



(12) 发明专利申请

(10) 申请公布号 CN 116802193 A

(43) 申请公布日 2023. 09. 22

(21) 申请号	202180049973.2	伊利亚·芬克尔斯坦
(22) 申请日	2021.05.28	穆罕默德·贾万马尔迪
(30) 优先权数据		杰弗里·绍布 郭洪哲 周家伟
	63/032,502 2020.05.29 US	约里·戈德史密斯
(85) PCT国际申请进入国家阶段日		克里斯蒂·约尔特 谢庆林
	2023.01.28	帕特里克·伯恩 尼克尔·伯恩
(86) PCT国际申请的申请数据		王念双 丹尼尔·雷普
	PCT/US2021/034713 2021.05.28	(74) 专利代理机构 北京集佳知识产权代理有限公司 11227
(87) PCT国际申请的公布数据		专利代理师 张珊珊 刘振佳
	W02021/243122 EN 2021.12.02	(51) Int.Cl.
(71) 申请人 得克萨斯州大学系统董事会		C07K 14/005 (2006.01)
地址 美国得克萨斯州		
申请人 达特茅斯大学理事会		
(72) 发明人 贾森·麦克莱伦 珍妮弗·梅纳德		
安德烈亚·沙塞		权利要求书5页 说明书51页
		序列表25页 附图13页

(54) 发明名称

经改造冠状病毒刺突(S)蛋白及其使用方法

(57) 摘要

本文中提供了经改造冠状病毒S蛋白,例如经改造SARS-CoV-2 S蛋白。在一些方面中,所述经改造S蛋白表现出增强的构象稳定性和/或抗原性。还提供了用于在筛选平台和/或在疫苗组合物中使用经改造蛋白质作为诊断剂的方法。

1. 经改造蛋白质,其包含含有与以下具有至少90%同一性的序列的经改造冠状病毒S蛋白胞外结构域:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位;其中所述经改造蛋白质相对于SEQ ID NO:1或2的序列包含以下替换:F817P、A892P、A899P、A942P、K986P和V987P。

2. 经改造蛋白质,其包含与以下具有至少90%同一性的经改造冠状病毒S蛋白胞外结构域:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位,

所述经改造蛋白质相对于SEQ ID NO:1或2的序列包含至少一个突变,所述至少一个突变包含:

- (1) 经改造二硫键;
- (2) 腔填充替换;和/或
- (3) 提供静电或极性相互作用的替换。

3. 经改造蛋白质,其包含与以下具有至少90%同一性的经改造冠状病毒S蛋白胞外结构域:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位,

所述经改造蛋白质相对于SEQ ID NO:1或2的序列包含至少一个突变,所述至少一个突变包含:

- (i) 在与以下相对应的位置处的替换:

T724, T752, T778, T961, I1013, H1058, S735, T859, I770, A1015, L727, S1021, Q901, S875, T912, H1088, L1141, V1040, L966, A766, T778, L938, V963, V911, N1108, V705, A893, N703, A672, A694, A1080, I1132, P862, T859, T547, N978, T961, S758, Q762, D1118, S659, S698, R1039, V722, A930, A903, Q913, S974, D979, P728, V951, V736, L858, S884, A893, P807, S875, T791, A879, G799, A924, V826, A899, Q779, F817, L865, T866, A892, A899, T912, A570, V963; T874, S1055, V729, A1022, L894, A713, L828, H1058, L822, A1056, Q965, S1003, A972, Q992, I980, A1078, V1133, H1088, T1120, I870, S1055, T1117, D1139, T1116, Y1138, I896, G885, Q901, F1103, P1112, G889, L1034, E819, S1055, A972, I980, I1081, N1135, E819, Q1054, Q957, I1130, V1040, H1088, V1104, R1000, A944, T724, A944, S730, S730, G769, A893, Q895, K921, L922, N978, A942, G946, S975, A890, S1003; 和/或

(ii) 与第829至851、675至686、673至684、1161至1208或1142至1208位相对应的缺失; 和/或

- (iii) 第673至686位氨基酸的两个氨基酸的替换。

4. 权利要求1至3中任一项所述的经改造蛋白质,其包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:S735C和T859C; I770C和A1015C; L727C和S1021C; V911C和N1108C; A672C和A694C; A1080C和I1132C; S659C和S698C; V722C和A930C; A903C和Q913C; S974C和D979C; P728C和V951C; V736C和L858C; S884C和A893C; P807C和S875C; T791C和A879C; G799C和A924C; A570C和V963C; T874C和S1055C; V729C和A1022C; L822C和A1056C; Q965C和S1003C; A972C和Q992C; I980C和Q992C; A1078C和V1133C; H1088C和T1120C; I870C和S1055C; T1117C和D1139C; T1116C和Y1138C; I896C和Q901C; G885C和Q901C; F1103C和P1112C; G889C和L1034C; E819C和S1055C; A972C和I980C; I1081C和N1135C; 或E819C和

Q1054C。

5. 权利要求4所述的经改造蛋白质,其包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:A903C和Q913C;S884C和A893C;T791C和A879C;Q965C和S1003C;或T1117C和D1139C。

6. 权利要求4所述的经改造蛋白质,其包含含有在与S884C和A893C相对应的位置处的成对半胱氨酸替换的经改造二硫键。

7. 权利要求6所述的经改造蛋白质,其还包含至少一个另外的经改造二硫键。

8. 权利要求7所述的经改造蛋白质,其还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:T791C和A879C;或G799C和A924C。

9. 权利要求5所述的经改造蛋白质,其包含含有在与A903C和Q913C相对应的位置处的成对半胱氨酸替换的经改造二硫键。

10. 权利要求9所述的经改造蛋白质,其还包含至少一个另外的经改造二硫键。

11. 权利要求10所述的经改造蛋白质,其还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:Q965C和S1003C;S884C和A893C;T791C和A879C;或G799C和A924C。

12. 权利要求11所述的经改造蛋白质,其还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:A903C和Q913C;和/或Q965C和S1003C。

13. 权利要求1至3中任一项所述的经改造蛋白质,其包含在与以下相对应的位置处的腔填充替换:

T724,I1013,H1058,Q901,S875,H1088,L1141,V1040,T778,L938,V963,R1039,V826,A899,Q779,L894,V1040,V1104,R1000,A944,S730,A890,D1118,或S1003。

14. 权利要求13所述的经改造蛋白质,其包含在与以下相对应的位置处的腔填充替换:T778、L938、V963或H1088。

15. 权利要求13所述的经改造蛋白质,其包含选自以下的腔填充替换:

T724M,I1013F,H1058W,Q901M,S875F,H1088W,L1141F,V1040F,T778L,L938F,V963L,R1039F,V826L,A899F,Q779M,L894F,H1058F,H1058Y,V1040Y,H1088Y,V1104I,R1000Y,R1000W,A944F,T724I,A944Y,S730L,A890V,D1118F,或S1003V。

16. 权利要求15所述的经改造蛋白质,其包含选自以下的腔填充替换:T778L、L938F、V963L或H1088Y。

17. 权利要求13所述的经改造蛋白质,其包含在与L938相对应的位置处的腔填充替换。

18. 权利要求17所述的经改造蛋白质,其包含L938F替换。

19. 权利要求17至18中任一项所述的经改造蛋白质,其还包含在与V963相对应的位置处的腔填充替换。

20. 权利要求19所述的经改造蛋白质,其包含V963L替换。

21. 权利要求1至3中任一项所述的经改造蛋白质,其包含选自以下的脯氨酸替换:F817P,L865P,T866P,A892P,A899P,T912P,A893P,Q895P,K921P,L922P,N978P,A942P,G946P,或S975P。

22. 权利要求21所述的经改造蛋白质,其包含选自以下的脯氨酸替换:F817P、A892P、A899P或A942P。

23. 权利要求21所述的经改造蛋白质,其包含脯氨酸替换F817P。
24. 权利要求23所述的经改造蛋白质,其还包含经改造二硫键。
25. 权利要求24所述的经改造蛋白质,其还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:S884C和A893C;或T791C和A879C。
26. 权利要求25所述的经改造蛋白质,其还包含含有在与S884C和A893C相对应的位置处的成对半胱氨酸替换的经改造二硫键。
27. 权利要求21所述的经改造蛋白质,其还包含在V987P和/或K986P处的另外的脯氨酸替换。
28. 权利要求21所述的经改造蛋白质,其包含脯氨酸替换A892P。
29. 权利要求28所述的经改造蛋白质,其还包含在A942P、A899P和/或F817P处的另外的脯氨酸替换。
30. 权利要求21所述的经改造蛋白质,其包含至少两个选自A892P、A942P、A899P和/或F817P的脯氨酸替换。
31. 权利要求30所述的经改造蛋白质,其包含至少三个选自A892P、A942P、A899P和/或F817P的脯氨酸替换。
32. 权利要求31所述的经改造蛋白质,其包含在A892P、A942P、A899P和F817P处的脯氨酸替换。
33. 权利要求28所述的经改造蛋白质,其包含脯氨酸替换A899P或T912P。
34. 权利要求28所述的经改造蛋白质,其包含脯氨酸替换A892P和T912P。
35. 权利要求1至3中任一项所述的经改造蛋白质,其包含在与以下相对应的位置处提供静电相互作用替换的替换:T752,T912,L966,L828,S730,T961,A766,P862,T859,Q957,或G769。
36. 权利要求35所述的经改造蛋白质,其包含在与T961、L966、T859或G769相对应的位置处的静电相互作用替换。
37. 权利要求36所述的经改造蛋白质,其包含T961D或T961E静电相互作用替换。
38. 权利要求38所述的经改造蛋白质,其包含T961D替换。
39. 权利要求37至38中任一项所述的经改造蛋白质,其还包含L966E替换。
40. 权利要求35所述的经改造蛋白质,其包含选自以下的静电相互作用替换:T752K,T912R,L828K,L828R,S730R,T961D,A766E,P862E,T859K,Q957E,或G769E。
41. 权利要求40所述的经改造蛋白质,其包含选自以下的静电相互作用替换:T961D、L966D、T859K或G769K。
42. 权利要求1至3中任一项所述的经改造蛋白质,其包含在与T778、A713或I1130相对应的位置处提供静电或极性相互作用替换的替换。
43. 权利要求35所述的经改造蛋白质,其包含选自以下的静电相互作用替换:T778Q、A713S或I1130Y。
44. 权利要求2或3所述的经改造蛋白质,其包含在位置和F817P处提供静电相互作用替换的替换。
45. 权利要求1至44中任一项所述的经改造蛋白质,其还包含在与L984、D985、K986和/或V987相对应的位置处的替换。

46. 权利要求44所述的经改造蛋白质,其还包含在与L984、D985、K986和/或V987相对应的位置处向甘氨酸或脯氨酸的替换。

47. 权利要求1至44中任一项所述的经改造蛋白质,其包含K986P和V987P替换。

48. 权利要求1至47中任一项所述的经改造蛋白质,其还包含在与A570、T572、F855和/或N856相对应的位置处的替换。

49. 权利要求48所述的经改造蛋白质,其还包含在与A570、T572、F855和/或N856相对应的位置处的腔填充替换。

50. 权利要求1至49中任一项所述的经改造蛋白质,其包含至少一个经改造二硫键、至少一个腔填充替换、至少一个脯氨酸替换以及至少一个静电相互作用替换的组合。

51. 权利要求1至50中任一项所述的经改造蛋白质,其与SEQ ID NO:1或2的第319至1208位具有至少95%同一性。

52. 权利要求1至50中任一项所述的经改造蛋白质,其包含与SEQ ID NO:1或2的第16至1208位具有95%同一性的经改造冠状病毒S蛋白胞外结构域。

53. 权利要求1至52中任一项所述的经改造蛋白质,其中所述经改造冠状病毒S蛋白胞外结构域包含消除弗林蛋白酶切割位点的突变。

54. 权利要求53所述的经改造蛋白质,其中所述消除弗林蛋白酶切割位点的突变包含在第682至685位处的GSAS替换。

55. 权利要求1至54中任一项所述的经改造蛋白质,其中所述蛋白质与三聚结构域融合或缀合。

56. 权利要求55所述的经改造蛋白质,其中所述蛋白质与三聚结构域融合。

57. 权利要求55所述的经改造蛋白质,其中所述三聚结构域相对于S蛋白胞外结构域位于C端。

58. 权利要求56所述的经改造蛋白质,其中所述三聚结构域包含T4纤维蛋白三聚结构域。

59. 权利要求1至54中任一项所述的经改造蛋白质,其中所述蛋白质与跨膜结构域融合或缀合。

60. 权利要求59所述的经改造蛋白质,其中所述蛋白质与跨膜结构域融合。

61. 权利要求59所述的经改造蛋白质,其中所述跨膜结构域包含冠状病毒刺突蛋白跨膜结构域。

62. 权利要求61所述的经改造蛋白质,其中所述跨膜结构域包含SARS-CoV-2跨膜结构域。

63. 经改造冠状病毒三聚体,其包含至少一个根据权利要求1至50中任一项所述的亚基。

64. 权利要求63所述的经改造三聚体,其中相对于野生型S蛋白亚基的三聚体,所述三聚体在融合前构象中是稳定的。

65. 权利要求63所述的经改造三聚体,其中所述三聚体包含至少一个在亚基之间的经改造二硫键。

66. 权利要求65所述的经改造三聚体,其中所述至少一个在亚基之间的经改造二硫键选自:V705C和A983C、T547C和N968C、T961C和S758C、和/或T961C和Q762C。

67. 药物组合物,其包含:可药用载体;以及(i)权利要求1至62中任一项所述的经改造蛋白质,或(ii)权利要求63至66中任一项所述的经改造三聚体。

68. 权利要求67所述的组合物,其还包含佐剂。

69. 核酸分子,其包含编码权利要求1至62中任一项所述的经改造蛋白质的氨基酸序列的核苷酸序列。

70. 权利要求69所述的核酸,其中所述核酸包含DNA表达载体。

71. 权利要求69所述的核酸,其中所述核酸包含mRNA。

72. 在对象中预防冠状病毒感染或与冠状病毒感染相关的疾病的方法,其包括向所述对象施用有效量的根据权利要求67至68中任一项所述的药物组合物或根据权利要求69至71中任一项所述的核酸分子。

73. 组合物,其包含与抗体结合的权利要求1至62中任一项所述的经改造蛋白质。

经改造冠状病毒刺突(S)蛋白及其使用方法

[0001] 相关申请的引用

[0002] 本申请要求于2020年5月29日提交的美国临时申请号63/032,502的优先权权益,其全部内容通过引用并入本文。

[0003] 关于联邦资助研究的声明

[0004] 本发明是在由国立卫生研究院(National Institutes of Health)授予的基金号R01 AI127521的政府支持下完成的。政府对本发明具有一定的权利。

[0005] 序列表的引用

[0006] 本申请包含已通过EFS-Web以ASCII格式提交并且在此通过引用整体并入的序列表。创建于2021年5月19日的所述ASCII拷贝命名为UTFBP1251W0_ST25.txt,并且大小为41.8千字节。

背景技术

1. 技术领域

[0007] 本公开内容一般地涉及医学、病毒学、免疫学和蛋白质工程领域。更特别地,本公开内容涉及经改造冠状病毒S蛋白及其在药物设计和疫苗制剂中的用途。

[0008] 2. 相关技术描述

[0009] 由冠状病毒SARS-CoV-2感染引起的疾病COVID-19的爆发已导致数百万例感染及超过10万例死亡。像在数年前导致SARS爆发的病毒SARS-CoV一样,SARS-CoV-2病毒使用其刺突蛋白以结合宿主细胞受体血管紧张素转换酶2(angiotensin-converting enzyme 2, ACE2)。刺突糖蛋白的受体结合结构域(receptor binding domain,RBD)与全长人ACE2蛋白之间相互作用。虽然SARS-CoV-2刺突蛋白的序列和结构是已知的(参见例如,Wrapp et al.2020),但是仍需要可用于鉴定药物候选物并用于刺激针对S蛋白的有效免疫应答的稳定的S蛋白。

发明内容

[0010] 在一些实施方案中,本公开内容提供了经改造蛋白质,其包含与以下具有至少90%同一性的经改造冠状病毒S蛋白胞外结构域:(a)SEQ ID NO:1或2的第14至1208位;(b)SEQ ID NO:1或2的第14至1160位;或者(c)SEQ ID NO:1或2的第319至1208位,所述经改造蛋白质相对于SEQ ID NO:1或2的序列包含至少一个突变,所述至少一个突变包含:

[0011] (1)经改造二硫键;

[0012] (2)腔填充替换;

[0013] (3)提供静电或极性相互作用的替换;和/或

[0014] (4)脯氨酸替换。

[0015] 在另一些方面中,经改造蛋白质包含至少一个经改造二硫键和至少一个腔填充替换。在又一些方面中,经改造蛋白质包含至少一个经改造二硫键和至少一个提供静电或极

性相互作用的替换。在又一些方面中,经改造蛋白质包含至少一个经改造二硫键和至少一个脯氨酸替换。在另一些方面中,经改造蛋白质包含至少一个腔填充替换和至少一个提供静电或极性相互作用的替换。在又一些方面中,经改造蛋白质包含至少一个腔填充替换和至少一个脯氨酸替换。在又一些方面中,经改造蛋白质包含至少一个提供静电或极性相互作用的替换和至少一个脯氨酸替换。

[0016] 在一些实施方案中,本公开内容提供了经改造蛋白质,其包含含有与以下具有至少90%同一性的序列的经改造冠状病毒S蛋白胞外结构域:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位;其中所述经改造蛋白质相对于SEQ ID NO:1或2的序列包含以下替换:F817P、A892P、A899P、A942P、K986P和V987P。在另一些方面中,经改造冠状病毒S蛋白与以下具有至少约91%、92%、93%、94%、95%、96%、97%或98%同一性:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位。

[0017] 在另一些实施方案中,本公开内容提供了经改造蛋白质,其包含与以下具有至少90%同一性的经改造冠状病毒S蛋白胞外结构域:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位,所述至少一个突变包含:

[0018] (i) 在与以下相对应的位置处的替换:

[0019] T724,T752,T778,T961,I1013,H1058,S735,T859,I770,A1015,L727,S1021,Q901,S875,T912,H1088,L1141,V1040,L966,A766,T778,L938,V963,V911,N1108,V705,A893,N703,A672,A694,A1080,I1132,P862,T547,N978,S758,Q762,D1118,S659,S698,R1039,V722,A930,A903,Q913,S974,D979,P728,V951,V736,L858,S884,P807,T791,A879,G799,A924,V826,A899,Q779,F817,L865,T866,A892,A899,A570,T874,S1055,V729,A1022,L894,A713,L828,L822,A1056,Q965,S1003,A972,Q992,1980,A1078,V1133,T1120,I870,T1117,D1139,T1116,Y1138,I896,G885,F1103,P1112,G889,L1034,E819,A972,I980,I1081,N1135,E819,Q1054,Q957,I1130,V1040,V1104,R1000,A944,T724,A944,S730,G769,Q895,K921,L922,A942,G946,S975,A890;以及

[0020] (ii) 与第829至851、675至686、673至684、1161至1208、或1142至1208位相对应的缺失;以及

[0021] (iii) 第673至686位氨基酸的两个氨基酸的替换。

[0022] 在另一些方面中,经改造冠状病毒S蛋白与以下具有至少约91%、92%、93%、94%、95%、96%、97%或98%同一性:(a) SEQ ID NO:1或2的第14至1208位;(b) SEQ ID NO:1或2的第14至1160位;或者(c) SEQ ID NO:1或2的第319至1208位。

[0023] 在一些方面中,经改造蛋白质包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:S735C和T859C;I770C和A1015C;L727C和S1021C;V911C和N1108C;A672C和A694C;A1080C和I1132C;S659C和S698C;V722C和A930C;A903C和Q913C;S974C和D979C;P728C和V951C;V736C和L858C;S884C和A893C;P807C和S875C;T791C和A879C;G799C和A924C;A570C和V963C;T874C和S1055C;V729C和A1022C;L822C和A1056C;Q965C和S1003C;A972C和Q992C;I980C和Q992C;A1078C和V1133C;H1088C和T1120C;I870C和S1055C;T1117C和D1139C;T1116C和Y1138C;I896C和Q901C;G885C和Q901C;F1103C和P1112C;G889C和

L1034C;E819C和S1055C;A972C和I980C;I1081C和N1135C;或E819C和Q1054C。

[0024] 在一些方面中,经改造蛋白质包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:A903C和Q913C、S884C和A893C、T791C和A879C、Q965C和S1003C、或T1117C和D1139C。在另一些方面中,经改造蛋白质包含含有在与S884C和A893C相对应的位置处的成对半胱氨酸替换的经改造二硫键。在另一些方面中,经改造蛋白质还包含至少一个另外的经改造二硫键。在另一些方面中,经改造蛋白质还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:T791C和A879C;或G799C和A924C。

[0025] 在另一些方面中,经改造蛋白质包含含有在与A903C和Q913C相对应的位置处的成对半胱氨酸替换的经改造二硫键。在另一些方面中,经改造蛋白质还包含至少一个另外的经改造二硫键。在另一些方面中,经改造蛋白质还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:Q965C和S1003C;S884C和A893C;T791C和A879C;或G799C和A924C。在另一些方面中,经改造蛋白质还包含含有在与以下相对应的位置处的成对半胱氨酸替换的经改造二硫键:A903C和Q913C;和/或Q965C和S1003C。

[0026] 在一些方面中,经改造蛋白质包含在与以下相对应的位置处的腔填充替换:

[0027] T724,I1013,H1058,Q901,S875,H1088,L1141,V1040,T778,L938,V963,R1039,V826,A899,Q779,L894,V1040,V1104,R1000,A944,S730,A890,D1118,或S1003。

[0028] 在另一些方面中,经改造蛋白质包含在与以下相对应的位置处的腔填充替换:T778、L938、V963或H1088。在另一些方面中,经改造蛋白质包含选自以下的腔填充替换:

[0029] T724M,I1013F,H1058W,Q901M,S875F,H1088w,L1141F,V1040F,T778L,L938F,V963L,R1039F,V826L,A899F,Q779M,L894F,H1058F,H1058Y,V1040Y,H1088Y,V1104I,R1000Y,R1000W,A944F,T724I,A944Y,S730L,A890V,D1118F,或S1003V。

[0030] 在另一些方面中,经改造蛋白质包含选自以下的腔填充替换:T778L、L938F、V963L或H1088Y。在另一些方面中,经改造蛋白质包含在与L938相对应的位置处的腔填充替换。在另一些方面中,经改造蛋白质包含L938F替换。在另一些方面中,经改造蛋白质还包含在与V963相对应的位置处的腔填充替换。在另一些方面中,经改造蛋白质包含V963L替换。

[0031] 在一些方面中,经改造蛋白质包含选自以下的脯氨酸替换:F817P,L865P,T866P,A892P,A899P,T912P,A893P,Q895P,K921P,L922P,N978P,A942P,G946P,或S975P。在另一些方面中,经改造蛋白质包含选自以下的脯氨酸替换:F817P、A892P、A899P或A942P。在另一些方面中,经改造蛋白质包含脯氨酸替换F817P。在另一些方面中,经改造蛋白质还包含经改造二硫键。在另一些方面中,经改造蛋白质还包含含有在与S884C和A893C;或T791C和A879C相对应的位置处的成对半胱氨酸替换的经改造二硫键。在另一些方面中,经改造蛋白质还包含含有在与S884C和A893C相对应的位置处的成对半胱氨酸替换的经改造二硫键。在一些方面中,经改造蛋白质还包含在V987P和/或K986P处的另外的脯氨酸替换。在一些方面中,经改造蛋白质包含脯氨酸替换A892P。在另一些方面中,经改造蛋白质还包含在A942P、A899P和/或F817P处的另外的脯氨酸替换。在一些方面中,经改造蛋白质包含至少两个选自A892P、A942P、A899P和/或F817P的脯氨酸替换。在另一些方面中,经改造蛋白质包含至少三个选自A892P、A942P、A899P和/或F817P的脯氨酸替换。在另一些方面中,经改造蛋白质包含在A892P、A942P、A899P和F817P处的脯氨酸替换。在另一些方面中,经改造蛋白质包含脯氨酸替换A899P或T912P。在又一些方面中,经改造蛋白质包含脯氨酸替换A892P和T912P。

[0032] 在一些方面中,经改造蛋白质包含在与T752,T912,L966,L828,S730,T961,A766,P862,T859,Q957,或G769相对应的位置处提供静电相互作用替换的替换。在另一些方面中,经改造蛋白质包含在与T961、L966、T859或G769相对应的位置处的静电相互作用替换。在又一些方面中,经改造蛋白质包含T961D或T961E静电相互作用替换。在又一些方面中,经改造蛋白质包含T961D替换。在一些方面中,经改造蛋白质还包含L966D或L966E替换,优选L966E替换。在另一些方面中,经改造蛋白质包含选自以下的静电相互作用替换:T752K,T912R,L966D,L828K,L828R,S730R,T961D,A766E,P862E,T859K,Q957E,或G769E。在另一些方面中,经改造蛋白质包含选自以下的静电相互作用替换:T961D、L966D、T859K或G769K。在又一些方面中,静电相互作用替换选自:T961D、T859K或G769K。在一些方面中,经改造蛋白质包含在与T778、A713或I1130相对应的位置处提供静电或极性相互作用替换的替换。在一些方面中,经改造蛋白质包含选自以下的静电相互作用替换:T778Q、A713S或I1130Y。在一些方面中,经改造蛋白质包含在位置和F817P处提供静电相互作用替换的替换。

[0033] 在一些方面中,经改造蛋白质还包含在与L984、D985、K986和/或V987相对应的位置处的替换。在另一些方面中,经改造蛋白质包含在与L984、D985、K986和/或V987相对应的位置处向甘氨酸或脯氨酸的替换。在一些方面中,经改造蛋白质包含K986P和V987P替换。在另一些方面中,经改造蛋白质还包含在与A570、T572、F855和/或N856相对应的位置处的替换。在另一些方面中,经改造蛋白质还包含在与A570、T572、F855和/或N856相对应的位置处的腔填充替换。在一些方面中,经改造蛋白质包含至少一个经改造二硫键和至少一个脯氨酸替换的组合。在另一些方面中,经改造蛋白质包含至少一个腔填充替换和至少一个脯氨酸替换的组合。在又一些方面中,经改造蛋白质包含至少一个脯氨酸替换和至少一个静电相互作用替换的组合。在一些方面中,经改造蛋白质包含至少一个经改造二硫键、至少一个腔填充替换、至少一个脯氨酸替换和至少一个静电相互作用替换的组合。在一些方面中,经改造蛋白质与SEQ ID NO:1或2的第319至1208位具有至少95%同一性。在一些方面中,经改造蛋白质包含与SEQ ID NO:1或2的第16至1208位具有95%同一性的经改造冠状病毒S蛋白胞外结构域。在一些方面中,经改造蛋白质包含含有消除弗林蛋白酶切割位点的突变的经改造冠状病毒S蛋白胞外结构域。在另一些方面中,经改造蛋白质的消除弗林蛋白酶切割位点的突变包含在第682至685位处的GSAS替换。

[0034] 本文中所述的替换中的任一者可被改造成任何已知的冠状病毒S蛋白变体,包括但不限于具有以下修饰中的任一者或更多者的冠状病毒S蛋白(参见SEQ ID NO:2):

[0035] L5F,S13I,L18F,T19R,T20N,P26S,Q52R,A67V,H69缺失,V70缺失,V70I,D80A,T95I,D138Y,Y144缺失,Y144V,W152C,E154K,R190S,D215G,L242缺失,A243缺失,L244缺失,D253G,W258L,K417N,K417T,L452R,S477N,T478K,E484K,E484Q,E484K,N501Y,A570D,D614G,H655Y,Q677H,P681R,P681H,A701V,T716I,F888L,D950N,S982A,T1027I,D1118H,和V1176F。这样的修饰的一些示例性组合在表5中提供。

[0036] 在一些方面中,经改造蛋白质与三聚结构域融合或缀合。在另一些方面中,蛋白质与三聚结构域融合。在一些方面中,三聚结构域相对于S蛋白胞外结构域位于C端。在一些方面中,三聚结构域包含T4纤维蛋白三聚结构域。在一些方面中,蛋白质与跨膜结构域融合或缀合。在一些方面中,蛋白质与跨膜结构域融合。在一些方面中,跨膜结构域包含冠状病毒刺突蛋白跨膜结构域。在一些方面中,跨膜结构域包含SARS-CoV或SARS-CoV-2跨膜结构域。

[0037] 在又一些实施方案中,本公开内容提供了经改造冠状病毒三聚体,其包含至少一个本公开内容的亚基。在一些方面中,相对于野生型S蛋白亚基的三聚体,所述三聚体在融合前构象中是稳定的。在一些方面中,三聚体包含至少一个在亚基之间的经改造二硫键。在另一些方面中,至少一个在亚基之间的经改造二硫键选自:V705C和A983C;T547C和N968C;T961C和S758C;和/或T961C和Q762C。

[0038] 在又一些实施方案中,本公开内容提供了药物组合物,其包含:可药用载体;以及(i)本公开内容的经改造蛋白质,或(ii)本公开内容的经改造三聚体。在一些方面中,组合物还包含佐剂。

[0039] 在另一些实施方案中,本公开内容提供了核酸分子,其包含编码本公开内容的经改造蛋白质的氨基酸序列的核苷酸序列。在一些方面中,核酸包含DNA表达载体。在另一些方面中,核酸包含mRNA。

[0040] 在又一些实施方案中,本公开内容提供了组合物,其包含与抗体结合的本公开内容的经改造蛋白质。

[0041] 在又一些实施方案中,本公开内容提供了在对象中治疗或预防冠状病毒感染的方法,所述方法包括向对象施用治疗有效量的针对对象的实施方案的经改造S蛋白。在一些方面中,所述方法在对象中刺激体液和/或细胞免疫应答。在一些方面中,所述方法降低了感染冠状病毒的对象的肺中的炎症。在一些方面中,对象感染了SARS-CoV或SARs-Cov-2。在一些方面中,对象是处于感染SARS-CoV或SARs-Cov-2的风险中的未感染对象。在一些方面中,对象患肺炎的风险提高。在一些方面中,所述方法还包括向对象施用另外的抗病毒治疗。

[0042] 在另一些实施方案中,本公开内容提供了在样品或对象中检测冠状病毒、冠状病毒S蛋白结合抗体和/或被冠状病毒感染的细胞的方法,其包括:(a)使对象或来自对象的样品与本公开内容的经改造S蛋白接触;以及(b)检测所述抗体或细胞与经改造S蛋白的结合。在一些方面中,样品是体液或活检物。在一些方面中,样品是血液、骨髓、痰、泪、唾液、黏液、血清、尿、粪便或鼻拭子。在一些方面中,检测包括免疫组织化学、流式细胞术、FACS、ELISA、RIA或Western印迹。在一些方面中,所述方法还包括第二次进行步骤(a)和(b),以及确定与第一次相比检测水平的变化。在一些方面中,经改造S蛋白还包含标记或是固定在表面上的。在另一些方面中,所述标记是肽标签、酶、磁粒、发色团、荧光分子、化学发光分子或染料。在一些方面中,经改造S蛋白与脂质体或纳米粒缀合。

[0043] 预期本文中所述的任何方法或组合物可相对于本文中所述的任何其他方法或组合物来实施。根据以下详细描述,本公开内容的其他目的、特征和优点将变得明显。然而,应理解,详细描述和具体实例虽然指出了本发明的一些具体实施方案,但仅通过举例说明的方式给出,因为根据该详细描述,在本公开内容的精神和范围内的多种改变和修改对于本领域技术人员而言将变得明显。

附图说明

[0044] 以下附图构成了本说明书的一部分,并且被包括在内以进一步说明本发明的某些方面。通过结合本文中给出的具体实施方案的详细描述来参考这些附图中的一个或多个可以更好地理解本发明。

[0045] 图1示出了针对SARS-CoV-2刺突稳定的示例性替换。融合前构象中三聚体SARS-

CoV-2刺突胞外结构域的侧视图(PDB ID:6VSB)。S1结构域示出为透明表面。各原聚体的S2结构域示出为带状图。各插图对应于四种类型的刺突变体(脯氨酸、盐桥、二硫化物、腔填充)之一。各插图中的侧链如下:左上(脯氨酸)、左下(二硫化物)、右上(盐桥)以及右下(腔填充)。

[0046] 图2A至2G示出了单替换刺突变体的特征。(图2A) SARS-CoV-2S-2P和单替换刺突变体的SDS-PAGE。分子量标准在左侧以kDa表示。(图2B至2D)按类型分组的经纯化刺突变体的尺寸排阻色谱(图2B,二硫化物变体;图2C,腔填充和盐桥;图2D,脯氨酸)。S-2P的代表性数据在各图上示出为黑色虚线。垂直虚线表示S-2P的特征峰保留体积。图2B上的顶部线是Q965C、S1003C。图2D上的顶部线是A942P,并且图2D上从顶部线起第二个是F817P。(图2E)四种变体的代表性负染色电子显微照片。(图2F)刺突变体热稳定性的差示扫描荧光法(differential scanning fluorimetry,DSF)分析。垂直虚线表示S-2P的第一表观解链温度。在约47℃下具有谷值的线是A942P。在约67℃下具有谷值的线是A892P。(图2G)通过定量生物层干涉术确定的培养基中单一变体的浓度。变体按类型进行排序。虚线表示S-2P的计算浓度,其用作对照用于比较。

[0047] 图3A至3D示出了多替换刺突变体的特征。(图3A) SARS-CoV-2Combo变体的SDS-PAGE。分子量标准在左侧以kDa表示。(图3B) S-2P、A892P和四种Combo变体的SEC迹线。垂直虚线表示S-2P的峰保留体积。(图3C) Combo变体热稳定性的DSF分析。左侧垂直虚线表示S-2P的第一表观解链温度,右侧垂直虚线示出Combo47(HexaPro)的第一表观解链温度。(图3D)经纯化Combo47的负染色电子显微照片。

[0048] 图4A至4H示出了与S-2P相比,HexaPro表现出增强的表达和稳定性。(图4A)在从FreeStyle 293细胞的2L培养物中纯化之后HexaPro的SEC迹线。(图4B)从FreeStyle 293细胞纯化的HexaPro的负染色电子显微照片。(图4C)在从ExpiCHO细胞的2L培养物中纯化之后HexaPro的SEC迹线。(图4D)ExpiCHO细胞的2L培养物的经纯化HexaPro的负染色电子显微照片。(图4E和4F)通过表面等离子体共振评估的S-2P(图4E)和HexaPro(图4F)与ACE2的结合。结合数据示出为黑色线,并且最佳拟合至1:1结合模型示出为红色线。(图4G和4H)通过负染色电子显微术评估蛋白质的稳定性。(图4G)和(图4H)中显微照片的顶行对应于S-2P,底行对应于HexaPro。

[0049] 图5A至5C示出了HexaPro的高分辨率cryo-EM结构。(图5A)三聚体HexaPro的EM密度图。(图5B)HexaPro(绿色条带)与S-2P(白色条带,PDB ID:6VSB)的比对。示出了采用1-RBD-向上构象(one-RBD-up conformation)的单独的原聚体。(图5C)HexaPro独有的四种脯氨酸替换的缩放视图。EM密度图示出为透明表面,单个原子示出为棒。

[0050] 图6示出了具有左移SEC峰的变体的负染色EM图像。

[0051] 图7示出了折叠良好的颗粒的负染色EM图像。

[0052] 图8A至8B示出了二硫化物和腔填充组合变体(Combo23)的表征。(图8A) S-2P、Combo23以及亲本变体S884C/A893C(二硫键)和L938F(腔填充)的SEC迹线。顶部线是Combo23。(图8B) S-2P、Combo23及其亲本变体的DSF解链温度分析。左侧垂直虚线表示S-2P的T_m,并且右侧垂直虚线表示S884C/A893C的T_m。

[0053] 图9示出了cryo-EM数据处理工作流程。

[0054] 图10示出了cryo-EM结构验证。在cryoSPARC v2.15中生成的FSC曲线和观察分布

图示出了2-RBD-向上(two-RBD-up)(左)和1-RBD-向上(one-RBD-up)(右)重构二者。根据局部分辨率显示各重构的Cryo-EM密度,其中通过密度的中心切片示出在右侧。

具体实施方式

[0055] 本文中提供了经改造冠状病毒刺突蛋白。在一些方面中,一些实施方案的蛋白质在于膜融合之前存在的构象中是稳定的。这样的经改造蛋白质可用于例如刺激抗冠状病毒S蛋白特异性免疫应答。在另一些方面中,经改造S蛋白可用于在样品中检测S蛋白结合抗体。因此,本文中提供的经改造蛋白质允许用于更有效的用于针对冠状病毒的疫苗接种的方法,以及能够实现用于检测抗冠状病毒抗体(例如,生物样品)的新的测定方法。

[0056] I. 定义

[0057] 应理解,上述一般性描述和以下详细描述二者仅是示例性和说明性的,并且不对所要求保护的发明具有限制性。在本申请中,除非另外特别说明,否则单数的使用包括复数。在本申请中,除非另有说明,否则“或/或者”的使用意指“和/或”。此外,术语“包括”以及其他形式例如“包含”和“含有”的使用不是限制性的。此外,除非另外特别说明,否则术语例如“要素”或“组分”涵盖构成一个单元的要素和组分以及构成多于一个亚单元的要素和组分二者。此外,术语“部分”的使用可以包括部分的一部分或整个部分。

[0058] 如本文中使用的,除非上下文另外明确指出,否则单数形式包括复数指代物。如本文在说明书中使用的,没有数量词修饰的名词可意指一个/种或更多个/种。如本文在权利要求书中使用的,当与词语“包含/包括”联合使用时,没有数量词修饰的名词可意指一个/种或多于一个/种。

[0059] 当提及可测量值例如量、持续时间等时,本文中使用的术语“约”意指涵盖来自指定值的多至 $\pm 10\%$ 的变化。除非另有说明,否则本说明书和权利要求书中使用的表示成分的量、特性例如分子量、反应条件等的所有数字在所有情况下应被理解为由术语“约”修饰。因此,除非相反地指出,否则以下说明书和所附权利要求书中阐述的数值参数为近似值,其可以根据由所公开的主题试图获得的期望特性而变化。至少,并不试图将等同原则的应用限制在权利要求书的范围内,每个数值参数应当至少根据报告的有效数字的数值并通过应用一般的舍入技术来解释。虽然陈述本发明宽范围的数值范围和参数是近似值,但是在具体实施例中陈述的数值是尽可能精确地报导的。然而,任何数值均固有地包括不可避免地由在其各自的测试测量中存在的标准偏差产生的某些误差。

[0060] 本文中使用的就指定组分而言的“基本上不含”在本文中用于意指没有指定组分被有目的地配制成组合物和/或仅作为污染物或以痕量存在。因此,由组合物的任何非预期的污染导致的指定组分的总量远低于 0.05% ,优选低于 0.01% 。最优选的是其中用标准分析方法不能检测到指定组分的量的组合物。

[0061] 术语“抗体”是指任何同种型的完整免疫球蛋白、或者其可与完整抗体竞争与靶抗原特异性结合的片段,并且包括例如嵌合的、人源化的、完全的人以及双特异性的抗体。“抗体”是一种抗原结合蛋白。完整抗体将通常包含至少两条全长重链和两条全长轻链,但是在一些情况下可以包含更少的链,例如天然存在于骆驼科中的抗体,其可以仅包含重链。抗体可以仅来源于单一来源,或者可以是“嵌合的”,即抗体的不同部分可以来源于两种不同的抗体,如下文进一步所述的。抗原结合蛋白、抗体或结合片段可在杂交瘤中产生、通过重组

DNA技术或通过完整抗体的酶促或化学切割。除非另有说明,否则术语“抗体”除了包含两条全长重链和两条全长轻链的抗体之外,还包括其衍生物、变体、片段和突变蛋白,这些的实例在下面描述。此外,除非明确排除,否则抗体分别包括单克隆抗体、双特异性抗体、微抗体、结构域抗体、合成抗体(本文中有时称为“抗体模拟物”)、嵌合抗体、人源化抗体、人抗体、抗体融合体(本文中有时称为“抗体缀合物”)及其片段。在一些实施方案中,该术语还涵盖肽抗体。

[0062] 天然存在的抗体结构单元通常包含四聚体。每个这样的四聚体通常由两个相同的多肽链对构成,每对具有一条全长“轻”链(在某些实施方案中,约25kDa)和一条全长“重”链(在某些实施方案中,约50至70kDa)。每条链的氨基末端部分通常包含具有约100至110个或更多个氨基酸的可变区,该可变区通常负责抗原识别。每条链的羧基末端部分通常限定可以负责效应物功能的恒定区。人轻链通常被分类为 κ 和 λ 轻链。重链通常被分类为 μ 、 δ 、 γ 、 α 或 ϵ ,并且将抗体的同种型分别限定为IgM、IgD、IgG、IgA和IgE。IgG具有数个亚类,包括但不限于IgG1、IgG2、IgG3和IgG4。IgM具有亚类,包括但不限于IgM1和IgM2。类似地,将IgA细分为亚类,包括但不限于IgA1和IgA2。通常来说,在全长轻链和重链中,可变区和恒定区通过具有约12个或更多个氨基酸的“J”区连接,其中重链还包含具有约10个更多个氨基酸的“D”区。参见,例如,Fundamental Immunology, Ch.7 (Paul, W., ed., 2nd ed. Raven Press, N.Y. (1989)) (出于所有目的,其通过引用整体并入)。每个轻链/重链对的可变区通常形成了抗原结合位点。

[0063] 术语“可变区”或“可变结构域”是指抗体的轻链和/或重链的一部分,其通常包含重链中约120至130个氨基末端氨基酸和轻链中约100至110个氨基末端氨基酸。在某些实施方案中,不同抗体的可变区在氨基酸序列上有很大差异,甚至在同种抗体之间也是如此。抗体的可变区通常确定了特定抗体对其靶标的特异性。

[0064] 可变区通常显示出由三个高变区(也称为互补决定区或CDR)连接的相对保守的框架区(framework region, FR)的相同一般结构。来自每对的两条链的CDR通常通过框架区进行排列,这可使得能够与特定的表位结合。从N端到C端,轻链和重链可变区二者通常包含结构域FR1、CDR1、FR2、CDR2、FR3、CDR3和FR4。每个结构域的氨基酸分配通常根据以下的定义: Kabat Sequences of Proteins of Immunological Interest (National Institutes of Health, Bethesda, Md. (1987 and 1991)), Chothia & Lesk, J. Mol Biol., 196:901-917 (1987) 或 Chothia et al., Nature, 342:878-883 (1989)。

[0065] 在某些实施方案中,在缺乏抗体轻链的情况下,抗体重链与抗原结合。在某些实施方案中,在缺乏抗体重链的情况下,抗体轻链与抗原结合。在某些实施方案中,在缺乏抗体轻链的情况下,抗体结合区与抗原结合。在某些实施方案中,在缺乏抗体重链的情况下,抗体结合区与抗原结合。在某些实施方案中,在缺乏其他可变区的情况下,单个可变区与抗原特异性地结合。

[0066] 在某些实施方案中,CDR的确定性描述和包含抗体之结合位点的残基的鉴定是通过辨析抗体的结构和/或辨析抗体-配体复合物的结构实现的。在某些实施方案中,这可以通过本领域技术人员已知的多种技术中的任一种,例如X射线晶体学来实现。在某些实施方案中,可以采用多种分析方法来鉴定或模拟CDR区。这样的方法的一些实例包括但不限于Kabat定义、Chothia定义、AbM定义和接触定义(contact definition)。

[0067] Kabat定义是对抗体中的残基进行编号的标准,并且通常用于鉴定CDR区。参见例如,Johnson&Wu,Nucleic Acids Res.,28:214-8(2000)。Chothia定义类似于Kabat定义,但是Chothia定义考虑了某些结构环区的位置。参见例如,Chothia et al.,J.Mol.Biol.,196:901-17(1986);Chothia et al.,Nature,342:877-83(1989)。AbM定义使用由Oxford Molecular Group产生的对抗体结构进行建模的集成计算机程序套件。参见例如,Martin et al.,Proc Natl Acad Sci (USA),86:9268-9272(1989);“AbMTM,A Computer Program for Modeling Variable Regions of Antibodies,”Oxford,UK;Oxford Molecular,Ltd。AbM定义使用知识数据库与从头计算方法(ab initio method)(例如由在PROTEINS,Structure,Function and Genetics Suppl.,3:194-198(1999)中的Samudrala et al.,“Ab Initio Protein Structure Prediction Using a Combined Hierarchical Approach,”所述的那些)的组合由一级序列对抗体的三级结构进行建模。接触定义是基于对可用的复杂晶体结构的分析。参见,例如,MacCallum et al.,J.Mol.Biol.,5:732-45(1996)。

[0068] 按照惯例,重链中的CDR区通常被称为H1、H2和H3,并且在从氨基末端到羧基末端的方向上顺序编号。轻链中的CDR区通常被称为L1、L2和L3,并且在从氨基末端到羧基末端的方向上顺序编号。

[0069] 术语“轻链”包括全长轻链及其具有足够可变区序列以赋予结合特异性的片段。全长轻链包含可变区结构域VL和恒定区结构域CL。轻链的可变区结构域位于多肽的氨基末端处。轻链包含 κ 链和 λ 链。

[0070] 术语“重链”包括全长重链及其具有足够可变区序列以赋予结合特异性的片段。全长重链包含可变区结构域VH和三个恒定区结构域CH1、CH2和CH3。VH结构域在多肽的氨基末端处,并且CH结构域在羧基末端处,其中CH3最靠近多肽的羧基末端。重链可以是任何同种型,包括IgG(包括IgG1、IgG2、IgG3和IgG4亚型)、IgA(包括IgA1和IgA2亚型)、IgM和IgE。

[0071] 双特异性或双功能性的抗体通常是具有两个不同重链/轻链对和两个不同结合位点的人工杂交抗体。双特异性抗体可以通过多种方法产生,包括但不限于杂交瘤的融合或Fab'片段的连接。参见例如,Songsivilai et al.,Clin.Exp.Immunol.,79:315-321(1990);Kostelny et al.,J.Immunol.,148:1547-1553(1992)。

[0072] 术语“抗原”是指能够诱导适应性免疫应答的物质。具体地,抗原是充当用于适应性免疫应答的受体的靶标的物质。通常来说,抗原是与抗原特异性受体结合但其本身不能在体内诱导免疫应答的分子。抗原通常是蛋白质和多糖,较低频率也是脂质。如本文中使用的,抗原还包括免疫原和半抗原。

[0073] “Fc”区包含两个含有抗体的CH1和CH2结构域的重链片段。两个重链片段通过两个或更多个二硫键以及通过CH3结构域的疏水相互作用保持在一起。

[0074] “Fv区”包含来自重链和轻链二者的可变区,但缺少恒定区。

[0075] “特异性地结合”特定多肽或特定多肽上的表位或者对其“具有特异性”的抗体是与该特定多肽或特定多肽上的表位结合而基本上不与任何其他多肽或多肽表位结合的抗体。例如,本发明的冠状病毒S蛋白特异性抗体对冠状病毒S蛋白具有特异性。在一些实施方案中,与冠状病毒S蛋白结合的抗体的解离常数(Kd)为 $\leq 100\text{nM}$ 、 $\leq 10\text{nM}$ 、 $\leq 1\text{nM}$ 、 $\leq 0.1\text{nM}$ 、 $\leq 0.01\text{nM}$ 或 $\leq 0.001\text{nM}$ (例如 10^{-8}M 或更小,例如 10^{-8}M 至 10^{-13}M ,例如 10^{-9}M 至 10^{-13}M)。

[0076] 当在竞争相同表位的抗原结合蛋白(例如,抗体或其抗原结合片段)的背景下使用时,术语“竞争”意指如通过这样的测定所确定的抗原结合蛋白之间的竞争:在该测定中,所测试的抗原结合蛋白(例如,抗体或其抗原结合片段)防止或抑制(例如,降低)参考抗原结合蛋白(例如,配体或参考抗体)与共同抗原(例如,冠状病毒S蛋白或其片段)的特异性结合。许多类型的竞争性结合测定可用于确定一种抗原结合蛋白是否与另一种竞争,例如:固相直接或间接放射免疫测定(radioimmunoassay,RIA),固相直接或间接酶免疫测定(enzyme immunoassay,EIA),夹心竞争测定(参见例如,Stahli et al.,1983,Methods in Enzymology 9:242-253);固相直接生物素-亲和素EIA(参见例如,Kirkland et al.,1986,J.Immunol.137:3614-3619)固相直接标记测定,固相直接标记夹心测定(参见例如,Harlow and Lane,1988,Antibodies,A Laboratory Manual,Cold Spring Harbor Press);使用1-125标记的固相直接标记RIA(参见例如,Morel et al.,1988,Molec.Immunol.25:7-15);固相直接生物素-亲和素EIA(参见例如Cheung,et al.,1990,Virology 176:546-552);以及直接标记RIA(Moldenhauer et al.,1990,Scand.J.Immunol.32:77-82)。通常来说,这样的测定涉及以下的使用:与固体表面结合的纯化抗原或携带这些之一的细胞,未经标记的测试抗原结合蛋白和经标记的参考抗原结合蛋白。竞争性抑制是通过在存在测试抗原结合蛋白的情况下确定与固体表面或细胞结合的标记的量来测量的。通常测试抗原结合蛋白是过量存在的。通过竞争测定鉴定的抗原结合蛋白(竞争抗原结合蛋白)包括与参考抗原结合蛋白结合相同表位的抗原结合蛋白,以及与足够接近参考抗原结合蛋白所结合的表位的邻近表位结合以产生空间位阻的抗原结合蛋白。关于用于确定竞争性结合的方法的另外的细节在本文中的实施例中提供。通常,当竞争性抗原结合蛋白过量存在时,其将抑制(例如,降低)参考抗原结合蛋白与共同抗原的特异性结合至少40%至45%、45%至50%、50%至55%、55%至60%、60%至65%、65%至70%、70%至75%或75%或更多。在一些情况下,结合被抑制至少80%至85%、85%至90%、90%至95%、95%至97%、或97%或更多。

[0077] 本文中使用的术语“表位”是指与抗体结合的抗原上特定氨基酸或原子组。表位可以是线性表位或构象表位。线性表位由来自抗原的连续氨基酸序列形成,并基于其一级结构与抗体相互作用。在另一方面,构象表位由抗原的氨基酸序列的不连续部分构成,并基于抗原的3D结构与抗体相互作用。一般来说,表位的长度为约5或6个氨基酸。如果两种抗体对抗原显示出竞争性结合,则它们可以与抗原中的同一表位结合。

[0078] 术语“宿主细胞”意指已经用核酸序列转化或能够用核酸序列转化并因此表达目的基因的细胞。该术语包括亲代细胞的子代,无论子代在形态或遗传组成上与原始亲代细胞是否相同,只要目的基因存在即可。

[0079] 术语“同一性”是指两个或更多个多肽分子或者两个或更多个核酸分子的序列之间的关系,如通过比对和比较序列来确定的。“百分比同一性”意指在所比较的分子中氨基酸或核苷酸之间相同残基的百分比,并且基于所比较的分子中最小者的尺寸进行计算。对于这些计算,比对中的空位(如果有的话)优选地通过特定的数学模型或计算机程序(即“算法”)来解决。可用于计算比对的核酸或多肽的同一性的方法包括在以下中描述的那些:Computational Molecular Biology,(Lesk,A.M.,ed.),1988,New York:Oxford University Press;Biocomputing Informatics and Genome Projects,(Smith,D.W.,ed.),1993,New York:Academic Press;Computer Analysis of Sequence Data,Part I,

(Griffin,A.M.,and Griffin,H.G.,eds.),1994,New Jersey:Humana Press;von Heinje, G.,1987,Sequence Analysis in Molecular Biology,New York:Academic Press; Sequence Analysis Primer, (Gribskov,M.and Devereux,J.,eds.),1991,New York: M.Stockton Press;以及Carillo et al.,1988,SIAM J.Applied Math.48:1073。

[0080] 在计算百分比同一性中,所比较的序列通常以给出序列之间的最大匹配的方式进行比对。可用于确定百分比同一性的计算机程序的一个实例是GCG程序包,其包括GAP (Devereux et al.,1984,Nucl.Acid Res.12:387;Genetics Computer Group,University of Wisconsin,Madison,Wis)。计算机算法GAP用于比对待确定序列百分比同一性的两个多肽或多核苷酸。将序列进行比对,以使它们各自的氨基酸或核苷酸最佳匹配(“匹配跨度”,如由算法确定的)。空位开放罚分(其以 $3 \times$ 平均对角线计算,其中“平均对角线”是所使用的比较矩阵的对角线的平均值;“对角线”是由特定的比较矩阵指定给每个完美氨基酸匹配的分数或数字)和空位延伸罚分(通常是空位开放罚分的1/10倍),以及比较矩阵(例如PAM 250或BLOSUM 62)与该算法一起使用。在某些实施方案中,该算法也使用标准比较矩阵(参见Dayhoff et al.,1978,Atlas of Protein Sequence and Structure5:345-352,对于PAM 250比较矩阵;Henikoff et al.,1992,Proc.Natl.Acad.Sci.U.S.A.89:10915-10919,对于BLOSUM 62比较矩阵)。

[0081] 可以在Needleman et al.,1970,J.Mol.Biol.48:443-453中找到可在使用GAP程序确定多肽或核苷酸序列的百分比同一性中使用的参数的一些实例。

[0082] 用于比对两个氨基酸序列的某些比对方案可导致两个序列的仅短的区域匹配,并且该小的比对区域可具有非常高的序列同一性,即使两个全长序列之间没有显著的关系。因此,如果期望的话,可以调整所选择的比对方法(GAP程序),以得到跨越靶多肽的至少50个或其他数目的连续氨基酸的比对。

[0083] 本文中使用的术语“连接”是指通过分子内相互作用(例如共价键、金属键和/或离子键)或分子间相互作用(例如氢键或非共价键)的缔合。

[0084] 术语“可操作地连接”是指元件的排列,其中如此描述的组件被配置以便执行其通常功能。因此,与多肽可操作地连接的给定信号肽指导从细胞中分泌多肽。在启动子的情况下,与编码序列可操作地连接的启动子将指导编码序列的表达。启动子或其他控制元件不必与编码序列连续,只要它们发挥功能以指导其表达即可。例如,在启动子序列与编码序列之间可以存在插入的未翻译但转录的序列,并且启动子序列仍然可以被认为与编码序列“可操作地连接”。

[0085] 除非明确地指出仅指替代方案或者替代方案相互排斥,否则权利要求书中术语“或/或者”的使用用于意指“和/或”,但是本公开内容支持指仅替代方案和“和/或”的定义。本文中使用的“另一”可意指至少第二个或更多个。

[0086] 术语“多核苷酸”或“核酸”包括单链和双链核苷酸聚合物二者。构成多核苷酸的核苷酸可以是核糖核苷酸或脱氧核糖核苷酸或任一类型核苷酸的经修饰形式。所述修饰包括碱基修饰(例如溴尿苷和肌苷衍生物)、核糖修饰(例如2',3'-双脱氧核糖)和核苷酸间键联修饰(例如硫代磷酸酯、二硫代磷酸酯、硒代磷酸酯、二硒代磷酸酯、苯胺硫代磷酸酯(phosphoroanilothioate)、苯胺基磷酸酯(phoshoraniladate)和氨基磷酸酯)。

[0087] 术语“多肽”或“蛋白质”意指具有天然蛋白质之氨基酸序列的大分子,即由天然存

在和非重组的细胞产生的蛋白质;或者其由经遗传改造或重组的细胞产生,并且包含具有天然蛋白质的氨基酸序列的分子,或者具有天然序列的一个或更多个氨基酸的缺失、添加和/或替换的分子。该术语还包括氨基酸聚合物,其中一个或更多个氨基酸是相应的天然存在氨基酸和聚合物的化学类似物。术语“多肽”和“蛋白质”具体涵盖具有抗原结合蛋白的一个或更多个氨基酸的缺失、添加和/或替换的冠状病毒S蛋白结合蛋白、抗体或序列。术语“多肽片段”是指与全长天然蛋白质相比具有氨基末端缺失、羧基末端缺失和/或内部缺失的多肽。与天然蛋白质相比,这样的片段也可以包含经修饰的氨基酸。在某些实施方案中,片段为约5至500个氨基酸长。例如,片段可以为至少5、6、8、10、14、20、50、70、100、110、150、200、250、300、350、400或450个氨基酸长。可用的多肽片段包括抗体的免疫功能性片段,包括结合结构域。在冠状病毒S蛋白结合抗体的情况下,可用的片段包括但不限于CDR区、重链和/或轻链的可变区、抗体链的一部分或仅其包含两个CDR的可变区,等等。

[0088] 可用于本发明中的可药用载体是常规的。E.W.Martin的Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, PA, 15th Edition (1975) 描述了适合于本文中公开的融合蛋白的药物递送的组合物和制剂。一般而言,载体的性质将取决于所采用的特定施用方式。例如,肠胃外制剂通常包含可注射流体,其包括药理学上和生理学上可接受的流体例如水、生理盐水、平衡盐溶液、右旋糖水溶液、甘油等作为载剂。针对固体组合物(例如,散剂、丸剂、片剂或胶囊剂形式),常规的无毒固体载体可包括例如药用级的甘露醇、乳糖、淀粉或硬脂酸镁。除了生物中性载体之外,待施用的药物组合物可包含少量无毒辅助物质,例如润湿剂或乳化剂、防腐剂和pH缓冲剂等,例如乙酸钠或脱水山梨醇单月桂酸酯。

[0089] 本文中使用的术语“对象”是指人或任何非人动物(例如,小鼠、大鼠、兔、狗、猫、牛、猪、羊、马或灵长类)。人包括出生前和出生后形式。在许多实施方案中,对象是人。对象可以是患者,其指的是为了疾病的诊断或治疗而呈现给医疗提供者的人。本文中使用的术语“对象”可与“个体”或“患者”互换。对象可患有或易患疾病或病症,但可显示或不显示疾病或病症的症状。

[0090] 本文中使用的术语“治疗有效量”或“有效剂量”是指药物有效治疗疾病或病症的剂量或浓度。例如,关于本文中公开的单克隆抗体或其抗原结合片段用于治疗病毒感染的用途。

[0091] 本文中使用的对病症的“治疗”及其变化形式包括预防或减轻病症、减缓病症的发作或病症的发生速度、降低发生病症的风险、预防与病症相关的症状或延缓其发生、减轻或终止与病症相关的症状、引起病症完全或部分消退、治愈病症,或其一些组合。

[0092] 本文中使用的“载体”是指这样的核酸分子,其被引入到宿主细胞中,从而产生经转化的宿主细胞。载体可以包含允许其在宿主细胞中复制的核酸序列,例如复制起点。载体还可以包含一种或更多种治疗性基因和/或选择标记基因以及本领域已知的其他遗传元件。载体可以转导、转化或感染细胞,从而使细胞表达除了细胞原有的核酸和/或蛋白质之外的核酸和/或蛋白质。载体任选地包含有助于实现核酸进入细胞的物质,例如病毒颗粒、脂质体、蛋白质包被等。

[0093] II. 冠状病毒刺突蛋白

[0094] SARS-CoV-2的刺突蛋白在病毒进入到宿主细胞中发挥关键作用,并因此是中和抗

体的主要靶标。刺突蛋白包含N端S1亚基和C端S2亚基,其负责受体结合和膜融合。S1亚基进一步被分为N端结构域、受体结合结构域(receptor-binding domain,RBD)、亚结构域1(subdomain 1,SD1)和亚结构域2(subdomain 2,SD2),并且S2亚基进一步被分为融合肽(fusion peptide,FP)、七肽重复1(heptad repeat 1,HR1)和七肽重复2(heptad repeat 2,HR2)。刺突通过其RBD与细胞受体结合,这引发了刺突的构象变化。经激活的刺突在S1/S2位点被蛋白酶(例如针对SARS-CoV和SARS-CoV-2的TMPRSS2)切割以释放S1亚基并暴露S2亚基上的FP。HR1和HR2重折叠成融合后构象以驱动膜融合³⁵。由于S1的功能性和较高的免疫原性,迄今为止,大多数针对冠状病毒表征的中和抗体靶向S1亚基。主要的挑战是S2构象在膜融合期间是高度动态的,使得难以制备刺突蛋白抗原并产生针对刺突的有效免疫应答(例如,产生中和抗体)。本文中通过刺突蛋白编码序列的突变表明了稳定刺突蛋白的策略。本文中分析并提供的突变在下表1至3中进行了详述。突变体蛋白质如实施例中所所述的表达,并确定了产生的蛋白质和三聚体复合物的量。

[0095]

表1. 刺突蛋白替换和突变。

名称	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
CL-1	T724M	腔	1.3 X (AUC)	99%	
CL-2	T752K	盐桥	<0.5X (SDS)		
CL-3	T778Q	H键	2.6X (AUC)	99%	<10%
CL-4	T961D	盐桥 (原聚体间)	1.8X (AUC)	99%	>99%
CL-5	I1013F	腔	0.8X (AUC)	99%	>99%
CL-6	H1058W	腔	<0.5X (SDS)	N.D.	
CL-7	S735C, T859C	DS	<0.5X (SDS)	N.D.	
CL-8	I770C, A1015C	DS	<0.5X (SDS)	N.D.	
CL-9	L727C, S1021C	DS	<0.5X (SDS)	N.D.	
CL-10	Q901M	腔 (以H键为代价)	0.9X (AUC)	99%	
CL-11	S875F	腔	<0.5X (SDS)	N.D.	
CL-12	T912R	盐桥	<0.5X (SDS)	N.D.	
CL-13	H1088W	腔	0.6X (AUC)	小于 33%	
CL-14	L1141F	腔	0.8X (AUC)	99%	
CL-15	V1040F	腔	<0.5X (SDS)	N.D.	
CL-16	L966D	盐桥	<0.5X (SEC)	N.D.	
CL-17	A766E	盐桥 (原聚体间)	<0.5X (SDS)	N.D.	
CL-18	缺失 (829至851)	移除柔性区域	<0.5X (SDS)	N.D.	
CL-19	T778L	腔	1.5X (AUC)	99%	
CL-20	L938F	腔	2.5X (AUC)	99%	
CL-21	V963L	腔	1.9X (AUC)	99%	
CL-22	V911C, N1108C	DS	0X (SDS)	N.D.	
CL-23	V705C-A893C	DS (原聚体间), 引入N-聚糖	<0.5X (SDS)	N.D.	
CL-24	N703Q/V705C-A893C	DS (原聚体间)	<0.5X (SDS)	N.D.	
CL-25	用GS替换 (673至686)	移除柔性区域	0X (SDS)	N.D.	

[0096]

名称	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
CL-26	用GS替换 (673至686) + A672C至A694C	移除柔性区域, DS (S1-S2)	<0.5X (SDS)	N.D.	
CL-48	A1080C/I1132C	DS	<0.5X (SDS)	N.D.	
CL-58	P862E	盐桥 (S1/S2间)	<0.5X (SDS)	N.D.	
CL-59	T859K	盐桥 (S1/S2间)	2.1X (AUC)	>95%	
CL-60	T547C/N978C	DS (S1/S2间)	0X (SDS)	N.D.	
CL-61	T961C/S758C	DS (原聚体间)	0X (SDS)	N.D.	
CL-62	T961C/Q762C	DS (原聚体间)	0X (SDS)	N.D.	
CL-63	D1118F	电荷移除、 π - π 堆叠	0.5X (SDS)	N.D.	
CL-64	S659C-S698C	DS (S1/S2间)	0.4X (AUC)		
CL-65	缺失HR2	移除柔性HR2 (I161至I208)	2.5X (AUC)	>99%	
CL-66	缺失柄	移除柔性柄区 (I142至I208)	2.6X (AUC)	>99%	
DW-1	R1039F	电荷移除、 π - π 堆叠	0.5X (SDS)	NA	
JM-1	V722C, A930C	二硫化物	0.1X (SDS)	N.D.	
JM-3	A903C, Q913C	二硫化物	2.3X (SDS)	>90%	
JM-6	S974C, D979C	二硫化物	0.3X (SDS)	N.D.	
JM-11	P728C, V951C	二硫化物	0X (SDS)	N/A	
JM-14	V736C, L858C	二硫化物	0X (SDS)	N.D.	
JM-15	S884C, A893C	二硫化物	2X (AUC)	>99%	>99%
JM-18	P807C, S875C	二硫化物	1.1X (AUC)	>99%	
JM-19	T791C, A879C	二硫化物	1.0X (SDS)	>99%	>99%
JM-25	G799C, A924C	二硫化物	1.2X (SDS)	>90%	
CL-49	V826L	腔	1.0X (SDS)	>90%	
CL-50	A899F	腔 (原聚体间)	0.3X (SDS)	N.D.	
CL-51	F817P	脯氨酸	2.3X (SDS)	>95%	>99%

[0097]

名称	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
CL-52	L865P/Q779M	脯氨酸/腔	0.1X (SDS)	N.D.	
CL-35	T866P	脯氨酸	<0.5X (SDS)		
CL-36	A892P	脯氨酸, 腔	1.5X (AUC)	>99%	
CL-37	A899P	脯氨酸, 腔	1.5X (AUC)	>99%	
CL-38	T912P	脯氨酸, 腔	2.5X	51%	
IG-1	A570C/V963C	二硫化物	0X (SDS)	N.D.	
CL-27	T874C, S1055C	DS	0.6X (BCA)	N.D.	
CL-28	L894F	腔 (原聚体间)	0.9X (BCA, AUC)	N.D.	
CL-29	A713S	H键	1X (BCA, AUC)	N.D.	
CL-30	V729C, A1022C	DS	0.4X (BCA), 0.1X (AUC)	N.D.	
CL-31	L828K	盐桥	0.8X (AUC)	>99%	
CL-32	L828R	盐桥	0.8X (BCA), 0.4X (AUC)	nd	
CL-33	H1058F	腔	0X (SDS)	>99%	
CL-34	H1058Y	腔, 可以是H键	0.3X (AUC)	nd	
JM-2	L822C, A1056C	二硫化物			
JM-4	Q965C, S1003C	二硫化物	2X (AUC), 2x (BCA)	100%	
JM-5	A972C, Q992C	二硫化物	约1X (AUC, %三聚体), 1.2x (BCA)	97.80%	
JM-7	I980C, Q992C	二硫化物	1.3X (BCA), 2X (AUC)		
JM-8	A1078C, V1133C	二硫化物	约0.5X (BCA), <1% (AUC)	nd	
JM-9	H1088C, T1120C	二硫化物			
JM-10	I870C, S1055C	二硫化物	约0.37X (BCA), 约1% AUC	nd	
JM-12	T1117C, D1139C	二硫化物	1.9X (BCA), 1X (三聚体 AUC)	75%	
JM-13	T1116C, Y1138C	二硫化物	0X (SDS)	N.D.	
JM-16	I896C, Q901C	二硫化物			
JM-17	G885C, Q901C	二硫化物	1.2X (BCA), 1.1x (AUC),	96%	
JM-20	F1103C, P1112C	二硫化物	0.45X (BCA), 15% (AUC)	nd	
JM-21	G889C, L1034C	二硫化物	0.3X (BCA), 5% (AUC)	nd	

[0098]

名称	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
JM-22	E819C, S1055C	二硫化物			
JM-23	A972C, I980C	二硫化物	1.3X	54%	
JM-24	I1081C, N1135C	二硫化物	0.5X (BCA), 0.3 X (AUC)	nd	
JM-26	E819C, Q1054C	二硫化物	0X (SDS)	nd	
JM-27	Q957E	盐桥 (原聚体间)	1.5X (SDS, BCA)	76%	
CL-44	V1040Y	腔	0.5X (BCA), 0.3X (AUC)	高	
CL-45	H1088Y	腔	1.0X (BCA), 1.6X (AUC)	高	
CL-46	V1104I	腔	0.7X (AUC)	>99%	
CL-47	I1130Y	H键 (原聚体间)	0X (SDS)	>99%	
JM-28	R1000Y	腔填充加与HRI的H键	1X (BCA), 0.3X (AUC)	nd	
JM-29	R1000W	腔填充	0.7x (BCA), 1X (AUC)	nd	
JM-30	A944F	腔填充	1X (BCA), 0.4x (AUC)	nd	
JM-31	A944F, T724I	腔填充	0.9X (BCA),	nd	
JM-32	A944Y	腔填充			
JM-33	S730L	腔填充	0X (SDS)	>99%	
JM-34	S730R	盐桥	0.15X (AUC)	nd	
JM-35	G769E	盐桥	3X (AUC)	高	
CL-53	A893P	脯氨酸	1.5 (SDS)	>95%	
CL-54	Q895P	脯氨酸	2.1 (SDS)	>95%	
CL-55	K921P	脯氨酸	1.1 (SDS)	N.D	
CL-56	L922P	脯氨酸	0.8 (SDS)	N.D	
CL-57	N978P	脯氨酸	0.9 (SDS)	N.D	
CL-39	A942P	脯氨酸	4.0X (AUC)	>99%	>99%
CL-40	G946P	脯氨酸	1.0X (SDS)		
CL-41	S975P	脯氨酸			
CL-42	A890V	腔	1.0X (SDS)		
CL-43	S1003V	腔			

[0099]

表2. 另外的突变。

	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
CL-1	T724M	腔	1.3 X (AUC)	99%	
CL-2	T752K	盐桥	<0.5X (SDS)		
CL-3	T778Q	H键	2.6X (AUC)	99%	<10%
CL-4	T961D	盐桥 (原聚体间)	1.8X (AUC)	99%	>99%
CL-5	I1013F	腔	0.8X (AUC)	99%	>99%
CL-6	H1058W	腔	<0.5X (SDS)	N.D.	
CL-7	S735C, T859C	DS	<0.5X (SDS)	N.D.	
CL-8	I770C, A1015C	DS	<0.5X (SDS)	N.D.	
CL-9	L727C, S1021C	DS	<0.5X (SDS)	N.D.	
CL-10	Q901M	腔 (以H键为代价)	0.9X (AUC)	99%	
CL-11	S875F	腔	<0.5X (SDS)	N.D.	
CL-12	T912R	盐桥	<0.5X (SDS)	N.D.	
CL-13	H1088W	腔	0.6X (AUC)	小于 33%	
CL-14	L1141F	腔	0.8X (AUC)	99%	
CL-15	V1040F	腔	<0.5X (SDS)	N.D.	
CL-16	L966D	盐桥	1.5X (AUC)	99%	
CL-17	A766E	盐桥 (原聚体间)	<0.5X (SDS)	N.D.	
CL-18	缺失 (829至851)	移除柔性区域	<0.5X (SDS)	N.D.	
CL-19	T778L	腔	1.5X (AUC)	99%	
CL-20	L938F	腔	2.5X (AUC)	99%	
CL-21	V963L	腔	1.9X (AUC)	99%	
CL-22	V911C, N1108C	DS	0X (SDS)	N.D.	
CL-23	V705C-A893C	DS (原聚体间), 引入N-聚糖	<0.5X (SDS)	N.D.	
CL-24	N703Q/V705C-A893C	DS (原聚体间)	<0.5X (SDS)	N.D.	
CL-25	用GS替换 (673至686)	移除柔性区域	0X (SDS)	N.D.	

[0100]

	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
CL-26	用GS替换 (673至686) + A672C至A694C	移除柔性区域, DS (S1-S2)	<0.5X (SDS)	N.D.	
CL-48	A1080C/I1132C	DS	<0.5X (SDS)	N.D.	
CL-58	P862E	盐桥 (S1/S2间)	<0.5X (SDS)	N.D.	
CL-59	T859K	盐桥 (S1/S2间)	2.1X (AUC)	>95%	
CL-60	T547C/N978C	DS (S1/S2间)	0X (SDS)	N.D.	
CL-61	T961C/S758C	DS (原聚体间)	0X (SDS)	N.D.	
CL-62	T961C/Q762C	DS (原聚体间)	0X (SDS)	N.D.	
CL-63	D1118F	电荷移除、 π - π 堆叠	0.5X (SDS)	N.D.	
CL-64	S659C-S698C	DS (S1/S2间)	0.4X (AUC)		
CL-65	缺失HR2	移除松散的HR2	2.5X (AUC)		
DW-1	R1039F	电荷移除、 π - π 堆叠	0.5X (SDS)	NA	
JM-1	V722C, A930C	二硫化物	0.1X (SDS)	N.D.	
JM-3	A903C, Q913C	二硫化物	2.3X (SDS)	>90%	
JM-6	S974C, D979C	二硫化物	0.3X (SDS)	N.D.	
JM-11	P728C, V951C	二硫化物	0X (SDS)	N/A	
JM-14	V736C, L858C	二硫化物	0X (SDS)	N.D.	
JM-15	S884C, A893C	二硫化物	2X (AUC)	>99%	>99%
JM-18	P807C, S875C	二硫化物	1.1X (AUC)	>99%	
JM-19	T791C, A879C	二硫化物	1.0X (SDS)	>99%	>99%
JM-25	G799C, A924C	二硫化物	1.2X (SDS)	>90%	
CL-49	V826L	腔	1.0X (SDS)	>90%	
CL-50	A899F	腔 (原聚体间)	0.3X (SDS)	N.D.	
CL-51	F817P	脯氨酸	2.3X (SDS)	>95%	>99%
CL-52	L865P/Q779M	脯氨酸/腔	0.1X (SDS)	N.D.	
CL-35	T866P	脯氨酸	<0.5X (SDS)		
CL-36	A892P	脯氨酸, 腔	1.5X (AUC)	>99%	

[0101]

	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
CL-37	A899P	脯氨酸, 腔	1.5X (AUC)	>99%	
CL-38	T912P	脯氨酸, 腔	2.5X	51%	
JG-1	A570C/V963C	二硫化物	0X (SDS)	N.D.	
CL-27	T874C, S1055C	DS	0.6X (BCA)	N.D.	
CL-28	L894F	腔 (原聚体间)	0.9X (BCA, AUC)	N.D.	
CL-29	A713S	H键	1X (BCA, AUC)	N.D.	
CL-30	V729C, A1022C	DS	0.4X (BCA), 0.1X (AUC)	N.D.	
CL-31	L828K	盐桥	0.8X (AUC)	>99%	
CL-32	L828R	盐桥	0.8X (BCA), 0.4X (AUC)	nd	
CL-33	H1058F	腔	0X (SDS)	>99%	
CL-34	H1058Y	腔, 可以是H键	0.3X (AUC)	nd	
JM-2	L822C, A1056C	二硫化物			
JM-4	Q965C, S1003C	二硫化物	2X (AUC), 2x (BCA)	100%	
JM-5	A972C, Q992C	二硫化物	约1X (AUC, %三聚体), 1.2x (BCA)	97.80%	
JM-7	I980C, Q992C	二硫化物	1.3X (BCA), 2X (AUC)		
JM-8	A1078C, V1133C	二硫化物	约0.5X (BCA), <1% (AUC)	nd	
JM-9	H1088C, T1120C	二硫化物			
JM-10	I870C, S1055C	二硫化物	约0.37X (BCA), 约1% AUC	nd	
JM-12	T1117C, D1139C	二硫化物	1.9X (BCA), 1X (三聚体 AUC)	75%	
JM-13	T1116C, Y1138C	二硫化物	0X (SDS)	N.D.	
JM-16	I896C, Q901C	二硫化物			
JM-17	G885C, Q901C	二硫化物	1.2X (BCA), 1.1x (AUC),	96%	
JM-20	F1103C, P1112C	二硫化物	0.45X (BCA), 15% (AUC)	nd	
JM-21	G889C, L1034C	二硫化物	0.3X (BCA), 5% (AUC)	nd	
JM-22	E819C, S1055C	二硫化物			
JM-23	A972C, I980C	二硫化物	1.3X	54%	
JM-24	I1081C, N1135C	二硫化物	0.5X (BCA), 0.3 X (AUC)	nd	

[0102]

	突变	策略	表达 (产率, % WT AUC)	三聚体的分数	EM (%三聚体)
JM-26	E819C, Q1054C	二硫化物	0X (SDS)	nd	
JM-27	Q957E	盐桥 (原聚体间)	1.5X (SDS, BCA)	76%	
CL-44	V1040Y	腔	0.5X (BCA), 0.3X (AUC)	高	
CL-45	H1088Y	腔	1.0X (BCA), 1.6X (AUC)	高	
CL-46	V1104I	腔	0.7X (AUC)	>99%	
CL-47	I1130Y	H键 (原聚体间)	0X (SDS)	>99%	
JM-28	R1000Y	腔填充与HRI的H键	1X (BCA), 0.3X (AUC)	nd	
JM-29	R1000W	腔填充	0.7x (BCA), 1X (AUC)	nd	
JM-30	A944F	腔填充	1X (BCA), 0.4x (AUC)	nd	
JM-31	A944F, T724I	腔填充	0.9X (BCA),	nd	
JM-32	A944Y	腔填充			
JM-33	S730L	腔填充	0X (SDS)	>99%	
JM-34	S730R	盐桥	0.15X (AUC)	nd	
JM-35	G769E	盐桥	3X (AUC)	高	
CL-53	A893P	脯氨酸	1.5 (SDS)	>95%	
CL-54	Q895P	脯氨酸	2.1 (SDS)	>95%	
CL-55	K921P	脯氨酸	1.1 (SDS)	N.D	
CL-56	L922P	脯氨酸	0.8 (SDS)	N.D	
CL-57	N978P	脯氨酸	0.9 (SDS)	N.D	
CL-39	A942P	脯氨酸	4.0X (AUC)	>99%	>99%
CL-40	G946P	脯氨酸	1.0X (SDS)		
CL-41	S975P	脯氨酸			
CL-42	A890V	腔	1.0X (SDS)		
CL-43	S1003V	腔			

[0103]

表3. 进一步的刺突蛋白突变。

	突变	策略	表达 (产率, % WT AUC)	三聚体的分数
NW-1	R983P	脯氨酸		
NW-2	L984P	脯氨酸	0.4 × (SDS)	NA
NW-3	D985P	脯氨酸	0.2 × (SDS)	NA
NW-4	986P	脯氨酸	2.1 × (SDS)	>90%
NW-5	987P	脯氨酸	1.5 × (SDS)	
δ Strep	移除一个Strep标签		1.2 × (SDS)	
δ TwinStrep	移除两个Strep标签		1 × (AUC)	
JM-36	T1027I	腔	0.9 × (AUC)	
δ RBD	用GGSG替换333至525	移除RBD	0.9 × (AUC)	
δ MPER	以1161为末端	缺失C端		
NW-6	984P,985P	脯氨酸	3 × (SDS)	
NW-7	984P,985G,986P	脯氨酸/甘氨酸	0.5	
NW-8	984P,985G,986G	脯氨酸/甘氨酸	0.5	
NW-9	984G,985P,986G	脯氨酸/甘氨酸	0.8	
NW-10	984G,985P,986P	脯氨酸/甘氨酸	1.2	是
NW-11	984G,985G,986P	脯氨酸/甘氨酸	0.5	
NW-12	984P,985P,986G	脯氨酸/甘氨酸	0.5	

表4 (第一部分). 寡聚物。

	突变	策略	表达 (产率, % WT AUC)	三聚体的分数
NW-1	R983P	脯氨酸		
NW-2	L984P	脯氨酸	0.4 × (SDS)	NA
NW-3	D985P	脯氨酸	0.2 × (SDS)	NA
NW-4	986P	脯氨酸	2.1 × (SDS)	>90%
NW-5	987P	脯氨酸	1.5 × (SDS)	
δ Strep	移除一个Strep标签		1.2 × (SDS)	
δ TwinStrep	移除两个Strep标签		1 × (AUC)	

[0104]

JM-36	T1027I	腔	0.9X(AUC)	
δ RBD	用GSG替换333至525	移除RBD	0.9X(AUC)	
δ MPER	以1161为末端	缺失C端		
NW-6	984P,985P	脯氨酸	3X(SDS)	
NW-7	984P,985G,986P	脯氨酸/甘氨酸	0.5	
NW-8	984P,985G,986G	脯氨酸/甘氨酸	0.5	
NW-9	984G,985P,986G	脯氨酸/甘氨酸	0.8	
NW-10	984G,985P,986P	脯氨酸/甘氨酸	1.2	是
NW-11	984G,985G,986P	脯氨酸/甘氨酸	0.5	
NW-12	984P,985P,986G	脯氨酸/甘氨酸	0.5	

表4（第二部分）.寡聚物。

唯一-#	序列 (5'-3')
0	GGTACCAGA (SEQ ID NO: 3)
P0001	CCGTCTCAGGCCGAGTTCGGTACC (SEQ ID NO: 4)
P0002	GGAAACAAGGCAACTTCAAGACCTGAGAGAATTC (SEQ ID NO: 5)
P0003	GCCGTCGATGTTCTTGAAACACGAAATTC (SEQ ID NO: 6)
P0004	GGGCTGTCTTCCTCTGTAGCTAGC (SEQ ID NO: 7)
P0005	GGTGTAGCGCATGATGCTCTGGCTAGC (SEQ ID NO: 8)
P0006	CAGCTCTGTGCTGAACGATATCCTGTCTAGA (SEQ ID NO: 9)
P0007	GGCTTCTGGAGGGTCCAGTCTAGA (SEQ ID NO: 10)
P0008	GGCCTCGGGATGTATCCGGATC (SEQ ID NO: 11)
P0009	GGGCAGGATCTCTGTTGTGCAAGCTAATTGTAAG (SEQ ID NO: 12)
P0010	CTTTACAATTAGCTGCACAACAGAGATCCTGCC (SEQ ID NO: 13)
P0011	GGATCTTGCCGATGCAAGAATTGAATGG (SEQ ID NO: 14)
P0012	CCAGTTCAATTCTTGTCATCGGCAAGATCC (SEQ ID NO: 15)
P0013	GCCATGCAGATGTGCTATAGATTCAACGGAATCGGCGTGACCTGC AACGTGCTGTATG (SEQ ID NO: 16)
P0014	CATACAGCAGTTGTCAGGTCACGCCGATTCCGTTGAATCTATAGCACATCTGCATGGC (SEQ ID NO: 17)
P0015	CTTCTGGAGGGTCCAGTCTAGACAGGATACAGTTCAGCACAGAGCAGATGGCGCCAAAT (SEQ ID NO: 18)
P0016	CAGAGATCCTGTGCGTGAGCATGACC (SEQ ID NO: 19)

[0107]

表5. SARS-CoV-2变体分类和定义。

变体名称 (Pango 键联)	刺突蛋白替换(参见SEQ ID NO: 2)
B.1.525	Q52R, A67V, V70I, Y144V, E484K, D614G, Q677H, F888L
B.1.526	L5F, T95I, D253G, S477N, E484K, D614G, A701V
B.1.617.1	T95I, E154K, L452R, E484Q, D614G, P681R
B.1.617.2	T19R, L452R, T478K, D614G, P681R, D950N
P.2	E484K, D614G, V1176F
B.1.1.7	H69del, V70del, Y144del, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
B.1.351	D80A, D215G, L242del, A243del, L244del, K417N, E484K, N501Y, D614G, A701V
B.1.427	S13I, W152C, W258L, L452R, D614G
B.1.429	S13I, P26S, W152C, L452R, D614G
P.1	L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F

[0108] III. 药物制剂

[0109] 本公开内容提供了药物组合物,其包含:经改造冠状病毒S蛋白、编码经改造冠状

病毒S蛋白的核酸分子、以及在其基因组物质中包含经改造冠状病毒S蛋白和/或编码经改造冠状病毒S蛋白的病毒载体。这样的组合物可用于刺激免疫应答,例如疫苗制剂的一部分。

[0110] 在编码经改造冠状病毒S蛋白的核酸分子用于药物组合物的情况下,该核酸分子可包含共价连接在一起的脱氧核糖核苷酸和/或核糖核苷酸或其类似物,或者由共价连接在一起的脱氧核糖核苷酸和/或核糖核苷酸或其类似物组成。本文中所述的核酸分子通常包含磷酸二酯键,但是在一些情况下包括核酸类似物,其可具有至少一个不同的键联,例如氨基磷酸酯、硫代磷酸酯、二硫代磷酸酯或O-甲基亚磷酰胺键联,以及肽核酸骨架和键联。可制备天然存在的多核苷酸和类似物的混合物;或者,可制备不同多核苷酸类似物的混合物,以及天然存在的多核苷酸和类似物的混合物。核酸分子可包含经修饰的核苷酸,例如甲基化核苷酸和核苷酸类似物。如果存在的话,可在聚合物组装之前或之后赋予对核苷酸结构的修饰。核苷酸序列可被非核苷酸组分中断。多核苷酸可在聚合之后被进一步修饰,例如通过与标记组分缀合。该术语还包括双链和单链分子二者。除非另有说明或要求,否则术语多核苷酸涵盖以下二者:双链形式和已知或预测构成双链形式的两个互补单链形式中的每一个。核酸分子由以下四种核苷酸碱基的特定序列构成:腺嘌呤(A)、胞嘧啶(C)、鸟嘌呤(G)、胸腺嘧啶(T)和尿嘧啶(U)(当多核苷酸是RNA时代替胸腺嘧啶)。因此,术语“核酸序列”是核酸分子的字母表示。除非另有说明,否则特定的核酸序列还隐含地涵盖其保守修饰的变体(例如简并密码子替换)和互补序列以及明确指出的序列。特别地,简并密码子替换可通过产生其中一个或更多个选定(或所有)密码子的第三位被混合碱基和/或脱氧肌苷残基替换的序列来实现。

[0111] 在一些实施方案中,本公开内容的核酸包含一个或更多个包含经修饰糖部分的经修饰核苷。包含一个或更多个糖被修饰的核苷的这样的化合物可具有期望的特性,例如相对于仅包含含有天然存在的糖部分的核苷的寡核苷酸增强的核酸酶稳定性或提高的与靶核酸的结合亲和力。在一些实施方案中,经修饰糖部分是经取代糖部分。在一些实施方案中,经修饰糖部分是糖替代物。这样的糖替代物可包含一个或更多个与经取代糖部分的那些对应的取代。

[0112] 在一些实施方案中,经修饰糖部分是经取代糖部分,其包含一个或更多个非桥接糖取代基,包括但不限于在2'位和/或5'位的取代基。适合于2'-位的糖取代基的一些实例包括但不限于:2'-F、2'-OCH₃ (“OMe”或“O-甲基”)、和2'-O(CH₂)₂OCH₃ (“MOE”)。在某些实施方案中,在2'位的糖取代基选自:烯丙基、氨基、叠氮基、硫代、O-烯丙基、O--C₁-C₁₀烷基、O--C₁-C₁₀经取代烷基;OCF₃、O(CH₂)₂SCH₃、O(CH₂)₂--O--N(R_m)(R_n)和O--CH₂--C(=O)--N(R_m)(R_n),其中R_m和R_n各自独立地为H或者经取代或未经取代的C₁-C₁₀烷基。在5'-位的糖取代基的一些实例包括但不限于:5'-甲基(R或S)、5'-乙烯基和5'-甲氧基。在一些实施方案中,经取代的糖包含多于一个的非桥接糖取代基,例如T-F-5'-甲基糖部分(对于另外的5',2'-双取代的糖部分和核苷,参见例如PCT国际申请W0 2008/101157)。

[0113] 包含2'-取代的糖部分的核苷被称为2'-取代的核苷。在一些实施方案中,2'-取代的核苷包含选自以下的2'-取代基:卤素、烯丙基、氨基、叠氮基、SH、CN、OCN、CF₃、OCF₃、O、S或N(R_m)-烷基;O、S或N(R_m)-烯基;O、S或N(R_m)-炔基;O-亚烷基-O-烷基、炔基、烷芳基、芳烷基、O-烷芳基、O-芳烷基、O(CH₂)₂SCH₃、O(CH₂)₂--O--N(R_m)(R_n)或O--CH₂--C(=O)--N(R_m)

(Rn), 其中Rm和Rn各自独立地为H、氨基保护基或者经取代或未经取代的C1-C10烷基。这些2'-取代基可进一步被一个或多个独立地选自以下的取代基取代:羟基、氨基、烷氧基、羧基、苄基、苯基、硝基(NO₂)、硫醇、硫代烷氧基(S-烷基)、卤素、烷基、芳基、烯基和炔基。

[0114] 在一些实施方案中,2'-取代的核苷包含选自以下的2'-取代基:F、NH₂、N₃、OCF₃、O--CH₃、O(CH₂)₃NH₂、CH₂—CH=CH₂、O--CH₂—CH=CH₂、OCH₂CH₂OCH₃、O(CH₂)₂SCH₃、O--(CH₂)₂--O--N(Rm)(Rn)、O(CH₂)₂O(CH₂)₂N(CH₃)₂和N-取代的乙酰胺(O--CH₂--C(=O)--N(Rm)(Rn), 其中Rm和Rn各自独立地为H、氨基保护基或者经取代或未经取代的C1-C10烷基。

[0115] 在一些实施方案中,2'-取代的核苷包含含有选自以下的2'-取代基的糖部分:F、OCF₃、O--CH₃、OCH₂CH₂OCH₃、O(CH₂)₂SCH₃、O(CH₂)₂--O--N(CH₃)₂、--O(CH₂)₂O(CH₂)₂N(CH₃)₂和O--CH₂--C(=O)--N(H)CH₃。

[0116] 在一些实施方案中,2'-取代的核苷包含含有选自以下的2'-取代基的糖部分:F、O--CH₃和OCH₂CH₂OCH₃。

[0117] 在一些实施方案中,本公开内容的核苷包含一个或多个未经修饰的核碱基。在某些实施方案中,本公开内容的核苷包含一个或多个经修饰的核碱基。

[0118] 在一些实施方案中,经修饰的核碱基选自:通用碱基、疏水碱基、泛主碱基(promiscuous base)、尺寸扩大的碱基(size-expanded base)和氟化碱基,如本文中所限定。5-取代的嘧啶,6-氮杂嘧啶以及N-2、N-6和O-6取代的嘌呤,包括2-氨基丙基腺嘌呤、5-丙炔基尿嘧啶;5-丙炔基胞嘧啶;5-羟甲基胞嘧啶,黄嘌呤,次黄嘌呤,2-氨基腺嘌呤,腺嘌呤和鸟嘌呤的6-甲基和其他烷基衍生物,腺嘌呤和鸟嘌呤的2-丙基和其他烷基衍生物,2-硫代尿嘧啶,2-硫代胸腺嘧啶和2-硫代胞嘧啶,5-卤代尿嘧啶和胞嘧啶,5-丙炔基CH₃尿嘧啶和胞嘧啶以及嘧啶碱基的其他炔基衍生物,6-偶氮尿嘧啶、胞嘧啶和胸腺嘧啶,5-尿嘧啶(假尿嘧啶),4-硫代尿嘧啶,8-卤、8-氨基、8-硫醇、8-硫代烷基、8-羟基和其他8-取代的腺嘌呤和鸟嘌呤,5-卤(特别是5-溴)、5-三氟甲基和其他5-取代的尿嘧啶和胞嘧啶,7-甲基鸟嘌呤和7-甲基腺嘌呤,2-F-腺嘌呤,2-氨基-腺嘌呤,8-氮杂鸟嘌呤和8-氮杂腺嘌呤,7-脱氮鸟嘌呤和7-脱氮腺嘌呤,3-脱氮鸟嘌呤和3-脱氮腺嘌呤,通用碱基、疏水碱基、泛主碱基、尺寸扩大的碱基和氟化碱基,如本文中所限定。另一些经修饰的核碱基包括三环嘧啶例如吩噻嗪胞苷([5,4-b][1,4]苯并噻嗪-2(3H)-酮)、吩噻嗪胞苷(1H-嘧啶并[5,4-b][1,4]苯并噻嗪-2(3H)-酮)、G-钳(G-clamp)例如经取代的吩噻嗪胞苷(例如,9-(2-氨基乙氧基)-H-嘧啶并[5,4-13][1,4]苯并噻嗪-2(3H)-酮)、吡啶胞苷(2H-嘧啶并[4,5-b]吡啶-2-酮)、吡啶吡啶胞苷(H-吡啶并[3',2':4,5]吡咯并[2,3-d]嘧啶-2-酮)。经修饰的核碱基还可包括其中嘌呤或嘧啶碱基被其他杂环取代的那些,例如7-脱氮-腺嘌呤、7-脱氮鸟苷、2-氨基吡啶和2-吡啶酮。另一些核碱基包括在美国专利3,687,808中公开的那些,在The Concise Encyclopedia Of Polymer Science And Engineering,Kroschwitz,J.I.,Ed.,John Wiley&Sons,1990,858-859中公开的那些;由Englisch et al.,1991公开的那些;以及由Sanghvi,Y.S.,1993公开的那些。

[0119] 教导上述经修饰的核碱基以及其他经修饰的核碱基中某些的制备的代表性美国专利包括但不限于美国专利3,687,808;4,845,205;5,130,302;5,134,066;5,175,273;5,367,066;5,432,272;5,457,187;5,459,255;5,484,908;5,502,177;5,525,711;5,552,

540;5,587,469;5,594,121;5,596,091;5,614,617;5,645,985;5,681,941;5,750,692;5,763,588;5,830,653和6,005,096,其中每个均通过引用整体并入本文。

[0120] 另外的修饰还可以在寡核苷酸的其他位置进行,特别是3'端核苷酸上的糖的3'位以及5'端核苷酸的5'位。例如,本公开内容的经配体缀合的寡核苷酸的一种另外的修饰涉及将一个或更多个另外的非配体部分或缀合物与寡核苷酸化学连接,所述一个或更多个另外的非配体部分或缀合物增强寡核苷酸的活性、细胞分布或细胞摄取。这样的部分包括但不限于脂质部分,例如:胆固醇部分(Letsinger et al.,1989);胆酸(Manoharan et al.,1994);硫醚,例如己基-5-三苯甲基硫醇(Manoharan et al.,1992;Manoharan et al.,1993);硫胆固醇(Oberhauser et al.,1992);脂族链,例如十二烷二醇或十一烷基基团(Saison-Behmoaras et al.,1991;Kabanov et al.,1990;Svinarchuk et al.,1993);磷脂,例如二-十六烷基-rac-甘油或三乙基铵1,2-二-0-十六烷基-rac-甘油-3-H-磷酸酯(Manoharan et al.,1995;Shea et al.,1990);多胺或聚乙二醇链(Manoharan et al.,1995);或金刚烷乙酸(Manoharan et al.,1995);棕榈基部分(Mishra et al.,1995);或者十八胺或己基氨基-羰基-氧基胆固醇部分(Crooke et al.,1996)。在一些方面中,编码经改造冠状病毒S蛋白的核酸分子是经修饰RNA,例如如经修饰mRNA。经修饰(m)RNA考虑了赋予mRNA提高的稳定性和低免疫原性,从而促进治疗上重要的蛋白质的表达的某些化学修饰。例如,就翻译能力而言,N1-甲基-假尿苷(N1m Ψ)优于数种其他核苷修饰及其组合。在一些实施方案中,本文中使用的(m)RNA分子有尿嘧啶可被假尿嘧啶(pseudouracil)(例如1-甲基-3'-假尿苷酰基(1-methyl-3'-pseudouridylyl)碱基)替换。在一些实施方案中,尿嘧啶中的一些被替换,但是在另一些实施方案中,所有尿嘧啶均被替换。(m)RNA可包含5'帽、5'UTR元件、任选地经密码子优化的开放阅读框、3'UTR元件和poly(A)序列和/或多腺苷酸化信号。

[0121] 无论是天然的还是经修饰的核酸分子均可作为裸核酸分子或在递送载剂(例如脂质纳米粒)中递送。脂质纳米粒可包含一种或更多种核酸,其以与脂质纳米粒的重量比为约5:1至约1:100存在。在一些实施方案中,核酸与脂质纳米粒的重量比为约5:1、2.5:1、1:1、1:5、1:10、1:15、1:20、1:25、1:30、1:35、1:40、1:45、1:50、1:60、1:70、1:80、1:90或1:100,或其中可推导出的任何值。

[0122] 在一些实施方案中,本文中使用的脂质纳米粒可包含一种、两种、三种、四种、五种、六种、七种、八种、九种或十种脂质。这些脂质可包括甘油三酯、磷脂、类固醇或固醇、聚乙二醇化脂质、或具有可电离基团的基团(例如烷基胺和一种或更多种疏水基团例如C6或更大的烷基)。

[0123] 在本公开内容的一些方面中,脂质纳米粒与一种或更多种类固醇或类固醇衍生物混合。在一些实施方案中,类固醇或类固醇衍生物包含任何类固醇或类固醇衍生物。如本文中使用的,在一些实施方案中,术语“类固醇”是一类具有四环17碳环状结构的化合物,其可另外包含一个或更多个取代基,包括烷基、烷氧基、羟基、氧代基、酰基,或者两个或更多个碳原子之间的双键。

[0124] 在本公开内容的一些方面中,脂质纳米粒与一种或更多种聚乙二醇化脂质(或PEG脂质)混合。在一些实施方案中,本公开内容包括使用已与PEG基团连接的任何脂质。在一些实施方案中,PEG脂质是甘油二酯,其还包含与甘油基连接的PEG链。在另一些实施方案中,

PEG脂质是包含与具有PEG链的接头基团连接的一个或多个的C6-C24长链烷基或烯基或C6-C24脂肪酸基团的化合物。PEG脂质的一些非限制性实例包括：经PEG修饰的磷脂酰乙醇胺和磷脂酸，经PEG神经酰胺缀合的、经PEG修饰的二烷基胺和经PEG修饰的1,2-二酰氧基丙烷-3-胺，经PEG修饰的二酰基甘油和二烷基甘油。在一些实施方案中，经PEG修饰的二硬脂酰磷脂酰乙醇胺或经PEG修饰的二肉豆蔻酰基-sn-甘油。在一些实施方案中，PEG修饰通过脂质的PEG组分的分子量来测量。在一些实施方案中，PEG修饰的分子量为约100至约15,000。在一些实施方案中，分子量为约200至约500、约400至约5,000、约500至约3,000、或约1,200至约3,000。PEG修饰的分子量为

[0125] 约100,200,400,500,600,800,1,000,1,250,1,500,1,750,2,000,2,250,2,500,2,750,3,000,3,500,4,000,4,500,5,000,6,000,7,000,8,000,9,000,10,000,12,500,至约15,000。可用于本公开内容的脂质的一些非限制性实例通过美国专利5,820,873、W0 2010/141069或美国专利8,450,298教导，其通过引用并入本文。

[0126] 在本公开内容的一些方面中，脂质纳米粒与一种或更多种磷脂混合。在一些实施方案中，任何脂质也包含磷酸基。在一些实施方案中，磷脂是包含一个或两个长链C6-C24烷基或烯基、甘油或鞘氨醇、一个或两个磷酸基、和任选的小有机分子的结构。在一些实施方案中，小有机分子是氨基酸、糖或经氨基取代的烷氧基，例如胆碱或乙醇胺。在一些实施方案中，磷脂是磷脂酰胆碱。在一些实施方案中，磷脂是二硬脂酰磷脂酰胆碱或二油酰磷脂酰乙醇胺。在一些实施方案中，使用了另一些两性离子脂质，其中两性离子脂质限定了带有正电荷和负电荷二者的脂质和脂质样分子。

[0127] 在本公开内容的一些方面中，提供了包含含有亲脂性和阳离子组分的化合物的脂质纳米粒，其中阳离子组分是可电离的。在一些实施方案中，阳离子可电离脂质包含在生理pH下质子化的一种或更多种基团，但在超过8、9、10、11或12的pH下可去质子化且不带电荷。可电离阳离子基团可包含能够在生理pH下形成阳离子基团的一种或更多种可质子化的胺。阳离子可电离脂质化合物还可进一步包含一种或更多种脂质组分，例如两种或更多种具有C6-C24烷基或烯基碳基团的脂肪酸。这些脂质基团可通过酯键连接，或者可通过迈克尔加成(Michael addition)进一步添加至硫原子。在一些实施方案中，这些化合物可以是树枝状聚合物(dendrimer)、树枝状物(dendron)、聚合物，或其组合。

[0128] 在本公开内容的一些方面中，提供了包含含有亲脂性和阳离子组分的化合物的组合物，其中阳离子组分是可电离的。在一些实施方案中，可电离阳离子脂质是指具有可获得电荷(pKa)的氮原子的脂质和脂质样分子。这些脂质在文献中可称为阳离子脂质。这些具有氨基的分子通常具有2至6个疏水链，通常是烷基或烯基，例如C6-C24烷基或烯基，但是可具有至少1个或多于6个尾部。

[0129] 在一些实施方案中，药物组合物中包封的具有核酸分子的脂质纳米粒的量为约0.1%w/w至约50%w/w、约0.25%w/w至约25%w/w、约0.5%w/w至约20%w/w、约1%w/w至约15%w/w、约2%w/w至约10%w/w、约2%w/w至约5%w/w、或约6%w/w至约10%w/w。在一些实施方案中，药物组合物中包封的具有核酸分子的脂质纳米粒的量为约0.1%w/w,0.25%w/w,0.5%w/w,1%w/w,2.5%w/w,5%w/w,7.5%w/w,10%w/w,15%w/w,20%w/w,25%,w/w,30%w/w,35%w/w,40%w/w,45%w/w,50%w/w,55%w/w,60%w/w,65%w/w,70%w/w,75%w/w,80%w/w,85%w/w,90%w/w,至约95%w/w,或者其中可推导出的任何范围。

[0130] 在一些方面中,本公开内容包括被配制成药物组合物的一种或更多种糖。在一些实施方案中,本文中使用的糖是糖类。这些糖类可用于在干燥过程期间充当保护药物组合物免受不稳定的冻干保护剂。这些水溶性赋形剂包括碳水化合物或糖类,例如,二糖(例如蔗糖、海藻糖或乳糖)、三糖(例如果糖、葡萄糖、包含棉子糖的半乳糖)、多糖(例如淀粉或纤维素)、或者糖醇(例如木糖醇、山梨糖醇或甘露糖醇)。在一些实施方案中,这些赋形剂在室温下是固体。糖醇的一些非限制性实例包括赤藓糖醇、苏糖醇、阿糖醇、木糖醇、核糖醇、甘露糖醇、山梨糖醇、半乳糖醇、岩藻糖醇、艾杜糖醇、肌醇、庚七醇、异麦芽酮糖醇、麦芽糖醇、乳糖醇、麦芽三糖醇、麦芽四糖醇或多糖醇。

[0131] 在一些实施方案中,药物组合物中糖的量为约25%w/w至约98%w/w、40%w/w至约95%w/w、50%w/w至约90%w/w、50%w/w至约70%w/w、或约80%w/w至约90%w/w。在一些实施方案中,药物组合物中糖的量为

[0132] 约10%w/w,15%w/w,20%w/w,25%w/w,30%w/w,35%w/w,40%w/w,45%w/w,50%w/w,52.5%w/w,55%w/w,57.5%w/w,60%w/w,62.5%w/w,65%w/w,67.5%w/w,70%w/w,75%w/w,80%w/w,82.5%w/w,85%w/w,87.5%w/w,90%w/w,至约95%w/w,或其中可推导出的任何范围。

[0133] 在一些实施方案中,可药用聚合物是共聚物。可药用聚合物还可包含离散的不同类型的聚合物亚基中的一种、两种、三种、四种、五种或六种亚基。这些聚合物亚基可包括聚氧丙烯、聚氧乙烯或类似的亚基。特别地,可药用聚合物可包含至少一种疏水性亚基和至少一种亲水性亚基。特别地,该共聚物可在疏水性单元的每一侧具有亲水性亚基。该共聚物可具有为聚氧乙烯的亲水性亚基和为聚氧丙烯的疏水性亚基。

[0134] 在一些实施方案中,表达盒被用于表达经改造冠状病毒S蛋白,以用于随后的纯化和递送至细胞/对象,或直接用于基于病毒的递送方法。本文中提供了包含一种或更多种编码经改造冠状病毒S蛋白的核酸的表达载体。

[0135] 表达需要在载体中提供适当的信号,并且包括多种来自病毒和哺乳动物来源二者的驱动经改造冠状病毒S蛋白在细胞中表达的调节元件(例如增强子/启动子)。在本申请通篇,术语“表达盒”意在包括含有编码基因产物的核酸的任何类型遗传构建体,其中核酸编码序列中的部分或全部能够被转录和翻译,即,在启动子的控制下。“启动子”是指由细胞的合成机制或引入的合成机制识别的DNA序列,是启动基因的特异性转录所需的。短语“在转录控制下”意指启动子相对于核酸处于正确的位置和方向中以控制RNA聚合酶启动和基因的表达。“表达载体”意在包括包含在遗传构建体中的能够复制的表达盒,并因此包含一个或更多个复制起点、转录终止信号、poly-A区、选择性标记和多用途克隆位点。

[0136] 术语启动子在此将用于指在RNA聚合酶II的起始位点周围簇集的一组转录控制模块。关于如何组织启动子的大多数想法来源于对数种病毒启动子的分析,包括HSV胸苷激酶(thymidine kinase,tk)和SV40早期转录单元的那些。通过更新近的工作而加强的这些研究已经示出启动子由离散的功能性模块构成,各自由约7至20bp的DNA组成,并含有转录激活蛋白或阻遏蛋白的一个或更多个识别位点。

[0137] 每个启动子中至少一个模块发挥定位RNA合成的起始位点的作用。其最知名的实例是TATA盒(TATA box),但是在缺少TATA盒的一些启动子(例如哺乳动物末端脱氧核糖核苷酸转移酶基因的启动子和SV40晚期基因的启动子)中,覆盖起始位点的离散元件本身有助于

固定起始位置。

[0138] 另外的启动子元件调节转录起始的频率。一般来说,这些位于起始位点上游30至110bp的区域中,但是最近已经示出许多启动子在起始位点下游也包含功能元件。启动子元件之间的间隔通常是柔性的,以使得当元件相对于彼此倒置或移动时保持启动子功能。在tk启动子中,启动子元件之间的间隔在活性开始下降之前可提高至相距50bp。根据启动子,显示出单个元件可合作地或独立地发挥作用以激活转录。

[0139] 在某些实施方案中,病毒启动子例如人巨细胞病毒(cytomegalovirus,CMV)即早期基因启动子、SV40早期启动子、劳斯肉瘤病毒长末端重复序列(Rous sarcoma virus long terminal repeat)、大鼠胰岛素启动子和甘油醛-3-磷酸脱氢酶可用于获得目的编码序列的高水平表达。还考虑使用本领域中公知的另一些病毒或哺乳动物细胞或细菌噬菌体启动子来实现目的编码序列的表达,前提是表达水平对于给定目的是足够的。通过使用具有公知特性的启动子,可优化转染或转化之后目的蛋白质的表达的水平 and 模式。此外,选择响应于特定生理信号而被调节的启动子可允许基因产物的诱导型表达。

[0140] 增强子是提高来自位于同一DNA分子上的远端位置之启动子的转录的遗传元件。增强子的组织很像启动子。换言之,它们由许多单个元件构成,其各自与一个或更多个转录蛋白结合。增强子与启动子之间的基本区别是可操作性。增强子区作为整体必须能够在一定距离处刺激转录;对于启动子区或其组成元件则不需要这样。另一方面,启动子必须具有在特定位点和特定方向上指导RNA合成起始的一个或更多个元件,而增强子缺乏这些特异性。启动子和增强子通常是重叠和连续的,通常看来具有非常类似的模块化组织。

[0141] 以下是可与表达构建体中编码目的基因的核酸组合使用的启动子/增强子和诱导型启动子/增强子的列表。另外地,任何启动子/增强子组合(根据真核启动子数据库(Eukaryotic Promoter Data Base)EPDB)也可用于驱动基因的表达。如果提供合适的细菌聚合酶作为递送复合体的一部分或作为另外的遗传表达构建体,则真核细胞可支持来自某些细菌启动子的胞质转录。

[0142] 启动子和/或增强子可以是例如:免疫球蛋白轻链、免疫球蛋白重链、T细胞受体、HLA DQ α 和/或DQ β 、 β -干扰素、白介素-2、白介素-2受体、MHC II类5、MHC II类HLA-Dra、 β -肌动蛋白、肌肉肌酸激酶(muscle creatine kinase,MCK)、前清蛋白(甲状腺素视黄质运载蛋白(transthyretin))、弹性蛋白酶I、金属硫蛋白(MTII)、胶原酶、白蛋白、甲胎蛋白、t-珠蛋白、 β -珠蛋白、c-fos、c-HA-ras、胰岛素、神经细胞黏附分子(neural cell adhesion molecule,NCAM)、 α 1-胰蛋白酶、H2B(TH2B)组蛋白、小鼠和/或I型胶原蛋白、葡萄糖调节蛋白(GRP94和GRP78)、大鼠生长激素、人血清淀粉样蛋白A(serum amyloid A,SAA)、肌钙蛋白I(troponin I,TN I)、血小板源性生长因子(platelet-derived growth factor,PDGF)、SV40、多瘤、逆转录病毒、乳头状瘤病毒、乙型肝炎病毒、人免疫缺陷病毒、巨细胞病毒(cytomegalovirus,CMV)和长臂猿白血病病毒。

[0143] 在使用cDNA插入物的情况下,通常将期望包含多腺苷酸化信号以实现基因转录物的适当多腺苷酸化。可使用任何多腺苷酸化序列,例如人生长激素和SV40多腺苷酸化信号。还考虑终止子作为表达盒的元件。这些元件可用于增强信息水平并使从盒到其他序列中的通读最小化。

[0144] 存在许多可将表达载体引入到细胞中的方式。在某些实施方案中,表达构建体包

含病毒或来源于病毒基因组的经改造构建体。某些病毒通过受体介导的胞吞作用进入细胞以整合到宿主细胞基因组中并稳定且高效地表达病毒基因的能力已使其成为将外来基因转移到哺乳动物细胞中的有吸引力的候选物。这些对外来DNA序列具有相对低的容纳力,并且具有受限的宿主谱。此外,它们在允许细胞(permissive cell)中的致癌潜力和细胞病变效应引起了安全性问题。它们可仅容纳高至8kB的外来遗传物质,但可容易地被引入到多种细胞系和实验动物中。

[0145] 用于体内递送的一种方法涉及使用腺病毒表达载体。“腺病毒表达载体”意在包括含有足以(a)支持构建体包装和(b)表达其中已经克隆的经改造冠状病毒S蛋白的腺病毒序列的那些构建体。在本发明上下文中,表达不需要合成基因产物。

[0146] 表达载体包括遗传改造形式的腺病毒。对腺病毒(36kB的线性双链DNA病毒)的遗传组织的了解允许用高至7kB的外来序列替换大片段腺病毒DNA。与逆转录病毒相反,宿主细胞的腺病毒感染不导致染色体整合,原因是腺病毒DNA可以以附加体方式复制而没有潜在的遗传毒性。此外,腺病毒在结构上是稳定的,并且在大量扩增之后没有检测到基因组重排。腺病毒可感染几乎所有上皮细胞,而不论其细胞周期阶段如何。迄今为止,腺病毒感染显示出仅与轻度疾病(例如人的急性呼吸系统疾病)相关。

[0147] 腺病毒特别适合用作基因转移载体,原因是其具有中等尺寸的基因组、易于操作、高滴度、宽靶细胞范围和高感染性。病毒基因组的两端均包含100至200个碱基对的反向重复序列(inverted repeat, ITR),其是病毒DNA复制和包装所必需的顺式元件。基因组的早期(E)和晚期(L)区包含不同的转录单元,其被病毒DNA复制的起始分开。E1区(E1A和E1B)编码负责调节病毒基因组和少数细胞基因的转录的蛋白质。E2区(E2A和E2B)的表达导致用于病毒DNA复制的蛋白质的合成。这些蛋白质参与DNA复制、晚期基因表达和宿主细胞关闭。晚期基因的产物(包括大多数病毒衣壳蛋白)仅在由主要晚期启动子(major late promoter, MLP)产生的单一初级转录物的显著加工之后表达。在感染的晚期期间MLP(位于16.8m.u.)特别有效,并且从该启动子产生的所有mRNA具有5'-三联前导(tripartite leader, TPL)序列,这使其成为用于翻译的优选mRNA。在一个系统中,重组腺病毒由穿梭载体与原病毒载体之间的同源重组产生。由于两种原病毒载体之间的可能的重组,因此可从该过程产生野生型腺病毒。因此,从单个噬斑(plaque)分离单个病毒克隆并检查其基因组结构是至关重要的。

[0148] 产生和繁殖复制缺陷型的现有腺病毒载体取决于独特的称为293的辅助细胞系,其通过Ad5 DNA片段从人胚胎肾细胞转化并组成型表达E1蛋白。由于E3区是腺病毒基因组非必要的(dispensable),因此现有腺病毒载体在293细胞的帮助下在E1、D3或这两个区域中携带外来DNA。在自然界中,腺病毒可包装约105%的野生型基因组,提供额外约2kb DNA的容量。加上在E1和E3区中可被替代的约5.5kb的DNA,现有腺病毒载体的最大容量在7.5kb或载体总长度的约15%以下。载体骨架中保留超过80%的腺病毒病毒基因组并且是载体携带的细胞毒性的来源。此外,E1缺失的病毒的复制缺陷是不完全的。

[0149] 辅助细胞系可来源于人细胞,例如人胚胎肾细胞、肌细胞、造血细胞或者其他入胚胎间充质或上皮细胞。或者,辅助细胞可来源于允许人腺病毒的其他哺乳动物物种的细胞。这样的细胞包括例如Vero细胞或其他猴胚胎间充质或上皮细胞。如上所述,优选的辅助细胞系是293。

[0150] 本公开内容的腺病毒是复制缺陷的或至少条件性复制缺陷的。腺病毒可以是42种不同的已知血型或A至F亚组中的任一种。亚组C的5型腺病毒是一种可用于获得用于本公开内容的条件性复制缺陷的腺病毒载体的示例性起始物质。

[0151] 另一些病毒载体可用作本公开内容中的表达构建体。可采用来源于病毒(例如痘苗病毒、腺相关病毒(adeno-associated virus, AAV)和疱疹病毒)的载体。它们为多种哺乳动物细胞提供了数个有吸引力的特征。

[0152] 在一些实施方案、一些具体实施方案中,载体是AAV载体。AAV是感染人和一些其他灵长类物种的小病毒。目前不知道AAV导致疾病。该病毒导致非常温和的免疫应答,使得进一步支持其明显缺乏致病性。在许多情况下,将AAV载体整合到宿主细胞基因组中,这对于某些应用可能是重要的,但也可具有不期望的结果。使用AAV的基因治疗载体可感染分裂细胞和静止细胞二者,并保持在外染色体状态而不整合到宿主细胞的基因组中,但是在天然病毒中存在病毒携带的基因进入到宿主基因组的一些整合。这些特征使AAV成为用于创建用于基因治疗的病毒载体和用于创建等基因人疾病模型的非常有吸引力的候选物。最近在视网膜中使用AAV用于基因治疗的人临床试验显示出了前景。AAV属于Dependoparvovirus属,其继而属于细小病毒(Parvoviridae)科。该病毒是小的(20nm)复制缺陷的、无包膜的病毒。

[0153] 野生型AAV由于许多特征吸引了来自基因治疗研究者的相当大的兴趣。在这些中主要的是病毒明显缺乏致病性。其还可感染非分裂细胞,并且具有在人染色体19的特定位点(命名为AAVS1)处稳定整合到宿主细胞基因组中的能力。该特征使其比存在随机插入和诱变(有时随后是癌症的发生)威胁的逆转录病毒稍微更可预测。AAV基因组最频繁地整合到所述位点中,而随机并入到基因组中以可忽略不计的频率发生。然而,AAV作为基因治疗载体发展通过从载体的DNA中去除rep和cap而消除了这种整合能力。将所期望的基因与驱动基因转录的启动子一起插入反向末端重复序列(ITR)之间,这有助于在通过宿主细胞DNA聚合酶复合物将单链载体DNA转化成双链DNA之后核内多联体形成。基于AAV的基因治疗载体在宿主细胞核中形成附加体多联体。在非分裂细胞中,这些多联体在宿主细胞的生命内保持完整。在分裂细胞中,AAV DNA通过细胞分裂而丢失,因为附加体DNA不与宿主细胞DNA一起复制。AAV DNA随机整合到宿主基因组中是可检测到的,但以非常低的频率发生。AAV还表现出非常低的免疫原性,显示限于产生中和抗体,而它们不诱导明确限定的细胞毒性响应。该特征与感染静止细胞的能力一起表现出其作为用于人基因治疗的载体优于腺病毒的优势。

[0154] AAV基因组是由约4.7千碱基长的正义或反义单链脱氧核糖核酸(single-stranded deoxyribonucleic acid, ssDNA)构建的。基因组包含在DNA链两端的反向末端重复序列(ITR)、以及两个开放阅读框(open reading frame, ORF):rep和cap。前者由编码AAV生命周期所需的Rep蛋白的四个重叠基因构成,并且后者包含衣壳蛋白VP1、VP2和VP3的重叠核苷酸序列,所述衣壳蛋白一起相互作用以形成二十面体对称的衣壳。

[0155] 反向末端重复(ITR)序列各包含145个碱基。它们被如此命名是因为它们的对称性,其显示是AAV基因组有效繁殖所需的。赋予它们该特性的这些序列的特征是它们形成发夹的能力,其有助于允许第二DNA链的不依赖引发酶的合成的所谓的自引发。ITR也显示出是将AAV DNA整合到宿主细胞基因组(人的第19条染色体)中和从其中挽救这两种情况所需

的,以及是使AAV DNA有效衣壳化以及产生完全组装的脱氧核糖核酸酶抗性AAV颗粒所需的。

[0156] 关于基因治疗,ITR看起来是紧挨着以下治疗基因的唯一所需的顺式序列:可以以反式被递送的结构(cap)和包装(rep)蛋白。在该假设下,建立了用于有效产生包含报道子或治疗基因的重组AAV(recombinant AAV,rAAV)载体的许多方法。然而,还公开了ITR不是对于有效复制和衣壳化所需的唯一顺式元件。一些研究组已经在rep基因的编码序列内鉴定了被命名为顺式作用Rep依赖性元件(cis-acting Rep-dependent element,CARE)的序列。CARE显示出当以顺式存在时增强了复制和衣壳化。

[0157] 在一些方面中,本公开内容提供了包含一种或更多种盐的药物组合物。盐可以是无机钾盐或钠盐,例如氯化钾、氯化钠、磷酸氢二钾、磷酸二氢钾、磷酸氢二钠或磷酸二氢钠。药物组合物可包含一种或更多种磷酸盐,如此以产生磷酸缓冲溶液。磷酸缓冲溶液可包含各种磷酸盐,以将溶液缓冲至pH约6.8、6.9、7.0、7.1、7.2、7.3、7.4、7.5、7.6、7.7、7.8、7.9或8.0,或其中可推导出的任何范围。

[0158] 在一些方面中,本公开内容包括被配制成药物组合物的一种或更多种赋形剂。“赋形剂”是指为相对惰性物质的可药用载体,其用于促进将API施用于或递送至对象,或者用于促进将API加工成可在药学上用于递送至对象的作用部位的药物制剂。此外,这些化合物可用作稀释剂以获得易于测量或向患者施用的剂量。赋形剂的一些非限制性实例包括聚合物、稳定剂、表面活性剂、表面改性剂、溶解度增强剂、缓冲剂、包封剂、抗氧化剂、防腐剂、非离子润湿剂或澄清剂、黏度提升剂(viscosity increasing agent)和吸收增强剂(absorption-enhancing agent)。

[0159] 在一个具体实施方案中,术语“可药用”意指由联邦或州政府的监管机构批准或者在美国药典或其他公认的药典中列出用于动物并且更特别地用于人。术语“载体”是指与治疗剂一起施用的稀释剂、赋形剂或载剂。这样的药用载体可以是无菌液体(例如水)并且可优选地包含辅料。当药物组合物通过注射(例如肌肉注射)施用,水是特定的载体。盐水溶液和右旋糖水溶液和甘油溶液也可用作液体载体,特别是用于可注射溶液。另一些合适的药用赋形剂包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、稻、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石、氯化钠、脱脂乳粉、甘油、丙二醇、水、乙醇等。

[0160] 如果期望的话,组合物还可包含少量的润湿剂或乳化剂、或pH缓冲剂。这些组合物可采取溶液剂、混悬剂、乳剂、片剂、丸剂、胶囊剂、散剂、持续释放制剂等形式。经口制剂可包含标准载体,例如药用级的甘露糖醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素、碳酸镁等。合适的试剂的一些实例在“Remington's Pharmaceutical Sciences”中描述。这样的组合物将包含预防或治疗有效量的抗体或其片段(优选以纯化的形式)以及合适量的载体,以便向患者提供用于适当施用的形式。制剂应适合于施用方式,其可以是经口、静脉内、动脉内、颊内、鼻内、雾化、支气管吸入或通过机械通气递送。

[0161] 如本文中所述,本公开内容的经改造蛋白质或编码经改造蛋白质的核酸可配制成用于肠胃外施用,例如,配制成用于通过皮内、静脉内、肌肉内、皮下、瘤内或甚至腹膜内途径注射。或者,制剂可通过表面途径直接施用于黏膜,例如,通过滴鼻剂、吸入或通过雾化器。可药用盐包括酸盐以及与以下形成的盐:无机酸,例如如盐酸或磷酸;或者有机酸,例如乙酸、草酸、酒石酸、扁桃酸等。与游离羧基形成的盐也可来源于:无机碱,例如如氢氧化钠、氢

氧化钾、氢氧化铵、氢氧化钙或氢氧化铁；以及有机碱，例如异丙胺、三甲胺、2-乙氨基乙醇、组氨酸、普鲁卡因等。

[0162] 通常来说，本公开内容的组合物的成分分开提供或以单位剂型混合在一起提供，例如作为在指示活性剂的量的气密密封容器（例如安瓿或药袋（sachette））中的干燥的冻干粉末或无水浓缩物提供。当组合物通过输注施用，其可用含有无菌药用级水或盐水的输液瓶进行分配。当组合物通过注射施用，可提供无菌注射用水或盐水的安瓿以便可在施用之前将成分混合。

[0163] 本公开内容的组合物可被配制为中性或盐形式。可药用盐包括：与阴离子形成的那些，所述阴离子例如来源于盐酸、磷酸、乙酸、草酸、酒石酸等的那些；以及和阳离子形成的那些，所述阳离子例如来源于钠、钾、铵、钙、铁氢氧化物、异丙胺、三乙胺、2-乙基氨基乙醇、组氨酸、普鲁卡因等的那些。

[0164] 剂量可以是单剂量方案或多剂量方案。多剂量可在初次免疫接种方案和/或在加强免疫接种方案中使用。在多剂量方案中，可通过相同或不同的途径给予多种剂量。多剂量将通常间隔至少1周（例如，约2周、约3周、约4周、约6周、约8周、约10周、约12周、约16周等）进行施用。

[0165] 本文中公开的组合物可用于治疗儿童和成人二者。因此，人对象可小于1岁、1至5岁、5至16岁、16至55岁、55至65岁、或至少65岁。

[0166] 优选的一些施用途径包括但不限于肌内、腹膜内、皮内、皮下、静脉内、动脉内和眼内注射。特别优选的一些施用途径包括肌内、皮内和皮下注射。

[0167] IV. 免疫检测方法

[0168] 在另一些实施方案中，本公开内容涉及用于结合、纯化、去除、定量和以其他方式一般地检测冠状病毒S蛋白的免疫检测方法。虽然这样的方法可以在传统意义上应用，但另一种用途将是用于疫苗储液的质量控制和监测，其中根据本公开内容的抗体可用于评估抗原的量或完整性（即长期稳定性）。或者，该方法可用于筛选多种抗体的适当/期望反应性谱。

[0169] 一些免疫检测方法包括酶联免疫吸附测定（enzyme linked immunosorbent assay, ELISA）、放射免疫测定（RIA）、免疫放射测定、荧光免疫测定、化学发光测定、生物发光测定和Western印迹等。特别地，还提供了用于检测和定量冠状病毒S蛋白的竞争性测定。科学文献例如，如Doolittle and Ben-Zeev(1999)、Gulbis and Galand(1993)、De Jager et al.(1993)和Nakamura et al.(1987)中已经描述了多种可用的免疫检测方法的步骤。通常来说，免疫结合方法包括获得怀疑包含冠状病毒S蛋白的样品并根据本公开内容使样品与第一抗体接触，视情况而定，这在有效允许免疫复合物形成的条件下进行。

[0170] 这些方法包括用于从样品中检测或纯化冠状病毒S蛋白或冠状病毒S蛋白的方法。抗体将优选地连接至固体支持物，例如柱基质形式的固体支持物，并且将怀疑包含冠状病毒S蛋白的样品应用于固定化抗体。不需要的成分将从柱上洗掉，使冠状病毒S蛋白表达细胞免疫复合至固定化抗体，其然后通过从柱上去除生物体或抗原来收集。

[0171] 免疫结合方法还包括用于在样品中检测和定量冠状病毒S蛋白或相关组分的量以及检测和定量在结合过程期间形成的任何免疫复合物的方法。在此，将获得怀疑包含冠状病毒S蛋白的样品，并使该样品与结合冠状病毒S蛋白或其组分的抗体接触，随后检测并定

量在特定条件下形成的免疫复合物的量。就抗原检测而言,所分析的生物样品可以是任何怀疑含有冠状病毒S蛋白的样品,例如组织切片或试样,均质化组织提取物,生物流体(例如,鼻拭子)包括血液和血清,或者分泌物例如粪便或尿。

[0172] 使所选生物样品与抗体在有效条件下接触并持续足以允许形成免疫复合物(初级免疫复合物)的时间通常是简单地将抗体组合物添加至样品并将混合物孵育足够长的时间以使抗体与冠状病毒S蛋白形成免疫复合物(即与之结合)的问题。在该时间之后,通常对样品-抗体组合物,例如组织切片、ELISA板、斑点印迹或Western印迹进行洗涤以去除任何非特异性结合的抗体种类,从而允许仅特异性结合在初级免疫复合物中的那些抗体被检测到。

[0173] 通常来说,免疫复合物形成的检测是本领域中公知的,并且可通过应用多种方法来实现。这些方法通常基于标记或标志物(例如那些放射性标签、荧光标签、生物标签和酶标签中的任一种)的检测。有关使用这样的标记的专利包括美国专利3,817,837、3,850,752、3,939,350、3,996,345、4,277,437、4,275,149和4,366,241。当然,如本领域中已知的,通过使用第二结合配体例如第二抗体和/或生物素/抗生物素蛋白配体结合布置,可发现另外的优点。

[0174] 用于检测的抗体可本身与可检测标记连接,其中然后将简单地检测该标记,从而允许确定组合物中初级免疫复合物的量。或者,可通过对第一抗体具有结合亲和力的第二结合配体来检测结合在初级免疫复合物中的该抗体。在这些情况下,第二结合配体可与可检测标记连接。第二结合配体本身通常是抗体,其因此可称为“第二”抗体。使初级免疫复合物与经标记的第二结合配体或抗体在有效条件下接触并持续足以允许形成次级免疫复合物的时间。然后,通常洗涤次级免疫复合物以去除任何非特异性结合的经标记的第二抗体或配体,并且随后检测次级免疫复合物中的剩余标记。

[0175] 另一些方法包括通过两步法检测初级免疫复合物。如上所述,将对抗体具有结合亲和力的第二结合配体(例如抗体)用于形成次级免疫复合物。在洗涤之后,再次使次级免疫复合物与对第二抗体具有结合亲和力的第三结合配体或抗体在有效条件下接触并持续足以允许形成免疫复合物(三级免疫复合物)的时间。第三配体或抗体与可检测标记连接,从而允许检测由此形成的三级免疫复合物。如果期望的话,该系统可提供信号放大。

[0176] 一种免疫检测方法使用两种不同的抗体。将第一生物素化抗体用于检测靶抗原,并且然后将第二抗体用于检测与复合生物素连接的生物素。在该方法中,首先将待测试样品在包含第一步抗体的溶液中孵育。如果存在靶抗原,则一些抗体与抗原结合形成生物素化抗体/抗原复合物。然后通过链霉抗生物素蛋白(或抗生物素蛋白)、生物素化DNA和/或互补生物素化DNA的连续溶液中孵育来扩大抗体/抗原复合物,其中每个步骤向抗体/抗原复合物添加另外的生物素位点。重复扩大步骤直至达到合适的扩大水平,此时将样品在包含针对生物素的第二步抗体的溶液中孵育。该第二步抗体是经标记的,如例如经酶标记的,所述酶可用于使用色原底物通过组织酶学来检测抗体/抗原复合物的存在。经过适当扩大,可产生宏观可见的缀合物。

[0177] 另一种已知的免疫检测方法利用免疫PCR(聚合酶链反应)方法。在与生物素化DNA一起孵育之前,PCR方法类似于Cantor方法,然而,作为使用多轮链霉抗生物素蛋白和生物素化DNA孵育的替代,将DNA/生物素/链霉抗生物素蛋白/抗体复合物用低pH或高盐缓冲液

洗掉,这释放抗体。然后将所得洗涤溶液用于用合适的引物与合适的对照进行PCR反应。至少在理论上,PCR的巨大扩增能力和特异性可用于检测单抗原分子。

[0178] A.ELISA

[0179] 免疫测定在其最简单和直接的意义上是结合测定。某些优选的免疫测定是本领域中已知的多种类型的酶联免疫吸附测定(ELISA)和放射免疫测定(RIA)。使用组织切片的免疫组织化学检测也特别有用。然而,将容易理解的是,检测不限于此类技术,并且也可使用western印迹、斑点印迹、FACS分析等。

[0180] 在一种示例性ELISA中,将本公开内容的抗体固定在表现出蛋白质亲和力的所选表面上,例如聚苯乙烯微量滴定板中的孔上。然后,向孔添加怀疑包含冠状病毒S蛋白的受试组合物。在结合并洗涤以去除非特异性结合的免疫复合物之后,可检测结合的抗原。检测可通过添加与可检测标记连接的另一抗冠状病毒S蛋白抗体来实现。这种类型的ELISA是简单的“夹心ELISA”。检测也可通过添加第二抗冠状病毒S蛋白抗体,随后添加对该第二抗体具有结合亲和力的第三抗体来实现,其中该第三抗体与可检测标记连接。

[0181] 在另一个示例性的ELISA中,将怀疑含有冠状病毒S蛋白(例如,潜在的受感染细胞)的样品固定在孔表面上,并随后使其与本公开内容的抗冠状病毒S蛋白抗体接触。在结合并洗涤以去除非特异性结合的免疫复合物之后,检测结合的抗冠状病毒S蛋白抗体。当初始抗冠状病毒S蛋白抗体与可检测标记连接时,可直接检测免疫复合物。同样,可使用对第一抗冠状病毒S蛋白抗体具有结合亲和力的第二抗体来检测免疫复合物,其中第二抗体与可检测标记连接。

[0182] 不论使用何种形式,ELISA都具有某些共同特征,例如包被、孵育和结合、洗涤以去除非特异性结合的物质以及检测结合的免疫复合物。以下对这些进行描述。

[0183] 在用抗原或抗体包被板中,通常用抗原或抗体的溶液孵育板的孔过夜或持续指定的一段时间。然后对板的孔进行洗涤以去除不完全吸附的物质。然后,将孔的任何剩余可用表面用相对于受试抗血清为抗原性中性的非特异性蛋白质“包被”。这些包括牛血清白蛋白(BSA)、酪蛋白或乳粉溶液。包被允许封闭固定表面上的非特异性吸附位点,并因此降低由抗血清在表面上的非特异性结合引起的背景。

[0184] 在ELISA中,可能更习惯于使用二级或三级检测方法而不是直接操作。因此,在蛋白质或抗体与孔结合,用非反应性物质包被以降低背景,并洗涤以去除未结合的物质之后,使固定表面与待测试的生物样品在有效允许免疫复合物(抗原/抗体)形成的条件下接触。然后,免疫复合物的检测需要经标记的第二结合配体或抗体,以及与经标记的第三抗体或第三结合配体联合的第二结合配体或抗体。

[0185] “在有效允许免疫复合物(抗原/抗体)形成的条件下”意指该条件优选包括用溶液(例如BSA、牛丙种球蛋白(bovine gamma globulin,BGG)或磷酸缓冲盐水(PBS)/吐温)来稀释抗原和/或抗体。这些添加的试剂还倾向于有助于降低非特异性背景。

[0186] “合适的”条件还意指孵育在足以允许有效结合的温度或时间下进行。孵育步骤通常在优选为约25℃至27℃的温度下进行约1至2至4小时左右,或者可在约4℃左右下过夜进行。

[0187] 在ELISA中所有孵育步骤之后,洗涤接触的表面以去除未复合的物质。一个优选的洗涤过程包括用溶液例如PBS/吐温或硼酸盐缓冲液洗涤。在受试样品与最初结合物质之间

形成特异性免疫复合物并随后进行洗涤之后,可确定甚至是微量的免疫复合物的出现。

[0188] 为了提供检测手段,第二或第三抗体具有缔合的标记以允许检测。优选地,这是在与合适的显色底物孵育之后产生显色的酶。因此,例如,期望在有利于发生进一步的免疫复合物形成的时间和条件下使初级和次级免疫复合物与脲酶、葡萄糖氧化酶、碱性磷酸酶或过氧化氢酶缀合的抗体接触或孵育(例如,在室温下在含有PBS的溶液(例如PBS-吐温)中孵育2小时)。

[0189] 在与经标记的抗体孵育并随后进行洗涤以去除未结合的物质之后,对标记的量进行定量,例如通过与显色底物(例如尿素或溴甲酚紫或2,2'-联氨基-二-(3-乙基-苯并噻唑啉-6-磺酸(ABTS)或 H_2O_2 (在过氧化物酶作为酶标记的情况下))一起孵育进行。然后通过测量产生的颜色的程度,例如使用可见光谱分光光度计实现定量。

[0190] B. Western印迹

[0191] Western印迹(或者,蛋白质免疫印迹)是用于在给定的组织匀浆或提取物样品中检测特定蛋白质的分析技术。其使用凝胶电泳通过多肽的长度(变性条件)或通过蛋白质的3-D结构(天然/非变性条件)来分离天然或变性的蛋白质。然后将蛋白质转移到膜(通常是硝酸纤维素或PVDF)上,在膜上使用对靶蛋白具有特异性的抗体对其进行探测(检测)。

[0192] 样品可取自整个组织或来自细胞培养物。在大多数情况下,首先使用搅拌器(用于较大的样品体积)、使用均化器(较小的体积)或通过声处理将实体组织机械破碎。细胞也可通过上述机械方法之一破裂。可使用多种洗涤剂、盐和缓冲液来促进细胞裂解和使蛋白质溶解。通常添加蛋白酶和磷酸酶抑制剂以防止样品被其自身的酶消化。组织制备通常在低温下进行以避免蛋白质变性。

[0193] 使用凝胶电泳分离样品的蛋白质。蛋白质的分离可通过等电点(pI)、分子量、电荷或这些因素的组合来进行。分离的本质取决于样品的处理和凝胶的性质。这是确定蛋白质的非常有用的方法。也可使用二维(two-dimensional, 2-D)凝胶,其将来自单个样品的蛋白质以两个维度散布。在第一维度中根据等电点(蛋白质具有中性净电荷时的pH)分离蛋白质,而在第二维度中根据其分子量分离蛋白质。

[0194] 为了使蛋白质易于进行抗体检测,将其从凝胶中移动到由硝酸纤维素或聚偏二氟乙烯(PVDF)制成的膜上。将膜放置在凝胶的顶部,并在其顶部放置一叠滤纸。将整叠放置在缓冲溶液中,该缓冲溶液通过毛细作用移动至纸上,从而使蛋白质随之移动。用于转移蛋白质的另一种方法称为电印迹,并且使用电流将蛋白质从凝胶拉到PVDF或硝酸纤维素膜中。蛋白质在维持其在凝胶中具有的组织的同时从凝胶中移动到膜上。作为该印迹过程的结果,蛋白质被暴露在薄的表面层上用于检测(参见下文)。选择这两种膜是因为其非特异性的蛋白质结合特性(即,同等良好地结合所有蛋白质)。蛋白质结合基于疏水性相互作用以及膜与蛋白质之间的带电荷相互作用。硝酸纤维素膜比PVDF便宜,但更加易脆而不能良好地经受反复探测。蛋白质从凝胶转移到膜上的均匀性和整体有效性可通过用考马斯亮蓝或丽春红S(Ponceau S)染料对膜进行染色来检查。一旦转移,就使用经标记的第一抗体或未经标记的第一抗体并随后使用经标记蛋白A或与第一抗体的Fc区结合的第二经标记抗体进行间接检测来检测蛋白质。

[0195] C. 免疫组织化学

[0196] 本公开内容的抗体还可与新鲜冷冻和/或福尔马林固定石蜡包埋的组织块二者联

合使用,所述组织块为了通过免疫组织化学(immunohistochemistry, IHC)进行研究而制备。由这些颗粒状试样制备组织块的方法已成功用于多种预后因子的先前IHC研究中,并且是本领域技术人员公知的(Brown et al., 1990; Abbondanzo et al., 1990; Allred et al., 1990)。

[0197] 简言之,冷冻切片可通过以下来制备:在室温下在小塑料囊中的磷酸缓冲盐水(PBS)中将50ng冷冻的“粉碎”组织再水合;通过离心使颗粒沉淀;将其重悬在黏性包埋介质(OCT)中;将囊倒置和/或通过离心再次沉淀;在-70℃异戊烷中速冻;切割塑料囊和/或取出冷冻的组织圆柱状物;将组织圆柱状物固定在低温恒温切片机卡盘上;和/或从囊上切下25至50个连续切片。或者,可将整个冷冻的组织样品用于连续切片切割。

[0198] 可通过类似的方法制备永久切片,其包括在塑料微量离心管中将50mg样品再水合;沉淀;重悬于10%福尔马林中固定4小时;洗涤/沉淀;重悬于温热的2.5%琼脂中;沉淀;在冰水中冷却以使琼脂硬化;从管中取出组织/琼脂块;将块浸入和/或包埋在石蜡中;和/或切割多至50个连续的永久切片。同样,整个组织样品可替换。

[0199] D. 免疫检测试剂盒

[0200] 在另一些实施方案中,本公开内容涉及与上述免疫检测方法一起使用的免疫检测试剂盒。由于抗体可用于检测冠状病毒S蛋白,因此该抗体可包括在试剂盒中。因此,免疫检测试剂盒在合适的容器装置中包含与冠状病毒S蛋白结合的第一抗体和任选地免疫检测试剂。

[0201] 在某些实施方案中,抗体可与固体支持物例如柱基质和/或微量滴定板的孔预先结合。试剂盒的免疫检测试剂可采取多种形式中的任一种,包括与给定抗体缔合或连接的那些可检测标记。也考虑了与第二结合配体缔合或连接的可检测标记。一些示例性的第二配体是对第一抗体具有结合亲和力的那些第二抗体。

[0202] 用于本发明试剂盒的其他合适的免疫检测试剂包括双组分试剂,其包含对第一抗体具有结合亲和力的第二抗体,以及对第二抗体具有结合亲和力的第三抗体,所述第三抗体与可检测标记连接。如上所述,许多示例性的标记在本领域中是已知的,并且所有这样的标记均可结合本公开内容使用。

[0203] 试剂盒还可包含适当等分的冠状病毒S蛋白组合物,无论是标记的还是未标记的,如可用于制备用于检测测定的标准曲线。试剂盒可包含以完全缀合形式、以中间体形式或作为将由试剂盒使用者进行缀合的独立部分的抗体-标记缀合物。试剂盒的组分可包装在水性介质中或以冻干形式包装。

[0204] 试剂盒的容器装置通常包括至少一个小瓶、试管、烧瓶、瓶、注射器或其他容器装置,其中可放置抗体,或者优选将抗体适当地等分。本公开内容的试剂盒通常还包括用于封闭限制式容纳抗体、抗原和任何其他试剂容器以用于商业销售的装置。这样的容器可包含其中保持期望小瓶的注塑或吹塑的塑料容器。

[0205] E. 流式细胞术和FACS

[0206] 本公开内容的抗体也可用于流式细胞术或FACS。流式细胞术是用于许多检测测定的基于激光或阻抗的技术,包括细胞计数、细胞分选、生物标志物检测和蛋白质改造。该技术将细胞悬浮在流体流中,并使其通过电子检测装置,该装置允许同时对多达数千个颗粒/秒的物理和化学特征进行多参数分析。流式细胞术常规用于诊断病症,特别是血癌,但是在

基础研究、临床实践和临床试验中也有许多其他应用。

[0207] 荧光激活细胞分选 (fluorescence-activated cell sorting, FACS) 是专业型的细胞术。它提供了用于根据每个细胞的特定光散射和荧光特征, 将生物细胞的异质混合物分选到两个或更多个容器 (每次一个细胞) 中的方法。一般而言, 该技术涉及在狭窄、快速流动的液体流的中心夹带细胞悬液。布置流动, 使得细胞之间相对于它们的直径存在大的间隔。振动机制使细胞流破碎成单独的微滴。就在流体破裂成微滴之前, 流通过荧光测量站, 在那里测量每个细胞的荧光。将充电环恰好置于水流破裂成微滴的位置。在紧邻荧光强度测量之前, 使电荷置于环上, 而相反的电荷则在微滴破裂形成水流时被捕获。然后带电荷的微滴通过静电偏转系统落下, 该系统基于其电荷将微滴转移到容器中。

[0208] 在某些实施方案中, 为了用于流式细胞术或 FACS, 本公开内容的抗体用荧光团标记, 并随后允许其与目的细胞结合, 所述细胞在流式细胞仪中进行分析或由 FACS 设备进行分选。

[0209] V. 实施例

[0210] 包括以下实施例以说明本发明的一些优选实施方案。本领域技术人员应理解, 以下实施例中公开的技术表示本发明人发现的在本发明的实践中运行良好的技术, 并因此可认为构成用于其实践的优选模式。然而, 根据本公开内容, 本领域技术人员应理解, 在不脱离本发明的精神和范围的情况下, 可在所公开的一些具体实施方案中进行许多改变, 并且仍然获得相同或类似的结果。

[0211] 实施例1-SARS-CoV-2刺突蛋白的改造

[0212] A. 方法

[0213] 针对融合前稳定的 nCoV 刺突变体的设计方案。将 SARS-CoV-2S-2P 变体用作所有后续设计的基础构建体²⁰。在哺乳动物表达质粒 pAH 中, S-2P 基础构建体包含: 在第 986 和 987 位残基处具有脯氨酸替换、在弗林蛋白酶切割位点 (第 682 至 685 位残基) 处具有“GSAS”替换的 SARS-CoV-2S (GenBank: MN908947) 的第 1 至 1208 位残基, T4 纤维蛋白的三聚折叠基序, HRV3C 蛋白酶识别位点, Twin-Strep 标签和八组氨酸标签。使用该质粒作为模板, 在 SARS-CoV-2S 基因中的所选位置处引入所期望的突变。基于 cryo-EM 结构, 考虑将 β -碳原子间隔小于 $4.6 \text{ \AA}$ 的残基对用于二硫化物设计。在融合前至融合后转变期间剧烈移动的区域特定地靶向例如 FP、CR 和 HR1。盐桥变体需要突变残基的带电荷基团预测在 $4.0 \text{ \AA}$ 以内。对于环中的残基, 允许距离比 $4.0 \text{ \AA}$ 稍长。为了较大的腔填充残基的潜在替换, 对具有与预先存在的内部腔相连的侧链的面向核心的残基进行了检查。在 FP、CR 或 HR1 中设计了脯氨酸替换, 并将其置于柔性环中或螺旋的 N 端处。所有期望的替换最大化地避免了: 1. 与邻近残基产生冲突, 2. 产生较大的疏水腔, 3. 失去与邻近残基的极性相互作用, 4. 不利的二面角, 5. 失去现有的或产生新的 N-糖基化位点。选择组合以测试相同类型的设计 (即二硫化物/二硫化物) 或不同类型的设计 (例如二硫化物/脯氨酸) 的对是否可导致针对刺突表达和稳定性的累加作用。避免了预测潜在地彼此干扰的替换, 例如可限制参与二硫键形成的相邻半胱氨酸的柔性的脯氨酸。

[0214] 蛋白质表达和纯化。使用聚乙烯亚胺将编码 S 变体的质粒瞬时转染到 FreeStyle293F 细胞 (Thermo Fisher) 中。在转染之后 4 天, 收获培养物并通过离心从细胞中

分离培养基。使上清液通过0.22 μ m过滤器,然后通过StrepTactin树脂(IBA)。在由2mM Tris pH 8.0、200mM NaCl和0.02% NaN₃构成的缓冲液中,使用Superose 6 10/300柱(GE Healthcare)通过尺寸排阻色谱进一步纯化S变体。对于初始纯化和表征,从40mL培养物中纯化单替换S变体和Combo S变体。从2L培养物中纯化HexaPro变体。

[0215] 差示扫描荧光法。在96孔qPCR板中,用最终浓度为5 $\times$ SYPRO Orange Protein Gel Stain(ThermoFisher)和0.25mg/mL刺突来制备溶液。使用Roche LightCycler 480II使用4.4 $^{\circ}$ /分钟的温度渐变速率(ramp rate)从25 $^{\circ}$ C升至95 $^{\circ}$ C来进行连续荧光测量(λ_{ex} =465nm, λ_{em} =580nm)。数据绘制为解链曲线的导数。

[0216] 负染色EM。将经纯化的2019-nCoV S变体在2mM Tris pH 8.0、200mM NaCl和0.02% NaN₃中稀释至0.04mg/mL的浓度。将各蛋白质沉积在已在Solarus 950等离子体清洗器(Gatan)中用4:1比例的O₂/H₂进行了30秒等离子体清洗的CF-400-CU网格(Electron Microscopy Sciences)上并使用甲胺钨酸盐(Nanoprobes)进行染色。在配备有Ceta 16M检测器(Thermo Fisher Scientific)的Talos F200C TEM显微镜中,以92,000 $\times$ 的放大率(对应于1.63 Å/pix的校准像素尺寸)对网格进行成像。S-2P和HexaPro的稳定性实验通过以下进行:在用液氮快速冷冻和解冻3轮之后、在室温下储存样品1至2天之后、或者在50 $^{\circ}$ C、55 $^{\circ}$ C或60 $^{\circ}$ C下孵育30分钟之后对如上样品进行成像。

[0217] 用于量化蛋白质表达的生物层干涉术。将FreeStyle293F细胞(Thermo Fisher)在3mL基本培养基中进行转染,并在转染之后4天收获。在离心之后,将上清液用由10mM HEPES pH 7.5、150mM NaCl、3mM EDTA、0.05%吐温20和1mg/mL牛血清白蛋白构成的缓冲液稀释5倍。使用Octet RED96e(FortéBio)将抗foldon IgG固定至抗人捕获(anti-human capture, AHC)传感器尖端(sensortip)(FortéBio)。将传感器尖端浸入含有单一刺突变体的孔中。通过测量浓度范围为10 μ g/mL至0.16 μ g/mL的经纯化S-2P的2倍系列稀释液来确定标准曲线。使用Octet Data Analysis软件v11.1将数据减去参考、在IgG捕获之后与基线进行比对并基于各相关曲线的初始斜率的线性拟合进行定量。

[0218] 表面等离子体共振。使用Biacore X100(GE Healthcare)和由10mM HEPES pH 8.0、150mM NaCl和0.05%吐温20构成的运行缓冲液,将经His标记的HexaPro固定至NiNTA传感器芯片(sensorchip)(GE Healthcare)至约800响应单位(response unit, RU)的水平。以250至15.6nM的浓度范围注射经纯化hACE2的系列稀释液。使用Biacore X100评价软件(GE Healthcare)将响应曲线拟合为1:1结合模型。

[0219] Cryo-EM样品制备和数据收集。将经纯化HexaPro在2mM Tris pH 8.0、200mM NaCl、0.02% NaN₃中稀释至0.35mg/mL的浓度,并且施加至经等离子体清洗的CF-400 1.2/1.3网格,随后在Vitrobot Mark IV(ThermoFischer)中印迹6秒,并冷冻投入液体乙烷中。使用配备有K3直接电子检测器(Gatan)的FEI Titan Krios(ThermoFischer)从单个网格收集显微照片。以对应于1.08 Å/pix的校准像素尺寸的46,296 $\times$ 放大率收集数据。数据收集参数的完整描述可见于表S5。

[0220] Cryo-EM数据处理。在Warp中进行运动校正、CTF估计和颗粒拾取³³。然后将颗粒导入cryoSPARC v2.15.0中用于2D分类、从头计算3D重构、异质3D细化以及非均匀同质细化³⁴。以Coot、Phenix和ISOLDE进行迭代模型构建和细化³⁵⁻³⁷。

[0221] B. 结果

[0222] 融合前稳定的SARS-CoV-2刺突的基于结构的设计。为了产生比原始S-2P变体²⁰更好表达且更稳定的融合前稳定的SARS-CoV-2刺突蛋白,分析了SARS-CoV-2S-2P cryo-EM结构(PDB:6VSB),并且基于I类融合蛋白功能和一般蛋白质稳定性原理的知识设计了替换物。这些策略包括以下:在融合前状态下引入二硫键以阻止在融合前至融合后转变期间的构象变化、引入盐桥以中和电荷不平衡、引入疏水残基以填充内部腔、以及引入脯氨酸以封闭螺旋或稳定环。克隆了100种单一S-2P变体并表征了其相对表达水平,并且对于表达良好的那些,表征了它们的单分散性、热稳定性和四级结构(表6)。考虑到S2亚基在融合前至融合后转变期间经历了大规模的重折叠,努力集中在稳定S2上。鉴定了各类别的替换,其在维持融合前构象的同时提高了表达(图1和2A)。总的来说,100种变体中有28种与S-2P相比表达更好并保留了融合前构象,如通过负染色EM所评估的。

[0223] 表6

[0224]

突变	策略	相对于S-2P的表达倍数变化
T547C, N978C	二硫化物	0 ^b
A570C, V963C	二硫化物	0 ^b
S659C, S698C	二硫化物	0.4 ^a
用GS替换(673至686)	移除柔性区域	0 ^b
用GS替换(673至686) + A672C, A694C	二硫化物, 移除柔性区域	<0.5 ^b
N703Q, V705C, A893C	二硫化物	<0.5 ^b
V705C, A893C	二硫化物	<0.5 ^b
A713S	H键	1.0 ^a
V722C, A930C	二硫化物	<0.1 ^b
T724M	腔填充	1.3 ^a
L727C, S1021C	二硫化物	<0.5 ^b
P728C, V951C	二硫化物	0 ^b
V729C, A1022C	二硫化物	<0.1 ^a
S730L	腔填充	0 ^b
S730R	盐桥	0.15 ^a
S735C, T859C	二硫化物	<0.5 ^b
V736C, L858C	二硫化物	0 ^b
T752K	盐桥	<0.5 ^b
A766E	盐桥	<0.5 ^b
G769E	盐桥	3.0 ^a
I770C, A1015C	二硫化物	<0.5 ^b
T778Q	氢键	2.6 ^a
T778L	腔填充	1.5 ^a
T791C, A879C	二硫化物	1.0 ^b
G799C, A924C	二硫化物	1.3 ^a
P807C, S875C	二硫化物	1.1 ^a

[0225]

突变	策略	相对于S-2P的 表达倍数变化
F817P	脯氨酸	2.9 ^a
E819C, S1055C	二硫化物	0 ^b
E819C, Q1054C	二硫化物	0 ^b
L822C, A1056C	二硫化物	0 ^b
V826L	腔填充	1.0 ^b
L828K	盐桥	0.8 ^a
L828R	盐桥	0.4 ^a
Δ(829-851)	移除柔性区域	<0.5 ^b
T859K	盐桥	2.1 ^a
P862E	盐桥	<0.5 ^b
L865P, Q779M	脯氨酸, 腔填充	<0.5 ^b
T866P	脯氨酸	<0.5 ^b
I870C, S1055C	二硫化物	0 ^b
T874C, S1055C	二硫化物	<0.5 ^b
S875F	腔填充	<0.5 ^b
S884C, A893C	二硫化物	2 ^a
G885C, Q901C	二硫化物	1.1 ^a
G889C, L1034C	二硫化物	<0.1 ^a
A890V	腔填充	1.0 ^b
A892P	脯氨酸, 腔填充	1.5 ^a
A893P	脯氨酸	1.5 ^b
L894F	腔填充	0.9 ^a
Q895P	脯氨酸	2.1 ^b
I896C, Q901C	二硫化物	0 ^b
A899F	腔填充	0.3 ^b
A899P	脯氨酸, 腔	1.5 ^a
Q901M	腔填充	0.9 ^a
A903C, Q913C	二硫化物	2.3 ^b
V911C, N1108C	二硫化物	0 ^b
T912R	盐桥	<0.5 ^b
T912P	脯氨酸, 腔填充	2.5 ^a
K921P	脯氨酸	1.1 ^b
L922P	脯氨酸	0.8 ^b
L938F	腔填充	2.5 ^a
A942P	脯氨酸	4.0 ^a
A944F	腔填充	1.0 ^a
A944F, T724I	腔填充	0.4 ^a
A944Y	腔填充	1.9 ^b
G946P	脯氨酸	1.0 ^b
Q957E	盐桥	1.0 ^a
T961D	盐桥	1.8 ^a
T961C, S758C	二硫化物	0 ^b
T961C, Q762C	二硫化物	0 ^b
V963L	腔填充	1.9 ^a
Q965C, S1003C	二硫化物	2 ^a

[0226]

突变	策略	相对于S-2P的表达倍数变化
A972C, Q992C	二硫化物	1 ^a
A972C, I980C	二硫化物	1.3 ^a
S974C, D979C	二硫化物	0.3 ^b
S975P	脯氨酸	2.2 ^b
N978P	脯氨酸	0.9 ^b
I980C, Q992C	二硫化物	2.0 ^a
R1000Y	腔填充+氢键	0.3 ^a
R1000W	腔填充	1.0 ^a
S1003V	腔填充	1.9 ^b
I1013F	腔填充	0.8 ^a
R1039F	电荷移除、 π - π 堆叠	0.5 ^b
V1040F	腔填充	<0.5 ^b
V1040Y	腔填充	0.3 ^a
H1058W	腔填充	<0.5 ^b
H1058F	腔填充	0 ^b
H1058Y	腔填充	0.3 ^a
A1078C, V1133C	二硫化物	<0.5 ^b
A1080C, I1132C	二硫化物	<0.5 ^b
I1081C, N1135C	二硫化物	0.3 ^a
H1088Y	腔填充	1.6 ^a
H1088W	腔填充	0.6 ^a
F1103C, P1112C	二硫化物	0.15 ^a
V1104I	腔填充	0.7 ^a
T1116C, Y1138C	二硫化物	0 ^b
T1117C, D1139C	二硫化物	1.0 ^a
D1118F	电荷移除、 π - π 堆叠	0.5 ^b
I1130Y	氢键	0 ^b
L1141F	腔填充	0.8 ^a
Δ HR2 (Δ 1161-1208)	移除柔性区域	2.5 ^a

[0227] ^a使用尺寸排阻三聚体峰的曲线下面积进行定量[0228] ^b使用SDS-PAGE条带强度进行定量

[0229] 单替换突变体。虽然二硫键的引入需要两次替换,但是出于这些目的,它们被认为是单替换。稳定I类融合蛋白(例如刺突)的一种常用策略是通过二硫键将经历构象变化的区域与不经历构象变化的区域共价连接。例如,Q965C/S1003C替换试图将HR1与中心螺旋连接,而G799C/A924C旨在将HR1与上游螺旋连接。与S-2P相比,这两种变体分别将蛋白质表达提高了3倍和2倍(图2B)。然而,与S-2P相比,两种变体的尺寸排阻色谱(size-exclusion chromatography, SEC)迹线显示出向左移动,表明正在运行的蛋白质比预期的更大,这与显示出部分错误折叠的刺突颗粒的负染色电子显微术(negative stain electron microscopy, nsEM)结果良好吻合。相比之下,S884C/A893C和T791C/A879C变体在SEC上以类似于S-2P的体积洗脱,并且通过nsEM显示为良好折叠的三聚体颗粒(图2E)。这些变体将同一 α -螺旋与两个不同的柔性环连接,所述柔性环针对相邻的原聚体进行包装(图1)。值得注

意的是, S884C/A893C与S-2P相比具有两倍高的表达, 并且还提高了热稳定性(图2F)。

[0230] 选择腔填充替换和盐桥的引入应改善蛋白质的稳定性, 而不干扰整体折叠。腔填充特别地有助于稳定RSV F和HIV-1Env的融合前构象^{15,22}。在此, 发现与S-2P相比, 许多腔填充和盐桥设计改善了蛋白质表达(图2G)。例如, L938F和T961D二者具有蛋白质产率的约2倍提高, 并且还维持了正确的刺突四级结构(图2C和2E), 但是如通过差示扫描荧光法(DSF)评估的两种变体的热稳定性与S-2P保持相同(图2F)。

[0231] 受先前使用脯氨酸替换的成功的启发, 尝试了14种单独变体, 其中将脯氨酸替换到融合肽(fusion peptide, FP)、连接区(connector region, CR)和HR1中的柔性环或螺旋N端中(表1和图1G)。如所预期的, 多脯氨酸变体加强了蛋白质表达并提高了热稳定性(图2D和2F)。与S-2P相反, 两个最成功的替换F817P和A942P分别表现出蛋白质产率的2.3倍和4倍提高。A942P替换还将解链温度(melting temperature, T_m)提高了约3°C, 并且两种变体通过nsEM均显示为良好折叠的三聚体(图2E)。

[0232] 多替换刺突变体。为了检查有益的单替换的潜在累加或协同作用, 在以下考虑下产生了初始组合(“Combo”)变体: 替换在空间上不应彼此接近, 并且每个构建体至多两个二硫键。包含两个二硫键的Combo变体通常具有单二硫化物变体的约50%的表达水平, 这表明两个替换彼此干扰(表2)。向单脯氨酸变体(F817P)添加一个二硫化物(S884C/A893C)也降低了表达水平, 但是刺突的四级结构维持良好(表7)。二硫键的有益作用在与腔填充变体L938F组合时最显著。Combo23(S884C、A893C、L938F)与其任一亲本变体相比具有更高的蛋白质产率, 但是与S884C/A893C相比, Combo23的T_m没有进一步提高。另外, 与单独的L938F相比, 将两个腔填充替换组合(Combo18)或将一个腔填充替换与一个脯氨酸替换混合(Combo20)提高了表达(表7)。

[0233] 表7。

[0234]

Combo #	突变	策略	相对于S-2P 的表达的倍 数变化
Combo1	A903C, Q913C, Q965C, S1003C	二硫化物+二硫化物	3.3
Combo2	S884C, A893C, A903C, Q913C	二硫化物+二硫化物	N.D.
Combo3	T791C, A879C, A903C, Q913C	二硫化物+二硫化物	N.D.
Combo4	G799C, A924C, A903C, Q913C	二硫化物+二硫化物	N.D.
Combo8	T791C, A879C, S884C, A893C	二硫化物+二硫化物	0.5
Combo9	G799C, A924C, S884C, A893C	二硫化物+二硫化物	0.4
Combo11	A892P, A899P	脯氨酸+脯氨酸	1.9
Combo12	A892P, T912P	脯氨酸+脯氨酸	2.7
Combo14	A892P, A942P	脯氨酸+脯氨酸	6.2
Combo16	A899P, A942P	脯氨酸+脯氨酸	5.1
Combo18	L938F, V963L	腔填充+腔填充	1.9
Combo19	L938F, A892P	腔填充+脯氨酸	3.0
Combo20	L938F, A899P	腔填充+脯氨酸	3.0
Combo21	F817P, L938F	脯氨酸+脯氨酸	3.9
Combo22	L938F, A942P	腔填充+脯氨酸	6.0
Combo23	S884C, A893C, L938F	二硫化物+腔填充	2.9
Combo24	T791C, A879C, L938F	二硫化物+腔填充	2.2
Combo26	L938F, A903C, Q913C	腔填充+二硫化物	2.0
Combo40	F817P, S884C, A893C	脯氨酸+二硫化物	2.0
Combo42	T791C, A879C, F817P	二硫化物+脯氨酸	1.4
Combo45	A892P, A899P, A942P	3×脯氨酸	6.2
Combo46	F817P, A892P, A899P	3×脯氨酸	3.8
Combo47	F817P, A892P, A899P, A942P	4×脯氨酸	9.8

[0235] 最显著的结果来自多脯氨酸替换的组合(图3A)。与单独的A892P相比,包含A892P和A942P的Combo14具有蛋白质产率的6.2倍提高(图3B和3C)。在将A899P添加至Combo14的情况下,Combo45显示出具有与Combo14相同的表达水平,但+1.2℃的T_m(图3C)。Combo46是A892P、A899P和F817P的组合,并且与A892P相比,Combo46具有蛋白质产率的3.4倍提高,并且T_m升高3.3℃。在初始S-2P之上,Combo47包含所有四种有益的脯氨酸替换,这不仅加强蛋白质表达为S-2P的9.8倍高,而且在T_m提高约5℃的情况下稳定蛋白质。最重要的是,如通过nsEM显示的,所有具有脯氨酸替换的Combo变体都是良好折叠的三聚体(图3E)。Combo47被重命名为HexaPro,因为其包含总共6个脯氨酸替换,并且是迄今为止最佳的构建体。

[0236] HexaPro大规模表达和应激测试(Stress Testing)。为了评估HexaPro作为潜在疫苗抗原或诊断试剂的生存力(viability),全面检查了其在FreeStyle 293细胞中的大规模产生、在ExpiCHO细胞中蛋白质表达的可行性、表位完整性和蛋白质稳定性。从2L的FreeStyle 293细胞产生了约12mg的HexaPro,或6mg/L,这表示超过S-2P的大于10倍的改善。大规模HexaPro制剂的SEC谱是单分散峰,其对应于三聚体的尺寸(图4A)。根据nsEM,HexaPro的四级结构也维持良好,与S-2P无法区分(图4B)。常规地,重组蛋白的工业生产依赖于CHO细胞而不是HEK293细胞,并因此通过瞬时转染研究了HexaPro在ExpiCHO细胞中的表达。ExpiCHO细胞产生了1.3mg HexaPro/40mL培养物,或32.5mg/L,并且蛋白质是良好折叠的(图4C和4D)。HexaPro与其天然受体人ACE2的结合动力学也与S-2P相当(图4F和4E),其中亲和力分别为13.3nM和11.3nM。重要的是,在3次冻融循环、室温下孵育2天或55℃下孵育

30分钟之后,HexaPro仍保持融合前构象中的折叠(图4G和4H)。相比之下,S-2P显示出在3次冻融循环之后聚集的迹象,并在50℃下30分钟之后开始展开。总的来说,这些数据表明HexaPro具有最佳特征,并表明其可以是用于SARS-CoV-2疫苗开发的有前途的候选物。

[0237] SARS-CoV-2S HexaPro的Cryo-EM结构。为了确定稳定替换不导致任何非预期的构象变化,确定了SARS-CoV-2S HexaPro的cryo-EM结构。从单个网格中,获得了S的两个不同构象的高分辨率3D重构:一个在向上构象中具有单个RBD,并且另一个在向上构象中具有两个RBD。在先前的SARS-CoV-2S-2P的结构表征期间没有观察到该2-RBD-向上构象,并且虽然很容易推测HexaPro中S2增强的稳定性允许观察到这种较低稳定的中间体,但是验证该假设将需要进一步的研究。观察到约三分之一(30.6%)的颗粒处于2RBD向上构象中,引起3.20 Å重构。在1-RBD-向上构象中捕获了剩余的颗粒,但是受体-可及RBD位置的一些灵活性促使去除针对该结构域缺乏清晰密度的1-RBD-向上颗粒的子集,产生最终的引起3.21 Å重构的85,675个颗粒的组(图5A)。1-RBD-向上HexaPro结构与先前确定的3.46 Å S-2P结构的比较在图5B中示出。该重构的相对高的分辨率也允许观察在含有稳定脯氨酸替换的所有四个位置处的密度,这确定了各替换均被适当地引入到刺突蛋白中,并且这些替换对S2亚基的构象没有任何有害的作用(图5C)。

[0238] C. 讨论

[0239] 通常来说,融合前稳定的I类病毒融合蛋白与其不稳定对应物相比诱导更强效的中和抗体,并作为更好的疫苗抗原发挥作用^{15,23}。为了响应于对针对COVID-19大流行的预防性对策的紧急需求,融合前稳定的SARS-CoV-2S-2P结构²⁰被用作设计旨在具有提高的表达或稳定性的100种单替换变体的指南。考虑到S2亚基(像HIV-1gp41或RSV F一样)经历大规模重折叠以促进膜融合,努力特别地集中在该刺突部分。采用的策略之一是引入二硫键,其中至少一个半胱氨酸在融合前与融合后状态之间改变构象的区域中。虽然该方法在HIV-1Env(SOSIP)和RSV F(DS-Cav1)的情况下是成功的^{23,24},但是引入到S2中的二硫化物通常具有有害作用。例如,亚基间二硫化物(例如S659C/S698C)将蛋白质表达降低了60%,并且Q965C/S1003C替换导致部分错误折叠的刺突(图2)。原聚体间二硫化物已显示出改善HIV-1Env的三聚体完整性和RSV F的稳定性^{25,26},但是相对于S-2P,原聚体间T961C/S758C替换消除了表达。相比之下,未发现稳定位于原聚体界面的柔性环是有益的。S884C/A893C和T791C/A879C二者均提高了热稳定性和表达,并产生天然的三聚体结构。将柔性环锚定到相对刚性的α-螺旋可有利于原聚体组装。

[0240] 在HIV-1gp120-gp41界面处引入盐桥不仅加强了蛋白质表达,而且增强了三聚体特异性抗体的结合,这表明天然四级结构的保持得到改善²²。基于类似的原理,T961D替换被引入以与来自相邻原聚体的Arg765形成静电相互作用(图1)。同样,G769E替换旨在与Arg1014形成原聚体间盐桥。两种变体均提高了表达并且也与良好折叠的三聚体刺突相似(图2)。除了盐桥之外,使蛋白质重折叠的对松散堆积的疏水核心的填充可有助于稳定融合前状态,如由先前在RSV F和HIV-1Env中的腔填充替换所示^{23,24,27}。在此,L983F替换被设计成填充部分由HR1、FP和β-发夹形成的腔。该变体具有表达的2倍提高(图2)并且当与二硫化物或脯氨酸替换配对时,显示出具有累加作用。

[0241] 在所发现的最佳单替换变体中是F817P和A942P,二者均显著改善了刺突的品质和

量(图2)。通过进一步将它们与A892P和A899P替换组合,产生了迄今为止最佳的构建体HexaPro。这些结果让人联想到先前针对HIV-1Env、RSV F、hMPV F、MERS-CoV S和Ebola GP成功应用的脯氨酸替换^{23,24,28-30}。另外,融合肽附近疏水残基的溶剂可及性是针对仅流感HA茎设计的问题³¹,并且类似地,该问题在此通过用Pro替换暴露的Phe817来解决(图5C)。A942P替换向连接区与HRX之间的柔性环施加了刚性,并且与被发现有助于稳定Ebola GP的T577P替换类似²⁸。

[0242] 在HexaPro cryo-EM数据集中,观察到三分之一的颗粒在2-RBD-向上构象中,先前对于SARS-CoV-2刺突没有观察到这一点,直至最近确定了包含四个疏水替换的经修饰刺突的结构,这使SD1更接近S2并因此使RBD在向上位置中³²。假设HexaPro中更稳定的S2允许捕获这种可在S1触发和解离之前短暂存在的相对不稳定的构象。这与在稳定的MERS-CoV S-2P刺突的结构中观察到的类似,其中甚至可观察到3-RBD-向上构象(3-RBD-up conformation)³⁰。HexaPro刺突也能够冷冻-解冻、室温储存和热应激之后保持融合前状态,这使得能够将HexaPro刺突开发为亚基疫苗抗原。此外,重组蛋白的工业生产通常通过在CHO细胞中大规模表达来进行。从1L Expi-CHO细胞获得32.5mg良好折叠的HexaPro,这为工业生产提供了可行性。HexaPro刺突还可通过每核酸分子产生更多抗原来改善基于DNA或mRNA的疫苗,从而在相同剂量下改善效力或在较低剂量下维持效力。

[0243] *****

[0244] 根据本公开内容,本文中公开和要求保护的所有方法无需过度实验就可进行和实施。尽管已根据一些优选实施方案描述了本发明的组合物和方法,但对于本领域技术人员来说明显的是,可对本文中所述方法以及所述方法的步骤或步骤顺序作出改变而不脱离本发明的概念、精神和范围。更具体地,显而易见的是,可用某些在化学和生理学两个方面均相关的试剂替代本文中所述的试剂,同时实现相同或类似的结果。对于本领域技术人员明显的所有这样的类似替代和改变被视为在由所附权利要求书所限定的本发明的精神、范围和概念内。

[0245] 参考文献

[0246] 以下参考文献就其提供补充本文中所阐述的那些操作或细节的示例性操作或其他细节而言特别地通过引用并入本文。

[0247] 1.Peiris,J.S.M.et al.Coronavirus as a possible cause of severe acute respiratory syndrome.Lancet 361,1319-1325(2003) .

[0248] 2.Zaki,A.M.,Van Boheemen,S.,Bestebroer,T.M.,Osterhaus,A.D.M.E.& Fouchier,R.A.M.Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.N.Engl.J.Med.367,1814-1820(2012) .

[0249] 3.Chan,J.F.W.et al.A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission:a study of a family cluster.Lancet 395,514-523(2020) .

[0250] 5.Li,F.Structure,Function,and Evolution of Coronavirus Spike Proteins.Annu.Rev.Virol.3,237-261(2016) .

[0251] 6.Siebert,D.N.,Bosch,B.J.,van der Zee,R.,de Haan,C.A.M.&Rottier,P.J.M.The Coronavirus Spike Protein Is a Class I Virus Fusion Protein:

Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* 77, 8801-8811 (2003) .

[0252] 7. Mille, J.K. & Whittaker, G.R. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15214-15219 (2014) .

[0253] 8. Hoffmann, M. et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181, 271-280.e8 (2020) .

[0254] 10. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270-273 (2020) .

[0255] 11. Walls, A.C. et al. Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.* 114, 11157-11162 (2017) .

[0256] 12. Wang, C. et al. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nat. Commun.* 11, 1-6 (2020) .

[0257] 13. Buchholz, U.J. et al. Contributions of the structural proteins of severe respiratory syndrome coronavirus to protective immunity. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9804-9809 (2004) .

[0258] 14. Hofmann, H. et al. S Protein of Severe Acute Respiratory Syndrome-Associated Coronavirus Mediates Entry into Hepatoma Cell Lines and Is Targeted by Neutralizing Antibodies in Infected Patients. *J. Virol.* 78, 6134-6142 (2004) .

[0259] 15. McLellan, J.S. et al. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* 342, 592-598 (2013) .

[0260] 16. Pallesen, J. et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7348-E7357 (2017) .

[0261] 17. Park, Y.J. et al. Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. *Nat. Struct. Mol. Biol.* 26, 1151-1157 (2019) .

[0262] 18. Li, Z. et al. The human coronavirus HCoV-229E S-protein structure and receptor binding. *Elife* 8, 1-22 (2019) .

[0263] 19. Wang, N. et al. Structural Definition of a Neutralization-Sensitive Epitope on the MERS-CoV S1-NTD. *Cell Rep.* 28, 3395-3405.e6 (2019) .

[0264] 20. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* (80-.). 367, 1260-1263 (2020) .

[0265] 21. Walls, A.C. et al. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281-292.e6 (2020) .

[0266] 22. Rutten, L. et al. A Universal Approach to Optimize the Folding and

Stability of Prefusion-Closed HIV-1 Envelope Trimers. *Cell Rep.* 23, 584-595 (2018) .

[0267] 23. Sanders, R.W. et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp 140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog.* 9, e1003618 (2013) .

[0268] 24. McLellan, J.S. et al. Structure-Based Design of a Fusion Glycoprotein Vaccine for Respiratory Syncytial Virus. *Science* (80-.). 342, 592-598 (2013) .

[0269] 25. Yang, L. et al. Structure-Guided Redesign Improves NFL HIV Env Trimer Integrity and Identifies an Inter-Protomer Disulfide Permitting Post-Expression Cleavage. *Front. Immunol.* 9, 1631 (2018) .

[0270] 26. Joyce, M.G. et al. Iterative structure-based improvement of a fusion-glycoprotein vaccine against RSV. *Nat. Struct. Mol. Biol.* 23, 811-820 (2016) .

[0271] 27. Krarup, A. et al. A highly stable prefusion RSV F vaccine derived from structural analysis of the fusion mechanism. *Nat. Commun.* 6, 8143 (2015) .

[0272] 28. Rutten, L. et al. Structure-Based Design of Prefusion-Stabilized Filovirus Glycoprotein Trimers. *Cell Rep.* 30, 4540-4550.e3 (2020) .

[0273] 29. Battles, M.B. et al. Structure and immunogenicity of pre-fusion-stabilized human metapneumovirus F glycoprotein. *Nat. Commun.* 8, 1528 (2017) .

[0274] 30. Pallesen, J. et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7348-E7357 (2017) .

[0275] 31. Impagliazzo, A. et al. A stable trimeric influenza hemagglutinin stem as a broadly protective immunogen. *Science* 349, 1301-1306 (2015) .

[0276] 32. Henderson, R. et al. Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. *bioRxiv* 2020.05.18.102087 (2020) . doi:10.1101/2020.05.18.102087

[0277] 33. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* 16, 1146-1152 (2019) .

[0278] 34. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290-296 (2017) .

[0279] 35. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 486-501 (2010) .

[0280] 36. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* 75, 861-877 (2019) .

[0281] 37. Croll, T.I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. D, Struct. Biol.* 74, 519-530 (2018) .

序列表

<110> Board of Regents, The University of Texas System
 The Trustees of Dartmouth College
 <120> 经改造冠状病毒刺突 (S) 蛋白及其使用方法
 <130> UTFB.P1251W0
 <140> 尚未知
 <141> 2021-05-28
 <150> US 63/032,502
 <151> 2020-05-29
 <160> 80
 <170> PatentIn version 3.5
 <210> 1
 <211> 1288
 <212> PRT
 <213> 人工序列
 <220>
 <223> 在C端处具有三聚结构域和标签的WA1 SARS-Cov-2胞外结构域
 <220>
 <221> MISC_FEATURE
 <222> (1)..(1208)
 <223> WA1 SARS-Cov-2胞外结构域
 <220>
 <221> MISC_FEATURE
 <222> (1209)..(1288)
 <223> 三聚结构域和标签
 <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

[0001]

	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				
[0002]	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
 [0003] 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
 [0004] 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 1055	Val Phe Leu His Val 1060	Thr Tyr Val 1065
	Pro Ala Gln Glu Lys Asn Phe 1070	Thr Thr Ala Pro Ala 1075	Ile Cys His 1080
	Asp Gly Lys Ala His Phe Pro 1085	Arg Glu Gly Val Phe 1090	Val Ser Asn 1095
	Gly Thr His Trp Phe Val Thr 1100	Gln Arg Asn Phe Tyr 1105	Glu Pro Gln 1110
	Ile Ile Thr Thr Asp Asn Thr 1115	Phe Val Ser Gly Asn 1120	Cys Asp Val 1125
	Val Ile Gly Ile Val Asn Asn 1130	Thr Val Tyr Asp Pro 1135	Leu Gln Pro 1140
	Glu Leu Asp Ser Phe Lys Glu 1145	Glu Leu Asp Lys Tyr 1150	Phe Lys Asn 1155
	His Thr Ser Pro Asp Val Asp 1160	Leu Gly Asp Ile Ser 1165	Gly Ile Asn 1170
[0005]	Ala Ser Val Val Asn Ile Gln 1175	Lys Glu Ile Asp Arg 1180	Leu Asn Glu 1185
	Val Ala Lys Asn Leu Asn Glu 1190	Ser Leu Ile Asp Leu 1195	Gln Glu Leu 1200
	Gly Lys Tyr Glu Gln Gly Ser 1205	Gly Tyr Ile Pro Glu 1210	Ala Pro Arg 1215
	Asp Gly Gln Ala Tyr Val Arg 1220	Lys Asp Gly Glu Trp 1225	Val Leu Leu 1230
	Ser Thr Phe Leu Gly Arg Ser 1235	Leu Glu Val Leu Phe 1240	Gln Gly Pro 1245
	Gly His His His His His 1250	His His Ser Ala Trp 1255	Ser His Pro 1260
	Gln Phe Glu Lys Gly Gly Gly 1265	Ser Gly Gly Gly Gly 1270	Ser Gly Gly 1275
	Ser Ala Trp Ser His Pro Gln 1280	Phe Glu Lys 1285	
	<210> 2		
	<211> 1208		
	<212> PRT		
	<213> 人工序列		
	<220>		
	<223> 野生型和变体的SARS-CoV-2胞外结构域共有序列		

	<220>	
	<221>	MISC_FEATURE
	<222>	(5)..(5)
	<223>	Xaa是Leu或Phe
	<220>	
	<221>	MISC_FEATURE
	<222>	(13)..(13)
	<223>	Xaa是Ser或Ile
	<220>	
	<221>	MISC_FEATURE
	<222>	(18)..(18)
	<223>	Xaa是Leu或Phe
	<220>	
	<221>	MISC_FEATURE
	<222>	(19)..(19)
	<223>	Xaa是Thr或Arg
	<220>	
	<221>	MISC_FEATURE
	<222>	(20)..(20)
	<223>	Xaa是Thr或Asn
	<220>	
	<221>	MISC_FEATURE
	<222>	(26)..(26)
	<223>	Xaa是Pro或Ser
	<220>	
	<221>	MISC_FEATURE
	<222>	(52)..(52)
	<223>	Xaa是Gln或Arg
[0006]	<220>	
	<221>	MISC_FEATURE
	<222>	(67)..(67)
	<223>	Xaa是Ala或Val
	<220>	
	<221>	MISC_FEATURE
	<222>	(69)..(69)
	<223>	Xaa是His或不存在
	<220>	
	<221>	MISC_FEATURE
	<222>	(70)..(70)
	<223>	Xaa是Val、Ile或不存在
	<220>	
	<221>	MISC_FEATURE
	<222>	(80)..(80)
	<223>	Xaa是Asp、Gly或Ala
	<220>	
	<221>	MISC_FEATURE
	<222>	(95)..(95)
	<223>	Xaa是Thr或Ile
	<220>	
	<221>	MISC_FEATURE
	<222>	(138)..(138)
	<223>	Xaa是Asp或Tyr
	<220>	
	<221>	MISC_FEATURE
	<222>	(142)..(142)
	<223>	Xaa是Gly或Asp
	<220>	
	<221>	MISC_FEATURE
	<222>	(144)..(144)
	<223>	Xaa是Tyr、Val或不存在

<220>
<221> MISC FEATURE
<222> (152)..(152)
<223> Xaa是Trp或Cys

<220>
<221> MISC FEATURE
<222> (154)..(154)
<223> Xaa是Glu或Lys

<220>
<221> MISC FEATURE
<222> (157)..(157)
<223> Xaa是Phe或Ser

<220>
<221> MISC FEATURE
<222> (190)..(190)
<223> Xaa是Arg或Ser

<220>
<221> MISC FEATURE
<222> (215)..(215)
<223> Xaa是Asp或Gly

<220>
<221> MISC FEATURE
<222> (242)..(242)
<223> Xaa是Leu或不存在

<220>
<221> MISC FEATURE
<222> (243)..(243)
<223> Xaa是Ala或不存在

[0007] <220>
<221> MISC FEATURE
<222> (244)..(244)
<223> Xaa是Leu或不存在

<220>
<221> MISC FEATURE
<222> (253)..(253)
<223> Xaa是Asp或Gly

<220>
<221> MISC FEATURE
<222> (258)..(258)
<223> Xaa是Trp或Leu

<220>
<221> MISC FEATURE
<222> (417)..(417)
<223> Xaa是Lys、Asn或Thr

<220>
<221> MISC FEATURE
<222> (452)..(452)
<223> Xaa是Leu或Arg

<220>
<221> MISC FEATURE
<222> (478)..(478)
<223> Xaa是Thr或Lys

<220>
<221> MISC FEATURE
<222> (484)..(484)
<223> Xaa是Glu、Lys或Gln

<220>
<221> MISC FEATURE
<222> (501)..(501)
<223> Xaa是Asn或Tyr

<220>
<221> MISC FEATURE
<222> (570)..(570)
<223> Xaa是Ala或Asp

<220>
<221> MISC FEATURE
<222> (614)..(614)
<223> Xaa是Asp或Gly

<220>
<221> MISC FEATURE
<222> (655)..(655)
<223> Xaa是His或Tyr

<220>
<221> MISC FEATURE
<222> (677)..(677)
<223> Xaa是Gln或His

<220>
<221> MISC FEATURE
<222> (681)..(681)
<223> Xaa是Pro、His或Arg

<220>
<221> MISC FEATURE
<222> (701)..(701)
<223> Xaa是Ala或Val

<220>
<221> MISC FEATURE
<222> (716)..(716)
<223> Xaa是Thr或Ile

[0008] <220>
<221> MISC FEATURE
<222> (791)..(791)
<223> Xaa是Thr或Ile

<220>
<221> MISC FEATURE
<222> (859)..(859)
<223> Xaa是Thr或Asn

<220>
<221> MISC FEATURE
<222> (888)..(888)
<223> Xaa是Phe或Leu

<220>
<221> MISC FEATURE
<222> (950)..(950)
<223> Xaa是Asp、His或Asn

<220>
<221> MISC FEATURE
<222> (982)..(982)
<223> Xaa是Ser或Ala

<220>
<221> MISC FEATURE
<222> (1027)..(1027)
<223> Xaa是Thr或Ile

<220>
<221> MISC FEATURE
<222> (1118)..(1118)
<223> Xaa是Asp或His

<220>
<221> MISC FEATURE
<222> (1176)..(1176)
<223> Xaa是Val或Phe

<400> 2

Met Phe Val Phe Xaa Val Leu Leu Pro Leu Val Ser Xaa Gln Cys Val
1 5 10 15

Asn Xaa Xaa Xaa Arg Thr Gln Leu Pro Xaa 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 Xaa Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
50 55 60

Phe His Xaa Ile Xaa Xaa Ser Gly Thr Asn Gly Thr Lys Arg Phe Xaa
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Xaa 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 Xaa Pro Phe Leu Xaa Val Xaa
130 135 140

[0009]

Tyr His Lys Asn Asn Lys Ser Xaa Met Xaa Ser Glu Xaa 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 Xaa 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 Xaa 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 Xaa Xaa Xaa His Arg Ser Tyr Leu Thr Pro Gly Xaa Ser Ser Ser
245 250 255

Gly Xaa 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
 Xaa Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
 420 425 430
 [0010] Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
 435 440 445
 Tyr Asn Tyr Xaa 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 Xaa Pro Cys
 465 470 475 480
 Asn Gly Val Xaa Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
 485 490 495
 Phe Gln Pro Thr Xaa 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 Xaa 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 Xaa 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 Xaa 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 Xaa Thr Asn Ser Xaa Arg Arg Ala Arg Ser Val Ala
 675 680 685
 Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Xaa Glu Asn Ser
 690 695 700
 Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Xaa 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
 [0011] 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 Xaa 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 Xaa 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 Xaa 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	Xaa	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	Xaa	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	Xaa	Lys	Met	Ser	Glu	Cys	Val	Leu	Gly	Gln	Ser	Lys		
	1025					1030					1035					
[0012]	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	Xaa	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	Xaa	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			
	1205			
	<210> 3			
	<211> 9			
	<212> DNA			
	<213> 人工序列			
	<220>			
	<223> 合成的寡核苷酸			
	<400> 3			
	ggtaccaga			9
	<210> 4			
	<211> 24			
	<212> DNA			
	<213> 人工序列			
	<220>			
	<223> 合成的寡核苷酸			
	<400> 4			
	ccgtctcagg ccgagttcgg tacc			24
[0013]	<210> 5			
	<211> 35			
	<212> DNA			
	<213> 人工序列			
	<220>			
	<223> 合成的寡核苷酸			
	<400> 5			
	ggaaacaagg caacttcaag aacctgagag aattc			35
	<210> 6			
	<211> 27			
	<212> DNA			
	<213> 人工序列			
	<220>			
	<223> 合成的寡核苷酸			
	<400> 6			
	gccgtcgatg ttcttgaaca cgaattc			27
	<210> 7			
	<211> 24			
	<212> DNA			
	<213> 人工序列			
	<220>			
	<223> 合成的寡核苷酸			
	<400> 7			
	gggtctgctt cctctgtagc tagc			24
	<210> 8			
	<211> 27			
	<212> DNA			
	<213> 人工序列			

	<220>		
	<223>	合成的寡核苷酸	
	<400>	8	
		ggtgtaggcg atgatgctct ggctagc	27
	<210>	9	
	<211>	31	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	9	
		cagctctgtg ctgaacgata tcctgtctag a	31
	<210>	10	
	<211>	24	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	10	
		ggcttctgga gggccagtc taga	24
	<210>	11	
	<211>	23	
	<212>	DNA	
	<213>	人工序列	
[0014]	<220>		
	<223>	合成的寡核苷酸	
	<400>	11	
		ggcctcgggg atgtatccgg atc	23
	<210>	12	
	<211>	34	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	12	
		gggcaggatc tctgttgtgc agctaattgt aaag	34
	<210>	13	
	<211>	34	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	13	
		ctttacaatt agctgcacaa cagagatcct gccc	34
	<210>	14	
	<211>	29	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	

	<400> 14 ggatcttgcc gatgcaagaa ttgaactgg	29
	<210> 15 <211> 29 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 15 ccagtccaat tcttgcatcg gcaagatcc	29
	<210> 16 <211> 58 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 16 gccatgcaga tgtgctatag attcaacgga atcggcgtga cctgcaacgt gctgtatg	58
	<210> 17 <211> 58 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
[0015]	<400> 17 catacagcac gttgcaggtc acgccgattc cgttgaatct atagcacatc tgcattggc	58
	<210> 18 <211> 60 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 18 cttctggagg gtccagtcta gacaggatac agttcagcac agagcagatg gcgccaaaat	60
	<210> 19 <211> 26 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 19 cagagatcct gtgcgtgagc atgacc	26
	<210> 20 <211> 26 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 20 ggtcatgctc acgcacagga tctctg	26

	<210> 21	
	<211> 30	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 21	
	caaactgcag gactgcgtga atcagaacgc	30
	<210> 22	
	<211> 30	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 22	
	gcgttctgat tcacgcagtc ctgcagtttg	30
	<210> 23	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 23	
	gaccaagacc agctgtgact gtacaatgt	29
[0016]	<210> 24	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 24	
	acattgtaca gtcacagctg gtcttggtc	29
	<210> 25	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 25	
	ttcaacggct gcacagttct cccac	25
	<210> 26	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 26	
	gtgggagaac tgtgcagccg ttgaa	25
	<210> 27	

[0017]	<211> 40		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的寡核苷酸		
	<400> 27		
	tcacatgtgg gtggacattt ggcgccggcg cctgcctgca	40	
	<210> 28		
	<211> 40		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的寡核苷酸		
	<400> 28		
	tgcaggcagg cgccggcgcc aaatgtccac ccacatgtga	40	
	<210> 29		
	<211> 30		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的寡核苷酸		
	<400> 29		
	aggcttagaa ggatcacaga gaatctggct	30	
	<210> 30		
	<211> 30		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的寡核苷酸		
	<400> 30		
	agccagattc tctgtgatcc ttctaagcct	30	
	<210> 31		
	<211> 25		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的寡核苷酸		
	<400> 31		
	tcagtacacc tgtgccctgc tgget	25	
	<210> 32		
	<211> 25		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成的寡核苷酸		
	<400> 32		
	agccagcagg gcacaggtgt actga	25	
	<210> 33		
	<211> 30		
	<212> DNA		
	<213> 人工序列		

	<220>		
	<223>	合成的寡核苷酸	
	<400>	33	
		aagcagatct acaagtgcc acctatcaag	30
	<210>	34	
	<211>	30	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	34	
		cttgataggt gggcacttgt agatctgctt	30
	<210>	35	
	<211>	25	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	35	
		tctgccctgc tgtgtggcac catca	25
	<210>	36	
	<211>	25	
	<212>	DNA	
	<213>	人工序列	
[0018]	<220>		
	<223>	合成的寡核苷酸	
	<400>	36	
		tgatggtgcc acacagcagg gcaga	25
	<210>	37	
	<211>	30	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	37	
		ctggctaaag ttaaagcagc cgaagtcctt	30
	<210>	38	
	<211>	30	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	38	
		aaggacttcg gctgctttaa cttagccag	30
	<210>	39	
	<211>	33	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	

	<400> 39 aaccagaagc tgatctgtaa ccagttcaat tct	33
	<210> 40 <211> 33 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 40 agaattgaac tggttacaga tcagcttctg gtt	33
	<210> 41 <211> 23 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 41 cggccagggt gagcttggtg aac	23
	<210> 42 <211> 23 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
[0019]	<400> 42 gttcaacaag ctcaccctgg ccg	23
	<210> 43 <211> 28 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 43 aggccatctg catgaaaaag gggatctg	28
	<210> 44 <211> 28 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 44 cagatcccct tttcatgca gatggcct	28
	<210> 45 <211> 24 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 45 gatcctcgat agggctccgc ttgc	24

	<210> 46	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 46	
	gcaagcggag ccctatcgag gatc	24
	<210> 47	
	<211> 28	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 47	
	gacaaaaata ccatggaggt gttcgccc	28
	<210> 48	
	<211> 28	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 48	
	gggcgaacac ctccatgcta ttttgtc	28
[0020]	<210> 49	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 49	
	tcgtcggtcg gcagaggtgg	20
	<210> 50	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 50	
	ccacctctgc cgaccgacga	20
	<210> 51	
	<211> 34	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 51	
	gccgcagaag tcaacgaatt tagactgtcc cagc	34
	<210> 52	

[0021]	<211>	34	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	52	
		gctgggacag tctaaattcg ttgacttctg cggc	34
	<210>	53	
	<211>	37	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	53	
		ggcttctgga gggtcagtg gagacaggat atcgttc	37
	<210>	54	
	<211>	37	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	54	
		gaacgatatc ctgtctccac tggaccctcc agaagcc	37
[0021]	<210>	55	
	<211>	38	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	55	
		cttcggcttc tggagggctc ggtctagaca ggataatcg	38
	<210>	56	
	<211>	38	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	56	
		cgatatcctg tctagaccgg accctccaga agccgaag	38
	<210>	57	
	<211>	35	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的寡核苷酸	
	<400>	57	
		ggacttcggc ttctggaggg ggcagtctag acagg	35
	<210>	58	
	<211>	35	
	<212>	DNA	
	<213>	人工序列	

	<220> <223> 合成的寡核苷酸	
	<400> 58 cctgtctaga ctgccccctc cagaagccga agtcc	35
	<210> 59 <211> 37 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 59 ggctcttggtc atgctcacgc acaggatctc tgttgtg	37
	<210> 60 <211> 37 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 60 cacaacagag atcctgtgcg tgagcatgac caagacc	37
	<210> 61 <211> 35 <212> DNA <213> 人工序列	
[0022]	<220> <223> 合成的寡核苷酸	
	<400> 61 gggcgttctg attcacgcag tcctgcagtt tgccc	35
	<210> 62 <211> 35 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 62 gggcaaactg caggactgcg tgaatcagaa cgccc	35
	<210> 63 <211> 28 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 63 cagttacgag tgcgacatcc ctatcggc	28
	<210> 64 <211> 24 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	

	<400> 64 ccctgaacac cctggtgaag cagc	24
	<210> 65 <211> 23 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 65 ccctgacag ctgttgggtc acg	23
	<210> 66 <211> 22 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 66 cagcacctcc agggatctgc cc	22
	<210> 67 <211> 76 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
[0023]	<400> 67 ctggttttca tacagcacgt tgcaggtcac gccgattccg ttgaatctat agcacatctg	60
	catggcaaag gggatc	76
	<210> 68 <211> 76 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 68 gatcccccttt gccatgcaga tgtgctatag attcaacgga atcggcgtga cctgcaacgt	60
	gctgtatgaa aaccag	76
	<210> 69 <211> 30 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 69 gcaggagctc ggcaaatacg agcagggatc	30
	<210> 70 <211> 25 <212> DNA <213> 人工序列	
	<220>	

[0024]	<223> 合成的寡核苷酸	
	<400> 70	
	gccagaagtc agatgctcaa ggggc	25
	<210> 71	
	<211> 45	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 71	
	cttctcgaac tgggggtggg accagggcgt atgatggtgg tgatg	45
	<210> 72	
	<211> 45	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 72	
	tgataatgac tcgagcgata attcactcct caggtgcagg ctgcc	45
	<210> 73	
	<211> 32	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 73	
	atgatggtgg tgatggtggt gatggcctgg gc	32
	<210> 74	
	<211> 36	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 74	
	cacaggaaac agctatgacc atgattacgc caagct	36
	<210> 75	
	<211> 46	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 75	
	gatccgcccc ctccactacc tcatcacttc tcgaactggg ggtggg	46
	<210> 76	
	<211> 46	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的寡核苷酸	
	<400> 76	

	cccccccca gttcgagaag tgatgaggta gtggaggggg cggatc	46
	<210> 77 <211> 48 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 77 ctcgaactgg gggtagggacc atcatcaatg atggtggtga tggtaggtg	48
	<210> 78 <211> 48 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
[0025]	<400> 78 caccaccatc accaccatca ttgatgatgg tcccaccccc agttcgag	48
	<210> 79 <211> 40 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 79 cagggcgctg gggtagagg agaggagtc ctggatcttg	40
	<210> 80 <211> 28 <212> DNA <213> 人工序列	
	<220> <223> 合成的寡核苷酸	
	<400> 80 tcctctaccc ccagcgccct gggcaaac	28

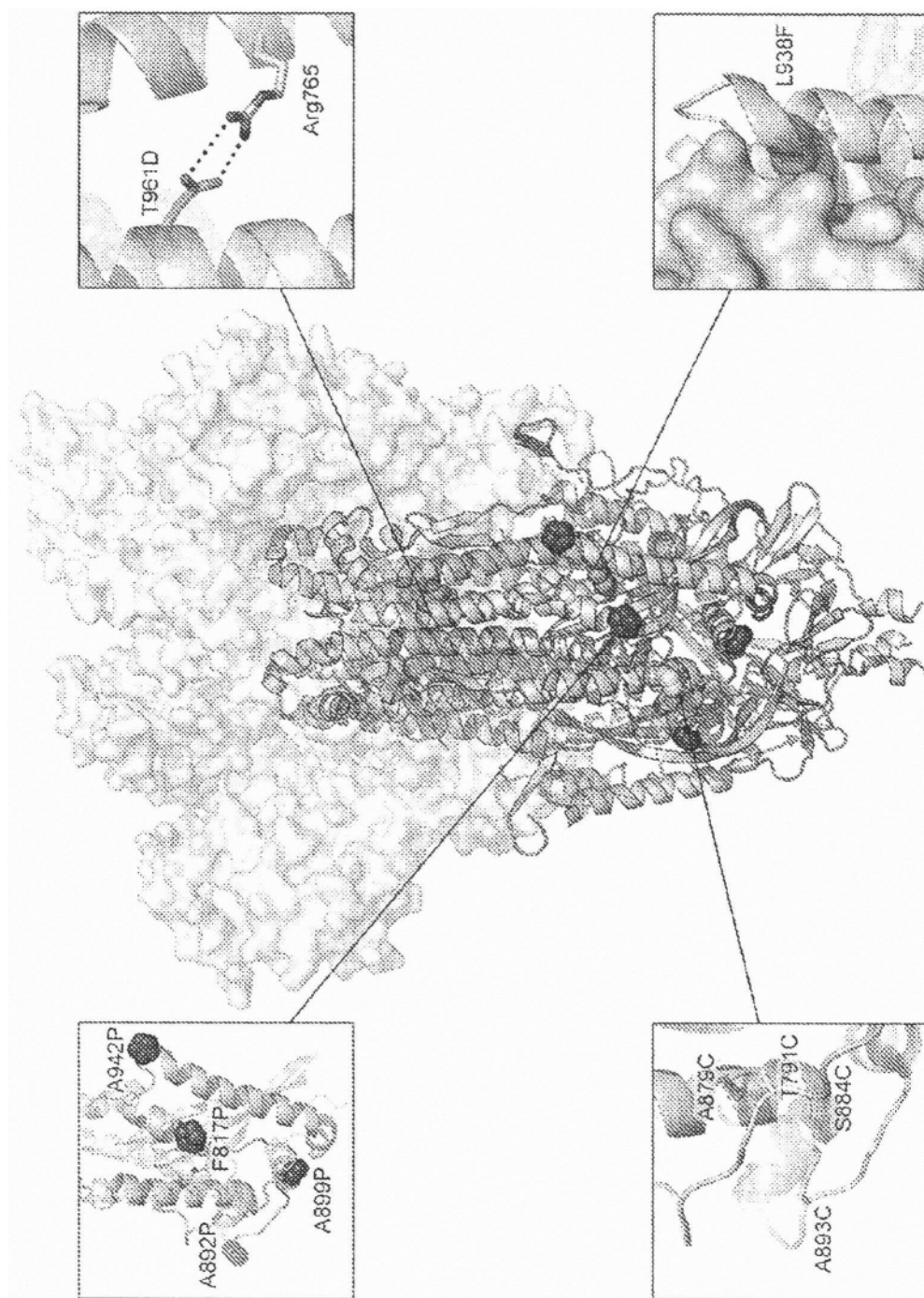


图1

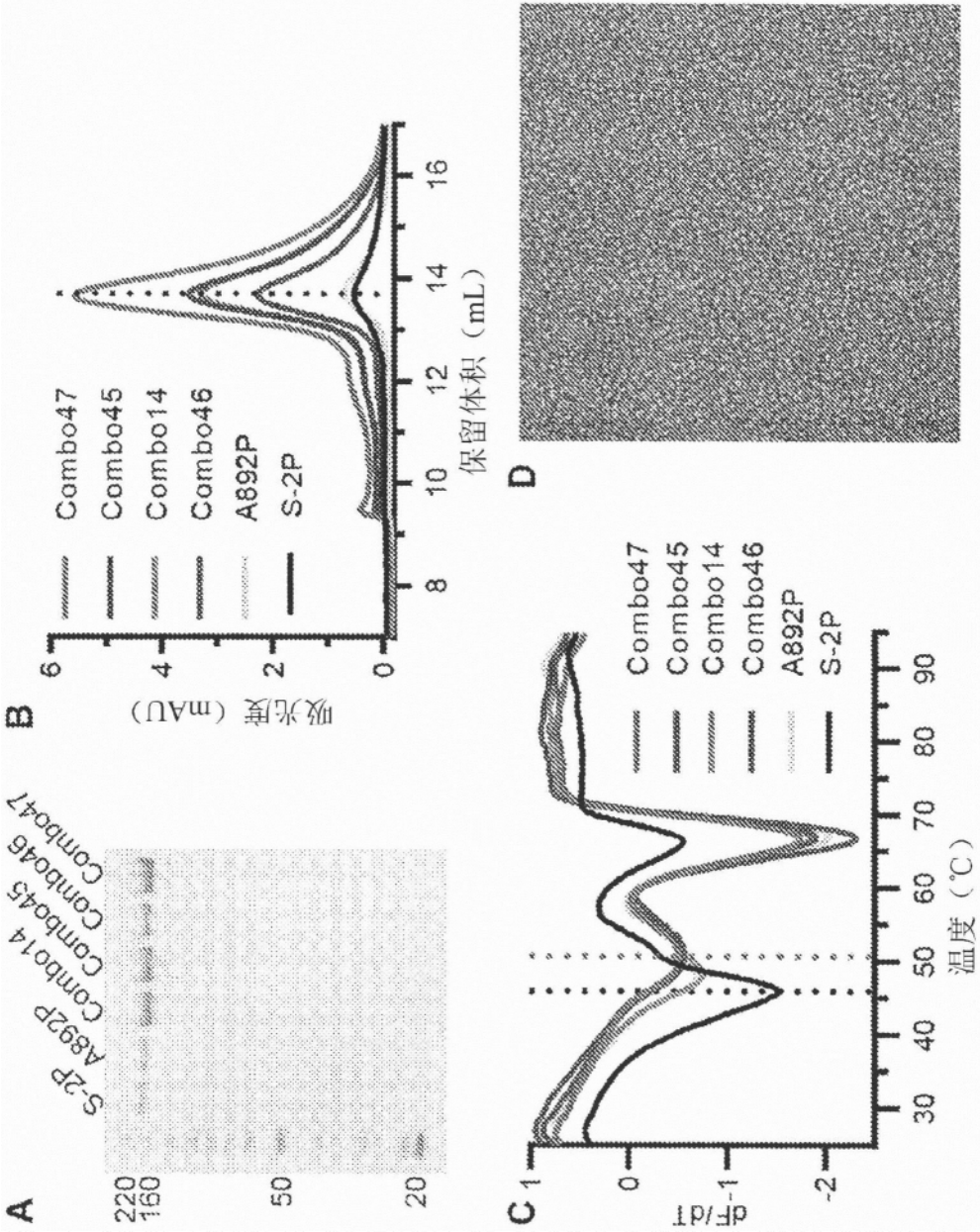


图3A-3D

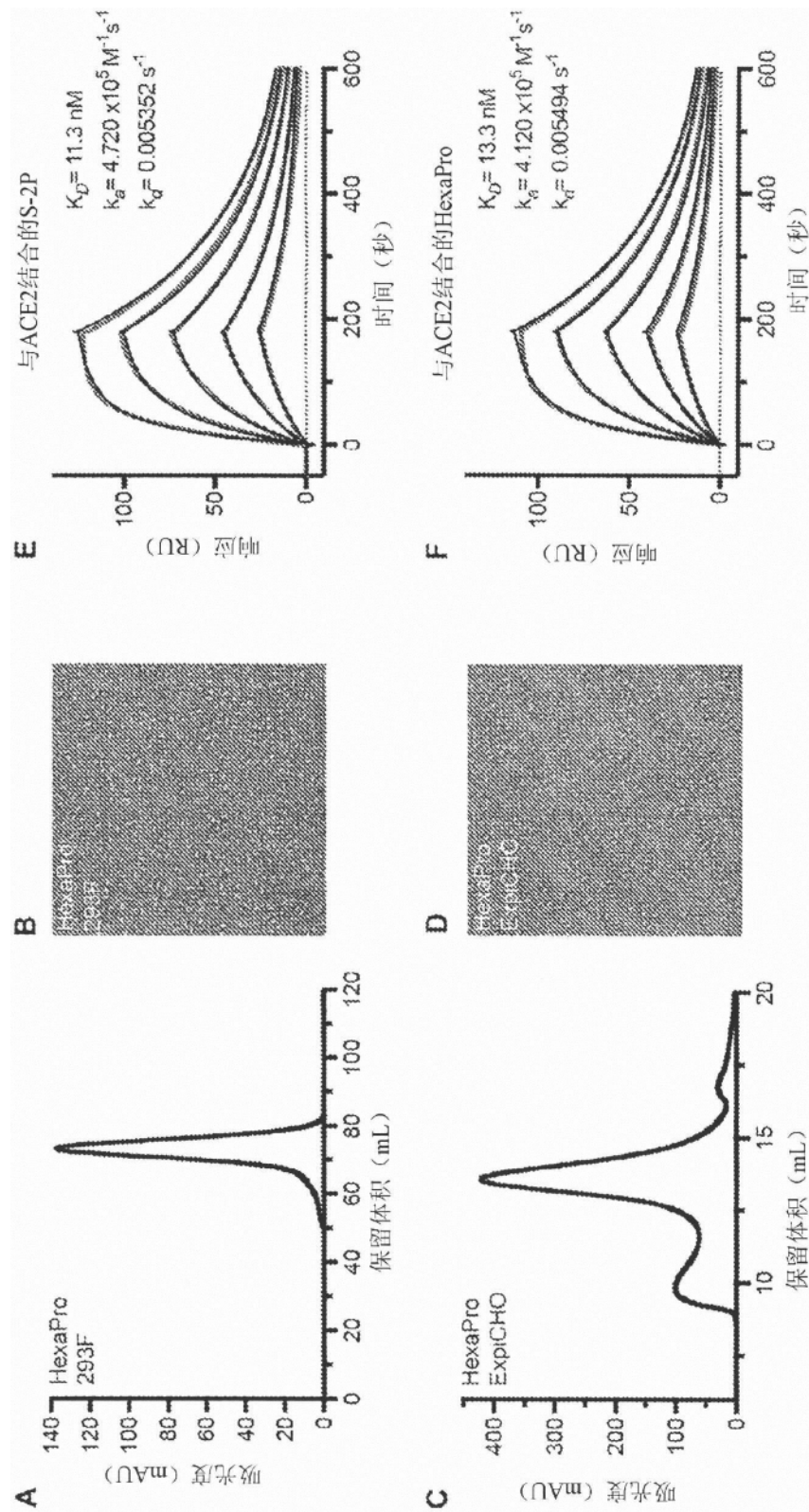


图4A-4F

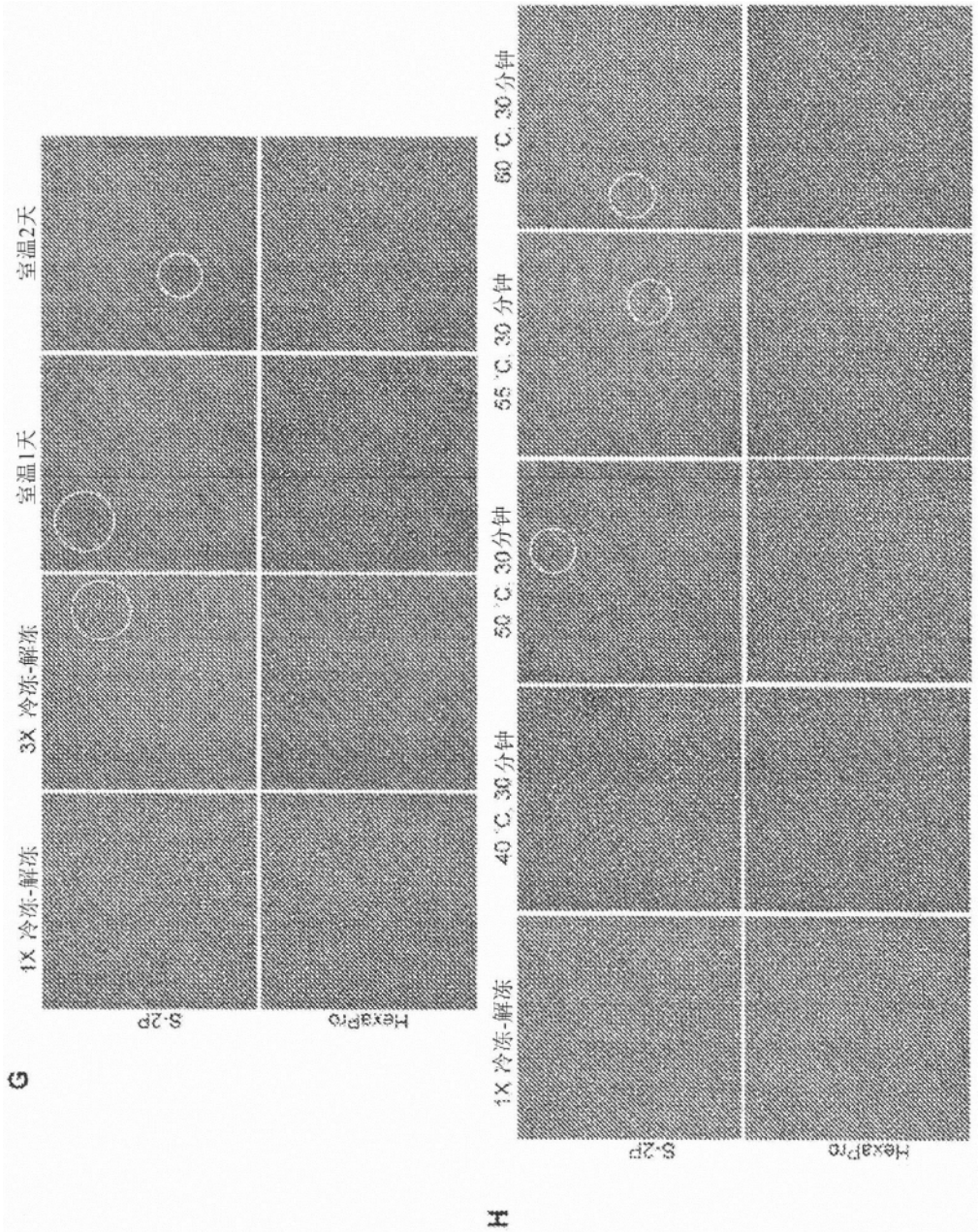


图4G-4H

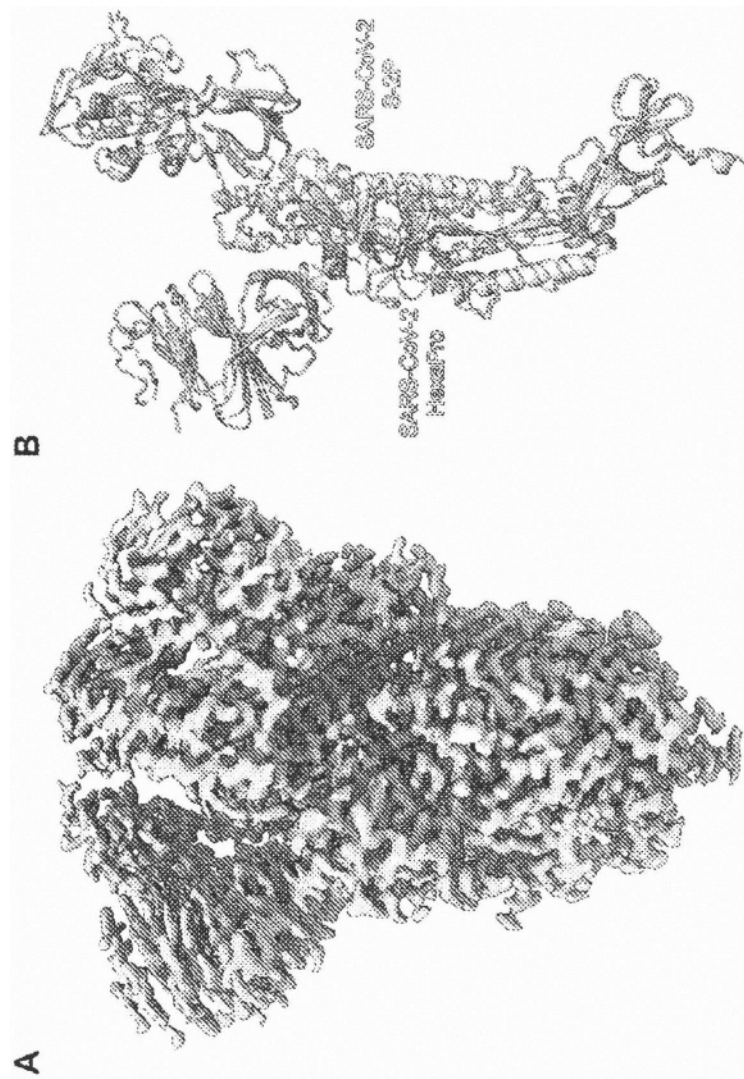


图5A-5B

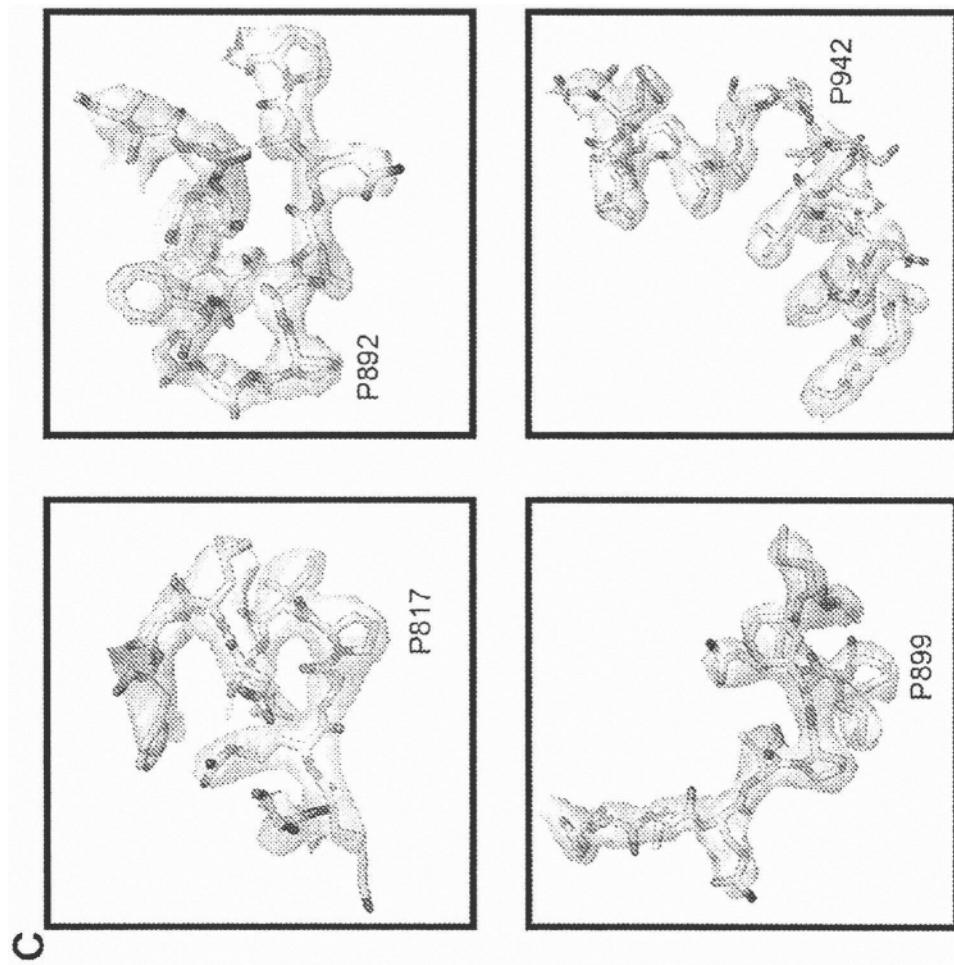


图5C

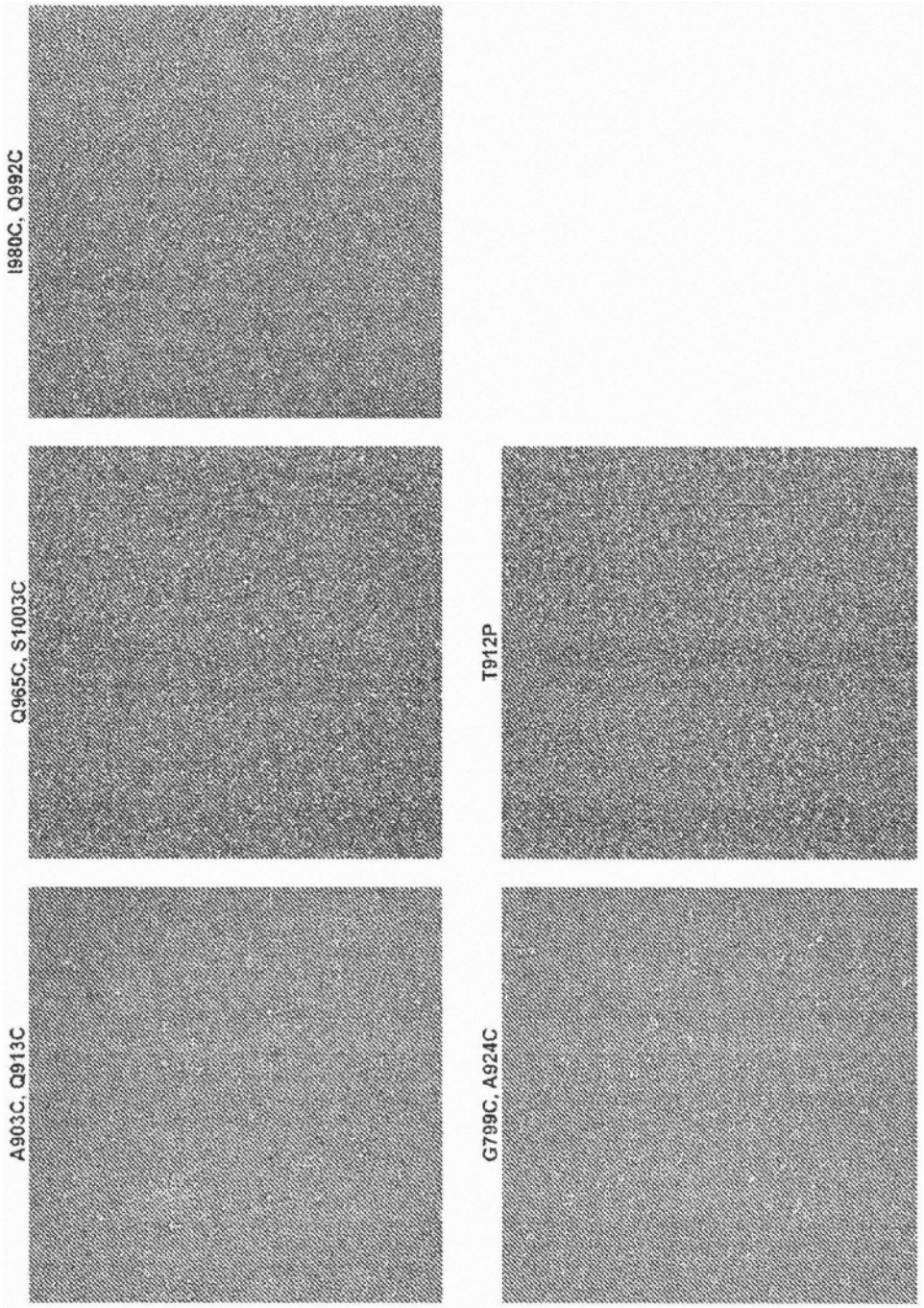


图6

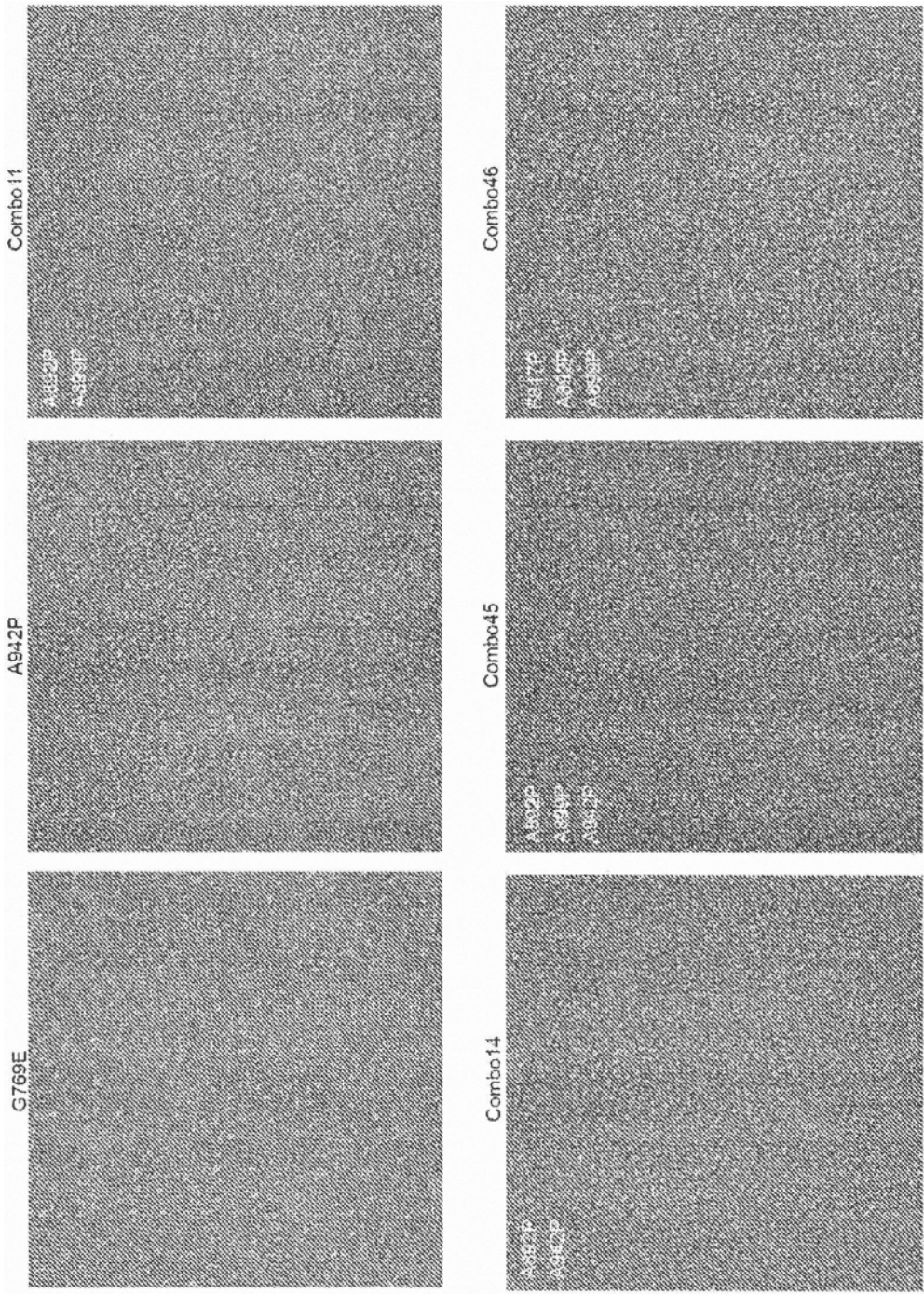


图7

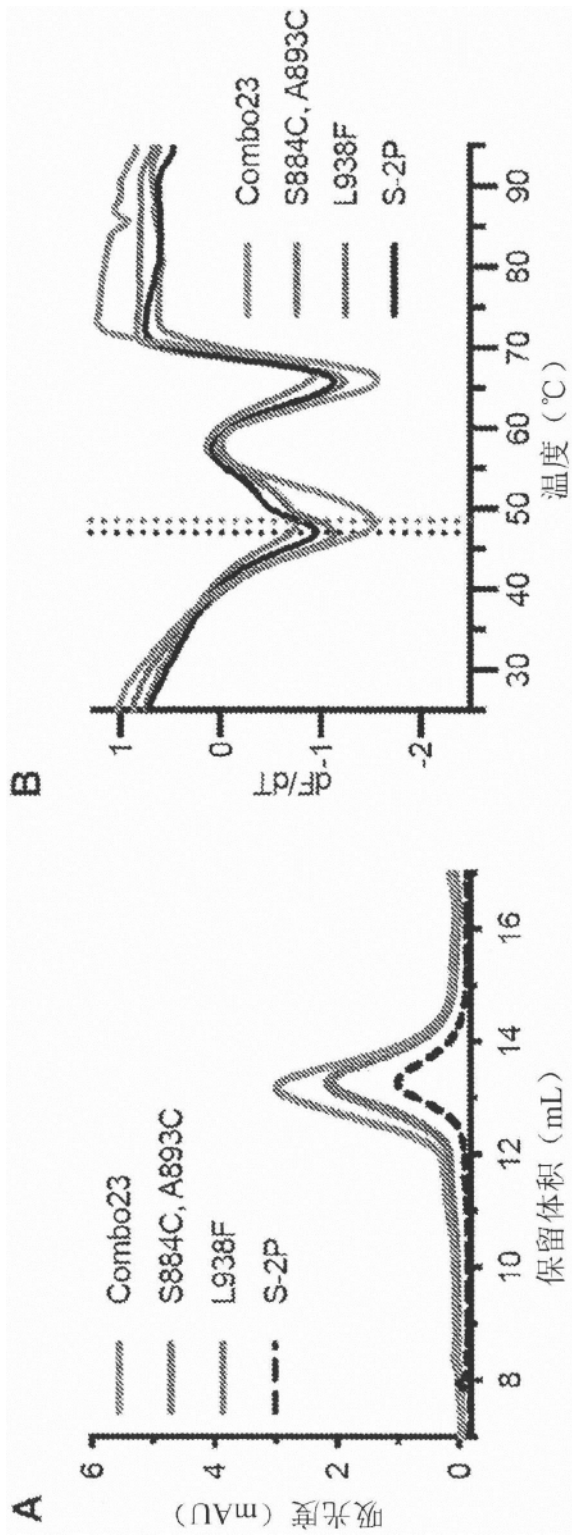


图8A-8B

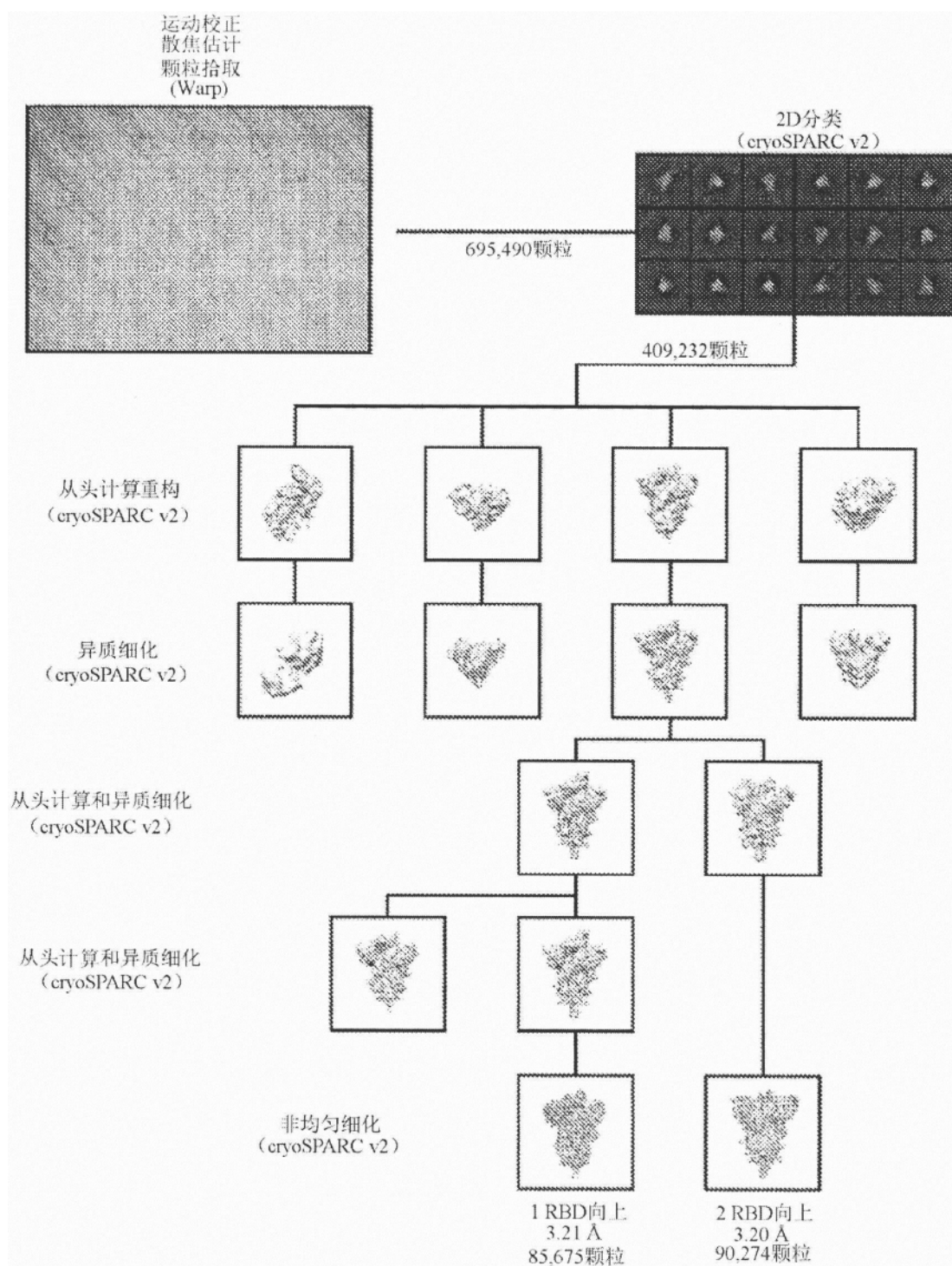


图9

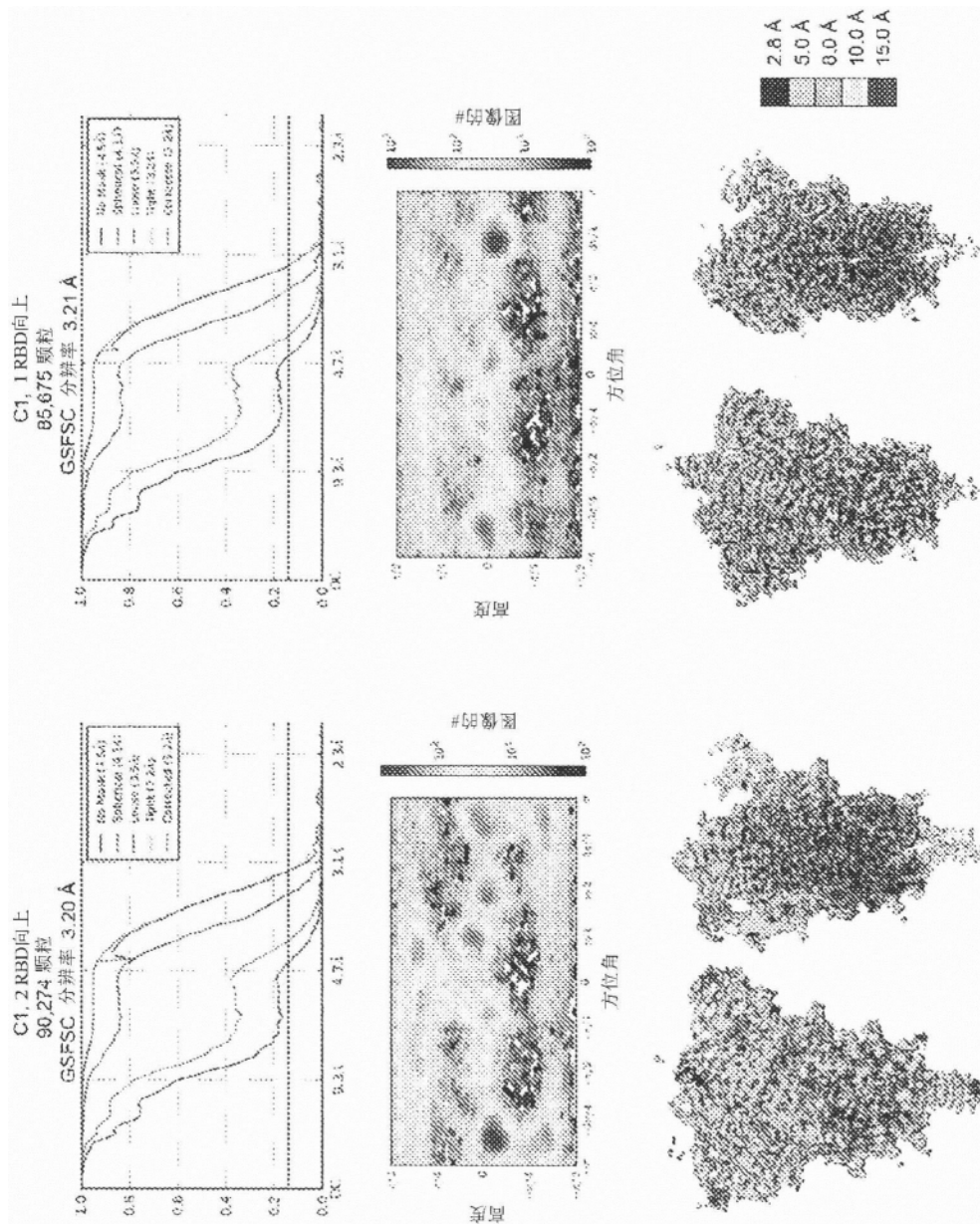


图10