(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115551545 A
(43)申请公布日 2022.12.30

(21)申请号 202180013506.4

(22)申请日 2021.02.06

(30)优先权数据
62/971,825 2020.02.07 US
63/016,175 2020.04.27 US
63/044,330 2020.06.25 US
63/063,137 2020.08.07 US

(85)PCT国际申请进入国家阶段日
2022.08.08

(86)PCT国际申请的申请数据
PCT/US2021/016979 2021.02.06

(87)PCT国际申请的公布数据
WO2021/159040 EN 2021.08.12

(71)申请人 摩登纳特斯有限公司
地址 美国马萨诸塞州

(72)发明人 纪尧姆·斯图尔特-琼斯
萨伊达·马哈古卜·埃尔巴希尔
安德烈亚·卡尔菲
米希尔·迈特卡

(74)专利代理机构 北京集佳知识产权代理有限公司 11227
专利代理师 张福誉 韩晓帆

(51)Int.Cl.
A61K 39/12(2006.01)
A61K 39/215(2006.01)
A61P 11/00(2006.01)
A61P 31/14(2006.01)

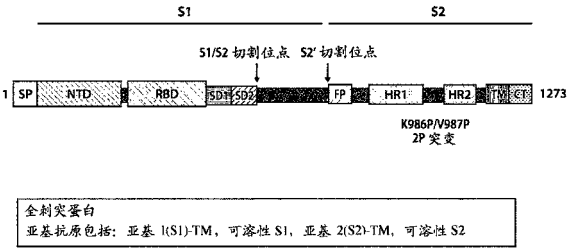
权利要求书9页 说明书181页
序列表239页 附图11页

(54)发明名称

SARS-COV-2 mRNA结构域疫苗

(57)摘要

本公开涉及冠状病毒核糖核酸(RNA)疫苗以及使用所述疫苗的方法和包含所述疫苗的组合物。所述RNA疫苗编码冠状病毒的结构域和亚基。



全刺突蛋白
亚基抗原包括：亚基 1(S1)-TM，可溶性 S1，亚基 2(S2)-TM，可溶性 S2

CN 115551545 A

1.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码SARS-CoV-2刺突蛋白的至少两个结构域并且小于全长刺突蛋白的开放阅读框(ORF)。

2.如权利要求1所述的mRNA,其中所述两个结构域之一是SARS-CoV-2刺突蛋白的N末端结构域(NTD)。

3.如权利要求1或2所述的mRNA,其中所述两个结构域之一是SARS-CoV-2刺突蛋白的受体结合结构域(RBD)。

4.如权利要求2或3所述的mRNA,其中所述ORF编码与所述NTD和/或所述RBD连接的跨膜结构域(TD)。

5.如权利要求4所述的mRNA,其中所述TD是流感血凝素跨膜结构域。

6.如权利要求3或4所述的mRNA,其中所述ORF包含NTD-RBD-TM。

7.如权利要求1-6中任一项所述的mRNA,其中所述至少两个结构域经由可切割或不可切割的接头连接。

8.如权利要求7所述的mRNA,其中所述不可切割的接头是甘氨酸-丝氨酸(GS)接头。

9.如权利要求8所述的mRNA,其中所述GS接头包含4-15个氨基酸。

10.如权利要求6所述的mRNA,其中所述接头是泛HLA DR结合表位(PADRE)。

11.如权利要求1-10中任一项所述的mRNA,其中所述ORF编码信号肽。

12.如权利要求11所述的mRNA,其中所述信号肽与所述NTD连接。

13.如权利要求11所述的mRNA,其中所述信号肽与所述RBD连接。

14.如权利要求11-13中任一项所述的mRNA,其中所述信号肽对SARS-CoV-2是异源的。

15.如权利要求1-3中任一项所述的mRNA,其中所述至少两个结构域是可溶的。

16.如权利要求1-14中任一项所述的mRNA,其中所述ORF编码运输信号结构域。

17.如权利要求16所述的mRNA,其中所述运输信号结构域是巨噬细胞标志物。

18.如权利要求16所述的mRNA,其中所述巨噬细胞标志物是CD86和/或CD11b。

19.如权利要求16所述的mRNA,其中所述运输信号结构域是VSV-G胞质尾(VSVGct)。

20.如权利要求1所述的mRNA,其中所述两个结构域之一是SARS-CoV-2刺突蛋白的第一重复七肽:HPPHCPC(HR1)。

21.如权利要求1所述的mRNA,其中所述两个结构域之一是SARS-CoV-2刺突蛋白的第二重复七肽:HPPHCPC(HR2)。

22.如权利要求20或21所述的mRNA,其中所述ORF编码与所述HR1和/或所述HR2连接的跨膜结构域(TD)。

23.如权利要求22所述的mRNA,其中所述TD是流感血凝素跨膜结构域。

24.如权利要求20-23中任一项所述的mRNA,其中所述ORF编码融合肽(FP)。

25.如权利要求20-23中任一项所述的mRNA,其中所述ORF编码CT尾。

26.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码SARS-CoV-2刺突蛋白的受体结合结构域(RBD)的开放阅读框(ORF)。

27.如权利要求26所述的mRNA,其中所述RBD是可溶的。

28.如权利要求26所述的mRNA,其中所述RBD与跨膜结构域、任选地流感血凝素跨膜结构域连接。

29.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码SARS-CoV-2刺突蛋

白的N末端结构域(NTD)的开放阅读框(ORF)。

30.如权利要求29所述的mRNA,其中所述NTD与SARS-CoV-2刺突蛋白的RBD连接以形成NTD-RBD融合蛋白。

31.如权利要求30所述的mRNA,其中所述NTD-RBD融合体与跨膜结构域(TM)、任选地流感血凝素跨膜结构域连接,以形成NTD-RBD-TM蛋白。

32.如权利要求30所述的mRNA,其中所述NTD-RBD融合体包含C末端截短。

33.如权利要求26-32中任一项所述的mRNA,其中所述NTD和/或所述RBD包括延伸区。

34.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码S1亚基的开放阅读框(ORF),其中所述S1亚基与跨膜结构域、任选地流感血凝素跨膜结构域连接。

35.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码S1亚基的开放阅读框(ORF),其中所述S1亚基经修饰以去除S蛋白的RBD或RBD的一部分。

36.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码与S2亚基连接的S1亚基的开放阅读框(ORF),其中所述S1亚基来自HKU1 S蛋白。

37.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码与S2亚基连接的S1亚基的开放阅读框(ORF),其中所述S1亚基来自OC43 S蛋白。

38.如权利要求34-37中任一项所述的mRNA,其中还包含运输信号,任选地选自巨噬细胞标志物,任选地CD86、CD11B和/或VSVGct。

39.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码SARS-CoV-2刺突蛋白的S2亚基的开放阅读框(ORF)。

40.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)和蛋白质跨膜结构域。

41.如权利要求40所述的mRNA,其中所述蛋白质跨膜结构域是流感血凝素跨膜结构域。

42.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的氨基(N)末端结构域(NTD)和跨膜结构域。

43.如权利要求42所述的mRNA,其中所述跨膜结构域是流感血凝素跨膜结构域。

44.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码融合蛋白的开放阅读框,所述融合蛋白包含与SARS-CoV-2刺突蛋白的受体结合结构域连接的SARS-CoV-2刺突蛋白的氨基(N)末端结构域。

45.如权利要求44所述的mRNA,其中所述融合蛋白还包含跨膜结构域。

45.如前述权利要求中任一项所述的mRNA,所述mRNA还包含5′帽,任选地7mG(5′)ppp(5′)NlmpNp。

46.如权利要求1-45中任一项所述的mRNA,所述mRNA还包含多聚A尾,任选地具有约100个核苷酸的长度。

47.如权利要求1-46中任一项所述的mRNA,其中所述mRNA包含化学修饰,任选地1-甲基假尿苷。

48.如权利要求1-47中任一项所述的mRNA,其中所述mRNA中的每个尿苷是1-甲基假尿苷。

49.一种组合物,所述组合物包含:

(a)信使核糖核酸(mRNA),其包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)和蛋白质跨膜结构域;和

(b)mRNA,其包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的氨基(N)末端结构域(NTD)和跨膜结构域。

50.如权利要求49所述的组合物,其中所述蛋白质跨膜结构域是流感血凝素跨膜结构域。

51.如权利要求50所述的组合物,其中(a)的所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少80%同一性的氨基酸序列。

52.如权利要求51所述的组合物,其中(a)的所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

53.如权利要求52所述的组合物,其中(a)的所述融合蛋白包含SEQ ID NO:77的氨基酸序列。

54.如权利要求50-53中任一项所述的组合物,其中(a)的所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70%同一性的核苷酸序列。

55.如权利要求54所述的组合物,其中(a)的所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

56.如权利要求55所述的组合物,其中(a)的所述开放阅读框包含SEQ ID NO:76的核苷酸序列。

57.如权利要求50-56中任一项所述的组合物,其中(b)的所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少80%同一性的氨基酸序列。

58.如权利要求57所述的组合物,其中(b)的所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

59.如权利要求58所述的组合物,其中(b)的所述融合蛋白包含SEQ ID NO:47的氨基酸序列。

60.如权利要求50-59中任一项所述的组合物,其中(b)的所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70%同一性的核苷酸序列。

61.如权利要求60所述的组合物,其中(b)的所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

62.如权利要求61所述的组合物,其中(b)的所述开放阅读框包含SEQ ID NO:46的核苷酸序列。

63.如权利要求49-62中任一项所述的组合物,其中(a)的所述mRNA与(b)的所述mRNA的比率为约1:1。

64.如权利要求1-48中任一项所述的mRNA,所述mRNA在还包含脂质纳米颗粒的组合物中。

65.如权利要求49-62中任一项所述的组合物,所述组合物还包含脂质纳米颗粒。

66.如权利要求64或65所述的组合物,其中所述mRNA在所述脂质纳米颗粒中。

67.如权利要求49-62中任一项所述的组合物,其中(a)的所述mRNA在脂质纳米颗粒中并且(b)的所述mRNA在脂质纳米颗粒中。

68.如权利要求67所述的组合物,其中(a)和(b)的所述mRNA在相同脂质纳米颗粒中,或者其中(a)和(b)的所述mRNA相对于彼此各自配制在单独纳米颗粒中。

69.如权利要求64-68中任一项所述的组合物,其中所述脂质纳米颗粒包含阳离子脂质。

70.如权利要求69所述的组合物,其中所述脂质纳米颗粒还包含中性脂质。

71.如权利要求69或70所述的组合物,其中所述脂质纳米颗粒还包含固醇。

72.如权利要求69-71中任一项所述的组合物,其中所述脂质纳米颗粒还包含聚乙二醇(PEG)修饰的脂质。

73.如权利要求69-72中任一项所述的组合物,其中所述脂质纳米颗粒包含可电离的阳离子脂质、中性脂质、固醇和PEG修饰的脂质。

74.如权利要求73所述的组合物,其中所述可电离的阳离子脂质是8((2羟基乙基)(6氧代6-(十一烷基氧基)己基)氨基)辛酸十七烷-9-基酯(化合物1)。

75.如权利要求73或74所述的组合物,其中所述中性脂质是1,2二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)。

76.如权利要求69-75中任一项所述的组合物,其中所述固醇是胆固醇。

77.如权利要求69-76中任一项所述的组合物,其中所述PEG修饰的脂质是1,2二肉豆蔻酰基-sn-甘油,甲氧基聚乙二醇(PEG2000 DMG)。

78.如权利要求69-77中任一项所述的组合物,其中所述脂质纳米颗粒包含20-60mol%可电离的阳离子脂质、5-25mol%中性脂质、25-55mol%固醇和0.5-15mol%PEG修饰的脂质。

79.如权利要求78所述的组合物,其中所述脂质纳米颗粒包含:

40-50mol%可电离的脂质,任选地45-50mol%、30-45mol%固醇,任选地35-40mol%、5-15mol%辅助脂质,任选地10-12mol%、1-5%PEG脂质,任选地1-3mol%或1.5-2.5mol%。

80.如权利要求79所述的组合物,其中所述脂质纳米颗粒包含:

40-50mol%化合物1,任选地45-50mol%、30-45mol%胆固醇,任选地35-40mol%、5-15mol%DSPC,任选地10-12mol%、1-5%PEG2000DMG,任选地1-3mol%或1.5-2.5mol%。

81.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应的量的前述权利要求中任一项所述的mRNA或组合物。

82.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的T细胞免疫反应的量的前述权利要求中任一项所述的mRNA或组合物。

83.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应和T细胞免疫反应的量的权利要求1-48中任一项所述的mRNA。

84.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码SARS-CoV-2刺突蛋白的至少一个结构域的开放阅读框(ORF)。

85.如权利要求84所述的mRNA,其中所述mRNA是权利要求1-48中的任一者。

86.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:77的氨基酸序列具

有至少80%同一性的氨基酸序列。

87.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:77的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

88.如权利要求84或85所述的mRNA,其中所述ORF编码SEQ ID NO:77的氨基酸序列。

89.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:76的核苷酸序列具有至少70%同一性的核苷酸序列。

90.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:76的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

91.如权利要求84或85所述的mRNA,其中所述ORF包含SEQ ID NO:76的核苷酸序列。

92.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:47的氨基酸序列具有至少80%同一性的氨基酸序列。

93.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:47的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

94.如权利要求84或85所述的mRNA,其中所述ORF编码SEQ ID NO:47的氨基酸序列。

95.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:46的核苷酸序列具有至少70%同一性的核苷酸序列。

96.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:46的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

97.如权利要求84或85所述的mRNA,其中所述ORF包含SEQ ID NO:46的核苷酸序列。

98.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:62的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:62的氨基酸序列。

99.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:61的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:61的核苷酸序列。

100.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:92、140或143的氨基酸序列具有至少80%同一性的氨基酸序列。

101.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:92、140或143的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

102.如权利要求84或85所述的mRNA,其中所述ORF编码SEQ ID NO:92、140或143的氨基酸序列。

103.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:91、139或142的核苷酸序列具有至少70%同一性的核苷酸序列。

104.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:91、139或142的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

105.如权利要求84或85所述的mRNA,其中所述ORF包含SEQ ID NO:91、139或142的核苷酸序列。

106.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:107的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:107的氨基酸序列。

107.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:106的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:106的核苷酸序列。

108.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:59、86、89、116、119或122中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:59、86、89、116、119或122中任一者的氨基酸序列。

109.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:58、85、88、115、118或121中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:58、85、88、115、118或121中任一者的核苷酸序列。

110.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码S1亚基抗原的开放阅读框(ORF),所述S1亚基抗原包含与SEQ ID NO:5的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:5的氨基酸序列。

111.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:3的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:3的核苷酸序列。

112.如权利要求84或85所述的mRNA,其中所述抗原包含与SEQ ID NO:17或146的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:17或146的氨基酸序列。

113.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:16或145的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:16或145的核苷酸序列。

114.如权利要求84或85所述的mRNA,其中所述ORF编码与SEQ ID NO:20、23、26、29、32或35中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:20、23、26、29、32或35中任一者的氨基酸序列。

115.如权利要求84或85所述的mRNA,其中所述ORF包含与SEQ ID NO:19、22、25、28、41

或34中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:19、22、25、28、31或34中任一者的核苷酸序列。

116.如权利要求84或85所述的mRNA，其中所述ORF编码与SEQ ID NO:38的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:38的氨基酸序列。

117.如权利要求84或85所述的mRNA，其中所述ORF包含与SEQ ID NO:37的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:37的核苷酸序列。

118.如权利要求84或85所述的mRNA，其中所述ORF编码与SEQ ID NO:41的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:41的氨基酸序列。

119.如权利要求84或85所述的mRNA，其中所述ORF包含与SEQ ID NO:40的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:40的核苷酸序列。

120.如前述权利要求中任一项所述的mRNA，所述mRNA还包含5′非翻译区(UTR)，所述5′非翻译区任选地包含SEQ ID NO:131或2的核苷酸序列。

121.如前述权利要求中任一项所述的mRNA，所述mRNA还包含3′非翻译区(UTR)，所述3′非翻译区任选地包含SEQ ID NO:132或4的核苷酸序列。

122.一种信使核糖核酸(mRNA)，所述信使核糖核酸(mRNA)包含编码抗原的开放阅读框(ORF)，所述抗原包含与SEQ ID NO:8或65的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:8或65的氨基酸序列。

123.如权利要求122所述的mRNA，其中所述ORF包含与SEQ ID NO:7或64的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:7或64的核苷酸序列。

124.一种信使核糖核酸(mRNA)，所述信使核糖核酸(mRNA)包含编码抗原的开放阅读框(ORF)，所述抗原包含与SEQ ID NO:11、14、68或71中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:11、14、68或71中任一者的氨基酸序列。

125.如权利要求124所述的mRNA，其中所述开放阅读框包含与SEQ ID NO:10、13、67或70中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:10、13、67或70中任一者的核苷酸序列。

126.一种信使核糖核酸(mRNA)，所述信使核糖核酸(mRNA)包含编码抗原的开放阅读框(ORF)，所述抗原包含与SEQ ID NO:44、50、74、80、83、101、104或113中任一者的氨基酸序列

具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:44、50、74、80、83、101、104或113中任一者的氨基酸序列。

127.如权利要求126所述的mRNA，其中所述开放阅读框包含与SEQ ID NO:43、49、73、79、82、100、103或112中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:43、49、73、79、82、100、103或112中任一者的核苷酸序列。

128.如权利要求108所述的mRNA，其中所述抗原包含与SEQ ID NO:95、98或110中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:95、98或110中任一者的氨基酸序列。

129.如权利要求127或128所述的mRNA，其中所述开放阅读框包含与SEQ ID NO:94、97或109中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:94、97或109中任一者的核苷酸序列。

130.如权利要求84-129中任一项所述的mRNA，所述mRNA在还包含脂质纳米颗粒的组合物中。

131.如权利要求130所述的组合物，其中所述mRNA在所述脂质纳米颗粒中。

132.如权利要求130-131中任一项所述的组合物，其中所述脂质纳米颗粒包含阳离子脂质。

133.如权利要求132所述的组合物，其中所述脂质纳米颗粒还包含中性脂质。

134.如权利要求130-133中任一项所述的组合物，其中所述脂质纳米颗粒还包含固醇。

135.如权利要求130-134中任一项所述的组合物，其中所述脂质纳米颗粒还包含聚乙二醇(PEG)修饰的脂质。

136.如权利要求130-135中任一项所述的组合物，其中所述脂质纳米颗粒包含可电离的阳离子脂质、中性脂质、固醇和PEG修饰的脂质。

137.如权利要求136所述的组合物，其中所述可电离的阳离子脂质是8((2羟基乙基)(6氧代6-(十一烷基氧基)己基)氨基)辛酸十七烷-9-基酯(化合物1)。

138.如权利要求136或137所述的组合物，其中所述中性脂质是1,2二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)。

139.如权利要求130-138中任一项所述的组合物，其中所述固醇是胆固醇。

140.如权利要求130-139中任一项所述的组合物，其中所述PEG修饰的脂质是1,2二肉豆蔻酰基-sn-甘油，甲氧基聚乙二醇(PEG2000 DMG)。

141.如权利要求130-140中任一项所述的组合物，其中所述脂质纳米颗粒包含20-60mol%可电离的阳离子脂质、5-25mol%中性脂质、25-55mol%固醇和0.5-15mol%PEG修饰的脂质。

142.如权利要求141所述的组合物，其中所述脂质纳米颗粒包含：

40-50mol%可电离的脂质，任选地45-50mol%、30-45mol%固醇，任选地35-40mol%、

5-15mol％辅助脂质，任选地10-12mol％、1-5％PEG脂质，任选地1-3mol％或1.5-2.5mol％。

143.如权利要求142所述的组合物，其中所述脂质纳米颗粒包含：

40-50mol％化合物1，任选地45-50mol％、30-45mol％胆固醇，任选地35-40mol％、5-15mol％DSPC，任选地10-12mol％、1-5％PEG2000DMG，任选地1-3mol％或1.5-2.5mol％。

144.如权利要求84-129中任一项所述的mRNA，其中所述mRNA包含化学修饰，任选地1-甲基假尿苷。

145.如权利要求84-129中任一项所述的mRNA，其中所述mRNA中的每个尿苷是1-甲基假尿苷。

146.如权利要求130-143中任一项所述的组合物，其中所述mRNA包含化学修饰，任选地1-甲基假尿苷。

147.如权利要求130-143中任一项所述的组合物，其中所述mRNA中的每个尿苷是1-甲基假尿苷。

148.一种方法，所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应的量的权利要求84-129中任一项所述的mRNA。

149.一种方法，所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的T细胞免疫反应的量的权利要求84-129中任一项所述的mRNA。

150.一种方法，所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应和T细胞免疫反应的量的权利要求84-129中任一项所述的mRNA。

151.一种方法，所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应的量的权利要求130-143中任一项所述的组合物。

152.一种方法，所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的T细胞免疫反应的量的权利要求130-143中任一项所述的组合物。

153.一种方法，所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应和T细胞免疫反应的量的权利要求130-143中任一项所述的组合物。

SARS-COV-2 mRNA结构域疫苗

[0001] 相关申请

[0002] 本申请根据35 U.S.C.§119(e)要求于2020年2月7日提交的美国临时申请号62/971,825、于2020年4月27日提交的美国临时申请号63/016,175、于2020年6月25日提交的美国临时申请号63/044,330和于2020年8月7日提交的美国临时申请号63/063,137的权益,每个申请都通过引用整体并入本文。

背景技术

[0003] 人类冠状病毒是冠状病毒科(Coronaviridae family)的高度接触传染性包膜正单链RNA病毒。已知冠状病毒科的两个亚科会引起人类疾病。最重要的是β-冠状病毒(betacoronavirus)。β-冠状病毒是轻度至中度上呼吸道感染的常见病原体。然而,新型冠状病毒感染(例如由最初于2019年12月从中国武汉市鉴定的冠状病毒引起的感染)的爆发与高死亡率死亡人数相关。这种最近鉴定的冠状病毒(称为严重急性呼吸综合征冠状病毒2(SARS-CoV-2),以前称为“2019新型冠状病毒”或“2019-nCoV”)已迅速感染数十万人。SARS-CoV-2病毒引起的大流行病已由世界卫生组织(World Health Organization,WHO)命名为COVID-19(冠状病毒疾病(Coronavirus Disease)2019)。SARS-CoV-2分离株(Wuhan-Hu-1)的第一个基因组序列由位于北京的中国CDC的研究人员于2020年1月10日在Virological发表,所述Virological是用于分析和解释病毒分子进化和流行病学的基于UK的讨论论坛。然后将所述序列于2020年1月12日保藏于GenBank,Genbank登录号为MN908947.1。

[0004] 目前,尚无针对COVID-19的特异性治疗或针对SARS-CoV-2感染的疫苗。与冠状病毒感染、特别是SARS-CoV-2大流行相关的持续的健康问题和死亡率在国际上受到极大关注。由SARS-CoV-2引起的公共健康危机强调了快速开发针对这些病毒的有效且安全的疫苗候选者的重要性。

发明内容

[0005] 在一些实施方案中,本文提供了包含一种或多种信使核糖核酸(mRNA)分子的组合物(例如疫苗),所述信使核糖核酸(mRNA)分子编码能够引发针对SARS-CoV-2抗原的强效中和抗体反应的高免疫原性抗原。本文所述的mRNA分子用于表达SARS-CoV-2冠状病毒刺突(S)蛋白的关键中和结构域,当个别地或组合用作免疫原性组合物或疫苗时,其有效诱导保护性免疫,以保护人免于天然病毒感染和/或在感染时减轻症状。

[0006] 已知β冠状病毒的包膜S蛋白决定病毒宿主向性和进入宿主细胞,并且对于SARS-CoV-2感染是至关重要的。S蛋白的组织在β冠状病毒(例如SARS-CoV-2、SARS-CoV、MERS-CoV、HKU1-CoV、MHV-CoV和NL63-CoV(包括两个亚基S1和S2,其分别介导附着和膜融合))中相似。S1亚基包括N末端结构域(NTD)和受体结合结构域(RBD)。

[0007] 亚基抗原的表达将免疫反应集中于特定亚基,对与其他相关病毒共有的抗原的其他结构域具有特异性的记忆B和T细胞的刺激最小。本文提供的数据表明,施用编码膜结合或可溶性SARS-CoV-2 S1亚基抗原的mRNA产生针对SARS-CoV-2 RBD抗原、NTD抗原、野生型

全长S蛋白和具有双脯氨酸突变以稳定融合前构象的S蛋白中的每一者的抗体滴度。如本文所示，在所有测试的剂量下，两剂量方案(即，包括加强剂量)有效诱导可识别并结合至SARS-CoV-2 WT S蛋白的抗体。令人惊讶的是，即使在S1亚基中未发现双脯氨酸突变(双脯氨酸突变发生在S2中，而在测试的免疫原中不存在S2)，当针对双脯氨酸稳定化形式的S蛋白测量时，诱导的滴度最高。

[0008] 另外，已知NTD和RBD两者都是结合中和病毒活性的抗体的位点。在SARS-CoV-2的情形下，RBD是刺突蛋白的受体结合位点，其结合血管收缩肽转化酶2(ACE2)。NTD的功能尚未完全了解，其似乎在结合糖部分和促进刺突蛋白从融合前构象向融合后构象的构象转变中起作用。无论如何，NTD和RBD结构域都诱导如本文所示的高结合抗体和中和抗体滴度。

[0009] 例如，相当令人惊讶的是，本文一些实施方案中提供的数据显示，尽管来自施用编码膜结合RBD抗原(RBD-TM)或膜结合NTD抗原(NTD-TM)的mRNA的血清显示对SARS-CoV-2 S1/S2刺突蛋白具有免疫原性，但两种mRNA(并且因此两种抗原)的50:50组合产生对SARS-CoV-2 S1/S2刺突蛋白出乎意料高的协同中和抗体滴度。

[0010] 因此，本公开的一些方面提供了包含编码SARS-CoV-2 S蛋白的功能结构域的mRNA的组合物，所述功能结构域能够诱导针对SARS-CoV-2的免疫反应，例如中和抗体反应。在一些实施方案中，所述mRNA配制在脂质纳米颗粒中。

[0011] 在一些方面，提供了包含编码SARS-CoV-2刺突蛋白的至少两个结构域并且小于全长刺突蛋白的开放阅读框(ORF)的mRNA。小于全长刺突蛋白的刺突蛋白是具有比全长刺突蛋白少至少一个氨基酸的刺突蛋白的一个或多个结构域和/或亚基，或具有以非天然顺序或序列连接在一起的一个或多个结构域的融合蛋白。在一些实施方案中，两个结构域之一是SARS-CoV-2刺突蛋白的N末端结构域(NTD)。在一些实施方案中，两个结构域之一是SARS-CoV-2刺突蛋白的受体结合结构域(RBD)。在一些实施方案中，所述ORF编码与所述NTD和/或RBD连接的跨膜结构域(TD)。在一些实施方案中，所述TD是流感血凝素跨膜结构域。在一些实施方案中，所述ORF包含NTD-RBD-TM。在一些实施方案中，所述至少两个结构域经由可切割或不可切割的接头连接。在一些实施方案中，所述不可切割的接头是甘氨酸-丝氨酸(GS)接头。在一些实施方案中，所述GS接头包含4-15个氨基酸。在一些实施方案中，所述接头是泛HLA DR结合表位(PADRE)。在一些实施方案中，所述ORF编码信号肽。在一些实施方案中，所述信号肽与所述NTD连接。在一些实施方案中，所述信号肽与所述RBD连接。在一些实施方案中，所述信号肽与SARS-CoV-2异源。在一些实施方案中，所述至少两个结构域是可溶的。在一些实施方案中，所述ORF编码运输信号结构域。在一些实施方案中，所述运输信号结构域是巨噬细胞标志物。在一些实施方案中，所述巨噬细胞标志物是CD86和/或CD11b。在一些实施方案中，所述运输信号结构域是VSV-G胞质尾(VSVGct)。在一些实施方案中，所述两个结构域之一是SARS-CoV-2刺突蛋白的第一重复七肽:HPPHCPC(HR1)。在一些实施方案中，所述两个结构域之一是SARS-CoV-2刺突蛋白的第二重复七肽:HPPHCPC(HR2)。在一些实施方案中，所述ORF编码与HR1和/或HR2连接的跨膜结构域(TD)。在一些实施方案中，所述TD是流感血凝素跨膜结构域。在一些实施方案中，所述ORF编码融合肽(FP)。在一些实施方案中，所述ORF编码CT尾。

[0012] 在一些方面，提供了包含编码SARS-CoV-2刺突蛋白的受体结合结构域(RBD)的开放阅读框(ORF)的mRNA。在一些实施方案中，所述RBD是可溶的。在一些实施方案中，所述RBD

与跨膜结构域、任选地流感血凝素跨膜结构域连接。

[0013] 本发明的一个或多个实施方案的细节在以下描述中阐述。本发明的其他特征或优点将从以下附图和若干实施方案的详细描述以及所附权利要求中显而易见。

附图说明

[0014] 图1由本发明的mRNA编码的野生型和2P刺突蛋白抗原的示意图;信号肽(SP),无填充;N末端结构域(NTD),有点的;受体结合结构域(RBD),向下对角条带;子结构域1(SD1),水平条带;子结构域2(SD2),波形;融合肽(FP),向上对角条带;七肽重复区1(HR1),编织形;七肽重复区2(HR2),对角砖形;(TM),垂直条带;和胞质尾区(CT),砖形。

[0015] 图2显示了由实施例1-3中所述的mRNA编码的抗原的示例性线性设计。

[0016] 图3显示了图2中绘示的抗原的序列比对。

[0017] 图4显示了由实施例4-6中所述的mRNA编码的抗原的示例性线性设计。

[0018] 图5显示了本文所述的各种S1亚基抗原的序列比对。

[0019] 图6显示了由实施例7和8中所述的mRNA编码的抗原的示例性线性设计。

[0020] 图7显示了中和与ELISA滴度的相关性。

[0021] 图8A-图8C显示了用LNP中的编码NTD-RBD-TM的mRNA对小鼠进行第1天初免和第21天加强剂量后第36天的血清IgG1滴度和IgG2a滴度。

具体实施方式

[0022] 严重急性呼吸综合征冠状病毒2(SARS-CoV-2)是新近出现的高发病率和死亡率的呼吸道病毒。与2002年出现的SARS-CoV和2012年出现的中东呼吸综合征冠状病毒(MERS-CoV)相比,SARS-CoV-2已经迅速在全世界传播。世界卫生组织(WHO)报告,截至2020年7月6日,COVID-19的当前爆发在世界范围内具有几乎1150万确诊病例,其中有超过530,000例死亡。COVID-19感染的新病例呈上升趋势,并且仍然在迅速增加。因此,开发了各种安全且有效的疫苗和药物来预防和治疗COVID-19并且减少COVID-19在全世界所具有的严重影响是至关重要的。使用多种方式制得的疫苗和药物以及具有改进的安全性和功效的疫苗是必要的。仍然需要加速针对冠状病毒疾病2019(COVID-19)的疫苗和治疗药物的先进设计和开发。

[0023] 在2020年1月7日,将严重急性呼吸综合征冠状病毒2(SARS-CoV-2)鉴定为于2019年12月出现在中国湖北省武汉市的新型肺炎的病原体(Lu H等人,(2020)J Med Virol.4月;92(4):401-402。)。不久以后,所述病毒在中国引起爆发并传播至世界各地。根据SARS-CoV-2的基因组结构的分析,其属于β-冠状病毒(CoV)(Chan等人,2020 Emerg Microbes Infect.;9(1):221-236)。

[0024] 冠状病毒表面上的关键蛋白是刺突蛋白。已设计大量mRNA构建体并在本文中公开。当在适当递送媒介物中配制时,编码刺突抗原、其亚基和结构域的mRNA能够诱导针对SARS-CoV-2的强烈免疫反应,从而产生有效且强效的mRNA疫苗。施用编码各种刺突蛋白抗原、特别是刺突蛋白亚基和结构域抗原的mRNA导致mRNA递送至免疫系统的免疫组织和细胞,在那里其快速翻译成蛋白抗原。其他免疫细胞(例如B细胞和T细胞)然后能够识别和启动,并且免疫反应发展出针对编码蛋白的免疫反应,并最终产生针对冠状病毒的持久保护

性反应。经由使用本文公开的编码刺突蛋白、其亚基和结构域的高效mRNA疫苗避免了低免疫原性(由于对免疫系统的呈递差或抗原的不正确折叠而导致的蛋白疫苗开发的缺点)。

[0025] 本公开提供了引发针对冠状病毒抗原的强效中和抗体的组合物(例如,mRNA疫苗)。在一些实施方案中,组合物包括编码至少一种(例如,一种、两种或更多种)冠状病毒抗原(例如SARS-CoV-2抗原)的mRNA。在一些实施方案中,所述mRNA编码刺突蛋白结构域,例如受体结合结构域(RBD)、N末端结构域(NTD)或RBD和NTD的组合。

[0026] 本公开的一些方面提供了包含编码融合蛋白的开放阅读框的信使核糖核酸(mRNA),所述融合蛋白包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)和蛋白质跨膜结构域,例如天然存在的或异源的跨膜结构域。

[0027] 在一些实施方案中,所述蛋白质跨膜结构域是流感血凝素跨膜结构域。

[0028] 在一些实施方案中,所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少80%同一性的氨基酸序列。

[0029] 在一些实施方案中,所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0030] 在一些实施方案中,所述融合蛋白包含SEQ ID NO:77的氨基酸序列。

[0031] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70%同一性的核苷酸序列。

[0032] 在一些实施方案中,其中所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0033] 在一些实施方案中,所述开放阅读框包含SEQ ID NO:76的核苷酸序列。

[0034] 本公开的其他方面提供了包含编码融合蛋白的开放阅读框的信使核糖核酸(mRNA),所述融合蛋白包含SARS-CoV-2刺突蛋白的氨基(N)末端结构域和跨膜结构域。

[0035] 在一些实施方案中,所述跨膜结构域是流感血凝素跨膜结构域。

[0036] 在一些实施方案中,所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少80%同一性的氨基酸序列。

[0037] 在一些实施方案中,所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0038] 在一些实施方案中,所述融合蛋白包含SEQ ID NO:47的氨基酸序列。

[0039] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70%同一性的核苷酸序列。

[0040] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0041] 在一些实施方案中,所述开放阅读框包含SEQ ID NO:46的核苷酸序列。

[0042] 本公开的其他方面提供了包含编码融合蛋白的开放阅读框的信使核糖核酸(mRNA),所述融合蛋白包含任选地经由接头与SARS-CoV-2刺突蛋白的氨基(N)末端结构域

连接的SARS-CoV-2刺突蛋白的受体结合结构域。

[0043] 在一些实施方案中,所述融合蛋白还包含跨膜结构域。

[0044] 在一些实施方案中,所述融合蛋白包含与SEQ ID NO:92的氨基酸序列具有至少80%同一性的氨基酸序列。

[0045] 在一些实施方案中,所述融合蛋白包含与SEQ ID NO:92的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0046] 在一些实施方案中,所述融合蛋白包含SEQ ID NO:92的氨基酸序列。

[0047] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:91的核苷酸序列具有至少70%同一性的核苷酸序列。

[0048] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:91的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0049] 在一些实施方案中,所述开放阅读框包含SEQ ID NO:91的核苷酸序列。

[0050] 在一些实施方案中,所述mRNA还包含5′非翻译区(UTR),所述5′非翻译区任选地包含SEQ ID NO:131或2的核苷酸序列。

[0051] 在一些实施方案中,所述mRNA还包含3′非翻译区(UTR),所述3′非翻译区任选地包含SEQ ID NO:132或4的核苷酸序列。

[0052] 在一些实施方案中,所述mRNA还包含5′帽,任选地7mG(5′)ppp(5′)NlmpNp。

[0053] 在一些实施方案中,所述mRNA还包含多聚A尾,所述多聚A尾任选地具有约100个核苷酸的长度。

[0054] 在一些实施方案中,所述mRNA包含化学修饰,任选地1-甲基假尿苷。

[0055] 本公开的一些方面提供了包含前述段落中任一者的mRNA的组合物。

[0056] 本公开的其他方面提供了包含前述段落中任一者的至少两种mRNA的组合物。

[0057] 本公开的其他方面提供了一种组合物,所述组合物包含:(a)包含编码融合蛋白的开放阅读框的信使核糖核酸(mRNA),所述融合蛋白包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)和蛋白质跨膜结构域;和(b)包含编码融合蛋白的开放阅读框的mRNA,所述融合蛋白包含SARS-CoV-2刺突蛋白的氨基(N)末端结构域和跨膜结构域。在一些实施方案中,(a)的mRNA对(b)的mRNA的比率为约1:1,例如1:2、1:3、2:1或3:1。在一些实施方案中,组合物的至少50%的mRNA为(a)的mRNA。例如,组合物的至少55%、60%、65%、70%、75%、80%、85%、90%或95%的mRNA为(a)的mRNA。在一些实施方案中,组合物的至少50%的mRNA为(b)的mRNA。例如,组合物的至少55%、60%、65%、70%、75%、80%、85%、90%或95%的mRNA为(b)的mRNA。

[0058] 在一些实施方案中,所述蛋白质跨膜结构域是流感血凝素跨膜结构域。

[0059] 在一些实施方案中,(a)的融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少80%同一性的氨基酸序列。

[0060] 在一些实施方案中,(a)的融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0061] 在一些实施方案中，(a)的融合蛋白包含SEQ ID NO:77的氨基酸序列。

[0062] 在一些实施方案中，(a)的开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70％同一性的核苷酸序列。

[0063] 在一些实施方案中，(a)的开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列。

[0064] 在一些实施方案中，(a)的开放阅读框包含SEQ ID NO:76的核苷酸序列。

[0065] 在一些实施方案中，(b)的融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少80％同一性的氨基酸序列。

[0066] 在一些实施方案中，(b)的融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列。

[0067] 在一些实施方案中，(b)的融合蛋白包含SEQ ID NO:47的氨基酸序列。

[0068] 在一些实施方案中，(b)的开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70％同一性的核苷酸序列。

[0069] 在一些实施方案中，(b)的开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列。

[0070] 在一些实施方案中，(b)的开放阅读框包含SEQ ID NO:46的核苷酸序列。

[0071] 在一些实施方案中，所述mRNA配制在脂质纳米颗粒中。

[0072] 在一些实施方案中，所述组合物还包含脂质纳米颗粒。

[0073] 在一些实施方案中，(a)的mRNA配制在脂质纳米颗粒中，并且(b)的mRNA配制在脂质纳米颗粒中。

[0074] 在一些实施方案中，所述脂质纳米颗粒包含阳离子脂质。

[0075] 在一些实施方案中，所述脂质纳米颗粒还包含中性脂质。

[0076] 在一些实施方案中，所述脂质纳米颗粒还包含固醇。

[0077] 在一些实施方案中，所述脂质纳米颗粒还包含聚乙二醇(PEG)修饰的脂质。

[0078] 在一些实施方案中，所述脂质纳米颗粒包含可电离的阳离子脂质、中性脂质、固醇和PEG修饰的脂质。

[0079] 在一些实施方案中，所述可电离的阳离子脂质是8((2羟基乙基)(6氧代6-(十一烷基氧基)己基)氨基)辛酸十七烷-9-基酯(化合物1)。

[0080] 在一些实施方案中，所述中性脂质是1,2二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)。

[0081] 在一些实施方案中，所述固醇是胆固醇。

[0082] 在一些实施方案中，所述PEG修饰的脂质是1,2二肉豆蔻酰基-sn-甘油-甲氧基聚乙二醇(PEG2000 DMG)。

[0083] 在一些实施方案中，所述脂质纳米颗粒包含20-60mol％可电离的阳离子脂质、5-25mol％中性脂质、25-55mol％固醇和0.5-15mol％PEG修饰的脂质。

[0084] 在一些实施方案中，所述脂质纳米颗粒包含：47mol％可电离的阳离子脂质；

11.5mol％中性脂质;38.5mol％固醇;和3.0mol％PEG修饰的脂质;48mol％可电离的阳离子脂质;11mol％中性脂质;38.5mol％固醇;和2.5mol％PEG修饰的脂质;49mol％可电离的阳离子脂质;10.5mol％中性脂质;38.5mol％固醇;和2.0mol％PEG修饰的脂质;50mol％可电离的阳离子脂质;10mol％中性脂质;38.5mol％固醇;和1.5mol％PEG修饰的脂质;或51mol％可电离的阳离子脂质;9.5mol％中性脂质;38.5mol％固醇;和1.0mol％PEG修饰的脂质。

[0085] 在一些实施方案中,所述脂质纳米颗粒包含47mol％化合物1;11.5mol％DSPC;38.5mol％胆固醇;和3.0mol％PEG2000 DMG;48mol％化合物1;11mol％DSPC;38.5mol％胆固醇;和2.5mol％PEG2000 DMG;49mol％化合物1;10.5mol％DSPC;38.5mol％胆固醇;和2.0mol％PEG2000 DMG;50mol％化合物1;10mol％DSPC;38.5mol％胆固醇;和1.5mol％PEG2000 DMG;或51mol％化合物1;9.5mol％DSPC;38.5mol％胆固醇;和1.0mol％PEG2000 DMG。

[0086] 本公开的其他方面提供了一种方法,所述方法包括向受试者施用在所述受试者中有效诱导针对SARS-CoV-2的中和抗体反应的量的前述权利要求中任一项所述的mRNA或组合物。

[0087] 本公开的其他方面提供了一种方法,所述方法包括向受试者施用在所述受试者中有效诱导针对SARS-CoV-2的T细胞免疫反应的量的前述权利要求中任一项所述的mRNA或组合物。

[0088] 本公开的一些方面提供了包含开放阅读框(ORF)的信使核糖核酸(mRNA),所述开放阅读框编码能够诱导针对SARS-CoV-2的免疫反应(例如中和抗体反应)的冠状病毒抗原,其中所述抗原包含SARS-CoV-2的蛋白质片段或功能性蛋白质结构域,任选地其中所述RNA配制在脂质纳米颗粒中。

[0089] 在一些实施方案中,所述抗原是功能性蛋白质结构域。

[0090] 在一些实施方案中,所述蛋白质结构域是SARS-CoV-2刺突蛋白的N末端结构域(NTD)。

[0091] 在一些实施方案中,所述NTD与跨膜结构域、任选地流感血凝素跨膜结构域连接。

[0092] 在一些实施方案中,所述抗原包含与SEQ ID NO:47的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:47的氨基酸序列。

[0093] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:46的核苷酸序列。

[0094] 在一些实施方案中,所述蛋白质结构域是SARS-CoV-2刺突蛋白的受体结合结构域(RBD)。

[0095] 在一些实施方案中,所述RBD是可溶的。

[0096] 在一些实施方案中,所述抗原包含与SEQ ID NO:62的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:62的氨基酸序列。

[0097] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:61的核苷酸序列具有至少

70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:61的核苷酸序列。

[0098] 在一些实施方案中，所述RBD与跨膜结构域、任选地流感血凝素跨膜结构域连接。

[0099] 在一些实施方案中，所述抗原包含与SEQ ID NO:77的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:77的氨基酸序列。

[0100] 在一些实施方案中，所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:76的核苷酸序列。

[0101] 在一些实施方案中，所述NTD与SARS-CoV-2刺突蛋白的RBD连接以形成NTD-RBD融合蛋白。

[0102] 在一些实施方案中，所述NTD-RBD融合体与跨膜结构域(TM)、任选地流感血凝素跨膜结构域连接，以形成NTD-RBD-TM蛋白。

[0103] 在一些实施方案中，所述抗原包含与SEQ ID NO:92的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:92的氨基酸序列。

[0104] 在一些实施方案中，所述开放阅读框包含与SEQ ID NO:91的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:91的核苷酸序列。

[0105] 在一些实施方案中，所述NTD-RBD融合体包含C末端截短。

[0106] 在一些实施方案中，所述抗原包含与SEQ ID NO:107的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:107的氨基酸序列。

[0107] 在一些实施方案中，所述开放阅读框包含与SEQ ID NO:106的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:106的核苷酸序列。

[0108] 在一些实施方案中，所述NTD和/或RBD包括延伸区。

[0109] 在一些实施方案中，所述抗原包含与SEQ ID NO:59、86、89、116、119或122中的任一者的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:59、86、89、116、119或122中的任一者的氨基酸序列。

[0110] 在一些实施方案中，所述开放阅读框包含与SEQ ID NO:58、85、88、115、118或121中的任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:58、85、88、115、118或121中的任一者的核苷酸序列。

[0111] 在一些实施方案中，所述蛋白质结构域是SARS-CoV-2刺突蛋白的S1亚基结构域。

[0112] 在一些实施方案中，所述S1亚基是可溶的。

[0113] 在一些实施方案中，所述抗原包含与SEQ ID NO:5的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序

列,任选地其中所述抗原包含SEQ ID NO:5的氨基酸序列。

[0114] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:3的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:3的核苷酸序列。

[0115] 在一些实施方案中,所述S1亚基与跨膜结构域、任选地流感血凝素跨膜结构域连接。

[0116] 在一些实施方案中,所述抗原包含与SEQ ID NO:17的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:17的氨基酸序列。

[0117] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:16的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:16的核苷酸序列。

[0118] 在一些实施方案中,所述S1亚基经修饰以去除S蛋白的RBD或RBD的一部分。

[0119] 在一些实施方案中,所述抗原包含与SEQ ID NO:20、23、26、29、32或35中的任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:20、23、26、29、32或35中的任一者的氨基酸序列。

[0120] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:19、22、25、28、41或34中的任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:19、22、25、28、31或34中的任一者的核苷酸序列。

[0121] 在一些实施方案中,所述S1亚基与S蛋白的S2亚基连接。

[0122] 在一些实施方案中,所述S2亚基来自SARS-CoV-2 S蛋白。

[0123] 在一些实施方案中,所述S1亚基来自HKU1 S蛋白。

[0124] 在一些实施方案中,所述抗原包含与SEQ ID NO:38的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:38的氨基酸序列。

[0125] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:37的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:37的核苷酸序列。

[0126] 在一些实施方案中,所述S1亚基来自OC43S蛋白。

[0127] 在一些实施方案中,所述抗原包含与SEQ ID NO:41的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:41的氨基酸序列。

[0128] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:40的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:40的核苷酸序列。

[0129] 在一些实施方案中,所述抗原还包含支架结构域,任选地选自铁蛋白、2,4-二氧四氢蝶啶(lumazine)合成酶和折叠子。

[0130] 在一些实施方案中,所述支架结构域是铁蛋白。

[0131] 在一些实施方案中,所述抗原包含与SEQ ID NO:8或65的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:8或65的氨基酸序列。

[0132] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:7或64的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:7或64的核苷酸序列。

[0133] 在一些实施方案中,所述支架结构域是2,4-二氧四氢蝶啶合成酶。

[0134] 在一些实施方案中,所述抗原包含与SEQ ID NO:11、14、68或71中的任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:11、14、68或71中的任一者的氨基酸序列。

[0135] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:10、13、67或70中的任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:10、13、67或70中的任一者的核苷酸序列。

[0136] 在一些实施方案中,所述支架结构域是折叠子。

[0137] 在一些实施方案中,所述抗原包含与SEQ ID NO:44、50、74、80、83、101、104或113中的任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:44、50、74、80、83、101、104或113中的任一者的氨基酸序列。

[0138] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:43、49、73、79、82、100、103或112中的任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:43、49、73、79、82、100、103或112中的任一者的核苷酸序列。

[0139] 在一些实施方案中,所述抗原还包含运输信号,任选地选自巨噬细胞标志物,任选地CD86、CD11B和/或VSVGct。

[0140] 在一些实施方案中,所述抗原包含与SEQ ID NO:95、98或110中的任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:95、98或110中的任一者的氨基酸序列。

[0141] 在一些实施方案中,所述开放阅读框包含与SEQ ID NO:94、97或109中的任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:94、97或109中的任一者的核苷酸序列。

[0142] 在一些实施方案中,所述mRNA配制在脂质纳米颗粒中。

[0143] 在一些实施方案中,所述脂质纳米颗粒包含阳离子脂质(任选地可电离的阳离子脂质)、中性脂质、固醇和/或聚乙二醇(PEG)修饰的脂质。可电离的阳离子脂质在本文中可

与可电离的脂质和阳离子脂质互换使用以指可电离的脂质。在一些实施方案中，所述脂质纳米颗粒包含40-50mol％可电离的脂质，任选地45-50mol％，例如45-46mol％、46-47mol％、47-48mol％、48-49mol％或49-50mol％，例如约45mol％、45.5mol％、46mol％、46.5mol％、47mol％、47.5mol％、48mol％、48.5mol％、49mol％或49.5mol％。在一些实施方案中，所述脂质纳米颗粒包含30-45mol％固醇，任选地35-40mol％，例如30-31mol％、31-32mol％、32-33mol％、33-34mol％、35-35mol％、35-36mol％、36-37mol％、38-38mol％、38-39mol％或39-40mol％。在一些实施方案中，所述脂质纳米颗粒包含5-15mol％辅助脂质，任选地10-12mol％，例如5-6mol％、6-7mol％、7-8mol％、8-9mol％、9-10mol％、10-11mol％、11-12mol％、12-13mol％、13-14mol％或14-15mol％。在一些实施方案中，所述脂质纳米颗粒包含1-5％PEG脂质，任选地1-3mol％，例如1.5-2.5mol％、1-2mol％、2-3mol％、3-4mol％或4-5mol％。

[0144] 在一些实施方案中，所述可电离的阳离子脂质是8((2羟基乙基)(6氧代6-(十一烷基氧基)己基)氨基)辛酸十七烷-9-基酯(化合物1)，所述中性脂质是1,2二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)，所述固醇是胆固醇，和/或所述PEG修饰的脂质是1,2二肉豆蔻酰基-sn-甘油-甲氧基聚乙二醇(PEG2000 DMG)。

[0145] 在一些实施方案中，所述脂质纳米颗粒包含20-60mol％可电离的阳离子脂质、5-25mol％中性脂质、25-55mol％固醇和0.5-15mol％PEG修饰的脂质。

[0146] 在一些实施方案中，所述脂质纳米颗粒包含：47mol％可电离的阳离子脂质；11.5mol％中性脂质；38.5mol％固醇；和3.0mol％PEG修饰的脂质；48mol％可电离的阳离子脂质；11mol％中性脂质；38.5mol％固醇；和2.5mol％PEG修饰的脂质；49mol％可电离的阳离子脂质；10.5mol％中性脂质；38.5mol％固醇；和2.0mol％PEG修饰的脂质；50mol％可电离的阳离子脂质；10mol％中性脂质；38.5mol％固醇；和1.5mol％PEG修饰的脂质；或51mol％可电离的阳离子脂质；9.5mol％中性脂质；38.5mol％固醇；和1.0mol％PEG修饰的脂质。

[0147] 在一些实施方案中，所述脂质纳米颗粒包含：47mol％化合物1；11.5mol％DSPC；38.5mol％胆固醇；和3.0mol％PEG2000 DMG；48mol％化合物1；11mol％DSPC；38.5mol％胆固醇；和2.5mol％PEG2000 DMG；49mol％化合物1；10.5mol％DSPC；38.5mol％胆固醇；和2.0mol％PEG2000 DMG；50mol％化合物1；10mol％DSPC；38.5mol％胆固醇；和1.5mol％PEG2000 DMG；或51mol％化合物1；9.5mol％DSPC；38.5mol％胆固醇；和1.0mol％PEG2000 DMG。

[0148] 国际申请号PCT/US2016/058327(公开号WO2017/070626)和国际申请号PCT/US2018/022777(公开号WO2018/170347)的全部内容通过引用并入本文。

[0149] SARS-Cov-2

[0150] SARS-CoV-2的基因组是单链正义RNA(+ssRNA)，大小为29.8-30kb，编码约9860个氨基酸(Chan等人，2000，同上；Kim等人，2020 Cell，5月14日；181(4)：914-921.e10.)。SARS-CoV-2是具有5′-帽和3′-多聚A尾的多顺反子mRNA。将SARS-CoV-2基因组组织成编码结构蛋白和非结构蛋白(Nsp)的特定基因。结构蛋白在基因组中的顺序是5′-复制酶(开放阅读框(ORF)1/ab)-结构蛋白[刺突(S)-包膜(E)-膜(M)-核衣壳(N)]-3′。冠状病毒的基因组包括可变数目的编码辅助蛋白、非结构蛋白和结构蛋白的开放阅读框(Song等人，2019

Viruses;11(1):第59页)。大部分抗原性肽位于结构蛋白中(Cui等人,2019 Nat.Rev.Microbiol.;17(3):181-192)。刺突表面糖蛋白(S)、小包膜蛋白(E)、基质蛋白(M)和核衣壳蛋白(N)是四种主要的结构蛋白。由于S-蛋白有助于细胞向性,其能够诱导中和抗体(NAb)和保护性免疫,因此可将其视为所有其他结构蛋白中冠状病毒疫苗开发中最重要的靶标之一。此外,氨基酸序列分析显示S-蛋白含有冠状病毒中的保守区,其可为通用疫苗开发的基础。

[0151] 抗原

[0152] 本发明的组合物(例如疫苗组合物)的特征在于设计用于编码目标抗原(例如,源自β冠状病毒结构蛋白的抗原,特别是源自SARS-CoV-2刺突蛋白的抗原)的核酸,特别是mRNA。本发明的组合物(例如疫苗组合物)不包含抗原本身,而是包含一旦递送至细胞、组织或受试者即编码抗原或抗原性序列的核酸,特别是mRNA。核酸分子(特别是mRNA)的递送是通过将所述核酸分子配制在适当载体或递送媒介物(例如,脂质纳米颗粒)中来实现,使得在施用至细胞、组织或受试者后,核酸由细胞摄取,所述细胞进而表达由核酸(例如mRNA)编码的蛋白质。如本文所用的术语"抗原"是指例如蛋白质(例如糖蛋白)、多肽、肽等的物质,其引发免疫反应,例如当存在于受试者中时(例如,当存在于人类或哺乳动物受试者中时)引发免疫反应。本发明至少部分基于以下理解:当mRNA编码的抗原从施用至细胞或受试者的mRNA表达时,可引起免疫系统产生针对表达的抗原的免疫反应,例如可触发针对表达抗原的抗体(例如结合和/或中和抗体)的产生,可触发对表达的抗原具特异性的B细胞和或T细胞反应,并且最终可引起针对随后遇到抗原或与抗原相关的病原体的保护性或预防性反应。优选的mRNA编码的抗原是"病毒抗原"。如本文所用的术语"病毒抗原"是指源自病毒、例如源自致病病毒的抗原。如本文所用的术语抗原可指全长蛋白,例如全长病毒蛋白,或者可指蛋白的片段(例如多肽或肽片段)、亚基或结构域,例如病毒蛋白亚基或结构域。

[0153] 许多蛋白质具有四级或三维结构,其由一个以上多肽或若干缔合成寡聚分子的多肽链组成。如本文所用的术语"亚基"是指单一蛋白质分子,例如,由加工新生蛋白质分子产生的多肽或多肽链,所述亚基与其他蛋白质分子(例如亚基或链)组装(或"共组装")形成蛋白质复合体。蛋白质可具有相对小数目的亚基,并且因此阐述为"寡聚的",或者可由大数目的亚基组成并且因此阐述为"多聚的"。寡聚或多聚蛋白质的亚基可为相同的、同源的或完全不相似的,并且专用于不同的任务。

[0154] 蛋白质或蛋白质亚基还可包含结构域。如本文所用的术语"结构域"是指蛋白质内不同的功能和/或结构单元。通常,"结构域"负责特定功能或相互作用,从而有助于蛋白质的总体作用。结构域可存在于各种生物环境中。相似结构域(即,共享结构、功能和/或序列同源性的结构域)可存在于单一蛋白质内或者可存在于具有相似或不同功能的不同蛋白质内。蛋白质结构域通常是给定蛋白质三级结构或序列的保守部分,其可独立于蛋白质或其亚基的其余部分起作用和存在。

[0155] 在结构和分子生物学中,相同、同源或相似的亚基或结构域可帮助对新近鉴定或新颖蛋白质进行分类,正如在SARS-CoV-2病毒基因组序列公开后立即进行的那样。

[0156] 如本文所用的术语抗原不同于术语"表位",所述表位是抗原的亚结构,例如多肽或碳水化合物结构,其可由抗原结合位点识别,但不足以诱导免疫反应。本领域阐述了以分离形式递送至受试者或免疫细胞的蛋白抗原,例如分离的蛋白质、多肽或肽抗原,然而,蛋

白抗原的设计、测试、验证和产生可为昂贵且耗时的，尤其是当大规模产生蛋白时。相反，mRNA技术可用于快速设计和测试编码各种抗原的mRNA构建体。此外，结合配制在适当递送媒介物(例如，脂质纳米颗粒)中的mRNA的快速产生可快速进行并且可快速大规模产生mRNA疫苗。潜在益处还来自于以下事实：由本发明的mRNA编码的抗原由受试者的细胞表达，例如由人体表达，并且因此，受试者(例如人体)用作“工厂”以产生抗原，其进而引发期望免疫反应。

[0157] 在优选方面，抗原是能够诱导免疫反应(例如，引起免疫系统产生针对抗原的抗体)的蛋白质。在本文中，除非另有说明，否则术语“抗原”的使用涵盖免疫原性蛋白以及衍生自免疫原性蛋白的多肽或肽，例如免疫原性片段(诱导(或能够诱导)对抗原的免疫反应的免疫原性片段)。应理解，术语“蛋白质”涵盖多肽和肽，并且术语“抗原”涵盖抗原性片段。其他分子可为抗原性的，例如细菌多糖或蛋白和多糖结构的组合，但对于本文包括的病毒疫苗，源自β冠状病毒SARS-CoV-2的病毒蛋白、病毒蛋白片段和设计和或突变的蛋白是本文特征性的抗原。

[0158] 核酸/mRNA

[0159] 本文所述的疫苗技术的特征在于设计用于编码目标抗原(例如β冠状病毒刺突蛋白抗原、其亚基、结构域或片段(例如抗原性片段))的核酸，特别是信使RNA(mRNA)。本发明的核酸(例如mRNA)优选配制在适当载体或递送媒介物(例如，脂质纳米颗粒)中，使得核酸(例如mRNA)适于体内使用。核酸(例如mRNA)当经适当配制时能够递送至受试者(例如人类受试者)内的细胞和/或组织，以完成由这些核酸编码的蛋白质的翻译。

[0160] 核酸分子是由携带遗传信息的连接核苷酸构成的大分子，并且通过指导蛋白质合成过程，指导大部分(如果并非全部)细胞功能。核酸包含核苷酸(核苷酸单体)的聚合物。因此，核酸也称为多核苷酸(也称为多核苷酸链)。核酸的两个主要类别是脱氧核糖核酸(DNA)和核糖核酸(RNA)。DNA构成所有自由生活的生物体和大部分病毒中的遗传物质。RNA是某些病毒的遗传物质，但也在所有活细胞中发现，在所述活细胞中，RNA在细胞过程中、最显著的是在蛋白质的制备中起重要作用。

[0161] 核苷是核酸(例如DNA和RNA)的结构亚基。核苷由共价附接至五碳碳水化合物核糖或“糖”(其为核糖或脱氧核糖)的含氮碱基(核碱基)、通常为嘧啶(胞嘧啶、胸腺嘧啶或尿嘧啶)或嘌呤(腺嘌呤或鸟嘌呤)构成。核苷酸由含氮碱基、糖(核糖或脱氧核糖)和一个至三个磷酸基团组成。本质上，核苷酸仅仅是具有一个或多个额外磷酸基团的核苷。

[0162] 核酸分子(DNA和RNA)由在一个核苷酸的糖碱基与相邻核苷酸的磷酸基团之间通过化学键(称为酯键)在链中彼此连接的核苷酸构成。糖在每个核苷酸的3′端，并且磷酸在每个核苷酸的5′端。附接至一个核苷酸上的糖的5′碳的磷酸基团与下一个核苷酸的3′碳上的游离羟基形成酯键。这些键称为磷酸二酯键，并且当合成分子时，糖-磷酸主链阐述为在5′至3′方向上延伸或生长。

[0163] 核酸的核碱基部分的特征在于嘌呤碱基：腺嘌呤(A)和鸟嘌呤(G)；和嘧啶碱基：胞嘧啶(C)、DNA中的胸腺嘧啶(T)和RNA中的尿嘧啶(U)。核酸的糖部分的特征在于DNA中的脱氧核糖、RNA中的核糖。五个核苷通常分别缩写为其单字母代码A、G、C、T和U。然而，胸苷更通常地写为“dT”(“d”代表“脱氧”)，因为其含有2′-脱氧呋喃核糖部分而非尿苷中发现的呋喃核糖环。这是由于胸苷在脱氧核糖核酸(DNA)而非核糖核酸(RNA)中发现。相反，尿苷在RNA

中而非DNA中发现。其余三个核苷可在RNA和DNA二者中发现。在RNA中，其将表示为A、C和G，而在DNA中，其将表示为dA、dC和dG。

[0164] 本领域技术人员应理解，除非另有说明，否则本申请中所阐释的核酸序列可在代表性DNA序列中列举“T”，但在序列表示mRNA的情况下，“T”将被取代为“U”。因此，所公开和由本文特定序列识别号鉴定的任何DNA还公开了与DNA互补的相应mRNA序列，其中DNA序列的每个“T”被“U”取代。

[0165] 核酸可为或可包括例如脱氧核糖核酸(DNA)、核糖核酸(RNA)(例如mRNA)、苏糖核酸(TNA)、二醇核酸(GNA)、肽核酸(PNA)、锁核酸(LNA，包括具有β-D-核糖构型的LNA、具有α-L-核糖构型的α-LNA(LNA的非对映异构体)、具有2′-氨基官能化的2′-氨基-LNA和具有2′-氨基官能化的2′-氨基-α-LNA)、乙烯核酸(ENA)、环己烯基核酸(CeNA)和/或其嵌合体和/或组合。

[0166] 本发明的特征在于信使RNA(mRNA)，特别是设计用于编码目标抗原(例如β冠状病毒刺突蛋白抗原、其亚基、结构域或片段(例如抗原性片段))的mRNA。信使RNA(mRNA)(一种RNA亚型)是对应于基因的遗传序列的单链RNA分子。在转录过程中产生mRNA，其中单链DNA通过RNA聚合酶解码，并且合成、即转录mRNA。mRNA在合成蛋白质、即翻译的过程中由核糖体读取。因此，信使RNA(mRNA)是编码(至少一种)蛋白质(天然存在的、非天然存在的或经修饰的氨基酸聚合物)的RNA，并且可经翻译以在体外、体内、原位或离体产生编码的蛋白质。

[0167] 本公开的组合物包含(至少一种)具有编码冠状病毒抗原的开放阅读框(ORF)的mRNA。在一些实施方案中，mRNA还包含5′UTR、3′UTR、多聚(A)尾和/或5′帽或帽类似物。开放阅读框(ORF)是一段连续的DNA或RNA，以起始密码子(例如，甲硫氨酸(ATG或AUG))开始并以终止密码子(例如TAA、TAG或TGA，或UAA、UAG或UGA)结束。ORF通常编码蛋白质。应理解，本文公开的序列还可包含额外元件，例如5′和3′UTR，但与ORF不同，那些元件不一定存在于本公开的mRNA中。还应理解，本发明的mRNA(例如本公开的β冠状病毒疫苗中特征性的mRNA)可包括任何5′非翻译区(UTR)和/或任何3′UTR。示例性UTR序列提供于序列表中(例如，SEQ ID NO:2、4、131和132)；然而，其他UTR序列可使用或替换本文所述的任何UTR序列。还可从本文提供的mRNA省略UTR。

[0168] 在一些实施方案中，组合物包含mRNA，所述mRNA包含与SEQ ID NO:45、75或90中的任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或100％同一性的核苷酸序列。在一些实施方案中，组合物包含mRNA，所述mRNA包含与表1至表15中的序列中的任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或100％同一性的核苷酸序列。

[0169] 在一些实施方案中，组合物包含mRNA，所述mRNA包含与SEQ ID NO:46、76或91中的任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或100％同一性的ORF。在一些实施方案中，组合物包含mRNA，所述mRNA包含与表1至表15中的序列中的任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或100％同一性的ORF。

[0170] 表1至表15中提供了本公开的组合物的冠状病毒抗原和编码所述冠状病毒抗原的

mRNA的示例性序列。

[0171] 应理解，本文所述mRNA编码的任何一种抗原可包含或不包含信号序列。

[0172] 编码的冠状病毒刺突(S)蛋白抗原

[0173] 已知的β冠状病毒的包膜刺突(S)蛋白决定病毒宿主向性和进入宿主细胞。冠状病毒刺突(S)蛋白是疫苗设计的选择抗原，因为其可诱导中和抗体和保护性免疫。S蛋白对于SARS-CoV-2感染至关重要。S蛋白的组织在β冠状病毒(例如SARS-CoV-2、SARS-CoV、MERS-CoV、HKU1-CoV、MHV-CoV和NL63-CoV)中是相似的。

[0174] 如本文所用的术语“刺突蛋白”是指形成从病毒(包括β冠状病毒)的包膜(病毒表面)突出的同三聚体的糖蛋白。三聚化刺突蛋白通过与宿主细胞表面上的受体结合、接着融合病毒和宿主细胞膜而促进病毒粒子进入宿主细胞。S蛋白是高度糖基化的大的I型跨膜融合蛋白，其由1,160至1,400个氨基酸组成，这取决于病毒的类型。β冠状病毒刺突蛋白包含约1100至1500个氨基酸，并且包含如图1所阐释的结构(即，结构域组成和组织)。SARS-CoV-2刺突(S)蛋白是疫苗设计的选择抗原，因为其可诱导中和抗体和保护性免疫。本发明的mRNA经设计用于产生SARS-CoV-2刺突蛋白(即，编码刺突蛋白，使得当mRNA递送至细胞或组织(例如受试者的细胞或组织)时刺突蛋白表达)，以及其抗原性变体。本领域技术人员将理解，尽管基本上全长或完整刺突蛋白对于病毒(例如β冠状病毒)实施其促进病毒进入宿主细胞的预期功能可为必需的，但当主要寻求引发针对刺突蛋白的免疫反应时，可耐受刺突蛋白结构和/或序列的一定量的变化。例如，可耐受例如从编码的刺突蛋白(例如编码的刺突蛋白抗原)的N末端或C末端的一个至几个、可能至多5个或至多10个氨基酸的少量截短，而不改变蛋白质的抗原性质。同样，可耐受编码的刺突蛋白(例如编码的刺突蛋白抗原)的一个至几个、可能至多5个或至多10个氨基酸(或更多)的变化(例如保守取代)，而不改变蛋白质的抗原性质。在示例性实施方案中，刺突蛋白(例如编码的刺突蛋白抗原)具有表1-15的序列中的任一者中所阐释的氨基酸序列(例如，源自如SEQ ID NO:125所阐释的氨基酸序列)。在其他实施方案中，与具有表1-15的序列中的任一者中所阐释的氨基酸序列(例如，源自如SEQ ID NO:125所阐释的氨基酸序列)的刺突蛋白相比(当与其比对时)，刺突蛋白(例如编码的刺突蛋白抗原)具有不大于100、不大于90、不大于80、不大于70、不大于60、不大于50、不大于40、不大于30、不大于20、不大于10或不大于5个氨基酸取代和/或缺失。在编码的刺突蛋白序列中发生微小变化的情况下，变体优选具有与参考刺突蛋白序列相同的活性和/或具有与参考刺突蛋白相同的免疫特异性，如例如在免疫测定(例如酶联免疫吸附测定(ELISA测定))中所确定。

[0175] 冠状病毒的S蛋白可分为两个重要的功能亚基，其中包括N末端S1亚基，其形成S蛋白的球状头部；和C末端S2区，其形成蛋白的茎并直接嵌入病毒包膜中。在与潜在宿主细胞相互作用时，S1亚基将识别并结合至宿主细胞上的受体，特别是血管收缩肽转化酶2(ACE2)受体，而作为S蛋白的最保守组分的S2亚基将负责将病毒的包膜与宿主细胞膜融合。(参见例如Shang等人，PLoS Pathog.2020年3月;16(3):e1008392。)。三聚化S蛋白三聚体的每个单体含有两个亚基S1和S2，分别介导附着和膜融合。参见例如图1。作为体内感染过程的一部分，两个亚基通过酶切割过程彼此分离。S蛋白首先通过弗林蛋白酶(furin)介导的切割在感染细胞中在S1/S2位点进行切割，在体内，随后在S1内的S2′位点发生丝氨酸蛋白酶介导的切割事件。在SARS-CoV2中，S1/S2切割位点在氨基酸676-TQTNSPRRAR/SVA-688(参考

SEQ ID NO:127)。S2′切割位点在氨基酸811-KPSKR/SFI-818(参考SEQ ID NO:126)。

[0176] 如本文所用,例如在设计由本发明的核酸(例如mRNA)编码的SARS-CoV-2 S蛋白抗原的情况下,术语“S1亚基”(例如,S1亚基抗原)是指刺突蛋白的N末端亚基,以S蛋白N末端开始并且以S1/S2切割位点结束,而术语“S2亚基”(例如,S2亚基抗原)是指刺突蛋白的C末端亚基,以S1/S2切割位点开始并且以刺突蛋白的C末端结束。如上文所述,本领域技术人员将理解,尽管基本上全长或完整刺突蛋白S1或S2亚基可分别为受体结合或膜融合所必需的,但当主要寻求引发针对刺突蛋白亚基的免疫反应时,可耐受S1或S2结构和/或序列的一定量的变化。例如,可耐受例如从编码的亚基(例如编码的S1或S2蛋白抗原)的N末端或C末端的一个至几个、可能至多4、5、6、7、8、9或10个氨基酸的少量截短,而不改变蛋白质的抗原性质。同样,可耐受编码的刺突蛋白亚基(例如编码的S1或S2蛋白抗原)的一个至几个、可能至多4、5、6、7、8、9或10个氨基酸(或更多个氨基酸)的变化(例如保守取代),而不改变蛋白质的抗原性质。在示例性实施方案中,刺突蛋白(例如编码的刺突蛋白抗原)具有表1-15的序列中的任一者中所阐释的氨基酸序列(例如,源自如SEQ ID NO:125所阐释的氨基酸序列)。在其他实施方案中,与包含具有如SEQ ID NO:125所阐释的氨基酸序列的刺突蛋白的氨基酸1-685或由氨基酸1-685组成的刺突蛋白S1亚基或包含氨基酸686-1273或由氨基酸686-1273组成的刺突蛋白S2亚基相比(当与其比对时),刺突蛋白亚基(例如编码的S1或S2蛋白抗原)具有不大于50、不大于40、不大于30、不大于20、不大于10或不大于5个氨基酸取代和/或缺失。当在编码的刺突蛋白亚基序列中发生微小变化时,变体优选具有与参考刺突蛋白亚基序列相同的活性和/或具有与参考刺突蛋白亚基相同的免疫特异性,如例如在免疫测定(例如酶联免疫吸附测定(ELISA测定))中所确定。

[0177] SARS-CoV-2刺突蛋白的S1和S2亚基还包括通过结构和功能容易辨别的结构域,其特征又在于设计由核酸疫苗(特别是本发明的mRNA疫苗)编码的抗原。在S1亚基内,结构域包括N末端结构域(NTD)和受体结合结构域(RBD),所述RBD结构域还包括受体结合基序(RBM)。野生型S1亚基还包括在NTD结构域的N末端的信号肽(SD)以及第一子结构域(SD1)和第二子结构域(SD2)。在S2亚基内,结构域包括融合肽(FP)、七肽重复区1(HR1)、七肽重复区2(HR2)、跨膜结构域(TM)和细胞质结构域,也称为胞质尾区(CT)(Lu R.等人,同上;Wan等人,J.Virol.2020年3月,94(7)e00127-20)。HR1和HR2结构域可称为SARS-CoV-2的“融合核心区”(Xia等人,2020 Cell Mol Immunol.1月;17(1):1-12。)。图1绘示SARS-CoV-2刺突蛋白中的结构域架构。S1亚基包括N末端结构域(NTD)、接头区、受体结合结构域(RBD)、第一子结构域(SD1)和第二子结构域(SD2)。S1亚基可经修饰以添加C末端跨膜结构域(TM),或者其可为可溶的。S2亚基尤其包括第一七肽重复区(HR1)、第二七肽重复区(HR2)、跨膜结构域(TM)和胞质尾区。可溶性S2亚基可在没有TM结构域的情况下产生。

[0178] S1的NTD和RBD是用于本发明的疫苗设计方法的良好抗原,因为所述结构域已显示为β冠状病毒感染个体中的中和抗体的靶标。如本文所用,例如,在抗原设计(所述抗原由本发明的mRNA编码并且例如由本发明的mRNA疫苗表达)的情形下,术语“N末端结构域”或“NTD”是指SARS-CoV-2 S1亚基内的结构域,其包含约290个氨基酸长度,与具有如SEQ ID NO:125所阐释的氨基酸序列的刺突蛋白的S1亚基的氨基酸1-290具有同一性。如本文所用,例如,在抗原设计(所述抗原由本发明的mRNA编码并且例如由本发明的mRNA疫苗表达)的情形下,术语“受体结合结构域”或“RBD”是指SARS-CoV-2的S1亚基内的结构域,其包含约175-

225个氨基酸长度，与具有如SEQ ID NO:125所阐释的氨基酸序列的刺突蛋白的S1亚基的氨基酸316-517具有同一性。如本文所用的术语“受体结合基序”是指RBD中直接接触ACE2受体的部分。预测表达的RBD特异性结合至血管收缩肽转化酶2(ACE2)作为其受体和/或与RBD结合和/或中和抗体(例如CR3022)特异性反应。

[0179] 本文提供的组合物包括mRNA，其可编码任何一个或多个全长或部分(截短或其他序列缺失)S蛋白亚基(例如，S1或S2亚基)、S蛋白亚基的一个或多个结构域或结构域的组合(例如，NTD、RBD或NTD-RBD融合体，有或无SD1和/或SD2)或全长或部分S1和S2蛋白亚基的嵌合体。本文考虑其他S蛋白亚基和/或结构域配置。

[0180] 图2和图6绘示源自SARS-CoV-2刺突蛋白的示例性结构域和亚基抗原。图2A和图2B分别绘示可溶性和跨膜RBD抗原。跨膜NTD抗原示于图2C中。图2D-图2F和图2I中所示的结构域抗原表示NTD和RBD的示例性融合蛋白，每种具有SP和TM结构域。构建体中的两者还具有末端运输结构域(CD86和/或CD11b)。所述结构域经由接头、特别是GS接头或PADRE接头(图2I)连接。具有在NTD结构域的N末端的RBD结构域的结构域构建体绘示于图2G和图2H中。每个构建体还可包括SP和/或TM结构域。

[0181] 编码的亚基抗原

[0182] 本公开的一些方面提供了包含编码SARS-CoV-2 S蛋白的(至少一个)亚基的mRNA的组合物。在一些实施方案中，mRNA编码S1亚基(例如全长或部分)。在其他实施方案中，mRNA编码S2亚基(例如全长或部分)。在其他实施方案中，mRNA编码嵌合S1-S2蛋白，其中一个亚基来自SARS-CoV-2 S蛋白，而另一个亚基来自另一生物体，例如病毒，如流感病毒。由本公开的mRNA编码的SARS-CoV-2亚基(S1和/或S2)可为可溶的或膜结合的(例如，与跨膜结构域连接)。基于S2的示例性抗原设计示于图6中。图6A绘示全长S2，包括FP、HR1、HR2、TM和CT结构域。由亚基之间的接头构成的一种形式的S2示于图6B中。无CT结构域的结构域抗原示于图6C和图6D中。

[0183] 可溶性亚基抗原

[0184] 可溶性蛋白存在于细胞的细胞质中或从细胞分泌(例如，不结合膜)。细胞分泌的可溶性抗原可由补体调理，并且由淋巴结中的滤泡树突细胞捕获，在这里它们可由对表达的蛋白质上存在的表位特异性的B细胞识别。亚基抗原的表达进一步允许将免疫反应集中于特定亚基，并且对与其他相关病毒共有的抗原的其他结构域具有特异性的记忆B和T细胞的刺激最小。不受限于理论，据认为SARS-CoV-2 S1亚基(包括NTD、RBD和在一些情况下SARS-CoV-2 S1亚基的插入多肽)以可溶形式的呈递，产生S1亚基特异性免疫反应。因此，在一些实施方案中，本文提供的mRNA编码可溶性SARS-CoV-2 S1亚基抗原和/或可溶性SARS-CoV-2 S2亚基抗原。可溶性SARS-CoV-2 S1亚基抗原和编码其的mRNA的非限制性实例提供于下文表1A和表1B中。本文提供了可溶性SARS-CoV-2亚基抗原的其他实例。

[0185] 表1A.可溶性亚基抗原

[0186]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
SARS-CoV-2 可溶性 S1 亚基	3	5

[0187] 表1B.可溶性亚基抗原

[0188]

SARS-CoV-2 可溶性 S1 亚基		
SEQ ID NO: 1 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 3 和 3' UTR SEQ ID NO: 4。		1
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAG AAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCC UGGUGAGCAGCCAGUGCGUGAACCUGAC CACCCGGACCCAGCUGCCACCAGCCUACA CCAACAGCUUCACCCGGGGCGUCUACUAC CCCGACAAGGUGUUCCGGAGCAGCGUCC UGCACAGCACCCAGGACCUGUUCCUGCCC UUCUUCAGCAACGUGACCUGGUUCCACG CCAUCCACGUGAGCGGCACCAACGGCACC AAGCGGUUCGACAACCCCGUGCUGCCCUU CAACGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGGAU CUUCGGCACCACCCUGGACAGCAAGACCC AGAGCCUGCUGAUCGUGAAUAACGCCAC CAACGUGGUGAUCAAGGUGUGCGAGUUC CAGUUCUGCAACGACCCCUUCCUGGGCGU GUACUACCACAAGAACAACAAGAGCUGG AUGGAGAGCGAGUUCCGGGUGUACAGCA GCGCCAACAACUGCACCUUCGAGUACGU GAGCCAGCCCUUCCUGAUGGACCUGGAG GGCAAGCAGGGCAACUUCAAGAACCUGC GGGAGUUCGUGUUCAAGAACAUCGACGG CUACUUCAAGAUCUACAGCAAGCACACCC CAAUCAACCUGGUGCGGGAUCUGCCCCA	3

[0189]

	GGGCUUCUCAGCCCUGGAGCCCCUGGUG GACCUGCCCAUCGGCAUCAACAUCACCCG GUUCCAGACCCUGCUGGCCCUGCACCGGA GCUACCUGACCCCAGGCGACAGCAGCAGC GGGUGGACAGCAGGCGCGGCUGCUUACU ACGUGGGCUACCUGCAGCCCCGGACCUUC CUGCUGAAGUACAACGAGAACGGCACCA UCACCGACGCCGUGGACUGCGCCCUGGAC CCUCUGAGCGAGACCAAGUGCACCCUGA AGAGCUUCACCGUGGAGAAGGGCAUCUA CCAGACCAGCAACUUCCGGGUGCAGCCCA CCGAGAGCAUCGUGCGGUUCCCCAACAUC ACCAACCUGUGCCCCUUCGGCGAGGUGU UCAACGCCACCCGGUUCGCCAGCGUGUAC GCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAA CAGCGCCAGCUUCAGCACCUUCAAGUGCU ACGGCGUGAGCCCCACCAAGCUGAACGAC CUGUGCUUCACCAACGUGUACGCCGACA GCUUCGUGAUCCGUGGCGACGAGGUGCG GCAGAUCGCACCCGGCCAGACAGGCAAG AUCGCCGACUACAACUACAAGCUGCCCGA CGACUUCACCGGCUGCGUGAUCGCCUGG AACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCU GUUCCGGAAGAGCAACCUGAAGCCCUUC GAGCGGGACAUCAGCACCGAGAUCUACC AAGCCGGCUCCACCCCUUGCAACGGCGUG GAGGGCUUCAACUGCUACUUCCCUCUGC AGAGCUACGGCUUCCAGCCCACCAACGGC GUGGGCUACCAGCCCUACCGGGUGGUGG UGCUGAGCUUCGAGCUGCUGCACGCCCCA GCCACCGUGUGUGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAA CUUCAACUUCAACGGCCUUACCGGCACCG GCGUGCUGACCGAGAGCAACAAGAAAUU CCUGCCCUUUCAGCAGUUCGGCCGGGACA UCGCCGACACCACCGACGCUGUGCGGGAU CCCCAGACCCUGGAGAUCCUGGACAUCAC CCCUUGCAGCUUCGGCGGCGUGAGCGUG AUCACCCCAGGCACCAACACCAGCAACCA GGUGGCCGUGCUGUACCAGGACGUGAAC UGCACCGAGGUGCCCGUGGCCAUCCACGC CGACCAGCUGACACCCACCUGGCGGGUCU ACAGCACCGGCAGCAACGUGUUCCAGACC CGGGCCGGUUGCCUGAUCGGCGCCGAGC	

[0190]

	ACGUGAACAACAGCUACGAGUGCGACAU CCCCAUCGGCGCCGGCAUCUGUGCCAGCU ACCAGACCCAGACCAAUUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU AGCUUCUUGCCCCUUGGGCCUCCCCCCAG CCCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGG GCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNS FTRGVYYPDKVFRSSVLHSTQDLFLPFFSNV TWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNAT NVVIKVCEFQFCNDPFLGVYYHKNNKSWM ESEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDL PQGFSALEPLVDLPIGINITRFQTLLALHRSYL TPGDSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDCALDPLSETKCTLKSFTVEK GIYQTSNFRVQPTESIVRFPNITNLCPFGEVFN ATRFASVYAWNRKRISNCVADYSVLYNSAS FSTFKCYGVSPTKLNDLCFTNVYADSFVIRG DEVRQIAPGQTGKIADYNYKLPDDFTGCVIA WNSNNLDSKVGGNYNYLYRLFRKSNLKPFE RDISTEIYQAGSTPCNGVEGFNCYFPLQSYGF QPTNGVGYQPYRVVVLSFELLHAPATVCGP KKSTNLVKNKCVNFNFNGLTGTGVLTESNK KFLPFQQFGRDIADTTDAVRDPQTLEILDITP CSFGGVSVITPGTNTSNQVAVLYQDVNCTE VPVAIHADQLTPTWRVYSTGSNVFQTRAGC LIGAEHVNNSYECDIPIGAGICASYQTQTNS	5
多聚 A 尾	100 nt	

[0191] 膜结合的亚基抗原

[0192] 膜结合蛋白锚定在细胞膜中(不可溶的)。不受限于理论,据认为抗原呈递细胞将嵌入的抗原携带至引流淋巴结以产生强烈免疫反应。引流淋巴结中出现的生发中心反应涉及$CD4^+T_{FH}$细胞与B细胞之间的延长接触,允许共刺激和局部细胞因子信号,例如IL-4和IL-21,其有利于对所呈递抗原具特异性的B细胞的复制和类别转换至IgG1的产生,其各自可促进长寿命的浆细胞和记忆B细胞的产生。因此,在一些实施方案中,mRNA编码膜结合的SARS-CoV-2 S1亚基抗原和/或膜结合的SARS-CoV-2 S2亚基抗原。在一些实施方案中,膜结合抗原(例如,S1亚基、S2亚基、NTD、RBD或其任何组合)与跨膜结构域(例如天然存在的跨膜结构域或异源跨膜结构域(源自异源蛋白质),其负责将蛋白质锚定在细胞膜中)连接。膜结合的SARS-CoV-2 S1亚基抗原和SARS-CoV-2 S2亚基抗原和编码其的mRNA的非限制性实例提供于下文表2A和表2B中。本文考虑其他膜结合的SARS-CoV-2 S1亚基抗原。

[0193] 表2A.膜结合的亚基抗原

[0194]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-666-TM)	16	17
与跨膜结构域连接的 SARS-CoV-2 S2 亚基(S2-TM)	145	146

[0195] 表2B.膜结合的亚基抗原

与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-666-TM)		
SEQ ID NO: 15 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 16 和 3' UTR SEQ ID NO: 4。		15
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC	2

[0196]

mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCCA	16

[0197]

	GCACCGAGAAGAGCAACAUCAUCCGGGG CUGGAUCUUCGGCACCACCCUGGACAGC AAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUG CGAGUUCCAGUUCUGCAACGACCCCUUC CUGGGCGUGUACUACCACAAGAACAACA AGAGCUGGAUGGAGAGCGAGUUCCGGG UGUACAGCAGCGCCAACAACUGCACCUU CGAGUACGUGAGCCAGCCCUUCCUGAUG GACCUGGAGGGCAAGCAGGGCAACUUCA AGAACCUGCGGGAGUUCGUGUUCAAGA ACAUCGACGGCUACUUCAAGAUCUACAG CAAGCACACCCCAAUCAACCUGGUGCGG GAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAU CAACAUCACCCGGUUCCAGACCCUGCUG GCCCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUG CAGCCCCGGACCUUCCUGCUGAAGUACA ACGAGAACGGCACCAUCACCGACGCCGU GGACUGCGCCCUGGACCCUCUGAGCGAG ACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAA CUUCCGGGUGCAGCCCACCGAGAGCAUC GUGCGGUUCCCCAACAUCACCAACCUGU GCCCCUUCGGCGAGGUGUUCAACGCCAC CCGGUUCGCCAGCGUGUACGCCUGGAAC CGGAAGCGGAUCAGCAACUGCGUGGCCG ACUACAGCGUGCUGUACAACAGCGCCAG CUUCAGCACCUUCAAGUGCUACGGCGUG AGCCCCACCAAGCUGAACGACCUGUGCU UCACCAACGUGUACGCCGACAGCUUCGU GAUCCGUGGCGACGAGGUGCGGCAGAUC GCACCCGGCCAGACAGGCAAGAUCGCCG ACUACAACUACAAGCUGCCCGACGACUU CACCGGCUGCGUGAUCGCCUGGAACAGC AACAACCUCGACAGCAAGGUGGGCGGCA ACUACAACUACCUGUACCGGCUGUUCCG GAAGAGCAACCUGAAGCCCUUCGAGCGG GACAUCAGCACCGAGAUCUACCAAGCCG GCUCCACCCCUUGCAACGGCGUGGAGGG CUUCAACUGCUACUUCCCUCUGCAGAGC	

[0198]

	UACGGCUUCCAGCCCACCAACGGCGUGG GCUACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAGCC ACCGUGUGUGGCCCCAAGAAGAGCACCA ACCUGGUGAAGAACAAGUGCGUGAACU UCAACUUCAACGGCCUUACCGGCACCGG CGUGCUGACCGAGAGCAACAAGAAAUUC CUGCCCUUUCAGCAGUUCGGCCGGGACA UCGCCGACACCACCGACGCUGUGCGGGA UCCCCAGACCCUGGAGAUCCUGGACAUC ACCCCUUGCAGCUUCGGCGGCGUGAGCG UGAUCACCCCAGGCACCAACACCAGCAA CCAGGUGGCCGUGCUGUACCAGGACGUG AACUGCACCGAGGUGCCCGUGGCCAUCC ACGCCGACCAGCUGACACCCACCUGGCG GGUCUACAGCACCGGCAGCAACGUGUUC CAGACCCGGGCCGGUUGCCUGAUCGGCG CCGAGCACGUGAACAACAGCUACGAGUG CGACAUCCCCAUCGGCGCCGGCAUCUGU GCCAGCUACCAGACCCAGACCAAUUCAU CUGGCGGAGGCAGCAUCCUGGCCAUCUA CAGCACCGUGGCCAGCAGCCUGGUGCUG CUGGUGAGCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU AGCUUCUUGCCCCUUGGGCCUCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTN SFTRGVYYPDKVFRSSVLHSTQDLFLPFFSN VTWFHAIHVSGTNGTKRFDNPVLPFNDGV YFASTEKSNIIRGWIFGTTLDSKTQSLLIVNN ATNVVIKVCEFQFCNDPFLGVYYHKNNKS WMESEFRVYSSANNCTFEYVSQPFLMDLE GKQGNFKNLREFVFKNIDGYFKIYSKHTPIN LVRDLPQGFSALEPLVDLPIGINITRFQTLLA LHRSYLTPGDSSSGWTAGAAAYYVGYLQP RTFLLKYNENGTITDAVDCALDPLSETKCT LKSFTVEKGIYQTSNFRVQPTESIVRFPNITN LCPFGEVFNATRFASVYAWNRKRISNCVAD YSVLYNSASFSTFKCYGVSPTKLNDLCFTN VYADSFVIRGDEVRQIAPGQTGKIADYNYK LPDDFTGCVIAWNSNNLDSKVGGNYNYLY	17

	RLFRKSNLKPFERDISTEIYQAGSTPCNGVE GFNCYFPLQSYGFQPTNGVGYQPYRVVVLS FELLHAPATVCGPKKSTNLVKNKCVNFNFN GLTGTGVLTESNKKFLPFQQFGRDIADTTD AVRDPQTLEILDITPCSFGGVSVITPGTNTSN QVAVLYQDVNCTEVPVAIHADQLTPTWRV YSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSGGGSILAIYSTVASSLV LLVSLGAISF	
多聚 A 尾	100 nt	
包含跨膜结构域的 SARS-CoV-2 S2 亚基(S2-TM)		
SEQ ID NO: 147 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 145 和 3' UTR SEQ ID NO: 4。		147
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAA GAAGAAAUAUAAGACCCCGGCGCCGCCA CC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCUGCG UGACCCUGACCCUGGUGCUGCUGGUGAA CAGCCAGGGCGCCGAGAACAGCGUGGCC UACAGCAACAACAGCAUCGCCAUCCCCA CCAACUUCACCAUCAGCGUGACCACCGA GAUUCUGCCCGUGAGCAUGACCAAGACC AGCGUGGACUGCACCAUGUACAUCUGCG GCGACAGCACCGAGUGCAGCAACCUGCU GCUGCAGUACGGCAGCUUCUGCACCCAG CUGAACCGGGCCCUGACCGGCAUCGCCG UGGAGCAGGACAAGAACACCCAGGAGGU GUUCGCCCAGGUGAAGCAGAUCUACAAG ACCCCUCCCAUCAAGGACUUCGGCGGCU UCAACUUCAGCCAGAUCCUGCCCGACCC CAGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACCC UAGCCGACGCCGGCUUCAUCAAGCAGUA CGGCGACUGCCUCGGCGACAUAGCCGCC CGGGACCUGAUCUGCGCCCAGAAGUUCA ACGGCCUGACCGUGCUGCCUCCCCUGCU GACCGACGAGAUGAUCGCCCAGUACACC AGCGCCCUGUUAGCCGGAACCAUCACCA GCGGCUGGACUUUCGGCGCUGGAGCCGC	145

[0199]

[0200]

	UCUGCAGAUCCCCUUCGCCAUGCAGAUG GCCUACCGGUUCAACGGCAUCGGCGUGA CCCAGAACGUGCUGUACGAGAACCAGAA GCUGAUCGCCAACCAGUUCAACAGCGCC AUCGGCAAGAUCCAGGACAGCCUGAGCA GCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCCAGGCC CUGAACACCCUGGUGAAGCAGCUGAGCA GCAACUUCGGCGCCAUCAGCAGCGUGCU GAACGACAUCCUGAGCCGGCUGGACCCU CCCGAGGCCGAGGUGCAGAUCGACCGGC UGAUCACUGGCCGGCUGCAGAGCCUGCA GACCUACGUGACCCAGCAGCUGAUCCGG GCCGCCGAGAUUCGGGCCAGCGCCAACC UGGCCGCCACCAAGAUGAGCGAGUGCGU GCUGGGCCAGAGCAAGCGGGUGGACUUC UGCGGCAAGGGCUACCACCUGAUGAGCU UUCCCCAGAGCGCACCCCACGGAGUGGU GUUCCUGCACGUGACCUACGUGCCCGCC CAGGAGAAGAACUUCACCACCGCCCCAG CCAUCUGCCACGACGGCAAGGCCCACUU UCCCCGGGAGGGCGUGUUCGUGAGCAAC GGCACCCACUGGUUCGUGACCCAGCGGA ACUUCUACGAGCCCCAGAUCAUCACCAC CGACAACACCUUCGUGAGCGGCAACUGC GACGUGGUGAUCGGCAUCGUGAACAACA CCGUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGACAA GUACUUCAAGAAUCACACCAGCCCCGAC GUGGACCUGGGCGACAUCAGCGGCAUCA ACGCCAGCGUGGUGAACAUCCAGAAGGA GAUCGAUCGGCUGAACGAGGUGGCCAAG AACCUGAACGAGAGCCUGAUCGACCUGC AGGAGCUGGGCAAGUACGAGCAGUACA UCAAGUGGCCCUGGUACAUCUGGCUGGG CUUCAUCGCCGGCCUGAUCGCCAUCGUG AUGGUGACCAUCAUGCUGUGCUGCAUGA CCAGCUGCUGCAGCUGCCUGAAGGGCUG UUGCAGCUGCGGCAGCUGCUGCAAGUUC GACGAGGACGACAGCGAGCCCGUGCUGA AGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU AGCUUCUUGCCCCUUGGGCCUCCCCCCA	4

[0201]

	GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGC	
相应的氨基酸序列	MYSMQLASCVTLTLVLLVNSQGAENSVAY SNNSIAIPTNFTISVTTEILPVSMTKTSVDCT MYICGDSTECSNLLLQYGSFCTQLNRALTGI AVEQDKNTQEVFAQVKQIYKTPPIKDFGGF NFSQILPDPSKPSKRSFIEDLLFNKVTLADA GFIKQYGDCLGDIAARDLICAQKFNGLTVL PPLLTDEMIAQYTSALLAGTITSGWTFGAG AALQIPFAMQMAYRFNGIGVTQNVLYENQ KLIANQFNSAIGKIQDSLSSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDIL SRLDPPEAEVQIDRLITGRLQSLQTYVTQQL IRAAEIRASANLAATKMSECVLGQSKRVDF CGKGYHLMSFPQSAPHGVVFLHVTYVPAQ EKNFTTAPAICHDGKAHFPREGVFVSNGTH WFVTQRNFYEPQIITTDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSP DVDLGDISGINASVVNIQKEIDRLNEVAKNL NESLIDLQELGKYEQYIKWPWYIWLGFIAG LIAIVMVTIMLCCMTSCCSCLKGCCSCGSC CKFDEDDSEPVLKGVKLHYT	146
多聚A尾	100 nt	

[0202] 亚基抗原截短和RBD缺失

[0203] 在一些实施方案中，组合物包含编码已经修饰以去除RBD或RBD的一部分的S1亚基的mRNA。S1亚基的截短为免疫系统提供较少的表位来识别，从而使免疫反应偏向于其余表位，这可选择针对对于病毒中和重要的特定表位的抗体。RBD的截短或部分缺失可防止表达的蛋白质或携带其的细胞与受体ACE2相互作用，使其更可能到达淋巴结并刺激所需免疫反应。此外，去除RBD可防止以前针对相关病毒产生的交叉反应性抗体造成的表位掩蔽，并因此将引发的免疫反应特异性地集中于所需抗原。另外，去除RBD可改变所表达的亚基的构象，允许对这些替代构象表位具特异性的B细胞摄取线性肽并且将线性肽呈递给T细胞，从而间接增强$CD4^+$T细胞对仍以天然构象存在的那些表位的反应。

[0204] 在一些实施方案中，组合物包含编码S1亚基的mRNA，所述S1亚基已被修饰以去除RBD或RBD的一部分，其中S2亚基含有聚糖。聚糖通过经由天冬酰胺残基的N-连接糖基化或丝氨酸或苏氨酸残基上的O-连接糖基化与蛋白质附接。在蛋白质的一些组分上存在聚糖屏蔽物可掩蔽肽表位，从而将抗体反应集中于其他暴露的肽表位。此外，糖基化蛋白还引发识别包被聚糖的抗体。识别聚糖表位的B细胞将摄取线性肽表位并将其呈递给$CD4^+$T细胞，从而加强$CD4^+$T细胞对整个蛋白中发现的线性表位的反应。

[0205] 截短的SARS-CoV-2 S1亚基抗原和编码其的mRNA的非限制性实例提供于下文表3A和表3B中。

[0206] 具有RBD缺失的SARS-CoV-2 S1亚基和编码其的mRNA的非限制性实例提供于下文

表4A和表4B中。

[0207] 表3A.亚基抗原截短

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
经截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-531-TM)	19	20
经截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-594-TM)	22	23
经 PolyG 截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-594-PolyG-TM)	25	26
经 PolyG/DS 截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-594-PolyG-DS-TM)	28	29
经截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-666-TM)	150	151

[0208]

[0209] 表3B.亚基抗原截短

经截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-531-TM)	
SEQ ID NO: 18 从 5'端至 3'端由以下组成:5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 19 和 3' UTR SEQ ID NO: 4。	18

[0210]

[0211]

化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAA GAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA构建体的ORF(不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUG GUGAGCAGCCAGUGCGUGAACCUGACCACC CGGACCCAGCUGCCACCAGCCUACACCAACA GCUUCACCCGGGGCGUCUACUACCCCGACAA GGUGUUCCGGAGCAGCGUCCUGCACAGCAC CCAGGACCUGUUCCUGCCCUUCUUCAGCAAC GUGACCUGGUUCCACGCCAUCCACGUGAGC GGCACCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGUACU UCGCCAGCACCGAGAAGAGCAACAUCAUCC GGGGCUGGAUCUUCGGCACCACCCUGGACA GCAAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUGCG AGUUCCAGUUCUGCAACGACCCCUUCCUGG GCGUGUACUACCACAAGAACAACAAGAGCU GGAUGGAGAGCGAGUUCCGGGUGUACAGCA GCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGGGAGU UCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCU GGUGCGGGAUCUGCCCCAGGGCUUCUCAGC CCUGGAGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCUGCUGG CCCUGCACCGGAGCUACCUGACCCCAGGCGA CAGCAGCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAGCCCCG GACCUUCCUGCUGAAGUACAACGAGAACGG CACCAUCACCGACGCCGUGGACUGCGCCCUG GACCCUCUGAGCGAGACCAAGUGCACCCUG AAGAGCUUCACCGUGGAGAAGGGCAUCUAC CAGACCAGCAACUUCCGGGUGCAGCCCACCG AGAGCAUCGUGCGGUUCCCCAACAUCACCA ACCUGUGCCCCUUCGGCGAGGUGUUCAACG CCACCCGGUUCGCCAGCGUGUACGCCUGGAA CCGGAAGCGGAUCAGCAACUGCGUGGCCGA CUACAGCGUGCUGUACAACAGCGCCAGCUU CAGCACCUUCAAGUGCUACGGCGUGAGCCCC	19

[0212]

	ACCAAGCUGAACGACCUGUGCUUCACCAAC GUGUACGCCGACAGCUUCGUGAUCCGUGGC GACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCG CCUGGAACAGCAACAACCUCGACAGCAAGG UGGGCGGCAACUACAACUACCUGUACCGGC UGUUCCGGAAGAGCAACCUGAAGCCCUUCG AGCGGGACAUCAGCACCGAGAUCUACCAAG CCGGCUCCACCCCUUGCAACGGCGUGGAGGG CUUCAACUGCUACUUCCCUCUGCAGAGCUAC GGCUUCCAGCCCACCAACGGCGUGGGCUACC AGCCCUACCGGGUGGUGGUGCUGAGCUUCG AGCUGCUGCACGCCCCAGCCACCGUGUGUGG CCCCAAGAAGAGCACCUCUGGCGGAGGCAG CAUCCUGGCCAUCUACAGCACCGUGGCCAGC AGCCUGGUGCUGCUGGUGAGCCUGGGCGCC AUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAG CUUCUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFT RGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFH AIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEF QFCNDPFLGVYYHKNNKSWMESEFRVYSSANN CTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGI NITRFQTLLALHRSYLTPGDSSSGWTAGAAAYY VGYLQPRTFLLKYNENGTITDAVDCALDPLSET KCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN LCPFGEVFNATRFASVYAWNRKRISNCVADYS VLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS FVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGC VIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPF ERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQ PTNGVGYQPYRVVVLSFELLHAPATVCGPKKS Tsgggsilaiystvasslvllvslgaisf	20
多聚A尾	100 nt	
经截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-594-TM)		

[0213]

SEQ ID NO: 21 从5'端至3'端由以下组成:5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 22 和 3' UTR SEQ ID NO: 4。		21
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAA GAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA构建体的ORF(不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUG GUGAGCAGCCAGUGCGUGAACCUGACCACC CGGACCCAGCUGCCACCAGCCUACACCAACA GCUUCACCCGGGGCGUCUACUACCCCGACAA GGUGUUCCGGAGCAGCGUCCUGCACAGCAC CCAGGACCUGUUCCUGCCCUUCUUCAGCAAC GUGACCUGGUUCCACGCCAUCCACGUGAGC GGCACCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGUACU UCGCCAGCACCGAGAAGAGCAACAUCAUCC GGGGCUGGAUCUUCGGCACCACCCUGGACA GCAAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUGCG AGUUCCAGUUCUGCAACGACCCCUUCCUGG GCGUGUACUACCACAAGAACAACAAGAGCU GGAUGGAGAGCGAGUUCCGGGUGUACAGCA GCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGGGAGU UCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCU GGUGCGGGAUCUGCCCCAGGGCUUCUCAGC CCUGGAGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCUGCUGG CCCUGCACCGGAGCUACCUGACCCCAGGCGA CAGCAGCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAGCCCCG GACCUUCCUGCUGAAGUACAACGAGAACGG CACCAUCACCGACGCCGUGGACUGCGCCCUG GACCCUCUGAGCGAGACCAAGUGCACCCUG AAGAGCUUCACCGUGGAGAAGGGCAUCUAC CAGACCAGCAACUUCCGGGUGCAGCCCACCG AGAGCAUCGUGCGGUUCCCCAACAUCACCA ACCUGUGCCCCUUCGGCGAGGUGUUCAACG CCACCCGGUUCGCCAGCGUGUACGCCUGGAA CCGGAAGCGGAUCAGCAACUGCGUGGCCGA	22

[0214]

	CUACAGCGUGCUGUACAACAGCGCCAGCUU CAGCACCUUCAAGUGCUACGGCGUGAGCCCC ACCAAGCUGAACGACCUGUGCUUCACCAAC GUGUACGCCGACAGCUUCGUGAUCCGUGGC GACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCG CCUGGAACAGCAACAACCUCGACAGCAAGG UGGGCGGCAACUACAACUACCUGUACCGGC UGUUCCGGAAGAGCAACCUGAAGCCCUUCG AGCGGGACAUCAGCACCGAGAUCUACCAAG CCGGCUCCACCCCUUGCAACGGCGUGGAGGG CUUCAACUGCUACUUCCCUCUGCAGAGCUAC GGCUUCCAGCCCACCAACGGCGUGGGCUACC AGCCCUACCGGGUGGUGGUGCUGAGCUUCG AGCUGCUGCACGCCCCAGCCACCGUGUGUGG CCCCAAGAAGAGCACCAACCUGGUGAAGAA CAAGUGCGUGAACUUCAACUUCAACGGCCU UACCGGCACCGGCGUGCUGACCGAGAGCAA CAAGAAAUUCCUGCCCUUUCAGCAGUUCGG CCGGGACAUCGCCGACACCACCGACGCUGUG CGGGAUCCCCAGACCCUGGAGAUCCUGGAC AUCACCCCUUGCAGCUUCGGCGGCUCUGGCG GAGGCAGCAUCCUGGCCAUCUACAGCACCG UGGCCAGCAGCCUGGUGCUGCUGGUGAGCC UGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAG CUUCUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFT RGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFH AIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEF QFCNDPFLGVYYHKNNKSWMESEFRVYSSANN CTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGI NITRFQTLLALHRSYLTPGDSSSGWTAGAAAYY VGYLQPRTFLLKYNENGTITDAVDCALDPLSET KCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN LCPFGEVFNATRFASVYAWNRKRISNCVADYS VLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS FVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGC	23

[0215]

	VIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPF ERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQ PTNGVGYQPYRVVVLSFELLHAPATVCGPKKS TNLVKNKCVNFNFNGLTGTGVLTESNKKFLPF QQFGRDIADTTDAVRDPQTLEILDITPCSFGGsgg gsilaiystvasslvllvslgaisf	
多聚 A 尾	100 nt	
经 PolyG 截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基 (S1-594-PolyG-TM)		
SEQ ID NO: 24 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 25 和 3' UTR SEQ ID NO: 4。		24
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAA GAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUG GUGAGCAGCCAGUGCGUGAACCUGACCACC CGGACCCAGCUGCCACCAGCCUACACCAACA GCUUCACCCGGGGCGUCUACUACCCCGACAA GGUGUUCCGGAGCAGCGUCCUGCACAGCAC CCAGGACCUGUUCCUGCCCUUCUUCAGCAAC GUGACCUGGUUCCACGCCAUCCACGUGAGC GGCACCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGUACU UCGCCAGCACCGAGAAGAGCAACAUCAUCC GGGGCUGGAUCUUCGGCACCACCCUGGACA GCAAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUGCG AGUUCCAGUUCUGCAACGACCCCUUCCUGG GCGUGUACUACCACAAGAACAACAAGAGCU GGAUGGAGAGCGAGUUCCGGGUGUACAGCA GCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGGGAGU UCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCU GGUGCGGGAUCUGCCCCAGGGCUUCUCAGC CCUGGAGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCUGCUGG CCCUGCACCGGAGCUACCUGACCCCAGGCGA CAGCAGCAGCGGGUGGACAGCAGGCGCGGC	25

[0216]

	UGCUUACUACGUGGGCUACCUGCAGCCCCG GACCUUCCUGCUGAAGUACAACGAGAACGG CACCAUCACCGACGCCGUGGACUGCGCCCUG GACCCUCUGAGCGAGACCAAGUGCACCCUG AAGAGCUUCACCGUGGGCAGCGGCGGCGGC AGCGGCGGAGGCAGCGGAGGAGGCAGCGGC GGAGGCAGUGGAGGCCAGCCCACCGAGAGC AUCGUGCGGUUCCCCAACAUCACCAACCUGU GCCCCUUCGGCGAGGUGUUCAACGCCACCCG GUUCGCCAGCGUGUACGCCUGGAACCGGAA GCGGAUCAGCAACUGCGUGGCCGACUACAG CGUGCUGUACAACAGCGCCAGCUUCAGCACC UUCAAGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGUGUAC GCCGACAGCUUCGUGAUCCGUGGCGACGAG GUGCGGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGCUGCCC GACGACUUCACCGGCUGCGUGAUCGCCUGG AACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGG GACAUCAGCACCGAGAUCUACCAAGCCGGC UCCACCCCUUGCAACGGCGUGGAGGGCUUC AACUGCUACUUCCCUCUGCAGAGCUACGGC UUCCAGCCCACCAACGGCGUGGGCUACCAGC CCUACCGGGUGGUGGUGCUGAGCUUCGAGC UGCUGCACGCCCCAGCCACCGUGUGUGGCCC CAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUAC CGGCACCGGCGUGCUGACCGAGAGCAACAA GAAAUUCCUGCCCUUUCAGCAGUUCGGCCG GGACAUCGCCGACACCACCGACGCUGUGCGG GAUCCCCAGACCCUGGAGAUCCUGGACAUC ACCCCUUGCAGCUUCGGCGGCUCUGGCGGA GGCAGCAUCCUGGCCAUCUACAGCACCGUG GCCAGCAGCCUGGUGCUGCUGGUGAGCCUG GGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAG CUUCUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
相应的	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFT RGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFH	26

氨基酸序列	AIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEF QFCNDPFLGVYYHKNNKSWMESEFRVYSSANN CTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGI NITRFQTLLALHRSYLTPGDSSSGWTAGAAAYY VGYLQPRTFLLKYNENGTITDAVDCALDPLSET KCTLKSFTVGSGGGSGGGSGGGSGGGSGGQPT ESIVRFPNITNLCPFGEVFNATRFASVYAWNRK RISNCVADYSVLYNSASFSTFKCYGVSPTKLND LCFTNVYADSFVIRGDEVRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYR LFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNC YFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAP ATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLT ESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDI TPCSFGGsgggsilaiystvasslvllvslgaisf	
多聚A尾	100 nt	

[0217]

经 PolyG/DS 截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基 (S1-594-PolyG-DS-TM)		
SEQ ID NO: 27 从5'端至3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 28 和 3' UTR SEQ ID NO: 4。		27
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAA GAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA构建体的ORF(不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUG GUGAGCAGCCAGUGCGUGAACCUGACCACC CGGACCCAGCUGCCACCAGCCUACACCAACA GCUUCACCCGGGGCGUCUACUACCCCGACAA GGUGUGCCGGAGCAGCGUCCUGCACAGCAC CCAGGACCUGUUCCUGCCCUUCUUCAGCAAC GUGACCUGGUUCCACGCCAUCCACGUGAGC GGCACCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGUACU UCGCCAGCACCGAGAAGAGCAACAUCAUCC GGGGCUGGAUCUUCGGCACCACCCUGGACA GCAAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUGCG AGUUCCAGUUCUGCAACGACCCCUUCCUGG GCGUGUACUACCACAAGAACAACAAGAGCU	28

[0218]

	GGAUGGAGAGCGAGUUCCGGGUGUACAGCA GCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGGGAGU UCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCU GGUGCGGGAUCUGCCCCAGGGCUUCUCAGC CCUGGAGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCUGCUGG CCCUGCACCGGAGCUACCUGACCCCAGGCGA CAGCAGCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAGCCCCG GACCUUCCUGCUGAAGUACAACGAGAACGG CACCAUCACCGACGCCGUGGACUGCGCCCUG GACCCUCUGAGCGAGACCAAGUGCACCCUG AAGAGCUUCACCGUGGGCAGCGGCGGCGGC AGCGGCGGAGGCAGCGGAGGAGGCAGCGGC GGAGGCAGUGGAGGCCAGCCCACCGAGAGC AUCGUGCGGUUCCCCAACAUCACCAACCUGU GCCCCUUCGGCGAGGUGUUCAACGCCACCCG GUUCGCCAGCGUGUACGCCUGGAACCGGAA GCGGAUCAGCAACUGCGUGGCCGACUACAG CGUGCUGUACAACAGCGCCAGCUUCAGCACC UUCAAGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGUGUAC GCCGACAGCUUCGUGAUCCGUGGCGACGAG GUGCGGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGCUGCCC GACGACUUCACCGGCUGCGUGAUCGCCUGG AACAGCAACAACCUCGACAGCAAGGUGGGC GGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGG GACAUCAGCACCGAGAUCUACCAAGCCGGC UCCACCCCUUGCAACGGCGUGGAGGGCUUC AACUGCUACUUCCCUCUGCAGAGCUACGGC UUCCAGCCCACCAACGGCGUGGGCUACCAGC CCUACCGGGUGGUGGUGCUGAGCUUCGAGC UGCUGCACGCCCCAGCCACCGUGUGUGGCCC CAAGAAGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGCCUUAC CGGCACCGGCGUGCUGACCGAGAGCAACAA GAAAUUCCUGCCCUUUUGCCAGUUCGGCCG GGACAUCGCCGACACCACCGACGCUGUGCGG	

[0219]

	GAUCCCCAGACCCUGGAGAUCCUGGACAUC ACCCCUUGCAGCUUCGGCGGCUCUGGCGGA GGCAGCAUCCUGGCCAUCUACAGCACCGUG GCCAGCAGCCUGGUGCUGCUGGUGAGCCUG GGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAG CUUCUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFT RGVYYPDKVCRSSVLHSTQDLFLPFFSNVTWFH AIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEF QFCNDPFLGVYYHKNNKSWMESEFRVYSSANN CTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGI NITRFQTLLALHRSYLTPGDSSSGWTAGAAAYY VGYLQPRTFLLKYNENGTITDAVDCALDPLSET KCTLKSFTVGSGGGSGGGSGGGSGGGSGGQPT ESIVRFPNITNLCPFGEVFNATRFASVYAWNRK RISNCVADYSVLYNSASFSTFKCYGVSPTKLND LCFTNVYADSFVIRGDEVRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYR LFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNC YFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAP ATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLT ESNKKFLPFCQFGRDIADTTDAVRDPQTLEILDI TPCSFGGsgggsilaiystvasslvllvslgaisf	29
多聚A尾	100 nt	
经截短并与跨膜结构域连接的 SARS-CoV-2 S1 亚基(S1-666-TM)		
SEQ ID NO: 149 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 150 和 3' UTR SEQ ID NO: 4。		149
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAA GAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA构建体的ORF	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUG GUGAGCAGCCAGUGCGUGAACCUGACCACC CGGACCCAGCUGCCACCAGCCUACACCAACA	150

[0220]

(不包括终止密码子)	GCUUCACCCGGGGCGUCUACUACCCCGACAA GGUGUUCCGGAGCAGCGUCCUGCACAGCAC CCAGGACCUGUUCCUGCCCUUCUUCAGCAAC GUGACCUGGUUCCACGCCAUCCACGUGAGC GGCACCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGUACU UCGCCAGCACCGAGAAGAGCAACAUCAUCC GGGGCUGGAUCUUCGGCACCACCCUGGACA GCAAGACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUGUGCG AGUUCCAGUUCUGCAACGACCCCUUCCUGG GCGUGUACUACCACAAGAACAACAAGAGCU GGAUGGAGAGCGAGUUCCGGGUGUACAGCA GCGCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGGGAGU UCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCU GGUGCGGGAUCUGCCCCAGGGCUUCUCAGC CCUGGAGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCUGCUGG CCCUGCACCGGAGCUACCUGACCCCAGGCGA CAGCAGCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAGCCCCG GACCUUCCUGCUGAAGUACAACGAGAACGG CACCAUCACCGACGCCGUGGACUGCGCCCUG GACCCUCUGAGCGAGACCAAGUGCACCCUG AAGAGCUUCACCGUGGAGAAGGGCAUCUAC CAGACCAGCAACUUCCGGGUGCAGCCCACCG AGAGCAUCGUGCGGUUCCCCAACAUCACCA ACCUGUGCCCCUUCGGCGAGGUGUUCAACG CCACCCGGUUCGCCAGCGUGUACGCCUGGAA CCGGAAGCGGAUCAGCAACUGCGUGGCCGA CUACAGCGUGCUGUACAACAGCGCCAGCUU CAGCACCUUCAAGUGCUACGGCGUGAGCCCC ACCAAGCUGAACGACCUGUGCUUCACCAAC GUGUACGCCGACAGCUUCGUGAUCCGUGGC GACGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCG CCUGGAACAGCAACAACCUCGACAGCAAGG UGGGCGGCAACUACAACUACCUGUACCGGC UGUUCCGGAAGAGCAACCUGAAGCCCUUCG	

[0221]

	AGCGGGACAUCAGCACCGAGAUCUACCAAG CCGGCUCCACCCCUUGCAACGGCGUGGAGGG CUUCAACUGCUACUUCCCUCUGCAGAGCUAC GGCUUCCAGCCCACCAACGGCGUGGGCUACC AGCCCUACCGGGUGGUGGUGCUGAGCUUCG AGCUGCUGCACGCCCCAGCCACCGUGUGUGG CCCCAAGAAGAGCACCAACCUGGUGAAGAA CAAGUGCGUGAACUUCAACUUCAACGGCCU UACCGGCACCGGCGUGCUGACCGAGAGCAA CAAGAAAUUCCUGCCCUUUCAGCAGUUCGG CCGGGACAUCGCCGACACCACCGACGCUGUG CGGGAUCCCCAGACCCUGGAGAUCCUGGAC AUCACCCCUUGCAGCUUCGGCGGCGUGAGC GUGAUCACCCCAGGCACCAACACCAGCAACC AGGUGGCCGUGCUGUACCAGGACGUGAACU GCACCGAGGUGCCCGUGGCCAUCCACGCCGA CCAGCUGACACCCACCUGGCGGGUCUACAGC ACCGGCAGCAACGUGUUCCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCGGCGC CGGCAUCUGUGCCAGCUACCAGACCCAGACC AAUUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAG CUUCUUGCCCCUUGGGCCUCCCCCCAGCCCC UCCUCCCCUUCCUGCACCCGUACCCCCGUGG UCUUUGAAUAAAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFT RGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFH AIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEF QFCNDPFLGVYYHKNNKSWMESEFRVYSSANN CTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGI NITRFQTLLALHRSYLTPGDSSSGWTAGAAAYY VGYLQPRTFLLKYNENGTITDAVDCALDPLSET KCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN LCPFGEVFNATRFASVYAWNRKRISNCVADYS VLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS FVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGC VIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPF ERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQ PTNGVGYQPYRVVVLSFELLHAPATVCGPKKS TNLVKNKCVNFNFNGLTGTGVLTESNKKFLPF	151

[0222]

	QQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQL TPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYE CDIPIGAGICASYQTQTNS	
多聚 A 尾	100 nt	

[0223] 表4A.亚基抗原RBD缺失

[0224]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
缺失 RBD 的 SARS-CoV-2 S1 亚基和添加聚糖的 S2 亚基	31	32
缺失 RBD 和 NTD 的 SARS-CoV-2 S1 亚基和添加聚糖的 S2 亚基	34	35

[0225] 表4B.亚基抗原RBD缺失

[0226]

缺失 RBD 的 SARS-CoV-2 S1 亚基和添加聚糖的 S2 亚基		
SEQ ID NO: 30 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 31 和 3' UTR SEQ ID NO: 4。		30
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAG AAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCC UGGUGAGCAGCCAGUGCGUGAACCUGAC CACCCGGACCCAGCUGCCACCAGCCUACA CCAACAGCUUCACCCGGGGCGUCUACUAC CCCGACAAGGUGUUCCGGAGCAGCGUCCU GCACAGCACCCAGGACCUGUUCCUGCCCU UCUUCAGCAACGUGACCUGGUUCCACGCC AUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCA ACGACGGCGUGUACUUCGCCAGCACCGAG AAGAGCAACAUCAUCCGGGGCUGGAUCU UCGGCACCACCCUGGACAGCAAGACCCAG AGCCUGCUGAUCGUGAAUAACGCCACCA ACGUGGUGAUCAAGGUGUGCGAGUUCCA	31

[0227]

	GUUCUGCAACGACCCCUUCCUGGGCGUGU ACUACCACAAGAACAACAAGAGCUGGAU GGAGAGCGAGUUCCGGGUGUACAGCAGC GCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGC AAGCAGGGCAACUUCAAGAACCUGCGGG AGUUCGUGUUCAAGAACAUCGACGGCUA CUUCAAGAUCUACAGCAAGCACACCCCAA UCAACCUGGUGCGGGAUCUGCCCCAGGGC UUCUCAGCCCUGGAGCCCCUGGUGGACCU GCCCAUCGGCAUCAACAUCACCCGGUUCC AGACCCUGCUGGCCCUGCACCGGAGCUAC CUGACCCCAGGCGACAGCAGCAGCGGGUG GACAGCAGGCGCGGCUGCUUACUACGUG GGCUACCUGCAGCCCCGGACCUUCCUGCU GAAGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACUGCGCCCUGGACCCUCU GAGCGAGACCAAGUGCACCCUGAAGAGC UUCACCGUGGAGAAGGGCAUCUACCAGA CCAGCAACUUCGGCGGCAGCGGCGGCGUG AGCGUGAUCACCCCAGGCACCAACACCAG CAACCAGGUGGCCGUGCUGUACCAGGAC GUGAACUGCACCGAGGUGCCCGUGGCCA UCCACGCCGACCAGCUGACACCCACCUGG CGGGUCUACAGCACCGGCAGCAACGUGU UCCAGACCCGGGCCGGUUGCCUGAUCGGC GCCGAGCACGUGAACAACAGCUACGAGU GCGACAUCCCCAUCGGCGCCGGCAUCUGU GCCAGCUACCAGACCCAGACCAAUUCACC CCGGAGGGCAAGGAGCGUGGCCAGCCAG AGCAUCAUCGCCUACACCAUGAGCCUGGG CGCCGAGAACAGCGUGGCCUACAGCAACA ACAGCAUCGCCAUCCCCACCAACUUCACC AUCAGCGUGACCACCGAGAUUCUGCCCGU GAGCAUGACCAAGACCAGCGUGGACUGC ACCAUGUACAUCUGCGGCGACAGCACCGA GUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCU GACCGGCAUCGCCGUGGAGCAGGACAAG AACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUCUACAAGACCCCUCCCAUCAAG GACUUCGGCGGCUUCAACUUCAGCCAGA UCCUGCCCGACCCCAGCAAGCCCAGCAAG	

[0228]

	CGGAGCUUCAUCGAGGACCUGCUGUUCA ACAAGGUGACCCUAGCCGACGCCGGCUUC AUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCC CAGAAGUUCAACGGCCUGACCGUGCUGCC UCCCCUGCUGACCGACGAGAUGAUCGCCC AGUACACCAGCGCCCUGUUAGCCGGAACC AUCACCAGCGGCUGGACUUUCGGCGCUG GAGCCGCUCUGCAGAUCCCCUUCGCCAUG CAGAUGGCCUACCGGUUCAACGGCAUCG GCGUGACCCAGAACGUGCUGUACGAGAA CCAGAAGCUGAUCGCCAACCAGUUCAACA GCGCCAUCGGCAAGAUCCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCU GCAGGACGUGGUGAACCAGAACGCCCAG GCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUG CUGAACGACAUCCUGAGCCGGCUGGACCC UCCCAACGCCACCGUGCAGAUCGACCGGC UGAUCACUGGCCGGCUGCAGAGCCUGCA GACCUACGUGACCCAGCAGCUGAUCCGGG CCGCCGAGAUUCGGGCCAGCGCCAACCUG GCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUG CGGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUGUU CCUGCACGUGACCUACGUGCCCGCCCAGG AGAAGAACUUCACCACCGCCCCAGCCAUC UGCCACGACGGCAAGGCCCACUUUCCCCG GGAGGGCGUGUUCGUGAGCAACGGCACC CACUGGUUCGUGACCCAGCGGAACUUCU ACGAGCCCCAGAUCAUCACCACCGACAAC ACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUA CGAUCCCCUGCAGCCCGAGCUGGACAGCU UCAAGGAGGAGCUGGACAAGUACUUCAA GAAUCACACCAGCCCCGACGUGGACCUGG GCGACAUCAGCGGCAUCAACGCCAGCGUG GUGAACAUCCAGAAGGAGAUCGAUCGGC UGAACGAGGUGGCCAAGAACCUGAACGA GAGCCUGAUCGACCUGCAGGAGCUGGGC AAGUACGAGCAGUACAUCAAGUGGCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGG	

	CCUGAUCGCCAUCGUGAUGGUGACCAUC AUGCUGUGCUGCAUGACCAGCUGCUGCA GCUGCCUGAAGGGCUGUUGCAGCUGCGG CAGCUGCUGCAAGUUCGACGAGGACGAC AGCGAGCCCGUGCUGAAGGGCGUGAAGC UGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU AGCUUCUUGCCCCUUGGGCCUCCCCCCAG CCCCUCCUCCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGG GCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNS FTRGVYYPDKVFRSSVLHSTQDLFLPFFSNV TWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNAT NVVIKVCEFQFCNDPFLGVYYHKNNKSWME SEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDL PQGFSALEPLVDLPIGINITRFQTLLALHRSYL TPGDSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDCALDPLSETKCTLKSFTVEK GIYQTSNFGGSGGVSVITPGTNTSNQVAVLY QDVNCTEVPVAIHADQLTPTWRVYSTGSNV FQTRAGCLIGAEHVNNSYECDIPIGAGICASY QTQTNSPRRARSVASQSIIAYTMSLGAENSV AYSNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTECSNLLLQYGSFCTQLNRALTG IAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF NFSQILPDPSKPSKRSFIEDLLFNKVTLADAG FIKQYGDCLGDIAARDLICAQKFNGLTVLPP LLTDEMIAQYTSALLAGTITSGWTFGAGAAL QIPFAMQMAYRFNGIGVTQNVLYENQKLIA NQFNSAIGKIQDSLSSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRLDPP NATVQIDRLITGRLQSLQTYVTQQLIRAAEIR ASANLAATKMSECVLGQSKRVDFCGKGYH LMSFPQSAPHGVVFLHVTYVPAQEKNFTTAP AICHDGKAHFPREGVFVSNGTHWFVTQRNF YEPQIITTDNTFVSGNCDVVIGIVNNTVYDPL QPELDSFKEELDKYFKNHTSPDVDLGDISGIN ASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLC CMTSCCSCLKGCCSCGSCCKFDEDDSEPVLK	32

[0229]

	GVKLHYT	
多聚 A 尾	100 nt	
缺失 RBD 和 NTD 的 SARS-CoV-2 S1 亚基和添加聚糖的 S2 亚基		
SEQ ID NO: 33 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 34 和 3' UTR SEQ ID NO: 4。		33
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAG AAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCC UGGUGAGCAGCCAGGGCACCAUCACCGAC GCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUC ACCGUGGAGAAGGGCAUCUACCAGACCA GCAACUUCGGCGGCAGCGGCGGCGUGAG CGUGAUCACCCCAGGCACCAACACCAGCA ACCAGGUGGCCGUGCUGUACCAGGACGU GAACUGCACCGAGGUGCCCGUGGCCAUCC ACGCCGACCAGCUGACACCCACCUGGCGG GUCUACAGCACCGGCAGCAACGUGUUCCA GACCCGGGCCGGUUGCCUGAUCGGCGCCG AGCACGUGAACAACAGCUACGAGUGCGA CAUCCCCAUCGGCGCCGGCAUCUGUGCCA GCUACCAGACCCAGACCAAUUCACCCCGG AGGGCAAGGAGCGUGGCCAGCCAGAGCA UCAUCGCCUACACCAUGAGCCUGGGCGCC GAGAACAGCGUGGCCUACAGCAACAACA GCAUCGCCAUCCCCACCAACUUCACCAUC AGCGUGACCACCGAGAUUCUGCCCGUGA GCAUGACCAAGACCAGCGUGGACUGCACC AUGUACAUCUGCGGCGACAGCACCGAGU GCAGCAACCUGCUGCUGAACUACACCAGC UUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAAC ACCCAGGAGGUGUUCGCCCAGGUGAAGC AGAUCUACAAGACCCCUCCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCCU GCCCGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAACAA GGUGACCCUAGCCGACGCCGGCUUCAUCA AGCAGUACGGCGACUGCCUCGGCGACAU AGCCGCCCGGGACCUGAUCUGCGCCCAGA	34

[0230]

[0231]

	AGUUCAACGGCCUGACCGUGCUGCCUCCC CUGCUGACCGACGAGAUGAUCGCCCAGU ACACCAGCGCCCUGUUAGCCGGAACCAUC ACCAGCGGCUGGACUUUCGGCGCUGGAG CCGCUCUGCAGAUCCCCUUCGCCAUGCAG AUGGCCUACCGGUUCAACGGCAUCGGCG UGACCCAGAACGUGCUGUACGAGAACCA GAAGCUGAUCGCCAACCAGUUCAACAGC GCCAUCGGCAAGAUCCAGGACAGCCUGA GCAGCACCGCUAGCGCCCUGGGCAAGCUG CAGGACGUGGUGAACCAGAACGCCCAGG CCCUGAACACCCUGGUGAAGCAGCUGAGC AGCAACUUCGGCGCCAUCAGCAGCGUGCU GAACGACAUCCUGAGCCGGCUGGACCCUC CCAACGCCACCGUGCAGAUCGACCGGCUG AUCACUGGCCGGCUGCAGAGCCUGCAGAC CUACGUGACCCAGCAGCUGAUCCGGGCCG CCGAGAUUCGGGCCAGCGCCAACCUGGCC GCCACCAAGAUGAGCGAGUGCGUGCUGG GCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCCCC AGAGCGCACCCCACGGAGUGGUGUUCCU GCACGUGACCUACGUGCCCGCCCAGGAGA AGAACUUCACCACCGCCCCAGCCAUCUGC CACGACGGCAAGGCCCACUUUCCCCGGGA GGGCGUGUUCGUGAGCAACGGCACCCAC UGGUUCGUGACCCAGCGGAACUUCUACG AGCCCCAGAUCAUCACCACCGACAACACC UUCGUGAGCGGCAACUGCGACGUGGUGA UCGGCAUCGUGAACAACACCGUGUACGA UCCCCUGCAGCCCGAGCUGGACAGCUUCA AGGAGGAGCUGGACAAGUACUUCAAGAA UCACACCAGCCCCGACGUGGACCUGGGCG ACAUCAGCGGCAUCAACGCCAGCGUGGU GAACAUCCAGAAGGAGAUCGAUCGGCUG AACGAGGUGGCCAAGAACCUGAACGAGA GCCUGAUCGACCUGCAGGAGCUGGGCAA GUACGAGCAGUACAUCAAGUGGCCCUGG UACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUG CUGUGCUGCAUGACCAGCUGCUGCAGCU GCCUGAAGGGCUGUUGCAGCUGCGGCAG CUGCUGCAAGUUCGACGAGGACGACAGC	

[0232]

	GAGCCCGUGCUGAAGGGCGUGAAGCUGC ACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU AGCUUCUUGCCCCUUGGGCCUCCCCCCAG CCCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGG GCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQGTITDAVDCALDPLSET KCTLKSFTVEKGIYQTSNFGGSGGVSVITPGT NTSNQVAVLYQDVNCTEVPVAIHADQLTPT WRVYSTGSNVFQTRAGCLIGAEHVNNSYEC DIPIGAGICASYQTQTNSPRRARSVASQSIIAY TMSLGAENSVAYSNNSIAIPTNFTISVTTEILP VSMTKTSVDCTMYICGDSTECSNLLLNYTSF CTQLNRALTGIAVEQDKNTQEVFAQVKQIY KTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLL FNKVTLADAGFIKQYGDCLGDIAARDLICAQ KFNGLTVLPPLLTDEMIAQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSSTASALGK LQDVVNQNAQALNTLVKQLSSNFGAISSVL NDILSRLDPPNATVQIDRLITGRLQSLQTYVT QQLIRAAEIRASANLAATKMSECVLGQSKRV DFCGKGYHLMSFPQSAPHGVVFLHVTYVPA QEKNFTTAPAICHDGKAHFPREGVFVSNGTH WFVTQRNFYEPQIITTDNTFVSGNCDVVIGIV NNTVYDPLQPELDSFKEELDKYFKNHTSPDV DLGDISGINASVVNIQKEIDRLNEVAKNLNES LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIV MVTIMLCCMTSCCSCLKGCCSCGSCCKFDE DDSEPVLKGVKLHYT	35
多聚A尾	100 nt	

[0233] 嵌合S1-S2亚基抗原

[0234] 在一些实施方案中,组合物包含编码嵌合蛋白(例如具有来自一种病毒的S蛋白的S1亚基和来自另一种不同病毒的S蛋白的S2亚基的嵌合S1-S2蛋白)的mRNA。例如,S2亚基可来自SARS-CoV-2,而S1亚基可来自HKU1。作为另一实例,S2亚基可来自SARS-CoV-2,而S1亚基可来自OC43。这些嵌合蛋白可能由通过先前暴露而产生的对HKU1或OC43的S1亚基具特异性的循环抗体调理,促进巨噬细胞和树突细胞有效地摄取SARS-CoV-2 S2亚基肽并将其交叉呈递给$CD4^+$T细胞。通过循环抗体的调理作用还促进由滤泡树突细胞的捕获,用于呈递给具有对SARS-CoV-2 S2亚基表位具特异性的受体的B细胞。嵌合S1/S2亚基构建体和编码其的mRNA的非限制性实例提供于下文表5A和表5B中。

[0235] 表5A.嵌合S1亚基-S2亚基抗原

[0236]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与 HKU1 S1 亚基连接的 SARS-CoV-2 S2 亚基	37	38
与 OC43 S1 亚基连接的 SARS-CoV-2 S2 亚基	40	41

[0237] 表5B.嵌合S1亚基-S2亚基抗原

[0238]

与 HKU1 S1 亚基连接的 SARS-CoV-2 S2 亚基		
SEQ ID NO: 36从5'端至3'端由以下组成:5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 37 和 3' UTR SEQ ID NO: 4。		36
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUA AGAAGAAAUAUAAGACCCCGGCGCCGC CACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCCUGAUCAUCUUCAUCCUGCCC ACCACCCUGGCCGUGAUCGGCGACUUC AACUGCACCAACAGCUUCAUCAACGAC UACAACAAGACCAUCCCUCGGAUCAGC GAAGACGUGGUGGACGUGUCCCUGGGC CUGGGCACCUACUACGUGCUGAACCGG GUGUACCUGAACACCACACUGCUGUUC ACCGGCUACUUCCCCAAGAGCGGCGCC AACUUCCGGGACCUGGCCCUGAAGGGC AGCAUCUACCUGAGCACCUUGUGGUAC AAGCCUCCCUUCCUGAGCGACUUCAAU	37

[0239]

	AACGGCAUCUUCUCUAAGGUGAAGAAC	
	ACCAAGCUGUACGUAAACAACACCCUG	
	UACAGCGAGUUCAGCACCAUCGUGAUC	
	GGCAGCGUGUUCGUCAACACCAGCUAC	
	ACCAUCGUGGUGCAGCCCCACAACGGC	
	AUCCUGGAGAUCACCGCCUGCCAGUAC	
	ACCAUGUGCGAGUACCCUCACACCGUG	
	UGCAAGAGCAAGGGCUCCAUCCGGAAC	
	GAGAGCUGGCACAUCGACAGCAGCGAG	
	CCGCUGUGCCUGUUCAAGAAGAACUUC	
	ACCUACAACGUGAGCGCCGACUGGCUG	
	UACUUCCACUUCUACCAGGAGCGGGGC	
	GUGUUCUACGCCUACUACGCCGACGUG	
	GGCAUGCCAACCACCUUCCUGUUCAGC	
	CUGUACCUGGGCACCAUCCUGAGCCAC	
	UACUACGUGAUGCCCCUGACCUGCAAC	
	GCCAUCAGCUCAAACACCGACAACGAG	
	ACCCUGGAGUACUGGGUGACUCCACUG	
	AGCCGGCGGCAGUACCUGCUGAACUUC	
	GACGAGCACGGCGUGAUCACCAACGCC	
	GUGGACUGCGCCCUGGACCCUCUGAGC	
	GAGACCAAGUGCACCCUGAAGAGCUUC	
	ACCGUGGAGAAGGGCAUCUACCAGACC	
	AGCGGCUUCACCGUGAAGCCCGUAGCC	
	ACCGUGUACCGGCGGAUCCCCAACCUG	
	CCCGACUGCGACAUCGACAACUGGCUG	
	AACAACGUCAGCGUGCCCAGCCCACUG	
	AACUGGGAGCGGCGGAUCUUCAGCAAC	
	UGCAACUUCAAUCUGAGCACCCUGCUG	
	CGGCUGGUGCACGUGGACAGCUUCAGC	
	UGCAACAACCUGGACAAGAGCAAGAUC	
	UUCGGUAGCUGCUUCAACAGCAUCACC	
	GUGGACAAGUUCGCCAUCCCUAACCGG	
	CGGCGGGACGAUCUGCAGCUGGGCAGC	
	AGCGGCUUCCUGCAGAGCAGCAACUAC	
	AAGAUCGACAUCAGCAGCUCAAGCUGC	
	CAGCUGUACUACAGCCUGCCCCUGGUG	
	AACGUGACCAUCAACAACUUCAACCCC	
	AGCAGCUGGAACCGGCGGUACGGCUUC	
	GGCAGCUUCAACCUGAGCAGCUACGAC	
	GUGGUGUACAGCGACCACUGCUUCAGC	
	GUGAACAGCGACUUCUGCCCCUGUGCC	
	GACCCUAGCGUGGUGAACAGCUGCGCC	

[0240]

	AAGAGCAAGCCUCCCAGCGCCAUUUGC CCCGCCGGCACCAAGUACCGGCACUGC GACCUGGACACCACCCUGUACGUGAAG AACUGGUGCCGGUGCAGCUGCCUGCCC GACCCCAUCAGCACCUACAGCCCCAAC ACCUGUCCCCAGAAGAAGGUGGUGGUG GGUAUCGGCGAGCACUGUCCCGGCCUG GGCAUCAACGAGGAGAAGUGCGGCACC CAGCUGAACCACAGCAGCUGCUUCUGU AGCCCCGACGCCUUCCUGGGCUGGAGC UUCGACAGCUGCAUCAGCAACAACCGG UGCAACAUCUUUAGCAACUUCAUCUUC AACGGAAUCAACAGCGGCACCACCUGC AGCAACGACCUGCUGUAUAGCAACACC GAGAUCAGCACCGGCGUGUGCGUGAAC UACGACCUGUACGGCAUCACCGGCCAG GGCAUCUUCAAGGAGGUGAGCGCCGCC UACUACAACAACUGGCAGAACCUGCUG UACGACAGCAACGGCAACAUCAUCGGC UUCAAGGACUUUCUGACCAACAAGACC UACACCAUCCUGCCCUGCUACAGCGGC GGCGUGAGCGUGAUCACCCCAGGCACC AACACCAGCAACCAGGUGGCCGUGCUG UACCAGGACGUGAACUGCACCGAGGUG CCCGUGGCCAUCCACGCCGACCAGCUG ACACCCACCUGGCGGGUCUACAGCACC GGCAGCAACGUGUUCCAGACCCGGGCC GGUUGCCUGAUCGGCGCCGAGCACGUG AACAACAGCUACGAGUGCGACAUCCCC AUCGGCGCCGGCAUCUGUGCCAGCUAC CAGACCCAGACCAAUUCACCCCGGAGG GCAAGGAGCGUGGCCAGCCAGAGCAUC AUCGCCUACACCAUGAGCCUGGGCGCC GAGAACAGCGUGGCCUACAGCAACAAC AGCAUCGCCAUCCCCACCAACUUCACC AUCAGCGUGACCACCGAGAUUCUGCCC GUGAGCAUGACCAAGACCAGCGUGGAC UGCACCAUGUACAUCUGCGGCGACAGC ACCGAGUGCAGCAACCUGCUGCUGCAG UACGGCAGCUUCUGCACCCAGCUGAAC CGGGCCCUGACCGGCAUCGCCGUGGAG CAGGACAAGAACACCCAGGAGGUGUUC GCCCAGGUGAAGCAGAUCUACAAGACC	

[0241]

	CCUCCCAUCAAGGACUUCGGCGGCUUC AACUUCAGCCAGAUCCUGCCCGACCCC AGCAAGCCCAGCAAGCGGAGCUUCAUC GAGGACCUGCUGUUCAACAAGGUGACC CUAGCCGACGCCGGCUUCAUCAAGCAG UACGGCGACUGCCUCGGCGACAUAGCC GCCCGGGACCUGAUCUGCGCCCAGAAG UUCAACGGCCUGACCGUGCUGCCUCCC CUGCUGACCGACGAGAUGAUCGCCCAG UACACCAGCGCCCUGUUAGCCGGAACC AUCACCAGCGGCUGGACUUUCGGCGCU GGAGCCGCUCUGCAGAUCCCCUUCGCC AUGCAGAUGGCCUACCGGUUCAACGGC AUCGGCGUGACCCAGAACGUGCUGUAC GAGAACCAGAAGCUGAUCGCCAACCAG UUCAACAGCGCCAUCGGCAAGAUCCAG GACAGCCUGAGCAGCACCGCUAGCGCC CUGGGCAAGCUGCAGGACGUGGUGAAC CAGAACGCCCAGGCCCUGAACACCCUG GUGAAGCAGCUGAGCAGCAACUUCGGC GCCAUCAGCAGCGUGCUGAACGACAUC CUGAGCCGGCUGGACCCUCCCGAGGCC GAGGUGCAGAUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUAC GUGACCCAGCAGCUGAUCCGGGCCGCC GAGAUUCGGGCCAGCGCCAACCUGGCC GCCACCAAGAUGAGCGAGUGCGUGCUG GGCCAGAGCAAGCGGGUGGACUUCUGC GGCAAGGGCUACCACCUGAUGAGCUUU CCCCAGAGCGCACCCCACGGAGUGGUG UUCCUGCACGUGACCUACGUGCCCGCC CAGGAGAAGAACUUCACCACCGCCCCA GCCAUCUGCCACGACGGCAAGGCCCAC UUUCCCCGGGAGGGCGUGUUCGUGAGC AACGGCACCCACUGGUUCGUGACCCAG CGGAACUUCUACGAGCCCCAGAUCAUC ACCACCGACAACACCUUCGUGAGCGGC AACUGCGACGUGGUGAUCGGCAUCGUG AACAACACCGUGUACGAUCCCCUGCAG CCCGAGCUGGACAGCUUCAAGGAGGAG CUGGACAAGUACUUCAAGAAUCACACC AGCCCCGACGUGGACCUGGGCGACAUC AGCGGCAUCAACGCCAGCGUGGUGAAC	

	AUCCAGAAGGAGAUCGAUCGGCUGAAC GAGGUGGCCAAGAACCUGAACGAGAGC CUGAUCGACCUGCAGGAGCUGGGCAAG UACGAGCAGUACAUCAAGUGGCCCUGG UACAUCUGGCUGGGCUUCAUCGCCGGC CUGAUCGCCAUCGUGAUGGUGACCAUC AUGCUGUGCUGCAUGACCAGCUGCUGC AGCUGCCUGAAGGGCUGUUGCAGCUGC GGCAGCUGCUGCAAGUUCGACGAGGAC GACAGCGAGCCCGUGCUGAAGGGCGUG AAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCC UAGCUUCUUGCCCCUUGGGCCUCCCCC CAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGC	4
相应的氨基酸序列	MFLIIFILPTTLAVIGDFNCTNSFINDYNKTI PRISEDVVDVSLGLGTYYVLNRVYLNTTL LFTGYFPKSGANFRDLALKGSIYLSTLWY KPPFLSDFNNGIFSKVKNTKLYVNNTLYSE FSTIVIGSVFVNTSYTIVVQPHNGILEITAC QYTMCEYPHTVCKSKGSIRNESWHIDSSEP LCLFKKNFTYNVSADWLYFHFYQERGVF YAYYADVGMPTTFLFSLYLGTILSHYYVM PLTCNAISSNTDNETLEYWVTPLSRRQYLL NFDEHGVITNAVDCALDPLSETKCTLKSFT VEKGIYQTSGFTVKPVATVYRRIPNLPDCD IDNWLNNVSVPSPLNWERRIFSNCNFNLST LLRLVHVDSFSCNNLDKSKIFGSCFNSITV DKFAIPNRRRDDLQLGSSGFLQSSNYKIDIS SSSCQLYYSLPLVNVTINNFNPSSWNRRYG FGSFNLSSYDVVYSDHCFSVNSDFCPCADP SVVNSCAKSKPPSAICPAGTKYRHCDLDT TLYVKNWCRCSCLPDPISTYSPNTCPQKK VVVGIGEHCPGLGINEEKCGTQLNHSSCFC SPDAFLGWSFDSCISNNRCNIFSNFIFNGIN SGTTCSNDLLYSNTEISTGVCVNYDLYGIT GQGIFKEVSAAYYNNWQNLLYDSNGNIIG FKDFLTNKTYTILPCYSGGVSVITPGTNTS NQVAVLYQDVNCTEVPVAIHADQLTPTW RVYSTGSNVFQTRAGCLIGAEHVNNSYEC DIPIGAGICASYQTQTNSPRRARSVASQSII AYTMSLGAENSVAYSNNSIAIPTNFTISVT	38

[0242]

	TEILPVSMTKTSVDCTMYICGDSTECSNLL LQYGSFCTQLNRALTGIAVEQDKNTQEVF AQVKQIYKTPPIKDFGGFNFSQILPDPSKPS KRSFIEDLLFNKVTLADAGFIKQYGDCLG DIAARDLICAQKFNGLTVLPPLLTDEMIAQ YTSALLAGTITSGWTFGAGAALQIPFAMQ MAYRFNGIGVTQNVLYENQKLIANQFNSA IGKIQDSLSSTASALGKLQDVVNQNAQAL NTLVKQLSSNFGAISSVLNDILSRLDPPEAE VQIDRLITGRLQSLQTYVTQQLIRAAEIRAS ANLAATKMSECVLGQSKRVDFCGKGYHL MSFPQSAPHGVVFLHVTYVPAQEKNFTTA PAICHDGKAHFPREGVFVSNGTHWFVTQR NFYEPQIITTDNTFVSGNCDVVIGIVNNTV YDPLQPELDSFKEELDKYFKNHTSPDVDL GDISGINASVVNIQKEIDRLNEVAKNLNES LIDLQELGKYEQYIKWPWYIWLGFIAGLIA IVMVTIMLCCMTSCCSCLKGCCSCGSCCK FDEDDSEPVLKGVKLHYT	
多聚 A 尾	100 nt	
与 OC43 S1 亚基连接的 SARS-CoV-2 S2 亚基		
SEQ ID NO: 39 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 40 和 3' UTR SEQ ID NO: 4。		39
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUA AGAAGAAAUAUAAGACCCCGGCGCCGC CACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCCUGAUCCUGCUGAUCAGCCUG CCCACCGCCUUCGCCGUGAUCGGCGAC CUGAAGUGCACCAGCGACAACAUCAAC GACAAGGACACCGGCCCACCACCCAUC AGCACCGACACCGUGGACGUGACCAAC GGCCUGGGCACCUACUACGUGCUGGAC CGGGUGUACCUGAACACCACCCUGUUC CUGAACGGCUACUACCCCACCAGCGGC AGCACCUACCGGAAUAUGGCCCUGAAG GGCAGCGUGCUGCUGAGCCGGCUGUGG UUCAAGCCACCAUUCCUGAGCGACUUC AUCAACGGAAUCUUCGCCAAGGUGAAG AACACCAAGGUGAUCAAGGACCGGGUG AUGUACAGCGAGUUCCCCGCCAUCACC	40

[0243]

[0244]

	AUUGGCAGUACCUUCGUGAACACCAGC UACAGCGUGGUGGUGCAGCCCCGGACC AUCAACAGCACCCAGGACGGCGACAAC AAGCUGCAGGGCCUGCUGGAGGUGAGC GUGUGCCAGUACAACAUGUGCGAGUAC CCUCAGACCAUCUGCCACCCCAACCUG GGCAACCACCGGAAGGAGCUGUGGCAC CUGGACACCGGCGUGGUGAGCUGCCUG UACAAGCGGAACUUCACCUACGACGUA AACGCCGACUACCUGUACUUCCACUUC UACCAGGAGGGCGGCACCUUCUACGCC UACUUCACCGACACGGGCGUGGUGACC AAGUUCCUGUUCAACGUGUACCUGGGC AUGGCCCUGAGCCACUACUACGUGAUG CCCCUGACCUGUAACAGCAAGCUGACC CUGGAGUACUGGGUGACCCCUCUGACC AGCCGGCAGUACCUGCUGGCCUUCAAC CAGGACGGCAUCAUCUUCAACGCCGUG GACUGCGCCCUGGACCCUCUGAGCGAG ACCAAGUGCACCCUGAAGAGCUUCACC GUGGAGAAGGGCAUCUACCAGACCAAC GGCUACACCGUGCAGCCCAUCGCCGAC GUGUACCGGCGGAAGCCCAACCUGCCC AACUGCAACAUCGAGGCCUGGCUGAAC GACAAGAGCGUGCCCUCGCCCCUGAAC UGGGAGCGGAAGACCUUCAGCAACUGC AACUUCAACAUGAGCAGUCUGAUGAGC UUCAUCCAGGCCGACAGCUUCACCUGC AACAACAUCGACGCCGCCAAGAUCUAC GGCAUGUGCUUCAGCAGCAUCACCAUC GACAAGUUUGCCAUCCCCAACGGCCGG AAGGUGGACCUGCAGCUGGGCAACCUG GGCUACCUGCAGAGCUUCAACUACCGG AUCGACACCACCGCCACCUCUUGCCAG CUGUACUACAACCUGCCCGCCGCCAAC GUGAGCGUGAGCCGGUUCAACCCCAGC ACCUGGAACAAGCGGUUCGGCUUCAUU GAGGACAGCGUGUUCAAACCCCGGCCC GCAGGAGUACUGACCAACCACGACGUG GUGUACGCCCAGCACUGCUUCAAGGCA CCCAAGAACUUCUGCCCCUGCAAGCUG AACGGCAGCUGUGUGGGCUCUGGCCCC GGUAAGAACAACGGCAUAGGGACUUGC	

[0245]

	CCGGCAGGGACCAACUACCUGACCUGC GACAACCUGUGCACACCCGACCCCAUC ACCUUCACCGGCACCUACAAGUGUCCC CAGACCAAGAGCCUGGUGGGCAUCGGC GAGCACUGCAGCGGCCUGGCCGUGAAG AGCGACUACUGCGGCGGCAACAGCUGC ACCUGUCGGCCCCAGGCCUUCCUGGGC UGGAGCGCCGACAGCUGCCUGCAGGGC GACAAGUGCAACAUCUUUGCCAACUUC AUCCUGCACGACGUGAACAGCGGCCUG ACCUGCAGCACCGACCUGCAGAAGGCC AACACCGACAUCAUCCUGGGCGUGUGC GUGAACUACGACUUGUACGGCAUCCUG GGCCAGGGCAUCUUCGUGGAGGUGAAC GCCACCUACUACAACAGCUGGCAGAAC CUGCUGUACGACAGCAACGGCAACCUG UACGGCUUCCGGGACUACAUCAUCAAC CGGACCUUCAUGAUCCGGAGCUGCUAC AGCGGCGGCGUGAGCGUGAUCACCCCA GGCACCAACACCAGCAACCAGGUGGCC GUGCUGUACCAGGACGUGAACUGCACC GAGGUGCCCGUGGCCAUCCACGCCGAC CAGCUGACACCCACCUGGCGGGUCUAC AGCACCGGCAGCAACGUGUUCCAGACC CGGGCCGGUUGCCUGAUCGGCGCCGAG CACGUGAACAACAGCUACGAGUGCGAC AUCCCCAUCGGCGCCGGCAUCUGUGCC AGCUACCAGACCCAGACCAAUUCACCC CGGAGGGCAAGGAGCGUGGCCAGCCAG AGCAUCAUCGCCUACACCAUGAGCCUG GGCGCCGAGAACAGCGUGGCCUACAGC AACAACAGCAUCGCCAUCCCCACCAAC UUCACCAUCAGCGUGACCACCGAGAUU CUGCCCGUGAGCAUGACCAAGACCAGC GUGGACUGCACCAUGUACAUCUGCGGC GACAGCACCGAGUGCAGCAACCUGCUG CUGCAGUACGGCAGCUUCUGCACCCAG CUGAACCGGGCCCUGACCGGCAUCGCC GUGGAGCAGGACAAGAACACCCAGGAG GUGUUCGCCCAGGUGAAGCAGAUCUAC AAGACCCCUCCCAUCAAGGACUUCGGC GGCUUCAACUUCAGCCAGAUCCUGCCC GACCCCAGCAAGCCCAGCAAGCGGAGC	

[0246]

	UUCAUCGAGGACCUGCUGUUCAACAAG	
	GUGACCCUAGCCGACGCCGGCUUCAUC	
	AAGCAGUACGGCGACUGCCUCGGCGAC	
	AUAGCCGCCCGGGACCUGAUCUGCGCC	
	CAGAAGUUCAACGGCCUGACCGUGCUG	
	CCUCCCCUGCUGACCGACGAGAUGAUC	
	GCCCAGUACACCAGCGCCCUGUUAGCC	
	GGAACCAUCACCAGCGGCUGGACUUUC	
	GGCGCUGGAGCCGCUCUGCAGAUCCCC	
	UUCGCCAUGCAGAUGGCCUACCGGUUC	
	AACGGCAUCGGCGUGACCCAGAACGUG	
	CUGUACGAGAACCAGAAGCUGAUCGCC	
	AACCAGUUCAACAGCGCCAUCGGCAAG	
	AUCCAGGACAGCCUGAGCAGCACCGCU	
	AGCGCCCUGGGCAAGCUGCAGGACGUG	
	GUGAACCAGAACGCCCAGGCCCUGAAC	
	ACCCUGGUGAAGCAGCUGAGCAGCAAC	
	UUCGGCGCCAUCAGCAGCGUGCUGAAC	
	GACAUCCUGAGCCGGCUGGACCCUCCC	
	GAGGCCGAGGUGCAGAUCGACCGGCUG	
	AUCACUGGCCGGCUGCAGAGCCUGCAG	
	ACCUACGUGACCCAGCAGCUGAUCCGG	
	GCCGCCGAGAUUCGGGCCAGCGCCAAC	
	CUGGCCGCCACCAAGAUGAGCGAGUGC	
	GUGCUGGGCCAGAGCAAGCGGGUGGAC	
	UUCUGCGGCAAGGGCUACCACCUGAUG	
	AGCUUUCCCCAGAGCGCACCCCACGGA	
	GUGGUGUUCCUGCACGUGACCUACGUG	
	CCCGCCCAGGAGAAGAACUUCACCACC	
	GCCCCAGCCAUCUGCCACGACGGCAAG	
	GCCCACUUUCCCCGGGAGGGCGUGUUC	
	GUGAGCAACGGCACCCACUGGUUCGUG	
	ACCCAGCGGAACUUCUACGAGCCCCAG	
	AUCAUCACCACCGACAACACCUUCGUG	
	AGCGGCAACUGCGACGUGGUGAUCGGC	
	AUCGUGAACAACACCGUGUACGAUCCC	
	CUGCAGCCCGAGCUGGACAGCUUCAAG	
	GAGGAGCUGGACAAGUACUUCAAGAAU	
	CACACCAGCCCCGACGUGGACCUGGGC	
	GACAUCAGCGGCAUCAACGCCAGCGUG	
	GUGAACAUCCAGAAGGAGAUCGAUCGG	
	CUGAACGAGGUGGCCAAGAACCUGAAC	
	GAGAGCCUGAUCGACCUGCAGGAGCUG	

[0247]

	GGCAAGUACGAGCAGUACAUCAAGUGG CCCUGGUACAUCUGGCUGGGCUUCAUC GCCGGCCUGAUCGCCAUCGUGAUGGUG ACCAUCAUGCUGUGCUGCAUGACCAGC UGCUGCAGCUGCCUGAAGGGCUGUUGC AGCUGCGGCAGCUGCUGCAAGUUCGAC GAGGACGACAGCGAGCCCGUGCUGAAG GGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCC UAGCUUCUUGCCCCUUGGGCCUCCCCC CAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGC	4
相应的氨基酸序列	MFLILLISLPTAFAVIGDLKCTSDNINDKDT GPPPISTDTVDVTNGLGTYYVLDRVYLNT TLFLNGYYPTSGSTYRNMALKGSVLLSRL WFKPPFLSDFINGIFAKVKNTKVIKDRVM YSEFPAITIGSTFVNTSYSVVVQPRTINSTQ DGDNKLQGLLEVSVCQYNMCEYPQTICHP NLGNHRKELWHLDTGVVSCLYKRNFTYD VNADYLYFHFYQEGGTFYAYFTDTGVVT KFLFNVYLGMALSHYYVMPLTCNSKLTLE YWVTPLTSRQYLLAFNQDGIIFNAVDCAL DPLSETKCTLKSFTVEKGIYQTNGYTVQPI ADVYRRKPNLPNCNIEAWLNDKSVPSPLN WERKTFSNCNFNMSSLMSFIQADSFTCNNI DAAKIYGMCFSSITIDKFAIPNGRKVDLQL GNLGYLQSFNYRIDTTATSCQLYYNLPAA NVSVSRFNPSTWNKRFGFIEDSVFKPRPAG VLTNHDVVYAQHCFKAPKNFCPCKLNGS CVGSGPGKNNGIGTCPAGTNYLTCDNLCT PDPITFTGTYKCPQTKSLVGIGEHCSGLAV KSDYCGGNSCTCRPQAFLGWSADSCLQG DKCNIFANFILHDVNSGLTCSTDLQKANT DIILGVCVNYDLYGILGQGIFVEVNATYYN SWQNLLYDSNGNLYGFRDYIINRTFMIRSC YSGGVSVITPGTNTSNQVAVLYQDVNCTE VPVAIHADQLTPTWRVYSTGSNVFQTRAG CLIGAEHVNNSYECDIPIGAGICASYQTQT NSPRRARSVASQSIIAYTMSLGAENSVAYS NNSIAIPTNFTISVTTEILPVSMTKTSVDCT MYICGDSTECSNLLLQYGSFCTQLNRALT GIAVEQDKNTQEVFAQVKQIYKTPPIKDFG	41

[0248]

	GFNFSQILPDPSKPSKRSFIEDLLFNKVTLA DAGFIKQYGDCLGDIAARDLICAQKFNGL TVLPPLLTDEMIAQYTSALLAGTITSGWTF GAGAALQIPFAMQMAYRFNGIGVTQNVL YENQKLIANQFNSAIGKIQDSLSSTASALG KLQDVVNQNAQALNTLVKQLSSNFGAISS VLNDILSRLDPPEAEVQIDRLITGRLQSLQT YVTQQLIRAAEIRASANLAATKMSECVLG QSKRVDFCGKGYHLMSFPQSAPHGVVFL HVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFV SGNCDVVIGIVNNTVYDPLQPELDSFKEEL DKYFKNHTSPDVDLGDISGINASVVNIQKE IDRLNEVAKNLNESLIDLQELGKYEQYIK WPWYIWLGFIAGLIAIVMVTIMLCCMTSC CSCLKGCCSCGSCCKFDEDDSEPVLKGVK LHYT	
多聚 A 尾	100 nt	

[0249] 编码的结构域抗原

[0250] 本公开的其他方面提供了包含编码S蛋白的SARS-CoV-2 S1亚基的(至少一个)子结构域的mRNA的组合物。所述子结构域可为N末端结构域(NTD)或受体结合结构域(RBD)(有或无SD1和/或SD2)。在一些实施方案中,mRNA编码NTD和RBD(有或无SD1和/或SD2)的组合(例如,非天然组合)。在一些实施方案中,所述NTD和/或RBD与跨膜结构域(有或无SD1和/或SD2)连接。在一些实施方案中,所述mRNA编码S蛋白的SARS-CoV-2 S1亚基的两个子结构域(NTD和RBD),其经突变以包含半胱氨酸残基。在一些实施方案中,此类突变导致二硫键的形成。举例来说,mRNA可编码包含F43C突变的NTD和包含Q563C突变的RBD,最终产生经由二硫键与RBD连接的NTD。

[0251] N末端结构域(NTD)构建体

[0252] 在一些实施方案中,本文提供的mRNA编码SARS-CoV-2 S蛋白的S1亚基的NTD。某些β冠状病毒的NTD会引发保护性水平的抗体。对其他β冠状病毒(例如MERS)的NTD具特异性的抗体通过防止膜融合和病毒进入起作用(Zhou H等人,Nat Commun.2019;3068),提供不同于防止病毒附着至ACE2的第二种中和机制。由本公开的mRNA编码的SARS-CoV-2 NTD可为可溶的或膜结合的。膜结合的SARS-CoV-2 NTD抗原和编码其的mRNA的非限制性实例提供于下文表6A和表6B中。

[0253] 表6A.膜结合的NTD抗原

[0254]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与跨膜结构域连接的 SARS-CoV-2 NTD (NTD-TM)	46	47

[0255] 表6B.膜结合的NTD抗原

[0256]

与跨膜结构域连接的 SARS-CoV-2 NTD (NTD-TM)		
SEQ ID NO: 45 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 46 和 3' UTR SEQ ID NO: 4。		45
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAG AAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCC UGGUGAGCAGCCAGUGCGUGAACCUGAC CACCCGGACCCAGCUGCCACCAGCCUACA CCAACAGCUUCACCCGGGGCGUCUACUAC CCCGACAAGGUGUUCCGGAGCAGCGUCCU GCACAGCACCCAGGACCUGUUCCUGCCCU UCUUCAGCAACGUGACCUGGUUCCACGCC AUCCACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCUUCA ACGACGGCGUGUACUUCGCCAGCACCGAG AAGAGCAACAUCAUCCGGGGCUGGAUCU UCGGCACCACCCUGGACAGCAAGACCCAG AGCCUGCUGAUCGUGAAUAACGCCACCA ACGUGGUGAUCAAGGUGUGCGAGUUCCA GUUCUGCAACGACCCCUUCCUGGGCGUGU	46

[0257]

	ACUACCACAAGAACAACAAGAGCUGGAU GGAGAGCGAGUUCCGGGUGUACAGCAGC GCCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGGGC AAGCAGGGCAACUUCAAGAACCUGCGGG AGUUCGUGUUCAAGAACAUCGACGGCUA CUUCAAGAUCUACAGCAAGCACACCCCAA UCAACCUGGUGCGGGAUCUGCCCCAGGGC UUCUCAGCCCUGGAGCCCCUGGUGGACCU GCCCAUCGGCAUCAACAUCACCCGGUUCC AGACCCUGCUGGCCCUGCACCGGAGCUAC CUGACCCCAGGCGACAGCAGCAGCGGGUG GACAGCAGGCGCGGCUGCUUACUACGUG GGCUACCUGCAGCCCCGGACCUUCCUGCU GAAGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACUCUGGCGGAGGCAGCA UCCUGGCCAUCUACAGCACCGUGGCCAGC AGCCUGGUGCUGCUGGUGAGCCUGGGCG CCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU AGCUUCUUGCCCCUUGGGCCUCCCCCCAG CCCCUCCUCCCCUUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGG GCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNS FTRGVYYPDKVFRSSVLHSTQDLFLPFFSNV TWFHAIHVSGTNGTKRFDNPVLPFNDGVYF ASTEKSNIIRGWIFGTTLDSKTQSLLIVNNAT NVVIKVCEFQFCNDPFLGVYYHKNNKSWME SEFRVYSSANNCTFEYVSQPFLMDLEGKQG NFKNLREFVFKNIDGYFKIYSKHTPINLVRDL PQGFSALEPLVDLPIGINITRFQTLLALHRSYL TPGDSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDSGGGSILAIYSTVASSLVLLV SLGAISF	47
多聚A尾	100 nt	

[0258] 受体结合结构域(RBD)构建体

[0259] 在其他实施方案中,本文提供的mRNA编码SARS-CoV-2 S蛋白的S1亚基的RBD。RBD结合宿主细胞上的ACE2受体,这介导病毒附着至细胞。附着对于病毒进入细胞和复制是必需的。因此,阻断病毒附着至细胞中的RBD靶向抗体反应有效地中和细胞外病毒颗粒,从而防止增殖并促进对中和的病毒颗粒的其他组分的进一步免疫反应。由本公开的mRNA编码的SARS-CoV-2 RBD可为可溶的或膜结合的(例如,与跨膜结构域连接)。

[0260] 可溶性RBD抗原

[0261] 在一些实施方案中，mRNA编码可溶性SARS-CoV-2 RBD。树突细胞通过胞饮作用对可溶性蛋白进行取样，并且在迁移至引流淋巴结时，将包含取样的蛋白质的线性肽呈递至$CD4^+$T细胞。这些$CD4^+$T细胞向已经识别、摄取和呈递来自RBD的表位的B细胞提供增殖信号，因此施用特异性RBD而没有SARS-CoV-2刺突蛋白的其他组分预计将免疫反应集中于RBD中存在的表位。可溶性SARS-CoV-2 RBD和编码其的mRNA的非限制性实例提供于下文表7A和表7B中。

[0262] 表7A.可溶性RBD抗原

[0263]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
SARS-CoV-2 可溶性 RBD	61	62

[0264] 表7B.可溶性RBD抗原

[0265]

SARS-CoV-2 可溶性 RBD		
SEQ ID NO: 60 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 61 和 3' UTR SEQ ID NO: 4。		60
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体	AUGUACAGCAUGCAGCUGGCUAGCUG	61

[0266]

的 ORF (不包括终止密码子)	CGUGACCCUGACCCUGGUGCUGCUGG UGAACAGCCAGCCCAACAUCACCAACC UGUGCCCCUUCGGCGAGGUGUUCAAC GCCACCCGGUUCGCCAGCGUGUACGCC UGGAACCGGAAGCGGAUCAGCAACUG CGUGGCCGACUACAGCGUGCUGUACA ACAGCGCCAGCUUCAGCACCUUCAAG UGCUACGGCGUGAGCCCCACCAAGCU GAACGACCUGUGCUUCACCAACGUGU ACGCCGACAGCUUCGUGAUCCGUGGC GACGAGGUGCGGCAGAUCGCACCCGG CCAGACAGGCAAGAUCGCCGACUACA ACUACAAGCUGCCCGACGACUUCACCG GCUGCGUGAUCGCCUGGAACAGCAAC AACCUCGACAGCAAGGUGGGCGGCAA CUACAACUACCUGUACCGGCUGUUCC GGAAGAGCAACCUGAAGCCCUUCGAG CGGGACAUCAGCACCGAGAUCUACCA AGCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUC UGCAGAGCUACGGCUUCCAGCCCACCA ACGGCGUGGGCUACCAGCCCUACCGG GUGGUGGUGCUGAGCUUCGAGCUGCU GCACGCCCCAGCCACCGUGUGUGGCCC CAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGC	4
相应的氨基酸序列	MYSMQLASCVTLTLVLLVNSQPNITNLC PFGEVFNATRFASVYAWNRKRISNCVAD YSVLYNSASFSTFKCYGVSPTKLNDLCFT NVYADSFVIRGDEVRQIAPGQTGKIADY NYKLPDDFTGCVIAWNSNNLDSKVGGN YNYLYRLFRKSNLKPFERDISTEIYQAGS TPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVVLSFELLHAPATVCGPK	62
多聚 A 尾	100 nt	

[0267] 膜结合的RBD抗原

[0268] 在一些实施方案中，mRNA编码膜结合的SARS-CoV-2 RBD。预计表达膜结合的RBD的细胞将这些膜结合的抗原携带至引流淋巴结，并促进RBD特异性B细胞对表位的有效识别。

由于B细胞表面含有许多表面结合的抗体并且表达细胞含有许多膜结合的RBD的拷贝，因此预计B细胞初始识别抗原后会发生B细胞受体的交联，从而经由亲合效应刺激强烈反应。膜结合的SARS-CoV-2 RBD和编码其的mRNA的非限制性实例提供于下文表8A和表8B中。

[0269] 表8A.膜结合的RBD抗原

[0270]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与跨膜结构域连接的 SARS-CoV-2 RBD (RBD-TM)	76	77

[0271] 表8B.膜结合的RBD抗原

[0272]

与跨膜结构域连接的 SARS-CoV-2 RBD (RBD-TM)		
SEQ ID NO: 75 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 76 和 3' UTR SEQ ID NO: 4。		75
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCUGCGUGACCCUGACCCUGGUGCUGCUGGUGAACAGCCAGCCCAACAUCACCAACCUGUGCCCCUUCGGCGAGGUGUUCAACGCCACCCGGUUCGCCAGCGUGUACGCCUGGAACCGGAAGCGGAUCAGCAACUGCGUGGCCGACUACAGCGUGCUGUACAACAGCGCCAGCUUCAGCACCUUCAAGUGCUACGGCGUGAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUCCGUGGCGACGAGGUGCGGCAGAUCGCACCCGG	76

[0273]

	CCAGACAGGCAAGAUCGCCGACUACA ACUACAAGCUGCCCGACGACUUCACCG GCUGCGUGAUCGCCUGGAACAGCAAC AACCUCGACAGCAAGGUGGGCGGCAA CUACAACUACCUGUACCGGCUGUUCC GGAAGAGCAACCUGAAGCCCUUCGAG CGGGACAUCAGCACCGAGAUCUACCA AGCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUC UGCAGAGCUACGGCUUCCAGCCCACCA ACGGCGUGGGCUACCAGCCCUACCGG GUGGUGGUGCUGAGCUUCGAGCUGCU GCACGCCCCAGCCACCGUGUGUGGCCC CAAGUCUGGCGGAGGCAGCAUCCUGG CCAUCUACAGCACCGUGGCCAGCAGCC UGGUGCUGCUGGUGAGCCUGGGCGCC AUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCCC CCAGCCCCUCCUCCCCUUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGC	4
相应的氨基酸序列	MYSMQLASCVTLTLVLLVNSQPNITNLC PFGEVFNATRFASVYAWNRKRISNCVAD YSVLYNSASFSTFKCYGVSPTKLNDLCFT NVYADSFVIRGDEVRQIAPGQTGKIADY NYKLPDDFTGCVIAWNSNNLDSKVGGN YNYLYRLFRKSNLKPFERDISTEIYQAGS TPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVVLSFELLHAPATVCGPKSGGGSI LAIYSTVASSLVLLVSLGAISF	77
多聚A尾	100 nt	

[0274] 结构域融合抗原

[0275] 在其他实施方案中，本文提供的mRNA编码SARS-CoV-2 NTD-RBD融合蛋白。例如，S蛋白的SARS-CoV-2 S1亚基的NTD和RBD可经由接头(例如短氨基酸(例如，甘氨酸-丝氨酸)接头)彼此连接，以允许结构域之间的柔性/铰接和空间。在另一个实施方案中，可使用包含抗原性表位(例如II类通用T细胞表位，例如PADRE)的接头。在一些实施方案中，跨膜区例如经由另一短氨基酸(例如，甘氨酸-丝氨酸或PADRE)接头与NTD-RBD融合体连接以获得柔性并允许膜与抗原之间存在合理距离。不受限于理论，据认为这种膜结合的串联配置在一个开放阅读框中呈递大部分(如果并非全部)已知中和表位和保护性表位。施用这种融合蛋白然后应将免疫反应集中于已知保护性表位，并减少对非保护性表位特异的抗体和T细胞的

不必要产生。此外,针对不同结构域的抗体可经由不同机制、例如通过阻断与宿主细胞的附着或防止结合的病毒经历膜融合并进入宿主细胞来中和病毒颗粒。因此,由包含不同结构域的融合蛋白引发的广泛反应在进化上可能更稳固,需要多个不同突变来逃避疫苗诱导的免疫。SARS-CoV-2 NTD-RBD融合蛋白和编码其的mRNA的非限制性实例提供于下文表9A和表9B中。

[0276] 接头

[0277] 根据本公开可使用多种接头。如本文提供的接头仅仅是将两个其他氨基酸序列人工连接在一起的氨基酸序列。本文所用的接头可为可切割的或不可切割的。可切割的接头允许将mRNA翻译成多肽,之后接头的切割允许独立地释放每个个别组分。不可切割的接头保持连接一个或多个蛋白质亚基,允许整个蛋白质实施需要组分亚基紧密接近的功能。此类接头的非限制性实例包括甘氨酸-丝氨酸(GS)接头(不可切割的);以及F2A接头、P2A接头、T2A接头和E2A接头(可切割的)。本文可使用其他连接。

[0278] 在一些实施方案中,所述接头是GS接头。GS接头是包括甘氨酸和丝氨酸氨基酸重复的多肽接头。其包含柔性和亲水性残基,并且可用于实施蛋白质亚基的融合,而不干扰蛋白质结构域的折叠和功能,并且不形成二级结构。在一些实施方案中,mRNA编码包含3至20个氨基酸长的GS接头的融合蛋白。例如,所述GS接头可具有3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个氨基酸的长度(或具有至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个氨基酸的长度)。在一些实施方案中,GS接头是(或至少是)15个氨基酸长(例如,GGSGGSGGSGGSGGG(SEQ ID NO:133))。在一些实施方案中,GS接头是(或至少是)8个氨基酸长(例如,GGGSGGGS(SEQ ID NO:134))。在一些实施方案中,GS接头是(或至少是)7个氨基酸长(例如,GGGSGGG(SEQ ID NO:135))。在一些实施方案中,GS接头是(或至少是)4个氨基酸长(例如,GGGS(SEQ ID NO:136))。在一些实施方案中,GS接头包含(GGGS)n(SEQ ID NO:136),其中n为1-5的任何整数。在一些实施方案中,GS接头是(或至少是)4个氨基酸长(例如,GSGG(SEQ ID NO:152))。在一些实施方案中,所述GS接头包含(GSGG)n(SEQ ID NO:152),其中n为1-5的任何整数。

[0279] 在一些实施方案中,接头是甘氨酸接头,例如具有3个氨基酸的长度(或至少3个氨基酸的长度)(例如,GGG)。

[0280] 在一些实施方案中,由mRNA疫苗编码的蛋白质包括一个以上接头,其可彼此相同或不同(例如,在同一S蛋白构建体中的GGGSGGG(SEQ ID NO:135)和GGGS(SEQ ID NO:136))。

[0281] 在一些实施方案中,接头包含编码泛HLA DR结合表位(PADRE)的mRNA(例如,AKFVAAWTLKAAA(SEQ ID NO:148))。PADRE是免疫显性辅助CD4T细胞表位和强效免疫原(参见,例如Alexander J.等人,J of Immuno.164(3):1625-33,通过引用并入本文)。

[0282] 表9A.结构域融合抗原

[0283]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与跨膜结构域连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM)	91	92
与跨膜结构域连接的 SARS-CoV-2 RBD-NTD (RBD-NTD-TM)	139	140

[0284]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与跨膜结构域连接的 SARS-CoV-2 NTD-PADRE-RBD (NTD-PADRE-RBD-TM)	142	143

[0285] 表9B.结构域融合抗原

[0286]

与跨膜结构域连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM)		
SEQ ID NO: 90 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 91 和 3' UTR SEQ ID NO: 4。		90
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUA AGAAGAAAUAUAAGACCCCGGCGCCGC CACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCC CUGGUGAGCAGCCAGUGCGUGAACCUG ACCACCCGGACCCAGCUGCCACCAGCCU ACACCAACAGCUUCACCCGGGGCGUCU ACUACCCCGACAAGGUGUUCCGGAGCA GCGUCCUGCACAGCACCCAGGACCUGU UCCUGCCCUUCUUCAGCAACGUGACCU GGUUCCACGCCAUCCACGUGAGCGGCA CCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGU ACUUCGCCAGCACCGAGAAGAGCAACA UCAUCCGGGGCUGGAUCUUCGGCACCA CCCUGGACAGCAAGACCCAGAGCCUGC UGAUCGUGAAUAACGCCACCAACGUGG UGAUCAAGGUGUGCGAGUUCCAGUUCU GCAACGACCCCUUCCUGGGCGUGUACU ACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCG CCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGC GGGAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCACA CCCCAAUCAACCUGGUGCGGGAUCUGC	91

[0287]

	CCCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAACA UCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCG ACAGCAGCAGCGGGUGGACAGCAGGCG CGGCUGCUUACUACGUGGGCUACCUGC AGCCCCGGACCUUCCUGCUGAAGUACA ACGAGAACGGCACCAUCACCGACGCCG UGGACGGAGGCGGAUCGGGAGGCGGAC CCAACAUCACCAACCUGUGCCCCUUCG GCGAGGUGUUCAACGCCACCCGGUUCG CCAGCGUGUACGCCUGGAACCGGAAGC GGAUCAGCAACUGCGUGGCCGACUACA GCGUGCUGUACAACAGCGCCAGCUUCA GCACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGA UCCGUGGCGACGAGGUGCGGCAGAUCG CACCCGGCCAGACAGGCAAGAUCGCCG ACUACAACUACAAGCUGCCCGACGACU UCACCGGCUGCGUGAUCGCCUGGAACA GCAACAACCUCGACAGCAAGGUGGGCG GCAACUACAACUACCUGUACCGGCUGU UCCGGAAGAGCAACCUGAAGCCCUUCG AGCGGGACAUCAGCACCGAGAUCUACC AAGCCGGCUCCACCCCUUGCAACGGCG UGGAGGGCUUCAACUGCUACUUCCCUC UGCAGAGCUACGGCUUCCAGCCCACCA ACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGC ACGCCCCAGCCACCGUGUGUGGCCCCA AGUCUGGCGGAGGCAGCAUCCUGGCCA UCUACAGCACCGUGGCCAGCAGCCUGG UGCUGCUGGUGAGCCUGGGCGCCAUCA GCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCC UAGCUUCUUGCCCCUUGGGCCUCCCCC CAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYT NSFTRGVYYPDKVFRSSVLHSTQDLFLPFF SNVTWFHAIHVSGTNGTKRFDNPVLPFND	92

	GVYFASTEKSNIIRGWIFGTTLDSKTQSLLI VNNATNVVIKVCEFQFCNDPFLGVYYHK NNKSWMESEFRVYSSANNCTFEYVSQPFL MDLEGKQGNFKNLREFVFKNIDGYFKIYS KHTPINLVRDLPQGFSALEPLVDLPIGINIT RFQTLLALHRSYLTPGDSSSGWTAGAAAY YVGYLQPRTFLLKYNENGTITDAVDgggsgg gPNITNLCPFGEVFNATRFASVYAWNRKRI SNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVRQIAPGQTG KIADYNYKLPDDFTGCVIAWNSNNLDSKV GGNYNYLYRLFRKSNLKPFERDISTEIYQA GSTPCNGVEGFNCYFPLQSYGFQPTNGVG YQPYRVVVLSFELLHAPATVCGPKsgggsilai ystvasslvllvslgaisf	
多聚 A 尾	100 nt	
与跨膜结构域连接的 SARS-CoV-2 RBD-NTD (RBD-NTD-TM)		
SEQ ID NO: 141 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 139 和 3' UTR SEQ ID NO: 4。		141
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUA AGAAGAAAUAUAAGACCCCGGCGCCGC CACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCUGC GUGACCCUGACCCUGGUGCUGCUGGUG AACAGCCAGCCCAACAUCACCAACCUG UGCCCCUUCGGCGAGGUGUUCAACGCC ACCCGGUUCGCCAGCGUGUACGCCUGG AACCGGAAGCGGAUCAGCAACUGCGUG GCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUAC GGCGUGAGCCCCACCAAGCUGAACGAC CUGUGCUUCACCAACGUGUACGCCGAC AGCUUCGUGAUCCGUGGCGACGAGGUG CGGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGCUG CCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGC AAGGUGGGCGGCAACUACAACUACCUG UACCGGCUGUUCCGGAAGAGCAACCUG	139

[0288]

[0289]

	AAGCCCUUCGAGCGGGACAUCAGCACC GAGAUCUACCAAGCCGGCUCCACCCCU UGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCUCUGCAGAGCUACGGCUUC CAGCCCACCAACGGCGUGGGCUACCAG CCCUACCGGGUGGUGGUGCUGAGCUUC GAGCUGCUGCACGCCCCAGCCACCGUG UGUGGCCCCAAGUCUGGCGGAGGCAGC GGAGGCGGAAGCCAGUGCGUGAACCUG ACCACCCGGACCCAGCUGCCACCAGCCU ACACCAACAGCUUCACCCGGGGCGUCU ACUACCCCGACAAGGUGUUCCGGAGCA GCGUCCUGCACAGCACCCAGGACCUGU UCCUGCCCUUCUUCAGCAACGUGACCU GGUUCCACGCCAUCCACGUGAGCGGCA CCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGU ACUUCGCCAGCACCGAGAAGAGCAACA UCAUCCGGGGCUGGAUCUUCGGCACCA CCCUGGACAGCAAGACCCAGAGCCUGC UGAUCGUGAAUAACGCCACCAACGUGG UGAUCAAGGUGUGCGAGUUCCAGUUCU GCAACGACCCCUUCCUGGGCGUGUACU ACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCG CCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGC GGGAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCACA CCCCAAUCAACCUGGUGCGGGAUCUGC CCCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAACA UCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCG ACAGCAGCAGCGGGUGGACAGCAGGCG CGGCUGCUUACUACGUGGGCUACCUGC AGCCCCGGACCUUCCUGCUGAAGUACA ACGAGAACGGCACCAUCACCGACGCCG UGGACUCUGGCGGAGGCAGCAUCCUGG CCAUCUACAGCACCGUGGCCAGCAGCC UGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	

[0290]

3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCC UAGCUUCUUGCCCCUUGGGCCUCCCCC CAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGC	4
相应的氨基酸序列	MYSMQLASCVTLTLVLLVNSQPNITNLCP FGEVFNATRFASVYAWNRKRISNCVADYS VLYNSASFSTFKCYGVSPTKLNDLCFTNV YADSFVIRGDEVRQIAPGQTGKIADYNYK LPDDFTGCVIAWNSNNLDSKVGGNYNYL YRLFRKSNLKPFERDISTEIYQAGSTPCNG VEGFNCYFPLQSYGFQPTNGVGYQPYRVV VLSFELLHAPATVCGPKSGGGSGGGSQCV NLTTRTQLPPAYTNSFTRGVYYPDKVFRSS VLHSTQDLFLPFFSNVTWFHAIHVSGTNGT KRFDNPVLPFNDGVYFASTEKSNIIRGWIF GTTLDSKTQSLLIVNNATNVVIKVCEFQFC NDPFLGVYYHKNNKSWMESEFRVYSSAN NCTFEYVSQPFLMDLEGKQGNFKNLREFV FKNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPGDS SSGWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDSGGGSILAIYSTVASSLVLLVSL GAISF	140
多聚 A 尾	100 nt	
与跨膜结构域连接的 SARS-CoV-2 NTD-PADRE-RBD (NTD-PADRE-RBD-TM)		
SEQ ID NO: 144 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 142 和 3' UTR SEQ ID NO: 4。		144
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUA AGAAGAAAUAUAAGACCCCGGCGCCGC CACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCC CUGGUGAGCAGCCAGUGCGUGAACCUG ACCACCCGGACCCAGCUGCCACCAGCCU ACACCAACAGCUUCACCCGGGGCGUCU ACUACCCCGACAAGGUGUUCCGGAGCA GCGUCCUGCACAGCACCCAGGACCUGU UCCUGCCCUUCUUCAGCAACGUGACCU	142

[0291]

	GGUUCCACGCCAUCCACGUGAGCGGCA CCAACGGCACCAAGCGGUUCGACAACC CCGUGCUGCCCUUCAACGACGGCGUGU ACUUCGCCAGCACCGAGAAGAGCAACA UCAUCCGGGGCUGGAUCUUCGGCACCA CCCUGGACAGCAAGACCCAGAGCCUGC UGAUCGUGAAUAACGCCACCAACGUGG UGAUCAAGGUGUGCGAGUUCCAGUUCU GCAACGACCCCUUCCUGGGCGUGUACU ACCACAAGAACAACAAGAGCUGGAUGG AGAGCGAGUUCCGGGUGUACAGCAGCG CCAACAACUGCACCUUCGAGUACGUGA GCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGC GGGAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCACA CCCCAAUCAACCUGGUGCGGGAUCUGC CCCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAACA UCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCG ACAGCAGCAGCGGGUGGACAGCAGGCG CGGCUGCUUACUACGUGGGCUACCUGC AGCCCCGGACCUUCCUGCUGAAGUACA ACGAGAACGGCACCAUCACCGACGCCG UGGACGGAGGCGCCAAGUUCGUGGCCG CCUGGACUCUGAAGGCCGCAGCCGGCG GACCCAACAUCACCAACCUGUGCCCCU UCGGCGAGGUGUUCAACGCCACCCGGU UCGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGACU ACAGCGUGCUGUACAACAGCGCCAGCU UCAGCACCUUCAAGUGCUACGGCGUGA GCCCCACCAAGCUGAACGACCUGUGCU UCACCAACGUGUACGCCGACAGCUUCG UGAUCCGUGGCGACGAGGUGCGGCAGA UCGCACCCGGCCAGACAGGCAAGAUCG CCGACUACAACUACAAGCUGCCCGACG ACUUCACCGGCUGCGUGAUCGCCUGGA ACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGC UGUUCCGGAAGAGCAACCUGAAGCCCU UCGAGCGGGACAUCAGCACCGAGAUCU	

[0292]

	ACCAAGCCGGCUCCACCCCUUGCAACG GCGUGGAGGGCUUCAACUGCUACUUCC CUCUGCAGAGCUACGGCUUCCAGCCCA CCAACGGCGUGGGCUACCAGCCCUACC GGGUGGUGGUGCUGAGCUUCGAGCUGC UGCACGCCCCAGCCACCGUGUGUGGCC CCAAGUCUGGCGGAGGCAGCAUCCUGG CCAUCUACAGCACCGUGGCCAGCAGCC UGGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCC UAGCUUCUUGCCCCUUGGGCCUCCCCC CAGCCCCUCCUCCCCUUCCUGCACCCGU ACCCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYT NSFTRGVYYPDKVFRSSVLHSTQDLFLPFF SNVTWFHAIHVSGTNGTKRFDNPVLPFND GVYFASTEKSNIIRGWIFGTTLDSKTQSLLI VNNATNVVIKVCEFQFCNDPFLGVYYHK NNKSWMESEFRVYSSANNCTFEYVSQPFL MDLEGKQGNFKNLREFVFKNIDGYFKIYS KHTPINLVRDLPQGFSALEPLVDLPIGINIT RFQTLLALHRSYLTPGDSSSGWTAGAAAY YVGYLQPRTFLLKYNENGTITDAVDGGAK FVAAWTLKAAAGGPNITNLCPFGEVFNAT RFASVYAWNRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRG DEVRQIAPGQTGKIADYNYKLPDDFTGCV IAWNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVVLSFELLHA PATVCGPKSGGGSILAIYSTVASSLVLLVS LGAISF	143
多聚 A 尾	100 nt	

[0293] 运输信号

[0294] 在一些实施方案中，mRNA编码与高尔基体(Golgi)运输信号连接的SARS-CoV-2 S蛋白结构域(例如，NTD、RBD或NTD-RBD融合体)。此类信号的非限制性实例包括巨噬细胞标志物，例如CD86和/或CD11b，其是高度表达的，并且细胞内区可控制从高尔基体至细胞表面的有效输出。本文可使用其他细胞运输信号(序列)，例如VSV-G胞质尾(VSVGct)。预计将编码的蛋白质更有效地运输至细胞表面会增加B细胞识别的抗原可用性，并因此促进产生针对编码的SARS-CoV-2 S蛋白结构域的抗体。与运输信号连接的SARS-CoV-2抗原和编码其的mRNA的非限制性实例提供于下文表10A和表10B中。

[0295] 表10A.与运输信号连接的结构域融合抗原

[0296]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与跨膜结构域和 huCD86 连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM-CD86)	94	95
与跨膜结构域和 huCD11B 连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM-CD11B)	97	98
与 VSVGct 连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-VSVGct)	109	110

[0297] 表10B.与运输信号连接的结构域融合抗原

[0298]

与跨膜结构域和 huCD86 连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM-CD86)		
SEQ ID NO: 93 从 5'端至 3'端由以下组成:5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 94 和 3' UTR SEQ ID NO: 4。		93
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCGACAACCCCGUGCUGCCCU	94

[0299]

	UCAACGACGGCGUGUACUUCGCCAGCACCGAG AAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGA UCAAGGUGUGCGAGUUCCAGUUCUGCAACGA CCCCUUCCUGGGCGUGUACUACCACAAGAACA ACAAGAGCUGGAUGGAGAGCGAGUUCCGGGU GUACAGCAGCGCCAACAACUGCACCUUCGAGU ACGUGAGCCAGCCCUUCCUGAUGGACCUGGAG GGCAAGCAGGGCAACUUCAAGAACCUGCGGG AGUUCGUGUUCAAGAACAUCGACGGCUACUU CAAGAUCUACAGCAAGCACACCCCAAUCAACC UGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCC CUGGAGCCCCUGGUGGACCUGCCCAUCGGCAU CAACAUCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCGACAGC AGCAGCGGGUGGACAGCAGGCGCGGCUGCUU ACUACGUGGGCUACCUGCAGCCCCGGACCUUC CUGCUGAAGUACAACGAGAACGGCACCAUCAC CGACGCCGUGGACGGAGGCGGAUCGGGAGGC GGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCG UGUACGCCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUACAACAG CGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUC ACCAACGUGUACGCCGACAGCUUCGUGAUCCG UGGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGCC CCAGCCACCGUGUGUGGCCCCAAGUCUGGCGG AGGCAGCAUCCUGGCCAUCUACAGCACCGUGG CCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGC GCCAUCAGCUUCAAGAAGAAGAAGCGGCCACG GAACUCCUACAAGUGCGGCACCAACACCAUGG AGCGGGAGGAGAGCGAGCAGACCAAGAAGCG GGAGAAGAUCCACAUUCCUGAACGGUCCGACG AAGCCCAGCGGGUGUUCAAGAGCAGCAAGACC AGCAGCUGCGACAAGAGCGACACCUGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGC UUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGUACCCCCGUGGUCU	4

	UUGAAUAAAGUCUGAGUGGGCGGC	
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRG WIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYV SQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYS KHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLL ALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTF LLKYNENGTITDAVDgggsgggPNITNLCPFGEVFNA TRFASVYAWNRKRISNCVADYSVLYNSASFSTFK CYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVCGPKsgggsilaiystvasslvllvslgaisfKK KKRPRNSYKCGTNTMEREESEQTKKREKIHIPERS DEAQRVFKSSKTSSCDKSDTCF	95
多聚 A 尾	100 nt	
与跨膜结构域和 huCD11B 连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM-CD11B)		
SEQ ID NO: 96 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 97 和 3' UTR SEQ ID NO: 4。		96
[0300] 化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAG AAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGU GAGCAGCCAGUGCGUGAACCUGACCACCCGGA CCCAGCUGCCACCAGCCUACACCAACAGCUUC ACCCGGGGCGUCUACUACCCCGACAAGGUGUU CCGGAGCAGCGUCCUGCACAGCACCCAGGACC UGUUCCUGCCCUUCUUCAGCAACGUGACCUGG UUCCACGCCAUCCACGUGAGCGGCACCAACGG CACCAAGCGGUUCGACAACCCCGUGCUGCCCU UCAACGACGGCGUGUACUUCGCCAGCACCGAG AAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGA UCAAGGUGUGCGAGUUCCAGUUCUGCAACGA CCCCUUCCUGGGCGUGUACUACCACAAGAACA ACAAGAGCUGGAUGGAGAGCGAGUUCCGGGU GUACAGCAGCGCCAACAACUGCACCUUCGAGU ACGUGAGCCAGCCCUUCCUGAUGGACCUGGAG GGCAAGCAGGGCAACUUCAAGAACCUGCGGG AGUUCGUGUUCAAGAACAUCGACGGCUACUU CAAGAUCUACAGCAAGCACACCCCAAUCAACC UGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCC	97

[0301]

	CUGGAGCCCCUGGUGGACCUGCCCAUCGGCAU CAACAUCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCGACAGC AGCAGCGGGUGGACAGCAGGCGCGGCUGCUU ACUACGUGGGCUACCUGCAGCCCCGGACCUUC CUGCUGAAGUACAACGAGAACGGCACCAUCAC CGACGCCGUGGACGGAGGCGGAUCGGGAGGC GGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCG UGUACGCCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUACAACAG CGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUC ACCAACGUGUACGCCGACAGCUUCGUGAUCCG UGGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGCC CCAGCCACCGUGUGUGGCCCCAAGUCUGGCGG AGGCAGCAUCCUGGCCAUCUACAGCACCGUGG CCAGCAGCCUGGUGCUGCUGGUGAGCCUGGGC GCCAUCAGCUUCAAGCGGCAGUACAAGGACAU GAUGAGCGAGGGAGGACCACCUGGCGCUGAG CCACAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGC UUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRG WIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYV SQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYS KHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLL ALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTF LLKYNENGTITDAVDgggsgggPNITNLCPFGEVFNA TRFASVYAWNRKRISNCVADYSVLYNSASFSTFK CYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVCGPKsgggsilaiystvasslvllvslgaisfKR	98

[0302]

	QYKDMMSEGGPPGAEPQ	
多聚 A 尾	100 nt	
与 VSVGct 连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-VSVGct)		
SEQ ID NO: 108 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 109 和 3' UTR SEQ ID NO: 4。		108
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAG AAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGU GAGCAGCCAGUGCGUGAACCUGACCACCCGGA CCCAGCUGCCACCAGCCUACACCAACAGCUUC ACCCGGGGCGUCUACUACCCCGACAAGGUGUU CCGGAGCAGCGUCCUGCACAGCACCCAGGACC UGUUCCUGCCCUUCUUCAGCAACGUGACCUGG UUCCACGCCAUCCACGUGAGCGGCACCAACGG CACCAAGCGGUUCGACAACCCCGUGCUGCCCU UCAACGACGGCGUGUACUUCGCCAGCACCGAG AAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGA UCAAGGUGUGCGAGUUCCAGUUCUGCAACGA CCCCUUCCUGGGCGUGUACUACCACAAGAACA ACAAGAGCUGGAUGGAGAGCGAGUUCCGGGU GUACAGCAGCGCCAACAACUGCACCUUCGAGU ACGUGAGCCAGCCCUUCCUGAUGGACCUGGAG GGCAAGCAGGGCAACUUCAAGAACCUGCGGG AGUUCGUGUUCAAGAACAUCGACGGCUACUU CAAGAUCUACAGCAAGCACACCCCAAUCAACC UGGUGCGGGAUCUGCCCCAGGGCUUCUCAGCC CUGGAGCCCCUGGUGGACCUGCCCAUCGGCAU CAACAUCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCGACAGC AGCAGCGGGUGGACAGCAGGCGCGGCUGCUU ACUACGUGGGCUACCUGCAGCCCCGGACCUUC CUGCUGAAGUACAACGAGAACGGCACCAUCAC CGACGCCGUGGACGGAGGCGGAUCGGGAGGC GGACCCAACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUCGCCAGCG UGUACGCCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUACAACAG CGCCAGCUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUGUGCUUC ACCAACGUGUACGCCGACAGCUUCGUGAUCCG UGGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG	109

	GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGCC CCAGCCACCGUGUGUGGCCCCAAGUCUGGCGG AGGCAGCAGCAGCAUCGCCAGCUUCUUCUUCA UCAUCGGGCUGAUCAUCGGCCUCUUCCUGGUG CUGCGGGUGGGCAUCCACCUGUGCAUCAAGCU GAAGCACACCAAGAAGAGACAGAUCUACACCG ACAUCGAGAUGAACCGGCUGGGCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGC UUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUC CUCCCCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRG WIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYV SQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYS KHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLL ALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTF LLKYNENGTITDAVDgggsgggPNITNLCPFGEVFNA TRFASVYAWNRKRISNCVADYSVLYNSASFSTFK CYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVCGPKsgggsSSIASFFFIIGLIIGLFLV LRVGIHLCIKLKHTKKRQIYTDIEMNRLGK	110
多聚 A 尾	100 nt	

[0303]

[0304] 结构域融合C末端截短

[0305] 在其他实施方案中，本文提供的mRNA编码SARS-CoV-2 NTD-RBD融合蛋白，其中C末端结构域的某一部分已经截短/缺失。在一个实施方案中，NTD-RBD融合蛋白的C末端结构域缺失13个(或至少13个)氨基酸。预计这些氨基酸的缺失会增加表位于抗体中的暴露，从而刺激对NTD和RBD结构域上存在的保护性表位的更稳固的免疫反应。

[0306] 具有C末端截短的SARS-CoV-2结构域融合抗原和编码其的mRNA的非限制性实例提供于下文表11A和表11B中。

[0307] 表11A.结构域融合C末端截短

[0308]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
具有 13 个氨基酸的 C 末端截短的 SARS-CoV-2 NTD-RBD (NTD-RBD-Δ13)	106	107

[0309] 表11B.结构域融合C末端截短

具有 13 个氨基酸的 C 末端截短的 SARS-CoV-2 NTD-RBD (NTD-RBD-Δ13)		
SEQ ID NO: 105 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 106 和 3' UTR SEQ ID NO: 4。		105
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGG	106

[0310]

[0311]

	UGUACAGCAGCGCCAACAACUGCACCUU CGAGUACGUGAGCCAGCCCUUCCUGAUG GACCUGGAGGGCAAGCAGGGCAACUUCA AGAACCUGCGGGAGUUCGUGUUCAAGA ACAUCGACGGCUACUUCAAGAUCUACAG CAAGCACACCCCAAUCAACCUGGUGCGG GAUCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGCAU CAACAUCACCCGGUUCCAGACCCUGCUG GCCCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCAGG CGCGGCUGCUUACUACGUGGGCUACCUG CAGCCCCGGACCUUCCUGCUGAAGUACA ACGAGAACGGCACCAUCACCGACGCCGU GGACGGAGGCGGAUCGGGAGGCGGACCC AACAUCACCAACCUGUGCCCCUUCGGCG AGGUGUUCAACGCCACCCGGUUCGCCAG CGUGUACGCCUGGAACCGGAAGCGGAUC AGCAACUGCGUGGCCGACUACAGCGUGC UGUACAACAGCGCCAGCUUCAGCACCUU CAAGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGUGU ACGCCGACAGCUUCGUGAUCCGUGGCGA CGAGGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACUACA AGCUGCCCGACGACUUCACCGGCUGCGU GAUCGCCUGGAACAGCAACAACCUCGAC AGCAAGGUGGGCGGCAACUACAACUACC UGUACCGGCUGUUCCGGAAGAGCAACCU GAAGCCCUUCGAGCGGGACAUCAGCACC GAGAUCUACCAAGCCGGCUCCACCCCUU GCAACGGCGUGGAGGGCUUCAACUGCUA CUUCCCUCUGCAGAGCUACGGCUUCCAG CCCACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAGCU GCUGCACGCCCCAGCCACCGUGUGUGGC CCCAAGUCUGGCGGAGGCAGCAUCCUGG CCAUCUACAGCCUGGGCUUCAUCGCCGG CCUGAUCGCCAUCGUGAUGGUGACCAUC AUGCUGUGCUGCAUGACCAGCUGCUGCA GCUGCCUGAAGGGCUGUUGCAGCUGCGG CAGCUGCUGCAAGUUCGACGAGGACGAC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCU	4

[0312]

	AGCUUCUUGCCCCUUGGGCCUCCCCCCA GCCCCUCCUCCCCUUCCUGCACCCGUAC CCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGC	
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTN SFTRGVYYPDKVFRSSVLHSTQDLFLPFFSN VTWFHAIHVSGTNGTKRFDNPVLPFNDGV YFASTEKSNIIRGWIFGTTLDSKTQSLLIVNN ATNVVIKVCEFQFCNDPFLGVYYHKNNKS WMESEFRVYSSANNCTFEYVSQPFLMDLE GKQGNFKNLREFVFKNIDGYFKIYSKHTPIN LVRDLPQGFSALEPLVDLPIGINITRFQTLLA LHRSYLTPGDSSSGWTAGAAAYYVGYLQP RTFLLKYNENGTITDAVDgggsgggPNITNLCP FGEVFNATRFASVYAWNRKRISNCVADYS VLYNSASFSTFKCYGVSPTKLNDLCFTNVY ADSFVIRGDEVRQIAPGQTGKIADYNYKLP DDFTGCVIAWNSNNLDSKVGGNYNYLYRL FRKSNLKPFERDISTEIYQAGSTPCNGVEGF NCYFPLQSYGFQPTNGVGYQPYRVVVLSFE LLHAPATVCGPKsgggsilaiysLGFIAGLIAIVM VTIMLCCMTSCCSCLKGCCSCGSCCKFDED D	107
多聚 A 尾	100 nt	

[0313] 结构域延伸

[0314] 在一些实施方案中，SARS-CoV-2 S蛋白结构域抗原包括“延伸”区，其包括与本领域理解的NTD结构域或RBD结构域相邻和/或侧接的序列。RBD_EXT系列涵盖SD1(子结构域1)。NTD_EXT系列涵盖NTD中的C末端螺旋。一些B细胞和抗体识别仅在正确折叠而非变性形式的SARS-CoV-2 S蛋白NTD和RBD中发现的构象表位。包括与NTD和RBD结构域相邻和/或侧接的序列不仅可为抗原提供额外B细胞表位，并且可潜在地导致那些结构域的更优化折叠，并用对可在任一结构域边缘上发现的表位具特异性的抗体刺激B细胞。此外，包括这些延伸序列可因此增加NTD或RBD与表达细胞膜之间的距离，从而增加两个结构域于抗体中的暴露，如果表达的蛋白质太靠近细胞表面，则抗体结合效率较低。最后，包括延伸序列会增加肽的集合，所述肽可潜在地由识别NTD或RBD表位的B细胞呈递给$CD4^+$T细胞，然后加工完整蛋白用于抗原呈递，从而增加NTD或RBD特异性B细胞接受足够T细胞辅助的机会。SARS-CoV-2结构域延伸和编码其的mRNA的非限制性实例提供于下文表12A和表12B中。

[0315] 表12.结构域延伸

[0316]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
SARS-CoV-2 NTD DS 延伸区连接至跨膜结构域(NTD-EXT-F43C-TM)	52	53
SARS-CoV-2 NTD DS 延伸区连接至跨膜结构域(NTD-F43C-EXT-TM)	55	56
SARS-CoV-2 NTD 延伸区连接至跨膜结构域(NTD-EXT-TM)	58	59
SARS-CoV-2 RBD 延伸区连接至跨膜结构域(RBD-EXT-TM)	85	86
SARS-CoV-2 RBD DS 延伸区连接至跨膜结构域(RBD-Q563D-EXT-TM)	88	89
SARS-CoV-2 NTD-RBD 延伸区连接至跨膜结构域(NTD-RBD-EXT-TM)	115	116
SARS-CoV-2 NTD 延伸区-RBD 连接至跨膜结构域(NTD-EXT-RBD-TM)	118	119
SARS-CoV-2 NTD 延伸区-RBD-延伸区连接至跨膜结构域 (NTD-EXT-RBD-EXT-TM)	121	122

[0317] 表12B.结构域延伸

[0318]

SARS-CoV-2 NTD DS 延伸区连接至跨膜结构域 (NTD-EXT-F43C-TM)		
SEQ ID NO: 51 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 52 和 3' UTR SEQ ID NO: 4。		51
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	

[0319]

5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUGCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUG GAUCUUCGGCACCACCCUGGACAGCA AGACCCAGAGCCUGCUGAUCGUGAAU AACGCCACCAACGUGGUGAUCAAGGU GUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAG AACAACAAGAGCUGGAUGGAGAGCG AGUUCCGGGUGUACAGCAGCGCCAAC AACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGG GAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCAC ACCCCAAUCAACCUGGUGCGGGAUCU GCCCCAGGGCUUCUCAGCCCUGGAGC CCCUGGUGGACCUGCCCAUCGGCAUC AACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCC CAGGCGACAGCAGCAGCGGGUGGACA GCAGGCGCGGCUGCUUACUACGUGGG CUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUC ACCGACGCCGUGGACUCUGGCGGAGG CAGCAUCCUGGCCAUCUACAGCACCG UGGCCAGCAGCCUGGUGCUGCUGGUG AGCCUGGGCGCCAUCAGCUUC	52
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA	4

	CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVCRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPG DSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDsgggsilaiystvasslvllvslgaisf	53
多聚 A 尾	100 nt	
SARS-CoV-2 NTD DS 延伸区连接至跨膜结构域 (NTD-F43C-EXT-TM)		
SEQ ID NO: 54 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 55 和 3' UTR SEQ ID NO: 4。		54
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUGCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUG GAUCUUCGGCACCACCCUGGACAGCA AGACCCAGAGCCUGCUGAUCGUGAAU AACGCCACCAACGUGGUGAUCAAGGU GUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAG AACAACAAGAGCUGGAUGGAGAGCG	55

[0320]

		AGUUCCGGGUGUACAGCAGCGCCAAC AACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGG GAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCAC ACCCCAAUCAACCUGGUGCGGGAUCU GCCCCAGGGCUUCUCAGCCCUGGAGC CCCUGGUGGACCUGCCCAUCGGCAUC AACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCC CAGGCGACAGCAGCAGCGGGUGGACA GCAGGCGCGGCUGCUUACUACGUGGG CUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUC ACCGACGCCGUGGACUGCGCCCUGGA CCCUCUGAGCGAGACCAAGUGCACCC UGAAGAGCUUCACCUCUGGCGGAGGC AGCAUCCUGGCCAUCUACAGCACCGU GGCCAGCAGCCUGGUGCUGCUGGUGA GCCUGGGCGCCAUCAGCUUC	
[0321]	3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
	相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVCRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPG DSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDCALDPLSETKCTLKSFTs gggsilaiystvasslvllvslgaisf	56
	多聚 A 尾	100 nt	
SARS-CoV-2 NTD 延伸区连接至跨膜结构域(NTD-EXT-TM)			
SEQ ID NO: 57 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 58 和 3' UTR SEQ ID NO:			57

[0322]

4。		
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUG GAUCUUCGGCACCACCCUGGACAGCA AGACCCAGAGCCUGCUGAUCGUGAAU AACGCCACCAACGUGGUGAUCAAGGU GUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAG AACAACAAGAGCUGGAUGGAGAGCG AGUUCCGGGUGUACAGCAGCGCCAAC AACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGG GAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCAC ACCCCAAUCAACCUGGUGCGGGAUCU GCCCCAGGGCUUCUCAGCCCUGGAGC CCCUGGUGGACCUGCCCAUCGGCAUC AACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCC CAGGCGACAGCAGCAGCGGGUGGACA GCAGGCGCGGCUGCUUACUACGUGGG CUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUC ACCGACGCCGUGGACUGCGCCCUGGA CCCUCUGAGCGAGACCAAGUGCACCC UGAAGAGCUUCACCUCUGGCGGAGGC	58

	AGCAUCCUGGCCAUCUACAGCACCGU GGCCAGCAGCCUGGUGCUGCUGGUGA GCCUGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPG DSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDCALDPLSETKCTLKSFTs gggsilaiystvasslvllvslgaisf	59
多聚 A 尾	100 nt	

[0323] **SARS-CoV-2 RBD 延伸区连接至跨膜结构域(RBD-EXT-TM)**

SEQ ID NO: 84 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 85 和 3' UTR SEQ ID NO: 4。		84
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCU GCGUGACCCUGACCCUGGUGCUGCUG GUGAACAGCCAGCGGGUGCAGCCCAC CGAGAGCAUCGUGCGGUUCCCCAACA UCACCAACCUGUGCCCCUUCGGCGAG GUGUUCAACGCCACCCGGUUCGCCAG CGUGUACGCCUGGAACCGGAAGCGGA UCAGCAACUGCGUGGCCGACUACAGC GUGCUGUACAACAGCGCCAGCUUCAG CACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUC ACCAACGUGUACGCCGACAGCUUCGU	85

[0324]

	GAUCCGUGGCGACGAGGUGCGGCAG AUCGCACCCGGCCAGACAGGCAAGAU CGCCGACUACAACUACAAGCUGCCCG ACGACUUCACCGGCUGCGUGAUCGCC UGGAACAGCAACAACCUCGACAGCAA GGUGGGCGGCAACUACAACUACCUGU ACCGGCUGUUCCGGAAGAGCAACCUG AAGCCCUUCGAGCGGGACAUCAGCAC CGAGAUCUACCAAGCCGGCUCCACCC CUUGCAACGGCGUGGAGGGCUUCAAC UGCUACUUCCCUCUGCAGAGCUACGG CUUCCAGCCCACCAACGGCGUGGGCU ACCAGCCCUACCGGGUGGUGGUGCUG AGCUUCGAGCUGCUGCACGCCCCAGC CACCGUGUGUGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGU GAACUUCAACUUCAACGGCCUUACCG GCACCGGCGUGCUGACCGAGAGCAAC AAGAAAUUCCUGCCCUUUCAGCAGUU CGGCCGGGACAUCGCCGACACCACCG ACGCUGUGCGGGAUCCCCAGACCCUG GAGAUCCUGGACAUCACCCCUUGCAG CUCUGGCGGAGGCAGCAUCCUGGCCA UCUACAGCACCGUGGCCAGCAGCCUG GUGCUGCUGGUGAGCCUGGGCGCCAU CAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
相应的氨基酸序列	mysmqlascvtltlvllvnsQRVQPTESIVRFPNIT NLCPFGEVFNATRFASVYAWNRKRISN CVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTG KIADYNYKLPDDFTGCVIAWNSNNLDS KVGGNYNYLYRLFRKSNLKPFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGFQP TNGVGYQPYRVVVLSFELLHAPATVCG PKKSTNLVKNKCVNFNFNGLTGTGVLT ESNKKFLPFQQFGRDIADTTDAVRDPQT LEILDITPCSsgggsilaiystvasslvllvslgaisf	86
多聚 A 尾	100 nt	

[0325]

SARS-CoV-2 RBD DS 延伸区连接至跨膜结构域 (RBD-Q563D-EXT-TM)		
SEQ ID NO: 87 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 88 和 3' UTR SEQ ID NO: 4。		87
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCU GCGUGACCCUGACCCUGGUGCUGCUG GUGAACAGCCAGCGGGUGCAGCCCAC CGAGAGCAUCGUGCGGUUCCCCAACA UCACCAACCUGUGCCCCUUCGGCGAG GUGUUCAACGCCACCCGGUUCGCCAG CGUGUACGCCUGGAACCGGAAGCGGA UCAGCAACUGCGUGGCCGACUACAGC GUGCUGUACAACAGCGCCAGCUUCAG CACCUUCAAGUGCUACGGCGUGAGCC CCACCAAGCUGAACGACCUGUGCUUC ACCAACGUGUACGCCGACAGCUUCGU GAUCCGUGGCGACGAGGUGCGGCAG AUCGCACCCGGCCAGACAGGCAAGAU CGCCGACUACAACUACAAGCUGCCCG ACGACUUCACCGGCUGCGUGAUCGCC UGGAACAGCAACAACCUCGACAGCAA GGUGGGCGGCAACUACAACUACCUGU ACCGGCUGUUCCGGAAGAGCAACCUG AAGCCCUUCGAGCGGGACAUCAGCAC CGAGAUCUACCAAGCCGGCUCCACCC CUUGCAACGGCGUGGAGGGCUUCAAC UGCUACUUCCCUCUGCAGAGCUACGG CUUCCAGCCCACCAACGGCGUGGGCU ACCAGCCCUACCGGGUGGUGGUGCUG AGCUUCGAGCUGCUGCACGCCCCAGC CACCGUGUGUGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGU GAACUUCAACUUCAACGGCCUUACCG GCACCGGCGUGCUGACCGAGAGCAAC AAGAAAUUCCUGCCCUUUUGCCAGUU CGGCCGGGACAUCGCCGACACCACCG	88

[0326]

	ACGCUGUGCGGGAUCCCCAGACCCUG GAGAUCCUGGACAUCACCCCUUGCAG CUCUGGCGGAGGCAGCAUCCUGGCCA UCUACAGCACCGUGGCCAGCAGCCUG GUGCUGCUGGUGAGCCUGGGCGCCAU CAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
相应的氨基酸序列	mysmqlascvtltlvllvnsQRVQPTESIVRFPNIT NLCPFGEVFNATRFASVYAWNRKRISN CVADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTG KIADYNYKLPDDFTGCVIAWNSNNLDS KVGGNYNYLYRLFRKSNLKPFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGFQP TNGVGYQPYRVVVLSFELLHAPATVCG PKKSTNLVKNKCVNFNFNGLTGTGVLT ESNKKFLPFCQFGRDIADTTDAVRDPQT LEILDITPCSsgggsilaiystvasslvllvslgaisf	89
多聚 A 尾	100 nt	
SARS-CoV-2 NTD-RBD 延伸区连接至跨膜结构域 (NTD-RBD-EXT-TM)		
SEQ ID NO: 114 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 115 和 3' UTR SEQ ID NO: 4。		114
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG	115

[0327]

	UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUG GAUCUUCGGCACCACCCUGGACAGCA AGACCCAGAGCCUGCUGAUCGUGAAU AACGCCACCAACGUGGUGAUCAAGGU GUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAG AACAACAAGAGCUGGAUGGAGAGCG AGUUCCGGGUGUACAGCAGCGCCAAC AACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGG GAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCAC ACCCCAAUCAACCUGGUGCGGGAUCU GCCCCAGGGCUUCUCAGCCCUGGAGC CCCUGGUGGACCUGCCCAUCGGCAUC AACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCC CAGGCGACAGCAGCAGCGGGUGGACA GCAGGCGCGGCUGCUUACUACGUGGG CUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUC ACCGACGCCGUGGACGGAGGCGGAUC GGGAGGCGGACAGCGGGUGCAGCCCA CCGAGAGCAUCGUGCGGUUCCCCAAC AUCACCAACCUGUGCCCCUUCGGCGA GGUGUUCAACGCCACCCGGUUCGCCA GCGUGUACGCCUGGAACCGGAAGCGG AUCAGCAACUGCGUGGCCGACUACAG CGUGCUGUACAACAGCGCCAGCUUCA GCACCUUCAAGUGCUACGGCGUGAGC CCCACCAAGCUGAACGACCUGUGCUU CACCAACGUGUACGCCGACAGCUUCG UGAUCCGUGGCGACGAGGUGCGGCA GAUCGCACCCGGCCAGACAGGCAAGA UCGCCGACUACAACUACAAGCUGCCC GACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCA AGGUGGGCGGCAACUACAACUACCUG UACCGGCUGUUCCGGAAGAGCAACCU GAAGCCCUUCGAGCGGGACAUCAGCA	

[0328]

	CCGAGAUCUACCAAGCCGGCUCCACC CCUUGCAACGGCGUGGAGGGCUUCAA CUGCUACUUCCCUCUGCAGAGCUACG GCUUCCAGCCCACCAACGGCGUGGGC UACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAG CCACCGUGUGUGGCCCCAAGAAGAGC ACCAACCUGGUGAAGAACAAGUGCG UGAACUUCAACUUCAACGGCCUUACC GGCACCGGCGUGCUGACCGAGAGCAA CAAGAAAUUCCUGCCCUUUCAGCAGU UCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCU GGAGAUCCUGGACAUCACCCCUUGCA GCUCUGGCGGAGGCAGCAUCCUGGCC AUCUACAGCACCGUGGCCAGCAGCCU GGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPG DSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDgggsgggQRVQPTESIVRF PNITNLCPFGEVFNATRFASVYAWNRKR ISNCVADYSVLYNSASFSTFKCYGVSPT KLNDLCFTNVYADSFVIRGDEVRQIAPG QTGKIADYNYKLPDDFTGCVIAWNSNN LDSKVGGNYNYLYRLFRKSNLKPFERDI STEIYQAGSTPCNGVEGFNCYFPLQSYG FQPTNGVGYQPYRVVVLSFELLHAPAT VCGPKKSTNLVKNKCVNFNFNGLTGTG VLTESNKKFLPFQQFGRDIADTTDAVRD	116

	PQTLEILDITPCSsgggsilaiystvasslvllvslgaisf	
多聚 A 尾	100 nt	
SARS-CoV-2 NTD 延伸区-RBD 连接至跨膜结构域 (NTD-EXT-RBD-TM)		
SEQ ID NO: 117 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 118 和 3' UTR SEQ ID NO: 4。		117
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUG GAUCUUCGGCACCACCCUGGACAGCA AGACCCAGAGCCUGCUGAUCGUGAAU AACGCCACCAACGUGGUGAUCAAGGU GUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAG AACAACAAGAGCUGGAUGGAGAGCG AGUUCCGGGUGUACAGCAGCGCCAAC AACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGG GAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCAC ACCCCAAUCAACCUGGUGCGGGAUCU GCCCCAGGGCUUCUCAGCCCUGGAGC CCCUGGUGGACCUGCCCAUCGGCAUC AACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCC CAGGCGACAGCAGCAGCGGGUGGACA	118

[0329]

[0330]

	GCAGGCGCGGCUGCUUACUACGUGGG CUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUC ACCGACGCCGUGGACUGCGCCCUGGA CCCUCUGAGCGAGACCAAGUGCACCC UGAAGAGCUUCACCGGAGGCGGAUC GGGAGGCGGACCCAACAUCACCAACC UGUGCCCCUUCGGCGAGGUGUUCAAC GCCACCCGGUUCGCCAGCGUGUACGC CUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUAC AACAGCGCCAGCUUCAGCACCUUCAA GUGCUACGGCGUGAGCCCCACCAAGC UGAACGACCUGUGCUUCACCAACGUG UACGCCGACAGCUUCGUGAUCCGUGG CGACGAGGUGCGGCAGAUCGCACCCG GCCAGACAGGCAAGAUCGCCGACUAC AACUACAAGCUGCCCGACGACUUCAC CGGCUGCGUGAUCGCCUGGAACAGCA ACAACCUCGACAGCAAGGUGGGCGGC AACUACAACUACCUGUACCGGCUGUU CCGGAAGAGCAACCUGAAGCCCUUCG AGCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGG CGUGGAGGGCUUCAACUGCUACUUCC CUCUGCAGAGCUACGGCUUCCAGCCC ACCAACGGCGUGGGCUACCAGCCCUA CCGGGUGGUGGUGCUGAGCUUCGAG CUGCUGCACGCCCCAGCCACCGUGUG UGGCCCCAAGUCUGGCGGAGGCAGCA UCCUGGCCAUCUACAGCACCGUGGCC AGCAGCCUGGUGCUGCUGGUGAGCCU GGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL	119

	GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPG DSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDCALDPLSETKCTLKSFT gggsgggPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKC YGVSPTKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYNYKLPDDFTGCVIA WNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYF PLQSYGFQPTNGVGYQPYRVVVLSFELL HAPATVCGPKsgggsilaiystvasslvllvslgaisf	
多聚 A 尾	100 nt	
SARS-CoV-2 NTD 延伸区-RBD-延伸区连接至跨膜结构域 (NTD-EXT-RBD-EXT-TM)		
SEQ ID NO: 120 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 121 和 3' UTR SEQ ID NO: 4。		120
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUG GAUCUUCGGCACCACCCUGGACAGCA AGACCCAGAGCCUGCUGAUCGUGAAU AACGCCACCAACGUGGUGAUCAAGGU GUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAG	121

[0331]

[0332]

	AACAACAAGAGCUGGAUGGAGAGCG AGUUCCGGGUGUACAGCAGCGCCAAC AACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGCA AGCAGGGCAACUUCAAGAACCUGCGG GAGUUCGUGUUCAAGAACAUCGACG GCUACUUCAAGAUCUACAGCAAGCAC ACCCCAAUCAACCUGGUGCGGGAUCU GCCCCAGGGCUUCUCAGCCCUGGAGC CCCUGGUGGACCUGCCCAUCGGCAUC AACAUCACCCGGUUCCAGACCCUGCU GGCCCUGCACCGGAGCUACCUGACCC CAGGCGACAGCAGCAGCGGGUGGACA GCAGGCGCGGCUGCUUACUACGUGGG CUACCUGCAGCCCCGGACCUUCCUGC UGAAGUACAACGAGAACGGCACCAUC ACCGACGCCGUGGACUGCGCCCUGGA CCCUCUGAGCGAGACCAAGUGCACCC UGAAGAGCUUCACCGGAGGCGGAUC GGGAGGCGGACAGCGGGUGCAGCCCA CCGAGAGCAUCGUGCGGUUCCCCAAC AUCACCAACCUGUGCCCCUUCGGCGA GGUGUUCAACGCCACCCGGUUCGCCA GCGUGUACGCCUGGAACCGGAAGCGG AUCAGCAACUGCGUGGCCGACUACAG CGUGCUGUACAACAGCGCCAGCUUCA GCACCUUCAAGUGCUACGGCGUGAGC CCCACCAAGCUGAACGACCUGUGCUU CACCAACGUGUACGCCGACAGCUUCG UGAUCCGUGGCGACGAGGUGCGGCA GAUCGCACCCGGCCAGACAGGCAAGA UCGCCGACUACAACUACAAGCUGCCC GACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCA AGGUGGGCGGCAACUACAACUACCUG UACCGGCUGUUCCGGAAGAGCAACCU GAAGCCCUUCGAGCGGGACAUCAGCA CCGAGAUCUACCAAGCCGGCUCCACC CCUUGCAACGGCGUGGAGGGCUUCAA CUGCUACUUCCCUCUGCAGAGCUACG GCUUCCAGCCCACCAACGGCGUGGGC UACCAGCCCUACCGGGUGGUGGUGCU GAGCUUCGAGCUGCUGCACGCCCCAG	

[0333]

	CCACCGUGUGUGGCCCCAAGAAGAGC ACCAACCUGGUGAAGAACAAGUGCG UGAACUUCAACUUCAACGGCCUUACC GGCACCGGCGUGCUGACCGAGAGCAA CAAGAAAUUCCUGCCCUUUCAGCAGU UCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCU GGAGAUCCUGGACAUCACCCCUUGCA GCUCUGGCGGAGGCAGCAUCCUGGCC AUCUACAGCACCGUGGCCAGCAGCCU GGUGCUGCUGGUGAGCCUGGGCGCCA UCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSAL EPLVDLPIGINITRFQTLLALHRSYLTPG DSSSGWTAGAAAYYVGYLQPRTFLLKY NENGTITDAVDCALDPLSETKCTLKSFT gggsgggQRVQPTESIVRFPNITNLCPFGEV FNATRFASVYAWNRKRISNCVADYSVL YNSASFSTFKCYGVSPTKLNDLCFTNVY ADSFVIRGDEVRQIAPGQTGKIADYNYK LPDDFTGCVIAWNSNNLDSKVGGNYNY LYRLFRKSNLKPFERDISTEIYQAGSTPC NGVEGFNCYFPLQSYGFQPTNGVGYQP YRVVVLSFELLHAPATVCGPKKSTNLV KNKCVNFNFNGLTGTGVLTESNKKFLPF QQFGRDIADTTDAVRDPQTLEILDITPCS sgggsilaiystvasslvllvslgaisf	122
多聚A尾	100 nt	

[0334] 结构域混合物

[0335] 在一些方面，本公开提供了包含编码SARS-CoV-2 S蛋白子结构域的mRNA的混合物的组合物。在一个实例中，组合物包含编码NTD(有或无SD1、SD2和/或跨膜结构域)的mRNA和编码RBD(有或无SD1、SD2和/或跨膜结构域)的mRNA的混合物。在一些实施方案中，组合物包

含编码与跨膜结构域连接的NTD(例如SEQ ID NO:47)的mRNA(例如SEQ ID NO:45或46)和编码与跨膜结构域连接的RBD(例如SEQ ID NO:77)的mRNA(例如SEQ ID NO:75或76)。

[0336] 组合物中一种mRNA对另一种mRNA的浓度的比率可为1:1(50:50)、1:2、1:3、1:4或1:5。在一些实施方案中,所述比率为1:1。例如,组合物可包含1:1比率的编码与跨膜结构域连接的NTD(例如SEQ ID NO:47)的mRNA(例如SEQ ID NO:45或46)对编码与跨膜结构域连接的RBD(例如SEQ ID NO:77)的mRNA(例如SEQ ID NO:75或76)。在一些实施方案中,所述比率为1:2。例如,组合物可包含1:2比率的编码与跨膜结构域连接的NTD(例如SEQ ID NO:47)的mRNA(例如SEQ ID NO:45或46)对编码与跨膜结构域连接的RBD(例如SEQ ID NO:77)的mRNA(例如SEQ ID NO:75或76)。另一个实例,组合物可包含1:2比率的编码与跨膜结构域连接的RBD(例如,SEQ ID NO:77)的mRNA(例如SEQ ID NO:75或76)对编码与跨膜结构域连接的NTD(例如,SEQ ID NO:47)的mRNA(例如,SEQ ID NO:45或46)。编码不同抗原的不同mRNA可刺激不同强度的免疫反应(Magini D等人,PLoS ONE.2016;11:e0161193),并且施用等摩尔比率的编码两种不同抗原的两种mRNA可引起对一者而非另一者的免疫反应(John S等人,Vaccine.2018;36:1689-1699)。对共递送的mRNA的比率的操纵可用于引发以相等效力靶向期望抗原的广泛免疫反应。

[0337] 编码的纳米颗粒抗原

[0338] 在一些实施方案中,本文提供的mRNA疫苗编码包含与支架结构域连接的冠状病毒抗原的融合蛋白。在一些实施方案中,支架结构域赋予由本公开的mRNA编码的抗原期望的性质。例如,支架结构域可例如通过改变抗原的结构、改变抗原的摄取和加工和/或使抗原结合至另一分子来改善抗原的免疫原性。在一些实施方案中,与抗原连接的支架结构域促进抗原自组装成病毒纳米颗粒或更大的蛋白质折叠免疫原。可如本文提供使用的支架结构域的非限制性实例包括铁蛋白结构域、2,4-二氧四氢蝶啶合成酶结构域、折叠子结构域和包封蛋白(encapsulin)结构域。可使用其他支架结构域。

[0339] 铁蛋白

[0340] 在一些实施方案中,铁蛋白结构域用作支架结构域。铁蛋白是主要功能为在细胞内储存铁的蛋白质。铁蛋白由二十四(24)个亚基组成,每个亚基由自组装成具有八面体对称性的四级结构的四α螺旋束构成(Cho K.J.等人,J Mol Biol.2009;390:83-98;Granier T.等人,J Biol Inorg Chem.2003;8:105-111;和Lawson D.M.等人,Nature.1991;349:541-544)。铁蛋白自组装成具有稳固的热稳定性和化学稳定性的纳米颗粒。预计以这种方式将抗原包封在铁蛋白纳米颗粒内既会延迟抗原的降解还会聚集个别抗原,其中每个纳米颗粒含有二十四(24)个抗原亚基。相同抗原的多个拷贝的聚集增强了由树突细胞的抗原摄取和迁移,以及更稳固的$CD4^+$和$CD8^+$T细胞反应(Kastenmüller K等人,J Clin Invest.2011;121(5):1782-96)。因此,铁蛋白纳米颗粒是非常适合抗原呈递和疫苗开发的平台。

[0341] 在一些实施方案中,本文提供的mRNA编码例如经由甘氨酸(例如,GGG)接头结构域与铁蛋白结构域连接的RBD。可使用其他接头。

[0342] 在其他实施方案中,本文提供的mRNA编码例如经由甘氨酸(例如,GGG)接头与铁蛋白结构域连接的S蛋白的S1结构域。如本文中别处指示,可使用其他接头。

[0343] 与铁蛋白结构域连接的SARS-CoV-2抗原和编码其的mRNA的非限制性实例提供于

下文表13A和表13B中。

[0344] 表13A.与铁蛋白结构域连接的抗原

[0345]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与铁蛋白连接的 SARS-CoV-2 S1 亚基(S1-铁蛋白)	7	8
与铁蛋白连接的 SARS-CoV-2 RBD (RBD-铁蛋白)	64	65

[0346] 表13B.与铁蛋白结构域连接的抗原

[0347]

与铁蛋白连接的 SARS-CoV-2 S1 亚基(S1-铁蛋白)		
SEQ ID NO: 6 从 5'端至 3'端由以下组成:5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 7 和 3' UTR SEQ ID NO: 4。		6
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGGUGAGCAGCCAGUGCGUGAACCUGACCACCCGGACCCAGCUGCCACCAGCCUACACCAACAGCUUCACCCGGGGCGUCUACUACCCCGACAAGGUGUUCCGGAGCAGCGUCCUGCACAGCACCCAGGACCUGUUCCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACCAAGCGGUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCCAGCACCGAGAAGAGCAACAUCAUCCGGGGCUGGAUCUUCGGCACCACCCUGGACAGCAAGACCCAGAGCCUGCUGAUCGUGAAUAACGCCACCAACGUGGUGAUCAAGGUGUGCGAGUUCCAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCCGGGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGAGCCAGCCCUUCCUGAUGGACCUGGAGGGCAAG	7

[0348]

	CAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACC CCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAAC AUCACCCGGUUCCAGACCCUGCUGGC CCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCA GGCGCGGCUGCUUACUACGUGGGCUA CCUGCAGCCCCGGACCUUCCUGCUGA AGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACUGCGCCCUGGACCC UCUGAGCGAGACCAAGUGCACCCUGA AGAGCUUCACCGUGGAGAAGGGCAUC UACCAGACCAGCAACUUCCGGGUGCA GCCCACCGAGAGCAUCGUGCGGUUCC CCAACAUCACCAACCUGUGCCCCUUC GGCGAGGUGUUCAACGCCACCCGGUU CGCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGAC UACAGCGUGCUGUACAACAGCGCCAG CUUCAGCACCUUCAAGUGCUACGGCG UGAGCCCCACCAAGCUGAACGACCUG UGCUUCACCAACGUGUACGCCGACAG CUUCGUGAUCCGUGGCGACGAGGUGC GGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGCU GCCCGACGACUUCACCGGCUGCGUGA UCGCCUGGAACAGCAACAACCUCGAC AGCAAGGUGGGCGGCAACUACAACUA CCUGUACCGGCUGUUCCGGAAGAGCA ACCUGAAGCCCUUCGAGCGGGACAUC AGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCU UCAACUGCUACUUCCCUCUGCAGAGC UACGGCUUCCAGCCCACCAACGGCGU GGGCUACCAGCCCUACCGGGUGGUGG UGCUGAGCUUCGAGCUGCUGCACGCC CCAGCCACCGUGUGUGGCCCCAAGAA GAGCACCAACCUGGUGAAGAACAAGU GCGUGAACUUCAACUUCAACGGCCUU ACCGGCACCGGCGUGCUGACCGAGAG	

[0349]

	CAACAAGAAAUUCCUGCCCUUUCAGC AGUUCGGCCGGGACAUCGCCGACACC ACCGACGCUGUGCGGGAUCCCCAGAC CCUGGAGAUCCUGGACAUCACCCCUU GCAGCUUCGGCGGCGUGAGCGUGAUC ACCCCAGGCACCAACACCAGCAACCA GGUGGCCGUGCUGUACCAGGACGUGA ACUGCACCGAGGUGCCCGUGGCCAUC CACGCCGACCAGCUGACACCCACCUG GCGGGUCUACAGCACCGGCAGCAACG UGUUCCAGACCCGGGCCGGUUGCCUG AUCGGCGCCGAGCACGUGAACAACAG CUACGAGUGCGACAUCCCCAUCGGCG CCGGCAUCUGUGCCAGCUACCAGACC CAGACCAAUUCAGGAGGAGGCAGCGG CGGCGAUAUCAUCAAGCUUCUGAACG AGCAAGUUAACAAGGAAAUGCAGAGC AGUAAUCUCUACAUGAGCAUGAGCAG CUGGUGCUACACCCACUCCCUGGACG GAGCAGGCCUCUUCCUGUUCGACCAC GCAGCCGAGGAGUACGAGCACGCUAA GAAGUUGAUCAUUUUCUUGAACGAGA ACAACGUGCCCGUGCAGCUAACGUCA AUCAGCGCACCUGAGCACAAGUUCGA GGGCCUGACCCAGAUCUUCCAGAAGG CCUACGAACACGAACAGCACAUCUCC GAGAGCAUCAACAAUAUUGUGGAUCA CGCUAUCAAGUCCAAGGACCACGCUA CCUUCAACUUCCUGCAGUGGUACGUG GCCGAGCAACAUGAGGAGGAGGUGCU GUUCAAGGACAUCCUGGACAAGAUCG AGCUGAUCGGUAAUGAGAAUCACGGC CUGUACCUGGCCGACCAGUACGUGAA GGGCAUCGCCAAGAGCCGGAAGUCAG GCUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAYT NSFTRGVYYPDKVFRSSVLHSTQDLFLPF FSNVTWFHAIHVSGTNGTKRFDNPVLPF	8

[0350]

	NDGVYFASTEKSNIIRGWIFGTTLDSKTQ SLLIVNNATNVVIKVCEFQFCNDPFLGVY YHKNNKSWMESEFRVYSSANNCTFEYV SQPFLMDLEGKQGNFKNLREFVFKNIDG YFKIYSKHTPINLVRDLPQGFSALEPLVD LPIGINITRFQTLLALHRSYLTPGDSSSGW TAGAAAYYVGYLQPRTFLLKYNENGTIT DAVDCALDPLSETKCTLKSFTVEKGIYQ TSNFRVQPTESIVRFPNITNLCPFGEVFNA TRFASVYAWNRKRISNCVADYSVLYNS ASFSTFKCYGVSPTKLNDLCFTNVYADS FVIRGDEVRQIAPGQTGKIADYNYKLPD DFTGCVIAWNSNNLDSKVGGNYNYLYR LFRKSNLKPFERDISTEIYQAGSTPCNGV EGFNCYFPLQSYGFQPTNGVGYQPYRVV VLSFELLHAPATVCGPKKSTNLVKNKCV NFNFNGLTGTGVLTESNKKFLPFQQFGR DIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAI HADQLTPTWRVYSTGSNVFQTRAGCLIG AEHVNNSYECDIPIGAGICASYQTQTNSg ggSGGDIIKLLNEQVNKEMQSSNLYMSM SSWCYTHSLDGAGLFLFDHAAEEYEHA KKLIIFLNENNVPVQLTSISAPEHKFEGLT QIFQKAYEHEQHISESINNIVDHAIKSKDH ATFNFLQWYVAEQHEEEVLFKDILDKIE LIGNENHGLYLADQYVKGIAKSRKSGS	
多聚 A 尾	100 nt	
与铁蛋白连接的 SARS-CoV-2 RBD (RBD-铁蛋白)		
SEQ ID NO: 63 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 64 和 3' UTR SEQ ID NO: 4。		63
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCUG CGUGACCCUGACCCUGGUGCUGCUGG UGAACAGCCAGCCCAACAUCACCAAC CUGUGCCCCUUCGGCGAGGUGUUCAA CGCCACCCGGUUCGCCAGCGUGUACG CCUGGAACCGGAAGCGGAUCAGCAAC	64

[0351]

	UGCGUGGCCGACUACAGCGUGCUGUA CAACAGCGCCAGCUUCAGCACCUUCA AGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUG GCGACGAGGUGCGGCAGAUCGCACCC GGCCAGACAGGCAAGAUCGCCGACUA CAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGC AACAACCUCGACAGCAAGGUGGGCGG CAACUACAACUACCUGUACCGGCUGU UCCGGAAGAGCAACCUGAAGCCCUUC GAGCGGGACAUCAGCACCGAGAUCUA CCAAGCCGGCUCCACCCCUUGCAACG GCGUGGAGGGCUUCAACUGCUACUUC CCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAG CUGCUGCACGCCCCAGCCACCGUGUG UGGCCCCAAGGGAGGAGGCAGCGGCG GCGAUAUCAUCAAGCUUCUGAACGAG CAAGUUAACAAGGAAAUGCAGAGCAG UAAUCUCUACAUGAGCAUGAGCAGCU GGUGCUACACCCACUCCCUGGACGGA GCAGGCCUCUUCCUGUUCGACCACGC AGCCGAGGAGUACGAGCACGCUAAGA AGUUGAUCAUUUUCUUGAACGAGAAC AACGUGCCCGUGCAGCUAACGUCAAU CAGCGCACCUGAGCACAAGUUCGAGG GCCUGACCCAGAUCUUCCAGAAGGCC UACGAACACGAACAGCACAUCUCCGA GAGCAUCAACAAUAUUGUGGAUCACG CUAUCAAGUCCAAGGACCACGCUACC UUCAACUUCCUGCAGUGGUACGUGGC CGAGCAACAUGAGGAGGAGGUGCUGU UCAAGGACAUCCUGGACAAGAUCGAG CUGAUCGGUAAUGAGAAUCACGGCCU GUACCUGGCCGACCAGUACGUGAAGG GCAUCGCCAAGAGCCGGAAGUCAGGC UCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC	4

[0352]

	CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	
相应的氨基酸序列	MYSMQLASCVTLTLVLLVNSQPNITNLC PFGEVFNATRFASVYAWNRKRISNCVAD YSVLYNSASFSTFKCYGVSPTKLNDLCFT NVYADSFVIRGDEVRQIAPGQTGKIADY NYKLPDDFTGCVIAWNSNNLDSKVGGN YNYLYRLFRKSNLKPFERDISTEIYQAGS TPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVVLSFELLHAPATVCGPKGGGSG GDIIKLLNEQVNKEMQSSNLYMSMSSWC YTHSLDGAGLFLFDHAAEEYEHAKKLIIF LNENNVPVQLTSISAPEHKFEGLTQIFQK AYEHEQHISESINNIVDHAIKSKDHATFN FLQWYVAEQHEEEVLFKDILDKIELIGNE NHGLYLADQYVKGIAKSRKSGS	65
多聚A尾	100 nt	

[0353] 2,4-二氧四氢蝶啶合成酶

[0354] 在一些实施方案中,2,4-二氧四氢蝶啶合成酶结构域用作支架结构域。2,4-二氧四氢蝶啶合成酶是一种在多种生物体(包括古细菌、细菌、真菌、植物和真细菌)中负责核黄素生物合成的倒数第二个催化步骤的酶。2,4-二氧四氢蝶啶合成酶由同寡聚体构成,所述同寡聚体根据其来源的种类而大小和亚基数不同,包括五聚体、十聚体和二十面六十聚体(icosahedral sixty-mer)。2,4-二氧四氢蝶啶合成酶单体长度为150个氨基酸,并且包括具有侧翼串联α-螺旋的β-折叠。已报导2,4-二氧四氢蝶啶合成酶的不同四级结构,阐释其形态多样性:从同五聚体至形成150 Å直径的衣壳的十二(12)个五聚体的对称装配。在2,4-二氧四氢蝶啶合成酶表面的抗原呈递导致以有序阵列展示的抗原的高局部浓度。此类重复结构使得B细胞受体能够交联,并经由亲合效应导致强烈免疫反应。

[0355] 在一些实施方案中,本文提供的mRNA编码例如经由甘氨酸-丝氨酸(例如,GGS)与2,4-二氧四氢蝶啶合成酶结构域连接的RBD。可使用其他接头。

[0356] 在其他实施方案中,本文提供的mRNA编码例如经由甘氨酸-丝氨酸(例如,GGS)接头与2,4-二氧四氢蝶啶合成酶结构域连接的S蛋白的S1结构域。如本文中别处指示,可使用其他接头。

[0357] 与折叠子结构域连接的SARS-CoV-2抗原和编码其的mRNA的非限制性实例提供于下文表14A和表14B中。

[0358] 表14A.与2,4-二氧四氢蝶啶合成酶结构域连接的抗原

[0359]

名称	SEQ ID NO:	
与 2,4-二氧四氢蝶啶合成酶 C 末端连接的 SARS-CoV-2 可溶性 S1 (LS-S1)	10	11
与 2,4-二氧四氢蝶啶合成酶 N 末端连接的 SARS-CoV-2 可溶性 S1 (S1-LS)	13	14
与 2,4-二氧四氢蝶啶合成酶 C 末端连接的 SARS-CoV-2 RBD (LS-RBD)	67	68
与 2,4-二氧四氢蝶啶合成酶 N 末端连接的 SARS-CoV-2 RBD (RBD-LS)	70	71

[0360] 表14B.与2,4-二氧四氢蝶啶合成酶结构域连接的抗原

[0361]

与 2,4-二氧四氢蝶啶合成酶 C 末端连接的 SARS-CoV-2 可溶性 S1 (LS-S1)		
SEQ ID NO: 9 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 10 和 3' UTR SEQ ID NO: 4。		9
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGGGCAUCCUGCCCAGCCCUGGCAUGCCCGCUCUGCUGAGCCUGGUGAGCCUGCUGAGCGUGCUGCUGAUGGGCUGCGUGGCUGAGACCGGCAUGCAGAUCUACGAGGGCAAGCUGACCGCAGAGGGCCUGCGGUUCGGCAUCGUGGCCAGCCGCGCCAACCACGCUCUGGUGGACC	10

[0362]

	GGCUUGUGGAGGGCGCUAUCGACGC CAUCGUGAGACACGGCGGCCGGGAA GAGGACAUCACCCUGGUGCGGGUGU GCGGCAGCUGGGAGAUUCCCGUCGC CGCCGGAGAACUGGCCCGGAAGGAG GACAUCGACGCCGUGAUCGCCAUCG GCGUGCUGUGCAGAGGCGCCACGCCC AGCUUCGACUACAUCGCCAGCGAGG UGAGCAAGGGCCUGGCCGACCUGAG CCUGGAGCUGCGGAAGCCCAUCACCU UCGGCGUGAUCACCGCCGACACCCUG GAGCAGGCCAUCGAGGCCGCAGGCA CCUGCCACGGCAACAAGGGCUGGGA AGCCGCCCUGUGCGCCAUCGAGAUG GCCAACCUGUUCAAGAGCCUGCGGG GCGGAAGUGGAGGCUCUGGUGGCAG CGGAGGAUCUGGCGGCGGCACCACCC GGACCCAGCUGCCACCAGCCUACACC AACAGCUUCACCCGGGGCGUCUACU ACCCCGACAAGGUGUUCCGGAGCAG CGUCCUGCACAGCACCCAGGACCUGU UCCUGCCCUUCUUCAGCAACGUGACC UGGUUCCACGCCAUCCACGUGAGCG GCACCAACGGCACCAAGCGGUUCGAC AACCCCGUGCUGCCCUUCAACGACGG CGUGUACUUCGCCAGCACCGAGAAG AGCAACAUCAUCCGGGGCUGGAUCU UCGGCACCACCCUGGACAGCAAGACC CAGAGCCUGCUGAUCGUGAAUAACG CCACCAACGUGGUGAUCAAGGUGUG CGAGUUCCAGUUCUGCAACGACCCCU UCCUGGGCGUGUACUACCACAAGAA CAACAAGAGCUGGAUGGAGAGCGAG UUCCGGGUGUACAGCAGCGCCAACA ACUGCACCUUCGAGUACGUGAGCCA GCCCUUCCUGAUGGACCUGGAGGGC AAGCAGGGCAACUUCAAGAACCUGC GGGAGUUCGUGUUCAAGAACAUCGA CGGCUACUUCAAGAUCUACAGCAAG CACACCCCAAUCAACCUGGUGCGGGA UCUGCCCCAGGGCUUCUCAGCCCUGG AGCCCCUGGUGGACCUGCCCAUCGGC AUCAACAUCACCCGGUUCCAGACCCU	

[0363]

	GCUGGCCCUGCACCGGAGCUACCUGA CCCCAGGCGACAGCAGCAGCGGGUG GACAGCAGGCGCGGCUGCUUACUAC GUGGGCUACCUGCAGCCCCGGACCUU CCUGCUGAAGUACAACGAGAACGGC ACCAUCACCGACGCCGUGGACUGCGC CCUGGACCCUCUGAGCGAGACCAAG UGCACCCUGAAGAGCUUCACCGUGG AGAAGGGCAUCUACCAGACCAGCAA CUUCCGGGUGCAGCCCACCGAGAGCA UCGUGCGGUUCCCCAACAUCACCAAC CUGUGCCCCUUCGGCGAGGUGUUCA ACGCCACCCGGUUCGCCAGCGUGUAC GCCUGGAACCGGAAGCGGAUCAGCA ACUGCGUGGCCGACUACAGCGUGCU GUACAACAGCGCCAGCUUCAGCACCU UCAAGUGCUACGGCGUGAGCCCCACC AAGCUGAACGACCUGUGCUUCACCA ACGUGUACGCCGACAGCUUCGUGAU CCGUGGCGACGAGGUGCGGCAGAUC GCACCCGGCCAGACAGGCAAGAUCGC CGACUACAACUACAAGCUGCCCGACG ACUUCACCGGCUGCGUGAUCGCCUG GAACAGCAACAACCUCGACAGCAAG GUGGGCGGCAACUACAACUACCUGU ACCGGCUGUUCCGGAAGAGCAACCU GAAGCCCUUCGAGCGGGACAUCAGC ACCGAGAUCUACCAAGCCGGCUCCAC CCCUUGCAACGGCGUGGAGGGCUUC AACUGCUACUUCCCUCUGCAGAGCU ACGGCUUCCAGCCCACCAACGGCGUG GGCUACCAGCCCUACCGGGUGGUGG UGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGCCCCAAGA AGAGCACCAACCUGGUGAAGAACAA GUGCGUGAACUUCAACUUCAACGGC CUUACCGGCACCGGCGUGCUGACCGA GAGCAACAAGAAAUUCCUGCCCUUU CAGCAGUUCGGCCGGGACAUCGCCG ACACCACCGACGCUGUGCGGGAUCCC CAGACCCUGGAGAUCCUGGACAUCA CCCCUUGCAGCUUCGGCGGCGUGAGC GUGAUCACCCCAGGCACCAACACCAG	

[0364]

	CAACCAGGUGGCCGUGCUGUACCAG GACGUGAACUGCACCGAGGUGCCCG UGGCCAUCCACGCCGACCAGCUGACA CCCACCUGGCGGGUCUACAGCACCGG CAGCAACGUGUUCCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGU GAACAACAGCUACGAGUGCGACAUC CCCAUCGGCGCCGGCAUCUGUGCCAG CUACCAGACCCAGACCAAUUCA	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MGILPSPGMPALLSLVSLLSVLLMGCVA ETGMQIYEGKLTAEGLRFGIVASRANH ALVDRLVEGAIDAIVRHGGREEDITLVR VCGSWEIPVAAGELARKEDIDAVIAIGV LCRGATPSFDYIASEVSKGLADLSLELR KPITFGVITADTLEQAIEAAGTCHGNKG WEAALCAIEMANLFKSLRGGSGGSGGS GGSGGGTTRTQLPPAYTNSFTRGVYYP DKVFRSSVLHSTQDLFLPFFSNVTWFHA IHVSGTNGTKRFDNPVLPFNDGVYFAST EKSNIIRGWIFGTTLDSKTQSLLIVNNAT NVVIKVCEFQFCNDPFLGVYYHKNNKS WMESEFRVYSSANNCTFEYVSQPFLMD LEGKQGNFKNLREFVFKNIDGYFKIYSK HTPINLVRDLPQGFSALEPLVDLPIGINIT RFQTLLALHRSYLTPGDSSSGWTAGAA AYYVGYLQPRTFLLKYNENGTITDAVD CALDPLSETKCTLKSFTVEKGIYQTSNF RVQPTESIVRFPNITNLCPFGEVFNATRF ASVYAWNRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVI RGDEVRQIAPGQTGKIADYNYKLPDDF TGCVIAWNSNNLDSKVGGNYNYLYRL FRKSNLKPFERDISTEIYQAGSTPCNGVE GFNCYFPLQSYGFQPTNGVGYQPYRVV VLSFELLHAPATVCGPKKSTNLVKNKC VNFNFNGLTGTGVLTESNKKFLPFQQF GRDIADTTDAVRDPQTLEILDITPCSFGG VSVITPGTNTSNQVAVLYQDVNCTEVP	11

	VAIHADQLTPTWRVYSTGSNVFQTRAG CLIGAEHVNNSYECDIPIGAGICASYQT QTNS	
多聚 A 尾	100 nt	
与 2,4-二氧四氢蝶啶合成酶 N 末端连接的 SARS-CoV-2 可溶性 S1 (S1-LS)		
SEQ ID NO: 12 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 13 和 3' UTR SEQ ID NO: 4。		12
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGC CGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGC CCCUGGUGAGCAGCCAGUGCGUGAA CCUGACCACCCGGACCCAGCUGCCAC CAGCCUACACCAACAGCUUCACCCGG GGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACC CAGGACCUGUUCCUGCCCUUCUUCAG CAACGUGACCUGGUUCCACGCCAUCC ACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCU UCAACGACGGCGUGUACUUCGCCAG CACCGAGAAGAGCAACAUCAUCCGG GGCUGGAUCUUCGGCACCACCCUGG ACAGCAAGACCCAGAGCCUGCUGAU CGUGAAUAACGCCACCAACGUGGUG AUCAAGGUGUGCGAGUUCCAGUUCU GCAACGACCCCUUCCUGGGCGUGUAC UACCACAAGAACAACAAGAGCUGGA UGGAGAGCGAGUUCCGGGUGUACAG CAGCGCCAACAACUGCACCUUCGAGU ACGUGAGCCAGCCCUUCCUGAUGGA CCUGGAGGGCAAGCAGGGCAACUUC AAGAACCUGCGGGAGUUCGUGUUCA AGAACAUCGACGGCUACUUCAAGAU CUACAGCAAGCACACCCCAAUCAACC UGGUGCGGGAUCUGCCCCAGGGCUU CUCAGCCCUGGAGCCCCUGGUGGACC UGCCCAUCGGCAUCAACAUCACCCGG	13

[0365]

[0366]

	UUCCAGACCCUGCUGGCCCUGCACCG GAGCUACCUGACCCCAGGCGACAGCA GCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAG CCCCGGACCUUCCUGCUGAAGUACAA CGAGAACGGCACCAUCACCGACGCCG UGGACUGCGCCCUGGACCCUCUGAGC GAGACCAAGUGCACCCUGAAGAGCU UCACCGUGGAGAAGGGCAUCUACCA GACCAGCAACUUCCGGGUGCAGCCCA CCGAGAGCAUCGUGCGGUUCCCCAAC AUCACCAACCUGUGCCCCUUCGGCGA GGUGUUCAACGCCACCCGGUUCGCCA GCGUGUACGCCUGGAACCGGAAGCG GAUCAGCAACUGCGUGGCCGACUAC AGCGUGCUGUACAACAGCGCCAGCU UCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGU GCUUCACCAACGUGUACGCCGACAGC UUCGUGAUCCGUGGCGACGAGGUGC GGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGC UGCCCGACGACUUCACCGGCUGCGUG AUCGCCUGGAACAGCAACAACCUCG ACAGCAAGGUGGGCGGCAACUACAA CUACCUGUACCGGCUGUUCCGGAAG AGCAACCUGAAGCCCUUCGAGCGGG ACAUCAGCACCGAGAUCUACCAAGCC GGCUCCACCCCUUGCAACGGCGUGGA GGGCUUCAACUGCUACUUCCCUCUGC AGAGCUACGGCUUCCAGCCCACCAAC GGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCU GCACGCCCCAGCCACCGUGUGUGGCC CCAAGAAGAGCACCAACCUGGUGAA GAACAAGUGCGUGAACUUCAACUUC AACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUG CCCUUUCAGCAGUUCGGCCGGGACA UCGCCGACACCACCGACGCUGUGCGG GAUCCCCAGACCCUGGAGAUCCUGG ACAUCACCCCUUGCAGCUUCGGCGGC GUGAGCGUGAUCACCCCAGGCACCA	

[0367]

	ACACCAGCAACCAGGUGGCCGUGCU GUACCAGGACGUGAACUGCACCGAG GUGCCCGUGGCCAUCCACGCCGACCA GCUGACACCCACCUGGCGGGUCUACA GCACCGGCAGCAACGUGUUCCAGACC CGGGCCGGUUGCCUGAUCGGCGCCG AGCACGUGAACAACAGCUACGAGUG CGACAUCCCCAUCGGCGCCGGCAUCU GUGCCAGCUACCAGACCCAGACCAAU UCAGGAGGAGGCUCCGGAGGCGGUA GCGCUGAGACCGGCAUGCAGAUCUA CGAGGGCAAGCUGACCGCAGAGGGC CUGCGGUUCGGCAUCGUGGCCAGCC GCGCCAACCACGCUCUGGUGGACCGG CUUGUGGAGGGCGCUAUCGACGCCA UCGUGAGACACGGCGGCCGGGAAGA GGACAUCACCCUGGUGCGGGUGUGC GGCAGCUGGGAGAUUCCCGUCGCCG CCGGAGAACUGGCCCGGAAGGAGGA CAUCGACGCCGUGAUCGCCAUCGGCG UGCUGUGCAGAGGCGCCACGCCCAGC UUCGACUACAUCGCCAGCGAGGUGA GCAAGGGCCUGGCCGACCUGAGCCU GGAGCUGCGGAAGCCCAUCACCUUC GGCGUGAUCACCGCCGACACCCUGGA GCAGGCCAUCGAGGCCGCAGGCACCU GCCACGGCAACAAGGGCUGGGAAGC CGCCCUGUGCGCCAUCGAGAUGGCCA ACCUGUUCAAGAGCCUGCGG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSA LEPLVDLPIGINITRFQTLLALHRSYLTP	14

[0368]

	GDSSSGWTAGAAAYYVGYLQPRTFLL KYNENGTITDAVDCALDPLSETKCTLKS FTVEKGIYQTSNFRVQPTESIVRFPNITN LCPFGEVFNATRFASVYAWNRKRISNC VADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTG KIADYNYKLPDDFTGCVIAWNSNNLDS KVGGNYNYLYRLFRKSNLKPFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGFQP TNGVGYQPYRVVVLSFELLHAPATVCG PKKSTNLVKNKCVNFNFNGLTGTGVLT ESNKKFLPFQQFGRDIADTTDAVRDPQT LEILDITPCSFGGVSVITPGTNTSNQVAV LYQDVNCTEVPVAIHADQLTPTWRVYS TGSNVFQTRAGCLIGAEHVNNSYECDIP IGAGICASYQTQTNSgggsgggsAETGMQI YEGKLTAEGLRFGIVASRANHALVDRL VEGAIDAIVRHGGREEDITLVRVCGSWE IPVAAGELARKEDIDAVIAIGVLCRGAT PSFDYIASEVSKGLADLSLELRKPITFGV ITADTLEQAIEAAGTCHGNKGWEAALC AIEMANLFKSLR	
多聚 A 尾	100 nt	
与 2,4-二氧四氢蝶啶合成酶 C 末端连接的 SARS-CoV-2 RBD (LS-RBD)		
SEQ ID NO: 66 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 67 和 3' UTR SEQ ID NO: 4。		66
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGC CGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGGGCAUCCUGCCCAGCCCUGGCAU GCCCGCUCUGCUGAGCCUGGUGAGCC UGCUGAGCGUGCUGCUGAUGGGCUG CGUGGCUGAGACCGGCAUGCAGAUC UACGAGGGCAAGCUGACCGCAGAGG GCCUGCGGUUCGGCAUCGUGGCCAG CCGCGCCAACCACGCUCUGGUGGACC GGCUUGUGGAGGGCGCUAUCGACGC CAUCGUGAGACACGGCGGCCGGGAA	67

[0369]

	GAGGACAUCACCCUGGUGCGGGUGU GCGGCAGCUGGGAGAUUCCCGUCGC CGCCGGAGAACUGGCCCGGAAGGAG GACAUCGACGCCGUGAUCGCCAUCG GCGUGCUGUGCAGAGGCGCCACGCCC AGCUUCGACUACAUCGCCAGCGAGG UGAGCAAGGGCCUGGCCGACCUGAG CCUGGAGCUGCGGAAGCCCAUCACCU UCGGCGUGAUCACCGCCGACACCCUG GAGCAGGCCAUCGAGGCCGCAGGCA CCUGCCACGGCAACAAGGGCUGGGA AGCCGCCCUGUGCGCCAUCGAGAUG GCCAACCUGUUCAAGAGCCUGCGGG GCGGAAGUGGAGGCUCUGGUGGCAG CGGAGGAUCUGGCGGCGGCCAGCCC AACAUCACCAACCUGUGCCCCUUCGG CGAGGUGUUCAACGCCACCCGGUUC GCCAGCGUGUACGCCUGGAACCGGA AGCGGAUCAGCAACUGCGUGGCCGA CUACAGCGUGCUGUACAACAGCGCC AGCUUCAGCACCUUCAAGUGCUACG GCGUGAGCCCCACCAAGCUGAACGAC CUGUGCUUCACCAACGUGUACGCCG ACAGCUUCGUGAUCCGUGGCGACGA GGUGCGGCAGAUCGCACCCGGCCAG ACAGGCAAGAUCGCCGACUACAACU ACAAGCUGCCCGACGACUUCACCGGC UGCGUGAUCGCCUGGAACAGCAACA ACCUCGACAGCAAGGUGGGCGGCAA CUACAACUACCUGUACCGGCUGUUCC GGAAGAGCAACCUGAAGCCCUUCGA GCGGGACAUCAGCACCGAGAUCUAC CAAGCCGGCUCCACCCCUUGCAACGG CGUGGAGGGCUUCAACUGCUACUUC CCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGA GCUGCUGCACGCCCCAGCCACCGUGU GUGGCCCCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUA	4

	AAGUCUGAGUGGGCGGC	
相应的氨基酸序列	MGILPSPGMPALLSLVSLLSVLLMGCVA ETGMQIYEGKLTAEGLRFGIVASRANH ALVDRLVEGAIDAIVRHGGREEDITLVR VCGSWEIPVAAGELARKEDIDAVIAIGV LCRGATPSFDYIASEVSKGLADLSLELR KPITFGVITADTLEQAIEAAGTCHGNKG WEAALCAIEMANLFKSLRGGSGGSGGS GGSGGGQPNITNLCPFGEVFNATRFASV YAWNRKRISNCVADYSVLYNSASFSTF KCYGVSPTKLNDLCFTNVYADSFVIRG DEVRQIAPGQTGKIADYNYKLPDDFTG CVIAWNSNNLDSKVGGNYNYLYRLFR KSNLKPFERDISTEIYQAGSTPCNGVEGF NCYFPLQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVCGPK	68
多聚 A 尾	100 nt	
与 2,4-二氧四氢蝶啶合成酶 N 末端连接的 SARS-CoV-2 RBD (RBD-LS)		
SEQ ID NO: 69 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 70 和 3' UTR SEQ ID NO: 4。		69
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGC CGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCU GCGUGACCCUGACCCUGGUGCUGCU GGUGAACAGCCAGCCCAACAUCACCA ACCUGUGCCCCUUCGGCGAGGUGUU CAACGCCACCCGGUUCGCCAGCGUGU ACGCCUGGAACCGGAAGCGGAUCAG CAACUGCGUGGCCGACUACAGCGUG CUGUACAACAGCGCCAGCUUCAGCAC CUUCAAGUGCUACGGCGUGAGCCCC ACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGU GAUCCGUGGCGACGAGGUGCGGCAG AUCGCACCCGGCCAGACAGGCAAGA UCGCCGACUACAACUACAAGCUGCCC GACGACUUCACCGGCUGCGUGAUCG	70

[0370]

[0371]

	CCUGGAACAGCAACAACCUCGACAGC AAGGUGGGCGGCAACUACAACUACC UGUACCGGCUGUUCCGGAAGAGCAA CCUGAAGCCCUUCGAGCGGGACAUC AGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGC UUCAACUGCUACUUCCCUCUGCAGA GCUACGGCUUCCAGCCCACCAACGGC GUGGGCUACCAGCCCUACCGGGUGG UGGUGCUGAGCUUCGAGCUGCUGCA CGCCCCAGCCACCGUGUGUGGCCCCA AGGGAGGAGGCUCCGGAGGCGGUAG CGCUGAGACCGGCAUGCAGAUCUAC GAGGGCAAGCUGACCGCAGAGGGCC UGCGGUUCGGCAUCGUGGCCAGCCG CGCCAACCACGCUCUGGUGGACCGGC UUGUGGAGGGCGCUAUCGACGCCAU CGUGAGACACGGCGGCCGGGAAGAG GACAUCACCCUGGUGCGGGUGUGCG GCAGCUGGGAGAUUCCCGUCGCCGCC GGAGAACUGGCCCGGAAGGAGGACA UCGACGCCGUGAUCGCCAUCGGCGU GCUGUGCAGAGGCGCCACGCCCAGCU UCGACUACAUCGCCAGCGAGGUGAG CAAGGGCCUGGCCGACCUGAGCCUG GAGCUGCGGAAGCCCAUCACCUUCG GCGUGAUCACCGCCGACACCCUGGAG CAGGCCAUCGAGGCCGCAGGCACCUG CCACGGCAACAAGGGCUGGGAAGCC GCCCUGUGCGCCAUCGAGAUGGCCA ACCUGUUCAAGAGCCUGCGG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MYSMQLASCVTLTLVLLVNSQPNITNL CPFGEVFNATRFASVYAWNRKRISNCV ADYSVLYNSASFSTFKCYGVSPTKLND LCFTNVYADSFVIRGDEVRQIAPGQTGK IADYNYKLPDDFTGCVIAWNSNNLDSK VGGNYNYLYRLFRKSNLKPFERDISTEI YQAGSTPCNGVEGFNCYFPLQSYGFQP	71

[0372]

	TNGVGYQPYRVVVLSFELLHAPATVCG PKgggsgggsAETGMQIYEGKLTAEGLRF GIVASRANHALVDRLVEGAIDAIVRHG GREEDITLVRVCGSWEIPVAAGELARKE DIDAVIAIGVLCRGATPSFDYIASEVSKG LADLSLELRKPITFGVITADTLEQAIEAA GTCHGNKGWEAALCAIEMANLFKSLR	
多聚A尾	100 nt	

[0373] 折叠子

[0374] 在一些实施方案中，折叠子结构域用作支架结构域。T4次要纤维蛋白(fibritin)的C末端结构域(折叠子)对于形成次要纤维蛋白三聚体结构是必需的，并且可用作人工三聚化结构域(参见例如Meier S.等人，Journal of Molecular Biology 2004年12月3日；344(4)：1051-1069；Tao Y等人，Structure 1997年6月15日；5(6)：789-98)。当与S蛋白胞外结构域融合时，折叠子结构域促进S蛋白的正确三聚化，从而避免蛋白质的错误折叠。导致S蛋白融合前构象产生的这种过程导致表达、构象均质性增加，并引发强效中和抗体反应。

[0375] 不受限于理论，认为此构型将导致折叠子在蛋白质的细胞内区上在很大程度上免疫原性沉默。与折叠子结构域连接的SARS-CoV-2抗原和编码其的mRNA的非限制性实例提供于下文表15A和表15B中。

[0376] 表15A.与折叠子结构域连接的抗原

[0377]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与折叠子结构域连接的SARS-CoV-2 NTD	43	44
与跨膜结构域和折叠子结构域连接的SARS-CoV-2 NTD	49	50
与折叠子结构域连接的SARS-CoV-2 RBD	73	74
与折叠子结构域和跨膜结构域连接的SARS-CoV-2 RBD (RBD-FD-TM)	79	80
与跨膜结构域和折叠子结构域连接的SARS-CoV-2 RBD (RBD-TM-FD)	82	83

[0378]

名称	SEQ ID NO:	
	mRNA ORF	蛋白质
与折叠子结构域和跨膜结构域连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-FD-TM)	100	101
与跨膜结构域和折叠子结构域连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM-FD)	103	104
与折叠子结构域连接的 SARS-CoV-2 NTD-RBD	112	113

[0379] 表15B.与折叠子结构域连接的抗原

与折叠子结构域连接的 SARS-CoV-2 NTD		
SEQ ID NO: 42 从 5'端至 3'端由以下组成:5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 43 和 3' UTR SEQ ID NO: 4。		42
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGG AUCUUCGGCACCACCCUGGACAGCAA GACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUG UGCGAGUUCCAGUUCUGCAACGACCC CUUCCUGGGCGUGUACUACCACAAGA ACAACAAGAGCUGGAUGGAGAGCGAG UUCCGGGUGUACAGCAGCGCCAACAA	43

[0380]

[0381]

	CUGCACCUUCGAGUACGUGAGCCAGC CCUUCCUGAUGGACCUGGAGGGCAAG CAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACC CCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAAC AUCACCCGGUUCCAGACCCUGCUGGC CCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCA GGCGCGGCUGCUUACUACGUGGGCUA CCUGCAGCCCCGGACCUUCCUGCUGA AGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACUCUGGCGGAGGCGG CAGCGCCAUCGGCGGCUACAUCCCCG AGGCCCCUAGAGACGGCCAGGCCUAC GUGCGGAAGGACGGCGAGUGGGUGCU GCUGAGCACCUUCCUGGGC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFLG VYYHKNNKSWMESEFRVYSSANNCTFE YVSQPFLMDLEGKQGNFKNLREFVFKNI DGYFKIYSKHTPINLVRDLPQGFSALEPL VDLPIGINITRFQTLLALHRSYLTPGDSSS GWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDsggggSAIGGYIPEAPRDGQAY VRKDGEWVLLSTFLG	44
多聚 A 尾	100 nt	
与跨膜结构域和折叠子结构域连接的 SARS-CoV-2 NTD		
SEQ ID NO: 48 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 49 和 3' UTR SEQ ID NO: 4。		48
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	

[0382]

5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGG AUCUUCGGCACCACCCUGGACAGCAA GACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUG UGCGAGUUCCAGUUCUGCAACGACCC CUUCCUGGGCGUGUACUACCACAAGA ACAACAAGAGCUGGAUGGAGAGCGAG UUCCGGGUGUACAGCAGCGCCAACAA CUGCACCUUCGAGUACGUGAGCCAGC CCUUCCUGAUGGACCUGGAGGGCAAG CAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACC CCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAAC AUCACCCGGUUCCAGACCCUGCUGGC CCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCA GGCGCGGCUGCUUACUACGUGGGCUA CCUGCAGCCCCGGACCUUCCUGCUGA AGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACUCUGGCGGAGGCAG CAUCCUGGCCAUCUACAGCACCGUGG CCAGCAGCCUGGUGCUGCUGGUGAGC CUGGGCGCCAUCAGCUUCGGCGGAGG CAGCGCCAUCGGCGGCUACAUCCCCG AGGCCCCUAGAGACGGCCAGGCCUAC GUGCGGAAGGACGGCGAGUGGGUGCU	49

	GCUGAGCACCUUCCUGGGCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFLG VYYHKNNKSWMESEFRVYSSANNCTFE YVSQPFLMDLEGKQGNFKNLREFVFKNI DGYFKIYSKHTPINLVRDLPQGFSALEPL VDLPIGINITRFQTLLALHRSYLTPGDSSS GWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDsgggsilaiystvasslvllvslgaisfgggS AIGGYIPEAPRDGQAYVRKDGEWVLLST FLGk	50
多聚 A 尾	100 nt	
与折叠子结构域连接的 SARS-CoV-2 RBD		
SEQ ID NO: 72 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 73 和 3' UTR SEQ ID NO: 4。		72
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCUG CGUGACCCUGACCCUGGUGCUGCUGG UGAACAGCCAGCCCAACAUCACCAAC CUGUGCCCCUUCGGCGAGGUGUUCAA CGCCACCCGGUUCGCCAGCGUGUACG CCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUA CAACAGCGCCAGCUUCAGCACCUUCA AGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUG GCGACGAGGUGCGGCAGAUCGCACCC GGCCAGACAGGCAAGAUCGCCGACUA CAACUACAAGCUGCCCGACGACUUCA	73

[0383]

		CCGGCUGCGUGAUCGCCUGGAACAGC AACAACCUCGACAGCAAGGUGGGCGG CAACUACAACUACCUGUACCGGCUGU UCCGGAAGAGCAACCUGAAGCCCUUC GAGCGGGACAUCAGCACCGAGAUCUA CCAAGCCGGCUCCACCCCUUGCAACG GCGUGGAGGGCUUCAACUGCUACUUC CCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAG CUGCUGCACGCCCCAGCCACCGUGUG UGGCCCCAAGUCUGGCGGAGGCGGCA GCGCCAUCGGCGGCUACAUCCCCGAG GCCCCUAGAGACGGCCAGGCCUACGU GCGGAAGGACGGCGAGUGGGUGCUGC UGAGCACCUUCCUGGGC	
	3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
[0384]	相应的氨基酸序列	mysmqlascvtltlvllvnsQPNITNLCPFGEVFN ATRFASVYAWNRKRISNCVADYSVLYN SASFSTFKCYGVSPTKLNDLCFTNVYAD SFVIRGDEVRQIAPGQTGKIADYNYKLP DDFTGCVIAWNSNNLDSKVGGNYNYLY RLFRKSNLKPFERDISTEIYQAGSTPCNG VEGFNCYFPLQSYGFQPTNGVGYQPYR VVVLSFELLHAPATVCGPKsggggSAIGG YIPEAPRDGQAYVRKDGEWVLLSTFLG	74
	多聚 A 尾	100 nt	
	与折叠子结构域和跨膜结构域连接的 SARS-CoV-2 RBD (RBD-FD-TM)		
	SEQ ID NO: 78 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 79 和 3' UTR SEQ ID NO: 4。		78
	化学	1-甲基假尿苷	
	帽	7mG(5')ppp(5')NlmpNp	
	5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
	mRNA 构建体	AUGUACAGCAUGCAGCUGGCUAGCUG CGUGACCCUGACCCUGGUGCUGCUGG	79

[0385]

的 ORF (不包括终止密码子)	UGAACAGCCAGCCCAACAUCACCAAC CUGUGCCCCUUCGGCGAGGUGUUCAA CGCCACCCGGUUCGCCAGCGUGUACG CCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUA CAACAGCGCCAGCUUCAGCACCUUCA AGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUG GCGACGAGGUGCGGCAGAUCGCACCC GGCCAGACAGGCAAGAUCGCCGACUA CAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGC AACAACCUCGACAGCAAGGUGGGCGG CAACUACAACUACCUGUACCGGCUGU UCCGGAAGAGCAACCUGAAGCCCUUC GAGCGGGACAUCAGCACCGAGAUCUA CCAAGCCGGCUCCACCCCUUGCAACG GCGUGGAGGGCUUCAACUGCUACUUC CCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAG CUGCUGCACGCCCCAGCCACCGUGUG UGGCCCCAAGUCUGGCGGAGGCGGCA GCGCCAUCGGCGGCUACAUCCCCGAG GCCCCUAGAGACGGCCAGGCCUACGU GCGGAAGGACGGCGAGUGGGUGCUGC UGAGCACCUUCCUGGGCGGAGGCAGC AUCCUGGCCAUCUACAGCACCGUGGC CAGCAGCCUGGUGCUGCUGGUGAGCC UGGGCGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	mysmqlascvtltlvllvnsQPNITNLCPFGEVFN ATRFASVYAWNRKRISNCVADYSVLYN SASFSTFKCYGVSPTKLNDLCFTNVYAD SFVIRGDEVRQIAPGQTGKIADYNYKLP DDFTGCVIAWNSNNLDSKVGGNYNYLY RLFRKSNLKPFERDISTEIYQAGSTPCNG VEGFNCYFPLQSYGFQPTNGVGYQPYR	80

	VVVLSFELLHAPATVCGPKsggggSAIGG YIPEAPRDGQAYVRKDGEWVLLSTFLGg gsilaiystvasslvllvslgaisf	
多聚 A 尾	100 nt	
与跨膜结构域和折叠子结构域连接的 SARS-CoV-2 RBD (RBD-TM-FD)		
SEQ ID NO: 81 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 82 和 3' UTR SEQ ID NO: 4。		81
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUACAGCAUGCAGCUGGCUAGCUG CGUGACCCUGACCCUGGUGCUGCUGG UGAACAGCCAGCCCAACAUCACCAAC CUGUGCCCCUUCGGCGAGGUGUUCAA CGCCACCCGGUUCGCCAGCGUGUACG CCUGGAACCGGAAGCGGAUCAGCAAC UGCGUGGCCGACUACAGCGUGCUGUA CAACAGCGCCAGCUUCAGCACCUUCA AGUGCUACGGCGUGAGCCCCACCAAG CUGAACGACCUGUGCUUCACCAACGU GUACGCCGACAGCUUCGUGAUCCGUG GCGACGAGGUGCGGCAGAUCGCACCC GGCCAGACAGGCAAGAUCGCCGACUA CAACUACAAGCUGCCCGACGACUUCA CCGGCUGCGUGAUCGCCUGGAACAGC AACAACCUCGACAGCAAGGUGGGCGG CAACUACAACUACCUGUACCGGCUGU UCCGGAAGAGCAACCUGAAGCCCUUC GAGCGGGACAUCAGCACCGAGAUCUA CCAAGCCGGCUCCACCCCUUGCAACG GCGUGGAGGGCUUCAACUGCUACUUC CCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCU ACCGGGUGGUGGUGCUGAGCUUCGAG CUGCUGCACGCCCCAGCCACCGUGUG UGGCCCCAAGUCUGGCGGAGGCAGCA UCCUGGCCAUCUACAGCACCGUGGCC AGCAGCCUGGUGCUGCUGGUGAGCCU GGGCGCCAUCAGCUUCGGCGGAGGCA	82

[0386]

	GCGCCAUCGGCGGCUACAUCCCCGAG GCCCCUAGAGACGGCCAGGCCUACGU GCGGAAGGACGGCGAGUGGGUGCUGC UGAGCACCUUCCUGGGCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	mysmqlascvtltlvllvnsQPNITNLCPFGEVFN ATRFASVYAWNRKRISNCVADYSVLYN SASFSTFKCYGVSPTKLNDLCFTNVYAD SFVIRGDEVRQIAPGQTGKIADYNYKLP DDFTGCVIAWNSNNLDSKVGGNYNYLY RLFRKSNLKPFERDISTEIYQAGSTPCNG VEGFNCYFPLQSYGFQPTNGVGYQPYR VVVLSFELLHAPATVCGPKsgggsilaiystvas slvllvslgaisfgggSAIGGYIPEAPRDGQAYVR KDGEWVLLSTFLGk	83
多聚 A 尾	100 nt	

[0387]

与折叠子结构域和跨膜结构域连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-FD-TM)	
SEQ ID NO: 99 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 100 和 3' UTR SEQ ID NO: 4。	99
化学	1-甲基假尿苷
帽	7mG(5')ppp(5')NlmpNp
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGG AUCUUCGGCACCACCCUGGACAGCAA	10 0

[0388]

	GACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUG UGCGAGUUCCAGUUCUGCAACGACCC CUUCCUGGGCGUGUACUACCACAAGA ACAACAAGAGCUGGAUGGAGAGCGAG UUCCGGGUGUACAGCAGCGCCAACAA CUGCACCUUCGAGUACGUGAGCCAGC CCUUCCUGAUGGACCUGGAGGGCAAG CAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACC CCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAAC AUCACCCGGUUCCAGACCCUGCUGGC CCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCA GGCGCGGCUGCUUACUACGUGGGCUA CCUGCAGCCCCGGACCUUCCUGCUGA AGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACGGAGGCGGAUCGGG AGGCGGACCCAACAUCACCAACCUGU GCCCCUUCGGCGAGGUGUUCAACGCC ACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCG UGGCCGACUACAGCGUGCUGUACAAC AGCGCCAGCUUCAGCACCUUCAAGUG CUACGGCGUGAGCCCCACCAAGCUGA ACGACCUGUGCUUCACCAACGUGUAC GCCGACAGCUUCGUGAUCCGUGGCGA CGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAAC UACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACA ACCUCGACAGCAAGGUGGGCGGCAAC UACAACUACCUGUACCGGCUGUUCCG GAAGAGCAACCUGAAGCCCUUCGAGC GGGACAUCAGCACCGAGAUCUACCAA GCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUC UGCAGAGCUACGGCUUCCAGCCCACC AACGGCGUGGGCUACCAGCCCUACCG GGUGGUGGUGCUGAGCUUCGAGCUGC	

[0389]

	UGCACGCCCCAGCCACCGUGUGUGGC CCCAAGUCUGGCGGAGGCGGCAGCGC CAUCGGCGGCUACAUCCCCGAGGCCC CUAGAGACGGCCAGGCCUACGUGCGG AAGGACGGCGAGUGGGUGCUGCUGAG CACCUUCCUGGGCGGAGGCAGCAUCC UGGCCAUCUACAGCACCGUGGCCAGC AGCCUGGUGCUGCUGGUGAGCCUGGG CGCCAUCAGCUUC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFLG VYYHKNNKSWMESEFRVYSSANNCTFE YVSQPFLMDLEGKQGNFKNLREFVFKNI DGYFKIYSKHTPINLVRDLPQGFSALEPL VDLPIGINITRFQTLLALHRSYLTPGDSSS GWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDgggsgggPNITNLCPFGEVFNAT RFASVYAWNRKRISNCVADYSVLYNSA SFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDD FTGCVIAWNSNNLDSKVGGNYNYLYRL FRKSNLKPFERDISTEIYQAGSTPCNGVE GFNCYFPLQSYGFQPTNGVGYQPYRVV VLSFELLHAPATVCGPKsggggSAIGGYIP EAPRDGQAYVRKDGEWVLLSTFLGggsil aiystvasslvllvslgaisf	10 1
多聚 A 尾	100 nt	
与跨膜结构域和折叠子结构域连接的 SARS-CoV-2 NTD-RBD (NTD-RBD-TM-FD)		
SEQ ID NO: 102 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 103 和 3' UTR SEQ ID NO: 4。		10 2
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU	2

	AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGG AUCUUCGGCACCACCCUGGACAGCAA GACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUG UGCGAGUUCCAGUUCUGCAACGACCC CUUCCUGGGCGUGUACUACCACAAGA ACAACAAGAGCUGGAUGGAGAGCGAG UUCCGGGUGUACAGCAGCGCCAACAA CUGCACCUUCGAGUACGUGAGCCAGC CCUUCCUGAUGGACCUGGAGGGCAAG CAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACC CCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAAC AUCACCCGGUUCCAGACCCUGCUGGC CCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCA GGCGCGGCUGCUUACUACGUGGGCUA CCUGCAGCCCCGGACCUUCCUGCUGA AGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACGGAGGCGGAUCGGG AGGCGGACCCAACAUCACCAACCUGU GCCCCUUCGGCGAGGUGUUCAACGCC ACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCG UGGCCGACUACAGCGUGCUGUACAAC AGCGCCAGCUUCAGCACCUUCAAGUG CUACGGCGUGAGCCCCACCAAGCUGA	10 3

[0390]

[0391]

	ACGACCUGUGCUUCACCAACGUGUAC GCCGACAGCUUCGUGAUCCGUGGCGA CGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAAC UACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACA ACCUCGACAGCAAGGUGGGCGGCAAC UACAACUACCUGUACCGGCUGUUCCG GAAGAGCAACCUGAAGCCCUUCGAGC GGGACAUCAGCACCGAGAUCUACCAA GCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUC UGCAGAGCUACGGCUUCCAGCCCACC AACGGCGUGGGCUACCAGCCCUACCG GGUGGUGGUGCUGAGCUUCGAGCUGC UGCACGCCCCAGCCACCGUGUGUGGC CCCAAGUCUGGCGGAGGCAGCAUCCU GGCCAUCUACAGCACCGUGGCCAGCA GCCUGGUGCUGCUGGUGAGCCUGGGC GCCAUCAGCUUCGGCGGAGGCAGCGC CAUCGGCGGCUACAUCCCCGAGGCCC CUAGAGACGGCCAGGCCUACGUGCGG AAGGACGGCGAGUGGGUGCUGCUGAG CACCUUCCUGGGCAAG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFLG VYYHKNNKSWMESEFRVYSSANNCTFE YVSQPFLMDLEGKQGNFKNLREFVFKNI DGYFKIYSKHTPINLVRDLPQGFSALEPL VDLPIGINITRFQTLLALHRSYLTPGDSSS GWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDgggsgggPNITNLCPFGEVFNAT RFASVYAWNRKRISNCVADYSVLYNSA SFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDD	10 4

	FTGCVIAWNSNNLDSKVGGNYNYLYRL FRKSNLKPFERDISTEIYQAGSTPCNGVE GFNCYFPLQSYGFQPTNGVGYQPYRVV VLSFELLHAPATVCGPKsgggsilaiystvasslvl lvslgaisfgggSAIGGYIPEAPRDGQAYVRKD GEWVLLSTFLGk	
多聚 A 尾	100 nt	
与折叠子结构域连接的 SARS-CoV-2 NTD-RBD		
SEQ ID NO: 111 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 112 和 3' UTR SEQ ID NO: 4。		111
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGCC GCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGCC CCUGGUGAGCAGCCAGUGCGUGAACC UGACCACCCGGACCCAGCUGCCACCA GCCUACACCAACAGCUUCACCCGGGG CGUCUACUACCCCGACAAGGUGUUCC GGAGCAGCGUCCUGCACAGCACCCAG GACCUGUUCCUGCCCUUCUUCAGCAA CGUGACCUGGUUCCACGCCAUCCACG UGAGCGGCACCAACGGCACCAAGCGG UUCGACAACCCCGUGCUGCCCUUCAA CGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGG AUCUUCGGCACCACCCUGGACAGCAA GACCCAGAGCCUGCUGAUCGUGAAUA ACGCCACCAACGUGGUGAUCAAGGUG UGCGAGUUCCAGUUCUGCAACGACCC CUUCCUGGGCGUGUACUACCACAAGA ACAACAAGAGCUGGAUGGAGAGCGAG UUCCGGGUGUACAGCAGCGCCAACAA CUGCACCUUCGAGUACGUGAGCCAGC CCUUCCUGAUGGACCUGGAGGGCAAG CAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCU ACUUCAAGAUCUACAGCAAGCACACC CCAAUCAACCUGGUGCGGGAUCUGCC CCAGGGCUUCUCAGCCCUGGAGCCCC UGGUGGACCUGCCCAUCGGCAUCAAC	112

[0392]

[0393]

	AUCACCCGGUUCCAGACCCUGCUGGC CCUGCACCGGAGCUACCUGACCCCAG GCGACAGCAGCAGCGGGUGGACAGCA GGCGCGGCUGCUUACUACGUGGGCUA CCUGCAGCCCCGGACCUUCCUGCUGA AGUACAACGAGAACGGCACCAUCACC GACGCCGUGGACGGAGGCGGAUCGGG AGGCGGACCCAACAUCACCAACCUGU GCCCCUUCGGCGAGGUGUUCAACGCC ACCCGGUUCGCCAGCGUGUACGCCUG GAACCGGAAGCGGAUCAGCAACUGCG UGGCCGACUACAGCGUGCUGUACAAC AGCGCCAGCUUCAGCACCUUCAAGUG CUACGGCGUGAGCCCCACCAAGCUGA ACGACCUGUGCUUCACCAACGUGUAC GCCGACAGCUUCGUGAUCCGUGGCGA CGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAAC UACAAGCUGCCCGACGACUUCACCGG CUGCGUGAUCGCCUGGAACAGCAACA ACCUCGACAGCAAGGUGGGCGGCAAC UACAACUACCUGUACCGGCUGUUCCG GAAGAGCAACCUGAAGCCCUUCGAGC GGGACAUCAGCACCGAGAUCUACCAA GCCGGCUCCACCCCUUGCAACGGCGU GGAGGGCUUCAACUGCUACUUCCCUC UGCAGAGCUACGGCUUCCAGCCCACC AACGGCGUGGGCUACCAGCCCUACCG GGUGGUGGUGCUGAGCUUCGAGCUGC UGCACGCCCCAGCCACCGUGUGUGGC CCCAAGUCUGGCGGAGGCGGCAGCGC CAUCGGCGGCUACAUCCCCGAGGCCC CUAGAGACGGCCAGGCCUACGUGCGG AAGGACGGCGAGUGGGUGCUGCUGAG CACCUUCCUGGGC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCC CCCAGCCCCUCCUCCCCUUCCUGCACC CGUACCCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL	11 3

[0394]

	PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFLG VYYHKNNKSWMESEFRVYSSANNCTFE YVSQPFLMDLEGKQGNFKNLREFVFKNI DGYFKIYSKHTPINLVRDLPQGFSALEPL VDLPIGINITRFQTLLALHRSYLTPGDSSS GWTAGAAAYYVGYLQPRTFLLKYNEN GTITDAVDgggsgggPNITNLCPFGEVFNAT RFASVYAWNRKRISNCVADYSVLYNSA SFSTFKCYGVSPTKLNDLCFTNVYADSF VIRGDEVRQIAPGQTGKIADYNYKLPDD FTGCVIAWNSNNLDSKVGGNYNYLYRL FRKSNLKPFERDISTEIYQAGSTPCNGVE GFNCYFPLQSYGFQPTNGVGYQPYRVV VLSFELLHAPATVCGPKsggggSAIGGYIP EAPRDGQAYVRKDGEWVLLSTFLG	
多聚 A 尾	100 nt	

[0395] 包封蛋白

[0396] 在一些实施方案中,包封蛋白结构域用作支架结构域。包封蛋白是从嗜热生物海栖热袍菌(Thermotoga maritima)分离的蛋白质笼形纳米颗粒。包封蛋白由相同31kDa单体的60个拷贝组装而成,所述单体具有薄的二十面体T=1对称的笼形结构,其内径和外径分别为20nm和24nm(Sutter M.等人,Nat Struct Mol Biol.2008;15:939-947)。尽管包封蛋白在海栖热袍菌中的确切功能尚未清楚地为人们所了解,但最近其晶体结构已得以解决,并且其功能假定为包封例如DyP(染料脱色过氧化物酶)和Flp(铁蛋白样蛋白)的蛋白质的细胞区室,所述蛋白质参与氧化应激反应30(Rahmanpour R.等人,FEBS J.2013;280:2097-2104)。使用包封蛋白进行纳米颗粒构建使得蛋白质抗原能够在纳米颗粒表面上展示,并且能够将例如mRNA的货物包封在纳米颗粒本身内。以前的基于包封蛋白纳米颗粒的疫苗已引发对表面展示抗原和货物蛋白本身的强烈免疫反应(Lagoutte P.等人,Vaccine.2018;36(25):3622-3628)。

[0397] 在一些实施方案中,本文提供的mRNA编码与包封蛋白结构域连接的S蛋白结构域(例如,S1、S2、RBD和/或NTD)。

[0398] 融合蛋白

[0399] 在一些实施方案中,本公开的组合物包括编码抗原性融合蛋白的mRNA。因此,编码的一种或多种抗原可包括接合在一起的两种或更多种蛋白质(例如,蛋白质和/或蛋白质片段)。或者,与蛋白抗原融合的蛋白质不促进对其自身的强烈免疫反应,而是促进对冠状病毒抗原的强烈免疫反应。在一些实施方案中,抗原性融合蛋白保留来自每个原始蛋白质的功能性质。

[0400] 在一些实施方案中,融合蛋白包含来自SARS-CoV-2刺突蛋白的受体结合结构域。

[0401] 在一些实施方案中,融合蛋白包含来自SARS-CoV-2刺突蛋白的N末端结构域。

[0402] 在一些实施方案中,融合蛋白包含跨膜结构域。在一些实施方案中,所述跨膜结构域可来自不为SARS-CoV-2的病毒。例如,所述跨膜结构域可来自流感血凝素跨膜结构域,其

已展现有效地将蛋白质锚定在细胞表面。

[0403] 变体

[0404] 在一些实施方案中，本公开的组合物包括编码冠状病毒抗原变体的RNA。抗原变体或其他多肽变体是指其氨基酸序列与野生型、天然或参照序列不同的分子。与天然或参照序列相比，抗原/多肽变体可在氨基酸序列内的某些位置处具有取代、缺失和/或插入。通常，变体与野生型、天然或参照序列具有至少50％的同一性。在一些实施方案中，变体与野生型、天然或参照序列共享至少80％或至少90％的同一性。

[0405] 由本公开的核酸编码的变体抗原/多肽可含有赋予多种期望性质中的任一者的氨基酸变化，例如，在受试者中增强其免疫原性、增强其表达和/或改善其稳定性或PK/PD性质。变体抗原/多肽可使用常规诱变技术制备，并且适当地进行测定以确定其是否具有期望性质。用以确定表达水平和免疫原性的测定是本领域众所周知的，并且示例性的此类测定在实施例部分中阐释。类似地，可使用本领域公认的技术，例如通过测定抗原在接种的受试者中随时间推移的表达和/或通过观察诱导的免疫反应的耐久性，来测量蛋白质变体的PK/PD性质。由变体核酸编码的蛋白质的稳定性可通过测定热稳定性或尿素变性时的稳定性来测量，或者可使用计算机预测来测量。用于此类实验和计算机测定的方法是本领域中已知的。

[0406] 在一些实施方案中，组合物包含mRNA或mRNA ORF，其包含本文提供的序列中的任一者的核苷酸序列(参见例如序列表)，或包含与本文提供的序列中的任一者的核苷酸序列至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一的核苷酸序列。

[0407] 术语“同一性”是指如通过比较序列确定的两种或更多种多肽(例如抗原)或多核苷酸(核酸)的序列之间的关系。同一性也指如通过两个或更多个氨基酸残基或核酸残基串之间的匹配数确定的序列之间或之中的序列相关性程度。同一性度量由特定数学模型或计算机程序(例如，“算法”)处理的具有空位比对(如果有的话)的两个或更多个序列中的较小者间的同一性匹配百分比。相关抗原或核酸的同一性可容易地通过已知方法计算。“同一性百分比(％)”当应用于多肽或多核苷酸序列时，定义为在比对序列并引入空位(必要时)以实现最大同一性百分比之后，候选氨基酸或核酸序列中与第二序列的氨基酸序列或核酸序列中的残基(氨基酸残基或核酸残基)同一的残基的百分数。用于比对的方法和计算机程序是本领域中众所周知的。应理解，同一性取决于同一性百分比的计算，但由于计算中引入的空位和罚分，同一性的值可能不同。通常，特定多核苷酸或多肽(例如抗原)的变体与所述特定参照多核苷酸或多肽具有如通过本文所述和本领域技术人员已知的序列比对程序和参数所确定的至少40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％但小于100％的序列同一性。用于比对的此类工具包括BLAST套件的那些工具(Stephen F.Altschul等人(1997)，“Gapped BLAST and PSI-BLAST:a new generation of protein database search programs”,Nucleic Acids Res.25:3389-3402)。另一种常用局部比对技术是基于Smith-Waterman算法(Smith,T.F.和Waterman,M.S.(1981)“Identification of common molecular subsequences.”J.Mol.Biol.147:195-197)。基于动态程序化的通用全局比对技术是Needleman-Wunsch算法(Needleman,S.B.和Wunsch,C.D.(1970)“A general method applicable to the

search for similarities in the amino acid sequences of two proteins." J.Mol.Biol.48:443-453)。最近,已开发出一种快速最佳全局序列比对算法(FOGSAA),据称其比其他最佳全局比对方法(包括Needleman-Wunsch算法)更快地产生核苷酸和蛋白质序列的全局比对。

[0408] 因此,编码相对于参照序列、特别是本文公开的多肽(例如抗原)序列含有取代、插入和/或添加、缺失和共价修饰的肽或多肽的多核苷酸包括在本公开的范围内。例如,序列标签或氨基酸(例如一个或多个赖氨酸)可添加至肽序列中(例如,在N末端或C末端)。序列标签可用于肽检测、纯化或定位。赖氨酸可用于增加肽的溶解度或允许生物素化。或者,位于肽或蛋白质的氨基酸序列的羧基和氨基末端区的氨基酸残基可任选地缺失,以提供截短的序列。根据序列的用途,例如作为可溶或连接至固体载体的更大序列的一部分的序列的表达,某些氨基酸(例如,C末端或N末端残基)可替代地缺失。在一些实施方案中,信号序列、终止序列、跨膜结构域、接头、多聚化结构域(例如,折叠子区)等(或编码其)的序列可用实现相同或相似功能的替代序列取代。在一些实施方案中,可例如通过引入更大的氨基酸填充蛋白质核心中的空腔以改善稳定性。在其他实施方案中,可用疏水性残基置换包埋的氢键网络以改进稳定性。在其他实施方案中,可去除糖基化位点并用适当残基置换。本领域技术人员可容易地鉴定此类序列。还应理解,本文提供的一些序列含有可在例如用于制备mRNA疫苗之前缺失的序列标签或末端肽序列(例如,在N末端或C末端)。

[0409] 如由本领域技术人员认识到,也将蛋白质片段、功能蛋白质结构域和同源蛋白质视为在目标冠状病毒抗原的范围内。例如,本文提供了参照蛋白质的任何蛋白质片段(意指比参照抗原序列短至少一个氨基酸残基、但在其他方面相同的多肽序列),条件是所述片段是免疫原性的并赋予针对冠状病毒的保护性免疫反应。除了与参照蛋白质相同但截短的变体之外,在一些实施方案中,如本文提供或参考的任何序列中所示,抗原包括2、3、4、5、6、7、8、9、10或更多个突变。抗原/抗原性多肽的长度范围可为约4、6或8个氨基酸至全长蛋白质。

[0410] 稳定化元件

[0411] 天然存在的真核mRNA分子除了例如5′-帽结构或3′-多聚(A)尾的其他结构特征之外,还可含有稳定化元件,包括但不限于在其5′端(5′UTR)和/或在其3′端(3′UTR)的非翻译区(UTR)。5′UTR和3′UTR二者都通常是从基因组DNA转录并且是成熟前mRNA的元件。通常在mRNA加工期间向转录的(成熟前)mRNA中添加成熟mRNA的特征性结构特征(例如5′-帽和3′-多聚(A)尾)。

[0412] 在一些实施方案中,组合物包括具有编码至少一种具有至少一个修饰的抗原性多肽的开放阅读框、至少一个5′末端帽的mRNA,并且配制在脂质纳米颗粒内。多核苷酸的5′-加帽可根据制造商的方案在体外转录反应期间使用以下化学RNA帽类似物同时完成,以产生5′-鸟苷帽结构:3′-O-Me-m7G(5′)ppp(5′)G[ARCA帽];G(5′)ppp(5′)A;G(5′)ppp(5′)G;m7G(5′)ppp(5′)A;m7G(5′)ppp(5′)G(New England BioLabs,Ipswich,MA)。经修饰的RNA的5′-加帽可在转录后使用牛痘病毒加帽酶完成,以产生"帽0"结构:m7G(5′)ppp(5′)G(NeW England BioLabs,Ipswich,MA)。帽1结构可使用牛痘病毒加帽酶和2′-O甲基转移酶产生,以产生:m7G(5′)ppp(5′)G-2′-O-甲基。帽2结构可从帽1结构、接着使用2′-O甲基转移酶对5′-倒数第三个核苷酸进行2′-O-甲基化来产生。帽3结构可从帽2结构、接着使用2′-O甲基转移酶对5′-倒数第四个核苷酸进行2′-O-甲基化来产生。酶可源自重组来源。

[0413] 3′-多聚(A)尾通常是添加至转录的mRNA的3′端的一段腺嘌呤核苷酸。在一些情况下,其可包含高达约400个腺嘌呤核苷酸。在一些实施方案中,3′-多聚(A)尾的长度对于个别mRNA的稳定性可为必需的要素。

[0414] 在一些实施方案中,组合物包括稳定化元件。稳定化元件可包括例如组蛋白茎环。已鉴定茎环结合蛋白(SLBP),一种32kDa的蛋白质。其与细胞核和细胞质二者中组蛋白信使3′端的组蛋白茎环结合。其表达水平受细胞周期的调控;其在S期达到峰值,此时组蛋白mRNA水平也升高。已显示所述蛋白质对于U7 snRNP对组蛋白前mRNA的有效3′端加工是至关重要的。SLBP在加工后继续与茎环结合,然后刺激成熟组蛋白mRNA在细胞质中翻译成组蛋白。SLBP的RNA结合结构域在后生动物和原生动物中是保守的;其与组蛋白茎环的结合取决于环的结构。最小结合位点包括相对于茎环的5′的至少三个核苷酸和3′的两个核苷酸。

[0415] 在一些实施方案中,mRNA包括编码区、至少一个组蛋白茎环和任选地多聚(A)序列或多聚腺苷酸化信号。所述多聚(A)序列或多聚腺苷酸化信号通常应增强编码的蛋白质的表达水平。在一些实施方案中,所述编码的蛋白质不为组蛋白、报告基因蛋白(例如荧光素酶、GFP、EGFP、β-半乳糖苷酶、EGFP)或标志物或选择蛋白(例如α-球蛋白、半乳糖激酶和黄嘌呤:鸟嘌呤磷酸核糖基转移酶(GPT))。

[0416] 在一些实施方案中,mRNA包括多聚(A)序列或多聚腺苷酸化信号与至少一个组蛋白茎环的组合,即使两者在性质上代表替代机制,但协同地作用以增加蛋白质表达超过用任一个别元件观察到的水平。多聚(A)和至少一种组蛋白茎环的组合的协同效应不依赖于元件的顺序或多聚(A)序列的长度。

[0417] 在一些实施方案中,mRNA不包括组蛋白下游元件(HDE)。“组蛋白下游元件”(HDE)包括在天然存在的茎环的3′的大约15至20个核苷酸的富含嘌呤的多核苷酸段,其代表U7 snRNA的结合位点,所述U7 snRNA参与组蛋白前mRNA加工成成熟组蛋白mRNA。在一些实施方案中,核酸不包括内含子。

[0418] mRNA可含有或可不含有增强子和/或启动子序列,其可经修饰或未经修饰,或者其可经活化或未经活化。在一些实施方案中,组蛋白茎环通常源自组蛋白基因,并且包括由间隔体分开的两个相邻的部分或完全反向互补序列的分子内碱基配对,所述间隔体由形成结构的环的短序列组成。未配对的环区通常不能与茎环元件中的任一者碱基配对。其由于是许多RNA二级结构的关键组分而更经常地出现于RNA中,但也可存在于单链DNA中。茎环结构的稳定性通常取决于长度、错配或膨出部的数目和配对区的碱基组成。在一些实施方案中,可产生摆动碱基配对(非沃森-克里克碱基配对(non-Watson-Crick base pairing))。在一些实施方案中,至少一种组蛋白茎环序列包含15至45个核苷酸的长度。

[0419] 在一些实施方案中,mRNA已去除一个或多个富含AU的序列。这些序列有时称为AURES,其是在3′UTR中发现的去稳定化序列。AURES可从RNA疫苗去除。或者,AURES可保留在RNA疫苗中。

[0420] 信号肽

[0421] 在一些实施方案中,组合物包含具有ORF的mRNA,所述ORF编码与冠状病毒抗原融合的信号肽。包含蛋白质N末端15-60个氨基酸的信号肽通常是分泌途径上跨膜易位所需要的,并且因此,普遍控制真核生物和原核生物中大部分蛋白质进入分泌途径。在真核生物中,新生前体蛋白(前蛋白)的信号肽将核糖体引导至粗内质网(ER)膜,并且起始生长的肽

链穿过其转运以进行加工。ER加工产生成熟蛋白质，其中信号肽通常通过宿主细胞的ER驻留信号肽酶从前体蛋白切割，或者其保持未切割并作为膜锚起作用。信号肽也可促进蛋白质靶向细胞膜。

[0422] 信号肽的长度可为15-60个氨基酸。例如，信号肽的长度可为15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59或60个氨基酸。在一些实施方案中，信号肽的长度为20-60、25-60、30-60、35-60、40-60、45-60、50-60、55-60、15-55、20-55、25-55、30-55、35-55、40-55、45-55、50-55、15-50、20-50、25-50、30-50、35-50、40-50、45-50、15-45、20-45、25-45、30-45、35-45、40-45、15-40、20-40、25-40、30-40、35-40、15-35、20-35、25-35、30-35、15-30、20-30、25-30、15-25、20-25或15-20个氨基酸。

[0423] 来自异源基因的信号肽(其调控自然界中除冠状病毒抗原外的基因的表达)是本领域中已知的，并且可测试其期望性质，然后并入本公开的核酸中。

[0424] 序列优化

[0425] 在一些实施方案中，对编码本公开的抗原的ORF进行密码子优化。密码子优化方法是本领域中已知的。例如，本文提供的任何一个或多个序列的ORF可为密码子优化的。在一些实施方案中，密码子优化可用于匹配靶标和宿主生物体中的密码子频率以确保正确折叠；偏向GC含量以增加mRNA稳定性或减少二级结构；最小化可损害基因构建或表达的串联重复密码子或碱基操作；定制转录和翻译控制区；插入或去除蛋白质运输序列；在编码的蛋白质中去除/添加翻译后修饰位点(例如，糖基化位点)；添加、去除或改组蛋白结构域；插入或缺失限制性位点；修饰核糖体结合位点和mRNA降解位点；调整翻译率以允许蛋白质的各种结构域正确折叠；或者减少或消除多核苷酸内的问题二级结构。密码子优化工具、算法和服务是本领域中已知的-非限制性实例包括来自GeneArt(Life Technologies)、DNA2.0(Menlo Park CA)和/或专有方法的服务。在一些实施方案中，使用优化算法优化开放阅读框(ORF)序列。

[0426] 在一些实施方案中，密码子优化的序列与天然存在或野生型序列ORF(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享小于95％的序列同一性。在一些实施方案中，密码子优化的序列与天然存在或野生型序列(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享小于90％的序列同一性。在一些实施方案中，密码子优化的序列与天然存在或野生型序列(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享小于85％的序列同一性。在一些实施方案中，密码子优化的序列与天然存在或野生型序列(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享小于80％的序列同一性。在一些实施方案中，密码子优化的序列与天然存在或野生型序列(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享小于75％的序列同一性。

[0427] 在一些实施方案中，密码子优化的序列与天然存在或野生型序列(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享介于65％与85％之间(例如，介于约67％与约85％之间或介于约67％与约80％之间)的序列同一性。在一些实施方案中，密码子优化的序列与天然存在或野生型序列(例如，编码冠状病毒抗原的天然存在或野生型mRNA序列)共享介于65％与75％或约80％之间的序列同一性。

[0428] 在一些实施方案中，密码子优化的序列编码的抗原与由非密码子优化的序列编码

的冠状病毒抗原的免疫原性一样，或比其免疫原性更高(例如，高至少10％、至少20％、至少30％、至少40％、至少50％、至少100％或至少200％)。

[0429] 经修饰的mRNA在转染至哺乳动物宿主细胞中时具有12-18小时之间或大于18小时、例如24小时、36小时、48小时、60小时、72小时或大于72小时的稳定性，并且能够由哺乳动物宿主细胞表达。

[0430] 在一些实施方案中，密码子优化的RNA可为其中G/C的水平增加的RNA。核酸分子(例如mRNA)的G/C含量可影响RNA的稳定性。具有增加量的鸟嘌呤(G)和/或胞嘧啶(C)残基的RNA可在功能上比含有大量腺嘌呤(A)和胸腺嘧啶(T)或尿嘧啶(U)核苷酸的mRNA更稳定。举例来说，WO02/098443公开了一种药物组合物，其含有通过翻译区中的序列修饰而稳定化的mRNA。由于遗传密码的简并性，修饰通过用促进更大RNA稳定性而不改变所得氨基酸的那些密码子取代现有密码子来起作用。所述方法限于RNA的编码区。

[0431] 化学上未经修饰的核苷酸

[0432] 在一些实施方案中，mRNA未经化学修饰，并且包含由腺苷、鸟苷、胞嘧啶和尿苷组成的标准核糖核苷酸。在一些实施方案中，本公开的核苷酸和核苷包含标准核苷残基，例如存在于转录的RNA中的那些核苷残基(例如A、G、C或U)。在一些实施方案中，本公开的核苷酸和核苷包含标准脱氧核糖核苷，例如存在于DNA中的那些脱氧核糖核苷(例如dA、dG、dC或dT)。

[0433] 化学修饰

[0434] 在一些实施方案中，本公开的组合物包含具有编码冠状病毒抗原的开放阅读框的mRNA，其中核酸包含可为标准的(未经修饰的)或如本领域中已知修饰的核苷酸和/或核苷。在一些实施方案中，本公开的核苷酸和核苷包含经修饰的核苷酸或核苷。此类经修饰的核苷酸和核苷可为天然存在的经修饰的核苷酸和核苷或非天然存在的经修饰的核苷酸和核苷。此类修饰可包括如本领域中公认的在核苷酸和/或核苷的糖、主链或核碱基部分的那些修饰。

[0435] 在一些实施方案中，本公开的天然存在的经修饰的核苷酸或核苷是本领域中通常已知或公认的核苷酸或核苷。此类天然存在的经修饰的核苷酸和核苷的非限制性实例尤其可见于广泛公认的MODOMICS数据库中。

[0436] 在一些实施方案中，本公开的非天然存在的经修饰的核苷酸或核苷是本领域中通常已知或公认的核苷酸或核苷。此类非天然存在的经修饰的核苷酸和核苷的非限制性实例尤其可见于公开的美国申请号PCT/US2012/058519;PCT/US2013/075177;PCT/US2014/058897;PCT/US2014/058891;PCT/US2014/070413;PCT/US2015/36773;PCT/US2015/36759;PCT/US2015/36771;或PCT/IB2017/051367，所述申请都通过引用并入本文。

[0437] 因此，本公开的核酸(例如，DNA核酸和RNA核酸，例如mRNA核酸)可包含标准核苷酸和核苷、天然存在的核苷酸和核苷、非天然存在的核苷酸和核苷或其任何组合。

[0438] 在一些实施方案中，本公开的核酸(例如，DNA核酸和RNA核酸，例如mRNA核酸)包含多种(一种以上)不同类型的标准和/或经修饰的核苷酸和核苷。在一些实施方案中，核酸的特定区含有一种、两种或更多种(任选地不同)类型的标准和/或经修饰的核苷酸和核苷。

[0439] 在一些实施方案中，相对于包含标准核苷酸和核苷的未经修饰的核酸，引入细胞或生物体的经修饰的RNA核酸(例如，经修饰的mRNA核酸)分别在细胞或生物体中展现降低

的降解。

[0440] 在一些实施方案中，相对于包含标准核苷酸和核苷的未经修饰的核酸，引入细胞或生物体中的经修饰的RNA核酸(例如，经修饰的mRNA核酸)可分别在细胞或生物体中展现降低的免疫原性(例如，降低的先天反应)。

[0441] 在一些实施方案中，核酸(例如RNA核酸，例如mRNA核酸)包含非天然修饰的核苷酸，其在核酸的合成期间或合成后引入以实现期望功能或性质。修饰可存在于核苷酸间键联、嘌呤或嘧啶碱基或糖上。修饰可用化学合成或用聚合酶在链的末端或链中的任何其他地方引入。核酸的任何区可经化学修饰。

[0442] 本公开提供了核酸(例如RNA核酸，例如mRNA核酸)的经修饰的核苷和核苷酸。“核苷”是指含有糖分子(例如戊糖或核糖)或其衍生物与有机碱基(例如嘌呤或嘧啶)或其衍生物(在本文中也称为“核碱基”)组合的化合物。“核苷酸”是指包括磷酸基团的核苷。经修饰的核苷酸可通过任何有用的方法，例如化学、酶促或重组方法合成，以包括一个或多个经修饰的或非天然的核苷。核酸可包含连接的核苷的一个或多个区。此类区可具有可变的主链键联。键联可为标准磷酸二酯键联，在这种情形下，核酸将包含核苷酸的区。

[0443] 经修饰的核苷酸碱基配对不仅涵盖标准腺苷-胸腺嘧啶、腺苷-尿嘧啶或鸟苷-胞嘧啶碱基对，而且还包括在核苷酸和/或包含非标准或经修饰的碱基的经修饰的核苷酸之间形成的碱基对，其中氢键供体和氢键受体的排列允许在非标准碱基与标准碱基之间或在两个互补非标准碱基结构之间、例如在具有至少一个化学修饰的那些核酸中的氢键。此类非标准碱基配对的一个实例是经修饰的核苷酸肌苷与腺嘌呤、胞嘧啶或尿嘧啶之间的碱基配对。碱基/糖或接头的任何组合可并入本公开的核酸中。

[0444] 在一些实施方案中，核酸(例如RNA核酸，例如mRNA核酸)中经修饰的核碱基包含1-甲基-假尿苷(m1ψ)、1-乙基-假尿苷(e1ψ)、5-甲氧基-尿苷(mo5U)、5-甲基-胞苷(m5C)和/或假尿苷(ψ)。在一些实施方案中，核酸(例如RNA核酸，例如mRNA核酸)中的经修饰的核碱基包含5-甲氧基甲基尿苷、5-甲硫基尿苷、1-甲氧基甲基假尿苷、5-甲基胞苷和/或5-甲氧基胞苷。在一些实施方案中，多核糖核苷酸包括至少两种(例如，2、3、4种或更多种)任何上文所提及的经修饰的核碱基的组合，包括但不限于化学修饰。

[0445] 在一些实施方案中，本公开的mRNA包含在核酸的一个或多个或所有尿苷位置的1-甲基-假尿苷(m1ψ)取代。

[0446] 在一些实施方案中，本公开的mRNA包含在核酸的一个或多个或所有尿苷位置的1-甲基-假尿苷(m1ψ)取代和在核酸的一个或多个或所有胞苷位置的5-甲基胞苷取代。

[0447] 在一些实施方案中，本公开的mRNA包含在核酸的一个或多个或所有尿苷位置的假尿苷(ψ)取代。

[0448] 在一些实施方案中，本公开的mRNA包含在核酸的一个或多个或所有尿苷位置的假尿苷(ψ)取代和在核酸的一个或多个或所有胞苷位置的5-甲基胞苷取代。

[0449] 在一些实施方案中，本公开的mRNA包含在核酸的一个或多个或所有尿苷位置的尿苷。

[0450] 在一些实施方案中，对于特定修饰，对mRNA进行均匀修饰(例如，完全修饰、贯穿整个序列中修饰)。例如，核酸可用1-甲基-假尿苷均匀修饰，这意味着mRNA序列中的所有尿苷残基都用1-甲基-假尿苷置换。类似地，通过用经修饰的残基(例如上文所阐释的那些残基)

置换,可针对序列中存在的任何类型的核苷残基对核酸进行均匀修饰。

[0451] 本公开的核酸可沿着分子的整个长度经部分或完全修饰。例如,在本公开的核酸中,或在其预定序列区中(例如,在包括或不包括多聚(A)尾的mRNA中),可均匀地修饰一种或多种或所有或给定类型的核苷酸(例如,嘌呤或嘧啶,或任何一种或多种或所有A、G、U、C)。在一些实施方案中,本公开的核酸中(或其序列区中)的所有核苷酸X都是经修饰的核苷酸,其中X可为核苷酸A、G、U、C中的任一者,或组合A+G、A+U、A+C、G+U、G+C、U+C、A+G+U、A+G+C、G+U+C或A+G+C中的任一者。

[0452] 所述核酸可含有约1%至约100%的经修饰的核苷酸(相对于总核苷酸含量,或相对于一种或多种类型的核苷酸,即A、G、U或C中的任何一者或多者)或任何居间百分比(例如,1%至20%、1%至25%、1%至50%、1%至60%、1%至70%、1%至80%、1%至90%、1%至95%、10%至20%、10%至25%、10%至50%、10%至60%、10%至70%、10%至80%、10%至90%、10%至95%、10%至100%、20%至25%、20%至50%、20%至60%、20%至70%、20%至80%、20%至90%、20%至95%、20%至100%、50%至60%、50%至70%、50%至80%、50%至90%、50%至95%、50%至100%、70%至80%、70%至90%、70%至95%、70%至100%、80%至90%、80%至95%、80%至100%、90%至95%、90%至100%和95%至100%)。应理解,未经修饰的A、G、U或C的存在占任何其余百分数。

[0453] mRNA可含有最少1%并且最多100%的经修饰的核苷酸或任何居间百分数,例如至少5%的经修饰的核苷酸、至少10%的经修饰的核苷酸、至少25%的经修饰的核苷酸、至少50%的经修饰的核苷酸、至少80%的经修饰的核苷酸或至少90%的经修饰的核苷酸。例如,所述核酸可含有经修饰的嘧啶,例如经修饰的尿嘧啶或胞嘧啶。在一些实施方案中,所述核酸中至少5%、至少10%、至少25%、至少50%、至少80%、至少90%或100%的尿嘧啶由经修饰的尿嘧啶(例如5-取代的尿嘧啶)置换。经修饰的尿嘧啶可由具有单一独特结构的化合物置换,或者可由具有不同结构(例如,2、3、4种或更多种独特结构)的多种化合物置换。在一些实施方案中,核酸中至少5%、至少10%、至少25%、至少50%、至少80%、至少90%或100%的胞嘧啶由经修饰的胞嘧啶(例如5-取代的胞嘧啶)置换。经修饰的胞嘧啶可由具有单一独特结构的化合物置换,或者可由具有不同结构(例如,2、3、4种或更多种独特结构)的多种化合物置换。

[0454] 非翻译区(UTR)

[0455] 本公开的mRNA可包含一个或多个充当或用作非翻译区的区或部分。当mRNA被设计为编码至少一种目标抗原时,核酸可包含这些非翻译区(UTR)中的一者或多者。核酸的野生型非翻译区经转录,而非经翻译。在mRNA中,5′UTR在转录起始位点开始,并继续至起始密码子,但不包括起始密码子;而3′UTR在终止密码子之后立即开始,并继续至转录终止信号。越来越多的证据表明UTR在核酸分子的稳定性和翻译方面所起的调控作用。UTR的调控特征可并入本公开的多核苷酸中,以尤其增强分子的稳定性。还可并入特定特征以确保在转录物被错误导向不期望器官部位的情形下转录物的受控下调。多种5′UTR和3′UTR序列是本领域中已知和可获得的。

[0456] 5′UTR是mRNA的在起始密码子(核糖体翻译的mRNA转录物的第一个密码子)正上游(5′)的区。5′UTR不编码蛋白质(是非编码的)。天然5′UTR具有在翻译起始中起作用的特征。它们带有印记如Kozak序列,通常已知所述Kozak序列参与核糖体起始许多基因翻译的过

程。Kozak序列具有共有CCR(A/G)CCAUGG(SEQ ID NO:128),其中R是起始密码子(AUG)上游三个碱基处的嘌呤(腺嘌呤或鸟嘌呤),起始密码子(AUG)后接另一个'G'。还已知5′UTR形成参与延伸因子结合的二级结构。

[0457] 在本公开的一些实施方案中,5′UTR是异源UTR,即,是在自然界中发现的与不同ORF相关的UTR。在另一个实施方案中,5′UTR是合成UTR,即,在自然界中不出现。合成的UTR包括已经突变以改善其性质的UTR,例如,增加基因表达的UTR,以及完全合成的UTR。示例性5′UTR包括非洲爪蟾属(Xenopus)或人类源a-球蛋白或b-球蛋白(8278063;9012219)、人类细胞色素b-245a多肽和羟基类固醇(17b)脱氢酶和烟草蚀纹病毒(US8278063、9012219)。还可使用CMV立即早期1(IE1)基因(US20140206753、WO2013/185069)、序列GGGAUCCUACC(SEQ ID NO:129)(WO2014144196)。在另一个实施方案中,TOP基因的5′UTR是缺乏5′TOP基序(寡嘧啶束)的TOP基因的5′UTR(例如,WO/2015101414、WO2015101415、WO/2015/062738、WO2015024667、WO2015024667);可使用源自核糖体蛋白大32(L32)基因的5′UTR元件(WO/2015101414、WO2015101415、WO/2015/062738)、源自羟基类固醇(17-β)脱氢酶4基因(HSD17B4)的5′UTR的5′UTR元件(WO2015024667)或源自ATP5A1的5′UTR的5′UTR元件(WO2015024667)。在一些实施方案中,使用内部核糖体进入位点(IRES)代替5′UTR。

[0458] 在一些实施方案中,本公开的5′UTR包含选自SEQ ID NO:131和SEQ ID NO:2的序列。

[0459] 3′UTR是mRNA的在终止密码子(mRNA转录物的发出翻译终止信号的密码子)正下游(3′)的区。3′UTR不编码蛋白质(是非编码的)。已知天然或野生型3′UTR具有嵌入其中的腺苷和尿苷的段。这些富AU的印记在具有高周转率的基因中特别普遍。富AU的元件(ARE)基于其序列特征和功能特性可分成三类(Chen等人,1995):I类ARE在富U区内含有AUUUA基序的若干分散的拷贝。C-Myc和MyoD含有I类ARE。II类ARE具有两个或更多个重叠UUAUUUA(U/A)(U/A)(SEQ ID NO:130)九聚体。含有这种类型的ARE的分子包括GM-CSF和TNF-a。III类ARE定义不太明确。这些富U区不含AUUUA基序。c-Jun和肌细胞生成素是这种类别的两个充分研究的实例。已知大多数结合至ARE的蛋白质使信使不稳定,而据记载ELAV家族的成员(最著名的是HuR)增加mRNA的稳定性。HuR结合至所有三个类别的ARE。将HuR特异性结合位点工程化至核酸分子的3′UTR中将导致HuR结合,并因此导致体内信使的稳定化。

[0460] 3′UTR富AU元件(ARE)的引入、去除或修饰可用于调节本公开的核酸(例如RNA)的稳定性。当工程化特定核酸时,可引入ARE的一个或多个拷贝以使本公开的核酸更不稳定,从而缩减翻译并减少所得蛋白质的产生。同样,可鉴定和去除或突变ARE以增加细胞内稳定性,从而增加所得蛋白质的翻译和产生。可在相关细胞系中使用本公开的核酸执行转染实验,并且可在转染后的不同时间点测定蛋白质产生。例如,可用不同ARE工程化分子,并且通过对相关蛋白质使用ELISA试剂盒并测定在转染后6小时、12小时、24小时、48小时和7天产生的蛋白质,来转染细胞。

[0461] 3′UTR可为异源的或合成的。关于3′UTR,球蛋白UTR(包括非洲爪蟾属β-球蛋白UTR和人类β-球蛋白UTR)是本领域中已知的(8278063、9012219、US20110086907)。通过首尾相连地克隆两个连续的人类β-球蛋白3′UTR,已开发在一些细胞类型中具有增强的稳定性的经修饰的β-球蛋白构建体,并且其是本领域中众所周知的(US2012/0195936、WO2014/071963)。另外,a2-球蛋白、a1-球蛋白、UTR和其突变体也是本领域中已知的

(WO2015101415、WO2015024667)。在非专利文献中mRNA构建体中阐述的其他3′UTR包括CYBA(Ferizi等人,2015)和白蛋白(Thess等人,2015)。其他示例性3′UTR包括牛或人类生长激素(野生型或经修饰的)(WO2013/185069、US20140206753、WO2014152774)、兔β球蛋白和乙型肝炎病毒(HBV)的3′UTR,α-球蛋白3′UTR和病毒VEEV 3′UTR序列也是本领域中已知的。在一些实施方案中,使用序列UUUGAAUU(WO2014144196)。在一些实施方案中,使用人类和小鼠核糖体蛋白的3′UTR。其他实例包括rps93′UTR(WO2015101414)、FIG4(WO2015101415)和人类白蛋白7(WO2015101415)。

[0462] 在一些实施方案中,本公开的3′UTR包含选自SEQ ID NO:132和SEQ ID NO:4的序列。

[0463] 本领域的普通技术人员将理解,异源或合成的5′UTR可与任何期望的3′UTR序列一起使用。例如,异源5′UTR可与合成3′UTR或异源3′UTR一起使用。

[0464] 非UTR序列也可用作核酸内的区或亚区。例如,内含子或内含子序列的部分可并入本公开的核酸的区中。内含子序列的并入可增加蛋白质产生以及核酸水平。

[0465] 特征的组合可包括在侧翼区中并且可包含在其他特征内。例如,ORF的侧翼可为5′UTR,其可含有强Kozak翻译起始信号;和/或3′UTR,其可包括用于以模板添加多聚A尾的寡(dT)序列。5′UTR可包含来自相同和/或不同基因的第一多核苷酸片段和第二多核苷酸片段,例如美国专利申请公开号20100293625和PCT/US2014/069155中所述的5′UTR,所述公开通过引用整体并入本文。

[0466] 应理解,来自任何基因的任何UTR可并入核酸的区中。此外,可利用任何已知基因的多个野生型UTR。提供不为野生型区的变体的人工UTR也在本公开的范围内。这些UTR或其部分可以与其所选自的转录物中相同的方向放置,或者可在方向或位置上改变。因此,5′或3′UTR可倒置、缩短、延长,与一个或多个其他5′UTR或3′UTR一起制备。如本文所用的术语“改变的”当关于UTR序列时,意指相对于参考序列,UTR已经以某种方式改变。例如,3′UTR或5′UTR可通过如上文教导的方向或位置的改变而相对于野生型或天然UTR改变,或者可通过包括额外核苷酸、核苷酸的缺失、核苷酸的交换或转座而改变。产生“改变的”UTR(无论3′或5′)的这些改变中的任一者包含变体UTR。

[0467] 在一些实施方案中,可使用双重、三重或四重UTR,例如5′UTR或3′UTR。如本文所用的“双重”UTR是其中串联或基本上串联编码相同UTR的两个拷贝的UTR。例如,可使用双重β-球蛋白3′UTR,如美国专利公开20100129877中所述,所述专利公开的内容通过引用整体并入本文。

[0468] 具有模式化UTR也在本公开的范围内。如本文所用的“模式化UTR”是反映重复或交替模式的那些UTR,例如重复一次、两次或超过3次的ABABAB或AABBAABBAABB或ABCABCABC或其变体。在这些模式中,每个字母A、B或C代表核苷酸水平的不同UTR。

[0469] 在一些实施方案中,侧翼区选自其蛋白质共享共同功能、结构、特征或性质的转录物家族。例如,目标多肽可属于在特定细胞、组织中或在发育期间的某个时间表达的蛋白质家族。来自这些基因中的任一者的UTR可交换为相同或不同蛋白质家族的任何其他UTR,以产生新的多核苷酸。如本文所用的“蛋白质家族”以最广泛的意义使用,指共享至少一种功能、结构、特征、定位、起源或表达模式的一组两种或更多种目标多肽。

[0470] 非翻译区还可包括翻译增强子元件(TEE)。作为一个非限制性实例,TEE可包括美

国申请号20090226470中阐述的那些TEE和本领域中已知的那些TEE,所述申请通过引用整体并入本文。

[0471] RNA的体外转录

[0472] 编码本文所述的多核苷酸的cDNA可使用体外转录(IVT)系统转录。RNA的体外转录是本领域中已知的,并且阐述于国际公开WO 2014/152027中,所述公开通过引用整体并入本文。在一些实施方案中,本公开的RNA是根据WO 2018/053209和WO 2019/036682中阐述的任何一种或多种方法制备,所述文献各自通过引用并入本文。

[0473] 在一些实施方案中,所述RNA转录物是使用非扩增的线性化DNA模板在体外转录反应中产生,以产生RNA转录物。在一些实施方案中,所述模板DNA是分离的DNA。在一些实施方案中,所述模板DNA是cDNA。在一些实施方案中,所述cDNA通过mRNA(例如但不限于冠状病毒mRNA)的逆转录形成。在一些实施方案中,用质粒DNA模板转染细胞,例如细菌细胞,例如大肠杆菌,例如DH-1细胞。在一些实施方案中,培养转染的细胞以复制质粒DNA,然后分离和纯化所述质粒DNA。在一些实施方案中,所述DNA模板包括RNA聚合酶启动子,例如位于目标基因5′并与其可操作连接的T7启动子。

[0474] 在一些实施方案中,体外转录模板编码5′非翻译(UTR)区,含有开放阅读框,并编码3′UTR和多聚(A)尾。体外转录模板的特定核酸序列组成和长度将取决于模板编码的mRNA。

[0475] “5′非翻译区”(UTR)是指不编码多肽的mRNA的位于起始密码子(即,核糖体翻译的mRNA转录物的第一个密码子)正上游(即,5′)的区。当产生RNA转录物时,5′UTR可包含启动子序列。此类启动子序列是本领域中已知的。应理解,此类启动子序列将不存在于本公开的疫苗中。

[0476] “3′非翻译区”(UTR)是指不编码多肽的mRNA的位于终止密码子(即,发出翻译终止信号的mRNA转录物的密码子)正下游(即,3′)的区。

[0477] “开放阅读框”是以起始密码子(例如,甲硫氨酸(ATG))开始并以终止密码子(例如,TAA、TAG或TGA)结束的DNA的连续段,并且编码多肽。

[0478] “多聚(A)尾”是mRNA的位于含有多个连续单磷酸腺苷的3′UTR的下游、例如正下游(即,3′)的区。多聚(A)尾可含有10至300个单磷酸腺苷。例如,多聚(A)尾可含有10、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290或300个单磷酸腺苷。在一些实施方案中,多聚(A)尾含有50至250个单磷酸腺苷。在相关生物学环境中(例如,在细胞中,在体内),多聚(A)尾用于保护mRNA免于例如在细胞质中的酶促降解,并帮助转录终止,和/或mRNA从细胞核输出和翻译。

[0479] 在一些实施方案中,核酸包括200至3,000个核苷酸。例如,核酸可包括200至500、200至1000、200至1500、200至3000、500至1000、500至1500、500至2000、500至3000、1000至1500、1000至2000、1000至3000、1500至3000或2000至3000个核苷酸。

[0480] 体外转录系统通常包含转录缓冲液、核苷酸三磷酸(NTP)、RNA酶抑制剂和聚合酶。

[0481] 所述NTP可在内部制造,可选自供应商,或者可如本文所述合成。所述NTP可选自但不限于本文所述的那些NTP,包括天然和非天然(经修饰的)NTP。

[0482] 任何数量的RNA聚合酶或变体可用于本公开的方法中。所述聚合酶可选自但不限于噬菌体RNA聚合酶,例如T7 RNA聚合酶、T3 RNA聚合酶、SP6 RNA聚合酶和/或突变型聚合

酶,例如但不限于能够并入经修饰的核酸和/或经修饰的核苷酸(包括化学修饰的核酸和/或核苷酸)的聚合酶。一些实施方案排除使用DNA酶。

[0483] 在一些实施方案中,RNA转录物通过酶促加帽来加帽。在一些实施方案中,RNA包含5′末端帽,例如7mG(5′)ppp(5′)NlmpNp。

[0484] 化学合成

[0485] 固相化学合成。本公开的核酸可使用固相技术全部或部分地制造。核酸的固相化学合成是一种自动化方法,其中分子固定于固体载体上,并且在反应物溶液中逐步合成。固相合成可用于在核酸序列中位点特异性引入化学修饰。

[0486] 液相化学合成。通过顺序添加单体结构单元来合成本公开的核酸可在液相中实施。

[0487] 合成方法的组合。上文论述的合成方法各自具有其自身的优点和限制。已尝试组合这些方法以克服所述限制。此类方法的组合在本公开的范围内。固相或液相化学合成与酶促连接的组合使用提供了产生不能通过单独化学合成获得的长链核酸的有效方式。

[0488] 核酸区或亚区的连接

[0489] 还可使用通过连接酶装配核酸。DNA或RNA连接酶经由磷酸二酯键的形成促进多核苷酸链的5′端和3′端的分子间连接。核酸(例如嵌合多核苷酸和/或环状核酸)可通过一个或多个区或亚区的连接来制备。DNA片段可通过连接酶催化的反应连接,以产生具有不同功能的重组DNA。两个寡脱氧核苷酸(一个具有5′磷酰基,并且另一个具有游离的3′羟基)用作DNA连接酶的底物。

[0490] 纯化

[0491] 本文所述的核酸的纯化可包括但不限于核酸净化、质量保证和质量控制。净化可通过本领域中已知的方法实施,所述方法例如但不限于AGENCOURT®珠粒(Beckman Coulter Genomics,Danvers,MA)、聚-T珠粒、LNATM寡-T捕获探针(EXIQON®公司,Vedbaek,Denmark)或基于HPLC的纯化方法,例如但不限于强阴离子交换HPLC、弱阴离子交换HPLC、反相HPLC(RP-HPLC)和疏水相互作用HPLC(HIC-HPLC)。术语“纯化的”当关于核酸使用时,例如“纯化的核酸”,是指与至少一种污染物分离的核酸。“污染物”是使另一种物质不适合、不纯或劣质的任何物质。因此,纯化的核酸(例如,DNA和RNA)以不同于其在自然界中发现的形式或环境存在,或以不同于其在进行处理或纯化方法之前存在的形式或环境存在。

[0492] 质量保证和/或质量控制检查可使用例如但不限于凝胶电泳、UV吸光度或分析型HPLC的方法执行。

[0493] 在一些实施方案中,核酸可通过包括但不限于逆转录酶-PCR的方法测序。

[0494] 定量

[0495] 在一些实施方案中,本公开的核酸可在外泌体中或当衍生自一种或多种体液时定量。体液包括外周血、血清、血浆、腹水、尿液、脑脊髓液(CSF)、痰、唾液、骨髓、滑液、房水、羊水、耳垢、母乳、支气管肺泡灌洗液、精液、前列腺液、考珀液(cowper′s fluid)或预射精液、汗液、粪便物、毛发、泪液、囊肿液、胸膜液和腹膜液、心包液、淋巴、食糜、乳糜、胆汁、间质液、月经、脓液、皮脂、呕吐物、阴道分泌物、粘膜分泌物、粪便水、胰液、窦腔灌洗液、支气管肺抽吸物、囊胚腔液和脐带血。或者,外泌体可从选自由以下组成的组的器官中取得:肺、心

脏、胰腺、胃、肠、膀胱、肾、卵巢、睾丸、皮肤、结肠、乳房、前列腺、脑、食管、肝和胎盘。

[0496] 可使用构建体特异性探针、血细胞计数法、qRT-PCR、实时PCR、PCR、流式细胞术、电泳、质谱或其组合来实施测定，而外泌体可使用免疫组织化学方法(例如酶联免疫吸附测定(ELISA))方法来分离。外泌体也可通过尺寸排阻色谱法、密度梯度离心、差速离心、纳米膜超滤、免疫吸附捕获、亲和纯化、微流体分离或其组合来分离。

[0497] 这些方法使得研究人员能够实时监测剩余或递送的核酸水平。这是可能的，因为在一些实施方案中，本公开的核酸由于结构或化学修饰而不同于内源形式。

[0498] 在一些实施方案中，核酸可使用例如但不限于紫外可见光谱(UV/Vis)的方法来定量。UV/Vis光谱仪的非限制性实例是NANODROP®光谱仪(ThermoFisher，Waltham，MA)。可对定量的核酸进行分析，以确定核酸是否可具有适当大小，检查核酸未发生降解。核酸的降解可通过以下方法检查，例如但不限于琼脂糖凝胶电泳；基于HPLC的纯化方法，例如但不限于强阴离子交换HPLC、弱阴离子交换HPLC、反相HPLC(RP-HPLC)和疏水相互作用HPLC(HIC-HPLC)；液相色谱-质谱(LCMS)、毛细管电泳(CE)和毛细管凝胶电泳(CGE)。

[0499] 脂质纳米颗粒(LNP)

[0500] 在一些实施方案中，本公开的mRNA配制在脂质纳米颗粒(LNP)中。脂质纳米颗粒通常包含可电离的阳离子脂质、非阳离子脂质、固醇和PEG脂质组分以及目标核酸货物。本公开的脂质纳米颗粒可使用本领域中通常已知的组分、组合物和方法产生，参见例如PCT/US2016/052352；PCT/US2016/068300；PCT/US2017/037551；PCT/US2015/027400；PCT/US2016/047406；PCT/US2016000129；PCT/US2016/014280；PCT/US2016/014280；PCT/US2017/038426；PCT/US2014/027077；PCT/US2014/055394；PCT/US2016/52117；PCT/US2012/069610；PCT/US2017/027492；PCT/US2016/059575和PCT/US2016/069491，所述专利都通过引用整体并入本文。

[0501] 本公开的疫苗通常配制在脂质纳米颗粒中。在一些实施方案中，所述脂质纳米颗粒包含至少一种可电离的阳离子脂质、至少一种非阳离子脂质、至少一种固醇和/或至少一种聚乙二醇(PEG)修饰的脂质。

[0502] 在一些实施方案中，所述脂质纳米颗粒包含40-50mol％可电离的脂质，任选地45-50mol％，例如45-46mol％、46-47mol％、47-48mol％、48-49mol％或49-50mol％，例如约45mol％、45.5mol％、46mol％、46.5mol％、47mol％、47.5mol％、48mol％、48.5mol％、49mol％或49.5mol％。

[0503] 在一些实施方案中，所述脂质纳米颗粒包含30-45mol％固醇，任选地35-40mol％，例如30-31mol％、31-32mol％、32-33mol％、33-34mol％、35-35mol％、35-36mol％、36-37mol％、38-38mol％、38-39mol％或39-40mol％。

[0504] 在一些实施方案中，所述脂质纳米颗粒包含5-15mol％辅助脂质，任选地10-12mol％，例如5-6mol％、6-7mol％、7-8mol％、8-9mol％、9-10mol％、10-11mol％、11-12mol％、12-13mol％、13-14mol％或14-15mol％。

[0505] 在一些实施方案中，所述脂质纳米颗粒包含1-5％PEG脂质，任选地1-3mol％，例如1.5-2.5mol％、1-2mol％、2-3mol％、3-4mol％或4-5mol％。

[0506] 在一些实施方案中，所述脂质纳米颗粒包含20-60mol％可电离的阳离子脂质。例如，所述脂质纳米颗粒可包含20-50mol％、20-40mol％、20-30mol％、30-60mol％、30-

50mol%、30-40mol%、40-60mol%、40-50mol%或50-60mol%可电离的阳离子脂质。在一些实施方案中，所述脂质纳米颗粒包含20mol%、30mol%、40mol%、50mol%或60mol%可电离的阳离子脂质。在一些实施方案中，所述脂质纳米颗粒包含35mol%、36mol%、37mol%、38mol%、39mol%、40mol%、41mol%、42mol%、43mol%、44mol%、45mol%%、46mol%、47mol%、48mol%、49mol%、50mol%、51mol%、52mol%、53mol%、54mol%或55mol%可电离的阳离子脂质。

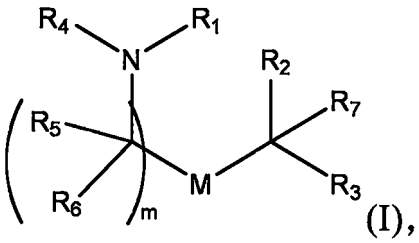
[0507] 在一些实施方案中，所述脂质纳米颗粒包含5-25mol%非阳离子脂质。例如，所述脂质纳米颗粒可包含5-20mol%、5-15mol%、5-10mol%、10-25mol%、10-20mol%、10-25mol%、15-25mol%、15-20mol%或20-25mol%非阳离子脂质。在一些实施方案中，所述脂质纳米颗粒包含5mol%、10mol%、15mol%、20mol%或25mol%非阳离子脂质。

[0508] 在一些实施方案中，所述脂质纳米颗粒包含25-55mol%固醇。例如，所述脂质纳米颗粒可包含25-50mol%、25-45mol%、25-40mol%、25-35mol%、25-30mol%、30-55mol%、30-50mol%、30-45mol%、30-40mol%、30-35mol%、35-55mol%、35-50mol%、35-45mol%、35-40mol%、40-55mol%、40-50mol%、40-45mol%、45-55mol%、45-50mol%或50-55mol%固醇。在一些实施方案中，所述脂质纳米颗粒包含25mol%、30mol%、35mol%、40mol%、45mol%、50mol%或55mol%固醇。

[0509] 在一些实施方案中，所述脂质纳米颗粒包含0.5-15mol%PEG修饰的脂质。例如，所述脂质纳米颗粒可包含0.5-10mol%、0.5-5mol%、1-15mol%、1-10mol%、1-5mol%、2-15mol%、2-10mol%、2-5mol%、5-15mol%、5-10mol%或10-15mol%。在一些实施方案中，所述脂质纳米颗粒包含0.5mol%、1mol%、2mol%、3mol%、4mol%、5mol%、6mol%、7mol%、8mol%、9mol%、10mol%、11mol%、12mol%、13mol%、14mol%或15mol%PEG修饰的脂质。

[0510] 在一些实施方案中，所述脂质纳米颗粒包含20-60mol%可电离的阳离子脂质、5-25mol%非阳离子脂质、25-55mol%固醇和0.5-15mol%PEG修饰的脂质。

[0511] 在一些实施方案中，本公开的可电离的阳离子脂质包含式(I)化合物：

[0512] 

(I)，

[0513] 或其盐或异构体，其中：

[0514] R_1选自由C_{5-30}烷基、C_{5-20}烯基、-R*YR″、-YR″和-R″M′R′组成的组；

[0515] R_2和R_3独立地选自由H、C_{1-14}烷基、C_{2-14}烯基、-R*YR″、-YR″和-R*OR″组成的组，或者R_2和R_3与其所附接的原子一起形成杂环或碳环；

[0516] R_4选自由C_{3-6}碳环、$-(CH_2)_nQ$、$-(CH_2)_nCHQR$、-CHQR、$-CQ(R)_2$和未被取代的C_{1-6}烷基组成的组，其中Q选自碳环、杂环、-OR、$-O(CH_2)_nN(R)_2$、-C(O)OR、-OC(O)R、$-CX_3$、$-CX_2H$、$-CXH_2$、-CN-、$N(R)_2$、$-C(O)N(R)_2$、$-N(R)C(O)R$、$-N(R)S(O)_2R$、$-N(R)C(O)N(R)_2$、$-N(R)C(S)N(R)_2$、$-N(R)R_8$、$-O(CH_2)_nOR$、$-N(R)C(=NR_9)N(R)_2$、$-N(R)C(=CHR_9)N(R)_2$、$-OC(O)N(R)_2$、$-N(R)C(O)OR$、$-N(OR)C(O)R$、$-N(OR)S(O)_2R$、$-N(OR)C(O)OR$、$-N(OR)C(O)N(R)_2$、$-N(OR)C(S)N(R)_2$、$-N(OR)C(=$

$NR_9)N(R)_2$、-N(OR)C(=CHR_9)$N(R)_2$、-C(=NR_9)$N(R)_2$、-C(=NR_9)R、-C(O)N(R)OR和-C(R)$N(R)_2$C(O)OR，并且每个n独立地选自1、2、3、4和5；

[0517] 每个R_5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0518] 每个R_6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0519] M和M′独立地选自-C(O)O-、-OC(O)-、-C(O)N(R′)-、-N(R′)C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR′)O-、-$S(O)_2$-、-S-S-、芳基和杂芳基；

[0520] R_7选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0521] R_8选自由C_{3-6}碳环和杂环组成的组；

[0522] R_9选自由H、CN、NO_2、C_{1-6}烷基、-OR、-$S(O)_2$R、-$S(O)_2N(R)_2$、C_{2-6}烯基、C_{3-6}碳环和杂环组成的组；

[0523] 每个R独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0524] 每个R′独立地选自由C_{1-18}烷基、C_{2-18}烯基、-R*YR″、-YR″和H组成的组；

[0525] 每个R″独立地选自由C_{3-14}烷基和C_{3-14}烯基组成的组；

[0526] 每个R*独立地选自由C_{1-12}烷基和C_{2-12}烯基组成的组；

[0527] 每个Y独立地是C_{3-6}碳环；

[0528] 每个X独立地选自由F、Cl、Br和I组成的组；并且

[0529] m选自5、6、7、8、9、10、11、12和13。

[0530] 在一些实施方案中，式(I)化合物的子集包括以下的那些化合物：其中当R_4是-$(CH_2)_n$Q、-$(CH_2)_n$CHQR、-CHQR或-CQ$(R)_2$时，则(i)当n为1、2、3、4或5时，Q不为-$N(R)_2$，或(ii)当n为1或2时，Q不为5元、6元或7元杂环烷基。

[0531] 在-些实施方案中，式(I)化合物的另一个子集包括以下的那些化合物：其中

[0532] R_1选自由C_{5-30}烷基、C_{5-20}烯基、-R*YR″、-YR″和-R″M′R′组成的组；

[0533] R_2和R_3独立地选自由H、C_{1-14}烷基、C_{2-14}烯基、-R*YR″、-YR″和-R*OR″组成的组，或者R_2和R_3与其所附接的原子一起形成杂环或碳环；

[0534] R_4选自由C_{3-6}碳环、-$(CH_2)_n$Q、-$(CH_2)_n$CHQR、-CHQR、-CQ$(R)_2$和未被取代的C_{1-6}烷基组成的组，其中Q选自C_{3-6}碳环、具有一个或多个选自N、O和S的杂原子的5至14元杂芳基、-OR、-$O(CH_2)_nN(R)_2$、-C(O)OR、-OC(O)R、-CX_3、-CX_2H、-CXH_2、-CN、-C(O)$N(R)_2$、-N(R)C(O)R、-N(R)$S(O)_2$R、-N(R)C(O)$N(R)_2$、-N(R)C(S)$N(R)_2$、-CRN$(R)_2$C(O)OR、-N(R)R_8、-$O(CH_2)_n$OR、-N(R)C(=NR_9)$N(R)_2$、-N(R)C(=CHR_9)$N(R)_2$、-OC(O)$N(R)_2$、-N(R)C(O)OR、-N(OR)C(O)R、-N(OR)$S(O)_2$R、-N(OR)C(O)OR、-N(OR)C(O)$N(R)_2$、-N(OR)C(S)$N(R)_2$、-N(OR)C(=NR_9)$N(R)_2$、-N(OR)C(=CHR_9)$N(R)_2$、-C(=NR_9)$N(R)_2$、-C(=NR_9)R、-C(O)N(R)OR和具有一个或多个选自N、O和S的杂原子的5至14元杂环烷基，其被一个或多个选自以下的取代基取代：氧代基(=O)、OH、氨基、单烷基氨基或二烷基氨基和C_{1-3}烷基，并且每个n独立地选自1、2、3、4和5；

[0535] 每个R_5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0536] 每个R_6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0537] M和M′独立地选自-C(O)O-、-OC(O)-、-C(O)N(R′)-、-N(R′)C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR′)O-、-$S(O)_2$-、-S-S-、芳基和杂芳基；

[0538] R_7选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0539] R_8选自由C_{3-6}碳环和杂环组成的组；

[0540] R_9选自由H、CN、NO_2、C_{1-6}烷基、-OR、-S(O)$_2$R、-S(O)$_2$N(R)$_2$、C_{2-6}烯基、C_{3-6}碳环和杂环组成的组；

[0541] 每个R独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0542] 每个R′独立地选自由C_{1-18}烷基、C_{2-18}烯基、-R*YR″、-YR″和H组成的组；

[0543] 每个R″独立地选自由C_{3-14}烷基和C_{3-14}烯基组成的组；

[0544] 每个R*独立地选自由C_{1-12}烷基和C_{2-12}烯基组成的组；

[0545] 每个Y独立地是C_{3-6}碳环；

[0546] 每个X独立地选自由F、Cl、Br和I组成的组；并且

[0547] m选自5、6、7、8、9、10、11、12和13，

[0548] 或其盐或异构体。

[0549] 在一些实施方案中，式(I)化合物的另一个子集包括以下的那些化合物：其中

[0550] R_1选自由C_{5-30}烷基、C_{5-20}烯基、-R*YR″、-YR″和-R″M′R′组成的组；

[0551] R_2和R_3独立地选自由H、C_{1-14}烷基、C_{2-14}烯基、-R*YR″、-YR″和-R*OR″组成的组，或者R_2和R_3与其所附接的原子一起形成杂环或碳环；

[0552] R_4选自由C_{3-6}碳环、-(CH$_2$)$_n$Q、-(CH$_2$)$_n$CHQR、-CHQR、-CQ(R)$_2$和未被取代的C_{1-6}烷基组成的组，其中Q选自C_{3-6}碳环、具有一个或多个选自N、O和S的杂原子的5至14元杂环、-OR、-O(CH$_2$)$_n$N(R)$_2$、-C(O)OR、-OC(O)R、-CX_3、-CX_2H、-CXH_2、-CN、-C(O)N(R)$_2$、-N(R)C(O)R、-N(R)S(O)$_2$R、-N(R)C(O)N(R)$_2$、-N(R)C(S)N(R)$_2$、-CRN(R)$_2$C(O)OR、-N(R)R_8、-O(CH$_2$)$_n$OR、-N(R)C(=NR_9)N(R)$_2$、-N(R)C(=CHR_9)N(R)$_2$、-OC(O)N(R)$_2$、-N(R)C(O)OR、-N(OR)C(O)R、-N(OR)S(O)$_2$R、-N(OR)C(O)OR、-N(OR)C(O)N(R)$_2$、-N(OR)C(S)N(R)$_2$、-N(OR)C(=NR_9)N(R)$_2$、-N(OR)C(=CHR_9)N(R)$_2$、-C(=NR_9)R、-C(O)N(R)OR和-C(=NR_9)N(R)$_2$，并且每个n独立地选自1、2、3、4和5；并且当Q是5至14元杂环并且(i)R_4是-(CH$_2$)$_n$Q，其中n为1或2，或(ii)R_4是-(CH$_2$)$_n$CHQR，其中n为1，或(iii)R_4是-CHQR和-CQ(R)$_2$时，则Q是5至14元杂芳基或8至14元杂环烷基；

[0553] 每个R_5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0554] 每个R_6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0555] M和M′独立地选自-C(O)O-、-OC(O)-、-C(O)N(R′)-、-N(R′)C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR′)O-、-S(O)$_2$-、-S-S-、芳基和杂芳基；

[0556] R_7选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0557] R_8选自由C_{3-6}碳环和杂环组成的组；

[0558] R_9选自由H、CN、NO_2、C_{1-6}烷基、-OR、-S(O)$_2$R、-S(O)$_2$N(R)$_2$、C_{2-6}烯基、C_{3-6}碳环和杂环组成的组；

[0559] 每个R独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0560] 每个R′独立地选自由C_{1-18}烷基、C_{2-18}烯基、-R*YR″、-YR″和H组成的组；

[0561] 每个R″独立地选自由C_{3-14}烷基和C_{3-14}烯基组成的组；

[0562] 每个R*独立地选自由C_{1-12}烷基和C_{2-12}烯基组成的组；

[0563] 每个Y独立地是C_{3-6}碳环；

[0564] 每个X独立地选自由F、Cl、Br和I组成的组；并且

[0565] m选自5、6、7、8、9、10、11、12和13，

[0566] 或其盐或异构体。

[0567] 在一些实施方案中，式(I)化合物的另一个子集包括以下的那些化合物：其中

[0568] R_1选自由C_{5-30}烷基、C_{5-20}烯基、-R*YR″、-YR″和-R″M′R′组成的组；

[0569] R_2和R_3独立地选自由H、C_{1-14}烷基、C_{2-14}烯基、-R*YR″、-YR″和-R*OR″组成的组，或者R_2和R_3与其所附接的原子一起形成杂环或碳环；

[0570] R_4选自由C_{3-6}碳环、$-(CH_2)_nQ$、$-(CH_2)_nCHQR$、-CHQR、$-CQ(R)_2$和未被取代的C_{1-6}烷基组成的组，其中Q选自C_{3-6}碳环、具有一个或多个选自N、O和S的杂原子的5至14元杂芳基、-OR、$-O(CH_2)_nN(R)_2$、$-C(O)OR$、$-OC(O)R$、$-CX_3$、$-CX_2H$、$-CXH_2$、-CN、$-C(O)N(R)_2$、$-N(R)C(O)R$、$-N(R)S(O)_2R$、$-N(R)C(O)N(R)_2$、$-N(R)C(S)N(R)_2$、$-CRN(R)_2C(O)OR$、$-N(R)R_8$、$-O(CH_2)_nOR$、$-N(R)C(=NR_9)N(R)_2$、$-N(R)C(=CHR_9)N(R)_2$、$-OC(O)N(R)_2$、$-N(R)C(O)OR$、$-N(OR)C(O)R$、$-N(OR)S(O)_2R$、$-N(OR)C(O)OR$、$-N(OR)C(O)N(R)_2$、$-N(OR)C(S)N(R)_2$、$-N(OR)C(=NR_9)N(R)_2$、$-N(OR)C(=CHR_9)N(R)_2$、$-C(=NR_9)R$、$-C(O)N(R)OR$和$-C(=NR_9)N(R)_2$，并且每个n独立地选自1、2、3、4和5；

[0571] 每个R_5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0572] 每个R_6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0573] M和M′独立地选自$-C(O)O-$、$-OC(O)-$、$-C(O)N(R′)-$、$-N(R′)C(O)-$、$-C(O)-$、$-C(S)-$、$-C(S)S-$、$-SC(S)-$、$-CH(OH)-$、$-P(O)(OR′)O-$、$-S(O)_2-$、-S-S-、芳基和杂芳基；

[0574] R_7选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0575] R_8选自由C_{3-6}碳环和杂环组成的组；

[0576] R_9选自由H、CN、NO_2、C_{1-6}烷基、-OR、$-S(O)_2R$、$-S(O)_2N(R)_2$、C_{2-6}烯基、C_{3-6}碳环和杂环组成的组；

[0577] 每个R独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0578] 每个R′独立地选自由C_{1-18}烷基、C_{2-18}烯基、-R*YR″、-YR″和H组成的组；

[0579] 每个R″独立地选自由C_{3-14}烷基和C_{3-14}烯基组成的组；

[0580] 每个R*独立地选自由C_{1-12}烷基和C_{2-12}烯基组成的组；

[0581] 每个Y独立地是C_{3-6}碳环；

[0582] 每个X独立地选自由F、Cl、Br和I组成的组；并且

[0583] m选自5、6、7、8、9、10、11、12和13，

[0584] 或其盐或异构体。

[0585] 在一些实施方案中，式(I)化合物的另一个子集包括以下的那些化合物：其中

[0586] R_1选自由C_{5-30}烷基、C_{5-20}烯基、-R*YR″、-YR″和-R″M′R′组成的组；

[0587] R_2和R_3独立地选自由H、C_{2-14}烷基、C_{2-14}烯基、-R*YR″、-YR″和-R*OR″组成的组，或者R_2和R_3与其所附接的原子一起形成杂环或碳环；

[0588] R_4是$-(CH_2)_nQ$或$-(CH_2)_nCHQR$，其中Q是$-N(R)_2$，并且n选自3、4和5；

[0589] 每个R_5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0590] 每个R_6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0591] M和M′独立地选自$-C(O)O-$、$-OC(O)-$、$-C(O)N(R′)-$、$-N(R′)C(O)-$、$-C(O)-$、$-C(S)-$、$-C(S)S-$、$-SC(S)-$、$-CH(OH)-$、$-P(O)(OR′)O-$、$-S(O)_2-$、-S-S-、芳基和杂芳基；

[0592] R_7选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0593] 每个R独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0594] 每个R′独立地选自由C_{1-18}烷基、C_{2-18}烯基、-R*YR″、-YR″和H组成的组；

[0595] 每个R″独立地选自由C_{3-14}烷基和C_{3-14}烯基组成的组；

[0596] 每个R*独立地选自由C_{1-12}烷基和C_{1-12}烯基组成的组；

[0597] 每个Y独立地是C_{3-6}碳环；

[0598] 每个X独立地选自由F、Cl、Br和I组成的组；并且

[0599] m选自5、6、7、8、9、10、11、12和13，

[0600] 或其盐或异构体。

[0601] 在一些实施方案中，式(I)化合物的另一个子集包括以下的那些化合物：其中

[0602] R_1选自由C_{5-30}烷基、C_{5-20}烯基、-R*YR″、-YR″和-R″M′R′组成的组；

[0603] R_2和R_3独立地选自由C_{1-14}烷基、C_{2-14}烯基、-R*YR″、-YR″和-R*OR″组成的组，或者R_2和R_3与其所附接的原子一起形成杂环或碳环；

[0604] R_4选自由-$(CH_2)_nQ$、-$(CH_2)_nCHQR$、-CHQR和-$CQ(R)_2$组成的组，其中Q是-$N(R)_2$，并且n选自1、2、3、4和5；

[0605] 每个R_5独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0606] 每个R_6独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0607] M和M′独立地选自-C(O)O-、-OC(O)-、-C(O)N(R′)-、-N(R′)C(O)-、-C(O)-、-C(S)-、-C(S)S-、-SC(S)-、-CH(OH)-、-P(O)(OR′)O-、-$S(O)_2$-、-S-S-、芳基和杂芳基；

[0608] R_7选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0609] 每个R独立地选自由C_{1-3}烷基、C_{2-3}烯基和H组成的组；

[0610] 每个R′独立地选自由C_{1-18}烷基、C_{2-18}烯基、-R*YR″、-YR″和H组成的组；

[0611] 每个R″独立地选自由C_{3-14}烷基和C_{3-14}烯基组成的组；

[0612] 每个R*独立地选自由C_{1-12}烷基和C_{1-12}烯基组成的组；

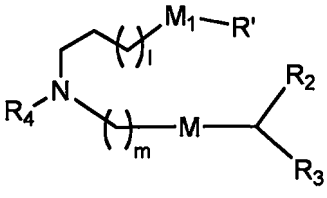
[0613] 每个Y独立地是C_{3-6}碳环；

[0614] 每个X独立地选自由F、Cl、Br和I组成的组；并且

[0615] m选自5、6、7、8、9、10、11、12和13，

[0616] 或其盐或异构体。

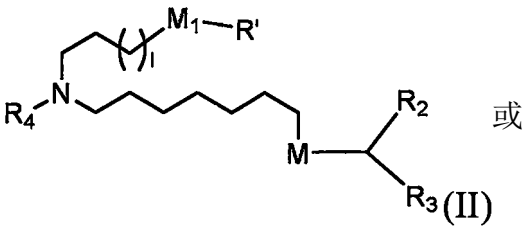
[0617] 在一些实施方案中，式(I)化合物的子集包括式(IA)化合物：

[0618] 

(IA)，

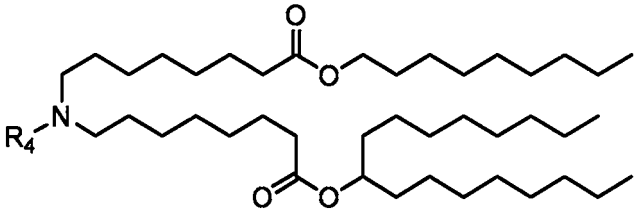
[0619] 或其盐或异构体，其中l选自1、2、3、4和5；m选自5、6、7、8和9；M_1是键或M′；R_4是未被取代的C_{1-3}烷基，或-$(CH_2)_nQ$，其中Q是OH、-NHC(S)$N(R)_2$、-NHC(O)$N(R)_2$、-N(R)C(O)R、-N(R)$S(O)_2R$、-N(R)R_8、-NHC(=NR_9)$N(R)_2$、-NHC(=CHR_9)$N(R)_2$、-OC(O)$N(R)_2$、-N(R)C(O)OR、杂芳基或杂环烷基；M和M′独立地选自-C(O)O-、-OC(O)-、-C(O)N(R′)-、-P(O)(OR′)O-、-S-S-、芳基和杂芳基；并且R_2和R_3独立地选自由H、C_{1-14}烷基和C_{2-14}烯基组成的组。

[0620] 在一些实施方案中，式(I)化合物的子集包括式(II)化合物：

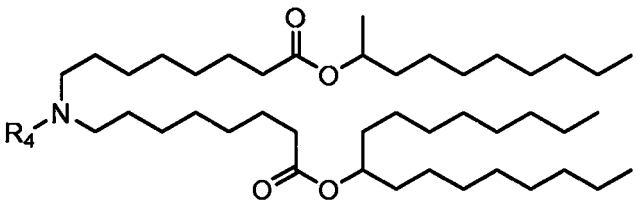
[0621]  (II)

或其盐或异构体，其中l选自1、2、3、4和5；M_1是键或M′；R_4是未被取代的C_{1-3}烷基，或$-(CH_2)_nQ$，其中n为2、3或4，并且Q为OH、-NHC(S)N(R)$_2$、-NHC(O)N(R)$_2$、-N(R)C(O)R、-N(R)S(O)$_2$R、-N(R)R_8、-NHC(=NR_9)N(R)$_2$、-NHC(=CHR_9)N(R)$_2$、-OC(O)N(R)$_2$、-N(R)C(O)OR、杂芳基或杂环烷基；M和M′独立地选自-C(O)O-、-OC(O)-、-C(O)N(R′)-、-P(O)(OR′)O-、-S-S-、芳基和杂芳基；并且R_2和R_3独立地选自由H、C_{1-14}烷基和C_{2-14}烯基组成的组。

[0622] 在一些实施方案中，式(I)化合物的子集包括式(IIa)、(IIb)、(IIc)或(IIe)的化合物：

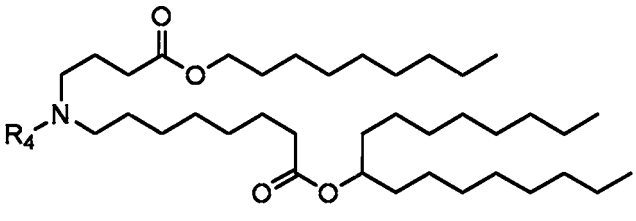


(IIa),

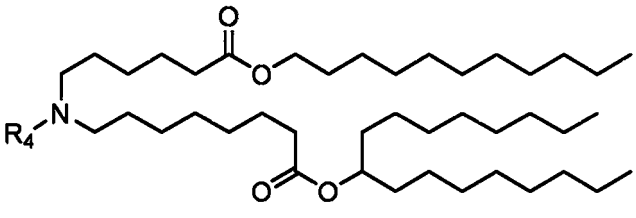


(IIb),

[0623]



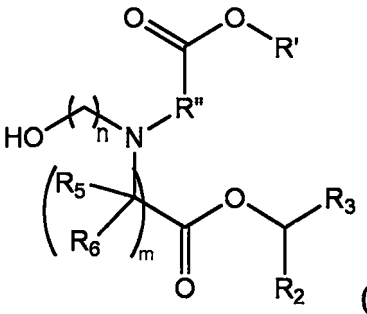
(IIc)，或



(IIe),

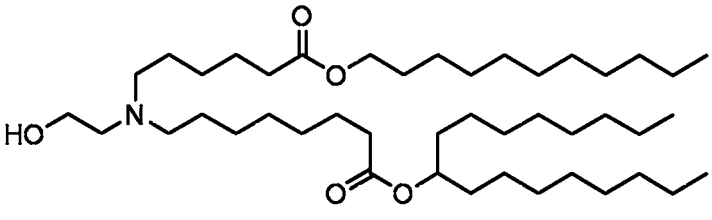
[0624] 或其盐或异构体，其中R_4如本文所述。

[0625] 在一些实施方案中，式(I)化合物的子集包括式(IId)化合物：

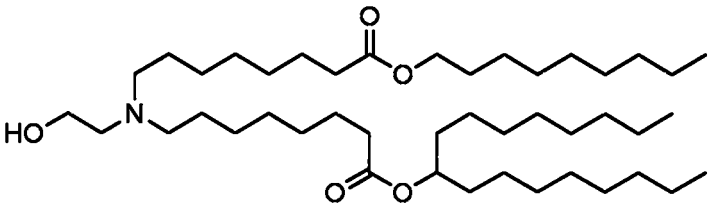
[0626] 
(IId),

[0627] 或其盐或异构体,其中n为2、3或4;并且m、R′、R″和R_2至R_6如本文所述。例如,R_2和R_3可各自独立地选自由C_{5-14}烷基和C_{5-14}烯基组成的组。

[0628] 在一些实施方案中,本公开的可电离的阳离子脂质包含具有以下结构的化合物:

[0629] 
(化合物 I)。

[0630] 在一些实施方案中,本公开的可电离的阳离子脂质包含具有以下结构的化合物:

[0631] 
(化合物 II)。

[0632] 在一些实施方案中,本公开的非阳离子脂质包含1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)、1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)、1,2-二亚油酰基-sn-甘油-3-磷酸胆碱(DLPC)、1,2-二肉豆蔻酰基-sn-甘油-磷酸胆碱(DMPC)、1,2-二油酰基-sn-甘油-3-磷酸胆碱(DOPC)、1,2-二棕榈酰基-sn-甘油-3-磷酸胆碱(DPPC)、1,2-二-十一烷酰基-sn-甘油-磷酸胆碱(DUPC)、1-棕榈酰基-2-油酰基-sn-甘油-3-磷酸胆碱(POPC)、1,2-二-O-十八碳烯基-sn-甘油-3-磷酸胆碱(18:0二醚PC)、1-油酰基-2胆固醇基半琥珀酰基-sn-甘油-3-磷酸胆碱(OChemsPC)、1-十六烷基-sn-甘油-3-磷酸胆碱(C16 Lyso PC)、1,2-二亚麻酰基-sn-甘油-3-磷酸胆碱、1,2-二花生四烯酰基-sn-甘油-3-磷酸胆碱、1,2-二-二十二碳六烯酰基-sn-甘油-3-磷酸胆碱、1,2-二植烷酰基-sn-甘油-3-磷酸乙醇胺(ME 16.0 PE)、1,2-二硬脂酰基-sn-甘油-3-磷酸乙醇胺、1,2-二亚油酰基-sn-甘油-3-磷酸乙醇胺、1,2-二亚麻酰基-sn-甘油-3-磷酸乙醇胺、1,2-二花生四烯酰基-sn-甘油-3-磷酸乙醇胺、1,2-二-二十二碳六烯酰基-sn-甘油-3-磷酸乙醇胺、1,2-二油酰基-sn-甘油-3-磷酸-外消旋-(1-甘油)钠盐(DOPG)、鞘磷脂和其混合物。

[0633] 在一些实施方案中,本公开的PEG修饰的脂质包含PEG修饰的磷脂酰乙醇胺、PEG修饰的磷脂酸、PEG修饰的神经酰胺、PEG修饰的二烷基胺、PEG修饰的二酰基甘油、PEG修饰的二烷基甘油和其混合物。在一些实施方案中,PEG修饰的脂质是DMG-PEG、PEG-c-DOMG(也称为PEG-DOMG)、PEG-DSG和/或PEG-DPG。

[0634] 在一些实施方案中,本公开的固醇包含胆固醇、粪生固醇、谷固醇、麦角固醇、菜油固醇、豆固醇、芸苔固醇、番茄碱、熊果酸、α-生育酚和其混合物。

[0635] 在一些实施方案中，本公开的LNP包含化合物1的可电离的阳离子脂质，其中非阳离子脂质是DSPC，结构脂质是胆固醇，并且PEG脂质是DMG-PEG。

[0636] 在一些实施方案中，所述脂质纳米颗粒包含45-55摩尔百分比(mol%)的可电离的阳离子脂质。例如，脂质纳米颗粒可包含45mol%、46mol%、47mol%、48mol%、49mol%、50mol%、51mol%、52mol%、53mol%、54mol%或55mol%可电离的阳离子脂质。

[0637] 在一些实施方案中，所述脂质纳米颗粒包含5-15mol%、5-10mol%或10-15mol%DSPC。例如，所述脂质纳米颗粒可包含5mol%、6mol%、7mol%、8mol%、9mol%、10mol%、11mol%、12mol%、13mol%、14mol%或15mol%DSPC。

[0638] 在一些实施方案中，所述脂质纳米颗粒包含35-40mol%胆固醇。例如，所述脂质纳米颗粒可包含35mol%、35.5mol%、36mol%、36.5mol%、37mol%、37.5mol%、38mol%、38.5mol%、39mol%、39.5mol%或40mol%胆固醇。

[0639] 在一些实施方案中，所述脂质纳米颗粒包含1-2mol%、1-3mol%、1-4mol%或1-5mol%DMG-PEG。例如，所述脂质纳米颗粒可包含1mol%、1.5mol%、2mol%、2.5mol%、3mol%或3.5mol%DMG-PEG。

[0640] 在一些实施方案中，所述脂质纳米颗粒包含50mol%可电离的阳离子脂质、10mol%DSPC、38.5mol%胆固醇和1.5mol%DMG-PEG。

[0641] 在一些实施方案中，所述脂质纳米颗粒包含49mol%可电离的阳离子脂质、10mol%DSPC、38.5mol%胆固醇和2.5mol%DMG-PEG。

[0642] 在一些实施方案中，所述脂质纳米颗粒包含49mol%可电离的阳离子脂质、11mol%DSPC、38.5mol%胆固醇和1.5mol%DMG-PEG。

[0643] 在一些实施方案中，所述脂质纳米颗粒包含48mol%可电离的阳离子脂质、11mol%DSPC、38.5mol%胆固醇和2.5mol%DMG-PEG。

[0644] 在一些实施方案中，本公开的LNP包含约2:1至约30:1的N:P比。

[0645] 在一些实施方案中，本公开的LNP包含约6:1的N:P比。

[0646] 在一些实施方案中，本公开的LNP包含约3:1的N:P比。

[0647] 在一些实施方案中，本公开的LNP包含约10:1至约100:1的可电离的阳离子脂质组分对RNA的重量/重量比。

[0648] 在一些实施方案中，本公开的LNP包含约20:1的可电离的阳离子脂质组分对RNA的重量/重量比。

[0649] 在一些实施方案中，本公开的LNP包含约10:1的可电离的阳离子脂质组分对RNA的重量/重量比。

[0650] 在一些实施方案中，本公开的LNP具有约50nm至约150nm的平均直径。

[0651] 在一些实施方案中，本公开的LNP具有约70nm至约120nm的平均直径。

[0652] 多价疫苗

[0653] 如本文提供的组合物可包括编码相同或不同种类的两种或更多种抗原的RNA或多种RNA。在一些实施方案中，组合物包括编码两种或更多种冠状病毒抗原的mRNA或多种mRNA。在一些实施方案中，所述RNA可编码1种、2种、3种、4种、5种、6种、7种、8种、9种、10种、11种、12种或更多种冠状病毒抗原。

[0654] 在一些实施方案中，可将两种或更多种不同的编码抗原的mRNA配制在相同脂质纳

米颗粒中。在其他实施方案中，可将两种或更多种不同的编码抗原的RNA配制在单独脂质纳米颗粒中(每个RNA配制在单一脂质纳米颗粒中)。然后，可将脂质纳米颗粒组合并作为单一疫苗组合物(例如，包含编码多种抗原的多种RNA)施用，或者可单独施用。

[0655] 组合疫苗

[0656] 如本文提供的组合物可包括编码相同或不同病毒株的两种或更多种抗原的mRNA或多种RNA。本文还提供了包含编码一种或多种冠状病毒和不同生物体的一种或多种抗原的RNA的组合疫苗。因此，本公开的疫苗可为靶向相同株/种类的一种或多种抗原或不同株/种类的一种或多种抗原的组合疫苗，例如诱导对在冠状病毒感染风险高的相同地理区域发现的生物体或个体在暴露于冠状病毒时有可能暴露于其的生物体的免疫性的抗原。

[0657] 药物制剂

[0658] 本文提供了用于预防或治疗例如人类和其他哺乳动物的冠状病毒的组合物(例如药物组合物)、方法、试剂盒和试剂。本文提供的组合物可用作治疗剂或预防剂。其可用于药物中以预防和/或治疗冠状病毒感染。

[0659] 在一些实施方案中，可将含有如本文所述的RNA的冠状病毒疫苗施用给受试者(例如哺乳动物受试者，例如人类受试者)，并且mRNA在体内翻译以产生抗原性多肽(抗原)。

[0660] 组合物(例如，包含RNA)的“有效量”至少部分地基于靶组织、靶细胞类型、施用方式、RNA的物理特征(例如，长度、核苷酸组成和/或修饰的核苷的程度)、疫苗的其他组分和其他决定因素，例如受试者的年龄、体重、身高、性别和一般健康状况。通常，有效量的组合物提供诱导或加强的免疫反应，其随受试者细胞中的抗原产生而变化。在一些实施方案中，有效量的含有具有至少一个化学修饰的mRNA的组合物比含有编码相同抗原或肽抗原的相应未经修饰的多核苷酸的组合物更有效。抗原产生增加可通过细胞转染(用RNA疫苗转染的细胞的百分数)增加、从多核苷酸的蛋白质翻译和/或表达增加、核酸降解减少(例如，如通过从经修饰的多核苷酸的蛋白质翻译的持续时间增加)或宿主细胞的抗原特异性免疫反应的改变来证明。

[0661] 术语“药物组合物”是指活性剂与惰性或活性载体的组合，使得组合物尤其适于体内或离体诊断或治疗用途。“药学上可接受的载体”在施用给受试者后或在施用给受试者时不会引起不期望的生理效应。药物组合物中的载体在其与活性成分相容并且能够稳定活性成分的意义上必须也是“可接受的”。一种或多种增溶剂可用作递送活性剂的药物载体。药学上可接受的载体的实例包括(但不限于)生物相容性媒介物、佐剂、添加剂和稀释剂，以实现可用作剂型的组合物。其他载体的实例包括胶态氧化硅、硬脂酸镁、纤维素和月桂基硫酸钠。额外适宜药物载体和稀释剂以及使用其的药物必需品阐述于Remington's Pharmaceutical Sciences中。

[0662] 在一些实施方案中，根据本公开的组合物(包含多核苷酸和其编码的多肽)可用于治疗或预防冠状病毒感染。组合物可作为主动免疫方案的一部分预防性或治疗性地施用给健康个体或在潜伏期(incubation phase)期间或在症状发作后的活动性感染期间的感染早期施用。在一些实施方案中，提供给细胞、组织或受试者的RNA的量可为有效用于免疫预防的量。

[0663] 组合物可与其他预防性或治疗性化合物一起施用。作为一个非限制性实例，预防性或治疗性化合物可为佐剂或加强剂。如本文所用，当提及预防性组合物(例如疫苗)时，术

语“加强剂”是指额外施用预防性(疫苗)组合物。加强剂(或加强疫苗)可在较早施用预防性组合物之后给予。初始施用预防性组合物与加强剂之间的施用时间可为(但不限于)1分钟、2分钟、3分钟、4分钟、5分钟、6分钟、7分钟、8分钟、9分钟、10分钟、15分钟、20分钟、35分钟、40分钟、45分钟、50分钟、55分钟、1小时、2小时、3小时、4小时、5小时、6小时、7小时、8小时、9小时、10小时、11小时、12小时、13小时、14小时、15小时、16小时、17小时、18小时、19小时、20小时、21小时、22小时、23小时、1天、36小时、2天、3天、4天、5天、6天、1周、10天、2周、3周、1个月、2个月、3个月、4个月、5个月、6个月、7个月、8个月、9个月、10个月、11个月、1年、18个月、2年、3年、4年、5年、6年、7年、8年、9年、10年、11年、12年、13年、14年、15年、16年、17年、18年、19年、20年、25年、30年、35年、40年、45年、50年、55年、60年、65年、70年、75年、80年、85年、90年、95年或超过99年。在示例性实施方案中,初始施用预防性组合物与加强剂之间的施用时间可为(但不限于)1周、2周、3周、1个月、2个月、3个月、6个月或1年。

[0664] 在一些实施方案中,组合物可经肌内、鼻内或真皮内施用,类似于本领域中已知的灭活疫苗的施用。

[0665] 根据感染的盛行率或未满足的医疗需要的程度或水平,组合物可用于各种环境。作为一个非限制性实例,RNA疫苗可用于治疗和/或预防多种传染病。RNA疫苗具有优越的性质,这是因为它们比市售疫苗产生大得多的抗体滴度、更好的中和免疫性、产生更耐久的免疫反应和/或更早产生反应。

[0666] 本文提供了药物组合物,其包括RNA和/或复合体,任选地与一种或多种药学上可接受的赋形剂组合。

[0667] 所述RNA可单独配制或施用,或与一种或多种其他组分联合配制或施用。例如,组合物可包含其他组分,包括但不限于佐剂。

[0668] 在一些实施方案中,组合物不包含佐剂(其不含佐剂)。

[0669] RNA可与一种或多种药学上可接受的赋形剂组合配制或施用。在一些实施方案中,疫苗组合物包含至少一种额外活性物质,例如治疗活性物质、预防活性物质或两者的组合。疫苗组合物可为无菌的、无热原的,或无菌和无热原的。在医药剂(例如疫苗组合物)的配制和/或制造中的一般考虑因素可见于例如Remington:The Science and Practice of Pharmacy,第21版,Lippincott Williams&Wilkins,2005(其通过引用整体并入本文)。

[0670] 在一些实施方案中,将组合物施用给人类、人类患者或受试者。出于本公开的目的,短语“活性成分”通常是指RNA疫苗或其中所含的多核苷酸,例如编码抗原的mRNA。

[0671] 本文所述的疫苗组合物的制剂可通过药理学领域中已知或以后开发的任何方法来制备。一般来说,此类制备方法包括使活性成分(例如mRNA)与赋形剂和/或一种或多种其他辅助成分缔合并且然后在必要时和/或需要时将产品分割、成型和/或包装成期望的单剂量或多剂量单元的步骤。

[0672] 根据本公开的药物组合物中的活性成分、药学上可接受的赋形剂和/或任何额外成分的相对量将根据所治疗的受试者的身份、体型和/或状况并且进一步根据组合物待施用的途径而变化。举例来说,所述组合物可包含介于0.1%与100%之间、例如介于0.5%与50%之间、介于1%与30%之间、介于5%与80%之间、至少80%(w/w)的活性成分。

[0673] 在一些实施方案中,使用一种或多种赋形剂配制mRNA以:(1)增加稳定性;(2)增加细胞转染;(3)允许持续或延迟释放(例如,来自储积制剂);(4)改变生物分布(例如,靶向特

定组织或细胞类型)；(5)增加编码的蛋白质在体内的翻译；和/或(6)改变编码的蛋白质(抗原)在体内的释放曲线。除了传统赋形剂(例如任何和所有溶剂、分散介质、稀释剂或其他液体媒介物、分散或悬浮助剂、表面活性剂、等渗剂、增稠剂或乳化剂、防腐剂)之外，赋形剂还可包括(但不限于)类脂质、脂质体、脂质纳米颗粒、聚合物、脂质复合体、核-壳纳米颗粒、肽、蛋白质、用RNA转染的细胞(例如，用于移植至受试者中)、玻尿酸酶、纳米颗粒模拟物和其组合。

[0674] 给药/施用

[0675] 本文提供了用于预防和/或治疗人类和其他哺乳动物的冠状病毒感染的组合物(例如，RNA疫苗)、方法、试剂盒和试剂。免疫组合物可用作治疗剂或预防剂。在一些实施方案中，组合物用于提供对冠状病毒感染的预防性保护。在一些实施方案中，组合物用于治疗冠状病毒感染。在一些实施方案中，组合物用于引发免疫效应细胞，例如离体激活外周血单核细胞(PBMC)，然后将其输注(再输注)至受试者中。

[0676] 受试者可为任何哺乳动物，包括非人类灵长类动物和人类受试者。通常，受试者是人类受试者。

[0677] 在一些实施方案中，将组合物(例如，RNA疫苗)以有效诱导抗原特异性免疫反应的量施用给受试者(例如，哺乳动物受试者，例如人类受试者)。编码冠状病毒抗原的RNA在体内表达和翻译以产生抗原，其然后刺激受试者的免疫反应。

[0678] 在施用本公开的组合物之后，可实现对冠状病毒的预防性保护。免疫组合物可施用一次、两次、三次、四次或更多次，但有可能施用一次疫苗即为足够的(任选地接着进行单次加强)。尽管不太期望，但可将组合物施用给受感染个体以实现治疗反应。可能需要相应地调整给药。

[0679] 在本公开的方面中，提供了在受试者中引发针对冠状病毒抗原(或多种抗原)的免疫反应的方法。在一些实施方案中，方法包括向受试者施用包含具有编码冠状病毒抗原的开放阅读框的mRNA的组合物，从而在受试者中诱导对冠状病毒抗原特异性的免疫反应，其中相对于用预防有效剂量的针对所述抗原的传统疫苗接种的受试者中的抗抗原抗体滴度，疫苗接种后受试者中的抗抗原抗体滴度增加。“抗抗原抗体”是特异性结合至抗原的血清抗体。

[0680] 预防有效剂量是在临床上可接受的水平上预防病毒感染的有效剂量。在一些实施方案中，有效剂量是疫苗的包装插页中列出的剂量。如本文所用，传统疫苗是指除本公开的mRNA疫苗之外的疫苗。例如，传统疫苗包括(但不限于)活微生物疫苗、杀灭的微生物疫苗、亚单位疫苗、蛋白质抗原疫苗、DNA疫苗、病毒样颗粒(VLP)疫苗等。在示例性实施方案中，传统疫苗是已经获得监管部门批准和/或在国家药物监管机构(例如美国食品药品管理局(FDA)或欧洲药物管理局(EMA))注册的疫苗。

[0681] 在一些实施方案中，相对于用预防有效剂量的针对冠状病毒的传统疫苗接种的受试者或未接种疫苗的受试者中的抗抗原抗体滴度，疫苗接种后的受试者中的抗抗原抗体滴度增加1log至10log。在一些实施方案中，相对于用预防有效剂量的针对冠状病毒的传统疫苗接种的受试者或未接种疫苗的受试者中的抗抗原抗体滴度，疫苗接种后的受试者中的抗抗原抗体滴度增加1log、2log、3log、4log、5log或10log。

[0682] 在本公开的其他方面，提供了在受试者中引发针对冠状病毒的免疫反应的方法。

所述方法包括向受试者施用包含含有编码冠状病毒抗原的开放阅读框的mRNA的组合物，从而在受试者中诱导对冠状病毒特异性的免疫反应，其中受试者中的免疫反应等同于以相对于组合物2倍至100倍剂量水平用针对冠状病毒的传统疫苗接种的受试者中的免疫反应。

[0683] 在一些实施方案中，受试者中的免疫反应等同于以相对于本公开的组合物两倍的剂量水平用传统疫苗接种的受试者中的免疫反应。在一些实施方案中，受试者中的免疫反应等同于以相对于本公开的组合物三倍的剂量水平用传统疫苗接种的受试者中的免疫反应。在一些实施方案中，受试者中的免疫反应等同于以相对于本公开的组合物4倍、5倍、10倍、50倍或100倍的剂量水平用传统疫苗接种的受试者中的免疫反应。在一些实施方案中，受试者中的免疫反应等同于以相对于本公开的组合物10倍至1000倍的剂量水平用传统疫苗接种的受试者中的免疫反应。在一些实施方案中，受试者中的免疫反应等同于以相对于本公开的组合物100倍至1000倍的剂量水平用传统疫苗接种的受试者中的免疫反应。

[0684] 在其他实施方案中，通过测定受试者中的[蛋白]抗体滴度来评价免疫反应。在其他实施方案中，测试来自经免疫受试者的血清或抗体中和病毒摄取或减少人类B淋巴细胞的冠状病毒转化的能力。在其他实施方案中，使用本领域中公认的技术来测量促进稳固T细胞反应的能力。

[0685] 本公开的其他方面提供了通过向受试者施用包含具有编码冠状病毒抗原的开放阅读框的mRNA的组合物，从而在受试者中诱导对冠状病毒抗原特异性的免疫反应，而在受试者中引发针对冠状病毒的免疫反应的方法，其中相对于在用预防有效剂量的针对冠状病毒的传统疫苗接种的受试者中诱导的免疫反应，所述受试者中诱导的免疫反应早2天至10周。在一些实施方案中，在以相对于本公开的组合物2倍至100倍的剂量水平用预防有效剂量的传统疫苗接种的受试者中，诱导所述受试者中的免疫反应。

[0686] 在一些实施方案中，相对于在用预防有效剂量的传统疫苗接种的受试者中诱导的免疫反应，所述受试者中诱导的免疫反应早2天、3天、1周、2周、3周、5周或10周。

[0687] 本文还提供了通过向受试者施用具有编码第一抗原的开放阅读框的mRNA而在受试者中引发针对冠状病毒的免疫反应的方法，其中RNA不包括稳定化元件，并且其中佐剂不与疫苗共配制或共施用。

[0688] 组合物可通过任何产生治疗有效结果的途径施用。这些途径包括(但不限于)真皮内、肌内、鼻内和/或皮下施用。本公开提供了包括向有需要的受试者施用RNA疫苗的方法。所需的确切量将因受试者而异，这取决于受试者的物种、年龄和一般状况、疾病的严重程度、特定组合物、其施用模式、其活性模式等。RNA通常配制成剂量单位形式以便于施用和剂量的均匀性。然而，应理解，RNA的总日用量可由主治医师在合理的医学判断范围内决定。对于任何特定患者的具体治疗有效、预防有效或适当的成像剂量水平将取决于多种因素，包括所治疗的病症和病症的严重程度；所采用的具体化合物的活性；所采用的具体组合物；患者的年龄、体重、一般健康状况、性别和饮食；施用时间、施用途径和所采用的具体化合物的排泄速率；治疗的持续时间；与所采用的具体化合物组合或同时使用的药物；以及医学领域众所周知的类似因素。

[0689] 如本文提供的RNA的有效量可低至20μg，例如以单一剂量或两个10μg剂量施用。在一些实施方案中，有效量为20μg-300μg或25μg-300μg的总剂量。例如，有效量可为20μg、25μg、30μg、35μg、40μg、45μg、50μg、55μg、60μg、65μg、70μg、75μg、80μg、85μg、90μg、95μg、100μ

g、110μg、120μg、130μg、140μg、150μg、160μg、170μg、180μg、190μg、200μg、250μg或300μg的总剂量。在一些实施方案中，有效量为20μg的总剂量。在一些实施方案中，有效量为25μg的总剂量。在一些实施方案中，有效量为50μg的总剂量。在一些实施方案中，有效量为75μg的总剂量。在一些实施方案中，有效量为100μg的总剂量。在一些实施方案中，有效量为150μg的总剂量。在一些实施方案中，有效量为200μg的总剂量。在一些实施方案中，有效量为250μg的总剂量。在一些实施方案中，有效量为300μg的总剂量。

[0690] 本文所述的RNA可配制成本文所述的剂型，例如鼻内、气管内或可注射(例如，静脉内、眼内、玻璃体内、肌内、真皮内、心内、腹膜内和皮下)。

[0691] 疫苗功效

[0692] 本公开的一些方面提供了组合物(例如，RNA疫苗)的制剂，其中以有效量配制RNA以在受试者中产生抗原特异性免疫反应(例如，产生对冠状病毒抗原特异性的抗体)。“有效量”是有效产生抗原特异性免疫反应的RNA的剂量。本文还提供了在受试者中诱导抗原特异性免疫反应的方法。

[0693] 如本文所用，对本公开的疫苗或LNP的免疫反应是在受试者中产生对疫苗中存在的(一种或多种)冠状病毒蛋白的体液和/或细胞免疫反应。出于本公开的目的，“体液”免疫反应是指由抗体分子(包括例如分泌性(IgA)或IgG分子)介导的免疫反应，而“细胞”免疫反应是由T淋巴细胞(例如，CD4+辅助细胞和/或CD8+ T细胞(例如CTL))和/或其他白细胞介导的免疫反应。细胞免疫的一个重要方面包括由细胞溶解性T细胞(CTL)引起的抗原特异性反应。CTL对与主要组织相容性复合体(MHC)编码的蛋白质缔合呈递并在细胞表面上表达的肽抗原具有特异性。CTL帮助诱导和促进细胞内微生物的破坏或感染此类微生物的细胞的溶解。细胞免疫的另一个方面包括由辅助T细胞引起的抗原特异性反应。辅助T细胞用于帮助刺激功能并且使非特异性效应细胞的活性集中于在表面上展示与MHC分子缔合的肽抗原的细胞。细胞免疫反应还导致产生细胞因子、趋化因子和由活化T细胞和/或其他白细胞产生的其他此类分子(包括源自CD4+和CD8+ T细胞的那些)。

[0694] 在一些实施方案中，抗原特异性免疫反应通过测量在施用如本文提供的组合物的受试者中产生的抗冠状病毒抗原抗体滴度来表征。抗体滴度是受试者内抗体(例如，对特定抗原或抗原的表位特异性的抗体)的量的测量结果。抗体滴度通常表示为提供阳性结果的最大稀释度的倒数。例如，酶联免疫吸附测定(ELISA)是用于测定抗体滴度的常见测定。

[0695] 可使用多种血清学测试来测量针对编码的目标抗原的抗体，例如SAR-CoV-2病毒或SAR-CoV-2病毒抗原，例如SAR-CoV-2刺突或S蛋白，或其结构域。这些测试包括血球凝集抑制测试、补体结合测试、荧光抗体测试、酶联免疫吸附测定(ELISA)和噬斑减少中和测试(PRNT)。这些测试中的每一者测量不同抗体活性。在示例性实施方案中，噬斑减少中和测试或PRNT(例如PRNT50或PRNT90)用作保护的血清学相关物。PRNT测量体外病毒中和的生物学参数，并且是某些类别的病毒中最具血清学病毒特异性的测试，与病毒感染的保护的血清水平充分相关。

[0696] PRNT的基本设计允许在试管或微量滴定板中发生病毒-抗体相互作用，并且然后通过将混合物平铺在病毒易感性细胞、优选哺乳动物来源的细胞上来测量抗体对病毒感染性的效应。用限制后代病毒传播的半固体培养基覆盖细胞。引发生产性感染的每个病毒产生局部感染区域(噬斑)，其可以多种方式检测。对噬斑计数并与病毒的起始浓度进行比较

以确定总病毒感染性的降低百分比。在PRNT中，通常在与标准量的病毒混合之前，将所测试的血清样品进行连续稀释。病毒的浓度保持恒定，使得当添加至易感细胞中并用半固体培养基覆盖时，可辨别个别噬斑并计数。以这种方式，可在任何选择的病毒活性降低百分比下计算每个血清样品的PRNT终点滴度。

[0697] 在旨在评价疫苗免疫原性的功能测定中，用于抗体滴定的血清样品稀释系列理想地应在低于“血清保护性”阈值滴度下开始。关于SARS-CoV-2中和抗体，“血清保护性”阈值滴度仍然未知；但在某些实施方案中，1:10的血清阳性阈值可被认为是血清保护阈值。

[0698] PRNT终点滴度表示为显示所需的噬斑计数减少百分比的最后血清稀释度的倒数。PRNT滴度可基于50％或更大的噬斑计数减少来计算(PRNT50)。对于疫苗血清，PRNT50滴度优于使用较高截止值(例如，PRNT90)的滴度，从滴定曲线的线性部分提供更准确的结果。

[0699] 存在用以计算PRNT滴度的若干方式。计算滴度的最简单和最广泛使用的方式是对噬斑计数，并且将滴度报告为基于输入噬斑的反滴定，显示输入噬斑计数减少＞50％的最后血清稀释度的倒数。使用来自若干血清稀释的曲线拟合方法可允许计算更精确的结果。存在多种计算机分析程序可用于此目的(例如，SPSS或GraphPad Prism)。

[0700] 在一些实施方案中，抗体滴度用于评价受试者是否已经感染或确定是否需要免疫。在一些实施方案中，抗体滴度用于确定自身免疫反应的强度，确定是否需要加强免疫，确定先前的疫苗是否有效，以及鉴定任何最近或先前的感染。根据本公开，抗体滴度可用于确定由组合物(例如，RNA疫苗)在受试者中诱导的免疫反应的强度。

[0701] 在一些实施方案中，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加至少1log。例如，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照可增加至少1.5、至少2、至少2.5或至少3log。在一些实施方案中，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加1、1.5、2、2.5或3log。在一些实施方案中，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加1-3log。例如，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照可增加1-1.5、1-2、1-2.5、1-3、1.5-2、1.5-2.5、1.5-3、2-2.5、2-3或2.5-3log。

[0702] 在一些实施方案中，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加至少2倍。例如，受试者中产生的抗冠状病毒抗原抗体滴度可相对于对照增加至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍或至少10倍。在一些实施方案中，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加2、3、4、5、6、7、8、9或10倍。在一些实施方案中，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加2-10倍。例如，受试者中产生的抗冠状病毒抗原抗体滴度相对于对照可增加2-10、2-9、2-8、2-7、2-6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9或9-10倍。

[0703] 在一些实施方案中，抗原特异性免疫反应测量为冠状病毒血清中和抗体滴度的几何平均滴度(GMT)的比率，称为几何平均比率(GMR)。几何平均滴度(GMT)是通过将所有值相乘并且取所述数的n次方根来计算的一组受试者的平均抗体滴度，其中n是具有可用数据的受试者的数目。

[0704] 在一些实施方案中，对照是在未施用组合物(例如RNA疫苗)的受试者中产生的抗冠状病毒抗原抗体滴度。在一些实施方案中，对照是在施用重组或纯化的蛋白疫苗的受试者中产生的抗冠状病毒抗原抗体滴度。重组蛋白疫苗通常包括在异源表达系统(例如细菌

或酵母)中产生或从大量病原生物体中纯化的蛋白抗原。

[0705] 在一些实施方案中,在鼠类模型中测量组合物(例如RNA疫苗)有效的能力。例如,可将组合物施用给鼠类模型,并且分析所述鼠类模型的中和抗体滴度的诱导。病毒攻击研究还可用于评价本公开的疫苗的功效。例如,可将组合物施用给鼠类模型,用病毒攻击鼠类模型,并测定鼠类模型的存活和/或免疫反应(例如中和抗体反应、T细胞反应(例如细胞因子反应))。

[0706] 在一些实施方案中,组合物(例如,RNA疫苗)的有效量是与重组蛋白疫苗的标准护理剂量相比减少的剂量。如本文所提供的"照护标准"是指医学或心理学治疗指南,并且可为一般性的或特异性的。"照护标准"基于科学证据和参与给定疾患的治疗的医学专业人士之间的合作来指定适当治疗。其是医师/临床医师对于某种类型的患者、疾病或临床情况应遵循的诊断和治疗过程。如本文所提供的"照护标准剂量"是指医师/临床医师或其他医学专业人士在遵循用于治疗或预防冠状病毒感染或相关疾患的照护标准指南的同时,将施用给受试者以治疗或预防冠状病毒感染或相关疾患的重组或纯化蛋白疫苗、或活的减毒或灭活疫苗、或VLP疫苗的剂量。

[0707] 在一些实施方案中,在施用有效量的组合物的受试者中产生的抗冠状病毒抗原抗体滴度等同于在施用照护标准剂量的重组或纯化蛋白疫苗、或活的减毒或灭活疫苗、或VLP疫苗的对照受试者中产生的抗冠状病毒抗原抗体滴度。

[0708] 疫苗功效可使用标准分析来评价(例如,参见Weinberg等人,J Infect Dis.2010年6月1日;201(11):1607-10)。例如,疫苗功效可通过双盲、随机化、临床对照试验来测量。疫苗功效可表示为未接种疫苗(ARU)和接种疫苗(ARV)的研究同类群组之间疾病罹患率(AR)的成比例降低,并且可使用下式从接种疫苗组中疾病的相对风险(RR)计算:

[0709] 功效=(ARU-ARV)/ARU×100;并且

[0710] 功效=(1-RR)×100。

[0711] 同样,疫苗有效性可使用标准分析来评估(例如,参见Weinberg等人,J Infect Dis.2010年6月1日;201(11):1607-10)。疫苗有效性是评价疫苗(其可能已经证明具有高疫苗功效)如何减少群体中的疾病。此量度可评价在自然现场条件下而非在受控临床试验中,疫苗接种程序而不仅仅疫苗本身的益处和不利效应的净平衡。疫苗有效性与疫苗功效(效力)成正比,但也受群体中的目标组的免疫程度以及影响住院治疗、流动就诊或费用的"真实世界"结果的其他非疫苗相关因素影响。例如,可使用回溯性病例对照分析,其中比较一组感染病例和适当对照之间的疫苗接种率。疫苗有效性可表示为比率差异,其中使用针对疫苗接种后仍发生感染的优势比(OR):

[0712] 有效性=(1-OR)×100。

[0713] 在一些实施方案中,相对于未接种疫苗的对照受试者,组合物(例如,RNA疫苗)的功效为至少60%。例如,相对于未接种疫苗的对照受试者,组合物的功效可为至少65%、至少70%、至少75%、至少80%、至少85%、至少95%、至少98%或100%。

[0714] 消除性免疫性。消除性免疫性是指防止病原体有效感染至宿主中的独特免疫状态。在一些实施方案中,有效量的本公开的组合物足以在受试者中提供消除性免疫性至少1年。例如,有效量的本公开的组合物足以在受试者中提供消除性免疫性至少2年、至少3年、至少4年或至少5年。在一些实施方案中,有效量的本公开的组合物足以在受试者中以相对

于对照低至少5倍的剂量提供消除性免疫性。例如,有效量可足以在受试者中以相对于对照低至少10倍、15倍或20倍的剂量提供消除性免疫性。

[0715] 可检测抗原。在一些实施方案中,有效量的本公开的组合物足以产生可检测水平的冠状病毒抗原,如在施用后1-72小时在受试者的血清中所测量。

[0716] 滴度。抗体滴度是受试者内抗体(例如,对特定抗原(例如,抗冠状病毒抗原)特异性的抗体)的数量的量度。抗体滴度通常表示为提供阳性结果的最大稀释度的倒数。例如,酶联免疫吸附测定(ELISA)是用于测定抗体滴度的常见测定。

[0717] 在一些实施方案中,有效量的本公开的组合物足以产生由针对冠状病毒抗原的中和抗体产生的1,000-10,000中和抗体滴度,如在施用后1-72小时在受试者的血清中所测量。在一些实施方案中,有效量足以产生由针对冠状病毒抗原的中和抗体产生的1,000-5,000中和抗体滴度,如在施用后1-72小时在受试者的血清中所测量。在一些实施方案中,有效量足以产生由针对冠状病毒抗原的中和抗体产生的5,000-10,000中和抗体滴度,如在施用后1-72小时在受试者的血清中所测量。

[0718] 在一些实施方案中,所述中和抗体滴度是至少100NT_{50}。例如,所述中和抗体滴度可为至少200、300、400、500、600、700、800、900或1000NT_{50}。在一些实施方案中,所述中和抗体滴度是至少10,000NT_{50}。

[0719] 在一些实施方案中,所述中和抗体滴度是至少100个中和单位/毫升(NU/mL)。例如,所述中和抗体滴度可为至少200、300、400、500、600、700、800、900或1000NU/mL。在一些实施方案中,所述中和抗体滴度是至少10,000NU/mL。

[0720] 在一些实施方案中,受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加至少1log。例如,受试者中产生的抗冠状病毒抗原抗体滴度相对于对照可增加至少2、3、4、5、6、7、8、9或10log。

[0721] 在一些实施方案中,受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加至少2倍。例如,受试者中产生的抗冠状病毒抗原抗体滴度相对于对照增加至少3、4、5、6、7、8、9或10倍。

[0722] 在一些实施方案中,几何平均值,即n个数的乘积的n次方根,通常用于阐述成比例生长。在一些实施方案中,几何平均值用于表征受试者中产生的抗体滴度。

[0723] 对照可为例如未接种疫苗的受试者,或施用活的减毒病毒疫苗、灭活病毒疫苗或蛋白质亚单位疫苗的受试者。

[0724] 额外实施方案

[0725] 以下编号的段落涵盖本公开的额外实施方案:

[0726] 1.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)和蛋白质跨膜结构域。

[0727] 2.如段落1所述的mRNA,其中所述蛋白质跨膜结构域是流感血凝素跨膜结构域。

[0728] 3.如段落2所述的mRNA,其中所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少80%同一性的氨基酸序列。

[0729] 4.如段落3所述的mRNA,其中所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基

酸序列。

[0730] 5.如段落4所述的mRNA,其中所述融合蛋白包含SEQ ID NO:77的氨基酸序列。

[0731] 6.如前述段落中任一项所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70%同一性的核苷酸序列。

[0732] 7.如段落6所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0733] 8.如段落7所述的mRNA,其中所述开放阅读框包含SEQ ID NO:76的核苷酸序列。

[0734] 9.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的氨基(N)末端结构域(NTD)和跨膜结构域。

[0735] 10.如段落9所述的mRNA,其中所述跨膜结构域是流感血凝素跨膜结构域。

[0736] 11.如段落10所述的mRNA,其中所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少80%同一性的氨基酸序列。

[0737] 12.如段落11所述的mRNA,其中所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0738] 13.如段落12所述的mRNA,其中所述融合蛋白包含SEQ ID NO:47的氨基酸序列。

[0739] 14.如前述段落中任一项所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70%同一性的核苷酸序列。

[0740] 15.如段落14所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0741] 16.如段落15所述的mRNA,其中所述开放阅读框包含SEQ ID NO:46的核苷酸序列。

[0742] 17.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码融合蛋白的开放阅读框,所述融合蛋白包含与SARS-CoV-2刺突蛋白的受体结合结构域连接的SARS-CoV-2刺突蛋白的氨基(N)末端结构域。

[0743] 18.如段落17所述的mRNA,其中所述融合蛋白还包含跨膜结构域。

[0744] 19.如段落18所述的mRNA,其中所述融合蛋白包含与SEQ ID NO:92的氨基酸序列具有至少80%同一性的氨基酸序列。

[0745] 20.如段落18所述的mRNA,其中所述融合蛋白包含与SEQ ID NO:92的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0746] 21.如段落20所述的mRNA,其中所述融合蛋白包含SEQ ID NO:92的氨基酸序列。

[0747] 22.如前述段落中任一项所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:91的核苷酸序列具有至少70%同一性的核苷酸序列。

[0748] 23.如段落22所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:91的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0749] 24.如段落23所述的mRNA,其中所述开放阅读框包含SEQ ID NO:91的核苷酸序列。

[0750] 25.如前述段落中任一项所述的mRNA,所述mRNA还包含5′非翻译区(UTR),所述5′非翻译区任选地包含SEQ ID NO:131或2的核苷酸序列。

[0751] 26.如前述段落中任一项所述的mRNA,所述mRNA还包含3′非翻译区(UTR),所述3′非翻译区任选地包含SEQ ID NO:132或4的核苷酸序列。

[0752] 27.如前述段落中任一项所述的mRNA,所述mRNA还包含5′帽,任选地7mG(5′)ppp(5′)NlmpNp。

[0753] 28.如前述段落中任一项所述的mRNA,所述mRNA还包含多聚A尾,任选地具有约100个核苷酸的长度。

[0754] 29.如前述段落中任一项所述的mRNA,其中所述mRNA包含化学修饰,任选地1-甲基假尿苷。

[0755] 30.一种组合物,所述组合物包含段落1-29中任一项所述的mRNA。

[0756] 31.一种组合物,所述组合物包含段落1-8中任一项所述的mRNA和段落9-16中任一项所述的mRNA。

[0757] 32.一种组合物,所述组合物包含段落17-29中任一项所述的mRNA。

[0758] 33.一种组合物,所述组合物包含:

[0759] (a)信使核糖核酸(mRNA),其包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)和蛋白质跨膜结构域;和

[0760] (b)mRNA,其包含编码融合蛋白的开放阅读框,所述融合蛋白包含SARS-CoV-2刺突蛋白的氨基(N)末端结构域和跨膜结构域。

[0761] 34.如段落33所述的组合物,其中所述蛋白质跨膜结构域是流感血凝素跨膜结构域。

[0762] 35.如段落34所述的组合物,其中(a)的所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少80%同一性的氨基酸序列。

[0763] 36.如段落35所述的组合物,其中(a)的所述融合蛋白包含与SEQ ID NO:77的氨基酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0764] 37.如段落36所述的组合物,其中(a)的所述融合蛋白包含SEQ ID NO:77的氨基酸序列。

[0765] 38.如段落34-37中任一项所述的组合物,其中(a)的所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70%同一性的核苷酸序列。

[0766] 39.如段落38所述的组合物,其中(a)的所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0767] 40.如段落39所述的组合物,其中(a)的所述开放阅读框包含SEQ ID NO:76的核苷酸序列。

[0768] 41.如段落34-40中任一项所述的组合物,其中(b)的所述融合蛋白包含与SEQ ID NO:47的氨基酸序列具有至少80%同一性的氨基酸序列。

[0769] 42.如段落41所述的组合物,其中(b)的所述融合蛋白包含与SEQ ID NO:47的氨基

酸序列具有至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列。

[0770] 43.如段落42所述的组合物,其中(b)的所述融合蛋白包含SEQ ID NO:47的氨基酸序列。

[0771] 44.如段落34-43中任一项所述的组合物,其中(b)的所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70%同一性的核苷酸序列。

[0772] 45.如段落44所述的组合物,其中(b)的所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列。

[0773] 46.如段落45所述的组合物,其中(b)的所述开放阅读框包含SEQ ID NO:46的核苷酸序列。

[0774] 47.如段落33-46中任一项所述的组合物,其中(a)的所述mRNA与(b)的所述mRNA的比率为约1:1。

[0775] 48.如段落1-29中任一项所述的mRNA,所述mRNA配制在脂质纳米颗粒中。

[0776] 49.如段落30-47中任一项所述的组合物,所述组合物还包含脂质纳米颗粒。

[0777] 50.如段落49所述的组合物,其中所述mRNA配制在所述脂质纳米颗粒中。

[0778] 51.如段落33-47中任一项所述的组合物,其中(a)的所述mRNA配制在脂质纳米颗粒中并且(b)的所述mRNA配制在脂质纳米颗粒中。

[0779] 52.如段落51所述的组合物,其中(a)和(b)的所述mRNA在相同脂质纳米颗粒中,或者其中(a)和(b)的所述mRNA相对于彼此各自配制在单独纳米颗粒中。

[0780] 53.如段落48所述的mRNA或如段落49-52中任一项所述的组合物,其中所述脂质纳米颗粒包含阳离子脂质。

[0781] 54.如段落53所述的mRNA或组合物,其中所述脂质纳米颗粒还包含中性脂质。

[0782] 55.如段落53或54所述的mRNA或组合物,其中所述脂质纳米颗粒还包含固醇。

[0783] 56.如段落53-55中任一项所述的mRNA或组合物,其中所述脂质纳米颗粒还包含聚乙二醇(PEG)修饰的脂质。

[0784] 57.如段落53-56中任一项所述的mRNA或组合物,其中所述脂质纳米颗粒包含可电离的阳离子脂质、中性脂质、固醇和PEG修饰的脂质。

[0785] 58.如段落57所述的mRNA或组合物,其中所述可电离的阳离子脂质是8((2羟基乙基)(6氧代6-(十一烷基氧基)己基)氨基)辛酸十七烷-9-基酯(化合物1)。

[0786] 59.如段落57或58所述的mRNA或组合物,其中所述中性脂质是1,2二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)。

[0787] 60.如段落57-59中任一项所述的mRNA或组合物,其中所述固醇是胆固醇。

[0788] 61.如段落57-60中任一项所述的mRNA或组合物,其中所述PEG修饰的脂质是1,2二肉豆蔻酰基-sn-甘油,甲氧基聚乙二醇(PEG2000 DMG)。

[0789] 62.如段落57-61中任一项所述的mRNA或组合物,其中所述脂质纳米颗粒包含20-60mol%可电离的阳离子脂质、5-25mol%中性脂质、25-55mol%固醇和0.5-15mol%PEG修饰的脂质。

[0790] 63.如段落62所述的mRNA或组合物,其中所述脂质纳米颗粒包含:

[0791] 47mol％可电离的阳离子脂质;11.5mol％中性脂质;38.5mol％固醇;和3.0mol％PEG修饰的脂质;

[0792] 48mol％可电离的阳离子脂质;11mol％中性脂质;38.5mol％固醇;和2.5mol％PEG修饰的脂质;

[0793] 49mol％可电离的阳离子脂质;10.5mol％中性脂质;38.5mol％固醇;和2.0mol％PEG修饰的脂质;

[0794] 50mol％可电离的阳离子脂质;10mol％中性脂质;38.5mol％固醇;和1.5mol％PEG修饰的脂质;或

[0795] 51mol％可电离的阳离子脂质;9.5mol％中性脂质;38.5mol％固醇;和1.0mol％PEG修饰的脂质。

[0796] 64.如段落63所述的mRNA或组合物,其中所述脂质纳米颗粒包含:

[0797] 47mol％化合物1;11.5mol％DSPC;38.5mol％胆固醇;和3.0mol％PEG2000 DMG;

[0798] 48mol％化合物1;11mol％DSPC;38.5mol％胆固醇;和2.5mol％PEG2000 DMG;

[0799] 49mol％化合物1;10.5mol％DSPC;38.5mol％胆固醇;和2.0mol％PEG2000 DMG;

[0800] 50mol％化合物1;10mol％DSPC;38.5mol％胆固醇;和1.5mol％PEG2000 DMG;或

[0801] 51mol％化合物1;9.5mol％DSPC;38.5mol％胆固醇;和1.0mol％PEG2000 DMG。

[0802] 65.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应的量的前述段落中任一项所述的mRNA或组合物。

[0803] 66.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的T细胞免疫反应的量的前述段落中任一项所述的mRNA或组合物。

[0804] 67.一种信使核糖核酸(mRNA),所述信使核糖核酸(mRNA)包含编码冠状病毒抗原的开放阅读框(ORF),所述冠状病毒抗原能够诱导针对SARS-CoV-2的免疫反应,例如中和抗体反应,其中所述抗原包含SARS-CoV-2的蛋白质片段或功能性蛋白质结构域,任选地其中所述RNA配制在脂质纳米颗粒中。

[0805] 68.如段落67所述的mRNA,其中所述抗原是功能性蛋白质结构域。

[0806] 69.如段落68所述的mRNA,其中所述蛋白质结构域是SARS-CoV-2刺突蛋白的N末端结构域(NTD)。

[0807] 70.如段落69所述的mRNA,其中所述NTD与跨膜结构域、任选地流感血凝素跨膜结构域连接。

[0808] 71.如段落70所述的mRNA,其中所述抗原包含与SEQ ID NO:47的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:47的氨基酸序列。

[0809] 72.如段落70或71所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:46的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:46的核苷酸序列。

[0810] 73.如段落68所述的mRNA,其中所述蛋白质结构域是SARS-CoV-2刺突蛋白的受体结合结构域(RBD)。

[0811] 74.如段落73所述的mRNA,其中所述RBD是可溶的。

[0812] 75.如段落74所述的mRNA,其中所述抗原包含与SEQ ID NO:62的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:62的氨基酸序列。

[0813] 76.如段落74或75所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:61的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:61的核苷酸序列。

[0814] 77.如段落73所述的mRNA,其中所述RBD连接至跨膜结构域,任选地流感血凝素跨膜结构域。

[0815] 78.如段落77所述的mRNA,其中所述抗原包含与SEQ ID NO:77的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:77的氨基酸序列。

[0816] 79.如段落77或78所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:76的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:76的核苷酸序列。

[0817] 80.如段落69所述的mRNA,其中所述NTD与SARS-CoV-2刺突蛋白的RBD连接以形成NTD-RBD融合蛋白。

[0818] 81.如段落80所述的mRNA,其中所述NTD-RBD融合体与跨膜结构域(TM)、任选地流感血凝素跨膜结构域连接,以形成NTD-RBD-TM蛋白。

[0819] 82.如段落81所述的mRNA,其中所述抗原包含与SEQ ID NO:92的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:92的氨基酸序列。

[0820] 83.如段落81或82所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:91的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:91的核苷酸序列。

[0821] 84.如段落80所述的mRNA,其中所述NTD-RBD融合体包含C末端截短。

[0822] 85.如段落84所述的mRNA,其中所述抗原包含与SEQ ID NO:107的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:107的氨基酸序列。

[0823] 86.如段落84或85所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:106的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:106的核苷酸序列。

[0824] 87.如前述段落中任一项所述的mRNA,其中所述NTD和/或RBD包括延伸区。

[0825] 88.如段落87所述的mRNA,其中所述抗原包含与SEQ ID NO:59、86、89、116、119或122中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:59、

86、89、116、119或122中任一者的氨基酸序列。

[0826] 89.如段落87或88所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:58、85、88、115、118或121中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:58、85、88、115、118或121中任一者的核苷酸序列。

[0827] 90.如段落68所述的mRNA,其中所述蛋白质结构域是SARS-CoV-2刺突蛋白的S1亚基结构域。

[0828] 91.如段落90所述的mRNA,其中所述S1亚基是可溶的。

[0829] 92.如段落91所述的mRNA,其中所述抗原包含与SEQ ID NO:5的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:5的氨基酸序列。

[0830] 93.如段落91或92所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:3的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:3的核苷酸序列。

[0831] 94.如段落90所述的mRNA,其中所述S1亚基连接至跨膜结构域、任选地与流感血凝素跨膜结构域。

[0832] 95.如段落94所述的mRNA,其中所述抗原包含与SEQ ID NO:17的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:17的氨基酸序列。

[0833] 96.如段落94或95所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:16的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:16的核苷酸序列。

[0834] 97.如段落90所述的mRNA,其中所述S1亚基已被修饰以去除S蛋白的RBD或RBD的一部分。

[0835] 98.如段落97所述的mRNA,其中所述抗原包含与SEQ ID NO:20、23、26、29、32或35中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:20、23、26、29、32或35中任一者的氨基酸序列。

[0836] 99.如段落97或98所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:19、22、25、28、41或34中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:19、22、25、28、31或34中任一者的核苷酸序列。

[0837] 100.如段落90所述的mRNA,其中所述S1亚基与S蛋白的S2亚基连接。

[0838] 101.如段落100所述的mRNA,其中所述S2亚基来自SARS-CoV-2 S蛋白,并且在一些实施方案中,其中所述S2亚基包含开放阅读框,所述开放阅读框包含与SEQ ID NO:145的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:145

的核苷酸序列。

[0839] 102.如段落101所述的mRNA,其中所述S1亚基来自HKU1 S蛋白。

[0840] 103.如段落102所述的mRNA,其中所述抗原包含与SEQ ID NO:38的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:38的氨基酸序列。

[0841] 104.如段落102或103所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:37的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:37的核苷酸序列。

[0842] 105.如段落101所述的mRNA,其中所述S1亚基来自OC43 S蛋白。

[0843] 106.如段落105所述的mRNA,其中所述抗原包含与SEQ ID NO:41的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:41的氨基酸序列。

[0844] 107.如段落105或106所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:40的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:40的核苷酸序列。

[0845] 108.如前述段落中任一项所述的mRNA,其中所述抗原还包含支架结构域,任选地选自铁蛋白、2,4-二氧四氢蝶啶合成酶和折叠子。

[0846] 109.如段落108所述的mRNA,其中所述支架结构域是铁蛋白。

[0847] 110.如段落109所述的mRNA,其中所述抗原包含与SEQ ID NO:8或65的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:8或65的氨基酸序列。

[0848] 111.如段落109或110所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:7或64的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:7或64的核苷酸序列。

[0849] 112.如段落108所述的mRNA,其中所述支架结构域是2,4-二氧四氢蝶啶合成酶。

[0850] 113.如段落112所述的mRNA,其中所述抗原包含与SEQ ID NO:11、14、68或71中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的氨基酸序列,任选地其中所述抗原包含SEQ ID NO:11、14、68或71中任一者的氨基酸序列。

[0851] 114.如段落112或113所述的mRNA,其中所述开放阅读框包含与SEQ ID NO:10、13、67或70中任一者的核苷酸序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的核苷酸序列,任选地其中所述开放阅读框包含SEQ ID NO:10、13、67或70中任一者的核苷酸序列。

[0852] 115.如段落108所述的mRNA,其中所述支架结构域是折叠子。

[0853] 116.如段落115所述的mRNA,其中所述抗原包含与SEQ ID NO:44、50、74、80、83、101、104或113中任一者的氨基酸序列具有至少80%、至少85%、至少90%、至少95%、至少

96％、至少97％、至少98％或至少99％同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:44、50、74、80、83、101、104或113中任一者的氨基酸序列。

[0854] 117.如段落115或116所述的mRNA，其中所述开放阅读框包含与SEQ ID NO:43、49、73、79、82、100、103或112中任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:43、49、73、79、82、100、103或112中任一者的核苷酸序列。

[0855] 118.如前述段落中任一项所述的mRNA，其中所述抗原还包含运输信号，任选地选自巨噬细胞标志物，任选地CD86、CD11B和/或VSVGct。

[0856] 119.如段落118所述的mRNA，其中所述抗原包含与SEQ ID NO:95、98或110中任一者的氨基酸序列具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的氨基酸序列，任选地其中所述抗原包含SEQ ID NO:95、98或110中任一者的氨基酸序列。

[0857] 120.如段落118或119所述的mRNA，其中所述开放阅读框包含与SEQ ID NO:94、97或109中任一者的核苷酸序列具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的核苷酸序列，任选地其中所述开放阅读框包含SEQ ID NO:94、97或109中任一者的核苷酸序列。

[0858] 121.如段落67-120中任一项所述的mRNA，所述mRNA配制在脂质纳米颗粒中。

[0859] 122.如段落121所述的mRNA，其中所述脂质纳米颗粒包含阳离子脂质、任选地可电离的阳离子脂质、中性脂质、固醇和/或聚乙二醇(PEG)修饰的脂质。

[0860] 123.如段落108所述的mRNA或组合物，其中所述可电离的阳离子脂质是8((2羟基乙基)(6氧代6-(十一烷基氧基)己基)氨基)辛酸十七烷-9-基酯(化合物1)，所述中性脂质是1,2二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)，所述固醇是胆固醇，和/或所述PEG修饰的脂质是1,2二肉豆蔻酰基-sn-甘油，甲氧基聚乙二醇(PEG2000 DMG)。

[0861] 124.如段落121-123中任一项所述的mRNA或组合物，其中所述脂质纳米颗粒包含20-60mol％可电离的阳离子脂质、5-25mol％中性脂质、25-55mol％固醇和0.5-15mol％PEG修饰的脂质。

[0862] 125.如段落124所述的mRNA或组合物，其中所述脂质纳米颗粒包含：

[0863] 47mol％可电离的阳离子脂质；11.5mol％中性脂质；38.5mol％固醇；和3.0mol％PEG修饰的脂质；

[0864] 48mol％可电离的阳离子脂质；11mol％中性脂质；38.5mol％固醇；和2.5mol％PEG修饰的脂质；

[0865] 49mol％可电离的阳离子脂质；10.5mol％中性脂质；38.5mol％固醇；和2.0mol％PEG修饰的脂质；

[0866] 50mol％可电离的阳离子脂质；10mol％中性脂质；38.5mol％固醇；和1.5mol％PEG修饰的脂质；或

[0867] 51mol％可电离的阳离子脂质；9.5mol％中性脂质；38.5mol％固醇；和1.0mol％PEG修饰的脂质。

[0868] 126.如段落125所述的mRNA或组合物，其中所述脂质纳米颗粒包含：

[0869] 47mol％化合物1;11.5mol％DSPC;38.5mol％胆固醇;和3.0mol％PEG2000 DMG;

[0870] 48mol％化合物1;11mol％DSPC;38.5mol％胆固醇;和2.5mol％PEG2000 DMG;

[0871] 49mol％化合物1;10.5mol％DSPC;38.5mol％胆固醇;和2.0mol％PEG2000 DMG;

[0872] 50mol％化合物1;10mol％DSPC;38.5mol％胆固醇;和1.5mol％PEG2000 DMG;或

[0873] 51mol％化合物1;9.5mol％DSPC;38.5mol％胆固醇;和1.0mol％PEG2000 DMG。

[0874] 127.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的中和抗体反应的量的段落67-126中任一项所述的mRNA。

[0875] 128.一种方法,所述方法包括向受试者施用有效地在所述受试者中诱导针对SARS-CoV-2的T细胞免疫反应的量的段落67-126中任一项所述的mRNA。

[0876] 实施例

[0877] 实施例1.表达数据

[0878] 本研究中使用的mRNA用于表达SARS-CoV-2冠状病毒刺突(S)蛋白的关键中和结构域,并且评价这些中和蛋白结构域在个别地或组合用作免疫原性组合物或疫苗以保护人免受活的和传播的天然病毒感染时是否可更有效地诱导保护性免疫性。由mRNA编码的蛋白质的线性设计如图2中所示。蛋白质还都含有源自流感的血凝素(HA)的羧基(C)末端跨膜结构域(TM)。

[0879] 已知NTD和RBD两者都是结合表现中和病毒活性的抗体的位点。在SARS-CoV-2的情形下,RBD是刺突蛋白的受体结合位点,并且结合血管收缩肽转化酶2(ACE2)。氨基(N)末端结构域NTD的功能尚未完全了解,似乎在结合糖部分和促进刺突蛋白从融合前构象向融合后构象的构象转变中起作用。参见Zhou H,Chen Y,Zhang S等人,Nat Commun.2019;10(1):3068。无论如何,NTD和RBD结构域两者都诱导如下文论述的高结合抗体和中和抗体滴度。

[0880] 使用对SARS-CoV-2刺突蛋白的受体结合结构域(RBD)(mAb1)和SARS-CoV-2刺突蛋白的N末端结构域(NTD)(Ab2)特异性的抗体,在表16和表17中显示mRNA RBD-TM疫苗(“SARS-CoV-2RBD-TM”;SEQ ID NO:75-77)、mRNA NTD-TM疫苗(“SARS-CoV-2 NTD-TM”;SEQ ID NO:45-47)和mRNA NTD-RBD-TM(“SARS-CoV-1NTD-RBD-TM”;SEQ ID NO:90-92)疫苗的表达数据。表16显示在24小时(hr)、48hr和72hr时,与WT SARS-CoV-2刺突蛋白mRNA相比,MFI*Freq的平均倍数差异(在稀释范围内)(表16)。

[0881] 表16.与S2P蛋白相比的总抗原表达(MFI*Freq)的倍数变化

[0882]

	单克隆抗体(mAb)	野生型(WT) S mRNA	RBD-TM mRNA	NTD-TM mRNA	NTD-RBD-TM mRNA	空白
24 小时(hr)	1 – 特异于 RBD	1.8	30.6	0.0	10.4	0.0
24 hr	2 – 特异于 NTD	0.7	0.0	9.7	2.5	0.0
48 hr	1	1.2	40.1	0.0	7.3	0.0
48 hr	2	0.8	0.0	19.0	6.0	0.0
72 hr	1	0.8	32.5	0.1	11.6	0.0
72 hr	2	0.7	0	14.4	4.1	0.0

[0883] RBD＝受体结合结构域

[0884] NTD＝N末端结构域

[0885] TM＝跨膜结构域

[0886] 实施例2.单一剂量后第21天的免疫原性数据和中和数据

[0887] 以下列剂量向小鼠施用mRNA NTD-TM和mRNA RBD-TM(阐述于实施例1中):0.001μg、0.01μg、0.1μg或1μg(N＝8)。以下列剂量向小鼠施用mRNA NTD-RBD-TM(阐述于实施例1中):0.1μg或1μg(N＝8)。向小鼠施用mRNA NTD-TM和mRNA RBD-TM的50:50混合物,其含有0.1μg每个mRNA,总共0.2μg mRNA,或1μg每个mRNA,总共2μg mRNA(N＝8)。然后在疫苗接种后第21天通过ELISA测量SARS-CoV-2刺突蛋白特异性IgG滴度(表17)、SARS-CoV-2 RBD特异性IgG滴度(表18)和SARS-CoV-2 NTD特异性IgG滴度(表19)。数据提供于表17-19中。0.1μg剂量的mRNA NTD-RBD-TM和0.2μg剂量的mRNA NTD-TM和mRNA RBD-TM组合物的50:50混合物引发可观察到的NTD特异性和RBD特异性IgG滴度,并且0.1μg剂量的RBD-TM和NTD-TM分别引发可测量的针对RBD和NTD抗原的IgG滴度。

[0888] 表17.SARS-CoV-2 S1/S2刺突蛋白特异性IgG滴度-平均值

[0889]

微克(每种mRNA构建体)	RBD-TM	NTD-TM	NTD-RBD-TM	50:50混合物
0.001	13	13	-	-
0.01	13	13	-	-
0.1	13	13	13	19
1	225	13	379	707

[0890] N＝8的平均值;仅PBS＝1.1

[0891] 表18.RBD结构域特异性IgG滴度-平均值

[0892]

微克(每种mRNA构建体)	RBD-TM	NTD-RBD-TM	50:50混合物
0.001	15	-	-
0.01	14	-	-
0.1	75	209	266
1	1947	7710	15530

[0893] N＝8的平均值;仅PBS＝1.1

[0894] 表19.NTD结构域特异性IgG滴度-平均值

[0895]

微克(每种mRNA构建体)	NTD-TM	NTD-RBD-TM	50:50混合物
0.001	13	-	-
0.01	14	-	-

[0896]

0.1	30	20	38
1	620	848	1075

[0897] N＝8的平均值;仅PBS＝1.1

[0898] 测量来自用1μg剂量的RBD-TM和NTD-RBD-TM组合物和2μg剂量的NTD-TM和RBD-TM组合物的50:50混合物接种的小鼠的血清的中和滴度,并且分析ELISA滴度与中和滴度之间

的相关性(图7)。

[0899] 由1μg剂量的NTD-RBD-TM组合物或2μg的NTD-TM和RBD-TM组合物的50:50混合物引起的滴度大于由1μg剂量的RBD-TM组合物引起的那些滴度(表20)。在刺突特异性IgG、RBD特异性IgG和NTD特异性IgG的中和滴度和ELISA滴度之间存在显著的相关性(图7)。

[0900] 表20.中和滴度-平均值

[0901]

微克(每种mRNA构建体)	RBD-TM	NTD-RBD-TM	50:50 混合物
1	238	504	-
2	-	-	421

[0902] N=8的平均值;仅PBS=1.1

[0903] 基于重组VSVΔG的SARS-CoV-2假病毒中和测定

[0904] 将密码子优化的野生型或D614G刺突基因(Wuhan-Hu-1株;NC_045512.2)克隆至pCAGGS载体中。为了产生基于VSVΔG的SARS-CoV-2假病毒,如先前所述(Whitt,2010),用刺突表达质粒和感染的VSVΔG-萤火虫-荧光素酶转染BHK-21/WI-2细胞。使用A549-hACE2-TMPRSS2细胞作为基于VSVΔG的SARS-CoV-2假病毒中和测定的靶细胞。制备编码hACE2-P2A-TMPRSS2的慢病毒,以产生A549-hACE2-TMPRSS2细胞,将其维持在补充有10%胎牛血清和1μg/mL嘌呤霉素(puromycin)的DMEM中。在37℃下用假病毒感染A549-hACE2-TMPRSS2细胞1小时。感染后除去接种病毒或病毒-抗体混合物。18小时后,将等体积的One-Glo试剂(Promega;E6120)添加至培养基中,使用BMG PHERastar-FS读板器进行读数。中和程序和数据分析与上文在基于慢病毒的假病毒中和测定中所述相同。参见Whitt,M.A.(2010).Journal of Virological Methods 169,365-374。

[0905] 实施例3.两个剂量后第36天的免疫原性数据

[0906] 接种疫苗第一剂量后第22天,再次将相同剂量的实施例2中所述的mRNA疫苗作为加强剂量施用给小鼠。从第36天血清通过ELISA测量在加强剂量后产生的针对RBD抗原、NTD抗原、野生型(WT)刺突(S)蛋白和S2P蛋白(具有双脯氨酸突变以稳定融合前构象的S蛋白)中的每一者的抗体的滴度,并显示如下。将由LNP中的mRNA编码的RBD-TM和NTD-TM的两种免疫原性组合物的50:50混合物在第22天以2μg或0.2μg总mRNA作为加强剂量施用给小鼠,并且在第36天测定滴度。参见表21。

[0907] 表21中显示的用由LNP中的mRNA编码的RBD-TM、NTD-TM或NTD-RBD-TM免疫的小鼠的WT S蛋白滴度指示,在诱导可识别并结合至SARS-CoV-2 WT S蛋白的抗体方面,两个剂量在所有测试的剂量下都是优异的。

[0908] 表21.SARS-CoV-2 WT S特异性IgG滴度-几何平均值

[0909]

微克(mRNA 构建体)	RBD-TM	NTD-TM	NTD-RBD-TM
	第 21 天(GMT)滴度		
0.001	13	13	-
0.01	13	13	-
0.1	13	13	14
1	225	13	379
	第 36 天(GMT)滴度		
0.001	13	13	-
0.01	295	13	-
0.1	1,017	39	3879
1	27,674	6317	24,413

[0910] N=8的平均值;仅PBS=1.1

[0911] 进一步分析来自用两个剂量的由LNP中的mRNA编码的RBD-TM、NTD-TM或NTD-RBD-TM免疫的小鼠的血清的抗体识别和结合SARS-CoV-2 S2P蛋白的能力。通过ELISA使用S2P作为板上的抗原测定SARS-CoV-2 S2P蛋白的滴度,并且其示于下表22中。当S2P为抗原时,这些免疫原中的每一者诱导比当WT S蛋白为ELISA抗原时高得多的抗体滴度。比较表21和表22。

[0912] 表22.SARS-CoV-2 S2P特异性IgG滴度-几何平均值

[0913]

微克(mRNA 构建体)	RBD-TM	NTD-TM	NTD-RBD-TM
	第 36 天(GMT)滴度		
0.001	88	16	-
0.01	1,416	280	-
0.1	3,892	11,687	25,298
1	107,611	192,040	223,140

[0914] N=8的平均值;仅PBS=1.1

[0915] 在表23中,免疫原是由LNP中的mRNA编码的RBD-TM和NTD-TM的50:50混合物,并且在一个剂量(第21天)和两个剂量(第36天)后测定对WT S、RBD、NTD和S2P的滴度。这些结果显示,当免疫原是RBD-TM和NTD-TM以50:50混合的组合时,与相同剂量的个别抗原诱导的抗体滴度相比,滴度显著增加。50:50混合物诱导对免疫抗原的良好滴度,但令人惊讶的是,对WT S蛋白的滴度甚至更佳并且对S2P蛋白的滴度极高。参见表23。

[0916] 表23.SARS-CoV-2抗原特异性IgG滴度-几何平均值

[0917]

免疫原	ELISA 板上检测抗原			
微克(mRNA) 50:50 混合物 RBD-TM LNP NTD-TM LNP	抗刺突蛋白 IgG	抗 RBD IgG	抗 NTD IgG	抗 S2P IgG

[0918]

第 21 天(GMT)滴度				
0.2	19	266	38	4,869
2	707	15,530	1,075	52,936
第 36 天(GMT)滴度				
0.2	7,259	15,957	4,869	74,718
2	75,871	186,740	52,936	2,726,948

[0919] N＝8的平均值;仅PBS＝1.1

[0920] 表24显示用RBD-TM和NTD-TM中的每一者作为编码那些抗原的mRNA进行免疫的结果。使用由mRNA免疫原编码的蛋白质作为ELISA板上的抗原,测量每组8只小鼠的几何平均滴度。同样,当抗原作为配制在LNP中的mRNA施用时,两种免疫原性组合物都诱导对免疫抗原的高滴度。在这种情形中,两个剂量在所有浓度下都产生优异的抗体反应。然而,以每微克(μg)计,这些抗原的50:50混合物诱导的抗体反应是单独施用抗原时的约10倍。比较表23和表24。

[0921] 表24.SARS-CoV-2 RBD和NTD结构域特异性IgG滴度-几何平均值

[0922]

微克(mRNA构建体)	RBD-TM板上抗原	NTD-TM板上抗原
	第 21 天(GMT)滴度	
0.001	15	13
0.01	14	14
0.1	75	30
1	1,947	620
	第 36 天(GMT)滴度	
0.001	25	13
0.01	1,357	123
0.1	4,678	6898
1	84,385	46,516

[0923] N＝8的平均值;仅PBS＝1.1

[0924] 在第1天和第21天,以0.1和1μg剂量将包含与RBD连接并由LNP中的mRNA编码的NTD的融合蛋白作为免疫原性组合物施用给8只小鼠的组。参见下表25。甚至编码NTD-RBD-TM的融合蛋白形式的mRNA对个别结构域也诱导非常好的滴度,所述滴度高于当单一结构域为免疫抗原时的滴度。对S2P蛋白的滴度是对WT S蛋白的滴度的约8倍。参见表25。

[0925] 表25.SARS-CoV-2 NTD-RBD-TM结构域特异性IgG滴度-几何平均值

[0926]

微克免疫原 NTD-RBD-TM (mRNA构建体)	RBD-TM 板上抗原	NTD-TM 板上抗原	野生型刺突 板上抗原	S2P 板上抗原
		第21天(GMT)滴度		
0.1	209	20	13	NA
1	7,710	848	379	NA
		第36天(GMT)滴度		
0.1	5,769	2,875	3,878	25,298
1	111,747	41,876	24,413	223,140

[0927] N＝8的平均值;仅PBS＝1.1

[0928] 中和数据示于表26中。由mRNA编码的S1-666-TM是使用S1子结构域的抗原,特别是SARS-CoV-2刺突蛋白的附接至跨膜结构域的残基1-666。

[0929] 表26.第1天和第22天两次免疫后第36天的平均中和滴度。

[0930]

微克 (每种mRNA构建体)	S1-666-TM	NTD-TM和RBD-TM的50:50混合物	RBD-TM	NTD-TM	NTD-RBD-TM
0.001	-	-	40	40	-
0.01	508	-	70.5	40	-
0.1	11,561	1,336	260	205	865

[0931]

1.0	-	25,208	7,402	4,669	57,891

[0932] N＝8的平均值;仅PBS＝1.1

[0933] 实施例4.S1-666-TM的免疫原性

[0934] 将由LNP中的mRNA编码的S1-666-TM(或刺突蛋白S的S1残基1-666)在第1天作为初免免疫施用给小鼠,并且在第22天作为加强剂量施用0.01μg和0.1μg(N＝8)组。从第21天(加强前)和第36天(加强后)血清通过ELISA测量加强剂量后产生的针对mRNA RBD、mRNA NTD和mRNA野生型(WT)刺突(S)蛋白(图1)中每一者的抗体滴度,并且示于下表27中。

[0935] 表27中显示的用由LNP中的mRNA编码的S1-666-TM免疫的小鼠的WT S蛋白滴度指示,在诱导可识别并结合至SARS-CoV-2 WT S蛋白的抗体方面,两个剂量在所有测试的剂量下都是优异的。令人惊讶的是,即使在S1中未发现2P突变(这是因为2P突变发生在S2中,而免疫原中不存在S2),当针对刺突蛋白的S2P形式测量时,诱导的滴度也最高。与其他构建体相似,NTD滴度需要第二剂量以变高。

[0936] 表27.SARS-CoV-2抗原特异性IgG滴度-几何平均值

[0937]

免疫原	ELISA 板上检测抗原			
微克(mRNA) S1-666-TM	抗刺突蛋白 IgG	抗 RBD IgG	抗 NTD IgG	抗 S2P IgG
第 21 天(GMT)滴度				
0.01	13	30	13	NA
0.1	175	868	49	NA
第 36 天(GMT)滴度				
0.01	2,450	12,952	112	186,561
0.1	17,066	69,969	9,015	874,609

[0938] N＝8的平均值;仅PBS＝1.1

[0939] 实施例5.RBD-TM、NTD-TM、NTD-RBD-TM和NTD-TM/RBD-TM组合物的50:50混合物在两个剂量后在第36天的免疫原性

[0940] 在这种重复实验中,在接种第一剂量后第22天,再次向小鼠施用上述实施例中所述相同剂量的mRNA疫苗作为加强剂量。然后在接种第一剂量后第36天通过ELISA测量SARS-CoV-2刺突蛋白特异性IgG滴度、SARS-CoV-2 S2P蛋白特异性IgG滴度、SARS-CoV-2 RBD特异性IgG滴度和SARS-CoV-2 NTD特异性IgG滴度。

[0941] 结果显示1μg和0.1μg剂量的mRNA RBD-TM、mRNA NTD-TM、mRNA NTD-RBD-TM组合物和含有1μg或0.1μg mRNA RBD-TM和mRNA NTD-TM组合物中的每一者的50:50混合物引发针对SARS-CoV-2刺突或SARS-CoV-2 S2P蛋白的高ELISA滴度。

[0942] 实施例6.免疫原性研究

[0943] 将mRNA NTD-TM和mRNA RBD-TM的50:50混合物的免疫原性以下列剂量施用给小鼠:0.2μg或2μg总mRNA(0.1μg或1μg每个mRNA)(N＝8)。在第1天施用初免剂量,并且在第22天施用加强剂量。在第36天,使用ELISA来评价抗体与SARS-CoV-2稳定化的融合前刺突蛋白(SARS-CoV-2 pre-S)的结合。将包括mRNA NTD-RBD-TM、mRNA RBD-TM和mRNA NTD-TM的下列疫苗组合物以下列剂量施用给小鼠:0.1μg和0.01μg(N＝8)。测定GMT数据并示于下表28中。

[0944] 表28.S蛋白结构域融合和组合诱导的SARS-CoV-2刺突滴度

[0945]

疫苗免疫原	**剂量 μg mRNA**	**第 21 天 GMT 滴度 剂量 1 后 3 周 S2P 包被在板上**	**第 36 天 GMT 滴度 剂量 2 后 3 周 S2P 包被在板上**
50:50 混合物 RBD-TM NTD-TM	1	5,913	301,881

[0946]

50:50 混合物 RBD-TM NTD-TM	0.1	71	13,423
NTD-RBD-TM	1.0	3,329	177,499
NTD-RBD-TM	0.1	80	48947
RBD-TM	1.0	3198	593,730
RBD-TM	0.1	73	31,090
NTD-TM	1.0	593	74,777
NTD-TM	0.1	21	23,916

[0947] 实施例7.NTD-RBD-TM的IgG2a和IgG1的比率的测定

[0948] 将NTD-RBD-TM、mRNA NTD-RBD-TM的组合物以下列剂量施用给小鼠:0.1μg和1μg。在第1天施用初免剂量,并且在第22天施用加强剂量。在第36天,评价S2P特异性IgG1和IgG2a的滴度。参见图7A-7C。截至第36天,在两个剂量水平下,IgG2a的滴度都高于IgG1的量。参见图7A。为了确定T细胞反应是否偏向Th1或Th2型反应,我们绘制了在第36天时间点的IgG2a/IgG1的比率。如图7B中所示,NTD-RBD-TM组合物诱导抗体免疫反应,所述抗体免疫反应明确地在Th1型反应中。Th2型反应在疫苗开发中是不利的,这是因为其与驱动疾病增强有关。

[0949] 实施例8.免疫原性研究

[0950] 将表29中列出的mRNA以下列剂量施用给小鼠:0.1μg和1μg(N=8)。在第1天施用初免剂量,并且在第22天施用加强剂量。第21天和第36天在S2P包被的板上测定血清IgG滴度。结果示于表29中。

[0951] 表29.

[0952]

制剂/材料	剂量(μg)	第 21 天	第 36 天
PBS		1.097	1.097
NTD-RBD-TM	1	4.237	6.036
	0.1	2.711	5.106

[0953]

S1-666-TM (SEQ ID NO: 16)	1	4.224	5.704
	0.1	1.944	4.633
混合(50:50) NTD-EXT-F43C-TM (SEQ ID NO: 55) + RBD-Q563D-EXT-TM (SEQ ID NO: 88)	1	3.939	5.442
	0.1	1.867	4.342
NTD-EXT-RBD-EXT-TM (SEQ ID NO: 121)	1	4.541	6.001
	0.1	3.694	4.885
S1-594-TM (SEQ ID NO: 22)	1	4.027	5.517
	0.1	3.313	4.780
S1-594-PolyG-DS-TM (SEQ ID NO: 28)	1	3.974	5.531
	0.1	2.551	4.425
S2-TM (SEQ ID NO: 145)	1	1.989	4.058
	0.1	1.097	2.015
RBD-NTD-TM (SEQ ID NO: 139)	1	4.523	6.062
	0.1	3.453	5.222
NTD-PADRE-RBD-TM (SEQ ID NO: 142)	1	4.489	6.040
	0.1	3.386	5.301

[0954] 额外序列

[0955] 应理解,本文所述的任何mRNA序列可包括5′UTR和/或3′UTR。所述UTR序列可选自以下序列,或者可使用其他已知UTR序列。还应理解,本文所述的任何mRNA构建体还可包含多聚(A)尾和/或帽(例如,7mG(5′)ppp(5′)NlmpNp)。此外,尽管本文所述的许多mRNA和编码的抗原序列包括信号肽和/或肽标签(例如,C末端His标签),但应理解,所指示的信号肽和/或肽标签可用不同信号肽和/或肽标签取代,或者可省略信号肽和/或肽标签。

[0956] 5′UTR:GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGAGCCACC(SEQ ID NO:131)

[0957] 5′UTR:GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGACCCCGGCGCCGCCACC(SEQ ID NO:2)

[0958] 3′UTR:UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC(SEQ ID NO:132)

[0959] 3′UTR:UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC(SEQ ID NO:4)

[0960]

SARS-CoV-2 野生型刺突(S)蛋白		
SEQ ID NO: 123 从 5'端至 3'端由以下组成：5' UTR SEQ ID NO: 2、mRNA ORF SEQ ID NO: 124 和 3' UTR SEQ ID NO: 4。		123
化学	1-甲基假尿苷	
帽	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGU AAGAAGAAAUAUAAGACCCCGGCGC CGCCACC	2
mRNA 构建体的 ORF (不包括终止密码子)	AUGUUCGUGUUCCUGGUGCUGCUGC CCCUGGUGAGCAGCCAGUGCGUGAA CCUGACCACCCGGACCCAGCUGCCAC CAGCCUACACCAACAGCUUCACCCGG GGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACC CAGGACCUGUUCCUGCCCUUCUUCAG CAACGUGACCUGGUUCCACGCCAUCC ACGUGAGCGGCACCAACGGCACCAA GCGGUUCGACAACCCCGUGCUGCCCU UCAACGACGGCGUGUACUUCGCCAG CACCGAGAAGAGCAACAUCAUCCGG GGCUGGAUCUUCGGCACCACCCUGG ACAGCAAGACCCAGAGCCUGCUGAU CGUGAAUAACGCCACCAACGUGGUG AUCAAGGUGUGCGAGUUCCAGUUCU GCAACGACCCCUUCCUGGGCGUGUAC UACCACAAGAACAACAAGAGCUGGA UGGAGAGCGAGUUCCGGGUGUACAG CAGCGCCAACAACUGCACCUUCGAGU	124

[0961]

	ACGUGAGCCAGCCCUUCCUGAUGGA CCUGGAGGGCAAGCAGGGCAACUUC AAGAACCUGCGGGAGUUCGUGUUCA AGAACAUCGACGGCUACUUCAAGAU CUACAGCAAGCACACCCCAAUCAACC UGGUGCGGGAUCUGCCCCAGGGCUU CUCAGCCCUGGAGCCCCUGGUGGACC UGCCCAUCGGCAUCAACAUCACCCGG UUCCAGACCCUGCUGGCCCUGCACCG GAGCUACCUGACCCCAGGCGACAGCA GCAGCGGGUGGACAGCAGGCGCGGC UGCUUACUACGUGGGCUACCUGCAG CCCCGGACCUUCCUGCUGAAGUACAA CGAGAACGGCACCAUCACCGACGCCG UGGACUGCGCCCUGGACCCUCUGAGC GAGACCAAGUGCACCCUGAAGAGCU UCACCGUGGAGAAGGGCAUCUACCA GACCAGCAACUUCCGGGUGCAGCCCA CCGAGAGCAUCGUGCGGUUCCCCAAC AUCACCAACCUGUGCCCCUUCGGCGA GGUGUUCAACGCCACCCGGUUCGCCA GCGUGUACGCCUGGAACCGGAAGCG GAUCAGCAACUGCGUGGCCGACUAC AGCGUGCUGUACAACAGCGCCAGCU UCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGU GCUUCACCAACGUGUACGCCGACAGC UUCGUGAUCCGUGGCGACGAGGUGC GGCAGAUCGCACCCGGCCAGACAGGC AAGAUCGCCGACUACAACUACAAGC UGCCCGACGACUUCACCGGCUGCGUG AUCGCCUGGAACAGCAACAACCUCG ACAGCAAGGUGGGCGGCAACUACAA CUACCUGUACCGGCUGUUCCGGAAG AGCAACCUGAAGCCCUUCGAGCGGG ACAUCAGCACCGAGAUCUACCAAGCC GGCUCCACCCCUUGCAACGGCGUGGA GGGCUUCAACUGCUACUUCCCUCUGC AGAGCUACGGCUUCCAGCCCACCAAC GGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCU GCACGCCCCAGCCACCGUGUGUGGCC CCAAGAAGAGCACCAACCUGGUGAA	

[0962]

	GAACAAGUGCGUGAACUUCAACUUC AACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUG CCCUUUCAGCAGUUCGGCCGGGACA UCGCCGACACCACCGACGCUGUGCGG GAUCCCCAGACCCUGGAGAUCCUGG ACAUCACCCCUUGCAGCUUCGGCGGC GUGAGCGUGAUCACCCCAGGCACCA ACACCAGCAACCAGGUGGCCGUGCU GUACCAGGACGUGAACUGCACCGAG GUGCCCGUGGCCAUCCACGCCGACCA GCUGACACCCACCUGGCGGGUCUACA GCACCGGCAGCAACGUGUUCCAGACC CGGGCCGGUUGCCUGAUCGGCGCCG AGCACGUGAACAACAGCUACGAGUG CGACAUCCCCAUCGGCGCCGGCAUCU GUGCCAGCUACCAGACCCAGACCAAU UCACCCCGGAGGGCAAGGAGCGUGG CCAGCCAGAGCAUCAUCGCCUACACC AUGAGCCUGGGCGCCGAGAACAGCG UGGCCUACAGCAACAACAGCAUCGCC AUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGC AUGACCAAGACCAGCGUGGACUGCA CCAUGUACAUCUGCGGCGACAGCACC GAGUGCAGCAACCUGCUGCUGCAGU ACGGCAGCUUCUGCACCCAGCUGAAC CGGGCCCUGACCGGCAUCGCCGUGGA GCAGGACAAGAACACCCAGGAGGUG UUCGCCCAGGUGAAGCAGAUCUACA AGACCCCUCCCAUCAAGGACUUCGGC GGCUUCAACUUCAGCCAGAUCCUGCC CGACCCCAGCAAGCCCAGCAAGCGGA GCUUCAUCGAGGACCUGCUGUUCAA CAAGGUGACCCUAGCCGACGCCGGCU UCAUCAAGCAGUACGGCGACUGCCU CGGCGACAUAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCU GACCGUGCUGCCUCCCCUGCUGACCG ACGAGAUGAUCGCCCAGUACACCAG CGCCCUGUUAGCCGGAACCAUCACCA GCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGC	

[0963]

	AGAUGGCCUACCGGUUCAACGGCAU CGGCGUGACCCAGAACGUGCUGUAC GAGAACCAGAAGCUGAUCGCCAACC AGUUCAACAGCGCCAUCGGCAAGAU CCAGGACAGCCUGAGCAGCACCGCUA GCGCCCUGGGCAAGCUGCAGGACGU GGUGAACCAGAACGCCCAGGCCCUG AACACCCUGGUGAAGCAGCUGAGCA GCAACUUCGGCGCCAUCAGCAGCGU GCUGAACGACAUCCUGAGCCGGCUG GACAAGGUGGAGGCCGAGGUGCAGA UCGACCGGCUGAUCACUGGCCGGCU GCAGAGCCUGCAGACCUACGUGACCC AGCAGCUGAUCCGGGCCGCCGAGAU UCGGGCCAGCGCCAACCUGGCCGCCA CCAAGAUGAGCGAGUGCGUGCUGGG CCAGAGCAAGCGGGUGGACUUCUGC GGCAAGGGCUACCACCUGAUGAGCU UUCCCCAGAGCGCACCCCACGGAGUG GUGUUCCUGCACGUGACCUACGUGC CCGCCCAGGAGAAGAACUUCACCACC GCCCCAGCCAUCUGCCACGACGGCAA GGCCCACUUUCCCCGGGAGGGCGUG UUCGUGAGCAACGGCACCCACUGGU UCGUGACCCAGCGGAACUUCUACGA GCCCCAGAUCAUCACCACCGACAACA CCUUCGUGAGCGGCAACUGCGACGU GGUGAUCGGCAUCGUGAACAACACC GUGUACGAUCCCCUGCAGCCCGAGCU GGACAGCUUCAAGGAGGAGCUGGAC AAGUACUUCAAGAAUCACACCAGCC CCGACGUGGACCUGGGCGACAUCAG CGGCAUCAACGCCAGCGUGGUGAAC AUCCAGAAGGAGAUCGAUCGGCUGA ACGAGGUGGCCAAGAACCUGAACGA GAGCCUGAUCGACCUGCAGGAGCUG GGCAAGUACGAGCAGUACAUCAAGU GGCCCUGGUACAUCUGGCUGGGCUU CAUCGCCGGCCUGAUCGCCAUCGUGA UGGUGACCAUCAUGCUGUGCUGCAU GACCAGCUGCUGCAGCUGCCUGAAG GGCUGUUGCAGCUGCGGCAGCUGCU GCAAGUUCGACGAGGACGACAGCGA	

[0964]

	GCCCGUGCUGAAGGGCGUGAAGCUG CACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCCUUCCUGCA CCCGUACCCCCGUGGUCUUUGAAUA AAGUCUGAGUGGGCGGC	4
相应的氨基酸序列	MFVFLVLLPLVSSQCVNLTTRTQLPPAY TNSFTRGVYYPDKVFRSSVLHSTQDLFL PFFSNVTWFHAIHVSGTNGTKRFDNPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLVRDLPQGFSA LEPLVDLPIGINITRFQTLLALHRSYLTP GDSSSGWTAGAAAYYVGYLQPRTFLL KYNENGTITDAVDCALDPLSETKCTLKS FTVEKGIYQTSNFRVQPTESIVRFPNITN LCPFGEVFNATRFASVYAWNRKRISNC VADYSVLYNSASFSTFKCYGVSPTKLN DLCFTNVYADSFVIRGDEVRQIAPGQTG KIADYNYKLPDDFTGCVIAWNSNNLDS KVGGNYNYLYRLFRKSNLKPFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGFQP TNGVGYQPYRVVVLSFELLHAPATVCG PKKSTNLVKNKCVNFNFNGLTGTGVLT ESNKKFLPFQQFGRDIADTTDAVRDPQT LEILDITPCSFGGVSVITPGTNTSNQVAV LYQDVNCTEVPVAIHADQLTPTWRVYS TGSNVFQTRAGCLIGAEHVNNSYECDIP IGAGICASYQTQTNSPRRARSVASQSIIA YTMSLGAENSVAYSNNSIAIPTNFTISVT TEILPVSMTKTSVDCTMYICGDSTECSN LLLQYGSFCTQLNRALTGIAVEQDKNT QEVFAQVKQIYKTPPIKDFGGFNFSQILP DPSKPSKRSFIEDLLFNKVTLADAGFIKQ YGDCLGDIAARDLICAQKFNGLTVLPPL LTDEMIAQYTSALLAGTITSGWTFGAG AALQIPFAMQMAYRFNGIGVTQNVLYE NQKLIANQFNSAIGKIQDSLSSTASALG KLQDVVNQNAQALNTLVKQLSSNFGAI SSVLNDILSRLDKVEAEVQIDRLITGRLQ	125

[0965]

	SLQTYVTQQLIRAAEIRASANLAATKMS ECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHD GKAHFPREGVFVSNGTHWFVTQRNFYE PQIITTDNTFVSGNCDVVIGIVNNTVYDP LQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNE SLIDLQELGKYEQYIKWPWYIWLGFIAG LIAIVMVTIMLCCMTSCCSCLKGCCSCG SCCKFDEDDSEPVLKGVKLHYT	
多聚A尾	100 nt	

[0966] 还应理解,本文所述的开放阅读框和/或相应氨基酸序列中的任一者可包括或排除信号序列。

[0967] 本文公开的所有参考文献、专利和专利申请都关于其各自所列举的主题通过引用并入,其在一些情形下可涵盖整个文件。

[0968] 除非明确指示相反情形,否则如本文在说明书和在权利要求书中所用的不定冠词“一个”和“一种”应理解为意指“至少一个/种”。还应理解,除非明确指示相反情形,否则在本文所要求的包括一个以上步骤或动作的任何方法中,所述方法的步骤或动作的顺序不一定受限于列举所述方法的步骤或动作的顺序。

[0969] 在权利要求书中以及在上文说明书中,所有过渡性短语(例如“包含”、“包括”、“携载”、“具有”、“含有”、“涉及”、“持有”、“由.....构成”等)应理解为是开放式的,即,意指包括但不限于。如美国专利局专利审查程序手册(United States Patent Office Manual of Patent Examining Procedures)第2111.03节所述,仅过渡性短语“由.....组成”和“基本上由.....组成”应分别为封闭式或半封闭式过渡性短语。

[0970] 数值前的术语“约”和“基本上”意指所列举的数值的±10%。

[0971] 在提供值的范围的情况下,本文具体考虑并阐述所述范围的上端和下端之间的每个值。

[0972] 国际申请号PCT/US2015/02740、PCT/US2016/043348、PCT/US2016/043332、PCT/US2016/058327、PCT/US2016/058324、PCT/US2016/058314、PCT/US2016/058310、PCT/US2016/058321、PCT/US2016/058297、PCT/US2016/058319和PCT/US2016/058314的整个内容通过引用并入本文。

序列表

<110> 摩登纳特斯有限公司(ModernaTX, Inc.)

<120> SARS-COV-2 mRNA结构域疫苗

<130> M1378.70157W000

<140> 尚未分配
<141> 同时随同提交

<150> US 63/063,137
<151> 2020-08-07

<150> US 63/044,330
<151> 2020-06-25

<150> US 63/016,175
<151> 2020-04-27

<150> US 62/971,825
<151> 2020-02-07

<160> 152

<170> PatentIn version 3.5

<210> 1
[0001] <211> 2216
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 1
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

[0002] ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

cuggagaucc uggacaucac cccuugcagc uucggcggcg ugagcgugau caccccaggc 1860

accaacacca gcaaccaggu ggccgugcug uaccaggacg ugaacugcac cgaggugccc 1920

guggccaucc acgccgacca gcugacaccc accuggcggg ucuacagcac cggcagcaac 1980

guguuccaga cccgggccgg uugccugauc ggcgccgagc acgugaacaa cagcuacgag 2040

ugcgacaucc ccaucggcgc cggcaucugu gccagcuacc agacccagac caauucauga 2100

uaauaggcug gagccucggu ggccuagcuu cuugccccuu gggccucccc ccagccccuc 2160

cuccccuucc ugcacccgua cccccguggu cuuugaauaa agucugagug ggcggc 2216

<210> 2
<211> 57
<212> RNA

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 2
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccacc 57

<210> 3
<211> 2040
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 3
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

[0003] auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agcccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau caccccuugc agcuucggcg gcgugagcgu gaucacccca 1800

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 1860

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 1920

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 1980

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 2040

[0004]

<210> 4
<211> 119
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 4
ugauaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60

cuccuccccu uccugcaccc guacccccgu ggucuuugaa uaaagucuga gugggcggc 119

<210> 5
<211> 680
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 5

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

[0005]

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

[0006] Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys

465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575

[0007] Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser
675 680

<210> 6
<211> 2729

<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 6

[0008]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

cuggagaucc uggacaucac cccuugcagc uucggcggcg ugagcgugau caccccaggc 1860

accaacacca gcaaccaggu ggccgugcug uaccaggacg ugaacugcac cgaggugccc 1920

guggccaucc acgccgacca gcugacaccc accuggcggg ucuacagcac cggcagcaac 1980

guguuccaga cccgggccgg uugccugauc ggcgccgagc acgugaacaa cagcuacgag 2040

ugcgacaucc ccaucggcgc cggcaucugu gccagcuacc agacccagac caauucagga 2100

ggaggcagcg gcggcgauau caucaagcuu cugaacgagc aaguuaacaa ggaaaugcag 2160

agcaguaauc ucuacaugag caugagcagc uggugcuaca cccacucccu ggacggagca 2220

ggccucuucc uguucgacca cgcagccgag gaguacgagc acgcuaagaa guugaucauu 2280

uucuugaacg agaacaacgu gcccgugcag cuaacgucaa ucagcgcacc ugagcacaag 2340

[0009] uucgagggcc ugacccagau cuuccagaag gccuacgaac acgaacagca caucuccgag 2400

agcaucaaca auauugugga ucacgcuauc aaguccaagg accacgcuac cuucaacuuc 2460

cugcaguggu acguggccga gcaacaugag gaggaggugc uguucaagga cauccuggac 2520

aagaucgagc ugaucgguaa ugagaaucac ggccuguacc uggccgacca guacgugaag 2580

ggcaucgcca agagccggaa gucaggcuca ugauaauagg cuggagccuc gguggccuag 2640

cuucuugccc cuugggccuc cccccagccc cuccuccccu uccugcaccc guacccccgu 2700

ggucuuugaa uaaagucuga gugggcggc 2729

<210> 7
<211> 2553
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 7
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

[0010]

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca agguggggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau cacccccuugc agcuucggcg gcgugagcgu gaucacccca 1800

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 1860

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 1920

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 1980

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 2040

ggaggaggca gcggcggcga uaucaucaag cuucugaacg agcaaguuaa caaggaaaug 2100

cagagcagua aucucuacau gagcaugagc agcuggugcu acacccacuc ccuggacgga 2160

gcaggccucu uccuguucga ccacgcagcc gaggaguacg agcacgcuaa gaaguugauc 2220

auuuucuuga acgagaacaa cgugcccgug cagcuaacgu caaucagcgc accugagcac 2280

aaguucgagg gccugaccca gaucuuccag aaggccuacg aacacgaaca gcacaucucc 2340

gagagcauca acaauauugu ggaucacgcu aucaagucca aggaccacgc uaccuucaac 2400

uuccugcagu gguacguggc cgagcaacau gaggaggagg ugcuguucaa ggacauccug 2460

gacaagaucg agcugaucgg uaaugagaau cacggccugu accuggccga ccaguacgug 2520

aagggcaucg ccaagagccg gaagucaggc uca 2553

<210> 8
<211> 851
<212> PRT
[0011] <213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 8

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

[0012]

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

[0013]

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

[0014] Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser Gly Gly Gly Ser Gly Gly Asp Ile
675 680 685

Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met Gln Ser Ser Asn
690 695 700

Leu Tyr Met Ser Met Ser Ser Trp Cys Tyr Thr His Ser Leu Asp Gly
705 710 715 720

Ala Gly Leu Phe Leu Phe Asp His Ala Ala Glu Glu Tyr Glu His Ala
725 730 735

Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val Pro Val Gln Leu
740 745 750

Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly Leu Thr Gln Ile
755 760 765

Phe Gln Lys Ala Tyr Glu His Glu Gln His Ile Ser Glu Ser Ile Asn

770 775 780

Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His Ala Thr Phe Asn
785 790 795 800

Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu Glu Val Leu Phe
805 810 815

Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu Asn His Gly
820 825 830

Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys Ser Arg Lys
835 840 845

Ser Gly Ser
850

<210> 9
<211> 2762
<212> RNA
<213> 人工序列(Artificial Sequence)

[0015] <220>
<223> 合成

<400> 9
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

ggcauccugc ccagcccugg caugcccgcu cugcugagcc uggugagccu gcugagcgug 120

cugcugaugg gcugcguggc ugagaccggc augcagaucu acgagggcaa gcugaccgca 180

gagggccugc gguucggcau cguggccagc cgcgccaacc acgcucuggu ggaccggcuu 240

guggagggcg cuaucgacgc caucgugaga cacggcggcc gggaagagga caucacccug 300

gugcgggugu gcggcagcug ggagauuccc gucgccgccg gagaacuggc ccggaaggag 360

gacaucgacg ccgugaucgc caucggcgug cugugcagag gcgccacgcc cagcuucgac 420

uacaucgcca gcgaggugag caagggccug gccgaccuga gccuggagcu gcggaagccc 480

aucaccuucg gcgugaucac cgccgacacc cuggagcagg ccaucgaggc cgcaggcacc 540

ugccacggca acaagggcug ggaagccgcc cugugcgcca ucgagauggc caaccuguuc 600

aagagccugc ggggcggaag uggaggcucu gguggcagcg gaggaucugg cggcggcacc 660

acccggaccc agcugccacc agccuacacc aacagcuuca cccggggcgu cuacuacccc 720

gacaaggugu uccggagcag cguccugcac agcacccagg accuguuccu gcccuucuuc 780

[0016]

agcaacguga ccugguucca cgccauccac gugagcggca ccaacggcac caagcgguuc 840

gacaaccccg ugcugcccuu caacgacggc guguacuucg ccagcaccga gaagagcaac 900

aucauccggg gcuggaucuu cggcaccacc cuggacagca agacccagag ccugcugauc 960

gugaauaacg ccaccaacgu ggugaucaag gugugcgagu uccaguucug caacgacccc 1020

uuccugggcg uguacuacca caagaacaac aagagcugga uggagagcga guuccgggug 1080

uacagcagcg ccaacaacug caccuucgag uacgugagcc agcccuuccu gauggaccug 1140

gagggcaagc agggcaacuu caagaaccug cgggaguucg uguucaagaa caucgacggc 1200

uacuucaaga ucuacagcaa gcacacccca aucaaccugg ugcgggaucu gccccagggc 1260

uucucagccc uggagccccu gguggaccug cccaucggca ucaacaucac ccgguuccag 1320

acccugcugg cccugcaccg gagcuaccug accccaggcg acagcagcag cggguggaca 1380

gcaggcgcgg cugcuuacua cgugggcuac cugcagcccc ggaccuuccu gcugaaguac 1440

aacgagaacg gcaccaucac cgacgccgug gacugcgccc uggacccucu gagcgagacc 1500

aagugcaccc ugaagagcuu caccguggag aagggcaucu accagaccag caacuuccgg 1560

gugcagccca ccgagagcau cgugcgguuc cccaacauca ccaaccugug ccccuucggc 1620

gagguguuca acgccacccg guucgccagc guguacgccu ggaaccggaa gcggaucagc 1680

aacugcgugg ccgacuacag cgugcuguac aacagcgcca gcuucagcac cuucaagugc 1740

uacggcguga gccccaccaa gcugaacgac cugugcuuca ccaacgugua cgccgacagc 1800

uucgugaucc guggcgacga ggugcggcag aucgcacccg gccagacagg caagaucgcc 1860

gacuacaacu acaagcugcc cgacgacuuc accggcugcg ugaucgccug gaacagcaac 1920

aaccucgaca gcaagguggg cggcaacuac aacuaccugu accggcuguu ccggaagagc 1980

aaccugaagc ccuucgagcg ggacaucagc accgagaucu accaagccgg cuccaccccu 2040

ugcaacggcg uggagggcuu caacugcuac uuccccucugc agagcuacgg cuuccagccc 2100

accaacggcg ugggcuacca gcccuaccgg guggugguge ugagcuucga gcugcugcac 2160

gccccagcca ccgugugugg ccccaagaag agcaccaacc uggugaagaa caagugcgug 2220

aacuucaacu ucaacggccu uaccggcacc ggcgugcuga ccgagagcaa caagaaauuc 2280

cugcccuuuc agcaguucgg ccgggacauc gccgacacca ccgacgcugu gcgggauccc 2340

cagacccugg agauccugga caucaccccu ugcagcuucg gcggcgugag cgugaucacc 2400

ccaggcacca acaccagcaa ccagguggcc gugcuguacc aggacgugaa cugcaccgag 2460

gugcccgugg ccauccacgc cgaccagcug acacccaccu ggcgggucua cagcaccggc 2520

agcaacgugu uccagacccg ggccgguugc cugaucggcg ccgagcacgu gaacaacagc 2580

uacgagugcg acaucccccau cggcgccggc aucugugcca gcuaccagac ccagaccaau 2640

ucaugauaau aggcuggagc cucgguggcc uagcuucuug ccccuugggc cucccccccag 2700

ccccuccucc ccuuccugca cccguacccc cguggucuuu gaauaaaguc ugagugggcg 2760

gc 2762

<210> 10
<211> 2586
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 10
augggcaucc ugcccagccc uggcaugccc gcucugcuga gccuggugag ccugcugagc 60

gugcugcuga ugggcugcgu ggcugagacc ggcaugcaga ucuacgaggg caagcugacc 120

gcagagggcc ugcgguucgg caucguggcc agccgcgcca accacgcucu gguggaccgg 180

[0017] cuuguggagg gcgcuaucga cgccaucgug agacacggcg gccgggaaga ggacaucacc 240

cuggugcggg ugugcggcag cugggagauu cccgucgccg ccggagaacu ggcccggaag 300

gaggacaucg acgccgugau cgccaucggc gugcugugca gaggcgccac gcccagcuuc 360

gacuacaucg ccagcgaggu gagcaagggc cuggccgacc ugagccugga gcugcggaag 420

cccaucaccu ucggcgugau caccgccgac acccuggagc aggccaucga ggccgcaggc 480

accugccacg gcaacaaggg cugggaagcc gcccugugcg ccaucgagau ggccaaccug 540

uucaagagcc ugcggggcgg aaguggaggc ucugguggca gcggaggauc uggcggcggc 600

accacccgga cccagcugcc accagccuac accaacagcu ucacccgggg cgucuacuac 660

cccgacaagg uguuccggag cagcguccug cacagcaccc aggaccuguu ccugcccuuc 720

uucagcaacg ugaccugguu ccacgccauc cacgugagcg gcaccaacgg caccaagcgg 780

uucgacaacc ccgugcugcc cuucaacgac ggcguguacu ucgccagcac cgagaagagc 840

aacaucaucc ggggcuggau cuucggcacc acccuggaca gcaagaccca gagccugcug 900

aucgugaaua acgccaccaa cguggugauc aaggugugcg aguuccaguu cugcaacgac 960

cccuuccugg gcguguacua ccacaagaac aacaagagcu ggauggagag cgaguuccgg 1020

guguacagca gcgccaacaa cugcaccuuc gaguacguga gccagcccuu ccugauggac 1080

cuggagggca agcagggcaa cuucaagaac cugcgggagu ucguguucaa gaacaucgac 1140

ggcuacuuca agaucuacag caagcacacc ccaaucaacc uggugcggga ucugccccag 1200

ggcuucucag cccuggagcc ccugguggac cugcccaucg gcaucaacau cacccgguuc 1260

cagacccugc uggcccugca ccggagcuac cugaccccag gcgacagcag cagcgggugg 1320

acagcaggcg cggcugcuua cuacgugggc uaccugcagc cccggaccuu ccugcugaag 1380

uacaacgaga acggcaccau caccgacgcc guggacugcg cccuggaccc ucugagcgag 1440

accaagugca cccugaagag cuucaccgug gagaagggca ucuaccagac cagcaacuuc 1500

cgggugcagc ccaccgagag caucgugcgg uuccccaaca ucaccaaccu gugccccuuc 1560

ggcgaggugu ucaacgccac ccgguucgcc agcguguacg ccuggaaccg gaagcggauc 1620

agcaacugcg uggccgacua cagcgugcug uacaacagcg ccagcuucag caccuucaag 1680

ugcuacggcg ugagccccac caagcugaac gaccugugcu ucaccaacgu guacgccgac 1740

agcuucguga uccguggcga cgaggugcgg cagaucgcac ccggccagac aggcaagauc 1800

gccgacuaca acuacaagcu gcccgacgac uucaccggcu gcgugaucgc cuggaacagc 1860

aacaaccucg acagcaaggu gggcggcaac uacaacuacc uguaccggcu guuccggaag 1920

[0018] agcaaccuga agcccuucga gcgggacauc agcaccgaga ucuaccaagc cggcuccacc 1980

ccuugcaacg gcguggaggg cuucaacugc uacuuccccuc ugcagagcua cggcuuccag 2040

cccaccaacg gcgugggcua ccagcccuac cggguggugg ugcugagcuu cgagcugcug 2100

cacgccccag ccaccgugug uggccccaag aagagcacca accuggugaa gaacaagugc 2160

gugaacuuca acuucaacgg ccuuaccggc accggcgugc ugaccgagag caacaagaaa 2220

uuccugcccu uucagcaguu cggccgggac aucgccgaca ccaccgacgc ugugcgggau 2280

ccccagaccc uggagauccu ggacaucacc ccuugcagcu ucggcggcgu gagcgugauc 2340

accccaggca ccaacaccag caaccaggug gccgugcugu accaggacgu gaacugcacc 2400

gaggugcccg uggccaucca cgccgaccag cugacaccca ccuggcgggu cuacagcacc 2460

ggcagcaacg uguuccagac ccgggccggu ugccugaucg gcgccgagca cgugaacaac 2520

agcuacgagu gcgacauccc caucggcgcc ggcaucugug ccagcuacca gacccagacc 2580

aauuca 2586

<210> 11
<211> 862
<212> PRT

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 11

Met Gly Ile Leu Pro Ser Pro Gly Met Pro Ala Leu Leu Ser Leu Val
1 5 10 15

Ser Leu Leu Ser Val Leu Leu Met Gly Cys Val Ala Glu Thr Gly Met
20 25 30

Gln Ile Tyr Glu Gly Lys Leu Thr Ala Glu Gly Leu Arg Phe Gly Ile
35 40 45

Val Ala Ser Arg Ala Asn His Ala Leu Val Asp Arg Leu Val Glu Gly
50 55 60

Ala Ile Asp Ala Ile Val Arg His Gly Gly Arg Glu Glu Asp Ile Thr
65 70 75 80

[0019] Leu Val Arg Val Cys Gly Ser Trp Glu Ile Pro Val Ala Ala Gly Glu
85 90 95

Leu Ala Arg Lys Glu Asp Ile Asp Ala Val Ile Ala Ile Gly Val Leu
100 105 110

Cys Arg Gly Ala Thr Pro Ser Phe Asp Tyr Ile Ala Ser Glu Val Ser
115 120 125

Lys Gly Leu Ala Asp Leu Ser Leu Glu Leu Arg Lys Pro Ile Thr Phe
130 135 140

Gly Val Ile Thr Ala Asp Thr Leu Glu Gln Ala Ile Glu Ala Ala Gly
145 150 155 160

Thr Cys His Gly Asn Lys Gly Trp Glu Ala Ala Leu Cys Ala Ile Glu
165 170 175

Met Ala Asn Leu Phe Lys Ser Leu Arg Gly Gly Ser Gly Gly Ser Gly
180 185 190

Gly Ser Gly Gly Ser Gly Gly Gly Thr Thr Arg Thr Gln Leu Pro Pro
195 200 205

Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val
210 215 220

Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe Leu Pro Phe
225 230 235 240

Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser Gly Thr Asn
245 250 255

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

Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe
275 280 285

Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile Val Asn Asn
290 295 300

Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe Cys Asn Asp
305 310 315 320

[0020]

Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser Trp Met Glu
325 330 335

Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr
340 345 350

Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln Gly Asn Phe
355 360 365

Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys
370 375 380

Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp Leu Pro Gln
385 390 395 400

Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile Gly Ile Asn
405 410 415

Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser Tyr Leu Thr
420 425 430

Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr
435 440 445

Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn
450 455 460

Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro Leu Ser Glu
465 470 475 480

Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln
485 490 495

Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val Arg Phe Pro
500 505 510

Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg
515 520 525

Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val
530 535 540

[0021]

Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys
545 550 555 560

Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn
565 570 575

Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile
580 585 590

Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro
595 600 605

Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp
610 615 620

Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys
625 630 635 640

Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln
645 650 655

Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe
660 665 670

Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln
675 680 685

Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala
690 695 700

Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys Asn Lys Cys
705 710 715 720

Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Glu
725 730 735

Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala
740 745 750

Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp
755 760 765

[0022] Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr
770 775 780

Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr
785 790 795 800

Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg
805 810 815

Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu
820 825 830

Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile
835 840 845

Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser
850 855 860

<210> 12
<211> 2714
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 12

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

[0023] gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

cuggagaucc uggacaucac cccuugcagc uucggcggcg ugagcgugau caccccaggc 1860

accaacacca gcaaccaggu ggccgugcug uaccaggacg ugaacugcac cgaggugccc 1920

guggccaucc acgccgacca gcugacaccc accuggcggg ucuacagcac cggcagcaac 1980

guguuccaga cccgggccgg uugccugauc ggcgccgagc acgugaacaa cagcuacgag 2040

ugcgacaucc ccaucggcgc cggcaucugu gccagcuacc agacccagac caauucagga 2100

ggaggcuccg gaggcgguag cgcugagacc ggcaugcaga ucuacgaggg caagcugacc 2160

gcagagggcc ugcgguucgg caucguggcc agccgcgcca accacgcucu gguggaccgg 2220

cuuguggagg gcgcuaucga cgccaucgug agacacggcg gccgggaaga ggacaucacc 2280

cuggugcggg ugugcggcag cugggagauu cccgucgccg ccggagaacu ggcccggaag 2340

gaggacaucg acgccgugau cgccaucggc gugcugugca gaggcgccac gcccagcuuc 2400

gacuacaucg ccagcgaggu gagcaagggc cuggccgacc ugagccugga gcugcggaag 2460

[0024]

cccaucaccu ucggcgugau caccgccgac acccuggagc aggccaucga ggccgcaggc 2520

accugccacg gcaacaaggg cugggaagcc gcccugugcg ccaucgagau ggccaaccug 2580

uucaagagcc ugcggugaua auaggcugga gccucggugg ccuagcuucu ugccccuugg 2640

gccucccccc agccccuccu ccccuuccug cacccguacc cccguggucu uugaauaaag 2700

ucugaguggg cggc 2714

<210> 13
<211> 2538
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 13
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

[0025] ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca agguggggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau cacccccuugc agcuucggcg gcgugagcgu gaucacccca 1800

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 1860

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 1920

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 1980

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 2040

ggaggaggcu ccggaggcgg uagcgcugag accggcaugc agaucuacga gggcaagcug 2100

accgcagagg gccugcgguu cggcaucgug gccagccgcg ccaaccacgc ucugguggac 2160

cggcuugugg agggcgcuau cgacgccauc gugagacacg gcggccggga agaggacauc 2220

acccuggugc gggugugcgg cagcugggag auucccgucg ccgccggaga acuggcccgg 2280

aaggaggaca ucgacgccgu gaucgccauc ggcgugcugu gcagaggcgc cacgcccagc 2340

uucgacuaca ucgccagcga ggugagcaag ggccuggccg accugagccu ggagcugcgg 2400

aagcccauca ccuucggcgu gaucaccgcc gacacccugg agcaggccau cgaggccgca 2460

ggcaccugcc acggcaacaa gggcugggaa gccgcccugu gcgccaucga gauggccaac 2520

cuguucaaga gccugcgg 2538

<210> 14
<211> 846
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0026]

<400> 14

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser

100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

[0027] Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

[0028]

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

[0029]

Ser Tyr Gln Thr Gln Thr Asn Ser Gly Gly Gly Ser Gly Gly Gly Ser
675 680 685

Ala Glu Thr Gly Met Gln Ile Tyr Glu Gly Lys Leu Thr Ala Glu Gly
690 695 700

Leu Arg Phe Gly Ile Val Ala Ser Arg Ala Asn His Ala Leu Val Asp
705 710 715 720

Arg Leu Val Glu Gly Ala Ile Asp Ala Ile Val Arg His Gly Gly Arg
725 730 735

Glu Glu Asp Ile Thr Leu Val Arg Val Cys Gly Ser Trp Glu Ile Pro
740 745 750

Val Ala Ala Gly Glu Leu Ala Arg Lys Glu Asp Ile Asp Ala Val Ile
755 760 765

Ala Ile Gly Val Leu Cys Arg Gly Ala Thr Pro Ser Phe Asp Tyr Ile
770 775 780

Ala Ser Glu Val Ser Lys Gly Leu Ala Asp Leu Ser Leu Glu Leu Arg
785 790 795 800

Lys Pro Ile Thr Phe Gly Val Ile Thr Ala Asp Thr Leu Glu Gln Ala
805 810 815

Ile Glu Ala Ala Gly Thr Cys His Gly Asn Lys Gly Trp Glu Ala Ala
820 825 830

Leu Cys Ala Ile Glu Met Ala Asn Leu Phe Lys Ser Leu Arg
835 840 845

<210> 15
<211> 2300
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 15

[0030]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

[0031] cuggagaucc uggacaucac cccuugcagc uucggcggcg ugagcgugau caccccaggc 1860

accaacacca gcaaccaggu ggccgugcug uaccaggacg ugaacugcac cgaggugccc 1920

guggccaucc acgccgacca gcugacaccc accuggcggg ucuacagcac cggcagcaac 1980

guguuccaga cccgggccgg uugccugauc ggcgccgagc acgugaacaa cagcuacgag 2040

ugcgacaucc ccaucggcgc cggcaucugu gccagcuacc agacccagac caauucaucu 2100

ggcggaggca gcauccuggc caucuacagc accguggcca gcagccuggu gcugcuggug 2160

agccugggcg ccaucagcuu cugauaauag gcuggagccu cgguggccua gcuucuugcc 2220

ccuugggccu ccccccagcc ccuccucccc uuccugcacc cguacccccg uggucuuuga 2280

auaaagucug agugggcggc 2300

<210> 16
<211> 2124
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 16
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

[0032]

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau caccccuugc agcuucggcg gcgugagcgu gaucacccca 1800

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 1860

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 1920

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 1980

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 2040

ucuggcggag gcagcauccu ggccaucuac agcaccgugg ccagcagccu ggugcugcug 2100

gugagccugg gcgccaucag cuuc 2124

<210> 17
<211> 707
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 17

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

[0033]

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

[0034] Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu

355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
[0035] 465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser Gly Gly Gly Ser Ile Leu Ala Ile
675 680 685

Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala
690 695 700

[0036]

Ile Ser Phe
705

<210> 18
<211> 1853
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 18
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccccgguu ccagaccccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gccccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

[0037] auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc ucuggcggag gcagcauccu ggccaucuac 1680

agcaccgugg ccagcagccu ggugcugcug gugagccugg gcgccaucag cuucugauaa 1740

uaggcuggag ccucgguggc cuagcuucuu gccccuuggg ccucccccca gccccuccuc 1800

cccuuccugc acccguaccc ccguggucuu ugaauaaagu cugaguggge ggc 1853

<210> 19
<211> 1677
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 19

auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

[0038] ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgcccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accucuggcg gaggcagcau ccuggccauc 1620

uacagcaccg uggccagcag ccuggugcug cuggugagcc ugggcgccau cagcuuc 1677

<210> 20
<211> 559
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 20

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

[0039]
Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

[0040]

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

[0041]

Lys Ser Thr Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val
530 535 540

Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
545 550 555

<210> 21
<211> 2042
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 21
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

[0042]

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

cuggagaucc uggacaucac cccuugcagc uucggcggcu cuggcggagg cagcauccug 1860

gccaucuaca gcaccguggc cagcagccug gugcugcugg ugagccuggg cgccaucagc 1920

uucugauaau aggcuggagc cucgguggcc uagcuucuug ccccuugggc cucccccccag 1980

ccccuccucc ccuuccugca cccguacccc cguggucuuu gaauaaaguc ugagugggcg 2040

gc 2042

<210> 22
<211> 1866
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 22
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

[0043] agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau caccccuugc agcuucggcg gcucuggcgg aggcagcauc 1800

cuggccaucu acagcaccgu ggccagcagc cuggugcugc uggugagccu gggcgccauc 1860

agcuuc 1866

<210> 23
<211> 622
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0044] <400> 23

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

[0045]

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

[0046] Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val

565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala
595 600 605

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
610 615 620

<210> 24
<211> 2066
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 24
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

[0047] acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu gggcagcggc ggcggcagcg gcggaggcag cggaggaggc 1020

agcggcggag gcaguggagg ccagcccacc gagagcaucg ugcgguuccc caacaucacc 1080

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1140

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1200

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1260

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1320

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1380

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1440

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1500

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1560

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1620

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagaagag caccaaccug 1680

gugaagaaca agugcgugaa cuucaacuuc aacggccuua ccggcaccgg cgugcugacc 1740

gagagcaaca agaaaauuccu gcccuuucag caguucggcc gggacaucgc cgacaccacc 1800

gacgcugugc gggaucccca gacccuggag auccuggaca ucaccccuug cagcuucggc 1860

[0048]

ggcucuggcg gaggcagcau ccuggccauc uacagcaccg uggccagcag ccuggugcug 1920

cuggugagcc ugggcgccau cagcuucuga uaauaggcug gagccucggu ggccuagcuu 1980

cuugccccuu gggccucccc ccagccccuc cuccccuucc ugcacccgua cccccguggu 2040

cuuugaauaa agucugagug ggcggc 2066

<210> 25
<211> 1890
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 25
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cgugggcagc ggcggcggca gcggcggagg cagcggagga 960

ggcagcggcg gaggcagugg aggccagccc accgagagca ucgugcgguu ccccaacauc 1020

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 1080

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1140

[0049] agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1200

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1260

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1320

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1380

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1440

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1500

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1560

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaagaa gagcaccaac 1620

cuggugaaga acaagugcgu gaacuucaac uucaacggcc uuaccggcac cggcgugcug 1680

accgagagca acaagaaauu ccugcccuuu cagcaguucg gccgggacau cgccgacacc 1740

accgacgcug ugcgggaucc ccagacccug gagauccugg acaucacccc uugcagcuuc 1800

ggcggcucug gcggaggcag cauccuggcc aucuacagca ccguggccag cagccuggug 1860

cugcugguga gccugggcgc caucagcuuc 1890

<210> 26
<211> 630

<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 26

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

[0050] Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr

195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

[0051] Ser Phe Thr Val Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly
305 310 315 320

Gly Ser Gly Gly Gly Ser Gly Gly Gln Pro Thr Glu Ser Ile Val Arg
325 330 335

Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala
340 345 350

Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn
355 360 365

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

Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe
385 390 395 400

Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg
405 410 415

Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys
420 425 430

Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn
435 440 445

Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe
450 455 460

Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile
465 470 475 480

Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys
485 490 495

Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly
500 505 510

Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala
515 520 525

Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys Asn
530 535 540

[0052]

Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu
545 550 555 560

Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg Asp
565 570 575

Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu Ile
580 585 590

Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Ser Gly Gly Gly Ser Ile
595 600 605

Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
610 615 620

Leu Gly Ala Ile Ser Phe
625 630

<210> 27
<211> 2066
<212> RNA

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 27

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccaccccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

gugugccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

[0053] aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu gggcagcggc ggcggcagcg gcggaggcag cggaggaggc 1020

agcggcggag gcaguggagg ccagcccacc gagagcaucg ugcgguuccc caacaucacc 1080

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1140

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1200

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1260

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1320

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1380

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1440

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1500

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1560

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1620

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagaagag caccaaccug 1680

gugaagaaca agugcgugaa cuucaacuuc aacggccuua ccggcaccgg cgugcugacc 1740

gagagcaaca agaaauuccu gcccuuuugc caguucggcc gggacaucgc cgacaccacc 1800

gacgcugugc gggaucccca gacccuggag auccuggaca ucaccccuug cagcuucggc 1860

ggcucuggcg gaggcagcau ccuggccauc uacagcaccg uggccagcag ccuggugcug 1920

cuggugagcc ugggcgccau cagcuucuga uaauaggcug gagccucggu ggccuagcuu 1980

cuugccccuu gggccucccc ccagccccuc cuccccuucc ugcacccgua cccccguggu 2040

cuuugaauaa agucugagug ggcggc 2066

<210> 28
<211> 1890
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0054] <400> 28
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aaggugugcc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgcccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cgugggcagc ggcggcggca gcggcggagg cagcggagga 960

ggcagcggcg gaggcagugg aggccagccc accgagagca ucgugcgguu ccccaacauc 1020

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 1080

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1140

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1200

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1260

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1320

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1380

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1440

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1500

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1560

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaagaa gagcaccaac 1620

cuggugaaga acaagugcgu gaacuucaac uucaacggcc uuaccggcac cggcgugcug 1680

[0055] accgagagca acaagaaauu ccugcccuuu ugccaguucg gccgggacau cgccgacacc 1740

accgacgcug ugcgggaucc ccagacccug gagauccugg acaucacccc uugcagcuuc 1800

ggcggcucug gcggaggcag cauccuggcc aucuacagca ccguggccag cagccuggug 1860

cugcugguga gccugggcgc caucagcuuc 1890

<210> 29
<211> 630
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 29

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Cys Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

[0056]

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly
305 310 315 320

Gly Ser Gly Gly Gly Ser Gly Gly Gln Pro Thr Glu Ser Ile Val Arg
325 330 335

Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala
340 345 350

Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn
355 360 365

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

[0057]

Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe
385 390 395 400

Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg
405 410 415

Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys
420 425 430

Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn
435 440 445

Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe
450 455 460

Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile
465 470 475 480

Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys
485 490 495

Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly
500 505 510

Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala
515 520 525

Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys Asn
530 535 540

Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu
545 550 555 560

Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Cys Gln Phe Gly Arg Asp
565 570 575

Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu Ile
580 585 590

Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Ser Gly Gly Gly Ser Ile
595 600 605

[0058]

Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
610 615 620

Leu Gly Ala Ile Ser Phe
625 630

<210> 30
<211> 3182
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 30
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gccccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu cggcggcagc 1020

ggcggcguga gcgugaucac cccaggcacc aacaccagca accagguggc cgugcuguac 1080

caggacguga acugcaccga ggugcccgug gccauccacg ccgaccagcu gacacccacc 1140

uggcgggucu acagcaccgg cagcaacgug uuccagaccc gggccgguug ccugaucggc 1200

[0059] gccgagcacg ugaacaacag cuacgagugc gacauccccа ucggcgccgg caucugugcc 1260

agcuaccaga cccagaccaa uucaccccgg agggcaagga gcguggccag ccagagcauc 1320

aucgccuaca ccaugagccu gggcgccgag aacagcgugg ccuacagcaa caacagcauc 1380

gccaucccca ccaacuucac caucagcgug accaccgaga uucugccccgu gagcaugacc 1440

aagaccagcg uggacugcac cauguacauc ugcggcgaca gcaccgagug cagcaaccug 1500

cugcugcagu acggcagcuu cugcacccag cugaaccggg cccugaccgg caucgccgug 1560

gagcaggaca agaacaccca ggagguguuc gcccagguga agcagaucua caagaccccu 1620

cccaucaagg acuucggcgg cuucaacuuc agccagaucc ugcccgaccc cagcaagccc 1680

agcaagcgga gcuucaucga ggaccugcug uucaacaagg ugacccuagc cgacgccggc 1740

uucaucaagc aguacggcga cugccucggc gacauagccg cccgggaccu gaucugcgcc 1800

cagaaguuca acggccugac cgugcugccu ccccugcuga ccgacgagau gaucgcccag 1860

uacaccagcg cccuguuagc cggaaccauc accagcggcu ggacuuucgg cgcuggagcc 1920

gcucugcaga uccccuucgc caugcagaug gccuaccggu ucaacggcau cggcgugacc 1980

cagaacgugc uguacgagaa ccagaagcug aucgccaacc aguucaacag cgccaucggc 2040

aagauccagg acagccugag cagcaccgcu agcgcccugg gcaagcugca ggacguggug 2100

aaccagaacg cccaggcccu gaacacccug gugaagcagc ugagcagcaa cuucggcgcc 2160

aucagcagcg ugcugaacga cauccugagc cggcuggacc cucccaacgc caccgugcag 2220

aucgaccggc ugaucacugg ccggcugcag agccugcaga ccuacgugac ccagcagcug 2280

auccgggccg ccgagauucg ggccagcgcc aaccuggccg ccaccaagau gagcgagugc 2340

gugcugggcc agagcaagcg gguggacuuc ugcggcaagg gcuaccaccu gaugagcuuu 2400

ccccagagcg caccccacgg agugguguuc cugcacguga ccuacgugcc cgcccaggag 2460

aagaacuuca ccaccgcccc agccaucugc cacgacggca aggcccacuu uccccgggag 2520

ggcguguucg ugagcaacgg cacccacugg uucgugaccc agcggaacuu cuacgagccc 2580

cagaucauca ccaccgacaa caccuucgug agcggcaacu gcgacguggu gaucggcauc 2640

gugaacaaca ccguguacga uccccugcag cccgagcugg acagcuucaa ggaggagcug 2700

gacaaguacu ucaagaauca caccagcccc gacguggacc ugggcgacau cagcggcauc 2760

aacgccagcg uggugaacau ccagaaggag aucgaucggc ugaacgaggu ggccaagaac 2820

cugaacgaga gccugaucga ccugcaggag cugggcaagu acgagcagua caucaagugg 2880

cccugguaca ucuggcuggg cuucaucgcc ggccugaucg ccaucgugau ggugaccauc 2940

[0060]

augcugugcu gcaugaccag cugcugcagc ugccugaagg gcuguugcag cugcggcagc 3000

ugcugcaagu ucgacgagga cgacagcgag cccgugcuga agggcgugaa gcugcacuac 3060

accugauaau aggcuggagc cucgguggcc uagcuucuug ccccuugggc cucccccccag 3120

ccccuccucc ccuuccugca cccguacccc cguggucuuu gaauaaaguc ugaguggggcg 3180

gc 3182

<210> 31
<211> 3006
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 31
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuucggcggc 960

agcggcggcg ugagcgugau caccccaggc accaacacca gcaaccaggu ggccgugcug 1020

uaccaggacg ugaacugcac cgaggugccc guggccaucc acgccgacca gcugacaccc 1080

[0061] accuggcggg ucuacagcac cggcagcaac guguuccaga cccgggccgg uugccugauc 1140

ggcgccgagc acgugaacaa cagcuacgag ugcgacaucc ccaucggcgc cggcaucugu 1200

gccagcuacc agacccagac caauucaccc cggagggcaa ggagcguggc cagccagagc 1260

aucaucgccu acaccaugag ccugggcgcc gagaacagcg uggccuacag caacaacagc 1320

aucgccaucc ccaccaacuu caccaucagc gugaccaccg agauucugcc cgugagcaug 1380

accaagacca gcguggacug caccauguac aucugcggcg acagcaccga gugcagcaac 1440

cugcugcugc aguacggcag cuucugcacc cagcugaacc gggcccugac cggcaucgcc 1500

guggagcagg acaagaacac ccaggaggug uucgcccagg ugaagcagau cuacaagacc 1560

ccucccauca aggacuucgg cggcuucaac uucagccaga uccugcccga ccccagcaag 1620

cccagcaagc ggagcuucau cgaggaccug cuguucaaca aggugacccu agccgacgcc 1680

ggcuucauca agcaguacgg cgacugccuc ggcgacauag ccgcccggga ccugaucugc 1740

gcccagaagu ucaacggccu gaccgugcug ccucccccugc ugaccgacga gaugaucgcc 1800

caguacacca gcgcccuguu agccggaacc aucaccagcg gcuggacuuu cggcgcugga 1860

gccgcucugc agaucccccuu cgccaugcag auggccuacc gguucaacgg caucggcgug 1920

acccagaacg ugcuguacga gaaccagaag cugaucgcca accaguucaa cagcgccauc 1980

ggcaagaucc aggacagccu gagcagcacc gcuagcgccc ugggcaagcu gcaggacgug 2040

gugaaccaga acgcccaggc ccugaacacc cuggugaagc agcugagcag caacuucggc 2100

gccaucagca gcgugcugaa cgacauccug agccggcugg acccucccaa cgccaccgug 2160

cagaucgacc ggcugaucac uggccggcug cagagccugc agaccuacgu gacccagcag 2220

cugauccggg ccgccgagau ucgggccagc gccaaccugg ccgccaccaa gaugagcgag 2280

ugcgugcugg gccagagcaa gcggguggac uucugcggca agggcuacca ccugaugagc 2340

uuuccccaga gcgcacccca cggaguggug uuccugcacg ugaccuacgu gcccgcccag 2400

gagaagaacu ucaccaccgc cccagccauc ugccacgacg gcaaggccca cuuuccccgg 2460

gagggcgugu ucgugagcaa cggcacccac ugguucguga cccagcggaa cuucuacgag 2520

ccccagauca ucaccaccga caacaccuuc gugagcggca acugcgacgu ggugaucggc 2580

aucgugaaca acaccgugua cgaucccccug cagcccgagc uggacagcuu caaggaggag 2640

cuggacaagu acuucaagaa ucacaccagc cccgacgugg accugggcga caucagcggc 2700

aucaacgcca gcguggugaa cauccagaag gagaucgauc ggcugaacga gguggccaag 2760

aaccugaacg agagccugau cgaccugcag gagcugggca aguacgagca guacaucaag 2820

[0062] uggccccuggu acaucuggcu gggcuucauc gccggccuga ucgccaucgu gauggugacc 2880

aucaugcugu gcugcaugac cagcugcugc agcugccuga agggcuguug cagcugcggc 2940

agcugcugca aguucgacga ggacgacagc gagcccgugc ugaagggcgu gaagcugcac 3000

uacacc 3006

<210> 32
<211> 1002
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 32

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

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

35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
[0063] 145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Gly Gly
305 310 315 320

Ser Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln
325 330 335

Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala
340 345 350

Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly
355 360 365

Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His
370 375 380

[0064]

Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys
385 390 395 400

Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val
405 410 415

Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn
420 425 430

Ser Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr
435 440 445

Ile Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser
450 455 460

Val Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn
465 470 475 480

Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu
485 490 495

Thr Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala
500 505 510

Gln Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly
515 520 525

Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg
530 535 540

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala
545 550 555 560

Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg
565 570 575

Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro
580 585 590

Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala
595 600 605

[0065]

Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln
610 615 620

Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val
625 630 635 640

Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe
645 650 655

Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser
660 665 670

Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu
675 680 685

Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser
690 695 700

Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Pro Pro Asn Ala Thr Val
705 710 715 720

Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr
725 730 735

Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn
740 745 750

Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg
755 760 765

Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser
770 775 780

Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln
785 790 795 800

Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala
805 810 815

His Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe
820 825 830

[0066] Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
835 840 845

Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn
850 855 860

Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu
865 870 875 880

Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly
885 890 895

Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile
900 905 910

Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp
915 920 925

Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr
930 935 940

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

945 950 955 960

Ile Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys
965 970 975

Cys Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro
980 985 990

Val Leu Lys Gly Val Lys Leu His Tyr Thr
995 1000

<210> 33
<211> 2378
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 33
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagg gcaccaucac cgacgccgug 120

[0067] gacugcgccc uggacccucu gagcgagacc aagugcaccc ugaagagcuu caccguggag 180

aagggcaucu accagaccag caacuucggc ggcagcggcg gcgugagcgu gaucacccca 240

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 300

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 360

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 420

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 480

ccccggaggg caaggagcgu ggccagccag agcaucaucg ccuacaccau gagccugggc 540

gccgagaaca gcguggccua cagcaacaac agcaucgcca uccccaccaa cuucaccauc 600

agcgugacca ccgagauucu gcccgugagc augaccaaga ccagcgugga cugcaccaug 660

uacaucugcg gcgacagcac cgagugcagc aaccugcugc ugaacuacac cagcuucugc 720

acccagcuga accgggcccu gaccggcauc gccguggagc aggacaagaa cacccaggag 780

guguucgccc aggugaagca gaucuacaag accccuccca ucaaggacuu cggcggcuuc 840

aacuucagcc agauccugcc cgaccccagc aagcccagca agcggagcuu caucgaggac 900

cugcuguuca acaaggugac ccuagccgac gccggcuuca ucaagcagua cggcgacugc 960

cucggcgaca uagccgcccg ggaccugauc ugcgcccaga aguucaacgg ccugaccgug 1020

cugccucccc ugcugaccga cgagaugauc gcccaguaca ccagcgcccu guuagccgga 1080

accaucacca gcggcuggac uuucggcgcu ggagccgcuc ugcagauccc cuucgccaug 1140

cagauggccu accgguucaa cggcaucggc gugacccaga acgugcugua cgagaaccag 1200

aagcugaucg ccaaccaguu caacagcgcc aucggcaaga uccaggacag ccugagcagc 1260

accgcuagcg cccugggcaa gcugcaggac guggugaacc agaacgccca ggcccugaac 1320

acccugguga agcagcugag cagcaacuuc ggcgccauca gcagcgugcu gaacgacauc 1380

cugagccggc uggacccucc caacgccacc gugcagaucg accggcugau cacuggccgg 1440

cugcagagcc ugcagaccua cgugacccag cagcugaucc gggccgccga gauucgggcc 1500

agcgccaacc uggccgccac caagaugagc gagugcgugc ugggccagag caagcgggug 1560

gacuucugcg gcaagggcua ccaccugaug agcuuucccc agagcgcacc ccacggagug 1620

guguuccugc acgugaccua cgugcccgcc caggagaaga acuucaccac cgccccagcc 1680

aucugccacg acggcaaggc ccacuuuccc cgggagggcg uguucgugag caacggcacc 1740

cacugguucg ugacccagcg gaacuucuac gagccccaga ucaucaccac cgacaacacc 1800

uucgugagcg gcaacugcga cguggugauc ggcaucguga acaacaccgu guacgauccc 1860

[0068] cugcagcccg agcuggacag cuucaaggag gagcuggaca aguacuucaa gaaucacacc 1920

agccccgacg uggaccuggg cgacaucagc ggcaucaacg ccagcguggu gaacauccag 1980

aaggagaucg aucggcugaa cgagguggcc aagaaccuga acgagagccu gaucgaccug 2040

caggagcugg gcaaguacga gcaguacauc aaguggcccu gguacaucug gcugggcuuc 2100

aucgccggcc ugaucgccau cgugauggug accaucaugc ugugcugcau gaccagcugc 2160

ugcagcugcc ugaagggcug uugcagcugc ggcagcugcu gcaaguucga cgaggacgac 2220

agcgagcccg ugcugaaggg cgugaagcug cacuacaccu gauaauaggc uggagccucg 2280

guggccuagc uucuugcccc uuggggccucc ccccagcccc uccuccccuu ccugcacccg 2340

uacccccgug gucuuugaau aaagucugag ugggcggc 2378

<210> 34
<211> 2202
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 34

auguucgugu uccuggugcu gcugccccug gugagcagcc agggcaccau caccgacgcc 60

guggacugcg cccuggaccc ucugagcgag accaagugca cccugaagag cuucaccgug 120

gagaagggca ucuaccagac cagcaacuuc ggcggcagcg gcggcgugag cgugaucacc 180

ccaggcacca acaccagcaa ccagguggcc gugcuguacc aggacgugaa cugcaccgag 240

gugcccgugg ccauccacgc cgaccagcug acacccaccu ggcgggucua cagcaccggc 300

agcaacgugu uccagacccg ggccgguugc cugaucggcg ccgagcacgu gaacaacagc 360

uacgagugcg acaucccccau cggcgccggc aucugugcca gcuaccagac ccagaccaau 420

ucaccccgga gggcaaggag cguggccagc cagagcauca ucgccuacac caugagccug 480

ggcgccgaga acagcguggc cuacagcaac aacagcaucg ccaucccccac caacuucacc 540

aucagcguga ccaccgagau ucugcccgug agcaugacca agaccagcgu ggacugcacc 600

auguacaucu gcggcgacag caccgagugc agcaaccugc ugcugaacua caccagcuuc 660

ugcacccagc ugaaccgggc ccugaccggc aucgccgugg agcaggacaa gaacacccag 720

gagguguucg cccaggugaa gcagaucuac aagaccccuc ccaucaagga cuucggcggc 780

uucaacuuca gccagauccu gcccgacccc agcaagccca gcaagcggag cuucaucgag 840

[0069] gaccugcugu ucaacaaggu gacccuagcc gacgccggcu ucaucaagca guacggcgac 900

ugccucggcg acauagccgc ccgggaccug aucugcgccc agaaguucaa cggccugacc 960

gugcugccuc cccugcugac cgacgagaug aucgcccagu acaccagcgc ccuguuagcc 1020

ggaaccauca ccagcggcug gacuuucggc gcuggagccg cucugcagau ccccuucgcc 1080

augcagaugg ccuaccgguu caacggcauc ggcgugaccc agaacgugcu guacgagaac 1140

cagaagcuga ucgccaacca guucaacagc gccaucggca agauccagga cagccugagc 1200

agcaccgcua gcgcccuggg caagcugcag gacgugguga accagaacgc ccaggcccug 1260

aacacccugg ugaagcagcu gagcagcaac uucggcgcca ucagcagcgu gcugaacgac 1320

auccugagcc ggcuggaccc ucccaacgcc accgugcaga ucgaccggcu gaucacuggc 1380

cggcugcaga gccugcagac cuacgugacc cagcagcuga uccgggccgc cgagauucgg 1440

gccagcgcca accuggccgc caccaagaug agcgagugcg ugcugggcca gagcaagcgg 1500

guggacuucu gcggcaaggg cuaccaccug augagcuuuc cccagagcgc accccacgga 1560

gugguguucc ugcacgugac cuacgugccc gcccaggaga agaacuucac caccgcccca 1620

gccaucugcc acgacggcaa ggcccacuuu ccccgggagg gcguguucgu gagcaacggc 1680

acccacuggu ucgugaccca gcggaacuuc uacgagcccc agaucaucac caccgacaac 1740

accuucguga gcggcaacug cgacguggug aucggcaucg ugaacaacac cguguacgau 1800

ccccugcagc ccgagcugga cagcuucaag gaggagcugg acaaguacuu caagaaucac 1860

accagccccg acguggaccu gggcgacauc agcggcauca acgccagcgu ggugaacauc 1920

cagaaggaga ucgaucggcu gaacgaggug gccaagaacc ugaacgagag ccugaucgac 1980

cugcaggagc ugggcaagua cgagcaguac aucaaguggc ccugguacau cuggcugggc 2040

uucaucgccg gccugaucgc caucgugaug gugaccauca ugcugugcug caugaccagc 2100

ugcugcagcu gccugaaggg cuguugcagc ugcggcagcu gcugcaaguu cgacgaggac 2160

gacagcgagc ccgugcugaa gggcgugaag cugcacuaca cc 2202

<210> 35
<211> 734
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 35

[0070]

```
Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Gly Thr
1               5                  10                  15

Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys
            20                  25                  30

Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser
        35                  40                  45

Asn Phe Gly Gly Ser Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn
    50                  55                  60

Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu
65                  70                  75                  80

Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val
                85                  90                  95

Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile
            100                 105                 110

Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly
```

115 120 125

Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg
130 135 140

Ala Arg Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu
145 150 155 160

Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro
165 170 175

Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met
180 185 190

Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr
195 200 205

Glu Cys Ser Asn Leu Leu Leu Asn Tyr Thr Ser Phe Cys Thr Gln Leu
210 215 220

Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln
[0071] 225 230 235 240

Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys
245 250 255

Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys
260 265 270

Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr
275 280 285

Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp
290 295 300

Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr
305 310 315 320

Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser
325 330 335

Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly
340 345 350

Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn
355 360 365

Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile
370 375 380

Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser
385 390 395 400

Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn
405 410 415

Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly
420 425 430

Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Pro Pro
435 440 445

Asn Ala Thr Val Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser
450 455 460

[0072]

Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg
465 470 475 480

Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly
485 490 495

Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser
500 505 510

Phe Pro Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr
515 520 525

Val Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His
530 535 540

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly
545 550 555 560

Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile
565 570 575

Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
580 585 590

Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser
595 600 605

Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp
610 615 620

Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
625 630 635 640

Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu
645 650 655

Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys
660 665 670

Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile
675 680 685

[0073]

Val Met Val Thr Ile Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys
690 695 700

Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp
705 710 715 720

Asp Ser Glu Pro Val Leu Lys Gly Val Lys Leu His Tyr Thr
725 730

<210> 36
<211> 4241
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 36
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uuccugauca ucuucauccu gcccaccacc cuggccguga ucggcgacuu caacugcacc 120

aacagcuuca ucaacgacua caacaagacc aucccucgga ucagcgaaga cgugguggac 180

gugucccugg gccugggcac cuacuacgug cugaaccggg uguaccugaa caccacacug 240

[0074]

cuguucaccg gcuacuuccc caagagcggc gccaacuucc gggaccuggc ccugaagggc 300

agcaucuacc ugagcaccuu gugguacaag ccucccuucc ugagcgacuu caauaacggc 360

aucuucucua aggugaagaa caccaagcug uacguaaaca acacccugua cagcgaguuc 420

agcaccaucg ugaucggcag cguguucguc aacaccagcu acaccaucgu ggugcagccc 480

cacaacggca uccuggagau caccgccugc caguacacca ugugcgagua cccucacacc 540

gugugcaaga gcaagggcuc cauccggaac gagagcuggc acaucgacag cagcgagccg 600

cugugccugu ucaagaagaa cuucaccuac aacgugagcg ccgacuggcu guacuuccac 660

uucuaccagg agcggggcgu guucuacgcc uacuacgccg acgugggcau gccaaccacc 720

uuccuguuca gccuguaccu gggcaccauc cugagccacu acuacgugau gccccugacc 780

ugcaacgcca ucagcucaaa caccgacaac gagacccugg aguacugggu gacuccacug 840

agccggcggc aguaccugcu gaacuucgac gagcacggcg ugaucaccaa cgccguggac 900

ugcgccccugg acccucugag cgagaccaag ugcacccuga agagcuucac cguggagaag 960

ggcaucuacc agaccagcgg cuucaccgug aagcccguag ccaccgugua ccggcggauc 1020

cccaaccugc ccgacugcga caucgacaac uggcugaaca acgucagcgu gcccagccca 1080

cugaacuggg agcggcggau cuucagcaac ugcaacuuca aucugagcac ccugcugcgg 1140

cuggugcacg uggacagcuu cagcugcaac aaccuggaca agagcaagau cuucgguagc 1200

ugcuucaaca gcaucaccgu ggacaaguuc gccaucccua accggcggcg ggacgaucug 1260

cagcugggca gcagcggcuu ccugcagagc agcaacuaca agaucgacau cagcagcuca 1320

agcugccagc uguacuacag ccugccccug gugaacguga ccaucaacaa cuucaacccc 1380

agcagcugga accggcggua cggcuucggc agcuucaacc ugagcagcua cgacguggug 1440

uacagcgacc acugcuucag cgugaacagc gacuucugcc ccugugccga cccuagcgug 1500

gugaacagcu gcgccaagag caagccuccc agcgccauuu gccccgccgg caccaaguac 1560

cggcacugcg accuggacac cacccuguac gugaagaacu ggugccggug cagcugccug 1620

cccgacccca ucagcaccua cagccccaac accugucccc agaagaaggu gguggugggu 1680

aucggcgagc acugucccgg ccugggcauc aacgaggaga agugcggcac ccagcugaac 1740

cacagcagcu gcuucuguag ccccgacgcc uuccugggcu ggagcuucga cagcugcauc 1800

agcaacaacc ggugcaacau cuuuagcaac uucaucuuca acggaaucaa cagcggcacc 1860

accugcagca acgaccugcu guauagcaac accgagauca gcaccggcgu gugcgugaac 1920

uacgaccugu acggcaucac cggccagggc aucuucaagg aggugagcgc cgccuacuac 1980

aacaacuggc agaaccugcu guacgacagc aacggcaaca ucaucggcuu caaggacuuu 2040

cugaccaaca agaccuacac cauccugccc ugcuacagcg gcggcgugag cgugaucacc 2100

ccaggcacca acaccagcaa ccagguggcc gugcuguacc aggacgugaa cugcaccgag 2160

gugcccgugg ccauccacgc cgaccagcug acacccaccu ggcgggucua cagcaccggc 2220

agcaacgugu uccagacccg ggccgguugc cugaucggcg ccgagcacgu gaacaacagc 2280

uacgagugcg acaucccccau cggcgccggc aucugugcca gcuaccagac ccagaccaau 2340

ucaccccgga gggcaaggag cguggccagc cagagcauca ucgccuacac caugagccug 2400

ggcgccgaga acagcguggc cuacagcaac aacagcaucg ccaucccccac caacuucacc 2460

aucagcguga ccaccgagau ucugcccgug agcaugacca agaccagcgu ggacugcacc 2520

auguacaucu gcggcgacag caccgagugc agcaaccugc ugcugcagua cggcagcuuc 2580

ugcacccagc ugaaccgggc ccugaccggc aucgccgugg agcaggacaa gaacacccag 2640

gagguguucg cccaggugaa gcagaucuac aagaccccuc ccaucaagga cuucggcggc 2700

uucaacuuca gccagauccu gcccgacccc agcaagccca gcaagcggag cuucaucgag 2760

[0075] gaccugcugu ucaacaaggu gacccuagcc gacgccggcu ucaucaagca guacggcgac 2820

ugccucggcg acauagccgc ccgggaccug aucugcgccc agaaguucaa cggccugacc 2880

gugcugccuc cccugcugac cgacgagaug aucgcccagu acaccagcgc ccuguuagcc 2940

ggaaccauca ccagcggcug gacuuucggc gcuggagccg cucugcagau ccccuucgcc 3000

augcagaugg ccuaccgguu caacggcauc ggcgugaccc agaacgugcu guacgagaac 3060

cagaagcuga ucgccaacca guucaacagc gccaucggca agauccagga cagccugagc 3120

agcaccgcua gcgcccuggg caagcugcag gacgugguga accagaacgc ccaggcccug 3180

aacacccugg ugaagcagcu gagcagcaac uucggcgcca ucagcagcgu gcugaacgac 3240

auccugagcc ggcuggaccc ucccgaggcc gaggugcaga ucgaccggcu gaucacuggc 3300

cggcugcaga gccugcagac cuacgugacc cagcagcuga uccgggccgc cgagauucgg 3360

gccagcgcca accuggccgc caccaagaug agcgagugcg ugcugggcca gagcaagcgg 3420

guggacuucu gcggcaaggg cuaccaccug augagcuuuc cccagagcgc accccacgga 3480

gugguguucc ugcacgugac cuacgugccc gcccaggaga agaacuucac caccgcccca 3540

gccaucugcc acgacggcaa ggcccacuuu ccccgggagg gcguguucgu gagcaacggc 3600

acccacuggu ucgugaccca gcggaacuuc uacgagcccc agaucaucac caccgacaac 3660

accuucguga gcggcaacug cgacguggug aucggcaucg ugaacaacac cguguacgau 3720

ccccugcagc ccgagcugga cagcuucaag gaggagcugg acaaguacuu caagaaucac 3780

accagccccg acguggaccu gggcgacauc agcggcauca acgccagcgu ggugaacauc 3840

cagaaggaga ucgaucggcu gaacgaggug gccaagaacc ugaacgagag ccugaucgac 3900

cugcaggagc ugggcaagua cgagcaguac aucaaguggc ccugguacau cuggcugggc 3960

uucaucgccg gccugaucgc caucgugaug gugaccauca ugcugugcug caugaccagc 4020

ugcugcagcu gccugaaggg cuguugcagc ugcggcagcu gcugcaaguu cgacgaggac 4080

gacagcgagc ccgugcugaa gggcgugaag cugcacuaca ccugauaaua ggcuggagcc 4140

ucgguggccu agcuucuugc cccuugggcc ucccccccagc cccuccuccc cuuccugcac 4200

ccguaccccc guggucuuug aauaaagucu gagugggcgg c 4241

<210> 37
<211> 4065
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
[0076] <223> 合成

<400> 37
auguuccuga ucaucuucau ccugcccacc acccuggccg ugaucggcga cuucaacugc 60

accaacagcu ucaucaacga cuacaacaag accaucccuc ggaucagcga agacguggug 120

gacguguccc ugggccuggg caccuacuac gugcugaacc ggguguaccu gaacaccaca 180

cugcuguuca ccggcuacuu ccccaagagc ggcgccaacu uccgggaccu ggcccugaag 240

ggcagcaucu accugagcac cuugugguac aagccucccu uccugagcga cuucaauaac 300

ggcaucuucu cuaaggugaa gaacaccaag cuguacguaa acaacacccu guacagcgag 360

uucagcacca ucgugaucgg cagcguguuc gucaacacca gcuacaccau cguggugcag 420

ccccacaacg gcauccugga gaucaccgcc ugccaguaca ccaugugcga guacccucac 480

accgugugca agagcaaggg cuccauccgg aacgagagcu ggcacaucga cagcagcgag 540

ccgcugugcc uguucaagaa gaacuucacc uacaacguga gcgccgacug gcuguacuuc 600

cacuucuacc aggagcgggg cguguucuac gccuacuacg ccgacguggg caugccaacc 660

accuuccugu ucagccugua ccugggcacc auccugagcc acuacuacgu gaugccccug 720

accugcaacg ccaucagcuc aaacaccgac aacgagaccc uggaguacug ggugacucca 780

cugagccggc ggcaguaccu gcugaacuuc gacgagcacg gcgugaucac caacgccgug 840

gacugcgccc uggacccucu gagcgagacc aagugcaccc ugaagagcuu caccguggag 900

aagggcaucu accagaccag cggcuucacc gugaagcccg uagccaccgu guaccggcgg 960

aucccccaacc ugcccgacug cgacaucgac aacuggcuga acaacgucag cgugcccagc 1020

ccacugaacu gggagcggcg gaucuucagc aacugcaacu ucaaucugag cacccugcug 1080

cggcuggugc acguggacag cuucagcugc aacaaccugg acaagagcaa gaucuucggu 1140

agcugcuuca acagcaucac cguggacaag uucgccaucc cuaaccggcg gcgggacgau 1200

cugcagcugg gcagcagcgg cuuccugcag agcagcaacu acaagaucga caucagcagc 1260

ucaagcugcc agcuguacua cagccugccc cuggugaacg ugaccaucaa caacuucaac 1320

cccagcagcu ggaaccggcg guacggcuuc ggcagcuuca accugagcag cuacgacgug 1380

guguacagcg accacugcuu cagcgugaac agcgacuucu gccccugugc cgacccuagc 1440

guggugaaca gcugcgccaa gagcaagccu cccagcgcca uuugccccgc cggcaccaag 1500

uaccggcacu gcgaccugga caccacccug uacgugaaga acuggugccg gugcagcugc 1560

cugcccgacc ccaucagcac cuacagcccc aacaccuguc cccagaagaa ggugguggug 1620

[0077] gguaucggcg agcacugucc cggccugggc aucaacgagg agaagugcgg cacccagcug 1680

aaccacagca gcugcuucug uagccccgac gccuuccugg gcuggagcuu cgacagcugc 1740

aucagcaaca accggugcaa caucuuuagc aacuucaucu ucaacggaau caacagcggc 1800

accaccugca gcaacgaccu gcuguauagc aacaccgaga ucagcaccgg cgugugcgug 1860

aacuacgacc uguacggcau caccggccag ggcaucuuca aggaggugag cgccgccuac 1920

uacaacaacu ggcagaaccu gcuguacgac agcaacggca acaucaucgg cuucaaggac 1980

uuucugacca acaagaccua caccauccug cccugcuaca gcggcggcgu gagcgugauc 2040

accccaggca ccaacaccag caaccaggug gccgugcugu accaggacgu gaacugcacc 2100

gaggugcccg uggccaucca cgccgaccag cugacaccca ccuggcgggu cuacagcacc 2160

ggcagcaacg uguuccagac ccgggccggu ugccugaucg gcgccgagca cgugaacaac 2220

agcuacgagu gcgacauccc caucggcgcc ggcaucugug ccagcuacca gacccagacc 2280

aauucacccc ggagggcaag gagcguggcc agccagagca ucaucgccua caccaugagc 2340

cugggcgccg agaacagcgu ggccuacagc aacaacagca ucgccauccc caccaacuuc 2400

accaucagcg ugaccaccga gauucugccc gugagcauga ccaagaccag cguggacugc 2460

accauguaca ucugcggcga cagcaccgag ugcagcaacc ugcugcugca guacggcagc 2520

uucugcaccc agcugaaccg ggcccugacc ggcaucgccg uggagcagga caagaacacc 2580

caggaggugu ucgcccaggu gaagcagauc uacaagaccc cucccaucaa ggacuucggc 2640

ggcuucaacu ucagccagau ccugcccgac cccagcaagc ccagcaagcg gagcuucauc 2700

gaggaccugc uguucaacaa ggugacccua gccgacgccg gcuucaucaa gcaguacggc 2760

gacugccucg gcgacauagc cgcccgggac cugaucugcg cccagaaguu caacggccug 2820

accgugcugc cucccccugcu gaccgacgag augaucgccc aguacaccag cgcccuguua 2880

gccggaacca ucaccagcgg cuggacuuuc ggcgcuggag ccgcucugca gaucccccuuc 2940

gccaugcaga uggccuaccg guucaacggc aucggcguga cccagaacgu gcuguacgag 3000

aaccagaagc ugaucgccaa ccaguucaac agcgccaucg gcaagaucca ggacagccug 3060

agcagcaccg cuagcgcccu gggcaagcug caggacgugg ugaaccagaa cgcccaggcc 3120

cugaacaccc uggugaagca gcugagcagc aacuucggcg ccaucagcag cgugcugaac 3180

gacauccuga gccggcugga cccucccgag gccgaggugc agaucgaccg gcugaucacu 3240

ggccggcugc agagccugca gaccuacgug acccagcagc ugauccgggc cgccgagauu 3300

cgggccagcg ccaaccuggc cgccaccaag augagcgagu gcgugcuggg ccagagcaag 3360

[0078] cggguggacu ucugcggcaa gggcuaccac cugaugagcu uuccccagag cgcaccccac 3420

ggaguggugu uccugcacgu gaccuacgug cccgcccagg agaagaacuu caccaccgcc 3480

ccagccaucu gccacgacgg caaggcccac uuuccccggg agggcguguu cgugagcaac 3540

ggcacccacu gguucgugac ccagcggaac uucuacgagc cccagaucau caccaccgac 3600

aacaccuucg ugagcggcaa cugcgacgug gugaucggca ucgugaacaa caccguguac 3660

gaucccccugc agcccgagcu ggacagcuuc aaggaggagc uggacaagua cuucaagaau 3720

cacaccagcc ccgacgugga ccugggcgac aucagcggca ucaacgccag cguggugaac 3780

auccagaagg agaucgaucg gcugaacgag guggccaaga accugaacga gagccugauc 3840

gaccugcagg agcugggcaa guacgagcag uacaucaagu ggcccuggua caucuggcug 3900

ggcuucaucg ccggccugau cgccaucgug auggugacca ucaugcugug cugcaugacc 3960

agcugcugca gcugccugaa gggcuguugc agcugcggca gcugcugcaa guucgacgag 4020

gacgacagcg agcccgugcu gaagggcgug aagcugcacu acacc 4065

<210> 38
<211> 1355
<212> PRT

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 38

Met Phe Leu Ile Ile Phe Ile Leu Pro Thr Thr Leu Ala Val Ile Gly
1 5 10 15

Asp Phe Asn Cys Thr Asn Ser Phe Ile Asn Asp Tyr Asn Lys Thr Ile
20 25 30

Pro Arg Ile Ser Glu Asp Val Val Asp Val Ser Leu Gly Leu Gly Thr
35 40 45

Tyr Tyr Val Leu Asn Arg Val Tyr Leu Asn Thr Thr Leu Leu Phe Thr
50 55 60

Gly Tyr Phe Pro Lys Ser Gly Ala Asn Phe Arg Asp Leu Ala Leu Lys
65 70 75 80

[0079] Gly Ser Ile Tyr Leu Ser Thr Leu Trp Tyr Lys Pro Pro Phe Leu Ser
85 90 95

Asp Phe Asn Asn Gly Ile Phe Ser Lys Val Lys Asn Thr Lys Leu Tyr
100 105 110

Val Asn Asn Thr Leu Tyr Ser Glu Phe Ser Thr Ile Val Ile Gly Ser
115 120 125

Val Phe Val Asn Thr Ser Tyr Thr Ile Val Val Gln Pro His Asn Gly
130 135 140

Ile Leu Glu Ile Thr Ala Cys Gln Tyr Thr Met Cys Glu Tyr Pro His
145 150 155 160

Thr Val Cys Lys Ser Lys Gly Ser Ile Arg Asn Glu Ser Trp His Ile
165 170 175

Asp Ser Ser Glu Pro Leu Cys Leu Phe Lys Lys Asn Phe Thr Tyr Asn
180 185 190

Val Ser Ala Asp Trp Leu Tyr Phe His Phe Tyr Gln Glu Arg Gly Val
195 200 205

Phe Tyr Ala Tyr Tyr Ala Asp Val Gly Met Pro Thr Thr Phe Leu Phe
210 215 220

Ser Leu Tyr Leu Gly Thr Ile Leu Ser His Tyr Tyr Val Met Pro Leu
225 230 235 240

Thr Cys Asn Ala Ile Ser Ser Asn Thr Asp Asn Glu Thr Leu Glu Tyr
245 250 255

Trp Val Thr Pro Leu Ser Arg Arg Gln Tyr Leu Leu Asn Phe Asp Glu
260 265 270

His Gly Val Ile Thr Asn Ala Val Asp Cys Ala Leu Asp Pro Leu Ser
275 280 285

Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly Ile Tyr
290 295 300

Gln Thr Ser Gly Phe Thr Val Lys Pro Val Ala Thr Val Tyr Arg Arg
305 310 315 320

[0080]

Ile Pro Asn Leu Pro Asp Cys Asp Ile Asp Asn Trp Leu Asn Asn Val
325 330 335

Ser Val Pro Ser Pro Leu Asn Trp Glu Arg Arg Ile Phe Ser Asn Cys
340 345 350

Asn Phe Asn Leu Ser Thr Leu Leu Arg Leu Val His Val Asp Ser Phe
355 360 365

Ser Cys Asn Asn Leu Asp Lys Ser Lys Ile Phe Gly Ser Cys Phe Asn
370 375 380

Ser Ile Thr Val Asp Lys Phe Ala Ile Pro Asn Arg Arg Arg Asp Asp
385 390 395 400

Leu Gln Leu Gly Ser Ser Gly Phe Leu Gln Ser Ser Asn Tyr Lys Ile
405 410 415

Asp Ile Ser Ser Ser Ser Cys Gln Leu Tyr Tyr Ser Leu Pro Leu Val
420 425 430

Asn Val Thr Ile Asn Asn Phe Asn Pro Ser Ser Trp Asn Arg Arg Tyr
435 440 445

Gly Phe Gly Ser Phe Asn Leu Ser Ser Tyr Asp Val Val Tyr Ser Asp
450 455 460

His Cys Phe Ser Val Asn Ser Asp Phe Cys Pro Cys Ala Asp Pro Ser
465 470 475 480

Val Val Asn Ser Cys Ala Lys Ser Lys Pro Pro Ser Ala Ile Cys Pro
485 490 495

Ala Gly Thr Lys Tyr Arg His Cys Asp Leu Asp Thr Thr Leu Tyr Val
500 505 510

Lys Asn Trp Cys Arg Cys Ser Cys Leu Pro Asp Pro Ile Ser Thr Tyr
515 520 525

Ser Pro Asn Thr Cys Pro Gln Lys Lys Val Val Val Gly Ile Gly Glu
530 535 540

[0081]

His Cys Pro Gly Leu Gly Ile Asn Glu Glu Lys Cys Gly Thr Gln Leu
545 550 555 560

Asn His Ser Ser Cys Phe Cys Ser Pro Asp Ala Phe Leu Gly Trp Ser
565 570 575

Phe Asp Ser Cys Ile Ser Asn Asn Arg Cys Asn Ile Phe Ser Asn Phe
580 585 590

Ile Phe Asn Gly Ile Asn Ser Gly Thr Thr Cys Ser Asn Asp Leu Leu
595 600 605

Tyr Ser Asn Thr Glu Ile Ser Thr Gly Val Cys Val Asn Tyr Asp Leu
610 615 620

Tyr Gly Ile Thr Gly Gln Gly Ile Phe Lys Glu Val Ser Ala Ala Tyr
625 630 635 640

Tyr Asn Asn Trp Gln Asn Leu Leu Tyr Asp Ser Asn Gly Asn Ile Ile
645 650 655

Gly Phe Lys Asp Phe Leu Thr Asn Lys Thr Tyr Thr Ile Leu Pro Cys
660 665 670

Tyr Ser Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn
675 680 685

Gln Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val
690 695 700

Ala Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr
705 710 715 720

Gly Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu
725 730 735

His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
740 745 750

Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser
755 760 765

[0082] Val Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu
770 775 780

Asn Ser Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe
785 790 795 800

Thr Ile Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr
805 810 815

Ser Val Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser
820 825 830

Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala
835 840 845

Leu Thr Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe
850 855 860

Ala Gln Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly
865 870 875 880

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

885 890 895

Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp
900 905 910

Ala Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala
915 920 925

Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro
930 935 940

Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu
945 950 955 960

Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu
965 970 975

Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly
980 985 990

Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln
[0083] 995 1000 1005

Phe Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr
1010 1015 1020

Ala Ser Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala
1025 1030 1035

Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly
1040 1045 1050

Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Pro
1055 1060 1065

Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu
1070 1075 1080

Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala Ala
1085 1090 1095

Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu
1100 1105 1110

Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly
1115 1120 1125

Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly Val Val
1130 1135 1140

Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn Phe Thr
1145 1150 1155

Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe Pro Arg
1160 1165 1170

Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val Thr Gln
1175 1180 1185

Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe
1190 1195 1200

Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr
1205 1210 1215

[0084]

Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu
1220 1225 1230

Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu
1235 1240 1245

Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys
1250 1255 1260

Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser
1265 1270 1275

Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys
1280 1285 1290

Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala
1295 1300 1305

Ile Val Met Val Thr Ile Met Leu Cys Cys Met Thr Ser Cys Cys
1310 1315 1320

Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys Lys Phe
1325 1330 1335

Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly Val Lys Leu His
1340 1345 1350

Tyr Thr
1355

<210> 39
<211> 4247
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 39

[0085]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uuccugaucc ugcugaucag ccugcccacc gccuucgccg ugaucggcga ccugaagugc 120

accagcgaca acaucaacga caaggacacc ggcccaccac ccaucagcac cgacaccgug 180

gacgugacca acggccuggg caccuacuac gugcuggacc ggguguaccu gaacaccacc 240

cuguuccuga acggcuacua ccccaccagc ggcagcaccu accggaauau ggcccugaag 300

ggcagcgugc ugcugagccg gcugugguuc aagccaccau uccugagcga cuucaucaac 360

ggaaucuucg ccaaggugaa gaacaccaag gugaucaagg accgggugau guacagcgag 420

uuccccgcca ucaccauugg caguaccuuc gugaacacca gcuacagcgu gguggugcag 480

ccccggacca ucaacagcac ccaggacggc gacaacaagc ugcagggccu gcuggaggug 540

agcgugugcc aguacaacau gugcgaguac ccucagacca ucugccaccc caaccugggc 600

aaccaccgga aggagcugug gcaccuggac accggcgugg ugagcugccu guacaagcgg 660

aacuucaccu acgacguaaa cgccgacuac cuguacuucc acuucuacca ggagggcggc 720

accuucuacg ccuacuucac cgacacgggc guggugacca aguuccuguu caacguguac 780

cugggcaugg cccugagcca cuacuacgug augccccuga ccuguaacag caagcugacc 840

cuggaguacu gggugacccc ucugaccagc cggcaguacc ugcuggccuu caaccaggac 900

ggcaucaucu ucaacgccgu ggacugcgcc cuggacccuc ugagcgagac caagugcacc 960

cugaagagcu ucaccgugga gaagggcauc uaccagacca acggcuacac cgugcagccc 1020

aucgccgacg uguaccggcg gaagcccaac cugcccaacu gcaacaucga ggccuggcug 1080

[0086]

aacgacaaga gcgugcccuc gccccugaac ugggagcgga agaccuucag caacugcaac 1140

uucaacauga gcagucugau gagcuucauc caggccgaca gcuucaccug caacaacauc 1200

gacgccgcca agaucuacgg caugugcuuc agcagcauca ccaucgacaa guuugccauc 1260

cccaacggcc ggaaggugga ccugcagcug ggcaaccugg gcuaccugca gagcuucaac 1320

uaccggaucg acaccaccgc caccucuugc cagcuguacu acaaccugcc cgccgccaac 1380

gugagcguga gccgguucaa ccccagcacc uggaacaagc gguucggcuu cauugaggac 1440

agcguguuca aaccccggcc cgcaggagua cugaccaacc acgacguggu guacgcccag 1500

cacugcuuca aggcacccaa gaacuucugc cccugcaagc ugaacggcag cugugugggc 1560

ucuggccccg guaagaacaa cggcauaggg acuugccccgg cagggaccaa cuaccugacc 1620

ugcgacaacc ugugcacacc cgaccccauc accuucaccg gcaccuacaa gugucccccag 1680

accaagagcc uggugggcau cggcgagcac ugcagcggcc uggccgugaa gagcgacuac 1740

ugcggcggca acagcugcac cugucggccc caggccuucc ugggcuggag cgccgacagc 1800

ugccugcagg gcgacaagug caacaucuuu gccaacuuca uccugcacga cgugaacagc 1860

ggccugaccu gcagcaccga ccugcagaag gccaacaccg acaucauccu gggcgugugc 1920

gugaacuacg acuuguacgg cauccugggc cagggcaucu ucguggaggu gaacgccacc 1980

uacuacaaca gcuggcagaa ccugcuguac gacagcaacg gcaaccugua cggcuuccgg 2040

gacuacauca ucaaccggac cuucaugauc cggagcugcu acagcggcgg cgugagcgug 2100

aucaccccag gcaccaacac cagcaaccag guggccgugc uguaccagga cgugaacugc 2160

accgaggugc ccguggccau ccacgccgac cagcugacac ccaccuggcg ggucuacagc 2220

accggcagca acguguucca gacccgggcc gguugccuga ucggcgccga gcacgugaac 2280

aacagcuacg agugcgacau ccccaucggc gccggcaucu gugccagcua ccagacccag 2340

accaauucac cccggagggc aaggagcgug gccagccaga gcaucaucgc cuacaccaug 2400

agccugggcg ccgagaacag cguggccuac agcaacaaca gcaucgccau ccccaccaac 2460

uucaccauca gcgugaccac cgagauucug cccgugagca ugaccaagac cagcguggac 2520

ugcaccaugu acaucugcgg cgacagcacc gagugcagca accugcugcu gcaguacggc 2580

agcuucugca cccagcugaa ccgggcccug accggcaucg ccguggagca ggacaagaac 2640

acccaggagg uguucgccca ggugaagcag aucuacaaga ccccucccau caaggacuuc 2700

ggcggcuuca acuucagcca gauccugccc gaccccagca agcccagcaa gcggagcuuc 2760

aucgaggacc ugcuguucaa caaggugacc cuagccgacg ccggcuucau caagcaguac 2820

ggcgacugcc ucggcgacau agccgccccgg gaccugaucu gcgcccagaa guucaacggc 2880

cugaccgugc ugccucccccu gcugaccgac gagaugaucg cccaguacac cagcgcccug 2940

uuagccggaa ccaucaccag cggcuggacu uucggcgcug gagccgcucu gcagaucccc 3000

uucgccaugc agauggccua ccgguucaac ggcaucggcg ugacccagaa cgugcuguac 3060

gagaaccaga agcugaucgc caaccaguuc aacagcgcca ucggcaagau ccaggacagc 3120

cugagcagca ccgcuagcgc ccugggcaag cugcaggacg uggugaacca gaacgcccag 3180

gcccugaaca cccuggugaa gcagcugagc agcaacuucg gcgccaucag cagcgugcug 3240

aacgacaucc ugagccggcu ggacccuccc gaggccgagg ugcagaucga ccggcugauc 3300

acuggccggc ugcagagccu gcagaccuac gugacccagc agcugauccg ggccgccgag 3360

auucgggcca gcgccaaccu ggccgccacc aagaugagcg agugcgugcu gggccagagc 3420

aagcgggugg acuucugcgg caagggcuac caccugauga gcuuucccca gagcgcaccc 3480

cacggagugg uguuccugca cgugaccuac gugcccgccc aggagaagaa cuucaccacc 3540

gccccagcca ucugccacga cggcaaggcc cacuuucccc gggagggcgu guucgugagc 3600

[0087] aacggcaccc acugguucgu gacccagcgg aacuucuacg agccccagau caucaccacc 3660

gacaacaccu ucgugagcgg caacugcgac guggugaucg gcaucgugaa caacaccgug 3720

uacgaucccc ugcagcccga gcuggacagc uucaaggagg agcuggacaa guacuucaag 3780

aaucacacca gccccgacgu ggaccugggc gacaucagcg gcaucaacgc cagcguggug 3840

aacauccaga aggagaucga ucggcugaac gagguggcca agaaccugaa cgagagccug 3900

aucgaccugc aggagcuggg caaguacgag caguacauca aguggcccug guacaucugg 3960

cugggcuuca ucgccggccu gaucgccauc gugaugguga ccaucaugcu gugcugcaug 4020

accagcugcu gcagcugccu gaagggcugu ugcagcugcg gcagcugcug caaguucgac 4080

gaggacgaca gcgagcccgu gcugaagggc gugaagcugc acuacaccug auaauaggcu 4140

ggagccucgg uggccuagcu ucuugccccu ugggccuccc cccagccccu ccuccccuuc 4200

cugcacccgu acccccgugg ucuuugaaua aagucugagu gggcggc 4247

<210> 40
<211> 4071
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>

<223> 合成

<400> 40

auguuccuga uccugcugau cagccugccc accgccuucg ccgugaucgg cgaccugaag 60

ugcaccagcg acaacaucaa cgacaaggac accggcccac cacccaucag caccgacacc 120

guggacguga ccaacggccu gggcaccuac uacgugcugg accgggugua ccugaacacc 180

acccuguucc ugaacggcua cuaccccacc agcggcagca ccuaccggaa uauggcccug 240

aagggcagcg ugcugcugag ccggcugugg uucaagccac cauuccugag cgacuucauc 300

aacggaaucu ucgccaaggu gaagaacacc aaggugauca aggaccgggu gauguacagc 360

gaguuccccg ccaucaccau uggcaguacc uucgugaaca ccagcuacag cgugguggug 420

cagccccgga ccaucaacag cacccaggac ggcgacaaca agcugcaggg ccugcuggag 480

gugagcgugu gccaguacaa caugugcgag uacccucaga ccaucugcca ccccaaccug 540

ggcaaccacc ggaaggagcu guggcaccug gacaccggcg uggugagcug ccuguacaag 600

cggaacuuca ccuacgacgu aaacgccgac uaccuguacu uccacuucua ccaggagggc 660

ggcaccuucu acgccuacuu caccgacacg ggcgugguga ccaaguuccu guucaacgug 720

[0088] uaccugggca uggcccugag ccacuacuac gugaugcccc ugaccuguaa cagcaagcug 780

acccuggagu acugggugac cccucugacc agccggcagu accugcuggc cuucaaccag 840

gacggcauca ucuucaacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 900

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccaacggcua caccgugcag 960

cccaucgccg acguguaccg gcggaagccc aaccugccca acugcaacau cgaggccugg 1020

cugaacgaca agagcgugcc cucgccccug aacugggagc ggaagaccuu cagcaacugc 1080

aacuucaaca ugagcagucu gaugagcuuc auccaggccg acagcuucac cugcaacaac 1140

aucgacgccg ccaagaucua cggcaugugc uucagcagca ucaccaucga caaguuugcc 1200

aucccccaacg gccggaaggu ggaccugcag cugggcaacc ugggcuaccu gcagagcuuc 1260

aacuaccgga ucgacaccac cgccaccucu ugccagcugu acuacaaccu gcccgccgcc 1320

aacgugagcg ugagccgguu caaccccagc accuggaaca agcgguucgg cuucauugag 1380

gacagcgugu ucaaaccccg gcccgcagga guacugacca accacgacgu gguguacgcc 1440

cagcacugcu ucaaggcacc caagaacuuc ugccccugca agcugaacgg cagcugugug 1500

ggcucuggcc ccgguaagaa caacggcaua gggacuugcc cggcagggac caacuaccug 1560

accugcgaca accugugcac acccgacccc aucaccuuca ccggcaccua caaguguccc 1620

cagaccaaga gccuggуggg caucggcgag cacugcagcg gccuggccgu gaagagcgac 1680

uacugcggcg gcaacagcug caccugucgg ccccaggccu uccugggcug gagcgccgac 1740

agcugccugc agggcgacaa gugcaacauc uuugccaacu ucauccugca cgacgugaac 1800

agcggccuga ccugcagcac cgaccugcag aaggccaaca ccgacaucau ccugggcgug 1860

ugcgugaacu acgacuugua cggcauccug ggccagggca ucuucgugga ggugaacgcc 1920

accuacuaca acagcuggca gaaccugcug uacgacagca acggcaaccu guacggcuuc 1980

cgggacuaca ucaucaaccg gaccuucaug auccggagcu gcuacagcgg cggcgugagc 2040

gugaucaccc caggcaccaa caccagcaac cagguggccg ugcuguacca ggacgugaac 2100

ugcaccgagg ugcccguggc cauccacgcc gaccagcuga cacccaccug gcgggucuac 2160

agcaccggca gcaacguguu ccagacccgg gccgguugcc ugaucggcgc cgagcacgug 2220

aacaacagcu acgagugcga caucccccauc ggcgccggca ucugugccag cuaccagacc 2280

cagaccaauu cacccggag ggcaaggagc guggccagcc agagcaucau cgccuacacc 2340

augagccugg gcgccgagaa cagcguggcc uacagcaaca acagcaucgc caucccccacc 2400

aacuucacca ucagcgugac caccgagauu cugcccguga gcaugaccaa gaccagcgug 2460

[0089] gacugcacca uguacaucug cggcgacagc accgagugca gcaaccugcu gcugcaguac 2520

ggcagcuucu gcacccagcu gaaccgggcc cugaccggca ucgccgugga gcaggacaag 2580

aacacccagg agguguucgc ccaggugaag cagaucuaca agaccccucc caucaaggac 2640

uucggcggcu ucaacuucag ccagauccug cccgacccca gcaagcccag caagcggagc 2700

uucaucgagg accugcuguu caacaaggug acccuagccg acgccggcuu caucaagcag 2760

uacggcgacu gccucggcga cauagccgcc cgggaccuga ucugcgccca gaaguucaac 2820

ggccugaccg ugcugccucc ccugcugacc gacgagauga ucgcccagua caccagcgcc 2880

cuguuagccg gaaccaucac cagcggcugg acuuucggcg cuggagccgc ucugcagauc 2940

cccuucgcca ugcagauggc cuaccgguuc aacggcaucg gcgugaccca gaacgugcug 3000

uacgagaacc agaagcugau cgccaaccag uucaacagcg ccaucggcaa gauccaggac 3060

agccugagca gcaccgcuag cgcccugggc aagcugcagg acguggugaa ccagaacgcc 3120

caggcccuga acacccuggu gaagcagcug agcagcaacu ucggcgccau cagcagcgug 3180

cugaacgaca uccugagccg gcuggacccu cccgaggccg aggugcagau cgaccggcug 3240

aucacuggcc ggcugcagag ccugcagacc uacgugaccc agcagcugau ccgggccgcc 3300

gagauucggg ccagcgccaa ccuggccgcc accaagauga gcgagugcgu gcugggccag 3360

agcaagcggg uggacuucug cggcaagggc uaccaccuga ugagcuuucc ccagagcgca 3420

ccccacggag ugguguuccu gcacgugacc uacgugcccg cccaggagaa gaacuucacc 3480

accgccccag ccaucugcca cgacggcaag gcccacuuuc cccgggaggg cguguucgug 3540

agcaacggca cccacugguu cgugacccag cggaacuucu acgagcccca gaucaucacc 3600

accgacaaca ccuucgugag cggcaacugc gacgugguga ucggcaucgu gaacaacacc 3660

guguacgauc cccugcagcc cgagcuggac agcuucaagg aggagcugga caaguacuuc 3720

aagaaucaca ccagccccga cguggaccug ggcgacauca gcggcaucaa cgccagcgug 3780

gugaacaucc agaaggagau cgaucggcug aacgaggugg ccaagaaccu gaacgagagc 3840

cugaucgacc ugcaggagcu gggcaaguac gagcaguaca ucaaguggcc cugguacauc 3900

uggcugggcu ucaucgccgg ccugaucgcc aucgugaugg ugaccaucau gcugugcugc 3960

augaccagcu gcugcagcug ccugaagggc uguugcagcu gcggcagcug cugcaaguuc 4020

gacgaggacg acagcgagcc cgugcugaag ggcgugaagc ugcacuacac c 4071

[0090]

<210> 41
<211> 1357
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 41

Met Phe Leu Ile Leu Leu Ile Ser Leu Pro Thr Ala Phe Ala Val Ile
1 5 10 15

Gly Asp Leu Lys Cys Thr Ser Asp Asn Ile Asn Asp Lys Asp Thr Gly
20 25 30

Pro Pro Pro Ile Ser Thr Asp Thr Val Asp Val Thr Asn Gly Leu Gly
35 40 45

Thr Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn Thr Thr Leu Phe Leu
50 55 60

Asn Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr Arg Asn Met Ala Leu
65 70 75 80

Lys Gly Ser Val Leu Leu Ser Arg Leu Trp Phe Lys Pro Pro Phe Leu

85 90 95

Ser Asp Phe Ile Asn Gly Ile Phe Ala Lys Val Lys Asn Thr Lys Val
100 105 110

Ile Lys Asp Arg Val Met Tyr Ser Glu Phe Pro Ala Ile Thr Ile Gly
115 120 125

Ser Thr Phe Val Asn Thr Ser Tyr Ser Val Val Val Gln Pro Arg Thr
130 135 140

Ile Asn Ser Thr Gln Asp Gly Asp Asn Lys Leu Gln Gly Leu Leu Glu
145 150 155 160

Val Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr Pro Gln Thr Ile Cys
165 170 175

His Pro Asn Leu Gly Asn His Arg Lys Glu Leu Trp His Leu Asp Thr
180 185 190

Gly Val Val Ser Cys Leu Tyr Lys Arg Asn Phe Thr Tyr Asp Val Asn
[0091] 195 200 205

Ala Asp Tyr Leu Tyr Phe His Phe Tyr Gln Glu Gly Gly Thr Phe Tyr
210 215 220

Ala Tyr Phe Thr Asp Thr Gly Val Val Thr Lys Phe Leu Phe Asn Val
225 230 235 240

Tyr Leu Gly Met Ala Leu Ser His Tyr Tyr Val Met Pro Leu Thr Cys
245 250 255

Asn Ser Lys Leu Thr Leu Glu Tyr Trp Val Thr Pro Leu Thr Ser Arg
260 265 270

Gln Tyr Leu Leu Ala Phe Asn Gln Asp Gly Ile Ile Phe Asn Ala Val
275 280 285

Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser
290 295 300

Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Asn Gly Tyr Thr Val Gln
305 310 315 320

Pro Ile Ala Asp Val Tyr Arg Arg Lys Pro Asn Leu Pro Asn Cys Asn
325 330 335

Ile Glu Ala Trp Leu Asn Asp Lys Ser Val Pro Ser Pro Leu Asn Trp
340 345 350

Glu Arg Lys Thr Phe Ser Asn Cys Asn Phe Asn Met Ser Ser Leu Met
355 360 365

Ser Phe Ile Gln Ala Asp Ser Phe Thr Cys Asn Asn Ile Asp Ala Ala
370 375 380

Lys Ile Tyr Gly Met Cys Phe Ser Ser Ile Thr Ile Asp Lys Phe Ala
385 390 395 400

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

Leu Gln Ser Phe Asn Tyr Arg Ile Asp Thr Thr Ala Thr Ser Cys Gln
420 425 430

[0092]

Leu Tyr Tyr Asn Leu Pro Ala Ala Asn Val Ser Val Ser Arg Phe Asn
435 440 445

Pro Ser Thr Trp Asn Lys Arg Phe Gly Phe Ile Glu Asp Ser Val Phe
450 455 460

Lys Pro Arg Pro Ala Gly Val Leu Thr Asn His Asp Val Val Tyr Ala
465 470 475 480

Gln His Cys Phe Lys Ala Pro Lys Asn Phe Cys Pro Cys Lys Leu Asn
485 490 495

Gly Ser Cys Val Gly Ser Gly Pro Gly Lys Asn Asn Gly Ile Gly Thr
500 505 510

Cys Pro Ala Gly Thr Asn Tyr Leu Thr Cys Asp Asn Leu Cys Thr Pro
515 520 525

Asp Pro Ile Thr Phe Thr Gly Thr Tyr Lys Cys Pro Gln Thr Lys Ser
530 535 540

Leu Val Gly Ile Gly Glu His Cys Ser Gly Leu Ala Val Lys Ser Asp
545 550 555 560

Tyr Cys Gly Gly Asn Ser Cys Thr Cys Arg Pro Gln Ala Phe Leu Gly
565 570 575

Trp Ser Ala Asp Ser Cys Leu Gln Gly Asp Lys Cys Asn Ile Phe Ala
580 585 590

Asn Phe Ile Leu His Asp Val Asn Ser Gly Leu Thr Cys Ser Thr Asp
595 600 605

Leu Gln Lys Ala Asn Thr Asp Ile Ile Leu Gly Val Cys Val Asn Tyr
610 615 620

Asp Leu Tyr Gly Ile Leu Gly Gln Gly Ile Phe Val Glu Val Asn Ala
625 630 635 640

Thr Tyr Tyr Asn Ser Trp Gln Asn Leu Leu Tyr Asp Ser Asn Gly Asn
645 650 655

[0093]

Leu Tyr Gly Phe Arg Asp Tyr Ile Ile Asn Arg Thr Phe Met Ile Arg
660 665 670

Ser Cys Tyr Ser Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr
675 680 685

Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val
690 695 700

Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr
705 710 715 720

Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly
725 730 735

Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala
740 745 750

Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala
755 760 765

Arg Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly
770 775 780

Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr
785 790 795 800

Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr
805 810 815

Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu
820 825 830

Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn
835 840 845

Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu
850 855 860

Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp
865 870 875 880

[0094] Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro
885 890 895

Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu
900 905 910

Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile
915 920 925

Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val
930 935 940

Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala
945 950 955 960

Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala
965 970 975

Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly
980 985 990

Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala

995 1000 1005

Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser
1010 1015 1020

Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln
1025 1030 1035

Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn
1040 1045 1050

Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu
1055 1060 1065

Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly
1070 1075 1080

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1085 1090 1095

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
[0095] 1100 1105 1110

Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1115 1120 1125

Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1130 1135 1140

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1145 1150 1155

Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1160 1165 1170

Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1175 1180 1185

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1190 1195 1200

Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1205 1210 1215

Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1220 1225 1230

Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1235 1240 1245

Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
1250 1255 1260

Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
1265 1270 1275

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr
1280 1285 1290

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1295 1300 1305

Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met Thr Ser
1310 1315 1320

[0096]

Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys
1325 1330 1335

Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly Val Lys
1340 1345 1350

Leu His Tyr Thr
1355

<210> 42
<211> 1157
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 42
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccegguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacucu ggcggaggcg gcagcgccau cggcggcuac 960

aucccegagg ccccuagaga cggccaggcc uacgugcgga aggacggcga gugggugcug 1020

[0097] cugagcaccu uccugggcug auaauaggcu ggagccucgg uggccuagcu ucuugccccu 1080

ugggccuccc cccagccccu ccucccccuuc cugcacccgu acccccgugg ucuuugaaua 1140

aagucugagu gggcggc 1157

<210> 43
<211> 981
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 43
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aaggguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ucuggcggag gcggcagcgc caucggcggc 900

uacaucccccg aggccccuag agacggccag gccuacgugc ggaaggacgg cgagugggug 960

cugcugagca ccuuccuggg c 981

<210> 44
<211> 327
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0098]

<400> 44

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser

100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

[0099] Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Ser Gly Gly Gly Gly Ser Ala Ile Gly Gly Tyr Ile Pro Glu
290 295 300

Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val
305 310 315 320

Leu Leu Ser Thr Phe Leu Gly
325

<210> 45
<211> 1130
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 45
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

[0100] agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacucu ggcggaggca gcauccuggc caucuacagc 960

accguggcca gcagccuggu gcugcuggug agccuggggcg ccaucagcuu cugauaauag 1020

gcuggagccu cgguggccua gcuucuugcc ccuugggccu ccccccagcc ccuccucccc 1080

uuccugcacc cguacccccg uggucuuuga auaaagucug agugggcggc 1130

<210> 46
<211> 954
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 46
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

[0101] ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ucuggcggag gcagcauccu ggccaucuac 900

agcaccgugg ccagcagccu ggugcugcug gugagccugg gcgccaucag cuuc 954

<210> 47
<211> 318
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 47

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

[0102] Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

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

275 280 285

Val Asp Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala
290 295 300

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
305 310 315

<210> 48
<211> 1238
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 48
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

[0103] gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccccguu ccagaccccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacucu ggcggaggca gcauccuggc caucuacagc 960

accguggcca gcagccuggu gcugcuggug agccuggggcg ccaucagcuu cggcggaggc 1020

agcgccaucg gcggcuacau ccccgaggcc ccuagagacg gccaggccua cgugcggaag 1080

gacggcgagu gggugcugcu gagcaccuuc cugggcaagu gauaauaggc uggagccucg 1140

guggccuagc uucuugcccc uugggccucc ccccagcccc uccuccccuu ccugcacccg 1200

uacccccgug gucuuugaau aaagucugag ugggcggc 1238

<210> 49
<211> 1062
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 49
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

[0104]

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ucuggcggag gcagcauccu ggccaucuac 900

agcaccgugg ccagcagccu ggugcugcug gugagccugg gcgccaucag cuucggcgga 960

ggcagcgcca ucggcggcua caucccccgag gccccuagag acggccaggc cuacgugcgg 1020

aaggacggcg agugggugcu gcugagcacc uuccugggca ag 1062

<210> 50
<211> 354
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 50

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

[0105]

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala
290 295 300

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Gly Gly
305 310 315 320

[0106] Gly Ser Ala Ile Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln
325 330 335

Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu
340 345 350

Gly Lys

<210> 51
<211> 1130
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 51
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

gugugccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccgguu ccagaccccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacucu ggcggaggca gcauccuggc caucuacagc 960

accguggcca gcagccuggu gcugcuggug agccuggggcg ccaucagcuu cugauaauag 1020

gcuggagccu cgguggccua gcuucuugcc ccuugggccu ccccccagcc ccuccucccc 1080

[0107] uuccugcacc cguacccccg uggucuuuga auaaagucug agugggcggc 1130

<210> 52
<211> 954
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 52
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aaggugugcc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agcccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ucuggcggag gcagcauccu ggccaucuac 900

agcaccgugg ccagcagccu ggugcugcug gugagccugg gcgccaucag cuuc 954

<210> 53
<211> 318
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 53

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

[0108]

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Cys Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

[0109]

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala
290 295 300

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
305 310 315

<210> 54
<211> 1181
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 54
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

gugugccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccccguu ccagacccug 780

[0110] cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccuc uggcggaggc agcauccugg ccaucuacag caccguggcc 1020

agcagccugg ugcugcuggu gagccugggc gccaucagcu ucugauaaua ggcuggagcc 1080

ucgguggccu agcuucuugc cccuugggcc ucccccccagc cccuccuccc cuuccugcac 1140

ccguaccccc guggucuuug aauaaagucu gagugggcgg c 1181

<210> 55
<211> 1005
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 55
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aaggugugcc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgcccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cucuggcgga ggcagcaucc uggccaucua cagcaccgug 960

gccagcagcc uggugcugcu ggugagccug ggcgccauca gcuuc 1005

[0111]

<210> 56
<211> 335
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 56

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Cys Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

[0112]

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val
305 310 315 320

Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
325 330 335

<210> 57
<211> 1181
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 57

[0113]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccccguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccuc uggcggaggc agcauccugg ccaucuacag caccguggcc 1020

agcagccugg ugcugcuggu gagccugggc gccaucagcu ucugauaaua ggcuggagcc 1080

ucgguggccu agcuucuugc cccuugggcc ucccccccagc cccuccuccc cuuccugcac 1140

ccguaccccc guggucuuug aauaaagucu gagugggcgg c 1181

<210> 58
<211> 1005
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 58
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

[0114] ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgcccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cucuggcgga ggcagcaucc uggccaucua cagcaccgug 960

gccagcagcc uggugcugcu ggugagccug ggcgccauca gcuuc 1005

<210> 59
<211> 335
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 59

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val

1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
[0115] 115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val
305 310 315 320

Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
325 330 335

[0116]

<210> 60
<211> 836
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 60
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaaguga 720

uaauaggcug gagccucggu ggccuagcuu cuugccccuu gggccucccc ccagccccuc 780

cucccccuucc ugcacccgua cccccguggu cuuugaauaa agucugagug ggcggc 836

<210> 61
<211> 660
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 61
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

[0117]

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuuccccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

<210> 62
<211> 220
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 62

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

[0118] Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
210 215 220

<210> 63
<211> 1349
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 63

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaaggga 720

ggaggcagcg gcggcgauau caucaagcuu cugaacgagc aaguuaacaa ggaaaugcag 780

agcaguaauc ucuacaugag caugagcagc uggugcuaca cccacucccu ggacggagca 840

[0119] ggccucuucc uguucgacca cgcagccgag gaguacgagc acgcuaagaa guugaucauu 900

uucuugaacg agaacaacgu gcccgugcag cuaacgucaa ucagcgcacc ugagcacaag 960

uucgagggcc ugacccagau cuuccagaag gccuacgaac acgaacagca caucuccgag 1020

agcaucaaca auauugugga ucacgcuauc aaguccaagg accacgcuac cuucaacuuc 1080

cugcaguggu acguggccga gcaacaugag gaggaggugc uguucaagga cauccuggac 1140

aagaucgagc ugaucgguaa ugagaaucac ggccuguacc uggccgacca guacgugaag 1200

ggcaucgcca agagccggaa gucaggcuca ugauaauagg cuggagccuc gguggccuag 1260

cuucuugccc cuugggccuc cccccagccc cuccuccccu uccugcaccc guaccccccgu 1320

ggucuuugaa uaaagucuga gugggcggc 1349

<210> 64
<211> 1173
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 64
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuuccccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

ggaggaggca gcggcggcga uaucaucaag cuucugaacg agcaaguuaa caaggaaaug 720

cagagcagua aucucuacau gagcaugagc agcuggugcu acacccacuc ccuggacgga 780

gcaggccucu uccuguucga ccacgcagcc gaggaguacg agcacgcuaa gaaguugauc 840

auuuucuuga acgagaacaa cgugcccgug cagcuaacgu caaucagcgc accugagcac 900

[0120]

aaguucgagg gccugaccca gaucuuccag aaggccuacg aacacgaaca gcacaucucc 960

gagagcauca acaauauugu ggaucacgcu aucaagucca aggaccacgc uaccuucaac 1020

uuccugcagu gguacguggc cgagcaacau gaggaggagg ugcuguucaa ggacauccug 1080

gacaagaucg agcugaucgg uaaugagaau cacggccugu accuggccga ccaguacgug 1140

aagggcaucg ccaagagccg gaagucaggc uca 1173

<210> 65
<211> 391
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 65

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

[0121]

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Gly Gly Gly Ser
210 215 220

Gly Gly Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met
225 230 235 240

Gln Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys Tyr Thr His
245 250 255

Ser Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala Ala Glu Glu
260 265 270

Tyr Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val
275 280 285

Pro Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly
290 295 300

Leu Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln His Ile Ser
305 310 315 320

Glu Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His
325 330 335

Ala Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu
340 345 350

Glu Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn
355 360 365

[0122] Glu Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala
370 375 380

Lys Ser Arg Lys Ser Gly Ser
385 390

<210> 66
<211> 1376
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 66
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

ggcauccugc ccagcccugg caugcccgcu cugcugagcc uggugagccu gcugagcgug 120

cugcugaugg gcugcguggc ugagaccggc augcagaucu acgagggcaa gcugaccgca 180

gagggccugc gguucggcau cguggccagc cgcgccaacc acgcucuggu ggaccggcuu 240

guggagggcg cuaucgacgc caucgugaga cacggcggcc gggaagagga caucacccug 300

gugcgggugu gcggcagcug ggagauuccc gucgccgccg gagaacuggc ccggaaggag 360

gacaucgacg ccgugaucgc caucggcgug cugugcagag gcgccacgcc cagcuucgac 420

uacaucgcca gcgaggugag caagggccug gccgaccuga gccuggagcu gcggaagccc 480

aucaccuucg gcgugaucac cgccgacacc cuggagcagg ccaucgaggc cgcaggcacc 540

ugccacggca acaagggcug ggaagccgcc cugugcgcca ucgagauggc caaccuguuc 600

aagagccugc ggggcggaag uggaggcucu gguggcagcg gaggaucugg cggcggccag 660

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 720

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 780

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 840

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 900

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 960

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 1020

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 1080

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 1140

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 1200

[0123] gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaaguga 1260

uaauaggcug gagccucggu ggccuagcuu cuugccccuu gggccucccc ccagccccuc 1320

cucccccuucc ugcacccgua cccccguggu cuuugaauaa agucugagug ggcggc 1376

<210> 67
<211> 1200
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 67
augggcaucc ugcccagccc uggcaugccc gcucugcuga gccuggugag ccugcugagc 60

gugcugcuga ugggcugcgu ggcugagacc ggcaugcaga ucuacgaggg caagcugacc 120

gcagagggcc ugcgguucgg caucguggcc agccgcgcca accacgcucu gguggaccgg 180

cuuguggagg gcgcuaucga cgccaucgug agacacggcg gccgggaaga ggacaucacc 240

cuggugcggg ugugcggcag cugggagauu cccgucgccg ccggagaacu ggcccggaag 300

gaggacaucg acgccgugau cgccaucggc gugcugugca gaggcgccac gcccagcuuc 360

gacuacaucg ccagcgaggu gagcaagggc cuggccgacc ugagccugga gcugcggaag 420

cccaucaccu ucggcgugau caccgccgac acccuggagc aggccaucga ggccgcaggc 480

accugccacg gcaacaaggg cugggaagcc gcccugugcg ccaucgagau ggccaaccug 540

uucaagagcc ugcggggcgg aaguggaggc ucugguggca gcggaggauc uggcggcggc 600

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 660

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 720

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 780

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 840

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 900

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 960

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 1020

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 1080

uacuucccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 1140

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 1200

[0124] <210> 68
<211> 400
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 68

Met Gly Ile Leu Pro Ser Pro Gly Met Pro Ala Leu Leu Ser Leu Val
1 5 10 15

Ser Leu Leu Ser Val Leu Leu Met Gly Cys Val Ala Glu Thr Gly Met
20 25 30

Gln Ile Tyr Glu Gly Lys Leu Thr Ala Glu Gly Leu Arg Phe Gly Ile
35 40 45

Val Ala Ser Arg Ala Asn His Ala Leu Val Asp Arg Leu Val Glu Gly
50 55 60

Ala Ile Asp Ala Ile Val Arg His Gly Gly Arg Glu Glu Asp Ile Thr
65 70 75 80

Leu Val Arg Val Cys Gly Ser Trp Glu Ile Pro Val Ala Ala Gly Glu
85 90 95

Leu Ala Arg Lys Glu Asp Ile Asp Ala Val Ile Ala Ile Gly Val Leu
100 105 110

Cys Arg Gly Ala Thr Pro Ser Phe Asp Tyr Ile Ala Ser Glu Val Ser
115 120 125

Lys Gly Leu Ala Asp Leu Ser Leu Glu Leu Arg Lys Pro Ile Thr Phe
130 135 140

Gly Val Ile Thr Ala Asp Thr Leu Glu Gln Ala Ile Glu Ala Ala Gly
145 150 155 160

Thr Cys His Gly Asn Lys Gly Trp Glu Ala Ala Leu Cys Ala Ile Glu
165 170 175

Met Ala Asn Leu Phe Lys Ser Leu Arg Gly Gly Ser Gly Gly Ser Gly
180 185 190

[0125]

Gly Ser Gly Gly Ser Gly Gly Gly Gln Pro Asn Ile Thr Asn Leu Cys
195 200 205

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
210 215 220

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
225 230 235 240

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

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
260 265 270

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
275 280 285

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
290 295 300

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
305 310 315 320

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
325 330 335

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
340 345 350

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
355 360 365

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
370 375 380

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
385 390 395 400

[0126]

<210> 69
<211> 1334
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 69
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaaggga 720

ggaggcuccg gaggcgguag cgcugagacc ggcaugcaga ucuacgaggg caagcugacc 780

gcagagggcc ugcgguucgg caucguggcc agccgcgcca accacgcucu gguggaccgg 840

cuuguggagg gcgcuaucga cgccaucgug agacacggcg gccgggaaga ggacaucacc 900

cuggugcggg ugugcggcag cugggagauu cccgucgccg ccggagaacu ggcccggaag 960

gaggacaucg acgccgugau cgccaucggc gugcugugca gaggcgccac gcccagcuuc 1020

gacuacaucg ccagcgaggu gagcaagggc cuggccgacc ugagccugga gcugcggaag 1080

cccaucaccu ucggcgugau caccgccgac acccuggagc aggccaucga ggccgcaggc 1140

accugccacg gcaacaaggg cugggaagcc gcccugugcg ccaucgagau ggccaaccug 1200

uucaagagcc ugcggugaua auaggcugga gccucggugg ccuagcuucu ugccccuugg 1260

gccucccccc agccccuccu ccccuuccug cacccguacc cccguggucu uugaauaaag 1320

ucugaguggg cggc 1334

<210> 70
<211> 1158
<212> RNA
<213> 人工序列(Artificial Sequence)

[0127] <220>
<223> 合成

<400> 70
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuucccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

ggaggaggcu ccggaggcgg uagcgcugag accggcaugc agaucuacga gggcaagcug 720

accgcagagg gccugcgguu cggcaucgug gccagccgcg ccaaccacgc ucugguggac 780

cggcuugugg agggcgcuau cgacgccauc gugagacacg gcggccggga agaggacauc 840

acccuggugc gggugugcgg cagcugggag auuccgucg ccgccggaga acuggcccgg 900

aaggaggaca ucgacgccgu gaucgccauc ggcgugcugu gcagaggcgc cacgcccagc 960

uucgacuaca ucgccagcga ggugagcaag ggccuggccg accugagccu ggagcugcgg 1020

aagcccauca ccuucggcgu gaucaccgcc gacacccugg agcaggccau cgaggccgca 1080

ggcaccugcc acggcaacaa gggcugggaa gccgcccugu gcgccaucga gauggccaac 1140

cuguucaaga gccugcgg 1158

<210> 71
<211> 386
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 71

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

[0128]

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Gly Gly Gly Ser
210 215 220

Gly Gly Gly Ser Ala Glu Thr Gly Met Gln Ile Tyr Glu Gly Lys Leu
225 230 235 240

[0129]

Thr Ala Glu Gly Leu Arg Phe Gly Ile Val Ala Ser Arg Ala Asn His
245 250 255

Ala Leu Val Asp Arg Leu Val Glu Gly Ala Ile Asp Ala Ile Val Arg
260 265 270

His Gly Gly Arg Glu Glu Asp Ile Thr Leu Val Arg Val Cys Gly Ser
275 280 285

Trp Glu Ile Pro Val Ala Ala Gly Glu Leu Ala Arg Lys Glu Asp Ile
290 295 300

Asp Ala Val Ile Ala Ile Gly Val Leu Cys Arg Gly Ala Thr Pro Ser
305 310 315 320

Phe Asp Tyr Ile Ala Ser Glu Val Ser Lys Gly Leu Ala Asp Leu Ser
325 330 335

Leu Glu Leu Arg Lys Pro Ile Thr Phe Gly Val Ile Thr Ala Asp Thr
340 345 350

Leu Glu Gln Ala Ile Glu Ala Ala Gly Thr Cys His Gly Asn Lys Gly
355 360 365

Trp Glu Ala Ala Leu Cys Ala Ile Glu Met Ala Asn Leu Phe Lys Ser
370 375 380

Leu Arg
385

<210> 72
<211> 947
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 72

[0130]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaagucu 720

ggcggaggcg gcagcgccau cggcggcuac aucccgagg ccccuagaga cggccaggcc 780

uacgugcgga aggacggcga gugggugcug cugagcaccu uccugggcug auaauaggcu 840

ggagccucgg uggccuagcu ucuugccccu ugggccuccc cccagccccu ccuccccuuc 900

cugcacccgu acccccgugg ucuuugaaua aagucugagu gggcggc 947

<210> 73
<211> 771
<212> RNA

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 73
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuucccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

[0131] ucuggcggag gcggcagcgc caucggcggc uacaucccCg aggccccuag agacggccag 720

gccuacgugc ggaaggacgg cgagugggug cugcugagca ccuuccuggg c 771

<210> 74
<211> 257
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 74

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

[0132]

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Ser Gly Gly Gly
210 215 220

Gly Ser Ala Ile Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln
225 230 235 240

Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu
245 250 255

Gly

<210> 75
<211> 920
<212> RNA

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 75
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

[0133] gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaagucu 720

ggcggaggca gcauccuggc caucuacagc accguggcca gcagccuggu gcugcuggug 780

agccugggcg ccaucagcuu cugauaauag gcuggagccu cgguggccua gcuucuugcc 840

ccuugggccu ccccccagcc ccuccucccc uuccugcacc cguacccccg uggucuuuga 900

auaaagucug agugggcggc 920

<210> 76
<211> 744
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 76
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuuccccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cgggugguggg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

ucuggcggag gcagcauccu ggccaucuac agcaccgugg ccagcagccu ggugcugcug 720

gugagccugg gcgccaucag cuuc 744

<210> 77
<211> 248
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 77

[0134]

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

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

115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Ser Gly Gly Gly
210 215 220

[0135] Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu
225 230 235 240

Val Ser Leu Gly Ala Ile Ser Phe
245

<210> 78
<211> 1025
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 78
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaagucu 720

ggcggaggcg gcagcgccau cggcggcuac aucccccgagg ccccuagaga cggccaggcc 780

uacgugcgga aggacggcga gugggugcug cugagcaccu uccugggcgg aggcagcauc 840

cuggccaucu acagcaccgu ggccagcagc cuggugcugc uggugagccu gggcgccauc 900

agcuucugau aauaggcugg agccucggug gccuagcuuc uugccccuug ggccucccccc 960

cagccccucc uccccuuccu gcacccguac ccccgugguc uuugaauaaa gucugagugg 1020

gcggc 1025

<210> 79
<211> 849
<212> RNA
<213> 人工序列(Artificial Sequence)

[0136]

<220>
<223> 合成

<400> 79
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuucccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

ucuggcggag gcggcagcgc caucggcggc uacaucccccg aggccccuag agacggccag 720

gccuacgugc ggaaggacgg cgagugggug cugcugagca ccuuccuggg cggaggcagc 780

auccuggcca ucuacagcac cguggccagc agccuggugc ugcuggugag ccugggcgcc 840

aucagcuuc 849

<210> 80
<211> 283
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 80

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

[0137]

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Ser Gly Gly Gly
210 215 220

Gly Ser Ala Ile Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln
225 230 235 240

Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu
245 250 255

Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu
260 265 270

[0138] Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
275 280

<210> 81
<211> 1028
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 81
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaagucu 720

ggcggaggca gcauccuggc caucuacagc accguggcca gcagccuggu gcugcuggug 780

agccuggggcg ccaucagcuu cggcggaggc agcgccaucg gcggcuacau ccccgaggcc 840

ccuagagacg gccaggccua cgugcggaag gacggcgagu gggugcugcu gagcaccuuc 900

cugggcaagu gauaauaggc uggagccucg guggccuagc uucuugcccc uugggccucc 960

ccccagcccc uccucccccuu ccugcacccg uacccccgug gucuuugaau aaagucugag 1020

ugggcggc 1028

<210> 82
<211> 852
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0139]

<400> 82
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuuccccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

ucuggcggag gcagcauccu ggccaucuac agcaccgugg ccagcagccu ggugcugcug 720

gugagccugg gcgccaucag cuucggcgga ggcagcgcca ucggcggcua caucccccgag 780

gccccuagag acggccaggc cuacgugcgg aaggacggcg agugggugcu gcugagcacc 840

uuccugggca ag 852

<210> 83
<211> 284
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 83

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

[0140]

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Ser Gly Gly Gly
210 215 220

Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu
225 230 235 240

Val Ser Leu Gly Ala Ile Ser Phe Gly Gly Gly Ser Ala Ile Gly Gly
245 250 255

Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp
260 265 270

Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly Lys
275 280

[0141]

<210> 84
<211> 1142
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 84
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cgggugcagc ccaccgagag caucgugcgg uuccccaaca ucaccaaccu gugccccuuc 180

ggcgaggugu ucaacgccac ccgguucgcc agcguguacg ccuggaaccg gaagcggauc 240

agcaacugcg uggccgacua cagcgugcug uacaacagcg ccagcuucag caccuucaag 300

ugcuacggcg ugagccccac caagcugaac gaccugugcu ucaccaacgu guacgccgac 360

agcuucguga uccguggcga cgaggugcgg cagaucgcac ccggccagac aggcaagauc 420

gccgacuaca acuacaagcu gcccgacgac uucaccggcu gcgugaucgc cuggaacagc 480

aacaaccucg acagcaaggu gggcggcaac uacaacuacc uguaccggcu guuccggaag 540

agcaaccuga agcccuucga gcgggacauc agcaccgaga ucuaccaagc cggcuccacc 600

ccuugcaacg gcguggaggg cuucaacugc uacuucccuc ugcagagcua cggcuuccag 660

cccaccaacg gcgugggcua ccagcccuac cggguggugg ugcugagcuu cgagcugcug 720

cacgccccag ccaccgugug uggccccaag aagagcacca accuggugaa gaacaagugc 780

gugaacuuca acuucaacgg ccuuaccggc accggcgugc ugaccgagag caacaagaaa 840

uuccugcccu uucagcaguu cggccgggac aucgccgaca ccaccgacgc ugugcgggau 900

ccccagaccc uggagauccu ggacaucacc ccuugcagcu cuggcggagg cagcauccug 960

gccaucuaca gcaccguggc cagcagccug gugcugcugg ugagccuggg cgccaucagc 1020

uucugauaau aggcuggagc cucgguggcc uagcuucuug ccccuugggc cucccccccag 1080

ccccuccucc ccuuccugca cccguacccc cguggucuuu gaauaaaguc ugagugggcg 1140

gc 1142

<210> 85
<211> 966
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
[0142] <223> 合成

<400> 85
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcgggugc agcccaccga gagcaucgug cgguucccca acaucaccaa ccugugcccc 120

uucggcgagg uguucaacgc cacccgguuc gccagcgugu acgccuggaa ccggaagcgg 180

aucagcaacu gcguggccga cuacagcgug cuguacaaca gcgccagcuu cagcaccuuc 240

aagugcuacg gcgugagccc caccaagcug aacgaccugu gcuucaccaa cguguacgcc 300

gacagcuucg ugauccgugg cgacgaggug cggcagaucg cacccggcca gacaggcaag 360

aucgccgacu acaacuacaa gcugcccgac gacuucaccg gcugcgugau cgccuggaac 420

agcaacaacc ucgacagcaa ggugggcggc aacuacaacu accuguaccg gcuguuccgg 480

aagagcaacc ugaagcccuu cgagcgggac aucagcaccg agaucuacca agccggcucc 540

accccuugca acggcgugga gggcuucaac ugcuacuucc cucugcagag cuacggcuuc 600

cagcccacca acggcguggg cuaccagccc uaccgggugg uggugcugag cuucgagcug 660

cugcacgccc cagccaccgu guguggcccc aagaagagca ccaaccuggu gaagaacaag 720

ugcgugaacu ucaacuucaa cggccuuacc ggcaccggcg ugcugaccga gagcaacaag 780

aaauuccugc ccuuucagca guucggccgg gacaucgccg acaccaccga cgcugugcgg 840

gaucccccaga cccuggagau ccuggacauc accccuugca gcucuggcgg aggcagcauc 900

cuggccaucu acagcaccgu ggccagcagc cuggugcugc uggugagccu gggcgccauc 960

agcuuc 966

<210> 86
<211> 322
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 86

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Arg Val Gln Pro Thr Glu Ser Ile Val Arg Phe
20 25 30

Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr
[0143] 35 40 45

Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys
50 55 60

Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe
65 70 75 80

Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr
85 90 95

Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln
100 105 110

Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu
115 120 125

Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu
130 135 140

Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg
145 150 155 160

Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr
165 170 175

Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr
180 185 190

Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr
195 200 205

Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro
210 215 220

Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys Asn Lys
225 230 235 240

Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr
245 250 255

Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg Asp Ile
260 265 270

[0144]

Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu Ile Leu
275 280 285

Asp Ile Thr Pro Cys Ser Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr
290 295 300

Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile
305 310 315 320

Ser Phe

<210> 87
<211> 1142
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 87
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cgggugcagc ccaccgagag caucgugcgg uuccccaaca ucaccaaccu gugccccuuc 180

ggcgaggugu ucaacgccac ccgguucgcc agcguguacg ccuggaaccg gaagcggauc 240

agcaacugcg uggccgacua cagcgugcug uacaacagcg ccagcuucag caccuucaag 300

ugcuacggcg ugagccccac caagcugaac gaccugugcu ucaccaacgu guacgccgac 360

agcuucguga uccguggcga cgaggugcgg cagaucgcac ccggccagac aggcaagauc 420

gccgacuaca acuacaagcu gcccgacgac uucaccggcu gcgugaucgc cuggaacagc 480

aacaaccucg acagcaaggu gggcggcaac uacaacuacc uguaccggcu guuccggaag 540

agcaaccuga agcccuucga gcgggacauc agcaccgaga ucuaccaagc cggcuccacc 600

ccuugcaacg gcguggaggg cuucaacugc uacuucccuc ugcagagcua cggcuuccag 660

cccaccaacg gcgugggcua ccagcccuac cggguggugg ugcugagcuu cgagcugcug 720

cacgccccag ccaccgugug uggccccaag aagagcacca accuggugaa gaacaagugc 780

gugaacuuca acuucaacgg ccuuaccggc accggcgugc ugaccgagag caacaagaaa 840

uuccugcccu uuugccaguu cggccgggac aucgccgaca ccaccgacgc ugugcgggau 900

[0145] ccccagaccc uggagauccu ggacaucacc ccuugcagcu cuggcggagg cagcauccug 960

gccaucuaca gcaccguggc cagcagccug gugcugcugg ugagccuggg cgccaucagc 1020

uucugauaau aggcuggagc cucgguggcc uagcuucuug ccccuugggc cucccccccag 1080

ccccuccucc ccuuccugca cccguacccc cguggucuuu gaauaaaguc ugaguggggcg 1140

gc 1142

<210> 88
<211> 966
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 88
auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcgggugc agcccaccga gagcaucgug cgguucccca acaucaccaa ccugugcccc 120

uucggcgagg uguucaacgc cacccgguuc gccagcgugu acgccuggaa ccggaagcgg 180

aucagcaacu gcguggccga cuacagcgug cuguacaaca gcgccagcuu cagcaccuuc 240

aagugcuacg gcgugagccc caccaagcug aacgaccugu gcuucaccaa cguguacgcc 300

gacagcuucg ugauccgugg cgacgaggug cggcagaucg cacccggcca gacaggcaag 360

aucgccgacu acaacuacaa gcugcccgac gacuucaccg gcugcgugau cgccuggaac 420

agcaacaacc ucgacagcaa ggugggcggc aacuacaacu accuguaccg gcuguuccgg 480

aagagcaacc ugaagcccuu cgagcgggac aucagcaccg agaucuacca agccggcucc 540

accccuugca acggcgugga gggcuucaac ugcuacuucc cucugcagag cuacggcuuc 600

cagcccacca acggcguggg cuaccagccc uaccgggugg uggugcugag cuucgagcug 660

cugcacgccc cagccaccgu guguggcccc aagaagagca ccaaccuggu gaagaacaag 720

ugcgugaacu ucaacuucaa cggccuuacc ggcaccggcg ugcugaccga gagcaacaag 780

aaauuccugc ccuuuugcca guucggccgg gacaucgccg acaccaccga cgcugugcgg 840

gaucccccaga cccuggagau ccuggacauc accccuugca gcucuggcgg aggcagcauc 900

cuggccaucu acagcaccgu ggccagcagc cuggugcugc uggugagccu gggcgccauc 960

agcuuc 966

[0146]

<210> 89
<211> 322
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 89

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Arg Val Gln Pro Thr Glu Ser Ile Val Arg Phe
20 25 30

Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr
35 40 45

Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys
50 55 60

Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe
65 70 75 80

Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr

85 90 95

Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln
100 105 110

Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu
115 120 125

Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu
130 135 140

Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg
145 150 155 160

Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr
165 170 175

Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr
180 185 190

Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr
[0147] 195 200 205

Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro
210 215 220

Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys Asn Lys
225 230 235 240

Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr
245 250 255

Glu Ser Asn Lys Lys Phe Leu Pro Phe Cys Gln Phe Gly Arg Asp Ile
260 265 270

Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu Ile Leu
275 280 285

Asp Ile Thr Pro Cys Ser Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr
290 295 300

Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile
305 310 315 320

Ser Phe

<210> 90
<211> 1748
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 90

[0148]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcagc 1560

auccuggcca ucuacagcac cguggccagc agccuggugc ugcuggugag ccugggcgcc 1620

aucagcuucu gauaauaggc uggagccucg guggccuagc uucuugcccc uugggccucc 1680

ccccagcccc uccucccccuu ccugcacccg uacccccgug gucuuugaau aaagucugag 1740

ugggcggc 1748

<210> 91
<211> 1572
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 91
[0149] auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

agcauccugg ccaucuacag caccguggcc agcagccugg ugcugcuggu gagccugggc 1560

gccaucagcu uc 1572

[0150]

<210> 92
<211> 524
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 92

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu

85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

[0151] Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

[0152]

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser
500 505 510

Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
515 520

<210> 93
<211> 1925
<212> RNA

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 93

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

[0153] aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcagc 1560

auccuggcca ucuacagcac cguggccagc agccuggugc ugcuggugag ccugggcgcc 1620

aucagcuuca agaagaagaa gcggccacgg aacuccuaca agugcggcac caacaccaug 1680

gagcgggagg agagcgagca gaccaagaag cgggagaaga uccacauucc ugaacggucc 1740

gacgaagccc agcggguguu caagagcagc aagaccagca gcugcgacaa gagcgacacc 1800

ugcuucugau aauaggcugg agccucggug gccuagcuuc uugccccuug ggccucccca 1860

cagccccucc uccccuuccu gcacccguac ccccgugguc uuugaauaaa gucugagugg 1920

gcggc 1925

<210> 94
<211> 1749
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 94
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

[0154]

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

agcauccugg ccaucuacag caccguggcc agcagccugg ugcugcuggu gagccugggc 1560

gccaucagcu ucaagaagaa gaagcggcca cggaacuccu acaagugcgg caccaacacc 1620

auggagcggg aggagagcga gcagaccaag aagcgggaga agauccacau uccugaacgg 1680

uccgacgaag cccagcgggu guucaagagc agcaagacca gcagcugcga caagagcgac 1740

accugcuuc 1749

[0155]

<210> 95
<211> 583
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 95

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

[0156]

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

[0157]

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser
500 505 510

Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Lys Lys Lys Lys
515 520 525

Arg Pro Arg Asn Ser Tyr Lys Cys Gly Thr Asn Thr Met Glu Arg Glu
530 535 540

Glu Ser Glu Gln Thr Lys Lys Arg Glu Lys Ile His Ile Pro Glu Arg
545 550 555 560

Ser Asp Glu Ala Gln Arg Val Phe Lys Ser Ser Lys Thr Ser Ser Cys
565 570 575

Asp Lys Ser Asp Thr Cys Phe
580

<210> 96
<211> 1805
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 96

[0158]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcagc 1560

auccuggcca ucuacagcac cguggccagc agccuggugc ugcuggugag ccugggcgcc 1620

aucagcuuca agcggcagua caaggacaug augagcgagg gaggaccacc uggcgcugag 1680

ccacagugau aauaggcugg agccucggug gccuagcuuc uugccccuug ggccucccccc 1740

cagccccucc uccccuuccu gcacccguac ccccgugguc uuugaauaaa gucugagugg 1800

[0159] gcggc 1805

<210> 97
<211> 1629
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 97
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

[0160]

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

agcauccugg ccaucuacag caccguggcc agcagccugg ugcugcuggu gagccugggc 1560

gccaucagcu ucaagcggca guacaaggac augaugagcg agggaggacc accuggcgcu 1620

gagccacag 1629

<210> 98
<211> 543
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 98

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

[0161]

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

[0162] Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys

485 490 495

Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser
500 505 510

Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Lys Arg Gln Tyr
515 520 525

Lys Asp Met Met Ser Glu Gly Gly Pro Pro Gly Ala Glu Pro Gln
530 535 540

<210> 99
<211> 1853
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 99
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

[0163] acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccccguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcggc 1560

agcgccaucg gcggcuacau ccccgaggcc ccuagagacg gccaggccua cgugcggaag 1620

gacggcgagu gggugcugcu gagcaccuuc cugggcggag gcagcauccu ggccaucuac 1680

agcaccgugg ccagcagccu ggugcugcug gugagccugg gcgccaucag cuucugauaa 1740

uaggcuggag ccucgguggc cuagcuucuu gccccuuggg ccucccccca gccccuccuc 1800

cccuuccugc acccguaccc ccguggucuu ugaauaaagu cugaguggggc ggc 1853

[0164]

<210> 100
<211> 1677
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 100
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

[0165] uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

ggcagcgcca ucggcggcua caucccccgag gccccuagag acggccaggc cuacgugcgg 1560

aaggacggcg agugggugcu gcugagcacc uuccuggggcg gaggcagcau ccuggccauc 1620

uacagcaccg uggccagcag ccuggugcug cuggugagcc ugggcgccau cagcuuc 1677

<210> 101
<211> 559
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 101

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

[0166] Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

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

260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

[0167] Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Gly Ser Ala Ile Gly Gly Tyr Ile Pro Glu Ala Pro
500 505 510

Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
515 520 525

Ser Thr Phe Leu Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val
530 535 540

Ala Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
545 550 555

<210> 102
<211> 1856
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 102
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

[0168]
uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcagc 1560

auccuggcca ucuacagcac cguggccagc agccuggugc ugcuggugag ccugggcgcc 1620

aucagcuucg gcggaggcag cgccaucggc ggcuacaucc ccgaggcccc uagagacggc 1680

caggccuacg ugcggaagga cggcgagugg gugcugcuga gcaccuuccu gggcaaguga 1740

[0169] uaauaggcug gagccucggu ggccuagcuu cuugccccuu gggccucccc ccagccccuc 1800

cuccccuucc ugcacccgua cccccguggu cuuugaauaa agucugagug ggcggc 1856

<210> 103
<211> 1680
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 103
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

[0170] uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

agcauccugg ccaucuacag caccguggcc agcagccugg ugcugcuggu gagccugggc 1560

gccaucagcu ucggcggagg cagcgccauc ggcggcuaca uccccgaggc cccuagagac 1620

ggccaggccu acgugcggaa ggacggcgag ugggugcugc ugagcaccuu ccugggcaag 1680

<210> 104
<211> 560
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 104

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

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

20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

[0171] Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

[0172]

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser
500 505 510

Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Gly Gly Gly Ser
515 520 525

Ala Ile Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
530 535 540

Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly Lys
545 550 555 560

<210> 105
<211> 1826
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0173] <400> 105
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gccccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcagc 1560

auccuggcca ucuacagccu gggcuucauc gccggccuga ucgccaucgu gauggugacc 1620

aucaugcugu gcugcaugac cagcugcugc agcugccuga agggcuguug cagcugcggc 1680

[0174]

agcugcugca aguucgacga ggacgacuga uaauaggcug gagccucggu ggccuagcuu 1740

cuugccccuu gggccucccc ccagccccuc cuccccuucc ugcacccgua cccccguggu 1800

cuuugaauaa agucugagug ggcggc 1826

<210> 106
<211> 1650
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 106
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

[0175] gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

agcauccugg ccaucuacag ccugggcuuc aucgccggcc ugaucgccau cgugauggug 1560

accaucaugc ugugcugcau gaccagcugc ugcagcugcc ugaagggcug uugcagcugc 1620

ggcagcugcu gcaaguucga cgaggacgac 1650

<210> 107
<211> 550
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 107

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

[0176]

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

[0177]

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Leu Gly Phe Ile Ala
500 505 510

Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met Thr
515 520 525

Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys
530 535 540

Lys Phe Asp Glu Asp Asp
545 550

[0178]

<210> 108
<211> 1826
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 108
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccacccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcagc 1560

[0179] agcagcaucg ccagcuucuu cuucaucauc gggcugauca ucggccucuu ccuggugcug 1620

cgggugggca uccaccugug caucaagcug aagcacacca agaagagaca gaucuacacc 1680

gacaucgaga ugaaccggcu gggcaaguga uaauaggcug gagccucggu ggccuagcuu 1740

cuugccccuu gggccucccc ccagccccuc cuccccuucc ugcacccgua cccccguggu 1800

cuuugaauaa agucugagug ggcggc 1826

<210> 109
<211> 1650
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 109
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

[0180] ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

agcagcagca ucgccagcuu cuucuucauc aucgggcuga ucaucggccu cuuccuggug 1560

cugcgggugg gcauccaccu gugcaucaag cugaagcaca ccaagaagag acagaucuac 1620

accgacaucg agaugaaccg gcugggcaag 1650

<210> 110
<211> 550
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 110

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

[0181]

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

[0182] Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly

450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Ser Ser Ser Ile Ala Ser Phe Phe Phe Ile Ile Gly
500 505 510

Leu Ile Ile Gly Leu Phe Leu Val Leu Arg Val Gly Ile His Leu Cys
515 520 525

Ile Lys Leu Lys His Thr Lys Lys Arg Gln Ile Tyr Thr Asp Ile Glu
530 535 540

Met Asn Arg Leu Gly Lys
545 550

[0183]

<210> 111
<211> 1775
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 111
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucaccccgg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggacc caacaucacc 960

aaccugugcc ccuucggcga gguguucaac gccacccggu ucgccagcgu guacgccugg 1020

aaccggaagc ggaucagcaa cugcguggcc gacuacagcg ugcuguacaa cagcgccagc 1080

uucagcaccu ucaagugcua cggcgugagc cccaccaagc ugaacgaccu gugcuucacc 1140

aacguguacg ccgacagcuu cgugauccgu ggcgacgagg ugcggcagau cgcacccggc 1200

cagacaggca agaucgccga cuacaacuac aagcugcccg acgacuucac cggcugcgug 1260

aucgccugga acagcaacaa ccucgacagc aaggugggcg gcaacuacaa cuaccuguac 1320

cggcuguucc ggaagagcaa ccugaagccc uucgagcggg acaucagcac cgagaucuac 1380

caagccggcu ccaccccuug caacggcgug gagggcuuca acugcuacuu cccucugcag 1440

agcuacggcu uccagcccac caacggcgug ggcuaccagc ccuaccgggu gguggugcug 1500

[0184]

agcuucgagc ugcugcacgc cccagccacc gugugguggcc ccaagucugg cggaggcggc 1560

agcgccaucg gcggcuacau ccccgaggcc ccuagagacg gccaggccua cgugcggaag 1620

gacggcgagu gggugcugcu gagcaccuuc cugggcugau aauaggcugg agccucggug 1680

gccuagcuuc uugccccuug ggccucccccc cagccccucc uccccuuccu gcacccguac 1740

ccccgugguc uuugaauaaa gucugagugg gcggc 1775

<210> 112
<211> 1599
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 112
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acccaacauc 900

accaaccugu gccccuucgg cgagguguuc aacgccaccc gguucgccag cguguacgcc 960

uggaaccgga agcggaucag caacugcgug gccgacuaca gcgugcugua caacagcgcc 1020

agcuucagca ccuucaagug cuacggcgug agccccacca agcugaacga ccugugcuuc 1080

[0185] accaacgugu acgccgacag cuucgugauc cguggcgacg aggugcggca gaucgcaccc 1140

ggccagacag gcaagaucgc cgacuacaac uacaagcugc ccgacgacuu caccggcugc 1200

gugaucgccu ggaacagcaa caaccucgac agcaaggugg gcggcaacua caacuaccug 1260

uaccggcugu uccggaagag caaccugaag cccuucgagc gggacaucag caccgagauc 1320

uaccaagccg gcuccacccc uugcaacggc guggagggcu ucaacugcua cuucccucug 1380

cagagcuacg gcuuccagcc caccaacggc gugggcuacc agcccuaccg ggugguggug 1440

cugagcuucg agcugcugca cgccccagcc accgugugug gccccaaguc uggcggaggc 1500

ggcagcgcca ucggcggcua caucccccgag gccccuagag acggccaggc cuacgugcgg 1560

aaggacggcg agugggugcu gcugagcacc uuccugggc 1599

<210> 113
<211> 533
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 113

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

[0186] Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

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

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
305 310 315 320

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
325 330 335

[0187] Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
340 345 350

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
355 360 365

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
370 375 380

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
385 390 395 400

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
405 410 415

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
420 425 430

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
435 440 445

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
450 455 460

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
465 470 475 480

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
485 490 495

Ser Gly Gly Gly Gly Ser Ala Ile Gly Gly Tyr Ile Pro Glu Ala Pro
500 505 510

Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
515 520 525

Ser Thr Phe Leu Gly
530

<210> 114
<211> 1973
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0188]

<400> 114
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcggaucgg gaggcggaca gcgggugcag 960

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1020

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1080

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1140

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1200

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1260

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1320

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1380

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1440

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1500

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1560

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1620

[0189] aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1680

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1740

cuggagaucc uggacaucac cccuugcagc ucuggcggag gcagcauccu ggccaucuac 1800

agcaccgugg ccagcagccu ggugcugcug gugagccugg gcgccaucag cuucugauaa 1860

uaggcuggag ccucgguggc cuagcuucuu gccccuuggg ccucccccca gccccuccuc 1920

cccuuccugc acccguaccc ccguggucuu ugaauaaagu cugaguggge ggc 1973

<210> 115
<211> 1797
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 115
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcggau cgggaggcgg acagcgggug 900

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 960

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1020

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1080

[0190] ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1140

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1200

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1260

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1320

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1380

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1440

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1500

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1560

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1620

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1680

acccuggaga uccuggacau caccccuugc agcucuggcg gaggcagcau ccuggccauc 1740

uacagcaccg uggccagcag ccuggugcug cuggugagcc ugggcgccau cagcuuc 1797

<210> 116
<211> 599
<212> PRT

<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 116

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

[0191] Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Gly Gly Gly Ser Gly Gly Gly Gln Arg Val Gln Pro Thr Glu
290 295 300

Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
305 310 315 320

[0192]

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
325 330 335

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
340 345 350

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
355 360 365

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
370 375 380

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
385 390 395 400

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
405 410 415

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
420 425 430

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
435 440 445

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
450 455 460

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
465 470 475 480

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
485 490 495

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn
500 505 510

Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly
515 520 525

Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln
530 535 540

[0193]

Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln
545 550 555 560

Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Ser Gly Gly Gly Ser
565 570 575

Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val
580 585 590

Ser Leu Gly Ala Ile Ser Phe
595

<210> 117
<211> 1799
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 117
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

[0194]

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgg aggcggaucg ggaggcggac ccaacaucac caaccugugc 1020

cccuucggcg agguguucaa cgccacccgg uucgccagcg uguacgccug gaaccggaag 1080

cggaucagca acugcguggc cgacuacagc gugcuguaca acagcgccag cuucagcacc 1140

uucaagugcu acggcgugag ccccaccaag cugaacgacc ugugcuucac caacguguac 1200

gccgacagcu ucgugauccg uggcgacgag gugcggcaga ucgcacccgg ccagacaggc 1260

aagaucgccg acuacaacua caagcugccc gacgacuuca ccggcugcgu gaucgccugg 1320

aacagcaaca accucgacag caaggugggc ggcaacuaca acuaccugua ccggcuguuc 1380

cggaagagca accugaagcc cuucgagcgg gacaucagca ccgagaucua ccaagccggc 1440

uccaccccuu gcaacggcgu ggagggcuuc aacugcuacu ucccucugca gagcuacggc 1500

uuccagccca ccaacggcgu gggcuaccag cccuaccggg ugguggugcu gagcuucgag 1560

cugcugcacg ccccagccac cguguguggc cccaagucug gcggaggcag cauccuggcc 1620

aucuacagca ccguggccag cagccuggug cugcugguga gccugggcgc caucagcuuc 1680

ugauaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 1740

cuccucccca uccugcaccc guacccccgu ggucuuugaa uaaagucuga gugggcggc 1799

<210> 118
<211> 1623
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 118

[0195]

auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgcccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cggaggcgga ucgggaggcg gacccaacau caccaaccug 960

ugccccuucg gcgagguguu caacgccacc cgguucgcca gcguguacgc cuggaaccgg 1020

aagcggauca gcaacugcgu ggccgacuac agcgugcugu acaacagcgc cagcuucagc 1080

accuucaagu gcuacggcgu gagccccacc aagcugaacg accugugcuu caccaacgug 1140

uacgccgaca gcuucgugau ccguggcgac gaggugcggc agaucgcacc cggccagaca 1200

ggcaagaucg ccgacuacaa cuacaagcug cccgacgacu ucaccggcug cgugaucgcc 1260

uggaacagca acaaccucga cagcaaggug ggcggcaacu acaacuaccu guaccggcug 1320

uuccggaaga gcaaccugaa gcccuucgag cgggacauca gcaccgagau cuaccaagcc 1380

ggcuccaccc cuugcaacgg cguggagggc uucaacugcu acuucccucu gcagagcuac 1440

ggcuuccagc ccaccaacgg cgugggcuac cagcccuacc gggugguggu gcugagcuuc 1500

gagcugcugc acgccccagc caccgugugu ggccccaagu cuggcggagg cagcauccug 1560

gccaucuaca gcaccguggc cagcagccug gugcugcugg ugagccuggg cgccaucagc 1620

uuc 1623

<210> 119
<211> 541
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 119

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

[0196] Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr

145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

[0197] Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Gly Gly Gly Ser Gly Gly Gly Pro Asn Ile Thr Asn Leu
305 310 315 320

Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr
325 330 335

Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val
340 345 350

Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser
355 360 365

Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser
370 375 380

Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr
385 390 395 400

Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly
405 410 415

Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly
420 425 430

Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro
435 440 445

Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro
450 455 460

Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr
465 470 475 480

Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val
485 490 495

[0198]

Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro
500 505 510

Lys Ser Gly Gly Gly Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser
515 520 525

Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe
530 535 540

<210> 120
<211> 2024
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 120
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccegguu ccagacccug 780

cuggccccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgg aggcggaucg ggaggcggac agcgggugca gcccaccgag 1020

[0199] agcaucgugc gguuccccaa caucaccaac cugugccccu ucggcgaggu guucaacgcc 1080

acccgguucg ccagcgugua cgccuggaac cggaagcgga ucagcaacug cguggccgac 1140

uacagcgugc uguacaacag cgccagcuuc agcaccuuca agugcuacgg cgugagcccc 1200

accaagcuga acgaccugug cuucaccaac guguacgccg acagcuucgu gauccguggc 1260

gacgaggugc ggcagaucgc acccggccag acaggcaaga ucgccgacua caacuacaag 1320

cugcccgacg acuucaccgg cugcgugauc gccuggaaca gcaacaaccu cgacagcaag 1380

gugggcggca acuacaacua ccuguaccgg cuguuccgga agagcaaccu gaagcccuuc 1440

gagcgggaca ucagcaccga gaucuaccaa gccggcucca ccccuugcaa cggcguggag 1500

ggcuucaacu gcuacuuccc ucugcagagc uacggcuucc agcccaccaa cggcgugggc 1560

uaccagcccu accggguggu ggugcugagc uucgagcugc ugcacgcccc agccaccgug 1620

uguggcccca agaagagcac caaccuggug aagaacaagu gcgugaacuu caacuucaac 1680

ggccuuaccg gcaccggcgu gcugaccgag agcaacaaga aauuccugcc cuuucagcag 1740

uucggccggg acaucgccga caccaccgac gcugugcggg aucccagac ccuggagauc 1800

cuggacauca ccccuugcag cucuggcgga ggcagcaucc uggccaucua cagcaccgug 1860

gccagcagcc uggugcugcu ggugagccug ggcgccauca gcuucugaua auaggcugga 1920

gccucggugg ccuagcuucu ugccccuugg gccucccccc agccccuccu ccccuuccug 1980

cacccguacc cccguggucu uugaauaaag ucugaguggg cggc 2024

<210> 121
<211> 1848
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 121
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

[0200]

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cggaggcgga ucgggaggcg gacagcgggu gcagcccacc 960

gagagcaucg ugcgguuccc caacaucacc aaccugugcc ccuucggcga gguguucaac 1020

gccacccggu ucgccagcgu guacgccugg aaccggaagc ggaucagcaa cugcguggcc 1080

gacuacagcg ugcuguacaa cagcgccagc uucagcaccu ucaagugcua cggcgugagc 1140

cccaccaagc ugaacgaccu gugcuucacc aacguguacg ccgacagcuu cgugauccgu 1200

ggcgacgagg ugcggcagau cgcacccggc cagacaggca agaucgccga cuacaacuac 1260

aagcugcccg acgacuucac cggcugcgug aucgccugga acagcaacaa ccucgacagc 1320

aaggugggcg gcaacuacaa cuaccuguac cggcuguucc ggaagagcaa ccugaagccc 1380

uucgagcggg acaucagcac cgagaucuac caagccggcu ccaccccuug caacggcgug 1440

gagggcuuca acugcuacuu cccucugcag agcuacggcu uccagcccac caacggcgug 1500

ggcuaccagc ccuaccgggu gguggugcug agcuucgagc ugcugcacgc cccagccacc 1560

gugugguggcc ccaagaagag caccaaccug gugaagaaca agugcgugaa cuucaacuuc 1620

aacggccuua ccggcaccgg cgugcugacc gagagcaaca agaaauuccu gcccuuucag 1680

caguucggcc gggacaucgc cgacaccacc gacgcugugc gggaucccca gacccuggag 1740

auccuggaca ucaccccuug cagcucuggc ggaggcagca uccuggccau cuacagcacc 1800

guggccagca gccuggugcu gcuggugagc cugggcgcca ucagcuuc 1848

<210> 122
<211> 616
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0201]

<400> 122

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

[0202]

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Gly Gly Gly Ser Gly Gly Gly Gln Arg Val Gln Pro Thr
305 310 315 320

Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly
325 330 335

Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg
340 345 350

Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser
355 360 365

Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu
370 375 380

Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg
385 390 395 400

Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala
405 410 415

Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala
420 425 430

Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr
435 440 445

[0203]

Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp
450 455 460

Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val
465 470 475 480

Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro
485 490 495

Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe
500 505 510

Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr
515 520 525

Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr
530 535 540

Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln
545 550 555 560

Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro
565 570 575

Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Ser Gly Gly Gly
580 585 590

Ser Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu
595 600 605

Val Ser Leu Gly Ala Ile Ser Phe
610 615

<210> 123
<211> 3995
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 123

[0204]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucacccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

[0205] cuggagaucc uggacaucac cccuugcagc uucggcggcg ugagcgugau caccccaggc 1860

accaacacca gcaaccaggu ggccgugcug uaccaggacg ugaacugcac cgaggugccc 1920

guggccaucc acgccgacca gcugacaccc accuggcggg ucuacagcac cggcagcaac 1980

guguuccaga cccgggccgg uugccugauc ggcgccgagc acgugaacaa cagcuacgag 2040

ugcgacaucc ccaucggcgc cggcaucugu gccagcuacc agacccagac caauucaccc 2100

cggagggcaa ggagcguggc cagccagagc aucaucgccu acaccaugag ccugggcgcc 2160

gagaacagcg uggccuacag caacaacagc aucgccaucc ccaccaacuu caccaucagc 2220

gugaccaccg agauucugcc cgugagcaug accaagacca gcguggacug caccauguac 2280

aucugcggcg acagcaccga gugcagcaac cugcugcugc aguacggcag cuucugcacc 2340

cagcugaacc gggcccugac cggcaucgcc guggagcagg acaagaacac ccaggaggug 2400

uucgcccagg ugaagcagau cuacaagacc ccucccauca aggacuucgg cggcuucaac 2460

uucagccaga uccugcccga ccccagcaag cccagcaagc ggagcuucau cgaggaccug 2520

cuguucaaca aggugacccu agccgacgcc ggcuucauca agcaguacgg cgacugccuc 2580

ggcgacauag ccgcccggga ccugaucugc gcccagaagu ucaacggccu gaccgugcug 2640

ccucccccugc ugaccgacga gaugaucgcc caguacacca gcgcccuguu agccggaacc 2700

aucaccagcg gcuggacuuu cggcgcugga gccgcucugc agaucccccuu cgccaugcag 2760

auggccuacc gguucaacgg caucggcgug acccagaacg ugcuguacga gaaccagaag 2820

cugaucgcca accaguucaa cagcgccauc ggcaagaucc aggacagccu gagcagcacc 2880

gcuagcgccc ugggcaagcu gcaggacgug gugaaccaga acgcccaggc ccugaacacc 2940

cuggugaagc agcugagcag caacuucggc gccaucagca gcgugcugaa cgacauccug 3000

agccggcugg acaaggugga ggccgaggug cagaucgacc ggcugaucac uggccggcug 3060

cagagccugc agaccuacgu gacccagcag cugauccggg ccgccgagau ucgggccagc 3120

gccaaccugg ccgccaccaa gaugagcgag ugcgugcugg gccagagcaa gcggguggac 3180

uucugcggca agggcuacca ccugaugagc uuuccccaga gcgcacccca cggaguggug 3240

uuccugcacg ugaccuacgu gcccgcccag gagaagaacu ucaccaccgc cccagccauc 3300

ugccacgacg gcaaggccca cuuuccccgg gagggcgugu ucgugagcaa cggcacccac 3360

ugguucguga cccagcggaa cuucuacgag ccccagauca ucaccaccga caacaccuuc 3420

gugagcggca acugcgacgu ggugaucggc aucgugaaca acaccgugua cgaucccccug 3480

cagcccgagc uggacagcuu caaggaggag cuggacaagu acuucaagaa ucacaccagc 3540

[0206]

cccgacgugg accugggcga caucagcggc aucaacgcca gcguggugaa cauccagaag 3600

gagaucgauc ggcugaacga gguggccaag aaccugaacg agagccugau cgaccugcag 3660

gagcugggca aguacgagca guacaucaag uggcccuggu acaucuggcu gggcuucauc 3720

gccggccuga ucgccaucgu gauggugacc aucaugcugu gcugcaugac cagcugcugc 3780

agcugccuga agggcuguug cagcugcggc agcugcugca aguucgacga ggacgacagc 3840

gagcccgugc ugaagggcgu gaagcugcac uacaccugau aauaggcugg agccucggug 3900

gccuagcuuc uugccccuug ggccucccccc cagccccucc uccccuuccu gcacccguac 3960

ccccgugguc uuugaauaaa gucugagugg gcggc 3995

<210> 124
<211> 3819
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 124
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagcccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgcccugg acccucugag cgagaccaag 900

[0207] ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau cacccccuugc agcuucggcg gcgugagcgu gaucacccca 1800

[0208]

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 1860

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 1920

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 1980

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 2040

ccccggaggg caaggagcgu ggccagccag agcaucaucg ccuacaccau gagccugggc 2100

gccgagaaca gcguggccua cagcaacaac agcaucgcca uccccaccaa cuucaccauc 2160

agcgugacca ccgagauucu gcccgugagc augaccaaga ccagcgugga cugcaccaug 2220

uacaucugcg gcgacagcac cgagugcagc aaccugcugc ugcaguacgg cagcuucugc 2280

acccagcuga accgggcccu gaccggcauc gccguggagc aggacaagaa cacccaggag 2340

guguucgccc aggugaagca gaucuacaag accccuccca ucaaggacuu cggcggcuuc 2400

aacuucagcc agauccugcc cgaccccagc aagcccagca agcggagcuu caucgaggac 2460

cugcuguuca acaaggugac ccuagccgac gccggcuuca ucaagcagua cggcgacugc 2520

cucggcgaca uagccgcccg ggaccugauc ugcgcccaga aguucaacgg ccugaccgug 2580

cugccucccc ugcugaccga cgagaugauc gcccaguaca ccagcgcccu guuagccgga 2640

accaucacca gcggcuggac uuucggcgcu ggagccgcuc ugcagauccc cuucgccaug 2700

cagauggccu accgguucaa cggcaucggc gugacccaga acgugcugua cgagaaccag 2760

aagcugaucg ccaaccaguu caacagcgcc aucggcaaga uccaggacag ccugagcagc 2820

accgcuagcg cccugggcaa gcugcaggac guggugaacc agaacgccca ggcccugaac 2880

acccugguga agcagcugag cagcaacuuc ggcgccauca gcagcgugcu gaacgacauc 2940

cugagccggc uggacaaggu ggaggccgag gugcagaucg accggcugau cacuggccgg 3000

cugcagagcc ugcagaccua cgugacccag cagcugaucc gggccgccga gauucgggcc 3060

agcgccaacc uggccgccac caagaugagc gagugcgugc ugggccagag caagcgggug 3120

gacuucugcg gcaagggcua ccaccugaug agcuuucccc agagcgcacc ccacggagug 3180

guguuccugc acgugaccua cgugcccgcc caggagaaga acuucaccac cgccccagcc 3240

aucugccacg acggcaaggc ccacuuuccc cgggagggcg uguucgugag caacggcacc 3300

cacugguucg ugacccagcg gaacuucuac gagccccaga ucaucaccac cgacaacacc 3360

uucgugagcg gcaacugcga cguggugauc ggcaucguga acaacaccgu guacgauccc 3420

cugcagcccg agcuggacag cuucaaggag gagcuggaca aguacuucaa gaaucacacc 3480

agccccgacg uggaccuggg cgacaucagc ggcaucaacg ccagcguggu gaacauccag 3540

aaggagaucg aucggcugaa cgagguggcc aagaaccuga acgagagccu gaucgaccug 3600

caggagcugg gcaaguacga gcaguacauc aaguggcccu gguacaucug gcugggcuuc 3660

aucgccggcc ugaucgccau cgugauggug accaucaugc ugugcugcau gaccagcugc 3720

ugcagcugcc ugaagggcug uugcagcugc ggcagcugcu gcaaguucga cgaggacgac 3780

agcgagcccg ugcugaaggg cgugaagcug cacuacacc 3819

<210> 125
<211> 1273
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 125

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
[0209] 20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

[0210]

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480

[0211]

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
675 680 685

Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
690 695 700

[0212] Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile
705 710 715 720

Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val
725 730 735

Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu
740 745 750

Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr
755 760 765

Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln
770 775 780

Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe
785 790 795 800

Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
805 810 815

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

820 825 830

Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp
835 840 845

Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu
850 855 860

Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly
865 870 875 880

Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
885 890 895

Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
900 905 910

Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn
915 920 925

[0213] Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala
930 935 940

Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn
945 950 955 960

Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val
965 970 975

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

Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val
995 1000 1005

Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn
1010 1015 1020

Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys
1025 1030 1035

Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro
1040 1045 1050

Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val
1055 1060 1065

Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His
1070 1075 1080

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn
1085 1090 1095

Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln
1100 1105 1110

Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val
1115 1120 1125

Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro
1130 1135 1140

Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1145 1150 1155

[0214]

His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn
1160 1165 1170

Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu
1175 1180 1185

Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu
1190 1195 1200

Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu
1205 1210 1215

Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met
1220 1225 1230

Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys
1235 1240 1245

Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro
1250 1255 1260

Val Leu Lys Gly Val Lys Leu His Tyr Thr
1265 1270

<210> 126
<211> 8
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 126

Lys Pro Ser Lys Arg Ser Phe Ile
1 5

<210> 127
<211> 13
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0215] <400> 127

Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
1 5 10

<210> 128
<211> 9
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 128
ccrccaugg 9

<210> 129
<211> 11
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 129
gggauccuac c 11

<210> 130
<211> 9
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 130
uuauuuaww 9

<210> 131
<211> 47
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 131
gggaaauaag agagaaaaga agaguaagaa gaaauauaag agccacc 47

<210> 132
<211> 119
<212> RNA
<213> 人工序列(Artificial Sequence)

[0216]

<220>
<223> 合成

<400> 132
ugauaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60

cuccuccccu uccugcaccc guacccccgu ggucuuugaa uaaagucuga gugggcggc 119

<210> 133
<211> 15
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 133

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Gly
1 5 10 15

<210> 134
<211> 8
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 134

Gly Gly Gly Ser Gly Gly Gly Ser
1 5

<210> 135
<211> 7
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 135

Gly Gly Gly Ser Gly Gly Gly
1 5

<210> 136
<211> 4
<212> PRT
<213> 人工序列(Artificial Sequence)

[0217] <220>
<223> 合成

<400> 136

Gly Gly Gly Ser
1

<210> 137
<211> 4
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 137

Ser Gly Gly Ser
1

<210> 138
<211> 5
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>

<223> 合成

<400> 138

Ser Gly Gly Gly Ser
1 5

<210> 139
<211> 1602
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 139

[0218]

auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagcccaaca ucaccaaccu gugccccuuc ggcgaggugu ucaacgccac ccgguucgcc 120

agcguguacg ccuggaaccg gaagcggauc agcaacugcg uggccgacua cagcgugcug 180

uacaacagcg ccagcuucag caccuucaag ugcuacggcg ugagccccac caagcugaac 240

gaccugugcu ucaccaacgu guacgccgac agcuucguga uccguggcga cgaggugcgg 300

cagaucgcac ccggccagac aggcaagauc gccgacuaca acuacaagcu gcccgacgac 360

uucaccggcu gcgugaucgc cuggaacagc aacaaccucg acagcaaggu gggcggcaac 420

uacaacuacc uguaccggcu guuccggaag agcaaccuga agcccuucga gcgggacauc 480

agcaccgaga ucuaccaagc cggcuccacc ccuugcaacg gcguggaggg cuucaacugc 540

uacuuccccuc ugcagagcua cggcuuccag cccaccaacg gcgugggcua ccagcccuac 600

cggguggugg ugcugagcuu cgagcugcug cacgccccag ccaccgugug uggccccaag 660

ucuggcggag gcagcggagg cggaagccag ugcgugaacc ugaccacccg gacccagcug 720

ccaccagccu acaccaacag cuucacccgg ggcgucuacu accccgacaa gguguuccgg 780

agcagcgucc ugcacagcac ccaggaccug uuccugcccu ucuucagcaa cgugaccugg 840

uuccacgcca uccacgugag cggcaccaac ggcaccaagc gguucgacaa ccccgugcug 900

cccuucaacg acggcgugua cuucgccagc accgagaaga gcaacaucau ccggggcugg 960

aucuucggca ccacccugga cagcaagacc cagagccugc ugaucgugaa uaacgccacc 1020

aacgugguga ucaaggugug cgaguuccag uucugcaacg accccuuccu gggcguguac 1080

uaccacaaga acaacaagag cuggauggag agcgaguucc ggguguacag cagcgccaac 1140

aacugcaccu ucgaguacgu gagccagccc uuccugaugg accuggaggg caagcagggc 1200

aacuucaaga accugcggga guucguguuc aagaacaucg acggcuacuu caagaucuac 1260

agcaagcaca ccccaaucaa ccuggugcgg gaucugcccc agggcuucuc agcccuggag 1320

ccccuggugg accugcccau cggcaucaac aucacccggu uccagacccu gcuggcccug 1380

caccggagcu accugacccc aggcgacagc agcagcgggu ggacagcagg cgcggcugcu 1440

uacuacgugg gcuaccugca gccccggacc uuccugcuga aguacaacga gaacggcacc 1500

aucaccgacg ccguggacuc uggcggaggc agcauccugg ccaucuacag caccguggcc 1560

agcagccugg ugcugcuggu gagccugggc gccaucagcu uc 1602

<210> 140
<211> 534
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 140

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

[0219]

Leu Val Asn Ser Gln Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu
20 25 30

Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys
35 40 45

Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala
50 55 60

Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn
65 70 75 80

Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly
85 90 95

Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp
100 105 110

Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp
115 120 125

Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu
130 135 140

Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile
145 150 155 160

Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu
165 170 175

Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr
180 185 190

Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu
195 200 205

Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Ser Gly Gly Gly
210 215 220

Ser Gly Gly Gly Ser Gln Cys Val Asn Leu Thr Thr Arg Thr Gln Leu
225 230 235 240

[0220] Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro Asp
245 250 255

Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe Leu
260 265 270

Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser Gly
275 280 285

Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn Asp
290 295 300

Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly Trp
305 310 315 320

Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile Val
325 330 335

Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe Cys
340 345 350

Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser Trp

355 360 365

Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr Phe
370 375 380

Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln Gly
385 390 395 400

Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly Tyr
405 410 415

Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp Leu
420 425 430

Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile Gly
435 440 445

Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser Tyr
450 455 460

Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala Ala
[0221] 465 470 475 480

Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr Asn
485 490 495

Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Ser Gly Gly Gly Ser Ile
500 505 510

Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
515 520 525

Leu Gly Ala Ile Ser Phe
530

<210> 141
<211> 1778
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 141
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

[0222]

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

cccaacauca ccaaccugug ccccuucggc gagguguuca acgccacccg guucgccagc 180

guguacgccu ggaaccggaa gcggaucagc aacugcgugg ccgacuacag cgugcuguac 240

aacagcgcca gcuucagcac cuucaagugc uacggcguga gccccaccaa gcugaacgac 300

cugugcuuca ccaacgugua cgccgacagc uucgugaucc guggcgacga ggugcggcag 360

aucgcacccg gccagacagg caagaucgcc gacuacaacu acaagcugcc cgacgacuuc 420

accggcugcg ugaucgccug gaacagcaac aaccucgaca gcaagguggg cggcaacuac 480

aacuaccugu accggcuguu ccggaagagc aaccugaagc ccuucgagcg ggacaucagc 540

accgagaucu accaagccgg cuccaccccu ugcaacggcg uggagggcuu caacugcuac 600

uucccucugc agagcuacgg cuuccagccc accaacggcg ugggcuacca gcccuaccgg 660

gugguggugc ugagcuucga gcugcugcac gccccagcca ccgugugugg ccccaagucu 720

ggcggaggca gcggaggcgg aagccagugc gugaaccuga ccacccggac ccagcugcca 780

ccagccuaca ccaacagcuu cacccggggc gucuacuacc ccgacaaggu guuccggagc 840

agcguccugc acagcaccca ggaccuguuc cugcccuucu ucagcaacgu gaccugguuc 900

cacgccaucc acgugagcgg caccaacggc accaagcggu ucgacaaccc cgugcugccc 960

uucaacgacg gcguguacuu cgccagcacc gagaagagca acaucauccg gggcuggauc 1020

uucggcacca cccuggacag caagacccag agccugcuga ucgugaauaa cgccaccaac 1080

guggugauca aggugugcga guuccaguuc ugcaacgacc ccuuccuggg cguguacuac 1140

cacaagaaca acaagagcug gauggagagc gaguuccggg uguacagcag cgccaacaac 1200

ugcaccuucg aguacgugag ccagcccuuc cugauggacc uggagggcaa gcagggcaac 1260

uucaagaacc ugcgggaguu cguguucaag aacaucgacg gcuacuucaa gaucuacagc 1320

aagcacaccc caaucaaccu ggugcgggau cugccccagg gcuucucagc ccuggagccc 1380

cugguggacc ugcccaucgg caucaacauc acccgguucc agacccugcu ggcccugcac 1440

cggagcuacc ugaccccagg cgacagcagc agcgggugga cagcaggcgc ggcugcuuac 1500

uacgugggcu accugcagcc ccggaccuuc cugcugaagu acaacgagaa cggcaccauc 1560

accgacgccg uggacucugg cggaggcagc auccuggcca ucuacagcac cguggccagc 1620

agccuggugc ugcuggugag ccugggcgcc aucagcuucu gauaauaggc uggagccucg 1680

guggccuagc uucuugcccc uugggccucc ccccagcccc uccuccccuu ccugcacccg 1740

uacccccgug gucuuugaau aaagucugag ugggcggc 1778

<210> 142
<211> 1602
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 142
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

[0223] agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agcccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ggaggcgcca aguucguggc cgccuggacu 900

cugaaggccg cagccggcgg acccaacauc accaaccugu gccccuucgg cgagguguuc 960

aacgccaccc gguucgccag cguguacgcc uggaaccgga agcggaucag caacugcgug 1020

gccgacuaca gcgugcugua caacagcgcc agcuucagca ccuucaagug cuacggcgug 1080

agccccacca agcugaacga ccugugcuuc accaacgugu acgccgacag cuucgugauc 1140

cguggcgacg aggugcggca gaucgcaccc ggccagacag gcaagaucgc cgacuacaac 1200

uacaagcugc ccgacgacuu caccggcugc gugaucgccu ggaacagcaa caaccucgac 1260

agcaaggugg gcggcaacua caacuaccug uaccggcugu uccggaagag caaccugaag 1320

cccuucgagc gggacaucag caccgagauc uaccaagccg gcuccacccc uugcaacggc 1380

guggagggcu ucaacugcua cuucccucug cagagcuacg gcuuccagcc caccaacggc 1440

gugggcuacc agcccuaccg ggugguggug cugagcuucg agcugcugca cgccccagcc 1500

accgugugug gccccaaguc uggcggaggc agcauccugg ccaucuacag caccguggcc 1560

agcagccugg ugcugcuggu gagccugggc gccaucagcu uc 1602

<210> 143
<211> 534
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 143

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

[0224] Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr

145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

[0225] Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Gly Gly Ala Lys Phe Val Ala Ala Trp Thr Leu Lys Ala Ala
290 295 300

Ala Gly Gly Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe
305 310 315 320

Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile
325 330 335

Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe
340 345 350

Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu
355 360 365

Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu
370 375 380

Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn
385 390 395 400

Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser
405 410 415

Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg
420 425 430

Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr
435 440 445

Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe
450 455 460

Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly
465 470 475 480

Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu
485 490 495

[0226]

His Ala Pro Ala Thr Val Cys Gly Pro Lys Ser Gly Gly Gly Ser Ile
500 505 510

Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
515 520 525

Leu Gly Ala Ile Ser Phe
530

<210> 144
<211> 1778
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 144
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacgga ggcgccaagu ucguggccgc cuggacucug 960

aaggccgcag ccggcggacc caacaucacc aaccugugcc ccuucggcga gguguucaac 1020

[0227] gccaccggu ucgccagcgu guacgccugg aaccggaagc ggaucagcaa cugcguggcc 1080

gacuacagcg ugcuguacaa cagcgccagc uucagcaccu ucaagugcua cggcgugagc 1140

cccaccaagc ugaacgaccu gugcuucacc aacguguacg ccgacagcuu cgugauccgu 1200

ggcgacgagg ugcggcagau cgcacccggc cagacaggca agaucgccga cuacaacuac 1260

aagcugcccg acgacuucac cggcugcgug aucgccugga acagcaacaa ccucgacagc 1320

aaggugggcg gcaacuacaa cuaccuguac cggcuguucc ggaagagcaa ccugaagccc 1380

uucgagcggg acaucagcac cgagaucuac caagccggcu ccaccccuug caacggcgug 1440

gagggcuuca acugcuacuu cccucugcag agcuacggcu uccagcccac caacggcgug 1500

ggcuaccagc ccuaccgggu gguggugcug agcuucgagc ugcugcacgc cccagccacc 1560

gugugguggcc ccaagucugg cggaggcagc auccuggcca ucuacagcac cguggccagc 1620

agccuggugc ugcuggugag ccugggcgcc aucagcuucu gauaauaggc uggagccucg 1680

guggccuagc uucuugcccc uugggccucc ccccagcccc uccucccccuu ccugcacccg 1740

uacccccgug gucuuugaau aaagucugag ugggcggc 1778

<210> 145
<211> 1785

<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 145

[0228]

auguacagca ugcagcuggc uagcugcgug acccugaccc uggugcugcu ggugaacagc 60

cagggcgccg agaacagcgu ggccuacagc aacaacagca ucgccauccc caccaacuuc 120

accaucagcg ugaccaccga gauucugccc gugagcauga ccaagaccag cguggacugc 180

accauguaca ucugcggcga cagcaccgag ugcagcaacc ugcugcugca guacggcagc 240

uucugcaccc agcugaaccg ggcccugacc ggcaucgccg uggagcagga caagaacacc 300

caggaggugu ucgcccaggu gaagcagauc uacaagaccc cucccaucaa ggacuucggc 360

ggcuucaacu ucagccagau ccugcccgac cccagcaagc ccagcaagcg gagcuucauc 420

gaggaccugc uguucaacaa ggugacccua gccgacgccg gcuucaucaa gcaguacggc 480

gacugccucg gcgacauagc cgcccgggac cugaucugcg cccagaaguu caacggccug 540

accgugcugc cucccccugcu gaccgacgag augaucgccc aguacaccag cgcccuguua 600

gccggaacca ucaccagcgg cuggacuuuc ggcgcuggag ccgcucugca gaucccccuuc 660

gccaugcaga uggccuaccg guucaacggc aucggcguga cccagaacgu gcuguacgag 720

aaccagaagc ugaucgccaa ccaguucaac agcgccaucg gcaagaucca ggacagccug 780

agcagcaccg cuagcgcccu gggcaagcug caggacgugg ugaaccagaa cgcccaggcc 840

cugaacaccc uggugaagca gcugagcagc aacuucggcg ccaucagcag cgugcugaac 900

gacauccuga gccggcugga cccucccgag gccgaggugc agaucgaccg gcugaucacu 960

ggccggcugc agagccugca gaccuacgug acccagcagc ugauccgggc cgccgagauu 1020

cgggccagcg ccaaccuggc cgccaccaag augagcgagu gcgugcuggg ccagagcaag 1080

cggguggacu ucugcggcaa gggcuaccac cugaugagcu uuccccagag cgcaccccac 1140

ggaguggugu uccugcacgu gaccuacgug cccgcccagg agaagaacuu caccaccgcc 1200

ccagccaucu gccacgacgg caaggcccac uuuccccggg agggcguguu cgugagcaac 1260

ggcacccacu gguucgugac ccagcggaac uucuacgagc cccagaucau caccaccgac 1320

aacaccuucg ugagcggcaa cugcgacgug gugaucggca ucgugaacaa caccguguac 1380

gaucccccugc agcccgagcu ggacagcuuc aaggaggagc uggacaagua cuucaagaau 1440

cacaccagcc ccgacgugga ccugggcgac aucagcggca ucaacgccag cguggugaac 1500

auccagaagg agaucgaucg gcugaacgag guggccaaga accugaacga gagccugauc 1560

gaccugcagg agcugggcaa guacgagcag uacaucaagu ggcccuggua caucuggcug 1620

ggcuucaucg ccggccugau cgccaucgug auggugacca ucaugcugug cugcaugacc 1680

agcugcugca gcugccugaa gggcuguugc agcugcggca gcugcugcaa guucgacgag 1740

gacgacagcg agcccgugcu gaagggcgug aagcugcacu acacc 1785

<210> 146
<211> 595
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 146

Met Tyr Ser Met Gln Leu Ala Ser Cys Val Thr Leu Thr Leu Val Leu
1 5 10 15

Leu Val Asn Ser Gln Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn
20 25 30

[0229]

Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile
35 40 45

Leu Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile
50 55 60

Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser
65 70 75 80

Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln
85 90 95

Asp Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys
100 105 110

Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu
115 120 125

Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu
130 135 140

Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly
145 150 155 160

Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys
165 170 175

Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile
180 185 190

Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp
195 200 205

Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met
210 215 220

Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu
225 230 235 240

Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile
245 250 255

[0230] Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp
260 265 270

Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu
275 280 285

Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser
290 295 300

Arg Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr
305 310 315 320

Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
325 330 335

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser
340 345 350

Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly
355 360 365

Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly Val Val Phe

370 375 380

Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala
385 390 395 400

Pro Ala Ile Cys His Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val
405 410 415

Phe Val Ser Asn Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr
420 425 430

Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys
435 440 445

Asp Val Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln
450 455 460

Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
465 470 475 480

His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala
[0231] 485 490 495

Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala
500 505 510

Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
515 520 525

Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala
530 535 540

Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met Thr
545 550 555 560

Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys
565 570 575

Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly Val Lys Leu
580 585 590

His Tyr Thr
595

<210> 147
<211> 1961
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 147

[0232]

gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uacagcaugc agcuggcuag cugcgugacc cugacccugg ugcugcuggu gaacagccag 120

ggcgccgaga acagcguggc cuacagcaac aacagcaucg ccaucccccac caacuucacc 180

aucagcguga ccaccgagau ucugcccgug agcaugacca agaccagcgu ggacugcacc 240

auguacaucu gcggcgacag caccgagugc agcaaccugc ugcugcagua cggcagcuuc 300

ugcacccagc ugaaccgggc ccugaccggc aucgccgugg agcaggacaa gaacacccag 360

gagguguucg cccaggugaa gcagaucuac aagaccccuc ccaucaagga cuucggcggc 420

uucaacuuca gccagauccu gcccgacccc agcaagccca gcaagcggag cuucaucgag 480

gaccugcugu ucaacaaggu gacccuagcc gacgccggcu ucaucaagca guacggcgac 540

ugccucggcg acauagccgc ccgggaccug aucugcgccc agaaguucaa cggccugacc 600

gugcugccuc cccugcugac cgacgagaug aucgcccagu acaccagcgc ccuguuagcc 660

ggaaccauca ccagcggcug gacuuucggc gcuggagccg cucugcagau ccccuucgcc 720

augcagaugg ccuaccgguu caacggcauc ggcgugaccc agaacgugcu guacgagaac 780

cagaagcuga ucgccaacca guucaacagc gccaucggca agauccagga cagccugagc 840

agcaccgcua gcgcccuggg caagcugcag gacgugguga accagaacgc ccaggcccug 900

aacacccugg ugaagcagcu gagcagcaac uucggcgcca ucagcagcgu gcugaacgac 960

auccugagcc ggcuggaccc ucccgaggcc gaggugcaga ucgaccggcu gaucacuggc 1020

cggcugcaga gccugcagac cuacgugacc cagcagcuga uccgggccgc cgagauucgg 1080

gccagcgcca accuggccgc caccaagaug agcgagugcg ugcugggcca gagcaagcgg 1140

guggacuucu gcggcaaggg cuaccaccug augagcuuuc cccagagcgc accccacgga 1200

gugguguucc ugcacgugac cuacgugccc gcccaggaga agaacuucac caccgcccca 1260

gccaucugcc acgacggcaa ggcccacuuu ccccgggagg gcguguucgu gagcaacggc 1320

acccacuggu ucgugaccca gcggaacuuc uacgagcccc agaucaucac caccgacaac 1380

accuucguga gcggcaacug cgacguggug aucggcaucg ugaacaacac cguguacgau 1440

ccccugcagc ccgagcugga cagcuucaag gaggagcugg acaaguacuu caagaaucac 1500

accagccccg acguggaccu gggcgacauc agcggcauca acgccagcgu ggugaacauc 1560

cagaaggaga ucgaucggcu gaacgaggug gccaagaacc ugaacgagag ccugaucgac 1620

cugcaggagc ugggcaagua cgagcaguac aucaaguggc ccugguacau cuggcugggc 1680

uucaucgccg gccugaucgc caucgugaug gugaccauca ugcugugcug caugaccagc 1740

ugcugcagcu gccugaaggg cuguugcagc ugcggcagcu gcugcaaguu cgacgaggac 1800

gacagcgagc ccgugcugaa gggcgugaag cugcacuaca ccugauaaua ggcuggagcc 1860

ucgguggccu agcuucuugc cccuugggcc ucccccccagc cccuccuccc cuuccugcac 1920

ccguaccccc guggucuuug aauaaagucu gagugggcgg c 1961

<210> 148
<211> 13
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

[0233]

<400> 148

Ala Lys Phe Val Ala Ala Trp Thr Leu Lys Ala Ala Ala
1 5 10

<210> 149
<211> 2216
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 149
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccaccaug 60

uucguguucc uggugcugcu gccccuggug agcagccagu gcgugaaccu gaccacccgg 120

acccagcugc caccagccua caccaacagc uucacccggg gcgucuacua ccccgacaag 180

guguuccgga gcagcguccu gcacagcacc caggaccugu uccugcccuu cuucagcaac 240

gugaccuggu uccacgccau ccacgugagc ggcaccaacg gcaccaagcg guucgacaac 300

cccgugcugc ccuucaacga cggcguguac uucgccagca ccgagaagag caacaucauc 360

cggggcugga ucuucggcac cacccuggac agcaagaccc agagccugcu gaucgugaau 420

[0234]

aacgccacca acguggugau caaggugugc gaguuccagu ucugcaacga ccccuuccug 480

ggcguguacu accacaagaa caacaagagc uggauggaga gcgaguuccg gguguacagc 540

agcgccaaca acugcaccuu cgaguacgug agccagcccu uccugaugga ccuggagggc 600

aagcagggca acuucaagaa ccugcgggag uucguguuca agaacaucga cggcuacuuc 660

aagaucuaca gcaagcacac cccaaucaac cuggugcggg aucugcccca gggcuucuca 720

gcccuggagc cccuggugga ccugcccauc ggcaucaaca ucaccgguu ccagacccug 780

cuggcccugc accggagcua ccugacccca ggcgacagca gcagcgggug gacagcaggc 840

gcggcugcuu acuacguggg cuaccugcag ccccggaccu uccugcugaa guacaacgag 900

aacggcacca ucaccgacgc cguggacugc gcccuggacc cucugagcga gaccaagugc 960

acccugaaga gcuucaccgu ggagaagggc aucuaccaga ccagcaacuu ccgggugcag 1020

cccaccgaga gcaucgugcg guuccccaac aucaccaacc ugugccccuu cggcgaggug 1080

uucaacgcca cccgguucgc cagcguguac gccuggaacc ggaagcggau cagcaacugc 1140

guggccgacu acagcgugcu guacaacagc gccagcuuca gcaccuucaa gugcuacggc 1200

gugagcccca ccaagcugaa cgaccugugc uucaccaacg uguacgccga cagcuucgug 1260

auccguggcg acgaggugcg gcagaucgca cccggccaga caggcaagau cgccgacuac 1320

aacuacaagc ugcccgacga cuucaccggc ugcgugaucg ccuggaacag caacaaccuc 1380

gacagcaagg ugggcggcaa cuacaacuac cuguaccggc uguuccggaa gagcaaccug 1440

aagcccuucg agcgggacau cagcaccgag aucuaccaag ccggcuccac cccuugcaac 1500

ggcguggagg gcuucaacug cuacuucccu cugcagagcu acggcuucca gcccaccaac 1560

ggcgugggcu accagcccua ccggguggug gugcugagcu ucgagcugcu gcacgcccca 1620

gccaccgugu guggccccaa gaagagcacc aaccugguga agaacaagug cgugaacuuc 1680

aacuucaacg gccuuaccgg caccggcgug cugaccgaga gcaacaagaa auuccugccc 1740

uuucagcagu ucggccggga caucgccgac accaccgacg cugugcggga uccccagacc 1800

cuggagaucc uggacaucac cccuugcagc uucggcggcg ugagcgugau caccccaggc 1860

accaacacca gcaaccaggu ggccgugcug uaccaggacg ugaacugcac cgaggugccc 1920

guggccaucc acgccgacca gcugacaccc accuggcggg ucuacagcac cggcagcaac 1980

guguuccaga cccgggccgg uugccugauc ggcgccgagc acgugaacaa cagcuacgag 2040

ugcgacaucc ccaucggcgc cggcaucugu gccagcuacc agacccagac caauucauga 2100

uaauaggcug gagccucggu ggccuagcuu cuugccccuu gggccucccc ccagccccuc 2160

cuccccuucc ugcacccgua cccccguggu cuuugaauaa agucugagug ggcggc 2216

<210> 150
<211> 2040
<212> RNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 150
auguucgugu uccuggugcu gcugccccug gugagcagcc agugcgugaa ccugaccacc 60

cggacccagc ugccaccagc cuacaccaac agcuucaccc ggggcgucua cuaccccgac 120

aagguguucc ggagcagcgu ccugcacagc acccaggacc uguuccugcc cuucuucagc 180

aacgugaccu gguuccacgc cauccacgug agcggcacca acggcaccaa gcgguucgac 240

aaccccgugc ugcccuucaa cgacggcgug uacuucgcca gcaccgagaa gagcaacauc 300

auccggggcu ggaucuucgg caccacccug gacagcaaga cccagagccu gcugaucgug 360

aauaacgcca ccaacguggu gaucaaggug ugcgaguucc aguucugcaa cgaccccuuc 420

[0235] cugggcgugu acuaccacaa gaacaacaag agcuggaugg agagcgaguu ccggguguac 480

agcagcgcca acaacugcac cuucgaguac gugagccagc ccuuccugau ggaccuggag 540

ggcaagcagg gcaacuucaa gaaccugcgg gaguucgugu ucaagaacau cgacggcuac 600

uucaagaucu acagcaagca caccccaauc aaccuggugc gggaucugcc ccagggcuuc 660

ucagccccugg agccccuggu ggaccugccc aucggcauca acaucacccg guuccagacc 720

cugcuggccc ugcaccggag cuaccugacc ccaggcgaca gcagcagcgg guggacagca 780

ggcgcggcug cuuacuacgu gggcuaccug cagccccgga ccuuccugcu gaaguacaac 840

gagaacggca ccaucaccga cgccguggac ugcgccccugg acccucugag cgagaccaag 900

ugcacccuga agagcuucac cguggagaag ggcaucuacc agaccagcaa cuuccgggug 960

cagcccaccg agagcaucgu gcgguucccc aacaucacca accugugccc cuucggcgag 1020

guguucaacg ccacccgguu cgccagcgug uacgccugga accggaagcg gaucagcaac 1080

ugcguggccg acuacagcgu gcuguacaac agcgccagcu ucagcaccuu caagugcuac 1140

ggcgugagcc ccaccaagcu gaacgaccug ugcuucacca acguguacgc cgacagcuuc 1200

gugauccgug gcgacgaggu gcggcagauc gcacccggcc agacaggcaa gaucgccgac 1260

uacaacuaca agcugcccga cgacuucacc ggcugcguga ucgccuggaa cagcaacaac 1320

cucgacagca aggugggcgg caacuacaac uaccuguacc ggcuguuccg gaagagcaac 1380

cugaagcccu ucgagcggga caucagcacc gagaucuacc aagccggcuc caccccuugc 1440

aacggcgugg agggcuucaa cugcuacuuc ccucugcaga gcuacggcuu ccagcccacc 1500

aacggcgugg gcuaccagcc cuaccgggug guggugcuga gcuucgagcu gcugcacgcc 1560

ccagccaccg uguguggccc caagaagagc accaaccugg ugaagaacaa gugcgugaac 1620

uucaacuuca acggccuuac cggcaccggc gugcugaccg agagcaacaa gaaauuccug 1680

cccuuucagc aguucggccg ggacaucgcc gacaccaccg acgcugugcg ggaucccccag 1740

acccuggaga uccuggacau caccccuugc agcuucggcg gcgugagcgu gaucacccca 1800

ggcaccaaca ccagcaacca gguggccgug cuguaccagg acgugaacug caccgaggug 1860

cccguggcca uccacgccga ccagcugaca cccaccuggc gggucuacag caccggcagc 1920

aacguguucc agacccgggc cgguugccug aucggcgccg agcacgugaa caacagcuac 1980

gagugcgaca uccccaucgg cgccggcauc ugugccagcu accagaccca gaccaauuca 2040

[0236]

<210> 151
<211> 680
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 151

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val
1 5 10 15

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu

85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

[0237] Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430

[0238]

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

[0239]

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser
675 680

<210> 152
<211> 4
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成

<400> 152

Gly Ser Gly Gly
1

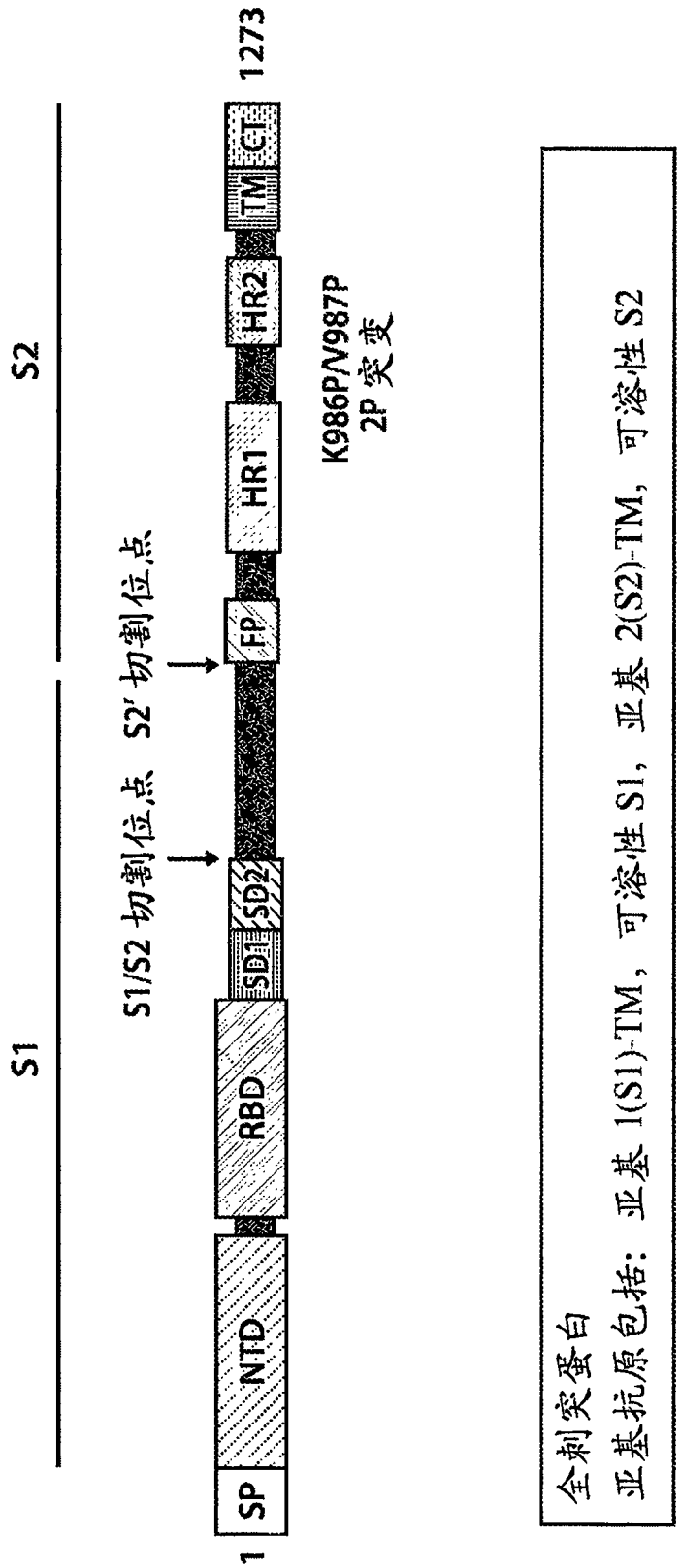

S1
S2
S1/S2 切割位点
S2′ 切割位点
1
SP
NTD
RBD
SD1
SD2
FP
HR1
HR2
TM
CT
1273
K986P/V987P
2P 突变
全刺突蛋白
亚基抗原包括：亚基 1(S1)-TM，可溶性 S1，亚基 2(S2)-TM，可溶性 S2

图1

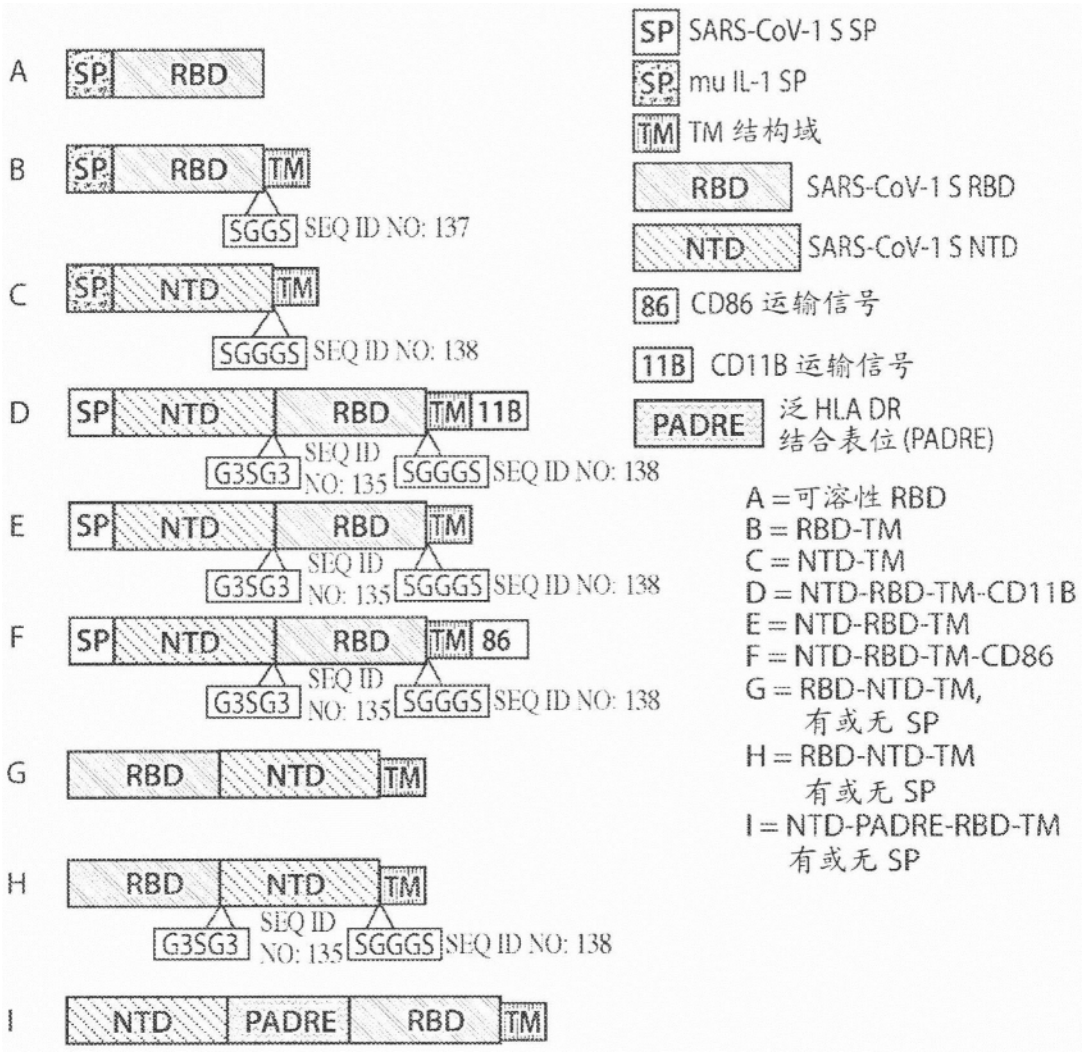

A
SP
RBD
B
SP
RBD
TM
SGGS
SEQ ID NO: 137
C
SP
NTD
TM
SGGGS
SEQ ID NO: 138
D
SP
NTD
RBD
TM
11B
G3SG3
SEQ ID NO: 135
SGGGS
SEQ ID NO: 138
E
SP
NTD
RBD
TM
G3SG3
SEQ ID NO: 135
SGGGS
SEQ ID NO: 138
F
SP
NTD
RBD
TM
86
G3SG3
SEQ ID NO: 135
SGGGS
SEQ ID NO: 138
G
RBD
NTD
TM
H
RBD
NTD
TM
G3SG3
SEQ ID NO: 135
SGGGS
SEQ ID NO: 138
I
NTD
PADRE
RBD
TM
SP SARS-CoV-1 S SP
SP mu IL-1 SP
TM TM 结构域
RBD SARS-CoV-1 S RBD
NTD SARS-CoV-1 S NTD
86 CD86 运输信号
11B CD11B 运输信号
PADRE 泛 HLA DR 结合表位 (PADRE)
A = 可溶性 RBD
B = RBD-TM
C = NTD-TM
D = NTD-RBD-TM-CD11B
E = NTD-RBD-TM
F = NTD-RBD-TM-CD86
G = RBD-NTD-TM, 有或无 SP
H = RBD-NTD-TM 有或无 SP
I = NTD-PADRE-RBD-TM 有或无 SP

图2

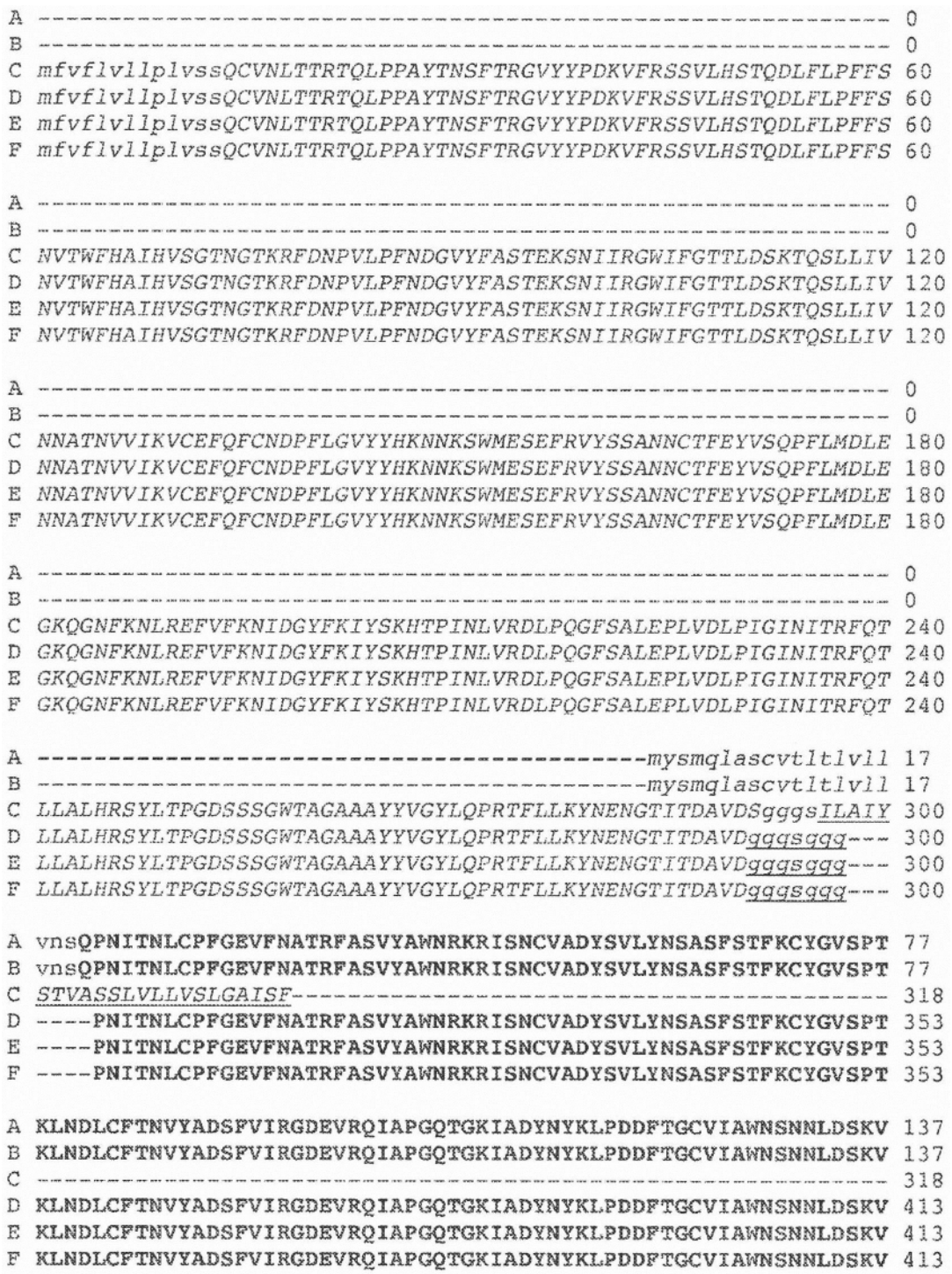

A ---------- 0
B ---------- 0
C mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
D mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
E mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
F mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
A ---------- 0
B ---------- 0
C NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
D NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
E NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
F NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
A ---------- 0
B ---------- 0
C NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
D NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
E NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
F NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
A ---------- 0
B ---------- 0
C GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
D GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
E GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
F GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
A ----------mysmqlascvtltlvll 17
B ----------mysmqlascvtltlvll 17
C LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDSgggsILAIY 300
D LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDgggsggg--- 300
E LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDgggsggg--- 300
F LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDgggsggg---- 300
A vnsQPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPT 77
B vnsQPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPT 77
C STVASSLVLLVSLGAISF---------- 318
D ----PNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPT 353
E ----PNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPT 353
F ----PNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPT 353
A KLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV 137
B KLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV 137
C ---------- 318
D KLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV 413
E KLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV 413
F KLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV 413

图3

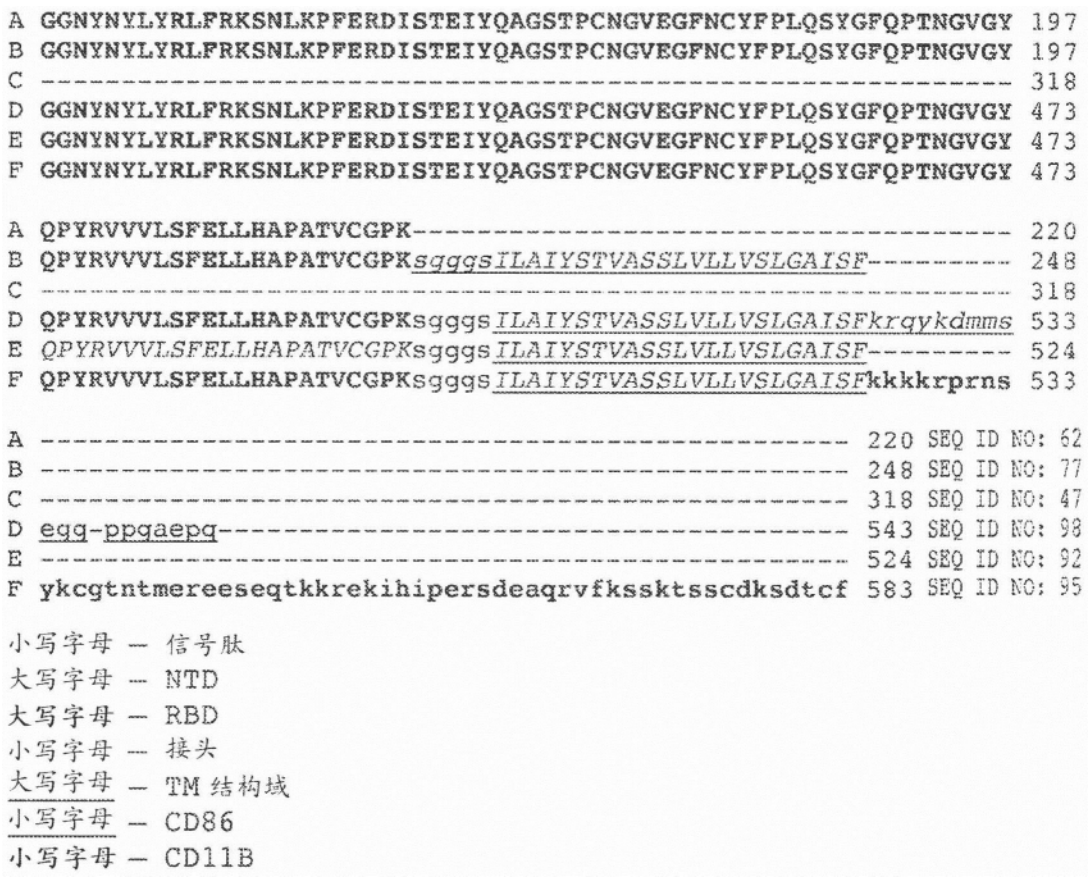

A GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY 197
B GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY 197
C ---------- 318
D GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY 473
E GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY 473
F GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY 473
A QPYRVVVLSFELLHAPATVCGPK---------- 220
B QPYRVVVLSFELLHAPATVCGPKsgggsILAIYSTVASSLVLLVSLGAISF---------- 248
C ---------- 318
D QPYRVVVLSFELLHAPATVCGPKsgggsILAIYSTVASSLVLLVSLGAISFkrqykdmms 533
E QPYRVVVLSFELLHAPATVCGPKsgggsILAIYSTVASSLVLLVSLGAISF---------- 524
F QPYRVVVLSFELLHAPATVCGPKsgggsILAIYSTVASSLVLLVSLGAISFkkkkrprns 533
A ---------- 220 SEQ ID NO: 62
B ---------- 248 SEQ ID NO: 77
C ---------- 318 SEQ ID NO: 47
D egg-ppgaepg---------- 543 SEQ ID NO: 98
E ---------- 524 SEQ ID NO: 92
F ykcgtntmereeseqtkkrekihipersdeaqrvfkssktsscdksdtcf 583 SEQ ID NO: 95
小写字母 — 信号肽
大写字母 — NTD
大写字母 — RBD
小写字母 — 接头
大写字母 — TM 结构域
小写字母 — CD86
小写字母 — CD11B

图3(续)

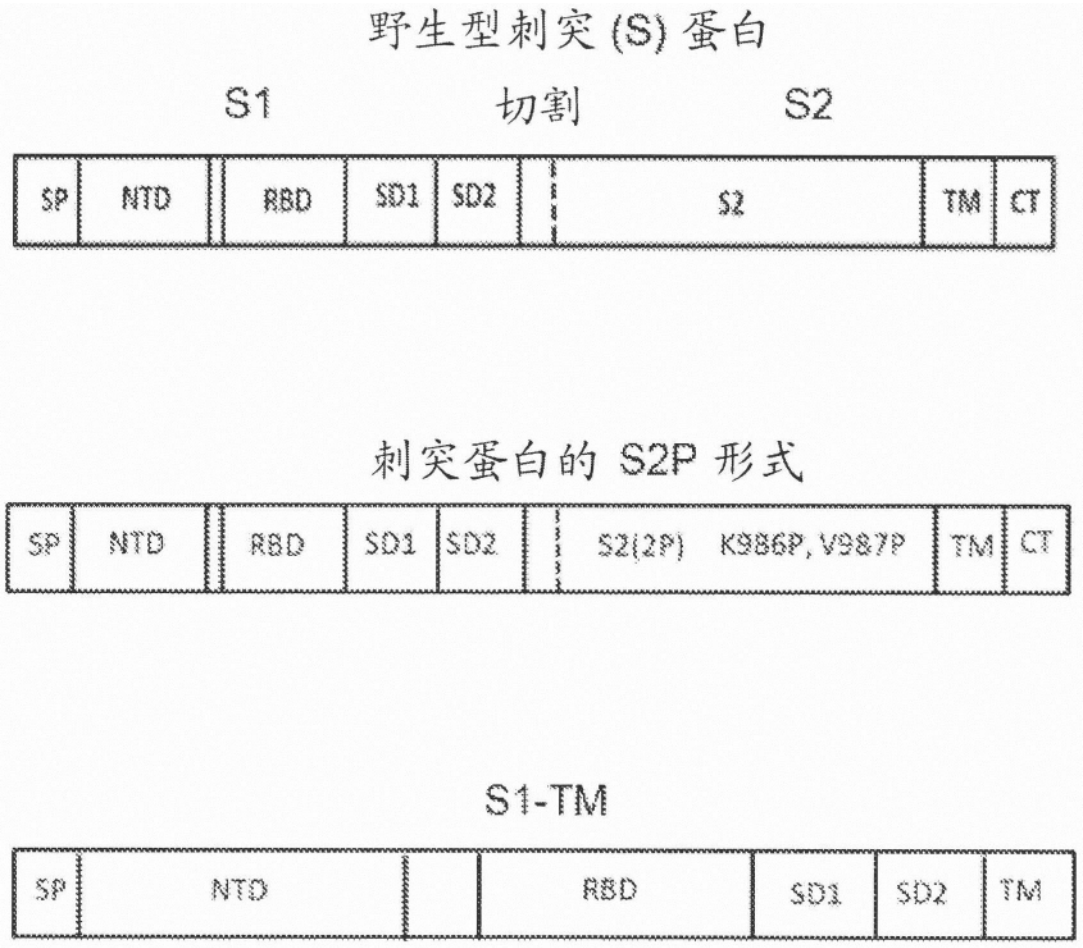

野生型刺突(S)蛋白
S1
切割
S2
SP
NTD
RBD
SD1
SD2
S2
TM
CT
刺突蛋白的 S2P 形式
SP
NTD
RBD
SD1
SD2
S2(2P)
K986P,V987P
TM
CT
S1-TM
SP
NTD
RBD
SD1
SD2
TM

图4

```
G mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
H mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVCRSSVLHSTQDLFLPFFS 60
I mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
J mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
K mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
L mfvflvllplvssQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS 60
  ************************************** *****************

G NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
H NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
I NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
J NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
K NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
L NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV 120
  ************************************************************

G NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
H NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
I NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
J NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
K NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
L NNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLE 180
  ************************************************************

G GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
H GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
I GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
J GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
K GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
L GKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQT 240
  ************************************************************

G LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
H LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
I LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
J LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
K LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
L LLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK 300
  ************************************************************

G CTLKSFTVgsgggsgggsgggsgggsggQPTESIVRFPNITNLCPFGEVFNATRFASVYA 360
H CTLKSFTVgsgggsgggsgggsgggsggQPTESIVRFPNITNLCPFGEVFNATRFASVYA 360
I CTLKSFTVEKGIYQ--------TSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA 352
J CTLKSFTVEKGIYQ--------TSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA 352
K CTLKSFTVEKGIYQ---------TSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA 352
L CTLKSFTVEKGIYQ---------TSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA 352
  ******** .*  .         ..   ********************************
```

图5

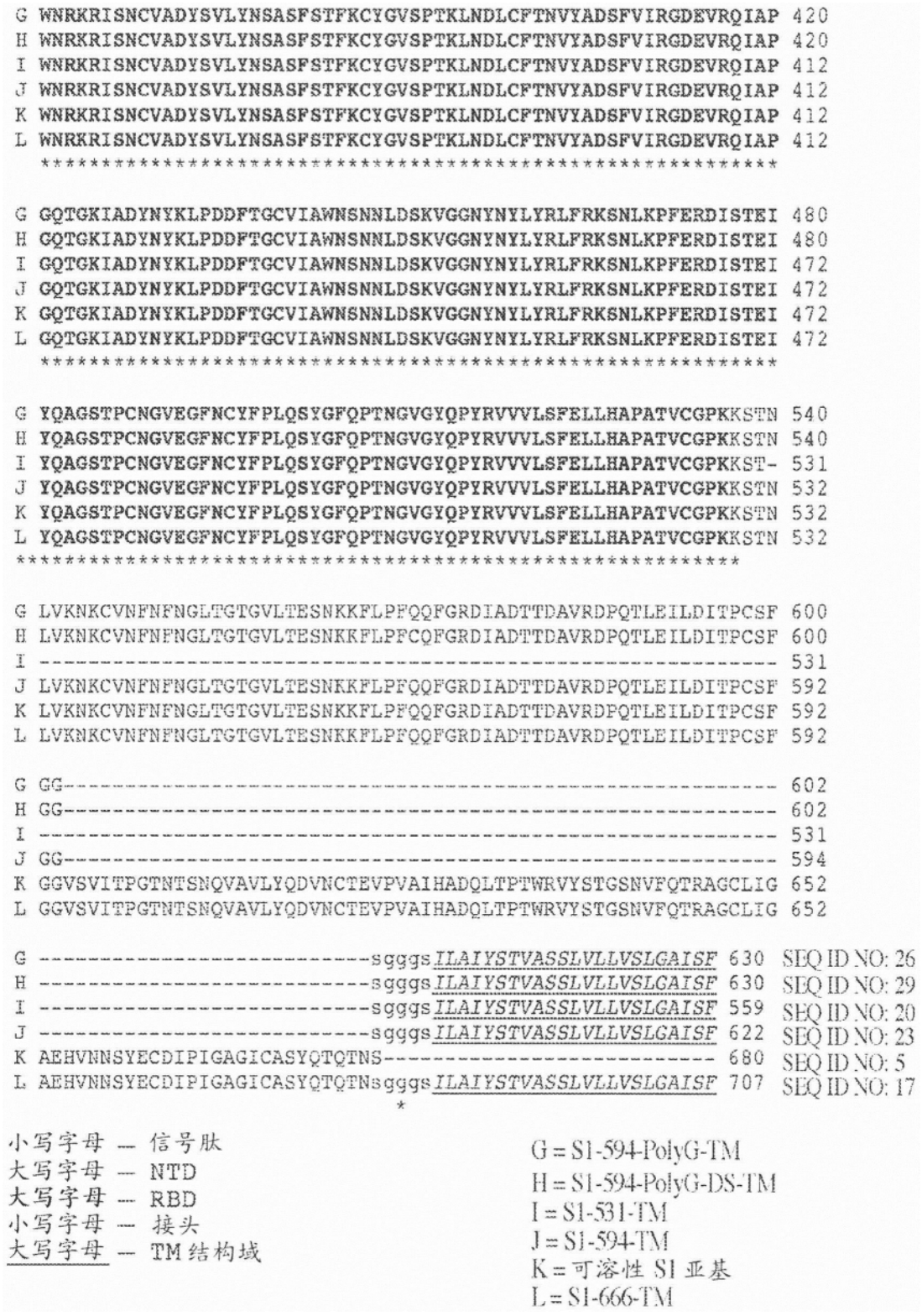

G WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP 420
H WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP 420
I WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP 412
J WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP 412
K WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP 412
L WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP 412
G GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI 480
H GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI 480
I GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI 472
J GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI 472
K GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI 472
L GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEI 472
G YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTN 540
H YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTN 540
I YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKST- 531
J YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTN 532
K YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTN 532
L YQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTN 532
G LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF 600
H LVKNKCVNFNFNGLTGTGVLTESNKKFLPFCQFGRDIADTTDAVRDPQTLEILDITPCSF 600
I ---------- 531
J LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF 592
K LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF 592
L LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSF 592
G GG---------- 602
H GG---------- 602
I ---------- 531
J GG---------- 594
K GGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG 652
L GGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIG 652
G ----------sgggsILAIYSTVASSLVLLVSLGAISF 630 SEQ ID NO: 26
H ----------sgggsILAIYSTVASSLVLLVSLGAISF 630 SEQ ID NO: 29
I ----------sgggsILAIYSTVASSLVLLVSLGAISF 559 SEQ ID NO: 20
J ----------sgggsILAIYSTVASSLVLLVSLGAISF 622 SEQ ID NO: 23
K AEHVNNSYECDIPIGAGICASYQTQTNS---------- 680 SEQ ID NO: 5
L AEHVNNSYECDIPIGAGICASYQTQTNsgggsILAIYSTVASSLVLLVSLGAISF 707 SEQ ID NO: 17
小写字母 — 信号肽
大写字母 — NTD
大写字母 — RBD
小写字母 — 接头
大写字母 — TM结构域
G = S1-594-PolyG-TM
H = S1-594-PolyG-DS-TM
I = S1-531-TM
J = S1-594-TM
K = 可溶性 S1 亚基
L = S1-666-TM

图5(续)

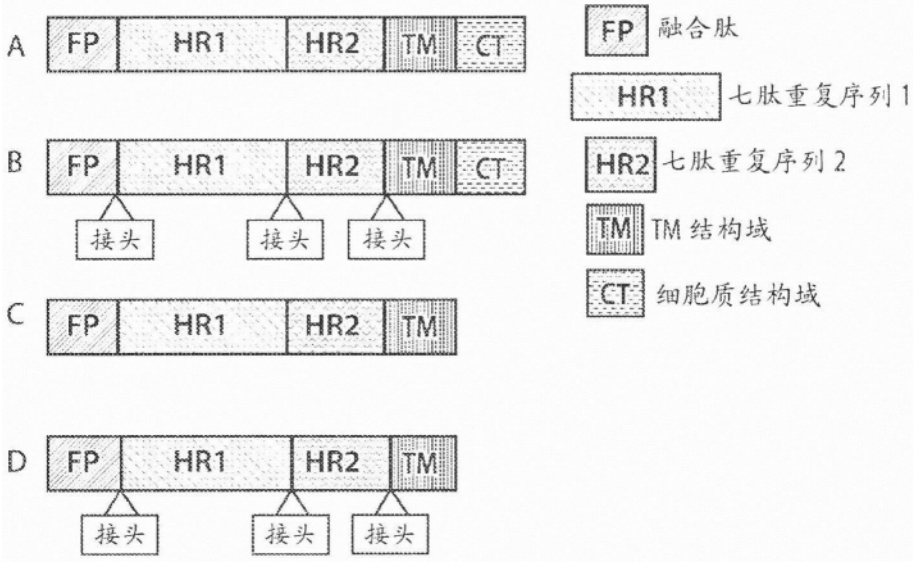

A
FP
HR1
HR2
TM
CT
B
接头
C
D
FP 融合肽
HR1 七肽重复序列 1
HR2 七肽重复序列 2
TM TM 结构域
CT 细胞质结构域

图6

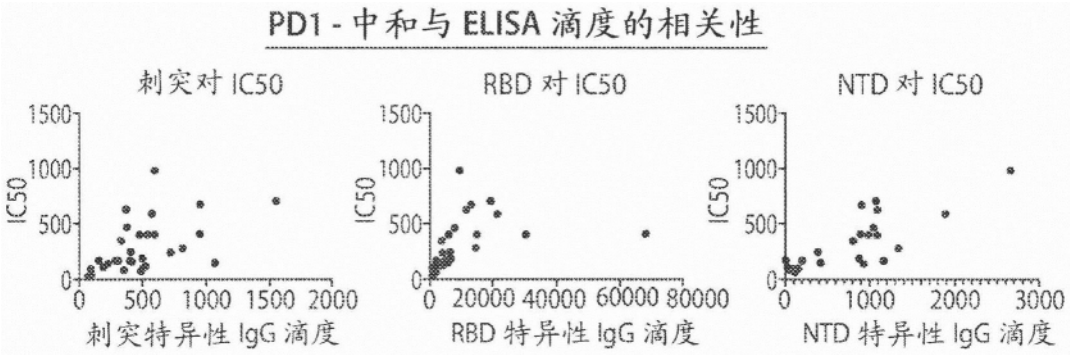

PD1 - 中和与 ELISA 滴度的相关性
刺突对 IC50
RBD 对 IC50
NTD 对 IC50
IC50
1500
1000
500
0
0 500 1000 1500 2000
0 20000 40000 60000 80000
0 1000 2000 3000
刺突特异性 IgG 滴度
RBD 特异性 IgG 滴度
NTD 特异性 IgG 滴度

图7

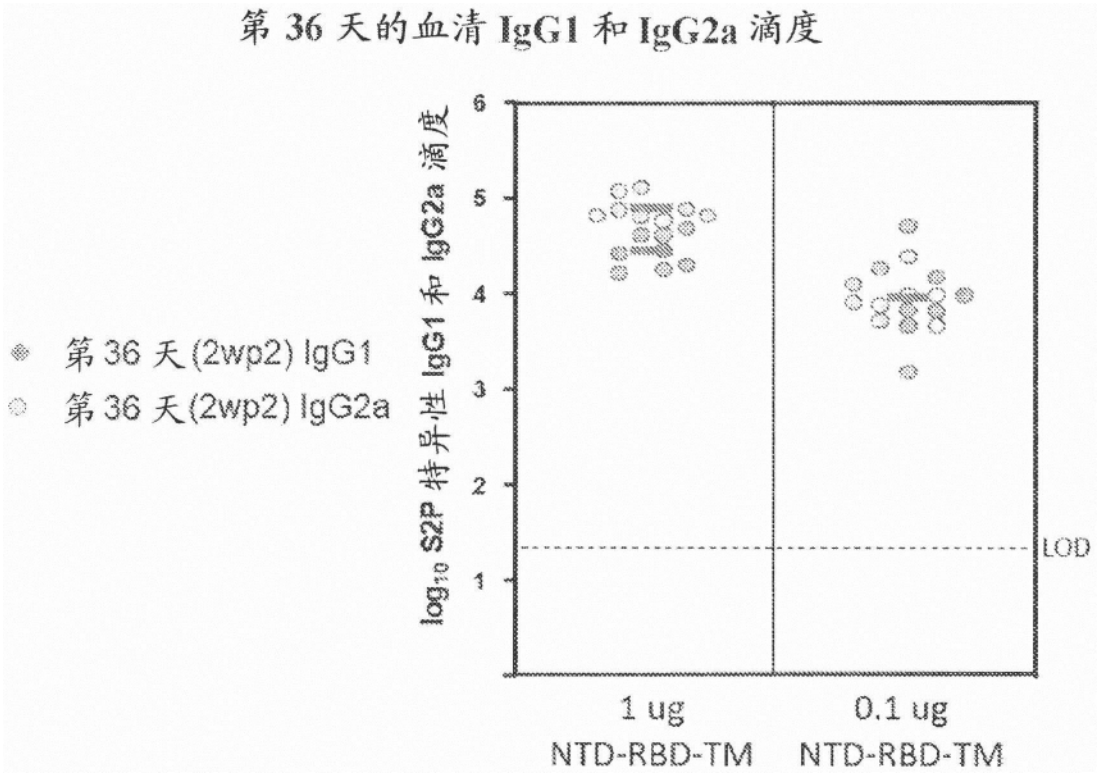

第 36 天的血清 IgG1 和 IgG2a 滴度
第 36 天(2wp2) IgG1
第 36 天(2wp2) IgG2a
log10 S2P 特异性 IgG1 和 IgG2a 滴度
6
5
4
3
2
1
LOD
1 ug
NTD-RBD-TM
0.1 ug
NTD-RBD-TM

图8A

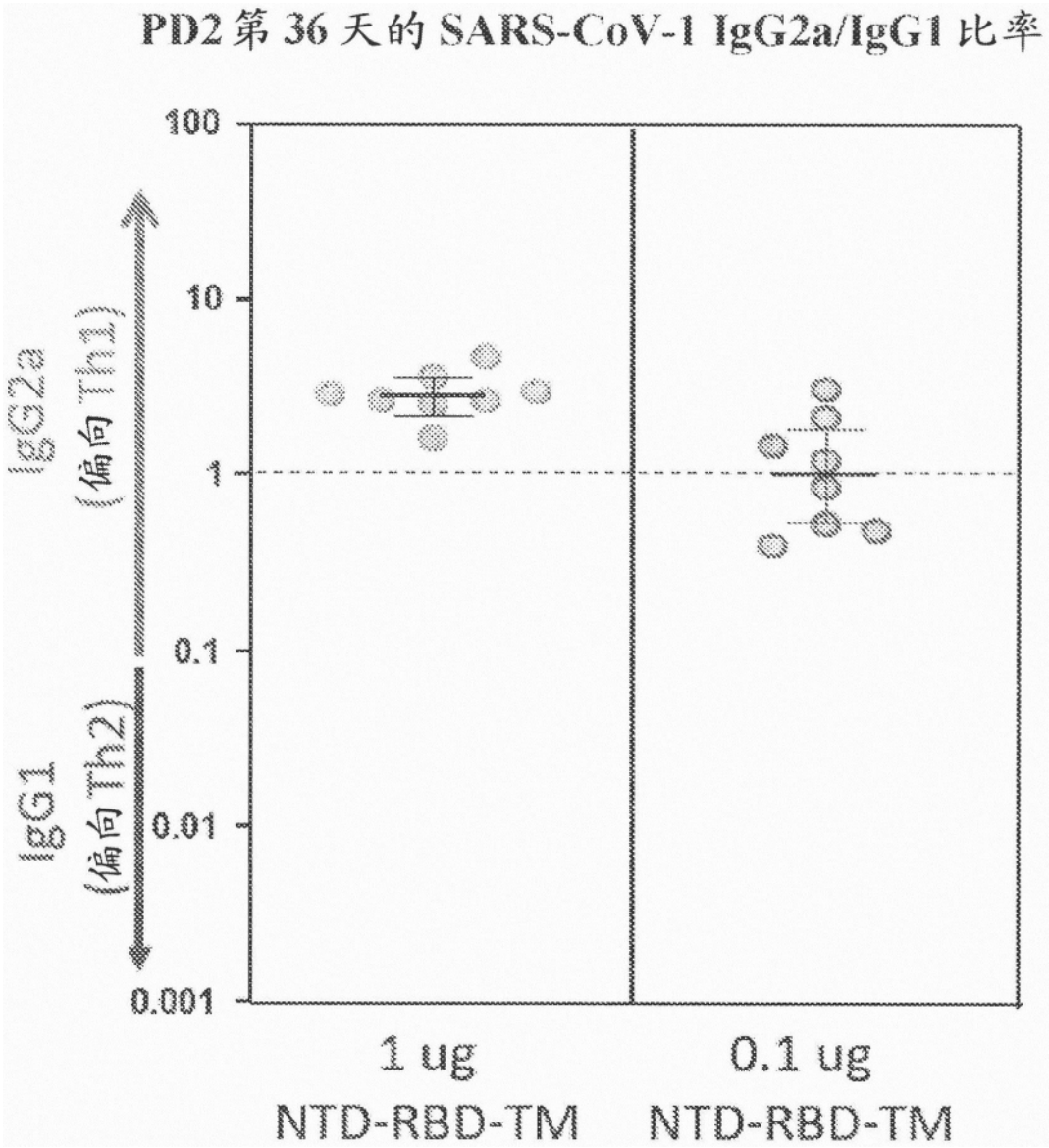

PD2第36天的SARS-CoV-1 IgG2a/IgG1比率
100
10
1
0.1
0.01
0.001
IgG2a
(偏向Th1)
IgG1
(偏向Th2)
1 ug
NTD-RBD-TM
0.1 ug
NTD-RBD-TM

图8B

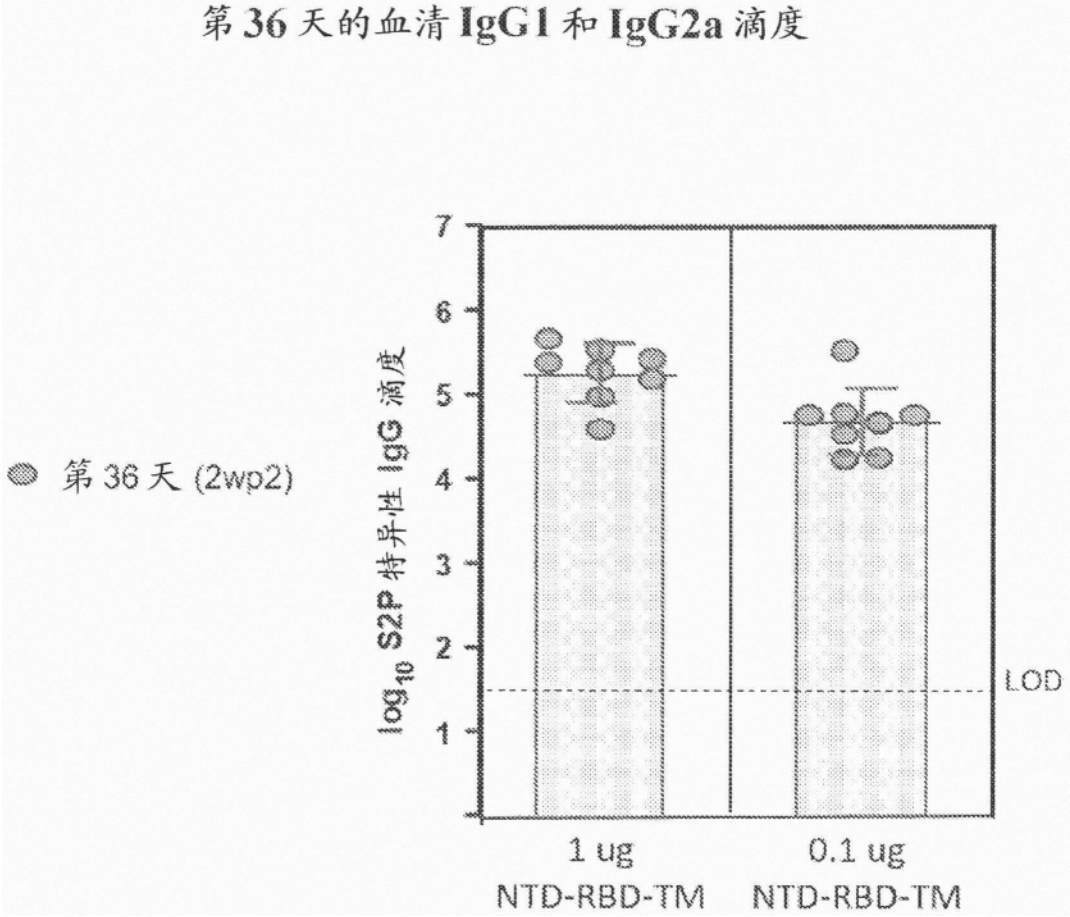

第36天的血清IgG1和IgG2a滴度
第36天 (2wp2)
log10 S2P 特异性 IgG 滴度
7
6
5
4
3
2
1
LOD
1 ug
NTD-RBD-TM
0.1 ug
NTD-RBD-TM

图8C