(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 109073639 A
(43)申请公布日 2018.12.21

(21)申请号 201780018856.3

(22)申请日 2017.02.02

(30)优先权数据
62/290,890 2016.02.03 US

(85)PCT国际申请进入国家阶段日
2018.09.20

(86)PCT国际申请的申请数据
PCT/US2017/016251 2017.02.02

(87)PCT国际申请的公布数据
WO2017/136575 EN 2017.08.10

(71)申请人 CG探索股份有限公司
地址 美国加利福尼亚州

(72)发明人 C·罗

(74)专利代理机构 北京坤瑞律师事务所 11494
代理人 封新琴

(51)Int.Cl.
G01N 33/53(2006.01)
A61K 39/145(2006.01)
A61K 39/155(2006.01)
A61P 31/16(2006.01)
C07K 14/115(2006.01)
C07K 14/285(2006.01)
C12N 15/44(2006.01)

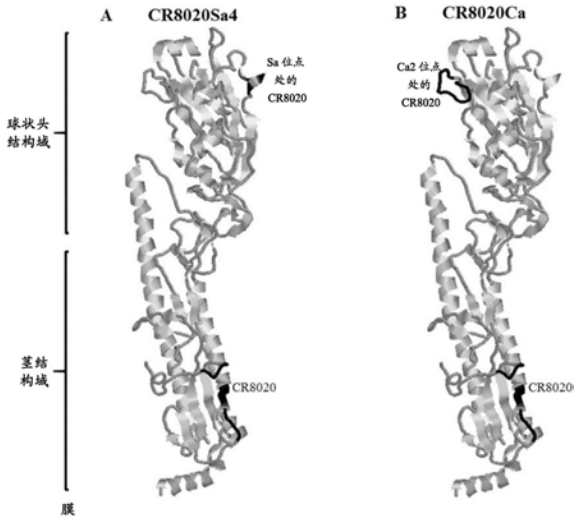
权利要求书2页 说明书52页
序列表76页 附图10页

(54)发明名称

具有异源表位和/或改变的成熟切割位点的流感血凝素的组合物

(57)摘要

公开了修饰形式的血凝素(HA)蛋白包括具有修饰的免疫显性区域和具有修饰的成熟切割位点的那些修饰形式的血凝素蛋白,以及含有它们的病毒和病毒样颗粒。



CN 109073639 A

1.一种甲型或乙型流感的修饰的血凝素(HA)蛋白,其中所述HA蛋白的免疫显性区域通过插入所述区域含有保守的替代流感表位或通过用流感的所述保守的替代表位替换所述区域含有所述表位。

2.权利要求1的修饰的HA蛋白,其中所述修饰的HA蛋白包含在活的或减毒的病毒或病毒样颗粒(VLP)中。

3.权利要求2的修饰的HA蛋白,其包含在活病毒中,并且所述修饰的HA蛋白进一步包含不被动物蛋白酶切割的修饰的成熟切割位点(MCS)。

4.权利要求3的修饰的HA蛋白,其中所述修饰的MCS已被合适的蛋白酶切割。

5.权利要求1-4中任一项的修饰的HA蛋白,其中流感的所述替代表位是来自流感HA蛋白的茎结构域的表位或是M2流感蛋白的表位。

6.权利要求1-4中任一项的修饰的HA蛋白,其中所述HA蛋白的多于一个免疫显性区域已被修饰为含有替代流感表位。

7.权利要求6的修饰的HA蛋白,其中所述多于一个免疫显性区域中的每一个中包含的所述替代流感表位中的至少一些是彼此不同的。

8.一种流感疫苗,其包含权利要求1-4中任一项的修饰的HA蛋白。

9.一种修饰的流感病毒,其中所述病毒包含已被修饰为含有不被动物蛋白酶切割的MCS的流感HA蛋白。

10.权利要求9的修饰的流感病毒,其中所述MCS已被合适的蛋白酶切割。

11.权利要求9的修饰的流感病毒,其中所述HA蛋白在至少一个免疫显性区域中进一步含有异源免疫原性表位。

12.权利要求11的修饰的流感病毒,其中所述异源免疫原性表位是肿瘤相关抗原的表位或除流感病毒之外的病毒的表位或替代流感表位。

13.权利要求11的修饰的流感病毒,其中所述HA蛋白的多于一个免疫显性区域已被修饰为含有异源表位。

14.权利要求13的修饰的流感病毒,其中所述多于一个免疫显性区域中的每一个中包含的所述异源表位中的至少一些是彼此不同的。

15.一种修饰的HA蛋白,其被修饰为含有不被动物蛋白酶切割且进一步含有插入的异源免疫原性表位的MCS,或替换至少一个免疫显性区域的MCS。

16.权利要求15的修饰的HA蛋白,其中所述MCS已被合适的蛋白酶切割,和/或其中所述异源免疫原性表位是肿瘤相关抗原的表位或除流感病毒之外的病毒的表位或替代流感表位。

17.一种产生抗流感抗体的方法,该方法包括向非人受试者给予权利要求8的流感疫苗。

18.用于在保护受试者免于流感的方法中使用的权利要求8的流感疫苗,该方法包括将所述疫苗给予所述受试者。

19.一种产生抗所需抗原的抗体的方法,该方法包括向受试者给予权利要求9-14中任一项的修饰的流感病毒。

20.一种编码流感HA蛋白的至少一个有效部分的核酸,其中编码流感蛋白的替代表位的核苷酸序列插入到或替换编码所述HA蛋白的免疫显性区域的核苷酸序列。

21.权利要求20的核酸,其中编码MCS的核苷酸序列被修饰为编码不被动物蛋白酶切割的MCS。

22.权利要求20的核酸,其中编码所述HA蛋白的免疫显性区域的多于一个区域已被修饰为含有编码替代流感表位的核苷酸序列。

23.权利要求21的核酸,其中编码所述HA蛋白的免疫显性区域的多于一个区域已被修饰为含有编码替代流感表位的核苷酸序列。

24.一种编码流感HA蛋白的至少一个有效部分的核酸,其中编码MCS的核苷酸序列被修饰为编码不被动物蛋白酶切割的MCS。

25.权利要求24的核酸,其中编码异源表位的核苷酸序列被插入到或替换编码所述HA蛋白的免疫显性区域的核苷酸序列。

26.权利要求25的核酸,其中编码所述HA蛋白的免疫显性区域的多于一个区域已被修饰为含有编码异源表位的核苷酸序列。

27.一种在编码一个或多个如权利要求20-26中任一项所述的核苷酸序列的重组宿主细胞中表达的核酸。

28.含有权利要求27的核酸的重组宿主细胞。

29.一种制备权利要求1的修饰的HA蛋白的方法,该方法包括培养包含对流感HA蛋白的至少一个有效部分进行编码的核酸的重组宿主细胞,其中在其中表达所述核酸并回收该修饰的HA蛋白的条件下,编码流感蛋白的替代表位的核苷酸序列被插入到或替换编码所述HA蛋白的免疫显性区域的核苷酸序列。

30.一种制备权利要求3的修饰的HA蛋白的方法,该方法包括培养包含对流感HA蛋白的至少一个有效部分进行编码的核酸的重组宿主细胞,其中在表达所述核酸并回收所述修饰的HA蛋白的条件下,编码流感蛋白的替代表位的核苷酸序列被插入到或替换编码所述HA蛋白的免疫显性区域的核苷酸序列,并且其中编码MCS的核苷酸序列被修饰为编码不被动物蛋白酶切割的MCS。

31.一种制备权利要求9的修饰的流感病毒的方法,该方法包括培养包含对流感HA蛋白的至少一个有效部分进行编码的核酸的重组宿主细胞,其中在表达所述核酸并回收所述修饰的流感病毒的条件下,编码MCS的核苷酸序列被修饰为编码不被动物蛋白酶切割的MCS。

32.一种制备权利要求11的修饰的流感病毒的方法,该方法包括培养包含对流感HA蛋白的至少一个有效部分进行编码的核酸的重组宿主细胞,其中在表达所述核酸并回收所述修饰的流感病毒的条件下,编码MCS的核苷酸序列被修饰为编码不被动物蛋白酶切割的MCS,并且其中编码所述HA蛋白的免疫显性区域的至少一个核苷酸序列已被修饰为含有编码异源表位的核苷酸序列。

33.一种产生抗流感抗体的方法,该方法包括向非人受试者给予权利要求20-27中任一项的核酸。

34.用于在保护受试者免于流感的方法中使用的权利要求20-27中任一项的核酸,该方法包括将所述核酸给予所述受试者。

35.一种产生抗所需抗原的抗体的方法,该方法包括向受试者给予权利要求20-27中任一项的核酸。

具有异源表位和/或改变的成熟切割位点的流感血凝素的组合物

[0001] 相关申请的交叉引用

[0002] 本申请要求2016年2月3日提交的标题为“具有异源表位和/或改变的成熟切割位点的流感血凝素的组合物及其使用方法(Compositions of Influenza Hemagglutinin with Heterologous Epitopes and/or Altered Maturation Cleavage Sites and Methods of Use Thereof)”的美国临时申请号62/290,894的优先权,将该申请的内容通过引用以其全文并入。

[0003] 通过引用序列表并入

[0004] 本申请与电子格式的序列表一起提交。序列表以2017年1月30日创建、大小为173,061字节的名为758252000140SeqList.TXT的文件提供。将在电子格式的序列表中的信息通过引用以其全文并入。

技术领域

[0005] 本发明属于免疫学和病毒学领域。更具体地,本发明涉及具有改变的和异源的表位和/或改变的成熟切割位点的重组流感血凝素(HA)蛋白的组合物。[1]

背景技术

[0006] 流感,通常称为“flu”,是由人或人畜共患流感病毒感染引起的传染性呼吸系统疾病。

[0007] 在每年的寒冷月份(通常称为“流感季节”)都会发生流感爆发。尽管大多数流感感染症状轻微,但这些年度流行病导致全球300-500万例严重疾病和多达500,000人死亡(万维网,who.int/mediacentre/factsheets/fs211/en/,世界卫生组织:流感(季节性))。每个世纪,当流感病毒感染一大部分人群并在世界范围内造成显著的发病率和死亡率时,有几种流感流行病。1918年造成全世界3%至5%的人口死亡的西班牙流感大流行是有史以来最致命的流行病。

[1]以下缩写是适用的。HA,血凝素;NA,神经氨酸酶;NAb,中和抗体;bNAb,广泛中和抗体;TEV,烟草蚀纹病毒;DNA,脱氧核糖核酸;cDNA,互补DNA;RNA,核糖核酸;kb,千碱基;kDa,千道尔顿;CHO细胞,中国仓鼠卵巢细胞;HEK 293细胞,人胚肾293细胞;VLP,病毒样颗粒;BEVS,杆状病毒表达载体系统;AcNPV,苜蓿银纹夜蛾核型多角体病毒;BmNPV,家蚕核型多角体病毒;TIPS,无滴定的感染细胞保存和放大;BIIC,杆状病毒感染的昆虫细胞;PCR,聚合酶链反应;FBS,胎牛血清;PDB,在万维网的蛋白质数据库(rcsb.org/pdb/home/home.do);MOI,感染复数;MCS,成熟切割位点;HPAI,高致病性禽流感;HI,HA抑制;WT,野生型;TIV,三价灭活流感疫苗;PBS,磷酸盐缓冲盐水。

[0008] 引发茎结构域反应性抗体的当前策略

[0009] 抗所有流感病毒的广泛中和抗体(bNAb)的引发是流感疫苗设计的最终目标。大多数分离的bNAb识别HA茎结构域中的保守表位。然而,天然流感病毒感染或当前流感疫苗引发的抗体主要识别HA球状头结构域中受体结合位点周围的抗原位点,这是由于这些抗原位

点的免疫显性。由于流感病毒株中这些抗原位点的高可变性，这些抗体通常是毒株特异性的。因此，目前的流感疫苗不会对流感病毒产生普遍免疫力。

[0010] 血凝素(HA)是完整的膜糖蛋白。它是流感病毒脂质双层包膜中最丰富的蛋白质。估计病毒粒子表面上约有500个HA分子。HA是病毒吸附和渗入宿主细胞所需的唯一蛋白质。它与宿主细胞受体结合，并使病毒粒子脂质包膜与宿主细胞膜之间能够融合。通过HA结合宿主细胞受体的能力，HA决定了流感病毒株的宿主范围。

[0011] HA在内质网中被合成为名为HA0的单链前体(Stevens,J.,等人,Science(2006)312:404-410)。HA0前体在其N-末端具有信号肽，在其C-末端具有膜锚序列。在HA被转运并锚定至宿主细胞膜的过程中，N-末端信号肽被去除。HA组装成内质网中相同亚基的三聚体，然后经由高尔基体网络输出至细胞表面。HA0在成熟切割位点(MCS)的C-末端被特异性宿主胰蛋白酶样蛋白酶切割，并转化为由两个二硫键连接的多肽HA1和HA2组成的成熟形式。HA1是HA0的较大N-末端部分，而HA2是HA0的较小C-末端部分。HA2具有带有短C-末端胞内尾的跨膜区域，并将HA锚定至膜。每种成熟HA的总分子量为约70kDa，其中HA1为约45kDa，且HA2为约25kDa。已经确定了来自许多流感病毒株的HA的细胞外部分的原子结构。成熟HA的结构模型示于图1中。所有HA，无论是HA0还是成熟形式，都具有相同的整体三聚体结构，即在茎上有球状头部。每个HA单体通过具有短的亲水性胞质尾的单个C-末端跨膜区域锚定在膜上。

[0012] 球状头结构域完全由HA1构成，并含有免疫显性表位。茎结构域主要由HA2构成，并含有免疫原性亚显性的保守区域。球状头结构域具有在其核心的八链β片层结构，连同表面环和螺旋。膜近端茎结构域由来自HA1和HA2二者的残基的三重卷曲螺旋结构的左手超螺旋构成。每种HA单体具有多个糖基化位点，所述多个糖基化位点的总碳水化合物的分子量为约13kDa或为总HA分子量的19％。大多数糖基化位于膜表面附近的茎结构域中。还通过胞质尾上的半胱氨酸残基的棕榈酰化来修饰HA(Veit,M.,等人,J.Virol(1991)65:2491-2500)。

[0013] 为了规避球状头结构域的免疫显性，已经基于没有球状头结构域的HA茎结构域(无头HA)设计了免疫原。由于难以产生具有适当三级结构的此类分子，这些尝试在很大程度上是不成功的(Krammer,F.和Palese,P.Curr Opin Virol(2013)3:521-530;Eckert,D.M.,和Kay,M.S.,PNAS(2010)107:13563-13564)。HA2本身已在大肠杆菌中以折叠成其最稳定的融合后低pH诱导的构象的可溶形式表达(Chen,J.,等人,PNAS(1995)92:12205-12209)。通过掺入设计的突变以使HA2的低pH构象不稳定，基于H3 HA的另一种HA2构建体(HA6)在大肠杆菌中表达并重折叠成所需的中性pH预融合构象(Bommakanti,G.,等人,PNAS(2010)107:13701-13706)。HA6在小鼠中具有高度免疫原性并且保护小鼠免受同源甲型流感病毒的感染。然而，来自HA6免疫小鼠的血清无法在体外中和病毒，这可能是由于所述测定仅检测识别球状头结构域的抗体的病毒中和活性的限制。已经描述了另一种“无头”HA茎结构域构建体，其具有HA2和茎结构域中HA1的区域(Steel,J.,等人,mBio(2010)1:e00018-10)。该免疫原保护小鼠免受同源流感攻击，并引发与同一组内的异源HA交叉反应的抗血清。与HA6一样，抗血清在体外不显示中和活性。这些设计都是基于通过消除免疫显性区域的蛋白质最小化。

[0014] 最近，通过对构建体文库进行基于结构的合理设计和反复bNAb选择，制备了稳定的三聚体H1 HA茎纯免疫原(无头迷你HA)(Impagliazzo,A.,等人,Science(2015)349:

1301-1306;Yassing,H.,M.,等人,Nat Med(2015)21:1065-1070;专利申请WO 2014/191435_A1)。这些茎纯迷你HA免疫原引发了抗HA茎结构域的预期抗体。保护用这些迷你HA免疫原免疫的小鼠和雪貂免受高致病性H5病毒的致命攻击。此外,迷你HA免疫原在食蟹猴中引发H5中和抗体。选择这些迷你HA免疫原以使其与特异性bNAb结合。需要针对不同类型的bNAb重复冗长的反复bNAb选择过程。尽管在这些迷你HA免疫原中保留了bNAb表位的构象,但总体结构不同于天然HA的茎结构域。由于这些结构变化,这些无头迷你HA免疫原不太可能组装成流感病毒或流感病毒样颗粒(VLP)。此外,这些无头迷你HA免疫原不具有用于宿主细胞结合的受体结合位点。

[0015] 嵌合HA已被设计用于引导对茎结构域的宿主免疫应答。由流行性H1N1感染诱导的大多数中和抗体与多种流感毒株的HA茎结构域和球状头结构域中的表位广泛交叉反应(Wrammert,J.,等人,J.Exp Med(2011)208:181-193)。采用在人体内不循环的来源于H5N1流感毒株的HA(第1组HA)进行免疫,大大增加了对第1组循环季节性毒株的HA茎特异性应答(Ellebedy,A.H.,等人,PNAS(2014)111:13133-13138;Nachbagauer,R.,等人,J Virol(2014)88:13260-13268;Whittle,J.R.R.,等人,J.Virol(2014)88:4047-4057)。在1957年、1968年和2009年每次出现流行性流感毒株后,现有的季节性病毒株在人群中被新型流行毒株替换。

[0016] 假设暴露于具有不同球状头结构域的新型流行性流感病毒株导致对茎结构域中保守表位的亲和力成熟记忆应答(Palese,P.和Wang,T.T.,mBio(2011)2:e00150-11)。为支持这一假设,H6、H9或H5球状头结构域和H1茎结构域的工程化重组嵌合HA生成了高滴度茎特异性中和抗体(Pica,N.,等人,PNAS(2012)109:2573-2578;Krammer,F.,等人,J.Virol(2013)87:6542-6550)。在这些情况下,H1 HA的球状头结构域被来自大多数人群固有的相同的第1组HA的H6、H9或H5的球状头结构域替换。这些替换是通过用H6、H9或H5的相应序列替代HA1的半胱氨酸52与半胱氨酸277之间的H1 HA序列来进行的。半胱氨酸52和半胱氨酸277在球状头结构域与茎结构域之间的铰链区中形成二硫键。用H3茎结构域制备类似的嵌合HA以开发用于第2组流感毒株的疫苗(Krammer,F.,等人,J.Virol(2014)88:2340-2343;Margine,I.,等人,J.Virol(2013)87:10435-10446)。用这些嵌合HA对动物进行免疫诱导了针对每组病毒具有广泛中和活性的茎特异性抗体。人群对循环H1(第1组)、H3(第2组)和乙型流感病毒株具有预先存在的免疫力。用嵌合HA进行疫苗接种提高了针对循环毒株的HA和嵌合HA常见的茎结构域的抗体水平。针对人类固有的嵌合HA上的新型球状头结构域仅诱导了初级应答。随后使用具有相同的茎结构域但不同的头结构域的第二嵌合HA进行加强,这进一步增加了茎特异性抗体水平。结果表明改变宿主暴露于在球状头结构域中的免疫显性表位可以增加针对茎结构域中免疫亚显性表位的广泛保护性免疫应答。

[0017] A/WSN/33(H1N1)血凝素的抗原性位点B的六氨基酸环可以被来自A/Japan/57(H2N2)和A/Hong Kong/8/68(H3N2)的HA的同源抗原性位点B残基替换(Li,S.,等人,J.Virol(1992)66:399-404)。这些替换不干扰HA的受体结合功能。将具有这些嵌合HA的重组流感病毒复制在MDCK(小猎犬肾上皮细胞)细胞培养物中。具有A/WSN/33(H1N1)和A/Hong Kong/8/68(H3N2)的嵌合HA的病毒诱导抗A/WSN/33(H1N1)和A/Hong Kong/8/68(H3N2)二者的抗体。这些结果表明HA的免疫显性抗原位点可以被来自不同毒株的其他HA的同源相应免疫显性抗原位点替换。这些替换改变了所得嵌合HA的抗原特异性。

[0018] 因此,在前面的段落中,一种流感毒株中的免疫显性表位被来自另一种毒株的同源免疫显性表位替换。

[0019] 另一种策略是抑制免疫显性表位并使宿主免疫应答重新聚焦于茎结构域。可以通过引入用于高糖基化的附加糖基化位点来屏蔽球状头结构域中的免疫显性抗原位点(Eggink,D.,等人,J.Virol(2014)88:699-704,Patent application US 2014/0004149_A1)。球状头结构域中的高糖基化不改变茎反应性抗体的结合亲和力。在用不同的季节性病毒攻击时,用高糖基化HA对小鼠进行免疫诱导了高滴度的茎反应性抗体以及对发病率和死亡率的预防。专利申请US 2013/0315929_A1披露了通过用不太可能作为表位一部分的其他氨基酸替换那些表位的一些残基,来抑制球状头结构域中的免疫显性表位的另一种方法。结果表明,屏蔽球状头结构域中的免疫显性表位可以增加针对茎结构域中免疫亚显性表位的广泛保护性免疫应答。

[0020] HA作为呈递外源表位的载体

[0021] 流感HA已被用作用于HIV-1包膜蛋白的V3环的表位的载体。从HIV-1包膜蛋白gp120到HA免疫显性抗原位点(位点A或位点B)的长度为12至22个残基的免疫显性表位肽的插入,产生了在球状头结构域中具有单个HIV-1表位的嵌合HA。没有去除位点A和位点B的残基。将免疫显性HIV-1表位插入位点中。嵌合HA在动物中诱导对HIV-1V3环表位的免疫应答(Kalyan,N.K.,等人,Vaccine(1994)12:753-760;美国专利US 5,591,823_A;Li,S.,等人,J.Virol(1993)67:6659-6666)。插入的表位的免疫原性似乎被HA增强,因为非常低剂量的嵌合HA蛋白足以诱导对插入的表位特异性的抗体。这些结果表明,可将来自HA以外的蛋白质的免疫显性外源表位插入HA的免疫显性抗原位点。这些具有插入的外源免疫显性表位的嵌合HA分子可以诱导对插入的外源免疫显性表位的免疫应答。

发明内容

[0022] 本发明涉及流感血凝素蛋白的修饰形式,涉及含有它的疫苗、病毒样颗粒和病毒以及用于其生产的重组方法和材料。通常,这些修饰包括用替代表位替换HA蛋白的球状头结构域的免疫显性区域,以产生抗体和/或修饰HA蛋白的茎结构域中的成熟切割位点(MCS)。切割MCS与融合肽之间HA0以产生对于细胞进入所必需的HA2的游离N-末端。通过改变MCS以防止其被动物蛋白酶切割,MCS已被切割的病毒允许感染宿主细胞,其中所述病毒可以繁殖,但其后代是不可感染的。因此,包括占据免疫显性位点的那些表位的任何表位的存在被扩增,而后代病毒不会进一步感染宿主。

[0023] 一方面,本发明涉及一种流感疫苗,其包含修饰的HA或含有它的病毒或病毒样颗粒,其中HA蛋白的免疫显性区域含有插入该区域中的相同流感毒株或另一种流感毒株的保守替代表位。这提供了更成功的免疫原性形式的插入替代表位,其在其天然位置是免疫亚显性的。替代表位是那些非来源于其他HA的球状头结构域的表位。这些替代表位通常来自HA茎结构域或来自非HA流感蛋白,如M2。

[0024] 另一方面,本发明涉及已被修饰为含有不被动物蛋白酶切割的MCS的修饰的流感病毒(其也可用作疫苗)。如上所述,这允许在不产生其感染形式的情况下扩增病毒。通过用替代或异源表位替换一个或多个免疫显性区域可以进一步修饰这种病毒,所述替代或异源表位可以是流感表位或外源表位,包括例如其他病毒、细菌或肿瘤相关抗原特有的表位。

[0025] 在另一些方面，本发明涉及重组材料和制备修饰的HA蛋白的蛋白质、病毒或病毒样颗粒的方法，以及使用这些蛋白质、病毒样颗粒或修饰的病毒产生抗体的方法。

附图说明

[0026] 图1示出了A/California/07/2009(猪源甲型流感病毒H1N1毒株)的成熟HA的细胞外部分的X射线晶体结构(PDB:3LZG)。该结构用于指导本发明的构建体设计。该HA的肽序列用于制备亲本野生型(WT)H1 HA构建体。图1A显示了HA三聚体，其中一个HA单体在带状图，且另外两个在骨架图中。示出了相对于膜的方向、球状头结构域的位置和茎结构域。HA1显示为浅色阴影带，且HA2显示为深色阴影带。图1B显示了具有HA1和HA2位置以及HA1和HA2的N-末端和C-末端的单个HA单体。融合肽被圈出。

[0027] 图2示出了映射至A/California/07/2009的H1 HA单体的H3 HA的免疫显性抗原位点的大概位置。图2A显示了如图1中的带状图。H3 HA免疫显性抗原位点A、B、C、D和E以深色阴影显示。图2B是在膜远端的球状头结构域上方的图2A的顶视图。

[0028] 图3示出了映射至A/California/07/2009的H1 HA单体的H1 HA的免疫显性抗原位点的大概位置。图3A显示了如图1中的带状图。H1 HA免疫显性抗原位点Ca1、Ca2、Cb、Sa和Sb以深色阴影显示。图3B是在膜远端的球状头结构域上方的图3A的顶视图。许多这些免疫显性抗原位点围绕如箭头所示的宿主细胞受体结合位点。

[0029] 图4示出了H1 HA构建体的示意设计。图4A是构建体的核酸特征的示意图。标记了用于亚克隆的限制性位点的位置。图4B是与图4A中所示核酸的相对比例的构建体的蛋白质特征的示意图。GP67ss代表N-末端的GP67分泌信号(信号肽)。HA1代表HA HA1。MCS代表HA成熟切割位点，其是HA1的一部分。HA2代表HA HA2。TEV代表TEV切割位点。Foldon代表foldon序列。His代表C-末端的10-组氨酸-标签(10xHis-标签)。

[0030] 图5示出了异源表位置于HA球状头结构域中的位置，以及改变的成熟切割位点所在的位置。亲本构建体肽序列(WT)与H3(A/Aichi/2/1968H3N2)和H1(A/Puerto Rico/8/1934 H1N1)的HA序列进行比对。每个序列的氨基酸位置在右侧基于HA0编号进行标记，其中去除信号序列后的第一个残基作为位置1。通过箭头和垂直线指示成熟切割位点(MCS)。H3 HA的免疫显性抗原位点A、B、C、D和E标记在H3序列上方。H1 HA的免疫显性抗原位点Ca1、Ca2、Cb、Sa和Sb标记在H1序列下方。WT中的下划线序列被序列下方所示的异源表位替代。改变的成熟切割位点示于成熟切割位点(MCS)下方。广泛中和抗体(bNAb)、FI6、1C9和CR8020的天然免疫亚显性抗原位点由序列上方的线分别标记。

[0031] 图6示出了HA单体的模型，其中复合CR8020表位位于球状头结构域的免疫显性抗原位点。HA单体示于如图1所示的带状图。被复合CR8020表位替换的免疫显性抗原位点以深色阴影显示。茎结构域中的原始天然CR8020表位也以深色阴影显示。图6A显示了CR8020Sa4的模型。图6B显示了CR8020Ca的模型。

[0032] 图7显示了本发明中描述的HA构建体的表达数据。通过Ni-NTA树脂(Ni下拉)收获和捕获用所示构建体的重组杆状病毒感染的SF9细胞的细胞培养基。用1xPBS洗涤Ni-NTA树脂以去除未结合的蛋白质。通过用Coomassie®染色的SDS-PAGE(图7A、图7B、图7C和图7D)或抗His Western印迹(图7E)分析与Ni-NTA树脂结合的蛋白质。箭头指示全长HA0。

[0033] 图8显示了用于本发明的纯化的HA构建体。通过Ni-NTA树脂(Ni下拉)收获和捕获

用所示构建体的重组杆状病毒感染的SF9细胞的细胞培养基。用1xPBS洗涤Ni-NTA树脂。用咪唑洗脱结合的蛋白质，并通过用Coomassie®染色的SDS-PAGE进行分析。箭头指示全长HA0。

具体实施方式

[0034] 由于本发明涉及HA蛋白的修饰，因此该蛋白质及其功能的进一步描述以及以上提供的蛋白质可能是有帮助的。

[0035] HA0的蛋白酶切割是甲型流感病毒传染性的先决条件。这些蛋白酶在宿主中的分布是组织嗜性的决定因素之一，并因此具有致病性。已在呼吸道和肺中鉴定出几种胰蛋白酶样蛋白酶，其切割大多数具有单碱基成熟切割位点的HA(Kido,H.,等人,J.Biol Chem(1992)267:13573-13579;Peitsch,C.,等人,J.Virol(2014)88:282-291;Zhirnov,O.,P.,等人,J.Virol(2002)76:8682-8689)。这些蛋白酶是丝氨酸蛋白酶。其中一种是胰蛋白酶样蛋白酶Clara，其首次从大鼠细支气管上皮Clara细胞中分离出来。切割HA0的这些蛋白酶的狭窄组织分布将流感病毒感染限制在哺乳动物的呼吸道和肺部。

[0036] 与HA成熟有关的蛋白酶尚未得到充分表征。与胰蛋白酶一样，这些蛋白酶将C-末端的肽键切割成碱基残基，如精氨酸(R)或赖氨酸(K)。成熟切割位点(MCS)位于HA1的C-末端，并且在MCS的C-末端的切割对于传染性是必需的。

[0037] H5和H7亚型的成熟切割位点的多碱基序列导致更广泛的细胞蛋白酶如弗林蛋白酶和枯草杆菌蛋白酶类蛋白酶的切割易感性，这与这些病毒在哺乳动物中的更广泛的组织嗜性和更高的致病性相关(Stieneke-Gröber,A.,等人,EMBO J(1992)11:2407-2414;Maines,T.R.,等人,J.Virol(2005)79:11788-11800)。许多高致病性禽流感(HPAI)亚型是H5和H7。在人类中的HPAI爆发期间，报告的病例死亡率高于流行性和季节性流感病毒的病例死亡率(Morens,D.M.,等人,Nature(2012)486:335-340)。MCS中的这些多碱基序列还导致禽类物种的多个器官中的病毒复制，从而导致这些禽类物种的高死亡率。

[0038] 季节性流感病毒和非致病性禽流感病毒的HA在细胞外被呼吸道和肺中限制其组织嗜性的的特定蛋白酶切割。另一方面，高致病性病毒的HA在细胞内被普遍存在的蛋白酶切割。这些高致病性病毒在各种组织中经历多周期复制以引起全身性感染(Steinhauer,D.A.,Virus.Virol(1999)258:1-20;Taubenberger,J.K.,PNAS(1998)95:9713-9715)。1918流行性人流感病毒毒株也利用广泛范围的细胞蛋白酶，并使用其神经氨酸酶(NA)招募纤溶酶原进行HA切割(Goto,H.and Kawaoka,Y.PNAS(1998)95:10224-10228;Chaipan,C.,等人,J.Virol(2009)83:3200-3211)。

[0039] HA0的切割产生具有对于病毒粒子与宿主细胞的融合所必需的新的游离N-末端的HA2。切割形式是HA的成熟形式，其具有受体结合和膜融合的全部功能。HA0具有结合受体但不介导膜融合的能力。HA的前体和成熟形式均存在于病毒粒子的表面。仅具有HA0的病毒没有融合活性，并且不会引起感染。

[0040] HA2的N末端十二个残基称为“融合肽”(Skehel,等人,Biochem Soc Trans(2001)29:623-626)。融合肽具有疏水序列。在HA0中，MCS和融合肽形成表面环。切割后，新生成的N-末端融合肽插入HA三聚体界面中。与宿主细胞受体结合后，附着的流感病毒粒子被宿主细胞内吞至核内体。成熟HA的融合潜能在核内体pH(依赖于特定的流感病毒株，在pH5至6之

间)下被激活。在低pH下HA结构的广泛变化导致融合肽向细胞膜的挤压。将融合肽插入宿主核内体膜中导致周围的宿主核内体膜与含有HA的C-末端膜锚区域的病毒膜融合。融合将病毒RNA区段释放至宿主细胞的细胞质中以进入发生病毒复制的宿主细胞核。

[0041] HA决定流感病毒株的宿主范围。HA与宿主细胞表面糖蛋白和糖脂的末端唾液酸结合。来自人流感毒株的HA几乎仅与由SAα2,6Gal终止的唾液酸寡糖结合,而来自禽流感和马流感毒株的HA结合SAα2,3Gal。人类主要具有SAα2,6Gal唾液酸寡糖,而禽类物种主要具有SAα2,3Gal唾液酸寡糖。如图3中示意性所示,HA上的宿主受体结合位点位于仅由HA1构成的球状头结构域中。流感病毒HA的受体特异性已被充分表征。已经映射了与SAα2,6Gal或SAα2,3Gal的识别有关的氨基酸残基。已显示受体结合口袋中的单个氨基酸残基取代改变了受体结合特异性(Rogers,G.N.,等人,Nature(1983)304:76-78)。1918流行性流感病毒的不同分离株由于HA的受体结合位点中的单个氨基酸取代而具有不同的受体结合特异性(Glaser,L.,等人,J.Virol(2005)79:11533-11536)。毒株A/South Carolina/1/18 HA优先结合人细胞受体SAα2,6Gal,而毒株A/New York/1/18 HA结合人细胞受体SAα2,6Gal和禽细胞受体SAα2,3Gal,从而揭示了宿主适应流感病毒的动态性质。

[0042] 尽管HA的受体结合位点中的单个突变使受体结合特异性从禽类改变为人类,但是人类中H5禽流感毒株的有效传播需要受体结合位点外的HA的其他几种变化以及其他流感蛋白的变化(Imai,M.,等人,Nature(2012)486:420-428;Herfst,S.,等人,Science(2012)336:1534-1541;Chen,L-M.,等人,Virol(2012)422:105-113;Russell,C.A.,等人,Science(2012)336:1541-1547)。连同其他变化,在H5 HA中仅四个氨基酸取代就足以使突变禽H5病毒能够通过飞沫在雪貂中传播。除了了解跨物种的流感病毒传播外,这些研究还建立了实验室程序来研究哺乳动物中非哺乳动物流感病毒的传播,这可用于评估通用流感疫苗候选物。

[0043] 抗流感抗体和HA的抗原位点的描述

[0044] 作为流感病毒脂质双层包膜中最丰富的蛋白质,HA是流感病毒的主要抗原,并且具有用于中和抗体的一级表位。大多数的人抗流感抗体是针对HA和NA的。通过疫苗接种引发的百分之六十(60%)的流感反应性抗体对HA起反应(Wrammert,J.,等人,Nature(2008)453:667-671)。大多数HA反应抗体识别球状头结构域中受体结合位点周围的抗原位点。这些抗体中的一些是病毒中和抗体。这些中和抗体(NAb)通常干扰HA与宿主细胞的结合并显示HA抑制(HI)活性。由于这些抗原位点的高可变性,这些中和抗体通常是毒株特异性的,因此缺乏所需的广泛中和活性(Wang,T.T.和Palese,P.,Nat Struct Mol Biol(2009)16:233-234)。

[0045] 流感病毒的广泛感染是病毒改变其抗原特性的能力的结果。流感病毒的抗原特性的变化是通过易错的病毒RNA依赖性RNA聚合酶复合物低保真度复制病毒基因组的结果。高突变率导致抗原漂移,即HA的抗原特性的逐渐变化。抗原漂移发生在所有类型的流感病毒中。HA基因的序列分析表明,尽管沉默的核酸序列变化遍布整个HA基因,但氨基酸序列的大部分变化位于HA1中(Palese,P.和Young,J.F.,Science(1982)215:1468-1474)。HA2比HA1更保守。

[0046] 当两种不同毒株的病毒同时共感染宿主时,病毒基因组的分段性质也导致病毒基因组区段的重配。新的病毒毒株从这种采用不同病毒表面蛋白的重配中形成,所述重配导

致抗原位移，即HA的抗原特性突然变化。抗原位移仅发生在甲型流感病毒中。抗原漂移和抗原位移使流感病毒能够逃脱现有抗体的中和作用。

[0047] H3流感

[0048] 35年前测定的流感HA的第一个详细结构阐明了HA的抗体结合位点，并提供了抗原漂移和抗原位移的分子解释(Wilson,I.A.,等人,Nature(1981)289:366-373;Wiley,D.C.,等人,Nature(1981)289:373-378;Wiley,D.C.和Skehel,J.J.,Ann Rev Biochem(1987)6:365-394)。

[0049] 抗原不同的病毒的HA序列的比较鉴定了H3 HA(第2组亚型HA)的球状头结构域中的免疫显性抗原位点。H3 HA的免疫显性抗原位点的描述由其他参考文献证实(Both,G.W.,等人,J.Virol(1983)48:52-60)。HA结构中这些抗原位点的大概位置示意性地示于图2中。这些抗原位点的序列位置如图5所示。位点A位于A/Aichi/2/1968 H3N2毒株的HA的残基133至148的表面环中(在图5中编号为H3)。该环称为140-环，从球状头结构域伸出，并位于受体结合口袋的下缘。位点B位于球状头结构域的顶部，并且包含残基187至196的表面α螺旋和沿着受体结合口袋的上缘的残基155至160的相邻表面环。位点C围绕Cys52与Cys277之间的二硫键。以Cys52为中心的环(残基46至55)与以Cys277为中心的环(残基271至280)的交叉在球状头结构域与茎结构域之间的铰链处形成凸起。位点D位于HA三聚体的HA单体亚基之间的界面区域中。它以球状头结构域核心中的八链β片层结构中的残基200至214的两个β链为中心。在其他β链的转角处的残基也可以是位点D的一部分。位点D主要埋在HA三聚体界面中。目前尚不清楚位点D如何作为抗原位点发挥作用。位点E位于球状头结构域侧面的位点A与位点C之间，并且由残基62至63、残基78至83的表面环和八链β片层的边缘上的残基119至122的β链构成。位点E的这些残基在球状头结构域侧面共同形成连续表面。比较1968年至2003年期间出现的流感病毒的氨基酸取代和抗原特性，表明了与主要抗原变化有关的取代仅位于位点A和位点B中(Smith,D.J.,等人,Science(2004)305:371-376;Koel,B.F.,等人,Science(2013)342:976-979)。位点C、D和E中的取代显现引起了小的抗原变化。结果表明，大多数毒株特异性中和抗体与球状头结构域的受体结合位点外围的位点A和位点B结合。

[0050] H1流感

[0051] 病毒毒株A/PR/8/34的抗原不同的第1组H1 HA的遗传分析鉴定了球状头结构域中的不同抗原位点，其称为Ca1、Ca2、Cb、Sa和Sb位点(Caton,A.J.,等人,Cell(1982)31:417-427;Gerhard,W.,等人,Nature(1981)290:713-717)。HA结构中这些抗原位点的大概位置示意性地示于图3中。这些抗原位点的序列位置如图5所示。Ca1位点位于八链β片层结构的β链的转角处。残基165至169的转角之一(在图5中编号为H1)是表面暴露的。Ca1位点的残基207位于连接对应于H3 HA的位点D的两个β链的转角处。H1 HA的Ca1位点通常对应于H3 HA的位点D。Ca2位点也由两个区段构成，这些区段在一级结构中分开但在三级结构中在一起。Ca2的一个区段位于H3 HA的位点A的相应位置处的残基136至141的表面环中。Ca2的另一个区段由长表面环上的残基220至221构成。在HA三聚体中Ca2位点位于HA单体的球状头结构域中的Ca1的对侧，但与另一个HA单体的Ca1位点相邻。在三聚体结构中一个HA单体的Ca1位点与另一个HA单体的Ca2位点形成连续表面。Cb位点是形成紧邻八链β片层结构的表面环的残基70至75的线性表位。它对应于H3 HA的位点E。位点Sa和Sb可以被认为是对应于H3 HA的位点B的亚位点。残基154至163的位点Sa与对应于H3 HA的位点B环的环重叠。Sa位点的另一个

区段是残基124至125的附近转角。位点Sb对应于H3 HA的位点Bα螺旋，并且大多数具有α螺旋残基188至194。

[0052] 广泛中和抗体和保守表位

[0053] 天然流感病毒感染或用三价灭活流感疫苗(TIV)接种也会引发低水平的抗保守HA表位的抗体，这些表位主要位于膜近端茎结构域(Ellebedy,A.H.,等人,PNAS(2014)111:13133-13138)。这些抗体识别许多流感病毒毒株中保守的表位。提供针对若干毒株的保护的那些抗体称为广泛中和抗体(bNAb)。这些bNAb通常不具有HA抑制活性，并且不阻止HA与宿主细胞表面受体结合。第一种交叉毒株抗体C179已在20年前得到表征(Okuno,Y.,等人,J.Virol(1993)67:2552-2558)。它识别HA1的残基318至322(TGLRN)和HA2的残基47至58(GITNKVNSVIEK)的构象表位，所述残基在H1和H2毒株中是保守的。C179抑制HA的融合活性，从而导致病毒中和。

[0054] 已经从感染流感病毒的患者中检测或分离出许多bNAb(Ekiert,D.C.,和Wilson,I.A.,Curr Opin Virol(2012)2:134-141)。已经证明bNAb可中和第1组、第2组或第1组和第2组两组的甲型流感病毒。已经鉴定和表征了由bNAb识别的许多抗原表位。这些表位是HA序列的线性区段或具有HA序列的多个线性区段的构象表位。bNAb的许多表位位于HA蛋白的变化较小的茎结构域中。这些表位包括但不限于融合肽和MCS周围的肽序列。

[0055] 许多bNAb及其表位已经在结构上得以表征。例如，bNAb FI6识别HA0中的MCS和融合肽的残基(Corti,D.,等人,Science(2011)333:850-856)。HA的肽图谱将两种肽鉴定为FI6表位，即由MCS和大部分融合肽组成的RKKRGLFGAIAGFIE以及HA2中的螺旋卷曲螺旋肽的KESTQKAIDGVTNKVNS组成。提出的FI6中和机制是通过阻断蛋白酶进入HA0的MCS来抑制膜融合以及防止HA成熟。另一种bNAb F10识别成熟形式的HA中融合肽周围的构象表位(Su,J.,等人,Nat Struct Mol Biol(2009)16:265-273)。F10可能通过阻止膜融合来抑制所有第1组甲型流感病毒。

[0056] 单克隆bNAb 1C9在体外抑制细胞融合(US Patent 8,540,995;Prabhu,N.,等人,J.Virol(2009)83:2553-2562)。1C9识别GLFGAIAGF的线性表位，即HPAI H5N1的H5 HA2的融合肽的N-末端(免疫表位数据库网址:iedb.org/assay details.php?assayId＝1599077)。1C9显示保护小鼠免受高致病性禽流感(HPAI)H5N1病毒感染。还从记忆B细胞分离出单克隆bNAb(Hu,W.,等人,Virol(2013)435:320-328)。这些单克隆抗体中的一些识别来自2009流行性H1N1流感病毒的HA2的FIEGGWTGMVDGWYGYHH的线性表位。该表位是融合肽的1C9表位的C-末端。HA融合肽的14个残基的序列在甲型和乙型流感病毒中高度保守。

[0057] 已经探究了HA融合肽的保守性质以开发通用流感疫苗。基于乙型流感病毒的HA0的高度保守序列(包括在断裂键两侧含有HA2的MCS和融合肽的HA1的最后9个氨基酸残基)，肽缀合疫苗引发了针对用乙型流感病毒的抗原不同的谱系的病毒毒株进行致命攻击的保护性免疫应答(Bianchi,E.,等人,J.Virol(2005)79:7380-7388)。

[0058] 单克隆bNAb CR6261识别茎结构域中HA2螺旋A和HA1残基中的高度保守区域(Ekiert,D.C.,等人,Science(2009)324:246-251)。CR6261通过阻止HA转化为融合后构象来中和第1组流感病毒。CR6261属于使用相同V_H1-69种系抗体重链的bNAb。另一种V_H1-69单克隆抗体CR8020中和第2组流感病毒。CR8020在与膜紧密接近(约15至20 Å)的茎结构域的基部结合HA，类似于识别来自HIV gp41亚基的膜近端外部区域(MPER)的那些抗HIV的抗体

(Ekiert,D.C.,等人,Science(2011)333:843-850)。CR8020表位的两个主要组分由融合肽的C-末端部分(HA2残基15至19,H3的EGMID)和茎结构域的基部附近的5链β片层的最外链(HA2残基30至36,H3的EGTGQAA)组成。这两个组分在H3的一级结构中相隔10个残基。与HA茎结构域结合的大多数bNAb中和甲型流感病毒的HA的第1组(H1、H2、H5、H6、H8、H9、H11、H12、H13和H16)或第2组(H3、H4、H7、H10、H14和H15)。这些抗体不抑制HA与宿主细胞的结合,但可阻止病毒膜和宿主细胞膜的融合。

[0059] HA受体结合位点是球状头结构域顶部的口袋(Wilson,I.A.,等人,Nature(1981)289:366-373;Wiley,D.C.和Skehel,J.J.,Ann Rev Biochem(1987)56:365-394)。该口袋由许多流感毒株中高度保守的氨基酸残基形成。口袋的边缘由免疫显性抗原位点(如以上所述的H3中的位点A和位点B)形成。来自健康人类受试者的人单克隆抗体(mAb)的克隆鉴定了bNAb,该bNAb识别与来自H1、H2和H3毒株的HA的球状头结构域中的受体结合位点紧密接近的保守残基(Krause,J.C.,等人,J.Virol(2011)85:10905-10908;Krause,J.C.,等人,J.Virol(2012)86:6334-6340)。结构研究显示,这些bNAb中的至少一些模拟唾液酸与受体结合口袋的相互作用(Whittle,J.R.R.,等人,PNAS(2011)108:14216-14221;Ekiert,D.C.,等人,Nature(2012)489:526-532)。

[0060] 茎反应性抗体在自然感染中很少见,并且在目前季节性流感疫苗的免疫中甚至更少。这些抗体中只有一部分是中和抗体。由于茎结构域的保守性质,大多数中和抗体是bNAb。从由人类受试者制备的抗体文库进行克隆目前可以常规地鉴定出这些罕见的广泛中和的茎反应性抗体(Kashyap,A.K.,等人,PNAS(2008)105:5986-5991;Wrammert,J.,等人,Nature(2008)453:667-671)。这些茎反应性抗体的罕见出现导致了如下假设,即茎结构域中的这些表位是免疫显性的,而球状头结构域中的表位是大多数抗体所针对的免疫显性(Krammer,F.和Palese,P.,Nature Rev Drug Disc(2015)14:167-182)。反复暴露于流行的季节性流感病毒或用季节性流感疫苗进行免疫导致了抗球状头结构域中的免疫显性抗原位点的抗体的产生。已经表明,球状头结构域中免疫显性表位的存在可能使二次应答偏离茎结构域(Russell,C.J.,N Engl J Med(2011)365:1541-1542)。已感染当前毒株或针对其接种疫苗的个体可能比那些天然免疫的个体更难以获得普遍应答。

[0061] M2蛋白及其表位

[0062] M2是在病毒包膜中形成同源四聚体质子通道的单程跨膜蛋白(Lamb.R.A.,等人,Cell(1985)40:627-633;Pielak,R.M.和Chou,J.J.,Biochim Biophys Acta.(2011)1808:522-529)。与病毒包膜中HA相比,M2存在的丰度要低得多(M2:HA的比例为1:10至1:100)。M2质子通道功能对于调节病毒内部的pH以将病毒蛋白释放至宿主细胞的细胞溶质中以及调节高尔基体腔的pH以将HA转运至细胞表面是重要的。甲型流感病毒M2(AM2)蛋白具有97个残基,其中有被称为M2e的残基1至23的细胞外N-末端结构域、残基24至46的跨膜(TM)结构域以及残基47至97的细胞内C-末端结构域。来自四个M2分子的TM结构域形成四螺旋束,它起到pH敏感性质子通道的作用,以在细胞进入期间调节跨病毒膜的pH,以及在病毒组装和退出期间调节感染细胞的跨反式高尔基体膜的pH。大的细胞质结构域对于稳定的四聚体形成至关重要,并且通过其与病毒粒子内壳的M1蛋白的结合在病毒组装中发挥作用。除了TM结构域中对于质子通道功能必需的HXXXW序列基序外,甲型、乙型和丙型流感病毒的M2蛋白几乎没有序列同源性。然而,AM2的10个N-末端细胞外残基在所有甲型流感病毒中都是保守

的。

[0063] 甲型流感病毒M2(AM2)是阻断AM2质子通道的抗病毒药物金刚烷胺和金刚乙胺的靶标。使质子通道闭合状态稳定的药物金刚乙胺与通道结构域C-末端附近的面向脂质的口袋结合。金刚烷胺是具有抗甲型流感病毒的活性,而不具有抗乙型流感病毒的活性的抗病毒药物。由于AM2的通道结构域突变所引起的广泛耐药性,已经停止使用这些通道阻断剂。这些突变中的许多产生了比野生型(WT)病毒传染性更小的稍微减毒的病毒。这些耐药的突变体可在没有药物选择压力的情况下恢复为WT。

[0064] M2蛋白是在感染流感病毒的宿主细胞表面上大量表达的完整膜蛋白(Lamb,R.A.,等人,Cell(1985)40:627-633)。表明了M2是对流感病毒反应的细胞毒性T淋巴细胞(CTL)的细胞表面抗原。甲型流感病毒的感染仅引发抗M2的低滴度抗体(Feng,J.,等人,Virol J.(2006)3:102-115)。M2e的高度结构保守可能部分由于不良的M2e特异性抗体应答,因此没有压力变化。抗M2抗体应答在具有M2蛋白的预先存在的抗体的个体中更稳健(Zhong,W.,等人,J.Infect Dis(2014)209:986-994)。由于感染2009流行性甲型H1N1流感病毒而诱导的抗M2抗体与季节性甲型流感病毒的M2蛋白交叉反应。用抗M2e抗体治疗小鼠显著延迟了疾病的进展并导致M2e逃逸突变体的分离,这表明了使用M2e作为针对甲型流感病毒感染的疫苗的潜力(Zharikova,D.,等人,J.Virol(2005)79:6644-6654)。

[0065] 已经描述了基于M2e与HA融合的DNA疫苗(Park,K.S.,等人,Vaccine(2011)29:5481-5487)。该融合蛋白具有一个人M2e肽和一个禽M2e肽,它们通过彼此之间的20个残基的接头位于HA蛋白的N-末端。证实了编码的M2e-HA融合蛋白的表达。用M2e-HA融合DNA疫苗免疫的小鼠显示了对M2e的增强的T细胞应答和完全免受异源禽流感病毒的致命攻击。

[0066] 重组融合蛋白包含与结核分枝杆菌HSP70(mHSP70)蛋白的C-末端基因融合的M2e的四个串联重复序列,显示了免受小鼠中多种流感病毒毒株(Ebrahimi,S.M.,等人,Virology(2012)430:63-72)。M2e肽SLLTEVETPIRNEWGCRCNDSSD已与阳离子脂质体递送运载体以及来自其他流感蛋白的肽和BM2缀合,所述BM2是作为疫苗的乙型流感病毒的M2e的同源物(专利申请US 2010/086584_A1)。该疫苗在小鼠中引发针对M2e的免疫应答。用含有M2e肽的疫苗接种的小鼠免受致命的流感病毒攻击。补充有M2e VLP的灭活流感疫苗在小鼠中赋予了针对抗原不同的甲型流感病毒的增强的且持久的交叉保护(Song,J-M.,等人,PNAS(2011)108:757-761)。

[0067] 免疫显性区域的修饰

[0068] 本发明包括基因修饰的重组流感血凝素(HA)基因和蛋白质,其通过用替代表位替换HA的球状头结构域的免疫显性区域或在其中插入替代表位以引导对这些表位的宿主免疫应答来产生。这些表位中的一些被bNAb识别为HA。本发明还涉及具有修饰的MCS的修饰的HA蛋白和基因以及这些修饰的组合。

[0069] 在一个实施方案中,修饰的重组HA具有被M2蛋白(M2e)的细胞外结构域替换的球状头结构域的免疫显性抗原位点,或M2e被插入免疫显性抗原位点中。

[0070] 在HA蛋白的三维结构的指导下,免疫显性区域的特定表面肽或这种表面肽被相同HA的其他区域的一个异源肽或几个异源肽、或来自流感病毒的另一种亚型或毒株的HA的一个异源肽或几个异源肽替换。这些异源肽可以插入免疫显性区域中。此外,HA的特定表面肽或多个肽被来自与HA无关的其他蛋白质的异源肽替换,或者将这种肽插入其中。这些异源

肽是天然蛋白质或人工设计的。目前已知的抗体可能无法识别它们。

[0071] HA的免疫显性区域位于HA的球状头结构域的表面上。这些待修饰的表面区域是表面暴露的螺旋、β链或环。这些表面区域优选为免疫显性抗原位点和表位，或这种抗原位点和表位的部分，或紧邻这种抗原位点和表位。

[0072] 在一些实施方案中，HA球状头结构域的免疫显性区域被HA的免疫亚显性表位(被bNAb识别为HA)的肽替换。在其他实施方案中，HA球状头结构域的免疫显性区域被免疫亚显性表位的肽替换，所述免疫亚显性表位被能中和流感病毒组的抗体识别。这些免疫亚显性表位选自HA茎结构域，或包括HA融合肽或HA成熟切割位点。

[0073] 本发明的重组修饰的HA蛋白在细胞培养物中表达，并以类似于重组野生型HA的水平分泌至细胞培养基中。这些设计适用于所有甲型、乙型和丙型流感病毒毒株的任何流感HA，包括但不限于感染人类的流感病毒、以及感染其他哺乳动物和禽类物种的流感病毒。在一个实施方案中，这些修饰的HA蛋白可用作免疫原以制备针对人或动物中的流感病毒感染的疫苗。

[0074] 可以通过定点诱变来改变这些免疫显性区域。通过HA蛋白的三维结构鉴定抗原位点的溶剂暴露的残基。抗原位点的一个溶剂暴露的残基或几个溶剂暴露的残基通过定点诱变来改变，或用含有这些残基的特定变化的肽替换。新位点具有与原始位点相同的二级结构。

[0075] 抗原位点被异源表位替换的示例性HA

[0076] 例如，基于如图3所示的HA结构，选择HA中的几个位置来说明通过来自HA的茎结构域的异源肽替换球状头结构域中的免疫显性抗原位点的可行性。

[0077] 残基119至122的肽KTSS(在图5中编号为WT)是Sa位点残基124至125附近的表面暴露的螺旋结构。Sa位点位于HA三聚体的球状头结构域的侧面并且远离受体结合位点。在该位点的修饰不太可能改变受体结合位点并因此不太可能改变通过修饰的HA的宿主受体结合。四残基肽KTSS被茎结构域的bNAb表位替换。在对应于H3 HA的位点A的Ca2位点处的残基137-142的HAGAKS是主要的免疫显性位点。残基153至164处的KKGNSYPKLSKS是在对应于H3 HA的位点B的Sa位点处的表面环。在一些构建体中，该肽在残基153至157处的KKGNS的N-末端部分被替换。在Sb位点的残基184至195处的TSADQQSLYQNA在HA球状头结构域的顶部形成螺旋。

[0078] Sa位点的表面环和Sb位点的螺旋在HA的球状头结构域中彼此相邻，并且对应于H3的位点B。Sa位点和Sb位点可共同容纳具有环和螺旋结构的构象表位。然而，由于这些位点靠近受体结合位点，因此这些肽的替换可能会破坏受体结合位点。在Ca2位点的残基213至224处的EIAIRPKVRDQE是两个HA单体之间界面附近的环。该环与对应于H3 HA的位点D的相邻HA单体的β链接触。环的一部分是表面暴露的。Ca2位点位于该肽内。在一些构建体中，该肽被部分替换。

[0079] 此外，异源肽可置于如图5所示的任何其他抗原位点以及如图1至图3所示的表面环或螺旋中或附近。此外，异源肽可以插入没有抗原位点残基的任何缺失的任何这些位置中。

[0080] bNAb CR8020表位位于茎结构域。如图5和图6所示，CR8020表位具有两个主要组分，其由茎结构域的基部附近的5链β片层的最外β链(HA2残基30至36，H3 HA的EGTGQAA，SEQ

ID NO:26)和融合肽的C-末端部分(HA2残基15中至19,H3 HA的EGMID,SEQ ID NO:25)组成。这两个组分在一级结构中被10个氨基酸残基分隔开。设计复合CR8020表位EGMIDYEGTGQAA(SEQ ID NO:27),其中所述两个表位组分通过酪氨酸(Y)连接。该复合CR8020表位用作异源肽以替换免疫显性位点。

[0081] 另一种bNAb 1C9表位是茎结构域中存在的HA2的N-末端中的融合肽。将修饰的1C9表位肽GIFGAIAGFIEG(SEQ ID NO:36)设计为在免疫显性位点展示的肽。该修饰的1C9肽被表示为I1C9,它在由bNAb 1C9识别的H5融合肽的位置2具有亮氨酸(L)的异亮氨酸(I)取代。这种I1C9融合肽存在于一些H6和H9病毒的H1 HA和HA中。将该I1C9肽置于不同的免疫显性位点。已采用I1C9将多个构建体制备成位置改变的抗原位点。

[0082] bNAb FI6表位是含有两个肽的构象表位,所述肽在HA前体HA0的三级结构上形成连续表面,但在一级结构中分开(图5)。一个FI6表位肽RKKRGLFGAIAGFIE是HA0的成熟切割位点和融合肽。另一种FI6肽KESTQKAIDGVTNKVNS在HA2中具有卷曲螺旋结构。已经制备了构建体,其中将这两种FI6表位肽置于球状头结构域中。FI6表位肽RKKRGLFGAIAGFIE被置于残基153至164的Sa位点,且卷曲螺旋FI6表位肽KESTQKAIDGVTNKVNS被置于残基184至195的Sb位点。在另一个构建体中,FI6表位肽RKKRGLFGAIAGFIE被置于没有第二个FI6表位肽的Sa位点。

[0083] 已经表征了许多bNAb表位。如上所述,它们中的任何一种都可以置于球状头结构域中的免疫显性抗原位点或表面环处或附近。

[0084] 在其他实施方案中,异源肽是M2蛋白的细胞外结构域的肽(M2e肽)。M2e肽在甲型流感病毒中是保守的并且变化较小。虽然来自甲型流感病毒和乙型流感病毒的M2e肽不同,但M2e肽在乙型流感病毒中也是保守的并且变化较小。

[0085] 在一些实施方案中,异源肽是人工设计的肽,其通过定点诱变对免疫显性抗原位点进行特定改变。在一些实施方案中,人工肽结合了免疫显性抗原位点和异源肽的期望特征。在一些实施方案中,人工肽具有与bNAb相互作用的残基和原始免疫显性抗原位点的残基,以维持HA球状头结构域的三维结构。基于HA三维结构或通过筛选随机产生的肽库合理设计这些人工肽。

[0086] 已经制备构建体以将M2e肽SLLTEVETPTRNGWECKCSDS置于Ca2位点或Sb位点,以使用杆状病毒表达系统或哺乳动物表达系统进行表达。可基于来自不同流感毒株的序列的共有序列进一步优化M2e肽。

[0087] 具有修饰的成熟切割位点的HA

[0088] 本发明包括如下实施方案,其中改变HA成熟切割位点(MCS)的蛋白酶易感性以使得所得的HA对不同种类的蛋白酶敏感,并且对那些切割所有毒株的天然HA成熟切割位点的那些胰蛋白酶样蛋白酶具有抗性。改变的MCS被设计为由流感病毒的已知天然宿主中不存在的特定蛋白酶识别。这使得所得的HA在流感病毒的所有天然宿主中对成熟具有抗性。在存在识别该改变的成熟切割位点的这种特异性蛋白酶的情况下,这些所得的HA被切割成含有HA1和HA2的成熟形式。在这种特异性蛋白酶存在下由所得HA制备的重组流感病毒形成了成熟HA,并对天然流感宿主具有感染性。然而,由于天然宿主中缺乏适当的蛋白酶,在感染的天然宿主中复制的病毒后代不具有感染性。

[0089] H1 HA分子通常在MCS中具有单个碱基残基。这种单一的碱基残基通常是在断裂键

的N-末端的精氨酸(R)残基。在H1 HA的MCS中添加多个碱基残基,如精氨酸(R)或赖氨酸(K),增加了含有修饰的HA的重组流感病毒的感染性。用含有多碱基残基的H5 MCS替换天然H1 MCS增加了含有修饰HA的重组流感病毒的感染性(Kong,W-p.,等人,PNAS(2006)103:15987-15991)。这些多碱基残基更容易被许多细胞内胰蛋白酶样蛋白酶切割。在一些实施方案中,H1 HA MCS通过用H5 HA MCS或多碱基残基替换进行修饰。H1 HA MCS也通过在HA1的最后一个残基(精氨酸)与HA2的第一个残基(甘氨酸)之间插入多碱基残基进行修饰。

[0090] 进一步公开了具有改变的MCS序列的基因修饰的重组流感HA基因和蛋白质。天然MCS被来自另一个HA的MCS的序列替换。所述改变包括用来自不同HA的另一个MCS替换HA的天然MCS,所述另一个MCS已知可能通过使所得HA更容易被未鉴定的宿主蛋白酶成熟切割,来使重组流感病毒更具感染性。在优选的实施方案中,H1 HA的天然MCS被H5 HA的MCS替换,或被精氨酸(R)和赖氨酸(K)的多碱基残基替换。

[0091] 几种哺乳动物蛋白酶如因子Xa和肠激酶具有完全位于断裂键的N-末端侧的其各自的切割识别位点。它们的切割识别位点也可用于替换天然HA成熟切割位点。与天然HA的情况一样,这些蛋白酶切割位点的断裂键的C-末端侧是甘氨酸(G)。凝血酶切割位点也在断裂键的C-末端侧具有甘氨酸(G)。凝血酶切割后,产生了新的甘氨酸(G)N-末端。所有这些蛋白酶都存在于许多天然流感宿主中。然而,这些蛋白酶通常不存在或以痕量存在于用于重组HA产生的细胞培养基或细胞系中。

[0092] 烟草蚀纹病毒(TEV)的TEV蛋白酶识别ENLYFQG的切割位点并在谷氨酸(Q)与甘氨酸(G)之间切割。通过TEV蛋白酶切割产生的游离N-末端是甘氨酸(G),这与HA成熟后HA2的N-末端残基相同。与天然HA的MCS不同,TEV切割位点没有碱基残基。在一些实施方案中,整个HA MCS被TEV切割位点替换,或者MCS的一些残基被TEV切割位点替换,或者仅被断裂键的精氨酸(R)N-末端替换,例如,被TEV切割位点替换。使用杆状病毒表达系统或哺乳动物表达系统表达具有被TEV切割位点替换的MCS的HA构建体,并且已经鉴定了与野生型HA表达水平相同的构建体。

[0093] 在切割修饰的MCS的TEV蛋白酶的存在下,所述修饰的HA经历成熟并且用融合肽的天然N-末端转化为HA1和HA2。成熟的修饰的重组HA具有结合相应宿主细胞受体并经历膜融合的能力。在不存在TEV蛋白酶的情况下,新合成的所述修饰的HA仍然是未切割的HA0前体。融合肽的游离N-末端的缺乏阻止所述HA进行膜融合。在感染的宿主细胞中产生的后代HA蛋白仍然是HA0,并且由于宿主中缺乏TEV蛋白酶而未加工成成熟的功能形式。新制备的具有HA0的病毒不具备膜融合能力,因此无感染性。

[0094] 抗原位点和成熟切割位点的组合

[0095] HA的球状结构域中的抗原位点的变化和HA的茎结构域中的MCS的变化可以是任何种类的组合。可以将特定的异源肽引入具有不同的改变的成熟切割位点的HA蛋白的球状头结构域中的抗原位点。此外,可以将两种或更多种异源肽在不同抗原位点引入单个HA蛋白中。在一个实施方案中,基因修饰的重组流感HA具有M2e肽,其替换HA的球状头结构域中的抗原位点以及作为MCS的TEV蛋白酶位点。在另一个实施方案中,基因修饰的重组流感H1 HA具有M2e肽,其替换HA的球状头结构域中的抗原位点以及H5 MCS(而非H1天然MCS)。

[0096] 生产方法

[0097] 为了产生修饰的HA,蛋白质由细胞以可溶形式分泌至细胞培养基中,并且该蛋白

质不是整合的膜蛋白或不附着于细胞膜或细胞表面。在一些实施方案中，天然HA信号序列用于昆虫细胞和哺乳动物细胞中的分泌。在其他实施方案中，天然HA信号序列被用于在昆虫细胞中分泌的昆虫细胞信号序列或用于在哺乳动物细胞中分泌的哺乳动物信号序列替换。将蛋白酶切割位点置于HA细胞外结构域而非HA跨膜结构域和细胞内结构域的C-末端，然后放置来源于噬菌体T4 fibritin的“foldon”序列以稳定HA三聚体，并放置C-末端His标签以便于纯化。将重组HA蛋白制成具有去除信号序列的全长未切割前体HA0，或在HA2的N-末端具有野生型融合肽的含有HA1和HA2亚基的成熟形式。在优选的实施方案中，最终纯化的基因修饰的HA形成了作为野生型HA的三聚体。

[0098] 设计的重组流感HA的基因通过从头基因合成来制备。基因合成技术已经完善并进行了综述(Kosuri,S.和Church,G.M.,Nat Methods(2014)11:499-507)。长度超过10kb的基因通常由商业提供者制备。基因合成提供了根据表达宿主或特定的基因工程需要修改特定HA序列的能力和密码子优化的机会。在合成的HA基因的特定位置设计限制位点，以促进不同构建体之间HA片段的交换。采用完善的方案将具有信号序列和蛋白酶位点、foldon序列和C-末端His标签的合成HA基因并入杆状病毒基因组中，使得由杆状病毒多角体蛋白启动子引导修饰的HA蛋白的表达。在其他实施方案中，将相同组的HA基因克隆至哺乳动物表达载体中以在哺乳动物细胞中表达。

[0099] 在一些实施方案中，使用杆状病毒载体对HA构建体进行密码子优化以在昆虫细胞中表达。在其他实施方案中，HA构建体经密码子优化以在哺乳动物细胞中表达，包括但不限于CHO细胞(中国仓鼠卵巢细胞)和HEK 293细胞(人胚肾293细胞)。对任何生物体的密码子优化通常如下进行：使用商业软件如来自DNASTAR公司(3801 Regent Street,Madison,WI 53705 USA)的Lasergene软件包，在线Web服务器如OPTIMIZER(在万维网，genomes.urv.es/OPTIMIZER/)，或通过基因合成服务提供商使用其专有算法。密码子优化考虑了G/C含量、隐蔽性剪接位点和RNA去稳定序列元件的消除，以及稳定的RNA二级结构的回避。密码子也是根据密码子使用数据库(在万维网，kazusa.or.jp/codon/)手动调整的。由于密码子简并性，可以在不改变编码的氨基酸序列的情况下改变基因序列。通过改变密码子而不改变氨基酸序列在特定位置引入限制位点。使用这些方法中的任何一种，在有或没有对特定生物体密码子进行优化的情况下，通常通过利用计算机算法或手动对蛋白质序列进行回译来生成基因序列。

[0100] 其他实施方案包括使用已建立的细胞培养方法，由从MCS改变为TEV蛋白酶识别位点的基因修饰的HA基因制备重组非感染性流感病毒的方法。已经开发了允许在使用或不使用辅助病毒的情况下由具有克隆cDNA的RNA区段或质粒产生流感病毒的反向遗传学系统(Luytjes,W.,等人,Cell(1989)59:1107-1113;Neumann,G.,等人,PNAS(1999)96:9345-9350;Fodor,E.,等人,J.Virol(1999)73:9679-9682;de Wit,E.,等人,J.Gen Virol(2007)88:1281-1287)。感染性流感病毒是通过用具有流感RNA的cDNA的流感RNA区段或质粒瞬时转染哺乳动物细胞来制备的。从细胞培养基分离出的这些病毒用于感染含胚鸡蛋，以产生用于制造疫苗的活的感染性流感病毒。一些流感RNA区段被外源基因替换。使用替换HA蛋白的丙型流感病毒HEF蛋白制备重组甲型流感病毒(Gao,Q.,等人,J.Virol(2008)82:6419-6426)。在这些实施方案中，将修饰的HA的cDNA或RNA与通过标准分子技术和基因合成制备的其他7种流感RNA或其他7种RNA区段的cDNA共转染至哺乳动物细胞。使用已建立的方法或

类似方法,可将多个相同或不同的HA蛋白包装在单个病毒粒子中(Uraki,R.,等人,J.Virol(2013)87:7874-7881)。

[0101] 由基因修饰的HA基因制备重组感染性流感病毒的方法使用如上所述的已建立的细胞培养方法。在这些实施方案中,将在MCS处具有H5 HA MCS或多碱基序列的基因修饰的HA基因与通过如上所述的标准反向遗传学系统制备的其他7种流感RNA或其他7种RNA区段的cDNA共转染至哺乳动物细胞。这些修饰的成熟切割位点提高了流感天然宿主或细胞培养物中HA成熟的效率(Kong,W-p.,等人,PNAS(2006)103:15987-15991)。成熟的功能性HA蛋白具有与宿主细胞结合以及与宿主细胞膜融合的能力。

[0102] 其他实施方案包括使用在哺乳动物细胞、昆虫细胞和植物细胞中建立的细胞培养方法由基因修饰的HA基因制备重组流感病毒样颗粒(VLP)的方法(Chen,B.J.,等人,J.Virol(2007)81:7111-7123;Smith,G.E.,等人,Vaccine(2013)31:4305-4313;D'Aoust,M.A.,等人,Plant Biotech(2010)8:607-619)。

[0103] 其他实施方案包括使用已建立的方法由基因修饰的HA基因制备DNA疫苗的方法(Jiang,Y.,等人,Antiviral Res(2007)75:234-241;Alexander,J.,等人,Vaccine(2010)28:664-672;Rao,S.S.,等人,PLoS ONE(2010)5:e9812)。

[0104] 本文公开的表达结果显示,可以修饰H1 HA的MCS而不影响所得HA的表达。已证明具有H5 HAMCS或多碱基MCS的构建体与野生型H1 HA的表达水平相同。此外,HA的MCS被不具有任何碱基残基的TEV蛋白酶切割位点替换。通过改变TEV蛋白酶切割位点的位置,已经制备了具有TEV切割位点作为MCS的构建体,该构建体与野生型H1 HA的表达水平相同。

[0105] 表达结果进一步表明,HA1的球状头结构域中的许多抗原位点可以被来自相同HA或不同HA的茎结构域的异源肽替换。一些构建体的表达与野生型HA的表达水平相同,而其他构建体表达更少。异源肽的性质和异源肽在球状头结构域中的位置对每种所得HA构建体的表达水平具有显著影响。可以针对其表达以相同的方式制备和测试更多构建体,以鉴定最佳表达构建体。

[0106] 此外,通过来自流感M2蛋白的M2e表位替换HA球状头结构域的某些免疫显性抗原位点来制备重组HA。可以想到,其他流感蛋白的任何表位可以用作替换肽来替换HA球状头结构域中的免疫显性抗原位点。此外,异源替换肽可来源于与流感病毒无关的另一种蛋白质。可以选择HA球状头结构域中的特定抗原位点用于替换肽来尽可能多地维持替换肽的天然结构特征,以将替换肽呈递给宿主免疫系统。

[0107] 杆状病毒表达系统

[0108] 自大约30年前其引入以来(美国专利4,745,051;Summers,M.D.和Smith,G.E.,(1987)A Manual ofMethods for Baculovirus Vectors and Insect Cell Culture Procedures.Texas Agricultural Experiment Station Bulletin No.1555.),杆状病毒表达载体系统(BEVS)已被用于表达许多不同类型的人和病毒蛋白,包括细胞内蛋白、膜蛋白和分泌蛋白。BEVS已被用于生产病毒样颗粒(VLP)。BEVS是使用昆虫细胞作为宿主的真核表达系统,其提供类似于哺乳动物细胞的蛋白质的翻译后修饰(Jarvis,D.L.,"Baculovirus Expression Vectors"于The Baculoviruses中,Miller,L.K.编(1997)pp.389-420Plenum Press,New York)。已经完善基于苜蓿银纹夜蛾核型多角体病毒(AcNPV)的BEVS。许多商业试剂盒可用于生产重组杆状病毒。BEVS已成功用于生产在美国销

售的重组疫苗。使用BEVS生产了两种FDA批准的疫苗:Cervarix™-一种抗宫颈癌的非感染性病毒样颗粒(VLP)形式的重组人乳头瘤病毒二价(16和18型)疫苗(专利申请WO 2010/012780 A1);和Flublok®(美国专利5,762,939 A和5,858,368 A)-抗流感的由膜结合血凝素前体(HA0)制成的三价流感疫苗。

[0109] 为了使用BEVS表达靶蛋白,首先将编码靶蛋白的目的基因亚克隆至转移载体中,该转移载体是含有在杆状病毒多角体蛋白基因侧翼的杆状病毒序列的大肠杆菌质粒。多角体蛋白是细胞培养中杆状病毒繁殖所不需要的野生型杆状病毒中丰富的病毒蛋白。转移载体在大肠杆菌中繁殖并使用标准分子生物学技术进行分离(Sambrook,J.,等人,(1989)Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Laboratory Press,Cold Spring Harbor)。将具有目的基因的转移载体与杆状病毒基因组DNA重组,以在大肠杆菌或昆虫细胞中产生重组杆状病毒基因组。重组杆状病毒基因组引导重组杆状病毒的产生和靶蛋白的表达。原始方法依赖于转移载体中和昆虫细胞中病毒基因组DNA中的多角体蛋白基因侧翼的相同序列之间的同源重组。将分离的转移载体DNA与杆状病毒基因组DNA共转染至昆虫细胞中。针对多角体蛋白的缺乏选择重组杆状病毒。目前的商业试剂盒,如来自Clontech的BacPAK™杆状病毒表达系统和来自Millipore的BacMagic™系统,使用线性化的杆状病毒基因组DNA,其在不重组的情况下不产生活病毒。这些试剂盒允许在高效且对非重组病毒的污染很小的情况下生产重组杆状病毒。还可以使用Gateway重组反应(BaculoDirect™杆状病毒表达系统,生命技术公司(Life Technologies))在体外对病毒基因组进行重组。制备重组病毒的另一种方法是通过大肠杆菌中的位点特异性转座(Bac-to-Bac系统,生命技术公司)。基于pFastBac的转移载体含有在大肠杆菌中制备杆粒(含有杆状病毒基因组的大质粒)的转座子(美国专利5,348,886)。在将杆粒转染至昆虫细胞之前,通过聚合酶链反应(PCR)可以容易地证实目标基因与杆粒的重组。

[0110] 重组杆状病毒在昆虫细胞培养物中繁殖。少量病毒用于感染昆虫细胞。几天后,收获含有扩增病毒的条件培养基作为病毒原液。这种扩增过程通常重复几次,以产生大量的病毒原液。病毒原液通常在黑暗中冷藏储存数月甚至数年。对病毒原液的5%-10%胎牛血清(FBS)的补充已用于保存病毒原液,并被认为可延长其保质期。病毒原液有时会冷冻以在-70℃下长期储存。病毒也以杆状病毒感染的昆虫细胞(BIIC)的形式被扩增和储存(Wasilko,D.J.,等人,Prot Exp Purif(2009)65:122-132)。感染的昆虫细胞在其裂解之前收获作为BIIC原液,并在标准细胞冷冻程序后冷冻。BIIC原液在液氮中或在-65℃至-85℃的超低温下储存很长一段时间,并用作病毒原液以感染昆虫细胞以进行蛋白质表达。冷冻BIIC原液比液体形式的病毒原液提供更长的储存时间。

[0111] BEVS常用的昆虫细胞系是来源于秋粘虫(草地贪夜蛾)的SF9和SF21细胞以及来源于粉纹夜蛾(粉纹夜蛾(Trichoplusia ni))的Hi5或粉纹夜蛾细胞。还使用了来源于桑蚕(家蚕),蜂窝蛾(大蜡螟)和舞毒蛾(舞毒蛾(Lymantria dispar))的物种的其他昆虫细胞。SF9、SF21和Hi5(或粉纹夜蛾)细胞已经适应于无血清培养基中的悬浮培养。这些细胞系和培养基可从许多商业来源获得。通常将这些细胞系的细胞培养物以1升或2升的小体积保存在22℃-28℃的温度范围下不补充气体的环境大气中的振荡培养箱的摇瓶中。细胞培养物在搅拌釜生物反应器或一次性生物反应器中按比例放大。生物反应器中用于大规模昆虫细胞培养的条件已经完善(WAVE生物反应器系统-细胞培养程序。通用医疗(GE

Healthcare))。补充氧气也常用于大规模昆虫细胞培养,以增加细胞密度。

[0112] 测试用于表达特定靶蛋白的一系列条件,以通过改变病毒与细胞的比率和感染后的收获时间,使用不同的细胞系优化靶蛋白的表达。SF9、SF21和Hi5(或粉纹夜蛾)细胞是用于BEVS蛋白表达的最常见细胞系。靶蛋白可以在一种细胞系中表达得比其他细胞系更好。一些报道显示Hi5(或粉纹夜蛾)细胞更多地表达某些分泌蛋白。测试病毒与细胞的比率,通常称为感染复数(MOI),以确定靶蛋白表达的最佳感染条件。在感染后的不同时间点采集培养样品。通过离心或过滤分离细胞和条件培养基。通过标准方法测定蛋白质表达水平。监测培养物的细胞密度、细胞活力和细胞大小,这些提供了细胞生长和培养条件的信息。针对培养基中营养物的消耗,有时监测培养物的葡萄糖水平、溶解氧和pH。选择导致最佳靶蛋白表达的条件用于蛋白质的大规模生产。

[0113] 尽管病毒原液的滴度和MOI的测定被广泛用来确定病毒扩增和利用BEVS进行蛋白质表达的感染条件,但上述TIPS方法提供了更快的方法来确定病毒扩增和蛋白质表达的最佳感染条件。昆虫细胞在感染重组杆状病毒后大小增大。细胞分裂也在感染后停止。重组杆状病毒在感染后约48小时引起细胞裂解。通过将细胞计数为培养物的细胞密度来监测细胞分裂。通过计数培养物中的活细胞和死细胞,以细胞活力的形式监测细胞裂解。活细胞大小以细胞直径衡量。细胞密度、细胞活力和活细胞大小通常由许多细胞计数仪器模型来测量。细胞密度和细胞活力也可以使用血细胞计数器来手动测量。细胞密度、细胞活力和活细胞大小共同提供感染动力学的信息。TIPS法使用感染动力学来确定使用特定病毒原液表达特定靶蛋白的最佳条件。它消除了用于MOI计算的病毒滴度的耗时测量。通常使用TIPS法实现一致的病毒扩增和蛋白质表达。

[0114] 除昆虫细胞外,感染重组杆状病毒的活昆虫已被用于表达许多分泌蛋白和膜蛋白。成熟HA已在烟青虫(烟芽夜蛾)的幼虫中表达(Kuroda,K.,等人,J.Virol(1989)63:1677-1685)。

[0115] 还开发了基于其他类型的杆状病毒如家蚕核型多角体病毒(BmNPV)的表达系统。BmNPV可能比AcNPV具有更好的生物安全性,因为BmNPV具有较窄的宿主范围并且不会在田间生长为害虫。基于BmNPV的杆状病毒表达系统已用于在细胞培养物和桑蚕(家蚕)幼虫中表达功能性蛋白质(Maeda,S.,等人,Nature(1985)315:592-594)。

[0116] BEVS通常用于生产蛋白质复合物和VLP。将两个或更多个基因克隆至单个载体中以表达多种蛋白质。将各自引导单一蛋白质表达的两种或更多种病毒原液用于昆虫细胞的共感染,以产生蛋白质复合物或VLP。已经使用杆状病毒表达系统制备了流感VLP(Bright,R.A.,等人,PLoS ONE(2008)3:e1501)。

[0117] 用于杆状病毒表达的HA转移载体

[0118] 有待于使用BEVS表达的基因通常在多角体蛋白启动子的控制下,该多角体蛋白启动子是在杆状病毒诱导的昆虫细胞死亡之前导致蛋白质表达的杆状病毒的强晚期启动子。其他早期或晚期启动子也用于蛋白质表达。为了通过昆虫细胞将重组靶蛋白分泌至细胞培养基中,编码信号肽的DNA区段在目的基因的框内的5'端被工程化。昆虫细胞的两种常用信号肽是蜜蜂蜂毒肽信号序列或AcNPV包膜表面糖蛋白GP67信号序列。信号序列数据库(在万维网,signalpeptide.de/index.php)列出了许多其他可能的信号肽以供考虑。分泌后,通过处理信号肽的细胞蛋白酶去除信号肽。尽管流感病毒不是昆虫病毒,但HA信号肽已被用

作分泌信号肽，用于BEVS中重组蛋白的分泌或膜插入。MKTIIALSYIFCLVFA的HA信号肽通常用于在昆虫细胞中表达重组跨膜G蛋白偶联受体(GPCR)(Rosenbaum,D.M.,等人,Science(2007)318:1266-1273;Zou,Y.,等人,PLoS One(2012)7:e46039)。已经使用其天然信号肽在昆虫细胞中表达HA。在该特定情况下，总表达低于使用昆虫细胞信号肽的总表达，但表达的HA是含有HA1和HA2的成熟形式(美国专利5,858,368A)。HA信号肽在HA中不保守。例如，前述HA信号肽不同于SEQ ID NO:7的H1 HA信号肽。HA信号肽的经验选择可以改善昆虫细胞中的HA表达。

[0119] 商业载体可用于通过同源重组，使用Bac-to-Bac系统或Gateway系统(生命技术公司)制备重组杆状病毒。这些转移载体的特征如启动子和信号序列可以通过标准分子生物学技术自定义。现在可以通过基因合成完全合成整个转移载体。基因合成允许设计具有特定序列和特征的转移载体。

[0120] 使用BEVS表达流感HA的方法已在实验室规模上完善(Stevens,J.,等人,Science(2004)303:1866-1870)。例如，转移载体含有来自在多角体蛋白启动子控制下的1918流感病毒的HA。1918流感病毒HA表达构建体在N-末端具有GP67信号肽而不是HA信号肽以进行分泌。将凝血酶切割位点引入细胞外结构域而非跨膜结构域的C-末端，然后放置来源于噬菌体T4fibritin的“foldon”序列以稳定HA三聚体，并放置C-末端His标签以便于纯化。可以通过凝血酶切割去除foldon序列和His-标签。HA表达转移载体的类似设计已被用于表达来自许多流感毒株的不同HA(Stevens,J.,等人,Science(2006)312:404-410;Xu,R.,等人,Science(2010)328:357-360;Whittle,J.R.R.,等人,PNAS(2011)108:14216-14221)。

[0121] 除了foldon之外，其他三聚化结构域可用于稳定重组HA的三聚体。例如，热稳定的HIV-1糖蛋白41(gp41)三聚化结构域或GCN4亮氨酸拉链序列的31个残基中的前16个被用于帮助截短的HA构建体的三聚化(Impagliazzo,A.,等人,Science(2015)349:1301-1306;Yassing,H.M.,等人,Nat Med(2015)21:1065-1070)。

[0122] 如美国专利5,762,939 A、5,858,368 A和6,245,532 B1中所公开的，已经使用商业规模的BEVS制备来自甲型流感病毒和乙型流感病毒的具有跨膜结构域和细胞内结构域的重组HA前体(HA0)。这些重组HA蛋白是已获得FDA批准作为流感疫苗的Flublok®中的成分。Flublok®中的HA0使用杆状病毒几丁质酶信号肽(称为61K信号肽)替换HA信号肽来制备。重组HA与昆虫细胞的外周膜结合。使用洗涤剂从膜中提取该重组HA并进一步纯化。

[0123] 哺乳动物表达

[0124] 中国仓鼠卵巢(CHO)细胞和人胚肾293(HEK293)细胞是用于重组蛋白的瞬时转染基因表达的常用哺乳动物细胞宿主。CHO细胞和HEK293细胞二者均具有悬浮细胞系和粘附细胞系。这些细胞通常在有或没有FBS的培养基中培养。当细胞在没有FBS的情况下培养时，通常使用化学成分确定的培养基。所有悬浮和粘附细胞系均可用于表达HA或其他分泌蛋白。

[0125] 瞬时转染是将DNA引入细胞进行蛋白质表达的完善方法。有许多商业转染试剂可供使用。与杆状病毒表达的表达优化过程类似，在转染后定期采集培养基样品，以使用标准方法分析蛋白质表达。通过HA构建体中的His标签的亲和捕获或使用许多市售的抗HA抗体的Western印迹，容易地检测HA蛋白。蛋白质表达通常在转染后约48小时开始，并且可在转染后数天内增加。有时在转染后添加补充剂以增加靶蛋白表达。其他哺乳动物细胞系如

COS-1,来源于非洲绿猴肾组织的成纤维细胞样细胞系也已用于HA表达。流感VLP通过共转染含有全部10种流感病毒编码的蛋白质的cDNA的多种质粒而产生(Mena,I.,等人,J.Virol(1996)70:5016-5024;Chen,B.J.,等人,J.Virol(2007)81:7111-7123)。

[0126] 将基因递送至哺乳动物细胞的另一种方法使用称为BacMam的重组杆状病毒,以将杆状病毒基因转移至哺乳动物细胞中(Boyce,F.M.和Bucher,N.L.,PNAS(1996)93:2348-2352;Dukkipatia.A.,等人,Protein Expr Purif(2008)62:160-170)。通过将哺乳动物表达盒并入杆状病毒表达载体中来修饰杆状病毒,以在哺乳动物细胞中进行转基因表达。重组BacMam杆状病毒通过杆状病毒的标准生产方法产生。将扩增的重组BacMam杆状病毒的病毒原液添加至CHO或HEK293培养物中,以将哺乳动物表达盒递送至CHO或HEK293细胞以进行蛋白质表达。BacMam平台能够轻松转导大量哺乳动物细胞。

[0127] 转染至哺乳动物细胞的基因表达载体可整合至细胞染色体中以建立用于表达靶蛋白的稳定细胞系。稳定的CHO细胞系是生产治疗性生物制剂如单克隆抗体的最常见宿主。该技术非常适合大规模生产治疗性蛋白质和疫苗。为了建立稳定的细胞系,基于表达载体上的选择标志物对用表达载体转染的细胞进行选择。通常将靶基因的多个拷贝整合至稳定细胞系的基因组中。

[0128] 用于哺乳动物表达的HA载体

[0129] 用于哺乳动物表达的HA表达构建体具有与杆状病毒表达相同的设计和氨基酸序列。针对CHO细胞表达或HEK293细胞表达对密码子进行了优化。杆状病毒密码子也适用于哺乳动物细胞,反之亦然。HA信号肽已用作信号肽,以在哺乳动物表达系统中分泌或膜插入重组蛋白。其他常用的哺乳动物信号肽包括人IL2信号肽、组织纤溶酶原激活物(tPA)信号肽、以及信号序列数据库(在万维网,signalpeptide.de/index.php)中发现的许多其他信号肽。许多哺乳动物表达载体可商购获得。每个载体通常具有用于高水平表达的增强子-启动子、用于mRNA稳定性的多腺苷酸化信号和转录终止序列、用于附加型复制的SV40起点、抗生素抗性基因和pUC起点(以用于在大肠杆菌中选择和维持)。用于哺乳动物表达的常用启动子包括CMV(巨细胞病毒)启动子、hEF1-HTLV启动子、包含延长因子-1α(EF-1α)核心启动子和人T细胞白血病病毒(HTLV)1型长末端重复序列的R区段和U5序列的部分(R-U5')的复合启动子、或者MPromDb(哺乳动物启动子数据库,在万维网,mpromdb.wistar.upenn.edu/)或真核启动子数据库(EPD,在万维网,epd.vital-it.ch/)中发现的其他启动子。转移载体通常具有用于产生稳定的细胞系的另一种选择标志物。

[0130] 提供以下实施例来说明而不是限制本发明。

[0131] 实施例1

[0132] 用于重组杆状病毒生产的H1 HA的转移载体的构建

[0133] 在该实施例中,描述了可以被修饰用于制备本发明的修饰的HA蛋白的亲本质粒。该质粒的基本轮廓显示在图4中,其中用于在杆状病毒中表达的标准质粒中的插入片段被这种插入片段的限制位点括起来,并且存在限制位点,其允许替换HA蛋白的MCS和替换HA的球状头结构域中的免疫显性区域。将编码的蛋白质与用于纯化的10个组氨酸(10xHis标签)的亲和标签融合,所述亲和标签可以在纯化后通过插入的蛋白酶切割位点的性能而被去除。该载体被命名为H1WT。

[0134] A/California/07/2009,即由WHO推荐用于2009流感疫苗生产(如在万维网

who.int/csr/resources/publications/swineflu/vaccine_recommendations/en/所述)的猪源甲型流感病毒H1N1毒株,它的HA序列(SEQ ID NO:1)用作亲本序列以并入异源肽。该HA序列的登录号为UniProt:C3W5X2。所有构建体具有如图4中示意性所示的相同的设计。N-末端HA信号肽(SEQ ID NO:7)被GP67信号肽MVSAIVLYVLLAAAAHSAFA(SEQ ID NO:2)替换。基于首次由Stevens,J.,等人,Science(2006)312:404-410描述的HA构建体,具有ILAIYSTVASSLVLVVSLGAISFWMCS序列的C-末端跨膜(TM)结构域和具有HA的NGSLQCRICI序列的细胞内结构域被TEV切割位点(SEQ ID NO:4)替换,然后被foldon(SEQ ID NO:5)和10xHis标签(SEQ ID NO:6),即GAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH(SEQ ID NO:3)替换。foldon稳定重组HA,而10xHis标签促进从细胞培养基中纯化HA。可以通过能识别位于HA序列与foldon序列之间的TEV切割位点的TEV蛋白酶去除foldon和10xHis标签。针对杆状病毒表达对该基因进行密码子优化,并通过标准方法(GENEWIZ,South Plainfield,NJ 07080,USA)合成该基因。

[0135] 如图4中示意性所示,将间隔的独特限制位点:ClaI位点、NsiI位点和BamHI位点引入HA序列以促进序列改变和交换。将得到的构建体GP67-H1WT-TEV-foldon-10His(表示为"H1WT"(SEQ ID NO:9))亚克隆至NcoI位点与HindIII位点之间的pFastBac质粒中。ClaI与NsiI位点之间的序列编码球状头结构域的大多数免疫显性抗原位点。NsiI与BamHI位点之间的序列编码MCS。H1WT构建体的表达受杆状病毒多角体蛋白启动子的控制。得到的质粒是用于表达A/California/07/2009的重组野生型HA的转移载体。H1WT构建体是亲本构建体,由其制备基因修饰的HA构建体。

[0136] 为了用异源肽制备基因修饰的HA,使用具有编码HA的特异性变化的合成DNA片段的上述限制位点替换H1WT的野生型片段。为了修饰MCS,H1WT的NsiI-BamHI片段被编码改变的成熟切割位点的DNA片段替换。为了改变HA1球状头结构域中的抗原位点,H1WT的ClaI-NsiI片段被编码具有异源肽的改变的抗原位点的DNA片段替换。通过简单交换不同质粒的限制片段,制备具有成熟切割位点和抗原位点的不同组合的构建体。通过DNA测序证实得到的构建体。

[0137] 实施例2

[0138] 设计和构建具有修饰的成熟切割位点的H1 HA构建体

[0139] 在该实施例中,包含MCS的NsiI与BamH1限制位点之间的H1WT的区段被修饰该MCS的替代序列替换。在几种情况下,TEV切割位点被包含在修饰的MCS中。

[0140] 将修饰的成熟切割位点引入实施例1中所述的H1 HA的转移载体中。具有修饰的成熟切割位点的构建体列于表1中。

[0141] 表1具有改变的成熟切割位点的H1 HA构建体的列表

[0142]

列表编号	构建体名称	构建体描述	SEQ ID编号
	H1MCS[a]	GP67ss-H1MCS-TEV-foldon-10His[b]	
1	H1WT[c]	GP67ss-H1WT-TEV-foldon-10His	SEQ ID NO: 9
2	H1H5cs	GP67ss-H1H5cs-TEV-foldon-10His	SEQ ID NO: 11
3	H1R5cs	GP67ss-H1R5cs-TEV-foldon-10His	SEQ ID NO: 14
4	H1IR5cs	GP67ss-H1IR5cs-TEV-foldon-10His	SEQ ID NO: 16
5	H1TEV1	GP67ss-H1TEV1-TEV-foldon-10His	SEQ ID NO: 18
6	H1TEV2	GP67ss-H1TEV2-TEV-foldon-10His	SEQ ID NO: 20
7	H1TEV3	GP67ss-H1TEV3-TEV-foldon-10His	SEQ ID NO: 22

[a]构建体名称基于成熟切割位点。H1表示H1 HA。MCS表示成熟切割位点。

[b]构建体由其序列特征描述。GP67ss表示GP67信号序列。TEV-foldon-10His表示构建体的C-末端处的TEV切割位点、foldon序列和10xHis标签的序列。构建体的示意图如图4所示。

[c] H1WT是具有天然MCS的野生型H1 HA构建体。

[0143] 这些修饰的位置示于图5。

[0144] 构建体H1H5cs(SEQ ID NO:11)具有被H5 MCS(SEQ ID NO:12)替换的野生型H1 HA(SEQ ID NO:9)的MCS。构建体H1R5cs(SEQ ID NO:14)具有由五个精氨酸替换的野生型H1的MCS。构建体H1IR5cs(SEQ ID NO:16)具有在保留所有天然HA1残基的同时插入HA1与HA2序列之间的四个精氨酸。合成具有变化的NsiI-BamHI片段并亚克隆至H1WT的NsiI与BamHI位点之间以替换野生型片段。

[0145] 为了测试H1 HA MCS是否可以被修饰至不含任何碱基残基的TEV切割位点,通过改变H1 HA MCS的残基数来制备几种构建体。构建体H1TEV1(SEQ ID NO:18)在H1 HA MCS处具有7个残基PSIQSRG,其被ENLYFQG的7个残基TEV切割位点(SEQ ID NO:4)替换。构建体H1TEV2(SEQ ID NO:20)具有由含有TEV切割位点的SPENLYFQG替换的H1 HA MCS的8个残基IPSIQSRG。构建体H1TEV3(SEQ ID NO:22)具有由TEV切割位点替换的H1 HA MCS(SEQ ID NO:8)的最后3个残基(QSR)。

[0146] 实施例3

[0147] 在球状头结构域中具有异源表位的H1 HA构建体的设计和构建

[0148] 在该实施例中,通过替代的核苷酸序列在ClaI与NsiI位点之间修饰碱基质粒和具有修饰的MCS的质粒,在所述替代的核苷酸序列中各种免疫显性位点被异源表位替换。

[0149] 将异源表位引入实施例1和实施例2中描述的构建体中。将异源表位插入HA球状头结构域的免疫显性抗原位点周围或将其用HA球状头结构域的免疫显性抗原位点替换。具有

修饰的抗原位点的构建体列于表2中。

[0150] 表2在球状头结构域中具有修饰的免疫显性抗原位点的H1 HA构建体的列表

[0151]

列表编号	构建体名称	构建体描述	SEQ ID No.
	H1MCS-AS[a]	GP67ss-H1MCS - AS - TEV-foldon-10His[b]	
8	H1H5cs-CR8020Ca	GP67ss-H1H5cs-CR8020Ca-TEV-foldon-10His	SEQ ID NO: 24
9	H1TEV2-CR8020Ca	GP67ss-H1TEV2-CR8020Ca-TEV-foldon-10His	SEQ ID NO: 29
10	H1TEV2-CR8020Sa3	GP67ss-H1TEV2-CR8020Sa3-TEV-foldon-10His	SEQ ID NO: 31
11	H1TEV2-CR8020Sa4	GP67ss-H1TEV2-CR8020Sa4-TEV-foldon-10His	SEQ ID NO: 33
12	H1TEV2-I1C9Ca1	GP67ss-H1TEV2-I1C9Ca1-TEV-foldon-10His	SEQ ID NO: 35

[0152]

列表编号	构建体名称	构建体描述	SEQ ID No.
	H1MCS-AS[a]	GP67ss-H1MCS - AS - TEV-foldon-10His[b]	
13	H1TEV2-I1C9Ca2	GP67ss-H1TEV2-I1C9Ca2-TEV-foldon-10His	SEQ ID NO: 38
14	H1TEV2-I1C9Sa3	GP67ss-H1TEV2-I1C9Sa3-TEV-foldon-10His	SEQ ID NO: 40
15	H1TEV2-I1C9Sa4	GP67ss-H1TEV2-I1C9Sa4-TEV-foldon-10His	SEQ ID NO: 42
16	H1H5cs-I1C9Sb	GP67ss-H1H5cs-I1C9Sb-TEV-foldon-10His	SEQ ID NO: 44
17	H1H5cs-I1C9Ca	GP67ss-H1H5cs-I1C9Ca-TEV-foldon-10His	SEQ ID NO: 46
18	H1H5cs-FI6Sab	GP67ss-H1H5cs-FI6Sab-TEV-foldon-10His	SEQ ID NO: 48
19	H1WT-FI6Sab	GP67ss-H1WT-FI6Sab-TEV-foldon-10His	SEQ ID NO: 52
20	H1TEV1-FI6Sa	GP67ss-H1TEV1-FI6Sa-TEV-foldon-10His	SEQ ID NO: 54
21	H1H5cs-M2eCa	GP67ss-H1H5cs-M2eCa2-TEV-foldon-10His	SEQ ID NO: 56
22	H1H5cs-M2eSb	GP67ss-H1H5cs-M2eSb-TEV-foldon-10His	SEQ ID NO: 59
23	H1TEV2-M2eCa	GP67ss-H1TEV2-M2eCa2-TEV-foldon-10His	SEQ ID NO: 61

[a]构建体名称基于成熟切割位点和抗原位点修饰。H1表示H1 HA。MCS表示成熟切割位点。AS表示具有异源表位的H1 HA的球状头结构域中改变的免疫显性抗原位点。每个AS由异源表位和放置异源表位的H1 HA免疫显性抗原位点表示。

[b]构建体由其序列特征描述。GP67ss表示GP67信号序列。TEV-foldon-10His表示构建体的C-末端处的TEV切割位点、foldon序列和10xHis标签。构建体的示意图如图4所示。

[0153] 这些修饰的位置示于图5。

[0154] 13个氨基酸残基的复合CR8020表位肽替换抗原位点Ca2周围的12个残基EIAIRPKVRDQE的表面环。合成具有复合CR8020表位肽取代的ClaI-NsiI片段并亚克隆在H1H5cs的ClaI与NsiI位点之间，以产生表示为H1H5cs-CR8020Ca(SEQ ID NO:24)的HA构建

体。为了制备H1TEV2-CR8020Ca(SEQ ID NO:29),分离H1H5cs-CR8020Ca的ClaI-NsiI片段并亚克隆至H1TEV2的ClaI与NsiI位点之间。

[0155] 构建体H1TEV2-CR8020Sa3(SEQ ID NO:31)具有替换Sa位点之一的残基KKGNS的复合CR8020表位肽,其对应于球状头结构域顶部的H3 HA的位点B环。合成具有CR8020修饰的ClaI-NsiI片段并亚克隆在H1TEV2的ClaI与NsiI位点之间以产生表示为H1TEV2-CR8020Sa3的HA构建体。

[0156] 构建体H1TEV2-CR8020Sa4(SEQ ID NO:33)具有复合CR8020表位肽,其替换Sa位点之一附近的残基KTSS的螺旋结构。该位置位于HA三聚体的球状头结构域的侧面并且远离受体结合位点。这种修饰不太可能改变受体结合。合成具有CR8020修饰的ClaI-NsiI片段并亚克隆在H1TEV2的ClaI与NsiI位点之间以产生表示为H1TEV2-CR8020Sa4的HA构建体。

[0157] 构建体CR8020Sa4和CR8020Ca的HA单体模型显示在图6中。每种HA单体在球状头结构域中具有复合CR8020表位肽,在茎结构域中具有天然CR8020表位。

[0158] 制备了几种具有GIFGAIAGFIEG(SEQ ID NO:36)的I1C9表位肽的HA构建体。一种构建体H1TEV2-I1C9Ca1(SEQ ID NO:35)具有I1C9肽,其在Ca2位点周围的残基217至224处替换RPKVRDQE。另一种构建体H1H5cs-I1C9Ca(SEQ ID NO:46)具有I1C9肽,其在Ca2的残基213至224处替换较长的肽EIAIRPKVRDQE。其他构建体具有I1C9肽,其分别替换对应于H3 HA的位点A的Ca2位点HAGAKS、Sa位点KKGNS和Sa位点KTSS。合成每个具有变化的ClaI-NsiI片段并亚克隆在H1TEV2的ClaI与NsiI位点之间,以产生分别表示为H1TEV2-I1C9Ca1(SEQ ID NO:35)、H1TEV2-I1C9Ca2(SEQ ID NO:38)、H1TEV2-1C9Sa3(SEQ ID NO:40)和H1TEV2-I1C9Sa4(SEQ ID NO:42)的HA构建体。合成具有I1C9Ca变化的ClaI-NsiI片段并亚克隆在H1H5cs的ClaI与NsiI位点之间,以产生HA构建体H1H5cs-I1C9Ca。I1C9肽也通过相同的方法替换H1H5cs构建体如H1H5cs-I1C9Sb(SEQ ID NO:44)的Sb位点的残基TSADQQSLYQNA。Sb位点对应于H3 HA的位点B螺旋。

[0159] ClaI-NsiI片段采用位于Sa位点的FI6表位肽RKKRGLFGAIAGFIE(SEQ ID NO:49)和位于Sb位点的卷曲螺旋FI6表位肽KESTQKAIDGVTNKVNS(SEQ ID NO:50)来进行基因合成。将具有FI6取代的ClaI-NsiI片段亚克隆在H1H5c和H1WT的ClaI与NsiI位点之间,分别产生了构建体H1H5cs-FI6Sab(SEQ ID NO:48)和H1WT-FI6Sab(SEQ ID NO:52)。另一个ClaI-NsiI片段采用位于Sa位点的FI6表位肽RKKRGLFGAIAGFIE(SEQ ID NO:49)来进行基因合成。该片段替换H1TEV1的ClaI-NsiI片段以产生构建体H1TEV1-FI6Sa(SEQ ID NO:54)。

[0160] 实施例4

[0161] 设计和构建在球形头结构域的抗原位点具有M2e肽的H1 HA构建体

[0162] 在该实施例中,同样,ClaI-NsiI片段的核苷酸序列被如下核苷酸序列替换,其中Ca2位点或Sb位点的一部分被编码M2e肽的核苷酸序列替换。具有M2e肽的构建体列于表2中。这些修饰的位置示于图5。

[0163] 基因合成产生了具有M2e肽的ClaI-NsiI片段(SEQ ID NO:57),所述M2e肽替换Ca2位点处的EIAIRPKVRDQE或替换Sb位点的残基TSADQQSLYQNA的螺旋。具有M2e肽的HA构建体首先在H1H5cs中制备,其中H1H5cs的ClaI-NsiI片段被含有M2e肽的单个ClaI-NsiI片段替换,分别得到了构建体H1H5cs-M2eCa(SEQ ID NO:56)和H1H5cs-M2eSb(SEQ ID NO:59)。为了将H1H5cs-M2eCa的ClaI-NsiI片段转移至H1TEV2构建体,分离H1H5cs-M2eCa的ClaI-NsiI

片段并亚克隆在H1TEV2的ClaI与NsiI位点之间,得到了名为H1TEV2-M2eCa的构建体(SEQ ID NO:61)。

[0164] 实施例5

[0165] 使用Bac-to-Bac杆状病毒表达系统产生重组杆状病毒

[0166] 该实施例描述了制备昆虫细胞的技术,所述昆虫细胞包含实施例1-4中制备的构建体中包含的表达系统。

[0167] 重组杆状病毒根据制造商的说明书由构建体的pFastBacbased转移载体产生(Bac-to-Bac杆状病毒表达系统,生命技术公司(Life Technology),卡尔斯巴德(Carlsbad),加利福尼亚州,美国)。简言之,将转移载体转化为大肠杆菌DH10Bac化学感受态细胞。在具有Blue-gal的平板上选择具有重组杆粒的白色菌落。根据制造商的说明书,通过标准碱性裂解微量制备方法分离Bacmid DNA。使用M13正向(-40)引物(GTTTTCCCAGTCACGAC)和M13反向引物(CAGGAAACAGCTATGAC)通过PCR鉴定重组杆粒。重组杆粒产生约4kb的PCR产物。

[0168] 使用Cellfectin®试剂、其他可商购的转染试剂或聚乙烯亚胺(PEI),将用重组杆粒转染附着于12孔板上的SF9昆虫细胞。转染后4-7天或当病毒感染引起SF9细胞增大时,收获细胞培养基作为P0病毒原液。每个P0病毒原液用于感染在摇瓶中密度为2x 10^6个细胞/ml的50ml SF9细胞以扩增病毒。使用Cedex细胞计数器(Roche Diagnostics Corporation,Indianapolis,IN,USA)监测培养物。收获时间由平均活细胞大小和细胞活力确定。当平均活细胞大小比未感染细胞大4-7μm,且感染后4-5天内活力范围为50%-70%时,通过离心去除细胞,收集细胞培养上清液并在4℃下黑暗中作为P1病毒原液保存。有时病毒原液通过0.2μm无菌过滤器过滤以保持无菌。

[0169] 为了进一步扩增病毒原液,用250μl P1病毒原液感染500ml的密度为2x10^6个细胞/ml的SF9细胞。使用Cedex细胞计数器监测培养物。当平均活细胞大小比未感染细胞大4-7μm,且感染后4-5天内活力范围为50%-70%时,通过离心去除细胞,收集细胞培养上清液并在4℃下黑暗中作为P2病毒原液保存。有时病毒原液通过0.2μm无菌过滤器过滤以保持无菌。

[0170] 实施例6

[0171] 重组HA蛋白的表达。

[0172] 该实施例证明了在实施例1-4中描述的各种构建体的昆虫细胞中的表达。根据具体构建体,获得不同水平的靶蛋白表达。

[0173] 为了检测重组HA的表达,收集用重组杆状病毒原液(通常是如实施例所述的P1或P2病毒原液)感染的SF9细胞的细胞培养基,并与Ni-NTA树脂(Qiagen,Germantown,MD USA)一起孵育,以通过重组HA的10xHis标签进行亲和捕获。用磷酸盐缓冲盐水(1xPBS)洗涤Ni-NTA树脂后,在存在或不存在还原剂的情况下,用凝胶加样缓冲液煮沸Ni-NTA树脂,以通过SDS-PAGE或Western印迹使用抗His抗体进行分析。在Coomassie®染色的凝胶上的约63kDa的蛋白质条带或抗His Western印迹表明了分泌的重组HA全长蛋白质的表达。

[0174] 如图7和图8所示,如预期检测到了亲本构建体H1 WT的H1 HA蛋白的表达。具有多碱基残基的改变的成熟切割位点的的HA蛋白:H1H5cs、H1R5cs和H1IR5cs与H1WT的表达水平相同。整个H1 HA成熟切割位点被TEV切割位点替换的H1TEV1的表达远低于H1WT。通过在成

熟切割位点保留更多的天然残基,H1TEV2和H1TEV3与H1WT相比表达良好(图7和图8)。并非所有取代都表达相同的水平。通过调节构建体中TEV切割位点的位置,达到了野生型HA的表达水平。

[0175] 图7和图8中还显示,将相同免疫亚显性bNAb表位置于不同免疫显性抗原位点导致了不同的表达水平。在Sa位点(H1TEV2-CR8020Sa4)或Ca2位点(H1H5cs-CR8020Ca)的复合CR8020表位肽的构建体与野生型HA的表达水平相同,而H1TEV2-CR8020Sa3构建体表达不良。因此,异源表位放置的确切位置对于所得HA的表达可能是重要的。免疫亚显性bNAb表位的性质也可能影响所得HA蛋白的表达水平。

[0176] 在几个抗原位点单独放置I1C9表位也显示了所得HA蛋白的不同表达水平(图7和图8)。构建体H1H5cs-I1C9Ca显示出高于构建体H1TEV2-I1C9Ca1的表达。在这两种构建体中被I1C9表位肽替换的抗原位点中的肽序列略有不同。与具有TEV切割位点如MCS的那些构建体类似,取代的位置影响表达水平。

[0177] 在Ca2位点具有M2e肽的构建体,即H1H5cs-M2eCa和H1TEV2-M2eCa,无论成熟切割位点如何,都得到与H1WT一样的表达(图7和图8)。然而,在Sb位点具有M2e肽的构建体H1H5cs-M2eSb表达不良。Sb部位呈螺旋状结构,这可能是正确折叠HA所需的。

[0178] 许多H1 HA构建体的成功表达证明,HA球状头结构域的免疫显性抗原位点可以被来自茎结构域的免疫显性bNAb表位或来自与HA无关的另一种蛋白质的M2e肽替换。这些结果揭示了球状头结构域的可塑性,并验证了使功能性HA在其球状头结构域中呈现异源表位的可行性。总之,实施例1-6说明了构建具有异源表位和/或改变的MCS的HA以及选择具有良好表达的修饰的HA的方法。

[0179] 实施例7

[0180] 重组HA蛋白的纯化

[0181] 该实施例证明了实施例1-4中描述的各种构建体的纯化。

[0182] 在摇瓶中,用体积比为1-4ml P1或P2病毒原液与1升细胞的杆状病毒原液感染50ml至1L密度为2x 10^6个细胞/ml的SF9细胞。使用Cedex细胞计数器监测培养物。收获时间由平均活细胞大小和细胞活力确定。当平均活细胞大小比未感染细胞大4-7μm,且感染后2至3天内活力为约80%时,通过离心去除细胞,收集条件细胞培养基并在4℃下保存。将条件细胞培养基与Ni-NTA树脂通过在旋转式烤架台上在4℃下摇动2-4小时,进行孵育。通过离心收集Ni-NTA树脂并用1xPBS洗涤。将洗涤的Ni-NTA树脂装入重力流动柱中,并进一步用补充有50mM咪唑的1xPBS洗涤。然后通过重力流用1xPBS中的400mM咪唑洗脱Ni-NTA树脂。将洗脱的HA蛋白缓冲液更换为1xPBS并通过超滤浓缩。通过如图8所示的SDS-PAGE分析纯化的HA蛋白。

[0183] 附录

[0184] 流感病毒、类型和宿主的描述

[0185] 流感病毒由病毒的正粘病毒科的三个属组成,它们是甲型、乙型和丙型流感病毒。每个属只有一种病毒,分别是甲型、乙型和丙型流感病毒。甲型、乙型和丙型流感病毒也分别称为甲型流感、乙型流感和丙型流感病毒。流感病毒是包膜的负义单链RNA病毒。病毒基因组在单独的RNA区段中编码。甲型和乙型流感病毒各有编码10种蛋白质的8个RNA区段。丙型流感病毒各有编码9种蛋白质的7个RNA区段。甲型流感病毒的10种特征蛋白是PB2、PB1和

PA聚合酶，血凝素(HA)，核蛋白(NP)，神经氨酸酶(NA)，基质蛋白M1和M2，非结构蛋白NS1和NS2(Webster,R.G.,等人,Microbiol Rev(1992)56:152-179)。丙型流感病毒具有血凝素-酯酶-融合(HEF)蛋白，其结合了HA和NA的功能(Herrler,G.,等人,J.Gen Virol(1988)69:839-846)。甲型、乙型和丙型流感病毒的分类基于其NP和基质蛋白的抗原差异。

[0186] 甲型流感病毒是直径为80-120nm的小颗粒，其由镶嵌有病毒编码的膜蛋白HA、NA和M2的宿主衍生的脂质双层包膜，由M1基质蛋白制成的内壳以及中心的各个RNA区段的病毒基因组的核壳体组成(Webster,R.G.,等人,Microbiol Rev(1992)56:152-179)。RNA区段被多个NP分子松散地包裹。三种病毒聚合酶蛋白(PB1、PB2和PA)的复合物位于核壳体的末端。所有RNA区段都是产生感染性病毒颗粒所必需的。

[0187] RNA区段1是通过凝胶电泳的最慢迁移RNA种类，其编码PB2 RNA聚合酶。RNA片段2通过使用来自相同RNA序列的不同阅读框编码PB1 RNA聚合酶，以及针对PB1-N40和PB1-F2蛋白的两个其他转录物。PB1-N40和PB1-F2蛋白诱导宿主细胞凋亡。RNA区段3编码PA RNA聚合酶，其与PB1和PB2形成RNA依赖性RNA聚合酶复合物。RNA区段4编码HA，即流感病毒粒子的主要表面抗原。每个病毒粒子具有约500个HA分子，其在病毒粒子表面上显示为均匀分布的突起(Ruigrok,R.W.H.,等人,J.Gen Virol(1984)65:799-902;Murti,K.G.和Webster,R.G.,Virol(1986)149:36-43)。由于与宿主受体结合，因此HA决定宿主范围。RNA区段5编码NP核蛋白。NP包裹病毒RNA并被转运至宿主细胞核。NP在感染细胞中大量合成，是流感病毒病毒粒子中第二丰富的蛋白质。它是宿主细胞毒性T细胞免疫应答的主要靶标。RNA区段6编码神经氨酸酶(NA)，其是流感病毒病毒粒子的第二主要表面抗原。NA是一种从糖蛋白或糖脂中切割末端唾液酸以从宿主细胞受体释放病毒颗粒的酶。其功能是病毒传播所必需的。每个病毒粒子在其表面具有约100个NA分子。NA形成四聚体并位于病毒粒子包膜上的离散斑块中(Murti,K.G.和Webster,R.G.,Virol(1986)149:36-43)。RNA区段7是双顺反子并且编码基质蛋白M1和M2二者。M1在病毒粒子包膜下面的病毒粒子核壳体周围形成壳，并且是流感病毒病毒粒子中最丰富的蛋白质。M2由M1的相同转录物的不同剪接形式构成。M2是完整的膜蛋白，并在宿主细胞中产生流感病毒的过程中用作质子通道以控制高尔基体网络的pH值。M2可被替代的拼接变体M42部分替换。需要约3000个基质蛋白分子来制备一个病毒粒子。RNA区段8编码用于病毒复制的两种非结构蛋白NS1和NS2。NS2来自与NS1不同的阅读框架。两种蛋白质在感染细胞中都很丰富，但未并入子代病毒粒子中。

[0188] 在流感病毒的三个属中，甲型流感病毒是最致命的人类病原体并且在人类中引起最严重的疾病。甲型流感病毒基于其表面糖蛋白HA和NA的抗原特性被分类为亚型。已经鉴定了总计18种HA亚型，命名为H1至H18，以及10种NA亚型，命名为N1至N10(Tong,S.,等人,PLoS Pathogens(2013)9:e1003657)。甲型流感亚型的常见命名法来源于HA亚型和NA亚型的不同组合，如H1N1、H3N2和H7N9(Bull World Health Organ(1980)58:585-591)。水禽是亚型H1至H16和N1至N9的天然宿主。最近在美国果蝠中发现了亚型H17N10和H18N10。在自然界中，当宿主共感染甲型流感病毒的两种亚型时，甲型流感病毒的分段基因组允许重新分类RNA区段。来自两种不同亚型的RNA区段可以包装在单个病毒粒子中以产生新的亚型。在这些甲型流感病毒的许多亚型中，只有H1N1、H2N2和H3N2已经开发出在人类中有效传播的能力。这三种亚型是所谓的季节性流感病毒的流行病毒亚型。因此，人群对许多甲型流感亚型的免疫反应是固有的。偶尔其他亚型甲型流感病毒跨越物种并感染人类，导致高死亡率

的流行性爆发。

[0189] 基于它们的系统发育关系,甲型流感HA亚型聚集成两个不同的组。组1包括16种亚型中的10种:H1、H2、H5、H6、H8、H9、H11、H12、H13和H16。组2占剩余6种亚型:H3、H4、H7、H10、H14和H15。

[0190] 乙型流感病毒在抗原上与甲型流感病毒不同。乙型流感病毒与甲型病毒共同传播并引起人类流行病。乙型流感病毒在没有已知动物宿主的情况下稳定适应于人类。与甲型流感病毒的HA的巨大遗传变异不同,在乙型流感病毒中仅报道了一种HA血清型,尽管三种谱系由HA基因的系统发育关系定义。乙型流感病毒似乎不与甲型流感病毒重组,因为没有关于重新分类甲型流感病毒与乙型流感病毒之间的RNA区段的报道。

[0191] 甲型和乙型流感病毒均使用在宿主细胞表面上的末端唾液酸作为细胞受体。两种病毒类型的HA具有相同的结构特征并与唾液酸受体结合以进入宿主细胞。

[0192] 甲型流感病毒和乙型流感病毒的亚型进一步分类为毒株。有许多不同的甲型流感和乙型流感病毒的毒株。每个流感季节都由几种甲型和乙型流感病毒的毒株控制,这些毒株通常与之前流感季节的毒株在遗传上存在差异。在流感季节的不同地理位置或时间分离的一种毒株的病毒(称为分离株)通常具有遗传变化。

[0193] 感染丙型流感病毒通常无症状或引起主要涉及儿童或年轻人的轻微疾病。丙型流感病毒伴随这散发病例和轻微局部爆发。与甲型和乙型流感病毒相比,丙型流感病毒的疾病负担要小得多。大部分人群显示出血清转化,这表明丙型流感病毒在人群中广泛传播。还从动物中分离出了丙型流感病毒。丙型流感病毒的血凝素-酯酶-融合(HEF)蛋白结合了甲型和乙型流感病毒的HA和NA的功能。与甲型和乙型流感病毒的HA不同,丙型流感病毒的HEF使用末端9-O-乙酰基-N-乙酰神经氨酸(9-O-Ac-NeuAc)作为细胞受体(Herrler,G.,等人,EMBO J.(1985)4:1503-1506)。

[0194] 病毒感染过程和生命周期的描述

[0195] 流感病毒通过飞沫或污染物(任何能够携带传染性生物的物体或物质)在人与人之间传播。病毒感染呼吸道的上皮细胞。与细胞表面受体结合后,附着的病毒粒子被宿主细胞内吞至核内体,在所述核内体中低pH引发HA的构象变化,导致其疏水性融合肽插入宿主细胞的液泡膜中,以及引发病毒和液泡膜的融合。融合将病毒粒子的内容物释放至受感染细胞的细胞质中。病毒核壳体迁移至宿主细胞核中,并且它们的相关聚合酶复合物开始mRNA的初级转录,以翻译病毒蛋白。同时,宿主mRNA的翻译被阻断。新合成的病毒RNA被包裹,病毒结构蛋白被合成并转运至宿主细胞表面,在所述宿主细胞表面它们整合至宿主细胞膜中。流感病毒从极化的上皮细胞如支气管上皮细胞的顶端表面发芽进入肺腔,因此通常是亲肺的(Roth,M.G.,等人,PNAS(1979)76:6430-6434;Nayak,D.P.,等人,Virus Res(2009)143:147-161)。已证实了猪与人之间传播流感病毒。个体感染后,免疫系统产生抗流感病毒的抗体。这是身体的主要保护来源。

[0196] 序列表

[0197] SEQ ID NO:1:肽序列:A/California/07/2009的HA的肽序列,猪源甲型流感病毒H1N1毒株(UniProt:C3W5X2)

[0198] MKAILVVLLYTFATANADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSN

KGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRY
SKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPK
GAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSG
YAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDY
HDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQIL
AIYSTVASSLVLVVSLGAISFWMCSNGSLQCRICI

[0199] SEQ ID NO:2:GP67分泌信号的肽序列(GP67ss)

[0200] MVSAIVLYVLLAAAAHSAFA

[0201] SEQ ID NO:3:TEV切割位点、foldon和10xHis标签的肽序列

[0202] GAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0203] SEQ ID NO:4:TEV切割位点的肽序列

[0204] ENLYFQG

[0205] SEQ ID NO:5:foldon的肽序列

[0206] GSGYIPEAPRDGQAYVRKDGEWVLLSTFL

[0207] SEQ ID NO:6:10xHis标签的肽序列

[0208] HHHHHHHHHH

[0209] SEQ ID NO:7:A/California/07/2009的HA的信号肽的肽序列(UniProt:C3W5X2)

[0210] MKAILVVLLYTFATANA

[0211] SEQ ID NO:8:H1 HA成熟切割位点(MCS)的肽序列

[0212] IPSIQSR

[0213] SEQ ID NO:9:GP67ss-H1WT-TEV-foldon-10His(H1WT)的肽序列

[0214] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQ
TPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQ
GSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERT
LDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIY
QGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0215] SEQ ID NO:10:GP67ss-H1WT-TEV-foldon-10His(H1WT)的核苷酸序列
ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA

AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATATCCCTAGCATCCAGAGCAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGG
AGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCCGATC
TGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTT
ACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCT
GGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATG
TGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTC
TACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAGCGAGGAGGC
CAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAAACCTGTATT
TTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTT
CTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0216] SEQ ID NO:11:GP67ss-H1H5cs-TEV-foldon-10His(H1H5cs)的肽序列

MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNI
AGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRYSKKFK
PEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINT
SLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYA
ADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYHD
SNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAEN
LYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0217] SEQ ID NO:12:H5成熟切割位点的肽序列

[0218] QRERRKKR

[0219] SEQ ID NO:13:GP67ss-H1H5cs-TEV-foldon-10His(H1H5cs)的核苷酸序列

[0220]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA

CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTT
TATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATG
CCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAAC
ACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGA
CGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACG
ACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGC
TTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAG
CGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAA
ACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGT
GAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0221] SEQ ID NO:14:GP67ss-H1R5cs-TEV-foldon-10His(H1R5cs)的肽序列

MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNI
AGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRYSKKFK
PEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINT
SLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPRRRRRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADL
KSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYHDSNV
KNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYF
QGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0222] SEQ ID NO:15:GP67ss-H1R5cs-TEV-foldon-10His(H1R5cs)的核苷酸序列

[0223]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATATCCCTAGGAGACGCAGAAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGG
AGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCCGATC
TGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTT
ACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCT

GGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATG
TGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTC
TACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAGCGAGGAGGC
CAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAAACCTGTATT
TTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTT
CTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0224] SEQ ID NO:16:GP67ss-H1IR5cs-TEV-foldon-10His(H1IR5cs)的肽序列

MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNI
AGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRYSKKFK
PEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINT
SLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRRRRRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGY
AADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYH
DSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAE
NLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0225] SEQ ID NO:17:GP67ss-H1IR5cs-TEV-foldon-10His(H1IR5cs)的核苷酸序列

[0226]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATATCCCTAGCATCCAGAGCAGGAGACGCAGAAGAGGCCTGTTTGGAGCCATCGCCGG
CTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGAT
ATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATG
AACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGA
CGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACC
ACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGC
TGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTA
TAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTG
AAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGAC

GGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0227] SEQ ID NO:18:GP67ss-H1TEV1-TEV-foldon-10His(H1TEV1)的肽序列

MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNI
AGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRYSKKFK
PEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINT
SLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADL
KSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYHDSNV
KNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYF
QGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0228] SEQ ID NO:19:GP67ss-H1TEV1-TEV-foldon-10His(H1TEV1)的核苷酸序列

[0229]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATATCGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGG
AGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCCGATC
TGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTT
ACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCT
GGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATG
TGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTC
TACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAGCGAGGAGGC
CAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAAACCTGTATT
TTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTT
CTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0230] SEQ ID NO:20:GP67ss-H1TEV2-TEV-foldon-10His(H1TEV2)的肽序列

MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNI
AGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRYSKKFK

PEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINT
SLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAAD
LKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYHDSN
VKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLY
FQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0231] SEQ ID NO:21:GP67ss-H1TEV2-TEV-foldon-10His(H1TEV2)的核苷酸序列

[0232]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGA
GGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCCG
ATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAG
TTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTT
CCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCA
ATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAG
TTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAGCGAGGA
GGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAAACCTGT
ATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAATGG
GTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0233] SEQ ID NO:22:GP67ss-H1TEV3-TEV-foldon-10His(H1TEV3)的肽序列

MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNI
AGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGSSRYSKKFK
PEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINT
SLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYA
ADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYHD
SNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAEN
LYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0234] SEQ ID NO:23:GP67ss-H1TEV3-TEV-foldon-10His(H1TEV3)的核苷酸序列

[0235]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATATCCCTAGCATCGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTT
TATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATG
CCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAAC
ACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGA
CGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACG
ACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGC
TTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAG
CGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAA
ACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGT
GAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0236] SEQ ID NO:24:GP67ss-H1H5cs-CR8020Ca-TEV-foldon-10His(H1H5cs-CR8020Ca)的肽序列

[0237] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPEGMIDYEGTGQAAGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTC
QTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVDGWYGYHH
QNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLE
NERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLES
TRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0238] SEQ ID NO:25:HA2残基15-19的CR8020表位的肽序列

[0239] EGMID

[0240] SEQ ID NO:26:HA2残基30-36的CR8020表位的肽序列

[0241] EGTGQAA

[0242] SEQ ID NO:27:复合C8020表位的肽序列

[0243] EGMIDYEGTGQAA

[0244] SEQ ID NO:28:GP67ss-H1H5cs-CR8020Ca-TEV-foldon-10His(H1H5cs-CR8020Ca)的核苷酸序列

[0245]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAGGCATGATTGATTACGAAGGCACAGGCCAGGCAGCCGGCAGAATGAATTACTACTGGACCCTGGTGGAA
CCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGC
TGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTA
ATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTG
AGACTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGG
CTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGAT
ATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATG
AACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGA
CGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACC
ACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGC
TGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTA
TAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTG
AAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGAC
GGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAAaagctt

[0246] SEQ ID NO:29:GP67ss-H1TEV2-CR8020Ca-TEV-foldon-10His(H1TEV2-CR8020Ca)的肽序列

[0247] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPEGMIDYEGTGQAAGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTC
QTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQN
EQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENE
RTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTR
IYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0248] SEQ ID NO:30:GP67ss-H1TEV2-CR8020Ca-TEV-foldon-10His(H1TEV2-CR8020Ca)

的核苷酸序列

[0249]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCcTCTGCAcCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCaCaAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAGGCATGATTGATTACGAAGGCACAGGCCAGGCAGCCGGCAGAATGAATTACTACTGGACCCTGGTGGAA
CCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGC
TGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTA
ATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTG
AGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTTTAT
TGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCG
CCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGtGAACAGCGTGATCGAGAAGATGAACACC
CAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGG
CTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACGACA
GCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTT
GAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAGCGA
GGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAAACC
TGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAA
TGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0250] SEQ ID NO:31:GP67ss-H1TEV2-CR8020Sa3-TEV-foldon-10His(H1TEV2-CR8020Sa3)的肽序列

[0251] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVEGMIDYEGTGQAAYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNA
DAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPV
HDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGW
YGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAEL
LVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDG
VKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0252] SEQ ID NO:32:GP67ss-H1TEV2-CR8020Sa3-TEV-foldon-10His(H1TEV2-CR8020Sa3)的核苷酸序列

[0253]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT

GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGGAAGGCATGATTGATTACGAA
GGCACAGGCCAGGCAGCCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTG
GGGAATCCACCACCCTAGCACAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCA
GCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAAT
TACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGC
ATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTC
AGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAG
TATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTT
TGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATG
AGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGC
GTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCT
GAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGA
GAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAG
GAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTA
CGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAA
TCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCT
TACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCA
CTGATAAAaagctt

[0254] SEQ ID NO:33:GP67ss-H1TEV2-CR8020Sa4-TEV-foldon-10His(H1TEV2-CR8020Sa4)的肽序列

[0255] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPEGMIDYEG
TGQAAWPNHDSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQN
ADAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTP
VHDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDG
WYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAE
LLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEID
GVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0256] SEQ ID NO:34:GP67ss-H1TEV2-CR8020Sa4-TEV-foldon-10His(H1TEV2-CR8020Sa4)的核苷酸序列

[0257]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA

TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTGAAGGCATGATTGATTACGAAGGCACAGGCCAGGCAGCCTGGCCTA
ATCACGATTCTAATAAGGGAGTGACAGCCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGG
CTGGTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCT
GTGGGGAATCCACCACCCTAGCACAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGG
GCAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATG
AATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATA
TGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCT
GTCAGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCT
AAGTATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCT
GTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGA
ATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAAC
AGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAA
CCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATG
AGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCC
AAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCAC
CTACGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAA
GAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAG
GCTTACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCA
TCACTGATAAaagctt

[0258] SEQ ID NO:35:GP67ss-H1TEV2-I1C9Ca1-TEV-foldon-10His(H1TEV2-I1C9Ca1)的肽序列

[0259] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPEIAIGIFGAIAGFIEGGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCN
TTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYH
HQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLL
ENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLE
STRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0260] SEQ ID NO:36:I1C9的肽序列

[0261] GIFGAIAGFIEG

[0262] SEQ ID NO:37:GP67ss-H1TEV2-I1C9Ca1-TEV-foldon-10His(H1TEV2-I1C9Ca1)的核苷酸序列

[0263]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA

TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTGGCATTTTCGGCGCTATCGCCGGCTTCATTGAGGGAGGCAGAATGAATTACTACTGGACC
CTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGA
GAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGG
GCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGC
ACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGC
CGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCG
GATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAG
ATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGT
GGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACT
ACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAAC
GGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAA
GTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCG
CTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAA
GACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagct
t

[0264] SEQ ID NO:38:GP67ss-H1TEV2-I1C9Ca2-TEV-foldon-10His(H1TEV2-I1C9Ca2)的肽序列

[0265] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPGIFGAIAGFIEGFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADA
YVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHD
CNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYG
YHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLV
LLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVK
LESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0266] SEQ ID NO:39:GP67ss-H1TEV2-I1C9Ca2-TEV-foldon-10His(H1TEV2-I1C9Ca2)的核苷酸序列

[0267]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT

GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTGGCATTTTCGGCGCTATCGCCGGCTTCATTGAGGGATTTTACAAGAACCTGATCTGGCTGGTGAAG
AAGGGCAACAGCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAAT
CCACCACCCTAGCACAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCA
GATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTAC
TGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGC
TATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACAC
CTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTG
AAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTTTGGAGC
CATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGG
GATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATC
GAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAA
GAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCC
TGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATC
GGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTA
CCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATC
AGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTT
CGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATA
Aaagctt

[0268] SEQ ID NO:40:GP67ss-H1TEV2-I1C9Sa3-TEV-foldon-10His(H1TEV2-I1C9Sa3)的肽序列

[0269] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVGIFGAIAGFIEGYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNAD
AYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVH
DCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWY
GYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELL
VLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGV
KLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0270] SEQ ID NO:41:GP67ss-H1TEV2-I1C9Sa3-TEV-foldon-10His(H1TEV2-I1C9Sa3)的核苷酸序列

[0271]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGGGCATTTTCGGCGCTATCGCC

GGCTTCATTGAGGGATACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGG
AATCCACCACCCTAGCACAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCA
GCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTAC
TACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATT
TGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGA
CACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTAT
GTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTTTGG
AGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGC
AGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTG
ATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAA
CAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAA
CCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAG
ATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGA
CTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCT
ATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTAC
GTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTG
ATAAaagctt

[0272] SEQ ID NO:42:GP67ss-H1TEV2-I1C9Sa4-TEV-foldon-10His(H1TEV2-I1C9Sa4)的肽序列

[0273] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPGIFGAIAG
FIEGWPNHDSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNA
DAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPV
HDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGW
YGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAEL
LVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDG
VKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0274] SEQ ID NO:43:GP67ss-H1TEV2-I1C9Sa4-TEV-foldon-10His(H1TEV2-I1C9Sa4)的核苷酸序列

[0275]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTGGCATTTTCGGCGCTATCGCCGGCTTCATTGAGGGATGGCCTAATC
ACGATTCTAATAAGGGAGTGACAGCCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTG
GTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTG
GGGAATCCACCACCCTAGCACAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCA

GCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAAT
TACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGC
ATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTC
AGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAG
TATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTT
TGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATG
AGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGC
GTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCT
GAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGA
GAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAG
GAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTA
CGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAA
TCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCT
TACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCA
CTGATAAaagctt

[0276] SEQ ID NO:44:GP67ss-H1H5cs-I1C9Sb-TEV-foldon-10His(H1H5cs-I1C9Sb)的肽序列

[0277] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSGIFGAIAGFIEGDAYVFVGS
SRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQ
TPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVDGWYGYHHQ
NEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLEN
ERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLEST
RIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0278] SEQ ID NO:45:GP67ss-H1H5cs-I1C9Sb-TEV-foldon-10His(H1H5cs-I1C9Sb)的核苷酸序列

[0279]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCGGCAT
TTTCGGCGCTATCGCCGGCTTCATtGAGGGAGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG

CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTT
TATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATG
CCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAAC
ACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGA
CGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACG
ACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGC
TTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAG
CGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAA
ACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGT
GAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0280] SEQ ID NO:46:GP67ss-H1H5cs-I1C9Ca-TEV-foldon-10His(H1H5cs-I1C9Ca)的肽序列

[0281] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPGIFGAIAGFIEGGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQ
TPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVDGWYGYHHQ
NEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLEN
ERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLEST
RIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0282] SEQ ID NO:47:GP67ss-H1H5cs-I1C9Ca-TEV-foldon-10His(H1H5cs-I1C9Ca)的核苷酸序列

[0283]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTGGCATTTTCGGCGCTATCGCCGGCTTCATTGAGGGAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCT
GGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGG
CTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCATTAATA
CCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACCAAGCTGAGA
CTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTT

TATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGcTATCACCACCAGAATGAGCAGGGATCCGGATATG
CCGCCGATCtGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAAC
ACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGA
CGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACCACG
ACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGC
TTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATAG
CGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAA
ACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGT
GAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0284] SEQ ID NO:48:GP67ss-H1H5cs-FI6Sab-TEV-foldon-10His(H1H5cs-FI6Sab)的肽序列

[0285] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVRKKRGLFGAIAGFIEYINDKGKEVLVLWGIHHPSKESTQKAIDGVTNKVNS
DAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPV
HDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVD
GWYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNA
ELLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEI
DGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0286] SEQ ID NO:49:FI6表位的成熟切割位点的肽序列

[0287] RKKRGLFGAIAGFIE

[0288] SEQ ID NO:50:FI6表位的卷曲螺旋肽的肽序列

[0289] KESTQKAIDGVTNKVNS

[0290] SEQ ID NO:51:GP67ss-H1H5cs-FI6Sab-TEV-foldon-10His(H1H5cs-FI6Sab)的核苷酸序列

[0291]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAGAAAGAAGAGAGGCCTGTTT
GGAGCCATCGCCGGCTTTATTGAGTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCC
TAGCAAGGAGTCTACACAGAAGGCCATTGATGGCGTTACAAATAAGGTCAATtCTGATGCCTATGTGTTTGTGGGCA
GCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAAT
TACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGC
ATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTC
AGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAG

TATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGG
CCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACC
AGAATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTG
AACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGA
GAACCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGA
ATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAAT
GCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGG
CACCTACGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCA
CAAGAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGT
CAGGCTTACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCA
TCATCACTGATAAaagctt

[0292] SEQ ID NO:52:GP67ss-H1WT-FI6Sab-TEV-foldon-10His(H1WT-FI6Sab)的肽序列

[0293] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVRKKRGLFGAIAGFIEYINDKGKEVLVLWGIHHPSKESTQKAIDGVTNKVNS
DAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPV
HDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAGFIEGGWTGMVDGWY
GYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELL
VLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGV
KLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0294] SEQ ID NO:53:GP67ss-H1WT-FI6Sab-TEV-foldon-10His(H1WT-FI6Sab)的核苷酸序列

[0295]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAGAAAGAAGAGAGGCCTGTTT
GGAGCCATCGCCGGCTTTATTGAGTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCC
TAGCAAGGAGTCTACACAGAAGGCCATTGATGGCGTTACAAATAAGGTCAATTCTGATGCCTATGTGTTTGTGGGCA
GCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAAT
TACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGC
ATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTC
AGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAG
TATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATATCCCTAGCATCCAGAGCAGAGGCCTGTTTGG
AGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGC
AGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTG

ATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAA
CAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAA
CCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAG
ATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGA
CTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCT
ATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTAC
GTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTG
ATAAaagctt

[0296] SEQ ID NO:54:GP67ss-H1TEV1-FI6Sa-TEV-foldon-10His(H1TEV1-FI6Sa)的肽序列

[0297] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVRKKRGLFGAIAGFIEGYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYV
FVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCN
TTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHH
QNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLE
NERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLES
TRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0298] SEQ ID NO:55:GP67ss-H1TEV1-FI6Sa-TEV-foldon-10His(H1TEV1-FI6Sa)的核苷酸序列

[0299]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAGAAAGAAGAGAGGCCTGTTC
GGCGCTATCGCCGGCTTCATTGAGGGATACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCA
CCCTAGCACAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACA
GCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACC
CTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGA
GAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGG
GCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGC
ACCAAGCTGAGACTGGCTACCGGTCTGAGAAATATCGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGG
CTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGAT
ATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGAACAGCGTGATCGAGAAGATG
AACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGA
CGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCCTGGACTACC

ACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGC
TGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTA
TAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTG
AAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGAC
GGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0300] SEQ ID NO:56:GP67ss-H1H5cs-M2eCa2-TEV-foldon-10His(H1H5cs-M2eCa)的肽序列

[0301] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPSLLTEVETPTRNGWECKCSDSGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTP
VHDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMV
DGWYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYN
AELLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREE
IDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0302] SEQ ID NO:57:M2e肽的肽序列

[0303] SLLTEVETPTRNGWECKCSDS

[0304] SEQ ID NO:58:GP67ss-H1H5cs-M2eCa2-TEV-foldon-10His(H1H5cs-M2eCa)的核苷酸序列

[0305]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTTCCCTGCTGACCGAGGTGGAGACCCCCACCAGGAACGGCTGGGAGTGCAAGTGCTCCGACTCCGGCAGAATG
AATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATA
TGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCT
GTCAGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCT
AAGTATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAG
AGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACC
ACCAGAATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAG
GTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAAT
CGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGG
AGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAAC

AATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAA
CGGCACCTACGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAA
GCACAAGAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGAC
GGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCA
CCATCATCACTGATAAaagctt

[0306] SEQ ID NO:59:GP67ss-H1H5cs-M2eSb-TEV-foldon-10His(H1H5cs-M2eSb)的肽序列

[0307] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSSLLTEVETPTRNGWECKCSD
SDAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTP
VHDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMV
DGWYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYN
AELLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREE
IDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0308] SEQ ID NO:60:GP67ss-H1H5cs-M2eSb-TEV-foldon-10His(H1H5cs-M2eSb)的核苷酸序列

[0309]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCTCCCT
GCTGACCGAGGTGGAGACCCCCACCAGGAACGGCTGGGAGTGCAAGTGCTCCGACTCCGATGCCTATGTGTTTGTGG
GCAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATG
AATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATA
TGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCT
GTCAGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCT
AAGTATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAG
AGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACC
ACCAGAATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAG
GTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAAT
CGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGG
AGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAAC
AATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAA
CGGCACCTACGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAA

GCACAAGAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGAC
GGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCA
CCATCATCACTGATAAaagctt

[0310] SEQ ID NO:61:GP67ss-H1TEV2-M2eCa2-TEV-foldon-10His(H1TEV2-M2eCa)的肽序列

[0311] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAP
LHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNH
DSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQNADAYVFVGS
SRYSKKFKPSLLTEVETPTRNGWECKCSDSGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTP
VHDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDG
WYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAE
LLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEID
GVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0312] SEQ ID NO:62:GP67ss-H1TEV2-M2eCa2-TEV-foldon-10His(H1TEV2-M2eCa)的核苷酸序列

[0313]

ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGATACACTGTGTATT
GGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAA
TCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTGCCCCTCTGCACCTGGGCAAATGTAATA
TTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACC
CCTAGCAGCGACAATGGCACCTGTTACCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGT
GTCCAGCTTCGAGAGATTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAG
CCGCCTGTCCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGCTACCCT
AAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGCTGTGGGGAATCCACCACCCTAGCACAAG
CGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCA
AGCCTTCCCTGCTGACCGAGGTGGAGACCCCCACCAGGAACGGCTGGGAGTGCAAGTGCTCCGACTCCGGCAGAATG
AATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTAGATA
TGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCT
GTCAGACACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCT
AAGTATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGGCCT
GTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAGA
ATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGtGAAC
AGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAA
CCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGCTCCTGGAGAATG
AGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCC
AAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCAC
CTACGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAA
GAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAG
GCTTACGTTCGTAAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTCACCATCATCACCACCATCACCATCA

TCACTGATAAaagctt

序列表

<110> CG探索公司(CG Discovery, Inc.)

LUO, Chun

<120> 具有异源表位和/或改变的成熟切割位点的流感血凝素的组合物及其使用方法

<130> 758252000140

<140> 尚未转让

<141> 同时与此一起

<150> US 62/290,894

<151> 2016-02-03

<160> 62

<170> 用于Windows 4.0版的FastSEQ

<210> 1

<211> 566

<212> PRT

<213> 甲型流感病毒

<400> 1

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

Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu Val Leu Val
180 185 190
Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu
195 200 205
Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser
210 215 220
Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln
225 230 235 240
Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys
245 250 255
Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe
260 265 270
Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro
275 280 285
Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn
290 295 300
Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys
305 310 315 320
Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg
325 330 335
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
340 345 350
Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
355 360 365
His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser
370 375 380
Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile
385 390 395 400
Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His
405 410 415
Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe
420 425 430
Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn
435 440 445
Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu
450 455 460
Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly
465 470 475 480
Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val

485 490 495
Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu
500 505 510
Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr
515 520 525
Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Val
530 535 540
Val Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu
545 550 555 560
Gln Cys Arg Ile Cys Ile
565

<210> 2
<211> 20
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 2
Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His
1 5 10 15
Ser Ala Phe Ala
20

<210> 3
<211> 49
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 3
Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu
1 5 10 15
Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val
20 25 30
Leu Leu Ser Thr Phe Leu Gly His His His His His His His His His
35 40 45
His

<210> 4
<211> 7
<212> PRT

<213> 人工序列
<220>
<223> 合成的
<400> 4
Glu Asn Leu Tyr Phe Gln Gly
1 5
<210> 5
<211> 29
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 5
Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
1 5 10 15
Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu
20 25
<210> 6
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 6
His His His His His His His His His His
1 5 10
<210> 7
<211> 17
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 7
Met Lys Ala Ile Leu Val Val Leu Leu Tyr Thr Phe Ala Thr Ala Asn
1 5 10 15
Ala
<210> 8
<211> 7
<212> PRT

<213> 人工序列
<220>
<223> 合成的
<400> 8
Ile Pro Ser Ile Gln Ser Arg
1 5
<210> 9
<211> 581
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 9
Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His
1 5 10 15
Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser
20 25 30
Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His
35 40 45
Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
50 55 60
Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
65 70 75 80
Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95
Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140
Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160
Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175
Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190
Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205

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

515 520 525

Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly

530 535 540

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

545 550 555 560

Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His His

565 570 575

His His His His His

580

<210> 10

<211> 1757

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 10

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420

caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540

agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600

accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660

ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720

tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780

tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840

atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900

agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960

tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020

atatccctag catccagagc agaggcctgt ttggagccat cgccggcttt attgagggag 1080

gatggaccgg aatggtggat ggctggtacg gctatcacca ccagaatgag cagggatccg 1140

gatatgccgc cgatctgaag tctacacaga acgccatcga cgagatcaca aacaaggtga 1200

acagcgtgat cgagaagatg aacacccagt ttacagctgt gggcaaggag ttcaaccacc 1260

tggagaagag aatcgagaac ctgaacaaga aagtggacga cggcttcctg gatatttgga 1320

cctacaatgc cgagctgctc gtgctcctgg agaatgagag aaccctggac taccacgaca 1380

gcaatgtgaa gaacctgtac gagaaggtga gaagccagct gaagaacaat gccaaggaga 1440
tcggcaacgg ctgctttgag ttctaccaca agtgtgacaa cacctgtatg gagtctgtga 1500
agaacggcac ctacgactac cctaagtata gcgaggaggc caagctgaat agagaggaga 1560
tcgacggcgt gaaactggaa agcacaagaa tctatcaggg cgctgaaaac ctgtattttc 1620
agggcggttc tggttacatc ccggaagctc cgcgtgacgg tcaggcttac gttcgtaaag 1680
acggtgaatg ggttctgctg tctaccttcc tgggtcacca tcatcaccac catcaccatc 1740
atcactgata aaagctt 1757

<210> 11
<211> 584
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 11

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

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val

225 230 235 240

Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro

245 250 255

Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg

260 265 270

Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser

275 280 285

Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly

290 295 300

Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile

305 310 315 320

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr

325 330 335

Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly Leu

340 345 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val

355 360 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr

370 375 380

Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn

385 390 395 400

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

405 410 415

Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys

420 425 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu

435 440 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn

450 455 460

Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala

465 470 475 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn

485 490 495

Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr

500 505 510

Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu
515 520 525
Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly
530 535 540
Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
545 550 555 560
Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His
565 570 575
His His His His His His His His
580

<210> 12
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 12
Gln Arg Glu Arg Arg Lys Lys Arg
1 5

<210> 13
<211> 1766
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 13
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780

tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atagccctca gagggagaga cgcaagaaga gaggcctgtt tggagccatc gccggcttta 1080
ttgagggagg atggaccgga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140
agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcacaa 1200
acaaggtgaa cagcgtgatc gagaagatga acacccagtt tacagctgtg ggcaaggagt 1260
tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320
atatttggac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
accacgacag caatgtgaag aacctgtacg agaaggtgag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680
ttcgtaaaga cggtgaatgg gttctgctgt ctaccttcct gggtcaccat catcaccacc 1740
atcaccatca tcactgataa aagctt 1766

<210> 14
<211> 581
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 14

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

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val

225 230 235 240

Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro

245 250 255

Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg

260 265 270

Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser

275 280 285

Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly

290 295 300

Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile

305 310 315 320

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr

325 330 335

Gly Leu Arg Asn Ile Pro Arg Arg Arg Arg Arg Gly Leu Phe Gly Ala

340 345 350

Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp

355 360 365

Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp

370 375 380

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

385 390 395 400

Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu

405 410 415

Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp

420 425 430

Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu
435 440 445
Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn
450 455 460
Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile
465 470 475 480
Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met
485 490 495
Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu
500 505 510
Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr
515 520 525
Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly
530 535 540
Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp
545 550 555 560
Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His His
565 570 575
His His His His His
580

<210> 15
<211> 1757
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 15

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720

tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atatccctag gagacgcaga agaggcctgt ttggagccat cgccggcttt attgagggag 1080
gatggaccgg aatggtggat ggctggtacg gctatcacca ccagaatgag cagggatccg 1140
gatatgccgc cgatctgaag tctacacaga acgccatcga cgagatcaca aacaaggtga 1200
acagcgtgat cgagaagatg aacacccagt ttacagctgt gggcaaggag ttcaaccacc 1260
tggagaagag aatcgagaac ctgaacaaga aagtggacga cggcttcctg gatatttgga 1320
cctacaatgc cgagctgctc gtgctcctgg agaatgagag aaccctggac taccacgaca 1380
gcaatgtgaa gaacctgtac gagaaggtga gaagccagct gaagaacaat gccaaggaga 1440
tcggcaacgg ctgctttgag ttctaccaca agtgtgacaa cacctgtatg gagtctgtga 1500
agaacggcac ctacgactac cctaagtata gcgaggaggc caagctgaat agagaggaga 1560
tcgacggcgt gaaactggaa agcacaagaa tctatcaggg cgctgaaaac ctgtattttc 1620
agggcggttc tggttacatc ccggaagctc cgcgtgacgg tcaggcttac gttcgtaaag 1680
acggtgaatg ggttctgctg tctaccttcc tgggtcacca tcatcaccac catcaccatc 1740
atcactgata aaagctt 1757

<210> 16
<211> 585
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 16

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

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

420 425 430

Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
435 440 445

Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
450 455 460

Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
465 470 475 480

Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
485 490 495

Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
500 505 510

Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
515 520 525

Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln
530 535 540

Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
545 550 555 560

Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His
565 570 575

His His His His His His His His His His
580 585

<210> 17

<211> 1769

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 17

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420

caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540

agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600

accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660

ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atatccctag catccagagc aggagacgca gaagaggcct gtttggagcc atcgccggct 1080
ttattgaggg aggatggacc ggaatggtgg atggctggta cggctatcac caccagaatg 1140
agcagggatc cggatatgcc gccgatctga agtctacaca gaacgccatc gacgagatca 1200
caaacaaggt gaacagcgtg atcgagaaga tgaacaccca gtttacagct gtgggcaagg 1260
agttcaacca cctggagaag agaatcgaga acctgaacaa gaaagtggac gacggcttcc 1320
tggatatttg gacctacaat gccgagctgc tcgtgctcct ggagaatgag agaaccctgg 1380
actaccacga cagcaatgtg aagaacctgt acgagaaggt gagaagccag ctgaagaaca 1440
atgccaagga gatcggcaac ggctgctttg agttctacca caagtgtgac aacacctgta 1500
tggagtctgt gaagaacggc acctacgact accctaagta tagcgaggag gccaagctga 1560
atagagagga gatcgacggc gtgaaactgg aaagcacaag aatctatcag ggcgctgaaa 1620
acctgtattt tcagggcggt tctggttaca tcccggaagc tccgcgtgac ggtcaggctt 1680
acgttcgtaa agacggtgaa tgggttctgc tgtctacctt cctgggtcac catcatcacc 1740
accatcacca tcatcactga taaaagctt 1769

<210> 18
<211> 581
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 18

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

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val

225 230 235 240

Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro

245 250 255

Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg

260 265 270

Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser

275 280 285

Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly

290 295 300

Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile

305 310 315 320

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr

325 330 335

Gly Leu Arg Asn Ile Glu Asn Leu Tyr Phe Gln Gly Leu Phe Gly Ala

340 345 350

Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp

355 360 365

Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp

370 375 380

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

385 390 395 400

Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu

405 410 415

Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp
420 425 430
Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu
435 440 445
Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn
450 455 460
Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile
465 470 475 480
Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met
485 490 495
Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu
500 505 510
Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr
515 520 525
Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly
530 535 540
Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp
545 550 555 560
Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His His
565 570 575
His His His His His
580

<210> 19
<211> 1757
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 19
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600

accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atatcgaaaa cctgtatttt caaggcctgt ttggagccat cgccggcttt attgagggag 1080
gatggaccgg aatggtggat ggctggtacg gctatcacca ccagaatgag cagggatccg 1140
gatatgccgc cgatctgaag tctacacaga acgccatcga cgagatcaca aacaaggtga 1200
acagcgtgat cgagaagatg aacacccagt ttacagctgt gggcaaggag ttcaaccacc 1260
tggagaagag aatcgagaac ctgaacaaga aagtggacga cggcttcctg gatatttgga 1320
cctacaatgc cgagctgctc gtgctcctgg agaatgagag aaccctggac taccacgaca 1380
gcaatgtgaa gaacctgtac gagaaggtga gaagccagct gaagaacaat gccaaggaga 1440
tcggcaacgg ctgctttgag ttctaccaca agtgtgacaa cacctgtatg gagtctgtga 1500
agaacggcac ctacgactac cctaagtata gcgaggaggc caagctgaat agagaggaga 1560
tcgacggcgt gaaactggaa agcacaagaa tctatcaggg cgctgaaaac ctgtattttc 1620
agggcggttc tggttacatc ccggaagctc cgcgtgacgg tcaggcttac gttcgtaaag 1680
acggtgaatg ggttctgctg tctaccttcc tgggtcacca tcatcaccac catcaccatc 1740
atcactgata aaagctt 1757

<210> 20
<211> 582
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 20

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

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

405 410 415

Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val
420 425 430

Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val
435 440 445

Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys
450 455 460

Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu
465 470 475 480

Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys
485 490 495

Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu
500 505 510

Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser
515 520 525

Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser
530 535 540

Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys
545 550 555 560

Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His
565 570 575

His His His His His His
580

<210> 21

<211> 1760

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 21

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420

caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540

agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atagccctga aaacctgtat tttcaaggcc tgtttggagc catcgccggc tttattgagg 1080
gaggatggac cggaatggtg gatggctggt acggctatca ccaccagaat gagcagggat 1140
ccggatatgc cgccgatctg aagtctacac agaacgccat cgacgagatc acaaacaagg 1200
tgaacagcgt gatcgagaag atgaacaccc agtttacagc tgtgggcaag gagttcaacc 1260
acctggagaa gagaatcgag aacctgaaca agaaagtgga cgacggcttc ctggatattt 1320
ggacctacaa tgccgagctg ctcgtgctcc tggagaatga gagaaccctg gactaccacg 1380
acagcaatgt gaagaacctg tacgagaagg tgagaagcca gctgaagaac aatgccaagg 1440
agatcggcaa cggctgcttt gagttctacc acaagtgtga caacacctgt atggagtctg 1500
tgaagaacgg cacctacgac taccctaagt atagcgagga ggccaagctg aatagagagg 1560
agatcgacgg cgtgaaactg gaaagcacaa gaatctatca gggcgctgaa aacctgtatt 1620
ttcagggcgg ttctggttac atcccggaag ctccgcgtga cggtcaggct tacgttcgta 1680
aagacggtga atgggttctg ctgtctacct tcctgggtca ccatcatcac caccatcacc 1740
atcatcactg ataaaagctt 1760

<210> 22
<211> 584
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 22

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

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
225 230 235 240

Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
245 250 255

Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
260 265 270

Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser
275 280 285

Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
290 295 300

Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
305 310 315 320

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
325 330 335

Gly Leu Arg Asn Ile Pro Ser Ile Glu Asn Leu Tyr Phe Gln Gly Leu
340 345 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
355 360 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
370 375 380

Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn
385 390 395 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
405 410 415
Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys
420 425 430
Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
435 440 445
Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn
450 455 460
Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala
465 470 475 480
Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn
485 490 495
Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
500 505 510
Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu
515 520 525
Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly
530 535 540
Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
545 550 555 560
Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His
565 570 575
His His His His His His His His
580

<210> 23
<211> 1766
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 23
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atatccctag catcgaaaac ctgtattttc aaggcctgtt tggagccatc gccggcttta 1080
ttgagggagg atggaccgga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140
agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcacaa 1200
acaaggtgaa cagcgtgatc gagaagatga acacccagtt tacagctgtg ggcaaggagt 1260
tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320
atatttggac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
accacgacag caatgtgaag aacctgtacg agaaggtgag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680
ttcgtaaaga cggtgaatgg gttctgctgt ctaccttcct gggtcaccat catcaccacc 1740
atcaccatca tcactgataa aagctt 1766

<210> 24
<211> 585
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 24

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

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95
Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140
Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160
Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175
Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190
Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205
Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220
Arg Tyr Ser Lys Lys Phe Lys Pro Glu Gly Met Ile Asp Tyr Glu Gly
225 230 235 240
Thr Gly Gln Ala Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu
245 250 255
Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro
260 265 270
Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile
275 280 285
Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys
290 295 300
Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr
305 310 315 320
Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala
325 330 335
Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly
340 345 350
Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met
355 360 365
Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly
370 375 380
Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr

385 390 395 400
Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala
405 410 415
Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
420 425 430
Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
435 440 445
Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
450 455 460
Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
465 470 475 480
Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
485 490 495
Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
500 505 510
Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
515 520 525
Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln
530 535 540
Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
545 550 555 560
Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His
565 570 575
His His His His His His His His His His
580 585

<210> 25
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 25
Glu Gly Met Ile Asp
1 5

<210> 26
<211> 7
<212> PRT
<213> 人工序列
<220>

<223> 合成的
<400> 26
Glu Gly Thr Gly Gln Ala Ala
1 5
<210> 27
<211> 13
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 27
Glu Gly Met Ile Asp Tyr Glu Gly Thr Gly Gln Ala Ala
1 5 10
<210> 28
<211> 1770
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 28
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aggcatgatt gattacgaag 720
gcacaggcca ggcagccggc agaatgaatt actactggac cctggtggaa cctggcgata 780
agatcacatt tgaggccacc ggaaatctgg tggtgcctag atatgcattt gctatggaga 840
gaaatgctgg ctctggcatc attatctctg ataccctgt gcacgactgt aataccacct 900
gtcagacacc taagggcgcc attaatacca gcctgccctt ccagaatatt caccctatca 960
ccatcggcaa gtgtcctaag tatgtgaaga gcaccaagct gagactggct accggtctga 1020
gaaatagccc tcagagggag agacgcaaga agagaggcct gtttggagcc atcgccggct 1080
ttattgaggg aggatggacc ggaatggtgg atggctggta cggctatcac caccagaatg 1140

agcagggatc cggatatgcc gccgatctga agtctacaca gaacgccatc gacgagatca 1200
caaacaaggt gaacagcgtg atcgagaaga tgaacaccca gtttacagct gtgggcaagg 1260
agttcaacca cctggagaag agaatcgaga acctgaacaa gaaagtggac gacggcttcc 1320
tggatatttg gacctacaat gccgagctgc tcgtgctcct ggagaatgag agaaccctgg 1380
actaccacga cagcaatgtg aagaacctgt acgagaaggt gagaagccag ctgaagaaca 1440
atgccaagga gatcggcaac ggctgctttg agttctacca caagtgtgac aacacctgta 1500
tggagtctgt gaagaacggc acctacgact accctaagta tagcgaggag gccaagctga 1560
atagagagga gatcgacggc gtgaaactgg aaagcacaag aatctatcag ggcgctgaaa 1620
acctgtattt tcagggcggt tctggttaca tcccggaagc tccgcgtgac ggtcaggctt 1680
acgttcgtaa agacggtgaa tgggttctgc tgtctacctt cctgggtcac catcatcacc 1740
accatcacca tcatcactga taaaaagctt 1770

<210> 29
<211> 583
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 29

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

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Gly Met Ile Asp Tyr Glu Gly

225 230 235 240

Thr Gly Gln Ala Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu

245 250 255

Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro

260 265 270

Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile

275 280 285

Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys

290 295 300

Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr

305 310 315 320

Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala

325 330 335

Thr Gly Leu Arg Asn Ser Pro Glu Asn Leu Tyr Phe Gln Gly Leu Phe

340 345 350

Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp

355 360 365

Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala

370 375 380

Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys

385 390 395 400

Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly

405 410 415

Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys

420 425 430

Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu

435 440 445

Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val

450 455 460

Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys

465 470 475 480

Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr
485 490 495
Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser
500 505 510
Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu
515 520 525
Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly
530 535 540
Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg
545 550 555 560
Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His
565 570 575
His His His His His His His
580

<210> 30
<211> 1763
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 30

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aggcatgatt gattacgaag 720
gcacaggcca ggcagccggc agaatgaatt actactggac cctggtggaa cctggcgata 780
agatcacatt tgaggccacc ggaaatctgg tggtgcctag atatgcattt gctatggaga 840
gaaatgctgg ctctggcatc attatctctg ataccctgt gcacgactgt aataccacct 900
gtcagacacc taagggcgcc attaatacca gcctgcccct ccagaatatt cacctatca 960
ccatcggcaa gtgtcctaag tatgtgaaga gcaccaagct gagactggct accggtctga 1020
gaaatagccc tgaaaacctg tattttcaag gcctgtttgg agccatcgcc ggctttattg 1080

agggaggatg gaccggaatg gtggatggct ggtacggcta tcaccaccag aatgagcagg 1140
gatccggata tgccgccgat ctgaagtcta cacagaacgc catcgacgag atcacaaaca 1200
aggtgaacag cgtgatcgag aagatgaaca cccagtttac agctgtgggc aaggagttca 1260
accacctgga gaagagaatc gagaacctga acaagaaagt ggacgacggc ttcctggata 1320
tttggaccta caatgccgag ctgctcgtgc tcctggagaa tgagagaacc ctggactacc 1380
acgacagcaa tgtgaagaac ctgtacgaga aggtgagaag ccagctgaag aacaatgcca 1440
aggagatcgg caacggctgc tttgagttct accacaagtg tgacaacacc tgtatggagt 1500
ctgtgaagaa cggcacctac gactacccta agtatagcga ggaggccaag ctgaatagag 1560
aggagatcga cggcgtgaaa ctggaaagca caagaatcta tcagggcgct gaaaacctgt 1620
attttcaggg cggttctggt tacatcccgg aagctccgcg tgacggtcag gcttacgttc 1680
gtaaagacgg tgaatgggtt ctgctgtcta ccttcctggg tcaccatcat caccaccatc 1740
accatcatca ctgataaaag ctt 1763

<210> 31
<211> 590
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 31

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

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

465 470 475 480
Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe
485 490 495
Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr
500 505 510
Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu
515 520 525
Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu
530 535 540
Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg
545 550 555 560
Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser
565 570 575
Thr Phe Leu Gly His His His His His His His His His His
580 585 590

<210> 32
<211> 1785
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 32

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgga aggcatgatt gattacgaag 540
gcacaggcca ggcagcctac cctaagctgt ctaagagcta catcaacgac aagggcaaag 600
aagtgctggt gctgtgggga atccaccacc ctagcacaag cgccgatcag cagagcctgt 660
accagaatgc cgatgcctat gtgtttgtgg gcagcagcag atacagcaaa aagttcaagc 720
ctgaaattgc cattagaccc aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaaatctg gtggtgccta 840
gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gataccctg 900
tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgccct 960
tccagaatat tcaccctatc accatcggca agtgtcctaa gtatgtgaag agcaccaagc 1020

tgagactggc taccggtctg agaaatagcc ctgaaaacct gtattttcaa ggcctgtttg 1080
gagccatcgc cggctttatt gagggaggat ggaccggaat ggtggatggc tggtacggct 1140
atcaccacca gaatgagcag ggatccggat atgccgccga tctgaagtct acacagaacg 1200
ccatcgacga gatcacaaac aaggtgaaca gcgtgatcga gaagatgaac acccagttta 1260
cagctgtggg caaggagttc aaccacctgg agaagagaat cgagaacctg aacaagaaag 1320
tggacgacgg cttcctggat atttggacct acaatgccga gctgctcgtg ctcctggaga 1380
atgagagaac cctggactac cacgacagca atgtgaagaa cctgtacgag aaggtgagaa 1440
gccagctgaa gaacaatgcc aaggagatcg gcaacggctg ctttgagttc taccacaagt 1500
gtgacaacac ctgtatggag tctgtgaaga acggcaccta cgactaccct aagtatagcg 1560
aggaggccaa gctgaataga gaggagatcg acggcgtgaa actggaaagc acaagaatct 1620
atcagggcgc tgaaaacctg tattttcagg gcggttctgg ttacatcccg gaagctccgc 1680
gtgacggtca ggcttacgtt cgtaaagacg gtgaatgggt tctgctgtct accttcctgg 1740
gtcaccatca tcaccaccat caccatcatc actgataaaa agctt 1785

<210> 33
<211> 591
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 33

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

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

Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg
465 470 475 480
Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu
485 490 495
Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly
500 505 510
Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu
515 520 525
Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala
530 535 540
Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro
545 550 555 560
Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
565 570 575
Ser Thr Phe Leu Gly His His His His His His His His His His
580 585 590

<210> 34
<211> 1787
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 34

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctgaag 420
gcatgattga ttacgaaggc acaggccagg cagcctggcc taatcacgat tctaataagg 480
gagtgacagc cgcctgtcct catgccggag ccaagtcctt ttacaagaac ctgatctggc 540
tggtgaagaa gggcaacagc taccctaagc tgtctaagag ctacatcaac gacaagggca 600
aagaagtgct ggtgctgtgg ggaatccacc accctagcac aagcgccgat cagcagagcc 660
tgtaccagaa tgccgatgcc tatgtgtttg tgggcagcag cagatacagc aaaaagttca 720
agcctgaaat tgccattaga cccaaagtga gagatcagga aggcagaatg aattactact 780
ggaccctggt ggaacctggc gataagatca catttgaggc caccggaaat ctggtggtgc 840
ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900
ctgtgcacga ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960

ccttccagaa tattcaccct atcaccatcg gcaagtgtcc taagtatgtg aagagcacca 1020
agctgagact ggctaccggt ctgagaaata gccctgaaaa cctgtatttt caaggcctgt 1080
ttggagccat cgccggcttt attgagggag gatggaccgg aatggtggat ggctggtacg 1140
gctatcacca ccagaatgag cagggatccg gatatgccgc cgatctgaag tctacacaga 1200
acgccatcga cgagatcaca aacaaggtga acagcgtgat cgagaagatg aacacccagt 1260
ttacagctgt gggcaaggag ttcaaccacc tggagaagag aatcgagaac ctgaacaaga 1320
aagtggacga cggcttcctg gatatttgga cctacaatgc cgagctgctc gtgctcctgg 1380
agaatgagag aaccctggac taccacgaca gcaatgtgaa gaacctgtac gagaaggtga 1440
gaagccagct gaagaacaat gccaaggaga tcggcaacgg ctgctttgag ttctaccaca 1500
agtgtgacaa cacctgtatg gagtctgtga agaacggcac ctacgactac cctaagtata 1560
gcgaggaggc caagctgaat agagaggaga tcgacggcgt gaaactggaa agcacaagaa 1620
tctatcaggg cgctgaaaac ctgtattttc agggcggttc tggttacatc ccggaagctc 1680
cgcgtgacgg tcaggcttac gttcgtaaag acggtgaatg ggttctgctg tctaccttcc 1740
tgggtcacca tcatcaccac catcaccatc atcactgata aaagctt 1787

<210> 35
<211> 586
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 35

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

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

450 455 460

Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn
465 470 475 480
Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys
485 490 495
Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro
500 505 510
Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val
515 520 525
Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe
530 535 540
Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala
545 550 555 560
Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly
565 570 575
His His His His His His His His His His
580 585

<210> 36
<211> 12
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 36
Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
1 5 10

<210> 37
<211> 1772
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 37
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt ggcattttcg 720
gcgctatcgc cggcttcatt gagggaggca gaatgaatta ctactggacc ctggtggaac 780
ctggcgataa gatcacattt gaggccaccg gaaatctggt ggtgcctaga tatgcatttg 840
ctatggagag aaatgctggc tctggcatca ttatctctga tacccctgtg cacgactgta 900
ataccacctg tcagacacct aagggcgcca ttaataccag cctgcccttc cagaatattc 960
accctatcac catcggcaag tgtcctaagt atgtgaagag caccaagctg agactggcta 1020
ccggtctgag aaatagccct gaaaacctgt attttcaagg cctgtttgga gccatcgccg 1080
gctttattga gggaggatgg accggaatgg tggatggctg gtacggctat caccaccaga 1140
atgagcaggg atccggatat gccgccgatc tgaagtctac acagaacgcc atcgacgaga 1200
tcacaaacaa ggtgaacagc gtgatcgaga agatgaacac ccagtttaca gctgtgggca 1260
aggagttcaa ccacctggag aagagaatcg agaacctgaa caagaaagtg gacgacggct 1320
tcctggatat ttggacctac aatgccgagc tgctcgtgct cctggagaat gagagaaccc 1380
tggactacca cgacagcaat gtgaagaacc tgtacgagaa ggtgagaagc cagctgaaga 1440
acaatgccaa ggagatcggc aacggctgct ttgagttcta ccacaagtgt gacaacacct 1500
gtatggagtc tgtgaagaac ggcacctacg actaccctaa gtatagcgag gaggccaagc 1560
tgaatagaga ggagatcgac ggcgtgaaac tggaaagcac aagaatctat cagggcgctg 1620
aaaacctgta ttttcagggc ggttctggtt acatcccgga agctccgcgt gacggtcagg 1680
cttacgttcg taaagacggt gaatgggttc tgctgtctac cttcctgggt caccatcatc 1740
accaccatca ccatcatcac tgataaaagc tt 1772

<210> 38
<211> 588
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 38

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His
1 5 10 15
Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser
20 25 30
Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His
35 40 45
Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
50 55 60

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

370 375 380
Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp
385 390 395 400
Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln
405 410 415
Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu
420 425 430
Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr
435 440 445
Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr
450 455 460
His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu
465 470 475 480
Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His
485 490 495
Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp
500 505 510
Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp
515 520 525
Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu
530 535 540
Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly
545 550 555 560
Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe
565 570 575
Leu Gly His His His His His His His His His His
580 585

<210> 39
<211> 1778
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 39
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctggcattt 480
tcggcgctat cgccggcttc attgagggat tttacaagaa cctgatctgg ctggtgaaga 540
agggcaacag ctaccctaag ctgtctaaga gctacatcaa cgacaagggc aaagaagtgc 600
tggtgctgtg gggaatccac caccctagca caagcgccga tcagcagagc ctgtaccaga 660
atgccgatgc ctatgtgttt gtgggcagca gcagatacag caaaaagttc aagcctgaaa 720
ttgccattag acccaaagtg agagatcagg aaggcagaat gaattactac tggaccctgg 780
tggaacctgg cgataagatc acatttgagg ccaccggaaa tctggtggtg cctagatatg 840
catttgctat ggagagaaat gctggctctg gcatcattat ctctgatacc cctgtgcacg 900
actgtaatac cacctgtcag acacctaagg gcgccattaa taccagcctg cccttccaga 960
atattcaccc tatcaccatc ggcaagtgtc ctaagtatgt gaagagcacc aagctgagac 1020
tggctaccgg tctgagaaat agccctgaaa acctgtattt tcaaggcctg tttggagcca 1080
tcgccggctt tattgaggga ggatggaccg gaatggtgga tggctggtac ggctatcacc 1140
accagaatga gcagggatcc ggatatgccg ccgatctgaa gtctacacag aacgccatcg 1200
acgagatcac aaacaaggtg aacagcgtga tcgagaagat gaacacccag tttacagctg 1260
tgggcaagga gttcaaccac ctggagaaga gaatcgagaa cctgaacaag aaagtggacg 1320
acggcttcct ggatatttgg acctacaatg ccgagctgct cgtgctcctg gagaatgaga 1380
gaaccctgga ctaccacgac agcaatgtga agaacctgta cgagaaggtg agaagccagc 1440
tgaagaacaa tgccaaggag atcggcaacg gctgctttga gttctaccac aagtgtgaca 1500
acacctgtat ggagtctgtg aagaacggca cctacgacta ccctaagtat agcgaggagg 1560
ccaagctgaa tagagaggag atcgacggcg tgaaactgga aagcacaaga atctatcagg 1620
gcgctgaaaa cctgtatttt cagggcggtt ctggttacat cccggaagct ccgcgtgacg 1680
gtcaggctta cgttcgtaaa gacggtgaat gggttctgct gtctaccttc ctgggtcacc 1740
atcatcacca ccatcaccat catcactgat aaaagctt 1778

<210> 40
<211> 589
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 40

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

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Gly Ile Phe Gly
165 170 175

Ala Ile Ala Gly Phe Ile Glu Gly Tyr Pro Lys Leu Ser Lys Ser Tyr
180 185 190

Ile Asn Asp Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His
195 200 205

Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Ala
210 215 220

Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu
225 230 235 240

Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr
245 250 255

Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr
260 265 270

Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala
275 280 285

Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr
290 295 300

Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln
305 310 315 320

Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser
325 330 335

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

Tyr Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly
355 360 365

Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu
370 375 380
Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile
385 390 395 400
Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr
405 410 415
Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile
420 425 430
Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr
435 440 445
Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp
450 455 460
Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln
465 470 475 480
Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr
485 490 495
His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr
500 505 510
Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile
515 520 525
Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn
530 535 540
Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp
545 550 555 560
Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr
565 570 575
Phe Leu Gly His His His His His His His His His His
580 585

<210> 41
<211> 1781
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 41
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtggg cattttcggc gctatcgccg 540
gcttcattga gggataccct aagctgtcta agagctacat caacgacaag ggcaaagaag 600
tgctggtgct gtggggaatc caccacccta gcacaagcgc cgatcagcag agcctgtacc 660
agaatgccga tgcctatgtg tttgtgggca gcagcagata cagcaaaaag ttcaagcctg 720
aaattgccat tagacccaaa gtgagagatc aggaaggcag aatgaattac tactggaccc 780
tggtggaacc tggcgataag atcacatttg aggccaccgg aaatctggtg gtgcctagat 840
atgcatttgc tatggagaga aatgctggct ctggcatcat tatctctgat acccctgtgc 900
acgactgtaa taccacctgt cagacaccta agggcgccat taataccagc ctgcccttcc 960
agaatattca ccctatcacc atcggcaagt gtcctaagta tgtgaagagc accaagctga 1020
gactggctac cggtctgaga aatagccctg aaaacctgta ttttcaaggc ctgtttggag 1080
ccatcgccgg ctttattgag ggaggatgga ccggaatggt ggatggctgg tacggctatc 1140
accaccagaa tgagcaggga tccggatatg ccgccgatct gaagtctaca cagaacgcca 1200
tcgacgagat cacaaacaag gtgaacagcg tgatcgagaa gatgaacacc cagtttacag 1260
ctgtgggcaa ggagttcaac cacctggaga agagaatcga gaacctgaac aagaaagtgg 1320
acgacggctt cctggatatt tggacctaca atgccgagct gctcgtgctc ctggagaatg 1380
agagaaccct ggactaccac gacagcaatg tgaagaacct gtacgagaag gtgagaagcc 1440
agctgaagaa caatgccaag gagatcggca acggctgctt tgagttctac cacaagtgtg 1500
acaacacctg tatggagtct gtgaagaacg gcacctacga ctaccctaag tatagcgagg 1560
aggccaagct gaatagagag gagatcgacg gcgtgaaact ggaaagcaca agaatctatc 1620
agggcgctga aaacctgtat tttcagggcg gttctggtta catcccggaa gctccgcgtg 1680
acggtcaggc ttacgttcgt aaagacggtg aatgggttct gctgtctacc ttcctgggtc 1740
accatcatca ccaccatcac catcatcact gataaaagct t 1781

<210> 42
<211> 590
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 42

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

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
50 55 60
Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
65 70 75 80
Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95
Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Gly Ile Phe Gly Ala Ile
130 135 140
Ala Gly Phe Ile Glu Gly Trp Pro Asn His Asp Ser Asn Lys Gly Val
145 150 155 160
Thr Ala Ala Cys Pro His Ala Gly Ala Lys Ser Phe Tyr Lys Asn Leu
165 170 175
Ile Trp Leu Val Lys Lys Gly Asn Ser Tyr Pro Lys Leu Ser Lys Ser
180 185 190
Tyr Ile Asn Asp Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His
195 200 205
His Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp
210 215 220
Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro
225 230 235 240
Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn
245 250 255
Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala
260 265 270
Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn
275 280 285
Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn
290 295 300
Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe
305 310 315 320
Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys
325 330 335
Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu Asn
340 345 350
Leu Tyr Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly

355 360 365

Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn
370 375 380

Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala
385 390 395 400

Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn
405 410 415

Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg
420 425 430

Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp
435 440 445

Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu
450 455 460

Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser
465 470 475 480

Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe
485 490 495

Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr
500 505 510

Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu
515 520 525

Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu
530 535 540

Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg
545 550 555 560

Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser
565 570 575

Thr Phe Leu Gly His His His His His His His His His His
580 585 590

<210> 43

<211> 1784

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 43

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctggca 420
ttttcggcgc tatcgccggc ttcattgagg gatggcctaa tcacgattct aataagggag 480
tgacagccgc ctgtcctcat gccggagcca agtcctttta caagaacctg atctggctgg 540
tgaagaaggg caacagctac cctaagctgt ctaagagcta catcaacgac aagggcaaag 600
aagtgctggt gctgtgggga atccaccacc ctagcacaag cgccgatcag cagagcctgt 660
accagaatgc cgatgcctat gtgtttgtgg gcagcagcag atacagcaaa aagttcaagc 720
ctgaaattgc cattagaccc aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaaatctg gtggtgccta 840
gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gataccctg 900
tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgccct 960
tccagaatat tcaccctatc accatcggca agtgtcctaa gtatgtgaag agcaccaagc 1020
tgagactggc taccggtctg agaaatagcc ctgaaaacct gtattttcaa ggcctgtttg 1080
gagccatcgc cggctttatt gagggaggat ggaccggaat ggtggatggc tggtacggct 1140
atcaccacca gaatgagcag ggatccggat atgccgccga tctgaagtct acacagaacg 1200
ccatcgacga gatcacaaac aaggtgaaca gcgtgatcga gaagatgaac acccagttta 1260
cagctgtggg caaggagttc aaccacctgg agaagagaat cgagaacctg aacaagaaag 1320
tggacgacgg cttcctggat atttggacct acaatgccga gctgctcgtg ctcctggaga 1380
atgagagaac cctggactac cacgacagca atgtgaagaa cctgtacgag aaggtgagaa 1440
gccagctgaa gaacaatgcc aaggagatcg gcaacggctg ctttgagttc taccacaagt 1500
gtgacaacac ctgtatggag tctgtgaaga acggcaccta cgactaccct aagtatagcg 1560
aggaggccaa gctgaataga gaggagatcg acggcgtgaa actggaaagc acaagaatct 1620
atcagggcgc tgaaaacctg tattttcagg gcggttctgg ttacatcccg gaagctccgc 1680
gtgacggtca ggcttacgtt cgtaaagacg gtgaatgggt tctgctgtct accttcctgg 1740
gtcaccatca tcaccaccat caccatcatc actgataaaa gctt 1784

<210> 44
<211> 584
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 44

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

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Gly Ile Phe Gly Ala

195 200 205

Ile Ala Gly Phe Ile Glu Gly Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val

225 230 235 240

Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro

245 250 255

Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg

260 265 270

Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser

275 280 285

Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly

290 295 300

Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile

305 310 315 320

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr

325 330 335

Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly Leu

340 345 350

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

<210> 45
<211> 1766
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 45
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cggcattttc ggcgctatcg ccggcttcat tgagggagat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atagccctca gagggagaga cgcaagaaga gaggcctgtt tggagccatc gccggcttta 1080
ttgagggagg atggaccgga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140
agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcacaa 1200
acaaggtgaa cagcgtgatc gagaagatga acacccagtt tacagctgtg ggcaaggagt 1260
tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320
atatttggac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
accacgacag caatgtgaag aacctgtacg agaaggtgag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680
ttcgtaaaga cggtgaatgg gttctgctgt ctaccttcct gggtcaccat catcaccacc 1740
atcaccatca tcactgataa aagctt 1766

<210> 46
<211> 584
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 46

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His
1 5 10 15
Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser
20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His
35 40 45
Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
50 55 60
Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
65 70 75 80
Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95
Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140
Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160
Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175
Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190
Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205
Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220
Arg Tyr Ser Lys Lys Phe Lys Pro Gly Ile Phe Gly Ala Ile Ala Gly
225 230 235 240
Phe Ile Glu Gly Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
245 250 255
Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
260 265 270
Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser
275 280 285
Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
290 295 300
Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
305 310 315 320
Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
325 330 335
Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly Leu

340 345 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
355 360 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
370 375 380

Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn
385 390 395 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
405 410 415

Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys
420 425 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
435 440 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn
450 455 460

Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala
465 470 475 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn
485 490 495

Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
500 505 510

Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu
515 520 525

Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly
530 535 540

Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
545 550 555 560

Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His
565 570 575

His His His His His His His His
580

<210> 47

<211> 1766

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 47

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagcctgg cattttcggc gctatcgccg 720
gcttcattga gggaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atagccctca gagggagaga cgcaagaaga gaggcctgtt tggagccatc gccggcttta 1080
ttgagggagg atggaccgga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140
agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcacaa 1200
acaaggtgaa cagcgtgatc gagaagatga acacccagtt tacagctgtg ggcaaggagt 1260
tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320
atatttggac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
accacgacag caatgtgaag aacctgtacg agaaggtgag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680
ttcgtaaaga cggtgaatgg gttctgctgt ctaccttcct gggtcaccat catcaccacc 1740
atcaccatca tcactgataa aagctt 1766

<210> 48
<211> 592
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 48

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His
1 5 10 15
Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Arg Lys Lys Arg

165 170 175

Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Tyr Ile Asn Asp Lys

180 185 190

Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Lys Glu

195 200 205

Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Asp

210 215 220

Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro

225 230 235 240

Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn

245 250 255

Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala

260 265 270

Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn

275 280 285

Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn

290 295 300

Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe

305 310 315 320

Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys

325 330 335

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

<210> 49
<211> 15
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 49

Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
1 5 10 15

<210> 50
<211> 17
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 50

Lys Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn
1 5 10 15

Ser

<210> 51
<211> 1790
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 51

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgag aaagaagaga ggcctgtttg 540
gagccatcgc cggctttatt gagtacatca acgacaaggg caaagaagtg ctggtgctgt 600
ggggaatcca ccaccctagc aaggagtcta cacagaaggc cattgatggc gttacaaata 660
aggtcaattc tgatgcctat gtgtttgtgg gcagcagcag atacagcaaa aagttcaagc 720
ctgaaattgc cattagaccc aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaaatctg gtggtgccta 840
gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gataccctg 900
tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgccct 960
tccagaatat tcaccctatc accatcggca agtgtcctaa gtatgtgaag agcaccaagc 1020
tgagactggc taccggtctg agaaatagcc ctcagaggga gagacgcaag aagagaggcc 1080
tgtttggagc catcgccggc tttattgagg gaggatggac cggaatggtg gatggctggt 1140
acggctatca ccaccagaat gagcagggat ccggatatgc cgccgatctg aagtctacac 1200

agaacgccat cgacgagatc acaaacaagg tgaacagcgt gatcgagaag atgaacaccc 1260
agtttacagc tgtgggcaag gagttcaacc acctggagaa gagaatcgag aacctgaaca 1320
agaaagtgga cgacggcttc ctggatattt ggacctacaa tgccgagctg ctcgtgctcc 1380
tggagaatga gagaaccctg gactaccacg acagcaatgt gaagaacctg tacgagaagg 1440
tgagaagcca gctgaagaac aatgccaagg agatcggcaa cggctgcttt gagttctacc 1500
acaagtgtga caacacctgt atggagtctg tgaagaacgg cacctacgac taccctaagt 1560
atagcgagga ggccaagctg aatagagagg agatcgacgg cgtgaaactg gaaagcacaa 1620
gaatctatca gggcgctgaa aacctgtatt ttcagggcgg ttctggttac atcccggaag 1680
ctccgcgtga cggtcaggct tacgttcgta aagacggtga atgggttctg ctgtctacct 1740
tcctgggtca ccatcatcac caccatcacc atcatcactg ataaaagctt 1790

<210> 52
<211> 589
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 52

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

Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Tyr Ile Asn Asp Lys
180 185 190
Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Lys Glu
195 200 205
Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Asp
210 215 220
Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro
225 230 235 240
Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn
245 250 255
Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala
260 265 270
Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn
275 280 285
Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn
290 295 300
Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe
305 310 315 320
Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys
325 330 335
Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ile Pro Ser Ile
340 345 350
Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly
355 360 365
Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu
370 375 380
Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile
385 390 395 400
Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr
405 410 415
Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile
420 425 430
Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr
435 440 445
Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp
450 455 460
Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln
465 470 475 480
Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr

485 490 495
His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr
500 505 510
Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile
515 520 525
Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn
530 535 540
Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp
545 550 555 560
Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr
565 570 575
Phe Leu Gly His His His His His His His His His His
580 585

<210> 53
<211> 1781
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 53
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgag aaagaagaga ggcctgtttg 540
gagccatcgc cggctttatt gagtacatca acgacaaggg caaagaagtg ctggtgctgt 600
ggggaatcca ccaccctagc aaggagtcta cacagaaggc cattgatggc gttacaaata 660
aggtcaattc tgatgcctat gtgtttgtgg gcagcagcag atacagcaaa aagttcaagc 720
ctgaaattgc cattagaccc aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaaatctg gtggtgccta 840
gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gataccctg 900
tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgccct 960
tccagaatat tcaccctatc accatcggca agtgtcctaa gtatgtgaag agcaccaagc 1020
tgagactggc taccggtctg agaaatatcc ctagcatcca gagcagaggc ctgtttggag 1080
ccatcgccgg ctttattgag ggaggatgga ccggaatggt ggatggctgg tacggctatc 1140

accaccagaa tgagcaggga tccggatatg ccgccgatct gaagtctaca cagaacgcca 1200
tcgacgagat cacaaacaag gtgaacagcg tgatcgagaa gatgaacacc cagtttacag 1260
ctgtgggcaa ggagttcaac cacctggaga agagaatcga gaacctgaac aagaaagtgg 1320
acgacggctt cctggatatt tggacctaca atgccgagct gctcgtgctc ctggagaatg 1380
agagaaccct ggactaccac gacagcaatg tgaagaacct gtacgagaag gtgagaagcc 1440
agctgaagaa caatgccaag gagatcggca acggctgctt tgagttctac cacaagtgtg 1500
acaacacctg tatggagtct gtgaagaacg gcacctacga ctaccctaag tatagcgagg 1560
aggccaagct gaatagagag gagatcgacg gcgtgaaact ggaaagcaca agaatctatc 1620
agggcgctga aaacctgtat tttcagggcg gttctggtta catcccggaa gctccgcgtg 1680
acggtcaggc ttacgttcgt aaagacggtg aatgggttct gctgtctacc ttcctgggtc 1740
accatcatca ccaccatcac catcatcact gataaaagct t 1781

<210> 54
<211> 585
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 54

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

165 170 175

Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Tyr Ile Asn Asp
180 185 190

Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr
195 200 205

Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe
210 215 220

Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile
225 230 235 240

Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr
245 250 255

Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu
260 265 270

Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly
275 280 285

Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln
290 295 300

Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His
305 310 315 320

Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu
325 330 335

Arg Leu Ala Thr Gly Leu Arg Asn Ile Glu Asn Leu Tyr Phe Gln Gly
340 345 350

Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met
355 360 365

Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly
370 375 380

Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr
385 390 395 400

Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala
405 410 415

Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
420 425 430

Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
435 440 445

Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
450 455 460

Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
465 470 475 480

Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
485 490 495
Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
500 505 510
Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
515 520 525
Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln
530 535 540
Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
545 550 555 560
Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His
565 570 575
His His His His His His His His His His
580 585

<210> 55
<211> 1769
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 55

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgag aaagaagaga ggcctgttcg 540
gcgctatcgc cggcttcatt gagggataca tcaacgacaa gggcaaagaa gtgctggtgc 600
tgtggggaat ccaccaccct agcacaagcg ccgatcagca gagcctgtac cagaatgccg 660
atgcctatgt gtttgtgggc agcagcagat acagcaaaaa gttcaagcct gaaattgcca 720
ttagacccaa agtgagagat caggaaggca gaatgaatta ctactggacc ctggtggaac 780
ctggcgataa gatcacattt gaggccaccg gaaatctggt ggtgcctaga tatgcatttg 840
ctatggagag aaatgctggc tctggcatca ttatctctga tacccctgtg cacgactgta 900
ataccacctg tcagacacct aagggcgcca ttaataccag cctgcccttc cagaatattc 960
accctatcac catcggcaag tgtcctaagt atgtgaagag caccaagctg agactggcta 1020
ccggtctgag aaatatcgaa aacctgtatt ttcaaggcct gtttggagcc atcgccggct 1080

ttattgaggg aggatggacc ggaatggtgg atggctggta cggctatcac caccagaatg 1140
agcagggatc cggatatgcc gccgatctga agtctacaca gaacgccatc gacgagatca 1200
caaacaaggt gaacagcgtg atcgagaaga tgaacaccca gtttacagct gtgggcaagg 1260
agttcaacca cctggagaag agaatcgaga acctgaacaa gaaagtggac gacggcttcc 1320
tggatatttg gacctacaat gccgagctgc tcgtgctcct ggagaatgag agaaccctgg 1380
actaccacga cagcaatgtg aagaacctgt acgagaaggt gagaagccag ctgaagaaca 1440
atgccaagga gatcggcaac ggctgctttg agttctacca caagtgtgac aacacctgta 1500
tggagtctgt gaagaacggc acctacgact accctaagta tagcgaggag gccaagctga 1560
atagagagga gatcgacggc gtgaaactgg aaagcacaag aatctatcag ggcgctgaaa 1620
acctgtattt tcagggcggt tctggttaca tcccggaagc tccgcgtgac ggtcaggctt 1680
acgttcgtaa agacggtgaa tgggttctgc tgtctacctt cctgggtcac catcatcacc 1740
accatcacca tcatcactga taaaagctt 1769

<210> 56
<211> 593
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 56

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

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

465 470 475 480
Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
485 490 495
Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys
500 505 510
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn
515 520 525
Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln
530 535 540
Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu
545 550 555 560
Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val
565 570 575
Leu Leu Ser Thr Phe Leu Gly His His His His His His His His His
580 585 590
His

<210> 57
<211> 21
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 57
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Cys
1 5 10 15
Lys Cys Ser Asp Ser
20

<210> 58
<211> 1793
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 58
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagccttc cctgctgacc gaggtggaga 720
cccccaccag gaacggctgg gagtgcaagt gctccgactc cggcagaatg aattactact 780
ggaccctggt ggaacctggc gataagatca catttgaggc caccggaaat ctggtggtgc 840
ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900
ctgtgcacga ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960
ccttccagaa tattcaccct atcaccatcg gcaagtgtcc taagtatgtg aagagcacca 1020
agctgagact ggctaccggt ctgagaaata gccctcagag ggagagacgc aagaagagag 1080
gcctgtttgg agccatcgcc ggctttattg agggaggatg gaccggaatg gtggatggct 1140
ggtacggcta tcaccaccag aatgagcagg gatccggata tgccgccgat ctgaagtcta 1200
cacagaacgc catcgacgag atcacaaaca aggtgaacag cgtgatcgag aagatgaaca 1260
cccagtttac agctgtgggc aaggagttca accacctgga gaagagaatc gagaacctga 1320
acaagaaagt ggacgacggc ttcctggata tttggaccta caatgccgag ctgctcgtgc 1380
tcctggagaa tgagagaacc ctggactacc acgacagcaa tgtgaagaac ctgtacgaga 1440
aggtgagaag ccagctgaag aacaatgcca aggagatcgg caacggctgc tttgagttct 1500
accacaagtg tgacaacacc tgtatggagt ctgtgaagaa cggcacctac gactacccta 1560
agtatagcga ggaggccaag ctgaatagag aggagatcga cggcgtgaaa ctggaaagca 1620
caagaatcta tcagggcgct gaaaacctgt attttcaggg cggttctggt tacatcccgg 1680
aagctccgcg tgacggtcag gcttacgttc gtaaagacgg tgaatgggtt ctgctgtcta 1740
ccttcctggg tcaccatcat caccaccatc accatcatca ctgataaaag ctt 1793

<210> 59
<211> 593
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 59

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

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Ser Leu Leu Thr Glu
195 200 205

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

Asp Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys
225 230 235 240

Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met
245 250 255

Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu
260 265 270

Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg
275 280 285

Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys
290 295 300

Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro
305 310 315 320

Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val
325 330 335

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

Arg Glu Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
355 360 365

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

<210> 60
<211> 1793
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 60
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag ctccctgctg accgaggtgg agacccccac caggaacggc tgggagtgca 660
agtgctccga ctccgatgcc tatgtgtttg tgggcagcag cagatacagc aaaaagttca 720
agcctgaaat tgccattaga cccaaagtga gagatcagga aggcagaatg aattactact 780
ggaccctggt ggaacctggc gataagatca catttgaggc caccggaaat ctggtggtgc 840
ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900
ctgtgcacga ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960
ccttccagaa tattcaccct atcaccatcg gcaagtgtcc taagtatgtg aagagcacca 1020
agctgagact ggctaccggt ctgagaaata gccctcagag ggagagacgc aagaagagag 1080
gcctgtttgg agccatcgcc ggctttattg agggaggatg gaccggaatg gtggatggct 1140
ggtacggcta tcaccaccag aatgagcagg gatccggata tgccgccgat ctgaagtcta 1200
cacagaacgc catcgacgag atcacaaaca aggtgaacag cgtgatcgag aagatgaaca 1260
cccagtttac agctgtgggc aaggagttca accacctgga gaagagaatc gagaacctga 1320
acaagaaagt ggacgacggc ttcctggata tttggaccta caatgccgag ctgctcgtgc 1380
tcctggagaa tgagagaacc ctggactacc acgacagcaa tgtgaagaac ctgtacgaga 1440
aggtgagaag ccagctgaag aacaatgcca aggagatcgg caacggctgc tttgagttct 1500
accacaagtg tgacaacacc tgtatggagt ctgtgaagaa cggcacctac gactacccta 1560
agtatagcga ggaggccaag ctgaatagag aggagatcga cggcgtgaaa ctggaaagca 1620
caagaatcta tcagggcgct gaaaacctgt attttcaggg cggttctggt tacatcccgg 1680
aagctccgcg tgacggtcag gcttacgttc gtaaagacgg tgaatgggtt ctgctgtcta 1740
ccttcctggg tcaccatcat caccaccatc accatcatca ctgataaaag ctt 1793

<210> 61
<211> 591
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 61

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

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Ser Leu Leu Thr Glu Val Glu Thr

225 230 235 240

Pro Thr Arg Asn Gly Trp Glu Cys Lys Cys Ser Asp Ser Gly Arg Met

245 250 255

Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu

260 265 270

Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg

275 280 285

Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys

290 295 300

Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro

305 310 315 320

Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val

325 330 335

Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu

340 345 350

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

<210> 62
<211> 1787
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 62
ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcatc gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatcac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggcaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagccttc cctgctgacc gaggtggaga 720
cccccaccag gaacggctgg gagtgcaagt gctccgactc cggcagaatg aattactact 780
ggaccctggt ggaacctggc gataagatca catttgaggc caccggaaat ctggtggtgc 840
ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900
ctgtgcacga ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960
ccttccagaa tattcaccct atcaccatcg gcaagtgtcc taagtatgtg aagagcacca 1020
agctgagact ggctaccggt ctgagaaata gccctgaaaa cctgtatttt caaggcctgt 1080
ttggagccat cgccggcttt attgagggag gatggaccgg aatggtggat ggctggtacg 1140
gctatcacca ccagaatgag cagggatccg gatatgccgc cgatctgaag tctacacaga 1200
acgccatcga cgagatcaca aacaaggtga acagcgtgat cgagaagatg aacacccagt 1260
ttacagctgt gggcaaggag ttcaaccacc tggagaagag aatcgagaac ctgaacaaga 1320
aagtggacga cggcttcctg gatatttgga cctacaatgc cgagctgctc gtgctcctgg 1380
agaatgagag aaccctggac taccacgaca gcaatgtgaa gaacctgtac gagaaggtga 1440
gaagccagct gaagaacaat gccaaggaga tcggcaacgg ctgctttgag ttctaccaca 1500
agtgtgacaa cacctgtatg gagtctgtga agaacggcac ctacgactac cctaagtata 1560
gcgaggaggc caagctgaat agagaggaga tcgacggcgt gaaactggaa agcacaagaa 1620
tctatcaggg cgctgaaaac ctgtattttc agggcggttc tggttacatc ccggaagctc 1680
cgcgtgacgg tcaggcttac gttcgtaaag acggtgaatg ggttctgctg tctaccttcc 1740
tgggtcacca tcatcaccac catcaccatc atcactgata aaagctt 1787

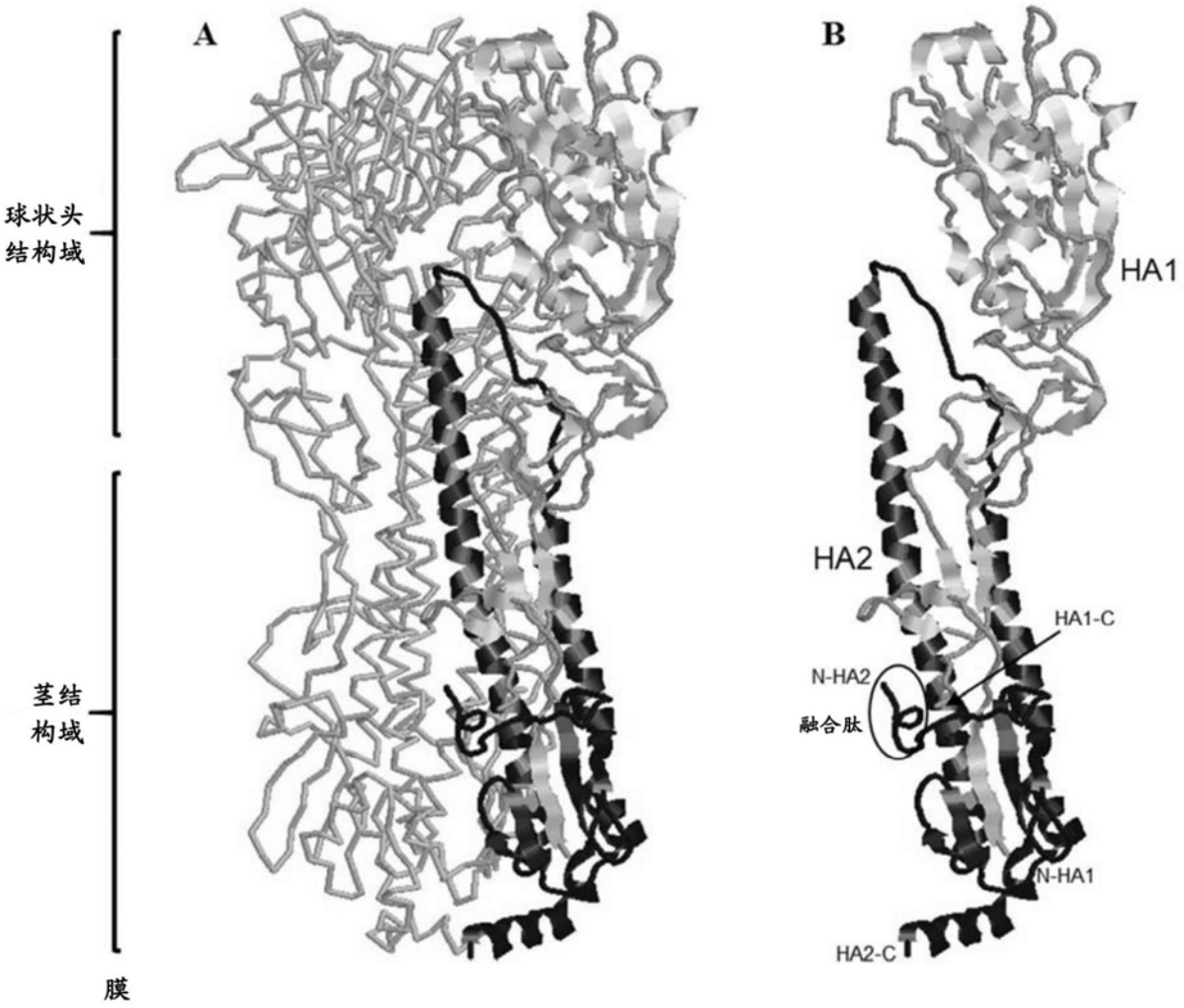


图1

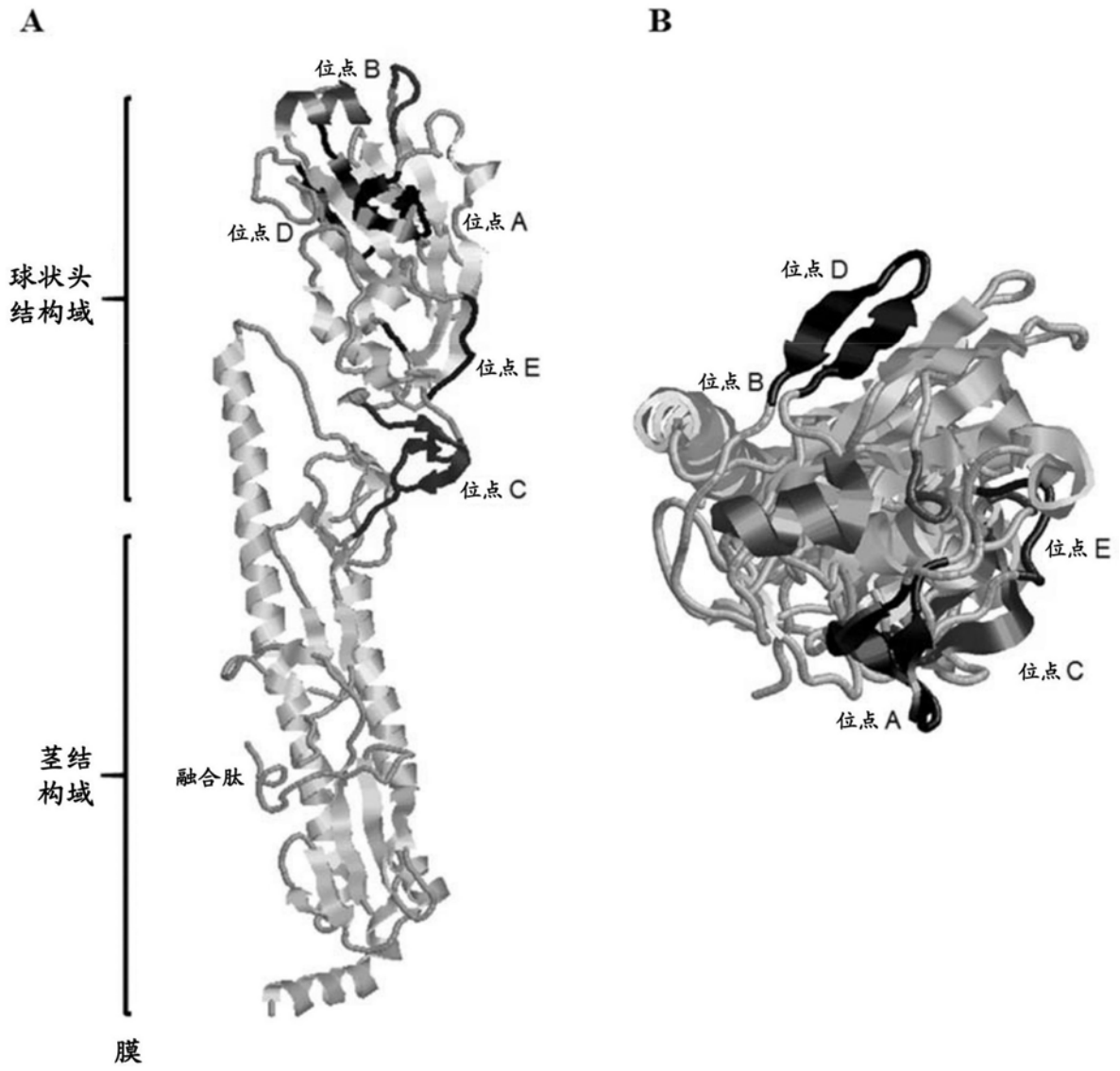


图2

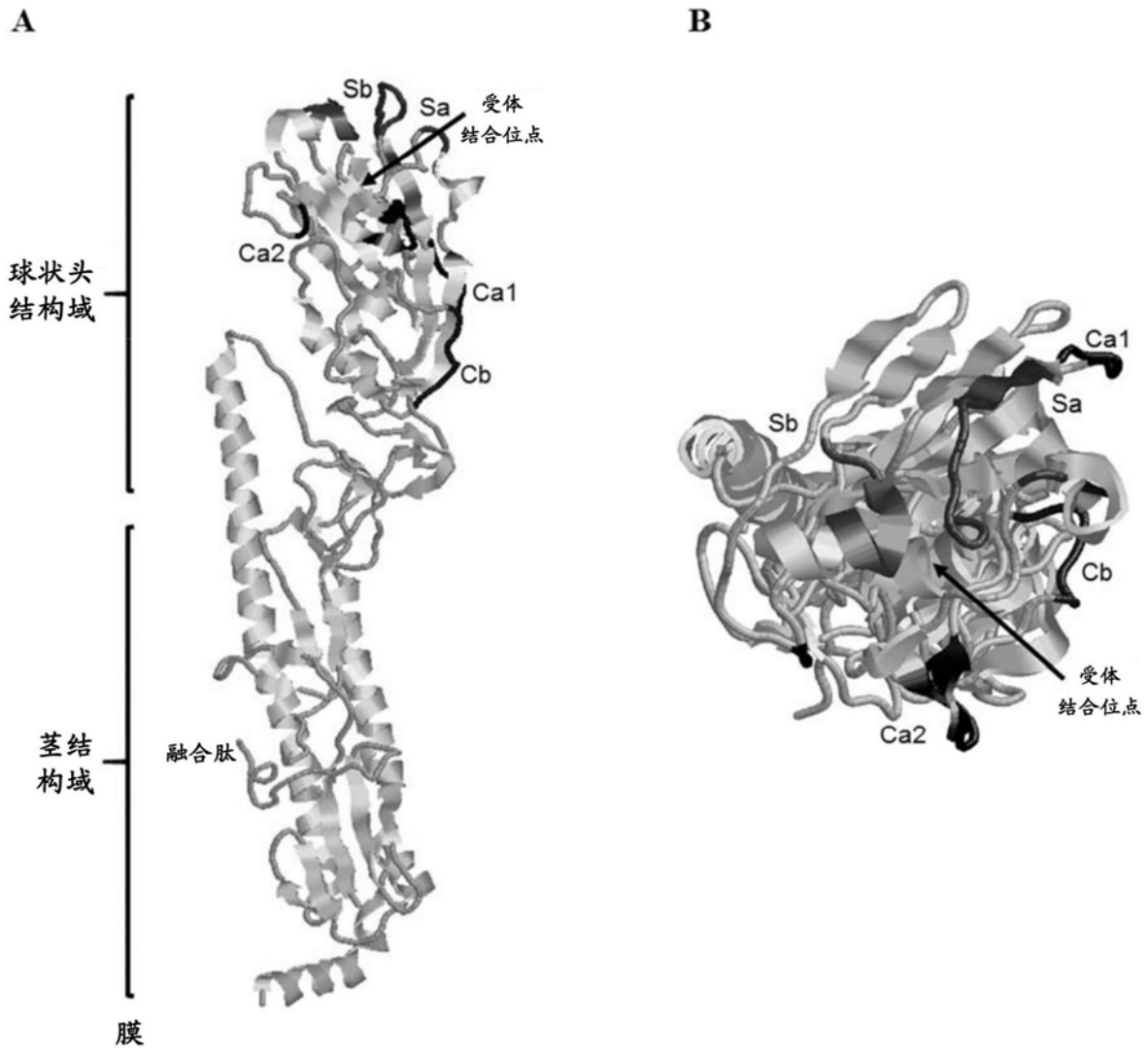

A
B
Sb
Sa
受体
结合位点
Ca2
Ca1
Cb
球状头
结构域
茎结
构域
融合肽
膜
Ca1
Sa
Sb
Cb
受体
结合位点
Ca2

图3

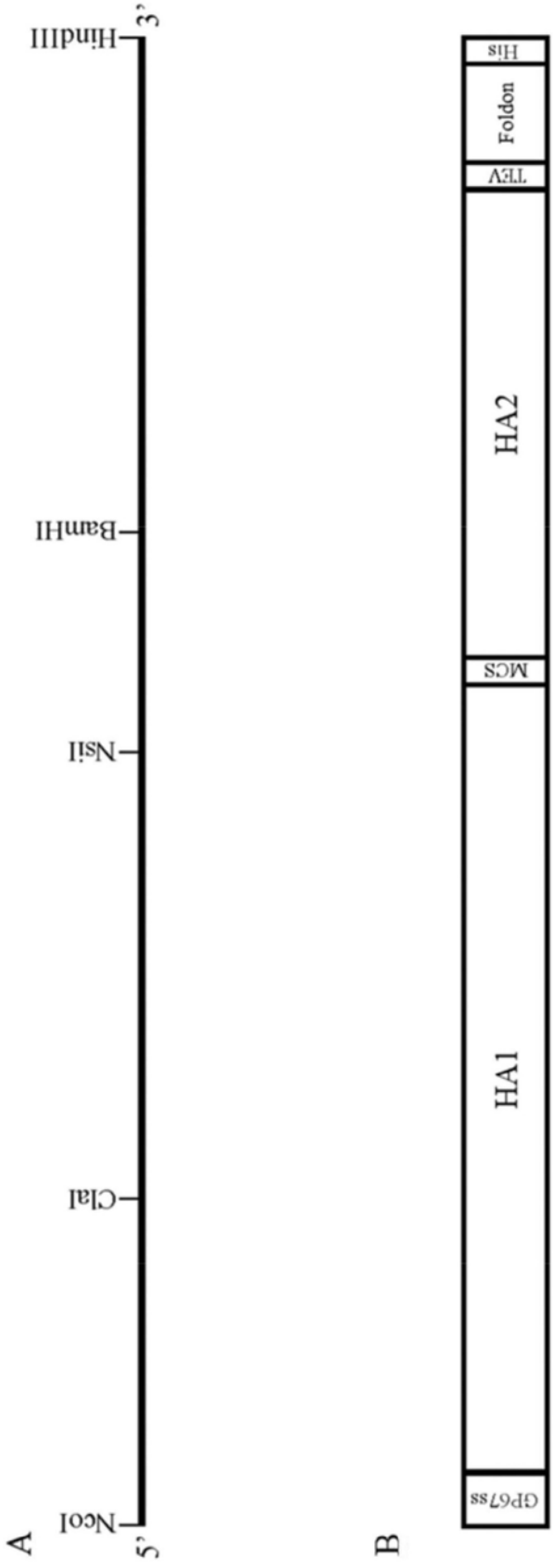

A
5'
NcoI
ClaI
NsiI
BamHI
HindIII
3'
B
GP67ss
HA1
MCS
HA2
TEV
Foldon
His

图4

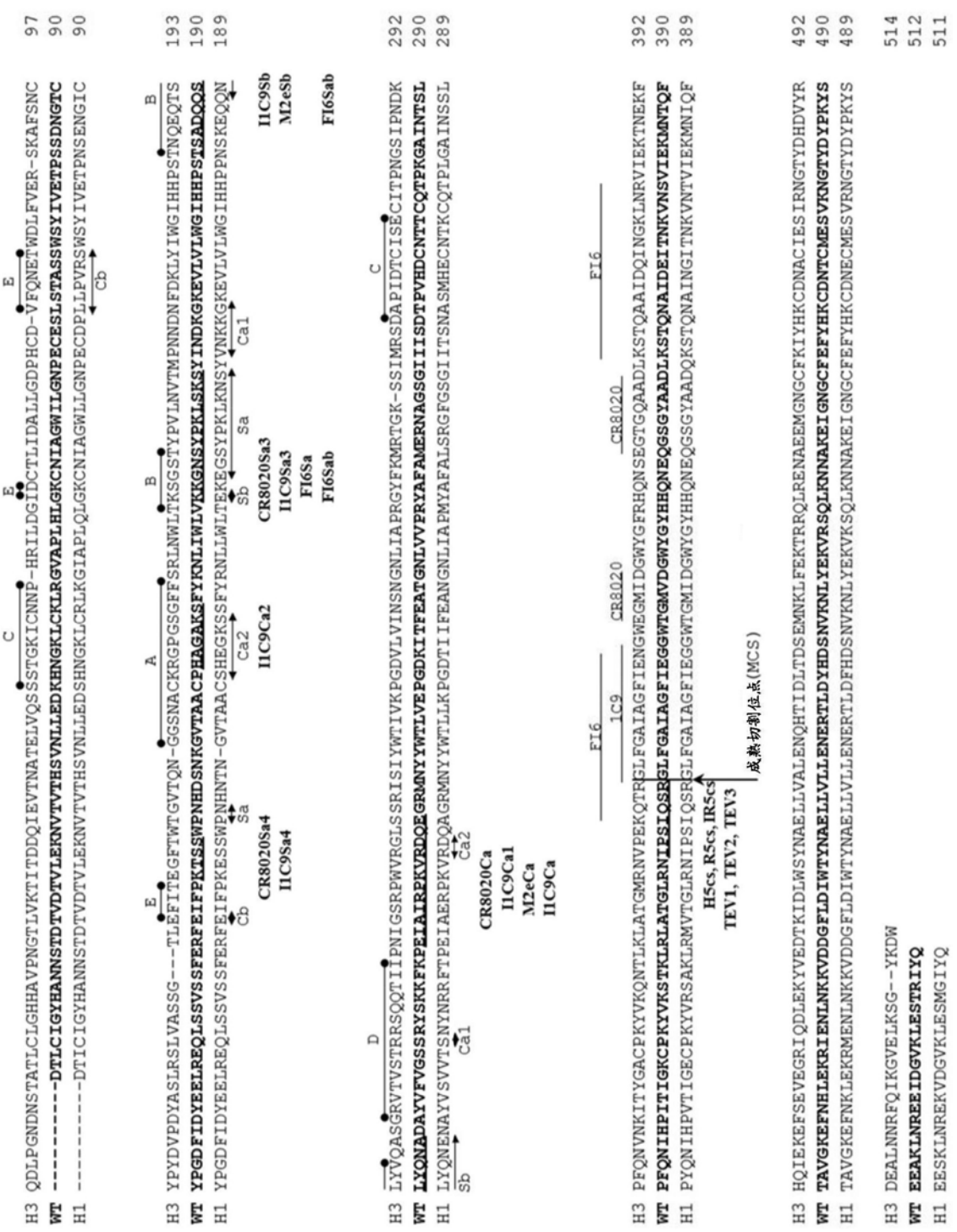

H3 QDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATELVQSSSTGKICNNP-HRILDGIDCTLIDALLGDPHCD-VFQNETWDLFVER-SKAFSNC 97
WT ----------DTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTC 90
H1 ----------DTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGIC 90
H3 YPYDVPDYASLRSLVASSG---TLEFITEGFTWTGVTQN-GGSNACKRGPGSGFFSRLNWLTKSGSTYPVLNVTMPNNDNFDKLYIWGIHHPSTNQEQTS 193
WT YPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQS 190
H1 YPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTN-GVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQN 189
H3 LYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGLSSRISIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTGK-SSIMRSDAPIDTCISECITPNGSIPNDK 292
WT LYQNADAVVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINTSL 290
H1 LYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLGAINSSL 289
H3 PFQNVNKTTYGACPKYVKQNTLKLATGMRNVPEKQTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKF 392
WT PFQNIHPTTIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQF 390
H1 PYQNIHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNTVIEKMNIQF 389
H3 HQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTRRQLRENAEEMGNGCFKIYHKCDNACIESIRNGTYDHDVYR 492
WT TAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYS 490
H1 TAVGKEFNKLEKRMENLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYPKYS 489
H3 DEALNNRFQIKGVELKSG--YKDW 514
WT EEAKLNREEIDGVKLESTRIYQ 512
H1 EESKLNREKVDGVKLESMGIYQ 511
A
B
C
D
E
Ca1
Ca2
Cb
Sa
Sb
CR8020Sa4
I1C9Sa4
CR8020Ca
I1C9Ca1
M2eCa
I1C9Ca
CR8020Sa3
I1C9Sa3
FI6Sa
FI6Sab
I1C9Sb
M2eSb
FI6Sab
CR8020
FI6
1C9
H5cs, R5cs, IR5cs
TEV1, TEV2, TEV3
成熟切割位点(MCS)

图5

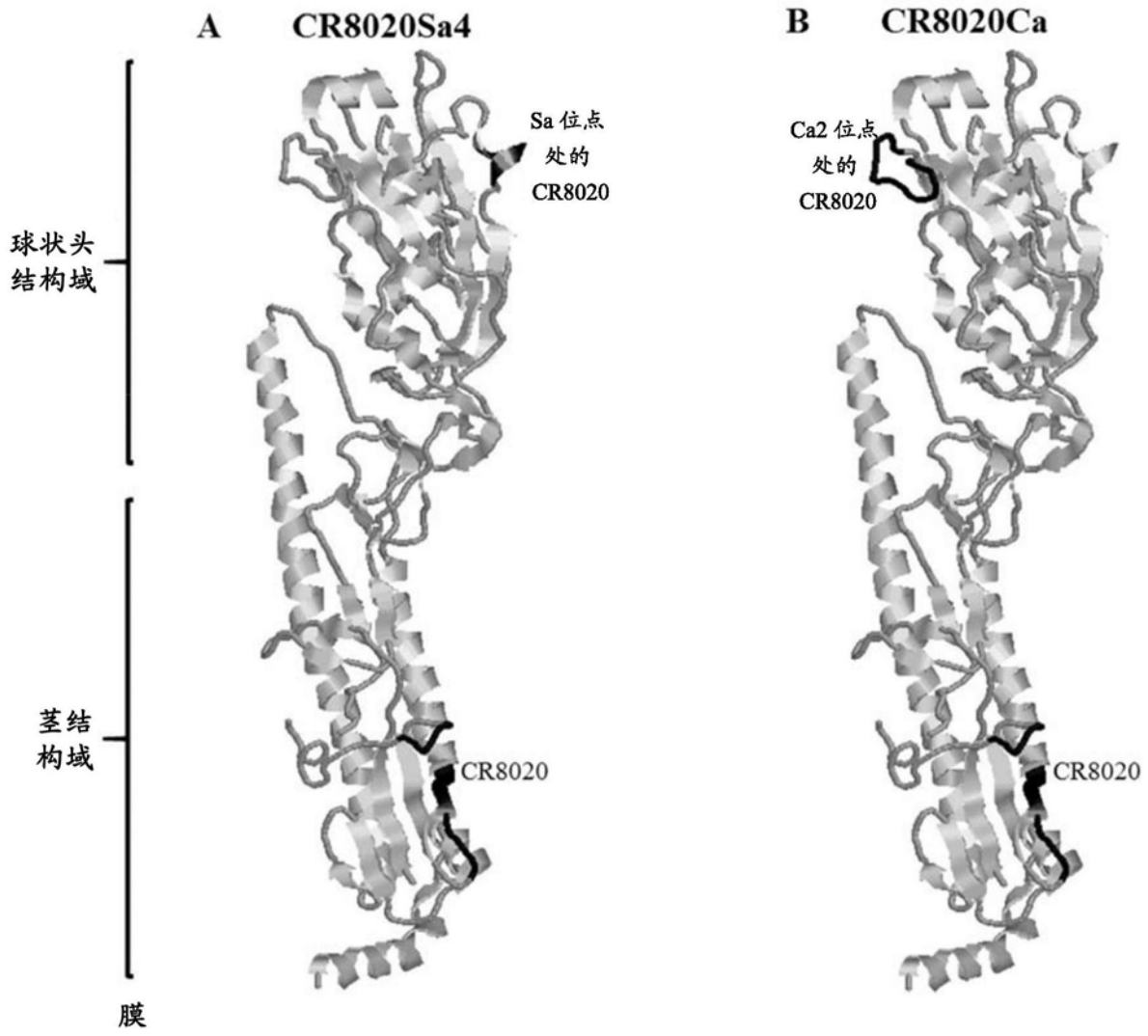
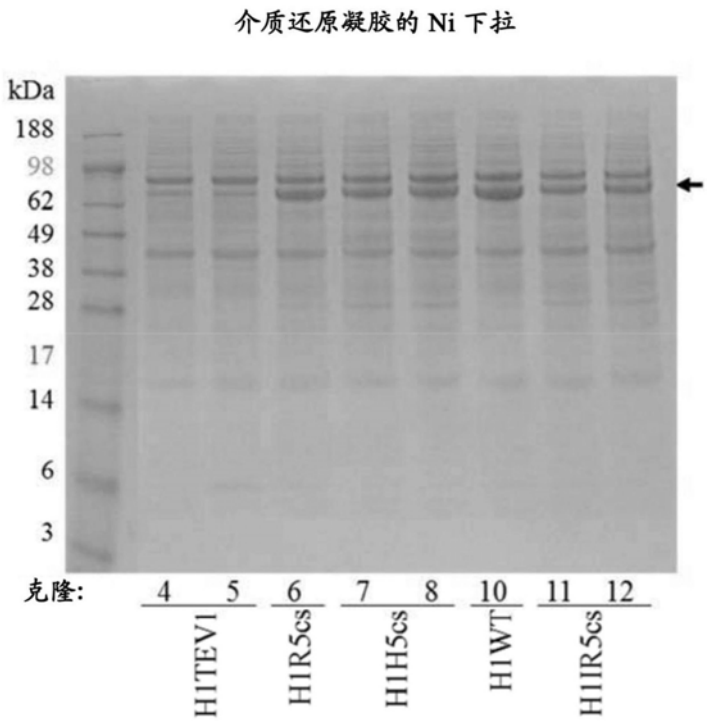

A
CR8020Sa4
Sa 位点
处的
CR8020
球状头
结构域
茎结
构域
CR8020
膜
B
CR8020Ca
Ca2 位点
处的
CR8020
CR8020

图6

A


介质还原凝胶的 Ni 下拉
kDa
188
98
62
49
38
28
17
14
6
3
克隆:
4 5
6
7 8
10
11 12
H1TEV1
H1R5cs
H1H5cs
H1WT
H1IR5cs

B

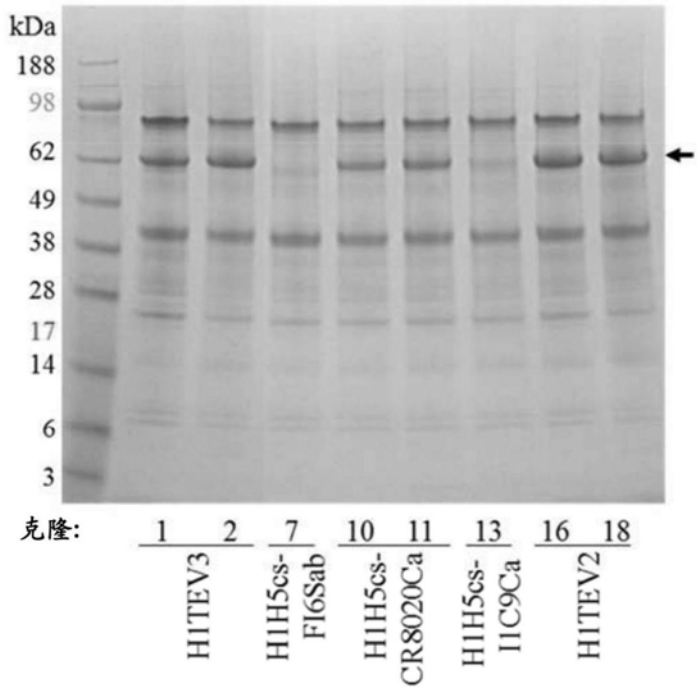
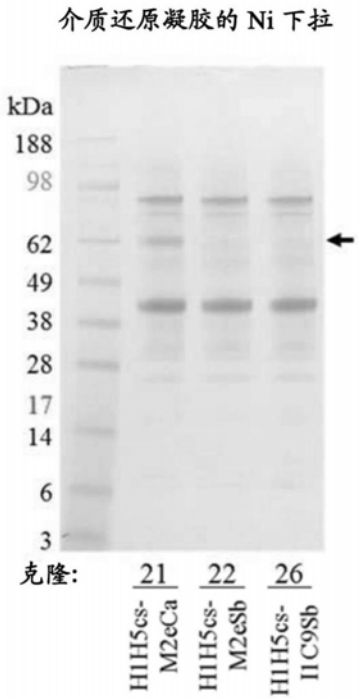

介质非还原凝胶的 Ni 下拉
kDa
188
98
62
49
38
28
17
14
6
3
克隆:
1 2
7
10 11
13
16 18
H1TEV3
H1H5cs-F16Sab
H1H5cs-CR8020Ca
H1H5cs-11C9Ca
H1TEV2

图7

C


介质还原凝胶的 Ni 下拉
kDa
188
98
62
49
38
28
17
14
6
3
克隆:
21
22
26
H1H5cs-M2eCa
H1H5cs-M2eSb
H1H5cs-I1C9Sb

D

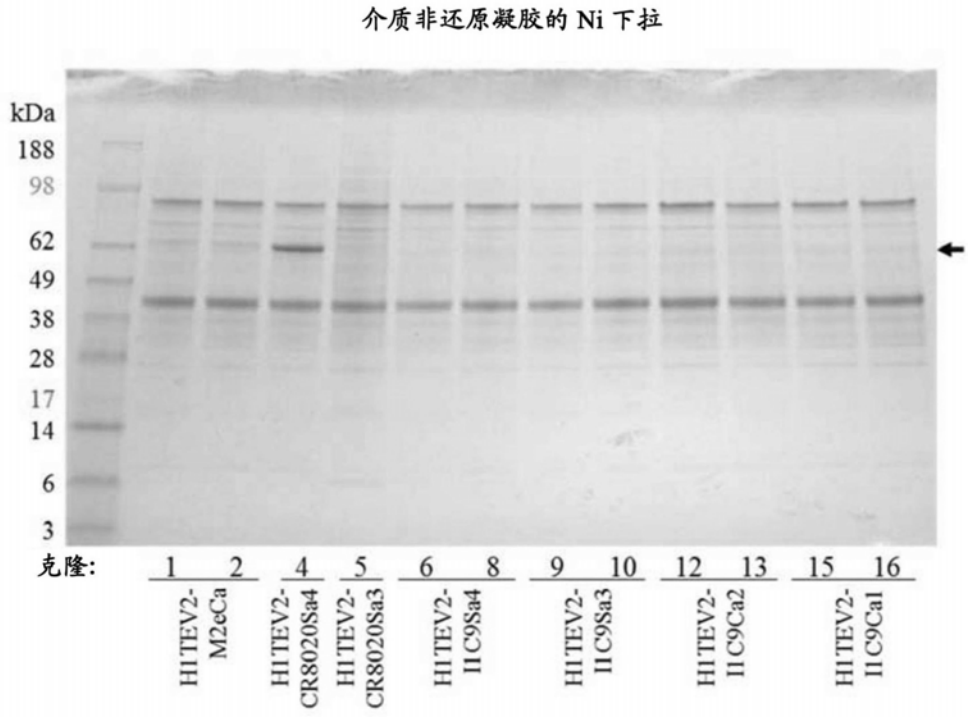

介质非还原凝胶的 Ni 下拉
kDa
188
98
62
49
38
28
17
14
6
3
克隆:
1
2
4
5
6
8
9
10
12
13
15
16
H1TEV2-M2eCa
H1TEV2-CR8020Sa4
H1TEV2-CR8020Sa3
H1TEV2-I1C9Sa4
H1TEV2-I1C9Sa3
H1TEV2-I1C9Ca2
H1TEV2-I1C9Ca1

图7(续)

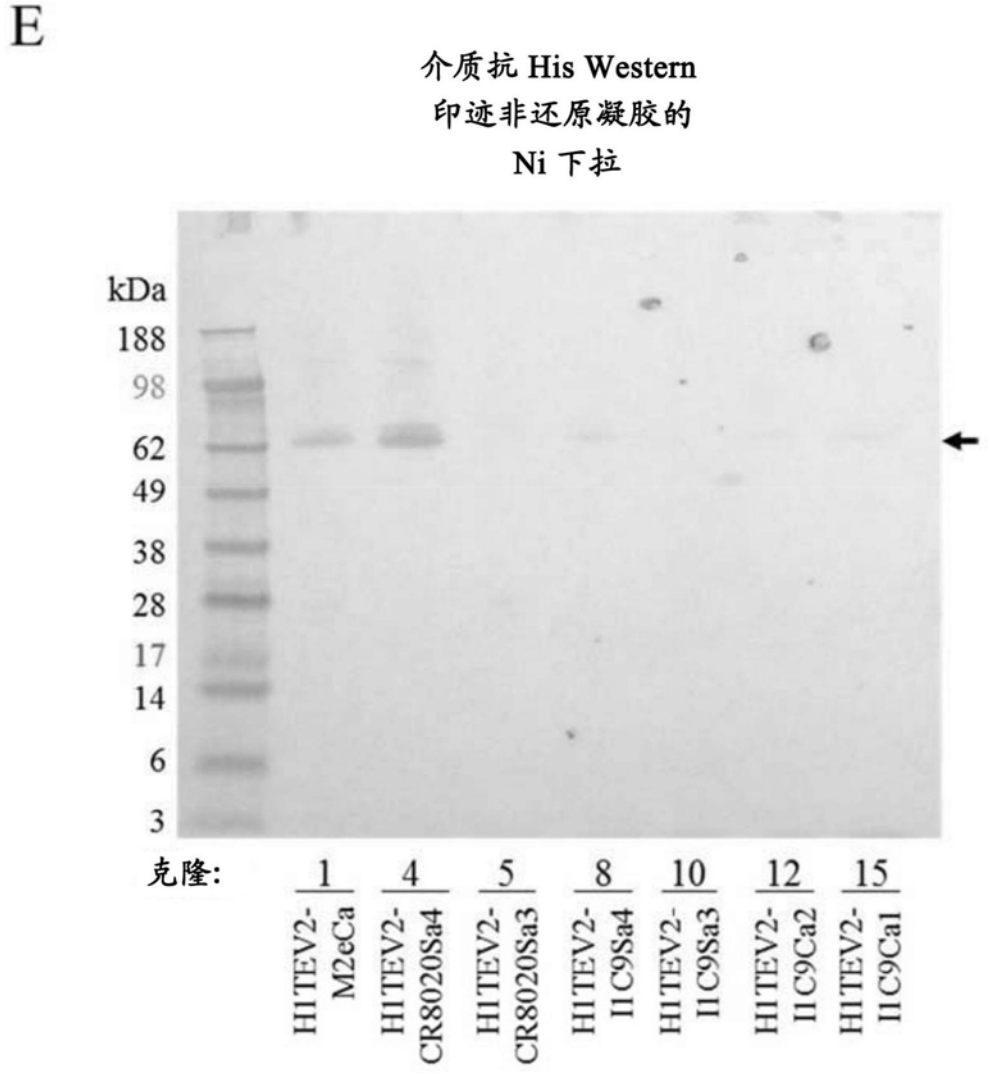

E
介质抗 His Western
印迹非还原凝胶的
Ni 下拉
kDa
188
98
62
49
38
28
17
14
6
3
克隆:
1
4
5
8
10
12
15
H1TEV2-M2eCa
H1TEV2-CR8020Sa4
H1TEV2-CR8020Sa3
H1TEV2-I1C9Sa4
H1TEV2-I1C9Sa3
H1TEV2-I1C9Ca2
H1TEV2-I1C9Ca1

图7(续)

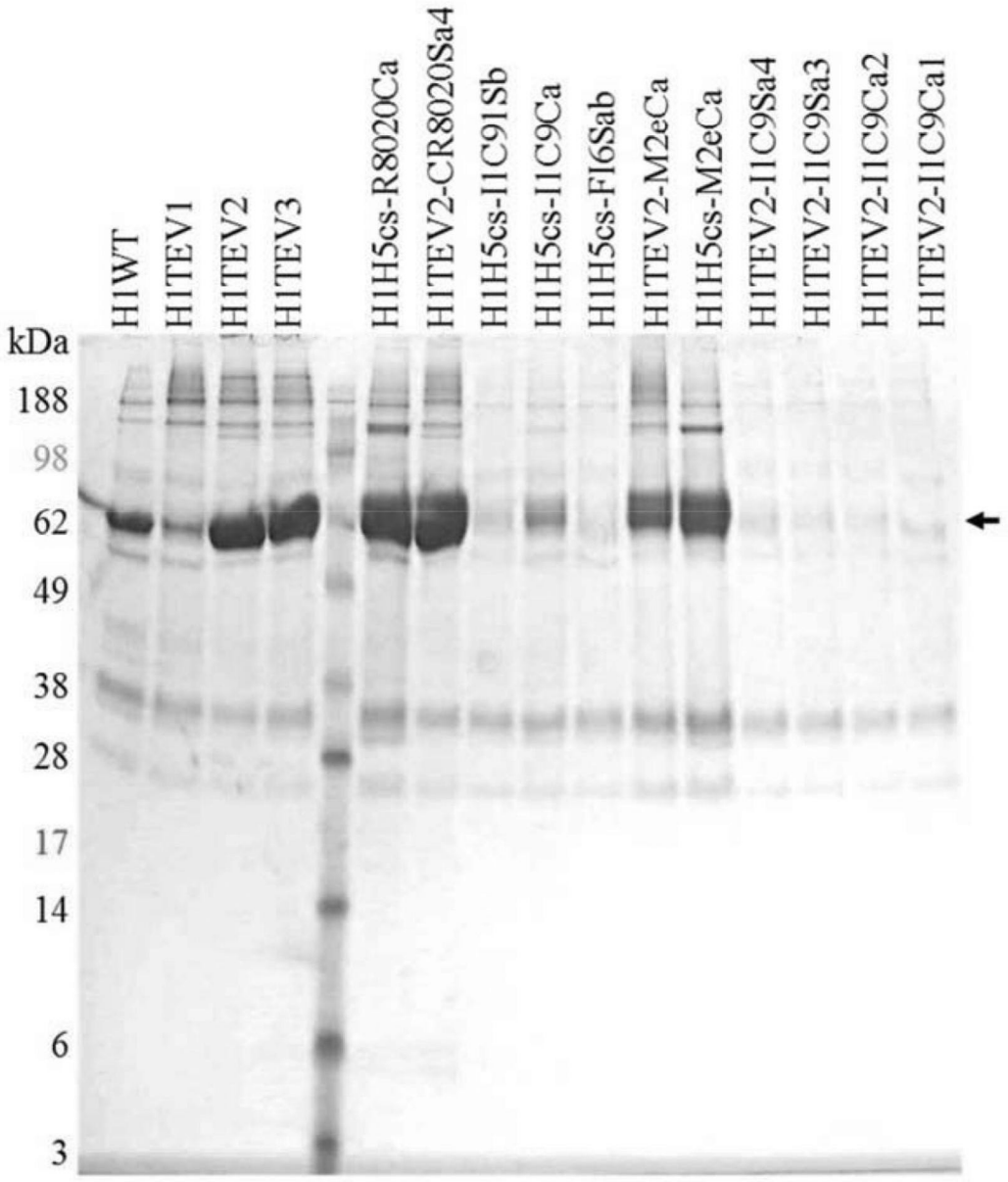

H1WT
H1TEV1
H1TEV2
H1TEV3
H1H5cs-R8020Ca
H1TEV2-CR8020Sa4
H1H5cs-I1C91Sb
H1H5cs-I1C9Ca
H1H5cs-FI6Sab
H1TEV2-M2eCa
H1H5cs-M2eCa
H1TEV2-I1C9Sa4
H1TEV2-I1C9Sa3
H1TEV2-I1C9Ca2
H1TEV2-I1C9Ca1
kDa
188
98
62
49
38
28
17
14
6
3

图8

专利名称(译)	具有异源表位和/或改变的成熟切割位点的流感血凝素的组合物		
公开(公告)号	CN109073639A	公开(公告)日	2018-12-21
申请号	CN201780018856.3	申请日	2017-02-02
[标]发明人	C罗		
发明人	C·罗		
IPC分类号	G01N33/53 A61K39/145 A61K39/155 A61P31/16 C07K14/115 C07K14/285 C12N15/44		
CPC分类号	A61P31/16 H01L23/50 H01L23/5223 H01L23/5227 H03F1/0222 H03F1/565 H03F3/193 H03F3/195 H03F2200/294 H03F2200/408 H03F2200/451 H03H7/0115 H04B1/0064 H01L23/66 C07K14/005 C12N2760/16122 C12N2760/16222 H04B1/40		
优先权	62/290890 2016-02-03 US		
外部链接	Espacenet SIPO		

摘要(译)

公开了修饰形式的血凝素(HA)蛋白包括具有修饰的免疫显性区域和具有修饰的成熟切割位点的那些修饰形式的血凝素蛋白，以及含有它们的病毒和病毒样颗粒。