(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110845622 A
(43)申请公布日 2020.02.28

(21)申请号 201810955073.4

(22)申请日 2018.08.21

(71)申请人 康码(上海)生物科技有限公司
地址 201321 上海市浦东新区芙蓉花路500弄8号楼3楼

(72)发明人 郭敏 代田纯 欧阳冬冬 丁晓辉 杨宁 于雪

(51)Int.Cl.
C07K 19/00(2006.01)
C12N 15/62(2006.01)
C12P 21/00(2006.01)

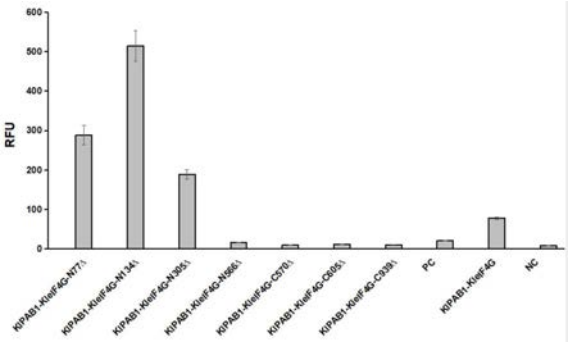
权利要求书2页 说明书14页
序列表55页 附图8页

(54)发明名称

不同结构域缺失融合蛋白的制备及其在提高蛋白质合成的应用

(57)摘要

本发明提供了不同结构域缺失的eIF4G与Pab1融合蛋白及其在改变蛋白质合成中的应用，具体地，不同结构域缺失融合蛋白可改变体外蛋白翻译效率。其中，eIF4G元件的RNA1和/或PABP结构域缺失能够显著的提高蛋白的表达。此外，本发明提供融合蛋白的制备方法及相应的体外蛋白合成体系及方法。



CN 110845622 A

1.一种融合蛋白,其特征在于,所述融合蛋白具有式Ia或式Ib结构:

S-A-B-C (Ia)

S-C-B-A (Ib);

式中,

A为Pab1元件;

B为无或连接肽;

C为不同结构域缺失的eIF4G元件;

S为无或任选的信号肽;以及

各“-”为肽键。

2.根据权利要求1所述的融合蛋白,其特征在于,所述的eIF4G元件的RNA1和/或PABP结构域缺失。

3.一种分离的多核苷酸,其特征在于,所述的多核苷酸编码权利要求1或2所述的融合蛋白。

4.一种载体或载体组合物,其特征在于,所述载体或载体组合物含有权利要求3所述的多核苷酸。

5.一种基因工程细胞,其特征在于,所述基因工程细胞基因组中整合有权利要求3所述的多核苷酸或含有权利要求4所述的载体或载体组合物。

6.一种权利要求5所述基因工程细胞的细胞提取物,其特征在于,所述细胞提取物包含权利要求1或2所述融合蛋白。

7.一种用于表达外源蛋白的体外蛋白合成体系,其特征在于,所述合成体系主要包括:

(a)权利要求6所述细胞提取物;

(b)用于合成蛋白质的底物;

(c)用于合成RNA的底物;

(d)无或含有RNA聚合酶。

8.如权利要求7所述体外蛋白合成体系,其特征在于,所述合成体系还包括(e)权利要求1或2所述的融合蛋白。

9.一种用于表达外源蛋白的体外蛋白合成体系,其特征在于,所述合成体系主要包括:

(a)细胞提取物;

(b)用于合成蛋白质的底物;

(c)用于合成RNA的底物;

(d)无或含有RNA聚合酶;

(e)权利要求1或2所述的融合蛋白。

10.如权利要求7-9任一项所述的体外蛋白合成体系,其特征在于,所述合成体系还包括:额外添加的eIF4G蛋白。

11.如权利要求7-9任一项所述的体外蛋白合成体系,其特征在于,所述合成体系还包括选自下组的一种或多种组分:镁离子、钾离子、缓冲剂、能量再生系统、聚乙二醇、二硫苏糖醇(DTT)和任选的溶剂,所述溶剂为水或水性溶剂。

12.如权利要求10所述的体外蛋白质合成体系,其特征在于,所述eIF4G蛋白由组成型或诱导型启动子诱导表达。

13.一种生产权利要求1或2所述融合蛋白的方法,其特征在于,所述方法主要包括以下步骤:

培养权利要求5所述的基因工程细胞,从而表达出权利要求1或2所述的融合蛋白。

14.根据权利要求13所述的方法,其特征在于,所述方法还包括:分离所述融合蛋白。

15.一种权利要求1或2所述的融合蛋白的用途,其特征在于,所述融合蛋白用于在蛋白质合成体系中改变内源或外源蛋白的表达。

16.一种合成外源蛋白的方法,其特征在于,所述方法包括以下步骤:

(i)提供一体外蛋白合成体系,其中所述的合成体系为权利要求7-12的任一合成体系;和

(ii)在适合表达蛋白的条件下,在编码所述外源蛋白的DNA模板存在下,孵育所述体外蛋白合成体系,从而表达所述的外源蛋白。

17.根据权利要求16所述的体外蛋白合成方法,其特征在于,所述方法还包括:(iii)分离或检测所述外源蛋白。

不同结构域缺失融合蛋白的制备及其在提高蛋白质合成的应用

技术领域

[0001] 本发明涉及基因工程领域，具体地，涉及不同结构域缺失融合蛋白的制备及其在提高蛋白质合成的应用。

[0002] 本申请是在申请号为2017106425174的在先申请的基础上，进一步得出；下文中所提在先申请均为申请号为2017106425174的专利申请。

背景技术

[0003] 蛋白质是细胞中的重要分子，几乎参与了细胞所有功能的执行。蛋白的序列和结构不同，决定了其功能的不同。在细胞内，蛋白可以作为酶类催化各种生化反应，可以作为信号分子协调生物体的各种活动，可以支持生物形态，储存能量，运输分子，并使生物体运动。在生物医学领域，蛋白质抗体作为靶向药物，是治疗癌症等疾病的重要手段。

[0004] 在细胞中，蛋白质翻译的调节在应对营养缺失等外界压力，细胞发育与分化等很多过程中发挥重要作用。蛋白质翻译的四个过程包括翻译起始、翻译延伸、翻译终止和核糖体再循环，其中翻译起始是受调控最多的一个过程。在翻译起始阶段，核糖体小亚基(40S)结合$(tRNA)_i^{Met}$，并在翻译起始因子的作用下识别mRNA 5'末端。小亚基向下游移动，并在起始密码子(ATG)位置与核糖体大亚基(60S)结合，形成完整核糖体，并进入翻译延伸阶段。

[0005] 在快速分裂的酵母细胞中，蛋白的合成速率大约为13,000个/秒。在体内，蛋白的合成速率受到核糖体数目的限制，细胞的平均核糖体数目约为200,000个，mRNA分子的数目约为15,000-60,000个。

[0006] 目前，经常实验的商业化体外蛋白表达系统包括大肠杆菌系统(E.coli extract，ECE)、兔网织红细胞(Rabbit reticulocyte Lysate，RRL)、麦胚(Wheat germ extract，WGE)、昆虫(Insect cell extract，ICE)和人源系统。无论自然界中细胞内还是细胞外的人造蛋白质合成体系都具有效率低，速度慢的特点，极大的限制了蛋白质合成的应用。因此，本领域迫切需要开发一种可以有效改变体外蛋白质合成效率的体外蛋白质合成体系。

发明内容

[0007] 本发明的目的在于提供一种可以有效改变体外蛋白质合成效率的体外蛋白质合成体系。本发明的不同结构域缺失融合蛋白，尤其是eIF4G元件的RNA1和/或PABP结构域缺失可显著提高无细胞体外蛋白质合成体系(尤其是酵母体外蛋白质合成体系)的体外蛋白质合成能力。

[0008] 本发明第一方面提供了一种融合蛋白，所述融合蛋白具有式Ia或式Ib结构：

[0009] S-A-B-C (Ia)

[0010] S-C-B-A (Ib)；

[0011] 式中，

[0012] A为Pab1元件；

[0013] B为无或连接肽；

[0014] C为不同结构域缺失的eIF4G元件；

[0015] S为无或任选的信号肽；以及

[0016] 各“-”为肽键。

[0017] 在另一优选例中，所述式Ia或Ib为从N端至C端的结构。

[0018] 在另一优选例中，所述元件A包括野生型和突变型的Pab1序列。

[0019] 在另一优选例中，所述的Pab1为来自酵母的Pab1。

[0020] 在另一优选例中，元件A具有SEQ ID NO.:4所示的序列或其活性片段，或者具有与SEQ ID NO:4所示氨基酸序列≥85％同源性(优选地，≥90％的同源性；等优选地≥95％的同源性；最优选地，≥97％的同源性，如98％以上，99％以上)且具有与SEQ ID NO.:4序列相同活性的多肽。

[0021] 在另一优选例中，所述元件C为eIF4G元件的RNA1和/或PABP结构域缺失。

[0022] 在另一优选例中，所述的eIF4G为来自酵母的eIF4G。

[0023] 在另一优选例中，元件C具有SEQ ID NO.:2所示的序列或其活性片段中不同结构域缺失的序列。

[0024] 在另一优选例中，所述融合蛋白是重组蛋白，较佳地为酵母表达的重组蛋白。

[0025] 在另一优选例中，所述酵母选自下组：克鲁维酵母、酿酒酵母之一或其组合。

[0026] 在另一优选例中，所述克鲁维酵母选自下组：乳酸克鲁维酵母、马克斯克鲁维酵母之一或其组合。

[0027] 在另一优选例中，所述元件A衍生自酵母的Pab1蛋白。

[0028] 在另一优选例中，所述元件C衍生自酵母的eIF4G蛋白。

[0029] 在另一优选例中，所述的连接肽的长度为0-50氨基酸，较佳地为10-40个氨基酸，更佳地为15-25个氨基酸。

[0030] 在另一优选例中，所述融合蛋白具有选自下组的一个或多个特性：

[0031] (a)改变外源蛋白表达效率；优选的，提高外源蛋白表达效率；

[0032] (b)改变体外翻译效率；优选的，提高体外翻译效率。

[0033] 在另一优选例中，所述外源蛋白选自下组：荧光素蛋白、荧光素酶(如萤火虫荧光素酶)、绿色荧光蛋白、黄色荧光蛋白、氨酰tRNA合成酶、甘油醛-3-磷酸脱氢酶、过氧化氢酶、肌动蛋白、抗体的可变区域、萤光素酶突变、α-淀粉酶、肠道菌素A、丙型肝炎病毒E2糖蛋白、胰岛素前体、干扰素αA、白细胞介素-1β、溶菌酶素、血清白蛋白、单链抗体段(scFV)、甲状腺素运载蛋白、酪氨酸酶、木聚糖酶、或其组合。

[0034] 本发明第二方面提供了一种分离的多核苷酸，所述的多核苷酸编码本发明第一方面所述的融合蛋白。

[0035] 在另一优选例中，所述的多核苷酸选自下组：DNA序列、RNA序列。

[0036] 在另一优选例中，所述的DNA序列选自下组：基因组序列、cDNA序列。

[0037] 在另一优选例中，所述的多核苷酸为mRNA或cDNA，并且所述多核苷酸具有式II所示结构：

[0038] A1-C1 (式I I)

[0039] 式中，

[0040] A1为编码上述A元件的核苷酸序列；

[0041] C1为编码上述C元件的核苷酸序列；

[0042] “-”为元件A1和元件C1之间的连接键。

[0043] 在另一优选例中，所述元件A1具有SEQ ID NO.:3所示的序列。

[0044] 在另一优选例中，所述元件C1具有SEQ ID NO.:1所示的序列中相应结构域所对应的核苷酸序列缺失的序列。

[0045] 本发明第三方面提供了一种载体或载体组合物，所述载体或载体组合物含有本发明第二方面所述的多核苷酸。

[0046] 本发明第四方面提供了一种基因工程细胞，所述基因工程细胞的基因组中整合有本发明第二方面所述的多核苷酸，或者所述基因工程细胞中含有本发明第三方面所述的载体或载体组合物。

[0047] 在另一优选例中，所述基因工程细胞是通过转入本发明第二方面所述的多核苷酸或本发明第三方面所述的表达载体并经同源重组而形成的，从而在基因组或染色体中整合有本发明第一方面所述融合蛋白的编码序列。

[0048] 在另一优选例中，所述基因工程细胞包括原核细胞、真核细胞。

[0049] 在另一优选例中，所述真核细胞包括高等真核细胞。

[0050] 在另一优选例中，所述基因工程细胞选自下组：人源细胞(如Hela细胞)、中国仓鼠卵巢细胞、昆虫细胞、麦胚细胞、兔网织红细胞、酵母细胞、或其组合。

[0051] 在另一优选例中，所述基因工程细胞为酵母细胞。

[0052] 在另一优选例中，所述酵母细胞选自下组：酿酒酵母、克鲁维酵母属酵母之一或其组合。

[0053] 在另一优选例中，所述克鲁维酵母属酵母选自下组：乳酸克鲁维酵母、马克斯克鲁维酵母之一或其组合。

[0054] 本发明第五方面提供了一种本发明第四方面所述基因工程细胞的细胞提取物，所述细胞提取物包含本发明第一方面所述融合蛋白。

[0055] 本发明第六方面提供了一种用于表达外源蛋白的体外蛋白合成体系，所述合成体系包括：

[0056] (a)本发明第五方面所述细胞提取物；

[0057] (b)用于合成蛋白质的底物；

[0058] (c)用于合成RNA的底物；

[0059] (d)不含或含有RNA聚合酶。

[0060] 在另一优选例中，所述合成体系还包括(e)本发明第一方面所述的融合蛋白。

[0061] 在另一优选例中，所述合成体系还包括：(f)额外添加的eIF4G蛋白。

[0062] 在另一优选例中，所述合成体系还包括选自下组的一种或多种组分：镁离子、钾离子、缓冲剂、能量再生系统、聚乙二醇、二硫苏糖醇(DTT)和任选的溶剂，所述溶剂为水或水性溶剂。

[0063] 在另一优选例中，所述细胞提取物来源于酵母细胞，所述酵母细胞选自下组的一种或多种来源的酵母：酿酒酵母、毕氏酵母、克鲁维酵母之一或其组合；较佳地，所述的酵母细胞包括：克鲁维酵母，更佳地为乳酸克鲁维酵母。

[0064] 在另一优选例中,所述细胞提取物为对酵母细胞的水性提取物。

[0065] 在另一优选例中,所述细胞提取物不含酵母内源性的长链核酸分子。

[0066] 在另一优选例中,所述的合成RNA的底物包括:核苷单磷酸、核苷三磷酸之一或其组合。

[0067] 在另一优选例中,所述的合成蛋白质的底物包括:1-20种天然氨基酸、以及非天然氨基酸。

[0068] 在另一优选例中,所述镁离子来源于镁离子源,所述镁离子源选自下组:醋酸镁、谷氨酸镁之一或其组合。

[0069] 在另一优选例中,所述钾离子来源于钾离子源,所述钾离子源选自下组:醋酸钾、谷氨酸钾之一或其组合。

[0070] 在另一优选例中,所述能量再生系统选自下组:磷酸肌酸/磷酸肌酸酶系统、糖酵解途径及其中间产物能量系统之一或其组合。

[0071] 在另一优选例中,所述缓冲剂选自下组:4-羟乙基哌嗪乙磺酸、三羟甲基氨基甲烷之一或其组合。

[0072] 在另一优选例中,所述eIF4G蛋白由组成型或诱导型启动子诱导表达。

[0073] 在另一优选例中,所述组成型或诱导型启动子来源于酵母。

[0074] 在另一优选例中,所述酵母选自下组:克鲁维酵母(Kluyveromyces)、酿酒酵母(Saccharomyces cerevisiae)或其组合。

[0075] 在另一优选例中,所述组成型或诱导型启动子选自下组:pScTEF1、pScPGK1、pKlTEF1、pKlPGK1、pScADH1、pScTPI1、pScTDH3、pKlADH1、pKlTPI1、pKlTDH3或其组合。其中p为promoter的简写,为启动子;Sc为酿酒酵母来源,Kl为克鲁维酵母来源,后续的字母为相应的启动子的基因名称。

[0076] 本发明第七方面提供了一种用于表达外源蛋白的体外蛋白合成体系,所述合成体系包括:

[0077] (a)细胞提取物;

[0078] (b)用于合成蛋白质的底物;

[0079] (c)用于合成RNA的底物;

[0080] (d)不含或含有RNA聚合酶;

[0081] (e)本发明第一方面所述的融合蛋白。

[0082] 在另一优选例中,所述合成体系还包括:(f)额外添加的eIF4G蛋白。

[0083] 在另一优选例中,所述合成体系还包括选自下组的一种或多种组分:镁离子、钾离子、缓冲剂、能量再生系统、聚乙二醇、二硫苏糖醇(DTT)和任选的溶剂,所述溶剂为水或水性溶剂。在另一优选例中,所述细胞提取物来源于酵母细胞,所述酵母细胞选自下组的一种或多种来源的酵母:酿酒酵母、毕氏酵母、克鲁维酵母之一或其组合;较佳地,所述的酵母细胞包括:克鲁维酵母,更佳地为乳酸克鲁维酵母。

[0084] 在另一优选例中,所述细胞提取物为对酵母细胞的水性提取物。

[0085] 在另一优选例中,所述细胞提取物不含酵母内源性的长链核酸分子。

[0086] 在另一优选例中,所述的合成RNA的底物包括:核苷单磷酸、核苷三磷酸之一或其组合。

[0087] 在另一优选例中，所述的合成蛋白质的底物包括：1-20种天然氨基酸、以及非天然氨基酸。

[0088] 在另一优选例中，所述镁离子来源于镁离子源，所述镁离子源选自下组：醋酸镁、谷氨酸镁之一或其组合。

[0089] 在另一优选例中，所述钾离子来源于钾离子源，所述钾离子源选自下组：醋酸钾、谷氨酸钾之一或其组合。

[0090] 在另一优选例中，所述能量再生系统选自下组：磷酸肌酸/磷酸肌酸酶系统、糖酵解途径及其中间产物能量系统之一或其组合。

[0091] 在另一优选例中，所述缓冲剂选自下组：4-羟乙基哌嗪乙磺酸、三羟甲基氨基甲烷之一或其组合。

[0092] 在另一优选例中，所述eIF4G蛋白由组成型或诱导型启动子诱导表达。

[0093] 在另一优选例中，所述组成型或诱导型启动子来源于酵母。

[0094] 在另一优选例中，所述酵母选自下组：克鲁维酵母(Kluyveromyces)、酿酒酵母(Saccharomyces cerevisiae)或其组合。

[0095] 在另一优选例中，所述组成型或诱导型启动子选自下组：pScTEF1、pScPGK1、pKlTEF1、pKlPGK1、pScADH1、pScTPI1、pScTDH3、pKlADH1、pKlTPI1、pKlTDH3或其组合。其中p为promoter的简写，为启动子；Sc为酿酒酵母来源，Kl为克鲁维酵母来源，后续的字母为相应的启动子的基因名称。

[0096] 本发明第八方面提供了一种生产本发明第一方面所述融合蛋白的方法，主要包括：培养本发明第四方面所述的基因工程细胞，从而表达出本发明第一方面所述的融合蛋白。和

[0097] 在另一优选例中，所述方法还包括：分离所述融合蛋白的步骤。

[0098] 本发明第九方面提供了一种本发明第一方面所述融合蛋白的用途，用于在蛋白质合成体系中改变内源或外源蛋白的表达。

[0099] 在另一优选例中，所述合成体系还包括额外的eIF4G蛋白。

[0100] 本发明第十方面提供了一种合成外源蛋白的方法，包括：

[0101] (i)提供一体外蛋白合成体系，其中所述的合成体系为第六方面或第七方面的任一合成体系；和

[0102] (ii)在适合表达蛋白的条件下，在编码所述外源蛋白的DNA模板存在下，孵育所述体外蛋白合成体系，从而表达所述的外源蛋白。

[0103] 在另一优选例中，所述的融合蛋白为额外添加的。

[0104] 在另一优选例中，所述的融合蛋白与所述的酵母体外蛋白质合成体系中的其他蛋白为来自相同酵母的提取物。

[0105] 在另一优选例中，所述的条件为反应温度为20-37℃，反应时间为1-72h。

[0106] 在另一优选例中，所述步骤(ii)还包括步骤(iii)：检测外源蛋白活性的表达活性Q1，并且在步骤(ii)相同条件下孵育野生型酵母菌株，检测所述外源蛋白的活性Q2，如果Q1显著高于Q2，则表明外源蛋白的表达效率显著提高。

[0107] 在另一优选例中，所述“显著高于”指Q1/Q2≥2，较佳地≥3，更佳地≥4。

[0108] 在另一优选例中，所述体外蛋白合成体系为真核体外蛋白合成体系，较佳的，为酵

母体外蛋白合成体系，更佳的，为基因改造的克鲁维酵母体外蛋白合成体系(优选乳酸克鲁维酵母体外蛋白合成体系)。

[0109] 在另一优选例中，所述外源蛋白的编码序列来自原核生物、真核生物。

[0110] 在另一优选例中，所述外源蛋白的编码序列来自动物、植物、病原体。

[0111] 在另一优选例中，所述外源蛋白的编码序列来自哺乳动物，较佳地灵长动物，啮齿动物，包括人、小鼠、大鼠。

[0112] 在另一优选例中，所述的外源蛋白的编码序列选自下组：编码荧光素蛋白、荧光素酶(如萤火虫荧光素酶)、绿色荧光蛋白、黄色荧光蛋白、氨酰tRNA合成酶、甘油醛-3-磷酸脱氢酶、过氧化氢酶、肌动蛋白、抗体的可变区域的外源DNA、萤光素酶突变体的DNA之一或其组合。

[0113] 在另一优选例中，所述外源蛋白选自下组：荧光素蛋白、或荧光素酶(如萤火虫荧光素酶)、绿色荧光蛋白、黄色荧光蛋白、氨酰tRNA合成酶、甘油醛-3-磷酸脱氢酶、过氧化氢酶、肌动蛋白、抗体的可变区域、萤光素酶突变、α-淀粉酶、肠道菌素A、丙型肝炎病毒E2糖蛋白、胰岛素前体、干扰素αA、白细胞介素-1β、溶菌酶素、血清白蛋白、单链抗体段(scFV)、甲状腺素运载蛋白、酪氨酸酶、木聚糖酶之一或其组合。

[0114] 应理解，在本发明范围内中，本发明的上述各技术特征和在下文(如实施例)中具体描述的各技术特征之间都可以互相组合，从而构成新的或优选的技术方案。限于篇幅，在此不再一一累述。

附图说明

[0115] 图1为pKM-KlPab1-KleIF4G-N77Δ的质粒图谱。

[0116] 图2为pKM-KlPab1-KleIF4G-N134Δ的质粒图谱。

[0117] 图3为pKM-KlPab1-KleIF4G-N305Δ的质粒图谱。

[0118] 图4为pKM-KlPab1-KleIF4G-N566Δ的质粒图谱。

[0119] 图5为pKM-KlPab1-KleIF4G-C570Δ的质粒图谱。

[0120] 图6为pKM-KlPab1-KleIF4G-C605Δ的质粒图谱。

[0121] 图7为pKM-KlPab1-KleIF4G-C939Δ的质粒图谱。

[0122] 图8为体外翻译活性测定示意图。

具体实施方式

[0123] 在先申请中，本发明人已经发现，将一个eIF4G拷贝与内源Pab1蛋白连接，形成一个新的融合蛋白后，能够显著提高酵母菌株的体外翻译效率。在本申请中，发明人构建了一系列不同结构域缺失的eIF4G拷贝，并与内源Pab1蛋白连接，形成一系列融合蛋白。体外蛋白翻译实验表明，不同结构域缺失的改造都会改变蛋白的表达，其中3个含有新的融合蛋白结构的酵母菌株活性高于含有完整KlPAB1-KleIF4G融合蛋白的菌株。

[0124] eIF4G蛋白

[0125] 真核生物中，多种翻译起始因子参与蛋白质翻译起始过程(表1)。其中eIF4F负责“帽子结构”的识别以及下游翻译起始因子和核糖体的招募。eIF4F由三个蛋白质亚基组成：eIF4E、eIF4G和eIF4A。eIF4E特异性结合“帽子结构”，将eIF4F锚定在mRNA 5’端非翻译区；

eIF4A是一种RNA解旋酶;eIF4G则几乎是整个翻译起始过程的支架蛋白,能与多种翻译起始因子相互作用,在下游因子招募过程中具有重要作用。酿酒酵母的eIF4G蛋白大体包括6个结构域(RNA1、RNA2、RNA3、PABP、eIF4E、HEAT/eIF4A),分别与mRNA,PABP,eIF4E和eIF4A蛋白结合。

[0126] 表1 翻译起始因子

翻译起始因子	亚基	基因	蛋白长度(AA)
eIF1		*SUI1*	108
eIF1A		*TIF11*	153
eIF2	α	*SUI2*	304
	β	*SUI3*	285
	γ	*GCD11*	527
eIF2B	α	*GCN3*	305
	β	*GCD7*	381
	γ	*GCD1*	578
	δ	*GCD2*	651
	ε	*GCD6*	712
eIF3	a	*RPG1/TIF32*	964
	b	*PRT1*	763
	c	*NIP1*	812
	g	*TIF35*	274
	i	*TIF34*	347
	j	*HCR1*	265
eIF4A		*TIF1*	395
		TIF2	395
eIF4B		*TIF3/STM1*	436
eIF4E		*CDC33*	213
eIF4G		*TIF4631*	952
		TIF4632	914
eIF5		*TIF5*	405
eIF5B		*FUN12*	1002

[0127]

[0128] 其中,乳酸克鲁维酵母(Kluyveromyces lactis)eIF4G的核苷酸序列如SEQ ID NO.:1所示;所述eIF4G的蛋白序列如SEQ ID NO.:2所示。

[0129] Pab1元件(Pab1蛋白)

[0130] Pab1是一个71kDa的RNA结合蛋白,由4个RRM(RNA recognition motif 1-4)结构域和1个CTD结构域(C末端结构域)组成。每个RRM结构域中都包含2个保守的RNP结构(RNP1/2),负责与RNA的结合。

[0131] 乳酸克鲁维酵母(Kluyveromyces lactis)的Pab1核苷酸序列如SEQ ID NO.:3所示;所述Pab1的蛋白序列如SEQ ID NO.:4所示。

[0132] 融合蛋白

[0133] 如本文所用,术语“融合蛋白”、“Pab1-eIF4G融合蛋白”可互换使用,指Pab1元件与

eIF4G元件或与不同结构域缺失的eIF4G元件融合形成的融合蛋白。在本发明的融合蛋白中,Pab1元件与eIF4G元件之间可以含有或不含有连接肽或柔性接头。此外,所述融合蛋白可以含有或不含有起始的Met;可以含有或不含有信号肽;以及含有(如6His、8His等)或不含有标签序列。

[0134] 体外蛋白质合成体系

[0135] 体外蛋白表达系统包括大肠杆菌系统(E.coli extract,ECE)、兔网织红细胞(Rabbit reticulocyte Lysate,RRL)、麦胚(Wheat germ extract,WGE)、昆虫(Insect cell extract,ICE)和人源系统。一种典型的体外蛋白质合成体系是酵母体外蛋白质合成体系。

[0136] 酵母(yeast)兼具培养简单、高效蛋白质折叠、和翻译后修饰的优势。其中酿酒酵母(Saccharomyces cerevisiae)和毕氏酵母(Pichia pastoris)是表达复杂真核蛋白质和膜蛋白的模式生物,酵母也可作为制备体外翻译系统的原料。

[0137] 克鲁维酵母(Kluyveromyces)是一种子囊孢子酵母,其中的马克斯克鲁维酵母(Kluyveromyces marxianus)和乳酸克鲁维酵母(Kluyveromyces lactis)是工业上广泛使用的酵母。例如乳酸克鲁维酵母是一种能够以乳酸作为其唯一的碳源和能源的酵母。与其他酵母相比,乳酸克鲁维酵母具有许多优点,如超强的分泌能力,良好的大规模发酵特性、食品安全的级别及同时具有蛋白翻译后修饰的能力等,其作为宿主系统表达药用蛋白也已显示出巨大的潜力。

[0138] 在本发明中,酵母体外蛋白质合成体系不受特别限制,一种优选的酵母体外蛋白质合成体系为克鲁维酵母表达系统(更佳地,乳酸克鲁维酵母表达系统)。

[0139] 在一优选实施方式中,本发明的酵母体外蛋白质合成体系为基因改造后的乳酸克鲁维酵母表达系统。

[0140] 在一优选实施方式中,本发明提供了一种表达外源蛋白的体外蛋白合成体系,所述合成体系主要包括:

[0141] (a)细胞提取物;

[0142] (b)用于合成蛋白质的底物;

[0143] (c)用于合成RNA的底物;

[0144] (d)无或含有RNA聚合酶。

[0145] 在另一优选例中,所述酵母提取物为含有融合蛋白的酵母提取物或为任意酵母提取物。

[0146] 在另一优选例中,所述合成体系还包括选自下组的一种或多种组分:镁离子、钾离子、缓冲剂、能量再生系统、聚乙二醇、二硫苏糖醇(DTT)和任选的溶剂,所述溶剂为水或水性溶剂。在另一优选例中,所述细胞提取物来源于酵母细胞,所述酵母细胞选自下组的一种或多种来源的酵母:酿酒酵母、毕氏酵母、克鲁维酵母之一或其组合;较佳地,所述的酵母细胞包括:克鲁维酵母,更佳地为乳酸克鲁维酵母。

[0147] 在另一优选例中,所述细胞提取物为对酵母细胞的水性提取物。

[0148] 在另一优选例中,所述细胞提取物不含酵母内源性的长链核酸分子。

[0149] 在另一优选例中,所述的合成RNA的底物包括:核苷单磷酸、核苷三磷酸之一或其组合。

[0150] 在另一优选例中，所述的合成蛋白质的底物包括：1-20种天然氨基酸、以及非天然氨基酸。

[0151] 在另一优选例中，所述RNA聚合酶没有特别限制，可以选自一种或多种RNA聚合酶，典型的RNA聚合酶为T7 RNA聚合酶。

[0152] 在另一优选例中，所述镁离子来源于镁离子源，所述镁离子源选自下组：醋酸镁、谷氨酸镁之一或其组合。

[0153] 在另一优选例中，所述钾离子来源于钾离子源，所述钾离子源选自下组：醋酸钾、谷氨酸钾之一或其组合。

[0154] 在另一优选例中，所述能量再生系统选自下组：磷酸肌酸/磷酸肌酸酶系统、糖酵解途径及其中间产物能量系统之一或其组合。

[0155] 在另一优选例中，所述缓冲剂选自下组：4-羟乙基哌嗪乙磺酸、三羟甲基氨基甲烷之一或其组合。

[0156] 在优选例中，所述体外蛋白质合成体系含有聚乙二醇(PEG)或其类似物。聚乙二醇或其类似物的浓度没有特别限制，通常，聚乙二醇或其类似物的浓度(w/v)为0.1-8％，较佳地，0.5-4％，更佳地，1-2％，以所述蛋白合成体系的总重量计。代表性的PEG选自下组：PEG3000、PEG8000、PEG6000、PEG3350之一或其组合。

[0157] 在另一优选例中，所述聚乙二醇包括分子量(Da)为200-10000的聚乙二醇，如PEG200、400、1500、2000、4000、6000、8000、10000等，较佳地，分子量为3000-10000的聚乙二醇。

[0158] 在一特别优选的实施方式中，本发明提供的体外蛋白合成体系包括：酵母细胞提取物，4-羟乙基哌嗪乙磺酸，醋酸钾，醋酸镁，腺嘌呤核苷三磷酸(ATP)，鸟嘌呤核苷三磷酸(GTP)，胞嘧啶核苷三磷酸(CTP)，胸腺嘧啶核苷三磷酸(TTP)，氨基酸混合物，磷酸肌酸，二硫苏糖醇(DTT)，磷酸肌酸激酶，RNA聚合酶。

[0159] 在本发明中，所述的细胞提取物不含完整的细胞，典型的细胞提取物包括用于蛋白翻译的核糖体、转运RNA、氨酰tRNA合成酶、蛋白质合成需要的起始因子和延伸因子以及终止释放因子。此外，细胞提取物中还含有一些源自细胞的细胞质中的其他蛋白，尤其是可溶性蛋白。

[0160] 在本发明中，所述的细胞提取物所含蛋白含量为20-100mg/ml，较佳为50-100mg/ml。所述的测定蛋白含量方法为考马斯亮蓝测定方法。

[0161] 在本发明中，所述的细胞提取物的制备方法不受限制，一种优选的制备方法包括以下步骤：

[0162] (i)提供细胞；

[0163] (ii)对细胞进行洗涤处理，获得经洗涤的细胞；

[0164] (iii)对经洗涤的细胞进行细胞破碎处理，从而获得细胞粗提物；

[0165] (iv)对所述细胞粗提物进行固液分离，获得液体部分，即为细胞提取物。

[0166] 在本发明中，所述的固液分离方式不受特别限制，一种优选的方式为离心。

[0167] 在本发明中，所述离心条件不受特别限制，一种优选的离心条件为5000-100000×g，较佳地，8000-30000×g。

[0168] 在本发明中，所述离心时间不受特别限制，一种优选的离心时间为0.5min-2h，较

佳地，20min-50min。

[0169] 在本发明中，所述离心的温度不受特别限制，优选的，所述离心在1-10℃下进行，较佳地，在2-6℃下进行。

[0170] 在本发明中，所述的洗涤处理方式不受特别限制，一种优选的洗涤处理方式为采用洗涤液在pH为7-8(较佳地，7.4)下进行处理，所述洗涤液没有特别限制，典型的所述洗涤液选自下组：4-羟乙基哌嗪乙磺酸钾、醋酸钾、醋酸镁、或其组合。

[0171] 在本发明中，所述细胞破碎处理的方式不受特别限制，一种优选的所述的细胞破碎处理包括高压破碎、冻融(如液氮低温)破碎。

[0172] 所述体外蛋白质合成体系中的核苷三磷酸混合物为腺嘌呤核苷三磷酸、鸟嘌呤核苷三磷酸、胞嘧啶核苷三磷酸和尿嘧啶核苷三磷酸。在本发明中，各种单核苷酸的浓度没有特别限制，通常每种单核苷酸的浓度为0.5-5mM，较佳地为1.0-2.0mM。

[0173] 所述体外蛋白质合成体系中的氨基酸混合物可包括天然或非天然氨基酸，可包括D型或L型氨基酸。代表性的氨基酸包括(但并不限于)20种天然氨基酸：甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苯丙氨酸、脯氨酸、色氨酸、丝氨酸、酪氨酸、半胱氨酸、蛋氨酸、天冬酰胺、谷氨酰胺、苏氨酸、天冬氨酸、谷氨酸、赖氨酸、精氨酸和组氨酸。每种氨基酸的浓度通常为0.01-0.5mM，较佳地0.02-0.2mM，如0.05、0.06、0.07、0.08mM。

[0174] 在优选例中，所述体外蛋白质合成体系还含有蔗糖。蔗糖的浓度没有特别限制，通常，蔗糖的浓度为0.03-40wt％，较佳地，0.08-10wt％，更佳地，0.1-5wt％，以所述蛋白合成体系的总重量计。

[0175] 一种特别优选的体外蛋白质合成体系，除了酵母细胞提取物之外，还含有以下组分：22mM pH为7.4的4-羟乙基哌嗪乙磺酸，30-150mM醋酸钾，1.0-5.0mM醋酸镁，1.5-4mM核苷三磷酸混合物，0.08-0.24mM的氨基酸混合物，25mM磷酸肌酸，1.7mM二硫苏糖醇，0.27mg/mL磷酸肌酸激酶，1％-4％聚乙二醇，0.5％-2％蔗糖，0.027-0.054mg/mL T7 RNA聚合酶。

[0176] 在一优选实施方式中，本发明的酵母体外蛋白质合成体系还含有：(e)本发明第一方面所述的融合蛋白，即Pab1-eIF4G融合蛋白。

[0177] 在一优选实施方式中，本发明的酵母体外蛋白质合成体系还包括eIF4G蛋白；其中，本发明的eIF4G蛋白通过来源于酵母(如酿酒酵母、克鲁维酵母等)组成型或诱导型的启动子(如pScTEF1、pScPGK1、pKlTEF1、pKlPGK1、pScADH1、pScTPI1、pScTDH3、pKlADH1、pKlTPI1、pKlTDH3等)进行诱导表达。

[0178] 在本发明中，含有本发明融合蛋白的酵母体外蛋白质合成体系可显著增强体外蛋白质的合成能力。此外，将本发明的融合蛋白和eIF4G蛋白联合使用的酵母体外蛋白质合成体系具有更高的体外蛋白质的合成能力。

[0179] 本发明的主要优点包括：

[0180] (a)本发明通过基因改造技术，借助高效的细胞转化平台，对细胞内基因进行改造，从而提高了翻译系统的蛋白合成效率。

[0181] (b)本发明对申请人之前发现的融合蛋白进一步改造，从而通过本发明的融合蛋白可改变体外蛋白质的合成能力。

[0182] 实施例1通过基因改造提高蛋白质合成的理论模型

[0183] 在先申请中，本发明人通过CRISPR-Cas9基因编辑技术，将一个完整的eIF4G蛋白

连接到K.lactis中内源PAB1(KlPAB1)蛋白C端,显著提高了无细胞体外翻译系统的效率。eIF4G蛋白含有多个结构域,分别与不同的RNA或者蛋白元件互作。其中有些结构域可能并不参与到体外翻译的过程中,所以本专利构建了一系列不同结构域缺失的eIF4G(KleIF4G-N77Δ(RNA1domain缺失),KleIF4G-N134Δ(RNA1+RNA1与PABP之间的区域缺失),KleIF4G-N305Δ(PABP+PABP之前的区域缺失),KleIF4G-N566Δ(eIF4E+eIF4E之前的区域缺失),KleIF4G-C570Δ(RNA2+之后的区域缺失),KleIF4G-C605Δ(HEAT/eIF4A+之后的区域缺失),KleIF4G-C939Δ(RNA3domain缺失)),并与PAB1蛋白连接,形成新的融合蛋白,测试其体外翻译的效率。

[0184] 对比例为PC组和NC组,其中,PC是不经过任何改造的野生型酵母菌株;NC为用水替换外源蛋白基因DNA模板。

[0185] 根据酿酒酵母中eIF4G基因报道的结构域序列,利用Blast预测克鲁维酵母eIF4G基因中的相应的6个结构域的序列。以在先申请(申请号为2017106425174)中的pKM-KlPab1-KleIF4G-DD质粒为模板,以不同的引物扩增不同结构域缺失的片段。不同结构域缺失的eIF4G的碱基序列和氨基酸序列参见序列表(SEQ ID NO.:5-SEQ ID NO.:18),在Pab1与eIF4G之间可以设置连接序列(linker),也可以不设置连接序列。

[0186] 实施例2不同结构域缺失的质粒构建

[0187] 本实施例质粒的构建是在在先申请已构建质粒的基础上完成,已构建质粒的构建方法详见在先申请(申请号为2017106425174)。

[0188] 1.KlPAB1-KleIF4G-N77Δ供体DNA质粒构建及扩增

[0189] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板,以引物PF:TTGGTGGAGGTGGATCTAACCAACCAGCGTACGGTG(SEQ ID NO.:19)和引物PR:AGATCCACCTCCACCAACAGTAG(SEQ ID NO.:20)进行PCR扩增。将扩增产物17μL,1μL DpnI,2μL 10×digestion buffer混合,37℃温浴3h。将DpnI处理后产物10μL加入100μL DH5α感受态细胞中,冰上放置30min,42℃热激45s后,加入1mL LB液体培养基37℃振荡培养1h,涂布于Amp抗性LB固体培养,37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-KlPab1-KleIF4G-N77Δ(图1)。

[0190] 以pKM-KlPab1-KleIF4G-N77Δ质粒为模板,以引物F:GTAAAACGACGGCCAGT(SEQ ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增,得到第一线性供体DNA。

[0191] 2.KlPAB1-KleIF4G-N134Δ供体DNA质粒构建及扩增

[0192] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板,以引物PF:TTGGTGGAGGTGGATCTGCTGTTTCAGCTAAACCAGCG(SEQ ID NO.:23)和引物PR:AGATCCACCTCCACCAACAGTAG(SEQ ID NO.:20)进行PCR扩增。将扩增产物17μL,1μL DpnI,2μL 10×digestion buffer混合,37℃温浴3h。将DpnI处理后产物10μL加入100μL DH5α感受态细胞中,冰上放置30min,42℃热激45s后,加入1mL LB液体培养基37℃振荡培养1h,涂布于Amp抗性LB固体培养,37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-KlPab1-KleIF4G-N134Δ(图2)。

[0193] 以pKM-KlPab1-KleIF4G-N134Δ质粒为模板,以引物F:GTAAAACGACGGCCAGT(SEQ ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增,得到第二线性供体DNA。

[0194] 3.KlPAB1-KleIF4G-N305Δ供体DNA质粒构建及扩增

[0195] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板，以引物PF:TTGGTGGAGGTGGATCTACTTTGGCCGAAAAATTGAGACTTAAGAG(SEQ ID NO.:24)和引物PR:AGATCCACCTCCACCAACAGTAG(SEQ ID NO.:20)进行PCR扩增。将扩增产物17μL，1μL DpnI，2μL 10×digestion buffer混合，37℃温浴3h。将DpnI处理后产物10μL加100μL DH5α感受态细胞中，冰上放置30min，42℃热激45s后，加入1mL LB液体培养基37℃振荡培养1h，涂布于Amp抗性LB固体培养，37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养，PCR检测阳性并测序确认后，提取质粒保存，命名为pKM-KlPab1-KleIF4G-N305Δ(图3)。

[0196] 以pKM-KlPab1-KleIF4G-N305Δ质粒为模板，以引物F:GTAAAACGACGGCCAGT(SEQ ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增，得到第三线性供体DNA。

[0197] 4.KlPAB1-KleIF4G-N566Δ供体DNA质粒构建及扩增

[0198] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板，以引物PF:TTGGTGGAGGTGGATCTGCTGTTTCAGCTAAACCAGCG(SEQ ID NO.:25)和引物PR:AGATCCACCTCCACCAACAGTAG(SEQ ID NO.:20)进行PCR扩增。将扩增产物17μL，1μL DpnI，2μL 10×digestion buffer混合，37℃温浴3h。将DpnI处理后产物10μL加入100μL DH5α感受态细胞中，冰上放置30min，42℃热激45s后，加入1mL LB液体培养基37℃振荡培养1h，涂布于Amp抗性LB固体培养，37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养，PCR检测阳性并测序确认后，提取质粒保存，命名为pKM-KlPab1-KleIF4G-N566Δ(图4)。

[0199] 以pKM-KlPab1-KleIF4G-N566Δ质粒为模板，以引物F:GTAAAACGACGGCCAGT(SEQ ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增，得到第四线性供体DNA。

[0200] 5.KlPAB1-KleIF4G-C570Δ供体DNA质粒构建及扩增

[0201] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板，以引物PF:TAAACTTGATTTTTTGACCTTGATCTTCATCTTGTCC(SEQ ID NO.:26)和引物PR:GAAGATCAAGGTCAAAAAATCAAGTTTAATCTTCGCCTCTTCCGCTTG(SEQ ID NO.:27)进行PCR扩增。将扩增产物17μL，1μL DpnI，2μL 10×digestion buffer混合，37℃温浴3h。将DpnI处理后产物10μL加入100μL DH5α感受态细胞中，冰上放置30min，42℃热激45s后，加入1mL LB液体培养基37℃振荡培养1h，涂布于Amp抗性LB固体培养，37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养，PCR检测阳性并测序确认后，提取质粒保存，命名为pKM-KlPab1-KleIF4G-C570Δ(图5)。

[0202] 以pKM-KlPab1-KleIF4G-C570Δ质粒为模板，以引物F:GTAAAACGACGGCCAGT(SEQ ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增，得到第五线性供体DNA。

[0203] 6.KlPAB1-KleIF4G-C605Δ供体DNA质粒构建及扩增

[0204] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板，以引物PF:TAAACTTGATTTTTTGACCTTGATCTTCATCTTGTCC(SEQ ID NO.:26)和引物PR:GATCAAGGTCAAAAAATCAAGTTTATCTATTAGATTTTCTATCATCCCTCTTTGAC(SEQ ID NO.:28)进行PCR扩增。将扩增产物17μL，1μL DpnI，2μL 10×digestion buffer混合，37℃温浴3h。将DpnI处理后产物10μL加入100μL DH5α感受态细胞中，冰上放置30min，42℃热激45s后，加入1mL LB液体培养基37℃振荡培养1h，涂布于Amp抗性LB固体培养，37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养，PCR检测阳性并测序确认后，提取质粒保存，命名为pKM-KlPab1-KleIF4G-C605Δ(图6)。

[0205] 以pKM-KlPab1-KleIF4G-C605Δ质粒为模板，以引物F:GTAAAACGACGGCCAGT(SEQ

ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增,得到第六线性供体DNA。

[0206] 7.KlPAB1-KleIF4G-C939Δ供体DNA质粒构建及扩增

[0207] 以在先申请中的pKM-KlPab1-KleIF4G-DD质粒为模板,以引物PF:TAAACTTGATTTTTTGACCTTGATCTTCATCTTGTCC(SEQ ID NO.:26)和引物PR:GAAGATCAAGGTCAAAAAATCAAGTTTATCTTTCCTCCAAAGCCCTCTTCAAG(SEQ ID NO.:29)进行PCR扩增。将扩增产物17μL,1μL DpnI,2μL 10×digestion buffer混合,37℃温浴3h。将DpnI处理后产物10μL加入100μL DH5α感受态细胞中,冰上放置30min,42℃热激45s后,加入1mL LB液体培养基37℃振荡培养1h,涂布于Amp抗性LB固体培养,37℃倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-KlPab1-KleIF4G-C939Δ(图7)。

[0208] 以pKM-KlPab1-KleIF4G-C939Δ质粒为模板,以引物F:GTAAAACGACGGCCAGT(SEQ ID NO.:21)和R:CAGGAAACAGCTATGAC(SEQ ID NO.:22)进行扩增,得到第七线性供体DNA。

[0209] 实施例3乳酸克鲁维酵母转化及阳性鉴定

[0210] i.将乳酸克鲁维酵母菌液在YPD固体培养基上划线并挑取单克隆,于25mL2×YPD液体培养基中振荡培养过夜,取2mL菌液于50mL液体2×YPD培养基中继续振荡培养2-8h。20℃条件下3000g离心5min收集酵母细胞,加入500μL无菌水重悬,同样条件下离心收集细胞。配制感受态细胞溶液(5%v/v甘油,10%v/v DMSO)并将酵母细胞溶解于500μL该溶液中。分装50μL至1.5mL离心管中,-80℃保存。

[0211] 将感受态细胞置于37℃融化15-30s,13000g离心2min并去除上清。配制转化缓冲液:PEG3350(50%(w/v))260μL,LiAc(1.0M)36μL,carrier DNA(5.0m g/mL)20μL,Cas9/gRNA质粒15μL,线性供体DNA 10μL,加入无菌水至最终体积360μL。热激后,13000g离心30s去除上清。加入1mL YPD液体培养基,培养2-3h,吸取200μL涂布于固体YPD(200μg/mL G418)培养基,培养2-3天至单菌落出现。

[0212] ii.在乳酸克鲁维酵母转化后的平板上挑取10-20个单克隆,置于1mL YPD(200μg/mL G418)液体培养基中振荡培养过夜,以菌液为模板,以引物KlPAB1-CICF1(KlPAB1序列内引物):TCTCCAGAAGAAGCTACCAAGGCTA(SEQ ID NO.:30)和引物KleIF4G-CICR1(KleIF4G序列内引物):TTCTCTTCGACAGCCTTCTTAGCAG(SEQ ID NO.:31);KleIF4G-CICF1(KleIF4G序列内引物):TACCCAAGTGACATTACGCCTCC(SEQ ID NO.:32)和KlPAB1-CICR1(KlPAB1序列内引物):TTGGAAGACCCCATTTTCATAGGGA(SEQ ID NO.:33)进行PCR扩增,对KlPAB1位点KleIF4G插入进行PCR检测,PCR结果阳性并经测序鉴定的菌株,确定为阳性菌株。

[0213] 实施例4改造菌株体外翻译活性测定

[0214] 将基因改造后的乳酸克鲁维酵母菌株制备成体外蛋白质合成体系,并加入增强绿色荧光蛋白(EGFP)基因DNA模板以测定改造菌株的蛋白翻译能力。将上述反应体系置于20-30℃的环境中,静置孵育约2-6h。反应结束后,立即放置于Envision 2120多功能酶标仪(Perkin Elmer),读数,检测EGFP信号强弱,相对荧光单位值(Relative Fluorescence Unit,RFU)作为活性单位。

[0215] 同样的,PC为未经改造的野生酵母菌株,将其制备成体外蛋白质合成体系,按照同样的方法进行测定蛋白翻译能力;NC是在制备的体外蛋白合成体系中不加入EGFP基因DNA,而加入相应体积的水。

[0216] 在改造的结构中,不同结构域改变的融合蛋白相较于KlPAB1-KleIF4G结构均有所变化(具体测定值见表2),KleIF4G N端结构域缺失的新融合蛋白KlPAB1-KleIF4G-N77Δ,KlPAB1-KleIF4G-N134Δ及KlPAB1-KleIF4G-N305Δ均表现出优于融合蛋白KlPAB1-KleIF4G的蛋白翻译效率。其中KlPAB1-KleIF4G-N134Δ最优,其体外蛋白翻译效率是KlPAB1-KleIF4G结构的约6.6倍,是野生型酵母菌株PC的25倍。这表明对KleIF4G的改造能够有效增强酵母体外蛋白质合成体系合成蛋白质的效率(图8)。

[0217]

不同结构域缺失	Data1	Data2	Data3	平均活性
KlPab1-KleIF4G-N77Δ	316	269	279	288
KlPab1-KleIF4G-N134Δ	470	536	537	514.33
KlPab1-KleIF4G-N305Δ	176	190	200	188.67
KlPab1-KleIF4G-N566Δ	15	17	15	15.67
KlPab1-KleIF4G-C570Δ	9	10	10	9.67
KlPab1-KleIF4G-C605Δ	10	11	12	11
KlPab1-KleIF4G-C939Δ	11	10	10	10.33
KlPab1_KleIF4G	75	77	81	77.67
PC	20	20	21	20.33
NC	7	9	9	8.33

[0218] 表2

[0219] 其中,PC:未经改造的野生酵母菌株;NC:表示negative control,阴性对照。

[0220] 上述实验结果表明:通过对乳酸克鲁维酵母KleIF4G基因的结构域进行改造,其与KlPab1形成的融合蛋白能够改变酵母体外蛋白质合成体系产生蛋白质的效率;尤其是eIF4G元件的RNA1和/或PABP结构域缺失,能够显著性的提高外源蛋白的表达。

[0221] 在本发明提及的所有文献都在本申请中引用作为参考,就如同每一篇文献被单独引用作为参考那样。此外应理解,在阅读了本发明的上述讲授内容之后,本领域技术人员可以对本发明作各种改动或修改,这些等价形式同样落于本申请所附权利要求书所限定的范围。

序列表

<110> 康码(上海)生物科技有限公司

<120> 不同结构域缺失融合蛋白的制备及其在提高蛋白质合成的应用

<130> 2018

<141> 2018-08-21

<160> 33

<170> SIPOSequenceListing 1.0

<210> 1

<211> 3066

<212> DNA

<213> 乳酸克鲁维酵母(Kluyveromyces lactis)

<400> 1

atgggcgaac ctacatccga tcagcaacca gctgttgaag ctccagttgt gcaggaggag 60
acaaccagtt ctccgcaaaa aaacagtgga tatgtcaaga atactgctgg aagcggtgct 120
cctagaaatg ggaaatatga tggtaacagg aagaactcta ggccttataa ccaaagaggt 180
aacaacaaca ataataatgg ttcttcctcg aataagcact atcaaaagta taaccaacca 240
gcgtacggtg tttctgcggg atacattccg aactacggcg tatcggcaga gtacaaccct 300
ctgtactata accagtacca acagcagcaa cagctgtacg ctgctgctta ccagactcca 360
atgagcggac aaggttatgt ccccccagta gtgtctccag ctgctgtttc agctaaacca 420
gcgaaggttg agattactaa caagtctggt gaacacatag atattgcttc cattgctcat 480
ccacatactc attctcattc tcaatctcat tcgcgtgcag ttccagtagt gtcgcctcca 540
gctaacgtta ccgtcgctgc tgctgtatca tcctctgtgt ctccatcagc ttctccagct 600
gtcaaagtac agagccctgc tgctaatggt aaggaacaat ctccagctaa gcctgaagaa 660
ccaaagaagg acactttaat tgtgaacgat ttcttggaac aagttaaaag acgcaaggct 720
gctttagctg ctaagaaggc tgtcgaagag aagggtcctg aggaaccgaa ggaatctgtc 780
gttggaactg acactgatgc aagcgttgat actaagacag ggcctacagc cactgaatct 840
gccaagtctg aagaagctca atcagaatca caagaaaaga ctaaggaaga ggctccagct 900
gagccaaaac cattgacttt ggccgaaaaa ttgagactta agaggatgga agctgcaaag 960
caagcttctg ctaagaccga ggaactaaag actgaagaat ctaagcctga agaaacaaag 1020
accgaggagc taaagactga agaatctaag cctgaagaaa caaagaccga ggagctaaag 1080
actgaagaaa caaagtccga ggaactaaag actgaagaac ctaaggcgga agaatcaaag 1140
gcggaagaac caaagcctga agaaccaaag accgaggaac cgacgactga acaaccaaag 1200
tcagatgaac caaagtcgga agaatcaaaa actgaagagc caaaaaccga ggtattaaag 1260
actgaagaac caaaatcgga agaatcaaag cctgcagaac caaagactga agaaacagca 1320
actgaagaaa cagcaactga agcaaacgcc gaagaaggtg aaccggctcc tgctggtccc 1380
gttgaaactc ctgctgatgt tgaaacaaaa cctcgagaag aggctgaagt tgaagacgat 1440
ggaaagatta ccatgaccga tttcctacag aagttgaaag aggtttctcc agttgatgat 1500
atttattcct tccaataccc aagtgacatt acgcctccaa atgatagata taaaaagaca 1560

agcattaaat atgcatacgg acctgatttc ttgtatcagt tcaaagaaaa ggtcgatgtt 1620
aaatacgatc cagcgtggat ggctgaaatg acgagtaaaa ttgtcatccc tcctaagaag 1680
cctggttcaa gcggaagagg cgaagataga tttagtaagg gtaaggttgg atctctaaga 1740
agtgaaggca gatcgggttc caggtccaac tcgaagaaga agtcaaagag ggatgataga 1800
aaatctaata gatcatacac ttccagaaag gaccgtgaaa gattcagaga ggaagaagtc 1860
gaagagccaa aggttgaggt tgccccattg gtcccaagtg ctaatagatg ggttcctaaa 1920
tctaagatga agaaaacaga agtcaagtta gctccagacg gaacagaact ttacgacgcg 1980
gaagaagcat caagaaagat gaagtcattg ctgaataaat tgacattaga aatgttcgaa 2040
cctatttctg atgatatcat gaagatcgct aaccaatcta gatgggaaga aaagggtgag 2100
actttgaaga ttgtcatcca acaaattttc aataaggcct gcgatgaacc tcattggtca 2160
tcaatgtacg cgcaattatg tggtaaggtc gttaaagact tagatgatag cattaaagac 2220
tcagaaaccc cagataagac tggttctcac ttggttttgc attacttagt ccaaagatgt 2280
caaactgaat tccaaacagg atggactgat caactaccta caaacgaaga cggtactcct 2340
ctacaacctg aaatgatgtc cgatgaatac tataagatgg ctgccgctaa gagaagaggt 2400
ttgggtttgg ttcgtttcat tggtttcttg taccgttcga acttattgac ttccagaatg 2460
gtcttcttct gtttcaagag actaatgaag gatattcaaa actctcctac tgaagatact 2520
ctagagtctg tatgtgaact tttggaaaca attggtgaac agttcgaagg tgctcgtatt 2580
caagttactg cagaagctgt cattgagggt tcaagcttgc tagacacact attcgaccaa 2640
ataaagaacg tgatcgaaaa tggtgacatc tccagcagaa tcaagtttaa gttgatcgac 2700
attgtcgaac taagagaaaa gaggaactgg aatagtaaaa ataagaacga tggtccaaag 2760
accattgctc aaattcacga agaagaagcc ttgaagaggg ctttggagga aagagaaaga 2820
gaaagagatc gccatgggtc cagaggtggt tccagacgta tgaatagcga gagaaactct 2880
tctagaagag atttctcctc tcattctcac agtcacaatc aaaatagaga cggtttcact 2940
actaccagat cgtcatcagt gagatattct gagccaaaga aggaagaaca agctccaact 3000
ccaactaaat cttctggtgg cgctgccaac atgtttgatg cattgatgga tgccgaagat 3060
gattaa 3066

<210> 2
<211> 1021
<212> PRT
<213> 乳酸克鲁维酵母(Kluyveromyces lactis)
<400> 2

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

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

370 375 380

Lys Pro Glu Glu Pro Lys Thr Glu Glu Pro Thr Thr Glu Gln Pro Lys
385 390 395 400

Ser Asp Glu Pro Lys Ser Glu Glu Ser Lys Thr Glu Glu Pro Lys Thr
405 410 415

Glu Val Leu Lys Thr Glu Glu Pro Lys Ser Glu Glu Ser Lys Pro Ala
420 425 430

Glu Pro Lys Thr Glu Glu Thr Ala Thr Glu Glu Thr Ala Thr Glu Ala
435 440 445

Asn Ala Glu Glu Gly Glu Pro Ala Pro Ala Gly Pro Val Glu Thr Pro
450 455 460

Ala Asp Val Glu Thr Lys Pro Arg Glu Glu Ala Glu Val Glu Asp Asp
465 470 475 480

Gly Lys Ile Thr Met Thr Asp Phe Leu Gln Lys Leu Lys Glu Val Ser
485 490 495

Pro Val Asp Asp Ile Tyr Ser Phe Gln Tyr Pro Ser Asp Ile Thr Pro
500 505 510

Pro Asn Asp Arg Tyr Lys Lys Thr Ser Ile Lys Tyr Ala Tyr Gly Pro
515 520 525

Asp Phe Leu Tyr Gln Phe Lys Glu Lys Val Asp Val Lys Tyr Asp Pro
530 535 540

Ala Trp Met Ala Glu Met Thr Ser Lys Ile Val Ile Pro Pro Lys Lys
545 550 555 560

Pro Gly Ser Ser Gly Arg Gly Glu Asp Arg Phe Ser Lys Gly Lys Val
565 570 575

Gly Ser Leu Arg Ser Glu Gly Arg Ser Gly Ser Arg Ser Asn Ser Lys
580 585 590

Lys Lys Ser Lys Arg Asp Asp Arg Lys Ser Asn Arg Ser Tyr Thr Ser
595 600 605

Arg Lys Asp Arg Glu Arg Phe Arg Glu Glu Glu Val Glu Glu Pro Lys
610 615 620

Val Glu Val Ala Pro Leu Val Pro Ser Ala Asn Arg Trp Val Pro Lys
625 630 635 640

Ser Lys Met Lys Lys Thr Glu Val Lys Leu Ala Pro Asp Gly Thr Glu
645 650 655

Leu Tyr Asp Ala Glu Glu Ala Ser Arg Lys Met Lys Ser Leu Leu Asn
660 665 670

Lys Leu Thr Leu Glu Met Phe Glu Pro Ile Ser Asp Asp Ile Met Lys
675 680 685

Ile Ala Asn Gln Ser Arg Trp Glu Glu Lys Gly Glu Thr Leu Lys Ile
690 695 700
Val Ile Gln Gln Ile Phe Asn Lys Ala Cys Asp Glu Pro His Trp Ser
705 710 715 720
Ser Met Tyr Ala Gln Leu Cys Gly Lys Val Val Lys Asp Leu Asp Asp
725 730 735
Ser Ile Lys Asp Ser Glu Thr Pro Asp Lys Thr Gly Ser His Leu Val
740 745 750
Leu His Tyr Leu Val Gln Arg Cys Gln Thr Glu Phe Gln Thr Gly Trp
755 760 765
Thr Asp Gln Leu Pro Thr Asn Glu Asp Gly Thr Pro Leu Gln Pro Glu
770 775 780
Met Met Ser Asp Glu Tyr Tyr Lys Met Ala Ala Ala Lys Arg Arg Gly
785 790 795 800
Leu Gly Leu Val Arg Phe Ile Gly Phe Leu Tyr Arg Ser Asn Leu Leu
805 810 815
Thr Ser Arg Met Val Phe Phe Cys Phe Lys Arg Leu Met Lys Asp Ile
820 825 830
Gln Asn Ser Pro Thr Glu Asp Thr Leu Glu Ser Val Cys Glu Leu Leu
835 840 845
Glu Thr Ile Gly Glu Gln Phe Glu Gly Ala Arg Ile Gln Val Thr Ala
850 855 860
Glu Ala Val Ile Glu Gly Ser Ser Leu Leu Asp Thr Leu Phe Asp Gln
865 870 875 880
Ile Lys Asn Val Ile Glu Asn Gly Asp Ile Ser Ser Arg Ile Lys Phe
885 890 895
Lys Leu Ile Asp Ile Val Glu Leu Arg Glu Lys Arg Asn Trp Asn Ser
900 905 910
Lys Asn Lys Asn Asp Gly Pro Lys Thr Ile Ala Gln Ile His Glu Glu
915 920 925
Glu Ala Leu Lys Arg Ala Leu Glu Glu Arg Glu Arg Glu Arg Asp Arg
930 935 940
His Gly Ser Arg Gly Gly Ser Arg Arg Met Asn Ser Glu Arg Asn Ser
945 950 955 960
Ser Arg Arg Asp Phe Ser Ser His Ser His Ser His Asn Gln Asn Arg
965 970 975
Asp Gly Phe Thr Thr Thr Arg Ser Ser Ser Val Arg Tyr Ser Glu Pro
980 985 990
Lys Lys Glu Glu Gln Ala Pro Thr Pro Thr Lys Ser Ser Gly Gly Ala

995 1000 1005

Ala Asn Met Phe Asp Ala Leu Met Asp Ala Glu Asp Asp

1010 1015 1020

<210> 3

<211> 1779

<212> DNA

<213> 乳酸克鲁维酵母(Kluyveromyces lactis)

<400> 3

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60
cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120
tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180
attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240
gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300
tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360
tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420
ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480
ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540
ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600
atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660
tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720
gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780
gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840
cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900
ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960
gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020
ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagcttaa 1779

<210> 4

<211> 592
<212> PRT
<213> 乳酸克鲁维酵母(Kluyveromyces lactis)
<400> 4

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

275 280 285

Glu Glu Leu Asn Gly Lys Glu Phe Lys Ser Gln Ala Leu Tyr Val Gly

290 295 300

Arg Ala Gln Lys Lys Tyr Glu Arg Ala Glu Glu Leu Lys Lys Gln Tyr

305 310 315 320

Glu Gln Tyr Arg Leu Glu Lys Leu Ala Lys Phe Gln Gly Val Asn Leu

325 330 335

Phe Ile Lys Asn Leu Asp Asp Ser Ile Asp Asp Glu Lys Leu Lys Glu

340 345 350

Glu Phe Ala Pro Tyr Gly Thr Ile Thr Ser Ala Arg Val Met Arg Asp

355 360 365

Gln Glu Gly Asn Ser Lys Gly Phe Gly Phe Val Cys Phe Ser Ser Pro

370 375 380

Glu Glu Ala Thr Lys Ala Met Thr Glu Lys Asn Gln Gln Ile Val Ala

385 390 395 400

Gly Lys Pro Leu Tyr Val Ala Ile Ala Gln Arg Lys Asp Val Arg Arg

405 410 415

Ser Gln Leu Ala Gln Gln Ile Gln Ala Arg Asn Gln Ile Arg Phe Gln

420 425 430

Gln Gln Gln Gln Gln Gln Ala Ala Ala Ala Ala Ala Gly Met Pro Gly

435 440 445

Gln Tyr Met Pro Gln Met Phe Tyr Gly Val Met Ala Pro Arg Gly Phe

450 455 460

Pro Gly Pro Asn Pro Gly Met Asn Gly Pro Met Gly Ala Gly Ile Pro

465 470 475 480

Lys Asn Gly Met Val Pro Pro Pro Gln Gln Phe Ala Gly Arg Pro Asn

485 490 495

Gly Pro Met Tyr Gln Gly Met Pro Pro Gln Asn Gln Phe Pro Arg His

500 505 510

Gln Gln Gln His Tyr Ile Gln Gln Gln Lys Gln Arg Gln Ala Leu Gly

515 520 525

Glu Gln Leu Tyr Lys Lys Val Ser Ala Lys Ile Asp Asp Glu Asn Ala

530 535 540

Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val

545 550 555 560

Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu

565 570 575

Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala

580 585 590

<210> 5
<211> 4701
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 5

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60
cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120
tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180
attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240
gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300
tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360
tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420
ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480
ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540
ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600
atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660
tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720
gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780
gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840
cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900
ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960
gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020
ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctaacc aaccagcgta cggtgtttct gcgggataca ttccgaacta cggcgtatcg 1920
gcagagtaca accctctgta ctataaccag taccaacagc agcaacagct gtacgctgct 1980
gcttaccaga ctccaatgag cggacaaggt tatgtccccc cagtagtgtc tccagctgct 2040

gtttcagcta aaccagcgaa ggttgagatt actaacaagt ctggtgaaca catagatatt 2100
gcttccattg ctcatccaca tactcattct cattctcaat ctcattcgcg tgcagttcca 2160
gtagtgtcgc ctccagctaa cgttaccgtc gctgctgctg tatcatcctc tgtgtctcca 2220
tcagcttctc cagctgtcaa agtacagagc cctgctgcta atggtaagga acaatctcca 2280
gctaagcctg aagaaccaaa gaaggacact ttaattgtga acgatttctt ggaacaagtt 2340
aaaagacgca aggctgcttt agctgctaag aaggctgtcg aagagaaggg tcctgaggaa 2400
ccgaaggaat ctgtcgttgg aactgacact gatgcaagcg ttgatactaa gacagggcct 2460
acagccactg aatctgccaa gtctgaagaa gctcaatcag aatcacaaga aaagactaag 2520
gaagaggctc cagctgagcc aaaaccattg actttggccg aaaaattgag acttaagagg 2580
atggaagctg caaagcaagc ttctgctaag accgaggaac taaagactga agaatctaag 2640
cctgaagaaa caaagaccga ggagctaaag actgaagaat ctaagcctga agaaacaaag 2700
accgaggagc taaagactga agaaacaaag tccgaggaac taaagactga agaacctaag 2760
gcggaagaat caaaggcgga agaaccaaag cctgaagaac caaagaccga ggaaccgacg 2820
actgaacaac caaagtcaga tgaaccaaag tcggaagaat caaaaactga agagccaaaa 2880
accgaggtat taaagactga agaaccaaaa tcggaagaat caaagcctgc agaaccaaag 2940
actgaagaaa cagcaactga agaaacagca actgaagcaa acgccgaaga aggtgaaccg 3000
gctcctgctg gtcccgttga aactcctgct gatgttgaaa caaaacctcg agaagaggct 3060
gaagttgaag acgatggaaa gattaccatg accgatttcc tacagaagtt gaaagaggtt 3120
tctccagttg atgatattta ttccttccaa tacccaagtg acattacgcc tccaaatgat 3180
agatataaaa agacaagcat taaatatgca tacggacctg atttcttgta tcagttcaaa 3240
gaaaaggtcg atgttaaata cgatccagcg tggatggctg aaatgacgag taaaattgtc 3300
atccctccta agaagcctgg ttcaagcgga agaggcgaag atagatttag taagggtaag 3360
gttggatctc taagaagtga aggcagatcg ggttccaggt ccaactcgaa gaagaagtca 3420
aagagggatg atagaaaatc taatagatca tacacttcca gaaaggaccg tgaaagattc 3480
agagaggaag aagtcgaaga gccaaaggtt gaggttgccc cattggtccc aagtgctaat 3540
agatgggttc ctaaatctaa gatgaagaaa acagaagtca agttagctcc agacggaaca 3600
gaactttacg acgcggaaga agcatcaaga aagatgaagt cattgctgaa taaattgaca 3660
ttagaaatgt tcgaacctat ttctgatgat atcatgaaga tcgctaacca atctagatgg 3720
gaagaaaagg gtgagacttt gaagattgtc atccaacaaa ttttcaataa ggcctgcgat 3780
gaacctcatt ggtcatcaat gtacgcgcaa ttatgtggta aggtcgttaa agacttagat 3840
gatagcatta aagactcaga aaccccagat aagactggtt ctcacttggt tttgcattac 3900
ttagtccaaa gatgtcaaac tgaattccaa acaggatgga ctgatcaact acctacaaac 3960
gaagacggta ctcctctaca acctgaaatg atgtccgatg aatactataa gatggctgcc 4020
gctaagagaa gaggtttggg tttggttcgt ttcattggtt tcttgtaccg ttcgaactta 4080
ttgacttcca gaatggtctt cttctgtttc aagagactaa tgaaggatat tcaaaactct 4140
cctactgaag atactctaga gtctgtatgt gaacttttgg aaacaattgg tgaacagttc 4200
gaaggtgctc gtattcaagt tactgcagaa gctgtcattg agggttcaag cttgctagac 4260
acactattcg accaaataaa gaacgtgatc gaaaatggtg acatctccag cagaatcaag 4320
tttaagttga tcgacattgt cgaactaaga gaaaagagga actggaatag taaaaataag 4380

aacgatggtc caaagaccat tgctcaaatt cacgaagaag aagccttgaa gagggctttg 4440
gaggaaagag aaagagaaag agatcgccat gggtccagag gtggttccag acgtatgaat 4500
agcgagagaa actcttctag aagagatttc tcctctcatt ctcacagtca caatcaaaat 4560
agagacggtt tcactactac cagatcgtca tcagtgagat attctgagcc aaagaaggaa 4620
gaacaagctc caactccaac taaatcttct ggtggcgctg ccaacatgtt tgatgcattg 4680
atggatgccg aagatgatta a 4701

<210> 6
<211> 1566
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 6

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

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

530 535 540

Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val
545 550 555 560

Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu
565 570 575

Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Thr Gln Asp Glu Val Gln
595 600 605

Gly Pro His Ala Gly Lys Ser Thr Val Gly Gly Gly Gly Ser Asn Gln
610 615 620

Pro Ala Tyr Gly Val Ser Ala Gly Tyr Ile Pro Asn Tyr Gly Val Ser
625 630 635 640

Ala Glu Tyr Asn Pro Leu Tyr Tyr Asn Gln Tyr Gln Gln Gln Gln Gln
645 650 655

Leu Tyr Ala Ala Ala Tyr Gln Thr Pro Met Ser Gly Gln Gly Tyr Val
660 665 670

Pro Pro Val Val Ser Pro Ala Ala Val Ser Ala Lys Pro Ala Lys Val
675 680 685

Glu Ile Thr Asn Lys Ser Gly Glu His Ile Asp Ile Ala Ser Ile Ala
690 695 700

His Pro His Thr His Ser His Ser Gln Ser His Ser Arg Ala Val Pro
705 710 715 720

Val Val Ser Pro Pro Ala Asn Val Thr Val Ala Ala Ala Val Ser Ser
725 730 735

Ser Val Ser Pro Ser Ala Ser Pro Ala Val Lys Val Gln Ser Pro Ala
740 745 750

Ala Asn Gly Lys Glu Gln Ser Pro Ala Lys Pro Glu Glu Pro Lys Lys
755 760 765

Asp Thr Leu Ile Val Asn Asp Phe Leu Glu Gln Val Lys Arg Arg Lys
770 775 780

Ala Ala Leu Ala Ala Lys Lys Ala Val Glu Glu Lys Gly Pro Glu Glu
785 790 795 800

Pro Lys Glu Ser Val Val Gly Thr Asp Thr Asp Ala Ser Val Asp Thr
805 810 815

Lys Thr Gly Pro Thr Ala Thr Glu Ser Ala Lys Ser Glu Glu Ala Gln
820 825 830

Ser Glu Ser Gln Glu Lys Thr Lys Glu Glu Ala Pro Ala Glu Pro Lys
835 840 845

Pro Leu Thr Leu Ala Glu Lys Leu Arg Leu Lys Arg Met Glu Ala Ala
850 855 860
Lys Gln Ala Ser Ala Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys
865 870 875 880
Pro Glu Glu Thr Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro
885 890 895
Glu Glu Thr Lys Thr Glu Glu Leu Lys Thr Glu Glu Thr Lys Ser Glu
900 905 910
Glu Leu Lys Thr Glu Glu Pro Lys Ala Glu Glu Ser Lys Ala Glu Glu
915 920 925
Pro Lys Pro Glu Glu Pro Lys Thr Glu Glu Pro Thr Thr Glu Gln Pro
930 935 940
Lys Ser Asp Glu Pro Lys Ser Glu Glu Ser Lys Thr Glu Glu Pro Lys
945 950 955 960
Thr Glu Val Leu Lys Thr Glu Glu Pro Lys Ser Glu Glu Ser Lys Pro
965 970 975
Ala Glu Pro Lys Thr Glu Glu Thr Ala Thr Glu Glu Thr Ala Thr Glu
980 985 990
Ala Asn Ala Glu Glu Gly Glu Pro Ala Pro Ala Gly Pro Val Glu Thr
995 1000 1005
Pro Ala Asp Val Glu Thr Lys Pro Arg Glu Glu Ala Glu Val Glu Asp
1010 1015 1020
Asp Gly Lys Ile Thr Met Thr Asp Phe Leu Gln Lys Leu Lys Glu Val
1025 1030 1035 1040
Ser Pro Val Asp Asp Ile Tyr Ser Phe Gln Tyr Pro Ser Asp Ile Thr
1045 1050 1055
Pro Pro Asn Asp Arg Tyr Lys Lys Thr Ser Ile Lys Tyr Ala Tyr Gly
1060 1065 1070
Pro Asp Phe Leu Tyr Gln Phe Lys Glu Lys Val Asp Val Lys Tyr Asp
1075 1080 1085
Pro Ala Trp Met Ala Glu Met Thr Ser Lys Ile Val Ile Pro Pro Lys
1090 1095 1100
Lys Pro Gly Ser Ser Gly Arg Gly Glu Asp Arg Phe Ser Lys Gly Lys
1105 1110 1115 1120
Val Gly Ser Leu Arg Ser Glu Gly Arg Ser Gly Ser Arg Ser Asn Ser
1125 1130 1135
Lys Lys Lys Ser Lys Arg Asp Asp Arg Lys Ser Asn Arg Ser Tyr Thr
1140 1145 1150
Ser Arg Lys Asp Arg Glu Arg Phe Arg Glu Glu Glu Val Glu Glu Pro

1155 1160 1165

Lys Val Glu Val Ala Pro Leu Val Pro Ser Ala Asn Arg Trp Val Pro
1170 1175 1180

Lys Ser Lys Met Lys Lys Thr Glu Val Lys Leu Ala Pro Asp Gly Thr
1185 1190 1195 1200

Glu Leu Tyr Asp Ala Glu Glu Ala Ser Arg Lys Met Lys Ser Leu Leu
1205 1210 1215

Asn Lys Leu Thr Leu Glu Met Phe Glu Pro Ile Ser Asp Asp Ile Met
1220 1225 1230

Lys Ile Ala Asn Gln Ser Arg Trp Glu Glu Lys Gly Glu Thr Leu Lys
1235 1240 1245

Ile Val Ile Gln Gln Ile Phe Asn Lys Ala Cys Asp Glu Pro His Trp
1250 1255 1260

Ser Ser Met Tyr Ala Gln Leu Cys Gly Lys Val Val Lys Asp Leu Asp
1265 1270 1275 1280

Asp Ser Ile Lys Asp Ser Glu Thr Pro Asp Lys Thr Gly Ser His Leu
1285 1290 1295

Val Leu His Tyr Leu Val Gln Arg Cys Gln Thr Glu Phe Gln Thr Gly
1300 1305 1310

Trp Thr Asp Gln Leu Pro Thr Asn Glu Asp Gly Thr Pro Leu Gln Pro
1315 1320 1325

Glu Met Met Ser Asp Glu Tyr Tyr Lys Met Ala Ala Ala Lys Arg Arg
1330 1335 1340

Gly Leu Gly Leu Val Arg Phe Ile Gly Phe Leu Tyr Arg Ser Asn Leu
1345 1350 1355 1360

Leu Thr Ser Arg Met Val Phe Phe Cys Phe Lys Arg Leu Met Lys Asp
1365 1370 1375

Ile Gln Asn Ser Pro Thr Glu Asp Thr Leu Glu Ser Val Cys Glu Leu
1380 1385 1390

Leu Glu Thr Ile Gly Glu Gln Phe Glu Gly Ala Arg Ile Gln Val Thr
1395 1400 1405

Ala Glu Ala Val Ile Glu Gly Ser Ser Leu Leu Asp Thr Leu Phe Asp
1410 1415 1420

Gln Ile Lys Asn Val Ile Glu Asn Gly Asp Ile Ser Ser Arg Ile Lys
1425 1430 1435 1440

Phe Lys Leu Ile Asp Ile Val Glu Leu Arg Glu Lys Arg Asn Trp Asn
1445 1450 1455

Ser Lys Asn Lys Asn Asp Gly Pro Lys Thr Ile Ala Gln Ile His Glu
1460 1465 1470

Glu Glu Ala Leu Lys Arg Ala Leu Glu Glu Arg Glu Arg Glu Arg Asp
1475 1480 1485
Arg His Gly Ser Arg Gly Gly Ser Arg Arg Met Asn Ser Glu Arg Asn
1490 1495 1500
Ser Ser Arg Arg Asp Phe Ser Ser His Ser His Ser His Asn Gln Asn
1505 1510 1515 1520
Arg Asp Gly Phe Thr Thr Thr Arg Ser Ser Ser Val Arg Tyr Ser Glu
1525 1530 1535
Pro Lys Lys Glu Glu Gln Ala Pro Thr Pro Thr Lys Ser Ser Gly Gly
1540 1545 1550
Ala Ala Asn Met Phe Asp Ala Leu Met Asp Ala Glu Asp Asp
1555 1560 1565

<210> 7

<211> 4530

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 7

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60
cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120
tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180
attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240
gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300
tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360
tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420
ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480
ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540
ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600
atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660
tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720
gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780
gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840
cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900
ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960
gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020
ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320

gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctgctg tttcagctaa accagcgaag gttgagatta ctaacaagtc tggtgaacac 1920
atagatattg cttccattgc tcatccacat actcattctc attctcaatc tcattcgcgt 1980
gcagttccag tagtgtcgcc tccagctaac gttaccgtcg ctgctgctgt atcatcctct 2040
gtgtctccat cagcttctcc agctgtcaaa gtacagagcc ctgctgctaa tggtaaggaa 2100
caatctccag ctaagcctga agaaccaaag aaggacactt taattgtgaa cgatttcttg 2160
gaacaagtta aaagacgcaa ggctgcttta gctgctaaga aggctgtcga agagaagggt 2220
cctgaggaac cgaaggaatc tgtcgttgga actgacactg atgcaagcgt tgatactaag 2280
acagggccta cagccactga atctgccaag tctgaagaag ctcaatcaga atcacaagaa 2340
aagactaagg aagaggctcc agctgagcca aaaccattga ctttggccga aaaattgaga 2400
cttaagagga tggaagctgc aaagcaagct tctgctaaga ccgaggaact aaagactgaa 2460
gaatctaagc ctgaagaaac aaagaccgag gagctaaaga ctgaagaatc taagcctgaa 2520
gaaacaaaga ccgaggagct aaagactgaa gaaacaaagt ccgaggaact aaagactgaa 2580
gaacctaagg cggaagaatc aaaggcggaa gaaccaaagc ctgaagaacc aaagaccgag 2640
gaaccgacga ctgaacaacc aaagtcagat gaaccaaagt cggaagaatc aaaaactgaa 2700
gagccaaaaa ccgaggtatt aaagactgaa gaaccaaaat cggaagaatc aaagcctgca 2760
gaaccaaaga ctgaagaaac agcaactgaa gaaacagcaa ctgaagcaaa cgccgaagaa 2820
ggtgaaccgg ctcctgctgg tcccgttgaa actcctgctg atgttgaaac aaaacctcga 2880
gaagaggctg aagttgaaga cgatggaaag attaccatga ccgatttcct acagaagttg 2940
aaagaggttt ctccagttga tgatatttat tccttccaat acccaagtga cattacgcct 3000
ccaaatgata gatataaaaa gacaagcatt aaatatgcat acggacctga tttcttgtat 3060
cagttcaaag aaaaggtcga tgttaaatac gatccagcgt ggatggctga aatgacgagt 3120
aaaattgtca tccctcctaa gaagcctggt tcaagcggaa gaggcgaaga tagatttagt 3180
aagggtaagg ttggatctct aagaagtgaa ggcagatcgg gttccaggtc caactcgaag 3240
aagaagtcaa agagggatga tagaaaatct aatagatcat acacttccag aaaggaccgt 3300
gaaagattca gagaggaaga agtcgaagag ccaaaggttg aggttgcccc attggtccca 3360
agtgctaata gatgggttcc taaatctaag atgaagaaaa cagaagtcaa gttagctcca 3420
gacggaacag aactttacga cgcggaagaa gcatcaagaa agatgaagtc attgctgaat 3480
aaattgacat tagaaatgtt cgaacctatt tctgatgata tcatgaagat cgctaaccaa 3540
tctagatggg aagaaaaggg tgagactttg aagattgtca tccaacaaat tttcaataag 3600
gcctgcgatg aacctcattg gtcatcaatg tacgcgcaat tatgtggtaa ggtcgttaaa 3660

gacttagatg atagcattaa agactcagaa accccagata agactggttc tcacttggtt 3720
ttgcattact tagtccaaag atgtcaaact gaattccaaa caggatggac tgatcaacta 3780
cctacaaacg aagacggtac tcctctacaa cctgaaatga tgtccgatga atactataag 3840
atggctgccg ctaagagaag aggtttgggt ttggttcgtt tcattggttt cttgtaccgt 3900
tcgaacttat tgacttccag aatggtcttc ttctgtttca agagactaat gaaggatatt 3960
caaaactctc ctactgaaga tactctagag tctgtatgtg aacttttgga aacaattggt 4020
gaacagttcg aaggtgctcg tattcaagtt actgcagaag ctgtcattga gggttcaagc 4080
ttgctagaca cactattcga ccaaataaag aacgtgatcg aaaatggtga catctccagc 4140
agaatcaagt ttaagttgat cgacattgtc gaactaagag aaaagaggaa ctggaatagt 4200
aaaaataaga acgatggtcc aaagaccatt gctcaaattc acgaagaaga agccttgaag 4260
agggctttgg aggaaagaga aagagaaaga gatcgccatg ggtccagagg tggttccaga 4320
cgtatgaata gcgagagaaa ctcttctaga agagatttct cctctcattc tcacagtcac 4380
aatcaaaata gagacggttt cactactacc agatcgtcat cagtgagata ttctgagcca 4440
aagaaggaag aacaagctcc aactccaact aaatcttctg gtggcgctgc caacatgttt 4500
gatgcattga tggatgccga agatgattaa 4530

<210> 8
<211> 1509
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 8

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

145 150 155 160
Phe Gly Glu Val Leu Ser Cys Lys Val Ala Leu Asp Glu Asn Gly Asn
165 170 175
Ser Arg Gly Phe Gly Phe Val His Phe Lys Glu Glu Ser Asp Ala Lys
180 185 190
Asp Ala Ile Glu Ala Val Asn Gly Met Leu Met Asn Gly Leu Glu Val
195 200 205
Tyr Val Ala Met His Val Pro Lys Lys Asp Arg Ile Ser Lys Leu Glu
210 215 220
Glu Ala Lys Ala Asn Phe Thr Asn Ile Tyr Val Lys Asn Ile Asp Val
225 230 235 240
Glu Thr Thr Asp Glu Glu Phe Glu Gln Leu Phe Ser Gln Tyr Gly Glu
245 250 255
Ile Val Ser Ala Ala Leu Glu Lys Asp Ala Glu Gly Lys Pro Lys Gly
260 265 270
Phe Gly Phe Val Asn Phe Val Asp His Asn Ala Ala Ala Lys Ala Val
275 280 285
Glu Glu Leu Asn Gly Lys Glu Phe Lys Ser Gln Ala Leu Tyr Val Gly
290 295 300
Arg Ala Gln Lys Lys Tyr Glu Arg Ala Glu Glu Leu Lys Lys Gln Tyr
305 310 315 320
Glu Gln Tyr Arg Leu Glu Lys Leu Ala Lys Phe Gln Gly Val Asn Leu
325 330 335
Phe Ile Lys Asn Leu Asp Asp Ser Ile Asp Asp Glu Lys Leu Lys Glu
340 345 350
Glu Phe Ala Pro Tyr Gly Thr Ile Thr Ser Ala Arg Val Met Arg Asp
355 360 365
Gln Glu Gly Asn Ser Lys Gly Phe Gly Phe Val Cys Phe Ser Ser Pro
370 375 380
Glu Glu Ala Thr Lys Ala Met Thr Glu Lys Asn Gln Gln Ile Val Ala
385 390 395 400
Gly Lys Pro Leu Tyr Val Ala Ile Ala Gln Arg Lys Asp Val Arg Arg
405 410 415
Ser Gln Leu Ala Gln Gln Ile Gln Ala Arg Asn Gln Ile Arg Phe Gln
420 425 430
Gln Gln Gln Gln Gln Gln Ala Ala Ala Ala Ala Ala Gly Met Pro Gly
435 440 445
Gln Tyr Met Pro Gln Met Phe Tyr Gly Val Met Ala Pro Arg Gly Phe
450 455 460

Pro Gly Pro Asn Pro Gly Met Asn Gly Pro Met Gly Ala Gly Ile Pro
465 470 475 480
Lys Asn Gly Met Val Pro Pro Pro Gln Gln Phe Ala Gly Arg Pro Asn
485 490 495
Gly Pro Met Tyr Gln Gly Met Pro Pro Gln Asn Gln Phe Pro Arg His
500 505 510
Gln Gln Gln His Tyr Ile Gln Gln Gln Lys Gln Arg Gln Ala Leu Gly
515 520 525
Glu Gln Leu Tyr Lys Lys Val Ser Ala Lys Ile Asp Asp Glu Asn Ala
530 535 540
Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val
545 550 555 560
Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu
565 570 575
Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala
580 585 590
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Thr Gln Asp Glu Val Gln
595 600 605
Gly Pro His Ala Gly Lys Ser Thr Val Gly Gly Gly Gly Ser Ala Val
610 615 620
Ser Ala Lys Pro Ala Lys Val Glu Ile Thr Asn Lys Ser Gly Glu His
625 630 635 640
Ile Asp Ile Ala Ser Ile Ala His Pro His Thr His Ser His Ser Gln
645 650 655
Ser His Ser Arg Ala Val Pro Val Val Ser Pro Pro Ala Asn Val Thr
660 665 670
Val Ala Ala Ala Val Ser Ser Ser Val Ser Pro Ser Ala Ser Pro Ala
675 680 685
Val Lys Val Gln Ser Pro Ala Ala Asn Gly Lys Glu Gln Ser Pro Ala
690 695 700
Lys Pro Glu Glu Pro Lys Lys Asp Thr Leu Ile Val Asn Asp Phe Leu
705 710 715 720
Glu Gln Val Lys Arg Arg Lys Ala Ala Leu Ala Ala Lys Lys Ala Val
725 730 735
Glu Glu Lys Gly Pro Glu Glu Pro Lys Glu Ser Val Val Gly Thr Asp
740 745 750
Thr Asp Ala Ser Val Asp Thr Lys Thr Gly Pro Thr Ala Thr Glu Ser
755 760 765
Ala Lys Ser Glu Glu Ala Gln Ser Glu Ser Gln Glu Lys Thr Lys Glu

770 775 780

Glu Ala Pro Ala Glu Pro Lys Pro Leu Thr Leu Ala Glu Lys Leu Arg
785 790 795 800

Leu Lys Arg Met Glu Ala Ala Lys Gln Ala Ser Ala Lys Thr Glu Glu
805 810 815

Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr Lys Thr Glu Glu Leu
820 825 830

Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr Lys Thr Glu Glu Leu Lys
835 840 845

Thr Glu Glu Thr Lys Ser Glu Glu Leu Lys Thr Glu Glu Pro Lys Ala
850 855 860

Glu Glu Ser Lys Ala Glu Glu Pro Lys Pro Glu Glu Pro Lys Thr Glu
865 870 875 880

Glu Pro Thr Thr Glu Gln Pro Lys Ser Asp Glu Pro Lys Ser Glu Glu
885 890 895

Ser Lys Thr Glu Glu Pro Lys Thr Glu Val Leu Lys Thr Glu Glu Pro
900 905 910

Lys Ser Glu Glu Ser Lys Pro Ala Glu Pro Lys Thr Glu Glu Thr Ala
915 920 925

Thr Glu Glu Thr Ala Thr Glu Ala Asn Ala Glu Glu Gly Glu Pro Ala
930 935 940

Pro Ala Gly Pro Val Glu Thr Pro Ala Asp Val Glu Thr Lys Pro Arg
945 950 955 960

Glu Glu Ala Glu Val Glu Asp Asp Gly Lys Ile Thr Met Thr Asp Phe
965 970 975

Leu Gln Lys Leu Lys Glu Val Ser Pro Val Asp Asp Ile Tyr Ser Phe
980 985 990

Gln Tyr Pro Ser Asp Ile Thr Pro Pro Asn Asp Arg Tyr Lys Lys Thr
995 1000 1005

Ser Ile Lys Tyr Ala Tyr Gly Pro Asp Phe Leu Tyr Gln Phe Lys Glu
1010 1015 1020

Lys Val Asp Val Lys Tyr Asp Pro Ala Trp Met Ala Glu Met Thr Ser
1025 1030 1035 1040

Lys Ile Val Ile Pro Pro Lys Lys Pro Gly Ser Ser Gly Arg Gly Glu
1045 1050 1055

Asp Arg Phe Ser Lys Gly Lys Val Gly Ser Leu Arg Ser Glu Gly Arg
1060 1065 1070

Ser Gly Ser Arg Ser Asn Ser Lys Lys Lys Ser Lys Arg Asp Asp Arg
1075 1080 1085

Lys Ser Asn Arg Ser Tyr Thr Ser Arg Lys Asp Arg Glu Arg Phe Arg
1090 1095 1100
Glu Glu Glu Val Glu Glu Pro Lys Val Glu Val Ala Pro Leu Val Pro
1105 1110 1115 1120
Ser Ala Asn Arg Trp Val Pro Lys Ser Lys Met Lys Lys Thr Glu Val
1125 1130 1135
Lys Leu Ala Pro Asp Gly Thr Glu Leu Tyr Asp Ala Glu Glu Ala Ser
1140 1145 1150
Arg Lys Met Lys Ser Leu Leu Asn Lys Leu Thr Leu Glu Met Phe Glu
1155 1160 1165
Pro Ile Ser Asp Asp Ile Met Lys Ile Ala Asn Gln Ser Arg Trp Glu
1170 1175 1180
Glu Lys Gly Glu Thr Leu Lys Ile Val Ile Gln Gln Ile Phe Asn Lys
1185 1190 1195 1200
Ala Cys Asp Glu Pro His Trp Ser Ser Met Tyr Ala Gln Leu Cys Gly
1205 1210 1215
Lys Val Val Lys Asp Leu Asp Asp Ser Ile Lys Asp Ser Glu Thr Pro
1220 1225 1230
Asp Lys Thr Gly Ser His Leu Val Leu His Tyr Leu Val Gln Arg Cys
1235 1240 1245
Gln Thr Glu Phe Gln Thr Gly Trp Thr Asp Gln Leu Pro Thr Asn Glu
1250 1255 1260
Asp Gly Thr Pro Leu Gln Pro Glu Met Met Ser Asp Glu Tyr Tyr Lys
1265 1270 1275 1280
Met Ala Ala Ala Lys Arg Arg Gly Leu Gly Leu Val Arg Phe Ile Gly
1285 1290 1295
Phe Leu Tyr Arg Ser Asn Leu Leu Thr Ser Arg Met Val Phe Phe Cys
1300 1305 1310
Phe Lys Arg Leu Met Lys Asp Ile Gln Asn Ser Pro Thr Glu Asp Thr
1315 1320 1325
Leu Glu Ser Val Cys Glu Leu Leu Glu Thr Ile Gly Glu Gln Phe Glu
1330 1335 1340
Gly Ala Arg Ile Gln Val Thr Ala Glu Ala Val Ile Glu Gly Ser Ser
1345 1350 1355 1360
Leu Leu Asp Thr Leu Phe Asp Gln Ile Lys Asn Val Ile Glu Asn Gly
1365 1370 1375
Asp Ile Ser Ser Arg Ile Lys Phe Lys Leu Ile Asp Ile Val Glu Leu
1380 1385 1390
Arg Glu Lys Arg Asn Trp Asn Ser Lys Asn Lys Asn Asp Gly Pro Lys

1395 1400 1405

Thr Ile Ala Gln Ile His Glu Glu Glu Ala Leu Lys Arg Ala Leu Glu

1410 1415 1420

Glu Arg Glu Arg Glu Arg Asp Arg His Gly Ser Arg Gly Gly Ser Arg

1425 1430 1435 1440

Arg Met Asn Ser Glu Arg Asn Ser Ser Arg Arg Asp Phe Ser Ser His

1445 1450 1455

Ser His Ser His Asn Gln Asn Arg Asp Gly Phe Thr Thr Thr Arg Ser

1460 1465 1470

Ser Ser Val Arg Tyr Ser Glu Pro Lys Lys Glu Glu Gln Ala Pro Thr

1475 1480 1485

Pro Thr Lys Ser Ser Gly Gly Ala Ala Asn Met Phe Asp Ala Leu Met

1490 1495 1500

Asp Ala Glu Asp Asp

1505

<210> 9

<211> 4017

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 9

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60

cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120

tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180

attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240

gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300

tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360

tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420

ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480

ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540

ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600

atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660

tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720

gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780

gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840

cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900

ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960

gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020

ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080

acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140

ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctactt tggccgaaaa attgagactt aagaggatgg aagctgcaaa gcaagcttct 1920
gctaagaccg aggaactaaa gactgaagaa tctaagcctg aagaaacaaa gaccgaggag 1980
ctaaagactg aagaatctaa gcctgaagaa acaaagaccg aggagctaaa gactgaagaa 2040
acaaagtccg aggaactaaa gactgaagaa cctaaggcgg aagaatcaaa ggcggaagaa 2100
ccaaagcctg aagaaccaaa gaccgaggaa ccgacgactg aacaaccaaa gtcagatgaa 2160
ccaaagtcgg aagaatcaaa aactgaagag ccaaaaaccg aggtattaaa gactgaagaa 2220
ccaaaatcgg aagaatcaaa gcctgcagaa ccaaagactg aagaaacagc aactgaagaa 2280
acagcaactg aagcaaacgc cgaagaaggt gaaccggctc ctgctggtcc cgttgaaact 2340
cctgctgatg ttgaaacaaa acctcgagaa gaggctgaag ttgaagacga tggaaagatt 2400
accatgaccg atttcctaca gaagttgaaa gaggtttctc cagttgatga tatttattcc 2460
ttccaatacc caagtgacat tacgcctcca aatgatagat ataaaaagac aagcattaaa 2520
tatgcatacg gacctgattt cttgtatcag ttcaaagaaa aggtcgatgt taaatacgat 2580
ccagcgtgga tggctgaaat gacgagtaaa attgtcatcc ctcctaagaa gcctggttca 2640
agcggaagag gcgaagatag atttagtaag ggtaaggttg gatctctaag aagtgaaggc 2700
agatcgggtt ccaggtccaa ctcgaagaag aagtcaaaga gggatgatag aaaatctaat 2760
agatcataca cttccagaaa ggaccgtgaa agattcagag aggaagaagt cgaagagcca 2820
aaggttgagg ttgccccatt ggtcccaagt gctaatagat gggttcctaa atctaagatg 2880
aagaaaacag aagtcaagtt agctccagac ggaacagaac tttacgacgc ggaagaagca 2940
tcaagaaaga tgaagtcatt gctgaataaa ttgacattag aaatgttcga acctatttct 3000
gatgatatca tgaagatcgc taaccaatct agatgggaag aaaagggtga gactttgaag 3060
attgtcatcc aacaaatttt caataaggcc tgcgatgaac ctcattggtc atcaatgtac 3120
gcgcaattat gtggtaaggt cgttaaagac ttagatgata gcattaaaga ctcagaaacc 3180
ccagataaga ctggttctca cttggttttg cattacttag tccaaagatg tcaaactgaa 3240
ttccaaacag gatggactga tcaactacct acaaacgaag acggtactcc tctacaacct 3300
gaaatgatgt ccgatgaata ctataagatg gctgccgcta agagaagagg tttgggtttg 3360
gttcgtttca ttggtttctt gtaccgttcg aacttattga cttccagaat ggtcttcttc 3420
tgtttcaaga gactaatgaa ggatattcaa aactctccta ctgaagatac tctagagtct 3480

gtatgtgaac ttttggaaac aattggtgaa cagttcgaag gtgctcgtat tcaagttact 3540
gcagaagctg tcattgaggg ttcaagcttg ctagacacac tattcgacca aataaagaac 3600
gtgatcgaaa atggtgacat ctccagcaga atcaagttta agttgatcga cattgtcgaa 3660
ctaagagaaa agaggaactg gaatagtaaa aataagaacg atggtccaaa gaccattgct 3720
caaattcacg aagaagaagc cttgaagagg gctttggagg aaagagaaag agaaagagat 3780
cgccatgggt ccagaggtgg ttccagacgt atgaatagcg agagaaactc ttctagaaga 3840
gatttctcct ctcattctca cagtcacaat caaaatagag acggtttcac tactaccaga 3900
tcgtcatcag tgagatattc tgagccaaag aaggaagaac aagctccaac tccaactaaa 3960
tcttctggtg gcgctgccaa catgtttgat gcattgatgg atgccgaaga tgattaa 4017

<210> 10
<211> 1338
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 10

Met Ser Asp Ile Thr Glu Lys Thr Ala Glu Gln Leu Glu Asn Leu Gln
1 5 10 15
Ile Asn Asp Asp Gln Gln Pro Ala Gln Ser Ala Ser Ala Pro Ser Thr
20 25 30
Ser Ala Ser Glu Ser Glu Ala Ser Ser Val Ser Lys Val Glu Asn Asn
35 40 45
Asn Ala Ser Leu Tyr Val Gly Glu Leu Asp Pro Asn Ile Thr Glu Ala
50 55 60
Leu Leu Tyr Asp Val Phe Ser Pro Leu Gly Pro Ile Ser Ser Ile Arg
65 70 75 80
Val Cys Arg Asp Ala Val Thr Lys Ala Ser Leu Gly Tyr Ala Tyr Val
85 90 95
Asn Tyr Thr Asp Tyr Glu Ala Gly Lys Lys Ala Ile Gln Glu Leu Asn
100 105 110
Tyr Ala Glu Ile Asn Gly Arg Pro Cys Arg Ile Met Trp Ser Glu Arg
115 120 125
Asp Pro Ala Ile Arg Lys Lys Gly Ser Gly Asn Ile Phe Ile Lys Asn
130 135 140
Leu His Pro Ala Ile Asp Asn Lys Ala Leu His Glu Thr Phe Ser Thr
145 150 155 160
Phe Gly Glu Val Leu Ser Cys Lys Val Ala Leu Asp Glu Asn Gly Asn
165 170 175
Ser Arg Gly Phe Gly Phe Val His Phe Lys Glu Glu Ser Asp Ala Lys
180 185 190
Asp Ala Ile Glu Ala Val Asn Gly Met Leu Met Asn Gly Leu Glu Val

195 200 205

Tyr Val Ala Met His Val Pro Lys Lys Asp Arg Ile Ser Lys Leu Glu
210 215 220

Glu Ala Lys Ala Asn Phe Thr Asn Ile Tyr Val Lys Asn Ile Asp Val
225 230 235 240

Glu Thr Thr Asp Glu Glu Phe Glu Gln Leu Phe Ser Gln Tyr Gly Glu
245 250 255

Ile Val Ser Ala Ala Leu Glu Lys Asp Ala Glu Gly Lys Pro Lys Gly
260 265 270

Phe Gly Phe Val Asn Phe Val Asp His Asn Ala Ala Ala Lys Ala Val
275 280 285

Glu Glu Leu Asn Gly Lys Glu Phe Lys Ser Gln Ala Leu Tyr Val Gly
290 295 300

Arg Ala Gln Lys Lys Tyr Glu Arg Ala Glu Glu Leu Lys Lys Gln Tyr
305 310 315 320

Glu Gln Tyr Arg Leu Glu Lys Leu Ala Lys Phe Gln Gly Val Asn Leu
325 330 335

Phe Ile Lys Asn Leu Asp Asp Ser Ile Asp Asp Glu Lys Leu Lys Glu
340 345 350

Glu Phe Ala Pro Tyr Gly Thr Ile Thr Ser Ala Arg Val Met Arg Asp
355 360 365

Gln Glu Gly Asn Ser Lys Gly Phe Gly Phe Val Cys Phe Ser Ser Pro
370 375 380

Glu Glu Ala Thr Lys Ala Met Thr Glu Lys Asn Gln Gln Ile Val Ala
385 390 395 400

Gly Lys Pro Leu Tyr Val Ala Ile Ala Gln Arg Lys Asp Val Arg Arg
405 410 415

Ser Gln Leu Ala Gln Gln Ile Gln Ala Arg Asn Gln Ile Arg Phe Gln
420 425 430

Gln Gln Gln Gln Gln Gln Ala Ala Ala Ala Ala Ala Gly Met Pro Gly
435 440 445

Gln Tyr Met Pro Gln Met Phe Tyr Gly Val Met Ala Pro Arg Gly Phe
450 455 460

Pro Gly Pro Asn Pro Gly Met Asn Gly Pro Met Gly Ala Gly Ile Pro
465 470 475 480

Lys Asn Gly Met Val Pro Pro Pro Gln Gln Phe Ala Gly Arg Pro Asn
485 490 495

Gly Pro Met Tyr Gln Gly Met Pro Pro Gln Asn Gln Phe Pro Arg His
500 505 510

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

820 825 830

Arg Tyr Lys Lys Thr Ser Ile Lys Tyr Ala Tyr Gly Pro Asp Phe Leu

835 840 845

Tyr Gln Phe Lys Glu Lys Val Asp Val Lys Tyr Asp Pro Ala Trp Met

850 855 860

Ala Glu Met Thr Ser Lys Ile Val Ile Pro Pro Lys Lys Pro Gly Ser

865 870 875 880

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

885 890 895

Arg Ser Glu Gly Arg Ser Gly Ser Arg Ser Asn Ser Lys Lys Lys Ser

900 905 910

Lys Arg Asp Asp Arg Lys Ser Asn Arg Ser Tyr Thr Ser Arg Lys Asp

915 920 925

Arg Glu Arg Phe Arg Glu Glu Glu Val Glu Glu Pro Lys Val Glu Val

930 935 940

Ala Pro Leu Val Pro Ser Ala Asn Arg Trp Val Pro Lys Ser Lys Met

945 950 955 960

Lys Lys Thr Glu Val Lys Leu Ala Pro Asp Gly Thr Glu Leu Tyr Asp

965 970 975

Ala Glu Glu Ala Ser Arg Lys Met Lys Ser Leu Leu Asn Lys Leu Thr

980 985 990

Leu Glu Met Phe Glu Pro Ile Ser Asp Asp Ile Met Lys Ile Ala Asn

995 1000 1005

Gln Ser Arg Trp Glu Glu Lys Gly Glu Thr Leu Lys Ile Val Ile Gln

1010 1015 1020

Gln Ile Phe Asn Lys Ala Cys Asp Glu Pro His Trp Ser Ser Met Tyr

1025 1030 1035 1040

Ala Gln Leu Cys Gly Lys Val Val Lys Asp Leu Asp Asp Ser Ile Lys

1045 1050 1055

Asp Ser Glu Thr Pro Asp Lys Thr Gly Ser His Leu Val Leu His Tyr

1060 1065 1070

Leu Val Gln Arg Cys Gln Thr Glu Phe Gln Thr Gly Trp Thr Asp Gln

1075 1080 1085

Leu Pro Thr Asn Glu Asp Gly Thr Pro Leu Gln Pro Glu Met Met Ser

1090 1095 1100

Asp Glu Tyr Tyr Lys Met Ala Ala Ala Lys Arg Arg Gly Leu Gly Leu

1105 1110 1115 1120

Val Arg Phe Ile Gly Phe Leu Tyr Arg Ser Asn Leu Leu Thr Ser Arg

1125 1130 1135

Met Val Phe Phe Cys Phe Lys Arg Leu Met Lys Asp Ile Gln Asn Ser
1140 1145 1150
Pro Thr Glu Asp Thr Leu Glu Ser Val Cys Glu Leu Leu Glu Thr Ile
1155 1160 1165
Gly Glu Gln Phe Glu Gly Ala Arg Ile Gln Val Thr Ala Glu Ala Val
1170 1175 1180
Ile Glu Gly Ser Ser Leu Leu Asp Thr Leu Phe Asp Gln Ile Lys Asn
1185 1190 1195 1200
Val Ile Glu Asn Gly Asp Ile Ser Ser Arg Ile Lys Phe Lys Leu Ile
1205 1210 1215
Asp Ile Val Glu Leu Arg Glu Lys Arg Asn Trp Asn Ser Lys Asn Lys
1220 1225 1230
Asn Asp Gly Pro Lys Thr Ile Ala Gln Ile His Glu Glu Glu Ala Leu
1235 1240 1245
Lys Arg Ala Leu Glu Glu Arg Glu Arg Glu Arg Asp Arg His Gly Ser
1250 1255 1260
Arg Gly Gly Ser Arg Arg Met Asn Ser Glu Arg Asn Ser Ser Arg Arg
1265 1270 1275 1280
Asp Phe Ser Ser His Ser His Ser His Asn Gln Asn Arg Asp Gly Phe
1285 1290 1295
Thr Thr Thr Arg Ser Ser Ser Val Arg Tyr Ser Glu Pro Lys Lys Glu
1300 1305 1310
Glu Gln Ala Pro Thr Pro Thr Lys Ser Ser Gly Gly Ala Ala Asn Met
1315 1320 1325
Phe Asp Ala Leu Met Asp Ala Glu Asp Asp
1330 1335

<210> 11
<211> 3234
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 11

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60
cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120
tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180
attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240
gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300
tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360
tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420
ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480

ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540
ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600
atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660
tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720
gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780
gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840
cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900
ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960
gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020
ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctggcg aagatagatt tagtaagggt aaggttggat ctctaagaag tgaaggcaga 1920
tcgggttcca ggtccaactc gaagaagaag tcaaagaggg atgatagaaa atctaataga 1980
tcatacactt ccagaaagga ccgtgaaaga ttcagagagg aagaagtcga agagccaaag 2040
gttgaggttg ccccattggt cccaagtgct aatagatggg ttcctaaatc taagatgaag 2100
aaaacagaag tcaagttagc tccagacgga acagaacttt acgacgcgga agaagcatca 2160
agaaagatga agtcattgct gaataaattg acattagaaa tgttcgaacc tatttctgat 2220
gatatcatga agatcgctaa ccaatctaga tgggaagaaa agggtgagac tttgaagatt 2280
gtcatccaac aaattttcaa taaggcctgc gatgaacctc attggtcatc aatgtacgcg 2340
caattatgtg gtaaggtcgt taaagactta gatgatagca ttaaagactc agaaacccca 2400
gataagactg gttctcactt ggttttgcat tacttagtcc aaagatgtca aactgaattc 2460
caaacaggat ggactgatca actacctaca aacgaagacg gtactcctct acaacctgaa 2520
atgatgtccg atgaatacta taagatggct gccgctaaga gaagaggttt gggtttggtt 2580
cgtttcattg gtttcttgta ccgttcgaac ttattgactt ccagaatggt cttcttctgt 2640
ttcaagagac taatgaagga tattcaaaac tctcctactg aagatactct agagtctgta 2700
tgtgaacttt tggaaacaat tggtgaacag ttcgaaggtg ctcgtattca agttactgca 2760
gaagctgtca ttgagggttc aagcttgcta gacacactat tcgaccaaat aaagaacgtg 2820

atcgaaaatg gtgacatctc cagcagaatc aagtttaagt tgatcgacat tgtcgaacta 2880
agagaaaaga ggaactggaa tagtaaaaat aagaacgatg gtccaaagac cattgctcaa 2940
attcacgaag aagaagcctt gaagagggct ttggaggaaa gagaaagaga aagagatcgc 3000
catgggtcca gaggtggttc cagacgtatg aatagcgaga gaaactcttc tagaagagat 3060
ttctcctctc attctcacag tcacaatcaa aatagagacg gtttcactac taccagatcg 3120
tcatcagtga gatattctga gccaaagaag gaagaacaag ctccaactcc aactaaatct 3180
tctggtggcg ctgccaacat gtttgatgca ttgatggatg ccgaagatga ttaa 3234

<210> 12
<211> 1077
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 12

Met Ser Asp Ile Thr Glu Lys Thr Ala Glu Gln Leu Glu Asn Leu Gln
1 5 10 15
Ile Asn Asp Asp Gln Gln Pro Ala Gln Ser Ala Ser Ala Pro Ser Thr
20 25 30
Ser Ala Ser Glu Ser Glu Ala Ser Ser Val Ser Lys Val Glu Asn Asn
35 40 45
Asn Ala Ser Leu Tyr Val Gly Glu Leu Asp Pro Asn Ile Thr Glu Ala
50 55 60
Leu Leu Tyr Asp Val Phe Ser Pro Leu Gly Pro Ile Ser Ser Ile Arg
65 70 75 80
Val Cys Arg Asp Ala Val Thr Lys Ala Ser Leu Gly Tyr Ala Tyr Val
85 90 95
Asn Tyr Thr Asp Tyr Glu Ala Gly Lys Lys Ala Ile Gln Glu Leu Asn
100 105 110
Tyr Ala Glu Ile Asn Gly Arg Pro Cys Arg Ile Met Trp Ser Glu Arg
115 120 125
Asp Pro Ala Ile Arg Lys Lys Gly Ser Gly Asn Ile Phe Ile Lys Asn
130 135 140
Leu His Pro Ala Ile Asp Asn Lys Ala Leu His Glu Thr Phe Ser Thr
145 150 155 160
Phe Gly Glu Val Leu Ser Cys Lys Val Ala Leu Asp Glu Asn Gly Asn
165 170 175
Ser Arg Gly Phe Gly Phe Val His Phe Lys Glu Glu Ser Asp Ala Lys
180 185 190
Asp Ala Ile Glu Ala Val Asn Gly Met Leu Met Asn Gly Leu Glu Val
195 200 205
Tyr Val Ala Met His Val Pro Lys Lys Asp Arg Ile Ser Lys Leu Glu

210 215 220

Glu Ala Lys Ala Asn Phe Thr Asn Ile Tyr Val Lys Asn Ile Asp Val
225 230 235 240

Glu Thr Thr Asp Glu Glu Phe Glu Gln Leu Phe Ser Gln Tyr Gly Glu
245 250 255

Ile Val Ser Ala Ala Leu Glu Lys Asp Ala Glu Gly Lys Pro Lys Gly
260 265 270

Phe Gly Phe Val Asn Phe Val Asp His Asn Ala Ala Ala Lys Ala Val
275 280 285

Glu Glu Leu Asn Gly Lys Glu Phe Lys Ser Gln Ala Leu Tyr Val Gly
290 295 300

Arg Ala Gln Lys Lys Tyr Glu Arg Ala Glu Glu Leu Lys Lys Gln Tyr
305 310 315 320

Glu Gln Tyr Arg Leu Glu Lys Leu Ala Lys Phe Gln Gly Val Asn Leu
325 330 335

Phe Ile Lys Asn Leu Asp Asp Ser Ile Asp Asp Glu Lys Leu Lys Glu
340 345 350

Glu Phe Ala Pro Tyr Gly Thr Ile Thr Ser Ala Arg Val Met Arg Asp
355 360 365

Gln Glu Gly Asn Ser Lys Gly Phe Gly Phe Val Cys Phe Ser Ser Pro
370 375 380

Glu Glu Ala Thr Lys Ala Met Thr Glu Lys Asn Gln Gln Ile Val Ala
385 390 395 400

Gly Lys Pro Leu Tyr Val Ala Ile Ala Gln Arg Lys Asp Val Arg Arg
405 410 415

Ser Gln Leu Ala Gln Gln Ile Gln Ala Arg Asn Gln Ile Arg Phe Gln
420 425 430

Gln Gln Gln Gln Gln Gln Ala Ala Ala Ala Ala Ala Gly Met Pro Gly
435 440 445

Gln Tyr Met Pro Gln Met Phe Tyr Gly Val Met Ala Pro Arg Gly Phe
450 455 460

Pro Gly Pro Asn Pro Gly Met Asn Gly Pro Met Gly Ala Gly Ile Pro
465 470 475 480

Lys Asn Gly Met Val Pro Pro Pro Gln Gln Phe Ala Gly Arg Pro Asn
485 490 495

Gly Pro Met Tyr Gln Gly Met Pro Pro Gln Asn Gln Phe Pro Arg His
500 505 510

Gln Gln Gln His Tyr Ile Gln Gln Gln Lys Gln Arg Gln Ala Leu Gly
515 520 525

Glu Gln Leu Tyr Lys Lys Val Ser Ala Lys Ile Asp Asp Glu Asn Ala
530 535 540
Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val
545 550 555 560
Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu
565 570 575
Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala
580 585 590
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Thr Gln Asp Glu Val Gln
595 600 605
Gly Pro His Ala Gly Lys Ser Thr Val Gly Gly Gly Gly Ser Gly Glu
610 615 620
Asp Arg Phe Ser Lys Gly Lys Val Gly Ser Leu Arg Ser Glu Gly Arg
625 630 635 640
Ser Gly Ser Arg Ser Asn Ser Lys Lys Lys Ser Lys Arg Asp Asp Arg
645 650 655
Lys Ser Asn Arg Ser Tyr Thr Ser Arg Lys Asp Arg Glu Arg Phe Arg
660 665 670
Glu Glu Glu Val Glu Glu Pro Lys Val Glu Val Ala Pro Leu Val Pro
675 680 685
Ser Ala Asn Arg Trp Val Pro Lys Ser Lys Met Lys Lys Thr Glu Val
690 695 700
Lys Leu Ala Pro Asp Gly Thr Glu Leu Tyr Asp Ala Glu Glu Ala Ser
705 710 715 720
Arg Lys Met Lys Ser Leu Leu Asn Lys Leu Thr Leu Glu Met Phe Glu
725 730 735
Pro Ile Ser Asp Asp Ile Met Lys Ile Ala Asn Gln Ser Arg Trp Glu
740 745 750
Glu Lys Gly Glu Thr Leu Lys Ile Val Ile Gln Gln Ile Phe Asn Lys
755 760 765
Ala Cys Asp Glu Pro His Trp Ser Ser Met Tyr Ala Gln Leu Cys Gly
770 775 780
Lys Val Val Lys Asp Leu Asp Asp Ser Ile Lys Asp Ser Glu Thr Pro
785 790 795 800
Asp Lys Thr Gly Ser His Leu Val Leu His Tyr Leu Val Gln Arg Cys
805 810 815
Gln Thr Glu Phe Gln Thr Gly Trp Thr Asp Gln Leu Pro Thr Asn Glu
820 825 830
Asp Gly Thr Pro Leu Gln Pro Glu Met Met Ser Asp Glu Tyr Tyr Lys

835 840 845

Met Ala Ala Ala Lys Arg Arg Gly Leu Gly Leu Val Arg Phe Ile Gly
850 855 860

Phe Leu Tyr Arg Ser Asn Leu Leu Thr Ser Arg Met Val Phe Phe Cys
865 870 875 880

Phe Lys Arg Leu Met Lys Asp Ile Gln Asn Ser Pro Thr Glu Asp Thr
885 890 895

Leu Glu Ser Val Cys Glu Leu Leu Glu Thr Ile Gly Glu Gln Phe Glu
900 905 910

Gly Ala Arg Ile Gln Val Thr Ala Glu Ala Val Ile Glu Gly Ser Ser
915 920 925

Leu Leu Asp Thr Leu Phe Asp Gln Ile Lys Asn Val Ile Glu Asn Gly
930 935 940

Asp Ile Ser Ser Arg Ile Lys Phe Lys Leu Ile Asp Ile Val Glu Leu
945 950 955 960

Arg Glu Lys Arg Asn Trp Asn Ser Lys Asn Lys Asn Asp Gly Pro Lys
965 970 975

Thr Ile Ala Gln Ile His Glu Glu Glu Ala Leu Lys Arg Ala Leu Glu
980 985 990

Glu Arg Glu Arg Glu Arg Asp Arg His Gly Ser Arg Gly Gly Ser Arg
995 1000 1005

Arg Met Asn Ser Glu Arg Asn Ser Ser Arg Arg Asp Phe Ser Ser His
1010 1015 1020

Ser His Ser His Asn Gln Asn Arg Asp Gly Phe Thr Thr Thr Arg Ser
1025 1030 1035 1040

Ser Ser Val Arg Tyr Ser Glu Pro Lys Lys Glu Glu Gln Ala Pro Thr
1045 1050 1055

Pro Thr Lys Ser Ser Gly Gly Ala Ala Asn Met Phe Asp Ala Leu Met
1060 1065 1070

Asp Ala Glu Asp Asp
1075

<210> 13

<211> 3573

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 13

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60

cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120

tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180

attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240
gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300
tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360
tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420
ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480
ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540
ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600
atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660
tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720
gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780
gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840
cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900
ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960
gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020
ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctggcg aacctacatc cgatcagcaa ccagctgttg aagctccagt tgtgcaggag 1920
gagacaacca gttctccgca aaaaaacagt ggatatgtca agaatactgc tggaagcggt 1980
gctcctagaa atgggaaata tgatggtaac aggaagaact ctaggcctta taaccaaaga 2040
ggtaacaaca acaataataa tggttcttcc tcgaataagc actatcaaaa gtataaccaa 2100
ccagcgtacg gtgtttctgc gggatacatt ccgaactacg gcgtatcggc agagtacaac 2160
cctctgtact ataaccagta ccaacagcag caacagctgt acgctgctgc ttaccagact 2220
ccaatgagcg gacaaggtta tgtcccccca gtagtgtctc cagctgctgt ttcagctaaa 2280
ccagcgaagg ttgagattac taacaagtct ggtgaacaca tagatattgc ttccattgct 2340
catccacata ctcattctca ttctcaatct cattcgcgtg cagttccagt agtgtcgcct 2400
ccagctaacg ttaccgtcgc tgctgctgta tcatcctctg tgtctccatc agcttctcca 2460
gctgtcaaag tacagagccc tgctgctaat ggtaaggaac aatctccagc taagcctgaa 2520

gaaccaaaga aggacacttt aattgtgaac gatttcttgg aacaagttaa aagacgcaag 2580
gctgctttag ctgctaagaa ggctgtcgaa gagaagggtc ctgaggaacc gaaggaatct 2640
gtcgttggaa ctgacactga tgcaagcgtt gatactaaga cagggcctac agccactgaa 2700
tctgccaagt ctgaagaagc tcaatcagaa tcacaagaaa agactaagga agaggctcca 2760
gctgagccaa aaccattgac tttggccgaa aaattgagac ttaagaggat ggaagctgca 2820
aagcaagctt ctgctaagac cgaggaacta aagactgaag aatctaagcc tgaagaaaca 2880
aagaccgagg agctaaagac tgaagaatct aagcctgaag aaacaaagac cgaggagcta 2940
aagactgaag aaacaaagtc cgaggaacta aagactgaag aacctaaggc ggaagaatca 3000
aaggcggaag aaccaaagcc tgaagaacca aagaccgagg aaccgacgac tgaacaacca 3060
aagtcagatg aaccaaagtc ggaagaatca aaaactgaag agccaaaaac cgaggtatta 3120
aagactgaag aaccaaaatc ggaagaatca aagcctgcag aaccaaagac tgaagaaaca 3180
gcaactgaag aaacagcaac tgaagcaaac gccgaagaag gtgaaccggc tcctgctggt 3240
cccgttgaaa ctcctgctga tgttgaaaca aaacctcgag aagaggctga agttgaagac 3300
gatggaaaga ttaccatgac cgatttccta cagaagttga aagaggtttc tccagttgat 3360
gatatttatt ccttccaata cccaagtgac attacgcctc caaatgatag atataaaaag 3420
acaagcatta aatatgcata cggacctgat ttcttgtatc agttcaaaga aaaggtcgat 3480
gttaaatacg atccagcgtg gatggctgaa atgacgagta aaattgtcat ccctcctaag 3540
aagcctggtt caagcggaag aggcgaagat taa 3573

<210> 14
<211> 1190
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 14

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

Asp Pro Ala Ile Arg Lys Lys Gly Ser Gly Asn Ile Phe Ile Lys Asn
130 135 140
Leu His Pro Ala Ile Asp Asn Lys Ala Leu His Glu Thr Phe Ser Thr
145 150 155 160
Phe Gly Glu Val Leu Ser Cys Lys Val Ala Leu Asp Glu Asn Gly Asn
165 170 175
Ser Arg Gly Phe Gly Phe Val His Phe Lys Glu Glu Ser Asp Ala Lys
180 185 190
Asp Ala Ile Glu Ala Val Asn Gly Met Leu Met Asn Gly Leu Glu Val
195 200 205
Tyr Val Ala Met His Val Pro Lys Lys Asp Arg Ile Ser Lys Leu Glu
210 215 220
Glu Ala Lys Ala Asn Phe Thr Asn Ile Tyr Val Lys Asn Ile Asp Val
225 230 235 240
Glu Thr Thr Asp Glu Glu Phe Glu Gln Leu Phe Ser Gln Tyr Gly Glu
245 250 255
Ile Val Ser Ala Ala Leu Glu Lys Asp Ala Glu Gly Lys Pro Lys Gly
260 265 270
Phe Gly Phe Val Asn Phe Val Asp His Asn Ala Ala Ala Lys Ala Val
275 280 285
Glu Glu Leu Asn Gly Lys Glu Phe Lys Ser Gln Ala Leu Tyr Val Gly
290 295 300
Arg Ala Gln Lys Lys Tyr Glu Arg Ala Glu Glu Leu Lys Lys Gln Tyr
305 310 315 320
Glu Gln Tyr Arg Leu Glu Lys Leu Ala Lys Phe Gln Gly Val Asn Leu
325 330 335
Phe Ile Lys Asn Leu Asp Asp Ser Ile Asp Asp Glu Lys Leu Lys Glu
340 345 350
Glu Phe Ala Pro Tyr Gly Thr Ile Thr Ser Ala Arg Val Met Arg Asp
355 360 365
Gln Glu Gly Asn Ser Lys Gly Phe Gly Phe Val Cys Phe Ser Ser Pro
370 375 380
Glu Glu Ala Thr Lys Ala Met Thr Glu Lys Asn Gln Gln Ile Val Ala
385 390 395 400
Gly Lys Pro Leu Tyr Val Ala Ile Ala Gln Arg Lys Asp Val Arg Arg
405 410 415
Ser Gln Leu Ala Gln Gln Ile Gln Ala Arg Asn Gln Ile Arg Phe Gln
420 425 430
Gln Gln Gln Gln Gln Gln Ala Ala Ala Ala Ala Ala Gly Met Pro Gly

435 440 445

Gln Tyr Met Pro Gln Met Phe Tyr Gly Val Met Ala Pro Arg Gly Phe
450 455 460

Pro Gly Pro Asn Pro Gly Met Asn Gly Pro Met Gly Ala Gly Ile Pro
465 470 475 480

Lys Asn Gly Met Val Pro Pro Pro Gln Gln Phe Ala Gly Arg Pro Asn
485 490 495

Gly Pro Met Tyr Gln Gly Met Pro Pro Gln Asn Gln Phe Pro Arg His
500 505 510

Gln Gln Gln His Tyr Ile Gln Gln Gln Lys Gln Arg Gln Ala Leu Gly
515 520 525

Glu Gln Leu Tyr Lys Lys Val Ser Ala Lys Ile Asp Asp Glu Asn Ala
530 535 540

Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val
545 550 555 560

Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu
565 570 575

Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Thr Gln Asp Glu Val Gln
595 600 605

Gly Pro His Ala Gly Lys Ser Thr Val Gly Gly Gly Gly Ser Gly Glu
610 615 620

Pro Thr Ser Asp Gln Gln Pro Ala Val Glu Ala Pro Val Val Gln Glu
625 630 635 640

Glu Thr Thr Ser Ser Pro Gln Lys Asn Ser Gly Tyr Val Lys Asn Thr
645 650 655

Ala Gly Ser Gly Ala Pro Arg Asn Gly Lys Tyr Asp Gly Asn Arg Lys
660 665 670

Asn Ser Arg Pro Tyr Asn Gln Arg Gly Asn Asn Asn Asn Asn Asn Gly
675 680 685

Ser Ser Ser Asn Lys His Tyr Gln Lys Tyr Asn Gln Pro Ala Tyr Gly
690 695 700

Val Ser Ala Gly Tyr Ile Pro Asn Tyr Gly Val Ser Ala Glu Tyr Asn
705 710 715 720

Pro Leu Tyr Tyr Asn Gln Tyr Gln Gln Gln Gln Gln Leu Tyr Ala Ala
725 730 735

Ala Tyr Gln Thr Pro Met Ser Gly Gln Gly Tyr Val Pro Pro Val Val
740 745 750

Ser Pro Ala Ala Val Ser Ala Lys Pro Ala Lys Val Glu Ile Thr Asn
755 760 765
Lys Ser Gly Glu His Ile Asp Ile Ala Ser Ile Ala His Pro His Thr
770 775 780
His Ser His Ser Gln Ser His Ser Arg Ala Val Pro Val Val Ser Pro
785 790 795 800
Pro Ala Asn Val Thr Val Ala Ala Ala Val Ser Ser Ser Val Ser Pro
805 810 815
Ser Ala Ser Pro Ala Val Lys Val Gln Ser Pro Ala Ala Asn Gly Lys
820 825 830
Glu Gln Ser Pro Ala Lys Pro Glu Glu Pro Lys Lys Asp Thr Leu Ile
835 840 845
Val Asn Asp Phe Leu Glu Gln Val Lys Arg Arg Lys Ala Ala Leu Ala
850 855 860
Ala Lys Lys Ala Val Glu Glu Lys Gly Pro Glu Glu Pro Lys Glu Ser
865 870 875 880
Val Val Gly Thr Asp Thr Asp Ala Ser Val Asp Thr Lys Thr Gly Pro
885 890 895
Thr Ala Thr Glu Ser Ala Lys Ser Glu Glu Ala Gln Ser Glu Ser Gln
900 905 910
Glu Lys Thr Lys Glu Glu Ala Pro Ala Glu Pro Lys Pro Leu Thr Leu
915 920 925
Ala Glu Lys Leu Arg Leu Lys Arg Met Glu Ala Ala Lys Gln Ala Ser
930 935 940
Ala Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr
945 950 955 960
Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr Lys
965 970 975
Thr Glu Glu Leu Lys Thr Glu Glu Thr Lys Ser Glu Glu Leu Lys Thr
980 985 990
Glu Glu Pro Lys Ala Glu Glu Ser Lys Ala Glu Glu Pro Lys Pro Glu
995 1000 1005
Glu Pro Lys Thr Glu Glu Pro Thr Thr Glu Gln Pro Lys Ser Asp Glu
1010 1015 1020
Pro Lys Ser Glu Glu Ser Lys Thr Glu Glu Pro Lys Thr Glu Val Leu
1025 1030 1035 1040
Lys Thr Glu Glu Pro Lys Ser Glu Glu Ser Lys Pro Ala Glu Pro Lys
1045 1050 1055
Thr Glu Glu Thr Ala Thr Glu Glu Thr Ala Thr Glu Ala Asn Ala Glu

1060 1065 1070

Glu Gly Glu Pro Ala Pro Ala Gly Pro Val Glu Thr Pro Ala Asp Val

1075 1080 1085

Glu Thr Lys Pro Arg Glu Glu Ala Glu Val Glu Asp Asp Gly Lys Ile

1090 1095 1100

Thr Met Thr Asp Phe Leu Gln Lys Leu Lys Glu Val Ser Pro Val Asp

1105 1110 1115 1120

Asp Ile Tyr Ser Phe Gln Tyr Pro Ser Asp Ile Thr Pro Pro Asn Asp

1125 1130 1135

Arg Tyr Lys Lys Thr Ser Ile Lys Tyr Ala Tyr Gly Pro Asp Phe Leu

1140 1145 1150

Tyr Gln Phe Lys Glu Lys Val Asp Val Lys Tyr Asp Pro Ala Trp Met

1155 1160 1165

Ala Glu Met Thr Ser Lys Ile Val Ile Pro Pro Lys Lys Pro Gly Ser

1170 1175 1180

Ser Gly Arg Gly Glu Asp

1185 1190

<210> 15

<211> 3678

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 15

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60

cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120

tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180

attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240

gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300

tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360

tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420

ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480

ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540

ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600

atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660

tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720

gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780

gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840

cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900

ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960

gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020

ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500
caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctggcg aacctacatc cgatcagcaa ccagctgttg aagctccagt tgtgcaggag 1920
gagacaacca gttctccgca aaaaaacagt ggatatgtca agaatactgc tggaagcggt 1980
gctcctagaa atgggaaata tgatggtaac aggaagaact ctaggcctta taaccaaaga 2040
ggtaacaaca acaataataa tggttcttcc tcgaataagc actatcaaaa gtataaccaa 2100
ccagcgtacg gtgtttctgc gggatacatt ccgaactacg gcgtatcggc agagtacaac 2160
cctctgtact ataaccagta ccaacagcag caacagctgt acgctgctgc ttaccagact 2220
ccaatgagcg gacaaggtta tgtcccccca gtagtgtctc cagctgctgt ttcagctaaa 2280
ccagcgaagg ttgagattac taacaagtct ggtgaacaca tagatattgc ttccattgct 2340
catccacata ctcattctca ttctcaatct cattcgcgtg cagttccagt agtgtcgcct 2400
ccagctaacg ttaccgtcgc tgctgctgta tcatcctctg tgtctccatc agcttctcca 2460
gctgtcaaag tacagagccc tgctgctaat ggtaaggaac aatctccagc taagcctgaa 2520
gaaccaaaga aggacacttt aattgtgaac gatttcttgg aacaagttaa aagacgcaag 2580
gctgctttag ctgctaagaa ggctgtcgaa gagaagggtc ctgaggaacc gaaggaatct 2640
gtcgttggaa ctgacactga tgcaagcgtt gatactaaga cagggcctac agccactgaa 2700
tctgccaagt ctgaagaagc tcaatcagaa tcacaagaaa agactaagga agaggctcca 2760
gctgagccaa aaccattgac tttggccgaa aaattgagac ttaagaggat ggaagctgca 2820
aagcaagctt ctgctaagac cgaggaacta aagactgaag aatctaagcc tgaagaaaca 2880
aagaccgagg agctaaagac tgaagaatct aagcctgaag aaacaaagac cgaggagcta 2940
aagactgaag aaacaaagtc cgaggaacta aagactgaag aacctaaggc ggaagaatca 3000
aaggcggaag aaccaaagcc tgaagaacca aagaccgagg aaccgacgac tgaacaacca 3060
aagtcagatg aaccaaagtc ggaagaatca aaaactgaag agccaaaaac cgaggtatta 3120
aagactgaag aaccaaaatc ggaagaatca aagcctgcag aaccaaagac tgaagaaaca 3180
gcaactgaag aaacagcaac tgaagcaaac gccgaagaag gtgaaccggc tcctgctggt 3240
cccgttgaaa ctcctgctga tgttgaaaca aaacctcgag aagaggctga agttgaagac 3300
gatggaaaga ttaccatgac cgatttccta cagaagttga aagaggtttc tccagttgat 3360

gatatttatt ccttccaata cccaagtgac attacgcctc caaatgatag atataaaaag 3420
acaagcatta aatatgcata cggacctgat ttcttgtatc agttcaaaga aaaggtcgat 3480
gttaaatacg atccagcgtg gatggctgaa atgacgagta aaattgtcat ccctcctaag 3540
aagcctggtt caagcggaag aggcgaagat agatttagta agggtaaggt tggatctcta 3600
agaagtgaag gcagatcggg ttccaggtcc aactcgaaga agaagtcaaa gagggatgat 3660
agaaaatcta atagataa 3678

<210> 16
<211> 1225
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 16

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

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

530 535 540

Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val
545 550 555 560

Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu
565 570 575

Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Thr Gln Asp Glu Val Gln
595 600 605

Gly Pro His Ala Gly Lys Ser Thr Val Gly Gly Gly Gly Ser Gly Glu
610 615 620

Pro Thr Ser Asp Gln Gln Pro Ala Val Glu Ala Pro Val Val Gln Glu
625 630 635 640

Glu Thr Thr Ser Ser Pro Gln Lys Asn Ser Gly Tyr Val Lys Asn Thr
645 650 655

Ala Gly Ser Gly Ala Pro Arg Asn Gly Lys Tyr Asp Gly Asn Arg Lys
660 665 670

Asn Ser Arg Pro Tyr Asn Gln Arg Gly Asn Asn Asn Asn Asn Asn Gly
675 680 685

Ser Ser Ser Asn Lys His Tyr Gln Lys Tyr Asn Gln Pro Ala Tyr Gly
690 695 700

Val Ser Ala Gly Tyr Ile Pro Asn Tyr Gly Val Ser Ala Glu Tyr Asn
705 710 715 720

Pro Leu Tyr Tyr Asn Gln Tyr Gln Gln Gln Gln Gln Leu Tyr Ala Ala
725 730 735

Ala Tyr Gln Thr Pro Met Ser Gly Gln Gly Tyr Val Pro Pro Val Val
740 745 750

Ser Pro Ala Ala Val Ser Ala Lys Pro Ala Lys Val Glu Ile Thr Asn
755 760 765

Lys Ser Gly Glu His Ile Asp Ile Ala Ser Ile Ala His Pro His Thr
770 775 780

His Ser His Ser Gln Ser His Ser Arg Ala Val Pro Val Val Ser Pro
785 790 795 800

Pro Ala Asn Val Thr Val Ala Ala Ala Val Ser Ser Ser Val Ser Pro
805 810 815

Ser Ala Ser Pro Ala Val Lys Val Gln Ser Pro Ala Ala Asn Gly Lys
820 825 830

Glu Gln Ser Pro Ala Lys Pro Glu Glu Pro Lys Lys Asp Thr Leu Ile
835 840 845

Val Asn Asp Phe Leu Glu Gln Val Lys Arg Arg Lys Ala Ala Leu Ala
850 855 860
Ala Lys Lys Ala Val Glu Glu Lys Gly Pro Glu Glu Pro Lys Glu Ser
865 870 875 880
Val Val Gly Thr Asp Thr Asp Ala Ser Val Asp Thr Lys Thr Gly Pro
885 890 895
Thr Ala Thr Glu Ser Ala Lys Ser Glu Glu Ala Gln Ser Glu Ser Gln
900 905 910
Glu Lys Thr Lys Glu Glu Ala Pro Ala Glu Pro Lys Pro Leu Thr Leu
915 920 925
Ala Glu Lys Leu Arg Leu Lys Arg Met Glu Ala Ala Lys Gln Ala Ser
930 935 940
Ala Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr
945 950 955 960
Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr Lys
965 970 975
Thr Glu Glu Leu Lys Thr Glu Glu Thr Lys Ser Glu Glu Leu Lys Thr
980 985 990
Glu Glu Pro Lys Ala Glu Glu Ser Lys Ala Glu Glu Pro Lys Pro Glu
995 1000 1005
Glu Pro Lys Thr Glu Glu Pro Thr Thr Glu Gln Pro Lys Ser Asp Glu
1010 1015 1020
Pro Lys Ser Glu Glu Ser Lys Thr Glu Glu Pro Lys Thr Glu Val Leu
1025 1030 1035 1040
Lys Thr Glu Glu Pro Lys Ser Glu Glu Ser Lys Pro Ala Glu Pro Lys
1045 1050 1055
Thr Glu Glu Thr Ala Thr Glu Glu Thr Ala Thr Glu Ala Asn Ala Glu
1060 1065 1070
Glu Gly Glu Pro Ala Pro Ala Gly Pro Val Glu Thr Pro Ala Asp Val
1075 1080 1085
Glu Thr Lys Pro Arg Glu Glu Ala Glu Val Glu Asp Asp Gly Lys Ile
1090 1095 1100
Thr Met Thr Asp Phe Leu Gln Lys Leu Lys Glu Val Ser Pro Val Asp
1105 1110 1115 1120
Asp Ile Tyr Ser Phe Gln Tyr Pro Ser Asp Ile Thr Pro Pro Asn Asp
1125 1130 1135
Arg Tyr Lys Lys Thr Ser Ile Lys Tyr Ala Tyr Gly Pro Asp Phe Leu
1140 1145 1150
Tyr Gln Phe Lys Glu Lys Val Asp Val Lys Tyr Asp Pro Ala Trp Met

1155 1160 1165

Ala Glu Met Thr Ser Lys Ile Val Ile Pro Pro Lys Lys Pro Gly Ser

1170 1175 1180

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

1185 1190 1195 1200

Arg Ser Glu Gly Arg Ser Gly Ser Arg Ser Asn Ser Lys Lys Lys Ser

1205 1210 1215

Lys Arg Asp Asp Arg Lys Ser Asn Arg

1220 1225

<210> 17

<211> 4680

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 17

atgtctgata ttactgaaaa aactgctgag caattggaaa acttgcagat caacgatgat 60
cagcaaccag ctcaatctgc cagtgctcca tccacttctg cttctgaaag cgaagcttct 120
tctgtttcta aggttgaaaa caacaacgct tcattgtacg ttggtgaatt ggatccaaac 180
attactgaag cattgttgta cgatgtgttt tcaccattgg gtccaatttc ctcgatccgt 240
gtttgtcgtg atgccgtcac caaggcttcg ttaggttacg cttacgttaa ctatactgat 300
tacgaagctg gtaagaaagc tattcaagaa ttgaactatg ctgaaatcaa cggtagacca 360
tgtagaatta tgtggtccga acgtgaccca gctatcagaa agaagggttc tggtaacatt 420
ttcatcaaga acttgcaccc agccattgac aacaaggctt tgcatgaaac tttctccact 480
ttcggtgaag tcttgtcttg taaagttgct ttagatgaga atggaaactc tagaggcttc 540
ggtttcgttc atttcaagga agaatccgat gctaaggatg ctattgaagc cgtcaacggt 600
atgttgatga acggtttgga agtttacgtt gccatgcacg ttccaaagaa ggaccgtatc 660
tccaagttgg aagaagccaa ggctaacttc accaacattt acgtcaagaa cattgacgtt 720
gaaaccactg acgaagagtt cgaacagttg ttctcccaat acggtgaaat tgtctctgct 780
gctttggaaa aggatgctga gggtaagcca aagggtttcg gtttcgttaa ctttgttgac 840
cacaacgccg ctgccaaggc cgttgaagag ttgaacggta aggaattcaa gtctcaagct 900
ttgtacgttg gcagagctca aaagaagtac gaacgtgctg aagaattgaa gaaacaatac 960
gaacaatacc gtttggaaaa attggctaag ttccaaggtg ttaacttgtt catcaagaac 1020
ttggacgatt ccatcgatga cgaaaaattg aaggaagaat tcgccccata cggtaccatc 1080
acctctgcta gagtcatgag agaccaagag ggtaactcta agggtttcgg tttcgtttgt 1140
ttctcttctc cagaagaagc taccaaggct atgaccgaaa agaaccaaca aattgttgcc 1200
ggtaagccat tgtacgttgc cattgctcaa agaaaggatg tcagaagatc ccaattggct 1260
caacaaattc aagccagaaa ccaaatcaga ttccaacaac agcaacaaca acaagctgct 1320
gccgctgctg ctggtatgcc aggccaatac atgccacaaa tgttctatgg tgttatggcc 1380
ccaagaggtt tcccaggtcc aaacccaggt atgaacggcc caatgggtgc cggtattcca 1440
aagaacggta tggtcccacc accacaacaa tttgctggta gaccaaacgg tccaatgtac 1500

caaggtatgc cacctcaaaa ccaattccca agacaccaac aacaacacta catccaacaa 1560
caaaagcaaa gacaagcctt gggtgaacaa ttgtacaaga aggtcagtgc caagattgac 1620
gacgaaaacg ccgctggtaa gatcaccggt atgatcttgg atctaccacc acagcaagtc 1680
atccaattgt tggacaacga cgaacaattt gaacagcaat tccaagaagc cttagctgct 1740
tacgaaaact tcaagaagga acaagaagct caagctggtg gaggtggatc tggaggtgga 1800
ggttcaactc aagatgaagt tcaaggtcca catgctggta agtctactgt tggtggaggt 1860
ggatctggcg aacctacatc cgatcagcaa ccagctgttg aagctccagt tgtgcaggag 1920
gagacaacca gttctccgca aaaaaacagt ggatatgtca agaatactgc tggaagcggt 1980
gctcctagaa atgggaaata tgatggtaac aggaagaact ctaggcctta taaccaaaga 2040
ggtaacaaca acaataataa tggttcttcc tcgaataagc actatcaaaa gtataaccaa 2100
ccagcgtacg gtgtttctgc gggatacatt ccgaactacg gcgtatcggc agagtacaac 2160
cctctgtact ataaccagta ccaacagcag caacagctgt acgctgctgc ttaccagact 2220
ccaatgagcg gacaaggtta tgtcccccca gtagtgtctc cagctgctgt ttcagctaaa 2280
ccagcgaagg ttgagattac taacaagtct ggtgaacaca tagatattgc ttccattgct 2340
catccacata ctcattctca ttctcaatct cattcgcgtg cagttccagt agtgtcgcct 2400
ccagctaacg ttaccgtcgc tgctgctgta tcatcctctg tgtctccatc agcttctcca 2460
gctgtcaaag tacagagccc tgctgctaat ggtaaggaac aatctccagc taagcctgaa 2520
gaaccaaaga aggacacttt aattgtgaac gatttcttgg aacaagttaa aagacgcaag 2580
gctgctttag ctgctaagaa ggctgtcgaa gagaagggtc ctgaggaacc gaaggaatct 2640
gtcgttggaa ctgacactga tgcaagcgtt gatactaaga cagggcctac agccactgaa 2700
tctgccaagt ctgaagaagc tcaatcagaa tcacaagaaa agactaagga agaggctcca 2760
gctgagccaa aaccattgac tttggccgaa aaattgagac ttaagaggat ggaagctgca 2820
aagcaagctt ctgctaagac cgaggaacta aagactgaag aatctaagcc tgaagaaaca 2880
aagaccgagg agctaaagac tgaagaatct aagcctgaag aaacaaagac cgaggagcta 2940
aagactgaag aaacaaagtc cgaggaacta aagactgaag aacctaaggc ggaagaatca 3000
aaggcggaag aaccaaagcc tgaagaacca aagaccgagg aaccgacgac tgaacaacca 3060
aagtcagatg aaccaaagtc ggaagaatca aaaactgaag agccaaaaac cgaggtatta 3120
aagactgaag aaccaaaatc ggaagaatca aagcctgcag aaccaaagac tgaagaaaca 3180
gcaactgaag aaacagcaac tgaagcaaac gccgaagaag gtgaaccggc tcctgctggt 3240
cccgttgaaa ctcctgctga tgttgaaaca aaacctcgag aagaggctga agttgaagac 3300
gatggaaaga ttaccatgac cgatttccta cagaagttga aagaggtttc tccagttgat 3360
gatatttatt ccttccaata cccaagtgac attacgcctc caaatgatag atataaaaag 3420
acaagcatta aatatgcata cggacctgat ttcttgtatc agttcaaaga aaaggtcgat 3480
gttaaatacg atccagcgtg gatggctgaa atgacgagta aaattgtcat ccctcctaag 3540
aagcctggtt caagcggaag aggcgaagat agatttagta agggtaaggt tggatctcta 3600
agaagtgaag gcagatcggg ttccaggtcc aactcgaaga agaagtcaaa gagggatgat 3660
agaaaatcta atagatcata cacttccaga aaggaccgtg aaagattcag agaggaagaa 3720
gtcgaagagc caaaggttga ggttgcccca ttggtcccaa gtgctaatag atgggttcct 3780
aaatctaaga tgaagaaaac agaagtcaag ttagctccag acggaacaga actttacgac 3840

gcggaagaag catcaagaaa gatgaagtca ttgctgaata aattgacatt agaaatgttc 3900
gaacctattt ctgatgatat catgaagatc gctaaccaat ctagatggga agaaaagggt 3960
gagactttga agattgtcat ccaacaaatt ttcaataagg cctgcgatga acctcattgg 4020
tcatcaatgt acgcgcaatt atgtggtaag gtcgttaaag acttagatga tagcattaaa 4080
gactcagaaa ccccagataa gactggttct cacttggttt tgcattactt agtccaaaga 4140
tgtcaaactg aattccaaac aggatggact gatcaactac ctacaaacga agacggtact 4200
cctctacaac ctgaaatgat gtccgatgaa tactataaga tggctgccgc taagagaaga 4260
ggtttgggtt tggttcgttt cattggtttc ttgtaccgtt cgaacttatt gacttccaga 4320
atggtcttct tctgtttcaa gagactaatg aaggatattc aaaactctcc tactgaagat 4380
actctagagt ctgtatgtga acttttggaa acaattggtg aacagttcga aggtgctcgt 4440
attcaagtta ctgcagaagc tgtcattgag ggttcaagct tgctagacac actattcgac 4500
caaataaaga acgtgatcga aaatggtgac atctccagca gaatcaagtt taagttgatc 4560
gacattgtcg aactaagaga aaagaggaac tggaatagta aaaataagaa cgatggtcca 4620
aagaccattg ctcaaattca cgaagaagaa gccttgaaga gggctttgga ggaaagataa 4680

<210> 18
<211> 1559
<212> PRT
<213> 人工序列(Artificial Sequence)
<400> 18

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

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

465 470 475 480
Lys Asn Gly Met Val Pro Pro Pro Gln Gln Phe Ala Gly Arg Pro Asn
485 490 495
Gly Pro Met Tyr Gln Gly Met Pro Pro Gln Asn Gln Phe Pro Arg His
500 505 510
Gln Gln Gln His Tyr Ile Gln Gln Gln Lys Gln Arg Gln Ala Leu Gly
515 520 525
Glu Gln Leu Tyr Lys Lys Val Ser Ala Lys Ile Asp Asp Glu Asn Ala
530 535 540
Ala Gly Lys Ile Thr Gly Met Ile Leu Asp Leu Pro Pro Gln Gln Val
545 550 555 560
Ile Gln Leu Leu Asp Asn Asp Glu Gln Phe Glu Gln Gln Phe Gln Glu
565 570 575
Ala Leu Ala Ala Tyr Glu Asn Phe Lys Lys Glu Gln Glu Ala Gln Ala
580 585 590
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Thr Gln Asp Glu Val Gln
595 600 605
Gly Pro His Ala Gly Lys Ser Thr Val Gly Gly Gly Gly Ser Gly Glu
610 615 620
Pro Thr Ser Asp Gln Gln Pro Ala Val Glu Ala Pro Val Val Gln Glu
625 630 635 640
Glu Thr Thr Ser Ser Pro Gln Lys Asn Ser Gly Tyr Val Lys Asn Thr
645 650 655
Ala Gly Ser Gly Ala Pro Arg Asn Gly Lys Tyr Asp Gly Asn Arg Lys
660 665 670
Asn Ser Arg Pro Tyr Asn Gln Arg Gly Asn Asn Asn Asn Asn Asn Gly
675 680 685
Ser Ser Ser Asn Lys His Tyr Gln Lys Tyr Asn Gln Pro Ala Tyr Gly
690 695 700
Val Ser Ala Gly Tyr Ile Pro Asn Tyr Gly Val Ser Ala Glu Tyr Asn
705 710 715 720
Pro Leu Tyr Tyr Asn Gln Tyr Gln Gln Gln Gln Gln Leu Tyr Ala Ala
725 730 735
Ala Tyr Gln Thr Pro Met Ser Gly Gln Gly Tyr Val Pro Pro Val Val
740 745 750
Ser Pro Ala Ala Val Ser Ala Lys Pro Ala Lys Val Glu Ile Thr Asn
755 760 765
Lys Ser Gly Glu His Ile Asp Ile Ala Ser Ile Ala His Pro His Thr
770 775 780

His Ser His Ser Gln Ser His Ser Arg Ala Val Pro Val Val Ser Pro
785 790 795 800
Pro Ala Asn Val Thr Val Ala Ala Ala Val Ser Ser Ser Val Ser Pro
805 810 815
Ser Ala Ser Pro Ala Val Lys Val Gln Ser Pro Ala Ala Asn Gly Lys
820 825 830
Glu Gln Ser Pro Ala Lys Pro Glu Glu Pro Lys Lys Asp Thr Leu Ile
835 840 845
Val Asn Asp Phe Leu Glu Gln Val Lys Arg Arg Lys Ala Ala Leu Ala
850 855 860
Ala Lys Lys Ala Val Glu Glu Lys Gly Pro Glu Glu Pro Lys Glu Ser
865 870 875 880
Val Val Gly Thr Asp Thr Asp Ala Ser Val Asp Thr Lys Thr Gly Pro
885 890 895
Thr Ala Thr Glu Ser Ala Lys Ser Glu Glu Ala Gln Ser Glu Ser Gln
900 905 910
Glu Lys Thr Lys Glu Glu Ala Pro Ala Glu Pro Lys Pro Leu Thr Leu
915 920 925
Ala Glu Lys Leu Arg Leu Lys Arg Met Glu Ala Ala Lys Gln Ala Ser
930 935 940
Ala Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr
945 950 955 960
Lys Thr Glu Glu Leu Lys Thr Glu Glu Ser Lys Pro Glu Glu Thr Lys
965 970 975
Thr Glu Glu Leu Lys Thr Glu Glu Thr Lys Ser Glu Glu Leu Lys Thr
980 985 990
Glu Glu Pro Lys Ala Glu Glu Ser Lys Ala Glu Glu Pro Lys Pro Glu
995 1000 1005
Glu Pro Lys Thr Glu Glu Pro Thr Thr Glu Gln Pro Lys Ser Asp Glu
1010 1015 1020
Pro Lys Ser Glu Glu Ser Lys Thr Glu Glu Pro Lys Thr Glu Val Leu
1025 1030 1035 1040
Lys Thr Glu Glu Pro Lys Ser Glu Glu Ser Lys Pro Ala Glu Pro Lys
1045 1050 1055
Thr Glu Glu Thr Ala Thr Glu Glu Thr Ala Thr Glu Ala Asn Ala Glu
1060 1065 1070
Glu Gly Glu Pro Ala Pro Ala Gly Pro Val Glu Thr Pro Ala Asp Val
1075 1080 1085
Glu Thr Lys Pro Arg Glu Glu Ala Glu Val Glu Asp Asp Gly Lys Ile

1090 1095 1100

Thr Met Thr Asp Phe Leu Gln Lys Leu Lys Glu Val Ser Pro Val Asp
1105 1110 1115 1120

Asp Ile Tyr Ser Phe Gln Tyr Pro Ser Asp Ile Thr Pro Pro Asn Asp
1125 1130 1135

Arg Tyr Lys Lys Thr Ser Ile Lys Tyr Ala Tyr Gly Pro Asp Phe Leu
1140 1145 1150

Tyr Gln Phe Lys Glu Lys Val Asp Val Lys Tyr Asp Pro Ala Trp Met
1155 1160 1165

Ala Glu Met Thr Ser Lys Ile Val Ile Pro Pro Lys Lys Pro Gly Ser
1170 1175 1180

Ser Gly Arg Gly Glu Asp Arg Phe Ser Lys Gly Lys Val Gly Ser Leu
1185 1190 1195 1200

Arg Ser Glu Gly Arg Ser Gly Ser Arg Ser Asn Ser Lys Lys Lys Ser
1205 1210 1215

Lys Arg Asp Asp Arg Lys Ser Asn Arg Ser Tyr Thr Ser Arg Lys Asp
1220 1225 1230

Arg Glu Arg Phe Arg Glu Glu Glu Val Glu Glu Pro Lys Val Glu Val
1235 1240 1245

Ala Pro Leu Val Pro Ser Ala Asn Arg Trp Val Pro Lys Ser Lys Met
1250 1255 1260

Lys Lys Thr Glu Val Lys Leu Ala Pro Asp Gly Thr Glu Leu Tyr Asp
1265 1270 1275 1280

Ala Glu Glu Ala Ser Arg Lys Met Lys Ser Leu Leu Asn Lys Leu Thr
1285 1290 1295

Leu Glu Met Phe Glu Pro Ile Ser Asp Asp Ile Met Lys Ile Ala Asn
1300 1305 1310

Gln Ser Arg Trp Glu Glu Lys Gly Glu Thr Leu Lys Ile Val Ile Gln
1315 1320 1325

Gln Ile Phe Asn Lys Ala Cys Asp Glu Pro His Trp Ser Ser Met Tyr
1330 1335 1340

Ala Gln Leu Cys Gly Lys Val Val Lys Asp Leu Asp Asp Ser Ile Lys
1345 1350 1355 1360

Asp Ser Glu Thr Pro Asp Lys Thr Gly Ser His Leu Val Leu His Tyr
1365 1370 1375

Leu Val Gln Arg Cys Gln Thr Glu Phe Gln Thr Gly Trp Thr Asp Gln
1380 1385 1390

Leu Pro Thr Asn Glu Asp Gly Thr Pro Leu Gln Pro Glu Met Met Ser
1395 1400 1405

Asp Glu Tyr Tyr Lys Met Ala Ala Ala Lys Arg Arg Gly Leu Gly Leu
1410 1415 1420
Val Arg Phe Ile Gly Phe Leu Tyr Arg Ser Asn Leu Leu Thr Ser Arg
1425 1430 1435 1440
Met Val Phe Phe Cys Phe Lys Arg Leu Met Lys Asp Ile Gln Asn Ser
1445 1450 1455
Pro Thr Glu Asp Thr Leu Glu Ser Val Cys Glu Leu Leu Glu Thr Ile
1460 1465 1470
Gly Glu Gln Phe Glu Gly Ala Arg Ile Gln Val Thr Ala Glu Ala Val
1475 1480 1485
Ile Glu Gly Ser Ser Leu Leu Asp Thr Leu Phe Asp Gln Ile Lys Asn
1490 1495 1500
Val Ile Glu Asn Gly Asp Ile Ser Ser Arg Ile Lys Phe Lys Leu Ile
1505 1510 1515 1520
Asp Ile Val Glu Leu Arg Glu Lys Arg Asn Trp Asn Ser Lys Asn Lys
1525 1530 1535
Asn Asp Gly Pro Lys Thr Ile Ala Gln Ile His Glu Glu Glu Ala Leu
1540 1545 1550
Lys Arg Ala Leu Glu Glu Arg
1555

<210> 19
<211> 36
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 19
ttggtggagg tggatctaac caaccagcgt acggtg 36
<210> 20
<211> 23
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 20
agatccacct ccaccaacag tag 23
<210> 21
<211> 17
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 21
gtaaaacgac ggccagt 17
<210> 22

<211> 17
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 22
caggaaacag ctatgac 17
<210> 23
<211> 38
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 23
ttggtggagg tggatctgct gtttcagcta aaccagcg 38
<210> 24
<211> 46
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 24
ttggtggagg tggatctact ttggccgaaa aattgagact taagag 46
<210> 25
<211> 38
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 25
ttggtggagg tggatctgct gtttcagcta aaccagcg 38
<210> 26
<211> 37
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 26
taaacttgat tttttgacct tgatcttcat cttgtcc 37
<210> 27
<211> 48
<212> DNA
<213> 人工序列(Artificial Sequence)
<400> 27
gaagatcaag gtcaaaaaat caagtttaat cttcgcctct tccgcttg 48
<210> 28
<211> 56
<212> DNA
<213> 人工序列(Artificial Sequence)

<400> 28

gatcaaggtc aaaaaatcaa gtttatctat tagattttct atcatccctc tttgac 56

<210> 29

<211> 53

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 29

gaagatcaag gtcaaaaaat caagtttatc tttcctccaa agccctcttc aag 53

<210> 30

<211> 25

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 30

tctccagaag aagctaccaa ggcta 25

<210> 31

<211> 25

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 31

ttctcttcga cagccttctt agcag 25

<210> 32

<211> 23

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 32

tacccaagtg acattacgcc tcc 23

<210> 33

<211> 25

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 33

ttggaagacc ccattttcat aggga 25

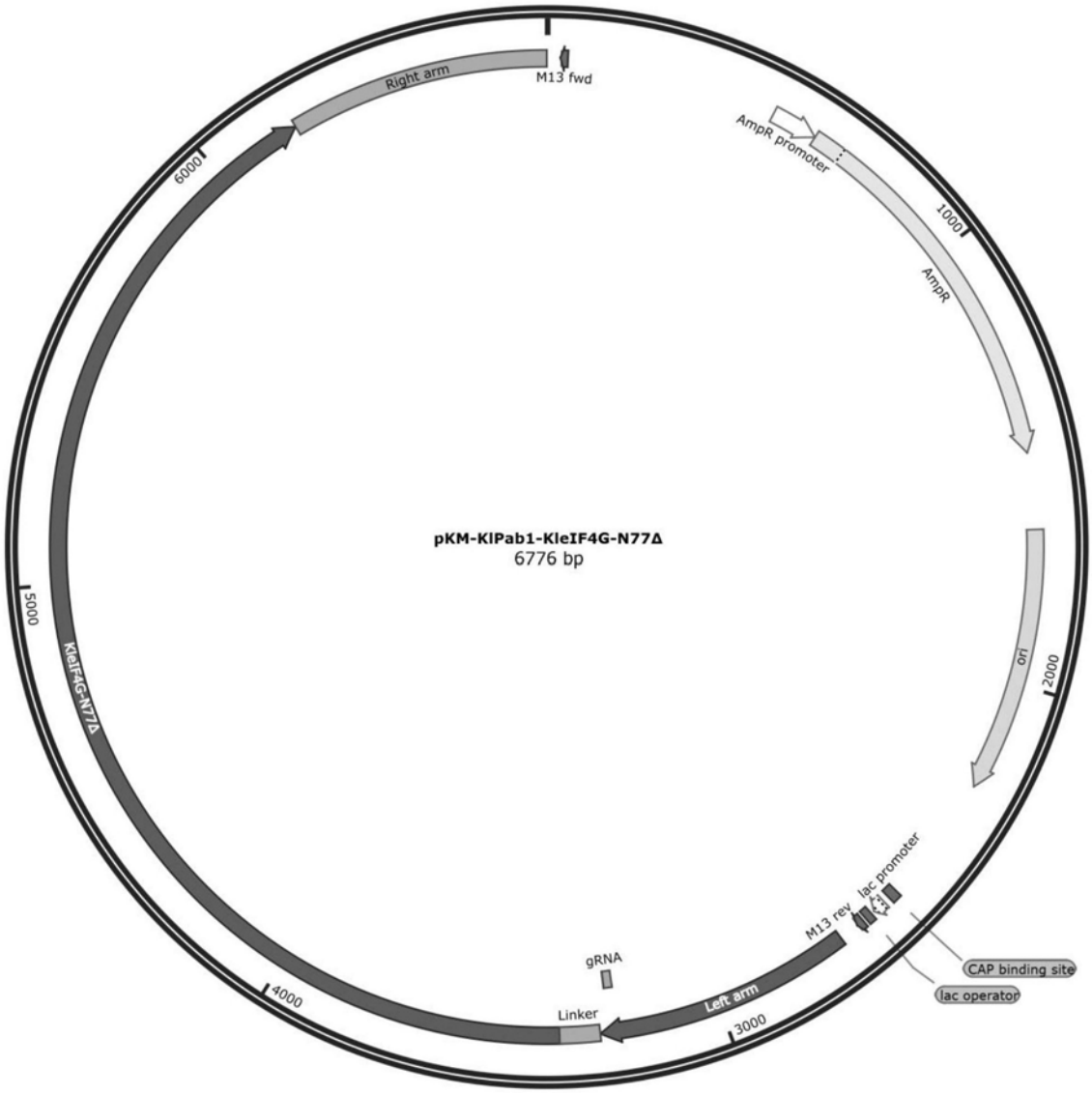

pKM-KlPab1-KleIF4G-N77Δ
6776 bp
Right arm
M13 fwd
AmpR promoter
AmpR
ori
lac promoter
M13 rev
CAP binding site
lac operator
Left arm
gRNA
Linker
KleIF4G-N77Δ
1000
2000
3000
4000
5000
6000

图1


pKM-KlPab1-KleIF4G-N134Δ
6605 bp
Right arm
M13 fwd
AmpR promoter
AmpR
ori
lac promoter
M13 rev
CAP binding site
lac operator
Left arm
Linker
gRNA
KleIF4G-N134Δ
1000
2000
3000
4000
5000
6000

图2

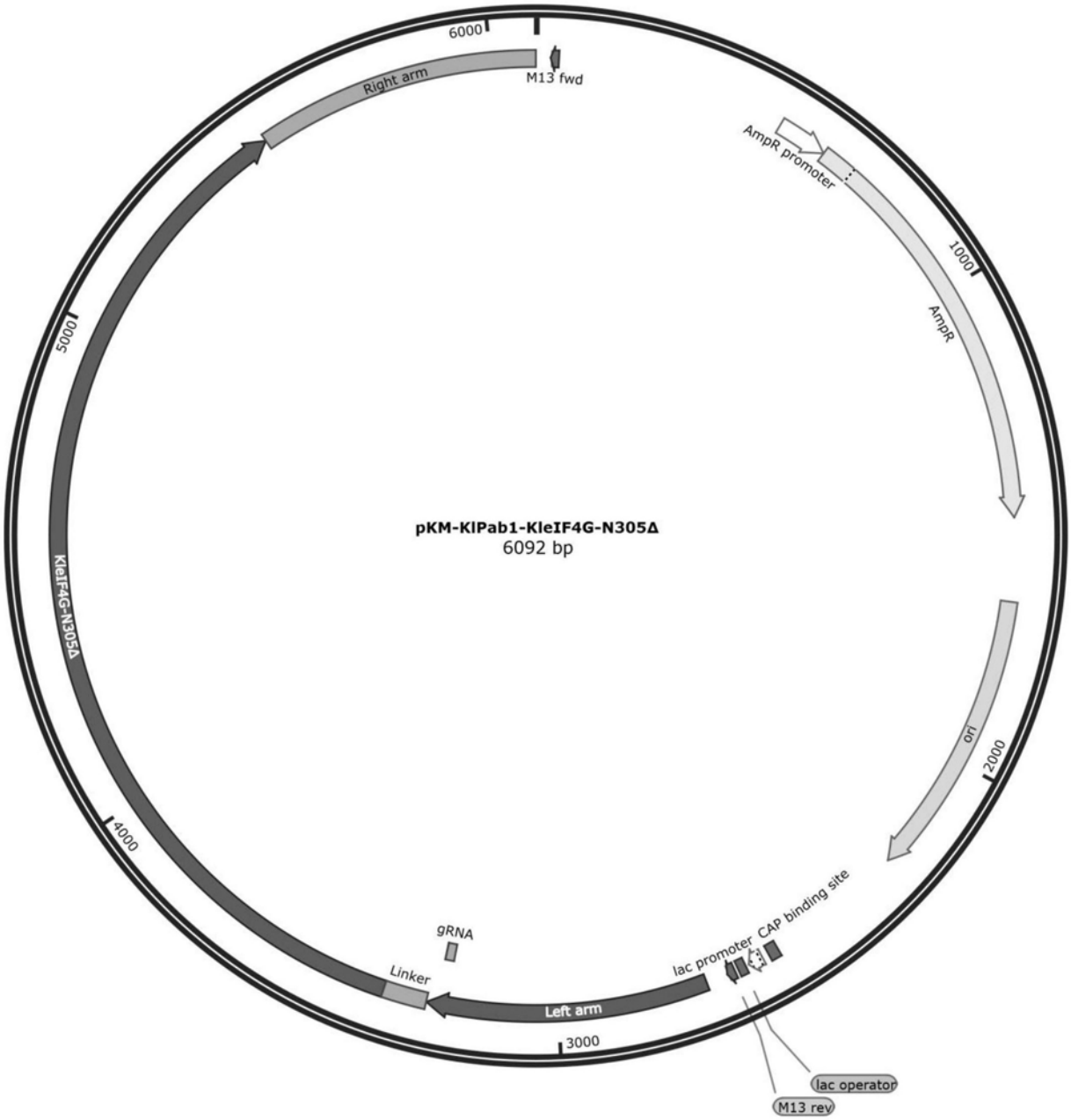

6000
Right arm
M13 fwd
AmpR promoter
1000
AmpR
pKM-KlPab1-KleIF4G-N305Δ
6092 bp
5000
KleIF4G-N305Δ
ori
2000
4000
gRNA
Linker
lac promoter
CAP binding site
Left arm
3000
lac operator
M13 rev

图3


pKM-KlPab1-KleIF4G-N566Δ
5309 bp
5000
Right arm
M13 fwd
AmpR promoter
1000
AmpR
ori
2000
lac operator
M13 rev | lac promoter
CAP binding site
3000
Left arm
Linker
gRNA
4000
KleIF4G-N566Δ

图4

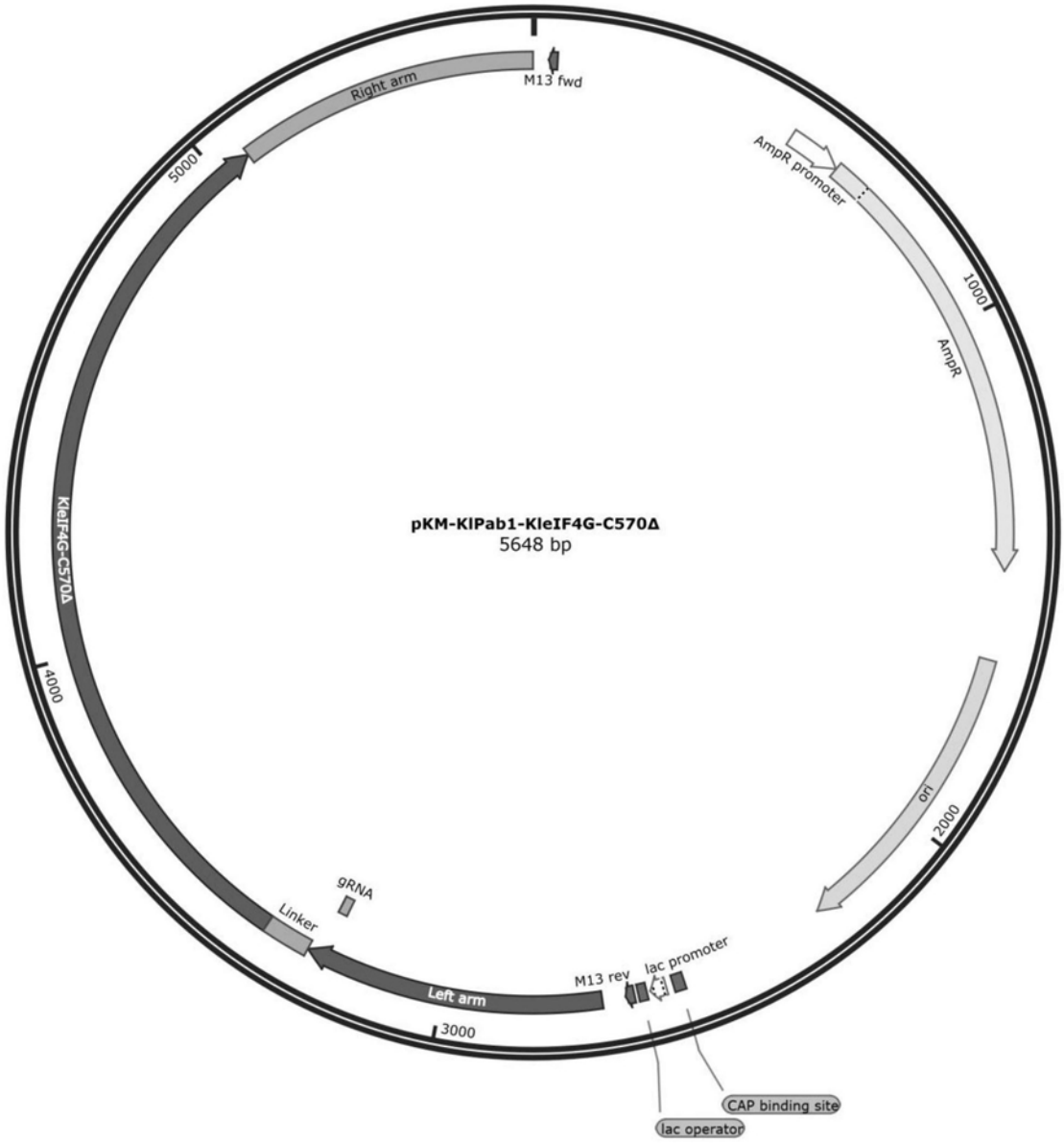

pKM-KlPab1-KleIF4G-C570Δ
5648 bp
Right arm
M13 fwd
AmpR promoter
AmpR
ori
M13 rev
lac promoter
CAP binding site
lac operator
Left arm
Linker
gRNA
KleIF4G-C570Δ
1000
2000
3000
4000
5000

图5

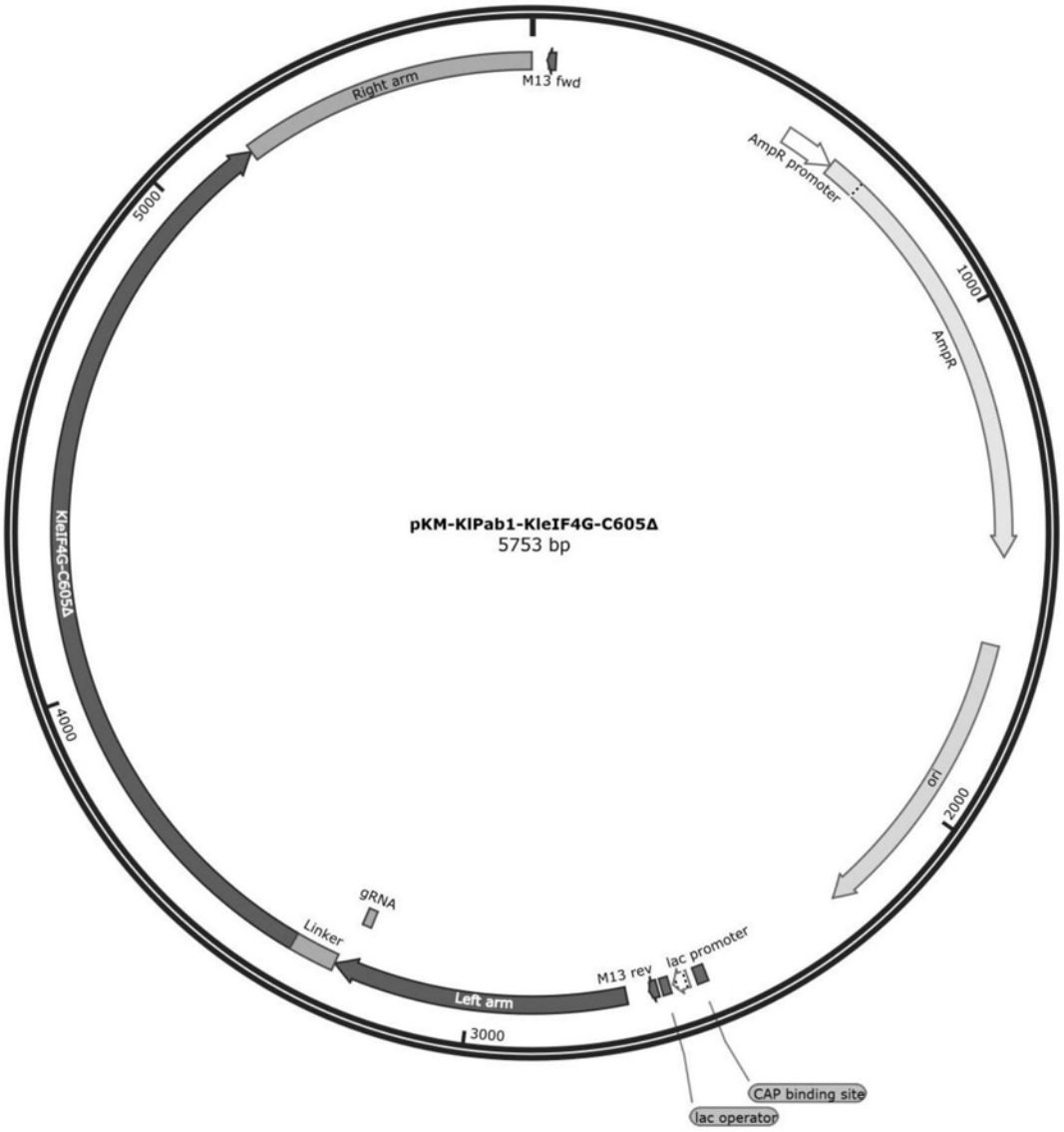

pKM-KlPab1-KleIF4G-C605Δ
5753 bp
Right arm
M13 fwd
AmpR promoter
AmpR
ori
gRNA
Linker
Left arm
M13 rev
lac promoter
CAP binding site
lac operator
KleIF4G-C605Δ
1000
2000
3000
4000
5000

图6

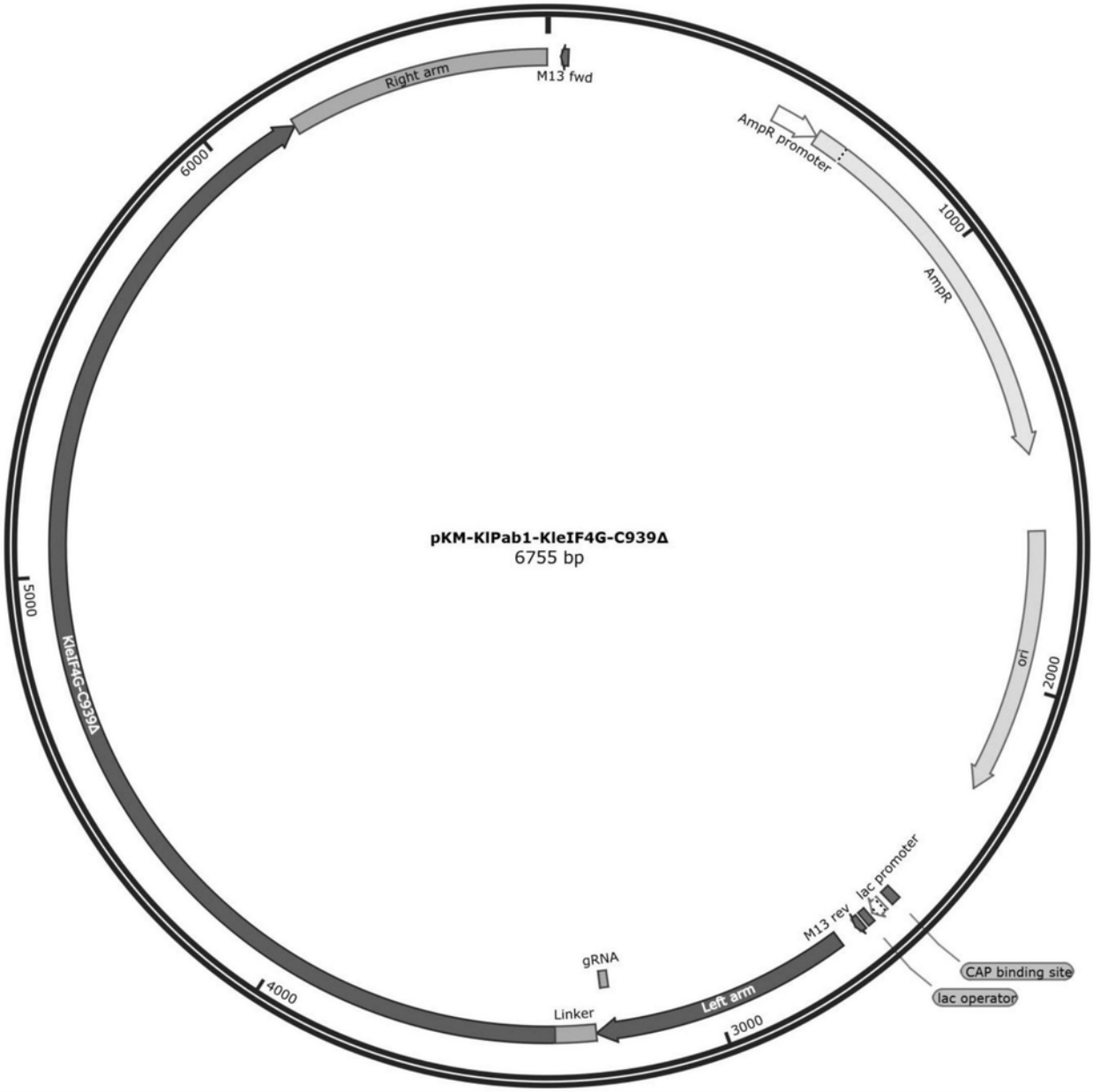

pKM-KlPab1-KleIF4G-C939Δ
6755 bp
Right arm
M13 fwd
AmpR promoter
AmpR
ori
lac promoter
M13 rev
CAP binding site
lac operator
Left arm
gRNA
Linker
KleIF4G-C939Δ
1000
2000
3000
4000
5000
6000

图7

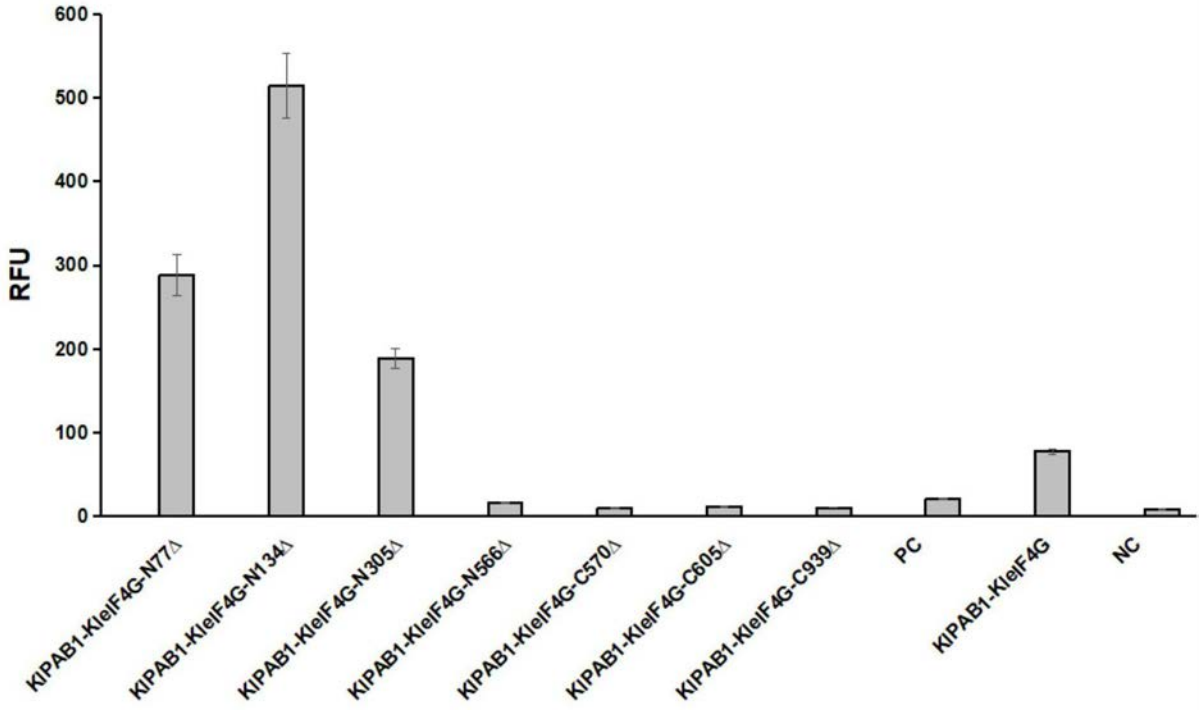

600
500
400
300
200
100
0
RFU
KlPAB1-KleIF4G-N77Δ
KlPAB1-KleIF4G-N134Δ
KlPAB1-KleIF4G-N305Δ
KlPAB1-KleIF4G-N566Δ
KlPAB1-KleIF4G-C570Δ
KlPAB1-KleIF4G-C605Δ
KlPAB1-KleIF4G-C939Δ
PC
KlPAB1-KleIF4G
NC

图8