ile
580 585 590
Gly Ile Ala Met Leu Phe Ser Gly Glu Val Asp Ile Asp His Ile Leu
595 600 605
Pro Trp Ser Lys Thr Leu Asp Asp Ser Gln Ala Asn Lys Leu Leu Cys
610 615 620
Leu Lys Ser Ala Asn Arg Gln Lys Arg Asn Arg Thr Pro Ser Asp Val
625 630 635 640
Pro Glu Trp Arg Asp Cys Tyr Glu Glu Val Leu Ala Arg Ala Ala Arg
645 650 655
Leu Pro Lys Asn Lys Arg Trp Arg Phe Ala Ala Asp Ala Met Gln Gln
660 665 670
Phe Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu Thr Asp Thr Gln
675 680 685
Tyr Leu Ser Arg Met Ala Phe Glu Tyr Leu Ser Ala Leu Phe Pro Ser
690 695 700
Glu Glu Ala Asp Lys Trp Gly Glu Leu Arg Gln Arg Lys Arg Val His
705 710 715 720
Val Val Pro Gly Arg Leu Thr Glu Leu Leu Arg Arg Asn Trp Gly Leu
725 730 735
Asn Thr Leu Leu Pro Asp His Asn Leu Gly Glu Met Ala Gln Glu Lys
740 745 750
Asn Arg Lys Asp His Arg His His Ala Ile Asp Ala Ala Val Ile Gly
755 760 765
Val Thr Ser Arg Ser Leu Leu Gln Arg Met Ser Gly Ala Ala Ala Arg
770 775 780
Leu Asp Asp Val Ala Phe Asp Asp Leu Val Arg Thr Val Val Lys Asp
785 790 795 800
Asn Pro Pro Trp Pro Gly Phe Arg Glu Glu Leu Leu Gly Cys Ile Asn
805 810 815
Arg Val Thr Val Ser His Lys Pro Asp His Gly Thr Val Ser Arg Ala
820 825 830
Ala Tyr Ala Gln Gly Lys Gly Gln Thr Ala Gly Lys Leu His Asn Asp
835 840 845
Thr Ala Tyr Gly Ile Thr Gly Leu Gln Asp Met Lys Gly Ser Pro Leu
850 855 860

Val Val Arg Arg Lys Pro Phe Met Ala Leu Glu Val Lys Asp Ile Ala
865 870 875 880
Ser Ile Arg Asp Thr Glu Leu Gln Ser Ala Leu Tyr Asp Ala Ile Gly
885 890 895
Ser Leu Thr Glu Lys Lys Ala Leu Gln Glu Ala Leu Val Arg Phe Arg
900 905 910
Asp Arg His Pro Gln Phe Lys Gly Ile Arg Arg Val Arg Val Leu Glu
915 920 925
Ala Leu Ser Val Ile Pro Ile Val Asp Asn Asn Gly Lys Ala Tyr Lys
930 935 940
Gly Tyr Lys Gly Asp Ala Asn Tyr Arg Tyr Glu Val Trp Glu Thr Leu
945 950 955 960
Asp Gly Arg Trp His Thr Glu Val Val Ser Met Phe Asp Ala His Gln
965 970 975
Pro Gly Trp Gln Ser Pro Phe His Arg Gln His Pro Ala Ala Arg Arg
980 985 990
Val Leu Lys Leu Gln Gln Asn Asp Met Val Ala Tyr Glu His Pro Ala
995 1000 1005
Asp Gly Tyr Thr Ile Ala Arg Val Val Lys Phe Ser Asn Asp Lys
1010 1015 1020
Arg Ile Tyr Phe Ala Ser His Arg Glu Ser Gly Ser Leu Lys Ala
1025 1030 1035
Arg Asp Ala Asp Lys Gly Asp Pro Phe Thr Tyr Phe Ala Lys Ala
1040 1045 1050
Asn Asn Gly Leu Arg Asp Ile Lys Cys Arg Gln Val Arg Ile Asp
1055 1060 1065
Ala Ala Gly Arg Val Phe Asp Pro Gly Pro Gln Asp Arg Glu Ala
1070 1075 1080
Arg Gln Ser Lys Gly Arg Ala
1085 1090

<210> 97
<211> 1108
<212> PRT
<213> Brevibacillus物种
<223> APG09748_RGN
<400> 97

Met Ala Ile Arg Ser Ile Lys Leu Lys Leu Lys Thr Arg Thr Gly Pro
1 5 10 15
Glu Ala Gln Asn Leu Arg Lys Gly Ile Trp Arg Thr His Arg Leu Leu

20 25 30

Asn Glu Gly Val Ala Tyr Tyr Met Lys Met Leu Leu Leu Phe Arg Gln
35 40 45

Glu Ser Thr Gly Gly Gln Thr Lys Lys Glu Leu Gln Glu Glu Leu Val
50 55 60

Arg His Ile Arg Glu Gln Gln Gln Lys Asn Arg Ala Asp Lys Asn Thr
65 70 75 80

Gln Ala Leu Pro Leu Asp Lys Ala Phe Ala Ala Leu Arg Gln Leu Tyr
85 90 95

Glu Leu Leu Val Pro Ser Ser Ile Gly Gln Ser Gly Asp Ala Gln Ile
100 105 110

Ile Ser Arg Lys Phe Leu Ser Pro Leu Val Asp Pro Asn Ser Glu Gly
115 120 125

Gly Lys Gly Thr Ser Lys Ala Gly Ala Lys Pro Thr Trp Gln Lys Lys
130 135 140

Lys Glu Ala Asn Asp Pro Thr Trp Glu Gln Asp Tyr Glu Lys Trp Lys
145 150 155 160

Lys Arg Arg Glu Glu Asp Pro Thr Ala Ser Val Ile Thr Thr Leu Glu
165 170 175

Glu Tyr Gly Ile Arg Pro Ile Phe Pro Leu Tyr Thr Asn Thr Val Ala
180 185 190

Asp Ile Ala Trp Leu Pro Leu Gln Ser Asn Gln Phe Val Arg Thr Trp
195 200 205

Asp Arg Asp Met Leu Gln Gln Ala Ile Glu Arg Leu Leu Ser Trp Glu
210 215 220

Ser Trp Asn Lys Arg Val Gln Glu Glu Tyr Ser Lys Leu Gln Glu Lys
225 230 235 240

Met Thr Gln Leu Asn Glu Gln Leu Glu Gly Gly Gln Glu Trp Ile Ser
245 250 255

Leu Leu Glu Gln Tyr Glu Glu Gln Arg Glu Gln Glu Leu Ile Glu Asn
260 265 270

Met Thr Ala Ala Asn Asp Lys Tyr Arg Ile Thr Lys Arg Gln Met Lys
275 280 285

Gly Trp Asn Glu Leu Tyr Glu Gln Trp Ser Thr Val Leu Pro Asn Ala
290 295 300

Ser His Glu Gln Tyr Arg Glu Ala Leu Lys Arg Val Gln Gln Arg Leu
305 310 315 320

Arg Gly Arg Phe Gly Asp Ala His Phe Phe Gln Tyr Leu Met Lys Glu
325 330 335

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

645 650 655
Gln Asn Arg Glu Arg Ala Asp Glu Leu Asp His Leu Ser Gln Ala Leu
660 665 670
Glu Lys Gln Lys Ser Leu Leu Asp Gln Thr Asp Arg Thr Phe Trp Asn
675 680 685
Gly Ile Val Cys Asp Leu Thr Asp Ala Leu Arg Glu Lys Glu Gly Gly
690 695 700
Trp Glu Gln Ala Val Val Gln Ile His Arg Lys Ala Glu Glu His Val
705 710 715 720
Gly Lys Val Val Gln Ala Trp Arg Lys Arg Phe Asp Ala Asp Glu Arg
725 730 735
Lys Gly Ile Ala Gly Leu Ser Met Trp Ser Ile Glu Glu Leu Asp Ser
740 745 750
Leu Arg Lys Leu Leu Ile Ser Trp Ser Arg Arg Thr Arg Asn Pro Arg
755 760 765
Glu Ile Asn Cys Phe Glu Gln Gly His Thr Ser His Gln Arg Leu Leu
770 775 780
Thr His Ile Gln Asn Val Lys Glu Asp Arg Leu Lys Gln Leu Ser His
785 790 795 800
Ala Ile Val Met Thr Ala Leu Gly Tyr Val Tyr Asp Glu Lys Lys Leu
805 810 815
Glu Trp Phe Ala Lys Tyr Pro Ala Cys Gln Val Ile Leu Phe Glu Asn
820 825 830
Leu Ser Gln Tyr Arg Ser Asn Met Asp Arg Ser Thr Lys Glu Asn Ser
835 840 845
Thr Leu Met Lys Trp Ala His Arg Ser Ile Pro Lys Tyr Val His Met
850 855 860
Gln Ala Glu Pro Tyr Gly Ile Gln Ile Gly Asp Val Arg Ala Glu Tyr
865 870 875 880
Ser Ser Arg Phe His Ala Lys Thr Gly Thr Pro Gly Ile Arg Cys Lys
885 890 895
Met Val Ser Gly His Asp Leu Gln Gly Arg Arg Phe Glu Asn Leu Gln
900 905 910
Lys Arg Leu Ile Ser Glu Gln Phe Leu Thr Glu Glu Gln Val Lys Gln
915 920 925
Leu Arg Pro Gly Asp Ile Val Pro Asp Asp Ser Gly Glu Trp Phe Met
930 935 940
Thr Leu Ser Asp Gly Ser Glu Gly Lys Glu Val Val Phe Leu Gln Ala
945 950 955 960

Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln Arg Tyr
965 970 975
Asn Glu Leu Phe Lys Val Ser Cys Arg Val Leu Ile Arg Gly Glu Glu
980 985 990
Glu Tyr Leu Ile Pro Lys Thr Lys Ser Val Gln Ala Lys Leu Gly Lys
995 1000 1005
Gly Leu Phe Val Lys Lys Thr Asp Thr Val Met Lys Asp Val Tyr
1010 1015 1020
Val Trp Asp Ser Gln Ala Lys Leu Lys Gly Lys Thr Thr Phe Thr
1025 1030 1035
Glu Glu Ser Glu Ser Pro Glu Gln Leu Glu Asp Phe Gln Glu Ile
1040 1045 1050
Ile Glu Glu Ala Glu Glu Ala Lys Gly Thr Tyr Arg Thr Leu Phe
1055 1060 1065
Arg Asp Pro Ser Gly Val Phe Phe Pro Glu Phe Val Trp Ser Thr
1070 1075 1080
Gln Lys Asp Phe Trp Ser Glu Val Lys Arg Arg Leu Tyr Gly Lys
1085 1090 1095
Leu Arg Glu Arg Phe Leu Met Lys Thr Arg
1100 1105

<210> 98
<211> 1429
<212> PRT
<213> 金黄杆菌属物种
<223> APG00771_RGN
<400> 98

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

100 105 110

Glu Pro Lys Leu Ala Trp Lys Asn Ile Glu Gly Lys Phe Ser Phe Leu
115 120 125

Phe Gln Asn Ser Phe Asn Glu Met Leu Glu Asp Phe Lys Ala Asn Glu
130 135 140

Gln Asp Leu Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Phe Leu Arg Lys
145 150 155 160

Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile Leu
165 170 175

Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu
180 185 190

Glu Glu Glu Asn Pro Asn Lys Leu Val Glu Phe Tyr Ser Leu Lys Val
195 200 205

Val Asp Val Leu Ala Asp Glu Pro Gln Lys Gly Lys Ser Glu Ile Trp
210 215 220

Tyr Ser Leu Val Leu Lys Asn Gly Trp Ile Tyr Arg Arg Ala Ser Lys
225 230 235 240

Ile Pro Leu Phe Glu Trp Lys Asp Lys Thr Arg Asp Phe Ile Val Thr
245 250 255

Thr Asp Leu Asn Asp Asp Gly Ser Ile Lys Lys Asp Lys Asp Gly Asn
260 265 270

Glu Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn Asp Trp Thr Leu Val
275 280 285

Lys Lys Lys Thr Glu Gln Glu Ile Glu Gln Ser His Lys Thr Val Gly
290 295 300

Thr Tyr Ile Tyr Glu Thr Leu Leu Leu Asn Pro Lys Gln Lys Ile Lys
305 310 315 320

Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Asp Glu Leu
325 330 335

Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Gln Asn
340 345 350

Asp Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn Glu
355 360 365

Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Leu Met
370 375 380

Asp Asp Leu Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser
385 390 395 400

Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn Gly
405 410 415

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

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

Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile Ala
1040 1045 1050
Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser Ala
1055 1060 1065
Lys Ser Asp Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys
1070 1075 1080
Phe Glu Lys Val Val Tyr His His Thr Gln Thr Gly Glu Lys Ile
1085 1090 1095
Glu Arg Glu Ile Pro Arg Gln Phe Leu Lys Pro Trp Glu Asn Phe
1100 1105 1110
Thr Val Asp Ala Lys His Asn Leu Glu Ser Ile Ile Val Ser Phe
1115 1120 1125
Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr Glu
1130 1135 1140
Lys Tyr Val Glu Lys Asp Gly Thr Lys Asn Lys Glu Arg Val Glu
1145 1150 1155
Gln Thr Gly Thr Asn Trp Ala Ile Arg Lys Pro Met His Lys Asp
1160 1165 1170
Thr Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro Lys
1175 1180 1185
Gly Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Thr Ser Phe
1190 1195 1200
Asp Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln Lys
1205 1210 1215
Ile Leu Lys Asn Tyr Leu Ala Phe Lys Asp Gly Asn Pro Glu Met
1220 1225 1230
Ala Phe Ser Pro Glu Gly Ile Asp Asp Leu Asn Lys Asn Ile Glu
1235 1240 1245
Lys Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val Arg
1250 1255 1260
Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Thr Gly Asn
1265 1270 1275
Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe
1280 1285 1290
Phe Ala Val Tyr Glu Asp Glu Lys Gly Lys Arg Ser Tyr Glu Thr
1295 1300 1305
Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Ser
1310 1315 1320
Val Val Asp Leu Lys Asp Thr Asn Asp Phe Tyr Leu Cys Pro Asn

1325 1330 1335
Asp Leu Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Val Asn
1340 1345 1350
Asn Val Asp Phe Lys Asp Ile Lys Lys Asp Lys Asn Glu Arg Ile
1355 1360 1365
Tyr Lys Val Val Ser Phe Ser Gly Ser Gln Ile Phe Phe Val Arg
1370 1375 1380
Gln Asp Ile Ala Thr Phe Ile Val Asn Lys Ala Glu Leu Ser Thr
1385 1390 1395
Leu Asn Lys Met Glu Arg Ala Ile Asp Gly Ser Met Ile Lys Glu
1400 1405 1410
Asn Cys Ile Lys Leu Asn Ile Asp Arg Leu Gly Asn Ile Ser Lys
1415 1420 1425
Ala

<210> 99
<211> 1112
<212> PRT
<213> 芽孢杆菌属物种
<223> APG02789_RGN
<400> 99
Met Ala Ile Arg Ser Ile Lys Leu Lys Met Lys Thr Asn Ser Gly Thr
1 5 10 15
Asp Ser Ile Tyr Leu Arg Lys Ala Leu Trp Arg Thr His Gln Leu Ile
20 25 30
Asn Glu Gly Ile Ala Tyr Tyr Met Asn Leu Leu Thr Leu Tyr Arg Gln
35 40 45
Glu Ala Ile Gly Asp Lys Thr Lys Glu Ala Tyr Gln Ala Glu Leu Ile
50 55 60
Asn Ile Ile Arg Asn Gln Gln Arg Asn Asn Gly Ser Ser Glu Glu His
65 70 75 80
Gly Ser Asp Gln Glu Ile Leu Ala Leu Leu Arg Gln Leu Tyr Glu Leu
85 90 95
Ile Ile Pro Ser Ser Ile Gly Glu Ser Gly Asp Ala Asn Gln Leu Gly
100 105 110
Asn Lys Phe Leu Tyr Pro Leu Val Asp Pro Asn Ser Gln Ser Gly Lys
115 120 125
Gly Thr Ser Asn Ala Gly Arg Lys Pro Arg Trp Lys Arg Leu Lys Glu
130 135 140
Glu Gly Asn Pro Asp Trp Glu Leu Glu Lys Lys Lys Asp Glu Glu Arg

145 150 155 160
Lys Ala Lys Asp Pro Thr Val Lys Ile Phe Asp Asn Leu Asn Lys Tyr
165 170 175
Gly Leu Leu Pro Leu Phe Pro Leu Phe Thr Asn Ile Gln Lys Asp Ile
180 185 190
Glu Trp Leu Pro Leu Gly Lys Arg Gln Ser Val Arg Lys Trp Asp Lys
195 200 205
Asp Met Phe Ile Gln Ala Ile Glu Arg Leu Leu Ser Trp Glu Ser Trp
210 215 220
Asn Arg Arg Val Ala Asp Glu Tyr Lys Gln Leu Lys Glu Lys Thr Glu
225 230 235 240
Ser Tyr Tyr Lys Glu His Leu Thr Gly Gly Glu Glu Trp Ile Glu Lys
245 250 255
Ile Arg Lys Phe Glu Lys Glu Arg Asn Met Glu Leu Glu Lys Asn Ala
260 265 270
Phe Ala Pro Asn Asp Gly Tyr Phe Ile Thr Ser Arg Gln Ile Arg Gly
275 280 285
Trp Asp Arg Val Tyr Glu Lys Trp Ser Lys Leu Pro Glu Ser Ala Ser
290 295 300
Pro Glu Glu Leu Trp Lys Val Val Ala Glu Gln Gln Asn Lys Met Ser
305 310 315 320
Glu Gly Phe Gly Asp Pro Lys Val Phe Ser Phe Leu Ala Asn Arg Glu
325 330 335
Asn Arg Asp Ile Trp Arg Gly His Ser Glu Arg Ile Tyr His Ile Ala
340 345 350
Ala Tyr Asn Gly Leu Gln Lys Lys Leu Ser Arg Thr Lys Glu Gln Ala
355 360 365
Thr Phe Thr Leu Pro Asp Ala Ile Glu His Pro Leu Trp Ile Arg Tyr
370 375 380
Glu Ser Pro Gly Gly Thr Asn Leu Asn Leu Phe Lys Leu Glu Glu Lys
385 390 395 400
Gln Lys Lys Asn Tyr Tyr Val Thr Leu Ser Lys Ile Ile Trp Pro Ser
405 410 415
Glu Glu Lys Trp Ile Glu Lys Glu Asn Ile Glu Ile Pro Leu Ala Pro
420 425 430
Ser Ile Gln Phe Asn Arg Gln Ile Lys Leu Lys Gln His Val Lys Gly
435 440 445
Lys Gln Glu Ile Ser Phe Ser Asp Tyr Ser Ser Arg Ile Ser Leu Asp
450 455 460

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

770 775 780

Asn Val Lys Asp Asp Arg Leu Lys Gln Met Ala Asn Leu Ile Ile Met
785 790 795 800

Thr Ala Leu Gly Phe Lys Tyr Asp Lys Glu Glu Lys Asp Arg Tyr Lys
805 810 815

Arg Trp Lys Glu Thr Tyr Pro Ala Cys Gln Ile Ile Leu Phe Glu Asn
820 825 830

Leu Asn Arg Tyr Leu Phe Asn Leu Asp Arg Ser Arg Arg Glu Asn Ser
835 840 845

Arg Leu Met Lys Trp Ala His Arg Ser Ile Pro Arg Thr Val Ser Met
850 855 860

Gln Gly Glu Met Phe Gly Leu Gln Val Gly Asp Val Arg Ser Glu Tyr
865 870 875 880

Ser Ser Arg Phe His Ala Lys Thr Gly Ala Pro Gly Ile Arg Cys His
885 890 895

Ala Leu Thr Glu Glu Asp Leu Lys Ala Gly Ser Asn Thr Leu Lys Arg
900 905 910

Leu Ile Glu Asp Gly Phe Ile Asn Glu Ser Glu Leu Ala Tyr Leu Lys
915 920 925

Lys Gly Asp Ile Ile Pro Ser Gln Gly Gly Glu Leu Phe Val Thr Leu
930 935 940

Ser Lys Arg Tyr Lys Lys Asp Ser Asp Asn Asn Glu Leu Thr Val Ile
945 950 955 960

His Ala Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln
965 970 975

Gln Asn Ser Glu Val Tyr Arg Val Pro Cys Gln Leu Ala Arg Met Gly
980 985 990

Glu Asp Lys Leu Tyr Ile Pro Lys Ser Gln Thr Glu Thr Ile Lys Lys
995 1000 1005

Tyr Phe Gly Lys Gly Ser Phe Val Lys Asn Asn Thr Glu Gln Glu
1010 1015 1020

Val Tyr Lys Trp Glu Lys Ser Glu Lys Met Lys Ile Lys Thr Asp
1025 1030 1035

Thr Thr Phe Asp Leu Gln Asp Leu Asp Gly Phe Glu Asp Ile Ser
1040 1045 1050

Lys Thr Ile Glu Leu Ala Gln Glu Gln Gln Lys Lys Tyr Leu Thr
1055 1060 1065

Met Phe Arg Asp Pro Ser Gly Tyr Phe Phe Asn Asn Glu Thr Trp
1070 1075 1080

Arg Pro Gln Lys Glu Tyr Trp Ser Ile Val Asn Asn Ile Ile Lys
1085 1090 1095
Ser Cys Leu Lys Lys Lys Ile Leu Ser Asn Lys Val Glu Leu
1100 1105 1110

<210> 100
<211> 1070
<212> PRT
<213> 芽孢杆菌属物种
<223> APG05083.1_RGN
<400> 100

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

225 230 235 240
Pro Phe Ala Ser Gly Asn Ala Ile Val Asn Lys Val Gly Tyr Cys Ser
245 250 255
Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
260 265 270
Gln Tyr Phe Ser Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285
Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Val Ile Leu Asp Glu
290 295 300
Met Phe Asn Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320
Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335
Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350
Leu Lys Ser Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
355 360 365
Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Ala Phe Ser Thr Leu Asp
370 375 380
Tyr Asp Thr Phe Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400
Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Cys Cys Tyr
405 410 415
Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430
Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445
Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Ala
450 455 460
Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Leu Phe Pro
465 470 475 480
Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495
Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510
His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Thr
515 520 525
Lys Ile Met Lys Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560
Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575
Ser Leu Lys Gln Ile Pro Ala Asn Thr Phe Phe Asn Glu Met Lys Lys
580 585 590
Glu Arg Ser Gly Ser Pro Val Leu Glu Ile Asp His Ile Leu Pro Tyr
595 600 605
Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620
Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640
Glu Gly Asn Lys Asp Trp Glu Ser Phe Glu Ser Tyr Val Arg Leu Asn
645 650 655
Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670
Leu Pro Arg Glu Ser Asn Met Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685
Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700
Phe Lys Glu Val Glu Gly Ser Thr Arg Lys Lys His Val Gln Thr Val
705 710 715 720
Asn Gly Ile Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735
Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750
Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765
Ile Lys Glu Ser Asn Lys Ser Ile Arg Lys Pro Tyr Phe Pro Leu Pro
770 775 780
Trp Val Gly Phe Arg Glu Glu Ile Leu Ser His Leu Ala Arg Gln Pro
785 790 795 800
Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Phe
805 810 815
Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
820 825 830
Ala His Glu Gln Thr Ile Met Lys Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Lys Arg Val Tyr Leu Lys Asp Ile Lys Phe Asp

850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Ile Glu Tyr Gly Lys Glu Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910
Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925
Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Ile Pro Ile Tyr Val
945 950 955 960
Met Asp Thr Val His Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Thr Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asn Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Asn
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Val
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Phe Val Arg Gln Glu Ile Arg Lys
1055 1060 1065
Asn Phe
1070

<210> 101
<211> 1071
<212> PRT
<213> 芽孢杆菌属物种
<223> APG07433.1_RGN
<400> 101
Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15
Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Arg
20 25 30

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

340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val
355 360 365

Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln
435 440 445

Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr
450 455 460

Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Asn His Asp Glu Arg Thr
515 520 525

Lys Ile Val Ser Ala Gln Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ser Glu His Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Glu Ile Pro Ala Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
610 615 620

Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
645 650 655

Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
660 665 670
Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685
Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
690 695 700
Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
705 710 715 720
Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735
Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750
Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
755 760 765
Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
770 775 780
Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
785 790 795 800
Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
805 810 815
Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
820 825 830
Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910
Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val
915 920 925
Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser

965 970 975
Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
1055 1060 1065
Asn Phe His
1070

<210> 102
<211> 1071
<212> PRT
<213> 芽孢杆菌属物种
<223> APG07513.1_RGN
<400> 102
Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15
Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30
Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45
Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60
Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80
Asp Ile Arg Asn Leu Leu Val Gln His Glu Ile Ile Ser Gln Lys Glu
85 90 95
Leu Thr Ser Leu Tyr Pro Leu Ser Lys Ser Ser Met Asp Ile Trp Asp
100 105 110
Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asp Arg Phe Glu Trp Ala
115 120 125
Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

Lys Ser Glu Leu Lys Asp Val Glu Thr Gly Lys Val Leu Ser Ser Ile
145 150 155 160
Gln Val Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175
Trp Met Lys Asn Ala Asp Cys Ser Lys Tyr Gly Lys Arg Arg Asn Ser
180 185 190
Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
195 200 205
Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe His Ser Ser Tyr Ala
210 215 220
Ser Val Asp Leu Gln Lys Thr Tyr Ile Gln Ile Trp Ala His Gln Leu
225 230 235 240
Pro Phe Ala Ser Gly Asn Ala Ile Val Asn Lys Val Gly Tyr Cys Ser
245 250 255
Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
260 265 270
Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285
Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
290 295 300
Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320
Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335
Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350
Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
355 360 365
Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
370 375 380
Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400
Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415
Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430
Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445
Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr

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

Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro
770 775 780
Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
785 790 795 800
Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Pro
805 810 815
Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser
820 825 830
Ala His Asp Gln Thr Val Met Lys Lys Gly Asp Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Glu Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910
Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925
Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
1055 1060 1065
Ile Phe Lys

1070

<210> 103

<211> 1072

<212> PRT

<213> 芽孢杆菌属物种

<223> APG08290.1_RGN

<400> 103

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

Leu Leu Lys Gly Lys Glu Arg Arg Ile Pro Lys Ala Thr Tyr Thr Phe
260 265 270
Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
275 280 285
Asp Phe Gln Pro Phe Thr Gln Glu Gln Lys Glu Ile Ile Leu Asp Lys
290 295 300
Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320
Ser Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
325 330 335
Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350
Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Lys Lys Val
355 360 365
Val Ala Asn Tyr Ala Glu Arg Thr Asn Glu Ala Phe Ser Thr Leu Asp
370 375 380
Tyr Asp Ala Ile Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400
Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415
Asp Asp Gln Leu Ile Glu Glu Leu Phe Thr Leu Ser Tyr Thr Lys Phe
420 425 430
Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445
Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
450 455 460
Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Leu Ile Pro
465 470 475 480
Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Ile Thr Gln Ala Arg
485 490 495
Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510
His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525
Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540
Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560
Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr

565 570 575
Ser Leu Lys Glu Ile Pro Pro Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590
Glu Arg Asn Gly Ser Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605
Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
610 615 620
Asp Glu Asn Arg Asn Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640
Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
645 650 655
Lys Leu Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Lys Thr Tyr
660 665 670
Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685
Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
690 695 700
Phe Lys Glu Val Glu Val Asn Leu Arg Lys Lys Arg Val Gln Thr Val
705 710 715 720
Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735
Asn Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750
Ala Cys Thr Asp His His Met Val Thr Arg Ile Thr Glu Tyr Tyr Gln
755 760 765
Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
770 775 780
Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
785 790 795 800
Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Ser
805 810 815
Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
820 825 830
Ala His Asp Gln Thr Ile Arg Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Lys Arg Val Arg Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Arg Lys Asn Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910
Asn Leu Ile Lys Arg Val Lys Ile Glu Gly Gln Thr Lys Ala Phe Val
915 920 925
Arg Glu Val Asn Gly Gly Val Ala Gln Asn Ser Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Lys Ser Gly
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990
Ser Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asn Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Asn Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Val
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Lys Gln Glu Thr Arg Arg
1055 1060 1065
Ile Phe Asn Arg
1070

<210> 104
<211> 1337
<212> PRT
<213> 肠球菌属物种
<223> APG05459.1_RGN
<400> 104

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

50 55 60

Lys Arg Thr Ala Arg Arg Arg Ile Ser Arg Arg Arg Asn Arg Leu Arg
65 70 75 80

Tyr Leu Gln Ala Phe Phe Glu Glu Ala Met Thr Ala Leu Asp Glu Asn
85 90 95

Phe Phe Ala Arg Leu Gln Glu Ser Phe Leu Val Pro Glu Asp Lys Lys
100 105 110

Trp His Arg His Pro Ile Phe Ala Lys Leu Glu Asp Glu Val Ala Tyr
115 120 125

His Glu Thr Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Ala Asp
130 135 140

Ser Ser Glu Gln Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Ile Val Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Lys Leu Ser Thr
165 170 175

Glu Asn Ile Ser Val Lys Glu Gln Phe Gln Gln Phe Met Ile Ile Tyr
180 185 190

Asn Gln Thr Phe Val Asn Gly Glu Ser Arg Leu Val Ser Ala Pro Leu
195 200 205

Pro Glu Ser Val Leu Ile Glu Glu Glu Leu Thr Glu Lys Ala Ser Arg
210 215 220

Thr Lys Lys Ser Glu Lys Val Leu Gln Gln Phe Pro Gln Glu Lys Ala
225 230 235 240

Asn Gly Leu Phe Gly Gln Phe Leu Lys Leu Met Val Gly Asn Lys Ala
245 250 255

Asp Phe Lys Lys Val Phe Gly Leu Glu Glu Glu Ala Lys Ile Thr Tyr
260 265 270

Ala Ser Glu Ser Tyr Glu Glu Asp Leu Glu Gly Ile Leu Ala Lys Val
275 280 285

Gly Asp Glu Tyr Ser Asp Val Phe Leu Ala Ala Lys Asn Val Tyr Asp
290 295 300

Ala Val Glu Leu Ser Thr Ile Leu Ala Asp Ser Asp Lys Lys Ser His
305 310 315 320

Ala Lys Leu Ser Ser Ser Met Ile Val Arg Phe Thr Glu His Gln Glu
325 330 335

Asp Leu Lys Lys Phe Lys Arg Phe Ile Arg Glu Asn Cys Pro Asp Glu
340 345 350

Tyr Asp Asn Leu Phe Lys Asn Glu Gln Lys Asp Gly Tyr Ala Gly Tyr
355 360 365

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

675 680 685
Asp Tyr Leu Ile Lys Asp Asp Gly Val Ser Lys His Tyr Asn Arg Asn
690 695 700
Phe Met Gln Leu Ile Asn Asp Ser Gln Leu Ser Phe Lys Asn Ala Ile
705 710 715 720
Gln Lys Ala Gln Ser Ser Glu His Glu Glu Thr Leu Ser Glu Thr Val
725 730 735
Asn Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Tyr Gln Ser
740 745 750
Leu Lys Ile Val Asp Glu Leu Val Ala Ile Met Gly Tyr Ala Pro Lys
755 760 765
Arg Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Ser Thr Gly
770 775 780
Lys Arg Arg Ser Ile Gln Arg Leu Lys Ile Val Glu Lys Ala Met Ala
785 790 795 800
Glu Ile Gly Ser Asn Leu Leu Lys Glu Gln Pro Thr Thr Asn Glu Gln
805 810 815
Leu Arg Asp Thr Arg Leu Phe Leu Tyr Tyr Met Gln Asn Gly Lys Asp
820 825 830
Met Tyr Thr Gly Asp Glu Leu Ser Leu His Arg Leu Ser His Tyr Asp
835 840 845
Ile Asp His Ile Ile Pro Gln Ser Phe Met Lys Asp Asp Ser Leu Asp
850 855 860
Asn Leu Val Leu Val Gly Ser Thr Glu Asn Arg Gly Lys Ser Asp Asp
865 870 875 880
Val Pro Ser Lys Glu Val Val Lys Asp Met Lys Ala Tyr Trp Glu Lys
885 890 895
Leu Tyr Ala Ala Gly Leu Ile Ser Gln Arg Lys Phe Gln Arg Leu Thr
900 905 910
Lys Gly Glu Gln Gly Gly Leu Thr Leu Glu Asp Lys Ala His Phe Ile
915 920 925
Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys Asn Val Ala Gly
930 935 940
Ile Leu Asp Gln Arg Tyr Asn Ala Asn Ser Lys Glu Lys Lys Val Gln
945 950 955 960
Ile Ile Thr Leu Lys Ala Ser Leu Thr Ser Gln Phe Arg Ser Ile Phe
965 970 975
Gly Leu Tyr Lys Val Arg Glu Val Asn Asp Tyr His His Gly Gln Asp
980 985 990

Ala Tyr Leu Asn Cys Val Val Ala Thr Thr Leu Leu Lys Val Tyr Pro
995 1000 1005
Asn Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Pro Lys Phe Gln
1010 1015 1020
Ala Phe Lys Glu Asn Lys Ala Thr Ala Lys Thr Ile Ile Tyr Thr
1025 1030 1035
Asn Leu Met Arg Phe Phe Thr Glu Asp Glu Pro Arg Phe Met Lys
1040 1045 1050
Asp Gly Glu Ile Leu Trp Ser Asn Ser Tyr Leu Lys Asn Ile Lys
1055 1060 1065
Lys Glu Leu Asn Tyr His Gln Met Asn Ile Val Lys Lys Val Glu
1070 1075 1080
Val Gln Lys Gly Gly Phe Ser Lys Glu Ser Ile Lys Pro Lys Gly
1085 1090 1095
Pro Ser Asn Lys Leu Ile Pro Val Lys Asn Gly Leu Asp Pro Gln
1100 1105 1110
Lys Tyr Gly Gly Phe Asp Ser Pro Val Val Ala Tyr Thr Val Leu
1115 1120 1125
Phe Thr His Glu Lys Gly Lys Lys Pro Leu Ile Lys Gln Glu Ile
1130 1135 1140
Leu Gly Ile Thr Ile Met Glu Lys Thr Arg Phe Glu Gln Asn Pro
1145 1150 1155
Ile Leu Phe Leu Glu Glu Lys Gly Phe Leu Arg Pro Arg Val Leu
1160 1165 1170
Met Lys Leu Pro Lys Tyr Thr Leu Tyr Glu Phe Pro Glu Gly Arg
1175 1180 1185
Arg Arg Leu Leu Ala Ser Ala Lys Glu Ala Gln Lys Gly Asn Gln
1190 1195 1200
Met Val Leu Pro Glu His Leu Leu Thr Leu Leu Tyr His Ala Lys
1205 1210 1215
Gln Cys Leu Leu Pro Asn Gln Ser Glu Ser Leu Ala Tyr Val Glu
1220 1225 1230
Gln His Gln Pro Glu Phe Gln Glu Ile Leu Glu Arg Val Val Asp
1235 1240 1245
Phe Ala Glu Val His Thr Leu Ala Lys Ser Lys Val Gln Gln Ile
1250 1255 1260
Val Lys Leu Phe Glu Ala Asn Gln Thr Ala Asp Val Lys Glu Ile
1265 1270 1275
Ala Ala Ser Phe Ile Gln Leu Met Gln Phe Asn Ala Met Gly Ala

1280 1285 1290
Pro Ser Thr Phe Lys Phe Phe Gln Lys Asp Ile Glu Arg Ala Arg
1295 1300 1305
Tyr Thr Ser Ile Lys Glu Ile Phe Asp Ala Thr Ile Ile Tyr Gln
1310 1315 1320
Ser Thr Thr Gly Leu Tyr Glu Thr Arg Arg Lys Val Val Asp
1325 1330 1335

<210> 105
<211> 1436
<212> PRT
<213> 肠球菌属物种
<223> APG04583.1_RGN
<400> 105

Met Ala Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly
1 5 10 15
Trp Ala Leu Val Gln Gln Asp Phe Glu Asn Lys Glu Gly Asn Ile Leu
20 25 30
Gly Met Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly Glu
35 40 45
Phe Gly Lys Gly Asn Ser Ile Ser Gln Thr Ala Glu Arg Thr Gly Tyr
50 55 60
Arg Gly Val Arg Arg Leu Arg Glu Arg His Leu Leu Arg Arg Glu Arg
65 70 75 80
Leu His Arg Val Leu His Leu Leu Gly Phe Leu Pro Lys His Tyr Asp
85 90 95
Glu Lys Ile Asp Phe Thr Gln Arg Phe Gly Lys Phe Ile Asn Gln Ala
100 105 110
Glu Pro Lys Leu Ala Phe Asp Ser Glu Phe Leu Phe Lys Asp Ser Phe
115 120 125
His Glu Met Leu Ala Asp Phe Lys Gln Asn Gln Pro Glu Phe Leu Lys
130 135 140
Asp Lys Asn Gly Glu Asp Cys Leu Val Pro Tyr Asp Trp Thr Ile Tyr
145 150 155 160
Tyr Leu Arg Lys Lys Ala Leu Thr Gln Lys Ile Glu Lys Tyr Glu Leu
165 170 175
Ala Trp Leu Ile Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu
180 185 190
Arg Gly Glu Glu Glu Lys Glu Asn Pro Asn Lys Leu Val Gly Phe His
195 200 205

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

515 520 525

Leu Ile Gln Ser Arg Leu Asp Lys Val Glu Tyr Ala Pro Lys Asp Phe

530 535 540

Leu Thr Tyr Glu Asn Glu Met Ala Leu Trp His Ile Ile Tyr Ser Ile

545 550 555 560

Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Asn Lys

565 570 575

Lys Gly Leu Asp Glu Val Thr Phe Val Glu Ala Phe Lys Lys Phe Pro

580 585 590

Pro Phe Lys Ser Asp Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys

595 600 605

Leu Leu Pro Leu Met Arg Phe Gly Thr Gln Trp Asn Trp Asp Asn Ile

610 615 620

Asp Gln Asn Ser Lys Glu Arg Ile Gly Lys Ile Leu Thr Gly Glu Tyr

625 630 635 640

Asp Glu Asn Ile Lys Gly Arg Val Arg Glu Lys Ala Arg His Leu Asn

645 650 655

Ser Glu Thr Asp Phe Gln Ala Leu Pro Leu Trp Leu Ala Gln Tyr Val

660 665 670

Val Tyr Gly Arg His Ser Glu Ala Asp Ile Ala Gly Lys Trp Asn Ser

675 680 685

Val Asp Asp Leu Lys Gln Phe Leu Asp Asp Phe Lys Gln His Ser Leu

690 695 700

Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Ala Val

705 710 715 720

Lys Asp Ile Trp Asn Phe Tyr Gly Lys Gly Ala Lys Asp Phe Phe Ser

725 730 735

Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Asn Thr Ala Asp Glu

740 745 750

Arg Lys Arg Ile Thr Thr Met Val Thr Asp Asn Glu Asn Thr Asn Leu

755 760 765

Arg Ile Lys Ala Leu Leu Ala Glu Met Ala Leu Asp Gln Asn Val Asp

770 775 780

Asn Val Arg Pro Tyr Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr

785 790 795 800

Glu Glu Gly Val Leu Asn Ala Glu Glu Asn Ile Asp Asp Asp Ile Leu

805 810 815

Lys Ile Ser Lys Thr Ala Gln Pro Ser Ala Thr Asp Leu Lys Arg Tyr

820 825 830

Lys Leu Trp Leu Glu Gln Lys Tyr Arg Ser Pro Tyr Thr Gly Gln Met
835 840 845
Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile
850 855 860
Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Met Ser Asn Lys Val Ile
865 870 875 880
Cys Glu Ala Ala Val Asn Lys Leu Lys Asp Asn Gln Ile Gly Leu Val
885 890 895
Phe Ile Lys Asn His His Gly Glu Val Val Asp Phe Gly Met Gly Lys
900 905 910
Gln Val Lys Ile Leu Glu Val Ser Asp Tyr Glu Asp Phe Val Lys Gln
915 920 925
Asn Tyr Asn Lys Asn Arg Gly Lys Arg Asn Lys Leu Leu Leu Glu Asp
930 935 940
Ile Pro Glu Lys Met Ile Glu Arg Gln Leu Asn Asp Thr Arg Tyr Ile
945 950 955 960
Ser Lys Tyr Ile Thr Gln Val Leu Ser Asn Ile Val Arg Asp Asp Lys
965 970 975
Glu Gly Ser Lys Asp Asp Gly Val Asn Ser Lys Asn Ile Val Pro Gly
980 985 990
Asn Gly Lys Ile Thr Thr Arg Leu Lys Gln Asp Trp Gly Leu Asn Asp
995 1000 1005
Val Trp Asn Asp Leu Val Leu Pro Arg Phe Glu Arg Met Asn Thr
1010 1015 1020
Leu Thr Asn Ser Asn Asp Phe Thr Ser Lys Asn Thr His Gly Lys
1025 1030 1035
Thr Ile Pro Thr Val Pro Ile Glu Leu Ser Lys Gly Phe Ser Lys
1040 1045 1050
Lys Arg Ile Asp His Arg His His Ala Met Asp Ala Leu Val Ile
1055 1060 1065
Ala Cys Ala Thr Arg Asp His Val Asn Leu Leu Asn Asn Glu Ser
1070 1075 1080
Ser Lys Ser Asp Thr Lys Arg Tyr Asp Leu Asn Arg Lys Leu Arg
1085 1090 1095
Lys Tyr Glu Lys Val Ala Tyr Asn Asp Pro Lys Thr Gly Glu Arg
1100 1105 1110
Ile Glu Lys Glu Val Pro Lys Asp Phe Ile Lys Pro Trp Glu Thr
1115 1120 1125
Phe Thr Glu Asp Thr Arg Thr Leu Leu Glu Asn Ile Val Ile Ser

1130 1135 1140

Phe Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr

1145 1150 1155

Glu Lys Ile Glu Asn Gly Lys Lys Val Lys Val Glu Gln Lys Gly

1160 1165 1170

Ile Asn Trp Ala Val Arg Lys Ala Leu His Lys Glu Thr Val Ser

1175 1180 1185

Gly Gln Val His Leu Asp Arg Ile Lys Val Ala Lys Gly Lys Ile

1190 1195 1200

Leu Thr Ala Thr Arg Lys Thr Leu Asp Ala Ser Phe Asn Glu Lys

1205 1210 1215

Thr Ile Glu Ser Ile Thr Asp Thr Gly Ile Gln Lys Ile Leu Leu

1220 1225 1230

Asn Tyr Leu Lys Ser Lys Asp Asn Asn Pro Glu Val Ala Phe Ser

1235 1240 1245

Pro Glu Gly Ile Glu Glu Leu Asn Lys Asn Ile Arg Leu Tyr Asn

1250 1255 1260

Asp Gly Lys Ala His Gln Pro Ile Leu Lys Val Arg Val Phe Glu

1265 1270 1275

Gln Gly Ser Lys Phe Thr Leu Gly Glu Thr Gly Asn Lys Thr Thr

1280 1285 1290

Lys Phe Val Glu Ala Ala Lys Gly Thr Asn Leu Phe Phe Gly Ile

1295 1300 1305

Tyr Glu Asp Lys Ser Gly Lys Arg Ser Tyr Glu Thr Ile Pro Leu

1310 1315 1320

Asn Ile Val Ile Glu Arg Gln Lys Gln Gly Leu Gln Ala Val Pro

1325 1330 1335

Glu Thr Asn Glu Lys Gly Lys Gln Leu Leu Phe Thr Leu Ser Pro

1340 1345 1350

Asn Asp Leu Val Tyr Val Pro Glu Glu Gly Val Phe Asp Glu Asn

1355 1360 1365

Asn Ile Lys Val Asp Arg Ile Tyr Lys Val Val Ser Phe Ser Thr

1370 1375 1380

Tyr Gln Cys Phe Phe Val Arg Asn Asp Val Ser Thr Ser Val Val

1385 1390 1395

Asn Lys Val Glu Tyr Ser Ala Leu Asn Lys Met Glu Lys Ser Ile

1400 1405 1410

Asp Asn Ile Met Ile Lys Glu Asn Cys Val Lys Leu Asn Val Asp

1415 1420 1425

Arg Leu Gly Lys Ile Ser Lys Ala
1430 1435

<210> 106
<211> 1456
<212> PRT
<213> 肠球菌属物种
<223> APG01688.1_RGN
<400> 106

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

245 250 255
Thr Thr Asp Leu Asn Asp Asp Arg Ser Val Lys Thr Asp Lys Glu Gly
260 265 270
Asn Glu Lys Arg Ser Phe Arg Ala Pro Ser Glu Asn Asp Trp Thr Leu
275 280 285
Val Lys Lys Lys Thr Glu Gln Glu Ile Asp Gln Ser His Lys Thr Val
290 295 300
Gly Thr Tyr Ile Tyr Glu Thr Leu Leu Leu Asn Pro Lys Gln Lys Ile
305 310 315 320
Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Asp Glu
325 330 335
Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Lys
340 345 350
Asn Asp Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn
355 360 365
Glu Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Leu
370 375 380
Met Asp Asp Leu Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser
385 390 395 400
Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn
405 410 415
Gly Ile Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro
420 425 430
Tyr Tyr Gln Glu Phe Arg Leu Trp Gln Trp Met Tyr Asn Leu Asn Ile
435 440 445
Tyr Arg Lys Asp Asp Glu Ala Asn Val Thr Lys Glu Phe Leu Asn Thr
450 455 460
Asn Lys Asp Phe Glu Ser Leu Phe Glu Phe Leu Asn Asn Arg Lys Glu
465 470 475 480
Ile Glu Gln Lys Pro Leu Ile Lys Phe Leu Leu Glu Gln Lys Asp Ile
485 490 495
Asn Lys Lys Leu Leu Asn Ala Glu Ala Glu Lys Tyr Arg Trp Asn Tyr
500 505 510
Val Glu Asp Lys Lys Tyr Pro Cys Asn Glu Thr Lys Thr Met Ile Ser
515 520 525
Ser Arg Leu Asp Lys Val Glu Asn Ile Ser Asp Asp Phe Leu Thr Arg
530 535 540
Asp Ile Glu Gln Lys Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys
545 550 555 560

Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Thr Arg Asn Asp Leu
565 570 575
Asp Glu Asn Ser Phe Ile Glu Ala Phe Lys Lys Phe Ser Pro Phe Lys
580 585 590
Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro
595 600 605
Leu Met Arg Leu Gly Lys Tyr Trp Tyr Glu Asp Glu Ile Val Lys His
610 615 620
Ser Asp Ile Tyr Phe Lys Asn Ile Glu Asn Leu Leu Gly Asp Phe Ser
625 630 635 640
Asn Arg Asp Lys Lys Ile Ser Glu Glu Asp Lys Glu Lys Trp Asn Lys
645 650 655
Ser Ile Asn Leu Lys Leu Gln Glu Glu Leu Lys Asp Phe Gln Thr Ala
660 665 670
Glu Ile Asp Leu Phe Gln Gly Leu Arg Leu His Ile Ala Gln Tyr Leu
675 680 685
Val Tyr Gly Arg His Ser Glu Ala Ser Met Ile Gly Lys Trp Asn Ser
690 695 700
Ala Glu Asp Leu Glu Glu Phe Leu Lys Asp Phe Lys Gln His Ser Leu
705 710 715 720
Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val
725 730 735
Lys Asp Ile Trp Leu Lys Tyr Gly Asn Gly Ala Lys Asp Phe Phe Asn
740 745 750
Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp Asp
755 760 765
Arg Lys Lys Leu Thr Asn Gln Ile Ser Glu Asn Glu Asn Thr Asn Phe
770 775 780
Arg Ile Lys Ala Leu Leu Ala Glu Met Met Asn Asp Ser Ser Val Glu
785 790 795 800
Asn Val Arg Pro Phe Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr
805 810 815
Glu Asp Asp Val Leu Lys Ser Asp Ile Glu Ile Glu Asp Asp Ile Leu
820 825 830
Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg Tyr
835 840 845
Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile
850 855 860
Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile

865 870 875 880
Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Val Ile
885 890 895
Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Glu
900 905 910
Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile Glu Leu Gly Phe Gly Lys
915 920 925
Ser Ile Lys Val Phe Glu Thr Lys Glu Tyr Glu Asp Phe Val Lys Lys
930 935 940
His Tyr Ala Asn Asn Gln Gly Lys Arg Asn Lys Leu Leu Met Glu Asp
945 950 955 960
Ile Pro Glu Lys Met Ile Glu Arg Gln Met Asn Asp Thr Arg Tyr Ile
965 970 975
Ser Lys Tyr Ile Ser Gly Val Leu Ser Asn Ile Val Arg Val Glu Asp
980 985 990
Gly Ser Asp Glu Gly Val Asn Ser Lys Asn Ile Val Pro Gly Asn Gly
995 1000 1005
Lys Ile Thr Thr Gln Leu Lys Gln Asp Trp Gly Leu Asn Asp Val
1010 1015 1020
Trp Asn Asp Leu Ile Leu Pro Arg Phe Glu Arg Met Asn Gln Leu
1025 1030 1035
Thr Asn Ser Lys Val Phe Thr Ala Trp Asn Glu Asn Tyr Gln Lys
1040 1045 1050
Phe Leu Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys
1055 1060 1065
Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile
1070 1075 1080
Ala Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser
1085 1090 1095
Ala Lys Ser Asp Thr Lys Arg Tyr Asp Leu Lys Lys Lys Ser Met
1100 1105 1110
Lys Phe Glu Lys Val Val Tyr Asn Asp Ala Lys Thr Gly Glu Lys
1115 1120 1125
Ile Glu Arg Glu Val Pro Lys Gln Phe Leu Lys Pro Trp Glu Asn
1130 1135 1140
Phe Thr Leu Asp Val Lys His Asn Leu Glu Thr Ile Ile Val Ser
1145 1150 1155
Phe Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr
1160 1165 1170

Glu Lys Tyr Val Glu Lys Asp Gly Thr Lys Asn Lys Glu Arg Val
1175 1180 1185
Glu Gln Thr Gly Thr Asn Trp Ala Ile Arg Lys Pro Met His Lys
1190 1195 1200
Asp Thr Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro
1205 1210 1215
Lys Gly Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Ser Ser
1220 1225 1230
Phe Asp Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln
1235 1240 1245
Lys Ile Leu Lys Asn Tyr Leu Ala Phe Lys Asp Gly Asn Pro Glu
1250 1255 1260
Leu Ala Phe Ser Pro Glu Gly Ile Asp Asp Leu Asn Lys Asn Ile
1265 1270 1275
Glu Lys Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val
1280 1285 1290
Arg Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Ser Gly
1295 1300 1305
Asn Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu
1310 1315 1320
Phe Phe Ala Val Tyr Glu Asp Glu Lys Gly Lys Arg Asn Tyr Glu
1325 1330 1335
Thr Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu
1340 1345 1350
Ser Val Val Asp Leu Lys Gly Thr Asn Asp Phe Tyr Leu Cys Pro
1355 1360 1365
Asn Asp Phe Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Ile
1370 1375 1380
Asn Asn Val Asp Phe Lys Asp Ile Lys Lys Glu Ile Asn Glu Arg
1385 1390 1395
Ile Tyr Lys Val Val Ser Phe Thr Gly Asn Arg Leu Ser Cys Ile
1400 1405 1410
Pro Tyr Met Val Ala Thr Thr Ile Val Asn Lys Leu Glu Phe Thr
1415 1420 1425
Gln Leu Asn Lys Ile Glu Phe Thr Lys Glu Lys Glu Ile Cys Ile
1430 1435 1440
Lys Leu Asn Val Asp Arg Leu Gly Asn Ile Ser Lys Ala
1445 1450 1455

<210> 107

<211> 1071
<212> PRT
<213> 芽孢杆菌属物种
<223> APG05733.1_RGN
<400> 107

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

Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285
Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
290 295 300
Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320
Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335
Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350
Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
355 360 365
Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
370 375 380
Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400
Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415
Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430
Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445
Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
450 455 460
Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Ile Ile Pro
465 470 475 480
Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495
Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510
His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525
Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540
Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560
Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575
Ser Leu Lys Lys Ile Pro Ala Asn Thr Phe Phe Asn Glu Leu Lys Lys

580 585 590

Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr

595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly

610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser

625 630 635 640

Glu Glu Asp Lys Glu Trp Glu Ser Phe Glu Ser Tyr Val Arg Ser Asn

645 650 655

Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr

660 665 670

Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr

675 680 685

Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys

690 695 700

Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val

705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys

725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val

740 745 750

Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln

755 760 765

Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro

770 775 780

Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro

785 790 795 800

Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Ser

805 810 815

Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser

820 825 830

Ala His Asp Gln Thr Val Met Lys Lys Gly Gly Ile Asp Lys Lys Gly

835 840 845

Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp

850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr

865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Lys Ser Lys

885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910
Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925
Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
1055 1060 1065
Ile Phe Lys
1070

<210> 108
<211> 1403
<212> PRT
<213> 金黄杆菌属物种
<223> APG06207.1_RGN
<400> 108

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

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

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

690 695 700

Ser Ala Gln Glu Arg Glu Asp Met Thr Lys Ser Ile Asn Glu Ala Thr
705 710 715 720

Ile Leu His Gln Lys Tyr Ala Thr Val Leu Arg Gln Glu Phe Gly Val
725 730 735

Val Asn Pro Ser Arg Asn Asp Ile Ile Arg Tyr Lys Leu Tyr Met Glu
740 745 750

Leu Ala Gly Asn Gly Tyr Lys Asp Leu Tyr Thr Asn Val Lys Ile Glu
755 760 765

Lys Glu Asn Ile Phe Thr Asp Lys Tyr Asp Ile Asp His Ile Ile Pro
770 775 780

Gln Ser Arg Phe Phe Asp Asp Ser Phe Ser Asn Lys Val Leu Val Pro
785 790 795 800

Arg Gln Ala Asn Leu Asp Lys Gly Asn Leu Thr Gly Tyr Asp Phe Met
805 810 815

Ser Asn Lys Gly Ala Glu Arg Glu Glu His Phe Leu Asn Val Ile Lys
820 825 830

Asp Leu Leu Glu Asn Gly Ser Ile Ser Lys Ala Lys Tyr Glu Lys Leu
835 840 845

Lys Lys Lys Gly Ile Glu Ile Gly Asp Gly Phe Ile Glu Arg Asp Leu
850 855 860

Arg Asp Thr Gln Tyr Ile Ala Arg Lys Ala Lys Glu Ile Leu Phe Glu
865 870 875 880

Ile Thr Asn Ser Val Ile Ser Thr Ser Gly Arg Ile Thr Asp Lys Leu
885 890 895

Arg Glu Asp Trp Asp Leu Val Asn Thr Met Lys Glu Leu Asn Leu Asp
900 905 910

Lys Tyr Arg Arg Leu Gly Leu Thr Glu Ile Val Ile Asn Ser Lys Gly
915 920 925

Glu Glu Lys Glu Val Ile Pro Asp Trp Thr Lys Arg Asn Asp His Arg
930 935 940

His His Ala Met Asp Ala Leu Thr Val Ala Phe Thr Thr Arg Asn His
945 950 955 960

Ile Gln Tyr Leu Asn Tyr Leu Asn Ala Arg Lys Asp Glu Lys His Lys
965 970 975

Glu His Gln Asn Ile Tyr Ala Ile Glu Asn Ile Ile Thr Glu Ile Ile
980 985 990

Glu Lys Lys Asn Gly Ser Lys Ser Arg Arg Phe Lys Thr Pro Met Asn
995 1000 1005

Arg Phe Arg His Glu Ala Lys Gln His Leu Lys Glu Val Leu Val
1010 1015 1020
Ser His Lys Thr Lys Asn Lys Val Val Thr Ser Asn Ile Asn Lys
1025 1030 1035
Thr Lys Lys Lys Arg Gly Ile His Lys Lys Ala Glu Leu Thr Pro
1040 1045 1050
Arg Gly Gln Leu His Lys Glu Thr Val Tyr Gly Ser Lys Lys Phe
1055 1060 1065
Leu Gln Ser Arg Glu Glu Lys Val Ser Ala Lys Phe Asp Tyr Glu
1070 1075 1080
Thr Ile Leu Met Val Val Asn Pro Ile His Arg Asn Ala Leu Leu
1085 1090 1095
Asn Arg Leu Lys Glu Tyr Gly Asn Asp Pro Lys Lys Ala Phe Ser
1100 1105 1110
Gly Lys Asn Ala Val Asn Lys Thr Pro Val Tyr Leu Ser Asp Ser
1115 1120 1125
Lys Ile Glu Val Leu Pro Glu Lys Val Thr Leu Ser Trp Phe Glu
1130 1135 1140
Thr Gly Tyr Thr Ile Arg Lys Ala Val Thr Pro Asp Asn Phe Lys
1145 1150 1155
Asp Tyr Lys Asn Leu Glu Lys Ile Thr Asp Met Gly Ile Lys Lys
1160 1165 1170
Ile Leu Lys Glu Arg Leu Asp Gln Phe Lys Gly Asn Ala Lys Glu
1175 1180 1185
Ala Phe Ser Asn Leu Asp Lys Glu Pro Ile Trp Leu Asn Glu Glu
1190 1195 1200
Lys Gly Ile Ala Ile Lys Thr Val Thr Ile Thr Gly Val Ser Asn
1205 1210 1215
Ala Glu Ser Leu His Phe Lys Lys Asp His Leu Gly Lys Glu Ile
1220 1225 1230
Leu Asp Glu Asn Gly Asn Lys Ile Pro Val Asp Phe Val Ser Thr
1235 1240 1245
Gly Asn Asn His His Leu Ala Ile Tyr Leu Asn Glu Glu Glu Lys
1250 1255 1260
Leu Asp Asp Lys Met Val Thr Phe Tyr Glu Ala Val Leu Arg Val
1265 1270 1275
Asn Gln Gly Leu Pro Val Ile Asp Lys Asn Tyr Asn Arg Glu Lys
1280 1285 1290
Gly Tyr Lys Phe Leu Met Thr Leu Lys Gln Asn Glu Met Phe Val

1295 1300 1305
Phe Pro Asn Glu Glu Phe Asp Pro Ser Glu Ile Asp Leu Leu Asp
1310 1315 1320
Glu Lys Asn Leu Glu Gln Ile Ser Gln Asn Leu Phe Arg Val Gln
1325 1330 1335
Lys Ile Ser Lys Val Gly Tyr Gly Asn Ser Phe Ile Arg Asp Phe
1340 1345 1350
Val Phe Arg His His Leu Glu Thr Thr Val Glu Asp Arg Lys Glu
1355 1360 1365
Leu Arg Asn Thr Thr Tyr Ile Gln Leu Lys Ser Leu Glu Gly Leu
1370 1375 1380
Arg Asn Ile Val Lys Val Arg Leu Asn His Leu Gly Lys Ile Val
1385 1390 1395
Gln Ile Gly Glu Tyr
1400

<210> 109
<211> 1404
<212> PRT
<213> 鞘氨醇杆菌属物种
<223> APG01647.1_RGN
<400> 109

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

Val Asn Lys Glu Asp Asp Gly Gln Leu Val Asp Gly Met Glu Val Ala
145 150 155 160
Lys Lys Leu Tyr His Glu Asn Leu Thr Pro Gly Glu Phe Val Tyr Arg
165 170 175
Leu Leu Lys Glu Gly Lys Lys Lys Ile Pro Asp Phe Tyr Arg Ser Asp
180 185 190
Leu Gln Arg Glu Leu Asp Leu Ile Trp Asp Phe Gln Phe Lys Tyr Tyr
195 200 205
Pro Asp Leu Leu Thr Gln Asp Leu Lys Glu Glu Leu Lys Gly Lys Gly
210 215 220
Leu Arg Ala Thr Ser Ala Ile Phe Trp Ser Lys Tyr Lys Phe Asn Thr
225 230 235 240
Ala Glu Asn Lys Gly Thr Arg Glu Glu Lys Lys Ile Gln Ala Tyr Gln
245 250 255
Trp Arg Lys Asp Gly Leu Thr Lys Gln Leu Glu Gln Glu Glu Val Ala
260 265 270
Tyr Ile Ile Ala Glu Ile Asn Gly Gln Leu Ala Asn Ser Ser Gly Tyr
275 280 285
Leu Gly Ala Ile Ser Asp Arg Ser Lys Glu Leu Tyr Phe Asn Lys Gln
290 295 300
Thr Val Gly Gln Phe Leu Phe Asn Gln Leu Gln His Asn Thr His Ala
305 310 315 320
Arg Val Lys Asn Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe
325 330 335
Glu Lys Ile Trp Ser Val Gln Ser Gln Tyr His Gln Gly Leu Thr Glu
340 345 350
Gln Leu Lys Glu Glu Ile Arg Asp Ile Ile Ile Phe Tyr Gln Arg Lys
355 360 365
Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Gln His
370 375 380
Glu Val Val Ile Asn Gly Lys Asn Lys Ile Val Gly Leu Arg Val Ala
385 390 395 400
Pro Lys Ser Ser Pro Ile Phe Gln Glu Phe Lys Ile Trp Gln Gln Ile
405 410 415
Asn Asn Val Lys Leu Arg Asn Lys Ile Thr Lys Glu Val Tyr Asn Leu
420 425 430
Ala Glu Glu Gln Lys Arg Ser Leu Phe Glu Thr Leu Asn Leu Lys Gly
435 440 445
Lys Leu Ser Ala Tyr Gln Val Leu Asn Leu Ile Glu Val Lys Pro Lys

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

Lys Asn Leu Tyr Thr Asn Glu Tyr Ile Gln Arg Lys Asp Leu Phe Thr
770 775 780
Asn Val Tyr Asp Ile Asp His Ile Ile Pro Gln Ser Arg Leu Phe Asp
785 790 795 800
Asp Ser Phe Ser Asn Lys Val Leu Val Pro Arg Asn Ile Asn Ile Glu
805 810 815
Lys Gly Asn Gln Thr Ala Phe Asp Tyr Val His Ala Lys Phe Gly Glu
820 825 830
Asp Gly Ile Glu Ala Tyr Glu Ala Arg Val Glu Arg Leu Phe Asn Leu
835 840 845
Lys Glu Glu Gly Val Ser Arg Ser Lys Tyr Lys Lys Leu Leu Met Arg
850 855 860
Gly Leu Asp Ile Gly Glu Gly Phe Ile Glu Arg Asp Leu Arg Asp Ser
865 870 875 880
Gln Tyr Ile Ala Lys Lys Ala Lys Ala Met Leu Phe Glu Ile Ala Pro
885 890 895
Ser Val Ile Ser Thr Ser Gly Ser Val Thr Asp Arg Leu Arg Glu Asp
900 905 910
Trp Gly Leu Val Ser Val Met Lys Glu Leu Asn Leu Pro Lys Phe Lys
915 920 925
Ala Val Gly Leu Thr Glu Tyr Leu Glu Thr Lys Asp Gly Asn Arg Lys
930 935 940
Glu Val Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala
945 950 955 960
Met Asp Ala Leu Thr Val Ala Phe Thr Lys His Ser His Ile His Tyr
965 970 975
Leu Asn His Leu Asn Ala Arg Lys Asn Glu Lys Ser Glu Phe Phe Ser
980 985 990
Thr Ile Lys Ala Ile Glu Ile Lys Glu Thr Tyr Val Glu Arg Asp Asp
995 1000 1005
Leu Gly Asn Arg Lys Arg Leu Phe Lys Glu Pro Ile Pro His Phe
1010 1015 1020
Arg Ser Ile Ala Lys Glu His Leu Glu Ser Val Leu Val Ser His
1025 1030 1035
Lys Ala Lys Asn Lys Val Val Thr Lys Asn Arg Asn Lys Ile Asp
1040 1045 1050
Gly Lys Lys Gln Ala Gln Glu Val Leu Thr Pro Arg Gly Gln Leu
1055 1060 1065
His Lys Glu Ser Ile Tyr Gly Lys Ile Leu Gln Tyr Ala Ser Lys

1070 1075 1080

Glu Glu Lys Ile Ser Ala Lys Phe Asp Ala Ala Thr Ile Ala Met

1085 1090 1095

Val Ser Asn Pro Arg Tyr Arg Ser Ala Leu Leu Asp Arg Leu Ser

1100 1105 1110

Ala Tyr Gly Asn Asp Pro Lys Lys Ala Phe Thr Gly Lys Asn Ser

1115 1120 1125

Pro Asn Lys Ser Pro Ile Tyr Leu Asp Gln Glu Lys Gln Leu Ala

1130 1135 1140

Val Pro Glu Lys Val Lys Leu Val Trp Leu Glu Glu Asp Tyr Thr

1145 1150 1155

Ile Arg Lys Glu Ile Gly Pro Asp Leu Lys Ile Glu Lys Val Ile

1160 1165 1170

Asp Gln Gly Val Lys Arg Ile Leu Glu Gln Arg Leu Thr Glu Tyr

1175 1180 1185

Gly Gly Asp Ser Lys Lys Ala Phe Ser Asp Leu Asp Asn Asn Pro

1190 1195 1200

Ile Trp Leu Asn Lys Glu Lys Gly Ile Ala Ile Lys Arg Val Thr

1205 1210 1215

Ile Ser Gly Val Lys Asn Ala Glu Ala Leu His Ile Lys Lys Asp

1220 1225 1230

His Leu Gly Gln Val Leu Lys Asp Lys Asp Gly Lys Pro Met Pro

1235 1240 1245

Val Asp Phe Val Ser Thr Gly Asn Asn His His Val Ala Ile Tyr

1250 1255 1260

Glu Asp Asp Asn Gly Ala Leu Gln Glu Arg Ile Val Pro Phe Tyr

1265 1270 1275

Glu Ala Val Ala Arg Val Asn Asn Gly Phe Pro Ile Ile Asp Lys

1280 1285 1290

His Tyr Asn Gln His Leu Gly Trp Lys Phe Leu Phe Ser Met Lys

1295 1300 1305

Gln Asn Glu Met Phe Leu Phe Pro Ser Glu Glu Phe Asp Pro Lys

1310 1315 1320

Gln Ile Asn Leu Leu Asp Arg Ser Asn Tyr Arg Leu Ile Ser Lys

1325 1330 1335

Asn Leu Phe Arg Val Gln Lys Phe Gly Glu Leu Ser Lys Ser Gly

1340 1345 1350

Phe Trp Phe Arg His His Leu Glu Thr Ser Val Asp Val Asn Lys

1355 1360 1365

Ser Leu Arg Ser Ile Ser Tyr Phe Asp Phe Tyr Ser Lys Asp Phe
1370 1375 1380
Met Lys Asn Ile Ala Lys Val Arg Ile Asn His Leu Gly Glu Ile
1385 1390 1395
Ile His Val Gly Glu Tyr
1400

<210> 110
<211> 1453
<212> PRT
<213> 金黄杆菌属物种
<223> APG08032.1_RGN
<400> 110

Met Met Ile Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile
1 5 10 15
Gly Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys Tyr Gly Glu Ile
20 25 30
Leu Gly Met Gly Ser Arg Ile Ile Pro Met Asp Gln Lys Val Leu Gly
35 40 45
Glu Phe Gly Ser Gly Asn Ser Val Ser Gln Thr Ala Asp Arg Thr Lys
50 55 60
Tyr Arg Gly Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu
65 70 75 80
Arg Leu His Arg Val Leu Asn Val Leu Asn Phe Leu Pro Ile His Tyr
85 90 95
Ala Ser Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Arg Glu Glu
100 105 110
Ile Glu Pro Lys Leu Val Tyr Asn Asn Asp Gly Phe Ile Phe Lys Asn
115 120 125
Ser Phe Glu Glu Met Leu Ser Asp Phe Arg Lys His Gln Pro Gln Ile
130 135 140
Leu Glu Asn Asp Lys Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu
145 150 155 160
Arg Lys Lys Ala Leu Thr Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp
165 170 175
Ile Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly
180 185 190
Glu Asp Ile Glu Glu Glu Lys Asp Lys Thr Phe Val Arg Leu Lys Val
195 200 205
Glu Lys Ile Val Asp Ser Gly Asp Asn Val Lys Gly Lys Ile Leu Tyr

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

Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys Ile Glu Tyr Glu Lys
530 535 540
Ala Leu Lys Ser Phe Ala Asn Lys Asn Asn Leu Asp Glu Val Ser Phe
545 550 555 560
Phe Glu Ala Phe Lys Lys Phe Pro Pro Phe Lys Asn Glu Tyr Gly Ser
565 570 575
Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Val Gly
580 585 590
Lys Tyr Trp Asp Glu Asp Glu Ile Val Lys Asn Ser Asp Thr Tyr Phe
595 600 605
Lys Asn Ile Glu Asp Leu Leu Gly Asn Ile Ala Arg Lys Glu Glu Asn
610 615 620
Ile Ser Asp Asp Asp Arg Lys Lys Tyr Asn Lys Thr Val Asn Leu Lys
625 630 635 640
Leu Arg Glu Glu Leu Glu Ile Phe Gln Asp Ala Glu Ile Ala Ser Phe
645 650 655
Gln Lys Leu Arg Leu His Ile Ala Gln Tyr Leu Val Tyr Gly Arg His
660 665 670
Ser Glu Ala Ser Ile Ile Gly Lys Trp Asn Ser Ala Asp Asp Leu Glu
675 680 685
Glu Phe Leu Lys Glu Phe Lys Gln His Ser Leu Arg Asn Pro Ile Val
690 695 700
Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp Met
705 710 715 720
Lys Tyr Gly Asn Gly Ala Lys Asp Phe Phe Asn Glu Ile His Ile Glu
725 730 735
Leu Gly Arg Glu Met Lys Leu Pro Ala Glu Asp Arg Lys Lys Leu Thr
740 745 750
Ser Gln Ile Ser Glu Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala Leu
755 760 765
Leu Ala Glu Met Met Asn Asp Ser Ser Val Glu Asn Val Arg Pro Phe
770 775 780
Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr Glu Asp Gly Val Leu
785 790 795 800
Asn Ser Gly Ile Glu Ile Glu Asp Glu Tyr Leu Lys Ile Ser Lys Thr
805 810 815
Ala Gln Pro Ser Ser Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu
820 825 830
Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys

835 840 845

Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile Ile Pro Lys Ser Arg

850 855 860

Tyr Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala Val

865 870 875 880

Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe

885 890 895

Ala Gly Ala Ile Val Glu Leu Gly Phe Gly Lys Asn Val Lys Val Phe

900 905 910

Glu Ile Glu Glu Tyr Glu Asp Phe Val Lys Lys His Tyr Ala Asn Asn

915 920 925

Arg Gly Lys Arg Asn Lys Leu Leu Leu Glu Glu Ile Pro Glu Lys Met

930 935 940

Ile Glu Arg Gln Leu Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser

945 950 955 960

Ser Val Leu Ser Asn Ile Val Arg Val Glu Asp Gly Thr Asp Glu Gly

965 970 975

Val Asn Ser Lys Asn Ile Val Pro Gly Asn Gly Lys Ile Thr Ile Gln

980 985 990

Leu Lys Gln Asp Trp Gly Leu Asn Asp Val Trp Asn Asp Leu Ile Leu

995 1000 1005

Pro Arg Phe Glu Arg Met Asn Gln Leu Thr Asn Ser Thr Asp Phe

1010 1015 1020

Thr Ala Trp Asn Lys Asn His Gln Lys Tyr Leu Pro Thr Val Pro

1025 1030 1035

Ile Glu Phe Ser Lys Gly Phe Ser Lys Lys Arg Ile Asp His Arg

1040 1045 1050

His His Ala Leu Asp Ala Leu Val Ile Ala Cys Thr Thr Lys Asp

1055 1060 1065

His Val Asn Leu Leu Asn Asn Gln Ser Ala Lys Ser Asp Thr Lys

1070 1075 1080

Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys Phe Glu Lys Val Val

1085 1090 1095

Tyr Asn His Thr Gln Thr Gly Glu Lys Ile Glu Arg Asp Ile Pro

1100 1105 1110

Lys Gln Phe Leu Lys Pro Trp Glu Thr Phe Thr Ile Asp Ala Lys

1115 1120 1125

Asn Arg Leu Glu Thr Ile Ile Val Ser Phe Lys Gln Asn Leu Arg

1130 1135 1140

Val Ile Asn Lys Ala Thr Asn His Tyr Glu Lys Tyr Val Glu Lys
1145 1150 1155
Glu Gly Val Lys Met Lys Glu Arg Val Glu Gln Thr Gly Thr Asn
1160 1165 1170
Trp Ala Ile Arg Lys Ser Leu His Glu Glu Thr Val Ser Gly Lys
1175 1180 1185
Ile Thr Leu Ser Trp Val Glu Thr Ser Lys Gly Glu Phe Ile Thr
1190 1195 1200
Ala Thr Arg Lys Pro Leu Asp Ser Ser Phe Thr Phe Glu Lys Ile
1205 1210 1215
Asn Lys Ile Thr Asp Thr Gly Ile Gln Lys Ile Leu Asn Asn Tyr
1220 1225 1230
Leu Glu Ala Lys Asp Asn Asn Pro Glu Leu Ala Phe Ser Ala Glu
1235 1240 1245
Gly Ile Glu Asp Leu Asn Lys Asn Ile Glu Lys Tyr Asn Asp Gly
1250 1255 1260
Lys Ser His Gln Pro Ile Ile Lys Ile Arg Leu Tyr Glu Lys Gly
1265 1270 1275
Lys Gly Arg Phe Val Leu Gly Gln Thr Gly Asn Lys Ile Asn Lys
1280 1285 1290
Tyr Val Gln Gly Ser Pro Asn Leu Phe Phe Ala Ile Tyr Lys Asp
1295 1300 1305
Glu Asn Gly Lys Lys Ile Phe Glu Ser Ile Arg Leu Asp Ile Val
1310 1315 1320
Ile Glu Arg Leu Lys Gln Gly Leu Gln Ala Ile Pro Glu Thr Asn
1325 1330 1335
Gln Asn Gly Val Ser Leu Tyr Gln Ser Leu Ser Pro Leu Asp Leu
1340 1345 1350
Val Tyr Ile Pro Thr Glu Tyr Glu Leu Glu Ser Pro His Ile Leu
1355 1360 1365
Asp Phe Ser Lys Leu Asn Lys Thr Gln Ile Thr Arg Leu Tyr Asn
1370 1375 1380
Thr Asn Asp Phe Ser Gly Val Thr Ala Tyr Phe Ser Gln Asn Ser
1385 1390 1395
Phe Ala Lys His Ile Tyr Pro Lys Glu Met Asp Leu Ser Trp Asn
1400 1405 1410
Glu Lys Lys Gln Lys Leu Ser Gly Ser Phe Asp Ser Lys Thr Ala
1415 1420 1425
Ser Tyr Asn Asn Thr Ser Ile Lys Asp Ile Phe Ile Lys Val Lys

1430 1435 1440

Val Asp Arg Leu Gly Asn Ile Ser Lys Ala

1445 1450

<210> 111

<211> 1093

<212> PRT

<213> 布鲁氏菌属物种

<223> APG05712.1_RGN

<400> 111

Met Lys Thr Leu Gly Ile Asp Leu Gly Thr Ala Ser Ile Gly Trp Ala
1 5 10 15

Ile Leu Asp Glu Ala Ser Ile Ile Ala Ser Gly Val Arg Ile Phe Ser
20 25 30

Gln Ser Glu Met Ala Gly Arg Asp Pro Gln Ser Lys Ala Ser Leu Ala
35 40 45

Val Ala Arg Arg Asp Ala Arg Gly Ala Arg Arg Arg Arg Asp Arg Tyr
50 55 60

Leu Lys Arg Arg Arg Arg Leu Leu Asp Leu Leu Thr Glu His Gly Leu
65 70 75 80

Ala Pro Ala Asp Glu Lys Ser Arg Lys Ala Leu Val Arg Glu Tyr Glu
85 90 95

Asp Gly Lys Asp Gly Asp Leu Ser Asn Ser Val Tyr Ala Leu Arg Ala
100 105 110

Arg Ala Leu Asp Glu Ala Leu Thr Pro Tyr Glu Ile Gly Arg Ala Ile
115 120 125

Phe Gln Leu Asn Gln Arg Arg Gly Phe Lys Ser Asn Arg Lys Thr Asp
130 135 140

Ser Asn Asp Pro Glu Gln Gly Lys Ile Ala Thr Ala Ile His Val Leu
145 150 155 160

Asp Ala Lys Met Asp Glu Asp Lys Ala Arg Thr Phe Gly Glu Trp Leu
165 170 175

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

Asp Gly Asp Ser Tyr Asp Phe Tyr Pro Ser Arg Ala Ala Leu Glu Arg
195 200 205

Glu Phe Asp Arg Leu Met Ala Ser Gln Lys Arg Phe His Pro Asp Leu
210 215 220

Leu Asn Ser Ser Val Ile Asp Asp Ile Arg Lys Val Val Phe His Gln
225 230 235 240

Arg Pro Leu Lys Pro Val Gln Pro Gly Lys Cys Ser Tyr Asn His Arg
245 250 255
Glu Ser Arg Leu Pro Lys Ala His Pro Leu Phe Gln Lys Phe Arg Leu
260 265 270
Leu Lys Glu Val Asn Glu Leu Glu Ile Val Gly Glu Asp Gln Arg Tyr
275 280 285
Val Lys Leu Thr Pro Ala Gln Arg Asp Val Leu Thr Leu Ala Leu Arg
290 295 300
Thr Gly Leu Thr Lys Gln Gly Arg Leu Pro Phe Ser Lys Leu Arg Ser
305 310 315 320
Ile Leu Lys Leu Gly Lys Glu Val Arg Phe Asn Lys Glu Lys Asp Asn
325 330 335
Arg Thr Asp Leu Glu Gly Asp Val Ile His Phe Arg Val Ser Arg Pro
340 345 350
Asp Cys Phe Gly Asn Arg Trp Ala Ala Met Pro Val Glu Glu Gln Ala
355 360 365
Ala Val Thr Glu Lys Leu Arg Thr Glu Pro Asp Tyr Ser Ala Leu Leu
370 375 380
Asp Trp Leu Lys Asn Glu Ala Gly Leu Asp Glu Ala His Ala Arg Ala
385 390 395 400
Val Ala Asp Thr Pro Val Pro Asp Gly Phe Gly Arg Met Gly Pro Ser
405 410 415
Ala Leu Ser Ala Leu Ala Asp Ala Met Glu His Glu Ile Asp Ala Gln
420 425 430
Gly Phe Val Ile Thr Glu Ala Glu Ala Ala Lys Arg Val Tyr Gly Arg
435 440 445
Thr Asn Ser Glu Ala Asp Pro Gly Arg Lys Gly Val Asp Gln Leu Pro
450 455 460
Lys Tyr Gln Glu Val Leu Gln Arg His Ile Pro Pro Gly Thr Gly Glu
465 470 475 480
Pro Asp Asp Pro Tyr Asp Glu Tyr Met Gly Arg Ile Thr Asn Pro Thr
485 490 495
Val His Ile Ala Leu Asn Gln Leu Arg Arg Leu Val Asn Ala Leu Ile
500 505 510
Arg Lys Tyr Gly Lys Pro Asn Lys Ile Ala Ile Glu Val Gly Arg Glu
515 520 525
Leu Lys Leu Asn Glu Lys Gln Arg Asn Glu Val Asn Arg Glu Ile Gly
530 535 540
Gln Asn Thr Arg Ala Ala Met Ala Arg Gly Gln Gln Leu Val Glu Ile

545 550 555 560
Phe Lys Gln Pro Asn Thr Gly Tyr Asn Arg Leu Arg Leu Glu Leu Trp
565 570 575
Glu Asp Leu Asn Arg Glu Gln Pro Leu Lys Arg Leu Cys Thr Tyr Cys
580 585 590
Gly Lys Ala Ile Ala Ala His Met Leu Phe Asn Gly Glu Thr Asp Ile
595 600 605
Asp His Ile Leu Pro Tyr Ser Lys Thr Leu Asp Asp Ser Lys Ala Asn
610 615 620
Arg Leu Leu Cys Cys Thr Pro Cys Asn Arg Glu Lys Lys Asn Tyr Ala
625 630 635 640
Pro Ala Asn Val Leu Gln Trp Arg Asp His Tyr Gly Glu Ile Leu Ala
645 650 655
Arg Ala Thr Ala Leu Pro Lys Asn Lys Gln Trp Arg Phe Ala Glu Asp
660 665 670
Ala Met Thr Arg Tyr Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu
675 680 685
Thr Asp Met Gln Tyr Ile Ser Arg Leu Ala Leu Thr Tyr Leu Ala His
690 695 700
Leu Tyr Asp Tyr Glu Glu Pro Asp Leu Asp Gly Val Tyr Lys Arg His
705 710 715 720
Asp Arg Val Arg Ala Leu Pro Gly Arg Met Thr Glu Met Leu Arg Arg
725 730 735
Gln Trp Ala Leu Asn Glu Leu Leu His Gly His Asn Leu Ala Gly Gly
740 745 750
Asp Gly Ala Lys Glu Lys Asn Arg Leu Asp His Arg His His Ala Ile
755 760 765
Asp Ala Ile Val Ile Ala Cys Thr Ser Gln Ser Leu Ile Asn Arg Leu
770 775 780
Ser Thr Ala Ala Gly Glu Ala Glu Glu Arg Gly Ala Ala Arg Val Val
785 790 795 800
Glu Arg Ile Asp Pro Pro Trp Pro Ser Phe Arg Glu Asp Val Arg Glu
805 810 815
Ala Val Asn Ala Ile Val Val Ser His Lys Pro Asp His Gly Thr Ala
820 825 830
Ser Arg Ser Gly Tyr Asp Lys Gly Arg Gly Gln Thr Ala Gly Lys Leu
835 840 845
His Asn Asp Thr Ala Tyr Gly Glu Thr Gly Glu Lys Asp His Asn Gly
850 855 860

Asn Asn Leu Val Val Arg Arg Ile Ala Ile Ser Asp Ile Lys Arg Ser
865 870 875 880
Ala Asp Ile Met Lys Ile Arg Thr Asn Ala His Gly His Ser Glu Leu
885 890 895
Arg Asp Arg Leu Tyr Glu Ala Thr Arg Asp Leu Glu Gly Lys Ala Phe
900 905 910
Glu Gln Ala Val Thr Ala Phe Val Lys His Asp Ala Lys Phe Lys Gly
915 920 925
Ile Arg His Val Arg Val Thr Glu Val Gln Asn Pro Val Trp Ile Thr
930 935 940
His Gly Gly Gly Lys Tyr Lys Lys Gly Tyr Leu Pro Gly Gly Asn Asp
945 950 955 960
Arg Phe Asp Val Trp Glu Leu Pro Asp Gly Lys Trp Asp Ala Glu Val
965 970 975
Val Thr Thr Phe Asp Ala His Arg Pro Asp Phe Thr Pro Arg Met Arg
980 985 990
Ile Glu His His Asn Ala Arg Lys Ile Met Ser Leu Lys Lys Gly Asp
995 1000 1005
Met Ile Ala Tyr Asp Asp Pro Asp Ser Gly Lys Arg Val Ile Ala
1010 1015 1020
Ile Val Arg Lys Phe Asp Gln Arg Asn Lys Gln Leu Tyr Leu Asp
1025 1030 1035
Pro His Asn Glu Ala Gly Asn Leu Asp Gln Arg Glu Lys Glu Lys
1040 1045 1050
Thr Tyr Lys Pro Leu Arg Pro Met Pro Asn Pro Leu Lys Lys Tyr
1055 1060 1065
Arg Pro Arg Gln Val Arg Val Asp Glu Ile Gly Gln Val Phe Asp
1070 1075 1080
Pro Gly Pro Trp Trp Glu Lys Arg Ser Asp
1085 1090

<210> 112
<211> 1054
<212> PRT
<213> 葡萄球菌属物种
<223> APG01658.1_RGN
<400> 112
Met Lys Glu Lys Tyr Ile Leu Gly Leu Asp Leu Gly Ile Thr Ser Val
1 5 10 15
Gly Tyr Gly Ile Ile Asn Phe Glu Thr Lys Lys Ile Ile Asp Ala Gly

20 25 30

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

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

645 650 655
Glu Leu Thr Ser Tyr Leu Lys Ala Tyr Phe Ser Ala Asn Asn Met Asp
660 665 670
Val Lys Val Lys Thr Ile Asn Gly Ser Phe Thr Asn His Leu Arg Lys
675 680 685
Val Trp Arg Phe Asp Lys Tyr Arg Asn His Gly Tyr Lys His His Ala
690 695 700
Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Leu Phe Lys Glu Asn
705 710 715 720
Lys Lys Leu Gln Asn Ala Asn Lys Ile Leu Glu Lys Pro Thr Ile Glu
725 730 735
Asn Asn Thr Lys Lys Val Thr Val Glu Lys Glu Glu Asp Tyr Asn Asn
740 745 750
Val Phe Glu Thr Pro Lys Leu Val Glu Asp Ile Lys Gln Tyr Arg Asp
755 760 765
Tyr Lys Phe Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln Leu Ile
770 775 780
Asn Asp Thr Leu Tyr Ser Thr Arg Met Lys Asp Glu His Asp Tyr Ile
785 790 795 800
Val Gln Thr Ile Thr Asp Ile Tyr Gly Lys Asp Asn Thr Asn Leu Lys
805 810 815
Lys Gln Phe Asn Lys Asn Pro Glu Lys Phe Leu Met Tyr Gln Asn Asp
820 825 830
Pro Lys Thr Phe Glu Lys Leu Ser Ile Ile Met Lys Gln Tyr Ser Asp
835 840 845
Glu Lys Asn Pro Leu Ala Lys Tyr Tyr Glu Glu Thr Gly Glu Tyr Leu
850 855 860
Thr Lys Tyr Ser Lys Lys Asn Asn Gly Pro Ile Val Lys Lys Ile Lys
865 870 875 880
Leu Leu Gly Asn Lys Val Gly Asn His Leu Asp Val Thr Asn Lys Tyr
885 890 895
Glu Asn Ser Thr Lys Lys Leu Val Lys Leu Ser Ile Lys Asn Tyr Arg
900 905 910
Phe Asp Val Tyr Leu Thr Glu Lys Gly Tyr Lys Phe Val Thr Ile Ala
915 920 925
Tyr Leu Asn Val Phe Lys Lys Asp Asn Tyr Tyr Tyr Ile Pro Lys Asp
930 935 940
Lys Tyr Gln Glu Leu Lys Glu Lys Lys Lys Ile Lys Asp Thr Asp Gln
945 950 955 960

Phe Ile Ala Ser Phe Tyr Lys Asn Asp Leu Ile Lys Leu Asn Gly Asp
965 970 975
Leu Tyr Lys Ile Ile Gly Val Asn Ser Asp Asp Arg Asn Ile Ile Glu
980 985 990
Leu Asp Tyr Tyr Asp Ile Lys Tyr Lys Asp Tyr Cys Glu Ile Asn Asn
995 1000 1005
Ile Lys Gly Glu Pro Arg Ile Lys Lys Thr Ile Gly Lys Lys Thr
1010 1015 1020
Glu Ser Ile Glu Lys Phe Thr Thr Asp Val Leu Gly Asn Leu Tyr
1025 1030 1035
Leu His Ser Thr Glu Lys Ala Pro Gln Leu Ile Phe Lys Arg Gly
1040 1045 1050
Leu

<210> 113
<211> 1076
<212> PRT
<213> 芽孢杆菌属物种
<223> APG06498.1_RGN
<400> 113

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

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

465 470 475 480
Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495
Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510
His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525
Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540
Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560
Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575
Ser Leu Lys Glu Ile Pro Pro Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590
Glu Arg Asn Gly Ser Ser Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605
Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
610 615 620
Asp Glu Asn Arg Asn Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640
Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
645 650 655
Lys Leu Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Lys Ala Tyr
660 665 670
Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685
Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
690 695 700
Phe Lys Glu Val Glu Val Asn Leu Arg Lys Lys Arg Val Gln Thr Val
705 710 715 720
Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735
Asn Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750
Ala Cys Thr Asp His His Met Val Thr Arg Ile Thr Glu Tyr Tyr Gln
755 760 765
Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
770 775 780

Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
785 790 795 800
Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Ser
805 810 815
Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
820 825 830
Ala His Asn Gln Met Ile Arg Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Lys Arg Val Arg Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Asp His Gly Lys Asn Leu Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910
Asn Leu Ile Lys Arg Val Lys Ile Glu Gly Gln Thr Lys Ala Phe Val
915 920 925
Arg Glu Val Asn Gly Gly Val Ala Gln Asn Ser Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Lys Ser Gly
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Asn Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
1055 1060 1065
Asn Phe His Glu Ile Lys Ile Asn
1070 1075

<210> 114

<211> 1108
<212> PRT
<213> 短芽孢杆菌属物种
<223> APG09106.1_RNG
<400> 114

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

Met Thr Ala Ala Asn Asp Lys Tyr Arg Ile Thr Lys Arg Gln Met Lys
275 280 285
Gly Trp Asn Glu Leu Tyr Glu Gln Trp Ser Thr Val Leu Pro Asn Ala
290 295 300
Ser His Glu Gln Tyr Arg Glu Ala Leu Lys Arg Val Gln Gln Arg Leu
305 310 315 320
Arg Gly Arg Phe Gly Asp Ala His Phe Phe Gln Tyr Leu Met Lys Glu
325 330 335
Glu His His Leu Ile Trp Lys Gly Asn Pro Gln Arg Ile His Tyr Phe
340 345 350
Val Ala Arg Asn Glu Leu Lys Lys Arg Leu Glu Glu Ala Lys Gln Asn
355 360 365
Ala Thr Met Thr Leu Pro Asp Ala Arg Lys His Pro Leu Trp Val Arg
370 375 380
Phe Asp Ala Arg Gly Gly Asn Leu Gln Asp Tyr Tyr Leu Thr Ala Glu
385 390 395 400
Ala Asp Asn Pro Arg Ser Arg Arg Phe Val Thr Phe Ser Gln Leu Ile
405 410 415
Trp Pro Asn Glu Ser Gly Trp Met Glu Lys Gln Asp Val Glu Val Glu
420 425 430
Leu Ala Leu Ser Lys Gln Phe Tyr Gln Gln Val Thr Leu Gln Lys Asn
435 440 445
Asp Lys Gly Lys Gln Glu Ile Glu Phe Lys Asp Lys Gly Ser Gly Ser
450 455 460
Thr Phe Ser Gly His Leu Gly Gly Ala Lys Leu Gln Leu Glu Arg Gly
465 470 475 480
Asp Leu Glu Lys Glu Glu Lys Asp Phe Glu Gly Gly Glu Ile Gly Ser
485 490 495
Val Tyr Leu Asn Ile Val Ile Asp Phe Glu Pro Leu Gln Glu Val Lys
500 505 510
Asn Gly Arg Leu Gln Ser Pro Tyr Gly Gln Val Leu Gln Leu Val Arg
515 520 525
Arg Pro Asn Glu Phe Pro Lys Val Thr Thr Tyr Lys Ser Glu Glu Leu
530 535 540
Val Glu Trp Met Lys Ala Ser Gln Asn His Ser Ser Gly Val Glu Ser
545 550 555 560
Leu Glu Ser Gly Phe Arg Val Met Ser Ile Asp Leu Gly Leu Arg Thr
565 570 575
Ala Ala Ala Thr Ser Ile Phe Ser Val Glu Glu Ser Asn Asp Ala Asn

580 585 590

Ala Ala Gly Phe Ser Tyr Trp Ile Glu Gly Thr Pro Leu Val Ala Val

595 600 605

His Lys Arg Ser Tyr Met Leu Lys Leu Pro Gly Glu Gln Val Glu Lys

610 615 620

Gln Val Arg Glu Lys Arg Asp Glu Arg Gln Asp Gln Gln Arg Arg Val

625 630 635 640

Arg Phe Gln Ile Arg Ile Leu Ser Gln Val Ile Arg Met Ala Lys Lys

645 650 655

Gln Asn Arg Glu Arg Ala Asp Glu Leu Asp His Leu Ser Gln Ala Leu

660 665 670

Glu Lys Gln Lys Ser Leu Leu Asp Gln Thr Asp Arg Thr Phe Trp Asn

675 680 685

Gly Ile Val Cys Asp Leu Thr Asp Ala Leu Arg Glu Lys Glu Gly Gly

690 695 700

Trp Glu Gln Ala Val Val Gln Ile His Arg Lys Ala Glu Glu His Val

705 710 715 720

Gly Lys Val Val Gln Ala Trp Arg Lys Arg Phe Asp Ala Asp Glu Arg

725 730 735

Lys Gly Ile Ala Gly Leu Ser Met Trp Ser Ile Glu Glu Leu Asp Ser

740 745 750

Leu Arg Lys Leu Leu Ile Ser Trp Ser Arg Arg Thr Arg Asn Pro Gln

755 760 765

Glu Ile Asn Arg Phe Glu Gln Gly His Thr Ser His Gln Arg Leu Leu

770 775 780

Thr His Ile Gln Asn Val Lys Glu Asp Arg Leu Lys Gln Leu Ser His

785 790 795 800

Ala Ile Val Met Thr Ala Leu Gly Tyr Val Tyr Asp Glu Lys Lys Leu

805 810 815

Glu Trp Phe Ala Lys Tyr Pro Ala Cys Gln Val Ile Leu Phe Glu Asn

820 825 830

Leu Ser Gln Tyr Arg Ser His Met Asp Arg Ser Thr Lys Glu Asn Ser

835 840 845

Thr Leu Met Lys Trp Ala His Arg Ser Ile Pro Lys Tyr Val His Met

850 855 860

Gln Ala Glu Pro Tyr Gly Ile Gln Ile Gly Asp Val Arg Ala Glu Tyr

865 870 875 880

Ser Ser Arg Phe His Ala Lys Thr Gly Thr Pro Gly Ile Arg Cys Lys

885 890 895

Met Val Lys Gly Gln Glu Leu Gln Gly Lys Arg Phe Glu Asn Leu Gln
900 905 910
Lys Arg Leu Val Ser Glu Gln Phe Leu Thr Glu Glu Gln Val Lys Gln
915 920 925
Leu Arg Pro Gly Asp Ile Val Pro Asp Asp Ser Gly Glu Trp Phe Met
930 935 940
Thr Leu Ser Asp Gly Ser Glu Gly Lys Glu Val Val Phe Leu Gln Ala
945 950 955 960
Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln Arg Tyr
965 970 975
Asn Glu Leu Phe Lys Val Ser Cys Arg Val Leu Ile Arg Gly Glu Glu
980 985 990
Glu Tyr Leu Ile Pro Lys Ala Lys Ser Val Gln Ala Lys Leu Gly Lys
995 1000 1005
Gly Leu Phe Val Lys Lys Thr Asp Thr Val Met Lys Asp Val Tyr
1010 1015 1020
Val Trp Asp Ser Gln Ala Lys Leu Lys Gly Lys Thr Thr Phe Thr
1025 1030 1035
Glu Glu Ser Glu Ser Pro Glu Gln Leu Glu Asp Phe Gln Glu Ile
1040 1045 1050
Ile Glu Glu Ala Glu Glu Ala Lys Gly Thr Tyr Arg Thr Leu Phe
1055 1060 1065
Arg Asp Pro Ser Gly Val Phe Phe Pro Glu Phe Val Trp Asn Thr
1070 1075 1080
Gln Lys Asp Phe Trp Ser Glu Val Lys Arg Arg Leu Tyr Gly Lys
1085 1090 1095
Leu Arg Glu Arg Phe Leu Met Lys Thr Arg
1100 1105

<210> 115
<211> 1150
<212> PRT
<213> 肠球菌属物种
<223> APG09882.1_RGN
<400> 115

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser
1 5 10 15
Val Ile Asp Glu Arg Thr Gly Asn Val Ile Asp Leu Gly Ile Arg Leu
20 25 30
Phe Ser Ala Lys Asn Ser Glu Lys Asn Leu Glu Arg Arg Thr Asn Arg

35 40 45

Gly Gly Arg Arg Leu Ile Arg Arg Lys Thr Asn Arg Leu Lys Asp Ala
50 55 60

Lys Lys Ile Leu Ala Ala Val Gly Phe Tyr Glu Asp Lys Ser Leu Lys
65 70 75 80

Asn Ser Cys Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Glu Pro Leu
85 90 95

Ser Lys Gly Glu Ile Tyr Lys Val Thr Leu His Ile Leu Lys Lys Arg
100 105 110

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

Gln Asp Tyr Lys Glu Gln Val Arg Glu Asn Ala Gln Leu Leu Thr Lys
130 135 140

Tyr Thr Pro Gly Gln Ile Gln Leu Gln Arg Leu Lys Glu Asn Asn Arg
145 150 155 160

Val Lys Thr Gly Ile Asn Ala Gln Gly Asn Tyr Gln Leu Asn Val Phe
165 170 175

Lys Val Ser Ala Tyr Ala Asn Glu Leu Ala Thr Ile Leu Lys Thr Gln
180 185 190

Gln Ala Phe Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe
195 200 205

Val Gln Pro Gly Ile Ala Glu Glu Ala Gly Leu Ile Tyr Arg Lys Arg
210 215 220

Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly
225 230 235 240

Arg Trp Ser Asp Phe Lys Lys Thr Gly Gln Pro Ala Thr Asn Ile Phe
245 250 255

Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly
260 265 270

Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn
275 280 285

Leu Lys Ile Asp Gly Glu Ile Pro Leu Ser Pro Glu Gln Lys Glu Tyr
290 295 300

Ile Leu Ala Glu Leu Met Thr Lys Glu Phe Thr Arg Phe Gly Val Asn
305 310 315 320

Asp Val Val Lys Leu Leu Gly Val Lys Lys Glu Arg Leu Ser Gly Trp
325 330 335

Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr
340 345 350

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

660 665 670

Glu Thr Arg Asn Ile Phe Asp Ser Glu Gln Val Glu Lys Phe Leu Ala

675 680 685

Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr

690 695 700

Leu Gln Ser Phe Phe Ala Asn Gln Glu Thr Lys Val Arg Val Val Asn

705 710 715 720

Gly Ser Phe Thr His Thr Leu Arg Lys Lys Trp Gly Ala Asp Leu Asp

725 730 735

Lys Thr Arg Glu Thr His His His His Ala Val Asp Ala Thr Leu Cys

740 745 750

Ala Val Thr Pro Phe Val Lys Val Ser Arg Tyr His Tyr Ala Val Lys

755 760 765

Glu Glu Thr Gly Glu Lys Val Met Arg Glu Ile Asp Phe Glu Thr Gly

770 775 780

Glu Ile Val Asp Glu Met Ser Tyr Arg Glu Phe Lys Lys Ser Lys Lys

785 790 795 800

Tyr Glu Arg Lys Thr Tyr Gln Val Lys Trp Pro Asn Phe Arg Glu Gln

805 810 815

Leu Lys Pro Val Asn Leu His Pro Arg Ile Lys Phe Ser His Gln Val

820 825 830

Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val

835 840 845

Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Gln Lys Ile

850 855 860

Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val

865 870 875 880

Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys

885 890 895

Asp Leu Asp Glu Lys Thr Tyr Glu Arg Leu Leu Ser Ile Ala Glu Thr

900 905 910

Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg

915 920 925

Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Asp Ile Pro

930 935 940

Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Leu Ile Arg Ser

945 950 955 960

Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys

965 970 975

Asp Ser Gln Gly Arg Pro Val Glu Lys Thr Lys Asn Gly Arg Lys Val
980 985 990
Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu
995 1000 1005
Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu
1010 1015 1020
Arg Phe Val Glu Gly Lys Tyr Gly Ile Thr Glu Lys Glu Tyr Met
1025 1030 1035
Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys
1040 1045 1050
Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp
1055 1060 1065
Ser Gln Cys Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn
1070 1075 1080
Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Met Pro Ala Glu
1085 1090 1095
Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
1100 1105 1110
Asn Ile Ala Lys Tyr Gly Lys Glu Gly Lys Lys Leu Arg Lys Phe
1115 1120 1125
Asn Thr Asp Ile Leu Gly Lys Lys His Tyr Leu Tyr Tyr Glu Lys
1130 1135 1140
Glu Pro Lys Asn Ile Ile Lys
1145 1150

<210> 116
<211> 1408
<212> PRT
<213> 鞘氨醇杆菌属物种
<223> APG02675.1_RGN
<400> 116

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

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

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

690 695 700

Ile Glu Leu Ala Arg Glu Leu Lys Lys Asn Ala Lys Glu Arg Glu Glu
705 710 715 720
Leu Thr Lys Ala Ile Asn Thr Ser Lys Glu Gln His Glu Lys Ile Ile
725 730 735
Lys Ile Leu Gln Thr Glu Asp Gly Ile Lys Asn Pro Thr Arg Asn Asp
740 745 750
Ile Ile Arg Phe Lys Leu Tyr Gln Glu Leu Lys Asn Asn Gly Tyr Lys
755 760 765
Asn Leu Tyr Thr Asn Glu Tyr Ile Gln Arg Lys Asp Leu Phe Thr Asn
770 775 780
Ile Tyr Asp Ile Asp His Ile Val Pro Gln Ser Arg Leu Phe Asp Asp
785 790 795 800
Ser Phe Ser Asn Lys Val Leu Val Pro Arg Asn Ile Asn Ile Glu Lys
805 810 815
Gly Asn Gln Thr Ala Phe Asp Tyr Val Arg Thr Lys Phe Gly Glu Asp
820 825 830
Gly Ile Glu Ala Tyr Glu Ala Arg Val Glu Arg Leu Phe Asn Leu Lys
835 840 845
Glu Glu Gly Val Ser Arg Ser Lys Tyr Lys Lys Leu Leu Met Arg Gly
850 855 860
Ser Asp Ile Gly Glu Gly Phe Ile Glu Arg Asp Leu Arg Asp Ser Gln
865 870 875 880
Tyr Ile Ala Lys Lys Ala Lys Ala Met Leu Phe Glu Ile Thr Pro Ser
885 890 895
Val Ile Ser Thr Ser Gly Ser Val Thr Asp Arg Leu Arg Glu Asp Trp
900 905 910
Gly Leu Val Ser Val Met Lys Glu Leu Asn Leu Pro Lys Phe Lys Ala
915 920 925
Val Gly Leu Thr Glu Tyr Leu Glu Thr Lys Asp Gly Asn Arg Lys Glu
930 935 940
Val Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala Met
945 950 955 960
Asp Ala Leu Thr Val Ala Phe Thr Lys His Ser His Ile Gln Tyr Leu
965 970 975
Asn His Leu Asn Ala Arg Lys Asn Glu Lys Ser Glu Phe Phe Ser Ser
980 985 990
Ile Lys Ala Ile Glu Ile Lys Glu Thr Tyr Val Glu Arg Asp Asp Leu
995 1000 1005

Gly Asn Arg Lys Arg Leu Phe Lys Glu Pro Ile Ala His Phe Arg
1010 1015 1020
Thr Ile Val Lys Glu His Leu Glu Arg Val Leu Val Ser His Lys
1025 1030 1035
Ala Lys Asn Lys Val Val Thr Lys Asn Arg Asn Lys Ile Asp Gly
1040 1045 1050
Lys Lys Glu Ala Gln Glu Val Leu Thr Pro Arg Gly Gln Leu His
1055 1060 1065
Lys Glu Thr Ile Tyr Gly Lys Ile Leu Gln Tyr Ala Ser Lys Glu
1070 1075 1080
Glu Lys Ile Ser Ala Lys Phe Asp Ala Ala Thr Ile Ala Met Val
1085 1090 1095
Ser Asn Pro Arg Tyr Arg Ser Ala Leu Leu Asp Arg Leu Ile Glu
1100 1105 1110
Tyr Gly Asn Asp Pro Lys Lys Ala Phe Thr Gly Lys Asn Ser Pro
1115 1120 1125
Ser Lys Ser Pro Ile Tyr Leu Asp Gln Gln Arg Gln Leu Ala Val
1130 1135 1140
Pro Glu Lys Val Lys Leu Val Trp Leu Glu Ala Asp Tyr Thr Ile
1145 1150 1155
Arg Lys Glu Ile Gly Pro Asp Leu Lys Ile Glu Lys Val Ile Asp
1160 1165 1170
Gln Gly Val Lys Arg Ile Leu Lys Gln Arg Leu Ser Val Tyr Asp
1175 1180 1185
Gly Asp Thr Lys Lys Ala Phe Ser Asp Leu Asp Asn Asn Pro Ile
1190 1195 1200
Trp Leu Asn Glu Glu Lys Gly Ile Ala Ile Lys Arg Val Thr Ile
1205 1210 1215
Ser Gly Val Lys Asn Ala Glu Ala Leu His Thr Lys Lys Asp His
1220 1225 1230
Leu Gly Gln Val Leu Glu Asp Lys Asp Gly Asn Thr Met Pro Val
1235 1240 1245
Asp Phe Val Ser Thr Gly Asn Asn His His Val Ala Ile Tyr Glu
1250 1255 1260
Asp Ser Glu Gly Asn Leu Tyr Asp Arg Val Val Ser Phe Tyr Glu
1265 1270 1275
Ala Val Glu Arg Ile Asn Gln Asn Leu Pro Ile Val Asp Arg His
1280 1285 1290
Tyr Lys Gln Glu Glu Gly Trp Lys Phe Leu Phe Ser Met Lys Gln

1295 1300 1305
Asn Glu Leu Phe Val Phe Pro Ser Ile Asn Phe Asp Pro Lys Glu
1310 1315 1320
Ile Asp Leu Leu Asp Lys Lys Asn Tyr Lys Glu Val Ser Lys His
1325 1330 1335
Leu Phe Arg Met Gln Lys Phe Ser Lys Val Glu Tyr Gly Asn Ser
1340 1345 1350
Ala Val Arg Asp Tyr Val Phe Arg His His Leu Glu Thr Ser Ile
1355 1360 1365
Ile Asp Thr Lys Glu Leu Arg Asp Ile Ala Tyr Lys Val Phe Lys
1370 1375 1380
Ser Ile Gly Glu Phe Lys Ser Phe Val Lys Ile Arg Thr Asn His
1385 1390 1395
Leu Gly Glu Ile Val His Leu Gly Glu Tyr
1400 1405

<210> 117
<211> 1369
<212> PRT
<213> 金黄杆菌属物种
<223> APG01405.1_RGN
<400> 117

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

Tyr Ser Leu Lys Ile Ile Asp Val Leu Ala Asp Glu Pro Gln Lys Gly
145 150 155 160
Lys Ser Asp Ile Trp Tyr Ser Leu Val Leu Glu Asn Gly Trp Ile Tyr
165 170 175
Arg Arg Ser Ser Lys Thr Ser Leu Leu Asp Trp Lys Asp Lys Ile Arg
180 185 190
Asp Phe Ile Val Thr Thr Asp Leu Asn Asp Asn Gly Ser Val Lys Thr
195 200 205
Asp Lys Glu Gly Asn Glu Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn
210 215 220
Asp Trp Thr Leu Val Lys Thr Lys Thr Glu Gln Glu Ile Asp Arg Ser
225 230 235 240
Arg Lys Thr Val Gly Thr Tyr Ile Tyr Glu Thr Leu Leu Gln Asn Pro
245 250 255
Lys Gln Lys Ile Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe
260 265 270
Tyr Lys Glu Glu Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His
275 280 285
Tyr Glu Leu Gln Ser Asp Asp Leu Tyr Asp Asp Cys Ile Arg Glu Leu
290 295 300
Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe
305 310 315 320
Val His Leu Phe Met Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu Lys
325 330 335
Ser Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr
340 345 350
Lys Asp Glu Asn Gly Val Glu His Thr Gln Phe Leu Lys Ala Ile Pro
355 360 365
Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp Ile Phe
370 375 380
Asn Leu Asn Ile Tyr Lys Lys Asp Asn Asp Asp Asn Val Thr Lys Glu
385 390 395 400
Phe Leu Ser Thr Thr Glu Asp Phe Glu Asn Leu Phe Glu Phe Leu Asn
405 410 415
Asn Arg Lys Glu Ile Asp Gln Lys Ala Leu Leu Lys His Phe Lys Leu
420 425 430
Asn Glu Lys Thr His Arg Trp Lys Tyr Val Glu Asp Lys Lys Tyr Pro
435 440 445
Cys Asn Glu Thr Lys Thr Met Ile Ser Glu Arg Leu Lys Lys Val Glu

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

Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu
770 775 780
Tyr Glu Ile Glu His Ile Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser
785 790 795 800
Leu Ser Asn Lys Ile Ile Cys Glu Ser Ala Val Asn Lys Leu Lys Asp
805 810 815
Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile
820 825 830
Glu Cys Gly Phe Gly Lys Arg Val Lys Val Phe Lys Ala Glu Glu Tyr
835 840 845
Glu Glu Phe Val Lys Lys His Tyr Ala Asn Asn Arg Gly Lys Arg Asn
850 855 860
Lys Leu Leu Leu Glu Asp Ile Pro Glu Lys Met Ile Glu Arg Gln Met
865 870 875 880
Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser Gly Ile Leu Ser His
885 890 895
Ile Val Arg Val Glu Asp Gly Thr Asp Glu Gly Val Asn Ser Lys Asn
900 905 910
Val Ile Pro Gly Asn Gly Lys Ile Thr Thr Gln Leu Arg Gln Asp Trp
915 920 925
Gly Leu Asn Asp Val Trp Asn Glu Leu Val Leu Pro Arg Phe Glu Arg
930 935 940
Met Asn Gln Leu Thr Asn Ser Thr Asp Phe Thr Ser Trp Asn Glu Asn
945 950 955 960
His Gln Lys Tyr Leu Pro Thr Val Pro Val Glu Phe Ser Lys Gly Phe
965 970 975
Ser Lys Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val
980 985 990
Ile Ala Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser
995 1000 1005
Ala Lys Ser Asp Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met
1010 1015 1020
Lys Phe Glu Lys Thr Val Tyr Lys Asp Pro Gln Thr Glu Lys Arg
1025 1030 1035
Ile Glu Arg Glu Val Pro Lys Tyr Phe Leu Lys Pro Trp Glu Thr
1040 1045 1050
Phe Thr Val Asp Ala Lys Asn Lys Leu Glu Thr Ile Ile Val Ser
1055 1060 1065
Phe Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr

1070 1075 1080
Glu Lys Tyr Val Asp Lys Asp Gly Val Lys Thr Lys Glu Arg Val
1085 1090 1095
Glu Gln Thr Gly Thr Asn Trp Ala Ile Arg Lys Pro Met His Lys
1100 1105 1110
Glu Thr Val Ser Gly Lys Ile Asp Leu Pro Trp Val Lys Val Pro
1115 1120 1125
Lys Gly Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Thr Ser
1130 1135 1140
Phe Asp Leu Lys Ala Ile Ala Ser Ile Thr Asp Thr Gly Ile Gln
1145 1150 1155
Lys Ile Leu Lys Asn Tyr Leu Glu Phe Lys Glu Ser Pro Glu Leu
1160 1165 1170
Ala Phe Ser Pro Glu Gly Ile Glu Asp Met Asn Lys Asn Ile Lys
1175 1180 1185
Lys Tyr Asn Gly Gly Lys Pro His Gln Pro Ile Ser Lys Val Arg
1190 1195 1200
Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Thr Gly Asn
1205 1210 1215
Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe
1220 1225 1230
Phe Ala Ile Tyr Glu Asp Arg Lys Gly Lys Arg Ser Tyr Glu Thr
1235 1240 1245
Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Ser
1250 1255 1260
Val Val Asp Leu Lys Asn Ile Asn Asp Phe Phe Leu Cys Pro Asn
1265 1270 1275
Asp Leu Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Gly Gly
1280 1285 1290
Ser Ile Glu Ile Lys Asn Ile Thr Lys Glu Lys Ser Glu Arg Phe
1295 1300 1305
Tyr Lys Val Val Ser Phe Ser Gly Ser Gln Ile Phe Phe Val Arg
1310 1315 1320
His Asp Ile Ala Val Ser Ile Val Asn Lys Gly Glu Phe Ser Thr
1325 1330 1335
Leu Asn Lys Met Glu Arg Ala Ile Asp Gly Ser Met Val Lys Glu
1340 1345 1350
Ser Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly Asn Val Leu Lys
1355 1360 1365

Ala

<210> 118

<211> 1440

<212> PRT

<213> 金黄杆菌属物种

<223> APG06250.1_RGN

<400> 118

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

Asp Leu Asn Asp Asp Gly Ser Val Lys Thr Asp Lys Asp Gly Asn Glu
260 265 270
Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn Asp Trp Thr Leu Val Lys
275 280 285
Lys Lys Thr Glu Gln Glu Ile Glu Gln Ser His Lys Thr Val Gly Met
290 295 300
Tyr Ile Tyr Glu Thr Leu Arg Ala Asn Pro Lys Gln Lys Ile Lys Gly
305 310 315 320
Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Glu Glu Leu Arg
325 330 335
Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Gln Asn Asp
340 345 350
Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn Glu Ala
355 360 365
His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Phe Ile Glu
370 375 380
Asp Ile Val Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser Ile
385 390 395 400
Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn Gly Ala
405 410 415
Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro Tyr Tyr
420 425 430
Gln Glu Phe Arg Leu Trp Gln Trp Ile Phe Asn Leu Asn Leu Tyr Lys
435 440 445
Lys Asp Asn Asp Glu Asn Val Ile Lys Glu Phe Leu Thr Thr Thr Gln
450 455 460
Asp Val Glu Asn Leu Phe Glu Phe Leu Asn Asn Arg Lys Glu Ile Asp
465 470 475 480
Gln Lys Ala Leu Leu Lys His Phe Lys Leu Ser Glu Lys Thr His Arg
485 490 495
Trp Asn Phe Val Glu Asp Lys Lys Tyr Pro Cys Asn Glu Thr Lys Thr
500 505 510
Met Ile Ala Thr Arg Leu Glu Lys Val Glu Asn Ile Ser Asp Asp Phe
515 520 525
Leu Thr Arg Glu Ile Glu Gln Lys Ile Trp His Ile Ile Tyr Ser Val
530 535 540
Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Leu Lys
545 550 555 560
His Asn Leu Asp Glu Asn Ser Phe Phe Glu Ala Phe Arg Lys Phe Pro

565 570 575
Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys
580 585 590
Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr Trp Asp Tyr Ala Asn Ile
595 600 605
Asp Gln Phe Ser Lys Gly Arg Ile Gln Lys Ile Ile Asn Gly Glu Tyr
610 615 620
Asp Glu Asn Ile Lys Asp Arg Val Arg Glu Lys Ala Val His Leu Thr
625 630 635 640
Ser Glu Asn Asp Phe Gln Gly Leu Gln Leu Trp Leu Ala Gln Tyr Ile
645 650 655
Val Tyr Gly Arg His Ser Glu Ala Ser Ile Ala Gly Lys Trp Asn Ser
660 665 670
Ala Asp Asp Leu Glu Glu Phe Leu Lys Asp Phe Lys His His Ser Leu
675 680 685
Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val
690 695 700
Lys Asp Ile Trp Leu Lys Tyr Gly Lys Gly Ala Lys Asp Phe Phe Asn
705 710 715 720
Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp Asp
725 730 735
Arg Lys Lys Leu Thr Asn Gln Ile Thr Glu Asn Glu Asn Thr Asn Leu
740 745 750
Arg Ile Lys Ala Leu Leu Ala Glu Met Met Asn Asp His Gly Val Glu
755 760 765
Asn Val Arg Pro Phe Ser Pro Gln Gln Gln Glu Ile Leu Lys Ile Tyr
770 775 780
Glu Asp Gly Val Leu Asn Ser Asp Ile Glu Ile Glu Asp Glu Tyr Leu
785 790 795 800
Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg Tyr
805 810 815
Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile
820 825 830
Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile
835 840 845
Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile
850 855 860
Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Gly
865 870 875 880

Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile Glu Leu Gly Phe Gly Lys
885 890 895
Ser Val Lys Val Phe Asp Ile Asp Glu Tyr Glu Asp Phe Val Lys Lys
900 905 910
His Tyr Ala Asn Asn Arg Ser Lys Arg Asn Lys Leu Leu Leu Glu Asp
915 920 925
Ile Pro Glu Lys Met Ile Glu Arg Gln Ile Asn Asp Thr Arg Tyr Ile
930 935 940
Ser Lys Tyr Ile Ser Gly Ile Leu Ser Asn Ile Val Arg Leu Glu Asp
945 950 955 960
Gly Ser Asp Glu Gly Ile Asn Ser Lys Asn Ile Val Pro Gly Asn Gly
965 970 975
Lys Ile Thr Thr Gln Leu Lys Lys Asp Trp Gly Leu Asn Asp Val Trp
980 985 990
Asn Asp Leu Ile Leu Pro Arg Phe Lys Arg Met Asn Gln Leu Thr Asn
995 1000 1005
Ser Thr Asp Phe Ile Ala Trp Asn Glu Asn Phe Gln Lys Phe Leu
1010 1015 1020
Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys Lys Arg
1025 1030 1035
Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile Ala Cys
1040 1045 1050
Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser Ala Lys
1055 1060 1065
Ser Glu Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys Phe
1070 1075 1080
Glu Arg Val Val Tyr Leu His Thr Gln Thr Gly Glu Lys Ile Glu
1085 1090 1095
Arg Glu Val Pro Lys His Phe Leu Lys Pro Trp Glu Asn Phe Thr
1100 1105 1110
Val Asp Val Lys His Asn Leu Asp Thr Val Ile Val Ser Phe Lys
1115 1120 1125
Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr Glu Lys
1130 1135 1140
Tyr Ala Glu Asn Gly Gly Ala Arg Asn Lys Ala Arg Val Glu Gln
1145 1150 1155
Lys Gly Ile Asn Trp Ala Ile Arg Lys Pro Met His Lys Asp Thr
1160 1165 1170
Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro Lys Gly

1175 1180 1185
Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Thr Ser Phe Asp
1190 1195 1200
Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln Lys Ile
1205 1210 1215
Leu Lys Asn Tyr Leu Met Leu Lys Asp Gly Asn Pro Glu Leu Ala
1220 1225 1230
Phe Ser Pro Glu Gly Ile Glu Asp Leu Asn Lys Ser Ile Glu Lys
1235 1240 1245
Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val Arg Val
1250 1255 1260
Phe Glu Leu Gly Ser Lys Phe Gln Ile Gly Gln Ala Gly Asn Lys
1265 1270 1275
Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe Phe
1280 1285 1290
Ala Val Tyr Glu Asp Lys Asn Gly Lys Arg Ser Tyr Glu Thr Ile
1295 1300 1305
Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Pro Val
1310 1315 1320
Val Asp Leu Arg Ser Thr Asn Asp Phe Tyr Leu Cys Pro Asn Asp
1325 1330 1335
Leu Val Tyr Ile Leu Ser Asp Asp Glu Arg Gly Asn Met Thr Asn
1340 1345 1350
Asn Asp Phe Glu Asn Leu Ser Asp Glu Gln Val Lys Arg Ile Tyr
1355 1360 1365
Lys Phe Val Ser Cys Thr Gly Gly Glu Gly His Phe Ile Pro Tyr
1370 1375 1380
Ala Asn Ala Thr Glu Ile Ile Lys Asn Glu Asn Gly Thr Asn Ser
1385 1390 1395
Lys Ser Glu Arg Met Gln Asn Phe Tyr Asp Gly Ser Cys Met Leu
1400 1405 1410
Asp Lys Asn Ser Lys Pro Ile Met Ile Lys Glu Asn Cys Val Lys
1415 1420 1425
Leu Lys Val Asp Arg Leu Gly Asn Ile Ser Lys Ile
1430 1435 1440

<210> 119
<211> 1071
<212> PRT
<213> 芽孢杆菌属物种

<223> APG06877.1_RGN

<400> 119

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

290 295 300
Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320
Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335
Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350
Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
355 360 365
Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
370 375 380
Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400
Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415
Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430
Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445
Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
450 455 460
Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Ile Ile Pro
465 470 475 480
Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495
Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510
His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525
Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540
Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560
Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575
Ser Leu Lys Lys Ile Ser Ala Asn Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590
Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620
Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser
625 630 635 640
Glu Glu Asp Lys Glu Trp Glu Ser Phe Glu Ser Tyr Val Arg Ser Asn
645 650 655
Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670
Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685
Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700
Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val
705 710 715 720
Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735
Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750
Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765
Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro
770 775 780
Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
785 790 795 800
Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Ser
805 810 815
Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser
820 825 830
Ala His Asp Gln Thr Val Met Lys Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845
Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Glu Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910
Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val

915 920 925
Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975
Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
1025 1030 1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
1040 1045 1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
1055 1060 1065
Ile Phe Lys
1070

<210> 120
<211> 1091
<212> PRT
<213> 根瘤菌属物种
<223> APG09053.1_RGN
<400> 120

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

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

405 410 415
Gly Arg Leu Gly Glu Thr Ala Leu Ser Ser Met Leu Asp Glu Met Lys
420 425 430
Ala Ala Val Ile Pro Glu Ala Glu Ala Ala Lys Asn Cys Gly Tyr Asp
435 440 445
His Ser Lys Leu Gly Glu Asp Arg Glu Glu Gly Glu Ala Phe Leu Pro
450 455 460
Pro Tyr Gln Glu Ile Leu Ser Arg Gln Ile Pro Pro Gly Thr Asn Asp
465 470 475 480
Pro Ala Asp Ile Tyr Asp Ile Arg Met Gly Arg Phe Thr Asn Pro Thr
485 490 495
Val His Ile Gly Leu Asn Gln Leu Arg Arg Val Val Asn Ala Leu Ile
500 505 510
Ala Arg His Gly Lys Pro Gln Phe Val Ser Leu Glu Leu Ala Arg Asp
515 520 525
Leu Gln Leu Ser Glu Lys Gln Lys Ala Glu Ala Asn Arg Ala Ile Ala
530 535 540
Lys Asn Thr Arg Glu Ala Ala Ala Arg Ser Lys Lys Leu Gly Glu Met
545 550 555 560
Gly Gln Leu Asp Thr Gly Tyr Asn Arg Leu Leu Leu Lys Leu Trp Glu
565 570 575
Glu Leu Asn Pro Ser Lys Pro Glu Asp Arg Val Cys Ile Tyr Ser Gly
580 585 590
Lys Pro Ile Gly Ile Asp Met Leu Phe Ser Ala Glu Val Asp Val Asp
595 600 605
His Ile Leu Pro Trp Ser Lys Thr Leu Asp Asp Ser Gln Ala Asn Lys
610 615 620
Leu Leu Cys Leu Lys Ser Ala Asn Arg Gln Lys Arg Asn Arg Ala Pro
625 630 635 640
Ala Asp Val Pro Glu Trp Arg Asp Arg Tyr Glu Glu Ile Leu Ala Arg
645 650 655
Ala Ala Arg Leu Ala Pro Asn Lys Arg Arg Arg Phe Ala Pro Asn Ala
660 665 670
Met Ala Glu Phe Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu Thr
675 680 685
Asp Thr Gln Tyr Leu Ser Arg Met Ala Arg Glu Tyr Leu Ser Cys Leu
690 695 700
Phe Pro Ser Ala Glu Ile Asp Arg Arg Gly Glu Arg Lys Lys Lys Ile
705 710 715 720

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

1025 1030 1035

Ala Arg Asp Ala Asp Lys Gln Asp Pro Phe Lys Tyr Phe Ser Lys

1040 1045 1050

Ser Ala Ala Gly Leu Lys Asp Ile Gln Cys Arg Gln Ile Arg Ile

1055 1060 1065

Asp Ala Ala Gly Arg Val Phe Asp Pro Gly Pro Gln Asp Arg Ala

1070 1075 1080

Ser Lys Ser Thr Arg Lys Thr Asn

1085 1090

<210> 121

<211> 1401

<212> PRT

<213> Myroides物种

<223> APG04293.1_RGN

<400> 121

Met Lys Arg Ile Leu Gly Ile Asp Leu Gly Thr Thr Ser Ile Gly Trp

1 5 10 15

Ala Tyr Val His Glu Ala Glu Asn Glu Asn Glu Ile Ser Ser Ile Glu

20 25 30

Lys Leu Gly Val Arg Val Asn Pro Leu Thr Val Asp Glu Gln Ile Asn

35 40 45

Phe Glu Lys Gly Lys Pro Ile Thr Thr Asn Ala Thr Arg Thr Leu Phe

50 55 60

Arg Ser Ala Arg Arg Asn Leu Gln Arg Phe Lys Leu Arg Arg Lys His

65 70 75 80

Leu Ile Glu Ile Leu Lys His Asn Asn Trp Ile Ser Lys Gln Thr Leu

85 90 95

Leu Thr Glu Asn Gly Asn His Thr Thr Phe Glu Thr Ile Ser Leu Arg

100 105 110

Asp Lys Ala Ala Thr Gln Lys Val Thr Leu Glu Glu Leu Ala Arg Val

115 120 125

Leu Leu Met Ile Asn Lys Lys Arg Gly Tyr Lys Ser Ser Arg Lys Val

130 135 140

Asn Asn Thr Glu Glu Gly Gln Leu Ile Asp Gly Met Glu Val Ala Lys

145 150 155 160

Lys Leu Tyr His Glu Gln Leu Thr Pro Gly Gln Leu Ser Leu Gln Leu

165 170 175

Ile Lys Gln Gly Val Lys Arg Leu Pro Asp Tyr Tyr Arg Ser Asp Leu

180 185 190

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

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

Lys Gly Asn Gln Thr Ala Tyr Asp Phe Ile Leu Ser Glu Tyr Gly Gln
820 825 830
Glu Lys Ala Asn Glu Phe Glu Thr Arg Ile Thr Asn Leu Tyr Ser Leu
835 840 845
Gly Lys Asp Glu Gly Ile Ser Lys Ser Lys Tyr Lys Lys Leu Leu Met
850 855 860
Gln Glu Ser Glu Ile Gly Lys Gly Phe Ile Glu Arg Asp Leu Arg Glu
865 870 875 880
Thr Gln Tyr Ile Ala Lys Lys Ala Lys Ser Leu Leu Leu Glu Ile Thr
885 890 895
Arg Asn Val Val Ser Thr Ser Gly Gly Ile Thr Asp Arg Leu Arg Glu
900 905 910
Asp Trp Gly Leu Val Asn Val Met Lys Glu Leu Asn Leu Lys Lys Phe
915 920 925
Arg Asp Ala Gly Leu Thr Glu Phe Val Glu Met Lys Asp Gly Asn Lys
930 935 940
Lys Glu Val Ile Ile Asp Trp Thr Lys Arg Asn Asp His Arg His His
945 950 955 960
Ala Met Asp Ala Leu Thr Ile Ala Phe Thr Lys His Asn His Ile Gln
965 970 975
Tyr Leu Asn His Leu Asn Ala Arg Lys Asn Thr Lys Asp Glu Leu His
980 985 990
Thr Asn Ile Ile Ala Ile Glu Lys Lys Glu Thr Thr Ile Ile Thr Asp
995 1000 1005
Glu Arg Gly Asn Lys Lys Arg Ile Phe Lys Glu Pro Ile Pro Asn
1010 1015 1020
Phe Arg Asn Ile Ala Lys Ser His Leu Glu Ala Ile Leu Val Ser
1025 1030 1035
His Lys Ala Lys Asn Lys Val Val Thr Lys Asn Ile Asn Lys Ile
1040 1045 1050
Ser Gly Lys Lys Ile Gly Gln Gln Thr Leu Thr Pro Arg Gly Gln
1055 1060 1065
Leu His Lys Glu Thr Val Tyr Gly Lys Ile Arg Gln Tyr Val Ser
1070 1075 1080
Lys Glu Glu Lys Val Gly Pro Lys Phe Thr Lys Glu Val Ile Glu
1085 1090 1095
Leu Val Ser Asn Pro Thr Tyr Arg Asp Leu Leu Leu Lys Arg Leu
1100 1105 1110
Gln Glu Asn Asn Asn Asp Pro Lys Lys Ala Phe Gly Gly Lys Asn

1115 1120 1125

Ala Leu Ser Lys Ser Pro Ile Tyr Ile Asn Leu Glu Lys Asn Ile

1130 1135 1140

Ile Val Pro Glu Val Val Lys Leu Val Trp Leu Glu Asp Asp Tyr

1145 1150 1155

Thr Ile Arg Lys Glu Ile Ser Pro Glu Leu Lys Ile Asp Lys Val

1160 1165 1170

Val Asp Glu Gly Val Lys Arg Ile Leu Gln Asn Arg Leu Ala Leu

1175 1180 1185

Tyr Asn Gly Asp Ala Lys Ile Ala Phe Ser Asp Leu Asp Asn Asn

1190 1195 1200

Pro Ile Trp Leu Ser Lys Asp Lys Gly Ile Thr Ile Lys Arg Val

1205 1210 1215

Thr Leu Ser Gly Val Lys Asn Ala Thr Ala Leu His Thr Lys Lys

1220 1225 1230

Asp His Leu Gly Gln Pro Ile Leu Asp His Lys Gly Gln Glu Ile

1235 1240 1245

Pro Val Asp Phe Ile Ser Thr Gly Asn Asn His His Val Ala Ile

1250 1255 1260

Tyr Glu Asp Lys Asn Gly Lys Leu Gln Glu Ser Val Val Ser His

1265 1270 1275

Phe Glu Ala Val Glu Arg Val Asn Gln Gln Leu Pro Ile Ile Asp

1280 1285 1290

Lys Thr Phe Asn Gln His Leu Gly Trp Lys Phe Leu Phe Thr Met

1295 1300 1305

Lys Gln Asn Glu Leu Phe Leu Phe Pro Ser Asp Asp Phe Asn Pro

1310 1315 1320

Asn Glu Ile Asp Leu Phe Asp Thr Arg Asn Tyr Ala Val Ile Ser

1325 1330 1335

Lys His Leu Phe Arg Ala Gln Lys Leu Ser Thr Lys Asp Tyr Ser

1340 1345 1350

Phe Arg His His Leu Glu Thr Asn Val Glu Asp Ile Lys Asp Leu

1355 1360 1365

Lys Thr Ile Thr Trp Arg Arg Glu Gly Leu Asn Gly Ile Gln Asn

1370 1375 1380

Ile Val Lys Ile Arg Thr Asn His Leu Gly Glu Ile Val His Ile

1385 1390 1395

Gly Glu Tyr

1400

<210> 122
<211> 1432
<212> PRT
<213> 金黄杆菌属物种
<223> APG01308.1_RGN
<400> 122

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

260 265 270

Lys Asp Trp Ile Ala Ile Lys Thr Lys Thr Glu Gln Glu Ile Asp Lys

275 280 285

Ser His Lys Thr Val Gly Val Tyr Ile Tyr Glu Thr Leu Leu Gln Asn

290 295 300

Pro Lys Gln Lys Ile Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys

305 310 315 320

Phe Tyr Lys Asp Glu Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe

325 330 335

His Gln Glu Leu Lys Asn Asp Asp Leu Tyr Asn Asp Cys Val Arg Glu

340 345 350

Leu Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys Glu

355 360 365

Phe Val His Leu Leu Met Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu

370 375 380

Arg Ser Gln Lys Ser Ser Ile Ser Asn Cys Ser Leu Glu Phe Arg Lys

385 390 395 400

Tyr Lys Asp Glu Asn Gly Val Glu His Ile Gln Tyr Leu Lys Ala Val

405 410 415

Pro Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp Ile

420 425 430

Phe Asn Leu Asn Ile Tyr Lys Arg Asp Asp Glu Gln Asn Pro Val Thr

435 440 445

Thr Glu Phe Leu Asn Thr Thr Thr Asp Ile Glu Asn Leu Phe Glu Phe

450 455 460

Leu Asn Asn Arg Lys Glu Val Asp Gln Lys Ala Leu Leu Lys His Phe

465 470 475 480

Lys Leu Asn Glu Lys Thr His Arg Trp Lys Tyr Val Glu Asp Lys Lys

485 490 495

Tyr Pro Cys Asn Glu Thr Lys Ser Met Ile Ser Glu Arg Leu Lys Lys

500 505 510

Val Glu Asn Ile Ala Asn Asp Phe Leu Thr Arg Glu Ile Glu Gln Lys

515 520 525

Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys Ile Glu Tyr Glu Lys

530 535 540

Ala Leu Lys Ser Phe Ala Lys Lys Asn Asn Leu Asp Glu Asn Ser Phe

545 550 555 560

Phe Glu Ala Phe Lys Lys Phe Pro Pro Phe Lys Ser Glu Tyr Gly Ser

565 570 575

Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Leu Gly
580 585 590
Lys Tyr Trp Asn Cys Glu Asn Ile Ser Asp Asp Ser Lys Glu Arg Ile
595 600 605
Gln Lys Ile Ile Asn Gly Glu Tyr Asp Glu Asn Ile Lys Asp Arg Val
610 615 620
Arg Glu Lys Ala Ile His Leu Thr Ser Glu Asn Asn Phe Gln Gly Leu
625 630 635 640
Gln Leu Trp Leu Ala Gln Tyr Val Val Tyr Asp Arg His Ser Glu Ala
645 650 655
Ser Met Ile Gly Lys Trp Asn Ser Ala Asp Asp Leu Glu Glu Phe Leu
660 665 670
Lys Glu Phe Lys Gln His Ser Leu Arg Asn Pro Ile Val Glu Gln Val
675 680 685
Ile Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp Leu Lys Tyr Gly
690 695 700
Asn Gly Ala Lys Asp Phe Phe Asn Glu Ile His Ile Glu Leu Gly Arg
705 710 715 720
Glu Met Lys Gln Thr Lys Asp Glu Arg Leu Asp Ala Thr Lys Arg Ile
725 730 735
Thr Glu Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala Leu Leu Ala Glu
740 745 750
Met Met Asn Asp Asn Ser Val Glu Asn Val Arg Pro Tyr Ser Pro Met
755 760 765
Gln Gln Glu Ile Leu Lys Ile Tyr Glu Asp Gly Val Leu Asn Ser Gly
770 775 780
Ile Glu Ile Glu Asp Glu Tyr Leu Lys Ile Ser Lys Thr Ala Gln Pro
785 790 795 800
Ser Pro Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu Gln Arg Tyr
805 810 815
Lys Ser Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys Leu Phe Thr
820 825 830
Pro Glu Tyr Glu Ile Glu His Ile Ile Pro Gln Ser Arg Tyr Phe Asp
835 840 845
Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala Val Asn Lys Leu
850 855 860
Lys Asp Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe Gly Gly Thr
865 870 875 880
Ile Val Glu Cys Gly Leu Gly Lys Asn Val Lys Val Phe Glu Val Asn

885 890 895
Glu Tyr Glu Asp Phe Val Lys Lys His Tyr Ala Asn Asn Arg Gly Lys
900 905 910
Arg Asn Lys Leu Leu Leu Glu Glu Ile Pro Glu Lys Met Ile Glu Arg
915 920 925
Gln Leu Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser Gly Val Leu
930 935 940
Ser Asn Ile Val Arg Val Glu Asp Gly Ser Asp Glu Gly Val Asn Ser
945 950 955 960
Lys Asn Ile Val Pro Gly Asn Gly Lys Ile Thr Thr Gln Leu Lys Gln
965 970 975
Asp Trp Gly Leu Asn Asp Val Trp Asn Asp Leu Ile Leu Pro Arg Phe
980 985 990
Glu Arg Met Asn Gln Leu Thr Asn Ser Thr Asp Phe Thr Ala Trp Asn
995 1000 1005
Glu Asn His Gln Lys Phe Leu Pro Thr Val Pro Ile Glu Phe Ser
1010 1015 1020
Lys Gly Phe Ser Lys Lys Arg Ile Asp His Arg His His Ala Leu
1025 1030 1035
Asp Ala Leu Val Ile Ala Cys Ala Thr Lys Asp His Ile Asn Leu
1040 1045 1050
Leu Asn Asn Gln Ser Ala Lys Ser Asp Thr Lys Arg Tyr Asp Leu
1055 1060 1065
Lys Lys Lys Leu Met Lys Phe Glu Lys Gly Val Tyr Asn His Pro
1070 1075 1080
Gln Thr Gly Glu Arg Ile Gln Arg Asp Val Pro Lys Gln Phe Leu
1085 1090 1095
Lys Pro Trp Glu Ser Phe Thr Ile Asp Ala Lys Asn Asn Leu Asp
1100 1105 1110
Lys Ile Ile Ile Ser Phe Lys Gln Asn Leu Arg Val Ile Asn Lys
1115 1120 1125
Ala Thr Asn Tyr Tyr Glu Lys Tyr Val Glu Lys Asn Gly Ile Lys
1130 1135 1140
Thr Lys Glu Arg Val Glu Gln Thr Gly Thr Asn Trp Ala Ile Arg
1145 1150 1155
Lys Pro Met His Lys Glu Thr Val Ser Gly Ile Val Asn Leu Pro
1160 1165 1170
Trp Val Lys Val Pro Lys Gly Lys Ile Leu Thr Ala Thr Arg Lys
1175 1180 1185

Ser Leu Asp Thr Thr Phe Asp Leu Lys Ser Ile Asn Ser Ile Thr
1190 1195 1200
Asp Thr Gly Ile Gln Lys Ile Leu Arg Asn Tyr Leu Glu Phe Lys
1205 1210 1215
Gly Ser Pro Glu Leu Ala Phe Ser Pro Glu Gly Ile Glu Asp Met
1220 1225 1230
Asn Lys Asn Ile Glu Lys Tyr Asn Asp Gly Lys Leu His Gln Pro
1235 1240 1245
Ile Asn Lys Val Arg Val Phe Glu Leu Gly Ser Lys Phe Gln Val
1250 1255 1260
Gly Gln Thr Gly Asn Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys
1265 1270 1275
Gly Thr Asn Leu Phe Phe Ala Val Tyr Glu Asp Lys Asn Gly Lys
1280 1285 1290
Arg Asn Tyr Glu Thr Ile Pro Leu Asn Ile Val Ile Glu Arg Gln
1295 1300 1305
Lys Gln Gly Leu Leu Ala Cys Pro Glu Ser Asn Glu Lys Gly Glu
1310 1315 1320
Lys Leu Leu Phe Gln Leu Ser Pro Asn Asp Phe Val Tyr Leu Thr
1325 1330 1335
Thr Glu Glu Glu Asn Asp Asn Ala Thr Pro Ile Asn Phe Ser Leu
1340 1345 1350
Leu Ser Lys Glu Gln Ile Asn Asn Leu Tyr Lys Ile Val Ser Phe
1355 1360 1365
Thr Gly Asn Arg Leu Tyr Gly Ile Pro Ile Cys Val Ala Thr Thr
1370 1375 1380
Ile Val Asn Lys Ala Glu Tyr Thr Gln Leu Asn Lys Ile Glu Phe
1385 1390 1395
Thr Lys Glu Lys Asp Leu Leu Leu Lys Leu Asn Val Asn Arg Leu
1400 1405 1410
Gly Asp Val Lys Thr Phe Thr Ala Asn Asp Ile Arg Lys Ile Phe
1415 1420 1425
Asn Arg Gln Glu
1430

<210> 123
<211> 1092
<212> PRT
<213> 芽孢杆菌属物种
<223> APG06646.1_RGN

<400> 123

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

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

610 615 620
Phe Gly His Glu Glu Asn Arg Trp Glu Asp Phe Val Ser Phe Ile Arg
625 630 635 640
Thr Met Lys Val Gly Lys Lys Lys Lys Asn Met Leu Leu Lys Lys Asn
645 650 655
Phe Asp Glu Glu Gln Glu Glu Gln Ile Leu Ser Arg Asn Leu Asn Asp
660 665 670
Thr Arg Tyr Ile Thr Arg Tyr Ile Lys Ser Phe Ile Glu Asp Asn Leu
675 680 685
Glu Phe Arg Thr Glu Glu Asn Lys Glu Gln Tyr Val His Thr Val Asn
690 695 700
Gly Ala Tyr Thr Ser Leu Met Arg Lys Arg Trp Gly Leu Asn Lys Asp
705 710 715 720
Arg Arg Gly Asn Asp Leu His His Ala Val Asp Ala Ala Ile Ile Ala
725 730 735
Val Ser Leu Pro Phe Lys Asn Lys Val Asn Ala Tyr Phe Lys Arg Gln
740 745 750
Glu Thr Gly Leu Ser Lys Leu Leu Asn Asn Lys Lys Asp Ile Phe Pro
755 760 765
Glu Pro Trp Arg Asn Phe Ile Lys Glu Leu Glu Ala Arg Met Ile Gln
770 775 780
Asp Pro Glu Lys Met Lys Arg Ala Leu Glu Ser Leu Glu Leu Glu Thr
785 790 795 800
Tyr Gly Glu Ile Phe Leu Asn Lys Leu Lys Pro Ile Phe Val Ser Arg
805 810 815
Met Pro Lys His Ser Ile Lys Gly Gln Ile His Glu Glu Thr Ile Arg
820 825 830
Arg Val Arg Gly Phe Thr Glu Glu Gly Phe Leu Val Thr Val Lys Lys
835 840 845
Thr Arg Leu Asp Gln Ile Pro Phe Asp Lys Asn Gly Asp Phe Pro Met
850 855 860
Tyr Gly Lys Glu Thr Asp Ile Lys Thr Tyr Met Ala Ile Lys Gln Arg
865 870 875 880
Tyr Leu Glu Tyr Gly Gln Asp Lys Gln Lys Ala Phe Ala Val Pro Leu
885 890 895
Arg Lys Pro Ser Lys Asn Pro Lys Asn Ala Pro Ile Val Arg Ser Val
900 905 910
Lys Ile Glu Gly Lys Ala Asn Arg Val Val Met Leu Asp Asp Lys Ala
915 920 925

Ala Ala Asp Asn Ala Ser Ile Val Arg Thr Glu Val Phe Arg His Lys
930 935 940
Lys Thr Gly Glu Tyr Tyr Leu Thr Pro Val Tyr Val Ala Asp Ile Leu
945 950 955 960
Ser Asn Lys Ile Pro Asp Arg Leu Ile Thr Ile Lys Lys Ser Tyr Ser
965 970 975
Asp Trp Asp Arg Ile Thr Asp Glu His Glu Tyr Leu Phe Ser Leu Tyr
980 985 990
Asn Asn Asp Leu Val Lys Ile Ile Leu Pro Lys Glu Lys Glu Thr Lys
995 1000 1005
Lys Tyr Thr Gly Gly Asn His Leu Trp Gln Glu Gly Phe Phe Tyr
1010 1015 1020
Phe Lys Gly Val Asp Ser Ser Asn Ala Gly Ile Lys Ile Ile Asn
1025 1030 1035
His Leu Asn Ser Phe Glu Ala Arg Ile Gly Thr Lys Arg Leu Ile
1040 1045 1050
Ala Phe Glu Lys Tyr Gln Val Asn Pro Leu Gly Glu Ile Asn Lys
1055 1060 1065
Val His Gly Glu Lys Arg Pro Gly Glu Leu Leu Asn Lys Glu Glu
1070 1075 1080
Ile Lys Glu Asn Arg Lys Asn Ile Ser
1085 1090

<210> 124
<211> 1735
<212> PRT
<213> 土地杆菌属物种
<223> APG06622_RGN
<400> 124

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

85 90 95

Lys Ile Glu Val Asp Phe Glu Arg Asn Asp Lys Arg Cys Gly Gln Phe
100 105 110

Lys Glu Gly Lys Glu Pro Lys Met Ala Tyr Leu Pro Thr Lys Asn Leu
115 120 125

Glu Asn Lys Phe Asp Phe Tyr Phe Glu Glu Ala Phe Asn Glu Met Ile
130 135 140

Glu Asp Leu Gln Lys Val Asn Pro Glu Ile Lys Asn Glu Lys Lys Lys
145 150 155 160

Arg Val Pro Lys Asp Trp Thr Ile Tyr Phe Leu Arg Asn Lys Ala Met
165 170 175

Ser Gln Lys Val Ser Leu Glu Glu Leu Ala Trp Val Leu Leu Ser Tyr
180 185 190

Asn Gln Lys Arg Gly Thr Asp Ser Glu Glu Ile Glu Asn Glu Asp Glu
195 200 205

Lys Ser Glu Ile Lys Glu Gln Ile Asp Leu Lys Val Ile Lys Ala Glu
210 215 220

Lys Lys Val Asp Val Asn Gly Ala Phe Phe Glu Ile Thr Leu Asn Asp
225 230 235 240

Ala Glu Asn Phe Lys Tyr Lys Glu Tyr Thr Thr Lys Gln Leu Thr Phe
245 250 255

Glu Asp Asp Val Lys Glu Val Thr Lys Ile Leu Lys Leu Asp Glu Ala
260 265 270

Gly Asn Ile Ile Asn Glu Lys Thr Ala Tyr Gln Ile Ser Asp Leu Tyr
275 280 285

Asn Leu Thr Leu Ala Asn Ile Glu His Thr Asp Ile Asn Asp Lys Lys
290 295 300

Ala Lys His Lys Tyr Asp Phe Ile Tyr Thr Asn Gly Trp Cys Thr Glu
305 310 315 320

Lys Lys Lys Glu Asn Phe Asp Val Thr Tyr Gln Lys Leu Glu Lys Gln
325 330 335

Ile Lys Glu Glu Ser Lys Val Ile Thr Glu Leu Phe Val Val Ser Asn
340 345 350

Lys Tyr Asp Phe Ala Gly Lys Pro Glu Ser Ile Ile Pro Ser Ile Lys
355 360 365

Leu Pro Asp Phe Asn Ser Glu Gly Ser Lys Asp Trp Thr Leu Leu Lys
370 375 380

Lys Lys Thr Glu Lys Glu Ile Val Lys Ser Glu Lys Thr Leu Ser Gly
385 390 395 400

Trp Ile Tyr Asp Val Leu Lys Asn Asp Val Asn Thr Gly Glu Arg Thr
405 410 415
Lys Ile Ile Gly Gly Leu Phe Gln Thr Ile Asp Arg Lys Phe Tyr Arg
420 425 430
Glu Glu Leu Tyr Gln Ile Ile Asn Thr Gln Arg Lys Phe His Thr Ser
435 440 445
Thr Leu Asp Asp Lys Asp Thr Phe Glu Lys Cys Val Lys Leu Leu Tyr
450 455 460
Pro His Asn Lys Asp His Ser Lys Thr Leu Leu Glu Asn Lys Gln Ala
465 470 475 480
Leu Thr Asn Leu Leu Ile Glu Asp Ile Leu Leu Tyr Gln Arg Pro Leu
485 490 495
Lys Ser Lys Lys Ser Glu Ile Ser Asp Cys Lys Tyr Glu Ile Asp Tyr
500 505 510
Trp Lys Glu Asp Val Asn Ile Ser Thr Gly Glu Val Val Glu Ile Pro
515 520 525
Val Tyr Lys Lys Ala Ile Phe Ala Ser His Pro Leu Phe Gln Glu Phe
530 535 540
Arg Ile Trp Asp Lys Ile His Asn Ile Lys Leu Ile Gln Leu Glu Ser
545 550 555 560
Lys Asp Ala Asp Gly Asn Thr Gln Thr Asn Val Asp Ile Thr Ser Gln
565 570 575
Tyr Phe Lys Pro Glu Asn Tyr Arg Ala Leu Phe Asn His Phe Asn Gly
580 585 590
Arg Ser Thr Val Ala Ile Asp Asp Phe Leu Lys Phe Cys Glu Glu Gln
595 600 605
Phe Lys Leu Gly Ile Gly Lys Lys Gly Glu Arg Lys Ile Leu Trp Asn
610 615 620
Tyr Pro Val Glu Glu Glu Phe Lys Gly Asn Glu Thr Arg Lys Gly Phe
625 630 635 640
Glu Asn Arg Phe Arg Lys Cys Gly Phe Asp Gly Phe Asp Asp Phe Met
645 650 655
Thr Gln Gln Lys Glu Ile Glu Leu Trp His Tyr Leu Tyr Ser Val Ser
660 665 670
Ser Ala Glu Arg Lys Lys Val Ser Gln Val Lys Asp Ser Ser Asp Pro
675 680 685
Lys Phe Gly Lys Thr Gly Ile Trp Asn Phe Phe Thr Lys Tyr Phe Lys
690 695 700
Asp Glu Asn Ile Ser Leu Glu Thr Leu Glu Lys Leu Cys Arg Asp Phe

705 710 715 720
Glu Thr Tyr Pro Lys Phe Thr Ser Lys Tyr Ala Ser Tyr Ser Val Lys
725 730 735
Ala Leu Asn Lys Met Leu Ala Val Met Arg Val Gly Lys Asn Phe Leu
740 745 750
Thr Ala Glu Ser Val Asn Glu Lys Trp Gln Glu Lys Tyr Ile Glu Arg
755 760 765
Ala Asp Thr Ile Ile Lys Lys Asn Asn Glu Ile Asp Trp Ser Ala Glu
770 775 780
Ser Ile Asp Leu Thr Asp Ile Ile Leu Asn Asp Val Asn Val Lys Asn
785 790 795 800
Gly Glu Ile Pro Phe Pro Lys Gly Leu Phe Asn Thr Phe Lys Ser Phe
805 810 815
Glu Thr Val Asp Asp Phe Glu Phe Leu Asn Leu Thr Gln Ala Ser Tyr
820 825 830
Phe Ile Tyr Gly Arg His Ser Glu Leu Ala Gln Ala Lys Tyr Trp Asn
835 840 845
Ser Pro Thr Lys Ile Arg Glu Glu Ile Thr Lys Glu Leu Lys His His
850 855 860
Ser Leu Asn Asn Pro Thr Ala Glu Lys Val Leu Lys Glu Thr Leu Lys
865 870 875 880
Val Val Ala Asp Ile Trp Glu Ile Tyr Gly Asn Ser Glu Glu Lys Phe
885 890 895
Phe Asn Glu Ile His Leu Glu Val Ala Arg Glu Leu Gln Lys Ser Asn
900 905 910
Gln Glu Lys Lys Asp Ile Ile Glu Arg Gln Lys Asn Asn Arg Ala Glu
915 920 925
Asn Asn Arg Ile Arg Asn Ile Leu Glu Glu Phe Leu Ser Ser Ser Pro
930 935 940
Tyr Asn Ala Lys Arg Gly Asn Gln Asp His Phe Glu Arg Leu Lys Ile
945 950 955 960
Val Glu Asp Gly Ala Lys Val Arg Ser Tyr Ser Asp Lys Glu Phe Tyr
965 970 975
Ser Asn Asn Lys Asp Lys Phe Ser Lys Ile Asp Ile Glu Val Ile Leu
980 985 990
Lys Lys Lys Lys Ile Thr Gln Ser Asp Phe Glu Lys Tyr Lys Leu Trp
995 1000 1005
Ile Glu Gln Gly Tyr Lys Ser Pro Tyr Thr Asp Gln Ile Ile Lys
1010 1015 1020

Leu Thr Asp Leu Phe Asp Gly Asn Lys Tyr Asn Val Asp His Ile
1025 1030 1035
Phe Ser Arg Ala Ala Ile Thr Asn Asp Ser Leu Asn Asn Lys Val
1040 1045 1050
Val Cys Glu Ala Val Ile Asn Arg Phe Lys Ser Tyr Lys Thr Gly
1055 1060 1065
Arg Glu Phe Ile Ala Gln Phe Gly Gly Lys Thr His Thr Ile Phe
1070 1075 1080
Asp Asp Lys Leu Asn Lys Asn Val Thr Phe Ser Leu Ile Asp Glu
1085 1090 1095
Asp Ala Tyr Val Ser Leu Val Lys Ser Gln Phe Lys Asp Ala Lys
1100 1105 1110
Arg Leu Ile Leu Leu Ser Lys Glu Val Pro Lys Ala Phe Thr Glu
1115 1120 1125
Ser Gln Leu Asn Asn Ala Lys Tyr Ile Ala Arg Lys Ala Met Glu
1130 1135 1140
Leu Leu Ser His Ile Val Arg Glu Glu Gly Glu Val Glu Phe Arg
1145 1150 1155
Ser Lys Asn Val Leu Pro Val Ser Gly Ala Ile Thr Asp Arg Leu
1160 1165 1170
Lys Lys Glu Trp Lys Ile Asn Gln Ile Trp Thr Glu Leu Leu Lys
1175 1180 1185
Pro Arg Phe Glu Arg Leu Asn Lys Ile His Asn Ser Glu Ala Phe
1190 1195 1200
Gly Ser Tyr Ala Ile Ser Lys Ser Gly His Gln Tyr Phe Asp Ile
1205 1210 1215
Asn Thr Lys Tyr Ile Leu Gly Glu Asn Asp Lys Phe Asp Leu Lys
1220 1225 1230
Arg Leu Asp His Arg His His Ala Leu Asp Ala Leu Ile Ile Ala
1235 1240 1245
Leu Cys Thr Asp Asn His Val Gln Tyr Phe Asn Asn Ile Asn Ser
1250 1255 1260
Gly Ile Thr Asn Lys Lys Lys Gly Lys Met Glu Ala Ile Lys Lys
1265 1270 1275
Gln Arg Ala Gly Ile Lys Arg Gln Ile Met Tyr Ser Glu Lys Asp
1280 1285 1290
Lys Glu Asn Pro Asn Glu Lys Val Trp Arg Tyr Met Leu Pro Gly
1295 1300 1305
Ser Phe Arg Asn Lys Glu Ser Val Asn Asn Glu Lys Asp Ser Val

1310 1315 1320
Val Asp Leu Ser Trp Glu Asn Asn Tyr Ser Ala Lys Glu Ser Asn
1325 1330 1335
Asp Tyr Arg Lys Thr Ile Leu Glu Cys Leu Asp Asn Cys Ile Val
1340 1345 1350
Ser Phe Lys Asn Asp Phe Lys Ile Val Ser Lys Ser Ser Asn Lys
1355 1360 1365
Tyr Glu Ser Tyr Tyr Asp Glu Asn Gly Asn Leu Arg Leu Asp Ser
1370 1375 1380
Asn Gly Lys Pro Gln Lys Glu Phe Ile Ser Gln Gln Asn Asn Asn
1385 1390 1395
Gln Lys His Trp Ser Val Arg Lys Pro Leu His Thr Asp Asn Pro
1400 1405 1410
Ser Ala Glu Ile Thr Leu Gln Phe Asp Arg Leu Lys Leu Ile Asp
1415 1420 1425
Asn Leu Gly Lys Val Asp Leu Ile Ile Glu Glu Lys Ile Lys Asp
1430 1435 1440
Ala Val Lys Glu Ile Leu Gly Met His Asn Asn Lys Ile Gly Lys
1445 1450 1455
Ala Glu Ser Tyr Leu Lys Lys Asn Pro Leu Ile Leu Glu Glu Lys
1460 1465 1470
Lys Ile Glu Tyr Thr Asp Phe Lys Ile Ser Asn Lys Lys Tyr Arg
1475 1480 1485
Lys Arg Gln Pro Ile Thr Asn Leu Ser Val Arg Thr Gly Thr Gly
1490 1495 1500
Ala Leu Pro Thr Tyr Asp Ser Val Val Asp Arg Ile Asn Lys Val
1505 1510 1515
Ser Asp Phe Val Leu Arg Asn Glu Leu Leu Gln His Leu Lys Ile
1520 1525 1530
Asn Asp Asn Asp Ile Asp Lys Ala Phe Ser Ile Glu Gly Ile Asn
1535 1540 1545
Glu Phe Asn Ser Asn Arg Lys Ile Pro Val Phe Arg Leu Pro Ile
1550 1555 1560
Ser Glu Ser Ala Thr Leu Lys Phe Lys Ile Gly His Lys Asn Ser
1565 1570 1575
Asn Ser His Lys Trp Val Glu Thr Gly Gly Asn Phe His Phe Arg
1580 1585 1590
Ile Ala Asp Glu Glu Asn Ser Arg Asp Tyr Glu Thr Ile Pro Leu
1595 1600 1605

Arg Ile Ala Ile Glu Met Glu Lys Asp Asn Ile Ile Arg Arg Gln
1610 1615 1620
Ile Val Asp Asn Glu Asn Glu Ile Val Leu Ser Pro Asn Asp Leu
1625 1630 1635
Val Tyr Val Pro Leu Glu Lys Asn Glu Ile Val Asp Phe Asn Asn
1640 1645 1650
Leu Thr Ile Gln Gln Ile Lys Arg Ile Tyr Lys Phe Val Phe Ser
1655 1660 1665
Thr Lys Lys Glu Ala His Phe Ile Pro Tyr Ser Ser Ser Ala Glu
1670 1675 1680
Ile Ile Lys Asn Glu Asn Gly Thr Asn Gly Tyr Ser Glu Arg Thr
1685 1690 1695
Lys Asp Phe Phe Asp Gly Asn Val Leu Gln Asp Asp Lys Gly Lys
1700 1705 1710
Pro Ile Gln Val Lys Ser Val Cys Trp Lys Leu Lys Val Asp Arg
1715 1720 1725
Leu Gly Asn Ile Ser Lys Ala
1730 1735

<210> 125
<211> 1224
<212> PRT
<213> Chitinophaga物种
<223> APG02787_RGN
<400> 125

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

115 120 125

Arg Arg Gly Tyr Arg Ser Ser Arg Ser Glu Glu Ser Ser Asp Gly Lys

130 135 140

Lys Asn Thr Ala Tyr Val Glu Glu Val Lys Gly Arg Tyr Glu Glu Leu

145 150 155 160

Lys Thr Arg Gly Ile Thr Ile Gly Gln His Phe Tyr Gln Lys Leu Gln

165 170 175

Thr Ala His Asn Arg Asn Glu Pro Tyr Arg Ile Lys Gln Gln Val Phe

180 185 190

Pro Arg Gln Ala Tyr Ile Asp Glu Phe Asp Ala Ile Ile Asn Gln Gln

195 200 205

Lys Asn His His Ser Ser Leu Thr Ala Asp Glu Ile Ala Phe Leu Arg

210 215 220

Asp Glu Ile Ile Tyr Met Gln Arg Pro Leu Lys Ser Gln Lys Gly Leu

225 230 235 240

Val Ser Ile Cys Glu Phe Glu Gly Gln Tyr Arg Lys Asn Ala Lys Asp

245 250 255

Ala Glu Val Phe Val Gly Pro Lys Val Ala Pro Arg Ser Ser Pro Leu

260 265 270

Ala Gln Val Cys Lys Ile Trp Glu Ser Ile Asn Asn Ile Thr Phe Thr

275 280 285

Asn Lys Arg Arg Glu Pro Tyr Pro Ile Asn Thr Asp Gln Lys Leu Arg

290 295 300

Leu Phe Asp Ala Leu Asn Lys Thr Glu Arg Leu Ser Gln Lys Ala Ile

305 310 315 320

Phe Phe Ile Leu Gly Ile Lys Pro Gly Asp Gly Tyr Thr Gly Asn Lys

325 330 335

Gln Thr Ala Lys Gly Ile Gln Gly Asn Leu Thr Asn Ala Ala Ile Ser

340 345 350

Thr Ile Ile Pro Glu Arg Lys Asp Leu Leu Asn Phe Glu Leu Lys Ile

355 360 365

Val Val Asp Glu Glu Asp Ala His Leu Ile Asn Arg Ser Asn Gly Glu

370 375 380

Ile Ile Ser Thr Val Gln Lys Lys Thr Val Ser Ala Asp Tyr Glu Arg

385 390 395 400

Ser Pro Phe Phe Arg Leu Trp His Val Ile Tyr Ser Ile Ser Asp Ile

405 410 415

Asp Lys Cys Ile Glu Thr Leu Met Glu Lys Phe Gly Leu Thr Lys Glu

420 425 430

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

740 745 750

Leu Met Asn Leu Gln Leu Pro Val Tyr Arg Asp Leu Gly Leu Thr Glu

755 760 765

Trp Lys Glu Tyr Thr Asp Arg Asp Gly Asn Leu His Arg Gln Glu Val

770 775 780

Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala Ile Asp

785 790 795 800

Ala Leu Thr Ile Ala Cys Thr Arg Gln Gly Phe Ile Gln Arg Ile Asn

805 810 815

Arg Leu Asn Ala Thr Asn Asn Arg Glu Glu Met Arg Ala Ser Leu Asn

820 825 830

Gly Asn Tyr Asp Pro Lys Lys Asn Leu Leu Glu Asn Tyr Ile Phe Ala

835 840 845

Leu Lys Pro Phe Ser Thr Lys Gln Val Glu Val Ala Ala Ala Gln Ile

850 855 860

Leu Ile Ser Phe Lys Pro Gly Lys Arg Val Ala Thr Leu Ser Lys Leu

865 870 875 880

Lys Ala Lys Gly Lys Asn Ser Gly Thr Gly Val Ile Val Pro Arg Gly

885 890 895

Pro Leu Ser Glu Glu Ser Val Tyr Gly Arg Val Arg Leu Met Asp Ser

900 905 910

Lys Lys Gln Pro Val Arg Tyr Leu Phe Glu His Pro Glu Leu Ile Phe

915 920 925

Lys Pro Tyr Ile Arg Val Leu Val Glu Glu Arg Leu Ala Ser Tyr Ser

930 935 940

Gln Asp Leu Lys Lys Ala Leu Ala Ser Leu Lys Lys Asp Pro Ile Tyr

945 950 955 960

Leu Asp Gly Thr Lys Thr Ile Ile Leu Glu Trp Gly Ser Cys Tyr Lys

965 970 975

Asp Glu Val Val Ile Lys Tyr Pro Leu Thr Asp Ile Lys Ala Lys Asp

980 985 990

Thr Gln Phe Ile Ile Asp Gly His Ile Arg Ala Ile Val Asn Glu Arg

995 1000 1005

Leu Ser Gln Tyr Lys Gly Asn Glu Lys Lys Ala Phe Glu His Pro

1010 1015 1020

Leu Tyr Ser Asp Ala Ala Cys Thr Lys Glu Ile Arg Ser Val Arg

1025 1030 1035

Cys Leu Thr Gly Leu Thr Ala Val Val Ser Val Lys Lys Asp Glu

1040 1045 1050

Ser Gly Arg Glu Ile Gly Phe Val Lys Pro Gly Asn Asn His His
1055 1060 1065
Ile Ala Leu Tyr Thr Asp Ala Glu Gly Gln Lys Arg Glu His Leu
1070 1075 1080
Cys Thr Phe Trp His Ala Val Glu Arg Lys Lys Trp Gly Met Pro
1085 1090 1095
Ala Val Ile Thr Gln Pro Gly Val Ile Trp Asp Asn Ile Leu Gln
1100 1105 1110
Lys Gln Asp Asn Phe Pro Gln Ala Phe Leu Glu Lys Leu Pro Leu
1115 1120 1125
Glu Ala Trp Thr Phe Asp Lys Ser Leu Gln Gln Asn Glu Met Phe
1130 1135 1140
Ile Val Gly Leu Thr Ile Glu Glu Val Gln Thr Cys Ile Asp Ser
1145 1150 1155
Phe Asp Tyr Ala Arg Leu Ser Lys His Leu Tyr Arg Ile Gln Lys
1160 1165 1170
Leu Ser Ser Thr Gly Tyr Lys Ile Val Leu Arg His His Leu Ala
1175 1180 1185
Thr Thr Ile Glu His Thr Arg Asp Met Ile Gly Ile Ala Ser Pro
1190 1195 1200
Ser Ser Phe Thr Gly Ile Lys Val Tyr Val Ser Lys Thr Gly Lys
1205 1210 1215
Ile Ser Met Pro His Asp
1220

<210> 126
<211> 1457
<212> PRT
<213> Mucilaginibacter物种
<223> APG06248_RGN
<400> 126

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

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

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

690 695 700

Val Glu Gln Val Val Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp
705 710 715 720
Leu Thr Tyr Gly Asn Gly Ala Lys Asn Phe Phe Asp Glu Ile His Ile
725 730 735
Glu Leu Ala Arg Glu Met Lys Asn Thr Met Glu Glu Arg Lys Arg Leu
740 745 750
Thr Ser Ile Val Thr Asp Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala
755 760 765
Leu Leu Val Glu Leu Phe Glu Asp Gly Asn Ile Glu Asn Val Arg Pro
770 775 780
Tyr Ser Pro Ile Gln His Glu Ala Leu Lys Ile Phe Glu Asp Gly Ile
785 790 795 800
Leu Asn Ser Gly Ile Glu Ile Asp Asp Asp Ile Leu Lys Ile Ser Lys
805 810 815
Thr Pro Gln Pro Thr Ser Thr Glu Leu Lys Arg Tyr Lys Leu Trp Leu
820 825 830
Glu Gln Lys Tyr Arg Ser Pro Tyr Thr Gly Asn Val Ile Pro Leu Asn
835 840 845
Arg Leu Phe Thr Thr Asp Tyr Glu Ile Glu His Ile Ile Pro Gln Ser
850 855 860
Arg Phe Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala
865 870 875 880
Ile Asn Lys Leu Lys Asp Lys Gln Leu Gly Leu Glu Phe Ile Lys Asn
885 890 895
His Ala Gly Thr Ile Val Glu Leu Gly Leu Gly Lys Thr Ala Arg Val
900 905 910
Phe Asp Glu Glu Arg Tyr Gln Asp Phe Val Lys Ala Asn Tyr Thr Lys
915 920 925
Asn Pro Gly Lys Arg Ser Lys Leu Met Leu Glu Glu Ile Pro Asp Lys
930 935 940
Met Ile Ala Arg Gln Leu Asn Asp Thr Arg Tyr Ile Ser Lys Phe Ile
945 950 955 960
Thr Ser Val Leu Ser Asn Ile Val Arg Ser Glu Thr Lys Asp Asp Gly
965 970 975
Val Asn Ser Lys Asn Val Leu Leu Cys Asn Gly Arg Ile Thr Ser Val
980 985 990
Leu Arg Gln Asp Trp Gly Met Asn Asp Ile Trp Asn Glu Leu Ile Leu
995 1000 1005

Pro Arg Phe Glu Arg Leu Asn Gln Ile Thr Asp Ser Thr Val Phe
1010 1015 1020
Thr Ala Trp Asn Glu Asn His Gln Lys Phe Leu Pro Thr Val Pro
1025 1030 1035
Met Gly Phe Ser Lys Gly Phe Gln Arg Lys Arg Ile Asp His Arg
1040 1045 1050
His His Ala Met Asp Ala Leu Ile Ile Ala Ser Cys Thr Arg Asp
1055 1060 1065
His Ile Asn Leu Leu Asn Asn Glu Tyr Ala Lys Ser Ser Lys Arg
1070 1075 1080
Phe Asp Leu Asn Arg Lys Leu Arg Lys Phe Glu Lys Val Ala Tyr
1085 1090 1095
Lys Ser Pro Val Asn Gly Gln Gln Ile Asp Lys Asp Ile Pro Thr
1100 1105 1110
Val Phe Ile Lys Pro Trp Gln Arg Phe Thr Ile Asp Ala Lys Cys
1115 1120 1125
Ser Leu Glu Asp Ile Ile Val Ser Phe Lys Gln Asn Leu Arg Val
1130 1135 1140
Ile Asn Lys Thr Thr Asn Tyr Phe Asp Gly Tyr Lys Asp Glu His
1145 1150 1155
Gly Ile Val Asn Phe Asp Asn Lys Gly Ala Pro Val Lys Gln Gln
1160 1165 1170
Arg Lys Gln Thr Lys Gly Asp Ser Trp Ala Ile Arg Lys Pro Met
1175 1180 1185
His Lys Asp Phe Val Tyr Gly Lys Val Asp Leu Lys His Ile Lys
1190 1195 1200
Ile Pro Lys Gly Lys Phe Ala Thr Ala Ile Arg Lys Ser Leu Asp
1205 1210 1215
Thr Thr Phe Asp Gln Lys Lys Ile Gln Thr Ile Thr Asp Thr Gly
1220 1225 1230
Val Gln Lys Ile Leu Ile Asn Tyr Leu Gln Ser Lys Asp Asn Lys
1235 1240 1245
Ala Glu Leu Ala Phe Ser Pro Glu Gly Ile Glu Asp Met Asn Gln
1250 1255 1260
Asn Ile Arg Ala Tyr Asn Asp Gly Lys Leu His Gln Pro Ile Lys
1265 1270 1275
Lys Val Arg Val Phe Glu Ile Gly Ser Arg Phe Pro Leu Gly Glu
1280 1285 1290
Thr Gly Asn Lys Lys Thr Lys Tyr Val Glu Thr Ala Lys Gly Thr

1295 1300 1305
Asn Leu Phe Phe Ala Ile Tyr Gln His Lys Asn Lys Arg Asn Tyr
1310 1315 1320
Glu Thr Val Pro Leu Asn Thr Val Ile Glu Arg Gln Lys Gln His
1325 1330 1335
Leu Ser Pro Val Pro Thr Glu Asp Asp Lys Gly Gly Thr Leu Leu
1340 1345 1350
Phe Tyr Leu Ser Pro Asn Asp Leu Val Tyr Val Pro Thr Pro Asp
1355 1360 1365
Glu Leu Gly Asn Asn Leu Gly Gly Leu Asp Ile Arg Asn Leu Asn
1370 1375 1380
Lys Glu Gln Arg Arg Arg Ile Tyr Lys Val Glu Lys Ala Ser Gly
1385 1390 1395
Val Glu Cys Tyr Phe Ile Arg Gln Asp Ile Ala Tyr Leu Ile Lys
1400 1405 1410
Ser Tyr Asp Ala Thr Thr Lys Ile Gly Glu Leu Glu Ser Gln Asn
1415 1420 1425
Lys Leu Gln Thr Thr Met Ser Ser Asp Arg Leu Lys Ile Ala Asp
1430 1435 1440
Val Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly Ile Ser Arg
1445 1450 1455

<210> 127
<211> 1662
<212> PRT
<213> 食酸菌属物种
<223> APG06007_RGN
<400> 127

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

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

405 410 415
Lys Arg Lys Phe Ala Ile Lys Gly Thr Gly Thr Lys Glu Leu Arg Glu
420 425 430
Glu His Ile Gly Arg Lys Val Val Pro Arg Ser His Pro Ile Phe Gln
435 440 445
Asp Phe Lys Val Trp Gln Gln Ile Asn Asn Val Arg Leu Ile Trp Asn
450 455 460
Thr Ala Asp Glu Lys Ile Asn Leu Phe Glu Asn Ala Glu Leu Ser Glu
465 470 475 480
Lys Leu Thr Gly Lys Thr Ile Glu Lys Val Lys Glu Leu Leu Phe Lys
485 490 495
Lys Leu Gln Gly Ser Lys Thr Leu Ser Trp Arg Thr Phe Val Lys Glu
500 505 510
Glu Leu Gly Leu Asp Thr Tyr Asp Glu Ile Glu Glu Lys Arg Lys Ser
515 520 525
Lys Val Ser Lys Lys Gln Gly Leu Asp Ile Glu Thr Gly Glu Ile Val
530 535 540
Asp Ser Phe Phe Ser Val Asn Phe Arg Lys Arg Lys Arg Asp Gly Thr
545 550 555 560
Phe Asp Asp Ile Lys Leu Lys Gly Asn Thr Thr Lys Ala Glu Leu Gln
565 570 575
Asn Ile Leu Lys Glu Lys Arg Ala Glu Trp Phe Asp Thr Leu His Asn
580 585 590
Gln Asn Lys Lys Val Thr Asn Leu Gln Leu Leu Trp Glu Ile Ile Tyr
595 600 605
Asp Ile Thr Asn Ser Asp Ala Thr Lys Val Ser Ser Ile Ile Lys Lys
610 615 620
His Phe Asp Phe Asp Glu Ser Val Cys Leu Thr Leu Ala Lys Val Gln
625 630 635 640
Phe Asp Asp Ala Gly Met Ala Asn Leu Ser Ala Lys Ala Ile Arg Asn
645 650 655
Leu Leu Pro Leu Met Ser Asp Gly Ala Asn Tyr Ser Lys Ile Thr Lys
660 665 670
Ala Lys Val Asp Glu Leu Ile Ser Leu Asn Thr Ser Glu Thr Asp Val
675 680 685
Asp Ala Asp Asp Lys Leu Asn Gly Leu Lys Asp Phe Val Thr Asp Arg
690 695 700
Lys Ala Arg Ile Arg Leu Ser Lys Phe Ser Asn Glu Asn Asp Phe Lys
705 710 715 720

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

1025 1030 1035

Met Leu Arg Glu Lys Trp Lys Leu Asn Asp Ile Trp Lys Glu Leu

1040 1045 1050

Val Gln Glu Arg Phe Leu Lys Phe Asn Asn Gly Glu Val Lys Asp

1055 1060 1065

Arg Leu Asn Leu Lys Tyr Thr Thr Lys Tyr Tyr Asn Asn Lys Thr

1070 1075 1080

Asn Gln Tyr Glu Asp Val Glu Lys Phe Ala Gly Met Ser Lys Arg

1085 1090 1095

Ile Asp His Arg His His Ala Leu Asp Ala Ile Ile Ile Ala Cys

1100 1105 1110

Thr Lys Gln Ser His Ile Gln Tyr Ile Asn Asn Leu Asn Ala Ile

1115 1120 1125

Asn Ser Ala Asp Ile Glu Asn Asp Asp Asn Lys Lys Thr Lys Tyr

1130 1135 1140

Lys Gly Ile Lys Thr Asp Val Cys Ile Gly Asn Ser Ser Ser Lys

1145 1150 1155

Phe Lys Phe Pro Trp Asp Glu Asn Lys Phe Ile Pro Glu Val Lys

1160 1165 1170

Gln Ala Leu Gln Asn Ile Leu Ile Ser His Lys Asn Ser Asn Val

1175 1180 1185

Leu Ile Ser Pro Ser Lys His Arg Asn Asn Lys Asp Ile Asn Ser

1190 1195 1200

Gly Lys Val Ala Ser Val Arg Gly Glu Leu His Leu Gln Thr Asn

1205 1210 1215

Tyr Ala Lys Lys Lys Tyr Phe Glu Asn Glu Lys Ile Asp Ile Val

1220 1225 1230

Lys Leu Ile Pro Leu Leu Phe Lys Lys Lys Phe Glu Asn Gln Asn

1235 1240 1245

Gln Thr Leu Val Arg Phe Lys Ser Phe Asp Glu Ile Ile Lys Glu

1250 1255 1260

Ala Val Leu Lys Glu Lys Tyr Gln Asn Val Leu Ile Ser Leu Phe

1265 1270 1275

Lys Asp Tyr Asp Ser Glu Lys Leu Ser Thr Gln Thr Thr Lys Asp

1280 1285 1290

Phe Ser Lys Lys Ile Leu Lys Arg Ile Glu Asp Glu Lys Leu Phe

1295 1300 1305

Ile His Glu Lys Ser Gly Glu Lys Met Thr Trp Leu Ser Thr Phe

1310 1315 1320

Thr Glu Lys Asn Ala Ser Val Arg Pro Met Gly Leu Ala Met Asp
1325 1330 1335
Leu Asn Ser Glu Lys Glu Ile Lys Asp Ile Ser Asn Pro Arg Leu
1340 1345 1350
Arg Arg Ile Gly Gln Tyr Arg Leu Asn Tyr Ile Asn Lys Lys Ile
1355 1360 1365
Ile Glu Ile Lys Ala Leu Lys Leu Pro Phe Lys Glu Glu Glu Asn
1370 1375 1380
Asn Arg Ile Arg Glu Val Arg Glu Ile Pro Leu Tyr Asn Asn Ala
1385 1390 1395
Ile Tyr Glu Leu Arg Ile Lys Gln Lys Asp Gly Lys Tyr His Trp
1400 1405 1410
Val Glu Ile Lys Asp Val Gln His Ser Asp Leu Glu Asn Ile Glu
1415 1420 1425
Tyr Ala Lys Gln Glu Thr Thr Ile Ala Val Lys Asn Lys Ile Lys
1430 1435 1440
Asp Phe Ser Leu Asp Glu Leu Lys Glu Arg Tyr Ile Ser Asn Pro
1445 1450 1455
Ile Phe Leu Ser Glu Lys Pro Thr Pro Val Lys Lys Ala Arg Gln
1460 1465 1470
Lys Ser Phe Phe Gln Asp Leu Tyr Gln Leu Asn Asp Gly Arg Tyr
1475 1480 1485
Val Tyr Ser Arg Asp Val Phe Met Thr Tyr Val Phe Val Pro Glu
1490 1495 1500
Asp Ser Asn Arg Lys Asp Ala Lys Arg Glu Leu Glu Phe Leu Lys
1505 1510 1515
Phe Phe Asp Ala Val Lys Ile Val Thr Leu Glu Lys Pro Glu Lys
1520 1525 1530
Ile Asp Tyr Gln Lys Leu Ile Lys Lys Glu Asp His Lys Leu Leu
1535 1540 1545
Phe Thr Leu Ser Lys Asn Asp Ile Val Tyr Ile Pro Glu Glu Ile
1550 1555 1560
Leu Thr Asp Glu Gln Ile Ala Glu Ile Asn Trp Asn Asp Lys Arg
1565 1570 1575
Asp Ile Val Pro Lys Leu Phe Val Val Lys Glu Met Thr Pro Ser
1580 1585 1590
Arg Asn Glu Ile Val Phe Gln His Leu Tyr Lys Ser Asp Ser Ile
1595 1600 1605
Arg Val Asn Gly Asp Glu Ala Lys Asn Ile Leu Asn Thr Glu Lys

1610 1615 1620
Leu Glu Glu Gln Ile Lys Tyr Gly Asp Thr Asn Met Trp Lys Arg
1625 1630 1635
Cys Ile Lys Val Asp Ile Asp Lys Leu Gly Thr Glu Ile Lys Pro
1640 1645 1650
Tyr Trp Glu Asn Asn Lys Ile Val Leu
1655 1660

<210> 128
<211> 1075
<212> PRT
<213> 芽孢杆菌属物种
<223> APG02874_RGN
<400> 128

Met Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp
1 5 10 15
Gly Val Ile Arg Met Gln Tyr Asn Asp Leu Leu Glu Arg Tyr Glu Lys
20 25 30
Gln Gly Ile Val Asp Ala Gly Val Arg Met Phe Asp Lys Ala Glu Ile
35 40 45
Pro Lys Thr Gly Ala Ser Leu Ala Ala Pro Arg Arg Glu Ala Arg Ser
50 55 60
Ser Arg Lys Arg Leu Arg Arg Arg Ser Gln Arg Lys Lys Ala Ile Arg
65 70 75 80
Tyr Leu Ile Ile Leu His Glu Leu Leu Ser Glu Asp Gln Leu Gln Gln
85 90 95
Leu Tyr Pro Leu Gln Lys Glu Ser Ile Asp Ile Trp Asp Ile Arg Leu
100 105 110
Glu Ser Leu Asp Arg Val Leu Asn Gln Thr Glu Trp Ala Arg Leu Leu
115 120 125
Ile His Leu Gly Gln Arg Arg Gly Phe Lys Ser Asn Arg Lys Ser Glu
130 135 140
Tyr Lys Glu Glu Glu Thr Gly Lys Val Leu Val Asn Ile Gln Glu Asn
145 150 155 160
Gln Lys Ala Leu Gln Gln Tyr Arg Thr Val Ala Glu Met Trp Met Lys
165 170 175
Asp Ala Lys Phe Ser Gln Phe Asp Lys Lys Arg Asn Ala Pro Gly Glu
180 185 190
Tyr Leu Phe Ser Val Ser Arg Tyr Asp Leu Gln Asn Glu Ile Gln Thr
195 200 205

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

515 520 525

Gln Ala Asn Tyr Glu Arg Asn Lys Gly Ala Ile Ser Val Leu Val Glu

530 535 540

Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp Ile Val Arg Tyr Lys Leu

545 550 555 560

Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr Ser Leu Lys Val Ile Pro

565 570 575

Pro Ser Val Phe Phe Gly Glu Leu Lys Arg Glu Arg Asn His Ile Pro

580 585 590

Thr Leu Glu Val Asp His Ile Leu Pro Tyr Ser Gln Ser Phe Met Asp

595 600 605

Asn Tyr His Asn Lys Val Leu Val Tyr Ser Asp Glu Asn Gln Lys Lys

610 615 620

Gly Asn Arg Ile Pro Ser Gln Tyr Leu Met Ala Val Ser Pro Glu Lys

625 630 635 640

Trp Thr Lys Phe Glu Glu Tyr Val Ser Ser Asn Lys Gly Phe Ser Lys

645 650 655

Gln Lys Arg Gln Tyr Leu Leu Lys Lys Asp Tyr Ser Ser Arg Glu Arg

660 665 670

Asp Ile Val Lys Glu Arg His Leu Val Asp Thr Arg Tyr Ala Thr Arg

675 680 685

Phe Met Lys Asn Phe Ile Glu Lys Thr Leu Leu Phe Lys Glu Ser Arg

690 695 700

Ser Ser Ser Leu Thr Lys Arg Val Gln Thr Val Asn Gly Lys Ile Thr

705 710 715 720

Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys Glu Arg Thr Glu Thr

725 730 735

Tyr Leu His His Ala Leu Asp Ala Ile Val Val Ala Cys Thr Asp Gln

740 745 750

His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln Glu Arg Glu Arg Asn

755 760 765

Val Arg Gln Pro Tyr Phe Pro Trp Pro Trp Glu Gly Phe Arg Asp Glu

770 775 780

Leu Leu Asn Tyr Leu Asn Arg Gln Pro Asp Ser Leu Glu Ile Ser Gln

785 790 795 800

Ser Ile Lys Gln Asn Leu Phe Leu Arg Asp Tyr Leu Met Val Ser Arg

805 810 815

Met Pro Lys His Ser Val Thr Gly Met Ser His Lys Gln Thr Ile Arg

820 825 830

Ser Lys Ala Cys Val Asp Glu Lys Gly Lys Val Lys Thr Thr Lys Arg
835 840 845
Val Leu Leu Gln Asp Ile Lys Phe Asp Ser Asn Gly Asp Phe Lys Met
850 855 860
Ile Gly Lys Glu Gln Asp Met Ala Thr Tyr Asn Ala Ile Lys Glu Arg
865 870 875 880
Phe Leu Ser Ser Gly Gly Asn Val Lys Ser Ala Phe Thr Glu Pro Leu
885 890 895
His Lys Pro Ser Lys Lys Gly Lys Gly Asn Val Ile Arg Lys Val Thr
900 905 910
Val Glu Asp Gln Thr Lys Ser Phe Val Arg Asp Val Asn Gly Gly Val
915 920 925
Ala Glu Asn Gly Ser Leu Val Arg Ile Asp Leu Phe Glu Lys Glu Gly
930 935 940
Met Tyr Tyr Val Val Pro Ile Tyr Val Gly Asp Thr Val Leu Lys Gln
945 950 955 960
Leu Pro Ile Lys Ile Ser Ala Ser Asn Arg Gly Tyr Tyr Lys Trp Val
965 970 975
Glu Leu Asp Glu Asn Tyr Arg Phe Lys Phe Ser Leu Tyr Ala His Asp
980 985 990
Leu Ile Arg Ile Val Thr Lys Asp Glu Asp Arg Phe Phe Tyr Phe Ser
995 1000 1005
Tyr Phe Asp Asn Asp Gly Asn Arg Ile Lys Phe Leu Asn Ile Asn
1010 1015 1020
Ser Pro Thr Gly Lys Asn Glu Asn Arg Tyr Gly Val Arg Ser Ile
1025 1030 1035
Lys Leu Ile Glu Lys Tyr Asn Val Ser Thr Leu Gly Asp Ile Ser
1040 1045 1050
Leu Val Lys Ser Glu Glu Arg Gln Leu Phe Gln Gly Met Lys Lys
1055 1060 1065
Lys Val Leu Asn Ile Ile Asn
1070 1075

<210> 129
<211> 1068
<212> PRT
<213> 芽孢杆菌属物种
<223> APG03850_RGN
<400> 129
Met Lys Tyr Val Leu Gly Leu Asp Ile Gly Ile Ala Ser Cys Gly Trp

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

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

625 630 635 640
Gln Gln Lys Leu Leu Leu Gln Tyr Phe Ser Tyr Thr Arg Glu Gln Glu
645 650 655
Phe Ile Lys Arg Asn Leu Asn Asp Thr Arg Tyr Ala Thr Ile Tyr Leu
660 665 670
Ser Thr Leu Ile Gln Gln His Leu Ile Phe Ser Glu Ser Ser Arg Lys
675 680 685
Lys Lys Val His Thr Val Ser Gly Ile Ile Thr Ser His Leu Arg Ser
690 695 700
Arg Trp Gly Phe Asn Lys Asp Arg Lys Glu Gly His Ile His His Ala
705 710 715 720
Leu Asp Ala Val Ile Val Ala Val Thr Ser Asp His Met Ile Gln Arg
725 730 735
Val Thr Lys Tyr Tyr Lys Leu Lys Glu Leu Asn Arg Asn Leu Gln Ala
740 745 750
Lys Arg Met Gln Phe Pro Glu Pro Trp Glu Gly Phe Arg Leu Glu Leu
755 760 765
Glu Ala Arg Ile Ser Pro Asn Thr Gln Gln Tyr Leu Lys Arg Ile Leu
770 775 780
Phe Lys Asn Tyr Ala Asp Val Asn Leu Ser Glu Ile Lys Pro Ile Phe
785 790 795 800
Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Glu Leu His Gln Glu
805 810 815
Thr Ile Arg Lys Leu Ile Gly Tyr Asn Glu Lys Gly Lys Val Leu Thr
820 825 830
Ala Ile Lys Thr Lys Leu Glu Asp Ile Pro Phe Asp Ala Asn Gly Asp
835 840 845
Phe Pro Met Tyr Gly Lys Glu Thr Asp Leu Tyr Thr Tyr Asn Ala Ile
850 855 860
Lys Glu Arg Tyr Leu Ser His Lys Lys Asp Lys Arg Lys Ser Phe Gln
865 870 875 880
Asp Pro Leu Tyr Lys Pro Thr Lys Ser Gly Glu Ile Gly Pro Leu Ile
885 890 895
Lys Ser Ile Lys Ile Met Asp Thr Arg Thr Ile Val Asn Pro Val Asn
900 905 910
Gln Gly Lys Gly Val Val Tyr Asn Ser Lys Ile Ala Arg Thr Asp Val
915 920 925
Phe Lys Lys Asp Glu Lys Tyr Tyr Leu Ile Pro Ile Tyr Thr Ile Asp
930 935 940

Leu Leu Lys Asn Ile Leu Pro Gln Lys Ala Ile Thr Ala Gly Lys Gly
945 950 955 960
Tyr Glu Asp Trp Thr Thr Ile Asp Pro Ser Phe Thr Phe Leu Phe Ser
965 970 975
Leu Phe Pro Asn Asp Leu Ile Gln Ile Val Pro Ser Lys Asn Lys Thr
980 985 990
Ile Lys Ala Arg Thr Thr Val Ser Lys Lys Glu Val Leu Leu Pro Ser
995 1000 1005
Leu Thr Gly Tyr Phe Lys Gly Val His Ser Gly Thr Ala Gly Ile
1010 1015 1020
Thr Val Glu Thr His Asp Gly Ser Val Ile Ala Asn Val Gly Ser
1025 1030 1035
Lys Gln Leu Leu Leu Phe Glu Lys Tyr Gln Val Asp Val Leu Gly
1040 1045 1050
His Tyr Thr Lys Ile Lys Glu Glu Lys Arg Ile Gly Met Val Ile
1055 1060 1065

<210> 130
<211> 1081
<212> PRT
<213> 芽孢杆菌属物种
<223> APG07553_RGN
<400> 130

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

130 135 140

Ser Ile Ser Glu Asn Gln Gln Ile Met Glu Gln Lys Asn Tyr Lys Thr
145 150 155 160
Ile Gly Glu Met Ile Leu Lys Asp Lys Lys Phe Glu Asn His Lys Arg
165 170 175
Asn Lys Asp Gly Thr Tyr Ile Gly Thr Val Thr Arg Gln Gln Leu Gln
180 185 190
Asp Glu Ile Gln Ile Ile Phe Asn Ala Gln Arg Leu Tyr Lys Asn Asn
195 200 205
Tyr Ala Thr Lys Glu Phe Glu Ser Ser Tyr Leu Glu Ile Trp Ala Ser
210 215 220
Gln Arg Pro Tyr Ala Ser Lys Asp Gln Ile Glu Lys Met Ile Gly Tyr
225 230 235 240
Cys Thr Leu Glu Gln Lys Glu Lys Arg Val Pro Arg Ala Ser Trp Ser
245 250 255
Phe Gln Tyr Phe Val Ala Leu His Thr Ile Asn Asn Leu Arg Leu Ile
260 265 270
Ser Lys Asp Arg Ile Glu Glu Leu Ser Phe Lys Glu Lys Lys Gln Ile
275 280 285
Met Asn Leu Ala Leu Glu Lys Pro Ile Val Lys Tyr Ile Asp Ile Arg
290 295 300
Lys Leu Leu Ser Ile Pro Asn Glu Leu His Phe Asn Ser Leu Leu Tyr
305 310 315 320
Ser Ala Asp Thr Val Asp Thr Thr Val Glu Asn Arg Lys Cys Ile Glu
325 330 335
Leu Lys Glu Tyr His Ser Ile Asn Lys Val Tyr Lys Gln Ile Tyr Gly
340 345 350
Lys Asn Ala Leu Asn Leu Leu Pro Ile Asp Tyr Asp Thr Ile Ala Tyr
355 360 365
Gly Leu Thr Ile Phe Lys Asp Asp Lys Asp Ile Leu Glu His Leu Lys
370 375 380
Asn Lys Tyr Val Asn Ala Lys Gly Lys Pro Ile Asn Asn Leu Ala Lys
385 390 395 400
Lys Thr Tyr Asp Asp Thr Phe Ile Gln Ala Leu Leu Thr Leu Asn Phe
405 410 415
Ser Lys Met Gly His Leu Ser Phe Lys Ala Leu Lys Asn Ile Ile Pro
420 425 430
Phe Leu Glu Glu Gly Leu Ser Tyr Asp Lys Ala Cys Glu Lys Ala Gly
435 440 445

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

755 760 765

Glu Ala Arg Ile Ser Pro Asn Thr Gln Gln Tyr Leu Lys Gly Leu Arg

770 775 780

Phe Lys Asn Tyr Ala Asp Val Asn Leu Cys Glu Ile Lys Pro Ile Phe

785 790 795 800

Val Ser Lys Met Pro Lys Arg Ser Ile Thr Gly Glu Leu His Gln Glu

805 810 815

Thr Ile Arg Lys Phe Ile Gly Tyr Asn Glu Lys Gly Lys Val Leu Thr

820 825 830

Ala Ile Lys Thr Lys Leu Glu Asp Ile Pro Phe Asp Ala Asn Gly Asp

835 840 845

Phe Pro Met Tyr Gly Lys Glu Thr Asp Leu Tyr Thr Tyr Asn Ser Ile

850 855 860

Lys Glu Arg Tyr Leu Ser His Lys Lys Asp Lys Arg Lys Ser Phe Gln

865 870 875 880

Glu Pro Leu Tyr Lys Pro Thr Lys Ser Gly Gly Ile Gly Pro Leu Ile

885 890 895

Lys Ser Ile Lys Ile Met Asp Thr Arg Thr Ile Val Asn Pro Val Asn

900 905 910

Gln Gly Lys Gly Val Val Tyr Asn Ser Lys Ile Ala Arg Thr Asp Val

915 920 925

Phe Lys Lys Asp Asp Lys Tyr Tyr Leu Ile Pro Ile Tyr Thr Ile Asp

930 935 940

Leu Met Lys Asn Ile Leu Pro Gln Lys Ala Ile Thr Ala Gly Lys Gly

945 950 955 960

Tyr Glu Asp Trp Ile Thr Ile Asp His Ser Phe Thr Phe Leu Phe Ser

965 970 975

Leu Phe Pro Asn Asp Leu Ile Lys Ile Val Pro Ser Lys Asn Lys Glu

980 985 990

Ile Lys Ala Arg Ser Thr Ser Ser Lys Lys Glu Ile Leu Leu Pro Ser

995 1000 1005

Leu Ile Gly Tyr Phe Lys Ser Val His Ser Gly Thr Ala Gly Ile

1010 1015 1020

Thr Val Glu Ser His Asp Gly Arg Phe Ile Ala Asn Val Gly Ser

1025 1030 1035

Lys Gln Leu Leu Leu Phe Glu Lys Tyr Gln Val Asp Val Leu Gly

1040 1045 1050

His Tyr Thr Lys Ile Lys Glu Glu Lys Arg Ile Gly Met Ala Thr

1055 1060 1065

Cys Asn Asp Asn Lys Lys Ser Thr Ala Phe Gly Ser Leu
1070 1075 1080

<210> 131
<211> 1404
<212> PRT
<213> 金黄杆菌属物种
<223> APG03031_RGN
<400> 131

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

245 250 255
Trp Lys Ala Ile Ile Trp Gln Arg Pro Leu Arg Thr Gln Lys Gly Leu
260 265 270
Val Gly Asn Cys Ile Tyr Glu Lys Asn Lys Lys Arg Val Ala Val Ser
275 280 285
His Pro Leu Tyr Glu Glu Tyr Arg Thr Trp Val Phe Ile Asn Asn Leu
290 295 300
Asn Ile Ile Pro Pro Phe Gly Glu Glu Lys Gln Asn Tyr Ile Ser Glu
305 310 315 320
Lys Ile Tyr Pro Leu Phe Tyr Lys Ser Lys Pro Asp Phe Glu Leu Ser
325 330 335
Asp Ile Asp Lys Gln Leu Lys Lys Asp Gly Ala Lys Arg Thr Ser Lys
340 345 350
His Asn Asp Lys Thr Lys Val Leu Ser Val Lys Leu Leu Lys Gln Phe
355 360 365
Gln Asp Val Phe Gly Glu Asp Trp Lys Gln Lys Leu Gln Trp Asp Ile
370 375 380
Asn Gly Asp Arg Asn Ser Gln Pro Glu Lys Lys Glu Lys Ser Leu Tyr
385 390 395 400
Thr Phe Glu Asp Leu Trp His Ile Leu Leu Thr Phe Asp Gly Gln Glu
405 410 415
Asn Leu Lys Lys Phe Ala Leu Glu Lys Leu Asn Leu Asp Asp Glu Lys
420 425 430
Ala Ile Lys Phe Ser Lys Ile Lys Leu Gln Gln Gly Tyr Ala Thr Leu
435 440 445
Ser Ile Ser Ala Ile Lys Lys Ile Leu Pro Tyr Leu Gln Arg Gly Phe
450 455 460
Leu Tyr Ser His Ala Val Tyr Met Ala Asn Leu Tyr Lys Met Leu Gly
465 470 475 480
Glu Thr Thr Ile Ser Glu Thr Leu Thr Asn His Phe Ser Asp Glu Ile
485 490 495
Leu Lys Ile Leu Glu Asn Asp Asp Phe Asn Arg Lys Asn Ser Ala Ile
500 505 510
Ile Asn Ser Leu Ile Ser Glu Met Leu Asn Asp Asp Asn Arg Tyr Tyr
515 520 525
Ile Ala Asp Asn Arg Asn Leu Asp Lys Thr Glu Glu Gln Ile Val Leu
530 535 540
Lys Ser Ile Thr Asp Asn Tyr Gly Leu Glu Thr Trp Gly Asn Phe Asp
545 550 555 560

Glu Glu Lys Lys Met Glu Ile Ile Asp Phe Val Ser Ser Lys Tyr Leu
565 570 575
Glu Phe Leu Lys Lys Thr Tyr Thr Glu Lys Gly Asn Leu Phe Ile Gln
580 585 590
Pro Glu Arg Ile His Asp Lys Ile Phe Asn Phe Leu Lys Glu Thr Tyr
595 600 605
His Val Ser Glu Asp Lys Ile Lys Tyr Leu Trp His Pro Ser Glu Gln
610 615 620
Glu Asn Tyr Ala Lys Ala Asp Glu Tyr Phe Glu Tyr Ser Ile Gly Ser
625 630 635 640
Lys Lys Tyr Tyr Val Lys Glu Glu Ser Thr Gln Asn Phe Leu Ser Arg
645 650 655
Asn Pro Asn Ala Glu Phe Asp Gly Arg Gln Leu Lys Leu Leu Gly Asn
660 665 670
Pro Glu Pro Leu Ser Lys Gly Phe Lys Asn Pro Met Ala Leu Lys Ser
675 680 685
Leu His Lys Leu Lys Asn Leu Val Asn Tyr Leu Leu Gln Thr Gly Lys
690 695 700
Ile Asp Glu Gly Thr Arg Val Val Ile Glu Ile Ala Arg Glu Leu Asn
705 710 715 720
Asp Lys Asn Lys Arg Lys Ala Ile Glu Asn Trp Gln Lys Asp Arg Glu
725 730 735
Lys Glu Asn Glu Ala Phe Arg Lys Glu Ile Glu Val Tyr Lys Glu Gln
740 745 750
Phe Pro Asn Ile Asn Leu Ile Asp Glu Asn Thr Leu Ile Arg Lys Ile
755 760 765
Arg Leu Trp His Glu Gln Asn Lys Ile Cys Leu Tyr Thr Gly Lys Thr
770 775 780
Ile Pro Phe Ser Glu Leu Ile Ile Gly Asn Lys Tyr Asp Phe Glu His
785 790 795 800
Thr Ile Pro Ala Ser Ile Ser Phe Asp Asn Glu Leu Lys Asn Leu Thr
805 810 815
Ile Ser Asp Ser Ser Tyr Asn Arg Leu Tyr Lys Gly Lys Lys Phe Pro
820 825 830
Thr Gln Leu Ser Asn Tyr Asp Ser Glu Gln Thr Ile Asn Gly Glu Ser
835 840 845
Cys Asn Ser Ile Leu Arg Asn Ile Glu Phe Ile Phe Gly Glu Arg Ile
850 855 860
Val Glu Tyr Lys Glu Ile Lys Gly Lys Thr Glu Lys Ile Val Lys Trp

865 870 875 880
Lys Lys Ile Asp Glu Leu Glu Lys Gln Phe Asp Glu Trp Lys Lys Lys
885 890 895
Ala Ser Tyr Ala Ser Thr Lys Glu Ile Lys Asp Asn Cys Ile Val Lys
900 905 910
Tyr His Thr Ile Lys Met Asp Leu Asp Tyr Leu Lys Ser Lys Leu Ala
915 920 925
Thr Phe Thr Ile Thr Glu Tyr Lys Ala Gly Trp Arg Asn Ser Gln Ile
930 935 940
Arg Asp Thr Gln Ile Ile Thr Lys Tyr Ala Leu Pro Tyr Phe Lys Thr
945 950 955 960
Leu Phe Lys Arg Val Ser Val Glu Lys Gly Gly Val Thr Asp Ile Phe
965 970 975
Lys Lys Val Tyr Lys Val Gln Ser Arg Asn Ser Lys Lys Asp Arg Ser
980 985 990
Val His Ser His His Ala Gln Asp Ala Ala Ile Leu Thr Leu Ile Pro
995 1000 1005
Asn Ala Phe His Arg Glu Arg Ile Ile Lys Ala Tyr Glu Asn Glu
1010 1015 1020
Ile Asp Asn Arg Thr Gly Lys Thr Tyr His Glu Gln Pro Leu Asp
1025 1030 1035
Trp Glu Asn Phe Ser Glu Lys Tyr Ile Leu Glu Leu Gln Asp Lys
1040 1045 1050
Val Leu Ile Asn Asn Leu Thr Asp Asn Arg Thr Ile Thr Gln Thr
1055 1060 1065
Tyr Lys Thr Val Arg Lys Arg Gly Lys Val Val Tyr Asp Ile Asp
1070 1075 1080
Glu Asn Gly Asn Ser Leu Lys Arg Val Ser Lys Gly Asp Thr Ile
1085 1090 1095
Arg Gly Gln Leu His Gly Glu Thr Phe Tyr Gly Ala Ile Lys Gln
1100 1105 1110
Pro Val Arg Asp Glu Asp Asn Lys Ile Leu Phe Asp Glu Asn Lys
1115 1120 1125
Lys Met Ile Leu Lys Asp Glu Ile Tyr Leu Ala Val Arg Lys Pro
1130 1135 1140
Leu Val Tyr Lys Lys Asp Ala Asn Ser Pro Gly Phe Lys Thr Leu
1145 1150 1155
Asp Glu Val Glu Lys Val Ile Val Asp Lys Gly Leu Phe Lys Met
1160 1165 1170

Ile Lys Lys Gln Val Glu Glu Ser Asp Phe Lys Thr Ala Leu Ile
1175 1180 1185
Asp Gly Val Tyr Met Leu Asn Glu Ser Gly Glu Lys Val Asn Lys
1190 1195 1200
Ile Arg Arg Ile Arg Cys Phe Glu Asn Gly Leu Lys Tyr Thr Thr
1205 1210 1215
Ala Ile Lys Val His Glu His Ser Phe Val Ser His Lys Glu Tyr
1220 1225 1230
Lys Gln Ser Thr Leu Ala Thr Asn Gly Glu Asn Thr Tyr Cys Leu
1235 1240 1245
Phe Tyr Lys Asn Asp Lys Gly Lys Ala Met Lys Ile Leu Ser Ile
1250 1255 1260
Val Asp Leu Ala Glu Leu Lys Leu Lys Asn Ile Gln Asn Leu Tyr
1265 1270 1275
Asn Glu Pro Glu Phe Ser Ser Phe Glu Val Gly Lys Gly Lys Asn
1280 1285 1290
Lys Tyr Lys Ile Pro Leu Tyr Ser Val Leu Lys Ser Gly Asp Lys
1295 1300 1305
Val Leu Phe Tyr Lys Glu Asn Ile Tyr Glu Leu Lys Asp Leu Glu
1310 1315 1320
Gln Lys Glu Leu Ser Asp Arg Met Phe Lys Met Tyr Gln Phe Glu
1325 1330 1335
Lys Asp Gly Arg Ile Lys Phe Arg His His Leu Ala Ala Gly Ile
1340 1345 1350
Asp Thr Glu Leu Lys Lys Glu Asn Pro Glu Asn Ser Ser Val Asn
1355 1360 1365
Phe Glu Glu Lys Gln Val Phe Leu Arg Leu Ser Gln Gly Gln Trp
1370 1375 1380
Asn Phe Ala Ile Asp Gly Lys Asp Phe Glu Ile Ser Leu Asp Gly
1385 1390 1395
Lys Ile Asn Trp Lys Phe
1400

<210> 132
<211> 1092
<212> PRT
<213> 芽孢杆菌属物种
<223> APG09208_RGN
<400> 132
Met Arg Tyr Ser Ile Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Ser

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

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

625 630 635 640
Thr Met Lys Val Gly Lys Lys Lys Lys Asp Met Leu Leu Lys Lys Asn
645 650 655
Phe Asp Glu Glu Gln Glu Glu Gln Ile Leu Ser Arg Asn Leu Asn Asp
660 665 670
Thr Arg Tyr Ile Thr Arg Tyr Ile Lys Ser Phe Ile Glu Asp Asn Leu
675 680 685
Glu Phe Arg Thr Glu Glu Asn Lys Lys Gln Tyr Val His Thr Val Asn
690 695 700
Gly Gly Tyr Thr Ser Leu Met Arg Lys Arg Trp Gly Leu Asn Lys Asp
705 710 715 720
Arg Arg Gly Asn Asp Leu His His Ala Val Asp Ala Ala Ile Ile Ala
725 730 735
Val Ser Leu Pro Phe Lys His Lys Val Asn Thr Tyr Phe Lys Arg Gln
740 745 750
Glu Ile Gly Leu Ser Asn Leu Ser Asn Asn Lys Lys Asp Ile Phe Pro
755 760 765
Glu Pro Trp Arg Asn Phe Ile Lys Glu Leu Glu Ala Arg Met Ile Gln
770 775 780
Asp Pro Glu Lys Met Lys Thr Ala Leu Glu Ser Leu Glu Leu Glu Thr
785 790 795 800
Tyr Gly Glu Ile Phe Leu Asn Lys Leu Lys Pro Ile Phe Val Ser Arg
805 810 815
Met Pro Lys His Ser Ile Lys Gly Gln Ile His Glu Glu Thr Ile Arg
820 825 830
Arg Val Arg Gly Leu Thr Glu Glu Gly Phe Leu Val Thr Val Lys Lys
835 840 845
Thr Arg Leu Asp Gln Ile Pro Phe Asp Lys Asn Gly Asp Phe Pro Met
850 855 860
Tyr Gly Lys Glu Thr Asp Ile Lys Thr Tyr Met Ala Ile Lys Gln Arg
865 870 875 880
Tyr Leu Glu Tyr Gly Gln Asp Lys Gln Lys Ala Phe Ala Ile Pro Leu
885 890 895
Arg Lys Pro Ser Lys Asn Pro Glu Asn Ala Pro Ile Val Arg Ser Val
900 905 910
Lys Ile Glu Gly Lys Ala Asn Arg Val Val Met Leu Asp Asp Lys Ala
915 920 925
Ala Ala Asp Asn Ala Ser Ile Ala Arg Thr Glu Val Phe Arg His Lys
930 935 940

Glu Thr Gly Glu Tyr Tyr Val Thr Pro Val Tyr Val Ala Asp Ile Val
945 950 955 960
Ser Asn Lys Ile Pro Asp Arg Leu Ile Thr Leu Lys Lys Pro Tyr Ser
965 970 975
Glu Trp Asp Arg Ile Thr Asp Glu Tyr Glu Tyr Leu Phe Ser Leu Tyr
980 985 990
Asn Asn Asp Leu Ile Lys Ile Val Leu Pro Lys Glu Lys Glu Thr Lys
995 1000 1005
Lys Tyr Thr Gly Gly Asn His Leu Trp Gln Glu Gly Phe Phe Tyr
1010 1015 1020
Phe Lys Gly Val Asp Ser Ser Asn Ala Ser Ile Lys Ile Ile Asp
1025 1030 1035
His Leu Asn Ser Phe Glu Val Arg Ile Gly Ile Arg Arg Leu Ile
1040 1045 1050
Ala Phe Glu Lys Tyr Gln Val Asn Pro Leu Gly Glu Ile Asn Lys
1055 1060 1065
Val His Gly Glu Lys Arg Pro Gly Glu Leu Leu Asn Glu Glu Glu
1070 1075 1080
Ile Lys Glu Asn Arg Lys Asn Ile Ser
1085 1090

<210> 133
<211> 1150
<212> PRT
<213> 肠球菌属物种
<223> APG05586_RGN
<400> 133

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

100 105 110

Gly Ile Ser Tyr Leu Asp Glu Asp Asp Ala Glu Ala Ala Lys Glu Ser

115 120 125

Gln Asp Tyr Lys Glu Gln Val Arg Lys Asn Ala Gln Leu Leu Thr Lys

130 135 140

Tyr Thr Pro Gly Gln Ile Gln Leu Gln Arg Leu Lys Glu Asn Asn Arg

145 150 155 160

Val Lys Thr Gly Ile Asn Gly Gln Gly His Tyr Gln Leu Asn Val Phe

165 170 175

Lys Val Ser Ala Tyr Ala Asp Glu Leu Ala Thr Ile Leu Lys Thr Gln

180 185 190

Gln Ala Leu Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe

195 200 205

Val Gln Pro Gly Ile Ala Glu Asn Ala Gly Leu Ile Tyr Arg Lys Arg

210 215 220

Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly

225 230 235 240

Arg Trp Ser Asp Phe Gln Lys Thr Gly Gln Pro Ala Ala Asn Ile Phe

245 250 255

Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly

260 265 270

Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn

275 280 285

Leu Lys Ile Asp Gly Glu Val Ser Leu Ser Pro Glu Gln Lys Glu Phe

290 295 300

Ile Leu Thr Glu Leu Met Thr Lys Glu Phe Ala Arg Phe Gly Val Asn

305 310 315 320

Asp Ile Ala Lys Leu Leu Gly Val Lys Lys Glu Gln Leu Ser Gly Trp

325 330 335

Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr

340 345 350

Arg Asn Trp Arg Lys Ile Phe Ala Glu Ala Gly Ile Asp Leu Ala Thr

355 360 365

Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn

370 375 380

Thr Glu Arg Glu Gly Val Glu Asn Thr Leu Ala Phe Glu Leu Pro Glu

385 390 395 400

Leu Ala Glu Pro Val Lys Ser Leu Val Leu Asp His Tyr Lys Glu Leu

405 410 415

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

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

Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys
1040 1045 1050
Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp
1055 1060 1065
Ser Gln Arg Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn
1070 1075 1080
Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Ile Pro Ala Glu
1085 1090 1095
Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
1100 1105 1110
Asn Ile Ala Lys Tyr Gly Lys Glu Gly Lys Lys Leu Arg Lys Phe
1115 1120 1125
Asn Thr Asp Ile Leu Gly Lys Lys His His Leu Ser Tyr Glu Lys
1130 1135 1140
Glu Pro Lys Asn Ile Ile Lys
1145 1150

<210> 134
<211> 1150
<212> PRT
<213> 肠球菌属物种
<223> APG08770_RGN
<400> 134

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

130 135 140
Tyr Thr Pro Gly Gln Ile Gln Leu Gln Arg Leu Lys Glu Asn Asn Arg
145 150 155 160
Val Lys Thr Gly Ile Asn Gly Gln Gly Asn Tyr Gln Leu Asn Val Phe
165 170 175
Lys Val Ser Ala Tyr Ala Asp Glu Leu Ala Thr Ile Leu Lys Thr Gln
180 185 190
Gln Ala Leu Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe
195 200 205
Val Gln Pro Gly Ile Ala Glu Asp Ala Gly Leu Ile Tyr Arg Lys Arg
210 215 220
Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly
225 230 235 240
Arg Trp Ser Asp Phe Gln Lys Thr Gly Gln Pro Ala Ala Asn Ile Phe
245 250 255
Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly
260 265 270
Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn
275 280 285
Leu Lys Ile Asp Gly Glu Gln Pro Leu Ser Pro Thr Gln Lys Glu Phe
290 295 300
Ile Leu Thr Glu Leu Met Thr Lys Glu Phe Ala Arg Phe Gly Val Asn
305 310 315 320
Asp Ile Ala Lys Leu Leu Gly Val Lys Lys Glu Gln Leu Ser Gly Trp
325 330 335
Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr
340 345 350
Arg Asn Trp Arg Lys Ile Phe Ala Glu Ala Gly Ile Asp Leu Ala Thr
355 360 365
Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn
370 375 380
Thr Glu Arg Glu Gly Val Glu Asn Thr Leu Ala Phe Glu Leu Pro Glu
385 390 395 400
Leu Ala Glu Pro Val Lys Ser Leu Val Leu Asp His Tyr Lys Glu Leu
405 410 415
Ser Gln Ser Ile Ser Thr Gln Ala Trp His Arg Phe Ser Leu Lys Thr
420 425 430
Leu His Leu Leu Ile Pro Glu Leu Ile Lys Ser Thr Ser Glu Gln Asn
435 440 445

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

755 760 765

Glu Glu Thr Gly Glu Lys Phe Met Arg Glu Ile Asp Val Glu Thr Gly

770 775 780

Glu Ile Leu Asp Glu Ile Pro Tyr Arg Glu Tyr Lys Lys Ala Lys His

785 790 795 800

Tyr Glu Arg Lys Thr Tyr Gln Val Lys Trp Ser Asn Phe Arg Glu Gln

805 810 815

Leu Lys Pro Leu Thr Ile His Pro Lys Ile Lys Phe Ser His Gln Val

820 825 830

Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val

835 840 845

Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Glu Lys Ile

850 855 860

Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val

865 870 875 880

Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys

885 890 895

Glu Phe Asp Glu Lys Thr Tyr Glu Leu Leu Val Thr Ile Ala Ala Thr

900 905 910

Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg

915 920 925

Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Gly Ile Pro

930 935 940

Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Val Ile Arg Ser

945 950 955 960

Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys

965 970 975

Asp Glu Lys Gly Arg Pro Val Glu Gln Thr Lys Asn Gly Arg Lys Val

980 985 990

Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu

995 1000 1005

Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu

1010 1015 1020

Arg Phe Val Glu Gly Glu Tyr Gly Ile Thr Glu Lys Glu Tyr Met

1025 1030 1035

Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys

1040 1045 1050

Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp

1055 1060 1065

Ser Gln Arg Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn
1070 1075 1080
Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Met Pro Ala Glu
1085 1090 1095
Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
1100 1105 1110
Asn Ile Ala Lys Tyr Gly Lys Glu Gly Asn Lys Leu Arg Lys Phe
1115 1120 1125
Asn Thr Asp Ile Leu Gly Lys Lys His His Leu Ser Tyr Glu Lys
1130 1135 1140
Glu Pro Lys Asn Ile Ile Lys
1145 1150

<210> 135
<211> 1053
<212> PRT
<213> 葡萄球菌属物种
<223> APG08167_RGN
<400> 135

Met Lys Arg His Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val
1 5 10 15
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Glu 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 Lys His Arg Leu
50 55 60
Gln Ser Val Lys Lys Leu Leu Phe Asp Tyr Asp Leu Leu Thr Asn 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 Asn Leu Ser Glu Val Glu Phe Ser Gly 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
Arg 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 Thr
145 150 155 160
Asp Ser Glu Val Arg Gly Ala Ile Asn Arg Phe Lys Thr Ser Asp Tyr

165 170 175
Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Glu 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 Lys Gly Ser Pro Phe Gly Trp Lys
210 215 220
Asp Ile Lys Glu Trp Tyr Glu Met Leu Ile 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 Ser Arg Asp Asp Asn
260 265 270
Glu Lys Leu Ala 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 Ser Lys Glu Ile Phe
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 Lys Phe Lys Ile Tyr His Asp Ile Lys Gly Ile Lys
325 330 335
Glu Arg Lys Glu Ile Ile Glu Asn Ala Glu Val Leu Asp Gln Ile Ala
340 345 350
Glu Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
355 360 365
Ala Lys Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
370 375 380
Lys Leu Thr Gly Tyr Thr Gly Thr His Ser Leu Ser Leu Lys Ala Ile
385 390 395 400
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Ser Asp Asn Gln Met Ala
405 410 415
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Met 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 Ser Asp Ile Ile Ile Glu Leu Ala Arg
465 470 475 480

Glu Lys Asn Ser Lys Asp Ala Gln Lys Leu 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 Asn
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 Asn Ser Lys Lys Asp Asn Arg Thr Pro Phe Gln Tyr Leu
580 585 590
Ser Ser Ser Asp Ser Lys Met Asp Tyr Glu Thr Phe Lys Lys His Ile
595 600 605
Leu Asn Leu Ser Lys Gly Lys Asp Arg Ile Ser Lys Lys Lys Lys Val
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 Glu Leu
645 650 655
Met Thr Leu Leu Arg Ser Tyr Phe Ser 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 Gln Gly Tyr Lys His His Ala Glu Asp
690 695 700
Ala Leu Ile Ile Ala Asn Ala Asp Tyr Val Phe Lys Glu Trp Lys Lys
705 710 715 720
Leu Asp Arg Ala Asn Lys Val Met Glu Asn Gln Thr Val Glu Glu Asn
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 Gln His Ile Lys Asp Phe Lys Asp
755 760 765
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Ser Asn Arg Glu Leu Val
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 Met Asn Gln 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
Gly Glu Lys Asn Pro Leu Tyr Lys Tyr His 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 Val 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 Phe
900 905 910
Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Ile
915 920 925
Lys Asn Leu Asp Val Ile Lys Lys Glu Asp Tyr Tyr Glu Val Asn Ser
930 935 940
Lys Cys Tyr Lys Glu Ala Lys Lys Leu Lys Gln Ile Ser Asp Gln Ala
945 950 955 960
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Val Asn Gly
965 970 975
Glu Leu Tyr Arg Ile Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
980 985 990
Glu Val Asn Met Val Asp Ile Thr Tyr Lys Glu Tyr Leu Glu Thr Met
995 1000 1005
Asn Asp Lys Arg Thr Pro Arg Ile Phe Lys Ala Ile Thr Gly 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 Gln Asn Pro Gln Met Ile Met Lys Arg
1040 1045 1050

<210> 136
<211> 1052
<212> PRT
<213> 葡萄球菌属物种
<223> APG01604_RGN
<400> 136

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

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

Lys Glu Tyr Leu Leu Glu Glu Arg Asp Ile Asn Lys Tyr Glu Val Gln
625 630 635 640
Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg
645 650 655
Glu Leu Thr Asn Tyr Leu Lys Ala Tyr Phe Ser Ala Asn Asp Met Asp
660 665 670
Val Lys Val Lys Thr Ile Asn Gly Ser Phe Thr Asp Tyr Leu Arg Lys
675 680 685
Val Trp Lys Phe Lys Lys Glu Arg Asn His Gly Tyr Lys His His Ala
690 695 700
Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Leu Phe Lys Glu Asn
705 710 715 720
Lys Lys Leu Lys Lys Ala Asn Ala Ile Leu Glu Gln Pro Ser Leu Asp
725 730 735
Asn Gly Lys Ser Asp Ala Thr Val Glu Asn Asp Asn Glu Tyr Val Glu
740 745 750
Thr Phe Ser Ile Pro Lys Gln Val Asn Asp Ile Lys Glu Phe Arg Asp
755 760 765
Phe Lys Phe Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln Leu Ile
770 775 780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Ile Glu Asn His Thr Phe Ile
785 790 795 800
Val Ser Pro Ile Thr Asn Ile Tyr Ser Lys Asp Asn Asp Glu Leu Lys
805 810 815
Lys Lys Phe Asn Lys Asn Pro Glu Lys Phe Leu Met Tyr Gln His Asp
820 825 830
Pro Lys Thr Phe Glu Lys Leu Glu Val Ile Met Lys Gln Tyr Ala Asn
835 840 845
Glu Lys Asn Pro Leu Ala Lys Tyr His Glu Glu Thr Gly Glu Tyr Leu
850 855 860
Thr Lys Tyr Ser Lys Lys Asn Asn Gly Pro Ile Val Lys Thr Ile Lys
865 870 875 880
Val Leu Gly Asp Lys Val Gly Lys His Leu Asp Val Thr His Lys Tyr
885 890 895
Lys Tyr Ser Asn Ser Lys Ile Val Lys Lys Thr Ile Asn Pro Tyr Arg
900 905 910
Phe Asp Val Tyr Leu Thr Asp Lys Gly Tyr Lys Phe Ile Thr Ile Ser
915 920 925
Tyr Leu Asp Val Leu Lys Lys Asp Asn Tyr Tyr Tyr Ile Leu Lys Glu

930 935 940
Lys Tyr Glu Glu Leu Lys Ile Lys Lys Ser Ile Ser Asp Thr Asp Gln
945 950 955 960
Phe Ile Gly Ser Phe Tyr Tyr Asn Asp Leu Ile Lys Ile Asn Asp Gln
965 970 975
Ile Phe Lys Val Val Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu
980 985 990
Leu Asp Leu Leu Asp Ile Ser Tyr Lys Glu Tyr Cys Lys Ile Asn Asn
995 1000 1005
Ile Lys Thr Asn Arg Ile Ile Lys Gly Ile Thr Lys Lys Ile Thr
1010 1015 1020
Asn Ile Glu Lys Phe Ser Thr Asp Val Leu Gly Asn Leu Tyr Lys
1025 1030 1035
Ala His Ser Asn His Pro Gln Leu Ile Phe Lys Gln Arg Asp
1040 1045 1050

<210> 137
<211> 1129
<212> PRT
<213> 链球菌属物种
<223> APG03021_RGN
<400> 137

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

Lys Thr Lys Thr Pro Gly Gln Ile Gln Leu Glu Arg Leu Glu Lys Tyr
145 150 155 160
Gly Gln Leu Arg Gly Asn Phe Thr Val Tyr Asp Glu Asn Gly Glu Ala
165 170 175
His Arg Leu Ile Asn Val Phe Ser Thr Ser Asp Tyr Lys Asn Glu Ala
180 185 190
Arg Lys Ile Leu Glu Ala Gln Ser Asn Tyr Asn Lys Gln Ile Thr Asp
195 200 205
Glu Phe Ile Glu Asp Tyr Ile Glu Ile Leu Thr Gln Lys Arg Lys Tyr
210 215 220
Tyr His Gly Pro Gly Asn Glu Lys Ser Arg Thr Asp Tyr Gly Arg Phe
225 230 235 240
Arg Thr Asp Gly Thr Thr Leu Glu Asn Ile Phe Gly Ile Leu Ile Gly
245 250 255
Lys Cys Ser Phe Tyr Pro Glu Glu Tyr Arg Ala Ser Lys Ala Ser Tyr
260 265 270
Thr Ala Gln Glu Phe Asn Phe Leu Asn Asp Leu Asn Asn Leu Lys Val
275 280 285
Pro Thr Glu Thr Gly Lys Leu Ser Thr Glu Gln Lys Glu Tyr Leu Val
290 295 300
Asp Phe Ala Lys Lys Ser Lys Ala Leu Gly Ala Ser Lys Leu Leu Lys
305 310 315 320
Glu Ile Ala Lys Ile Val Asp Cys Ser Val Asp Asp Ile Lys Gly Tyr
325 330 335
Arg Val Asp Asn Lys Asp Lys Pro Asp Leu His Thr Phe Glu Pro Tyr
340 345 350
Arg Lys Leu Lys Phe Asn Leu Ser Ser Ile Asp Ile Asp Glu Leu Ser
355 360 365
Arg Glu Thr Leu Asp Lys Leu Ala Asp Ile Leu Thr Leu Asn Thr Glu
370 375 380
Arg Glu Gly Ile Glu Asp Thr Ile Lys Arg Asn Leu Pro Ser Gln Phe
385 390 395 400
Thr Glu Glu Gln Ile Ser Glu Ile Val Gln Ile Arg Lys Asn Gln Ser
405 410 415
Ser Ala Phe Asn Lys Gly Trp His Ser Phe Ser Ala Lys Leu Met Asn
420 425 430
Glu Leu Ile Pro Glu Leu Tyr Val Thr Ser Glu Glu Gln Met Thr Ile
435 440 445
Leu Thr Arg Leu Glu Lys Phe Lys Val Asn Lys Lys Ser Ser Lys Asn

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

Gln Val Val Asn Lys Glu Thr Gly Glu Ile Leu Ser Ile Ser Asp Asp
770 775 780
Lys Tyr Lys Glu Leu Val Phe Gln Pro Pro Tyr Gln Gly Phe Val Asn
785 790 795 800
Thr Ile Ser Ser Lys Gly Phe Glu Asp Glu Ile Leu Phe Ser Tyr Gln
805 810 815
Val Asp Ser Lys Phe Asn Arg Lys Val Ser Asp Ala Thr Ile Tyr Ser
820 825 830
Thr Arg Lys Ala Lys Leu Gly Lys Asp Lys Lys Asp Glu Thr Tyr Val
835 840 845
Leu Gly Lys Ile Lys Asp Ile Tyr Ser Gln Asp Gly Phe Asp Thr Phe
850 855 860
Ile Lys Arg Tyr Lys Lys Asp Lys Thr Gln Phe Leu Met Tyr Gln Lys
865 870 875 880
Asp Pro Leu Thr Trp Glu Asn Val Ile Glu Val Ile Leu Arg Asp Tyr
885 890 895
Pro Ser Glu Lys Leu Ser Glu Asp Gly Lys Lys Thr Val Lys Cys Asn
900 905 910
Pro Phe Glu Glu Tyr Arg Arg Glu Asn Gly Leu Ile Cys Lys Tyr Ser
915 920 925
Lys Lys Gly Asn Gly Thr Pro Ile Lys Ser Leu Lys Tyr Tyr Asp Lys
930 935 940
Lys Leu Gly Asn Cys Ile Asp Ile Thr Pro Glu Lys Ser Lys Asn Arg
945 950 955 960
Val Val Leu Arg Gln Ile Ser Pro Trp Arg Ala Asp Ile Tyr Phe Asn
965 970 975
Leu Glu Thr Leu Lys Tyr Glu Leu Met Gly Leu Lys Tyr Ser Asp Leu
980 985 990
Ser Phe Glu Lys Gly Thr Gly Lys Tyr His Ile Ser Gln Glu Lys Tyr
995 1000 1005
Asp Ala Ile Arg Glu Lys Glu Gly Ile Gly Lys Lys Ser Glu Phe
1010 1015 1020
Lys Phe Thr Leu Tyr Arg Asn Asp Leu Ile Leu Ile Lys Asp Thr
1025 1030 1035
Leu Asn Asn Cys Glu Arg Met Leu Arg Phe Gly Ser Lys Asn Asp
1040 1045 1050
Thr Ser Lys His Tyr Val Glu Leu Lys Pro Leu Glu Lys Gly Thr
1055 1060 1065
Phe Asp Ser Glu Glu Glu Ile Leu Pro Val Leu Gly Lys Val Ala

1070 1075 1080
Lys Ser Gly Gln Phe Ile Lys Gly Leu Asn Lys Thr Asn Ile Ser
1085 1090 1095
Ile Tyr Lys Val Arg Thr Asp Val Leu Gly Asn Lys Phe Phe Ile
1100 1105 1110
Lys Lys Glu Gly Asp Lys Pro Lys Leu Asp Phe Lys Asn Asn Asn
1115 1120 1125
Lys

<210> 138
<211> 1359
<212> PRT
<213> 片球菌属物种
<223> APG06015_RGN
<400> 138

Met Glu Lys Val Pro Tyr Asn Ile Gly Leu Asp Ile Gly Thr Ser Ser
1 5 10 15
Ile Gly Phe Ala Ala Thr Asp Asn Leu Asn Lys Pro Ile Arg Ala Lys
20 25 30
Gly Lys Thr Val Ile Gly Val Arg Leu Phe Glu Glu Gly Lys Thr Ala
35 40 45
Ala Asp Arg Arg Gly Phe Arg Thr Thr Arg Arg Arg Leu Ser Arg Arg
50 55 60
Lys Trp Arg Leu Arg Leu Leu Asp Glu Ile Phe Asp Lys Glu Met Ala
65 70 75 80
Lys Val Asp Asn Thr Phe Phe Ala Arg Leu Lys Glu Ser Asn Leu Ser
85 90 95
Pro Lys Asp Ala Asn Lys Lys Tyr Leu Gly Ser Leu Leu Phe Pro Glu
100 105 110
Lys Lys Asp Phe Lys Phe Tyr Glu Asp Tyr Pro Thr Ile Tyr His Leu
115 120 125
Arg Tyr Ala Leu Met His Glu Lys Arg Gln Phe Asp Ile Arg Glu Val
130 135 140
Tyr Leu Ala Met His His Ile Ile Lys Tyr Arg Gly Asn Phe Leu Asn
145 150 155 160
Ser Ala Pro Met Asn Ser Phe Lys Thr Gln Asp Phe Asp Phe Val Ala
165 170 175
Lys Phe Lys Lys Leu Asn Glu Leu Phe Glu Ser Ile Asp Ala Glu His
180 185 190
Glu Thr Lys Phe Asp Ile Glu Asn Ile Ser Lys Phe Arg Asp Ile Met

195 200 205

Leu Asn Gln Asp Ile Arg Lys Leu Asp Arg Lys Lys Gln Ala Ala Lys

210 215 220

Ile Leu Ile Leu Asp Ser Thr Asp Lys Thr Ala Lys Lys Ile Asn Asn

225 230 235 240

Lys Ile Ala Thr Ala Val Ala Asn Ser Ala Leu Gly Tyr Lys Phe Ala

245 250 255

Leu Asp Ala Ile Leu Lys Leu Asp Val Glu Glu Ser Lys Asp Trp Ser

260 265 270

Ile Ser Leu Asn Asp Glu Glu Ile Asp Ser Ile Leu Asp Asn Leu Thr

275 280 285

Ser Asp Leu Asp Ala Glu Arg Ile Glu Ile Ile Glu Ile Leu Arg Asp

290 295 300

Leu Tyr Ser His Ile Ala Leu Asn Glu Ile Val Pro Asn Gly Gln Ser

305 310 315 320

Leu Ser Lys Ser Met Met Asp Lys Tyr Asp Lys His His Ala Asp Leu

325 330 335

Asp Val Leu Lys Lys Val Ile Ser Asn Met Asp Asp Arg Lys Lys Ala

340 345 350

Lys Ser Leu Lys Asn Ile Tyr Asn Gln Tyr Val Gly Lys Thr Asn Asp

355 360 365

Lys Val Leu Asp Lys Asp Glu Phe Tyr Lys Gln Ile Gln Lys Asn Leu

370 375 380

Asp Lys Ser Glu Asp Ala Met Lys Ile Val Asn Glu Ile Glu Leu Asp

385 390 395 400

Gln Phe Met Pro Lys Gln Arg Thr Ser Gln Asn Gly Val Ile Pro His

405 410 415

Gln Leu His Gln Lys Glu Leu Asp Glu Ile Ile Glu Asn Gln Lys Gln

420 425 430

Tyr Tyr Pro Phe Leu Ala Glu Pro Asn Pro Asn Glu Lys Arg Lys Pro

435 440 445

His Ala Gln Phe Lys Leu Asp Glu Leu Ile Ala Phe Lys Ile Pro Tyr

450 455 460

Tyr Val Gly Pro Leu Ile Thr Lys Glu Glu Gln Gln Ala Gln Ser Gly

465 470 475 480

Ala Lys Phe Ala Trp Met Lys Arg Lys Gln Asp Gly Val Ile Thr Pro

485 490 495

Trp Asn Phe Asp Glu Lys Val Asp Arg Met Ala Ser Ala Asn Glu Phe

500 505 510

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

820 825 830

Gly Lys Asp Ile Tyr Ser Gly Glu Thr Ile Asn Ile Asp Gln Leu Asn

835 840 845

Asn Tyr Asp Ile Asp His Ile Leu Pro Gln Ala Phe Ile Lys Asp Asp

850 855 860

Ser Leu Asp Asn Arg Val Leu Thr Ser Arg Asp Leu Asn Asn Gly Lys

865 870 875 880

Ser Asp Ser Val Pro Val Lys Gln Phe Gly Ala Asn Met Lys Ser Phe

885 890 895

Trp Met Arg Leu Gln Ala His Gly Leu Ile Ser Lys Arg Lys Leu Asn

900 905 910

Asn Leu Met Thr Asp Pro Asp Ser Ile Gly Lys Tyr Thr Met Gln Gly

915 920 925

Phe Val Arg Arg Gln Leu Val Glu Thr Ser Gln Val Ile Lys Leu Thr

930 935 940

Ala Asn Ile Leu Gly Ala Val Tyr Gly Glu Asn Thr Asp Ile Val Glu

945 950 955 960

Ile Pro Ala Lys Leu Thr His Gln Met Arg Glu Lys Phe Asn Leu Tyr

965 970 975

Lys Val Arg Glu Val Asn Asp Tyr His His Ala Phe Asp Ala Tyr Leu

980 985 990

Thr Thr Phe Val Gly Asn Tyr Leu Phe Lys Arg Tyr Pro Lys Leu Arg

995 1000 1005

Pro Tyr Phe Val Tyr Gly Asp Phe Lys Met Thr Asp Asn Ala Leu

1010 1015 1020

Lys Gly Met Arg Arg Phe Asn Phe Leu His Asp Leu Lys Asp Asp

1025 1030 1035

Glu Val Leu Val Asp Asn Glu Thr Gly Glu Val Leu Trp Glu Gly

1040 1045 1050

Gln Lys Ser Ile Glu Glu Leu Lys Lys Ile Tyr Gly Tyr Lys Phe

1055 1060 1065

Met Leu Thr Thr His Glu Ala Tyr Ile Asn Asn Gly Ala Met Tyr

1070 1075 1080

Asp Gln Thr Val Tyr Pro Ala Lys Ser Asn Lys Lys Leu Ile Pro

1085 1090 1095

Ser Lys Asn Asn Arg Pro Thr Asn Ile Tyr Gly Gly Phe Ser Gly

1100 1105 1110

Asn Lys Asp Ala Tyr Leu Ala Ile Val Lys Ile Leu Gly Lys Lys

1115 1120 1125

Gly Asp Thr Tyr Lys Val Val Gly Val Pro Arg Arg Phe Ser Asp
1130 1135 1140
Glu Leu Leu Asn Ala Tyr Asn Asp Asp Tyr Ser Glu Tyr Leu Asn
1145 1150 1155
Lys Leu His Asp Val Leu Val Thr Gln Phe Thr Lys Lys Gln Arg
1160 1165 1170
Asn Lys Lys Thr Gly Lys Ile Ile Lys Lys Ile Glu Asn Phe Glu
1175 1180 1185
Ile Val Val Gly Lys Val Asn Tyr Arg Gln Arg Ile Ile Asp Arg
1190 1195 1200
Asn Gln Ser Phe Met Leu Gly Ser Ser Thr Tyr Lys Tyr Asn Thr
1205 1210 1215
Lys Gln Leu Val Leu Thr Glu Ser Ser Met Lys Ile Leu Ser Asn
1220 1225 1230
Phe Lys His Ala Ser Asp Asp Asn Leu Asn Leu Val Tyr Asp Glu
1235 1240 1245
Ile Leu Glu Lys Val Asn Glu Ser Phe Glu Leu Tyr Asp Lys Asn
1250 1255 1260
Gly Phe Arg Lys Lys Leu Asn Asn Asn Ile Asp Lys Phe Arg Lys
1265 1270 1275
Leu Pro Cys Lys Ser Ile Phe Glu Lys Asn Lys Leu Ile Ser Val
1280 1285 1290
Gly Lys Cys Glu Ile Leu Phe Gln Ile Leu Asn Gly Leu His Ala
1295 1300 1305
Asn Ala Thr Met Gly Asp Leu Lys Gly Ile Gly Phe Ser Thr Pro
1310 1315 1320
Phe Gly Phe Met Gln Asp Ala Ser Gly Ile Lys Leu Asn Lys Asp
1325 1330 1335
Ser Met Leu Leu Tyr Gln Ser Pro Thr Gly Leu Phe Glu Arg Lys
1340 1345 1350
Val Lys Leu Ser Asp Leu
1355

<210> 139
<211> 1488
<212> PRT
<213> 魏斯氏菌属物种
<223> APG09344_RGN
<400> 139
Met Ala Lys Gln Lys Tyr Thr Val Gly Leu Asp Ile Gly Thr Asn Ser

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

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

625 630 635 640
Glu Lys Val Leu Pro Lys Asn Ser Leu Leu Tyr Gln Glu Phe Ile Val
645 650 655
Asn Asn Glu Leu Ile Ser Ser Gly Tyr Tyr Leu Asn Gly Arg Lys Glu
660 665 670
Tyr Phe Leu Pro Glu Gln Arg Gln Arg Ile Ile Thr Arg Leu Phe Lys
675 680 685
Arg Tyr Arg Lys Val Ser Ala Lys Gln Met Leu Asp Phe Leu Lys Asn
690 695 700
Glu Tyr Gln Ile Lys Leu Asp Asp Asp Pro Lys Lys Leu Phe Gly Leu
705 710 715 720
Asp Thr Phe Val Lys Val Pro Ser Tyr Asn Gly Ser Leu Ser Thr Val
725 730 735
Val Asp Leu Thr Asn Lys Met Gly Ile Ser Glu Asp Leu Ile Ala Asn
740 745 750
Tyr Pro Glu Arg Phe Glu Glu Ile Ile Lys Trp Gln Thr Ile Phe Glu
755 760 765
Asp Lys Lys Ile Leu Lys Lys Thr Ile Arg Glu Val Asn His Asn Lys
770 775 780
Trp Ser Gly Met Leu Thr Ala Glu Gln Val Asn Lys Leu Ser Lys Arg
785 790 795 800
His Tyr Thr Gly Trp Gly Arg Leu Ser Gln Lys Leu Leu Asp Gly Val
805 810 815
Lys Thr Ser Gln Gly Arg Thr Ile Ile Glu Ser Leu Arg Glu Gly Ile
820 825 830
Cys Asp Asn Phe Met Arg Leu Ile Glu Asp Glu Lys Ile Ala Glu Phe
835 840 845
Ile Ala Asn Ala Gln Val Ala Gly Ala Lys Asp Gly Val Leu Ala Tyr
850 855 860
Asp Leu Val Asp Gln Leu Pro Gly Ser Pro Ala Ile Lys Lys Gly Ile
865 870 875 880
Trp Gln Ser Leu Arg Val Leu Gln Glu Leu Glu His Tyr Leu Gly Arg
885 890 895
Asp Ala Ile Gly Lys Val Val Ile Glu Met Ser Arg Asp Asp Glu Ser
900 905 910
Ser Arg Arg Thr Lys Ala Arg Lys Gln Gln Leu Glu Glu Phe Tyr Lys
915 920 925
Lys Phe Lys Glu Ser Thr Gly Ser Asp Val Thr Ala Glu Leu Lys Arg
930 935 940

Glu Leu Ser Glu Lys Asp Ala Lys Glu Ile Asp Asp Glu Lys Val Phe
945 950 955 960
Leu Tyr Phe Leu Gln Asn Gly Lys Ser Met Tyr Ser Gly Asp Gln Leu
965 970 975
Tyr Leu Ser Arg Leu Ala Asp Tyr Gln Val Asp His Ile Val Pro Gln
980 985 990
Thr Tyr Ile Lys Asp Asp Ser Leu Asp Asn Lys Val Leu Val Thr Gln
995 1000 1005
Arg Asp Asn Gln Asn Lys Gly Gly Asp Thr Pro Ser His Thr Val
1010 1015 1020
Val Thr Arg Met Asn Ser Tyr Trp Asp Met Leu Ala Lys Asn Gly
1025 1030 1035
Gln Val Ser Pro Arg Lys Leu Ala Asn Leu Lys Lys Gly Thr Ile
1040 1045 1050
Thr Leu Lys Gln Arg Glu Gly Phe Ile Asn Arg Gln Leu Val Glu
1055 1060 1065
Asn Arg Gln Ile Ile Lys His Val Ala Asn Ile Leu Thr Asn Tyr
1070 1075 1080
Tyr Thr Gly Thr Asp Thr Leu Val Leu Thr Pro Lys Ser Gly Leu
1085 1090 1095
Thr Ser Gln Leu Arg Ser Gly Leu Ile Tyr Glu Leu Asn Pro Ala
1100 1105 1110
Phe Asp Glu Ala Lys Ala Leu Asp Arg Gly Lys Ala Tyr Gln Val
1115 1120 1125
Glu Arg Tyr Thr Thr Val Lys Leu His Asp Arg Phe Val Lys Asn
1130 1135 1140
Arg Lys Leu Asn Asp Tyr His His Ala His Asp Ala Tyr Leu Asn
1145 1150 1155
Ala Phe Val Ala Gln Tyr Val Tyr Gln Glu His Pro Glu Trp Lys
1160 1165 1170
Asn Ala Trp Val Tyr Gly Lys Tyr Pro Arg Asn Gly Gln Ala Asp
1175 1180 1185
Phe Gly Lys Trp Ala Thr Gln Arg Lys Gln Lys Ser Leu Gln Leu
1190 1195 1200
Leu Ser Ser Met Ala Asn Asp Val Trp Asn Leu Glu Asp Pro Asp
1205 1210 1215
Thr Met Glu Lys Thr Ile Leu Asn Arg Asp Glu Thr Phe Glu Gln
1220 1225 1230
Met Arg Gln Thr Leu Cys Tyr Arg Asn Ile Asn Ile Val Lys Lys

1235 1240 1245
Leu Glu Thr Gln Thr Gly Lys Phe Gly Asp Glu Ser Val Tyr Lys
1250 1255 1260
Lys Gly Asn Lys Ala Asp His Tyr Ser Leu Gly Leu Lys Arg Lys
1265 1270 1275
Tyr Pro Pro Gln Lys Tyr Gly Gly Thr Lys Gly Ala Ile Ser Ala
1280 1285 1290
Val Thr Ala Leu Val Lys Asp Asn Ser Gly Arg Val Ile Pro Val
1295 1300 1305
Ser Ile Ser Ala Ser Asp Tyr Asp Ser Tyr Gln Asn Ala Ser Asp
1310 1315 1320
Lys Leu Ala Trp Leu Leu Met Arg Lys Lys Asn Ile Ala Glu Ile
1325 1330 1335
Leu Val Ala Ser Leu Pro Lys Tyr Thr Lys Tyr Glu Leu Pro Ser
1340 1345 1350
Gly Ala Leu Arg Leu Leu Ala Ser Tyr Gln Glu Ala Gln Ser Ala
1355 1360 1365
Val Glu Leu Pro Met Leu Asn Leu Gly Leu Met Asp Asp Glu Ala
1370 1375 1380
Lys Val Leu Asn Leu Tyr Asp Gln Leu Ala Gln Phe Ile Ala Asp
1385 1390 1395
Asn Lys Leu Phe Thr Asp Lys Lys Ile Ala Leu Leu Lys Gly Glu
1400 1405 1410
Met Arg Val Ala Phe Glu Gly Leu Glu Glu Phe Lys Asp Lys Gln
1415 1420 1425
Lys Val Ile Asp Glu Leu Leu Gly Val Thr Asn Gly Ser Asn Gln
1430 1435 1440
Gly Leu Gln Ala Leu Asn Ser Ile Gly Leu Gly Thr Ser Asn Gln
1445 1450 1455
Arg Leu Lys Ser Gly Asn Thr Met Thr Asn Gly Val Thr Leu Ile
1460 1465 1470
Asn Glu Ser Val Thr Gly Leu Tyr Ser Thr Arg Thr Thr Tyr Asn
1475 1480 1485

<210> 140
<211> 1337
<212> PRT
<213> 肠球菌属物种
<223> APG07991_RGN
<400> 140

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

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

Glu Asp Ile Ile Lys Ile Leu Thr Ile Phe Glu Asp Arg Gln Arg Ile
625 630 635 640
Arg Thr Gln Leu Ser Thr Phe Lys Gly Gln Phe Ser Glu Glu Val Leu
645 650 655
Lys Lys Leu Glu Arg Lys His Tyr Thr Gly Trp Gly Arg Leu Ser Lys
660 665 670
Lys Leu Ile Asn Gly Ile Tyr Asp Lys Glu Ser Gly Lys Thr Ile Leu
675 680 685
Asp Tyr Leu Ile Lys Asp Asp Gly Val Ser Lys His Tyr Asn Arg Asn
690 695 700
Phe Met Gln Leu Ile Asn Asp Ser Gln Leu Ser Phe Lys Asn Ala Ile
705 710 715 720
Gln Lys Ala Gln Ser Ser Glu His Glu Glu Thr Leu Ser Glu Thr Val
725 730 735
Asn Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Tyr Gln Ser
740 745 750
Leu Lys Ile Val Asp Glu Leu Val Ala Ile Met Gly Tyr Ala Pro Lys
755 760 765
Arg Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Ser Thr Gly
770 775 780
Lys Arg Arg Ser Ile Gln Arg Leu Lys Ile Val Glu Lys Ala Met Ala
785 790 795 800
Glu Ile Gly Ser Asn Leu Leu Lys Glu Gln Pro Thr Thr Asn Glu Gln
805 810 815
Leu Arg Asp Thr Arg Leu Phe Leu Tyr Tyr Met Gln Asn Gly Lys Asp
820 825 830
Met Tyr Thr Gly Asp Glu Leu Ser Leu His Arg Leu Ser His Tyr Asp
835 840 845
Ile Asp His Ile Ile Pro Gln Ser Phe Met Lys Asp Asp Ser Leu Asp
850 855 860
Asn Leu Val Leu Val Gly Ser Thr Glu Asn Arg Gly Lys Ser Asp Asp
865 870 875 880
Val Pro Ser Lys Glu Val Val Lys Lys Met Lys Ala Tyr Trp Glu Lys
885 890 895
Leu Tyr Ala Ala Gly Leu Ile Ser Gln Arg Lys Phe Gln Arg Leu Thr
900 905 910
Lys Gly Glu Gln Gly Gly Leu Thr Leu Glu Asp Lys Ala His Phe Ile
915 920 925
Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys Asn Val Ala Gly

930 935 940

Ile Leu Asp Gln Arg Tyr Asn Ala Lys Ser Lys Glu Lys Lys Val Gln
945 950 955 960

Ile Ile Thr Leu Lys Ala Ser Leu Thr Ser Gln Phe Arg Ser Ile Phe
965 970 975

Gly Leu Tyr Lys Val Arg Glu Val Asn Asp Tyr His His Gly Gln Asp
980 985 990

Ala Tyr Leu Asn Cys Val Val Ala Thr Thr Leu Leu Lys Val Tyr Pro
995 1000 1005

Asn Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Pro Lys Phe Gln
1010 1015 1020

Ala Phe Lys Glu Asn Lys Ala Thr Ala Lys Ala Ile Ile Tyr Thr
1025 1030 1035

Asn Leu Leu Arg Phe Phe Thr Glu Asp Glu Pro Arg Phe Thr Lys
1040 1045 1050

Asp Gly Glu Ile Leu Trp Ser Asn Ser Tyr Leu Lys Thr Ile Lys
1055 1060 1065

Lys Glu Leu Asn Tyr His Gln Met Asn Ile Val Lys Lys Val Glu
1070 1075 1080

Val Gln Lys Gly Gly Phe Ser Lys Glu Ser Ile Lys Pro Lys Gly
1085 1090 1095

Pro Ser Asn Lys Leu Ile Pro Val Lys Asn Gly Leu Asp Pro Gln
1100 1105 1110

Lys Tyr Gly Gly Phe Asp Ser Pro Val Val Ala Tyr Thr Val Leu
1115 1120 1125

Phe Thr His Glu Lys Gly Lys Lys Pro Leu Ile Lys Gln Glu Ile
1130 1135 1140

Leu Gly Ile Thr Ile Met Glu Lys Thr Arg Phe Glu Gln Asn Pro
1145 1150 1155

Ile Leu Phe Leu Glu Glu Lys Gly Phe Leu Arg Pro Arg Val Leu
1160 1165 1170

Met Lys Leu Pro Lys Tyr Thr Leu Tyr Glu Phe Pro Glu Gly Arg
1175 1180 1185

Arg Arg Leu Leu Ala Ser Ala Lys Glu Ala Gln Lys Gly Asn Gln
1190 1195 1200

Met Val Leu Pro Glu His Leu Leu Thr Leu Leu Tyr His Ala Lys
1205 1210 1215

Gln Cys Leu Leu Pro Asn Gln Ser Glu Ser Leu Ala Tyr Val Glu
1220 1225 1230

Gln His Gln Pro Glu Phe Gln Glu Ile Leu Glu Arg Val Val Asp
1235 1240 1245
Phe Ala Glu Val His Thr Leu Ala Lys Ser Lys Val Gln Gln Ile
1250 1255 1260
Val Lys Leu Phe Glu Ala Asn Gln Thr Ala Asp Val Lys Glu Ile
1265 1270 1275
Ala Ala Ser Phe Ile Gln Leu Met Gln Phe Asn Ala Met Gly Ala
1280 1285 1290
Pro Ser Thr Phe Lys Phe Phe Gln Lys Asp Ile Glu Arg Ala Arg
1295 1300 1305
Tyr Thr Ser Ile Lys Glu Ile Phe Asp Ala Thr Ile Ile Tyr Gln
1310 1315 1320
Ser Thr Thr Gly Leu Tyr Glu Thr Arg Arg Lys Val Val Asp
1325 1330 1335

<210> 141
<211> 1340
<212> PRT
<213> 肠球菌属物种
<223> APG01868_RGN
<400> 141

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

145 150 155 160
Leu Leu Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Glu Leu Asn Thr
165 170 175
Glu Asn Ser Ser Val Thr Glu Thr Phe Arg Gln Phe Leu Ser Thr Tyr
180 185 190
Asn Gln Gln Phe Ser Glu Ala Gly Asp Lys Gln Thr Glu Lys Leu Asp
195 200 205
Glu Ala Val Asp Cys Ser Phe Val Phe Thr Glu Lys Met Ser Lys Thr
210 215 220
Lys Lys Ala Glu Thr Leu Leu Lys Tyr Phe Pro His Glu Lys Ser Asn
225 230 235 240
Gly Tyr Leu Ser Gln Phe Ile Lys Leu Met Val Gly Asn Gln Gly Asn
245 250 255
Phe Lys Asn Val Phe Gly Leu Glu Glu Glu Ala Lys Leu Gln Phe Ser
260 265 270
Lys Glu Thr Tyr Glu Glu Asp Leu Gln Glu Leu Leu Glu Lys Ile Gly
275 280 285
Asp Asp Tyr Ile Asp Leu Phe Val Gln Ala Lys Asn Val Tyr Asp Ala
290 295 300
Val Leu Leu Ser Glu Ile Leu Ser Asp Ser Thr Lys Asn Thr Arg Ala
305 310 315 320
Lys Leu Ser Ala Gly Met Ile Arg Arg Tyr Asp Ala His Lys Glu Asp
325 330 335
Leu Val Leu Leu Lys Arg Phe Val Lys Glu Asn Leu Pro Lys Lys Tyr
340 345 350
Arg Ala Phe Phe Gly Asp Asn Ser Val Asn Gly Tyr Ala Gly Tyr Ile
355 360 365
Glu Gly His Ala Thr Gln Glu Asp Phe Tyr Lys Phe Val Lys Lys Glu
370 375 380
Leu Thr Gly Ile Arg Gly Ser Glu Val Phe Leu Thr Lys Ile Glu Gln
385 390 395 400
Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Val Ile Pro
405 410 415
His Gln Ile His Leu Thr Glu Leu Arg Ala Ile Ile Ala Asn Gln Lys
420 425 430
Lys His Tyr Pro Phe Leu Lys Glu Glu Gln Glu Lys Leu Glu Ser Leu
435 440 445
Leu Thr Phe Lys Ile Pro Tyr Tyr Val Gly Pro Leu Ala Lys Lys Gln
450 455 460

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

770 775 780
Gly Leu Lys Ser Ser Arg Pro Arg Leu Lys Ala Leu Glu Glu Ser Leu
785 790 795 800
Lys Asp Phe Gly Ser Gln Leu Leu Lys Glu Tyr Pro Thr Asp Asn Ser
805 810 815
Ser Leu Gln Lys Asp Arg Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg
820 825 830
Asp Met Tyr Thr Gly Ala Pro Leu Asp Ile His Arg Leu Ser Asp Tyr
835 840 845
Asp Ile Asp His Ile Ile Pro Arg Ser Phe Thr Thr Asp Asn Ser Ile
850 855 860
Asp Asn Lys Val Leu Val Ser Ser Lys Glu Asn Arg Leu Lys Lys Asp
865 870 875 880
Asp Val Pro Ser Glu Lys Val Val Lys Lys Met Arg Ser Phe Trp Tyr
885 890 895
Asp Leu Tyr Ser Ser Lys Leu Ile Ser Lys Arg Lys Leu Asp Asn Leu
900 905 910
Thr Lys Ile Lys Leu Thr Glu Glu Asp Lys Ala Gly Phe Ile Lys Arg
915 920 925
Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gly Ile Leu
930 935 940
His His Arg Phe Asn Lys Ala Glu Asp Thr Asn Glu Pro Ile Arg Lys
945 950 955 960
Val Arg Ile Ile Thr Leu Lys Ser Ala Leu Val Ser Gln Phe Arg Asn
965 970 975
Arg Phe Gly Ile Tyr Lys Val Arg Glu Ile Asn Glu Tyr His His Ala
980 985 990
His Asp Ala Tyr Leu Asn Gly Val Ile Ala Leu Ala Leu Leu Lys Lys
995 1000 1005
Tyr Pro Gln Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Leu Lys
1010 1015 1020
Phe Asn Ala His Lys Ala Asn Lys Ala Thr Val Lys Lys Glu Phe
1025 1030 1035
Tyr Ser Asn Ile Met Lys Phe Phe Glu Ser Asp Thr Pro Val Cys
1040 1045 1050
Asp Glu Asn Gly Glu Ile Phe Trp Asp Lys Ser Lys Ser Ile Ala
1055 1060 1065
Gln Val Lys Lys Val Ile Asn His His His Met Asn Ile Val Lys
1070 1075 1080

Lys Thr Glu Ile Gln Lys Gly Gly Phe Ser Lys Glu Thr Val Glu
1085 1090 1095
Pro Lys Lys Asp Ser Ser Lys Leu Leu Pro Arg Lys Asn Asn Trp
1100 1105 1110
Asp Pro Ala Lys Tyr Gly Gly Leu Gly Ser Pro Asn Val Ala Tyr
1115 1120 1125
Thr Val Ala Phe Thr Tyr Glu Lys Gly Lys Ala Arg Lys Arg Thr
1130 1135 1140
Asn Ala Leu Glu Gly Ile Thr Ile Met Glu Arg Glu Ala Phe Glu
1145 1150 1155
Gln Ser Pro Val Leu Phe Leu Lys Asn Lys Gly Tyr Glu Gln Ala
1160 1165 1170
Glu Ile Glu Met Lys Leu Pro Lys Tyr Ala Leu Phe Glu Leu Glu
1175 1180 1185
Asn Gly Arg Lys Arg Met Val Ala Ser Asn Lys Glu Ala Gln Lys
1190 1195 1200
Ala Asn Ser Phe Leu Leu Pro Glu His Leu Val Thr Leu Leu Tyr
1205 1210 1215
His Ala Lys Gln Tyr Asp Glu Ile Ser His Lys Glu Ser Phe Asp
1220 1225 1230
Tyr Val Asn Glu His His Lys Glu Phe Ser Glu Val Phe Ala Arg
1235 1240 1245
Val Leu Glu Phe Ala Gly Lys Tyr Thr Leu Ala Glu Lys Asn Ile
1250 1255 1260
Glu Lys Leu Glu Lys Ile Tyr Lys Glu Asn Gln Thr Asp Asp Leu
1265 1270 1275
Ala Lys Leu Ala Ser Ser Phe Val Asn Leu Met Gln Phe Asn Ala
1280 1285 1290
Met Gly Ala Pro Ala Asp Phe Lys Phe Phe Asp Val Thr Ile Pro
1295 1300 1305
Arg Lys Arg Tyr Thr Ser Leu Thr Glu Ile Trp Gln Ser Thr Ile
1310 1315 1320
Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Arg Met
1325 1330 1335
Gly Lys
1340

<210> 142
<211> 1327
<212> PRT

<213> 肠球菌属物种

<223> APG02998_RGN

<400> 142

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

Val Phe Asp Ala Ala Lys Ser Val Tyr Asp Ala Ile Ser Leu Ala Asn
290 295 300
Ile Leu Lys Val Thr Asp Ala Ala Thr Arg Ala Lys Leu Ser Ser Ser
305 310 315 320
Met Val Ala Arg Phe Lys Glu His Lys Glu Asp Leu Lys Ala Leu Lys
325 330 335
Gln Phe Val Arg Asp Asn Leu Pro Asp Glu Tyr Asp Asp Leu Phe Lys
340 345 350
Asn Lys Lys Val Ala Gly Tyr Ala Gly Tyr Ile Asp Gly Asp Ala Thr
355 360 365
Gln Glu Glu Phe Tyr Lys Tyr Leu Lys Lys Ala Leu Ala Lys Thr Thr
370 375 380
Gly Ala Glu Tyr Phe Leu Ala Lys Ile Glu Gln Glu Asp Phe Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Val Ile Pro Tyr Gln Leu His Leu
405 410 415
Glu Glu Leu Lys Ala Ile Ile Arg Asn Gln Lys Pro Tyr Tyr Pro Phe
420 425 430
Leu Asp Glu Glu Lys Ile Ser Gln Leu Met Thr Phe Arg Ile Pro Tyr
435 440 445
Tyr Val Gly Pro Leu Ala Gln Gly Gln Gly Asn Phe Ala Trp Leu Thr
450 455 460
Arg Lys Glu Ser Gly Lys Ile Thr Pro Trp Asn Leu Asn Glu Lys Val
465 470 475 480
Asp Ile Glu Lys Ser Ala Thr Asp Phe Val Glu Arg Met Thr Asn Asn
485 490 495
Asp Ser Tyr Leu Pro Met Glu Lys Val Leu Pro Lys His Ser Leu Leu
500 505 510
Tyr Glu Lys Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Arg Tyr Met
515 520 525
Asp Asp Asn Gly Arg Val Gln Asn Phe Ser Ser Lys Glu Lys Gln Gln
530 535 540
Ile Val Asn Asp Leu Phe Lys Gln Gln Arg Lys Val Lys Lys Glu Ile
545 550 555 560
Leu Glu Ala Phe Ile Lys Asn Glu Tyr Gly Ile Glu Asn Pro Lys Val
565 570 575
Glu Gly Ile Glu Lys Ala Phe Asn Ala Ser Leu Gly Thr Tyr His Asp
580 585 590
Leu Val Lys Leu Gly Ile Arg Pro Ala Leu Phe Glu Gln Pro Glu Tyr

595 600 605

Glu Gln Gln Phe Glu Gln Ile Val Lys Ile Leu Thr Val Phe Glu Asp
610 615 620

Arg Lys Met Arg Arg Lys Gln Leu Glu Gln Phe Ser Asp Leu Leu Thr
625 630 635 640

Glu Asp Glu Gln Lys Gln Leu Glu Arg Lys His Tyr Lys Gly Trp Gly
645 650 655

Arg Leu Ser Ala Lys Leu Ile His Gly Ile Val Asp Gln Lys Thr Gln
660 665 670

Lys Thr Ile Leu Asp Tyr Leu Ile Asp Asp Asp Asp Leu Pro Arg Asn
675 680 685

Arg Asn Arg Asn Phe Met Gln Leu Ile Asn Asp Glu Asn Leu Ser Phe
690 695 700

Lys Glu Glu Ile Glu Lys Ile Ala Phe Asp Asn Asp Lys Ser Thr Glu
705 710 715 720

Glu Ile Val Gln Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
725 730 735

Leu Gln Ser Leu Lys Ile Val Glu Glu Ile Ile Glu Ile Met Gly Glu
740 745 750

Leu Pro Thr Asn Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr
755 760 765

Ala Gln Gly Asn Arg Ala Ser Gln Ala Arg Met Lys Tyr Leu Glu Glu
770 775 780

Ser Ile Lys Lys Leu Gly Ser Ser Ile Leu Glu Asp Glu Pro Ile Ser
785 790 795 800

Lys Asp Ala Asn Leu Leu Arg Asn Asp Arg Leu Phe Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Thr Gly Asp Glu Leu Asp Ile Asn Asn
820 825 830

Leu Ser Ser Tyr Asp Ile Asp His Ile Ile Pro Gln Ser Phe Val Lys
835 840 845

Asp Asp Ser Ile Asp Asn Arg Val Leu Thr Thr Gln Arg Ile Asn Arg
850 855 860

Gly Lys Ser Asn Thr Val Pro Ala Glu Ser Val Val Lys Asp Met Arg
865 870 875 880

Ala Thr Trp Glu Arg Leu Tyr Ala Ser Asn Leu Ile Ser Lys Lys Lys
885 890 895

Leu Asn Asn Leu Met Lys Ala Val Asn Gly Gly Leu Thr Glu Glu Asp
900 905 910

Lys Ala Tyr Phe Ile Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925
Lys Asn Val Ala Gln Ile Leu His Gln Lys Tyr Asn Asn Glu Gln Ser
930 935 940
Ser Glu Lys Pro Val Arg Val Val Thr Leu Lys Ser Ala Leu Ala Ser
945 950 955 960
Gln Phe Arg Lys Asp Phe Ser Leu Tyr Lys Ile Arg Glu Leu Asn Asp
965 970 975
Tyr His His Ala Gln Asp Ala Tyr Leu Asn Gly Val Ile Ala Gln Ala
980 985 990
Leu Leu Lys Val Tyr Pro Lys Leu Glu Pro Glu Phe Val Tyr Gly Glu
995 1000 1005
Tyr Gln Lys Val Ser Ile Arg Ala Leu Asn Lys Ala Thr Ala Lys
1010 1015 1020
Lys Glu Thr Tyr Ser Asn Ile Met Lys Phe Phe Thr Asn Asp Glu
1025 1030 1035
Val Leu Thr Asn Lys Glu Thr Gly Glu Val Leu Trp Asn Arg Gln
1040 1045 1050
Arg Asp Ile Lys Thr Ile Lys Lys Val Met Asn Tyr His Gln Met
1055 1060 1065
Asn Ile Val Lys Lys Val Glu Ile Gln Thr Gly Arg Phe Thr Lys
1070 1075 1080
Glu Ser Ile Gln Pro Lys Gly Pro Ser Lys Lys Leu Ile Ala Arg
1085 1090 1095
Lys Asn Asn Trp Ala Pro Val Asn Tyr Gly Gly Phe Asp Ser Pro
1100 1105 1110
Thr Val Ala Tyr Ser Val Ile Ile Thr His Glu Lys Gly Lys Ala
1115 1120 1125
Ala Lys Val Val Gln Gln Leu Val Gly Ile Lys Ile Leu Glu Arg
1130 1135 1140
Gln Ala Phe Glu Gln Asn Glu Val Ala Phe Leu Glu Gly Lys Gly
1145 1150 1155
Phe Ile His Pro Lys Val Gln Leu Lys Leu Pro Lys Tyr Ser Leu
1160 1165 1170
Tyr Gln Phe Ala Asp Gly Arg Arg Arg Leu Leu Ala Ser Ala Asp
1175 1180 1185
Glu Ala Gln Lys Gly Asn Gln Met Val Leu Pro Thr Glu Leu Ile
1190 1195 1200
Glu Leu Leu Tyr His Ala Lys His Val Asn Asp Ser Ser Gly Glu

1205 1210 1215

Ser Leu Thr Tyr Leu Glu Asn His Arg Thr Lys Phe Ala Glu Leu

1220 1225 1230

Leu Glu Val Ile Leu Gln Phe Thr Asn Arg Tyr Val Gly Ala Glu

1235 1240 1245

Asp Arg Gln Lys Thr Val Arg Asp Leu Tyr Glu Lys Asn Gln Asp

1250 1255 1260

Ala Asp Met Lys Glu Leu Ala Pro Ser Phe Val Gln Leu Leu Gln

1265 1270 1275

Leu Asn Lys Gln Gly Ala Pro Ala Ser Phe Lys Phe Phe Gly Glu

1280 1285 1290

Thr Ile Glu Arg Ser Arg Tyr Lys Asn Met Thr Glu Leu Asn Lys

1295 1300 1305

Ala Thr Phe Ile Asn Gln Ser Ile Thr Gly Leu Tyr Glu Thr Gln

1310 1315 1320

Arg Arg Leu Val

1325