



(12) 发明专利申请

(10) 申请公布号 CN 115515627 A

(43) 申请公布日 2022. 12. 23

(21) 申请号 202180027866.X

(22) 申请日 2021.02.11

(30) 优先权数据

62/972,886 2020.02.11 US

(85) PCT国际申请进入国家阶段日

2022.10.11

(86) PCT国际申请的申请数据

PCT/US2021/017709 2021.02.11

(87) PCT国际申请的公布数据

WO2021/163365 EN 2021.08.19

(71) 申请人 美国政府(由卫生和人类服务部的
部长所代表)

地址 美国马里兰州

申请人 德克萨斯州大学系统董事会
达特茅斯大学董事会

(72) 发明人 B·格雷汉姆 K·科比特

O·阿比奥纳 G·哈钦森

J·麦克莱伦 D·沃尔普 N·王

(74) 专利代理机构 北京华睿卓成知识产权代理
事务所(普通合伙) 11436

专利代理师 程淼 刘海

(51) Int.Cl.

A61K 39/12 (2006.01)

A61P 31/14 (2006.01)

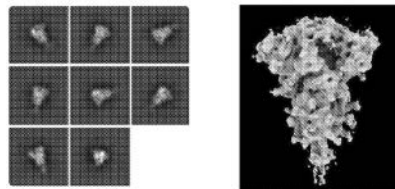
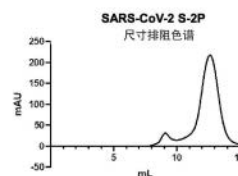
权利要求书2页 说明书35页
序列表36页 附图9页

(54) 发明名称

SARS-CoV-2疫苗

(57) 摘要

本发明公开了稳定在融合前构象中的SARS-CoV-2S胞外域三聚体、编码这些蛋白质的核酸分子和载体,以及它们的用途和生产方法。在若干实施方案中,所述SARS-CoV-2S胞外域三聚体和/或核酸分子可用于使受试者产生对SARS-CoV-2S的免疫应答,例如抑制所述受试者体内SARS-CoV-2感染的免疫应答。



1. 一种免疫原, 包含:

重组SARS-CoV-2S胞外域三聚体, 其包含的原聚体包含与SEQ ID NO:2的残基16-1208至少95%相同的氨基酸序列并且在SEQ ID NO:2的位置986和987处包含脯氨酸取代, 所述脯氨酸取代使所述S胞外域三聚体稳定在融合前构象中。

2. 根据权利要求1所述的免疫原, 其中所述重组SARS-CoV-2S胞外域三聚体中的原聚体包含与SEQ ID NO:2的残基16-1208至少98%相同的氨基酸序列并且包含两个氨基酸取代。

3. 根据权利要求2所述的免疫原, 其中所述重组SARS-CoV-2S胞外域三聚体中的原聚体包含与SEQ ID NO:2的残基16-1208至少99%相同的氨基酸序列并且包含两个氨基酸取代。

4. 根据权利要求1所述的免疫原, 其中所述重组SARS-CoV-2S胞外域三聚体中的原聚体包含如SEQ ID NO:2的残基16-1208所示的氨基酸序列。

5. 根据前述权利要求中任一项所述的免疫原, 其中所述一个或两个氨基酸取代是相对于如SEQ ID NO:1所示天然SARS-CoV-2S序列的K986P和V987P取代。

6. 根据前述权利要求中任一项所述的免疫原, 其中所述重组SARS-CoV-2S胞外域三聚体的原聚体还包含一个或更多个额外的氨基酸取代, 其使所述重组SARS-CoV-2S胞外域三聚体稳定在所述融合前构象中。

7. 根据前述权利要求中任一项所述的免疫原, 其中所述重组SARS-CoV-2S胞外域三聚体中的原聚体还包含N501Y、K417N和E484K取代中的一个或更多个。

8. 根据前述权利要求中任一项所述的免疫原, 其中所述胞外域中所述原聚体的C-末端残基通过肽接头连接至三聚体化域, 或直接连接至所述三聚体化域。

9. 根据权利要求8所述的免疫原, 其中所述三聚体化域是T4次要纤维蛋白三聚体化域。

10. 根据权利要求8或9所述的免疫原, 其中与所述T4次要纤维蛋白三聚体化域连接的所述原聚体包含与SEQ ID NO:2的残基16-1235至少95%相同的氨基酸序列, 并且包含使所述S胞外域三聚体稳定在所述融合前构象中的氨基酸取代。

11. 根据权利要求9所述的免疫原, 其中与所述T4次要纤维蛋白三聚体化域连接的所述原聚体包含SEQ ID NO:2的残基16-1235。

12. 根据前述权利要求中任一项所述的免疫原, 其中所述S胞外域的S1/S2蛋白酶切割位点被突变以抑制蛋白酶切割。

13. 根据前述权利要求中任一项所述的免疫原, 其中所述重组SARS-CoV-2S胞外域三聚体是可溶的。

14. 根据权利要求1-7中任一项所述的免疫原, 其中所述胞外域中所述原聚体的C-末端残基通过肽接头连接至跨膜域, 或直接连接至所述跨膜域。

15. 根据权利要求14所述的免疫原, 其中与所述跨膜域连接的所述原聚体包含与SEQ ID NO:3的残基16-1273至少95%相同的氨基酸序列, 并且包含使所述S胞外域三聚体稳定在所述融合前构象中的氨基酸取代。

16. 根据权利要求14所述的免疫原, 其中与所述跨膜域连接的所述原聚体包含如SEQ ID NO:3的残基16-1273所示的氨基酸序列。

17. 根据权利要求1-7中任一项所述的免疫原, 其中所述原聚体的C-末端残基通过肽接头连接至蛋白质纳米颗粒亚基, 或直接连接至所述蛋白质纳米颗粒亚基。

18. 一种蛋白质纳米颗粒, 包含前述权利要求中任一项所述的免疫原。

19. 一种类病毒颗粒, 包含权利要求1-15中任一项所述的免疫原。
20. 一种分离的核酸分子, 编码前述权利要求中任一项所述的重组SARS-CoV-S胞外域三聚体的原聚体。
 21. 根据权利要求20所述的核酸分子, 与启动子可操作地连接。
 22. 一种载体, 包含权利要求20或权利要求21所述的核酸分子。
 23. 根据权利要求22所述的载体, 其中所述载体是病毒载体。
 24. 一种免疫原性组合物, 包含前述权利要求中任一项所述的免疫原、蛋白质纳米颗粒、类病毒颗粒、核酸分子或载体, 以及药学上可接受的递体。
 25. 一种生产稳定在融合前构象中的重组SARS-CoV-2S胞外域三聚体的方法, 包含:
在宿主细胞中表达权利要求20-23中任一项所述的核酸分子或载体以产生所述重组SARS-CoV-2S胞外域三聚体; 和
纯化所述重组SARS-CoV-2S胞外域三聚体。
 26. 根据权利要求25所述的方法产生的所述重组SARS-CoV-2S胞外域三聚体。
 27. 一种使受试者产生对SARS-CoV-2S胞外域的免疫应答的方法, 包含向所述受试者施用有效量的权利要求1-24中任一项所述免疫原、蛋白质纳米颗粒、类病毒颗粒、核酸分子、载体或免疫原性组合物, 以产生所述免疫应答。
 28. 根据权利要求27所述的方法, 其中所述免疫应答抑制SARS-CoV-2感染。
 29. 根据权利要求27或28所述的方法, 其中产生所述免疫应答抑制所述SARS-CoV-2在所述受试者中复制。
 30. 根据权利要求1-24中任一项所述的免疫原、蛋白质纳米颗粒、类病毒颗粒、核酸分子、载体或免疫原性组合物用于在受试者中诱导对SARS-CoV-2S胞外域的免疫应答的用途。

SARS-CoV-2疫苗

[0001] 相关申请的交叉引用

[0002] 本申请要求于2020年2月11日提交的美国临时申请第62/972,886号的权益,该申请以全文引用的方式并入本文。

技术领域

[0003] 本公开内容涉及稳定在融合前构象中的重组SARS-CoV-2刺突(S)蛋白及其作为免疫原的用途。

背景技术

[0004] 冠状病毒是有包膜的正义单链RNA病毒。它们拥有已知RNA病毒中最大的基因组(26-32kb),并在系统发育上分为四个属(α 、 β 、 γ 、 δ),而 β 冠状病毒进一步细分为四个谱系(A、B、C、D)。冠状病毒感染包括人类在内的多种鸟类和哺乳动物物种。

[0005] 2019年,一种新型冠状病毒(世界卫生组织定名为SARS-CoV-2)被确定为冠状病毒疫情的病原体。高病死率、模糊的流行病学定义以及缺乏针对冠状病毒的预防或治疗措施,导致迫切需要有效的疫苗和相关治疗剂。

发明内容

[0006] 本文公开了包含原聚体的重组SARS-CoV-2 S胞外域三聚体,所述原聚体包含一个或多个氨基酸取代,所述氨基酸取代使所述S蛋白三聚体在所述融合前构象中稳定。

[0007] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含原聚体,所述原聚体包含与SEQ ID NO:2的残基16-1208至少95%(例如至少96%、至少97%、至少98%或至少99%)相同的氨基酸序列和SEQ ID NO:2位置986和987处的脯氨酸取代,其中所述脯氨酸使所述S胞外域三聚体稳定在融合前构象中。与天然SARS-CoV-2 S胞外域序列相比,位置986和987处的脯氨酸是氨基酸取代,例如K986P和V987P取代。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含含有SEQ ID NO:2残基16-1208的原聚体。

[0008] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体的原聚体还包含一个或多个额外的氨基酸取代或缺失,例如使所述重组SARS-CoV-2 S胞外域三聚体稳定在融合前构象中的氨基酸取代,或抑制或防止蛋白酶在所述S胞外域S1/S2蛋白酶切割位点切割的氨基酸取代。

[0009] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至三聚体化域(例如T4次要纤维蛋白(Fibrin)三聚体化域)。在另外的实施方案中,所述重组SARS-CoV-2 S胞外域三聚体的原聚体可以是膜锚定的,例如,通过与跨膜域连接锚定。

[0010] 在另外的实施方案中,所述重组SARS-CoV-2 S胞外域三聚体可以包含在自组装蛋白纳米颗粒上,例如铁蛋白纳米颗粒或基于合成蛋白的纳米颗粒。还提供了编码所公开重组SARS-CoV-2 S胞外域三聚体的原聚体的核酸分子,以及包括所述核酸分子的载体,以及产生所公开重组SARS-CoV-2 S胞外域三聚体的方法。

[0011] 还提供了包括所述重组SARS-CoV-2 S胞外域三聚体的免疫原性组合物,所述组合物适合施用于受试者,并且也可以包含在单位剂型中。所述组合物还可以包括佐剂。所述重组SARS-CoV-2 S胞外域三聚体也可以与递体缀合,以利于向免疫系统的呈递。公开了通过向受试者施用有效量的所公开重组SARS-CoV-2 S胞外域三聚体、核酸分子或载体,诱导所述受试者免疫应答的方法,以及抑制或防止所述受试者SARS-CoV-2感染的方法。

[0012] 通过参照附图对下述若干实施方案进行详细描述,本公开的前述和其他特征及优点变得更加明显。

附图说明

[0013] 图1A-1D.通过K986P和V987P氨基酸取代使SARS-CoV-2 S蛋白稳定在融合前构象中。(图1A) SARS-CoV-2 S一级结构示意图。SS=信号序列,NTD=N-末端域,RBD=受体结合域,S1/S2=S1/S2蛋白酶切割位点,FP=融合肽,HR1=七肽重复序列1,CH=中心螺旋,CD=连接子域,HR2=七肽重复序列2,TM=跨膜域,CT=胞质尾。箭头表示蛋白酶切割位点。(图1B) SARS-CoV-2 S-2P蛋白(SEQ ID NO:2)的尺寸排阻色谱产生单个大峰,表明DNA质粒的高水平蛋白质表达和均匀的蛋白质群。根据蛋白质大小,按预期洗脱峰值级分。SARS-CoV-2刺突蛋白的2D类别平均值(图1C)和4.7埃结构(图1D)仅显示融合前构象。

[0014] 图2A-2D.用SARS-CoV-2WT或S-2P免疫后多个小鼠品系的抗体反应。BALB/cJ(图2A、图2D)、C57BL/6J(图2B)或B6C3F1/J(图2C)小鼠在第0周和第3周用PBS,0.01 μ g、0.1 μ g或1 μ g佐以Sigma佐剂系统(SAS)的SARS-CoV-2 S WT或SARS-CoV-2 S-2P免疫,并在初免后2周(空心圆圈)和加强后2周(实心圆圈)收集血清。通过ELISA评估SARS-CoV-2 S-2P免疫小鼠血清的SARS-CoV-2 S特异性IgG(图2A-2C)。评估S WT和S-2P免疫的BALB/cJ小鼠的加强后血清中针对同型SARS-CoV-2假病毒的中和抗体(图2D)。双因素方差分析(Two-way ANOVA)和多重比较测试用于比较每个剂量水平内和加强后剂量之间的初免后和加强后结合抗体反应(图2A-2C),并比较S WT与S-2P在每个剂量下引发的中和抗体以及剂量对中和活性的影响(图2D)。虚线表示测定检出限。灰色虚线= p 值 <0.05 ,灰色线= p 值 <0.01 ,黑色虚线= p 值 <0.001 ,黑线= p 值 <0.0001 。

[0015] 图3A-3B.SARS-CoV-2 S WT和SARS-CoV-2 S-2P保护小鼠免受病毒复制的能力。BALB/cJ小鼠在第0周和第3周用PBS,0.01 μ g、0.1 μ g或1 μ g佐以SAS的SARS-CoV-2 S WT或SARS-CoV-2 S-2P进行免疫。加强后4周,小鼠接受了小鼠适应的SARS-CoV-2攻击。攻击后两天,在病毒载量峰值时,收集肺(图3A)和鼻甲(图3B)用于通过噬斑测定评估病毒载量。通过进行单因素方差分析(one-way ANOVA)和多重比较测试对各组进行比较。虚线表示测定检出限。灰色虚线= p 值 <0.05 ,灰色线= p 值 <0.01 。注意:由于与实验无关的死亡,0.01 μ g S-2P免疫小鼠未受到攻击(N/T)。

[0016] 图4A-4E.具有N-连接的聚糖和生物缀合标签(SpyTag)的二氧四氢嘧啶合酶(LuS)纳米颗粒支架和铁蛋白纳米颗粒支架,可在哺乳动物细胞中很好地表达成组装的纳米颗粒。(图4A)示意图显示了用于通过异肽键进行生物缀合的单独CnaB2域标签("SpyTag")和剩余的CnaB2域("SpyCatcher"),所述异肽键作为通过SpyTag和SpyCatcher生物缀合对共价连接分子的方式。(图4B)设计表达构建体以在哺乳动物细胞中产生具有SpyTag的激活的纳米颗粒,所述SpyTag用于在纳米颗粒表面上缀合抗原。上图显示了DNA构建体。在可切割

的信号肽之后,将SpyTag设置在纳米颗粒序列的N-端。His标签和Strep标签位于LuS纳米颗粒的C端。在纳米颗粒序列中设计了N连接的糖基化位点以促进蛋白质表达。下图显示了LuS-N71-SpyTag单体和铁蛋白-N96-SpyTag单体和组装纳米颗粒的预期结构。聚糖和SpyTag都位于所述纳米颗粒表面。(图4C)尺寸排阻色谱图证实了所述纳米颗粒的正确尺寸。将样品装载到Superdex 200Increase10/300GL色谱柱上,PBS为流动相。(图4D)在存在或不存在PNGase F的情况下LuS-N71-SpyTag和铁蛋白-N96-SpyTag的SDS-PAGE。标记了PNGase F的位置。铁蛋白产生多个条带的原因可能是蛋白水解切割和不完全糖基化。(图4E) LuS-N71-SpyTag和铁蛋白-N96-SpyTag的负染色EM图像(左图)和2D类平均值(右图)显示正确组装了预期尺寸的纯化纳米颗粒。

[0017] 图5A-5E. SARS-CoV-2 S三聚体与LuS-SpyTag的缀合显示在LuS-N71-SpyLinked-CoV-2刺突纳米颗粒表面上的SARS-CoV-2刺突三聚体。(图5A)示意图显示了SpyTag偶联的LuS与SpyCatcher偶联的SARS-CoV-2刺突三聚体缀合以制备LuS-N71-SpyLinked-CoV-2刺突纳米颗粒。(图5B) LuS-N71-SpyTag、SARS-CoV-2刺突-SpyCatcher和LuS-N71-SpyLinked-CoV-2刺突的缀合产物在Superdex 200Increase 10/300GL色谱柱的SEC曲线图,流动相为PBS。(图5C)在存在DTT的情况下, LuS-N71-SpyTag(第1道)、SARS-CoV-2刺突-SpyCatcher(第2道)和LuS-N71-SpyTag与SARS-CoV-2刺突-SpyCatcher缀合混合物(第3道)的SDS-PAGE。缀合混合物(第3道)显示LuS-N71-SpyTag稍过量的缀合LuS-N71-SpyLinked-CoV-2刺突纳米颗粒。(图5D) LuS-N71-SpyLinked-CoV-2刺突纳米颗粒在SEC纯化后的负染色EM,显示有代表性的显微照片(左图)和2D类平均值(右图)。(图5E) LuS-N71-SpyLinked-CoV-2刺突纳米颗粒与RBD靶向抗体CR3022 IgG结合的表面等离子体共振(SPR)反应曲线,其中IgG与溶液中的芯片和纳米颗粒偶联。分析了一系列纳米颗粒浓度,其中与所述纳米颗粒偶联的SARS CoV-2刺突浓度范围为200nM至1.56nM。提供了所观测 k_a 值。

[0018] 图6A-6C. LuS-N71-SpyLinked-CoV-2刺突的免疫原性。(图6A) SARS-CoV-2刺突免疫原的免疫程序示意图。(图6B) 抗-SARS-CoV-2刺突ELISA滴度的血清评估。免疫组采用颜色编码。垂直虚线将初免后免疫原剂量组和周数分开。起始血清稀释度倒数(100)用水平虚线表示。每只动物的ELISA滴度显示为单个点。三角形点设定为最大测定的ELISA滴度。几何平均值表示为黑色水平线。请注意,用0.08mg LuS-N71-SpyTag免疫的三只动物在第5周显示出高ELISA滴度,与该对照组的三只动物相同,所述对照组显示出可检测的中和作用。(图6C) 每只动物在第5周中和滴度显示为单个点,并且几何平均值由黑色水平线表示,为每组提供数值。免疫组用颜色编码,如图6B所示。测定检出限(滴度=40)用水平虚线表示。P值由双尾曼-惠特尼(Mann-Whitney)检验确定。*表示 $P \leq 0.05$,**表示 $P \leq 0.01$,***表示 $P \leq 0.001$,****表示 $P \leq 0.0001$ 。

[0019] 序列表

[0020] 随附序列表中列出的核酸序列和氨基酸序列使用如37C.F.R.1.822定义的核苷酸碱基的标准字母缩写和氨基酸三字母代码表示。仅显示了每个核酸序列的一条链,但是互补链被理解为包含在对所示链的任何引用中。序列表以ASCII文本文件的形式提交,文件名为“Sequence.txt”(约~88kb),创建于2021年2月11日,通过引用并入本文。

具体实施方式

[0021] 本公开提供了SARS-CoV-2刺突糖蛋白(S)胞外域三聚体,其稳定在融合前构象中并且可用于例如引发受试者的中和免疫应答。

[0022] I. 术语概述

[0023] 除非另有说明,否则按照常规用法使用技术术语。常用分子生物学术语的定义参见Benjamin Lewin, *Genes X*, published by Jones&Bartlett Publishers, 2009; 和Meyers et al. (eds.), *The Encyclopedia of Cell Biology and Molecular Medicine*, published by Wiley-VCH, in 16 volumes, 2008; 以及其他类似的参考资料。

[0024] 如本文所用,单数形式“一个”(a)、“一个”(an)和“所述”既指单数也指复数,除非上下文另有明确指示。例如,术语“一个抗原”包括单个或多个抗原并且可以被认为等同于短语“至少一个抗原”。如本文所用,术语“包含”是指“包括”。还应理解的是,针对核酸或多肽给出的任何和所有碱基大小或氨基酸大小,以及所有分子量或分子量值都是近似的,并且旨在描述,除非另有说明。尽管可以使用与本文所述的方法和材料相似或等效的许多方法和材料,但下文描述了特别合适的方法和材料。如有冲突,以本说明书(包括术语解释)为准。此外,材料、方法和实例仅是说明性的而不是限制性的。为了便于查看各种实施方案,提供以下术语解释:

[0025] **佐剂**:用于增强抗原性的媒介物。在一些实施方案中,佐剂可以包括其上吸附抗原的矿物质(明矾、氢氧化铝或磷酸盐)的悬浮液;或油包水乳液,例如,其中抗原溶液在矿物油中乳化(弗氏(Freund)不完全佐剂),有时包括杀死的分枝杆菌(弗氏完全佐剂)以进一步增强抗原性(抑制抗原降解和/或使得巨噬细胞涌入)。在一些实施方案中,在所公开免疫原性组合中使用的佐剂是卵磷脂和卡波姆均聚物的组合(例如可从Advanced Bioadjuvants, LLC商购的ADJUPLEX™佐剂,另参见Wegmann, *Clin Vaccine Immunol*, 22(9): 1004-1012, 2015)。用于所公开免疫原性组合的其他佐剂包括QS21纯化植物提取物、Matrix M、AS01、MF59和ALFQ佐剂。免疫刺激性寡核苷酸(例如包括CpG基序的寡核苷酸)也可以用作佐剂。佐剂包括生物分子(“生物佐剂”),例如共刺激分子。示例性佐剂包括IL-2、RANTES、GM-CSF、TNF- α 、IFN- γ 、G-CSF、LFA-3、CD72、B7-1、B7-2、OX-40L、4-1BBL和Toll样受体(TLR)激动剂,例如TLR-9激动剂。佐剂的其他说明可以参见例如Singh(ed.) *Vaccine Adjuvants and Delivery Systems*. Wiley-Interscience, 2007)。佐剂可以与所公开免疫原组合使用。

[0026] **施用**:通过选择的途径将药剂(例如所公开免疫原)引入受试者。施用可以是局部的或全身的。例如,如果选择的途径是鼻内,则通过将所述组合物引入所述受试者的鼻道来施用药剂(例如包含稳定在融合前构象中的重组SARS-CoV-2 S胞外域三聚体的免疫原)。示例性施用途径包括但不限于口服途径、注射(例如皮下、肌肉内、皮内、腹膜内和静脉内)途径、舌下途径、直肠途径、经皮(例如,局部)途径、鼻内途径、阴道途径和吸入途径。

[0027] **氨基酸取代**:用不同的氨基酸替代多肽中的一个氨基酸。

[0028] **抗体**:特异性结合和识别分析物(抗原)(如SARS-CoV-2 S蛋白、其抗原片段或所述抗原的二聚体或多聚体)的免疫球蛋白、抗原结合片段或其衍生物。术语“抗体”在本文中以最广义的含义使用并且涵盖各种抗体结构,包括但不限于单克隆抗体、多克隆抗体、多特异性抗体(例如双特异性抗体)和抗体片段,只要它们表现出所需的抗原-结合活性即可。抗体

的非限制性实例包括,例如,完整的免疫球蛋白及其保留抗原结合亲和力的变体和片段。抗体片段的实例包括但不限于Fv、Fab、Fab'、Fab'-SH、F(ab')₂;双价抗体;线性抗体;单链抗体分子(例如scFv);以及由抗体片段形成的多特异性抗体。抗体片段包括通过修饰整个抗体产生或使用重组DNA方法从头合成的抗原结合片段(参见例如,Kontermann and Dubel (Ed), *Antibody Engineering*, Vols.1-2, 2nd Ed., Springer Press, 2010)。

[0029] 递体:一种抗原可以连接的免疫原性分子。当连接至递体时,抗原可能变得更具免疫原性。选择递体以增强抗原的免疫原性和/或引发在诊断、分析和/或治疗方面有益的针对递体的抗体。可用的递体包括聚合物递体,可以是天然的(例如,细菌或病毒的蛋白质)、含有一个或更多个可附接反应物部分的官能团的半合成或合成材料。

[0030] 保守变体:“保守”氨基酸取代是那些基本上不影响或降低蛋白质功能(例如当施用至受试者时蛋白质诱导免疫应答的能力)的取代。术语保守变异还包括使用取代的氨基酸代替未取代的亲本氨基酸。此外,改变、添加或删除被编码序列中的单个氨基酸或小百分比的氨基酸(例如小于5%,在一些实施方案中小于1%)的缺失或添加是保守变异,其中所述改变能够用化学上相似的氨基酸取代氨基酸。

[0031] 以下六组是被认为是彼此保守取代氨基酸的实例:

[0032] 1) 丙氨酸(A)、丝氨酸(S)、苏氨酸(T);

[0033] 2) 天冬氨酸(D)、谷氨酸(E);

[0034] 3) 天冬酰胺(N)、谷氨酰胺(Q);

[0035] 4) 精氨酸(R)、赖氨酸(K);

[0036] 5) 异亮氨酸(I)、亮氨酸(L)、蛋氨酸(M)、缬氨酸(V);和

[0037] 6) 苯丙氨酸(F)、酪氨酸(Y)、色氨酸(W)。

[0038] 非保守取代是降低所述重组SARS-CoV-2 S胞外域三聚体的活性或功能(例如当施用于受试者时诱导免疫应答的能力)的取代。例如,如果氨基酸残基对于蛋白质的功能是必不可少的,那么即使是保守取代也可能会破坏所述活性。因此,保守取代不会改变所关注蛋白质的基本功能。

[0039] 对照:参考标准。在一些实施方案中,对照是从健康患者获得的阴性对照样品。在其他实施方案中,对照是从诊断患有SARS-CoV-2感染(例如SARS-CoV-2)的患者获得的阳性对照样品。在其他实施方案中,对照是历史对照或标准参考值或值范围(例如先前测试的对照样品,例如预后或结果已知的SARS-CoV-2感染的患者组,或代表基线值或正常值的样品组)。

[0040] 测试样品和对照之间的差异可以增大,或相反地减少。所述差异可以是质的差异或量的差异,例如统计学上显著的差异。在一些实例中,差异是相对于对照增加或减少至少约5%,例如至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、至少约90%、至少约100%、至少约150%、至少约200%、至少约250%、至少约300%、至少约350%、至少约400%、至少约500%或大于500%。

[0041] 冠状病毒:一个正义单链RNA病毒家族,已知会造成严重的呼吸道疾病。目前已知感染人类的冠状病毒家族病毒来自α冠状病毒属和β冠状病毒属。此外,有人认为γ冠状病毒属和δ冠状病毒属可能在未来感染人类。

[0042] β冠状病毒的非限制性实例包括SARS-CoV-2、中东呼吸综合征冠状病毒(MERS-

CoV)、严重急性呼吸综合征冠状病毒(SARS-CoV)、人类冠状病毒HKU1(HKU1-CoV)、人类冠状病毒OC43(OC43-CoV)、鼠肝炎病毒(MHV-CoV)、蝙蝠SARS样冠状病毒WIV1(WIV1-CoV)和人类冠状病毒HKU9(HKU9-CoV)。 α 冠状病毒的非限制性实例包括人类冠状病毒229E(229E-CoV)、人类冠状病毒NL63(NL63-CoV)、猪流行性腹泻病毒(PEDV)和传染性胃肠炎冠状病毒(TGEV)。 δ 冠状病毒的非限制性实例是猪 δ 冠状病毒(SDCV)。

[0043] 病毒基因组被加帽、聚腺苷酸化并覆盖有核衣壳蛋白。冠状病毒病毒体包括含有称为刺突(S)蛋白的I型融合糖蛋白的病毒包膜。大多数冠状病毒具有共同的基因组组织,复制酶基因包括在基因组的5'-部分中,结构基因包括在基因组的3'-部分中。

[0044] 简并变体:在本公开的语境中,“简并变体”是指编码多肽的多核苷酸,所述多肽包括由于遗传密码而简并的序列。有20种天然氨基酸,其中大部分由一个以上的密码子决定。因此,所有编码肽的简并核苷酸序列都包括在内,只要所述核苷酸序列编码肽的氨基酸序列不变。

[0045] 有效量:足以引发受试者所需反应(如免疫应答)的药剂(如免疫原)的量。应当理解的是,可能需要多次施用所公开免疫原以获得针对所关注抗原的保护性免疫应答,和/或施用所公开免疫原作为初免加强方案中的“初免”,其中加强免疫原可以不同于初免免疫原。因此,所公开免疫原的有效量可以是足以引发受试者致敏免疫应答的免疫原的量,所述致敏免疫应答随后可以用相同或不同的免疫原加强以引发保护性免疫应答。

[0046] 在一个实例中,所需的反应是抑制或减少或预防SARS-CoV-2感染。无需完全消除或减少或预防SARS-CoV-2感染,该方法也有效。例如,施用有效量的免疫原可诱导免疫应答,与合适的对照相比,所述免疫应答将SARS-CoV-2感染(例如,通过细胞感染或感染SARS-CoV-2的受试者数量或百分比来测量)减少所需的程度,例如至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少98%或甚至至少100%(消除或预防可检测到的SARS-CoV-2感染)。

[0047] 表位:抗原决定簇。这些是分子上具有抗原性的特定化学基团或肽序列,因此它们会引发特异性免疫应答,例如,表位是B和/或T细胞对其作出反应的抗原区域。抗体可以与特定的抗原表位结合,例如SARS-CoV-2 S胞外域上的表位。表位可以由连续氨基酸或通过蛋白质三级折叠并列的非连续氨基酸形成。

[0048] 表达:核酸序列的转录或翻译。例如,基因当其DNA被转录成RNA或RNA片段时表达,在某些实例中,会被加工成mRNA。基因也可以在其mRNA被翻译成氨基酸序列(例如蛋白质或蛋白质片段)时表达。在特定的实例中,异源性基因在转录成RNA时被表达。在另一实例中,异源性基因在其RNA被翻译成氨基酸序列时被表达。术语“表达”在本文中用于表示转录或翻译。调控表达可以包括对转录、翻译、RNA转运和加工、中间分子(例如mRNA)降解的控制,或通过特定蛋白质分子产生后的激活、失活、区室化或降解的控制。

[0049] 表达控制序列:调控与其可操作连接的异源性核酸序列表达的核酸序列。当所述表达控制序列控制和调控核酸序列的转录和视情况而定的翻译时,表达控制序列可操作连接至所述核酸序列。因此,表达控制序列可以包括适当的启动子、增强子、转录终止子、蛋白质编码基因前面的起始密码子(ATG)、内含子剪接信号、为正确翻译mRNA该基因正确阅读框的维持序列,和终止密码子。术语“控制序列”旨在至少包括其存在可以影响表达的组分,并且还可以包括其存在是有利的附加组分,例如前导序列和融合伴侣序列。表达控制序列可

以包括启动子。

[0050] 启动子是足以指导转录的最小序列。还包括那些启动子元件,足以使启动子依赖性基因表达对于细胞类型特异性、组织特异性可控或可被外部信号或药剂诱导;这些元件可能位于基因的5'或3'区域。包括组成型和诱导型启动子。例如,当在细菌系统中克隆时,可以使用诱导型启动子,例如细菌噬菌体λ的pL、plac、ptrp、ptac (ptrp-lac杂合启动子)等。在一个实施方案中,当在哺乳动物细胞系统中克隆时,可以使用衍生自哺乳动物细胞基因组的启动子(例如金属硫蛋白启动子)或衍生自哺乳动物病毒的启动子(例如逆转录病毒长末端重复序列;腺病毒晚期启动子;牛痘病毒7.5K启动子)。通过重组DNA或合成技术产生的启动子也可用于转录所述核酸序列。

[0051] 表达载体:包含重组多核苷酸的载体,所述重组多核苷酸包含与待表达核苷酸序列可操作地连接的表达控制序列。表达载体包含足以表达的顺式作用元件;用于表达的其他元件可由宿主细胞或在体外表达系统中提供。表达载体包括本领域已知的所有载体,例如掺入重组多核苷酸的粘粒、质粒(例如,裸露的或含在脂质体中的质粒)和病毒(例如,慢病毒、逆转录病毒、腺病毒和腺相关病毒)。

[0052] 铁蛋白:一种储存铁并以受控方式释放铁的蛋白质。几乎所有生物都产生这种蛋白质。铁蛋白多肽组装成24个蛋白质亚基的球状蛋白质复合物,24个亚基中的每一个都包含单个铁蛋白多肽。在一些实例中,铁蛋白用于形成在其表面呈递抗原(例如,SARS-CoV-2 S胞外域三聚体)的纳米颗粒。

[0053] 异源性的:源于不同的遗传来源。与细胞异源的核酸分子来源于表达其细胞以外的遗传来源。在一个特定的非限制性实例中,编码重组SARS-CoV-2 S胞外域的异源性核酸分子在细胞(例如哺乳动物细胞)中表达。将异源性核酸分子引入细胞或生物体的方法是本领域熟知的,例如注射含有编码所公开免疫原的核酸的纳米颗粒,或用所述核酸转化,包括电穿孔、脂质体转染、粒子枪加速和同源性重组。

[0054] 宿主细胞:可以在其中繁殖载体并表达载体DNA的细胞。所述细胞可以是原核细胞或真核细胞。该术语还包括受试者宿主细胞的任何后代细胞。应当理解的是,所有后代细胞可能与亲代细胞不同,因为在复制过程中可能会发生突变。然而,当使用术语“宿主细胞”时,包括此类后代细胞。

[0055] 免疫应答:免疫系统细胞(如B细胞、T细胞或单核细胞)对刺激的反应。在一个实施方案中,所述反应对特定抗原是特异性的(“抗原特异性反应”)。在一个实施方案中,免疫应答是T细胞反应,例如CD4+反应或CD8+反应。在另一个实施方案中,所述反应是B细胞反应,并产生特异性抗体。

[0056] 免疫原:可以在动物体内引发免疫应答的化合物、组合物或物质(例如,重组SARS-CoV-2 S胞外域三聚体),包括注射或吸收到动物体内的组合物。向受试者施用免疫原可产生针对所关注病原体的保护性免疫。

[0057] 免疫原性组合物:包含所公开重组SARS-CoV-2 S胞外域三聚体的组合物,当施用于受试者时,其诱导针对SARS-CoV-2可测量的CTL反应,或诱导针对SARS-CoV-2可测量的B细胞反应(例如产生抗体)。它还涉及编码所公开重组SARS-CoV-2 S胞外域三聚体的原聚体的分离核酸分子和载体,其可用于表达所述原聚体(并因此用于引发针对重组SARS-CoV-2 S胞外域三聚体的免疫应答)。对于体内使用,免疫原性组合物通常包括在药学上可接受的

递体中的重组SARS-CoV-2 S胞外域三聚体或编码所述重组SARS-CoV-2 S胞外域三聚体的原聚体的核酸分子,并且还可以包括其他药剂,例如佐剂。

[0058] 抑制或治疗疾病:抑制疾病或病症的全面发展,例如,针对有患SARS-CoV-2感染等疾病风险的受试者。“治疗”是指在疾病或病理状况开始发展后改善其体征或症状的治疗干预。提及疾病或病理状况的术语“改善”是指任何可观察到的有益治疗效果。抑制疾病可以包括预防或降低所述疾病的风险,例如预防或降低病毒感染的风险。可以证明有益效果,例如,通过延迟易感受试者疾病临床症状的发作、减轻一些或所有临床疾病症状的严重性、减缓疾病的进展、减少病毒载量、改善所述受试者整体健康或幸福感状况证明,或通过特定疾病特有的其他参数证明。“预防性”治疗是对未表现出疾病迹象或仅表现出早期迹象的受试者进行治疗,目的是降低发生病理学症状的风险。

[0059] 分离的:“分离的”生物组分已与其他生物组分(例如所述成分天然存在的其他生物组分,例如其他染色体和染色体外DNA、RNA和蛋白质)基本分离或纯化。已“分离”的蛋白质、肽、核酸和病毒包括通过标准纯化方法纯化的蛋白质、肽、核酸和病毒。分离的不需要绝对纯度,可以包括至少50%分离(例如至少75%、80%、90%、95%、98%、99%,甚至99.9%分离)的蛋白质、肽、核酸或病毒分子。

[0060] 接头和连接:一种双功能分子,可用于将两个分子连接成一个连续分子。肽接头的非限制性实例包括甘氨酸-丝氨酸肽接头。除非上下文另有说明,否则提及“连接”第一多肽和第二多肽,或“连接”在一起的两个多肽,或与第二多肽“连接”的第一多肽,是指共价连接(例如通过肽接头使得所述第一多肽和所述第二多肽形成连续的多肽链)。如果涉及肽接头,则所述第一多肽和所述第二多肽的共价键可以连接至肽接头的N-末端和C-末端。通常,这种连接是使用分子生物学技术来完成的,以遗传操作编码通过肽接头连接至所述第二多肽的所述第一多肽的DNA。

[0061] 天然蛋白质、序列或二硫键:未经例如选择性突变修饰的多肽、序列或二硫键。例如,选择性突变以将抗原的抗原性集中到靶标表位,或将二硫键引入天然蛋白质中不存在的蛋白质。天然蛋白质或天然序列也称为野生型蛋白质或野生型序列。非天然二硫键是天然蛋白质中不存在的二硫键,例如,由于通过基因工程将一个或多个半胱氨酸残基引入蛋白质中而在所述蛋白质中形成的二硫键。

[0062] 核酸分子:核苷酸的聚合形式,可包括RNA、cDNA、基因组DNA,以及上述合成形式和混合聚合物的有义链和反义链。核苷酸是指核糖核苷酸、脱氧核苷酸或任一类型核苷酸的修饰形式。如本文所用,术语“核酸分子”与“核酸”和“多核苷酸”同义。除非另有说明,核酸分子的长度通常至少为10个碱基。该术语包括单链和双链形式的DNA。多核苷酸可以包括通过天然存在的和/或非天然存在的核苷酸键连接在一起的天然存在和修饰的核苷酸之一或两者。“cDNA”是指单链或双链形式的与mRNA互补或相同的DNA。“编码”是指多核苷酸(例如基因、cDNA或mRNA)中特定核苷酸序列的固有特性,用作在生物过程中合成其他聚合物和大分子的模板,这些聚合物和大分子具有限定的核苷酸(即rRNA、tRNA和mRNA)序列或限定的氨基酸序列以及由此产生的生物学特性。

[0063] 可操作地连接:当第一核酸序列与第二核酸序列有功能关系时,所述第一核酸序列与所述第二核酸序列可操作地连接。例如,如果启动子影响编码序列的转录或表达,则所述启动子与所述编码序列可操作地连接。通常,可操作连接的核酸序列是连续的,并且在需

要连接两个蛋白质编码区时位于相同的阅读框中。

[0064] 药学上可接受的递体:使用的药学上可接受的递体是常规递体。Remington's Pharmaceutical Sciences, E.W. Martin, Mack Publishing Co., Easton, PA, 19th Edition, 1995, 说明了适用于药物递送所公开免疫原的组合物和配制物。

[0065] 一般而言,递体的性质将取决于所采用的特定施用方式。例如,肠胃外配制物通常包含可注射流体,其包括作为媒介物的药学和生理学上可接受的流体,例如水、生理盐水、平衡盐溶液、葡萄糖水溶液、甘油等。对于固体组合物(例如粉剂、丸剂、片剂或胶囊剂),常规的无毒固体递体可以包括例如药物级的甘露醇、乳糖、淀粉或硬脂酸镁。除了生物中性递体之外,待施用的药物组合物(例如免疫原性组合物)可以含有少量无毒辅助物质,例如润湿剂或乳化剂、防腐剂和pH缓冲剂等,例如乙酸钠或山梨糖醇单月桂酸酯。在特定实施方案中,适合于向受试者施用的递体可以是无菌的,和/或悬浮的或以其他方式含在单位剂型中,所述单位剂型含有适合诱导所需免疫应答的组合物一个或更多个测量的剂量。它也可能随附有用于治疗的药物。单位剂型可以,例如在含有无菌内容物的密封小瓶或用于受试者注射的注射器中,或被冻干用于随后的溶解和施用,或者是固体或控释剂型。

[0066] 多肽:任意氨基酸链,而不管长度或翻译后的修饰(例如糖基化或磷酸化)。“多肽”适用于氨基酸聚合物,包括天然存在的氨基酸聚合物和非天然存在的氨基酸聚合物以及其中一个或更多个氨基酸残基是非天然氨基酸(例如相应的天然氨基酸的人工化学模拟物)的氨基酸聚合物。“残基”是指通过酰胺键或酰胺键模拟物掺入多肽中的氨基酸或氨基酸模拟物。多肽具有氨基末端(N-末端)和羧基末端(C-末端)。“多肽”与肽或蛋白质可互换使用,并且在本文中用于指氨基酸残基的聚合物。

[0067] 初免-加强疫苗接种:一种免疫疗法,包括施用第一免疫原性组合物(初免疫苗),然后向受试者施用第二免疫原性组合物(加强疫苗)以诱导免疫应答。在一些实例中,所述初免疫苗和/或加强疫苗包括表达免疫应答所针对抗原的载体(例如病毒载体、RNA或DNA载体)。在施用初免疫苗的合适时间间隔后向所述受试者施用加强疫苗,并且本文公开了此类时间跨度的实例。在一些实施方案中,所述初免疫苗、加强疫苗或两者另外包括佐剂。在一个非限制性实例中,所述初免疫苗是基于DNA的疫苗(或基于基因递送的其他疫苗),而加强疫苗是基于蛋白质亚基或蛋白质纳米颗粒的疫苗。

[0068] 蛋白质纳米颗粒:一种多亚基、自组装、基于蛋白质的多面体形结构。所述亚基各自由蛋白质(例如糖基化多肽)组成,并且任选地由以下单个或更多个特征组成:核酸、辅基、有机和无机化合物。在一些实施方案中,所公开三聚体刺突蛋白的原聚体可以融合或缀合至蛋白质纳米颗粒的亚基以在每个蛋白质纳米颗粒上提供三聚体刺突的多个拷贝。蛋白质纳米颗粒的非限制性实例包括铁蛋白纳米颗粒(参见例如Zhang, Y. Int. J. Mol. Sci., 12: 5406-5421, 2011, 通过引用并入本文)、纳米隔室纳米颗粒(参见例如, Sutter et al., Nature Struct. and Mol. Biol., 15: 939-947, 2008, 通过引用并入本文)、硫氧化酶还原酶(SOR)纳米颗粒(参见,例如, Urich et al., Science, 311: 996-1000, 2006, 通过引用并入本文)、二氧四氢喋呤合酶纳米颗粒(参见例如, Zhang et al., J. Mol. Biol., 306: 1099-1114, 2001), 和丙酮酸脱氢酶纳米颗粒(参见,例如, Izard et al., PNAS 96: 1240-1245, 1999, 通过引用并入本文)。铁蛋白、纳米隔室、SOR、二氧四氢喋呤合酶和丙酮酸脱氢酶是单体蛋白质,它们自组装成球状蛋白质复合物,在某些情况下,分别由24、60、24、60和60个蛋白质亚

基组成。Heinze et al., J Phys Chem B., 120 (26) :5945-52, 2016; Hsia et al., Nature, 535 (7610) :136-9, 2016; 和 King, et al., Nature, 510 (7503) :103-8, 2014, 说明了另外的蛋白质纳米颗粒结构; 每个文献都以引用的方式并入本文。

[0069] 重组: 重组核酸分子是具有以下序列的分子: 非天然存在的序列, 例如, 包括一个或更多个核酸取代、缺失或插入; 和/或由两个原本分开的序列片段人工组合而成的序列。这种人工组合可以通过化学合成或更常见地通过人工操作 (例如通过基因工程技术) 分离的核酸片段来实现。重组病毒是一种包括基因组的病毒, 该基因组包括重组核酸分子。重组蛋白是具有非天然存在序列或具有由两个原本分开的序列片段人工组合而成序列的蛋白。在若干实施方案中, 所述重组蛋白由已被引入宿主细胞 (例如细菌细胞或真核细胞) 或重组病毒基因组的异源性 (例如重组) 核酸编码。

[0070] SARS-CoV-2: 也称为2019-nCoV或2019新型冠状病毒, SARS-CoV-2是一种正义单链RNA病毒, 属于β冠状病毒, 已成为严重急性呼吸道感染的高度致死原因。病毒基因组被加帽、聚腺苷酸化并覆盖有核衣壳蛋白。SARS-CoV-2病毒体包括带有大刺突糖蛋白的病毒包膜。与大多数冠状病毒一样, SARS-CoV-2基因组具有共同的基因组组织, 复制酶基因包括在5'-三分之二基因组中, 而结构基因则包括在3'-三分之一基因组中。SARS-CoV-2基因组以5'-刺突(S)-包膜(E)-膜(M)和核衣壳(N)-3'的顺序编码标准的结构蛋白基因集。SARS-CoV-2感染的症状包括发烧和呼吸系统疾病, 例如干咳和呼吸急促。严重感染病例可发展为严重肺炎、多器官衰竭和死亡。从暴露到出现症状的时间约为2至14天。

[0071] 检测病毒感染的标准方法可用于检测SARS-CoV-2感染, 包括但不限于评估患者症状和背景以及基因检测, 例如逆转录聚合酶链反应(rRT-PCR)。该测试可以在患者样品 (例如呼吸或血液样品) 上进行。

[0072] SARS-CoV-2刺突(S)蛋白: 一种I类融合糖蛋白, 最初合成为大小约为1270个氨基酸的前体蛋白。单个前体S多肽形成同源三聚体并在高尔基体内进行糖基化以及去除信号肽的加工。S多肽包括由大约在位置685/68之间的蛋白酶切割位点隔开的S1和S2蛋白。在该位点的切割产生分开的S1和S2多肽链, 它们在同源三聚体中作为S1/S2原聚体保持缔合。S1亚基位于病毒膜的远端, 含有介导病毒与其宿主受体附着的受体结合域(RBD)。S2亚基含有融合蛋白机制, 例如融合肽、两个七重复序列(HR1和HR2)和融合糖蛋白典型的中心螺旋、跨膜域和胞质尾域。

[0073] 所公开SARS-CoV-2 S蛋白及其片段中使用的编号是相对于SARS-CoV-2的S蛋白的, 其序列如SEQ ID NO:1所述, 并保藏为NCBI参考编号YP_009724390.1, 以全文引用的方式并入本文。

[0074] SARS-CoV-2刺突(S)蛋白融合前构象: SARS-CoV-2 S蛋白在分泌系统中加工成成熟的SARS-CoV-2 S蛋白后且在触发融合事件之前, 其胞外域采用的结构构象, 所述融合事件使得SARS-CoV-2 S转变为融合后构象。本文公开了示例性SARS-CoV-2 S蛋白在融合前构象中的三维结构 (参见实施例1)。

[0075] 与天然SARS-CoV-2 S序列相比, “稳定在融合前构象中”的SARS-CoV-2 S胞外域三聚体包含一个或更多个氨基酸取代、缺失或插入, 与由相应天然SARS-CoV-2 S序列形成的SARS-CoV-2胞外域三聚体相比, 所述氨基酸取代、缺失或插入使所述融合前构象保留更多。通过一个或更多个氨基酸取代、缺失或插入对融合前构象的“稳定”可以是例如能量稳定

(例如,相对于融合后开放构象降低融合前构象的能量)和/或动力学稳定(例如,降低从融合前构象到融合后构象的转变速率)。此外,与相应的天然SARS-CoV-2 S序列相比,SARS-CoV-2 S胞外域三聚体在融合前构象中稳定可以包括增强对变性的抵抗力。确定SARS-CoV-2 S胞外域三聚体是否在融合前构象中的方法包括(但不限于)负染色电子显微镜技术。

[0076] 序列同一性:氨基酸序列之间的相似性以序列之间的相似性来表示,也称为序列同一性。序列同一性经常以百分比同一性来衡量。百分比越高,两个序列越相似。当使用标准方法比对时,多肽的同源物、直向同源物或变体将具有较高度度的序列同一性。

[0077] 用于比较的序列比对方法是本领域众所周知的。各种程序和比对算法在以下文献中进行说明:Smith&Waterman,Adv. Appl. Math. 2:482,1981;Needleman&Wunsch, J. Mol. Biol. 48:443,1970;Pearson&Lipman,Proc. Natl. Acad. Sci. USA 85:2444,1988;Higgins&Sharp,Gene,73:237-44,1988;Higgins&Sharp,CABIOS 5:151-3,1989;Corpet et al.,Nuc. Acids Res. 16:10881-90,1988;Huang et al. Computer Appls. in the Biosciences 8,155-65,1992;和Pearson et al.,Meth. Mol. Bio. 24:307-31,1994。Altschul et al.,J. Mol. Biol. 215:403-10,1990,详细考虑了序列比对方法和同源性计算。

[0078] 多肽(例如SARS-CoV-2 S胞外域)的同源物和变体通常的特征在于,在与所关注氨基酸序列的全长比对中,拥有至少约75%,例如至少约80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性。当通过该方法评估时,与参考序列相似性更高的蛋白质显示出的百分比同一性将增大,例如至少80%、至少85%、至少90%、至少95%、至少98%或至少99%的序列同一性。当比较少于整个序列的序列同一性时,同源物和变体通常在10-20个氨基酸的短窗口内具有至少80%的序列同一性,并且可能具有至少85%或至少90%或95%的序列同一性,这取决于它们与参考序列的相似性。在此类短窗口上确定序列同一性的方法可在互联网上的NCBI网站获得。

[0079] 如本文所用,提及“至少90%同一性”或类似说法是指与特定的参考序列具有“至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或甚至100%的同一性”。

[0080] 信号肽:一种短的氨基酸序列(例如,长度约为10-35个氨基酸),可将新合成的分泌蛋白或膜蛋白引导至膜并穿过膜(例如,内质网膜)。信号肽通常位于多肽的N端并被信号肽酶移除。信号肽序列通常含有三个常见的结构特征:N-末端极性碱性区(n区)、疏水核心和亲水c区。

[0081] 单链SARS-CoV-2 S胞外域:重组SARS-CoV-2 S胞外域在单个连续多肽链中包括SARS-CoV-2 S₁和S₂蛋白。单链SARS-CoV-2 S胞外域可以三聚体化形成SARS-CoV-2 S胞外域三聚体。单个SARS-CoV-2 S胞外域包括防止蛋白酶在S₁/S₂切割位点切割的突变。因此,当在细胞中产生时,所述SARS-CoV-2 S多肽不会被切割成分开的S₁和S₂多肽链。

[0082] 可溶性蛋白质:在室温下能够溶解在水性液体中并保持溶解状态的蛋白质。蛋白质的溶解度可能会根据水基液体中蛋白质的浓度、液体的缓冲条件、液体中其他溶质的浓度(例如盐和蛋白质的浓度)以及液体的热量而改变。在若干实施方案中,可溶性蛋白质是室温下在磷酸盐缓冲盐水(pH 7.4)中溶解浓度为至少0.5mg/ml并保持溶解状态至少48小时的蛋白质。

[0083] 受试者:活体的多细胞脊椎动物生物,类别包括人类和非人类哺乳动物有机体,例如非人类灵长类动物、猪、骆驼、蝙蝠、羊、牛、狗、猫、啮齿动物等。在一个实例中,受试者是人类。在另一个实例中,选择的受试者需要抑制SARS-CoV-2感染。例如,所述受试者要么未感染SARS-CoV-2且有感染风险,要么已感染需要治疗。

[0084] T4次要纤维蛋白三聚体化域(T4 Fibritin trimerization domain):也称为“foldon”域,T4次要纤维蛋白三聚体化域包含天然形成三聚体结构的氨基酸序列。在一些实例中,T4次要纤维蛋白三聚体化域可以连接至所公开重组SARS-CoV-2 S蛋白胞外域的C端。在一个实例中,T4次要纤维蛋白三聚体化域包含如(GYIPEAPRDGQAYVRKDGWVLLSTF)(SEQ ID NO:6)所示的氨基酸序列。在一些实施方案中,蛋白酶切割位点(例如凝血酶切割位点)可以包括在所述重组SARS-CoV-2 S胞外域C端和T4次要纤维蛋白三聚体化域之间,以便根据需要(例如,在表达和纯化重组SARS-CoV-2 S胞外域之后)去除三聚体化域。

[0085] 跨膜域:插入脂质双层(例如细胞或病毒或类病毒颗粒的脂质双层)的氨基酸序列。跨膜域可用于将抗原锚定到膜上。在一些实例中,跨膜域是SARS-CoV-2 S跨膜域。

[0086] 疫苗:诱导受试者预防性或治疗性免疫应答的药物组合物。在某些情况下,免疫应答是保护性免疫应答。通常,疫苗会诱导针对病原体(例如病毒病原体)的抗原或与病理状况相关细胞组成物的抗原特异性免疫应答。疫苗可以包括多核苷酸(例如编码所公开抗原的核酸)、肽或多肽(例如所公开抗原)、病毒、细胞或一种或更多种细胞组成物。在一个非限制性实例中,与对照相比,疫苗诱导的免疫应答降低与SARS-CoV-2感染相关症状的严重性和/或降低病毒载量。在另一个非限制性实例中,与对照相比,疫苗诱导的免疫应答减少和/或预防SARS-CoV-2感染。

[0087] 载体:含带有启动子的DNA或RNA分子的实体,所述启动子可操作地连接至所关注抗原的编码序列并且可以表达所述编码序列。非限制性实例包括裸露或包装的(脂质和/或蛋白质)DNA、裸露或包装的RNA、可能没有复制能力的病毒或细菌或其他微生物的亚组分或可能有复制能力的病毒或细菌或其他微生物的亚组分。载体有时被称为构建体。重组DNA载体是具有重组DNA的载体。载体可以包括使其在宿主细胞中复制的核酸序列,例如复制起点。载体还可以包括一种或更多种可选择的标记基因和本领域已知的其他遗传元件。病毒载体是具有至少一些衍生自一种或更多种病毒的核酸序列的重组核酸载体。

[0088] 类病毒颗粒(VLP):一种非复制的病毒外壳,衍生自多种病毒中的任一种。VLP通常由一种或更多种病毒蛋白组成,例如但不限于被称为衣壳蛋白、包被蛋白、壳蛋白、表面蛋白和/或包膜蛋白的那些蛋白,或衍生自这些蛋白的颗粒形成多肽。在适当的表达系统中重组表达蛋白质后,VLP可以自发形成。在重组表达病毒蛋白后,VLP的存在可以使用常规技术检测,例如通过电子显微镜技术、生物物理表征等。此外,可以通过已知技术(例如密度梯度离心)分离VLP,并通过特征密度条带进行确定。例如,参见Baker et al. (1991) *Biophys.J.*60:1445-1456;和Hagensee et al. (1994) *J.Virol.*68:4503-4505;Vincente, *J Invertebr Pathol.*,2011;Schneider-Ohrum and Ross, *Curr.Top.Microbiol.Immunol.*, 354:53073,2012。

[0089] II.重组SARS-CoV-2刺突蛋白

[0090] 本文公开了包含原聚体的重组SARS-CoV-2 S胞外域三聚体,所述原聚体包含一个或更多个氨基酸取代,其抑制所述SARS-CoV-2 S蛋白从融合前构象到融合后构象的构象变

化,因此将所述SARS-CoV-2 S胞外域三聚体稳定在融合前构象中。与在融合前构象中不稳定的相应SARS-CoV-2 S胞外域三聚体相比,所述重组SARS-CoV-2 S胞外域三聚体产生更强的免疫应答。

[0091] 天然SARS-CoV-2 S蛋白(包括胞外域和TM域和CT域)的示例性序列如SEQ ID NO:1 (NCBI参考编号YP_009724390.1,通过引用并入本文)所示:

[0092] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH
VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIVKVECFQCNDPFLGVYYH
KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSAL
EPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTF
KCYGVSPTKLNLDCTFNVYADSFVIRGDEVQRDIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYR
LFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKST
NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQTLEILDITPCSGGGVSVITPGTNTSNQV
AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNSPRRAR
SVASQSI IAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNSLLLQYGSFCTQ
LNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGD
CLGDIAARDLICAQKFNGLTVLPPLLTDEMQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVL
YENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDLKVEAEVQI
DRLITGRLLQSLQTYVTQQILIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSMFPQSAPHGVVFLHVTYVPA
QEKNF T TAPAI CHD GKAHFPREGV FVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDS
FKEELDKYFKNHTSPDVLGDISGINASVNNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0093] 本文所用的SARS-CoV-2 S蛋白残基的氨基酸编号是参考如SEQ ID NO:1所示的SARS-CoV-2 S序列。参考如SEQ ID NO:1所示的SARS-CoV-2 S蛋白序列,SARS-CoV-2 S蛋白的胞外域包括大约残基16-1208。残基1-15是信号肽,在细胞加工过程中被去除。所述S1/S2切割位点位于位置685/686处。HR1大约位于残基915-983处。中心螺旋大约位于残基988-1029处。HR2大约位于残基1162-1194处。S2胞外域的C-末端大约位于残基1208处。在一些实施方案中,融合前稳定的SARS-CoV-2 S胞外域三聚体的原聚体可具有HR2(例如位置1194)或胞外域(例如位置1208)或位置1194-1208之一的C-末端残基(例如,其可连接至三聚体化域,或跨膜域)。S蛋白的位置编号可能因SARS-CoV-2毒株而异,但可以比对序列以确定相关的结构域和切割位点。应当理解的是,可以在所公开免疫原中去除或修饰所述胞外域N-末端和C-末端上的一些残基(例如至多10个),而不会降低所述S胞外域三聚体作为免疫原的效用。

[0094] 在一些实施方案中,所述免疫原包含重组SARS-CoV-2 S胞外域三聚体,其包含的原聚体在HR1域和中心螺旋域之间的边界处或附近包含一个或更多个(例如两个,例如两个连续的)氨基酸取代,使所述S胞外域三聚体稳定在融合前构象中,其中所述氨基酸取代是甘氨酸取代和/或脯氨酸取代。在一些此类实施方案中,使所述S胞外域稳定在融合前构象中的一个或更多个(例如两个,例如两个连续的)氨基酸取代位于HR1的C-末端残基位置15氨基酸的N-末端和所述中心螺旋N-末端残基位置5氨基酸的C-末端之间。在一些实施方案

中,使所述SARS-CoV-2 S胞外域三聚体稳定在融合前构象中的一个或多个(例如两个,例如两个连续的)氨基酸取代位于三聚体中S胞外域原聚体的残基975至995(例如981-992)之间,其中所述氨基酸取代是甘氨酸取代和/或脯氨酸取代。在一些实施方案中,通过在三聚体中S胞外域原聚体位置D985、K986和/或V987处的甘氨酸取代和/或脯氨酸取代,将SARS-CoV-2 S胞外域三聚体稳定在融合前构象中。

[0095] 在一些实施方案中,所述免疫原包含重组SARS-CoV-2 S胞外域三聚体,其包含的原聚体在HR1域和中心螺旋域之间的边界处或附近包含一个或多个(例如两个,例如两个连续的)脯氨酸取代,使所述S胞外域三聚体稳定在融合前构象中。在一些此类实施方案中,使所述S胞外域稳定在融合前构象中的一个或多个(例如两个,例如两个连续的)脯氨酸取代位于HR1的C-末端残基位置15氨基酸的N-末端和所述中心螺旋N-末端残基位置5氨基酸的C-末端之间。

[0096] 在一些实施方案中,使所述SARS-CoV-2 S胞外域三聚体稳定在融合前构象中的一个或多个(例如两个,例如两个连续的)脯氨酸取代位于三聚体中S胞外域原聚体的残基975至995(例如981-992)之间。在一些实施方案中,通过三聚体中S胞外域原聚体中的K986P和V987P取代(“2P”),将所述SARS-CoV-2 S胞外域三聚体稳定在融合前构象中。在一些实施方案中,通过在三聚体中S胞外域原聚体位置D985、K986或V987处的一个或两个脯氨酸取代,将所述SARS-CoV-2 S胞外域三聚体稳定在融合前构象中。

[0097] 在一些实施方案中,通过一个或多个脯氨酸取代(例如K986P和V987P取代)稳定在融合前构象中的重组SARS-CoV-2 S胞外域三聚体的原聚体包含一种或更多种用于稳定在融合前构象中的额外修饰。

[0098] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体中原聚体胞外域的C-末端残基可以连接至三聚体化域以促进所述原聚体的三聚体化,并使所述原聚体的膜近端方面稳定在三聚体构型中。外源多聚化域(促进可溶性重组蛋白的三聚体稳定)的非限制性实例包括:GCN4亮氨酸拉链(Harbury et al.1993Science 262:1401-1407)、肺表面活性蛋白的三聚体化基序(Hoppe et al.1994FEBS Lett 344:191-195)、胶原蛋白(McAlinden et al.2003J Biol Chem 278:42200-42207)和噬菌体T4次要纤维蛋白(Miroshnikov et al.1998Protein Eng 11:329-414),其中任何一种都可以与本文所述重组SARS-CoV-2 S胞外域连接(例如,通过与S2胞外域的C端连接)以促进所述重组SARS-CoV-2胞外域的三聚体化。

[0099] 在一些实例中,所述S2胞外域C-末端残基可以连接至T4次要纤维蛋白域。在具体的实例中,所述T4次要纤维蛋白域可以包括氨基酸序列GYIPEAPRDGQAYVRKDGWVLLSTF (SEQ ID NO:6),其采用 β -螺旋桨构象,并且可以自主折叠和三聚体化(Tao et al.1997Structure 5:789-798)。

[0100] 任选地,异源性三聚体域通过肽接头(例如氨基酸接头)连接至所述重组SARS-CoV-2 S胞外域。可以使用的肽接头非限制性实例包括甘氨酸接头、丝氨酸接头和甘氨酸-丝氨酸接头。

[0101] SARS-CoV-2 S胞外域,包括用于稳定在融合前构象中的双脯氨酸取代并与T4次要纤维蛋白三聚体化域连接,其示例性序列如SEQ ID NO:2所示:

[0102] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH

VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLI VNNATNVVIVKVEFQFCNDPFLGVYYH
KNNKSWMESEFRVYSSANNCTFEYVVSQPFLMDLEGKQGNFNKLNREFVFKNIDGYFKIYSKHTPINLVRDLPPQGSAL
EPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLVNSASFSTF
KCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYR
LFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKST
NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRAR
SVASQSIIAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQ
LNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQYGD
CLGDIAARDLICAQKFNGLTVLPPLLTDEMIAYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVL
YENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQI
DRLITGRQLSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPA
QEKNFTTAPAI CHDGAHFPPREGV FVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDS
FKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGGYIPEAPRDGQAYVR
KDGEWVLLSTF

[0103] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体包含SEQ ID NO:2的胞外域序列。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体包含SEQ ID NO:2残基16-1208。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体包含与SEQ ID NO:2胞外域序列至少90% (例如至少95%、至少98%或至少99%) 相同的序列,其中所述SARS-CoV-2 S胞外域三聚体通过本文提供的一个或更多个修饰(例如K986P和V987P取代) 稳定在融合前构象中。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体包含与SEQ ID NO:2的残基16-1208至少90% (例如至少95%、至少98%或至少99%) 相同的序列,其中所述SARS-CoV-2 S胞外域三聚体通过本文提供的一个或更多个修饰(例如K986P和V987P取代) 稳定在融合前构象中。

[0104] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体包含SEQ ID NO:2的胞外域序列,所述原聚体各自连接至三聚体化域,例如T4次要纤维蛋白三聚体化域。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体连接至包含SEQ ID NO:2的残基16-1235的三聚体化域。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体连接至三聚体化域,所述三聚体化域包含与SEQ ID NO:2的残基16-1235至少90% (例如至少95%、至少98%或至少99%) 相同的序列,其中所述SARS-CoV-2 S胞外域三聚体通过本文提供的一个或更多个修饰(例如K986P和V987P取代) 稳定在融合前构象中。

[0105] 在一些实施方案中,例如,对于其中SARS-CoV-2 S胞外域三聚体被表达成减毒病毒疫苗或类病毒颗粒,或通过重组核酸(如mRNA) 表达的实施方案,所述SARS-CoV-2 S胞外域三聚体可以是膜锚定的。在此类实施方案中,三聚体中的原聚体通常各自包含与跨膜域(例如SARS-CoV-2 S蛋白的跨膜域(和任选的胞质尾)) 的C-末端连接。在一些实施方案中,一种或更多种肽接头(例如甘氨酸-丝氨酸接头,例如10个氨基酸的甘氨酸-丝氨酸肽接头) 可用于将所述重组SARS-CoV-2 S胞外域原聚体连接至跨膜域。只要与跨膜域连接的重组

SARS-CoV-2 S胞外域三聚体保留所需的特性(例如,SARS-CoV-2 S融合前构象),与跨膜域连接的原聚体可以包括本文提供的任何稳定突变(或它们的组合)。

[0106] SARS-CoV-2 S蛋白(包括胞外域、TM域和CT域),包括用于稳定在融合前构象中的双脯氨酸取代,所述蛋白的示例性序列如SEQ ID NO:3所示:

[0107] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH
VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIRGWIFGTTLDSKTQSLLIVNNATNVVIKVFCEQFCNDPFLGVYYH
KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGFKNLRVFNIDGYFKIYSKHTPINLVRDLPPQGSAL
EPLVDLPIGINITRFQTLALHRSYLTGDSGSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLVNSASFSTF
KCYGVSPTKLNLDLCTNVDYADSFVIRGDEVQRVAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNVNYLYR
LFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKST
NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNSPRRAR
SVASQSI IAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQ
LNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTADAGFIKQYGD
CLGDIAARDLICAQKFNGLTVLPPLLTDEMIQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVL
YENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQI
DRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPA
QEKNFITAPAI CHDGAHFREGVVFVSNHWFVTQRNFYEPQIITDNTFVSGNCDVIGIVNNTVYDPLQPELDS
FKEELDKYFKNHTSPDVLDGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCCLGKCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0108] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体包含SEQ ID NO:3的胞外域序列,所述原聚体各自连接至跨膜域和/或胞质尾。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体与包含SEQ ID NO:3残基16-1273的跨膜域连接。在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体包含的原聚体与跨膜域连接,所述跨膜域包含与SEQ ID NO:3残基16-1273至少90%(例如至少95%、至少98%或至少99%)相同的序列,其中所述SARS-CoV-2 S胞外域三聚体通过本文提供的一个或多个修饰(例如K986P和V987P取代)稳定在融合前构象中。

[0109] 在一些实施方案中,所述SARS-CoV-2 S胞外域三聚体可以由三个单链SARS-CoV-2 S胞外域原聚体组成,每个原聚体都包括一条包括S1蛋白和S2胞外域的多肽链。单链SARS-CoV-2 S胞外域原聚体可以通过使S1/S2蛋白酶切割位点突变来产生,以防止切割和形成不同的S1和S2多肽链。在一些实施方案中,单链SARS-CoV-2 S胞外域原聚体中的S1和S2多肽通过接头(例如肽接头)连接。可以使用的肽接头的实例包括甘氨酸接头、丝氨酸接头和甘氨酸-丝氨酸接头。本文公开的任何稳定突变(或其组合)都可以包括在单链SARS-CoV-2 S胞外域原聚体中,只要由此类原聚体组成的SARS-CoV-2 S胞外域三聚体保留所需的特性(例如,融合前构象)即可。

[0110] 单链SARS-CoV-2 S胞外域,包括用于稳定在融合前构象中的双脯氨酸取代,并连接至T4次要纤维蛋白三聚体化域,所述胞外域的示例性序列如SEQ ID NO:4所示:

[0111] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH

VSGTNGTKRFDNPVLPFNDGVYFASTEKSNI IRGWIFGTTLDSKTQSLLI VNNATNVVIVKVEFQFCNDPFLGVYYH
 KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGFKNLREFVFNIDGYFKIYSKHTPINLVRDLPPQGSAL
 EPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
 CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLVNSASFSTF
 KCYGVSPTKLNLDLCTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYR
 LFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKST
 NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQGFGRDIAADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
 AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGNSVNFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNSPGGSV
 ASQSI IAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLN
 RALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDCL
 GDIAARDLICAQKFNGLTVLPPLLTDEMQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
 NQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDR
 LITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPAQE
 KNFTTAPAICHGKAHFPPREGVFSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK
 EELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGGYIPEAPRDGQAYVRKD
 GEWLLSTF

[0112] 在一些实施方案中,所述重组单链SARS-CoV-2 S胞外域三聚体包含的原聚体包含 SEQ ID NO:4的胞外域序列,所述胞外域序列连接至三聚体化域(例如T4次要纤维蛋白三聚体化域)。在一些实施方案中,与跨膜域连接的重组单链SARS-CoV-2 S胞外域三聚体包含的原聚体包含SEQ ID NO:4残基16-1233。在一些实施方案中,与跨膜域连接的所述重组单链SARS-CoV-2 S胞外域三聚体包含的原聚体包含与SEQ ID NO:4残基16-1233至少90%(例如至少95%、至少98%或至少99%)相同的序列,其中所述SARS-CoV-2 S胞外域三聚体通过本文提供的一个或更多个修饰(例如K986P和V987P取代)稳定在融合前构象中。

[0113] 单链SARS-CoV-2 S蛋白(包括胞外域、TM域和CT域),包括用于稳定在融合前构象中的双脯氨酸取代,所述蛋白的示例性序列如SEQ ID NO:5所示:

[0114] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH
 VSGTNGTKRFDNPVLPFNDGVYFASTEKSNI IRGWIFGTTLDSKTQSLLI VNNATNVVIVKVEFQFCNDPFLGVYYH
 KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGFKNLREFVFNIDGYFKIYSKHTPINLVRDLPPQGSAL
 EPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
 CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLVNSASFSTF
 KCYGVSPTKLNLDLCTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYR
 LFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKST
 NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQGFGRDIAADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
 AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGNSVNFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNSPGGSV
 ASQSI IAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLN
 RALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDCL
 GDIAARDLICAQKFNGLTVLPPLLTDEMQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
 NQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDR
 LITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPAQE

KNFTTAPAICHGKAHFPPREGVVFVSNNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK
EELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLCCMTSCCSCCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0115] 在一些实施方案中,所述重组单链SARS-CoV-2 S胞外域三聚体包含的原聚体包含SEQ ID NO:5的胞外域序列,所述原聚体各自连接至跨膜域和/或胞质尾。在一些实施方案中,所述重组单链SARS-CoV-2 S胞外域三聚体包含的原聚体连接至包含SEQ ID NO:5残基16-1271的跨膜域。在一些实施方案中,所述重组单链SARS-CoV-2 S胞外域三聚体包含的原聚体连接至跨膜域,所述跨膜域包含与SEQ ID NO:5残基16-1271至少90% (例如至少95%、至少98%或至少99%) 相同的序列,其中所述SARS-CoV-2 S胞外域三聚体通过本文提供的一个或更多个修饰(例如K986P和V987P取代)稳定在融合前构象中。

[0116] 在一些实施方案中,所述重组SARS-CoV-2 S胞外域三聚体中原聚体包含用于融合前稳定的K986P和V987P取代,并且还包含N501Y、K417N和E484K取代中的一个或更多个。例如,所述重组SARS-CoV-2 S胞外域三聚体中原聚体包含K986P和V987P取代,还包括N501Y取代,K417N取代,E484K取代,N501Y和K417N取代,K417N和E484K取代,N501Y和E484K取代,或N501Y、K417N和E484K取代。

[0117] 所述重组SARS-CoV-2 S胞外域三聚体及其变体可以使用重组技术生产,或化学或酶促合成。

[0118] 所述重组SARS-CoV-2 S胞外域三聚体的类似物和变体可用于本公开的方法和系统。通过使用重组DNA技术,可以通过改变底层DNA来制备重组SARS-CoV-2 S胞外域三聚体的变体。产生变体的重组SARS-CoV-2 S胞外域三聚体结构的所有此类变化或改变都包括在本公开的范围。此类变体包括一个或更多个氨基酸残基的插入、取代或缺失,糖基化变体,未糖基化的重组SARS-CoV-2 S胞外域三聚体,有机盐和无机盐,重组SARS-CoV-2 S胞外域三聚体的共价修饰衍生物,或它们的前体。此类变体可以保持所述重组SARS-CoV-2 S胞外域三聚体的一种或更多种功能性生物活性,例如与细胞表面受体结合。可以例如通过聚乙二醇化修饰重组SARS-CoV-2 S胞外域三聚体,以延长蛋白质在受体中的半衰期,和/或使蛋白质更稳定地递送至受试者。

[0119] 在一些实施方案中,用于本公开的重组SARS-CoV-2 S胞外域三聚体通过用以下其他侧链替换20个遗传编码氨基酸(或D-氨基酸)的一个或更多个天然存在的侧链来修饰:例如,烷基,低级烷基,4-、5-、6-至7-元环烷基,酰胺,酰胺低级烷基、酰胺二(低级烷基),低级烷氧基,羟基,羧基及其低级酯衍生物,以及4-、5-、6至7-元杂环。例如,可以制备脯氨酸类似物,其中脯氨酸残基的环大小从5-元环变为4-、6-或7-元环。环状基团可以是饱和的或不饱和的,如果是不饱和的,则可以是芳族基团或非芳族基团。杂环基团可以包含一个或更多个氮、氧和/或硫杂原子。

[0120] 蛋白质纳米颗粒

[0121] 在一些实施方案中,提供了一种蛋白质纳米颗粒(例如自组装蛋白纳米颗粒),包括在其表面上展示的重组SARS-CoV-2 S胞外域三聚体。自组装蛋白纳米颗粒的非限制性实例包括铁蛋白纳米颗粒、纳米隔室纳米颗粒、硫氧化还原酶(SOR)纳米颗粒和二氧四氢喋啶合酶纳米颗粒,它们分别由包括铁蛋白、纳米隔室蛋白、SOR蛋白和二氧四氢喋啶合酶的单体亚基组装而成。额外的蛋白质纳米颗粒结构在以下文献中进行了说明:Heinze et al.,J

Phys Chem B., 120 (26) :5945-52, 2016; Hsia et al., Nature, 535 (7610) :136-9, 2016; 和 King, et al., Nature, 510 (7503) :103-8, 2014; 每个文献以引用的方式并入本文。

[0122] 在若干实施方案中,为了构建这种蛋白质纳米颗粒,编码SARS-CoV-2 S胞外域三聚体的原聚体的核酸可以融合到编码蛋白质纳米颗粒(例如铁蛋白、纳米隔室蛋白、SOR蛋白或二氧四氢嘧啶合酶蛋白)亚基的核酸,并在适当条件下在细胞中表达。融合蛋白自组装成任何可以纯化的纳米颗粒。

[0123] 在若干实施方案中,为了构建这样的蛋白质纳米颗粒,纯化的SARS-CoV-2 S胞外域三聚体可以连接(例如,通过生物缀合)至纯化的自组装蛋白纳米颗粒(例如铁蛋白、纳米隔室蛋白、SOR蛋白或二氧四氢嘧啶合酶蛋白)的亚基,并纯化所得的纳米颗粒/S三聚体。

[0124] 在一些实施方案中,所述SARS-CoV-2 S胞外域三聚体包括在自组装蛋白质纳米笼中,所述纳米笼以类似于病毒的方式引导其自身在小囊泡内从细胞中释放出来,例如, Votteler et al., “Designed proteins induce the formation of nanocage-containing extracellular vesicles”, Nature 540, 292-29, 2016所述。这种混合生物材料可以将其膜与靶细胞融合并递送其内容物,从而将货物从一个细胞转移到另一个细胞。

[0125] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接到铁蛋白亚基以构建铁蛋白纳米颗粒。铁蛋白纳米颗粒及其用于免疫目的(例如,用于针对流感抗原进行免疫)的用途已在本领域中公开(参见,例如, Kanekiyo et al., Nature, 499: 102-106, 2013, 以全文引用的方式并入本文)。铁蛋白是一种球状蛋白质,存在于所有动物、细菌和植物中,主要通过水合铁离子和质子进出矿化核的转运来控制多核 $\text{Fe(III)}_2\text{O}_3$ 形成的速率和位置。所述球状铁蛋白纳米颗粒由单体亚基组成,这些亚基是分子量约为17-20kDa的多肽。一个此类单体铁蛋白亚基的氨基酸序列实例表示为:

[0126] ESQVRQQFSKDIEKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLN ENNVPVQLTISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIAKSKDHATFNFLQWYVAEQHEEEVLFKDIL DKIELIGNENHGLYLADQYVKGIAKSRKS (SEQ ID NO:7)

[0127] 每个单体亚基具有螺旋束的拓扑结构,其中包括四个反平行的螺旋基序,第五个较短螺旋(C-末端螺旋)大致垂直于所述4螺旋束的长轴。按照惯例,螺旋分别从N-末端标记为“A、B、C、D和E”。N-末端序列与衣壳三重轴相邻并延伸至表面,而E螺旋在四重轴处堆积在一起,C-末端延伸到衣壳核中。这种堆积的结果在衣壳表面产生了两个孔。预计这些孔中的一个或两个代表水合铁扩散进出所述衣壳的点。在产生后,这些单体亚基蛋白自组装成球状铁蛋白。因此,球状铁蛋白包含24个单体亚基蛋白,并具有类衣壳结构,其具有432种对称结构。构建铁蛋白纳米颗粒的方法是本领域普通技术人员已知的,并且在本文中进一步进行说明(参见例如Zhang, Int. J. Mol. Sci., 12:5406-5421, 2011, 以全文引用的方式并入本文)。

[0128] 在具体实例中,铁蛋白多肽是大肠杆菌(*E. coli*)铁蛋白、幽门螺杆菌(*Helicobacter pylori*)铁蛋白、人轻链铁蛋白、牛蛙铁蛋白或它们的杂合体,例如大肠杆菌(*E. coli*)-人杂合铁蛋白、大肠杆菌(*E. coli*)-牛蛙杂合铁蛋白或人-牛蛙杂交铁蛋白。用于制备铁蛋白纳米颗粒(包括重组SARS-CoV-2 S胞外域)的铁蛋白多肽的示例性氨基酸序列和编码铁蛋白多肽的核酸序列可以见于GENBANK®,例如登录号ZP_03085328、ZP_06990637、EJB64322.1、AAA35832、NP_000137、AAA49532、AAA49525、AAA49524和AAA49523,

它们于2015年4月10日通过引用的方式整体并入本文。在一些实施方案中,所述重组SARS-CoV-2 S胞外域可以连接至铁蛋白亚基,其包括与如SEQ ID NO:8所示的氨基酸序列至少80%(例如至少85%、至少90%、至少95%或至少97%)相同的氨基酸序列。

[0129] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至二氧四氢嘧啶合酶亚基以构建二氧四氢嘧啶合酶纳米颗粒。球状二氧四氢嘧啶合酶纳米颗粒由单体亚基组成;一个此类二氧四氢嘧啶合酶亚基的序列实例如下列氨基酸序列所示:

[0130] MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELAR KEDIDAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALS AI EMANLFKSLR (SEQ ID NO:8) .

[0131] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至二氧四氢嘧啶合酶亚基,其包含与SEQ ID NO:8所示的氨基酸序列至少80%(例如至少85%、至少90%、至少95%,或至少97%)相同的氨基酸序列。

[0132] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至纳米隔室纳米颗粒亚基以构建纳米隔室纳米颗粒。球状纳米隔室纳米颗粒由单体亚基组成;一个此类纳米隔室蛋白亚基的序列实例如下列氨基酸序列所示:

[0133] MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLGEVEVLSDENEVVKWG LRKSLPLIELRATFTLDLWELDNLERGKPNVDLSLEETVRKVAEFEDVIFRGCEKSGVKLLSFEERKIECGST PKDLLEAIVRALSIFSKDIEGPYTLVINTDRWINFLKEEAGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGG DFKLILGQDLSIGYEDREKDAVRLFITETFTFQVNPALILLKF (SEQ ID NO:9) .

[0134] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至纳米隔室蛋白亚基,其包括与SEQ ID NO:9所示的氨基酸序列至少80%(例如至少85%、至少90%、至少95%或至少97%)相同的氨基酸序列。纳米隔室蛋白是细菌蛋白的保守家族,也称为类林可霉素(linocin)蛋白,其形成大型蛋白质组装体,作为最小隔室以包装酶。所述纳米隔室蛋白组装体由单体亚基组成,这些亚基是分子量约为30kDa的多肽。在产生后,所述单体亚基自组装成球状纳米隔室组装体,包括60个单体亚基,或在某些情况下,包括180个单体亚基。构建纳米隔室蛋白纳米颗粒的方法是本领域普通技术人员已知的,并且在本文中进一步描述(参见,例如,Sutter et al.,Nature Struct.and Mol.Biol.,15:939-947,2008,以全文引用的方式并入本文)。在具体实例中,所述纳米隔室蛋白多肽是细菌纳米隔室蛋白,例如海栖热袍菌(*Thermotoga maritime*)或强烈炽热球菌(*Pyrococcus furiosus*)或红串红球菌(*Rhodococcus erythropolis*)或黄色黏球菌(*Myxococcus xanthus*)纳米隔室蛋白。

[0135] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至硫氧化还原酶(SOR)亚基以构建重组SOR纳米颗粒。在一些实施方案中,所述SOR亚基可以包括如下所示的氨基酸序列:

[0136] MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLGEVEVLSDENEVVKWG LRKSLPLIELRATFTLDLWELDNLERGKPNVDLSLEETVRKVAEFEDVIFRGCEKSGVKLLSFEERKIECGST PKDLLEAIVRALSIFSKDIEGPYTLVINTDRWINFLKEEAGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGG DFKLILGQDLSIGYEDREKDAVRLFITETFTFQVNPALILLKF (SEQ ID NO:10) .

[0137] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域三聚体的原聚体可以连接至

SOR亚基,其包括与如SEQ ID NO:10所示的氨基酸序列至少80% (例如至少85%、至少90%、至少95%,或至少97%)相同的氨基酸序列。

[0138] SOR蛋白是微生物蛋白(例如来自嗜酸热古菌双能酸菌(*Acidianus ambivalens*)的蛋白),其形成24个亚基蛋白组装体。构建SOR纳米颗粒的方法是本领域普通技术人员已知的(参见,例如,Urich et al.,*Science*,311:996-1000,2006,以全文引用的方式并入本文)。用于制备SOR纳米颗粒的SOR蛋白氨基酸序列的实例在Urich et al.,*Science*,311:996-1000,2006中进行说明,其以全文引用的方式并入本文。

[0139] 出于产生的目的,在一些实施方案中,连接至纳米颗粒亚基的重组SARS-CoV-2 S胞外域可以包括在细胞加工过程中被切割的N-末端信号肽。例如,连接至蛋白质纳米颗粒亚基的重组SARS-CoV-2 S胞外域原聚体可在其N-末端包括信号肽,包括例如天然冠状病毒S信号肽。

[0140] 蛋白质纳米颗粒可以在适当的细胞(例如,HEK 293 Freestyle细胞)中表达,并且融合蛋白从自组装成纳米颗粒的细胞中分泌。可以使用已知技术纯化纳米颗粒,例如通过几种不同的色谱程序,例如Mono Q(阴离子交换),然后是尺寸排阻(SUPEROSE®6)色谱。

[0141] 若干实施方案包括铁蛋白、纳米隔室蛋白、SOR蛋白或二氧四氢喋啶合酶蛋白的单体亚基,或其能够引导单体亚基自组装成球状蛋白质的任何部分。任何已知铁蛋白、纳米隔室蛋白、SOR蛋白或二氧四氢喋啶合酶蛋白的单体亚基氨基酸序列可用于与重组SARS-CoV-2 S胞外域或其免疫原性片段产生融合蛋白,只要所述单体亚基能够自组装成在其表面展示重组SARS-CoV-2 S胞外域或其免疫原性片段的纳米颗粒。

[0142] 所述融合蛋白不需要包含铁蛋白、纳米隔室蛋白、SOR蛋白或二氧四氢喋啶合酶蛋白的单体亚基多肽的全长序列。可以使用单体亚基多肽的部分或区域,只要该部分包含指导单体亚基自组装成球状蛋白质的氨基酸序列。

[0143] III. 多核苷酸和表达

[0144] 还提供了编码任何所公开重组S胞外域三聚体的原聚体的多核苷酸。这些多核苷酸包括编码所述原聚体的DNA序列、cDNA序列和RNA序列,以及包括DNA、cDNA和RNA序列的载体,例如用于免疫的DNA或RNA载体。用于构建多种功能等同核酸(例如序列不同但编码相同蛋白质序列的核酸)的遗传密码,或编码包括所述核酸序列的缀合物或融合蛋白的遗传密码。

[0145] 编码SARS-CoV-2 S蛋白的示例性核酸序列如SEQ ID NO:11所示:

[0146] ATGTTTGTCTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAA
TTACCCCTGCATACACTAATTCTTTCACACGTGGTGTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTTTACA
TTCAACTCAGGACTTGTTCTTACCTTTCTTTCCAATGTTACTTGGTTCATGCTATACATGTCTCTGGGACCAATG
GTACTAAGAGGTTTGATAACCCTGTCTACCATTTAATGATGGTGTATTATTTGCTTCCACTGAGAAGTCTAACATA
ATAAGAGGCTGGATTTTTGGTACTACTTTAGATTGGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGT
TGTTATTAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGGTGTTTATTACCACAAAAACAACAAAAGTT
GGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATG
GACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAT
ATATTCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATT
TGCCAATAGGTATTAACATCACTAGGTTTCAAACCTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCT

TCTTCAGGTTGGACAGCTGGTGTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTTAAAATATAA
TGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAAATCCT
TCACTGTAGAAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTTCTAAT
ATTACAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTGTCATCTGTTTATGCTTGGAACAGGAAGAG
AATCAGCAACTGTGTTGCTGATTATTCTGTCTATATAATTCCGCATCATTTTCCACTTTTAAGTGTATGGAGTGT
CTCCTACTAAATTAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGA
CAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTAT
AGCTTGGAAATCTAACAATCTTGATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTA
ATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAAATGGTGTGAAAGGT
TTTAATTGTTACTTTCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTGGTTACCAACCATACAGAGTAGT
AGTACTTTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACA
AATGTGTCAATTTCAACTTCAATGGTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTC
CAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTAC
ACCATGTTCTTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTATCAGG
ATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTATTATTCTACAGGT
TCTAATGTTTTTCAAACACGTGCAGGCTGTTAATAGGGGCTGAACATGTCAACAACCTCATATGAGTGTGACATAACC
CATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGTCAAT
CCATCATTGCCTACACTATGTCACCTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATAACCACA
AATTTTACTATTAGTGTACCACAGAAATCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACAT
TTGTGGTGATTCAACTGAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAACCGTGCTTTAA
CTGGAATAGCTGTTGAACAAGACAAAAACCCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACCA
ATTAAGATTTTGGTGGTTTTAATTTTTTCAAAATATTACCAGATCCATCAAACCAAGCAAGAGGTCATTTATTGA
AGATCTACTTTTCAAACAAAGTGACACTTGACAGATGCTGGCTTCATCAAACAATATGGTATTGCCTTGGTGATATTG
CTGCTAGAGACCTCATTGTGCACAAAAGTTAACGGCCTTACTGTTTTGCCACCTTGTCTCACAGATGAAATGATT
GCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTGGACCTTGGTGCAGGTGCTGCATTACAAAT
ACCATTTGCTATGCAAAATGGCTTATAGGTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAAT
TGATTGCCAACCAATTTAATAGTGTATTGGCAAAAATCAAGACTCACTTTCTTCCACAGCAAGTGCACCTTGAAAA
CTTCAAGATGTGGTCAACCAAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACCTTAGCTCCAATTTTGGTGAAT
TTCAAGTGTTTTAAATGATATCCTTTCACGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAG
GCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTT
GCTGCTACTAAAATGTCAGAGTGTGACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTAT
GTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCA
CAACTGCTCCTGCCATTTGTCATGATGAAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACAC
TGGTTTGTAAACACAAAGGAATTTTTATGAACCACAAAATCATTACTACAGACAACACATTTGTGTCTGGTAACTGTGA
TGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAG
ATAAATATTTTAAAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAAC
ATTCAAAAAGAAATTGACCGCCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGG
AAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGAATGCCATAGTAATGG
TGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAA

TTTGATGAAGACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAA

[0147] 可以对上文所示示例性SARS-CoV-2 S原聚体的DNA序列进行修饰,从而引入本文公开实现融合前稳定的氨基酸取代和缺失。

[0148] 在若干实施方案中,所述核酸分子编码所述原聚体的前体,当在适当的细胞中表达时,该前体被加工成所公开SARS-CoV-2 S胞外域原聚体,该原聚体可以自组装成相应的重组SARS-CoV-2 S胞外域三聚体。例如,所述核酸分子可以编码重组SARS-CoV-2 S胞外域,其包括用于进入细胞分泌系统的N-末端信号序列,该序列在细胞中重组SARS-CoV-2 S胞外域的加工期间被蛋白水解切割。

[0149] 在若干实施方案中,所述核酸分子编码前体SARS-CoV-2 S多肽,当在适当的细胞中表达时,该前体被加工成包括S1和S2多肽的所公开重组SARS-CoV-2 S胞外域原聚体,其中所述重组SARS-CoV-2 S胞外域原聚体包括本文所述的稳定化修饰,并且可任选地连接至三聚体化域,例如T4次要纤维蛋白三聚体化域。

[0150] 示例性核酸可以通过克隆技术制备。适当的克隆和测序技术实例,以及足以指导技术人员通过许多克隆实践的说明是已知的(例如,参见Sambrook et al. (Molecular Cloning:A Laboratory Manual,4th ed,Cold Spring Harbor,New York,2012)和Ausubel et al. (In Current Protocols in Molecular Biology,John Wiley&Sons,New York,through supplement104,2013)。

[0151] 核酸也可以通过扩增方法制备。扩增方法包括聚合酶链式反应(PCR)、连接酶链式反应(LCR)、基于转录的扩增系统(TAS)、自主序列复制系统(3SR)。多种克隆方法、宿主细胞和体外扩增方法是本领域技术人员熟知的。

[0152] 编码所公开重组SARS-CoV-2 S胞外域原聚体的多核苷酸可以包括重组DNA,该重组DNA被整合到载体(例如表达载体)中,进入自主复制的质粒或病毒或原核细胞或真核细胞的基因组DNA,或作为独立于其他序列的单独分子(例如cDNA)存在。所述核苷酸可以是核糖核苷酸、脱氧核糖核苷酸或任一核苷酸的修饰形式。该术语包括单链和双链形式的DNA。

[0153] 编码所公开重组SARS-CoV-2 S胞外域原聚体的多核苷酸序列可以可操作地连接至表达控制序列。连接与编码序列可操作地连接的表达控制序列,从而在与所述表达控制序列相容的条件下实现编码序列的表达。所述表达控制序列包括但不限于适当的启动子、增强子、转录终止子、蛋白质编码基因前面的起始密码子(即ATG)、内含子剪接信号、为正确翻译mRNA该基因正确阅读框的维持序列,和终止密码子。

[0154] 通过将DNA转移到合适的宿主细胞中,可以在体外表达编码所公开重组S胞外域原聚体的DNA序列。所述细胞可以是原核细胞或真核细胞。该术语还包括受试者宿主细胞的任何后代细胞。应当理解的是,所有后代细胞可能与亲代细胞不同,因为在复制过程中可能会发生突变。稳定转移(意味着外源性DNA持续在宿主中存在)的方法是本领域已知的。

[0155] 宿主可以包括微生物、酵母、昆虫和哺乳动物有机体。在原核细胞中表达具有真核序列或病毒序列的DNA序列的方法是本领域熟知的。合适宿主细胞的非限制性实例包括细菌、古细菌(archea)、昆虫细胞、真菌(例如酵母)细胞、植物细胞和动物细胞(例如哺乳动物细胞,例如人细胞)。使用的示例性细胞包括大肠杆菌(*Escherichia coli*)、枯草芽孢杆菌(*Bacillus subtilis*)、酿酒酵母(*Saccharomyces cerevisiae*)、鼠伤寒沙门氏菌(*Salmonella typhimurium*)、SF9细胞、C129细胞、293细胞、脉孢菌(*Neurospora*)和永生化

哺乳动物骨髓和淋巴样细胞系。哺乳动物细胞培养繁殖技术是众所周知的(参见,例如, Helgason and Miller (Eds.), 2012, Basic Cell Culture Protocols (Methods in Molecular Biology), 4th Ed., Humana Press)。常用哺乳动物宿主细胞系实例是VERO和HeLa细胞, CHO细胞以及WI38、BHK和COS细胞系,但也可以使用以下细胞系,例如旨在提供更高表达、所需糖基化模式或其他特征的细胞。在一些实施方案中,宿主细胞包括HEK293细胞或其衍生物,例如GnTI^{-/-}细胞(ATCC®第CRL-3022号),或HEK-293F细胞。

[0156] 可以通过常规技术用重组DNA转化宿主细胞。在宿主是原核生物的情况下,例如但不限于大肠杆菌,可以由指数生长期后收获的细胞制备能够吸收DNA的感受态细胞,随后使用标准程序通过CaCl₂方法处理。或者,可以使用MgCl₂或RbCl。如果需要,也可以在形成宿主细胞的原生质体之后进行转化,或通过电穿孔进行转化。

[0157] 当宿主是真核生物时,可以使用诸如磷酸钙共沉淀的DNA转染方法、常规的机械方法(例如显微注射、电穿孔、插入包裹在脂质体中的质粒,或病毒载体)。真核细胞也可以与编码所公开抗原的多核苷酸序列和编码可选择表型的第二外源DNA分子(如单纯疱疹胸苷激酶基因)共转化。另一种方法是使用真核病毒载体(例如猿猴病毒40(SV40)或牛乳头瘤病毒),瞬时感染或转化真核细胞并表达蛋白质(参见例如, Viral Expression Vectors, Springer Press, Muzyczka ed., 2011)。适当的表达系统(例如质粒和载体),用于在细胞中产生蛋白质,包括高等真核细胞(如COS、CHO、HeLa和骨髓瘤细胞系)。

[0158] 在一个非限制性实例中,使用pVRC8400载体表达所公开免疫原(Barouch et al., J. Virol, 79, 8828-8834, 2005进行说明,其通过引用并入本文)。

[0159] 可以对编码所公开重组SARS-CoV-2 S胞外域原聚体的核酸进行修饰,而不会降低其生物活性。可以进行一些修饰以促进靶向分子的克隆、表达或整合到融合蛋白中。此类修饰是本领域技术人员熟知的,并且包括例如终止密码子、在氨基末端添加以提供起始位点的甲硫氨酸、位于任一末端以产生方便定位的限制性位点的额外氨基酸,或有助于进行纯化步骤的额外氨基酸(如聚组氨酸)。

[0160] 在一些实施方案中,所公开重组SARS-CoV-2 S胞外域原聚体可以在以下情况于细胞中表达:重组SARS-CoV-2 S胞外域原聚体可以自组装成从细胞分泌到细胞培养基中的三聚体。在此类实施方案中,每个重组SARS-CoV-2 S胞外域原聚体都含有使得蛋白质进入分泌系统的前导序列(信号肽),其中所述信号肽被切割,并且原聚体形成三聚体,然后被分泌到细胞培养基。可以将培养基离心并从上清液纯化重组SARS-CoV-2 S胞外域三聚体。

[0161] IV. 病毒载体

[0162] 编码所公开重组SARS-CoV-2 S胞外域三聚体的原聚体的核酸分子可以包括在病毒载体中,例如,用于在宿主细胞中表达免疫原,或用于如本文所公开的对受试者进行免疫。在一些实施方案中,将所述病毒载体作为初免-加强疫苗接种的一部分施用于受试者。在若干实施方案中,所述病毒载体包括在疫苗中,例如用于初免-加强疫苗接种的初次疫苗或加强疫苗。

[0163] 在若干实例中,所述病毒载体可以具有复制能力。例如,所述病毒载体可以在病毒基因组中具有不抑制病毒在宿主细胞中复制的突变。所述病毒载体也可以有条件地具有复制能力。在其他实例中,所述病毒载体在宿主细胞中有复制能力缺陷。

[0164] 已经构建了许多病毒载体,可用于表达所公开的抗原,包括多瘤病毒,即SV40

(Madzak et al., 1992, J. Gen. Virol., 73:1533-1536)、腺病毒 (Berkner, 1992, Cur. Top. Microbiol. Immunol., 158:39-6; Berliner et al., 1988, Bio Techniques, 6: 616-629; Gorziglia et al., 1992, J. Virol., 66:4407-4412; Quantin et al., 1992, Proc. Natl. Acad. Sci. USA, 89:2581-2584; Rosenfeld et al., 1992, Cell, 68:143-155; Wilkinson et al., 1992, Nucl. Acids Res., 20:2233-2239; Stratford-Perricaudet et al., 1990, Hum. Gene Ther., 1:241-256)、牛痘病毒 (Mackett et al., 1992, Biotechnology, 24:495-499)、腺相关病毒 (Muzyczka, 1992, Curr. Top. Microbiol. Immunol., 158:91-123; On et al., 1990, Gene, 89:279-282)、疱疹病毒 (包括HSV和EBV) (Margolskee, 1992, Curr. Top. Microbiol. Immunol., 158:67-90; Johnson et al., 1992, J. Virol., 66:2952-2965; Fink et al., 1992, Hum. Gene Ther. 3:11-19; Breakfield et al., 1987, Mol. Neurobiol., 1:337-371; Fresse et al., 1990, Biochem. Pharmacol., 40:2189-2199)、辛德比斯病毒 (H. Herweijer et al., 1995, Human Gene Therapy 6:1161-1167; 美国专利第5,091,309号和第5,2217,879号)、 α 病毒 (S. Schlesinger, 1993, Trends Biotechnol. 11:18-22; I. Frolov et al., 1996, Proc. Natl. Acad. Sci. USA 93:11371-11377) 和禽类逆转录病毒 (Brandyopadhyay et al., 1984, Mol. Cell Biol., 4:749-754; Petropoulos et al., 1992, J. Virol., 66:3391-3397)、鼠类逆转录病毒 (Miller, 1992, Curr. Top. Microbiol. Immunol., 158:1-24; Miller et al., 1985, Mol. Cell Biol., 5:431-437; Sorge et al., 1984, Mol. Cell Biol., 4:1730-1737; Mann et al., 1985, J. Virol., 54:401-407) 和人类逆转录病毒 (Page et al., 1990, J. Virol., 64:5370-5276; Buchschalcher et al., 1992, J. Virol., 66:2731-2739)。杆状病毒 (苜蓿银纹夜蛾多核型多角体病毒 (*Autographa californica multinuclear polyhydrosis virus*); AcMNPV) 载体也是本领域已知的, 并且可以从商业来源获得 (例如 PharMingen, San Diego, Calif.; Protein Sciences Corp., Meriden, Conn.; Stratagene, La Jolla, Calif.)。

[0165] 在若干实施方案中, 所述病毒载体可以包括表达所公开重组SARS-CoV-2 S胞外域三聚体的原聚体的腺病毒载体。不同来源、亚型或亚型混合的腺病毒可用作腺病毒载体的病毒基因组源。非人类腺病毒 (例如, 猿腺病毒、黑猩猩腺病毒、大猩猩腺病毒、禽腺病毒、犬腺病毒、羊腺病毒或牛腺病毒) 可用于生成腺病毒载体。例如, 猿腺病毒可以用作腺病毒载体的病毒基因组源。猿腺病毒可以属于血清型1、3、7、11、16、18、19、20、27、33、38、39、48、49、50或任何其他猿腺病毒血清型。可以通过使用本领域已知的任何合适缩写来指代猿腺病毒, 例如SV、SAdV、SAV或sAV。在一些实例中, 猿腺病毒载体是血清型3、7、11、16、18、19、20、27、33、38或39的猿腺病毒载体。在一个实例中, 使用黑猩猩血清型C Ad3载体 (参见, 例如, Peruzzi et al., Vaccine, 27:1293-1300, 2009)。人腺病毒可用作腺病毒载体的病毒基因组源。人腺病毒可以属于各种亚组或血清型。例如, 腺病毒可以属于A亚组 (例如血清型12、18和31)、B亚组 (例如血清型3、7、11、14、16、21、34、35和50)、C亚组 (例如, 血清型1、2、5和6)、亚组D (例如, 血清型8、9、10、13、15、17、19、20、22、23、24、25、26、27、28、29、30、32、33、36-39和42-48)、E亚组 (例如血清型4)、F亚组 (例如, 血清型40和41)、未分类的血清型 (例如, 血清型49和51), 或任何其他腺病毒血清型。本领域普通技术人员熟悉复制型和复制缺陷型腺病毒载体 (包括单复制和多重复制缺陷型腺病毒载体)。复制缺陷型腺病毒载体的实

例,包括多重复制缺陷型腺病毒载体,公开于美国专利第5,837,511号、第5,851,806号、第5,994,106号、第6,127,175号、第6,482,616号,和第7,195,896号,以及国际专利申请号WO 94/28152、WO 95/02697、WO 95/16772、WO 95/34671、WO 96/22378、WO97/12986、WO 97/21826和WO 03/022311。

[0166] V. 类病毒颗粒

[0167] 在一些实施方案中,提供了包括所公开重组SARS-CoV-2 S胞外域三聚体的类病毒颗粒(VLP)。通常,此类VLP包括重组SARS-CoV-2 S胞外域三聚体,即通过C-末端跨膜域膜锚定的膜,例如,三聚体中重组SARS-CoV-2 S胞外域原聚体各自可以连接至SARS-CoV-2 S蛋白的跨膜域和胞质尾。VLP缺乏病毒复制所需的病毒组分,因此代表了一种高度减毒、无法复制的病毒形式。然而,VLP可以展示一种多肽(例如,重组SARS-CoV-2 S胞外域三聚体),类似于在传染性病毒颗粒上表达的多肽,并且当施用于受试者时可以引发对SARS-CoV-2的免疫应答。类病毒颗粒及其生产方法是本领域普通技术人员已知和熟悉的,并且几种病毒的病毒蛋白已知能够形成VLP,包括人乳突病毒、HIV(Kang et al.,*Biol.Chem.*380:353-64(1999))、塞姆利基森林(Semliki-Forest)病毒(Notka et al.,*Biol.Chem.*380:341-52(1999))、人多瘤病毒(Goldmann et al.,*J.Virol.*73:4465-9(1999))、轮状病毒(Jiang et al.,*Vaccine* 17:1005-13(1999))、细小病毒(Casal,*Biotechnology and Applied Biochemistry*,Vol 29,Part 2,pp 141-150(1999))、犬细小病毒(Hurtado et al.,*J.Virol.*70:5422-9(1996))、戊型肝炎病毒(Li et al.,*J.Virol.*71:7207-13(1997))和新城疫(Newcastle disease)病毒。这种VLP的形成可以通过任何合适的技术来检测。本领域已知用于检测介质中VLP的合适技术实例包括例如电子显微镜技术、动态光散射法(DLS)、选择性色谱分离法(例如离子交换、疏水相互作用和/或尺寸排阻色谱分离VLP)和密度梯度离心法。

[0168] VI. 免疫原性组合物

[0169] 还提供了包含所公开免疫原(例如所公开重组SARS-CoV-2 S胞外域三聚体或编码所公开重组SARS-CoV-2 S胞外域三聚体的原聚体的核酸分子)和药学上可接受的递体的免疫原性组合物。此类药物组合物可以通过本领域普通技术人员已知的多种施用方式施用于受试者,例如肌肉内、皮内、皮下、静脉内、动脉内、关节内、腹膜内、鼻内、舌下、扁桃体、口咽或其他胃肠外和粘膜途径。制备可施用组合物的实际方法对于本领域技术人员来说是已知的或显而易见的,并且在诸如Remington's Pharmaceutical Sciences,19th Ed.,Mack Publishing Company,Easton,Pennsylvania,1995等出版物中进行了更详细的描述。

[0170] 因此,本文所述免疫原可以与药学上可接受的递体一起配制以有助于保持生物活性,同时还利于在可接受的温度范围内储存期间提高稳定性。可能的递体包括但不限于生理平衡培养基,磷酸盐缓冲盐溶液,水,乳液(例如油/水或水/油乳液),各种类型的润湿剂、防冻添加剂或稳定剂(如蛋白质、肽或水解产物(例如白蛋白、明胶)),糖(例如蔗糖、乳糖、山梨糖醇),氨基酸(例如谷氨酸钠)或其他保护剂。所得水溶液可以按原样包装使用或冻干。冻干制剂在单次剂量或多次剂量施用前与无菌溶液合并。

[0171] 配制的组合物,尤其是液体配制物,可含有防止储存期间降解或使其最小化的抑菌剂,包括但不限于有效浓度(通常 $\leq 1\%w/v$)的苯甲醇、苯酚、间甲酚、氯丁醇、对羟基苯甲酸甲酯和/或对羟基苯甲酸丙酯。某些患者可能禁用抑菌剂;因此,冻干配制物可以在含有

或不含有这种组分的溶液中复原。

[0172] 本公开的免疫原性组合物可以含有作为接近生理条件所需的药学上可接受的媒介物物质,例如pH调节剂和缓冲剂、渗透调节剂、润湿剂等,例如乙酸钠、乳酸钠、氯化钠、氯化钾、氯化钙、脱水山梨糖醇单月桂酸酯和油酸三乙醇胺。

[0173] 所述免疫原性组合物可以任选地包括佐剂以增强宿主的免疫应答。合适的佐剂是例如toll样受体激动剂、明矾、 $AlPO_4$ 、水凝胶、脂质-A及其衍生物或变体、油乳剂、皂苷、中性脂质体、含有疫苗和细胞因子的脂质体、非离子型嵌段共聚物和趋化因子。含有聚氧乙烯(POE)和聚氧丙烯(POP)的非离子型嵌段聚合物,例如POE-POP-POE嵌段共聚物、MPL™(3-O-脱酰基单磷酸脂质A;Corixa,Hamilton,IN)和IL-12(Genetics Institute,Cambridge,MA),以及本领域熟知的许多其他合适的佐剂,可以用作佐剂(Newman et al.,1998,Critical Reviews in Therapeutic Drug Carrier Systems 15:89-142)。这些佐剂的优势在于它们有助于以非特异性方式刺激免疫系统,从而增强对药物产品的免疫应答。

[0174] 在某些情况下,可能需要将所公开免疫原与诱导对其他药剂保护性反应的其他药物产品(例如疫苗)结合。例如,包括本文所述重组SARS-CoV-2 S胞外域三聚体的组合物可以与免疫实施咨询委员会(ACIP;cdc.gov/vaccines/acip/index.html)针对目标年龄组(例如,大约1到6个月大的婴儿)推荐的其他疫苗同时或依次施用,例如流感疫苗或水痘带状疱疹疫苗。因此,包括本文所述重组SARS-CoV-2 S胞外域三聚体的公开免疫原可以与针对例如乙型肝炎(HepB)、白喉、破伤风和百日咳(DTaP)、肺炎球菌(PCV)、流感嗜血杆菌(*Haemophilus influenzae*) b型(Hib)、脊髓灰质炎、流感和轮状病毒的疫苗同时或依次施用。

[0175] 在一些实施方案中,所述组合物可以作为无菌组合物提供。所述药物组合物通常含有有效量的所公开免疫原并且可以通过常规技术制备。通常,每剂免疫原性组合物中免疫原的量被选定为诱导免疫应答而没有明显不良副作用的量。在一些实施方案中,所述组合物可以以单位剂型提供,用于在受试者中诱导免疫应答。单位剂型含有用于施用至受试者合适的单一预选剂量,或两个或更多个预选单位剂量的合适标记或测量的多倍剂量,和/或用于施用单位剂量或其多倍剂量的计量机制。在其他实施方案中,所述组合物还包括佐剂。

[0176] VII. 诱导免疫应答的方法

[0177] 所公开免疫原(例如,重组SARS-CoV-2 S胞外域三聚体、编码所公开重组SARS-CoV-2 S胞外域三聚体的原聚体的核酸分子(如RNA分子)或载体,或包含所公开重组SARS-CoV-2 S胞外域三聚体的蛋白质纳米颗粒或类病毒颗粒)可以施用于受试者以诱导受试者对SARS-CoV-2 S蛋白的免疫应答。在特定实例中,所述受试者是人。免疫应答可以是保护性免疫应答,例如抑制随后SARS-CoV-2感染的反应。引发免疫应答也可用于治疗或抑制SARS-CoV-2感染和与SARS-CoV-2感染相关的疾病。

[0178] 可以选择有SARS-CoV-2感染或其感染风险(例如由于暴露或可能暴露于SARS-CoV-2)的受试者进行免疫。在施用所公开免疫原后,可以监测所述受试者感染或与SARS-CoV-2感染相关的症状。

[0179] 旨在用本公开免疫原和方法免疫的典型受试者包括人,以及非人类灵长类动物和其他动物。为了根据本公开的方法确定用于免疫的受试者,采用公认的筛选方法来确定与

目标或疑似疾病或病症相关的风险因素,或确定受试者现有疾病或病症的状态。这些筛选方法包括,例如,确定可能与目标或疑似疾病或病症相关的环境、家庭、职业和其他此类风险因素的常规检查,以及诊断方法,例如各种ELISA和其他检测和/或表征冠状病毒感染的免疫测定方法。这些和其他常规方法使临床医生能够来选择需要使用本公开的方法和药物组合物进行免疫的患者。

[0180] 施用所公开免疫原可以用于预防或治疗目的。当以预防的目的提供时,所述免疫原在任何症状出现之前提供,例如在感染之前。预防性施用免疫原用于预防或改善任何后续感染过程。当以治疗的目的提供时,所述免疫原在感染症状出现时或之后提供,例如,在出现SARS-CoV-2感染症状后或诊断为SARS-CoV-2感染后。因此,可以在预期暴露于SARS-CoV-2之前提供免疫原,以便在暴露或疑似暴露于SARS-CoV-2后,或在实际开始感染后,减轻感染和/或相关疾病症状预期的严重程度、持续时间或程度。

[0181] 本文所述免疫原及其免疫原性组合物以有效量提供给受试者,从而有效诱导或增强所述受试者(优选人)针对SARS-CoV-2 S蛋白的免疫应答。所公开免疫原的实际剂量将根据诸如疾病适应症和受试者的特定状态(例如所述受试者的年龄、体型、健康状况、症状严重程度、易感因素等)、施用的时间和途径、同时施用的其他药物或治疗方案,以及用于引发受试者所需活性或生物反应的组合物的特定药理学等因素而变化。可以调整剂量方案以实现最佳的预防或治疗反应。

[0182] 包含一种或更多种所公开免疫原的免疫原性组合物可用于协同(或初免-加强)疫苗接种方案或组合配制物。在某些实施方案中,新型组合免疫原性组合物和协同免疫方案采用单独的免疫原或配制物,每种都旨在引发抗病毒免疫应答,例如对SARS-CoV-2S蛋白的免疫应答。引发抗病毒免疫应答的单独免疫原性组合物可以组合成多价免疫原性组合物,在单个免疫步骤中施用于受试者,或者它们可以在协同(或初免-加强)免疫方案中单独施用(在单价免疫原性组合物中)。

[0183] 可以有几次加强,每次加强可以使用不同的所公开免疫原。在一些实例中,加强可能使用与另一次加强或初免相同的免疫原。初免和加强可以以单剂或多剂进行,例如可以在数天、数周或数月内向受试者施用两剂、三剂、四剂、五剂、六剂或更多剂。也可以进行多次加强,例如一到五次(例如,1、2、3、4或5次加强)或更多次。在一系列连续免疫中可以使用不同的剂量。例如,在初次免疫中使用相对较大的剂量,然后在加强免疫中使用相对较小的剂量。

[0184] 在一些实施方案中,加强可以在初免后约两周、约三至八周或约四周或在初免后约几个月进行。在一些实施方案中,加强可以在初免后约5个月、约6个月、约7个月、约8个月、约10个月、约12个月、约18个月、约24个月,或在初免后更长或更短时间施用。也可以在适当的时间点进行周期性的额外加强来增强受试者的“免疫记忆”。所选择的疫苗接种参数(例如配制物、剂量、方案等)的适当性,可以通过从受试者取等分试样的血清并在免疫程序中测定抗体滴度来确定。此外,可以监测受试者的临床状况以获得所需疗效,例如预防感染或改善疾病状态(例如降低病毒载量)。如果这种监测表明疫苗接种是次优的,则可以用额外剂量的免疫原性组合物对所述受试者进行加强,并且可以以预期增强免疫应答的方式修改疫苗接种参数。

[0185] 在一些实施方案中,初免-加强方法可以包括针对受试者的DNA-初免和蛋白质-加

强疫苗接种方案。所述方法可以包括两次或更多次施用核酸分子或蛋白质。

[0186] 对于蛋白质治疗剂,通常每个人剂量将包含1-1000 μ g蛋白质,例如约1 μ g至约100 μ g,例如约1 μ g至约50 μ g,例如约1 μ g、约2 μ g、约5 μ g、约10 μ g、约15 μ g、约20 μ g、约25 μ g、约30 μ g、约40 μ g或约50 μ g。

[0187] 在免疫原性组合物的使用量基于受试者群体(例如,婴儿或老年人)来选择。可以通过标准研究确定特定组合物的最佳量,这些研究涉及观察受试者的抗体滴度和其他反应。应当理解的是,在免疫原性组合物中,有效量的所公开免疫原(例如所公开重组SARS-CoV-2 S胞外域三聚体、病毒载体或核酸分子),可以包括例如在初次-加强施用方案中,通过施用单次剂量对引发免疫应答无效,但在施用多剂量时有有效的量。

[0188] 在施用本公开的免疫原后,受试者的免疫系统通常通过产生对包括在免疫原中的SARS-CoV-2 S胞外域三聚体特异的抗体来作出反应。这样的反应表明免疫有效剂量被递送至受试者。

[0189] 在一些实施方案中,受试者的抗体反应将在评估有效剂量/免疫方案的情况下确定。在大多数情况下,评估从受试者获得血清或血浆的抗体滴度就足够了。关于是否进行加强接种和/或改变个体施用治疗剂的量的决定可以至少部分地基于抗体滴度水平。抗体滴度水平可以基于例如免疫结合测定法,该测定法测量血清中与抗原结合的抗体浓度,所述抗原包括例如包括在所述免疫原中的重组SARS-CoV-2 S胞外域三聚体。

[0190] 无需完全消除或减少或预防SARS-CoV-2感染,这些方法就可有效。例如,与没有免疫原时的SARS-CoV-2感染相比,用一种或更多种所公开免疫原引发对SARS-CoV-2的免疫应答可以将SARS-CoV-2感染减少或抑制所需的程度,例如至少10%、至少20%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少98%,或甚至至少100%(消除或预防可检测到的被感染细胞)。在另外的实例中,可以通过所公开的方法减少或抑制SARS-CoV-2复制。无需完全消除SARS-CoV-2复制,所述方法就可有效。例如,与没有免疫应答时SARS-CoV-2复制相比,使用一种或更多种所公开免疫原引发的免疫应答可以将SARS-CoV-2复制减少到所需的程度,例如至少10%、至少20%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少98%或甚至至少100%(消除或防止可检测到的SARS-CoV-2复制)。

[0191] 在一些实施方案中,将所公开免疫原与佐剂同时施用于所述受试者。在其他实施方案中,所公开免疫原在施用佐剂之后并在足以诱导免疫应答的时间内施用给受试者。

[0192] 一种施用核酸的方法是用质粒DNA直接免疫,例如用哺乳动物表达质粒。通过核酸构建体进行免疫是本领域众所周知的,并且在例如美国专利第5,643,578号(描述了通过引入编码所需抗原的DNA以引发细胞介导反应或体液反应来使脊椎动物免疫的方法)、美国专利第5,593,972号和美国专利第5,817,637号(描述了将编码抗原的核酸序列可操作地连接至能够表达的调控序列)中进行了教导,并且。美国专利第5,880,103号说明了几种将编码免疫原性肽或其他抗原的核酸递送至有机体的方法。这些方法包括脂质体递送核酸(或递送合成肽本身),以及免疫刺激构建体或ISCOMSTM(在混合胆固醇和Quil ATM(皂苷)时自发形成的30-40nm大小带负电的笼状结构)。保护性免疫已在使用ISCOMSTM作为抗原递送媒介物的多种感染(包括弓形虫病和爱泼斯坦-巴尔(Epstein-Barr)病毒诱导的肿瘤)实验模型中产生(Mowat and Donachie, Immunol. Today 12:383, 1991)。已发现封装在ISCOMSTM中低至

1 μ g剂量的抗原产生I类介导的CTL反应(Takahashi et al., Nature 344:873,1990)。

[0193] 在一些实施方案中,质粒DNA疫苗用于在受试者中表达所公开的免疫原。例如,可以将编码所公开免疫原的核酸分子施用于受试者以诱导对包括在免疫原中的SARS-CoV-2 S蛋白产生免疫应答。在一些实施方案中,所述核酸分子可以包括在用于DNA免疫的质粒载体上,例如pVRC8400载体(说明见Barouch et al., J.Virol, 79, 8828-8834, 2005, 其通过引用并入本文)。

[0194] 在使用核酸进行免疫的另一种方法中,所公开重组SARS-CoV-2 S胞外域或重组SARS-CoV-2 S胞外域三聚体可以由减毒病毒宿主或载体或细菌载体表达。重组牛痘病毒、腺相关病毒(AAV)、疱疹病毒、逆转录病毒、巨细胞病毒或其他病毒载体可用于表达肽或蛋白质,从而引发CTL反应。例如,用于免疫方案的牛痘载体和方法在美国专利第4,722,848号中有所说明。BCG(卡介苗)提供了另一种用于表达肽的载体(参见Stover, Nature 351:456-460, 1991)。

[0195] 在一个实施方案中,将编码所公开重组SARS-CoV-2 S胞外域或SARS-CoV-2 S胞外域三聚体的核酸直接引入细胞中。例如,可以通过标准方法将核酸装载到金微球上,并通过Bio-Rad公司的HELIOS™基因枪等设备引入皮肤。核酸可以是“裸露的”,由受强启动子控制的质粒组成。通常,尽管DNA也可以直接注射到其他部位,但它被注射到肌肉中。注射剂量通常为约0.5 μ g/kg至约50mg/kg,通常为约0.005mg/kg至约5mg/kg(参见,例如,美国专利第5,589,466号)。

[0196] 例如,可以通过标准方法将核酸装载到金微球上,并通过Bio-Rad公司的HELIOS™基因枪等设备引入皮肤。核酸可以是“裸露的”,由受强启动子控制的质粒组成。通常,尽管DNA也可以直接注射到其他部位,但它被注射到肌肉中。注射剂量通常为约0.5 μ g/kg至约50mg/kg,通常为约0.005mg/kg至约5mg/kg(参见,例如,美国专利第5,589,466号)。

[0197] 在另一实施方案中,基于mRNA的免疫方案可用于将编码所公开重组SARS-CoV-2 S胞外域的核酸直接递送到细胞中。在一些实施方案中,基于mRNA的核酸疫苗可以为前述方法提供有效的替代疫苗。mRNA疫苗排除了DNA整合到宿主基因组中的安全问题,并且可以直接在宿主细胞的细胞质中翻译。此外,简单无细胞、体外合成RNA避免了与病毒载体相关的生产复杂化难题。可用于递送编码所公开重组SARS-CoV-2 S胞外域核酸的两种示例性RNA疫苗接种形式包括常规的非扩增性mRNA免疫(参见例如Petsch et al., “Protective efficacy of in vitro synthesized, specific mRNA vaccines against influenza A virus infection”, Nature biotechnology, 30(12):1210-6, 2012)和自扩增mRNA免疫(例如,参见Geall et al., “Nonviral delivery of self-amplifying RNA vaccines”, PNAS, 109(36):14604-14609, 2012; Magini et al., “Self-Amplifying mRNA Vaccines Expressing Multiple Conserved Influenza Antigens Confer Protection against Homologous and Heterosubtypic Viral Challenge”, PLoS One, 11(8):e0161193, 2016; 和 Brito et al., “Self-amplifying mRNA vaccines”, Adv Genet., 89:179-233, 2015)。

[0198] 向受试者施用有效量的一种或更多种所公开免疫原诱导所述受试者的中和免疫应答。为了评估中和活性,在所述受试者免疫后,可以在适当的时间点收集所述受试者的血清,冷冻并储存用于中和测试。测定中和活性的方法包括但不限于空斑减少中和(PRNT)测定、微量中和测定、基于流式细胞术的测定、单周期感染测定。在一些实施方案中,血清中和

活性可以使用SARS-CoV-2假病毒(类似于用于SARS-CoV的假病毒)来测定(Martin et al., Vaccine 26,6338,2008;Yang et al.,Nature 428,561,2004;Naldini et al.,PNAS 93,11382,1996;Yang et al.,PNAS 102,797,2005)。

[0199] 实施例

[0200] 提供以下实施例以说明某些实施方案的特定特征,但权利要求的范围不应限于例示的那些特征。

[0201] 实施例1

[0202] 融合前稳定的SARS-CoV-2 S蛋白

[0203] 这个实施例描述了重组SARS-CoV-2 S胞外域三聚体的形成,该三聚体稳定在融合前构象中。

[0204] 对SARS-CoV-2 S蛋白的序列进行了研究,以揭示其构造细节。由此,评估了使用氨基酸取代将S蛋白稳定在其融合前构象中的可能性。确定了两种突变对将SARS-CoV-2 S蛋白稳定在融合前构象中特别有效:K986P和V987P。具有K986P和V987P取代的SARS-CoV-2 S被称为“S-2P”。这两种脯氨酸取代位于SARS-CoV-2 S2的顶部(膜远端),在中央螺旋和HR1之间,并防止融合前到融合后的构象变化。图1A显示了SARS-CoV-2 S域的示意图。

[0205] 融合前的SARS-CoV-2 S蛋白(具有K986P和V987P)表达为具有C-末端T4次要纤维蛋白三聚体化域的可溶性蛋白(无TM和CT)。具有K986P和V987P取代的SARS-CoV-2 S三聚体的原聚体序列(包括信号肽和T4次要纤维蛋白三聚体化域)如SEQ ID NO:2所示。所表达蛋白质(其C-末端连接至三聚体化域)包括纯化标签和检测标签,其包括HRV3C切割位点、6xHis标签和Twin-Strep标签。在验证序列后,将表达质粒瞬时转染到FreeStyle293细胞中。六天后收获培养物,并使用蛋白质C-末端上的亲和标签,使上清液通过Ni²⁺-NTA和StrepTactin树脂,纯化分泌的蛋白质。然后将纯化的蛋白质通过尺寸排阻色谱柱以评估它们的寡聚状态(图1B)并分离对应于三聚体胞外域的单分散级分。然后通过SDS-PAGE评估蛋白质表达水平。

[0206] 当这些突变结合时,与相应的野生型蛋白相比,所述SARS-CoV-2 S蛋白的融合前稳定性初步表现为表达水平增加。

[0207] 双脯氨酸突变SARS-CoV-2 S变体的构象通过负染色电子显微镜技术评估(图1C和图1D)。具有双脯氨酸突变的S变体是同源的,并形成预期融合前形状三聚体。这些胞外域三聚体中的每个都被纯化为单峰,并形成典型融合前构象的三聚体。相反,具有天然序列的相应S蛋白形成混合构象的三聚体,其中一些三聚体为典型的融合前构象,而另一些则为典型细长的融合后构象。

[0208] 实施例2

[0209] 融合前稳定的SARS-CoV-2 S蛋白在动物模型中引发中和免疫应答

[0210] 此实施例描述了使用融合前稳定的SARS-CoV-2 S蛋白作为免疫原在动物模型中引发对SARS-CoV-2感染的中和免疫应答。

[0211] 如实施例1所述,制备可溶性融合前稳定的SARS-CoV-2 S蛋白。作为对照,没有K986P和V987P取代的SARS-CoV-2 S表达为具有C-末端T4次要纤维蛋白三聚体化域的可溶性蛋白(没有TM和CT)。未进行工程化取代的SARS-CoV2 S序列称为SARS-CoV-2 S WT。

[0212] 三个不同的小鼠品系BALB/cJ、C57BL/6J和B6C3F1/J小鼠在第0周和第3周用PBS,

0.01 μ g、0.1 μ g或1 μ g佐以Sigma佐剂系统(SAS)的可溶性SARS-CoV-2 S WT或可溶性SARS-CoV-2 S-2P进行免疫,并在初免后两周和加强后两周收集血清。通过ELISA评估SARS-CoV-2 S-2P免疫小鼠血清的SARS-CoV-2 S特异性IgG(图2A-2C)。评估SWT和S-2P免疫的BALB/cJ小鼠加强后血清中针对同型SARS-CoV-2假病毒的中和抗体(图2D)。结果表明,所述可溶性SARS-CoV-2 S-2P在初免和加强条件后引发剂量依赖性S特异性结合抗体,并且1 μ g的可溶性SARS-CoV-2 S-2P免疫原在动物模型中引发强中和抗体反应。

[0213] 评估了可溶性SARS-CoV-2 S WT和可溶性SARS-CoV-2 S-2P免疫保护小鼠免受病毒复制的能力。BALB/cJ小鼠在第0周和第3周用PBS,0.01 μ g、0.1 μ g或1 μ g佐以SAS的可溶性SARS-CoV-2 S WT或可溶性SARS-CoV-2 S-2P进行免疫。增强后4周,小鼠接受了小鼠适应的SARS-CoV-2的攻击(说明见:Dinnon et al. “a mouse adapted SARS-CoV-2 model for the evaluation of COVID-19 medical countermeasures”, BioRxiv., 2020.05.06.081497, 通过引用并入本文)。攻击后两天,在病毒载量达到峰值时,收集肺(图3A)和鼻甲(图3B)用于通过噬斑测定评估病毒载量。通过单因素方差分析和多重比较检验对小组进行比较。结果表明,0.1 μ g和1 μ g条件消除了上呼吸道和下呼吸道中的病毒复制;0.01 μ g S WT没有保护作用,表明这是S WT的突破剂量。由于与实验无关的死亡,0.01 μ g S-2P免疫的小鼠没有受到攻击(N/T)。

[0214] 实施例3

[0215] 含融合前稳定的SARS-CoV-2 S蛋白作为免疫原的蛋白质纳米颗粒

[0216] 本实施例说明了与蛋白质纳米颗粒支架缀合的融合前稳定的SARS-CoV-2 S胞外域三聚体及其作为免疫原的用途。

[0217] 聚糖修饰LuS-纳米颗粒支架和铁蛋白-纳米颗粒支架。为了构建纳米颗粒呈递抗原的可靠平台,选择嗜嗜热菌(*Aquifex aeolicus*) 二氧四氢嘧啶合酶(LuS)和幽门螺杆菌(*Helicobacter pylori*)铁蛋白作为纳米颗粒支架,以及称为SpyTag:SpyCatcher系统的异肽键缀合系统(Brune, K.D. et al. Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization. *Sci Rep* 6, 19234, 2016) 以在纳米颗粒表面展示抗原。所述SpyTag:SpyCatcher系统具有高度特异性和异肽键稳定性,已用于在纳米颗粒表面缀合抗原(图4A)(参见Zakeri, B. et al. “Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin”, *Proc Natl Acad Sci U S A* 109, E690-697, (2012); Brune, K.D. et al. Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization. *Sci Rep* 6, 19234, (2016))。LuS和铁蛋白已在临床研究中用作纳米颗粒免疫原的支架。铁蛋白和LuS的N-末端都暴露于纳米颗粒表面,因此可用于SpyTag或SpyCatcher附着(图4B)。LuS的C-末端也能在纳米颗粒表面附着,并可用于显示纯化标签。设计了表达SpyTag或SpyCatcher与LuS或铁蛋白的融合蛋白的哺乳动物表达构建体。所述构建体包括用于纯化的His标签和Strep标签,以及用于将表达的蛋白质分泌到培养基中的信号肽(图4B)。

[0218] 最初的构建体产生了用于纳米颗粒SpyTag融合蛋白或纳米颗粒SpyCatcher融合蛋白的低水平可溶性蛋白。为了提高蛋白质溶解度和表达,聚糖被添加到纳米颗粒表面。设计了一组具有SpyTag和SpyCatcher并添加了N-连接糖基化位点的LuS和铁蛋白构建体(表1

和表2)。对于LuS构建体,在位置71(PDB 1HQK编号)处添加了的糖基化位点。对于铁蛋白构建体,测试了两个潜在的糖基化位点(位置96和148)。添加N-连接糖基化位点促进了可溶性纳米颗粒在哺乳动物细胞培养上清液中的表达。其中三个构建体高产率地产生了组装良好的纳米颗粒,在N-末端具有N71和SpyTag的LuS(以下称为LuS-N71-SpyTag)、具有N96和SpyTag的铁蛋白,以及具有S148(N146上的聚糖)和SpyTag的铁蛋白(表1)。在两种铁蛋白构建体中,具有N96和SpyTag的铁蛋白具有更高的产率并被选择用于进一步研究(称为铁蛋白-N96-SpyTag)。尺寸排阻色谱(SEC)和电子显微镜技术(EM)分析表明,LuS-N71-SpyTag在溶液中形成了均匀的纳米颗粒群(图4C和图4E)。铁蛋白-N96-SpyTag样品主要由完整的纳米颗粒和一些少量未组装的物质组成(图4C和图4D)。负染色电子显微镜技术(EM)图像表明,两种纳米颗粒都以预期的尺寸很好地组装(图4E)。二维类平均值揭示了纳米颗粒更详细的结构特征,这与先前公布的两种纳米颗粒的结构一致。这些数据表明铁蛋白纳米颗粒和LuS纳米颗粒与SpyTag和糖基化位点添加相容。这些改变具有良好的耐受性,能够实现稳固的纳米颗粒组装。为了验证LuS-SpyTag纳米颗粒和铁蛋白-SpyTag纳米颗粒的糖基化,进行了PNGase F消化并使用SDS-PAGE评估聚糖切割(图4D)。在PNGase F存在的情况下,两种纳米颗粒都显示出条带位移,表明纳米颗粒上存在N连接的聚糖,并且通过酰胺酶消化作用将其去除。虽然LuS-N71-SpyTag中的聚糖切割不同,但在铁蛋白-N96-SpyTag中不太明显,这可能是由于铁蛋白-N96-SpyTag的不完全糖基化和SDS-PAGE上的多个铁蛋白条带。在一些研究中观察到铁蛋白在SDS-PAGE上呈现单条带,但在其他情况下呈现多条带,可能是由于C-末端的蛋白酶切割或不完全糖基化。然而,这些不同尺寸的铁蛋白分子正确组装成具有如SEC和EM所示预期大小的纳米颗粒(图4C、图4E)。

[0219] 表1. 具有SpyTag的LuS纳米颗粒和铁蛋白纳米颗粒。

构建体 ID	SpyTag	SpyCatcher	聚糖的位置	表达水平 (mg/L)
二氧四氢嘧啶合酶				
LuS-SpyTag 无聚糖	x		无	<0.1
LuS-N71-SpyTag*	x		71	3.0
LuS-C-SpyCatcher 无聚糖		x	无	<0.1
LuS-C71-SpyCatcher		x	71	<0.1
LuS-N-SpyCatcher 无聚糖		x	无	<0.1
[0220] LuS-N71-SpyCatcher		x	71	<0.1
铁蛋白				
铁蛋白-SpyTag 无聚糖	x		无	<0.1
铁蛋白-N96-SpyTag*	x		96	2.5
铁蛋白-S148-SpyTag*	x		146	1.0
铁蛋白-SpyCatcher 无聚糖		X	无	<0.1
铁蛋白-N96-SpyCatcher		X	96	<0.1
铁蛋白-S148-SpyCatcher		X	146	<0.1

[0221] 表2. 用于蛋白质表达的构建体氨基酸序列。

构建体名称	氨基酸序列
LuS-N71-SpyTag	MDSKGS SQKGSRL LLLLVSNLLLPQGVVGAHVMV DAYKPTKSGSAMQIYEGKLTAEGLRFGI VASRFNHALVDRLVEGAIDAI VRHGGREEDITLVRVPGSWEI PVAAGELARKENISAVIAIGVLI RGATPHFDYIASEVSKGLADLSLELRKPIITFGVITADTLEQA IERAGTKHGKNGWEAALS AIEMA NLFKSLRGGLVPRGSHHHHHHSAWSHPQFEK (SEQ ID NO: 12)
铁蛋白-N96-SpyTag	MDSKGS SQKGSRL LLLLVSNLLLPQGVVQHHHHHHHSAWSHPQFEKGGLVPRGGAHVMVDA YKPTKGGSGDPMLSKDI I KLLNEQVNKEMQSNLYMSMSWCYTHSLDGAGLFLFDHAAEEYEH AKKLI I FLNENNVVQLT SISAPEHKFEGLTQIFQKAYEHEQNI SESINNI VDHAIKSKDHATFN FLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRKS (SEQ ID NO: 13)
SARS-CoV-2 刺突- SpyCatcher*	MGWSC I I LFLVATATGVHSAPELLGGPSVFLFPKPKD TLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG QPREPQVYTLPPSRDELTKNQVSLYCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFL YSKLTVDKSRWQQGNV FSCSVMEALHNHYTQKSLSLSPGKGGGGSGGGGSGGGGSAPEL LGGPSVFLFPKPKD TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNST YRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQV LTKLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLT SKLTVDKSRWQQGNV FSCSVMH EALHNHYTQKSLSLSPGKGGGGSGGGGSGGLEVL FQGPQC VNLTRTQLPPAYTNSFTRGVYYPDKV FRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPFVLPFNDGVYFAST EKSNI IRGWIFG TTLDSKTQSL L I VNNATNVV I K VCEFFQFCNDPFLGVY YHKNNKSWMESEFRVYSSANNCTFEYVS QPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP IGINITR FQTL LALHRSYLT PGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCT LKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRKRSNCVADYSV LYNSASFSTFKCYGVSP TKLNDLCFTN VYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGC VIAWNSNNLDSKVGNYNYLYRFRKSNLKPFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGF QPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESNKKFLPF QQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHAD QLTPTWRVYSTGSNVFQTRAGCLIGA EHVNNSEYCDIPIGAGICASYQTQ TNSPGSASSVASQSI IAYTMSLGAENSVAYSNN SIAI PTNFTI SVTTEILPVSMTKTSVDCTMYICGDSTEC SNLLLQYG SFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLL FNKVTLADAGFIKQYGDCLGDI AARDL ICAQKFNGLTVLPPLLTDEMI AQYTSALLAGTITSGWT FGAGAALQIPFAMQMA YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSL S STASALGKLQDVV NQNAQALNTLVKQLSSNFGAIS SVLNDILSR LDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAE IRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAIC HDGKAHFPREGVFSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDS FKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGGSY IPEAPRDGQAYVRKDGEWVLLSTFLGRSGGGLVPQQSGDSATHIKFSKRDEDEGKELAGATMELRD SSGKTI STWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQQQVTVNGKATKGD AHI (SEQ ID NO: 14)

[0222]

[0223] *此氨基酸序列包括单链Fc纯化标签(参见参考文献38)。

[0224] 通过SpyTag:SpyCatcher将SARS-CoV-2刺突三聚体与LuS纳米颗粒缀合,使刺突三聚体均匀地展示在纳米颗粒表面上。与C-末端SpyCatcher(SEQ ID NO:14)融合的SARS-CoV-2刺突被表达、纯化并缀合至纯化的LuS-N71-SpyTag纳米颗粒(图5A-图5C)。对于该构建体,融合前SARS-CoV-2 S胞外域三聚体含有原聚体,所述原聚体包括用于去除S1/S2切割位点的GSAS取代、用于融合前稳定的K986P和V987P取代以及C-末端T4噬菌体次要纤维蛋白三聚体化域。就纯化而言,所述蛋白质包括单链Fc标签(有关单链Fc标签的说明,请参见Zhou,T.et al.Structure-Based Design with Tag-Based Purification and In-Process Biotinylation Enable Streamlined Development of SARS-CoV-2spike Molecular Probes.bioRxiv,2020.2006.2022.166033)。将缀合混合物加载到SEC色谱柱上以从未缀合的LuS-N71-SpyTag和SARS-CoV-2刺突-SpyCatcher中纯化缀合的纳米颗粒产物LuS-N71-SpyLinked-CoV刺突(图5B)。SDS-PAGE分析显示缀合产物具有预期的分子量,缀合后未观察到未缀合的刺突-SpyCatcher(图5C)。

[0225] 为了估计缀合效率,测量缀合纳米颗粒产物的SDS-PAGE凝胶图像上每个条带强度(图5C)作为每种组分质量的替代。考虑到各组分的分子量,计算出样品中各组分与总蛋白质的摩尔比。基于此,估计全部LuS纳米颗粒亚基的91%与刺突三聚体缀合。负染色EM显示LuS-N71-SpyLinked-CoV-2刺突纳米颗粒表现出预期的尺寸,刺突三聚体显示在LuS纳米颗粒表面(图5D)。SPR测量显示LuS-N71-SpyLinked-SARS-CoV-2刺突与CR3022结合(ter Meulen, J. et al. Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. PLoS Med 3, e237, 2006; Yuan, M. et al. A highly conserved cryptic epitope in the receptor-binding domains of SARS-CoV-2 and SARS-CoV. Science, 2020), CR3022是一种靶向受体结合域(RBD)的抗体,表明使用LuS-SpyTag:SpyCatcher系统成功地进行了刺突三聚体的纳米颗粒呈递(图5E)。

[0226] SpyLinked-纳米颗粒展示增加了SARS-CoV-2刺突引发中和抗体的潜力。为了评估免疫原性,在第0周和第3周给小鼠注射LuS-N71-SpyLinked-CoV-2刺突纳米颗粒或可溶性刺突三聚体(通过实施例1中的2P突变稳定的三聚体)或模拟(LuS-N71-SpyTag)纳米颗粒(图6A)。每次免疫后两周收集血清样品。第一次免疫后,在0.08 μ g的最低免疫原剂量下,刺突纳米颗粒免疫血清表现出的抗SARS-CoV-2刺突ELISA几何平均滴度为5,116,而10份三聚体刺突免疫血清中只有1份表现出可测量的滴度(图6B);在第二次免疫后,刺突纳米颗粒免疫血清的滴度显著增加,大约增加了25倍。在第2周和第5周,用更高剂量的刺突纳米颗粒(0.4和2.0 μ g)进行免疫的滴度增加得更多。相比之下,刺突三聚体剂量的增加更显著地提高了ELISA滴度,其中两只小鼠的2.0 μ g刺突三聚体免疫血清达到测定检测上限,滴度为1,638,400(图6B)。

[0227] 此外,假病毒中和分析显示,LuS-N71-SpyLinked-CoV-2刺突纳米颗粒在0.08、0.4和2 μ g的免疫剂量下引发有效的中和反应,几何平均ID50滴度分别为412、1820和1501(图6C)。相比之下,两种剂量的可溶性三聚体刺突在0.4和2 μ g条件下引发中和滴度,几何平均ID50分别为49和315,在0.08 μ g剂量下没有可测量的中和反应。实质上,0.08 μ g的刺突纳米颗粒引发了更高的中和反应,但在统计学上与2 μ g三聚体刺突不能区分。这表明,按重量计,刺突纳米颗粒的免疫原性比刺突本身高~25倍,表明存在显著的“剂量保留”效应。总体而言,与三聚体形式相比,在LuS纳米颗粒表面呈递SARS-CoV-2刺突显著提高了其免疫原性,并且引发有效中和反应需要的免疫原剂量较低。

[0228] 显然,所述方法或组合物的精确细节可以在不背离所述实施方案精神的情况下进行改变或修改。我们要求所有落入权利要求范围和精神内的修改和变化。

序列表

<110> 美国政府(由卫生和人类服务部的部长所代表);德克萨斯州大学系统董事会;
达特茅斯大学董事会

<120> SARS-CoV-2疫苗

<130> 4239-104089-02

<150> US62/972886

<151> 2020-02-11

<160> 14

<170> PatentIn 版本 3.5

<210> 1

<211> 1273

<212> PRT

<213> 冠状病毒

<400> 1

```

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
 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 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
 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
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
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
Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile		
705	710	715
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
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
 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
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> 2		
<211> 1235		
<212> PRT		
<213> 人工序列		
<220>		
<223> 变异SARS-CoV-2刺突蛋白		
<400> 2		
Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val		
1	5	10
Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe		
	20	25
Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu		
	35	40
His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp		
	50	55
Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp		
65	70	75
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
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
 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
 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
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 Pro Pro 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
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 Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp		
1205	1210	1215
Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser		
1220	1225	1230
Thr Phe		
1235		
<210> 3		
<211> 1273		
<212> PRT		
<213> 人工序列		
<220>		
<223> 变异SARS-CoV-2刺突蛋白		
<400> 3		
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
 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
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
 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
 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
 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 Pro	Pro 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
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> 4
 <211> 1233
 <212> PRT
 <213> 人工序列
 <220>
 <223> 变异SARS-CoV-2刺突蛋白
 <400> 4
 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

530	535	540
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu		
545	550	555
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val		
	565	570
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe		
	580	585
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val		
	595	600
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
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val		
	645	650
Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala		
	660	665
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Gly Gly Ser Val Ala Ser Gln		
	675	680
Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala		
690	695	700
Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val		
705	710	715
Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys		
	725	730
Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu		
	740	745
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile		
	755	760
Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys		
	770	775
Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe		
785	790	795
Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile		
	805	810
Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile		
	820	825
Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile		
	835	840
		845

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

	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
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
 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 Gly Gly Ser Val Ala Ser Gln
 675 680 685
 Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala
 690 695 700
 Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val
 705 710 715 720
 Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys
 725 730 735
 Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu
 740 745 750
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile
 755 760 765
 Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys
 770 775 780
 Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe

785	790	795	800
Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile			
	805	810	815
Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile			
	820	825	830
Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile			
	835	840	845
Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr			
	850	855	860
Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile			
865	870	875	880
Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe			
	885	890	895
Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn			
	900	905	910
Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala			
	915	920	925
Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly			
	930	935	940
Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu			
945	950	955	960
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn			
	965	970	975
Asp Ile Leu Ser Arg Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp			
	980	985	990
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln			
	995	1000	1005
Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala			
	1010	1015	1020
Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val			
	1025	1030	1035
Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser			
	1040	1045	1050
Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala			
	1055	1060	1065
Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly			
	1070	1075	1080
Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr			
	1085	1090	1095

His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile
 1100 1105 1110
 Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile
 1115 1120 1125
 Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu
 1130 1135 1140
 Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr
 1145 1150 1155
 Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser
 1160 1165 1170
 Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala
 1175 1180 1185
 Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys
 1190 1195 1200
 Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe
 1205 1210 1215
 Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu Cys
 1220 1225 1230
 Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys
 1235 1240 1245
 Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu
 1250 1255 1260
 Lys Gly Val Lys Leu His Tyr Thr
 1265 1270

<210> 6

<211> 26

<212> PRT

<213> 人工序列

<220>

<223> T4次要纤维蛋白三聚体化域

<400> 6

Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys
 1 5 10 15
 Asp Gly Glu Trp Val Leu Leu Ser Thr Phe
 20 25

<210> 7

<211> 173

<212> PRT

<213> 人工序列

<220>

<223> 蛋白质纳米颗粒亚基

<400> 7

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

<210> 8

<211> 154

<212> PRT

<213> 人工序列

<220>

<223> 蛋白质纳米颗粒亚基

<400> 8

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

130	135	140			
Pro Lys Asp Leu Leu Glu Ala Ile Val Arg Ala Leu Ser Ile Phe Ser					
145	150	155	160		
Lys Asp Gly Ile Glu Gly Pro Tyr Thr Leu Val Ile Asn Thr Asp Arg					
	165	170	175		
Trp Ile Asn Phe Leu Lys Glu Glu Ala Gly His Tyr Pro Leu Glu Lys					
	180	185	190		
Arg Val Glu Glu Cys Leu Arg Gly Gly Lys Ile Ile Thr Thr Pro Arg					
	195	200	205		
Ile Glu Asp Ala Leu Val Val Ser Glu Arg Gly Gly Asp Phe Lys Leu					
	210	215	220		
Ile Leu Gly Gln Asp Leu Ser Ile Gly Tyr Glu Asp Arg Glu Lys Asp					
225	230	235	240		
Ala Val Arg Leu Phe Ile Thr Glu Thr Phe Thr Phe Gln Val Val Asn					
	245	250	255		
Pro Glu Ala Leu Ile Leu Leu Lys Phe					
	260	265			

<210> 11

<211> 3822

<212> DNA

<213> 冠状病毒

<400> 11

```

atgtttgttt ttcttgttt attgccacta gtctctagtc agtgtgtaa tcttacaacc 60
agaactcaat taccccctgc atacactaat tctttcacac gtggtgttta ttaccctgac 120
aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc tttcttttcc 180
aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat 240
aaccctgtcc taccatttaa tgatggtgtt tattttgctt cactgagaa gtctaacata 300
ataagaggct ggatttttgg tactacttta gattcgaaga cccagtcctt acttattggt 360
aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420
ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480
tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggacctgaa 540
ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat tgatggttat 600
tttaaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660
tcggtcttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720
ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg ttggacagct 780
gggtctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840
gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaag 900
tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960
caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc ttttggtgaa 1020

```

gtttttaacg ccaccagatt tgcactctgtt tatgcttgga acaggaagag aatcagcaac 1080
tgtgttgctg attattctgt cctatataat tccgcatcat tttccacttt taagtgttat 1140
ggagtgtctc ctactaaatt aatgatctc tgctttacta atgtctatgc agattcattt 1200
gtaattagag gtgatgaagt cagacaaatc gctccagggc aaactggaaa gattgctgat 1260
tataattata aattaccaga tgattttaca ggctgcgta tagcttgga ttctaacaat 1320
cttgattcta aggttgggtg taattataat tacctgtata gattgttttag gaagtctaata 1380
ctcaaaccctt ttgagagaga tatttcaact gaaatctatc aggccggtag cacaccttgt 1440
aatgggtgtg aaggttttaa ttgttacttt cttttacaat catatggttt ccaaccact 1500
aatgggtgtg gttaccaacc atacagagta gtagtacttt cttttgaaact tctacatgca 1560
ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa atgtgtcaat 1620
ttcaacttca atggtttaaaggcacaggt gttcttactg agtctaacaa aaagtttctg 1680
cctttccaac aatttggcag agacattgct gacactactg atgctgtccg tgatccacag 1740
acacttgaga ttcttgacat tacaccatgt tcttttgggt gtgtcagtggt tataacacca 1800
ggaacaaata cttctaacca ggttgcgtgtt ctttatcagg atgttaactg cacagaagtc 1860
cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc tacagttct 1920
aatgtttttc aaacacgtgc aggctgttta ataggggctg aacatgtcaa caactcatat 1980
gagtgtgaca taccatttgg tgcaggtata tgcgctagtt atcagactca gactaattct 2040
cctcggcggg cacgtagtgt agctagtcaa tccatcattg cctacactat gtcacttgggt 2100
gcagaaaatt cagttgctta ctctaataac tctattgcca taccacaaa ttttactatt 2160
agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga ttgtacaatg 2220
tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg cagtttttgt 2280
acacaattaa accgtgcttt aactggaata gctgttgaac aagacaaaa cacccaagaa 2340
gtttttgcac aagtcaacaa aatttacaac acaccacaa ttaaagattt tgggtggttt 2400
aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt tattgaagat 2460
ctacttttca acaaagtac acttgcagat gctggcttca tcaacaata tgggtgattgc 2520
cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg ctttactgtt 2580
ttgccacctt tgctcacaga tgaatgatt gctcaataca cttctgact gttagcgggt 2640
acaatcactt ctggttgac ctttgggtgca ggtgctgcat tacaataacc atttgctatg 2700
caaatggctt ataggtttaa tggatttga gttacacaga atgttctcta tgagaacca 2760
aaattgattg ccaaccaatt taatagtgtc attggcaaaa ttcaagactc actttcttcc 2820
acagcaagtg cacttgaaa acttcaagat gtggtaacc aaaatgcaca agctttaaac 2880
acgcttgta aacaacttag ctccaatttt ggtgcaattt caagtgtttt aatgatatac 2940
ctttcacgtc ttgacaaaagt tgaggctgaa gtgcaattg ataggttgat cacaggcaga 3000
cttcaaagt ttgcagacata tgtgactcaa caattaatta gagctgcaga aatcagagct 3060
tctgctaate ttgctgctac taaaatgtca gagtgtgtac ttggacaatc aaaaagagtt 3120
gatttttgtg gaaagggtta tcatcttatg tcttccctc agtcagcacc tcatggtgta 3180
gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac tgctctgccc 3240
atgtgtcatg atggaaaagc aacttttct cgtgaaggtg tctttgtttc aatggcaca 3300
cactggtttg taacacaaaag gaatttttat gaaccacaaa tcattactac agacaacaca 3360

tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt ttatgatcct 3420
 ttgcaacctg aattagactc attcaaggag gagttagata aatattttaa gaatcataca 3480
 tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt aaacattcaa 3540
 aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct catcgatctc 3600
 caagaacttg gaaagtatga gcagtatata aatggccat ggtacatttg gctaggtttt 3660
 atagctggct tgattgccat agtaatgggtg acaattatgc tttgctgtat gaccagttgc 3720
 tgtagttgtc tcaagggctg ttgttctgtg ggatcctgct gcaaatttga tgaagacgac 3780
 tctgagccag tgctcaaagg agtcaaatta cattacacat aa 3822

<210> 12

<211> 226

<212> PRT

<213> 人工序列

<220>

<223> 蛋白质纳米颗粒亚基

<400> 12

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

Glu Met Ala Asn Leu Phe Lys Ser Leu Arg Gly Gly Leu Val Pro Arg
 195 200 205
 Gly Ser His His His His His His Ser Ala Trp Ser His Pro Gln Phe
 210 215 220
 Glu Lys
 225
 <210> 13
 <211> 244
 <212> PRT
 <213> 人工序列
 <220>
 <223> 蛋白质纳米颗粒亚基
 <400> 13
 Met Asp Ser Lys Gly Ser Ser Gln Lys Gly Ser Arg Leu Leu Leu Leu
 1 5 10 15
 Leu Val Val Ser Asn Leu Leu Leu Pro Gln Gly Val Val Gly Gln His
 20 25 30
 His His His His His His His Ser Ala Trp Ser His Pro Gln Phe Glu
 35 40 45
 Lys Gly Gly Leu Val Pro Arg Gly Gly Ala His Ile Val Met Val Asp
 50 55 60
 Ala Tyr Lys Pro Thr Lys Gly Gly Gly Ser Gly Asp Pro Met Leu Ser
 65 70 75 80
 Lys Asp Ile Ile Lys Leu Leu Asn Glu Gln Val Asn Lys Glu Met Gln
 85 90 95
 Ser Ser Asn Leu Tyr Met Ser Met Ser Ser Trp Cys Tyr Thr His Ser
 100 105 110
 Leu Asp Gly Ala Gly Leu Phe Leu Phe Asp His Ala Ala Glu Glu Tyr
 115 120 125
 Glu His Ala Lys Lys Leu Ile Ile Phe Leu Asn Glu Asn Asn Val Pro
 130 135 140
 Val Gln Leu Thr Ser Ile Ser Ala Pro Glu His Lys Phe Glu Gly Leu
 145 150 155 160
 Thr Gln Ile Phe Gln Lys Ala Tyr Glu His Glu Gln Asn Ile Ser Glu
 165 170 175
 Ser Ile Asn Asn Ile Val Asp His Ala Ile Lys Ser Lys Asp His Ala
 180 185 190
 Thr Phe Asn Phe Leu Gln Trp Tyr Val Ala Glu Gln His Glu Glu Glu
 195 200 205

Val Leu Phe Lys Asp Ile Leu Asp Lys Ile Glu Leu Ile Gly Asn Glu
 210 215 220
 Asn His Gly Leu Tyr Leu Ala Asp Gln Tyr Val Lys Gly Ile Ala Lys
 225 230 235 240
 Ser Arg Lys Ser
 <210> 14
 <211> 1820
 <212> PRT
 <213> 人工序列
 <220>
 <223> 变异SARS-CoV-2刺突蛋白
 <400> 14
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1 5 10 15
 Val His Ser Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe
 20 25 30
 Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val
 35 40 45
 Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe
 50 55 60
 Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro
 65 70 75 80
 Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr
 85 90 95
 Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val
 100 105 110
 Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala
 115 120 125
 Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg
 130 135 140
 Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Tyr Cys Leu Val Lys Gly
 145 150 155 160
 Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro
 165 170 175
 Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser
 180 185 190
 Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln
 195 200 205
 Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His

210	215	220
Tyr Thr Gln Lys Ser Leu	Ser Leu Ser Pro Gly Lys Gly Gly Gly	
225	230	235
Ser Gly Gly Gly Gly Ser	Gly Gly Gly Gly Gly Ser Gly Gly Gly Ser	240
	245	250
Ala Pro Glu Leu Leu Gly	Gly Pro Ser Val Phe Leu Phe Pro Pro Lys	255
	260	265
Pro Lys Asp Thr Leu Met	Ile Ser Arg Thr Pro Glu Val Thr Cys Val	270
	275	280
Val Val Asp Val Ser His	Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr	285
	290	300
Val Asp Gly Val Glu Val	His Asn Ala Lys Thr Lys Pro Arg Glu Glu	310
305	310	315
Gln Tyr Asn Ser Thr Tyr	Arg Val Val Ser Val Leu Thr Val Leu His	320
	325	330
Gln Asp Trp Leu Asn Gly	Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys	335
	340	345
Ala Leu Pro Ala Pro Ile	Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln	350
	355	360
Pro Arg Glu Pro Gln Val	Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu	365
	370	375
Thr Lys Asn Gln Val Ser	Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro	380
385	390	395
Ser Asp Ile Ala Val Glu	Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn	400
	405	410
Tyr Lys Thr Thr Pro Pro	Val Leu Asp Ser Asp Gly Ser Phe Phe Leu	415
	420	425
Thr Ser Lys Leu Thr Val	Asp Lys Ser Arg Trp Gln Gln Gly Asn Val	430
	435	440
Phe Ser Cys Ser Val Met	His Glu Ala Leu His Asn His Tyr Thr Gln	445
	450	455
Lys Ser Leu Ser Leu Ser	Pro Gly Lys Gly Gly Ser Gly Gly Gly Gly	460
465	470	475
Ser Gly Gly Leu Glu Val	Leu Phe Gln Gly Pro Gln Cys Val Asn Leu	480
	485	490
Thr Thr Arg Thr Gln Leu	Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg	495
	500	505
Gly Val Tyr Tyr Pro Asp	Lys Val Phe Arg Ser Ser Val Leu His Ser	510
	515	520
		525

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

835	840	845
Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys		
850	855	860
Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile		
865	870	875
Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile		
885	890	895
Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile		
900	905	910
Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn		
915	920	925
Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg		
930	935	940
Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly		
945	950	955
Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln		
965	970	975
Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser		
980	985	990
Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser		
995	1000	1005
Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn Gly		
1010	1015	1020
Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu		
1025	1030	1035
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala		
1040	1045	1050
Val Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys		
1055	1060	1065
Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser		
1070	1075	1080
Asn Gln Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val		
1085	1090	1095
Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val		
1100	1105	1110
Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu		
1115	1120	1125
Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro		
1130	1135	1140

Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser	1145	1150	1155
Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr	1160	1165	1170
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn	1175	1180	1185
Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu	1190	1195	1200
Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met	1205	1210	1215
Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln	1220	1225	1230
Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile	1235	1240	1245
Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln Val	1250	1255	1260
Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe	1265	1270	1275
Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg	1280	1285	1290
Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp	1295	1300	1305
Ala Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala	1310	1315	1320
Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val	1325	1330	1335
Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser	1340	1345	1350
Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala	1355	1360	1365
Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg	1370	1375	1380
Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln	1385	1390	1395
Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln	1400	1405	1410
Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp	1415	1420	1425
Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln			

1430	1435	1440
Leu Ser Ser Asn Phe Gly	Ala Ile Ser Ser Val	Leu Asn Asp Ile
1445	1450	1455
Leu Ser Arg Leu Asp Pro	Pro Glu Ala Glu Val	Gln Ile Asp Arg
1460	1465	1470
Leu Ile Thr Gly Arg Leu	Gln Ser Leu Gln Thr	Tyr Val Thr Gln
1475	1480	1485
Gln Leu Ile Arg Ala Ala	Glu Ile Arg Ala Ser	Ala Asn Leu Ala
1490	1495	1500
Ala Thr Lys Met Ser Glu	Cys Val Leu Gly Gln	Ser Lys Arg Val
1505	1510	1515
Asp Phe Cys Gly Lys Gly	Tyr His Leu Met Ser	Phe Pro Gln Ser
1520	1525	1530
Ala Pro His Gly Val Val	Phe Leu His Val Thr	Tyr Val Pro Ala
1535	1540	1545
Gln Glu Lys Asn Phe Thr	Thr Ala Pro Ala Ile	Cys His Asp Gly
1550	1555	1560
Lys Ala His Phe Pro Arg	Glu Gly Val Phe Val	Ser Asn Gly Thr
1565	1570	1575
His Trp Phe Val Thr Gln	Arg Asn Phe Tyr Glu	Pro Gln Ile Ile
1580	1585	1590
Thr Thr Asp Asn Thr Phe	Val Ser Gly Asn Cys	Asp Val Val Ile
1595	1600	1605
Gly Ile Val Asn Asn Thr	Val Tyr Asp Pro Leu	Gln Pro Glu Leu
1610	1615	1620
Asp Ser Phe Lys Glu Glu	Leu Asp Lys Tyr Phe	Lys Asn His Thr
1625	1630	1635
Ser Pro Asp Val Asp Leu	Gly Asp Ile Ser Gly	Ile Asn Ala Ser
1640	1645	1650
Val Val Asn Ile Gln Lys	Glu Ile Asp Arg Leu	Asn Glu Val Ala
1655	1660	1665
Lys Asn Leu Asn Glu Ser	Leu Ile Asp Leu Gln	Glu Leu Gly Lys
1670	1675	1680
Tyr Glu Gln Gly Ser Gly	Tyr Ile Pro Glu Ala	Pro Arg Asp Gly
1685	1690	1695
Gln Ala Tyr Val Arg Lys	Asp Gly Glu Trp Val	Leu Leu Ser Thr
1700	1705	1710
Phe Leu Gly Arg Ser Gly	Gly Gly Leu Val Pro	Gln Gln Ser Gly
1715	1720	1725

Asp Ser Ala Thr His Ile Lys Phe Ser Lys Arg Asp Glu Asp Gly		
1730	1735	1740
Lys Glu Leu Ala Gly Ala Thr Met Glu Leu Arg Asp Ser Ser Gly		
1745	1750	1755
Lys Thr Ile Ser Thr Trp Ile Ser Asp Gly Gln Val Lys Asp Phe		
1760	1765	1770
Tyr Leu Tyr Pro Gly Lys Tyr Thr Phe Val Glu Thr Ala Ala Pro		
1775	1780	1785
Asp Gly Tyr Glu Val Ala Thr Ala Ile Thr Phe Thr Val Asn Glu		
1790	1795	1800
Gln Gly Gln Val Thr Val Asn Gly Lys Ala Thr Lys Gly Asp Ala		
1805	1810	1815
His Ile		
1820		

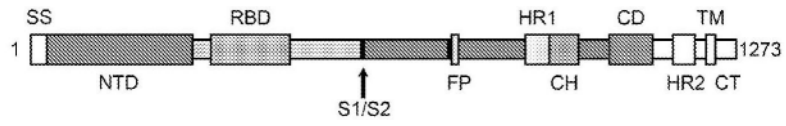


图1A

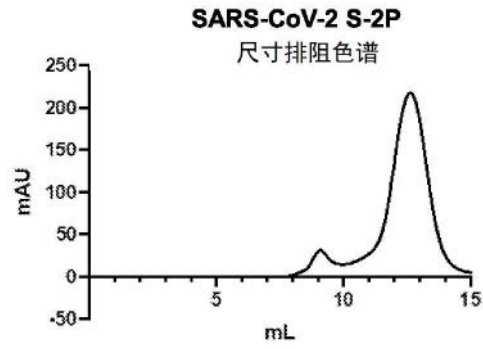


图1B

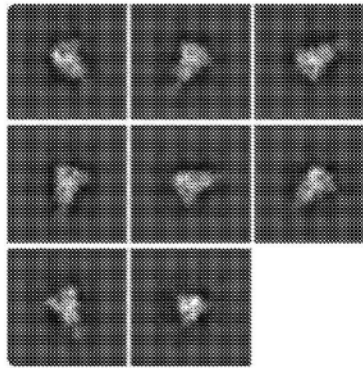


图1C

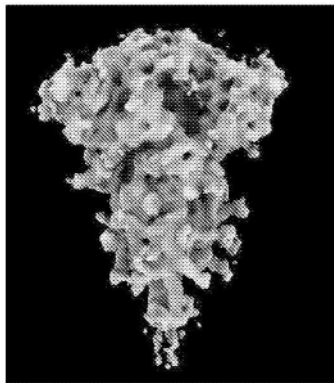


图1D

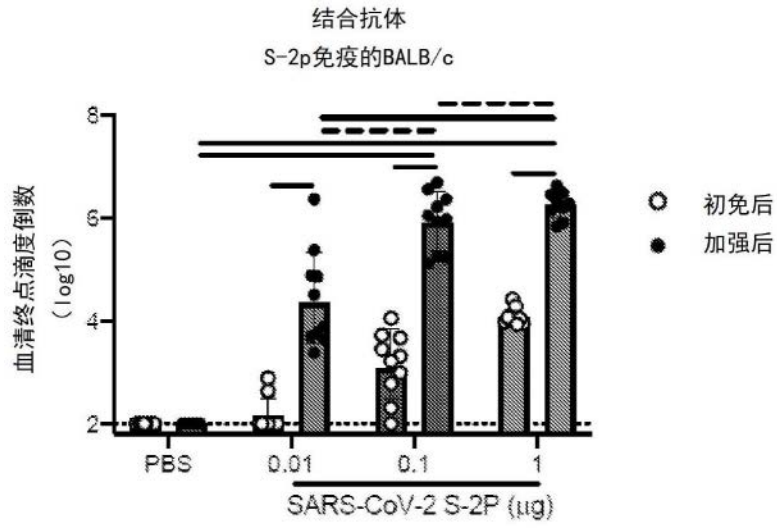


图2A

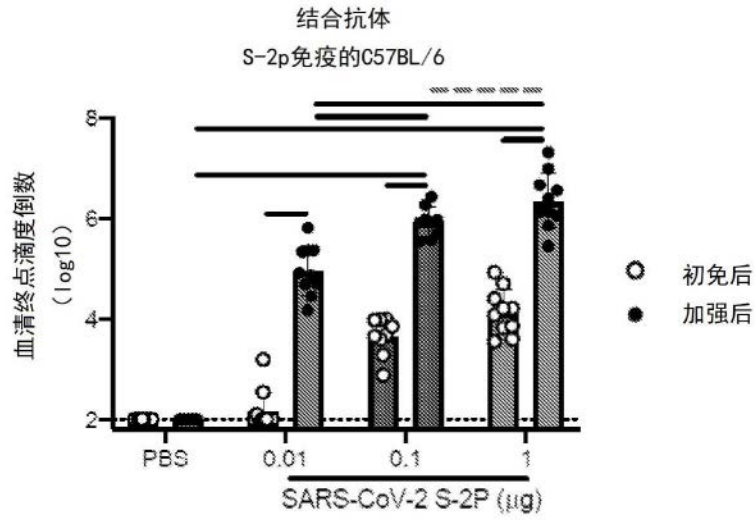


图2B

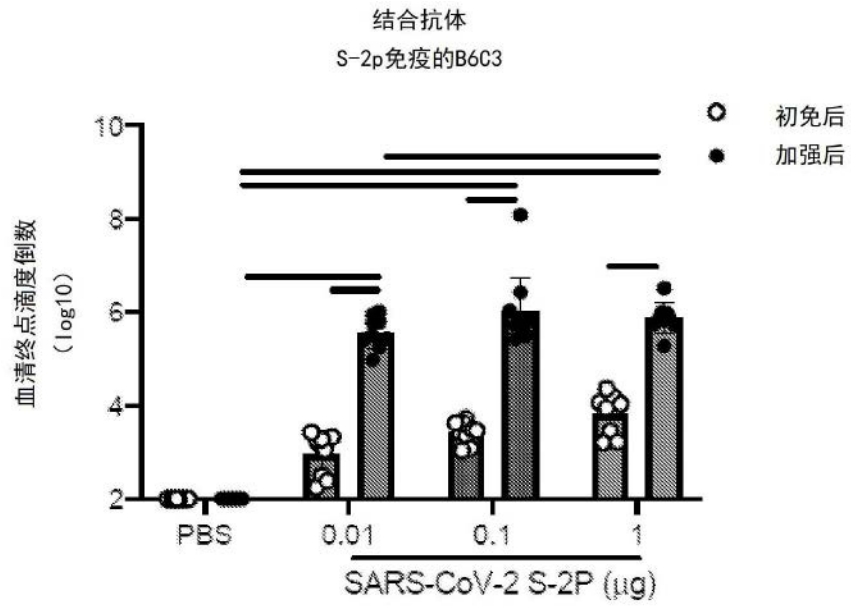


图2C

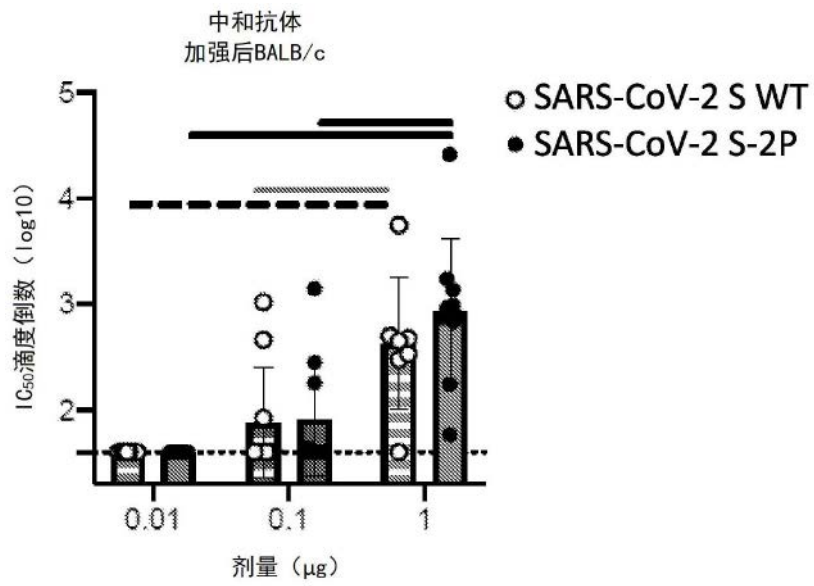


图2D

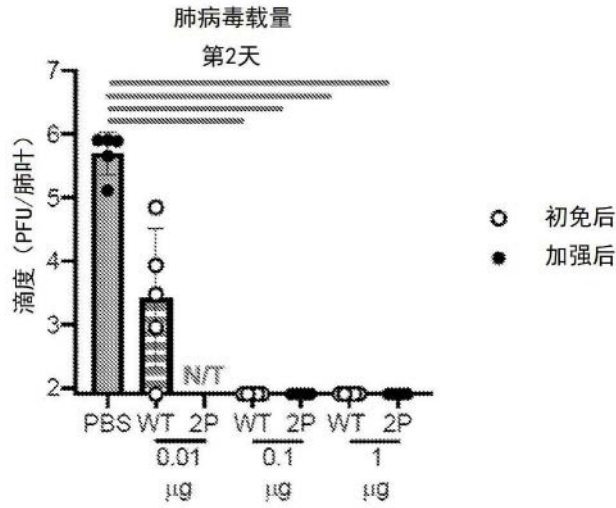


图3A

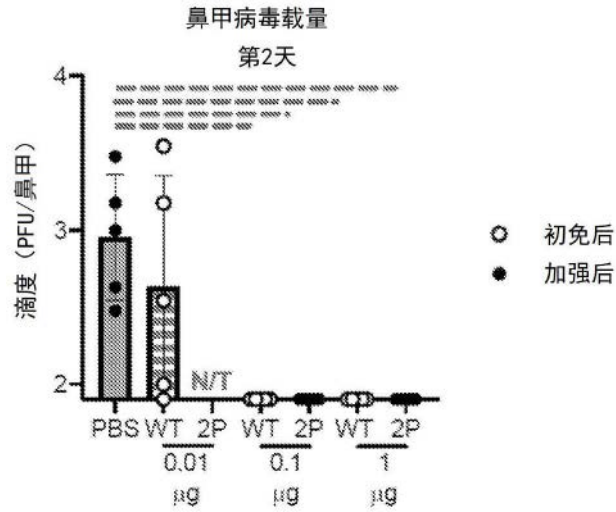


图3B

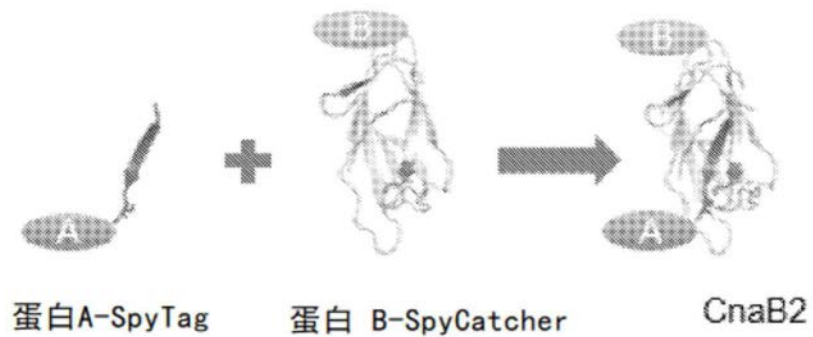


图4A

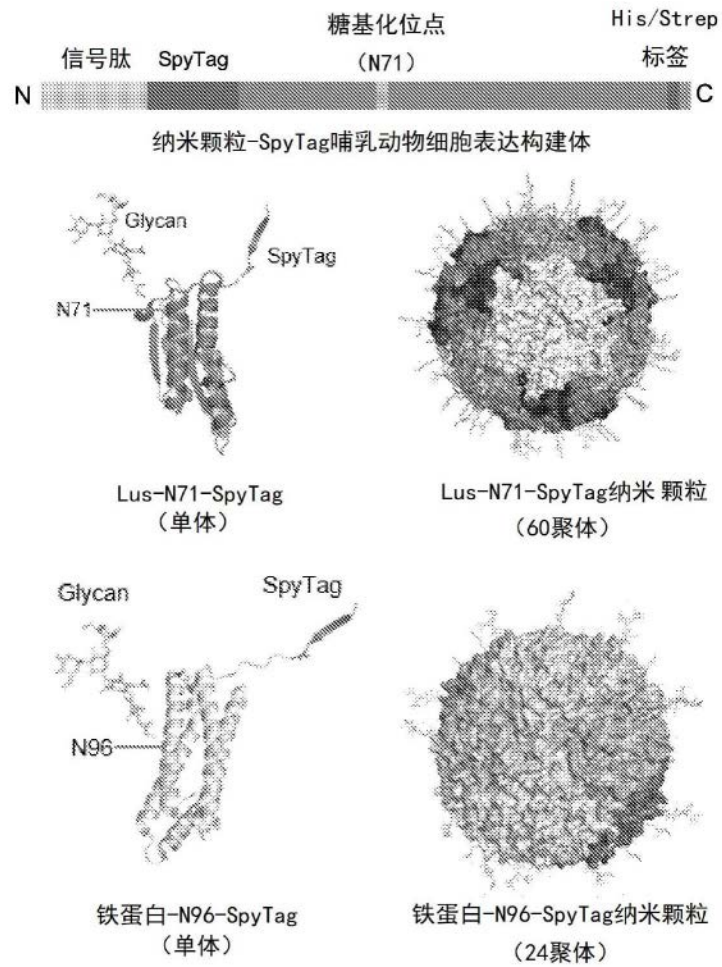


图4B

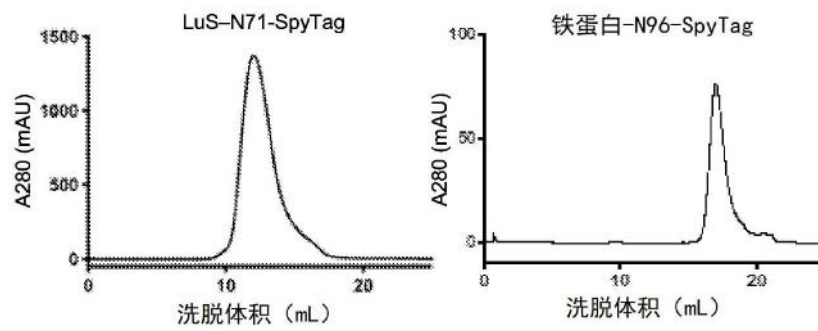


图4C

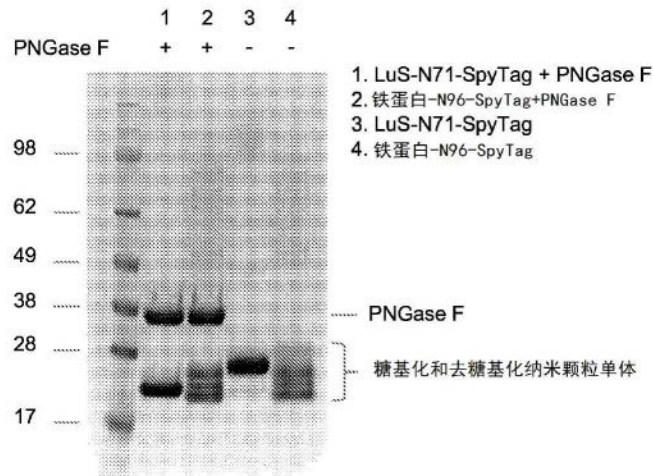


图4D

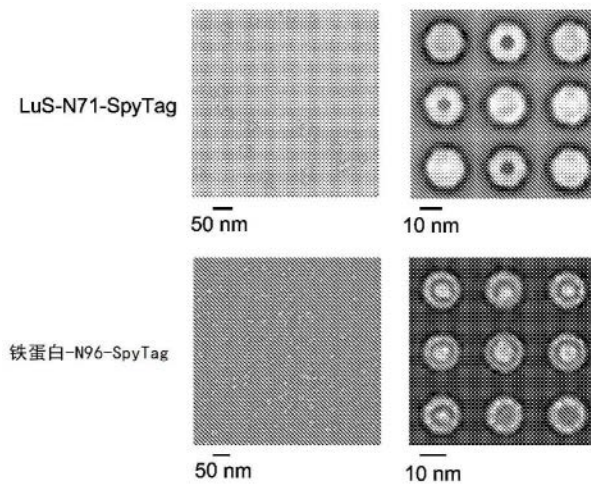


图4E

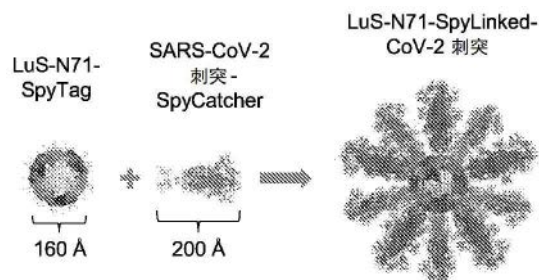


图5A

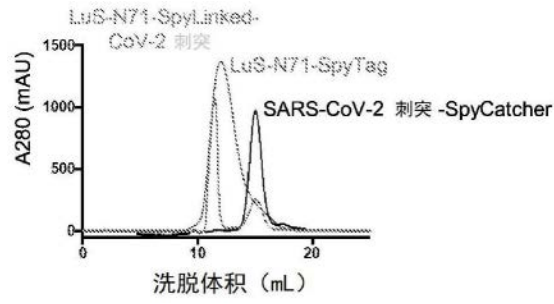


图5B

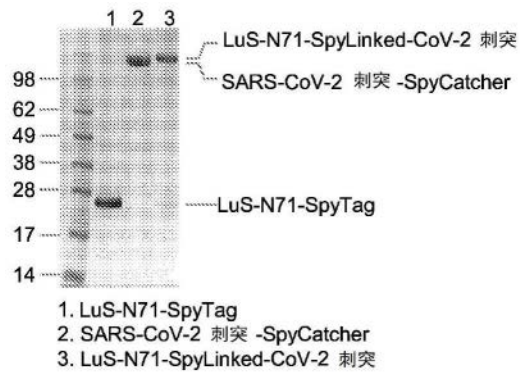


图5C

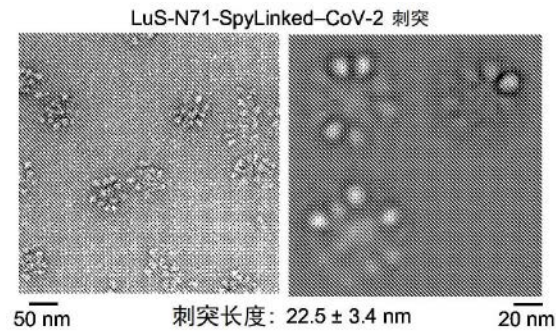


图5D

LuS-N71-SpyLinked-CoV-2 刺突的SPR

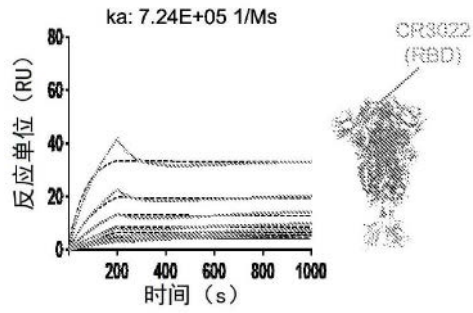


图5E

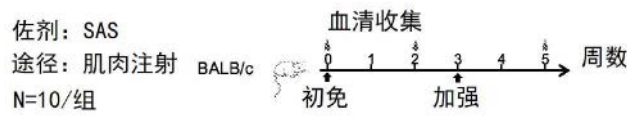


图6A

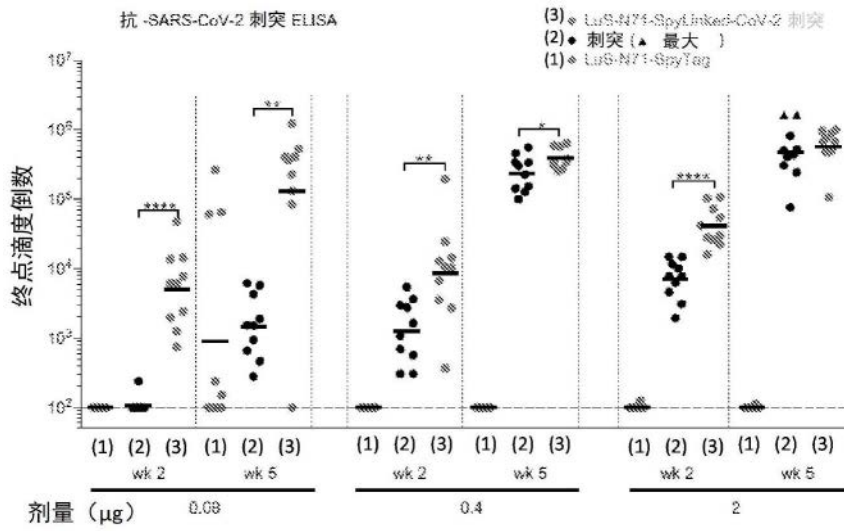


图6B

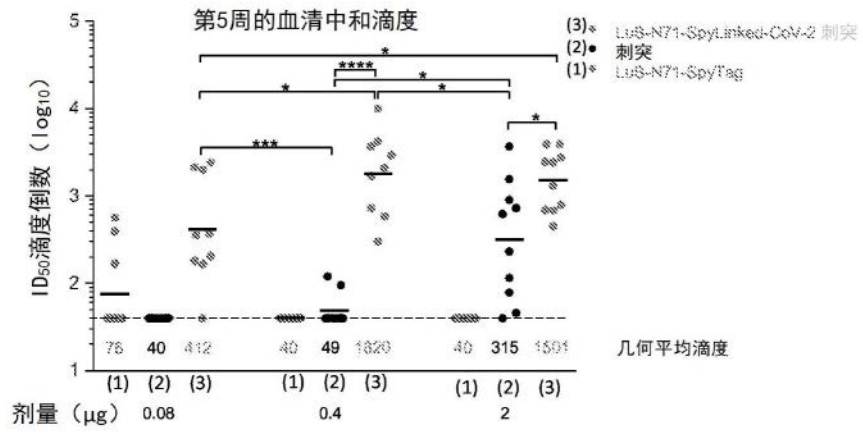


图6C