(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117715928 A
(43)申请公布日 2024.03.15

(21)申请号 202280050778.6
(22)申请日 2022.06.14
(30)优先权数据
63/210,456 2021.06.14 US
(85)PCT国际申请进入国家阶段日
2024.01.18
(86)PCT国际申请的申请数据
PCT/US2022/033452 2022.06.14
(87)PCT国际申请的公布数据
WO2022/266113 EN 2022.12.22
(71)申请人 桥生物基因疗法研究有限责任公司
地址 美国北卡罗来纳州
(72)发明人 D·W·斯科特
(74)专利代理机构 北京市柳沈律师事务所 11105
专利代理师 邹宗亮
(51)Int.Cl.
C07K 14/47(2006.01)
A61K 48/00(2006.01)

权利要求书5页 说明书34页
序列表68页

(54)发明名称

用于结节性硬化的基因疗法

(57)摘要

本公开提供了用于治疗结节性硬化的基因疗法组合物和方法。特别地,本公开提供了包含重组腺相关病毒(rAAV)的组合物及其使用方法,所述重组腺相关病毒包含AAV衣壳蛋白和编码凝聚的马铃薯球蛋白(cTuberin)的AAV表达盒。

CN 117715928 A

1.一种凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基419至932。

2.如权利要求1所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基947-988。

3.如权利要求1或2所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基1205-1271。

4.如权利要求1至3中任一项所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基1336-1497。

5.如权利要求1所述的cTuberin,其中所述C端区域包含与SEQ ID NO:10-12中的一个具有至少90%同一性的氨基酸序列。

6.如权利要求1所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基933至1109。

7.如权利要求6所述的cTuberin,其中C端结构域包含与SEQ ID NO:8具有至少90%同一性的氨基酸序列。

8.如权利要求1至7中任一项所述的cTuberin,其中N端结构域包含与SEQ ID NO:5具有至少90%同一性的氨基酸序列。

9.一种凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述C端区域包含与SEQ ID NO:7具有至少90%同一性的氨基酸序列,并且其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基451至932。

10.如权利要求9所述的cTuberin,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基419至932。

11.如权利要求9或10所述的cTuberin,其中所述C端区域包含与SEQ ID NO:8具有至少90%同一性的氨基酸序列。

12.如权利要求9或10所述的cTuberin,其中所述C端区域包含与SEQ ID NO:9具有至少90%同一性的氨基酸序列。

13.如权利要求9所述的cTuberin,其中所述N端区域包含与SEQ ID NO:4具有至少90%同一性的氨基酸序列。

14.如权利要求9至12中任一项所述的cTuberin,其中所述N端区域包含与SEQ ID NO:5具有至少90%同一性的氨基酸序列。

15.如权利要求9所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:14具有至少90%同一性的氨基酸序列。

16.如权利要求9所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:15具有至少90%同一性的氨基酸序列。

17.如权利要求9所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:16具有至少90%同一性的氨基酸序列。

18.一种凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述C端区域包含与SEQ ID NO:10-12中的一个具有至少90%同一性的氨基酸序列,并且其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残

基451至932。

19.如权利要求18所述的cTuberin,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932。

20.如权利要求18或19所述的cTuberin,其中所述C端区域包含与SEQ ID NO:10具有至少90%同一性的氨基酸序列。

21.如权利要求18或19所述的cTuberin,其中所述C端区域包含与SEQ ID NO:11具有至少90%同一性的氨基酸序列。

22.如权利要求18或19所述的cTuberin,其中所述C端区域包含与SEQ ID NO:12具有至少90%同一性的氨基酸序列。

23.如权利要求18所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:17具有至少90%同一性的氨基酸序列。

24.如权利要求18所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:18具有至少90%同一性的氨基酸序列。

25.如权利要求18所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:19具有至少90%同一性的氨基酸序列。

26.如权利要求18至25中任一项所述的cTuberin,其中所述N端区域包含与SEQ ID NO:5具有至少90%同一性的氨基酸序列。

27.如权利要求1至26中任一项所述的cTuberin,其中所述cTuberin在所述N端区域与所述C端区域之间包含间隔序列。

28.如权利要求27所述的cTuberin,其中所述间隔序列包含SEQ ID NO:2的序列。

29.如权利要求28所述的cTuberin,其中所述间隔序列包含SEQ ID NO:3的序列。

30.一种核酸分子,其编码如权利要求1至29中任一项所述的cTuberin。

31.如权利要求30所述的核酸分子,其中所述核酸分子经密码子优化以在人靶细胞中表达。

32.如权利要求31所述的核酸分子,其中所述人靶细胞是脑细胞、心脏细胞、肾细胞、皮肤细胞或肺细胞。

33.如权利要求30至32中任一项所述的核酸分子,其中所述核酸分子与调节控制序列可操作地连接。

34.如权利要求33所述的核酸分子,其中所述调节控制序列包含人巨细胞病毒(CMV)启动子、鸡β-肌动蛋白(CBA)启动子、劳斯肉瘤病毒(RSV)LTR启动子/增强子、SV40启动子、二氢叶酸还原酶启动子、磷酸甘油激酶启动子、CMV立即/早期基因增强子/CBA启动子、突触蛋白启动子或神经胶质纤维酸性蛋白(GFAP)启动子。

35.如权利要求33所述的核酸分子,其中所述调节控制序列包含人巨细胞病毒(CMV)立即/早期基因增强子/鸡β-肌动蛋白(CBA)启动子和土拨鼠肝炎病毒转录后调控元件(WPRE)。

36.如权利要求33所述的核酸分子,其中所述调节控制序列包含β-葡萄糖醛酸酶(GUSB)启动子。

37.如权利要求30至36中任一项所述的核酸分子,其中所述核酸分子与SEQ ID NO.21-26中的任一个具有至少90%序列同一性。

38.一种编码cTuberin的核酸分子,所述cTuberin包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基451至932;并且其中所述核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子的调节控制序列可操作地连接。

39.如权利要求38所述的核酸分子,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基451至1515。

40.如权利要求38或39所述的核酸分子,其中所述C端区域包含与SEQ ID NO:6具有至少90%同一性的氨基酸序列。

41.如权利要求38至40中任一项所述的核酸分子,其中所述N端区域包含与SEQ ID NO:4具有至少90%同一性的氨基酸序列。

42.一种核酸分子,其包含腺相关病毒(AAV)表达盒,所述AAV表达盒从5'至3'包含:

i)5'AAV反向末端重复序列(ITR);

ii)如权利要求25至35中任一项所述的核酸分子;以及

iii)3'AAV ITR。

43.如权利要求42所述的核酸分子,其中所述5'ITR和/或所述3'ITR来源于AAV2。

44.如权利要求42或43所述的核酸分子,其中所述5'AAV ITR序列包含与SEQ ID NO:27具有至少90%同一性的核酸序列。

45.如权利要求42至44中任一项所述的核酸分子,其中所述3'AAV ITR序列包含与SEQ ID NO:28具有至少90%同一性的核酸序列。

46.如权利要求42至45中任一项所述的核酸分子,其中所述AAV表达盒还包含聚腺苷酸化序列。

47.如权利要求42至46中任一项所述的核酸分子,其中所述AAV表达盒还包含Kozak序列。

48.一种质粒,其包含如权利要求30至47中任一项所述的核酸分子。

49.一种宿主细胞,其包含如权利要求30至47中任一项所述的核酸分子或如权利要求48所述的质粒。

50.一种组合物,其包含如权利要求30至47中任一项所述的核酸分子、如权利要求48所述的质粒或如权利要求49所述的宿主细胞。

51.一种产生重组腺相关病毒(rAAV)的方法,所述方法包括:使宿主细胞与如权利要求30至47中任一项所述的核酸分子或如权利要求48所述的质粒接触。

52.一种重组腺相关病毒(rAAV),其通过如权利要求51所述的方法产生。

53.一种重组腺相关病毒(rAAV),其包含:AAV衣壳蛋白;以及如权利要求30至47中任一项所述的核酸分子。

54.如权利要求52或53所述的rAAV,其中所述rAAV包含AAV1衣壳蛋白、AAV2衣壳蛋白、AAV3衣壳蛋白、AAV4衣壳蛋白、AAV5衣壳蛋白、AAV6衣壳蛋白、AAV7衣壳蛋白、AAV8衣壳蛋白、AAV9衣壳蛋白、AAV10衣壳蛋白、AAVrh10衣壳蛋白、AAV11衣壳蛋白和/或AAV12衣壳蛋白。

55.一种在靶细胞中表达cTuberin的方法,其包括:使所述靶细胞与如权利要求30至47中任一项所述的核酸分子、如权利要求48所述的质粒、如权利要求50所述的组合物或如权

利要求52至54中任一项所述的rAAV接触,从而在所述靶细胞中表达cTuberin。

56.如权利要求55所述的方法,其中所述接触步骤在体外、离体或体内进行。

57.如权利要求56所述的方法,其中所述接触步骤在有需要的受试者体内进行。

58.如权利要求57所述的方法,其中所述接触步骤包括向所述受试者施用治疗有效量的所述核酸分子、所述质粒、所述组合物或所述rAAV。

59.一种治疗患有结节性硬化综合征(TSC)的受试者的方法,其包括:向所述受试者施用治疗有效量的如权利要求1至29中任一项所述的cTuberin、如权利要求30至47中任一项所述的核酸分子、一种或多种包含如权利要求30至47中任一项所述的核酸分子的细胞外囊泡(EV)、如权利要求48所述的质粒、如权利要求50所述的组合物或如权利要求52至54中任一项所述的rAAV,从而治疗所述受试者中的TSC。

60.一种治疗患有肾癌的受试者的方法,其包括:向所述受试者施用治疗有效量的如权利要求1至29中任一项所述的cTuberin、如权利要求30至47中任一项所述的核酸分子、一种或多种包含如权利要求30至47中任一项所述的核酸分子的细胞外囊泡(EV)、如权利要求48所述的质粒、如权利要求50所述的组合物或如权利要求52至54中任一项所述的rAAV,从而治疗所述受试者中的肾癌。

61.如权利要求57至60中任一项所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV是通过血管内、肾动脉或静脉内、肺内、小脑延髓池内、脑内、鞘内、静脉内、心室内、脑室内、腹膜内或经皮施用。

62.如权利要求57至61中任一项所述的方法,其中所述受试者患有肾血管平滑肌脂肪瘤。

63.如权利要求62所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV靶向所述肾血管平滑肌脂肪瘤。

64.如权利要求57至63中任一项所述的方法,其中所述受试者表现出淋巴管平滑肌瘤病(LAM)。

65.如权利要求64所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV靶向所述LAM。

66.如权利要求57至65中任一项所述的方法,其中所述受试者患有脑功能障碍。

67.如权利要求66所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV被提供给蛛网膜下腔。

68.如权利要求57至67中任一项所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV被施用于脑细胞、心脏细胞、肾细胞、皮肤细胞或肺细胞。

69.如权利要求57至68中任一项所述的方法,其中所述受试者被施用雷帕霉素。

70.如权利要求57至69中任一项所述的方法,其中所述受试者是人。

71.如权利要求57至70中任一项所述的方法,其中所述受试者的年龄小于18岁。

72.如权利要求71所述的方法,其中所述受试者是婴儿。

73.如权利要求57至72中任一项所述的方法,其中所述受试者已被诊断为患有结节性硬化综合征。

74.如权利要求57至73中任一项所述的方法,其中所述受试者在TSC2基因中具有突变。

75.如权利要求74所述的方法,其中所述受试者在所述TSC2基因的外显子33、外显子37

和/或外显子38中具有突变。

76.如权利要求57至75中任一项所述的方法,其中所述受试者具有以下中的一种或多种:皮质结节、室管膜下结节和室管膜下巨细胞星形细胞瘤。

用于结节性硬化的基因疗法

[0001] 相关申请的交叉引用

[0002] 本申请要求2021年6月14日提交的美国临时申请第63/210,456号的优先权权益,该临时申请的内容以引用的方式整体并入本文中。

[0003] 以电子方式提交的文本文件的描述

[0004] 随本文以电子方式提交的文本文件的内容以引用的方式整体并入本文中:序列表的计算机可读格式副本(文件名:BGTR_003_01WO_SeqList_ST25.txt,创建日期:2022年6月14日,文件大小约170,338字节)。

背景技术

[0005] 结节性硬化综合征(Tuberous sclerosis complex;TSC),也称为结节性硬化,是一种多系统、常染色体显性遗传疾病,其可导致在脑中和其他重要器官中(如肾、心脏、肝、眼睛、肺和皮肤)中生长的非癌性肿瘤。在脑中,这类肿瘤可引起发育迟缓、孤独症、癫痫和脑积水。与结节性硬化相关的危及生命的病症包括肾血管平滑肌脂肪瘤(其可引起内出血)和淋巴管平滑肌瘤病(LAM)(其可损害呼吸)。

[0006] 结节性硬化是由TSC1基因和/或TSC2基因中的一个或多个突变引起。由TSC2中的突变引起的结节性硬化更严重且更普遍。TSC1和TSC2是分别编码蛋白质错构瘤蛋白(hamartin)和马铃薯球蛋白(tuberin)的肿瘤生长抑制基因。马铃薯球蛋白和错构瘤蛋白形成整合多种信号以主要通过抑制mTORC1复合物来调节雷帕霉素(rapamycin)的哺乳动物靶蛋白(mTOR)信号传导的蛋白质复合物。此外,马铃薯球蛋白还含有GTP酶活化结构域(GAP)结构域,其下调mTORC1活化剂Rheb。

[0007] 目前对结节性硬化的治疗涉及施用雷帕霉素和其类似物。然而,这些药物必须连续施用并且具有显著的副作用,包括受损的脑发育和免疫抑制。由于mTORC1活性的过度抑制,雷帕霉素和其类似物的施用也可能会引起不良事件。此外,一些患者对雷帕霉素没有反应,或者最初起反应并且接着变得抗性。

[0008] 因此,对可用于治疗结节性硬化的组合物和方法,特别是基于基因疗法的组合物和方法存在未满足的需求。

发明内容

[0009] 本公开提供了用于治疗结节性硬化综合征的组合物和方法。在一些实施方案中,本公开提供了凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的某些氨基酸残基。在一些实施方案中,本公开的cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451至932。在一些实施方案中,本公开的cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419至932。在一些实施方案中,本公开的cTuberin包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中C端区域包含与SEQ ID NO:7具有至少90%同一性的氨基酸序列,并且缺乏SEQ ID NO:1的氨基酸残基451

至932。在一些实施方案中,本公开的cTuberin包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中C端区域包含与SEQ ID NO:10-12中的一个具有至少90%同一性的氨基酸序列,并且其中cTuberin缺乏SEQ ID NO:1的氨基酸残基451至932。

[0010] 除了cTuberin蛋白之外,本公开提供了编码本文公开的任一种cTuberin蛋白的核酸分子以及被配置为引起本文公开的任一种cTuberin蛋白在给定细胞中表达的组合物。例如,本公开提供了腺相关病毒(AAV)表达盒,其从5′至3′包含:5’AAV反向末端重复序列(ITR);本文公开的任一种核酸分子;和3’AAV ITR。本公开还提供了重组AAV(rAAV),其包含:AAV衣壳蛋白以及本文公开的任一种核酸分子或AAV表达盒。此外,本公开提供了包含药物组合物的组合物,该药物组合物包含本文公开的任一种cTuberin蛋白、任一种核酸分子、任一种质粒、任一种宿主细胞或任一种rAAV。

[0011] 在另一方面,本公开提供了在靶细胞中表达本文公开的任一种cTuberin的方法,其包括:使本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种细胞外囊泡(EV)或本文公开的任一种组合物与靶细胞接触,从而在靶细胞中表达cTuberin。

[0012] 在另一方面,本公开提供了治疗有需要的受试者的结节性硬化的方法,其包括:向受试者施用治疗有效量的本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物,从而治疗受试者的结节性硬化。

具体实施方式

[0013] 由于与人马铃薯球蛋白的5.4kb cDNA相比,AAV载体的相对小的插入能力(~4.7kb),使用基于AAV的基因疗法来治疗结节性硬化或由TSC2中的突变引起的TSC而变得复杂。本公开提供了人马铃薯球蛋白基因(TSC2)的凝聚形式,其小到足以整合到AAV载体中并且编码凝聚的马铃薯球蛋白(cTuberin)。特别地,本公开提供了包含重组腺相关病毒(rAAV)的组合物,其包含AAV衣壳蛋白和编码cTuberin的AAV表达盒;以及其使用方法,包括在治疗结节性硬化中。

[0014] 应当理解,本文使用的术语仅用于描述特定实施方案的目的,而不旨在限制。

[0015] 除非另外定义,否则本文所用的所有技术和科学术语都具有与本申请所属领域中的普通技术人员通常所理解相同的含义。虽然可以在本申请的实践或测试中使用与本文所述的方法和材料类似或等效的任何方法和材料,但在本文中描述代表性方法和材料。

[0016] 当在本申请(包括权利要求书)中使用时,术语“一个/种”和“该/所述”是指“一个或多个”。因此,例如,提及“一种载剂”包括一种或多种载剂、两种或更多种载剂等的混合物,并且提及“所述方法”包括提及本领域技术人员已知的等同步骤和/或方法等。

[0017] 在本说明书中,除非另外指明,否则任何浓度范围、百分比范围、比率范围或整数范围应理解为包括所述范围内的任何整数的值,并且在适当时包括其分数(如整数的十分之一和百分之一)。术语“约”,当紧接在数字或数值之前时,意指数字或数值范围加或减0%至10%。

[0018] 同样如本文所用,“和/或”是指并且涵盖相关列出项目中的一个或多个的任何和所有可能的组合,以及当以替代方案(“或”)解释时缺乏组合。替代方案(例如,“或”)的使用

应理解为意指替代方案中的任一者、两者或其任何组合。

[0019] 如本文所用,“载剂”包括任何和所有溶剂、分散介质、媒介物、包衣、稀释剂、抗细菌剂和抗真菌剂、等渗剂和吸收延迟剂、缓冲剂、载剂溶液、悬浮液、胶体等。

[0020] 除非另外指明,否则术语“药学上可接受的”用于将部分(例如盐、剂型或赋形剂)表征为适于根据合理的医学判断使用。通常,药学上可接受的部分具有超过该部分可能具有的任何有害作用的一种或多种益处。有害作用可包括例如过度毒性、刺激、过敏反应和其他问题与并发症。

[0021] 本文所用的“治疗(treatment)”、“治疗(treating)”、“减轻”和“改善”可互换使用。这些术语是指获得有益或期望结果的方法,包括但不限于治疗益处和/或预防益处。治疗益处是指治疗下的一种或多种疾病、病症或症状的任何治疗相关的改善或效果。在一个实施方案中,术语“治疗”包括:(1)预防或延迟患者中发展的状态、障碍或病症的临床症状的出现,该患者可能患有或易患该状态、障碍或病症但尚未经历或显示该状态、障碍或病症的临床或亚临床症状;(2)抑制该状态、障碍或病症(例如,阻止、减少或延迟疾病的发展、或在维持治疗的情况下其至少一种临床或亚临床症状的复发);和(3)减轻病症(例如,通过引起该状态、障碍或病症或其临床或亚临床症状中的至少一种的消退、或减轻其严重程度)。例如,有益的临床结果包括但不限于延迟或减缓肿瘤或错构瘤的侵袭性或生长,以及改善与这类肿瘤或错构瘤相关的症状。例如,在肾血管平滑肌脂肪瘤的情况下,可以通过磁共振成像(MRI)监测肿瘤大小,并且可以根据标准程序,如用于监测使用雷帕霉素治疗结节性硬化的那些程序,分析由于施用本文公开的任一种组合物而导致的细胞大小的收缩。治疗还包括与未接受治疗的受试者相比降低受试者的死亡率或增加受试者的寿命。

[0022] 术语“有效量”或“治疗有效量”是指足以实现例如实现有益或所需结果(如治疗结节性硬化或其症状)的结果的剂的量。治疗有效量可根据以下中的一种或多种而变化:所治疗的受试者和疾病状况、受试者的体重和年龄、疾病状况的严重程度、施用方式等。治疗有效量可以是足以治疗结节性硬化和/或改善、减轻结节性硬化的严重程度、消除结节性硬化的一种或多种症状和/或延迟其发作的量。在一些实施方案中,治疗有效量可以是足以在受试者中表达马铃薯球蛋白(例如,缺乏一个或多个突变的马铃薯球蛋白,如本文提供的凝聚的马铃薯球蛋白)的量。

[0023] 术语“受试者”、“个体”和“患者”在本文中可互换使用,是指脊椎动物,如哺乳动物。哺乳动物可以是例如小鼠、大鼠、兔、猫、狗、猪、绵羊、马、非人灵长类动物(例如食蟹猴、黑猩猩)或人。还涵盖体内获得或体外培养的受试者的组织、细胞或其衍生物。在一些实施方案中,受试者是人。人受试者可以是成人、青少年(例如,12岁至18岁)、儿童(例如,2岁至14岁)、婴儿(例如,1个月至24个月大)或新生儿(至多1个月大)。在一些实施方案中,成人是约60岁或更年长的年长者,如约65岁或更年长。在一些实施方案中,受试者是孕妇或打算怀孕的妇女。在一些实施方案中,受试者小于18岁。

[0024] “腺相关病毒(AAV)表达盒”是包装到重组AAV载体中的核酸,并且包含编码侧接有5′反向末端重复序列(ITR)和3’ ITR的一个或多个转基因的序列。

[0025] 如本文所用,术语“病毒载体(virus vector)”、“病毒载体(viral vector)”和“基因递送载体”是指起核酸递送媒介物作用的病毒颗粒,并且其包含包装在病毒粒子内的核酸分子(例如AAV表达盒)。示例性病毒载体包括腺相关病毒载体(AAV)。

[0026] 如本文所用,术语“腺相关病毒”(AAV)包括但不限于1型AAV(例如,血清型1的AAV,也称为AAV1)、2型AAV(例如,AAV2)、3型AAV(例如,AAV3,包括3A和3B型:AAV3A和AAV3B)、4型AAV(例如,AAV4)、5型AAV(例如,AAV5)、6型AAV(例如,AAV6)、7型AAV(例如,AAV7)、8型AAV(例如,AAV8)、9型AAV(例如,AAV9)、10型AAV(例如,AAV10)、11型AAV(例如,AAV11)、12型AAV(例如,AAV12)、13型AAV(例如,AAV13)、rh32.33型AAV(例如,AAVrh32.33)、rh8型AAV(例如,AAVrh8)、rh10型AAV(例如,AAVrh10)、rh74型AAV(例如,AAVrh74)、hu.68型AAV(例如,AAVhu.68)、禽AAV(例如,AAAV)、牛AAV(例如,BAAV)、犬AAV、马AAV、绵羊AAV、蛇AAV、鬃狮蜥AAV、AAV2i8、AAV2g9、AAV-LK03、AAV7m8、AAV Anc80、AAV PHP.B,以及现在已知或以后发现的任何其他AAV。

[0027] 如本文所用,“序列同一性”是指两个最佳比对的多核苷酸或多肽序列在组分(例如核苷酸或氨基酸)的比对窗口中不变的程度。测试序列和参考序列的比对片段的“同一性分数”是两个比对序列共有的相同组分的数量除以参考序列片段(即,整个参考序列或参考序列的较小限定部分)中组分的总数。“百分比同一性”是同一性分数乘以100。可以使用计算机程序和数学算法确定两个序列之间的同一性(同源性)程度。可以使用在www.ebi.ac.uk/Tools/msa/clustalo可获得的比对程序Clustal Omega,使用默认参数计算百分比同一性。参见Sievers等人,“Fast,scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.”(2011年10月11日)Molecular systems biology 7:539。出于计算对于一个序列的同一性的目的,不包括如标签的延伸部分。

[0028] 如本文所用,当核酸序列(例如,编码序列)和调节序列以将核酸序列的表达或转录置于调节序列的影响或控制之下的方式共价连接时,它们被称为“可操作地连接”。如果需要将核酸序列翻译成功能性蛋白质,如果5′调节序列中启动子的诱导导致编码序列的转录,并且如果两个DNA序列之间的连接的性质不会(1)导致引入移码突变,(2)干扰启动子区域引导编码序列转录的能力,或(3)干扰相应RNA转录物翻译成蛋白质的能力,那么称两个DNA序列可操作地连接。

[0029] 如本文所用,“密码子优化”是指修饰核酸序列以改变单个核酸而相应的所编码氨基酸的无任何改变。以这种方式修饰的序列在本文中称为“密码子优化的”。进行密码子优化的方法描述于美国专利第7,561,972号、第7,561,973号和第7,888,112号中,其各自出于所有目的以引用的方式整体并入本文中。在一些实施方案中,翻译起始位点周围的序列可以转化为共有Kozak序列,如Kozak等人,Nucleic Acids Res.15(20):8125-81 48(1987)中进一步所描述,其出于所有目的以引用的方式整体并入本文中。

[0030] 凝聚的马铃薯球蛋白(cTuberin)

[0031] 如本文所用,凝聚的马铃薯球蛋白或cTuberin是指与天然马铃薯球蛋白序列相比具有一个或多个氨基酸残基缺失的重组马铃薯球蛋白。在一些实施方案中,天然马铃薯球蛋白是人马铃薯球蛋白。在一些实施方案中,天然马铃薯球蛋白序列具有SEQ ID NO:1的氨基酸序列,该序列包括1807个氨基酸残基。在一些实施方案中,本文提供的cTuberin缺乏SEQ ID NO:1的至少一个氨基酸残基。在一些实施方案中,本文提供的cTuberin从N端与C端区域之间的区域缺乏SEQ ID NO:1的至少一个氨基残基。

[0032] 在一些实施方案中,人马铃薯球蛋白的氨基酸和核酸序列分别见于NCBI登录号

NP_000539.2和GenBank登录号X75621.1。在一些实施方案中,人马铃薯球蛋白的氨基酸序列包括但不限于马铃薯球蛋白同种型4(NCBI登录号NP_001070651.1)、马铃薯球蛋白同种型5(NCBI登录号NP_001107854.1)、马铃薯球蛋白同种型6(NCBI登录号NP_001305756.1)、马铃薯球蛋白同种型7(NCBI登录号NP_001305758.1)、马铃薯球蛋白同种型8(NCBI登录号NP_001305760.1)、马铃薯球蛋白同种型9(NCBI登录号NP_001305761.1)、马铃薯球蛋白同种型X7(NCBI登录号XP_024306181.1)、马铃薯球蛋白同种型X8(NCBI登录号XP_005255586.2)、马铃薯球蛋白同种型X9(NCBI登录号XP_016879105.1)、马铃薯球蛋白同种型X10(NCBI登录号XP_005255588.2)、马铃薯球蛋白同种型X11(NCBI登录号XP_016879106.1)、马铃薯球蛋白同种型X12(NCBI登录号XP_016879107.1)等。

[0033] 在一些实施方案中,cTuberin包含能够结合错构瘤蛋白的N端区域。在一些实施方案中,cTuberin包含C端GTP酶活化蛋白(GAP)区域。在一些实施方案中,在cTuberin中缺失的一个或多个氨基酸残基位于能够结合错构瘤蛋白的N端区域与C端GTP酶活化蛋白(GAP)区域之间。

[0034] 在一些实施方案中,cTuberin包含能够结合错构瘤蛋白的N端区域和C端GTP酶活化蛋白(GAP)区域或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4或5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如约100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4或5具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4或5的氨基酸序列或由其组成。

[0035] 在一些实施方案中,C端区域包含与SEQ ID NO:7-12中的任一个具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含与SEQ ID NO:7-12中的任一个具有至少约90%同一性的氨基酸序列或由组成。在一些实施方案中,C端区域包含SEQ ID NO:7-12中任一个的氨基酸序列或由其组成。

[0036] 在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约75%(例如,约80%,约85%,约90%,约95%,约96%,约97%,约98%,约99%或100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:7具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:7具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:7的氨基酸序列或由其组成。

[0037] 在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:7具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:7具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:5的氨基酸序列或由

其组成;并且C端区域包含SEQ ID NO:7的氨基酸序列或由其组成。

[0038] 在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:8具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:8具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:8的氨基酸序列或由其组成。

[0039] 在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:8具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:8具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:5的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:8的氨基酸序列或由其组成。

[0040] 在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:9具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:9具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:9的氨基酸序列或由其组成。

[0041] 在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:9具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:9具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:5的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:9的氨基酸序列或由其组成。

[0042] 在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:10具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:10具有至少约90%同一性

的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:10的氨基酸序列或由其组成。

[0043] 在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:10具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:10具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:5的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:10的氨基酸序列或由其组成。

[0044] 在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:11具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:11具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:11的氨基酸序列或由其组成。

[0045] 在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:11具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:11具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:5的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:11的氨基酸序列或由其组成。

[0046] 在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:12具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:12具有至少约90%同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:12的氨基酸序列或由其组成。

[0047] 在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:12具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含与SEQ ID NO:5具有至少约90%

同一性的氨基酸序列或由组成;并且C端区域包含与SEQ ID NO:12具有至少约90％同一性的氨基酸序列或由其组成。在一些实施方案中,N端区域包含SEQ ID NO:5的氨基酸序列或由其组成;并且C端区域包含SEQ ID NO:12的氨基酸序列或由其组成。

[0048] 本公开提供了包含以下或由以下组成的cTuberin:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932。本公开还提供了包含以下或由以下组成的cTuberin:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932。当表示氨基酸范围时,这些范围包括端点(例如,缺乏氨基酸残基“419-932”或“419至932”的cTuberin是指缺乏SEQ ID NO:1的氨基酸419和932以及位于它们之间的所有氨基酸残基的cTuberin)。

[0049] 在一些实施方案中,cTuberin还缺乏SEQ ID NO:1的氨基酸残基947-988。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932和947-988。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932和氨基酸残基947-988。

[0050] 在一些实施方案中,cTuberin还缺乏SEQ ID NO:1的氨基酸残基1205-1271。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932和1205-1271。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932和1205-1271。

[0051] 在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932、947-988和1205-1271。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932、947-988和1205-1271。

[0052] 在一些实施方案中,cTuberin还缺乏SEQ ID NO:1的氨基酸残基1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932和1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932和1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932、947-988、1205-1271和1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932、947-988、1205-1271和1336-1497。

[0053] 在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932、1205-1271和1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932、1205-1271和1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932、947-988和1336-1497。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基451-932、947-988和1336-1497。

[0054] 在一些实施方案中,cTuberin还缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基933-1109。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932和933-1109。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932和933-1109。在一些实施方案中,cTuberin缺乏SEQ ID NO:1的氨基酸残基419-1109。

[0055] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1109。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1139。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基

451-1514。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白的氨基酸残基419-932。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-1109。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-1139。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-1514。

[0056] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的一个或多个外显子的氨基酸残基,如外显子25、30和/或33。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子25的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子30的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子33的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子25和30的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子25、30和33的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932以及外显子25的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932以及外显子25和30的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932以及外显子25、30和33的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932以及外显子25的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932以及外显子25和30的氨基酸残基。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932以及外显子25、30和33的氨基酸残基。

[0057] 本公开提供了凝聚的马铃薯球蛋白(cTuberin),其包含以下或由以下组成:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该C端区域包含以下或由以下组成:与SEQ ID NO:7具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列,并且其中cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。本公开还提供了凝聚的马铃薯球蛋白(cTuberin),其包含以下或由以下组成:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该C端区域包含与SEQ ID NO:7具有至少约90%同一性的氨基酸序列或由其组成,并且其中该cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。本公开还提供了凝聚的马铃薯球蛋白(cTuberin),其包含以下或由以下组成:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该C端区域包含SEQ ID NO:7的氨基酸序列或由其组成,并且其中该cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0058] 在一些实施方案中,C端区域包含与SEQ ID NO:8具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,例如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含与SEQ ID NO:8具有至少约90%同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含SEQ ID NO:8的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0059] 在一些实施方案中,C端区域包含与SEQ ID NO:9具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,例如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含以下或由以下组成:与SEQ ID NO:9具有至少约90%同一性的氨基酸序列,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含SEQ ID NO:9的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0060] 在一些实施方案中,C端区域包含与SEQ ID NO:7具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,如100%)同一性的氨基酸序列或由其组成,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932,并且N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%,约95%,约96%,约97%,约98%,约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含与SEQ ID NO:7具有至少约90%同一性的氨基酸序列或由其组成,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932,并且N端区域包含与SEQ ID NO:4具有至少90%同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含SEQ ID NO:7的氨基酸序列或由其组成,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932,并且N端区域包含SEQ ID NO:4的氨基酸序列或由其组成。

[0061] 在一些实施方案中,C端区域包含与SEQ ID NO:7具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,如100%)同一性的氨基酸序列或由其组成,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932,并且N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%,约95%,约96%,约97%,约98%,约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含与SEQ ID NO:7具有至少约90%同一性的氨基酸序列或由其组成,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932,并且N端区域包含与SEQ ID NO:5具有至少90%同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含SEQ ID NO:7的氨基酸序列或由其组成,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932,并且N端区域包含SEQ ID NO:5的氨基酸序列或由其组成。

[0062] 在另一方面,本公开提供了凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该C端区域包含与SEQ ID NO:10-12中的一个具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成。本公开提供了凝聚的马铃薯球蛋白(cTuberin),其包含以下或由以下组成:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该C端区域包含与SEQ ID NO:10-12中的一个具有至少约90%同一性的氨基酸序列或由其组成。本公开还提供了凝聚的马铃薯球蛋白(cTuberin),其包含以下或由以下组成:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该C端区域包含SEQ ID NO:10-12中的任一个的氨基酸序列或

由其组成。

[0063] 在一些实施方案中,C端区域包含与SEQ ID NO:10-12中的一个具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含与SEQ ID NO:10-12中的一个具有至少约90%同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含SEQ ID NO:10-12中的任一个的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0064] 在一些实施方案中,C端区域包含与SEQ ID NO:10具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,例如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含与SEQ ID NO:10具有至少约90%同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含SEQ ID NO:10的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0065] 在一些实施方案中,C端区域包含与SEQ ID NO:11具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,例如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含与SEQ ID NO:11具有至少约90%同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含SEQ ID NO:11的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0066] 在一些实施方案中,C端区域包含与SEQ ID NO:12具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,例如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含与SEQ ID NO:12具有至少约90%同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。在一些实施方案中,C端区域包含SEQ ID NO:12的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932。

[0067] 在一些实施方案中,C端区域包含与SEQ ID NO:10-12中的一个具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更高,如100%)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932,并且N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%,约95%,约96%,约97%,约98%,约99%或更大,如100%)同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含与SEQ ID NO:10-12中

的一个具有至少约90％同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932,并且N端区域包含与SEQ ID NO:5具有至少约90％同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含SEQ ID NO:10-12中的任一个的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932,并且N端区域包含SEQ ID NO:5的氨基酸序列或由其组成。

[0068] 在一些实施方案中,C端区域包含与SEQ ID NO:10-12中的一个具有至少约75％(例如,至少约80％、约85％、约90％、约95％、约96％、约97％、约98％、约99％或更高,如100％)同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932,并且N端区域包含与SEQ ID NO:4具有至少约75％(例如,至少约80％、约85％、约90％,约95％,约96％,约97％,约98％,约99％或更大,如100％)同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含与SEQ ID NO:10-12中的一个具有至少约90％同一性的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932,并且N端区域包含与SEQ ID NO:4具有至少约90％同一性的氨基酸序列或由其组成。在一些实施方案中,C端区域包含SEQ ID NO:10-12中的任一个的氨基酸序列或由其组成,并且cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932,并且N端区域包含SEQ ID NO:4的氨基酸序列或由其组成。

[0069] 在一些实施方案中,cTuberin在N端区域与C端区域之间包含多肽间隔区。在一些实施方案中,多肽间隔区包含SEQ ID NO:2(SGGG)的序列或由其组成。在一些实施方案中,多肽间隔区包含SEQ ID NO:3(SGGGSGGG SGGGSGGG)的序列或由其组成。

[0070] 在一些实施方案中,cTuberin包含与SEQ ID NO:14-19中的任一个具有至少约75％(例如,至少约80％、约85％、约90％、约95％、约96％、约97％、约98％、约99％或更大,如100％)同一性的氨基酸序列或由其组成。在一些实施方案中,cTuberin包含与SEQ ID NO:14-19中的任一个具有至少约90％同一性的氨基酸序列或由组成。在一些实施方案中,cTuberin包含SEQ ID NO:14-19中任一个的氨基酸序列或由其组成。

[0071] 编码cTuberin的AAV表达盒

[0072] 除了本文所述的氨基酸序列和相应的cTuberin以外,本公开还提供了编码本文公开的任一种cTuberin蛋白的核酸分子。在一些实施方案中,核酸分子经密码子优化以在人靶细胞中表达。在一些实施方案中,人靶细胞是脑细胞、心脏细胞、肾细胞、皮肤细胞或肺细胞。

[0073] 在一些实施方案中,核酸分子与调节控制序列可操作地连接。在一些实施方案中,调节控制序列包含人巨细胞病毒(CMV)启动子、鸡β-肌动蛋白(CBA)启动子、劳斯肉瘤病毒(RSV)LTR启动子/增强子、SV40启动子、二氢叶酸还原酶启动子、磷酸甘油激酶启动子、CMV立即/早期基因增强子/CBA启动子、突触蛋白启动子、CMV-IE启动子/增强子、神经胶质纤维酸性蛋白(GFAP)启动子或其组合。在一些实施方案中,调节控制序列包含CMV立即/早期基因增强子/CBA启动子和土拨鼠肝炎病毒转录后调控元件(WPRE)。在一些实施方案中,调节控制序列包含β-葡萄糖醛酸酶(GUSB)启动子。关于GUSB启动子的其他细节提供于Shipley等人,Analysis of the 5'Flanking Region of the Humanβ-Glucuronidase Gene,Genomics 10,1009-1018(1991)中,其内容以引用的方式整体并入本文中。

[0074] 在一些实施方案中,核酸分子包含与SEQ ID NO:20-26中的任一个的至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)序列同一性或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:20-26中的任一个的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:20的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:21的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:22的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:23的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:24的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:25的序列或由其组成。在一些实施方案中,核酸分子包含SEQ ID NO:26的序列或由其组成。

[0075] 本公开提供了编码cTuberin的核酸分子,该cTuberin包含以下或由以下组成:(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中该cTuberin至少缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932;并且其中该核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0076] 在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1109的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1139的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1514的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ IDNO:1)的氨基酸残基419-1109的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-1139的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-1514的cTuberin,并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子25的氨基酸残基的cTuberin。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子30的氨基酸残基的cTuberin。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子33的氨基酸残基的cTuberin。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子25和30的氨基酸残基的cTuberin。在一些实施方案中,核酸分子编码缺乏人马铃薯球蛋白(SEQ ID NO:1)的外显子25、30和33的氨基酸残基的cTuberin。

[0077] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109、419-1139、451-1109、451-1139、419-1514、451-1514、419-1515或451-1515;C端区域包含与SEQ ID NO:6具有至少约75%(例如,至少约80%、约85%、

约90％、约95％、约96％、约97％、约98％、约99％或更大,如100％)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1514或451-1515;C端区域包含与SEQ ID NO:6具有至少约90％同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1514或451-1515;C端区域包含SEQ ID NO:6的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0078] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109、419-1139、451-1109或451-1139;C端区域包含与SEQ ID NO:7具有至少约75％(例如,至少约80％、约85％、约90％、约95％、约96％、约97％、约98％、约99％或更大,如100％)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109、419-1139、451-1109或451-1139;C端区域包含与SEQ ID NO:7具有至少约90％同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109、419-1139、451-1109或451-1139;C端区域包含SEQ ID NO:7的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0079] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109或451-1109;C端区域包含与SEQ ID NO:8具有至少约75％(例如,至少约80％、约85％、约90％、约95％、约96％、约97％、约98％、约99％或更大,如100％)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109或451-1109;C端区域包含与SEQ ID NO:8具有至少约90％同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、451-932、419-1109或451-1109;C端区域包含SEQ ID NO:8的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0080] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932;C端区域包含与SEQ ID NO:9具有至少约75％(例如,至少约80％、约85％、约90％、约95％、约96％、约97％、约98％、约99％或更大,如100％)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932;C端区域包含与SEQ ID NO:9具有至少约90％同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的

氨基酸残基419-932或451-932;C端区域包含SEQ ID NO:9的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0081] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932;C端区域包含与SEQ ID NO:10-12中的一个具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932;C端区域包含与SEQ ID NO:10-12中的一个具有至少约90%同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932或451-932;C端区域包含SEQ ID NO:10-12中的一个的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0082] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932、451-1109、451-1139、451-1514、451-1515;N端区域包含与SEQ ID NO:4具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932、451-1109、451-1139、451-1514、451-1515;N端区域包含与SEQ ID NO:4具有至少约90%同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-932、451-1109、451-1139、451-1514或451-1515;N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0083] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、419-1109、419-1139、419-1514或419-1515;N端区域包含与SEQ ID NO:5具有至少约75%(例如,至少约80%、约85%、约90%、约95%、约96%、约97%、约98%、约99%或更大,如100%)同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、419-1109、419-1139、419-1514或419-1515;N端区域包含与SEQ ID NO:5具有至少约90%同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基419-932、419-1109、419-1139、419-1514或419-1515;N端区域包含SEQ ID NO:5的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0084] 在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基451-1514或451-1515;C端区域包含与SEQ ID NO:6具有至少90%同一性的氨基酸序列或由

其组成;N端区域包含与SEQ ID NO:4具有至少90％同一性的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。在一些实施方案中,cTuberin缺乏人马铃薯球蛋白(SEQ ID NO:1)的氨基酸残基或451-1514或451-1515;C端区域包含SEQ ID NO:6的氨基酸序列或由其组成;N端区域包含SEQ ID NO:4的氨基酸序列或由其组成;并且核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子或由其组成的调节控制序列可操作地连接。

[0085] 在一些实施方案中,核酸分子包含与SEQ ID NO:20-26中的任一个具有至少约75％(例如,至少约80％、约85％、约90％、约95％、约96％、约97％、约98％、约99％或更大,如100％)序列同一性的核酸序列。在一些实施方案中,核酸分子包含与SEQ ID NO:20-26中的任一个具有至少约90％序列同一性的核酸序列。

[0086] 在一些实施方案中,核酸分子包含腺相关病毒(AAV)表达盒,该AAV表达盒从5'至3'包含:5'AAV反向末端重复序列(ITR);本文公开的核酸分子中的任一种;和3'AAV ITR。在一些实施方案中,5'ITR和/或3'ITR来源于AAV2。

[0087] 在一些实施方案中,本文公开的AAV表达盒包含顺式作用的5'和3'反向末端重复序列,B.J.Carter,"Handbook of Parvoviruses"版,P.Tijsser,CRC Press,第155-168页(1990)中进一步描述,该文献出于所有目的以引用的方式整体并入本文中。AAV ITR序列可以从任何已知的AAV获得,包括本文公开的目前鉴定的哺乳动物AAV类型。

[0088] 在一些实施方案中,AAV表达盒包含来自以下各者的5'ITR和/或3'ITR:1型AAV、2型AAV、3型AAV(包括3A 3B型)、4型AAV、5型AAV、6型AAV、7型AAV、8型AAV、9型AAV、10型AAV、11型AAV、12型AAV、13型AAV、rh32.33型AAV、rh8型AAV、rh10型AAV、rh74型AAV、hu.68型AAV、禽AAV、牛AAV、犬AAV、马AAV、绵羊AAV、蛇AAV、鬃狮蜥AAV、AAV2i8、AAV2g9、AAV-LK03、AAV7m8、AAV Anc80或AAV PHP.B。在一些实施方案中,AAV表达盒包含来自AAV2的5'ITR、来自AAV2的3'ITR或其组合。在一些实施方案中,AAV表达盒包含来源于AAV2的5'ITR、来源于AAV2的3′ITR或其组合。

[0089] 在一些实施方案中,5′AAV ITR序列包含与SEQ ID NO:27的序列具有至少80％(例如,至少约85％、约90％、约95％、约96％、约97％、约98％、约99％、约99.5％或更大,如100％,包括位于其间的所有值和子范围)同一性的核酸序列或由其组成。在一些实施方案中,5′AAV ITR序列包含SEQ ID NO:27的序列或由其组成。

[0090] 在一些实施方案中,3'AAV ITR序列包含与SEQ ID NO:28的序列具有至少80％(例如,至少约85％、约90％、约95％、约96％、约97％、约98％、约99％、约99.5％或更大,如100％,包括位于其间的所有值和子范围)同一性的核酸序列或由其组成。在一些实施方案中,3′AAV ITR序列包含SEQ ID NO:28的核酸序列或由其组成。

[0091] 在一些实施方案中,本文公开的AAV表达盒包含与转基因可操作地连接的另外的表达控制元件。表达控制元件包括例如合适的转录起始、终止和增强子序列;有效的RNA加工信号,如剪接和聚腺苷酸化(polyA)信号;稳定细胞质mRNA的序列;增强翻译效率的序列;增强蛋白质稳定性的序列;以及增强所编码产物分泌的序列。

[0092] 在一些实施方案中,本文公开的AAV表达盒包含内含子。在一些实施方案中,内含子位于启动子/增强子序列与转基因之间。在一些实施方案中,内含子来源于SV-40,并被称为SV-40T内含子序列。在一些实施方案中,本文公开的AAV表达盒包含内部核糖体进入位点

(IRES)。在一些实施方案中,本文公开的AAV表达盒包含编码2A自裂解肽的核酸。说明性2A自裂解肽包括P2A、E2A、F2A和T2A。在一些实施方案中,本文公开的AAV表达盒包含Sambrook等人和其中引用的参考文献中描述的元件,例如第3.183.26和16.1716.27页,以及Ausubel等人,Current Protocols in Molecular Biology,John Wiley&Sons,New York,1989,其各自出于所有目的以引用的方式整体并入本文中。

[0093] 在一些实施方案中,本文公开的AAV表达盒包含土拨鼠肝炎病毒转录后元件(WPRE)。(参见例如Wang和Verma,Proc.Natl.Acad.Sci.,USA,96:3906-3910(1999))。在一些实施方案中,本文公开的AAV表达盒包含乙型肝炎病毒转录后调控元件(HBVPRE)或RNA转运元件(RTE)。在一些实施方案中,WPRE或HBVPRE序列是美国专利第6,136,597号或第6,287,814号中公开的WPRE或HBVPRE序列中的任一种,这两篇专利均以引用的方式整体并入本文中。

[0094] 在一些实施方案中,本文公开的AAV表达盒包含分别涉及转录和翻译起始的5'非转录和5'非翻译序列,如TATA盒、加帽序列、CAAT序列、增强子元件等。在一些实施方案中,本文公开的AAV表达盒包含增强子序列或上游活化因子序列。在一些实施方案中,本文公开的AAV表达盒包含5'前导序列或信号序列。

[0095] 在一些实施方案中,本文公开的AAV表达盒包含组成型启动子。组成型启动子的示例包括但不限于逆转录病毒劳斯肉瘤病毒(Rous sarcoma virus;RSV)LTR启动子(任选地具有RSV增强子)、巨细胞病毒(CMV)启动子(任选地具有CMV增强子)、SV40启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子和EFla启动子。

[0096] 在一些实施方案中,本文公开的AAV表达盒包含诱导型启动子。诱导型启动子的非限制性示例包括锌诱导型绵羊金属硫蛋白(MT)启动子、地塞米松(dexamethasone;Dex)诱导型小鼠乳腺肿瘤病毒(MMTV)启动子、T7聚合酶启动子系统、蜕皮激素昆虫启动子、四环素可抑制型系统、四环素诱导型系统、RU486诱导型系统和雷帕霉素诱导型系统。其他类型的诱导型启动子包括受特定生理状态调节的启动子,该生理状态例如温度、急性期、细胞的特定分化状态或特定细胞周期阶段。

[0097] 在一些实施方案中,本文公开的AAV表达盒包含天然启动子或其片段,或与编码cTuberin的转基因可操作地连接的天然表达控制元件。在一些实施方案中,本文公开的AAV表达盒包含赋予组织特异性基因表达能力的调节序列。在一些情况下,组织特异性调节序列结合以组织特异性方式诱导转录的组织特异性转录因子。组织特异性调节序列的示例包括但不限于以下组织特异性启动子:神经元启动子,如神经元特异性烯醇化酶(NSE)启动子;神经丝轻链基因启动子;和神经元特异性vgf基因启动子。

[0098] 在一些实施方案中,AAV表达盒包含一个或多个启动子。在一些实施方案中,AAV表达盒包含鸡β-肌动蛋白启动子。在一些实施方案中,AAV表达盒包含CB6启动子。在一些实施方案中,CB6启动子包含与SEQ ID NO:34的序列具有至少约80%(例如,至少约85%、约90%、约95%、约96%、约97%、约98%、约99%、约99.5%或更大,如100%,包括位于其间的所有值和子范围)同一性的核酸序列。在一些实施方案中,CB6启动子包含SEQ ID NO:34的核酸序列或由其组成。

[0099] 在一些实施方案中,AAV表达盒包含CMV-IE增强子。在一些实施方案中,增强子是CMV-IE增强子。在一些实施方案中,CMV-IE增强子包含与SEQ ID NO:33的序列具有至少约

80％(例如,至少约85％、约90％、约95％、约96％、约97％、约98％、约99％、约99.5％或更大,如约100％,包括位于其间的所有值和子范围)同一性的核酸序列。在一些实施方案中,CMV-IE增强子包含SEQ ID NO:33的核酸序列或由其组成。

[0100] 在一些实施方案中,AAV表达盒包含共有序列,如Kozak序列(例如,转录为RNA Kozak序列的DNA序列)。在一些实施方案中,AAV表达盒包含Kozak序列。在一些实施方案中,Kozak序列包含与SEQ ID NO:35的序列具有至少约80％(例如,至少约85％、约90％、约95％、约96％、约97％、约98％、约99％、约99.5％或更大,如100％,包括位于其间的所有值和子范围)同一性的核酸序列。在一些实施方案中,Kozak序列包含SEQ ID NO:35的核酸序列或由其组成。

[0101] 在一些实施方案中,本文公开的AAV表达盒包含一个或多个微小RNA(miRNA)的一个或多个结合位点。在一些实施方案中,AAV表达盒包含能够调节cTuberin转基因的组织特异性表达的miRNA结合位点。在一些实施方案中,能够调节cTuberin转基因的组织特异性表达的miRNA结合位点是miR-122结合位点、miR-133a或miR-1结合位点。例如,cTuberin转基因在肝脏中的表达可通过并入miR-122的结合位点来抑制,使得由转基因表达的mRNA结合肝脏中的miR-122并被其抑制。cTuberin转基因在心脏中的表达可通过并入miR-133a或miR-1的结合位点来抑制,使得由转基因表达的mRNA结合心脏中的miR-133a或miR-1并被其抑制。在一些实施方案中,mRNA中的miRNA靶位点在5′非翻译区(UTR)、3'UTR或编码区中。此外,可以设计cTuberin转基因,使得多个miRNA通过识别相同或多个位点来调节mRNA。多个miRNA结合位点的存在可导致多个RNA诱导的沉默复合物(RISC)的协同作用并提供对表达的高效抑制。靶位点序列可包含总共至少5个、10个或更多个核苷酸,如5-100或10-60个核苷酸。靶位点序列可包含靶基因结合位点的序列的至少5个核苷酸。在一些实施方案中,AAV表达盒包含miR-1结合位点、miR-133a结合位点、miR-122结合位点或其任何组合。

[0102] 在一些实施方案中,AAV表达盒包含聚腺苷酸化(polyA)序列。PolyA信号可以来源于许多合适的物种,包括但不限于SV-40、人和牛。在一些实施方案中,polyA序列是β-珠蛋白polyA序列,如哺乳动物β-珠蛋白polyA序列。在一些实施方案中,polyA序列是人polyA序列或牛β-珠蛋白polyA序列。在一些实施方案中,AAV表达盒包含兔β-珠蛋白polyA序列。在一些实施方案中,兔β-珠蛋白polyA序列包含SEQ ID NO:36的核酸序列或由其组成。

[0103] 在一些实施方案中,AAV表达盒从5'至3'包含:(i)5′基于AAV2的ITR,(ii)CMV-IE增强子,(iii)CB6启动子,(iv)编码本文公开的任一种cTuberin蛋白的转基因,(v)聚腺苷酸化序列,和(vi)3′基于AAV2的ITR。在一些实施方案中,AAV表达盒从5'至3'包含:(i)5′基于AAV2的ITR,其包含SEQ ID NO:27的核酸序列;(ii)CMV-IE增强子,其包含SEQ ID NO:33的核酸序列,(iii)CB6启动子,其包含SEQ ID NO:34的核酸序列;(iv)编码本文公开的任一种cTuberin蛋白的转基因;(v)聚腺苷酸化序列;和(vi)3′基于AAV2的ITR,其包含SEQ ID NO:28的核酸序列。

[0104] 在一些实施方案中,AAV表达盒从5′至3′包含:(i)5′基于AAV2的ITR,(ii)CB6启动子,(iii)编码本文公开的任一种cTuberin蛋白的转基因,(iv)聚腺苷酸化序列,和(v)3′基于AAV2的ITR。在一些实施方案中,AAV表达盒从5′至3′包含:(i)5′基于AAV2的ITR,其包含SEQ ID NO:27的核酸序列;(ii)CB6启动子,其包含SEQ ID NO:34的核酸序列;(iii)编码本文公开的任一种cTuberin蛋白的转基因;(iv)聚腺苷酸化序列;和(v)3'基于AAV2的ITR,其

包含SEQ ID NO:28的核酸序列。

[0105] 在一些实施方案中,AAV表达盒从5'至3'包含:(i)5'基于AAV2的ITR,(ii)GUSB启动子,(iii)编码本文公开的任一种cTuberin蛋白的转基因,(iv)聚腺苷酸化序列,和(v)3'基于AAV2的ITR。在一些实施方案中,AAV表达盒从5′至3′包含:(i)5′基于AAV2的ITR,其包含SEQ ID NO:27的核酸序列;(ii)GUSB启动子;(iii)编码本文公开的任一种cTuberin蛋白的转基因;(iv)聚腺苷酸化序列;和(v)3′基于AAV2的ITR,其包含SEQ ID NO:28的核酸序列。

[0106] 用于治疗结节性硬化的重组腺相关病毒(rAAV)

[0107] 本公开还提供了包含本文公开的任一种核酸分子的质粒,以及包含本文公开的任一种核酸分子或质粒的宿主细胞。

[0108] 本公开还提供了产生重组腺相关病毒(rAAV)的方法。在一些实施方案中,产生rAAV的方法包括使宿主细胞与本文公开的任一种核酸分子或质粒接触。因此,本公开还提供了通过本文公开的产生rAAV的方法产生的重组腺相关病毒(rAAV)。

[0109] 本公开还提供了rAAV。在一些实施方案中,rAAV包含AAV衣壳蛋白以及本文公开的任一种核酸分子或AAV表达盒。本公开还提供了组合物,其包含本文公开的任一种cTuberin蛋白、任一种核酸分子、任一种质粒、任一种宿主细胞或任一种rAAV。

[0110] 在一些实施方案中,rAAV包含1型AAV、2型AAV、3型AAV(包括3A 3B型)、4型AAV、5型AAV、6型AAV、7型AAV、8型AAV、9型AAV、10型AAV、11型AAV、12型AAV、13型AAV、rh32.33型AAV、rh8型AAV、rh10型AAV、rh74型AAV、hu.68型AAV、禽AAV、牛AAV、犬AAV、马AAV、绵羊AAV、蛇AAV、鬃狮蜥AAV、AAV2i8、AAV2g9、AAV-LK03、AAV7m8、AAV Anc80或AAV PHP.B衣壳蛋白。

[0111] 在一些实施方案中,rAAV包含AAV9衣壳蛋白、AAV8衣壳蛋白和/或AAVrh10衣壳蛋白。在一些实施方案中,rAAV包含AAV9衣壳蛋白。在一些实施方案中,rAAV包含AAV8衣壳蛋白。在一些实施方案中,rAAV包含AAVrh10衣壳蛋白。在一些实施方案中,rAAV是假型AAV,其包含一种血清型的AAV衣壳蛋白和来源于不同血清型的AAV ITR。在一些实施方案中,rAAV包含嵌合AAV衣壳或人源化AAV衣壳。在一些实施方案中,rAAV是自身互补AAV(scAAV)。在一些实施方案中,rAAV是单链AAV。

[0112] 在一些实施方案中,rAAV颗粒的制备涉及培养含有以下各者的宿主细胞:编码AAV衣壳蛋白或其片段的核酸序列;功能性rep基因;由AAV反向末端重复序列(ITR)和编码本文公开的任一种cTuberin蛋白的AAV表达盒构成的重组AAV载体;和足够的辅助功能以允许重组AAV载体包装到AAV衣壳蛋白中。在一些实施方案中,将在宿主细胞中培养以将rAAV载体包装在AAV衣壳中的组分以反式提供给宿主细胞。在一些实施方案中,任何一种或多种所需组分(例如,重组AAV载体、rep序列、cap序列和/或辅助功能)由稳定的宿主细胞提供,该宿主细胞已被工程化以含有一种或多种所需组分。

[0113] 在一些实施方案中,稳定的宿主细胞将含有受诱导型启动子或组成型启动子控制的所需组分。在一些实施方案中,选定的稳定宿主细胞含有在组成型启动子控制下的选定组分和在一种或多种诱导型启动子控制下的其他选定组分。例如,可以产生来源于293细胞(其含有在组成型启动子控制下的E1辅助功能),但含有在诱导型启动子控制下的rep和/或cap蛋白的稳定宿主细胞。产生本文公开的rAAV所需的重组AAV载体、rep序列、cap序列和辅助功能可以使用任何合适的遗传元件(例如载体)递送至包装宿主细胞。关于制备rAAV颗粒

的方法的其他细节提供于Sambrook等人,Molecular Cloning:ALaboratory Manual,Cold Spring Harbor Press,Cold Spring Harbor,N.Y.;K.Fisher等人,J.Virol.,70:520-532(1993)和美国专利第5,478,745号中,其各自的内容出于所有目的以其全文并入本文中。

[0114] 在一些实施方案中,使用三重转染方法产生重组AAV,如美国专利第6,001,650号中所述,其内容出于所有目的以其全文并入本文中。在一些实施方案中,通过用待包装到AAV颗粒中的重组AAV载体(包含编码cTuberin的AAV表达盒)、AAV辅助功能载体和辅助功能载体转染宿主细胞来产生重组AAV。AAV辅助功能载体编码"AAV辅助功能"序列(即,rep和cap),其反式作用用于生产性AAV复制和衣壳化。AAV辅助功能载体的非限制性示例包括分别描述于美国专利第6,001,650号和第6,156,303号中的pHLP19和pRep6cap6载体,其各自的内容出于所有目的以其全文并入本文中。辅助功能载体编码AAV依赖其复制的非AAV源性病毒和/或细胞功能(即,"辅助功能")的核苷酸序列。辅助功能包括AAV复制所需的那些功能,包括但不限于涉及AAV基因转录的活化、阶段特异性AAV mRNA剪接、AAV DNA复制、帽表达产物的合成以及AAV衣壳组装的那些部分。基于病毒的辅助功能可来源于任何已知的辅助病毒,如腺病毒、疱疹病毒(除了1型单纯疱疹病毒)和痘苗病毒。

[0115] 在一些实施方案中,使用杆状病毒载体产生重组AAV。杆状病毒载体用于在昆虫细胞(例如,草地夜蛾(Spodoptera frugiperda)(Sf9)细胞)中产生重组AAV。关于产生编码cTuberin的AAV的其他细节可见于美国专利公开案第2020/0079824号中,其内容出于所有目的以引用的方式整体并入本文中。

[0116] 药物组合物

[0117] 本公开还提供了药物组合物,其包含:(a)本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种宿主细胞或本文公开的任一种rAAV;和(b)药学上可接受的载剂。

[0118] 在一些实施方案中,本文公开的组合物包含至少一种药学上可接受的载剂、赋形剂和/或媒介物,例如溶剂、缓冲剂、溶液、分散介质、包衣、抗细菌剂、抗真菌剂、等渗剂和吸收延迟剂。在一些实施方案中,药学上可接受的载剂、赋形剂和/或媒介物包含盐水、缓冲盐水、右旋糖、水、甘油、无菌等渗水性缓冲液或其组合。在一些实施方案中,药学上可接受的载剂、赋形剂和/或媒介物包含磷酸盐缓冲盐水、无菌盐水、乳糖、蔗糖、磷酸钙、葡聚糖、琼脂、果胶、花生油、芝麻油、药物级甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素、碳酸镁、多元醇(例如,甘油、丙二醇和液体聚乙二醇等)或其合适的混合物。在一些实施方案中,本文公开的组合物还包含乳化剂或润湿剂,或pH缓冲剂。此类物种可以少量(例如,按组合物的重量计小于10%,如按组合物的重量计小于5%、按组合物的重量计2%、按组合物的重量计1%或更少)存在。

[0119] 在一些实施方案中,本文公开的组合物还包含一种或多种其他药物成分,如一种或多种防腐剂或化学稳定剂。防腐剂和化学稳定剂的示例包括但不限于氯丁醇、山梨酸钾、山梨酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸酯、乙基香草醛、甘油、苯酚、对氯苯酚和白蛋白。在一些实施方案中,本文公开的组合物还可包含抗细菌剂和/或抗真菌剂,如对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸和硫柳汞;等渗剂,如糖和氯化钠;和/或延迟吸收的剂,如单硬脂酸铝和明胶。在一些实施方案中,本文公开的组合物包含表面活性剂,如普朗尼克(pluronic)F68(泊洛沙姆(Poloxamer)188,也称为LUTROL® F68)。

[0120] 在一些实施方案中,配制本文公开的组合物以减少组合物中AAV颗粒的聚集,特别是在存在高rAAV浓度(例如,~10^{13}GC/ml或更高)时。用于减少rAAV聚集的方法包括添加表面活性剂、pH调节和盐浓度调节,如Wright,等人,Molecular Therapy(2005)12,171-178中进一步描述,其内容出于所有目的以引用的方式整体并入本文中。

[0121] 在一些实施方案中,药物组合物为可注射溶液或分散体的形式,如水溶液或分散体。在一些实施方案中,药物组合物是用于临时制备无菌可注射溶液或分散体的无菌粉末。分散体可在水、甘油、液体聚乙二醇、油或其任何组合中制备。递送载体(如脂质体、纳米胶囊、微粒、微球、脂质颗粒、囊泡等)可用于引入本文公开的药物组合物。

[0122] 本公开还提供了包含本文公开的任一种cTuberin蛋白或本文公开的任何核酸分子的细胞外囊泡(EV)。在一些实施方案中,药物组合物包含细胞外囊泡(EV),其包含本文公开的任一种cTuberin蛋白或本文公开的任何核酸分子。细胞外囊泡(包括但不限于外来体、微泡、微粒、循环微泡、脱落微泡、纳米囊泡、纳米颗粒、凋亡小体和膜囊泡)是从几乎所有细胞类型脱落的质膜的片段,范围例如在20纳米(nm)至10微米(μm)。在一些实施方案中,使用过滤、差速离心、超速离心、在梯度(蔗糖,OptiPrep™)中漂浮囊泡和利用针对膜蛋白的抗体的免疫亲和捕获来分离和纯化EV。其他细节在以下中提供:Simpson R J,Mathivanan S(2012)Extracellular Microvesicles:The Need for Internationally Recognised Nomenclature and Stringent Purification Criteria.J Proteomics;van der Pol等人,Classification,functions,and clinical relevance of extracellular vesicles,Pharmacol Rev.2012年7月;64(3):676-705;Raposo和Stoorvogel,Extracellular vesicles:exosomes,microvesicles,and friends,J Cell Biol.2013年2月18日;200(4):373-83;和Witwer等人,Standardization of sample collection,isolation and analysis methods in extracellular vesicle research,JExtracell Vesicles.2013年5月27日,其各自的内容出于所有目的以引用的方式整体并入本文中。

[0123] 治疗结节性硬化的方法

[0124] 本公开还提供了在靶细胞中表达本文公开的任一种cTuberin的方法,其包括:使本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物与靶细胞接触,从而在靶细胞中表达cTuberin。

[0125] 此外,本公开提供了抑制mTORC1和/或Rheb活性的方法,其包括:使本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物与靶细胞接触,从而抑制靶细胞中的mTORC1和/或Rheb活性。

[0126] 在一些实施方案中,靶细胞是脑细胞、心脏细胞、肾细胞、皮肤细胞、肺细胞或其任何组合。在一些实施方案中,接触步骤在体外、离体或体内进行。在一些实施方案中,接触步骤在有需要的受试者体内进行。在一些实施方案中,接触步骤包括向受试者施用治疗有效量的核酸分子、质粒、rAAV或组合物。

[0127] 本公开还提供了治疗有需要的受试者的结节性硬化或其症状的方法,其包括:向受试者施用治疗有效量的本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物,从而治疗受试者的结节性硬化。在一个相关方面,本公开提供了改善有需要的受试者的结节性硬化的一种或多种症状、减轻其严重性、消除其和/或延迟其发作的方法,该方法包括:向受试者施用治疗有效

量的本文公开的任一种核酸分子中、本文公开任一种的质粒中、本文公开的任一种rAAV或本文公开的任一种组合物,从而改善受试者的结节性硬化的一种或多种症状、减轻其严重性、消除其和/或延迟其发作。

[0128] 本公开提供了治疗有需要的受试者的肾血管平滑肌脂肪瘤或其症状的方法,其包括:向受试者施用治疗有效量的本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物,从而治疗受试者的肾血管平滑肌脂肪瘤。在一些实施方案中,肾血管平滑肌脂肪瘤的治疗引起肾血管平滑肌脂肪瘤的消退、收缩、消除或延迟生长。

[0129] 本公开另外提供了治疗有需要的受试者的淋巴管平滑肌瘤病(LAM)或其症状的方法,其包括:向受试者施用治疗有效量的本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物,从而治疗受试者的LAM。在一个相关方面,本公开提供了改善有需要的受试者的淋巴管平滑肌瘤病(LAM)的一种或多种症状、减轻其严重性、消除其和/或延迟其发作的方法,该方法包括:向受试者施用治疗有效量的本文公开的任一种核酸分子中、本文公开任一种的质粒中、本文公开的任一种rAAV或本文公开的任一种组合物,从而改善受试者的LAM的一种或多种症状、减轻其严重性、消除其和/或延迟其发作。

[0130] 本公开提供了治疗有需要的受试者的脑功能障碍的方法,其包括:向受试者施用治疗有效量的本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV、本文公开的任一种EV或本文公开的任一种组合物,从而治疗受试者的脑功能障碍。在一个相关方面,本公开提供了改善有需要的受试者的脑功能障碍的一种或多种症状、减轻其严重性、消除其和/或延迟其发作的方法,该方法包括:向受试者施用治疗有效量的本文公开的任一种核酸分子中、本文公开任一种的质粒中、本文公开的任一种rAAV或本文公开的任一种组合物,从而改善受试者的脑功能障碍的一种或多种症状、减轻其严重性、消除其和/或延迟其发作。

[0131] 在一些实施方案中,结节性硬化与马铃薯球蛋白活性的降低有关、相关或是由其引起。在一些实施方案中,受试者患有结节性硬化。在一些实施方案中,受试者处于发展结节性硬化的至少一种症状的风险中。

[0132] 在一些实施方案中,受试者已被诊断患有结节性硬化。在一些实施方案中,受试者基于临床标准被诊断为患有结节性硬化,该临床标准如癫痫发作的存在;延迟发育;皮肤上的白斑(低黑色素斑);鉴定心脏肿瘤横纹肌瘤;鉴定脑、心脏、肝脏或肾脏中的肿瘤;检查皮肤的各种皮肤特征;手指甲、脚指甲的指甲纤维瘤;牙齿和牙龈的牙窝和/或牙龈纤维瘤;眼睛的视网膜病变;面部血管纤维瘤;和/或低黑色素斑的存在。在一些实施方案中,受试者是人受试者。在一些实施方案中,受试者小于18岁,如12至18岁、8至12岁、6至12岁、2至18岁、0至2岁,或其中的任何范围。在一些实施方案中,受试者是新生儿或婴儿。在一些实施方案中,基于临床标准诊断婴儿患有结节性硬化,该临床标准如在出生时存在心脏横纹肌瘤或在生命的前六个月存在婴儿痉挛。

[0133] 在一些实施方案中,结节性硬化与TSC2基因中的突变有关、相关或是由其引起。在一些实施方案中,受试者在TSC2基因中具有突变。在一些实施方案中,受试者在体内的至少一个细胞中的TSC2的一个或两个等位基因中具有突变。在一些实施方案中,突变是遗传性

种系突变。在一些实施方案中,突变是体细胞突变。在一些实施方案中,TSC2的一个等位基因中的突变是遗传性突变,并且TSC2的第二个等位基因中的突变是体细胞突变。在一些实施方案中,处于发展结节性硬化的至少一种症状的风险中的受试者是在TSC2的一个等位基因中具有遗传性种系突变的受试者。在一些实施方案中,处于发展结节性硬化的至少一种症状的风险中的受试者是其一个或两个父母是一个或多个突变的马铃薯球蛋白等位基因的携带者的受试者。

[0134] 在一些实施方案中,受试者在体内的至少一个细胞中的TSC2的两个等位基因中具有突变。在一些实施方案中,受试者在脑、心脏、肾、皮肤、肺和/或其他器官中的至少一个细胞中的TSC2的两个等位基因中具有突变。在一些实施方案中,受试者的TSC2中的突变是纯合的或复合杂合的。

[0135] 在一些实施方案中,受试者在TSC1基因中具有突变。在一些实施方案中,受试者具有TSC1基因中的突变和TSC2基因中的突变。

[0136] 在一些实施方案中,通过测试来源于受试者的生物样品中本文所述的TSC2中的任何一种或多种突变的存在,将受试者诊断为患有结节性硬化或处于发展结节性硬化的风险中。在一些实施方案中,TSC2基因中的突变可以是任何氨基酸修饰,如氨基酸插入、缺失、剪接位点突变和/或氨基酸取代。关于TSC2基因中的突变的细节进一步描述于Reyna-Fabián,M.E.,Sci Rep 10,6589(2020),Gilbert JR,等人,Neurogenetics.1998年8月;1(4):267-72;Avgeris,S.,Sci Rep 7,16697(2017);和Rosset C,等人,Genet Mol Biol.2017;40(1):69-79中,其各自的内容出于所有目的以其全文并入本文中。在一些实施方案中,受试者在TSC2基因的外显子33、外显子37和/或外显子38中具有突变。在一些实施方案中,受试者在TSC2基因的外显子33中具有突变。

[0137] 在一些实施方案中,施用治疗有效量的本文提供的治疗剂(例如,治疗有效量的本文公开的任一种核酸分子、本文公开的任一种质粒、本文公开的任一种rAAV或本文公开的任一种组合物)减轻本文公开的结节性硬化的任一种症状的严重性。在一些实施方案中,施用治疗有效量的本文提供的治疗剂延迟本文公开的结节性硬化的任一种症状的发作。在一些实施方案中,施用治疗有效量的本文提供的治疗剂消除本文公开的结节性硬化的任一种症状的症状。在一些实施方案中,施用治疗有效量的本文提供的治疗剂改善本文公开的结节性硬化的任一种症状的症状。在一些实施方案中,结节性硬化的症状是存在以下中的任何一种或多种:脑、心脏、肝脏、肾脏,眼睛或皮肤中的肿瘤或错构瘤;室管膜下生长或结节;室管膜下巨细胞星形细胞瘤;皮质结节;脑功能障碍;癫痫发作;延迟发育;低黑色素性斑;心脏肿瘤横纹肌瘤;舌纤维瘤;牙窝和/或牙龈纤维瘤;视网膜病变;面部血管纤维瘤;肾血管平滑肌脂肪瘤;淋巴管平滑肌瘤病(LAM);内出血;孤独症;癫痫;脑积水;和/或存在低黑色素性斑。在一些实施方案中,结节性硬化的症状是Randle SC.,Pediatr Ann.2017年4月1日;46(4):e166-e171,Uysal SP,Turk J Med Sci.2020年11月3日;50(SI-2):1665-1676,和Henske EP等人Nat Rev Dis Primers.2016May 26;2:16035中所描述的任一者,其各自的内容出于所有目的以引用的方式整体并入本文中。

[0138] 在一些实施方案中,施用与受试者的肿瘤或错构瘤的大小减小有关、相关,或引起此。在一些实施方案中,施用与同未施用本文公开的组合物的患有结节性硬化的对照受试者相比的受试者的肿瘤或错构瘤的大小减小至少约5%(例如,至少约10%、约15%、约

20％、约25％、约30％、约35％、约40％、约45％、约50％、约55％、约60％、约65％、约70％、约75％、约80％、约85％、约90％、约95％或更大,如100％)有关、相关,或引起此。

[0139] 在一些实施方案中,施用与受试者的肿瘤或错构瘤的数量减少有关、相关,或引起此。在一些实施方案中,施用与同未施用本文公开的组合物的患有结节性硬化的对照受试者相比的受试者的肿瘤或错构瘤的数量减少至少约5％(例如,至少约10％、约15％、约20％、约25％、约30％、约35％、约40％、约45％、约50％、约55％、约60％、约65％、约70％、约75％、约80％、约85％、约90％、约95％或更大,如100％)有关、相关,或引起此。

[0140] 不受理论的束缚,认为虽然向患有结节性硬化的受试者施用雷帕霉素与由于mTORC1的过度抑制引起的副作用相关,但使用本文公开的组合物和方法的cTuberin的表达可能不会引起mTORC1的过度抑制。这是因为患有与TSC2中的突变相关的结节性硬化的受试者可能具有正常水平的错构瘤蛋白,此将确保mTORC1的正常抑制。此外,虽然雷帕霉素和其类似物可仅抑制mTORC1,但使用所公开的组合物的cTuberin的表达除了抑制mTORC1之外还导致mTORC1非依赖性Rheb依赖性病理作用。因此,本文公开的组合物和方法可以比用于结节性硬化的基于雷帕霉素的疗法更有效。

[0141] 在一些实施方案中,该方法包括施用治疗有效量的rAAV,其中该治疗有效量在约10^5个基因组拷贝至10^{20}个基因组拷贝/千克(kg)的范围内。例如约10^6个基因组拷贝/kg、约10^7个基因组拷贝/kg、约10^8个基因组拷贝/kg、约10^9个基因组拷贝/kg、约10^{10}个基因组拷贝/kg、约10^{11}个基因组拷贝/kg、约10^{12}个基因组拷贝/kg、约10^{13}个基因组拷贝/kg、约10^{14}个基因组拷贝/kg、约10^{15}个基因组拷贝/kg、约10^{16}个基因组拷贝/kg、约10^{17}个基因组拷贝/kg、约10^{18}个基因组拷贝/kg、或约10^{19}个基因组拷贝/kg,包括位于其间的所有值和子范围。在一些实施方案中,该方法包括施用治疗有效量的rAAV,其中该治疗有效量在10^{10}个基因组拷贝至10^{14}个基因组拷贝/千克的范围内。在一些实施方案中,该方法包括施用治疗有效量的rAAV,其中该治疗有效量在10^9个基因组拷贝至10^{15}个基因组拷贝/千克的范围内。

[0142] 在一些实施方案中,治疗有效量在10^5至10^{20}个基因组拷贝/受试者的范围内,例如约10^6个基因组拷贝/受试者、约10^7个基因组拷贝/受试者、约10^8个基因组拷贝/受试者、约10^9个基因组拷贝/受试者、约10^{10}个基因组拷贝/受试者、约10^{11}个基因组拷贝/受试者、约10^{12}个基因组拷贝/受试者、约10^{13}个基因组拷贝/受试者、约10^{14}个基因组拷贝/受试者、约10^{15}个基因组拷贝/受试者、约10^{16}个基因组拷贝/受试者、约10^{17}个基因组拷贝/受试者、约10^{18}个基因组拷贝/受试者、或约10^{19}个基因组拷贝/受试者,包括位于其间的所有值和子范围。在一些实施方案中,治疗有效量在约10^9至10^{16}个基因组拷贝/受试者的范围内。

[0143] 在一些实施方案中,治疗有效量以约1微升(μl)至约100mL溶液的体积施用,例如约10μl、约50μl、约100μl、约125μl、约150μl、约175μl、约200μl、约250μl、约300μl、约350μl、约400μl、约450μl、约500μl、约550μl、约600μl、约650μl、约700μl、约750μl、约800μl、约850μl、约900μl、约950μl、约1毫升(mL)、约20mL、约30mL、约40mL、约50mL、约60mL、约70mL、约80mL、约90mL或约100mL,包括位于其间的所有值和子范围。所用体积可取决于rAAV的剂量和施用途径。例如,对于鞘内或脑内施用,可以使用在约1μl至约10μl、或约10μl至约100μl范围内的体积。对于静脉内施用,可以使用在约10μl至约100μl、或约100μl至1mL、或约1mL至约10mL、或更多的范围内的体积。

[0144] 在一些实施方案中,可以采用超过一次施用(例如,两次、三次、四次或更多次施

用)以在不同的时间间隔(例如,每天、每周、每月、每年等)的时间段内实现期望的基因表达水平。

[0145] 在一些实施方案中,施用是通过注射到中枢神经系统中。可以使用的其他施用模式包括皮肤、口服、直肠、经粘膜、鼻内、吸入(例如,经由气雾剂)、经颊(例如,舌下)、阴道、鞘内、眼内、经皮、子宫内(或卵内)、肠胃外(例如,静脉内、皮下、皮内、肌内[包括施用至骨骼、隔膜和/或心肌]、皮内、胸膜内、脑内和关节内)、局部(例如,至皮肤和粘膜表面,包括气道表面,和经皮施用)、淋巴管内等,以及直接组织或器官注射(例如至肝脏、骨骼肌、心肌、隔膜肌肉或脑)。在一些实施方案中,该施用是通过脑室内或颅内注射。在一些实施方案中,cTuberin、核酸分子、质粒、组合物或rAAV是通过血管内、肾动脉或静脉内、肺内、小脑延髓池内、脑内、鞘内、静脉内、心室内、脑室内、腹膜内或经皮施用。

[0146] 在一些实施方案中,该方法包括通过血管内注射(例如,注射到肾动脉或静脉中)向患有肾血管平滑肌脂肪瘤的受试者施用治疗有效量的cTuberin、核酸分子、质粒、组合物或rAAV。在一些实施方案中,治疗有效量的cTuberin、核酸分子、质粒、组合物或rAAV靶向肾血管平滑肌脂肪瘤。在一些实施方案中,该方法包括通过血管内注射向患有淋巴管平滑肌瘤病(LAM)的受试者施用治疗有效量的cTuberin、核酸分子、质粒、组合物或rAAV。在一些实施方案中,治疗有效量的cTuberin、核酸分子、质粒、组合物或rAAV靶向LAM。在一些实施方案中,受试者患有脑功能障碍。在一些实施方案中,cTuberin、核酸分子、质粒、组合物或rAAV被提供给蛛网膜下腔。

[0147] 在一些实施方案中,cTuberin、核酸分子、质粒、组合物或rAAV被提供(例如施用)给脑细胞、心脏细胞、肾细胞、皮肤细胞或肺细胞。在一些实施方案中,将核酸分子、质粒、宿主细胞、rAAV或组合物施用到受试者的血流中。施用到血流中可以通过注射到静脉、动脉或任何其他血管中。在一些实施方案中,核酸分子、质粒、宿主细胞、rAAV或组合物是通过血管内施用。在一些实施方案中,核酸分子、质粒、宿主细胞、rAAV或组合物是通过静脉内施用。

[0148] 在一些实施方案中,将核酸分子、质粒、宿主细胞、rAAV或组合物递送到脑组织、脑膜、神经元细胞、神经胶质细胞、星形胶质细胞、少突胶质细胞、脑脊液(CSF)、间质间隙等。在一些实施方案中,使用神经外科技术,如通过立体定向注射,用针、导管或相关装置,通过注射到心室区域以及纹状体(例如,纹状体的尾状核或壳核)和神经肌肉接头或小脑小叶,可以将重组AAV直接递送到脊髓或脑。

[0149] 在一些实施方案中,施用可包括通过超过途径同时地或在不同时间点施用核酸分子、质粒、宿主细胞、rAAV或组合物。

[0150] 在一些实施方案中,本文公开的方法包括向受试者施用另一种辅助疗法。在一些实施方案中,辅助疗法包括施用抗癫痫发作药物。抗癫痫药物的非限制性示例包括卡马西平(Carbamazepine)、苯妥英(Phenytoin)、丙戊酸、奥卡西平(Oxcarbazepine)、拉莫三嗪(Lamotrigine)、加巴喷丁(Gabapentin)、托吡酯(Topiramate)和苯巴比妥(Phenobarbital)以及唑尼沙胺(Zonisamide)。在一些实施方案中,辅助疗法包括施用雷帕霉素或其类似物。可以向受试者依序或同时施用辅助疗法。

[0151] 本公开还提供了包含一种或多种剂(例如,本文公开的任一种核酸分子中、本文公开任一种的质粒中、本文公开的任一种rAAV或本文公开的任一种组合物)的试剂盒。在一些实施方案中,试剂盒是用于治疗、诊断或研究应用的药物或诊断或研究试剂盒。试剂盒可以

包括一个或多个容纳本文公开的剂和使用说明书的容器。在某些实施方案中,试剂盒中的剂是适合于特定应用和施用剂的方法的药物制剂和剂量。在一些实施方案中,容器是注射器、小瓶、管、局部施加装置、IV针管和袋、或另一容器。

[0152] 在一些实施方案中,试剂盒含有:第一药物组合物,其包含编码cTuberin的核酸分子(例如,在本文所述的rAAV中)或cTuberin;和第二药物组合物,其包含一种或多种用于治疗结节性硬化的药物,如雷帕霉素和其类似物。在一些实施方案中,试剂盒包括用于依序或同时施用两种组合物,或在施用前混合两种药物组合物的说明书。

[0153] ********

[0154] 本文所使用的小节标题仅出于组织性目的并且不解释为限制所描述的主题。本文中引用的所有文件或文件的部分,包括但不限于专利、专利申请、文章、书籍和论文,均出于任何目的特此明确地以引用的方式整体并入。在一个或多个并入的文献或文献的部分定义与本申请中术语的定义相矛盾的术语的情况下,以本申请中出现的定义为准。然而,本文所引用的任何参考文献、文章、出版物、专利、专利出版物和专利申请,并不也不应被视为承认或任何形式的建议,它们构成有效的现有技术或形成世界上任何国家的公知常识的一部分。

[0155] 除非上下文另有说明,否则特别意图本文描述的各种特征可以任何组合使用。

[0156] 实施例

[0157] 实施例1:新的凝聚的马铃薯球蛋白(cTuberin)构建体的产生

[0158] 通过将不同的核酸构建体克隆到合适的表达质粒中产生包含特定氨基酸残基的缺失的各种cTuberin蛋白。构建体和氨基酸缺失列于下表1中。

[0159]

表1:

质粒	上游反应元件(URE)	启动子	N端区域		C端区域		从变体中包括的SEQ ID NO: 1的C端起的氨基酸的编号	相对于SEQ ID NO: 1缺失的氨基酸	cTuberin的完全序列	编码cTuberin的核酸的完全序列
			相对于SEQ ID NO: 1的氨基酸残基	序列	相对于SEQ ID NO: 1的氨基酸残基	序列				
V2	CMV-IE	CB6	1-450	SEQ ID NO: 4	1140-1807	SEQ ID NO: 7	668	451-1139	SEQ ID NO: 14	SEQ ID NO: 21
V3	CMV-IE	CB6	1-418	SEQ ID NO: 5	1110-1807	SEQ ID NO: 8	698	419-1109	SEQ ID NO: 15	SEQ ID NO: 22
V4	CMV-IE	CB6	1-418	SEQ ID NO: 5	933-1807	SEQ ID NO: 9	875	419-932	SEQ ID NO: 16	SEQ ID NO: 23
V5		CB6	1-418	SEQ ID NO: 5	933-1807	SEQ ID NO: 9	875	419-932	SEQ ID NO: 16	SEQ ID NO: 23
V6	CMV-IE	CB6	1-418	SEQ ID NO: 5	933-946 和 989-180 /	SEQ ID NO: 10	875 Δ25	419-932, 947-988	SEQ ID NO: 17	SEQ ID NO: 24
V7	CMV-IE	CB6	1-418	SEQ ID NO: 5	933-946, 989-1204, 1272-1807	SEQ ID NO: 11	875 Δ25 Δ30	419-932, 947-988, 1205-1271	SEQ ID NO: 18	SEQ ID NO: 25
V8	CMV-IE	CB6	1-418	SEQ ID NO: 5	933-946, 989-1204, 1272-1335, 1498-1807	SEQ ID NO: 12	875 Δ25 Δ30 Δ33	419-932, 947-988, 1205-1271, 1336-1497	SEQ ID NO: 19	SEQ ID NO: 26
V9		GUSB	1-450	SEQ ID NO: 4	1516-1807	SEQ ID NO: 6	292	451-1515	SEQ ID NO: 13	SEQ ID NO: 20

[0160]

V10		GUSB	1-450	SEQ ID NO: 4	1140-1807	SEQ ID NO: 7	668	451-1139	SEQ ID NO: 14	SEQ ID NO: 21
V11		GUSB	1-418	SEQ ID NO: 5	1110-1807	SEQ ID NO: 8	698	419-1109	SEQ ID NO: 15	SEQ ID NO: 22
V12		GUSB	1-418	SEQ ID NO: 5	933-1807	SEQ ID NO: 9	875	419-932	SEQ ID NO: 16	SEQ ID NO: 23
V13		GUSB	1-418	SEQ ID NO: 5	933-1807	SEQ ID NO: 9	875	419-932	SEQ ID NO: 16	SEQ ID NO: 23
V14		GUSB	1-418	SEQ ID NO: 5	933-946 和 989-1807	SEQ ID NO: 10	875 Δ25	419-932, 947-988	SEQ ID NO: 17	SEQ ID NO: 24
V15		GUSB	1-418	SEQ ID NO: 5	933-946, 989-1204, 1272-1807	SEQ ID NO: 11	875 Δ25 Δ30	419-932, 947-988, 1205-1271	SEQ ID NO: 18	SEQ ID NO: 25
V16		GUSB	1-418	SEQ ID NO: 5	933-946, 989-1204, 1272-1335, 1498-1807	SEQ ID NO: 12	875 Δ25 Δ30 Δ33	419-932, 947-988, 1205-1271, 1336-1497	SEQ ID NO: 19	SEQ ID NO: 26

[0161] 产生包含AAV表达盒的AAV载体,这些表达盒包含编码表1中列出的各种cTuberin的核酸序列。AAV表达盒包含CB6启动子,其具有或不具有CMV-IE上游反应元件(URE)或可操作地连接于编码表1所指示的各变体的核酸序列的GUSB启动子。使用三重转染方法将AAV载

体包装到AAV颗粒中。在一些情况下,AAV是1型AAV(例如,AAV1)或9型AAV(例如,AAV9)。在一些情况下,载体包括一个或多个ITR元件,如一个或多个AAV2 ITR元件。

[0162] 如上所述的AAV载体可以通过任何有用的方法制备。在一些实施方案中,使用转染(例如三重转染)的细胞培养物(例如,HEK293悬浮培养物)制备AAV载体以产生感兴趣的产物。后续处理可包括一个或多个溶解、离子交换色谱、过滤(包括超滤)、亲和色谱和/或稀释步骤。可以使用各种测序方法确认载体的序列。可以使用PCR扩增方法或其他方法测定滴度(例如,gc/mL)。

[0163] 实施例2:cTuberin在细胞培养物中的表达

[0164] 用包含对应于实施例1中描述的cTuberin的核酸序列的载体质粒转染细胞(例如,COS-7细胞)。24小时后,用抗马铃薯球蛋白/TSC2抗体通过免疫印迹(例如,蛋白质印迹)检测cTuberin的表达。

[0165] 为了测试cTuberin的活性,用各种AAV构建体转染细胞(例如COS-7细胞),这些AAV构建体包括例如GFP、pAAV-CBA-cTSC2、TSC2-FLAG、pAAV-CBA-cTSC2+TSC1-FLAG、TSC1-FLAG+TSC2-FLAG和TSC1-FLAG载体。通过免疫印迹(例如,蛋白质印迹)检测磷酸化的S6(pS6)、S6和GAPDH的表达水平。虽然pS6激酶水平在不存在马铃薯球蛋白活性的情况下通常升高,但是用本文提供的质粒转染的细胞可显示较低的pS6水平,这将指示pS6激酶活性降低。

[0166] 实施例3:cTuberin在体内的表达

[0167] 使用缺乏马铃薯球蛋白的小鼠,如Onda等人描述的$Tcs2^{c/c}$ floxed小鼠(Onda等人,J.Clin.Invest.104(6):687-695,1999)进一步评估本公开的载体的功效。响应于Cre重组酶,$Tsc2^{c/c}$等位基因转化为无效等位基因,并且lacZ等位基因表达β-半乳糖苷酶。已经观察到这种小鼠具有健康且正常的寿命。

[0168] 使用实施例1中描述的选定的载体进行脑室内(ICV)和/或眶后(RO)注射。ICV注射在小鼠的生命周期早期进行,如出生后第(postnatal day;PND)0、1、2、3、4或5天。RO注射在小鼠的稍后生命周期中进行,如在PND 7-35之间,如在PND 21。注射滴度可以在$1x10^{11}$个基因组拷贝/毫升(gc/mL)至$1x10^{14}$gc/mL的范围内。监测存活期。使用苏木精和曙红(H&E)染色和/或对pS6的免疫组织化学分析来研究缺乏马铃薯球蛋白的小鼠和已处理的小鼠的脑。

[0169] 在一个研究中,评估cTuberin载体(如实施例1中所述的那些)对存活期的影响。在PND0或PND3对$Tcs2^{c/c}$小鼠ICV注射携带Cre的载体(如AAV1-CBA-Cre);在P21 RO注射实施例1中所述的cTuberin载体;或未注射。在PND3,脑脊液(CSF)屏障可能比在P0的穿透性略低,使得可观察到脑中较少的马铃薯球蛋白损失。监测存活期。注射cTuberin载体的小鼠的存活期预期比注射Cre载体的小鼠存活得更长。注射cTuberin载体的小鼠的存活期可与未注射的小鼠相当(例如,>175天)。

[0170] 在另一研究中,评估cTuberin载体(如实施例1中所述的那些)的功效。在PND0或PND3对所有小鼠ICV注射携带Cre的载体(如AAV1-CBA-Cre)。接着,在PND21对第一组小鼠RO注射实施例1中所述的cTuberin载体,并且对第二组小鼠不再进行注射。预期注射cTuberin载体的小鼠的存活期比未注射cTuberin载体的小鼠更长。在处死后(例如,在PND27),使用H&E染色或对pS6的IHC来对缺乏马铃薯球蛋白的小鼠和cTuberin处理的小鼠的脑进行研究。对于缺乏马铃薯球蛋白的小鼠,可以观察到室管膜细胞增殖、海马中锥体细胞的增大、室管膜下结节和多个室管膜下结节与增殖。对于cTuberin处理的小鼠,脑组织可能看起来

更类似于正常的未注射(例如,对照)的脑。

[0171] 在另一研究中,评估cTuberin载体(如实施例1中所述的那些)与抗惊厥剂组合的功效。TSC2突变与较早发作和较高频率的癫痫发作相关。因此,cTuberin载体与抗惊厥剂的组合可有效治疗具有TSC2的受试者。在P0或P3对所有小鼠ICV注射携带Cre的载体(如AAV1-CBA-Cre)。接着,在PND3或稍后对第一组小鼠RO注射实施例1中所述的cTuberin载体,并且对第二组小鼠不再进行注射。一组注射cTuberin载体的小鼠用抗惊厥剂(如氨己烯酸)(例如,200mg/kg)处理,并且另一组小鼠不用抗惊厥剂处理。注射cTuberin载体并用抗惊厥剂处理的小鼠的存活期预期比未注射cTuberin载体且未用药物处理的小鼠、未注射cTuberin载体且用药物处理的小鼠、和注射cTuberin载体且未用抗惊厥剂处理的小鼠更长。

[0172] 在其他研究中,体内测试cTuberin载体(如实施例1中描述的那些)对皮下注射于NOD-SCID II2R γ (NSG)小鼠中的淋巴管平滑肌瘤病(LAM)肿瘤的功效。将表达Fluc的TSC2无效、永生化的血管平滑肌脂肪瘤细胞悬浮于血清培养基中,与基质胶混合,并且皮下植入NSG小鼠的背部。数周(例如,4周)后,对小鼠腹膜内注射Fluc底物D-荧光素(LUCNA-1G)并用光谱和麻醉系统检测信号。通过生物发光在例如第1、4、6、9和14周监测肿瘤体积。在例如第4和9周对肿瘤注射或不注射cTuberin载体。到第14周,注射cTuberin载体的肿瘤预期停止大小增加,而未注射的肿瘤可能会继续体积扩大。

[0173] 另外的实验细节和相关研究描述于例如2018年5月17日提交的国际专利公开案第PCT/US2018/033247号中,该申请以引用的方式整体并入本文中。

[0174] 编号的实施方案

[0175] 实施方案1.一种凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基419至932。

[0176] 实施方案2.如实施方案1所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基947-988。

[0177] 实施方案3.如实施方案1或2所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基1205-1271。

[0178] 实施方案4.如实施方案1至3中任一项所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基1336-1497。

[0179] 实施方案5.如实施方案1所述的cTuberin,其中所述C端区域包含与SEQ ID NO:10-12中的一个具有至少90%同一性的氨基酸序列。

[0180] 实施方案6.如实施方案1所述的cTuberin,其中所述cTuberin还缺乏SEQ ID NO:1的氨基酸残基933至1109。

[0181] 实施方案7.如实施方案6所述的cTuberin,其中C端结构域包含与SEQ ID NO:8具有至少90%同一性的氨基酸序列。

[0182] 实施方案8.如实施方案1至7中任一项所述的cTuberin,其中N端结构域包含与SEQ ID NO:5具有至少90%同一性的氨基酸序列。

[0183] 实施方案9.一种凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述C端区域包含与SEQ ID NO:7具有至少90%同一性的氨基酸序列,并且其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残

基451至932。

[0184] 实施方案10.如实施方案9所述的cTuberin,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基419至932。

[0185] 实施方案11.如实施方案9或10所述的cTuberin,其中所述C端区域包含与SEQ ID NO:8具有至少90％同一性的氨基酸序列。

[0186] 实施方案12.如实施方案9或10所述的cTuberin,其中所述C端区域包含与SEQ ID NO:9具有至少90％同一性的氨基酸序列。

[0187] 实施方案13.如实施方案9所述的cTuberin,其中所述N端区域包含与SEQ ID NO:4具有至少90％同一性的氨基酸序列。

[0188] 实施方案14.如实施方案9至12中任一项所述的cTuberin,其中所述N端区域包含与SEQ ID NO:5具有至少90％同一性的氨基酸序列。

[0189] 实施方案15.如实施方案9所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:14具有至少90％同一性的氨基酸序列。

[0190] 实施方案16.如实施方案9所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:15具有至少90％同一性的氨基酸序列。

[0191] 实施方案17.如实施方案9所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:16具有至少90％同一性的氨基酸序列。

[0192] 实施方案18.一种凝聚的马铃薯球蛋白(cTuberin),其包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述C端区域包含与SEQ ID NO:10-12中的一个具有至少90％同一性的氨基酸序列,并且其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基451至932。

[0193] 实施方案19.如实施方案18所述的cTuberin,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基419-932。

[0194] 实施方案20.如实施方案18或19所述的cTuberin,其中所述C端区域包含与SEQ ID NO:10具有至少90％同一性的氨基酸序列。

[0195] 实施方案21.如实施方案18或19所述的cTuberin,其中所述C端区域包含与SEQ ID NO:11具有至少90％同一性的氨基酸序列。

[0196] 实施方案22.如实施方案18或19所述的cTuberin,其中所述C端区域包含与SEQ ID NO:12具有至少90％同一性的氨基酸序列。

[0197] 实施方案23.如实施方案18所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:17具有至少90％同一性的氨基酸序列。

[0198] 实施方案24.如实施方案18所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:18具有至少90％同一性的氨基酸序列。

[0199] 实施方案25.如实施方案18所述的cTuberin,其中所述cTuberin包含与SEQ ID NO:19具有至少90％同一性的氨基酸序列。

[0200] 实施方案26.如实施方案18至25中任一项所述的cTuberin,其中所述N端区域包含与SEQ ID NO:5具有至少90％同一性的氨基酸序列。

[0201] 实施方案27.如实施方案1至26中任一项所述的cTuberin,其中所述cTuberin在所述N端区域与所述C端区域之间包含间隔序列。

[0202] 实施方案28.如实施方案27所述的cTuberin,其中所述间隔序列包含SEQ ID NO:2的序列。

[0203] 实施方案29.如实施方案28所述的cTuberin,其中所述间隔序列包含SEQ ID NO:3的序列。

[0204] 实施方案30.一种核酸分子,其编码如实施方案1至29中任一项所述的cTuberin。

[0205] 实施方案31.如实施方案30所述的核酸分子,其中所述核酸分子经密码子优化以在人靶细胞中表达。

[0206] 实施方案32.如实施方案31所述的核酸分子,其中所述人靶细胞是脑细胞、心脏细胞、肾细胞、皮肤细胞或肺细胞。

[0207] 实施方案33.如实施方案30至32中任一项所述的核酸分子,其中所述核酸分子与调节控制序列可操作地连接。

[0208] 实施方案34.如实施方案33所述的核酸分子,其中所述调节控制序列包含人巨细胞病毒(CMV)启动子、鸡β-肌动蛋白(CBA)启动子、劳斯肉瘤病毒(RSV)LTR启动子/增强子、SV40启动子、二氢叶酸还原酶启动子、磷酸甘油激酶启动子、CMV立即/早期基因增强子/CBA启动子、突触蛋白启动子或神经胶质纤维酸性蛋白(GFAP)启动子。

[0209] 实施方案35.如实施方案33所述的核酸分子,其中所述调节控制序列包含人巨细胞病毒(CMV)立即/早期基因增强子/鸡β-肌动蛋白(CBA)启动子和土拨鼠肝炎病毒转录后调控元件(WPRE)。

[0210] 实施方案36.如实施方案33所述的核酸分子,其中所述调节控制序列包含β-葡萄糖醛酸酶(GUSB)启动子。

[0211] 实施方案37.如实施方案30至36中任一项所述的核酸分子,其中所述核酸分子与SEQ ID NOs.21-26中的任一个具有至少90%序列同一性。

[0212] 实施方案38.一种编码cTuberin的核酸分子,所述cTuberin包含(i)能够结合错构瘤蛋白的N端区域,和(ii)C端GTP酶活化蛋白(GAP)区域,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基451至932;并且其中所述核酸分子与包含β-葡萄糖醛酸酶(GUSB)启动子的调节控制序列可操作地连接。

[0213] 实施方案39.如实施方案38所述的核酸分子,其中所述cTuberin缺乏SEQ ID NO:1的氨基酸残基451至1515。

[0214] 实施方案40.如实施方案38或39所述的核酸分子,其中所述C端区域包含与SEQ ID NO:6具有至少90%同一性的氨基酸序列。

[0215] 实施方案41.如实施方案38至40中任一项所述的核酸分子,其中所述N端区域包含与SEQ ID NO:4具有至少90%同一性的氨基酸序列。

[0216] 实施方案42.一种核酸分子,其包含腺相关病毒(AAV)表达盒,所述AAV表达盒从5’至3’包含:

[0217] i)5’AAV反向末端重复序列(ITR);

[0218] ii)如实施方案25至35中任一项所述的核酸分子;以及

[0219] iii)3’AAV ITR。

[0220] 实施方案43.如实施方案42所述的核酸分子,其中所述5’ITR和/或所述3’ITR来源于AAV2。

[0221] 实施方案44.如实施方案42或43所述的核酸分子,其中所述5'AAV ITR序列包含与SEQ ID NO:27具有至少90%同一性的核酸序列。

[0222] 实施方案45.如实施方案42至44中任一项所述的核酸分子,其中所述3'AAV ITR序列包含与SEQ ID NO:28具有至少90%同一性的核酸序列。

[0223] 实施方案46.如实施方案42至45中任一项所述的核酸分子,其中所述AAV表达盒还包含聚腺苷酸化序列。

[0224] 实施方案47.如实施方案42至46中任一项所述的核酸分子,其中所述AAV表达盒还包含Kozak序列。

[0225] 实施方案48.一种质粒,其包含如实施方案30至47中任一项所述的核酸分子。

[0226] 实施方案49.一种宿主细胞,其包含如实施方案30至47中任一项所述的核酸分子或如实施方案48所述的质粒。

[0227] 实施方案50.一种组合物,其包含如实施方案30至47中任一项所述的核酸分子、如实施方案48所述的质粒或如实施方案49所述的宿主细胞。

[0228] 实施方案51.一种产生重组腺相关病毒(rAAV)的方法,所述方法包括:使宿主细胞与如实施方案30至47中任一项所述的核酸分子或如实施方案48所述的质粒接触。

[0229] 实施方案52.一种重组腺相关病毒(rAAV),其通过如实施方案51所述的方法产生。

[0230] 实施方案53.一种重组腺相关病毒(rAAV),其包含:AAV衣壳蛋白;和如实施方案30至47中任一项所述的核酸分子。

[0231] 实施方案54.如实施方案52或53所述的rAAV,其中所述rAAV包含AAV1衣壳蛋白、AAV2衣壳蛋白、AAV3衣壳蛋白、AAV4衣壳蛋白、AAV5衣壳蛋白、AAV6衣壳蛋白、AAV7衣壳蛋白、AAV8衣壳蛋白、AAV9衣壳蛋白、AAV10衣壳蛋白、AAVrh10衣壳蛋白、AAV11衣壳蛋白和/或AAV12衣壳蛋白。

[0232] 实施方案55.一种在靶细胞中表达cTuberin的方法,其包括:使所述靶细胞与如实施方案30至47中任一项所述的核酸分子、如实施方案48所述的质粒、如实施方案50所述的组合物或如实施方案52至54中任一项所述的rAAV接触,从而在所述靶细胞中表达cTuberin。

[0233] 实施方案56.如实施方案55所述的方法,其中所述接触步骤在体外、离体或体内进行。

[0234] 实施方案57.如实施方案56所述的方法,其中所述接触步骤在有需要的受试者体内进行。

[0235] 实施方案58.如实施方案57所述的方法,其中所述接触步骤包括向所述受试者施用治疗有效量的所述核酸分子、所述质粒、所述组合物或所述rAAV。

[0236] 实施方案59.一种治疗患有结节性硬化综合征(TSC)的受试者的方法,其包括:向所述受试者施用治疗有效量的如实施方案1至29中任一项所述的cTuberin、如实施方案30至47中任一项所述的核酸分子、一种或多种包含如实施方案30至47中任一项所述的核酸分子的细胞外囊泡(EV)、如实施方案48所述的质粒、如实施方案50所述的组合物或如实施方案52至54中任一项所述的rAAV,从而治疗所述受试者中的TSC。

[0237] 实施方案60.一种治疗患有肾癌的受试者的方法,其包括:向所述受试者施用治疗有效量的如实施方案1至29中任一项所述的cTuberin、如实施方案30至47中任一项所述的

核酸分子、一种或多种包含如实施方案30至47中任一项所述的核酸分子的细胞外囊泡(EV)、如实施方案48所述的质粒、如实施方案50所述的组合物或如实施方案52至54中任一项所述的rAAV,从而治疗所述受试者中的肾癌。

[0238] 实施方案61.如实施方案57至60中任一项所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV是通过血管内、肾动脉或静脉内、肺内、小脑延髓池内、脑内、鞘内、静脉内、心室内、脑室内、腹膜内或经皮施用。

[0239] 实施方案62.如实施方案57至61中任一项所述的方法,其中所述受试者患有肾血管平滑肌脂肪瘤。

[0240] 实施方案63.如实施方案62所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV靶向所述肾血管平滑肌脂肪瘤。

[0241] 实施方案64.如实施方案57至63中任一项所述的方法,其中所述受试者表现出淋巴管平滑肌瘤病(LAM)。

[0242] 实施方案65.如实施方案64所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV靶向所述LAM。

[0243] 实施方案66.如实施方案57至65中任一项所述的方法,其中所述受试者患有脑功能障碍。

[0244] 实施方案67.如实施方案66所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV被提供给蛛网膜下腔。

[0245] 实施方案68.如实施方案57至67中任一项所述的方法,其中所述cTuberin、所述核酸分子、所述质粒、所述组合物或所述rAAV被施用于脑细胞、心脏细胞、肾细胞、皮肤细胞或肺细胞。

[0246] 实施方案69.如实施方案57至68中任一项所述的方法,其中所述受试者被施用雷帕霉素。

[0247] 实施方案70.如实施方案57至69中任一项所述的方法,其中所述受试者是人。

[0248] 实施方案71.如实施方案57至70中任一项所述的方法,其中所述受试者的年龄小于18岁。

[0249] 实施方案72.如实施方案71所述的方法,其中所述受试者是婴儿。

[0250] 实施方案73.如实施方案57至72中任一项所述的方法,其中所述受试者已被诊断为患有结节性硬化综合征。

[0251] 实施方案74.如实施方案57至73中任一项所述的方法,其中所述受试者在TSC2基因中具有突变。

[0252] 实施方案75.如实施方案74所述的方法,其中所述受试者在所述TSC2基因的外显子33、外显子37和/或外显子38中具有突变。

[0253] 实施方案76.如实施方案57至75中任一项所述的方法,其中所述受试者具有以下中的一种或多种:皮质结节、室管膜下结节和室管膜下巨细胞星形细胞瘤。

序列表

<110> 桥生物基因疗法研究有限责任公司(BridgeBio Gene Therapy Research, Inc.)

<120> 用于结节性硬化的基因疗法

<130> BGTR-003/01WO 340389-2014

<150> US 63/210,456

<151> 2021-06-14

<160> 36

<170> PatentIn version 3.5

<210> 1

<211> 1807

<212> PRT

<213> 智人(Homo sapiens)

<400> 1

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

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

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

Asp Ile Ile Ile Lys Ala Leu Pro Val Leu Val Val Lys Leu Thr His
820 825 830
Ile Ser Ala Thr Ala Ser Met Ala Val Pro Leu Leu Glu Phe Leu Ser
835 840 845
Thr Leu Ala Arg Leu Pro His Leu Tyr Arg Asn Phe Ala Ala Glu Gln
850 855 860
Tyr Ala Ser Val Phe Ala Ile Ser Leu Pro Tyr Thr Asn Pro Ser Lys
865 870 875 880
Phe Asn Gln Tyr Ile Val Cys Leu Ala His His Val Ile Ala Met Trp
885 890 895
Phe Ile Arg Cys Arg Leu Pro Phe Arg Lys Asp Phe Val Pro Phe Ile
900 905 910
Thr Lys Gly Leu Arg Ser Asn Val Leu Leu Ser Phe Asp Asp Thr Pro
915 920 925
Glu Lys Asp Ser Phe Arg Ala Arg Ser Thr Ser Leu Asn Glu Arg Pro
930 935 940
Lys Ser Leu Arg Ile Ala Arg Pro Pro Lys Gln Gly Leu Asn Asn Ser
945 950 955 960
Pro Pro Val Lys Glu Phe Lys Glu Ser Ser Ala Ala Glu Ala Phe Arg
965 970 975
Cys Arg Ser Ile Ser Val Ser Glu His Val Val Arg Ser Arg Ile Gln
980 985 990
Thr Ser Leu Thr Ser Ala Ser Leu Gly Ser Ala Asp Glu Asn Ser Val
995 1000 1005
Ala Gln Ala Asp Asp Ser Leu Lys Asn Leu His Leu Glu Leu Thr
1010 1015 1020
Glu Thr Cys Leu Asp Met Met Ala Arg Tyr Val Phe Ser Asn Phe
1025 1030 1035
Thr Ala Val Pro Lys Arg Ser Pro Val Gly Glu Phe Leu Leu Ala
1040 1045 1050
Gly Gly Arg Thr Lys Thr Trp Leu Val Gly Asn Lys Leu Val Thr
1055 1060 1065
Val Thr Thr Ser Val Gly Thr Gly Thr Arg Ser Leu Leu Gly Leu
1070 1075 1080
Asp Ser Gly Glu Leu Gln Ser Gly Pro Glu Ser Ser Ser Ser Pro
1085 1090 1095
Gly Val His Val Arg Gln Thr Lys Glu Ala Pro Ala Lys Leu Glu
1100 1105 1110
Ser Gln Ala Gly Gln Gln Val Ser Arg Gly Ala Arg Asp Arg Val

1115 1120 1125

Arg Ser Met Ser Gly Gly His Gly Leu Arg Val Gly Ala Leu Asp

1130 1135 1140

Val Pro Ala Ser Gln Phe Leu Gly Ser Ala Thr Ser Pro Gly Pro

1145 1150 1155

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

1160 1165 1170

Arg Val Pro Val Gln Glu Lys Thr Asn Leu Ala Ala Tyr Val Pro

1175 1180 1185

Leu Leu Thr Gln Gly Trp Ala Glu Ile Leu Val Arg Arg Pro Thr

1190 1195 1200

Gly Asn Thr Ser Trp Leu Met Ser Leu Glu Asn Pro Leu Ser Pro

1205 1210 1215

Phe Ser Ser Asp Ile Asn Asn Met Pro Leu Gln Glu Leu Ser Asn

1220 1225 1230

Ala Leu Met Ala Ala Glu Arg Phe Lys Glu His Arg Asp Thr Ala

1235 1240 1245

Leu Tyr Lys Ser Leu Ser Val Pro Ala Ala Ser Thr Ala Lys Pro

1250 1255 1260

Pro Pro Leu Pro Arg Ser Asn Thr Val Ala Ser Phe Ser Ser Leu

1265 1270 1275

Tyr Gln Ser Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp

1280 1285 1290

Ala Asp Ser Ala Val Val Met Glu Glu Gly Ser Pro Gly Glu Val

1295 1300 1305

Pro Val Leu Val Glu Pro Pro Gly Leu Glu Asp Val Glu Ala Ala

1310 1315 1320

Leu Gly Met Asp Arg Arg Thr Asp Ala Tyr Ser Arg Ser Ser Ser

1325 1330 1335

Val Ser Ser Gln Glu Glu Lys Ser Leu His Ala Glu Glu Leu Val

1340 1345 1350

Gly Arg Gly Ile Pro Ile Glu Arg Val Val Ser Ser Glu Gly Gly

1355 1360 1365

Arg Pro Ser Val Asp Leu Ser Phe Gln Pro Ser Gln Pro Leu Ser

1370 1375 1380

Lys Ser Ser Ser Ser Pro Glu Leu Gln Thr Leu Gln Asp Ile Leu

1385 1390 1395

Gly Asp Pro Gly Asp Lys Ala Asp Val Gly Arg Leu Ser Pro Glu

1400 1405 1410

Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly Glu Ser Ala
1415 1420 1425
Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro Glu Gly
1430 1435 1440
Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg Pro
1445 1450 1455
Arg Gly Tyr Thr Ile Ser Asp Ser Ala Pro Ser Arg Arg Gly Lys
1460 1465 1470
Arg Val Glu Arg Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn
1475 1480 1485
Ala Glu Lys Val Pro Gly Ile Asn Pro Ser Phe Val Phe Leu Gln
1490 1495 1500
Leu Tyr His Ser Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile
1505 1510 1515
Leu Leu Pro Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln Leu
1520 1525 1530
Leu Asp Gln Ile Pro Ser Tyr Asp Thr His Lys Ile Ala Val Leu
1535 1540 1545
Tyr Val Gly Glu Gly Gln Ser Asn Ser Glu Leu Ala Ile Leu Ser
1550 1555 1560
Asn Glu His Gly Ser Tyr Arg Tyr Thr Glu Phe Leu Thr Gly Leu
1565 1570 1575
Gly Arg Leu Ile Glu Leu Lys Asp Cys Gln Pro Asp Lys Val Tyr
1580 1585 1590
Leu Gly Gly Leu Asp Val Cys Gly Glu Asp Gly Gln Phe Thr Tyr
1595 1600 1605
Cys Trp His Asp Asp Ile Met Gln Ala Val Phe His Ile Ala Thr
1610 1615 1620
Leu Met Pro Thr Lys Asp Val Asp Lys His Arg Cys Asp Lys Lys
1625 1630 1635
Arg His Leu Gly Asn Asp Phe Val Ser Ile Val Tyr Asn Asp Ser
1640 1645 1650
Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn Phe
1655 1660 1665
Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val
1670 1675 1680
Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser
1685 1690 1695
Val Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg

1700 1705 1710
Gln Met Ala Leu His Ala Asn Met Ala Ser Gln Val His His Ser
1715 1720 1725
Arg Ser Asn Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg
1730 1735 1740
Leu Arg His Ile Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala
1745 1750 1755
Ala Tyr Ser Asn Pro Ser Leu Pro Leu Val His Pro Pro Ser His
1760 1765 1770
Ser Lys Ala Pro Ala Gln Thr Pro Ala Glu Pro Thr Pro Gly Tyr
1775 1780 1785
Glu Val Gly Gln Arg Lys Arg Leu Ile Ser Ser Val Glu Asp Phe
1790 1795 1800
Thr Glu Phe Val
1805

<210> 2
<211> 4
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 多肽间隔区
<400> 2
Ser Gly Gly Gly
1

<210> 3
<211> 16
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 多肽间隔区
<400> 3
Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly
1 5 10 15

<210> 4
<211> 450
<212> PRT
<213> 智人(Homo sapiens)
<400> 4
Met Ala Lys Pro Thr Ser Lys Asp Ser Gly Leu Lys Glu Lys Phe Lys

1 5 10 15
Ile Leu Leu Gly Leu Gly Thr Pro Arg Pro Asn Pro Arg Ser Ala Glu
20 25 30
Gly Lys Gln Thr Glu Phe Ile Ile Thr Ala Glu Ile Leu Arg Glu Leu
35 40 45
Ser Met Glu Cys Gly Leu Asn Asn Arg Ile Arg Met Ile Gly Gln Ile
50 55 60
Cys Glu Val Ala Lys Thr Lys Lys Phe Glu Glu His Ala Val Glu Ala
65 70 75 80
Leu Trp Lys Ala Val Ala Asp Leu Leu Gln Pro Glu Arg Pro Leu Glu
85 90 95
Ala Arg His Ala Val Leu Ala Leu Leu Lys Ala Ile Val Gln Gly Gln
100 105 110
Gly Glu Arg Leu Gly Val Leu Arg Ala Leu Phe Phe Lys Val Ile Lys
115 120 125
Asp Tyr Pro Ser Asn Glu Asp Leu His Glu Arg Leu Glu Val Phe Lys
130 135 140
Ala Leu Thr Asp Asn Gly Arg His Ile Thr Tyr Leu Glu Glu Glu Leu
145 150 155 160
Ala Asp Phe Val Leu Gln Trp Met Asp Val Gly Leu Ser Ser Glu Phe
165 170 175
Leu Leu Val Leu Val Asn Leu Val Lys Phe Asn Ser Cys Tyr Leu Asp
180 185 190
Glu Tyr Ile Ala Arg Met Val Gln Met Ile Cys Leu Leu Cys Val Arg
195 200 205
Thr Ala Ser Ser Val Asp Ile Glu Val Ser Leu Gln Val Leu Asp Ala
210 215 220
Val Val Cys Tyr Asn Cys Leu Pro Ala Glu Ser Leu Pro Leu Phe Ile
225 230 235 240
Val Thr Leu Cys Arg Thr Ile Asn Val Lys Glu Leu Cys Glu Pro Cys
245 250 255
Trp Lys Leu Met Arg Asn Leu Leu Gly Thr His Leu Gly His Ser Ala
260 265 270
Ile Tyr Asn Met Cys His Leu Met Glu Asp Arg Ala Tyr Met Glu Asp
275 280 285
Ala Pro Leu Leu Arg Gly Ala Val Phe Phe Val Gly Met Ala Leu Trp
290 295 300
Gly Ala His Arg Leu Tyr Ser Leu Arg Asn Ser Pro Thr Ser Val Leu
305 310 315 320

Pro Ser Phe Tyr Gln Ala Met Ala Cys Pro Asn Glu Val Val Ser Tyr
325 330 335
Glu Ile Val Leu Ser Ile Thr Arg Leu Ile Lys Lys Tyr Arg Lys Glu
340 345 350
Leu Gln Val Val Ala Trp Asp Ile Leu Leu Asn Ile Ile Glu Arg Leu
355 360 365
Leu Gln Gln Leu Gln Thr Leu Asp Ser Pro Glu Leu Arg Thr Ile Val
370 375 380
His Asp Leu Leu Thr Thr Val Glu Glu Leu Cys Asp Gln Asn Glu Phe
385 390 395 400
His Gly Ser Gln Glu Arg Tyr Phe Glu Leu Val Glu Arg Cys Ala Asp
405 410 415
Gln Arg Pro Glu Ser Ser Leu Leu Asn Leu Ile Ser Tyr Arg Ala Gln
420 425 430
Ser Ile His Pro Ala Lys Asp Gly Trp Ile Gln Asn Leu Gln Ala Leu
435 440 445
Met Glu
450

<210> 5
<211> 418
<212> PRT
<213> 智人(Homo sapiens)
<400> 5

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

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

<210> 6

<211> 292

<212> PRT
<213> 智人(Homo sapiens)
<400> 6

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

Thr Glu Phe Val
290

<210> 7

<211> 668

<212> PRT

<213> 智人(Homo sapiens)

<400> 7

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

Asp Ile Leu Gly Asp Pro Gly Asp Lys Ala Asp Val Gly Arg Leu Ser
260 265 270
Pro Glu Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly Glu Ser
275 280 285
Ala Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro Glu Gly
290 295 300
Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg Pro Arg
305 310 315 320
Gly Tyr Thr Ile Ser Asp Ser Ala Pro Ser Arg Arg Gly Lys Arg Val
325 330 335
Glu Arg Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala Glu Lys
340 345 350
Val Pro Gly Ile Asn Pro Ser Phe Val Phe Leu Gln Leu Tyr His Ser
355 360 365
Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile Leu Leu Pro Asn Glu
370 375 380
Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu Asp Gln Ile Pro Ser
385 390 395 400
Tyr Asp Thr His Lys Ile Ala Val Leu Tyr Val Gly Glu Gly Gln Ser
405 410 415
Asn Ser Glu Leu Ala Ile Leu Ser Asn Glu His Gly Ser Tyr Arg Tyr
420 425 430
Thr Glu Phe Leu Thr Gly Leu Gly Arg Leu Ile Glu Leu Lys Asp Cys
435 440 445
Gln Pro Asp Lys Val Tyr Leu Gly Gly Leu Asp Val Cys Gly Glu Asp
450 455 460
Gly Gln Phe Thr Tyr Cys Trp His Asp Asp Ile Met Gln Ala Val Phe
465 470 475 480
His Ile Ala Thr Leu Met Pro Thr Lys Asp Val Asp Lys His Arg Cys
485 490 495
Asp Lys Lys Arg His Leu Gly Asn Asp Phe Val Ser Ile Val Tyr Asn
500 505 510
Asp Ser Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn
515 520 525
Phe Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val
530 535 540
Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser Val
545 550 555 560
Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg Gln Met

565 570 575

Ala Leu His Ala Asn Met Ala Ser Gln Val His His Ser Arg Ser Asn
580 585 590

Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg His Ile
595 600 605

Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser Asn Pro
610 615 620

Ser Leu Pro Leu Val His Pro Pro Ser His Ser Lys Ala Pro Ala Gln
625 630 635 640

Thr Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly Gln Arg Lys Arg
645 650 655

Leu Ile Ser Ser Val Glu Asp Phe Thr Glu Phe Val
660 665

<210> 8

<211> 698

<212> PRT

<213> 智人(Homo sapiens)

<400> 8

Ala Lys Leu Glu Ser Gln Ala Gly Gln Gln Val Ser Arg Gly Ala Arg
1 5 10 15

Asp Arg Val Arg Ser Met Ser Gly Gly His Gly Leu Arg Val Gly Ala
20 25 30

Leu Asp Val Pro Ala Ser Gln Phe Leu Gly Ser Ala Thr Ser Pro Gly
35 40 45

Pro Arg Thr Ala Pro Ala Ala Lys Pro Glu Lys Ala Ser Ala Gly Thr
50 55 60

Arg Val Pro Val Gln Glu Lys Thr Asn Leu Ala Ala Tyr Val Pro Leu
65 70 75 80

Leu Thr Gln Gly Trp Ala Glu Ile Leu Val Arg Arg Pro Thr Gly Asn
85 90 95

Thr Ser Trp Leu Met Ser Leu Glu Asn Pro Leu Ser Pro Phe Ser Ser
100 105 110

Asp Ile Asn Asn Met Pro Leu Gln Glu Leu Ser Asn Ala Leu Met Ala
115 120 125

Ala Glu Arg Phe Lys Glu His Arg Asp Thr Ala Leu Tyr Lys Ser Leu
130 135 140

Ser Val Pro Ala Ala Ser Thr Ala Lys Pro Pro Pro Leu Pro Arg Ser
145 150 155 160

Asn Thr Val Ala Ser Phe Ser Ser Leu Tyr Gln Ser Ser Cys Gln Gly

165 170 175
Gln Leu His Arg Ser Val Ser Trp Ala Asp Ser Ala Val Val Met Glu
180 185 190
Glu Gly Ser Pro Gly Glu Val Pro Val Leu Val Glu Pro Pro Gly Leu
195 200 205
Glu Asp Val Glu Ala Ala Leu Gly Met Asp Arg Arg Thr Asp Ala Tyr
210 215 220
Ser Arg Ser Ser Ser Val Ser Ser Gln Glu Glu Lys Ser Leu His Ala
225 230 235 240
Glu Glu Leu Val Gly Arg Gly Ile Pro Ile Glu Arg Val Val Ser Ser
245 250 255
Glu Gly Gly Arg Pro Ser Val Asp Leu Ser Phe Gln Pro Ser Gln Pro
260 265 270
Leu Ser Lys Ser Ser Ser Ser Pro Glu Leu Gln Thr Leu Gln Asp Ile
275 280 285
Leu Gly Asp Pro Gly Asp Lys Ala Asp Val Gly Arg Leu Ser Pro Glu
290 295 300
Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly Glu Ser Ala Ala
305 310 315 320
Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro Glu Gly Pro Leu
325 330 335
Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg Pro Arg Gly Tyr
340 345 350
Thr Ile Ser Asp Ser Ala Pro Ser Arg Arg Gly Lys Arg Val Glu Arg
355 360 365
Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala Glu Lys Val Pro
370 375 380
Gly Ile Asn Pro Ser Phe Val Phe Leu Gln Leu Tyr His Ser Pro Phe
385 390 395 400
Phe Gly Asp Glu Ser Asn Lys Pro Ile Leu Leu Pro Asn Glu Ser Gln
405 410 415
Ser Phe Glu Arg Ser Val Gln Leu Leu Asp Gln Ile Pro Ser Tyr Asp
420 425 430
Thr His Lys Ile Ala Val Leu Tyr Val Gly Glu Gly Gln Ser Asn Ser
435 440 445
Glu Leu Ala Ile Leu Ser Asn Glu His Gly Ser Tyr Arg Tyr Thr Glu
450 455 460
Phe Leu Thr Gly Leu Gly Arg Leu Ile Glu Leu Lys Asp Cys Gln Pro
465 470 475 480

Asp Lys Val Tyr Leu Gly Gly Leu Asp Val Cys Gly Glu Asp Gly Gln
485 490 495
Phe Thr Tyr Cys Trp His Asp Asp Ile Met Gln Ala Val Phe His Ile
500 505 510
Ala Thr Leu Met Pro Thr Lys Asp Val Asp Lys His Arg Cys Asp Lys
515 520 525
Lys Arg His Leu Gly Asn Asp Phe Val Ser Ile Val Tyr Asn Asp Ser
530 535 540
Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn Phe Val
545 550 555 560
His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val Ser Leu
565 570 575
Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser Val Ala Lys
580 585 590
Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg Gln Met Ala Leu
595 600 605
His Ala Asn Met Ala Ser Gln Val His His Ser Arg Ser Asn Pro Thr
610 615 620
Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg His Ile Lys Arg
625 630 635 640
Leu Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser Asn Pro Ser Leu
645 650 655
Pro Leu Val His Pro Pro Ser His Ser Lys Ala Pro Ala Gln Thr Pro
660 665 670
Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly Gln Arg Lys Arg Leu Ile
675 680 685
Ser Ser Val Glu Asp Phe Thr Glu Phe Val
690 695

<210> 9
<211> 875
<212> PRT
<213> 智人(Homo sapiens)
<400> 9

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

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

355 360 365

Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu Val Glu Pro Pro Gly

370 375 380

Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp Arg Arg Thr Asp Ala

385 390 395 400

Tyr Ser Arg Ser Ser Ser Val Ser Ser Gln Glu Glu Lys Ser Leu His

405 410 415

Ala Glu Glu Leu Val Gly Arg Gly Ile Pro Ile Glu Arg Val Val Ser

420 425 430

Ser Glu Gly Gly Arg Pro Ser Val Asp Leu Ser Phe Gln Pro Ser Gln

435 440 445

Pro Leu Ser Lys Ser Ser Ser Ser Pro Glu Leu Gln Thr Leu Gln Asp

450 455 460

Ile Leu Gly Asp Pro Gly Asp Lys Ala Asp Val Gly Arg Leu Ser Pro

465 470 475 480

Glu Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly Glu Ser Ala

485 490 495

Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro Glu Gly Pro

500 505 510

Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg Pro Arg Gly

515 520 525

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

530 535 540

Arg Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala Glu Lys Val

545 550 555 560

Pro Gly Ile Asn Pro Ser Phe Val Phe Leu Gln Leu Tyr His Ser Pro

565 570 575

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

580 585 590

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

595 600 605

Asp Thr His Lys Ile Ala Val Leu Tyr Val Gly Glu Gly Gln Ser Asn

610 615 620

Ser Glu Leu Ala Ile Leu Ser Asn Glu His Gly Ser Tyr Arg Tyr Thr

625 630 635 640

Glu Phe Leu Thr Gly Leu Gly Arg Leu Ile Glu Leu Lys Asp Cys Gln

645 650 655

Pro Asp Lys Val Tyr Leu Gly Gly Leu Asp Val Cys Gly Glu Asp Gly

660 665 670

Gln Phe Thr Tyr Cys Trp His Asp Asp Ile Met Gln Ala Val Phe His
675 680 685
Ile Ala Thr Leu Met Pro Thr Lys Asp Val Asp Lys His Arg Cys Asp
690 695 700
Lys Lys Arg His Leu Gly Asn Asp Phe Val Ser Ile Val Tyr Asn Asp
705 710 715 720
Ser Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn Phe
725 730 735
Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val Ser
740 745 750
Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser Val Ala
755 760 765
Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg Gln Met Ala
770 775 780
Leu His Ala Asn Met Ala Ser Gln Val His His Ser Arg Ser Asn Pro
785 790 795 800
Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg His Ile Lys
805 810 815
Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser Asn Pro Ser
820 825 830
Leu Pro Leu Val His Pro Pro Ser His Ser Lys Ala Pro Ala Gln Thr
835 840 845
Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly Gln Arg Lys Arg Leu
850 855 860
Ile Ser Ser Val Glu Asp Phe Thr Glu Phe Val
865 870 875

<210> 10
<211> 833
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 10

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

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

355 360 365

Glu Glu Lys Ser Leu His Ala Glu Glu Leu Val Gly Arg Gly Ile Pro

370 375 380

Ile Glu Arg Val Val Ser Ser Glu Gly Gly Arg Pro Ser Val Asp Leu

385 390 395 400

Ser Phe Gln Pro Ser Gln Pro Leu Ser Lys Ser Ser Ser Ser Pro Glu

405 410 415

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

420 425 430

Val Gly Arg Leu Ser Pro Glu Val Lys Ala Arg Ser Gln Ser Gly Thr

435 440 445

Leu Asp Gly Glu Ser Ala Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg

450 455 460

Gly Gln Pro Glu Gly Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser

465 470 475 480

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

485 490 495

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

500 505 510

Ser Asn Ala Glu Lys Val Pro Gly Ile Asn Pro Ser Phe Val Phe Leu

515 520 525

Gln Leu Tyr His Ser Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile

530 535 540

Leu Leu Pro Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu

545 550 555 560

Asp Gln Ile Pro Ser Tyr Asp Thr His Lys Ile Ala Val Leu Tyr Val

565 570 575

Gly Glu Gly Gln Ser Asn Ser Glu Leu Ala Ile Leu Ser Asn Glu His

580 585 590

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

595 600 605

Glu Leu Lys Asp Cys Gln Pro Asp Lys Val Tyr Leu Gly Gly Leu Asp

610 615 620

Val Cys Gly Glu Asp Gly Gln Phe Thr Tyr Cys Trp His Asp Asp Ile

625 630 635 640

Met Gln Ala Val Phe His Ile Ala Thr Leu Met Pro Thr Lys Asp Val

645 650 655

Asp Lys His Arg Cys Asp Lys Lys Arg His Leu Gly Asn Asp Phe Val

660 665 670

Ser Ile Val Tyr Asn Asp Ser Gly Glu Asp Phe Lys Leu Gly Thr Ile
675 680 685
Lys Gly Gln Phe Asn Phe Val His Val Ile Val Thr Pro Leu Asp Tyr
690 695 700
Glu Cys Asn Leu Val Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu
705 710 715 720
Val Asp Thr Ser Val Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe
725 730 735
Val Ala Arg Gln Met Ala Leu His Ala Asn Met Ala Ser Gln Val His
740 745 750
His Ser Arg Ser Asn Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala
755 760 765
Arg Leu Arg His Ile Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala
770 775 780
Ala Tyr Ser Asn Pro Ser Leu Pro Leu Val His Pro Pro Ser His Ser
785 790 795 800
Lys Ala Pro Ala Gln Thr Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val
805 810 815
Gly Gln Arg Lys Arg Leu Ile Ser Ser Val Glu Asp Phe Thr Glu Phe
820 825 830
Val

<210> 11
<211> 766
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 11

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

85 90 95
Thr Ser Val Gly Thr Gly Thr Arg Ser Leu Leu Gly Leu Asp Ser Gly
100 105 110
Glu Leu Gln Ser Gly Pro Glu Ser Ser Ser Ser Pro Gly Val His Val
115 120 125
Arg Gln Thr Lys Glu Ala Pro Ala Lys Leu Glu Ser Gln Ala Gly Gln
130 135 140
Gln Val Ser Arg Gly Ala Arg Asp Arg Val Arg Ser Met Ser Gly Gly
145 150 155 160
His Gly Leu Arg Val Gly Ala Leu Asp Val Pro Ala Ser Gln Phe Leu
165 170 175
Gly Ser Ala Thr Ser Pro Gly Pro Arg Thr Ala Pro Ala Ala Lys Pro
180 185 190
Glu Lys Ala Ser Ala Gly Thr Arg Val Pro Val Gln Glu Lys Thr Asn
195 200 205
Leu Ala Ala Tyr Val Pro Leu Leu Thr Gln Gly Trp Ala Glu Ile Leu
210 215 220
Val Arg Arg Pro Thr Gly Val Ala Ser Phe Ser Ser Leu Tyr Gln Ser
225 230 235 240
Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp Ala Asp Ser Ala
245 250 255
Val Val Met Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu Val Glu
260 265 270
Pro Pro Gly Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp Arg Arg
275 280 285
Thr Asp Ala Tyr Ser Arg Ser Ser Ser Val Ser Ser Gln Glu Glu Lys
290 295 300
Ser Leu His Ala Glu Glu Leu Val Gly Arg Gly Ile Pro Ile Glu Arg
305 310 315 320
Val Val Ser Ser Glu Gly Gly Arg Pro Ser Val Asp Leu Ser Phe Gln
325 330 335
Pro Ser Gln Pro Leu Ser Lys Ser Ser Ser Ser Pro Glu Leu Gln Thr
340 345 350
Leu Gln Asp Ile Leu Gly Asp Pro Gly Asp Lys Ala Asp Val Gly Arg
355 360 365
Leu Ser Pro Glu Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly
370 375 380
Glu Ser Ala Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro
385 390 395 400

Glu Gly Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg
405 410 415
Pro Arg Gly Tyr Thr Ile Ser Asp Ser Ala Pro Ser Arg Arg Gly Lys
420 425 430
Arg Val Glu Arg Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala
435 440 445
Glu Lys Val Pro Gly Ile Asn Pro Ser Phe Val Phe Leu Gln Leu Tyr
450 455 460
His Ser Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile Leu Leu Pro
465 470 475 480
Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu Asp Gln Ile
485 490 495
Pro Ser Tyr Asp Thr His Lys Ile Ala Val Leu Tyr Val Gly Glu Gly
500 505 510
Gln Ser Asn Ser Glu Leu Ala Ile Leu Ser Asn Glu His Gly Ser Tyr
515 520 525
Arg Tyr Thr Glu Phe Leu Thr Gly Leu Gly Arg Leu Ile Glu Leu Lys
530 535 540
Asp Cys Gln Pro Asp Lys Val Tyr Leu Gly Gly Leu Asp Val Cys Gly
545 550 555 560
Glu Asp Gly Gln Phe Thr Tyr Cys Trp His Asp Asp Ile Met Gln Ala
565 570 575
Val Phe His Ile Ala Thr Leu Met Pro Thr Lys Asp Val Asp Lys His
580 585 590
Arg Cys Asp Lys Lys Arg His Leu Gly Asn Asp Phe Val Ser Ile Val
595 600 605
Tyr Asn Asp Ser Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln
610 615 620
Phe Asn Phe Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn
625 630 635 640
Leu Val Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr
645 650 655
Ser Val Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg
660 665 670
Gln Met Ala Leu His Ala Asn Met Ala Ser Gln Val His His Ser Arg
675 680 685
Ser Asn Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg
690 695 700
His Ile Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser

705 710 715 720
Asn Pro Ser Leu Pro Leu Val His Pro Pro Ser His Ser Lys Ala Pro
725 730 735
Ala Gln Thr Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly Gln Arg
740 745 750
Lys Arg Leu Ile Ser Ser Val Glu Asp Phe Thr Glu Phe Val
755 760 765

<210> 12
<211> 604
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 12

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

195 200 205

Leu Ala Ala Tyr Val Pro Leu Leu Thr Gln Gly Trp Ala Glu Ile Leu
210 215 220

Val Arg Arg Pro Thr Gly Val Ala Ser Phe Ser Ser Leu Tyr Gln Ser
225 230 235 240

Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp Ala Asp Ser Ala
245 250 255

Val Val Met Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu Val Glu
260 265 270

Pro Pro Gly Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp Arg Arg
275 280 285

Thr Asp Ala Tyr Ser Arg Ser Phe Val Phe Leu Gln Leu Tyr His Ser
290 295 300

Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile Leu Leu Pro Asn Glu
305 310 315 320

Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu Asp Gln Ile Pro Ser
325 330 335

Tyr Asp Thr His Lys Ile Ala Val Leu Tyr Val Gly Glu Gly Gln Ser
340 345 350

Asn Ser Glu Leu Ala Ile Leu Ser Asn Glu His Gly Ser Tyr Arg Tyr
355 360 365

Thr Glu Phe Leu Thr Gly Leu Gly Arg Leu Ile Glu Leu Lys Asp Cys
370 375 380

Gln Pro Asp Lys Val Tyr Leu Gly Gly Leu Asp Val Cys Gly Glu Asp
385 390 395 400

Gly Gln Phe Thr Tyr Cys Trp His Asp Asp Ile Met Gln Ala Val Phe
405 410 415

His Ile Ala Thr Leu Met Pro Thr Lys Asp Val Asp Lys His Arg Cys
420 425 430

Asp Lys Lys Arg His Leu Gly Asn Asp Phe Val Ser Ile Val Tyr Asn
435 440 445

Asp Ser Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn
450 455 460

Phe Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val
465 470 475 480

Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser Val
485 490 495

Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg Gln Met
500 505 510

Ala Leu His Ala Asn Met Ala Ser Gln Val His His Ser Arg Ser Asn
515 520 525
Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg His Ile
530 535 540
Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser Asn Pro
545 550 555 560
Ser Leu Pro Leu Val His Pro Pro Ser His Ser Lys Ala Pro Ala Gln
565 570 575
Thr Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly Gln Arg Lys Arg
580 585 590
Leu Ile Ser Ser Val Glu Asp Phe Thr Glu Phe Val
595 600

<210> 13
<211> 758
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 13

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

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

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

<210> 14
<211> 1134

<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 14

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

Ile Tyr Asn Met Cys His Leu Met Glu Asp Arg Ala Tyr Met Glu Asp
275 280 285
Ala Pro Leu Leu Arg Gly Ala Val Phe Phe Val Gly Met Ala Leu Trp
290 295 300
Gly Ala His Arg Leu Tyr Ser Leu Arg Asn Ser Pro Thr Ser Val Leu
305 310 315 320
Pro Ser Phe Tyr Gln Ala Met Ala Cys Pro Asn Glu Val Val Ser Tyr
325 330 335
Glu Ile Val Leu Ser Ile Thr Arg Leu Ile Lys Lys Tyr Arg Lys Glu
340 345 350
Leu Gln Val Val Ala Trp Asp Ile Leu Leu Asn Ile Ile Glu Arg Leu
355 360 365
Leu Gln Gln Leu Gln Thr Leu Asp Ser Pro Glu Leu Arg Thr Ile Val
370 375 380
His Asp Leu Leu Thr Thr Val Glu Glu Leu Cys Asp Gln Asn Glu Phe
385 390 395 400
His Gly Ser Gln Glu Arg Tyr Phe Glu Leu Val Glu Arg Cys Ala Asp
405 410 415
Gln Arg Pro Glu Ser Ser Leu Leu Asn Leu Ile Ser Tyr Arg Ala Gln
420 425 430
Ser Ile His Pro Ala Lys Asp Gly Trp Ile Gln Asn Leu Gln Ala Leu
435 440 445
Met Glu Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
450 455 460
Gly Gly Gly Ala Leu Asp Val Pro Ala Ser Gln Phe Leu Gly Ser Ala
465 470 475 480
Thr Ser Pro Gly Pro Arg Thr Ala Pro Ala Ala Lys Pro Glu Lys Ala
485 490 495
Ser Ala Gly Thr Arg Val Pro Val Gln Glu Lys Thr Asn Leu Ala Ala
500 505 510
Tyr Val Pro Leu Leu Thr Gln Gly Trp Ala Glu Ile Leu Val Arg Arg
515 520 525
Pro Thr Gly Asn Thr Ser Trp Leu Met Ser Leu Glu Asn Pro Leu Ser
530 535 540
Pro Phe Ser Ser Asp Ile Asn Asn Met Pro Leu Gln Glu Leu Ser Asn
545 550 555 560
Ala Leu Met Ala Ala Glu Arg Phe Lys Glu His Arg Asp Thr Ala Leu
565 570 575
Tyr Lys Ser Leu Ser Val Pro Ala Ala Ser Thr Ala Lys Pro Pro Pro

580 585 590

Leu Pro Arg Ser Asn Thr Val Ala Ser Phe Ser Ser Leu Tyr Gln Ser

595 600 605

Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp Ala Asp Ser Ala

610 615 620

Val Val Met Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu Val Glu

625 630 635 640

Pro Pro Gly Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp Arg Arg

645 650 655

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

660 665 670

Ser Leu His Ala Glu Glu Leu Val Gly Arg Gly Ile Pro Ile Glu Arg

675 680 685

Val Val Ser Ser Glu Gly Gly Arg Pro Ser Val Asp Leu Ser Phe Gln

690 695 700

Pro Ser Gln Pro Leu Ser Lys Ser Ser Ser Ser Pro Glu Leu Gln Thr

705 710 715 720

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

725 730 735

Leu Ser Pro Glu Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly

740 745 750

Glu Ser Ala Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro

755 760 765

Glu Gly Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg

770 775 780

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

785 790 795 800

Arg Val Glu Arg Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala

805 810 815

Glu Lys Val Pro Gly Ile Asn Pro Ser Phe Val Phe Leu Gln Leu Tyr

820 825 830

His Ser Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile Leu Leu Pro

835 840 845

Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu Asp Gln Ile

850 855 860

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

865 870 875 880

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

885 890 895

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

<210> 15
<211> 1102
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 15

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

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

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

930 935 940
Tyr Asn Asp Ser Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln
945 950 955 960
Phe Asn Phe Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn
965 970 975
Leu Val Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr
980 985 990
Ser Val Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg
995 1000 1005
Gln Met Ala Leu His Ala Asn Met Ala Ser Gln Val His His Ser
1010 1015 1020
Arg Ser Asn Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg
1025 1030 1035
Leu Arg His Ile Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala
1040 1045 1050
Ala Tyr Ser Asn Pro Ser Leu Pro Leu Val His Pro Pro Ser His
1055 1060 1065
Ser Lys Ala Pro Ala Gln Thr Pro Ala Glu Pro Thr Pro Gly Tyr
1070 1075 1080
Glu Val Gly Gln Arg Lys Arg Leu Ile Ser Ser Val Glu Asp Phe
1085 1090 1095
Thr Glu Phe Val
1100

<210> 16
<211> 1309
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 16
Met Ala Lys Pro Thr Ser Lys Asp Ser Gly Leu Lys Glu Lys Phe Lys
1 5 10 15
Ile Leu Leu Gly Leu Gly Thr Pro Arg Pro Asn Pro Arg Ser Ala Glu
20 25 30
Gly Lys Gln Thr Glu Phe Ile Ile Thr Ala Glu Ile Leu Arg Glu Leu
35 40 45
Ser Met Glu Cys Gly Leu Asn Asn Arg Ile Arg Met Ile Gly Gln Ile
50 55 60
Cys Glu Val Ala Lys Thr Lys Lys Phe Glu Glu His Ala Val Glu Ala

65 70 75 80
Leu Trp Lys Ala Val Ala Asp Leu Leu Gln Pro Glu Arg Pro Leu Glu
85 90 95
Ala Arg His Ala Val Leu Ala Leu Leu Lys Ala Ile Val Gln Gly Gln
100 105 110
Gly Glu Arg Leu Gly Val Leu Arg Ala Leu Phe Phe Lys Val Ile Lys
115 120 125
Asp Tyr Pro Ser Asn Glu Asp Leu His Glu Arg Leu Glu Val Phe Lys
130 135 140
Ala Leu Thr Asp Asn Gly Arg His Ile Thr Tyr Leu Glu Glu Glu Leu
145 150 155 160
Ala Asp Phe Val Leu Gln Trp Met Asp Val Gly Leu Ser Ser Glu Phe
165 170 175
Leu Leu Val Leu Val Asn Leu Val Lys Phe Asn Ser Cys Tyr Leu Asp
180 185 190
Glu Tyr Ile Ala Arg Met Val Gln Met Ile Cys Leu Leu Cys Val Arg
195 200 205
Thr Ala Ser Ser Val Asp Ile Glu Val Ser Leu Gln Val Leu Asp Ala
210 215 220
Val Val Cys Tyr Asn Cys Leu Pro Ala Glu Ser Leu Pro Leu Phe Ile
225 230 235 240
Val Thr Leu Cys Arg Thr Ile Asn Val Lys Glu Leu Cys Glu Pro Cys
245 250 255
Trp Lys Leu Met Arg Asn Leu Leu Gly Thr His Leu Gly His Ser Ala
260 265 270
Ile Tyr Asn Met Cys His Leu Met Glu Asp Arg Ala Tyr Met Glu Asp
275 280 285
Ala Pro Leu Leu Arg Gly Ala Val Phe Phe Val Gly Met Ala Leu Trp
290 295 300
Gly Ala His Arg Leu Tyr Ser Leu Arg Asn Ser Pro Thr Ser Val Leu
305 310 315 320
Pro Ser Phe Tyr Gln Ala Met Ala Cys Pro Asn Glu Val Val Ser Tyr
325 330 335
Glu Ile Val Leu Ser Ile Thr Arg Leu Ile Lys Lys Tyr Arg Lys Glu
340 345 350
Leu Gln Val Val Ala Trp Asp Ile Leu Leu Asn Ile Ile Glu Arg Leu
355 360 365
Leu Gln Gln Leu Gln Thr Leu Asp Ser Pro Glu Leu Arg Thr Ile Val
370 375 380

His Asp Leu Leu Thr Thr Val Glu Glu Leu Cys Asp Gln Asn Glu Phe
385 390 395 400
His Gly Ser Gln Glu Arg Tyr Phe Glu Leu Val Glu Arg Cys Ala Asp
405 410 415
Gln Arg Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
420 425 430
Gly Gly Phe Arg Ala Arg Ser Thr Ser Leu Asn Glu Arg Pro Lys Ser
435 440 445
Leu Arg Ile Ala Arg Pro Pro Lys Gln Gly Leu Asn Asn Ser Pro Pro
450 455 460
Val Lys Glu Phe Lys Glu Ser Ser Ala Ala Glu Ala Phe Arg Cys Arg
465 470 475 480
Ser Ile Ser Val Ser Glu His Val Val Arg Ser Arg Ile Gln Thr Ser
485 490 495
Leu Thr Ser Ala Ser Leu Gly Ser Ala Asp Glu Asn Ser Val Ala Gln
500 505 510
Ala Asp Asp Ser Leu Lys Asn Leu His Leu Glu Leu Thr Glu Thr Cys
515 520 525
Leu Asp Met Met Ala Arg Tyr Val Phe Ser Asn Phe Thr Ala Val Pro
530 535 540
Lys Arg Ser Pro Val Gly Glu Phe Leu Leu Ala Gly Gly Arg Thr Lys
545 550 555 560
Thr Trp Leu Val Gly Asn Lys Leu Val Thr Val Thr Thr Ser Val Gly
565 570 575
Thr Gly Thr Arg Ser Leu Leu Gly Leu Asp Ser Gly Glu Leu Gln Ser
580 585 590
Gly Pro Glu Ser Ser Ser Ser Pro Gly Val His Val Arg Gln Thr Lys
595 600 605
Glu Ala Pro Ala Lys Leu Glu Ser Gln Ala Gly Gln Gln Val Ser Arg
610 615 620
Gly Ala Arg Asp Arg Val Arg Ser Met Ser Gly Gly His Gly Leu Arg
625 630 635 640
Val Gly Ala Leu Asp Val Pro Ala Ser Gln Phe Leu Gly Ser Ala Thr
645 650 655
Ser Pro Gly Pro Arg Thr Ala Pro Ala Ala Lys Pro Glu Lys Ala Ser
660 665 670
Ala Gly Thr Arg Val Pro Val Gln Glu Lys Thr Asn Leu Ala Ala Tyr
675 680 685
Val Pro Leu Leu Thr Gln Gly Trp Ala Glu Ile Leu Val Arg Arg Pro

690 695 700

Thr Gly Asn Thr Ser Trp Leu Met Ser Leu Glu Asn Pro Leu Ser Pro
705 710 715 720

Phe Ser Ser Asp Ile Asn Asn Met Pro Leu Gln Glu Leu Ser Asn Ala
725 730 735

Leu Met Ala Ala Glu Arg Phe Lys Glu His Arg Asp Thr Ala Leu Tyr
740 745 750

Lys Ser Leu Ser Val Pro Ala Ala Ser Thr Ala Lys Pro Pro Pro Leu
755 760 765

Pro Arg Ser Asn Thr Val Ala Ser Phe Ser Ser Leu Tyr Gln Ser Ser
770 775 780

Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp Ala Asp Ser Ala Val
785 790 795 800

Val Met Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu Val Glu Pro
805 810 815

Pro Gly Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp Arg Arg Thr
820 825 830

Asp Ala Tyr Ser Arg Ser Ser Ser Val Ser Ser Gln Glu Glu Lys Ser
835 840 845

Leu His Ala Glu Glu Leu Val Gly Arg Gly Ile Pro Ile Glu Arg Val
850 855 860

Val Ser Ser Glu Gly Gly Arg Pro Ser Val Asp Leu Ser Phe Gln Pro
865 870 875 880

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

Gln Asp Ile Leu Gly Asp Pro Gly Asp Lys Ala Asp Val Gly Arg Leu
900 905 910

Ser Pro Glu Val Lys Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly Glu
915 920 925

Ser Ala Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro Glu
930 935 940

Gly Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg Pro
945 950 955 960

Arg Gly Tyr Thr Ile Ser Asp Ser Ala Pro Ser Arg Arg Gly Lys Arg
965 970 975

Val Glu Arg Asp Ala Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala Glu
980 985 990

Lys Val Pro Gly Ile Asn Pro Ser Phe Val Phe Leu Gln Leu Tyr His
995 1000 1005

Ser Pro Phe Phe Gly Asp Glu Ser Asn Lys Pro Ile Leu Leu Pro
1010 1015 1020
Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu Asp Gln
1025 1030 1035
Ile Pro Ser Tyr Asp Thr His Lys Ile Ala Val Leu Tyr Val Gly
1040 1045 1050
Glu Gly Gln Ser Asn Ser Glu Leu Ala Ile Leu Ser Asn Glu His
1055 1060 1065
Gly Ser Tyr Arg Tyr Thr Glu Phe Leu Thr Gly Leu Gly Arg Leu
1070 1075 1080
Ile Glu Leu Lys Asp Cys Gln Pro Asp Lys Val Tyr Leu Gly Gly
1085 1090 1095
Leu Asp Val Cys Gly Glu Asp Gly Gln Phe Thr Tyr Cys Trp His
1100 1105 1110
Asp Asp Ile Met Gln Ala Val Phe His Ile Ala Thr Leu Met Pro
1115 1120 1125
Thr Lys Asp Val Asp Lys His Arg Cys Asp Lys Lys Arg His Leu
1130 1135 1140
Gly Asn Asp Phe Val Ser Ile Val Tyr Asn Asp Ser Gly Glu Asp
1145 1150 1155
Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn Phe Val His Val
1160 1165 1170
Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val Ser Leu Gln
1175 1180 1185
Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser Val Ala Lys
1190 1195 1200
Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg Gln Met Ala
1205 1210 1215
Leu His Ala Asn Met Ala Ser Gln Val His His Ser Arg Ser Asn
1220 1225 1230
Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg His
1235 1240 1245
Ile Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser
1250 1255 1260
Asn Pro Ser Leu Pro Leu Val His Pro Pro Ser His Ser Lys Ala
1265 1270 1275
Pro Ala Gln Thr Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly
1280 1285 1290
Gln Arg Lys Arg Leu Ile Ser Ser Val Glu Asp Phe Thr Glu Phe

1295 1300 1305

Val

<210> 17

<211> 1267

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 17

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

Val Thr Leu Cys Arg Thr Ile Asn Val Lys Glu Leu Cys Glu Pro Cys
245 250 255
Trp Lys Leu Met Arg Asn Leu Leu Gly Thr His Leu Gly His Ser Ala
260 265 270
Ile Tyr Asn Met Cys His Leu Met Glu Asp Arg Ala Tyr Met Glu Asp
275 280 285
Ala Pro Leu Leu Arg Gly Ala Val Phe Phe Val Gly Met Ala Leu Trp
290 295 300
Gly Ala His Arg Leu Tyr Ser Leu Arg Asn Ser Pro Thr Ser Val Leu
305 310 315 320
Pro Ser Phe Tyr Gln Ala Met Ala Cys Pro Asn Glu Val Val Ser Tyr
325 330 335
Glu Ile Val Leu Ser Ile Thr Arg Leu Ile Lys Lys Tyr Arg Lys Glu
340 345 350
Leu Gln Val Val Ala Trp Asp Ile Leu Leu Asn Ile Ile Glu Arg Leu
355 360 365
Leu Gln Gln Leu Gln Thr Leu Asp Ser Pro Glu Leu Arg Thr Ile Val
370 375 380
His Asp Leu Leu Thr Thr Val Glu Glu Leu Cys Asp Gln Asn Glu Phe
385 390 395 400
His Gly Ser Gln Glu Arg Tyr Phe Glu Leu Val Glu Arg Cys Ala Asp
405 410 415
Gln Arg Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
420 425 430
Gly Gly Phe Arg Ala Arg Ser Thr Ser Leu Asn Glu Arg Pro Lys Ser
435 440 445
Ser Arg Ile Gln Thr Ser Leu Thr Ser Ala Ser Leu Gly Ser Ala Asp
450 455 460
Glu Asn Ser Val Ala Gln Ala Asp Asp Ser Leu Lys Asn Leu His Leu
465 470 475 480
Glu Leu Thr Glu Thr Cys Leu Asp Met Met Ala Arg Tyr Val Phe Ser
485 490 495
Asn Phe Thr Ala Val Pro Lys Arg Ser Pro Val Gly Glu Phe Leu Leu
500 505 510
Ala Gly Gly Arg Thr Lys Thr Trp Leu Val Gly Asn Lys Leu Val Thr
515 520 525
Val Thr Thr Ser Val Gly Thr Gly Thr Arg Ser Leu Leu Gly Leu Asp
530 535 540
Ser Gly Glu Leu Gln Ser Gly Pro Glu Ser Ser Ser Ser Pro Gly Val

545 550 555 560
His Val Arg Gln Thr Lys Glu Ala Pro Ala Lys Leu Glu Ser Gln Ala
565 570 575
Gly Gln Gln Val Ser Arg Gly Ala Arg Asp Arg Val Arg Ser Met Ser
580 585 590
Gly Gly His Gly Leu Arg Val Gly Ala Leu Asp Val Pro Ala Ser Gln
595 600 605
Phe Leu Gly Ser Ala Thr Ser Pro Gly Pro Arg Thr Ala Pro Ala Ala
610 615 620
Lys Pro Glu Lys Ala Ser Ala Gly Thr Arg Val Pro Val Gln Glu Lys
625 630 635 640
Thr Asn Leu Ala Ala Tyr Val Pro Leu Leu Thr Gln Gly Trp Ala Glu
645 650 655
Ile Leu Val Arg Arg Pro Thr Gly Asn Thr Ser Trp Leu Met Ser Leu
660 665 670
Glu Asn Pro Leu Ser Pro Phe Ser Ser Asp Ile Asn Asn Met Pro Leu
675 680 685
Gln Glu Leu Ser Asn Ala Leu Met Ala Ala Glu Arg Phe Lys Glu His
690 695 700
Arg Asp Thr Ala Leu Tyr Lys Ser Leu Ser Val Pro Ala Ala Ser Thr
705 710 715 720
Ala Lys Pro Pro Pro Leu Pro Arg Ser Asn Thr Val Ala Ser Phe Ser
725 730 735
Ser Leu Tyr Gln Ser Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser
740 745 750
Trp Ala Asp Ser Ala Val Val Met Glu Glu Gly Ser Pro Gly Glu Val
755 760 765
Pro Val Leu Val Glu Pro Pro Gly Leu Glu Asp Val Glu Ala Ala Leu
770 775 780
Gly Met Asp Arg Arg Thr Asp Ala Tyr Ser Arg Ser Ser Ser Val Ser
785 790 795 800
Ser Gln Glu Glu Lys Ser Leu His Ala Glu Glu Leu Val Gly Arg Gly
805 810 815
Ile Pro Ile Glu Arg Val Val Ser Ser Glu Gly Gly Arg Pro Ser Val
820 825 830
Asp Leu Ser Phe Gln Pro Ser Gln Pro Leu Ser Lys Ser Ser Ser Ser
835 840 845
Pro Glu Leu Gln Thr Leu Gln Asp Ile Leu Gly Asp Pro Gly Asp Lys
850 855 860

Ala Asp Val Gly Arg Leu Ser Pro Glu Val Lys Ala Arg Ser Gln Ser
865 870 875 880
Gly Thr Leu Asp Gly Glu Ser Ala Ala Trp Ser Ala Ser Gly Glu Asp
885 890 895
Ser Arg Gly Gln Pro Glu Gly Pro Leu Pro Ser Ser Ser Pro Arg Ser
900 905 910
Pro Ser Gly Leu Arg Pro Arg Gly Tyr Thr Ile Ser Asp Ser Ala Pro
915 920 925
Ser Arg Arg Gly Lys Arg Val Glu Arg Asp Ala Leu Lys Ser Arg Ala
930 935 940
Thr Ala Ser Asn Ala Glu Lys Val Pro Gly Ile Asn Pro Ser Phe Val
945 950 955 960
Phe Leu Gln Leu Tyr His Ser Pro Phe Phe Gly Asp Glu Ser Asn Lys
965 970 975
Pro Ile Leu Leu Pro Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln
980 985 990
Leu Leu Asp Gln Ile Pro Ser Tyr Asp Thr His Lys Ile Ala Val Leu
995 1000 1005
Tyr Val Gly Glu Gly Gln Ser Asn Ser Glu Leu Ala Ile Leu Ser
1010 1015 1020
Asn Glu His Gly Ser Tyr Arg Tyr Thr Glu Phe Leu Thr Gly Leu
1025 1030 1035
Gly Arg Leu Ile Glu Leu Lys Asp Cys Gln Pro Asp Lys Val Tyr
1040 1045 1050
Leu Gly Gly Leu Asp Val Cys Gly Glu Asp Gly Gln Phe Thr Tyr
1055 1060 1065
Cys Trp His Asp Asp Ile Met Gln Ala Val Phe His Ile Ala Thr
1070 1075 1080
Leu Met Pro Thr Lys Asp Val Asp Lys His Arg Cys Asp Lys Lys
1085 1090 1095
Arg His Leu Gly Asn Asp Phe Val Ser Ile Val Tyr Asn Asp Ser
1100 1105 1110
Gly Glu Asp Phe Lys Leu Gly Thr Ile Lys Gly Gln Phe Asn Phe
1115 1120 1125
Val His Val Ile Val Thr Pro Leu Asp Tyr Glu Cys Asn Leu Val
1130 1135 1140
Ser Leu Gln Cys Arg Lys Asp Met Glu Gly Leu Val Asp Thr Ser
1145 1150 1155
Val Ala Lys Ile Val Ser Asp Arg Asn Leu Pro Phe Val Ala Arg

1160 1165 1170
Gln Met Ala Leu His Ala Asn Met Ala Ser Gln Val His His Ser
1175 1180 1185
Arg Ser Asn Pro Thr Asp Ile Tyr Pro Ser Lys Trp Ile Ala Arg
1190 1195 1200
Leu Arg His Ile Lys Arg Leu Arg Gln Arg Ile Cys Glu Glu Ala
1205 1210 1215
Ala Tyr Ser Asn Pro Ser Leu Pro Leu Val His Pro Pro Ser His
1220 1225 1230
Ser Lys Ala Pro Ala Gln Thr Pro Ala Glu Pro Thr Pro Gly Tyr
1235 1240 1245
Glu Val Gly Gln Arg Lys Arg Leu Ile Ser Ser Val Glu Asp Phe
1250 1255 1260
Thr Glu Phe Val
1265

<210> 18
<211> 1200
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 18

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

130 135 140
Ala Leu Thr Asp Asn Gly Arg His Ile Thr Tyr Leu Glu Glu Glu Leu
145 150 155 160
Ala Asp Phe Val Leu Gln Trp Met Asp Val Gly Leu Ser Ser Glu Phe
165 170 175
Leu Leu Val Leu Val Asn Leu Val Lys Phe Asn Ser Cys Tyr Leu Asp
180 185 190
Glu Tyr Ile Ala Arg Met Val Gln Met Ile Cys Leu Leu Cys Val Arg
195 200 205
Thr Ala Ser Ser Val Asp Ile Glu Val Ser Leu Gln Val Leu Asp Ala
210 215 220
Val Val Cys Tyr Asn Cys Leu Pro Ala Glu Ser Leu Pro Leu Phe Ile
225 230 235 240
Val Thr Leu Cys Arg Thr Ile Asn Val Lys Glu Leu Cys Glu Pro Cys
245 250 255
Trp Lys Leu Met Arg Asn Leu Leu Gly Thr His Leu Gly His Ser Ala
260 265 270
Ile Tyr Asn Met Cys His Leu Met Glu Asp Arg Ala Tyr Met Glu Asp
275 280 285
Ala Pro Leu Leu Arg Gly Ala Val Phe Phe Val Gly Met Ala Leu Trp
290 295 300
Gly Ala His Arg Leu Tyr Ser Leu Arg Asn Ser Pro Thr Ser Val Leu
305 310 315 320
Pro Ser Phe Tyr Gln Ala Met Ala Cys Pro Asn Glu Val Val Ser Tyr
325 330 335
Glu Ile Val Leu Ser Ile Thr Arg Leu Ile Lys Lys Tyr Arg Lys Glu
340 345 350
Leu Gln Val Val Ala Trp Asp Ile Leu Leu Asn Ile Ile Glu Arg Leu
355 360 365
Leu Gln Gln Leu Gln Thr Leu Asp Ser Pro Glu Leu Arg Thr Ile Val
370 375 380
His Asp Leu Leu Thr Thr Val Glu Glu Leu Cys Asp Gln Asn Glu Phe
385 390 395 400
His Gly Ser Gln Glu Arg Tyr Phe Glu Leu Val Glu Arg Cys Ala Asp
405 410 415
Gln Arg Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
420 425 430
Gly Gly Phe Arg Ala Arg Ser Thr Ser Leu Asn Glu Arg Pro Lys Ser
435 440 445

Ser Arg Ile Gln Thr Ser Leu Thr Ser Ala Ser Leu Gly Ser Ala Asp
450 455 460
Glu Asn Ser Val Ala Gln Ala Asp Asp Ser Leu Lys Asn Leu His Leu
465 470 475 480
Glu Leu Thr Glu Thr Cys Leu Asp Met Met Ala Arg Tyr Val Phe Ser
485 490 495
Asn Phe Thr Ala Val Pro Lys Arg Ser Pro Val Gly Glu Phe Leu Leu
500 505 510
Ala Gly Gly Arg Thr Lys Thr Trp Leu Val Gly Asn Lys Leu Val Thr
515 520 525
Val Thr Thr Ser Val Gly Thr Gly Thr Arg Ser Leu Leu Gly Leu Asp
530 535 540
Ser Gly Glu Leu Gln Ser Gly Pro Glu Ser Ser Ser Ser Pro Gly Val
545 550 555 560
His Val Arg Gln Thr Lys Glu Ala Pro Ala Lys Leu Glu Ser Gln Ala
565 570 575
Gly Gln Gln Val Ser Arg Gly Ala Arg Asp Arg Val Arg Ser Met Ser
580 585 590
Gly Gly His Gly Leu Arg Val Gly Ala Leu Asp Val Pro Ala Ser Gln
595 600 605
Phe Leu Gly Ser Ala Thr Ser Pro Gly Pro Arg Thr Ala Pro Ala Ala
610 615 620
Lys Pro Glu Lys Ala Ser Ala Gly Thr Arg Val Pro Val Gln Glu Lys
625 630 635 640
Thr Asn Leu Ala Ala Tyr Val Pro Leu Leu Thr Gln Gly Trp Ala Glu
645 650 655
Ile Leu Val Arg Arg Pro Thr Gly Val Ala Ser Phe Ser Ser Leu Tyr
660 665 670
Gln Ser Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp Ala Asp
675 680 685
Ser Ala Val Val Met Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu
690 695 700
Val Glu Pro Pro Gly Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp
705 710 715 720
Arg Arg Thr Asp Ala Tyr Ser Arg Ser Ser Ser Val Ser Ser Gln Glu
725 730 735
Glu Lys Ser Leu His Ala Glu Glu Leu Val Gly Arg Gly Ile Pro Ile
740 745 750
Glu Arg Val Val Ser Ser Glu Gly Gly Arg Pro Ser Val Asp Leu Ser

755 760 765

Phe Gln Pro Ser Gln Pro Leu Ser Lys Ser Ser Ser Ser Pro Glu Leu

770 775 780

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

785 790 795 800

Gly Arg Leu Ser Pro Glu Val Lys Ala Arg Ser Gln Ser Gly Thr Leu

805 810 815

Asp Gly Glu Ser Ala Ala Trp Ser Ala Ser Gly Glu Asp Ser Arg Gly

820 825 830

Gln Pro Glu Gly Pro Leu Pro Ser Ser Ser Pro Arg Ser Pro Ser Gly

835 840 845

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

850 855 860

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

865 870 875 880

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

885 890 895

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

900 905 910

Leu Pro Asn Glu Ser Gln Ser Phe Glu Arg Ser Val Gln Leu Leu Asp

915 920 925

Gln Ile Pro Ser Tyr Asp Thr His Lys Ile Ala Val Leu Tyr Val Gly

930 935 940

Glu Gly Gln Ser Asn Ser Glu Leu Ala Ile Leu Ser Asn Glu His Gly

945 950 955 960

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

965 970 975

Leu Lys Asp Cys Gln Pro Asp Lys Val Tyr Leu Gly Gly Leu Asp Val

980 985 990

Cys Gly Glu Asp Gly Gln Phe Thr Tyr Cys Trp His Asp Asp Ile Met

995 1000 1005

Gln Ala Val Phe His Ile Ala Thr Leu Met Pro Thr Lys Asp Val

1010 1015 1020

Asp Lys His Arg Cys Asp Lys Lys Arg His Leu Gly Asn Asp Phe

1025 1030 1035

Val Ser Ile Val Tyr Asn Asp Ser Gly Glu Asp Phe Lys Leu Gly

1040 1045 1050

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

1055 1060 1065

Leu Asp Tyr Glu Cys Asn Leu Val Ser Leu Gln Cys Arg Lys Asp
1070 1075 1080
Met Glu Gly Leu Val Asp Thr Ser Val Ala Lys Ile Val Ser Asp
1085 1090 1095
Arg Asn Leu Pro Phe Val Ala Arg Gln Met Ala Leu His Ala Asn
1100 1105 1110
Met Ala Ser Gln Val His His Ser Arg Ser Asn Pro Thr Asp Ile
1115 1120 1125
Tyr Pro Ser Lys Trp Ile Ala Arg Leu Arg His Ile Lys Arg Leu
1130 1135 1140
Arg Gln Arg Ile Cys Glu Glu Ala Ala Tyr Ser Asn Pro Ser Leu
1145 1150 1155
Pro Leu Val His Pro Pro Ser His Ser Lys Ala Pro Ala Gln Thr
1160 1165 1170
Pro Ala Glu Pro Thr Pro Gly Tyr Glu Val Gly Gln Arg Lys Arg
1175 1180 1185
Leu Ile Ser Ser Val Glu Asp Phe Thr Glu Phe Val
1190 1195 1200

<210> 19
<211> 1038
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 19

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

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

420 425 430

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

435 440 445

Ser Arg Ile Gln Thr Ser Leu Thr Ser Ala Ser Leu Gly Ser Ala Asp

450 455 460

Glu Asn Ser Val Ala Gln Ala Asp Asp Ser Leu Lys Asn Leu His Leu

465 470 475 480

Glu Leu Thr Glu Thr Cys Leu Asp Met Met Ala Arg Tyr Val Phe Ser

485 490 495

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

500 505 510

Ala Gly Gly Arg Thr Lys Thr Trp Leu Val Gly Asn Lys Leu Val Thr

515 520 525

Val Thr Thr Ser Val Gly Thr Gly Thr Arg Ser Leu Leu Gly Leu Asp

530 535 540

Ser Gly Glu Leu Gln Ser Gly Pro Glu Ser Ser Ser Ser Pro Gly Val

545 550 555 560

His Val Arg Gln Thr Lys Glu Ala Pro Ala Lys Leu Glu Ser Gln Ala

565 570 575

Gly Gln Gln Val Ser Arg Gly Ala Arg Asp Arg Val Arg Ser Met Ser

580 585 590

Gly Gly His Gly Leu Arg Val Gly Ala Leu Asp Val Pro Ala Ser Gln

595 600 605

Phe Leu Gly Ser Ala Thr Ser Pro Gly Pro Arg Thr Ala Pro Ala Ala

610 615 620

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

625 630 635 640

Thr Asn Leu Ala Ala Tyr Val Pro Leu Leu Thr Gln Gly Trp Ala Glu

645 650 655

Ile Leu Val Arg Arg Pro Thr Gly Val Ala Ser Phe Ser Ser Leu Tyr

660 665 670

Gln Ser Ser Cys Gln Gly Gln Leu His Arg Ser Val Ser Trp Ala Asp

675 680 685

Ser Ala Val Val Met Glu Glu Gly Ser Pro Gly Glu Val Pro Val Leu

690 695 700

Val Glu Pro Pro Gly Leu Glu Asp Val Glu Ala Ala Leu Gly Met Asp

705 710 715 720

Arg Arg Thr Asp Ala Tyr Ser Arg Ser Phe Val Phe Leu Gln Leu Tyr

725 730 735

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

<210> 20

<211> 2306
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 20

gctagcacta gtaccatggc gaaaccgacc agcaaagata gcggcctgaa agaaaaattt 60
aaaattctgc tgggcctggg caccccgcgc ccgaacccgc gcagcgcgga aggcaaacag 120
accgaattta ttattaccgc ggaaattctg cgcgaactga gcatggaatg cggcctgaac 180
aaccgcattc gcatgattgg ccagatttgc gaagtggcga aaaccaaaaa atttgaagaa 240
catgcggtgg aagcgctgtg gaaagcggtg gcggatctgc tgcagccgga acgcccgctg 300
gaagcgcgcc atgcggtgct ggcgctgctg aaagcgattg tgcagggcca gggcgaacgc 360
ctgggcgtgc tgcgcgcgct gtttttaaa gtgattaaag attatccgag caacgaagat 420
ctgcatgaac gcctggaagt gtttaaagcg ctgaccgata acggccgcca tattacctat 480
ctggaagaag aactggcgga ttttgtgctg cagtggatgg atgtgggcct gagcagcgaa 540
tttctgctgg tgctggtgaa cctggtgaaa tttaacagct gctatctgga tgaatatatt 600
gcgcgcatgg tgcagatgat ttgcctgctg tgcgtgcgca ccgcgagcag cgtggatatt 660
gaagtgagcc tgcaggtgct ggatgcggtg gtgtgctata actgcctgcc ggcggaaagc 720
ctgccgctgt ttattgtgac cctgtgccgc accattaacg tgaaagaact gtgcgaaccg 780
tgctggaaac tgatgcgcaa cctgctgggc acccatctgg gccatagcgc gatttataac 840
atgtgccatc tgatggaaga tcgcgcgtat atggaagatg cgccgctgct gcgcggcgcg 900
gtgtttttg tgggcatggc gctgtggggc gcgcatcgcc tgtatagcct gcgcaacagc 960
ccgaccagcg tgctgccgag cttttatcag gcgatggcgt gcccgaacga agtggtgagc 1020
tatgaaattg tgctgagcat tacccgcctg attaaaaaat atcgcaaaga actgcaggtg 1080
gtggcgtggg atattctgct gaacattatt gaacgcctgc tgcagcagct gcagaccctg 1140
gatagcccgg aactgcgcac cattgtgcat gatctgctga ccaccgtgga agaactgtgc 1200
gatcagaacg aatttcatgg cagccaggaa cgctattttg aactggtgga acgctgcgcg 1260
gatcagcgcc cggaaagcag cctgctgaac ctgattagct atcgcgcgca gagcattcat 1320
ccggcgaaag atggctggat tcagaacctg caggcgctga tggaatctgg tgggggtagc 1380
ggaggagggt cagggggcgg cagtggaggc ggaaaaccga ttctgctgcc gaacgaaagc 1440
cagagctttg aacgcagcgt gcagctgctg gatcagattc cgagctatga tacccataaa 1500
attgcggtgc tgtatgtggg cgaaggccag agcaacagcg aactggcgat tctgagcaac 1560
gaacatggca gctatcgcta taccgaattt ctgaccggcc tgggccgcct gattgaactg 1620
aaagattgcc agccggataa agtgtatctg ggcggcctgg atgtgtgcgg cgaagatggc 1680
cagtttacct attgctggca tgatgatatt atgcaggcgg tgtttcatat tgcgaccctg 1740
atgccgacca aagatgtgga taaacatcgc tgcgataaaa aacgccatct gggcaacgat 1800
tttgtgagca ttgtgtataa cgatagcggc gaagatttta aactgggcac cattaaaggc 1860
cagtttaact ttgtgcatgt gattgtgacc ccgctggatt atgaatgcaa cctggtgagc 1920
ctgcagtgcc gcaaagatat ggaaggcctg gtggatacca gcgtggcgaa aattgtgagc 1980

gatcgcaacc tgccgtttgt ggcgcgccag atggcgctgc atgcgaacat ggcgagccag 2040
gtgcatcata gccgcagcaa cccgaccgat atttatccga gcaaatggat tgcgcgcctg 2100
cgccatatta aacgcctgcg ccagcgcatt tgcgaagaag cggcgtatag caacccgagc 2160
ctgccgctgg tgcatccgcc gagccatagc aaagcgccgg cgcagacccc ggcggaaccg 2220
accccgggct atgaagtggg ccagcgcaaa cgcctgatta gcagcgtgga agattttacc 2280
gaatttgtgt aggcggccgc ctcgag 2306

<210> 21
<211> 3402
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 21

atggcaaaac ctacatcaaa agatagcgga ctgaaagaga agttcaaaat cctgctggga 60
ctggggacac cacgccctaa tccacggtcc gccgagggca agcagaccga gttcatcatc 120
acagccgaga tcctgcgcga gctgtctatg gagtgcggcc tgaacaatcg gatcagaatg 180
atcggccaga tttgcgaggt ggccaagacc aagaagtttg aggagcacgc agtggaggcc 240
ctgtggaagg cagtggccga tctgctgcag cctgagagac cactggaggc aaggcacgcc 300
gtgctggccc tgctgaaggc catcgtgcag ggacagggag agcgcctggg cgtgctgcgg 360
gccctgttct ttaaagtgat caaggactac cctagcaacg aggatctgca cgagagactg 420
gaggtgttca aggccctgac cgataatggc aggcacatca catatctgga ggaggagctg 480
gccgactttg tgctgcagtg gatggatgtg ggcctgagct ccgagttcct gctggtgctg 540
gtgaacctgg tgaagtttaa ttcctgctac ctggacgagt atatcgcccg catggtgcag 600
atgatctgcc tgctgtgcgt gcggaccgcc tctagcgtgg acatcgaggt gtctctgcag 660
gtgctggatg ccgtggtgtg ctacaactgt ctgcccgccg agagcctgcc tctgttcatc 720
gtgaccctgt gcagaacaat caatgtgaag gagctgtgcg agccttgttg gaagctgatg 780
aggaacctgc tgggcacaca cctgggacac agcgccatct acaatatgtg ccacctgatg 840
gaggaccgcg cctatatgga ggatgcacca ctgctgaggg gagccgtgtt ctttgtggga 900
atggcactgt ggggagcaca cagactgtac tccctgagga actctccaac cagcgtgctg 960
ccctcttttt accaggctat ggcctgtcca aatgaggtgg tgtcttatga gatcgtgctg 1020
agcatcacac gcctgatcaa gaagtatcgg aaggagctgc aggtggtggc ctgggacatc 1080
ctgctgaaca tcatcgagcg cctgctgcag cagctgcaga ccctggacag cccagagctg 1140
aggacaatcg tgcacgatct gctgaccaca gtggaggagc tgtgcgacca gaatgagttc 1200
cacggctccc aggagcggta ctttgagctg gtggagcggt gcgcagatca gaggccagag 1260
tcctctctgc tgaacctgat ctcctatcgg gcccagtcta tccaccctgc caaggacggc 1320
tggattcaga atctgcaggc cctgatggag agcggaggag gctccggagg aggctctgga 1380
ggcggcagcg gcggcggcgg cgccctggac gtgccggcct cccagttcct gggcagtgcc 1440
acttctccag gaccacggac tgcaccagcc gcgaaacctg agaaggcctc agctggcacc 1500
cgggttcctg tgcaggagaa gacgaacctg gcggcctatg tgcccctgct gacccagggc 1560

tgggcggaga tcctggtccg gaggcccaca gggaacacca gctggctgat gagcctggag 1620
aacccgctca gccctttctc ctcggacatc aacaacatgc ccctgcagga gctgtctaac 1680
gccctcatgg cggctgagcg cttcaaggag caccgggaca cagccctgta caagtcactg 1740
tcggtgccgg cagccagcac ggccaaaccc cctcctctgc ctcgctccaa cacagtggcc 1800
tctttctcct ccctgtacca gtccagctgc caaggacagc tgcacaggag cgtttcctgg 1860
gcagactccg ccgtggtcat ggaggaggga agtccgggcg aggttcctgt gctggtggag 1920
cccccagggt tggaggacgt tgaggcagcg ctaggcatgg acaggcgcac ggatgcctac 1980
agcaggtcgt cctcagtctc cagccaggag gagaagtcgc tccacgcgga ggagctggtt 2040
ggcaggggca tccccatcga gcgagtcgtc tcctcggagg gtggccggcc ctctgtggac 2100
ctctccttcc agccctcgca gcccctgagc aagtccagct cctctcccga gctgcagact 2160
ctgcaggaca tcctcgggga ccctggggac aaggccgacg tgggccggct gagccctgag 2220
gttaaggccc ggtcacagtc agggaccctg gacggggaaa gtgctgcctg gtcggcctcg 2280
ggcgaagaca gtcggggcca gcccgagggt cccttgcctt ccagctcccc ccgctcgccc 2340
agtggcctcc ggccccgagg ttacaccatc tccgactcgg ccccatcacg caggggcaag 2400
agagtagaga gggacgcctt aaagagcaga gccacagcct ccaatgcaga gaaagtgcca 2460
ggcatcaacc ccagtttcgt gttcctgcag ctctaccatt ccccccttctt tggcgacgag 2520
tcaaacaagc caatcctgct gcccaatgag tcacagtcct ttgagcggtc ggtgcagctc 2580
ctcgaccaga tcccatcata cgacacccac aagatcgccg tcctgtatgt tggagaaggc 2640
cagagcaaca gcgagctcgc catcctgtcc aatgagcatg gctcctacag gtacacggag 2700
ttcctgacgg gcctgggccg gctcatcgag ctgaaggact gccagccgga caaggtgtac 2760
ctgggaggcc tggacgtgtg tggtgaggac ggccagttca cctactgctg gcacgatgac 2820
atcatgcaag ccgtcttcca catcgccacc ctgatgccca ccaaggacgt ggacaagcac 2880
cgctgcgaca agaagcgcca cctgggcaac gactttgtgt ccattgtcta caatgactcc 2940
ggtgaggact tcaagcttgg caccatcaag ggccagttca actttgtcca cgtgatcgtc 3000
accccgctgg actacgagtg caacctggtg tccctgcagt gcaggaaaga catggagggc 3060
cttgtggaca ccagcgtggc caagatcgtg tctgaccgca acctgccctt cgtggcccgc 3120
cagatggccc tgcacgcaaa tatggcctca caggtgcatc atagccgctc caaccccacc 3180
gatatctacc cctccaagtg gattgcccgg ctccgccaca tcaagcggct ccgccagcgg 3240
atctgcgagg aagccgccta ctccaacccc agcctacctc tggtgcaccc tccgtcccat 3300
agcaaagccc ctgcacagac tccagccgag cccacacctg gctatgaggt gggccagcgg 3360
aagcgcctca tctcctcggt ggaggacttc accgagtttg tg 3402

<210> 22
<211> 3306
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 22
atggccaaac caacaagcaa agattcaggc ttgaaggaga agtttaagat tctgttggga 60

ctgggaacac cgaggccaaa tcccaggtct gcagagggta aacagacgga gtttatcatc 120
accgcggaaa tactgagaga actgagcatg gaatgtggcc tcaacaatcg catccggatg 180
atagggcaga tttgtgaagt cgcaaaaacc aagaaatttg aagagcacgc agtggaagca 240
ctctggaagg cggtcgcgga tctgttgcag ccggagcggc cgctggaggc ccggcacgcg 300
gtgctggctc tgctgaaggc catcgtgcag gggcagggcg agcgtttggg ggtcctcaga 360
gccctcttct ttaaggtcat caaggattac ccttccaacg aagaccttca cgaaaggctg 420
gaggttttca aggccctcac agacaatggg agacacatca cctacttgga ggaagagctg 480
gctgactttg tcctgcagtg gatggatgtt ggcttgtcct cggaattcct tctggtgctg 540
gtgaacttgg tcaaattcaa tagctgttac ctcgacgagt acatcgcaag gatggttcag 600
atgatctgtc tgctgtgcgt ccggaccgcg tcctctgtgg acatagaggt ctccctgcag 660
gtgctggacg ccgtggtctg ctacaactgc ctgccggctg agagcctccc gctgttcatc 720
gttaccctct gtcgcaccat caacgtcaag gagctctgcg agccttgctg gaagctgatg 780
cggaacctcc ttggcaccca cctgggccac agcgccatct acaacatgtg ccacctcatg 840
gaggacagag cctacatgga ggacgcgccc ctgctgagag gagccgtgtt ttttgtgggc 900
atggctctct ggggagccca ccggctctat tctctcagga actcgccgac atctgtgttg 960
ccatcatttt accaggccat ggcatgtccg aacgaggtgg tgtcctatga gatcgtcctg 1020
tccatcacca ggctcatcaa gaagtatagg aaggagctcc aggtggtggc gtgggacatt 1080
ctgctgaaca tcatcgaacg gctccttcag cagctccaga ccttggacag cccggagctc 1140
aggaccatcg tccatgacct gttgaccacg gtggaggagc tgtgtgacca gaacgagttc 1200
cacgggtctc aggagagata ctttgaactg gtggagagat gtgcggacca gaggagcgga 1260
ggaggctccg gaggaggctc tggaggcggc agcggcggcg gcggcgccct ggacgtgccg 1320
gcctcccagt tcctgggcag tgccacttct ccaggaccac ggactgcacc agccgcgaaa 1380
cctgagaagg cctcagctgg cacccgggtt cctgtgcagg agaagacgaa cctggcggcc 1440
tatgtgcccc tgctgaccca gggctgggcg gagatcctgg tccggaggcc cacagggaac 1500
accagctggc tgatgagcct ggagaacccg ctcagccctt tctcctcgga catcaacaac 1560
atgcccctgc aggagctgtc taacgccctc atggcggctg agcgcttcaa ggagcaccgg 1620
gacacagccc tgtacaagtc actgtcggtg ccggcagcca gcacggccaa acccccctcct 1680
ctgcctcgct ccaacacagt ggcctctttc tcctccctgt accagtccag ctgccaagga 1740
cagctgcaca ggagcgtttc ctgggcagac tccgccgtgg tcatggagga gggaagtccg 1800
ggcgaggttc ctgtgctggt ggagccccca gggttggagg acgttgaggc agcgctaggc 1860
atggacaggc gcacggatgc ctacagcagg tcgtcctcag tctccagcca ggaggagaag 1920
tcgctccacg cggaggagct ggttggcagg ggcatcccca tcgagcgagt cgtctcctcg 1980
gagggtggcc ggccctctgt ggacctctcc ttccagccct cgcagcccct gagcaagtcc 2040
agctcctctc ccgagctgca gactctgcag gacatcctcg gggaccctgg ggacaaggcc 2100
gacgtgggcc ggctgagccc tgaggttaag gcccggtcac agtcagggac cctggacggg 2160
gaaagtgctg cctggtcggc ctcgggcgaa gacagtcggg gccagcccga gggtcccttg 2220
ccttccagct ccccccgctc gcccagtggc ctccggcccc gaggttacac catctccgac 2280
tcggccccat cacgcagggg caagagagta gagagggacg ccttaaagag cagagccaca 2340
gcctccaatg cagagaaagt gccaggcatc aaccccagtt tcgtgttcct gcagctctac 2400

cattcccccct tctttggcga cgagtcaaac aagccaatcc tgctgcccaa tgagtcacag 2460
tcctttgagc ggtcggtgca gctcctcgac cagatcccat catacgacac ccacaagatc 2520
gccgtcctgt atgttggaga aggccagagc aacagcgagc tcgccatcct gtccaatgag 2580
catggctcct acaggtacac ggagttcctg acgggcctgg gccggctcat cgagctgaag 2640
gactgccagc cggacaaggt gtacctggga ggcctggacg tgtgtggtga ggacggccag 2700
ttcacctact gctggcacga tgacatcatg caagccgtct tccacatcgc caccctgatg 2760
cccaccaagg acgtggacaa gcaccgctgc gacaagaagc gccacctggg caacgacttt 2820
gtgtccattg tctacaatga ctccggtgag gacttcaagc ttggcaccat caagggccag 2880
ttcaactttg tccacgtgat cgtcaccccg ctggactacg agtgcaacct ggtgtccctg 2940
cagtgcagga aagacatgga gggccttgtg gacaccagcg tggccaagat cgtgtctgac 3000
cgcaacctgc ccttcgtggc ccgccagatg gccctgcacg caaatatggc ctcacaggtg 3060
catcatagcc gctccaaccc caccgatatc taccctcca agtggattgc ccggctccgc 3120
cacatcaagc ggctccgcca gcggatctgc gaggaagccg cctactccaa ccccagccta 3180
cctctggtgc accctccgtc ccatagcaaa gcccctgcac agactccagc cgagcccaca 3240
cctggctatg aggtgggcca gcggaagcgc ctcatctcct cggtggagga cttcaccgag 3300
tttgtg 3306

<210> 23
<211> 3927
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 23

atggccaaac caacaagcaa agattcaggc ttgaaggaga agtttaagat tctgttggga 60
ctgggaacac cgaggccaaa tcccaggtct gcagagggta aacagacgga gtttatcatc 120
accgcggaaa tactgagaga actgagcatg gaatgtggcc tcaacaatcg catccggatg 180
atagggcaga tttgtgaagt cgcaaaaacc aagaaatttg aagagcacgc agtggaagca 240
ctctggaagg cggtcgcgga tctgttgcag ccggagcggc cgctggaggc ccggcacgcg 300
gtgctggctc tgctgaaggc catcgtgcag gggcagggcg agcgtttggg ggtcctcaga 360
gccctcttct ttaaggtcat caaggattac ccttccaacg aagaccttca cgaaaggctg 420
gaggttttca aggccctcac agacaatggg agacacatca cctacttgga ggaagagctg 480
gctgactttg tcctgcagtg gatggatgtt ggcttgtcct cggaattcct tctggtgctg 540
gtgaacttgg tcaaattcaa tagctgttac ctcgacgagt acatcgcaag gatggttcag 600
atgatctgtc tgctgtgcgt ccggaccgcg tcctctgtgg acatagaggt ctccctgcag 660
gtgctggacg ccgtggtctg ctacaactgc ctgccggctg agagcctccc gctgttcatc 720
gttaccctct gtcgcaccat caacgtcaag gagctctgcg agccttgctg gaagctgatg 780
cggaacctcc ttggcaccca cctgggccac agcgccatct acaacatgtg ccacctcatg 840
gaggacagag cctacatgga ggacgcgccc ctgctgagag gagccgtgtt ttttgtgggc 900
atggctctct ggggagccca ccggctctat tctctcagga actcgccgac atctgtgttg 960

ccatcatttt accaggccat ggcatgtccg aacgaggtgg tgtcctatga gatcgtcctg 1020
tccatcacca ggctcatcaa gaagtatagg aaggagctcc aggtggtggc gtgggacatt 1080
ctgctgaaca tcatcgaacg gctccttcag cagctccaga ccttggacag cccggagctc 1140
aggaccatcg tccatgacct gttgaccacg gtggaggagc tgtgtgacca gaacgagttc 1200
cacgggtctc aggagagata ctttgaactg gtggagagat gtgcggacca gaggagcgga 1260
ggaggctccg gaggaggctc tggaggcggc agcggcggcg gcttcagggc ccggagtact 1320
agtctcaacg agagacccaa gagtctgagg atagccagac cccccaaaca aggcttgaat 1380
aactctccac ccgtgaaaga attcaaggag agctctgcag ccgaggcctt ccggtgccgc 1440
agcatcagtg tgtctgaaca tgtggtccgc agcaggatac agacgtccct caccagtgcc 1500
agcttggggt ctgcagatga gaactccgtg gcccaggctg acgatagcct gaaaaacctc 1560
cacctggagc tcacggaaac ctgtctggac atgatggctc gatacgtctt ctccaacttc 1620
acggctgtcc cgaagaggtc tcctgtgggc gagttcctcc tagcgggtgg caggaccaaa 1680
acctggctgg ttgggaacaa gcttgtcact gtgacgacaa gcgtgggaac cgggacccgg 1740
tcgttactag gcctggactc gggggagctg cagtccggcc cggagtcgag ctccagcccc 1800
ggggtgcatg tgagacagac caaggaggcg ccggccaagc tggagtccca ggctgggcag 1860
caggtgtccc gtggggcccg ggatcgggtc cgttccatgt cgggggggca tggtcttcga 1920
gttggcgccc tggacgtgcc ggcctcccag ttcctgggca gtgccacttc tccaggacca 1980
cggactgcac cagccgcgaa acctgagaag gcctcagctg gcacccgggt tcctgtgcag 2040
gagaagacga acctggcggc ctatgtgccc ctgctgaccc agggctgggc ggagatcctg 2100
gtccggaggc ccacagggaa caccagctgg ctgatgagcc tggagaaccc gctcagccct 2160
ttctcctcgg acatcaacaa catgcccctg caggagctgt ctaacgccct catggcggct 2220
gagcgcttca aggagcaccg ggacacagcc ctgtacaagt cactgtcggt gccggcagcc 2280
agcacggcca aacccccctcc tctgcctcgc tccaacacag tggcctcttt ctcctccctg 2340
taccagtcca gctgccaagg acagctgcac aggagcgttt cctgggcaga ctccgccgtg 2400
gtcatggagg agggaagtcc gggcgaggtt cctgtgctgg tggagccccc agggttggag 2460
gacgttgagg cagcgctagg catggacagg cgcacggatg cctacagcag gtcgtcctca 2520
gtctccagcc aggaggagaa gtcgctccac gcggaggagc tggttggcag gggcatcccc 2580
atcgagcgag tcgtctcctc ggagggtggc cggccctctg tggacctctc cttccagccc 2640
tcgcagcccc tgagcaagtc cagctcctct cccgagctgc agactctgca ggacatcctc 2700
ggggaccctg gggacaaggc cgacgtgggc cggctgagcc ctgaggttaa ggcccggtca 2760
cagtcaggga ccctggacgg ggaaagtgct gcctggtcgg cctcgggcga agacagtcgg 2820
ggccagcccg agggtccctt gccttccagc tccccccgct cgcccagtgg cctccggccc 2880
cgaggttaca ccatctccga ctcggcccca tcacgcaggg gcaagagagt agagagggac 2940
gccttaaaga gcagagccac agcctccaat gcagagaaag tgccaggcat caaccccagt 3000
ttcgtgttcc tgcagctcta ccattccccc ttctttggcg acgagtcaaa caagccaatc 3060
ctgctgccca atgagtcaca gtcctttgag cggtcggtgc agctcctcga ccagatccca 3120
tcatacgaca cccacaagat cgccgtcctg tatgttggag aaggccagag caacagcgag 3180
ctcgccatcc tgtccaatga gcatggctcc tacaggtaca cggagttcct gacgggcctg 3240
ggccggctca tcgagctgaa ggactgccag ccggacaagg tgtacctggg aggcctggac 3300

gtgtgtggtg aggacggcca gttcacctac tgctggcacg atgacatcat gcaagccgtc 3360
ttccacatcg ccaccctgat gcccaccaag gacgtggaca agcaccgctg cgacaagaag 3420
cgccacctgg gcaacgactt tgtgtccatt gtctacaatg actccggtga ggacttcaag 3480
cttggcacca tcaagggcca gttcaacttt gtccacgtga tcgtcacccc gctggactac 3540
gagtgcaacc tggtgtccct gcagtgcagg aaagacatgg agggccttgt ggacaccagc 3600
gtggccaaga tcgtgtctga ccgcaacctg cccttcgtgg cccgccagat ggccctgcac 3660
gcaaatatgg cctcacaggt gcatcatagc cgctccaacc ccaccgatat ctacccctcc 3720
aagtggattg cccggctccg ccacatcaag cggctccgcc agcggatctg cgaggaagcc 3780
gcctactcca accccagcct acctctggtg caccctccgt cccatagcaa agcccctgca 3840
cagactccag ccgagcccac acctggctat gaggtgggcc agcggaagcg cctcatctcc 3900
tcggtggagg acttcaccga gtttgtg 3927

<210> 24
<211> 3801
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 24

atggccaaac caacaagcaa agattcaggc ttgaaggaga agtttaagat tctgttggga 60
ctgggaacac cgaggccaaa tcccaggtct gcagagggta aacagacgga gtttatcatc 120
accgcggaaa tactgagaga actgagcatg gaatgtggcc tcaacaatcg catccggatg 180
atagggcaga tttgtgaagt cgcaaaaacc aagaaatttg aagagcacgc agtggaagca 240
ctctggaagg cggtcgcgga tctgttgcag ccggagcggc cgctggaggc ccggcacgcg 300
gtgctggctc tgctgaaggc catcgtgcag gggcagggcg agcgtttggg ggtcctcaga 360
gccctcttct ttaaggtcat caaggattac ccttccaacg aagaccttca cgaaaggctg 420
gaggttttca aggccctcac agacaatggg agacacatca cctacttgga ggaagagctg 480
gctgactttg tcctgcagtg gatggatgtt ggcttgtcct cggaattcct tctggtgctg 540
gtgaacttgg tcaaattcaa tagctgttac ctcgacgagt acatcgcaag gatggttcag 600
atgatctgtc tgctgtgcgt ccggaccgcg tcctctgtgg acatagaggt ctccctgcag 660
gtgctggacg ccgtggtctg ctacaactgc ctgccggctg agagcctccc gctgttcatc 720
gttaccctct gtcgcaccat caacgtcaag gagctctgcg agccttgctg gaagctgatg 780
cggaacctcc ttggcaccca cctgggccac agcgccatct acaacatgtg ccacctcatg 840
gaggacagag cctacatgga ggacgcgccc ctgctgagag gagccgtgtt ttttgtgggc 900
atggctctct ggggagccca ccggctctat tctctcagga actcgccgac atctgtgttg 960
ccatcatttt accaggccat ggcatgtccg aacgaggtgg tgtcctatga gatcgtcctg 1020
tccatcacca ggctcatcaa gaagtatagg aaggagctcc aggtggtggc gtgggacatt 1080
ctgctgaaca tcatcgaacg gctccttcag cagctccaga ccttggacag cccggagctc 1140
aggaccatcg tccatgacct gttgaccacg gtggaggagc tgtgtgacca gaacgagttc 1200
cacgggtctc aggagagata ctttgaactg gtggagagat gtgcggacca gaggagcgga 1260

ggaggctccg gaggaggctc tggaggcggc agcggcggcg gcttcagggc ccggagtact 1320
agtctcaacg agagacccaa gagtagcagg atacagacgt ccctcaccag tgccagcttg 1380
gggtctgcag atgagaactc cgtggcccag gctgacgata gcctgaaaaa cctccacctg 1440
gagctcacgg aaacctgtct ggacatgatg gctcgatacg tcttctccaa cttcacggct 1500
gtcccgaaga ggtctcctgt gggcgagttc ctcctagcgg gtggcaggac caaaacctgg 1560
ctggttggga acaagcttgt cactgtgacg acaagcgtgg gaaccgggac ccggtcgtta 1620
ctaggcctgg actcggggga gctgcagtcc ggcccggagt cgagctccag ccccggggtg 1680
catgtgagac agaccaagga ggcgccggcc aagctggagt cccaggctgg gcagcaggtg 1740
tcccgtgggg cccgggatcg ggtccgttcc atgtcggggg gccatggtct tcgagttggc 1800
gccctggacg tgccggcctc ccagttcctg ggcagtgcca cttctccagg accacggact 1860
gcaccagccg cgaaacctga gaaggcctca gctggcaccc gggttcctgt gcaggagaag 1920
acgaacctgg cggcctatgt gcccctgctg acccagggct gggcggagat cctggtccgg 1980
aggcccacag ggaacaccag ctggctgatg agcctggaga acccgctcag ccctttctcc 2040
tcggacatca acaacatgcc cctgcaggag ctgtctaacg ccctcatggc ggctgagcgc 2100
ttcaaggagc accgggacac agccctgtac aagtcactgt cggtgccggc agccagcacg 2160
gccaaacccc ctcctctgcc tcgctccaac acagtggcct ctttctcctc cctgtaccag 2220
tccagctgcc aaggacagct gcacaggagc gtttcctggg cagactccgc cgtggtcatg 2280
gaggagggaa gtccgggcga ggttcctgtg ctggtggagc ccccagggtt ggaggacgtt 2340
gaggcagcgc taggcatgga caggcgcacg gatgcctaca gcaggtcgtc ctcagtctcc 2400
agccaggagg agaagtcgct ccacgcggag gagctggttg gcaggggcat ccccatcgag 2460
cgagtcgtct cctcggaggg tggccggccc tctgtggacc tctccttcca gccctcgcag 2520
cccctgagca agtccagctc ctctcccgag ctgcagactc tgcaggacat cctcggggac 2580
cctggggaca aggccgacgt gggccggctg agccctgagg ttaaggcccg gtcacagtca 2640
gggaccctgg acggggaaag tgctgcctgg tcggcctcgg gcgaagacag tcggggccag 2700
cccgagggtc ccttgccttc cagctccccc cgctcgccca gtggcctccg gccccgaggt 2760
tacaccatct ccgactcggc cccatcacgc aggggcaaga gagtagagag ggacgcctta 2820
aagagcagag ccacagcctc caatgcagag aaagtgccag gcatcaaccc cagtttcgtg 2880
ttcctgcagc tctaccattc ccccttcttt ggcgacgagt caaacaagcc aatcctgctg 2940
cccaatgagt cacagtcctt tgagcggtcg gtgcagctcc tcgaccagat cccatcatac 3000
gacacccaca agatcgccgt cctgtatgtt ggagaaggcc agagcaacag cgagctcgcc 3060
atcctgtcca atgagcatgg ctcctacagg tacacggagt tcctgacggg cctgggccgg 3120
ctcatcgagc tgaaggactg ccagccggac aaggtgtacc tgggaggcct ggacgtgtgt 3180
ggtgaggacg gccagttcac ctactgctgg cacgatgaca tcatgcaagc cgtcttccac 3240
atcgccaccc tgatgcccac caaggacgtg gacaagcacc gctgcgacaa gaagcgccac 3300
ctgggcaacg actttgtgtc cattgtctac aatgactccg gtgaggactt caagcttggc 3360
accatcaagg gccagttcaa ctttgtccac gtgatcgtca ccccgctgga ctacgagtgc 3420
aacctggtgt ccctgcagtg caggaaagac atggagggcc ttgtggacac cagcgtggcc 3480
aagatcgtgt ctgaccgcaa cctgcccttc gtggcccgcc agatggccct gcacgcaaat 3540
atggcctcac aggtgcatca tagccgctcc aaccccaccg atatctaccc ctccaagtgg 3600

attgcccggc tccgccacat caagcggctc cgccagcgga tctgcgagga agccgcctac 3660
tccaacccca gcctacctct ggtgcaccct ccgtcccata gcaaagcccc tgcacagact 3720
ccagccgagc ccacacctgg ctatgaggtg ggccagcgga agcgcctcat ctcctcggtg 3780
gaggacttca ccgagtttgt g 3801

<210> 25
<211> 3600
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 25

atggccaaac caacaagcaa agattcaggc ttgaaggaga agtttaagat tctgttggga 60
ctgggaacac cgaggccaaa tcccaggtct gcagagggta aacagacgga gtttatcatc 120
accgcggaaa tactgagaga actgagcatg gaatgtggcc tcaacaatcg catccggatg 180
atagggcaga tttgtgaagt cgcaaaaacc aagaaatttg aagagcacgc agtggaagca 240
ctctggaagg cggtcgcgga tctgttgcag ccggagcggc cgctggaggc ccggcacgcg 300
gtgctggctc tgctgaaggc catcgtgcag gggcagggcg agcgtttggg ggtcctcaga 360
gccctcttct ttaaggtcat caaggattac ccttccaacg aagaccttca cgaaaggctg 420
gaggttttca aggccctcac agacaatggg agacacatca cctacttgga ggaagagctg 480
gctgactttg tcctgcagtg gatggatgtt ggcttgtcct cggaattcct tctggtgctg 540
gtgaacttgg tcaaattcaa tagctgttac ctcgacgagt acatcgcaag gatggttcag 600
atgatctgtc tgctgtgcgt ccggaccgcg tcctctgtgg acatagaggt ctccctgcag 660
gtgctggacg ccgtggtctg ctacaactgc ctgccggctg agagcctccc gctgttcatc 720
gttaccctct gtcgcaccat caacgtcaag gagctctgcg agccttgctg gaagctgatg 780
cggaacctcc ttggcaccca cctgggccac agcgccatct acaacatgtg ccacctcatg 840
gaggacagag cctacatgga ggacgcgccc ctgctgagag gagccgtgtt ttttgtgggc 900
atggctctct ggggagccca ccggctctat tctctcagga actcgccgac atctgtgttg 960
ccatcatttt accaggccat ggcatgtccg aacgaggtgg tgtcctatga gatcgtcctg 1020
tccatcacca ggctcatcaa gaagtatagg aaggagctcc aggtggtggc gtgggacatt 1080
ctgctgaaca tcatcgaacg gctccttcag cagctccaga ccttggacag cccggagctc 1140
aggaccatcg tccatgacct gttgaccacg gtggaggagc tgtgtgacca gaacgagttc 1200
cacgggtctc aggagagata ctttgaactg gtggagagat gtgcggacca gaggagcgga 1260
ggaggctccg gaggaggctc tggaggcggc agcggcggcg gcttcagggc ccggagtact 1320
agtctcaacg agagacccaa gagtagcagg atacagacgt ccctcaccag tgccagcttg 1380
gggtctgcag atgagaactc cgtggcccag gctgacgata gcctgaaaaa cctccacctg 1440
gagctcacgg aaacctgtct ggacatgatg gctcgatacg tcttctccaa cttcacggct 1500
gtcccgaaga ggtctcctgt gggcgagttc ctcctagcgg gtggcaggac caaaacctgg 1560
ctggttggga acaagcttgt cactgtgacg acaagcgtgg gaaccgggac ccggtcgtta 1620
ctaggcctgg actcggggga gctgcagtcc ggcccggagt cgagctccag ccccggggtg 1680

catgtgagac agaccaagga ggcgccggcc aagctggagt cccaggctgg gcagcaggtg 1740
tcccgtgggg cccgggatcg ggtccgttcc atgtcggggg gccatggtct tcgagttggc 1800
gccctggacg tgccggcctc ccagttcctg ggcagtgcca cttctccagg accacggact 1860
gcaccagccg cgaaacctga gaaggcctca gctggcaccc gggttcctgt gcaggagaag 1920
acgaacctgg cggcctatgt gcccctgctg acccagggct gggcggagat cctggtccgg 1980
aggcccacag gggtggcctc tttctcctcc ctgtaccagt ccagctgcca aggacagctg 2040
cacaggagcg tttcctgggc agactccgcc gtggtcatgg aggagggaag tccgggcgag 2100
gttcctgtgc tggtggagcc cccagggttg gaggacgttg aggcagcgct aggcatggac 2160
aggcgcacgg atgcctacag caggtcgtcc tcagtctcca gccaggagga gaagtcgctc 2220
cacgcggagg agctggttgg caggggcatc cccatcgagc gagtcgtctc ctcggagggt 2280
ggccggccct ctgtggacct ctccttccag ccctcgcagc ccctgagcaa gtccagctcc 2340
tctcccgagc tgcagactct gcaggacatc ctcggggacc ctggggacaa ggccgacgtg 2400
ggccggctga gccctgaggt taaggcccgg tcacagtcag ggaccctgga cggggaaagt 2460
gctgcctggt cggcctcggg cgaagacagt cggggccagc ccgagggtcc cttgccttcc 2520
agctcccccc gctcgcccag tggcctccgg ccccgaggtt acaccatctc cgactcggcc 2580
ccatcacgca ggggcaagag agtagagagg gacgccttaa agagcagagc cacagcctcc 2640
aatgcagaga aagtgccagg catcaacccc agtttcgtgt tcctgcagct ctaccattcc 2700
cccttctttg gcgacgagtc aaacaagcca atcctgctgc ccaatgagtc acagtccttt 2760
gagcggtcgg tgcagctcct cgaccagatc ccatcatacg acacccacaa gatcgccgtc 2820
ctgtatgttg gagaaggcca gagcaacagc gagctcgcca tcctgtccaa tgagcatggc 2880
tcctacaggt acacggagtt cctgacgggc ctgggccggc tcatcgagct gaaggactgc 2940
cagccggaca aggtgtacct gggaggcctg gacgtgtgtg gtgaggacgg ccagttcacc 3000
tactgctggc acgatgacat catgcaagcc gtcttccaca tcgccaccct gatgcccacc 3060
aaggacgtgg acaagcaccg ctgcgacaag aagcgccacc tgggcaacga ctttgtgtcc 3120
attgtctaca atgactccgg tgaggacttc aagcttggca ccatcaaggg ccagttcaac 3180
tttgtccacg tgatcgtcac cccgctggac tacgagtgca acctggtgtc cctgcagtgc 3240
aggaaagaca tggagggcct tgtggacacc agcgtggcca agatcgtgtc tgaccgcaac 3300
ctgcccttcg tggcccgcca gatggccctg cacgcaaata tggcctcaca ggtgcatcat 3360
agccgctcca accccaccga tatctacccc tccaagtgga ttgcccggct ccgccacatc 3420
aagcggctcc gccagcggat ctgcgaggaa gccgcctact ccaaccccag cctacctctg 3480
gtgcaccctc cgtcccatag caaagcccct gcacagactc cagccgagcc cacacctggc 3540
tatgaggtgg gccagcggaa gcgcctcatc tcctcggtgg aggacttcac cgagtttgtg 3600

<210> 26
<211> 3114
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 26

atggccaaac caacaagcaa agattcaggc ttgaaggaga agtttaagat tctgttggga 60
ctgggaacac cgaggccaaa tcccaggtct gcagagggta aacagacgga gtttatcatc 120
accgcggaaa tactgagaga actgagcatg gaatgtggcc tcaacaatcg catccggatg 180
atagggcaga tttgtgaagt cgcaaaaacc aagaaatttg aagagcacgc agtggaagca 240
ctctggaagg cggtcgcgga tctgttgcag ccggagcggc cgctggaggc ccggcacgcg 300
gtgctggctc tgctgaaggc catcgtgcag gggcagggcg agcgtttggg ggtcctcaga 360
gccctcttct ttaaggtcat caaggattac ccttccaacg aagaccttca cgaaaggctg 420
gaggttttca aggccctcac agacaatggg agacacatca cctacttgga ggaagagctg 480
gctgactttg tcctgcagtg gatggatgtt ggcttgtcct cggaattcct tctggtgctg 540
gtgaacttgg tcaaattcaa tagctgttac ctcgacgagt acatcgcaag gatggttcag 600
atgatctgtc tgctgtgcgt ccggaccgcg tcctctgtgg acatagaggt ctccctgcag 660
gtgctggacg ccgtggtctg ctacaactgc ctgccggctg agagcctccc gctgttcatc 720
gttaccctct gtcgcaccat caacgtcaag gagctctgcg agccttgctg gaagctgatg 780
cggaacctcc ttggcaccca cctgggccac agcgccatct acaacatgtg ccacctcatg 840
gaggacagag cctacatgga ggacgcgccc ctgctgagag gagccgtgtt ttttgtgggc 900
atggctctct ggggagccca ccggctctat tctctcagga actcgccgac atctgtgttg 960
ccatcatttt accaggccat ggcatgtccg aacgaggtgg tgtcctatga gatcgtcctg 1020
tccatcacca ggctcatcaa gaagtatagg aaggagctcc aggtggtggc gtgggacatt 1080
ctgctgaaca tcatcgaacg gctccttcag cagctccaga ccttggacag cccggagctc 1140
aggaccatcg tccatgacct gttgaccacg gtggaggagc tgtgtgacca gaacgagttc 1200
cacgggtctc aggagagata ctttgaactg gtggagagat gtgcggacca gaggagcgga 1260
ggaggctccg gaggaggctc tggaggcggc agcggcggcg gcttcagggc ccggagtact 1320
agtctcaacg agagacccaa gagtagcagg atacagacgt ccctcaccag tgccagcttg 1380
gggtctgcag atgagaactc cgtggcccag gctgacgata gcctgaaaaa cctccacctg 1440
gagctcacgg aaacctgtct ggacatgatg gctcgatacg tcttctccaa cttcacggct 1500
gtcccgaaga ggtctcctgt gggcgagttc ctcctagcgg gtggcaggac caaaacctgg 1560
ctggttggga acaagcttgt cactgtgacg acaagcgtgg gaaccgggac ccggtcgtta 1620
ctaggcctgg actcgggga gctgcagtcc ggcccggagt cgagctccag ccccggggtg 1680
catgtgagac agaccaagga ggcgccggcc aagctggagt cccaggctgg gcagcaggtg 1740
tcccgtgggg cccgggatcg ggtccgttcc atgtcggggg gccatggtct tcgagttggc 1800
gccctggacg tgccggcctc ccagttcctg ggcagtgcca cttctccagg accacggact 1860
gcaccagccg cgaaacctga gaaggcctca gctggcaccc gggttcctgt gcaggagaag 1920
acgaacctgg cggcctatgt gcccctgctg acccagggct gggcggagat cctggtccgg 1980
aggcccacag gggtggcctc tttctcctcc ctgtaccagt ccagctgcca aggacagctg 2040
cacaggagcg tttcctgggc agactccgcc gtggtcatgg aggagggaag tccgggcgag 2100
gttcctgtgc tggtggagcc cccagggttg gaggacgttg aggcagcgct aggcatggac 2160
aggcgcacgg atgcctacag caggagtttc gtgttcctgc agctctacca ttccccccttc 2220
tttggcgacg agtcaaacaa gccaatcctg ctgcccaatg agtcacagtc ctttgagcgg 2280
tcggtgcagc tcctcgacca gatcccatca tacgacaccc acaagatcgc cgtcctgtat 2340

gttggagaag gccagagcaa cagcgagctc gccatcctgt ccaatgagca tggctcctac 2400
aggtacacgg agttcctgac gggcctgggc cggctcatcg agctgaagga ctgccagccg 2460
gacaaggtgt acctgggagg cctggacgtg tgtggtgagg acggccagtt cacctactgc 2520
tggcacgatg acatcatgca agccgtcttc cacatcgcca ccctgatgcc caccaaggac 2580
gtggacaagc accgctgcga caagaagcgc cacctgggca acgactttgt gtccattgtc 2640
tacaatgact ccggtgagga cttcaagctt ggcaccatca agggccagtt caactttgtc 2700
cacgtgatcg tcaccccgct ggactacgag tgcaacctgg tgtccctgca gtgcaggaaa 2760
gacatggagg gccttgtgga caccagcgtg gccaagatcg tgtctgaccg caacctgccc 2820
ttcgtggccc gccagatggc cctgcacgca aatatggcct cacaggtgca tcatagccgc 2880
tccaacccca ccgatatcta cccctccaag tggattgccc ggctccgcca catcaagcgg 2940
ctccgccagc ggatctgcga ggaagccgcc tactccaacc ccagcctacc tctggtgcac 3000
cctccgtccc atagcaaagc ccctgcacag actccagccg agcccacacc tggctatgag 3060
gtgggccagc ggaagcgcct catctcctcg gtggaggact tcaccgagtt tgtg 3114

<210> 27
<211> 130
<212> DNA
<213> 未知(Unknown)
<220>
<223> 5’ ITR
<400> 27
ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg cccgggcttt 60
gcccgggcgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggttcct 130

<210> 28
<211> 130
<212> DNA
<213> 未知(Unknown)
<220>
<223> 3’ ITR
<400> 28
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60
ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120
gagcgcgcag 130

<210> 29
<211> 42
<212> PRT
<213> 智人(Homo sapiens)
<400> 29
Leu Arg Ile Ala Arg Pro Pro Lys Gln Gly Leu Asn Asn Ser Pro Pro

1 5 10 15
Val Lys Glu Phe Lys Glu Ser Ser Ala Ala Glu Ala Phe Arg Cys Arg
20 25 30
Ser Ile Ser Val Ser Glu His Val Val Arg
35 40
<210> 30
<211> 67
<212> PRT
<213> 智人(Homo sapiens)
<400> 30
Asn Thr Ser Trp Leu Met Ser Leu Glu Asn Pro Leu Ser Pro Phe Ser
1 5 10 15
Ser Asp Ile Asn Asn Met Pro Leu Gln Glu Leu Ser Asn Ala Leu Met
20 25 30
Ala Ala Glu Arg Phe Lys Glu His Arg Asp Thr Ala Leu Tyr Lys Ser
35 40 45
Leu Ser Val Pro Ala Ala Ser Thr Ala Lys Pro Pro Pro Leu Pro Arg
50 55 60
Ser Asn Thr
65
<210> 31
<211> 162
<212> PRT
<213> 智人(Homo sapiens)
<400> 31
Ser Ser Ser Val Ser Ser Gln Glu Glu Lys Ser Leu His Ala Glu Glu
1 5 10 15
Leu Val Gly Arg Gly Ile Pro Ile Glu Arg Val Val Ser Ser Glu Gly
20 25 30
Gly Arg Pro Ser Val Asp Leu Ser Phe Gln Pro Ser Gln Pro Leu Ser
35 40 45
Lys Ser Ser Ser Ser Pro Glu Leu Gln Thr Leu Gln Asp Ile Leu Gly
50 55 60
Asp Pro Gly Asp Lys Ala Asp Val Gly Arg Leu Ser Pro Glu Val Lys
65 70 75 80
Ala Arg Ser Gln Ser Gly Thr Leu Asp Gly Glu Ser Ala Ala Trp Ser
85 90 95
Ala Ser Gly Glu Asp Ser Arg Gly Gln Pro Glu Gly Pro Leu Pro Ser
100 105 110

Ser Ser Pro Arg Ser Pro Ser Gly Leu Arg Pro Arg Gly Tyr Thr Ile
115 120 125
Ser Asp Ser Ala Pro Ser Arg Arg Gly Lys Arg Val Glu Arg Asp Ala
130 135 140
Leu Lys Ser Arg Ala Thr Ala Ser Asn Ala Glu Lys Val Pro Gly Ile
145 150 155 160
Asn Pro

<210> 32
<211> 177
<212> PRT
<213> 智人(Homo sapiens)
<400> 32

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

<210> 33
<211> 382
<212> DNA
<213> 未知(Unknown)

<220>
<223> CMV-IE URE
<400> 33
ctagtcgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc 60
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac 120
cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa 180
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag 240
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc 300
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct 360
acgtattagt catcgctatt ac 382
<210> 34
<211> 260
<212> DNA
<213> 未知(Unknown)
<220>
<223> CB6启动子
<400> 34
ccacgttctg cttcactctc cccatctccc ccccctcccc acccccaatt ttgtatttat 60
ttatttttta attattttgt gcagcgatgg gggcgggggg ggggggcgcg cgccaggcgg 120
ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg cagccaatca 180
gagcggcgcg ctccgaaagt ttccttttat ggcgaggcgg cggcggcggc ggccctataa 240
aaagcgaagc gcgcggcggg 260
<210> 35
<211> 6
<212> DNA
<213> 未知(Unknown)
<220>
<223> Kozak序列
<400> 35
gccacc 6
<210> 36
<211> 56
<212> DNA
<213> 未知(Unknown)
<220>
<223> 兔polyA序列
<400> 36
aataaaggaa atttattttc attgcaatag tgtgttggaa ttttttgtgt ctctca 56