(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0008707
(43) 공개일자 2023년01월16일

(51) 국제특허분류(Int. Cl.)
A61K 39/12 (2006.01) *A61K 39/00* (2006.01)
A61K 39/215 (2006.01) *A61P 31/14* (2006.01)
A61P 37/04 (2006.01) *C07K 14/005* (2006.01)
C12N 15/86 (2006.01)
(52) CPC특허분류
A61K 39/12 (2013.01)
A61K 39/215 (2013.01)
(21) 출원번호 10-2022-7037855
(22) 출원일자(국제) 2021년03월30일
심사청구일자 없음
(85) 번역문제출일자 2022년10월28일
(86) 국제출원번호 PCT/IB2021/000190
(87) 국제공개번호 WO 2021/198769
국제공개일자 2021년10월07일
(30) 우선권주장
63/002,237 2020년03월30일 미국(US)
63/070,150 2020년08월25일 미국(US)

(71) 출원인
배리에이션 바이오테크놀로지스 아이엔씨.
캐나다, 온타리오 케이1브이1씨1, 오타와, 2엔디 플로워, 310 헌트클럽 로드 이스트
(72) 발명자
앤더슨, 데이비드, 이벤더
미국, 메사추세츠 02458, 뉴톤, 엘리엇 메모리얼 로드 34
플루키거, 앤-캐서린
프랑스, 세인트 제니스 레 올레레스 69290, 비스 루 드 라 가렌느 11
(74) 대리인
특허법인 동원

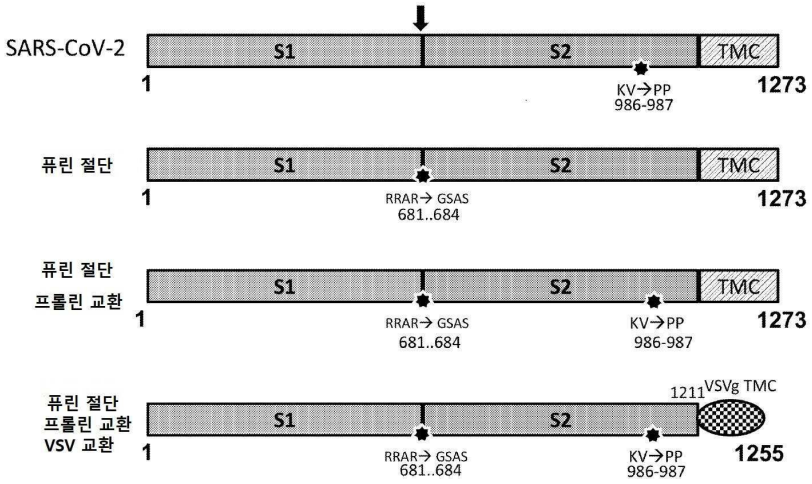
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **코로나바이러스 치료용 백신 조성물**

(57) 요 약

본 발명은 코로나바이러스 감염을 예방 및/또는 치료하는데 유용한 조성물 및 방법을 제공한다. 본 명세서에 기술된 바와 같이, 그 조성물 및 방법은, 하나 이상의 몰로니 뮤린 백혈병 바이러스(MMLV) 코어 단백질을 포함하여 구성되고 그리고 예를 들어, SARS-Cov-2 스파이크 단백질로부터의 하나 이상의 코로나바이러스 에피토프를 포함하는, 바이러스-유사 입자(VLP)를 포함하는 면역원성 조성물의 개발을 기반으로 한다

대 표 도 - 도3



(52) CPC특허분류
A61P 31/14 (2018.01)
A61P 37/04 (2018.01)
C07K 14/005 (2013.01)
C12N 15/86 (2013.01)
A61K 2039/5256 (2013.01)
A61K 2039/5258 (2013.01)
C12N 2740/13022 (2013.01)
C12N 2740/13023 (2013.01)
C12N 2740/13034 (2013.01)

명 세 서

청구범위

청구항 1

SEQ ID NO: 1의 아미노산 서열과 적어도 95% 동일성을 갖는 뮤린 백혈병 바이러스(MLV)에서 발견되는 Gag 단백질인 제1 폴리펩타이드;

베타 코로나바이러스로부터의 스파이크 당단백질인 적어도 하나의 추가 폴리펩타이드; 및

약제학적으로 허용되는 담체;를 포함하여 구성되는 바이러스-유사 입자(VLP)를 포함하여 구성되는, 면역원성 조성물:.

청구항 2

제1항에 있어서, 적어도 하나의 추가 폴리펩타이드가 SARS-CoV-2, SARS-CoV 또는 MERS-CoV로부터의 스파이크 당단백질인, 면역원성 조성물.

청구항 3

제2항에 있어서, 상기 스파이크 당단백질이 야생형 단백질인, 면역원성 조성물.

청구항 4

제2항에 있어서, 상기 스파이크 당단백질이 변형된 단백질인, 면역원성 조성물.

청구항 5

제2항에 있어서, 2개의 스파이크 당단백질을 포함하여 구성되는, 면역원성 조성물.

청구항 6

제2항에 있어서, 3개의 스파이크 당단백질을 포함하여 구성되는, 면역원성 조성물.

청구항 7

제4항에 있어서, 상기 변형된 단백질이 퓨린 절단 부위에 결실을 갖는, 면역원성 조성물.

청구항 8

제7항에 있어서, 상기 변형된 단백질이 VSV로부터의 막횡단 부위를 갖는, 면역원성 조성물.

청구항 9

제4항에 있어서, 상기 변형된 단백질이 라이신 잔기 및 프롤린 잔기로 대체된 발린 잔기를 갖는, 면역원성 조성물.

청구항 10

제9항에 있어서, 상기 변형된 단백질이 VSV로부터의 막횡단 부위를 갖는, 면역원성 조성물.

청구항 11

제4항에 있어서, 상기 변형된 단백질이 프롤린 잔기로 대체된 라이신 잔기 및 발린 잔기를 갖고 퓨린 절단 부위에 결실을 갖는, 면역원성 조성물.

청구항 12

제4항에 있어서, 상기 변형된 단백질이 프롤린 잔기로 대체된 라이신 잔기 및 발린 잔기를 갖고, 그리고 퓨린

절단 부위에서 결실을 가지며, 그리고 VSV로부터 유래한막횡단 부위를 갖는, 면역원성 조성물.

청구항 13

제1항에 있어서, 상기 추가 폴리펩타이드가 SEQ ID NO: 22의 아미노산 서열을 갖는, 면역원성 조성물.

청구항 14

제1항에 있어서, 애주번트를 추가로 포함하여 구성되는, 면역원성 조성물.

청구항 15

제14항에 있어서, 상기 애주번트가 사이토카인, 겔형 애주번트, 미생물 애주번트, 오일-에멀젼 및 유화제 기반 애주번트, 미립자 애주번트, 합성 애주번트, 폴리머 애주번트 및/또는 이들의 조합으로 구성되는 군으로부터 선택되는, 면역원성 조성물.

청구항 16

제15항에 있어서, 상기 미립자 애주번트가 알루미늄 염인, 면역원성 조성물.

청구항 17

제1항에 있어서, 상기 VLP가, SEQ ID NO: 3의 염기 서열을 포함하여 구성되는 제1 벡터 및 SEQ ID NO: SEQ ID NO: 9, 12, 15, 18, 21, 24 또는 27의 염기 서열을 포함하여 구성되는 제2 벡터로 숙주 세포를 공동-형질 감염시키는 단계, 및

벡터에 의해 암호화된 단백질의 발현을 허용하는 조건 하에 적합한 배지에서 숙주 세포를 배양하는 단계;에 의해 생성된, 면역원성 조성물.

청구항 18

제1항에 있어서, 상기 VLP가, SEQ ID NO: 3의 염기 서열을 포함하여 구성되는 제1 벡터; SEQ ID NO: 6의 염기 서열을 포함하여 구성되는 제2 벡터; SEQ ID NO: 9의 염기 서열을 포함하여 구성되는 제3 벡터; 및 SEQ ID NO: 12의 염기 서열을 포함하는 제4 벡터;로 숙주 세포를 공동-형질 감염시키는 단계; 및 벡터에 의해 암호화된 단백질의 발현을 허용하는 조건 하에 적합한 배지에서 숙주 세포를 배양하는 단계;에 의해 생성된, 면역원성 조성물.

청구항 19

제1항에 있어서, 상기 VLP가 SEQ ID NO: 3의 염기 서열을 포함하여 구성되는 제1 벡터 및 SEQ ID NO: 25의 염기 서열을 포함하여 구성되는 제2 벡터로 숙주 세포를 공동-형질 감염시키는 단계; 및 벡터에 의해 암호화된 단백질의 발현을 허용하는 조건 하에 적합한 배지에서 숙주 세포를 배양하는 단계;에 의해 생성된, 면역원성 조성물.

청구항 20

제1항의 약제학적 조성물을 대상체에게 투여하는 것을 포함하여 구성되는, 코로나바이러스 감염을 갖거나 감염 위험이 있는 대상체의 치료 방법.

발명의 설명

기 술 분 야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 2020년 3월 30일에 출원되고, 그 전체 내용이 본 명세서에 합되는, 미국 가출원 63/002,237호 및 2020년 8월 25일에 출원된 미국 가출원 63/070,150호의 이익을 주장하는 출원이다.

[0003] 서열 목록

[0004] 본 출원은 ASCII 포멧으로 전자적으로 제출된 서열 목록을 포함하며, 그 전체가 여기에 통합된다. 상기 ASCII

사본은 2021년 4월 28일에 만들어졌으며, 2007801-0140 SL.txt로 명명되고 크기가 184,699 바이트이다.

[0005] 기술분야

[0006] 본 발명은 백신 분야, 특히 코로나바이러스에 대한 바이러스 유사 입자에 관한 것이다.

배 경 기 술

[0007] 코로나바이러스는 직경이 160-180 nm 범위이고 양성 가닥 RNA 게놈을 함유하는 구형의 외피 바이러스이다. 약 30,000개 염기의 게놈을 가진 코로나바이러스는 알려진 RNA 바이러스 중 가장 큰 것으로 생각된다. 인플루엔자 바이러스와 마찬가지로 그들은 코로나바이러스 계열의 다른 구성원과 유전적으로 재조합하는 능력을 가지고 있다. 코로나바이러스는 4가지 주요 속으로 나뉜다. 코로나바이러스는 많은 동물에서 감염성 기관지염 바이러스, 고양이 감염성 복막염 바이러스 및 전염성 위장염 바이러스와 같은 여러 심각한 질병의 원인 병원체로 믿어진다. 코로나바이러스는 또한 일반 감기로부터 중증 호흡기 감염에 이르기까지 인간에게 다양한 질병을 유발한다. 4가지 인간 코로나바이러스, HCoV-OC43, HCoV-HKU1(베타코로나바이러스) 및 HCoV-NL63, HCoV-229E(알파코로나바이러스)는 일반 감기의 15%-30%에 기여한다[Fung et al (2019) Annu. Rev. Microbiol. 73:2-529-557]. 최근 몇 년 동안 베타 코로나바이러스는 인간에게 세 가지 중요한 질병 발병의 원인이 되었다.

[0008] 2000년대 초, SARS-CoV로 알려진 베타 코로나바이러스는 중증 급성 호흡기 증후군(SARS)이라고 하는 호흡기 질환의 발병을 일으켰다. 주요 증상은 발열, 마른 기침, 두통, 숨가쁨, 호흡곤란이었다. 감염된 사람들 중 많은 사람들이 바이러스성 폐렴으로 발전하여 하기도 감염을 일으켰다. SARS는 전염성이 매우 높으며 기침이나 재채기로 인한 비말 또는 분변 오염과 같은 다른 방법을 통해 전파된다. SARS는 전체 사례의 약 9.14%에서 치명적이었다. SARS의 세계적인 발병은 2003년 7월에 억제되었으며 2004년 이후 보고된 사례가 없다(Peeri et al Int. J. Epi, 2020년 2월 10일).

[0009] 2012년에, 현재 중동 호흡기 증후군 코로나바이러스(MERS-CoV)로 알려진 또 다른 새로운 코로나바이러스가 사우디아라비아에서 나타났다. MERS-CoV는 또한 베타 코로나바이러스이다. 이후 MERS-CoV 감염 사례가 보고되었으며 중동, 유럽, 아시아, 북미 등 27개국으로 확산되었다. MERS-CoV 감염은 발열, 기침, 숨가쁨 등의 증상을 보이는 중증 급성 호흡기 질환으로 제시되었다. 보고된 MERS-CoV 감염 사례의 약 34%가 사망으로 이어졌다. 보고된 사례 중 경미한 호흡기 질환이 있는 환자와 관련된 사례는 소수에 불과했다.

[0010] 2019년 말, 중국 우한에서 호흡기 감염이 나타났는데, 이는 SARS-CoV-2라는 새로운 코로나바이러스 변종에 의해 유발된 것으로 빠르게 확인되었다. COVID-19로 알려진 감염은 전염성이 강하고 특히 노인 환자에게 심각한 폐렴을 유발한다. 사망률은 국가별로 크게 다르며 이탈리아의 13.7%에서 일본의 1.9%까지 추정된다. 2021년 3월 현재 미국의 치사율은 약 1.8%이다(JJohns Hopkins Coronavirus Research Centre, Update as of March 30, 2021). COVID-19는 전 세계적으로 빠르게 확산되어 인류의 건강에 중대한 위협이 되었고 경제 활동이 크게 위축되었다. 2021년 2월 1일 기준으로 1억 명이 넘는 사람들이 COVID-19에 감염되었고 2백만 명이 넘는 사람들이 사망했다.

[0011] 2020년 말에 COVID-19에 대한 여러 백신이 긴급 사용을 위해 승인되었다. 이 백신은 스파이크 단백질로 알려진 SARS-CoV-2 표면의 단백질을 표적으로 하는 새로운 플랫폼을 활용하며 때로는 처음으로 인간에게 사용된다. 이 백신은 임상 시험에서 매우 효과적인 것으로 나타났지만 제조상의 어려움과 경우에 따라 초저온 보관 요구 사항으로 인해 세계 여러 지역에서 유통이 느렸다. 또한 몇 가지 새로운 백신이 안전한 것으로 입증되었지만 일부는 특정 국가에서 사용을 제한하는 드물지만 치명적인 부작용과 관련이 있다.

[0012] 2020년 하반기 동안 COVID-19 질병을 유발하는 SARS-CoV-2 변종이 나타났다. B.1.1.7(영국 변종이라고도 함), 501Y.V2(남아프리카 변종이라고도 함) 및 P.1(브라질 변종이라고도 함)의 세 가지 변종이 등장한 국가에서 빠르게 지배종이 되었다. 이러한 변종는 안지오텐신 전환 효소(angiotensin-converting enzyme) 2(ACE2)로 알려진 수용체에 대한 바이러스 수용체 결합 도메인의 증가된 결합 친화성으로 인해 매우 전염성이 있는 것으로 입증되었다. 새로운 변이의 급속한 확산과 새로운 변이의 출현 가능성은 SARS-CoV-2의 원래 변종에 대해 개발된 최근 승인된 백신의 재감염 및 효과에 대한 상당한 우려를 불러일으켰다.

[0013] 이에 따라, SARS-CoV-2에 대한 강력한 면역성을 유도하면서도 안전하고 보관 및 유통이 용이한 새로운 백신의 개발이 시급한 실정이다. 또한, SARS-CoV-2에 대한 백신이 광범위한 면역을 제공하여 돌연변이 형태의 바이러스로부터 환자를 보호할 수 있도록 해야 한다.

발명의 내용

해결하려는 과제

[0014] 따라서, 코로나바이러스 항원에 대한 광범위한 면역을 제공하는 인간 코로나바이러스에 대한 백신이 필요하다.

과제의 해결 수단

[0015] 본 발명은 인간 코로나바이러스에 의한 감염의 예방에 유용한 방법 및 조성물을 제공한다. 보다 구체적으로, 본 발명은 코로나바이러스에 의해 유발된 감염의 예방, 치료 및/또는 진단에 유용한 인간 코로나바이러스로부터 항원을 발현하는 바이러스 유사 입자(VLP)의 생산 방법, 및 이를 포함하는 조성물을 제공한다.

[0016] 본 발명은 하나 이상의 몰로니 뮤린 백혈병 바이러스(Moloney Murine leukemia virus, MMLV) 코어 단백질을 포함하여 구성되고 지질 이중층 막(lipid bilayer membrane)에 의해 둘러싸인 바이러스-유사 입자를 제공한다. VLP는 바이러스 중화 항체의 유도에서 역할을 하는 인간 코로나바이러스로부터의 하나 이상의 외피 폴리펩타이드(예를 들어, 하나 이상의 코로나바이러스 폴리펩타이드 에피토프)를 포함한다.

[0017] 일부 구현예에서, 본 발명은 야생형(wild type) 인간 코로나바이러스 외피 당단백질을 포함하는 외피를 갖는 VLP를 제공한다. 일부 구현예에서, 폴리펩타이드는 SARS-CoV로부터 유래한다. 일부 구현예에서, 폴리펩타이드는 MERS-CoV로부터 유래한다. 일부 구현예에서, 폴리펩타이드는 SARS-CoV-2에서 유래한다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 하나 보다 많은 폴리펩타이드를 포함한다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 3가지 모두로부터의 폴리펩타이드를 포함한다.

[0018] 일부 구현예에서, 본 발명은 변형된(modified) 인간 코로나바이러스 외피 당단백질을 포함하여 구성되는 외피를 갖는 VLP를 제공한다. 한 구현예에서, 본 발명은 미성숙 형태(premature conformation)의 코로나바이러스 폴리펩타이드를 포함하는 외피를 갖는 VLP의 생산을 포함한다. 특정 구체예에서, 변형된 외피 당단백질은 퓨린 절단 부위(furin cleavage site)가 결여되어 있다. 일부 구현예에서, 퓨린 절단 부위가 결여된 폴리펩타이드는 SARS-CoV로부터 유래한다. 일부 구현예에서, 퓨린 절단 부위가 결여된 폴리펩타이드는 MERS-CoV로부터 유래한다. 일부 구현예에서, 퓨린 절단 부위가 결여된 폴리펩타이드는 SARS-CoV-2로부터 유래한다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 하나 보다 많은 폴리펩타이드를 포함하며, 여기서 폴리펩타이드는 퓨린 절단 부위가 없다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 3개 모두로부터의 폴리펩타이드를 포함하며, 여기서 폴리펩타이드는 퓨린 절단 부위가 없다.

[0019] 또 다른 구현예에서, 본 발명은 변형된 아미노산 서열을 갖는 코로나바이러스 폴리펩타이드를 포함하는 외피를 갖는 VLP의 생산을 포함한다. 특정 구체예에서, 변형된 외피 당단백질은 프롤린 잔기(proline residues)에 대해 치환된 라이신 및 발린 잔기를 갖는다. 일부 구현예에서, 프롤린 치환(proline substitution)을 갖는 폴리펩타이드는 SARS-CoV로부터 유래한다. 일부 구현예에서, 프롤린 치환을 갖는 폴리펩타이드는 MERS-CoV로부유래한다. 일부 구현예에서, 프롤린 치환을 갖는 폴리펩타이드는 SARS-CoV-2로부터 유래한다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 하나보다 많은 폴리펩타이드를 포함하며, 여기서 폴리펩타이드는 프롤린 치환을 갖는다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 3가지 모두로부터의 폴리펩타이드를 포함하며, 여기서 폴리펩타이드는 프롤린 치환을 갖는다.

[0020] 다른 구현예에서, 본 발명은 변형된 아미노산 서열 및 미성숙 형태를 갖는 코로나바이러스 폴리펩타이드를 포함하는 외피를 갖는 VLP의 생산을 포함한다. 특정 구현예에서, 변형된 외피 당단백질은 프롤린 잔기로 치환된 라이신 및 발린 잔기를 갖고 퓨린 절단 부위가 없다. 일부 구현예에서, 프롤린 치환을 갖고 퓨린 절단 부위가 없는 폴리펩타이드는 SARS-CoV에서 유래한다. 일부 구현예에서, 프롤린 치환을 갖고 퓨린 절단 부위가 결여된 폴리펩타이드는 MERS-CoV로부터 유래한다. 일부 구현예에서, 프롤린 치환을 갖고 퓨린 절단 부위가 결여된 폴리펩타이드는 SARS-CoV-2로부터 유래한다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 하나보다 많은 폴리펩타이드를 포함하며, 여기서 폴리펩타이드는 프롤린 치환을 갖고 퓨린 절단 부위가 없다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 3개 모두로부터의 폴리펩타이드를 포함하며, 여기서 폴리펩타이드는 프롤린 치환을 갖고 퓨린 절단 부위가 없다.

[0021] 다른 구현예에서, 변형된 외피 당단백질은 막횡단 부위(transmembrane domain)가 다른 바이러스의 막횡단 부위로 대체되도록 변형되었다. 특히 바람직한 구현예에서, VLP는 단리된(isolated) 코로나바이러스 S 단백질을 포함하여 구성되는 변형된 외피 당단백질을 가지며, 이 단백질의 막횡단 부위 및 세포질 꼬리는 수포성 구내염 바

이러스(VSV)로부터의 막횡단 부위 및 세포질 꼬리로 대체되었다. 일부 구현예에서, VSV로부터의 막횡단 부위 및 세포질 꼬리를 갖는 폴리펩타이드는 SARS-CoV로부터 유래한다. 일부 구현예에서, VSV로부터의 막횡단 부위 및 세포질 꼬리를 갖는 폴리펩타이드는 MERS-CoV로부터 유래한다. 일부 구현예에서, VSV로부터 유래한막횡단 부위 및 세포질 꼬리를 갖는 폴리펩타이드는 SARS-CoV-2로부터 유래한다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 하나보다 많은 폴리펩타이드를 포함하고, 여기서 폴리펩타이드는 VSV의 막횡단 부위 및 세포질 꼬리를 갖는다. 일부 구현예에서, VLP는 SARS-CoV, MERS-CoV 및 SARS-CoV-2 중 3개 모두로부터 유래한 폴리펩타이드를 포함하고, 여기서 폴리펩타이드는 VSV로부터의 막횡단 부위 및 세포질 꼬리를 갖는다. 일부 구현예에서, VLPS는 SARS-CoV, MERS-CoV 및 SARS-CoV-2로부터의 하나 이상의 폴리펩타이드를 포함하고, 이들 중 하나 이상이 본 명세서에 기술된 바와 같이 변형되고 VSV로부터의 막횡단 부위 및 세포질 꼬리를 갖는다.

[0022] 바람직한 구현예에서, 본 발명은 변형된 아미노산 서열 및 미성숙 형태를 갖는 SAR-CoV-2 스파이크 폴리펩타이드를 포함하는 외피를 갖는 VLP의 생산을 포함한다. 변형된 외피 당단백질은 프롤린 잔기로 치환된 라이신 및 발린 잔기를 가지며 퓨린 절단 부위가 없다. 또한, 변형된 스파이크 당단백질은 막횡단 부위 및 세포질 꼬리가 수포성 구내염 바이러스(VSV)의 막횡단 부위 및 세포질 꼬리로 대체되도록 추가로 변형되었다.

[0023] 본 발명은 하나 이상의 변형된 인간 코로나바이러스 외피 단백질 및 하나 이상의 야생형 인간 코로나바이러스 단백질을 포함하는 2가 및 3가 VLP를 추가로 제공한다.

[0024] 본 발명의 다른 특징, 목적 및 이점은 다음의 상세한 설명에서 명백하다. 그러나, 상세한 설명은 본 발명의 구현예를 나타내면서 제한이 아니라 단지 예시로서 제공된다는 것을 이해해야 한다. 본 발명의 범위 내에서 다양한 변경 및 수정은 상세한 설명으로부터 이 분야의 통상의 기술자에게 명백할 것이다.

도면의 간단한 설명

[0025] 도면은 제한이 아니라 설명을 위한 것이다.

도 1은 SARS-CoV-2 외피의 구조를 도시한 도면이다.

도 2는 SARS-CoV, SARS CoV-2 및 MERS-CoV의 S1/S2 도메인을 보여준다.

도 3은 예시적인 대안적 COVID-S 컨스트럭트(construct)를 예시하는 다이어그램이다. 도 3은 SEQ ID NO:43으로서 'RRAR'과 SEQ ID NO:4로서 'GSAS'를 개시한다.

서열 목록

다음은 본 명세서에 언급된 서열의 목록이다:

SEQ ID NO: 1은 MMLV-Gag 아미노산 서열이다.

MGQTVTTPLSLTLGHWKDVERIAHNQSVDVKKRRWVTFCSAEWPTFNVGWPRDGTFNRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVH
PKPPPPLPPSAPSLPLEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLLTEDPPPYRDPRPPPSDRDGNGGEATPAGEAPDPSPMASRLRGRREP
PVADSTTSQAFPLRAGGNGQLQYWPFSSSDLYNWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLLGTLLTGEEKQRVLLEARKAVRGDDGRPTQLPN
EVDAAFPLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNLAKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRK
LERLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRRHREMSKLLATVVSGQKQDRQGGERRRSQLDRDQCAYCKEK
GHWAKDCPKKPRGPRGPRPQTSLLTLDD

SEQ ID NO: 2는 MMLV-Gag 염기 서열이다

ATGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTCGAGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGT
TGGGTTACCTTCTGCTCTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACCTTTAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTC
TTTTCACCTGGCCCGCATGGACACCCAGACCAGGTCCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCCTGGGTCAAGCCCTTTGTACAC
CCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCTCTCCCCCTTGAACCTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCT
TCTCTAGGCGCCAAACCTAAACCTCAAGTTCTTTCTGACAGTGGGGGGCCGCTCATCGACCTACTTACAGAAGACCCCCCGCCTTATAGGGACCCAAGACCA
CCCCCTTCCGACAGGGACGGAAATGGTGGAGAAGCGACCCCTGCGGGAGAGGCACCGGACCCCTCCCCAATGGCATCTCGCCTACGTGGGAGACGGGAGCCC
CCTGTGGCCGACTCCACTACCTCGCAGGCATTCCCCCTCCGCGCAGGAGGAAACGGACAGCTTCAATACTGGCCGTTCTCCTCTTCTGACCTTTACAACTGG
AAAAATAATAACCCTTCTTTTTCTGAAGATCCAGGTAAACTGACAGCTCTGATCGAGTCTGTTCTCATCACCCATCAGCCCACCTGGGACGACTGTCAGCAG
CTGTTGGGGACTCTGCTGACCGGAGAAGAAAAACAACGGGTGCTCTTAGAGGCTAGAAAGGCGGTGCGGGGCGATGATGGGCGCCCCACTCAACTGCCCAAT
GAAGTCGATGCCGCTTTTCCCCTCGAGCGCCCAGACTGGGATTACACCACCCAGGCAGGTAGGAACCACCTAGTCCACTATCGCCAGTTGCTCCTAGCGGGT
CTCCAAAACGCGGGCAGAAGCCCCACCAATTTGGCCAAGGTAAAAGGAATAACACAAGGGCCCAATGAGTCTCCCTCGGCCTTCCTAGAGAGACTTAAGGAA

GCCTATCGCAGGTACACTCCTTATGACCCTGAGGACCCAGGGCAAGAAACTAATGTGTCTATGTCTTTCATTTGGCAGTCTGCCCCAGACATTGGGAGAAAG
TTAGAGAGGTTAGAAGATTTAAAAAACAAGACGCTTGGAGATTTGGTTAGAGAGGCAGAAAAGATCTTTAATAAACGAGAAACCCCGGAAGAAAGAGAGGAA
CGTATCAGGAGAGAAACAGAGGAAAAAGAAGAACGCCGTAGGACAGAGGATGAGCAGAAAGAGAAAGAAAGAGATCGTAGGAGACATAGAGAGATGAGCAAG
CTATTGGCCACTGTCGTTAGTGGACAGAAACAGGATAGACAGGGAGGAGAACGAAGGAGGTCCCAACTCGATCGCGACCAGTGTGCCTACTGCAAAGAAAAG
GGGCACTGGGCTAAAGATTGTCCCAAGAAACCACGAGGACCTCGGGGACCAAGACCCCAGACCTCCCTCCTGACCCTAGATGAC

SEQ ID NO: 3은 코돈 최적화된 MMLV-Gag 염기 서열이다

ATGGGACAGACCGTCACAACACCCCTGAGCCTGACCCTGGGACATTGGAAAGACGTGGAGAGGATCGCACATAACCAGAGCGTGGACGTGAAGAAACGGAGA
TGGGTCACATTCTGCAGTGCTGAGTGGCCAACTTTTAATGTGGGATGGCCCCGAGACGGCACTTTCAACAGGGATCTGATCACCCAGGTGAAGATCAAGGTC
TTTAGCCCAGGACCTCACGGACATCCAGACCAGGTGCCTTATATCGTCACCTGGGAGGCACTGGCCTTCGATCCCCCTCCATGGGTGAAGCCATTTGTCCAC
CCAAAACCACCTCCACCACTGCCTCCAAGTGCCCCTTCACTGCCACTGGAACCACCCCGGAGCACACCACCCCGCAGCTCCCTGTATCCTGCTCTGACTCCA
TCTCTGGGCGCAAAGCCAAAACCACAGGTGCTGAGCGACTCCGGAGGACCACTGATTGACCTGCTGACAGAGGACCCCCCACCATACCGAGATCCTCGGCCT
CCACCAAGCGACCGCGATGGAAATGGAGGAGAGGCTACTCCTGCCGGCGAAGCCCCTGACCCATCTCCAATGGCTAGTAGGCTGCGCGGCAGGCGCGAGCCT
CCAGTGGCAGATAGCACCACATCCCAGGCCTTCCCTCTGAGGGCTGGGGGAAATGGGCAGCTCCAGTATTGGCCATTTTCTAGTTCAGACCTGTACAACTGG
AAGAACAATAACCCCTCTTTCAGTGAGGACCCCGGCAAACTGACCGCCCTGATCGAATCCGTGCTGATTACCCATCAGCCCACATGGGACGATTGTCAGCAG
CTCCTGGGCACCCTGCTGACCGGAGAGGAAAAGCAGCGCGTGCTGCTGGAGGCTCGCAAAGCAGTCCGAGGGGACGATGGACGGCCCACACAGCTCCCTAAT
GAGGTGGACGCCGCTTTTCCACTGGAAAGACCCGACTGGGATTATACTACCCAGGCAGGGAGAAACCACCTGGTCCATTACAGGCAGCTCCTGCTGGCAGGC
CTGCAGAATGCCGGGAGATCCCCCACCAACCTGGCCAAGGTGAAAGGCATCACACAGGGGCCTAATGAGTCACCAAGCGCCTTTCTGGAGAGGCTGAAGGAA
GCTTACCGACGGTATACCCCATACGACCCTGAGGACCCCGGACAGGAAACAAACGTCTCCATGTCTTTCATCTGGCAGTCTGCCCCAGACATTGGGCGGAAG
CTGGAGAGACTGGAAGACCTGAAGAACAAGACCCTGGGCGACCTGGTGCGGGAGGCTGAAAAGATCTTCAACAAACGGGAGACCCCCGAGGAAAGAGAGGAA
AGGATTAGAAGGGAAACTGAGGAAAAGGAGGAACGCCGACGGACCGAGGACGAACAGAAGGAGAAAGAAACGAGATCGGCGGCGGCACCGGGAGATGTCAAAG
CTGCTGGCCACCGTGGTCAGCGGACAGAAACAGGACAGACAGGGAGGAGAGCGACGGAGAAGCCAGCTCGACAGGGATCAGTGCGCATACTGTAAGGAAAAA
GGCCATTGGGCCAAGGATTGCCCCAAAAAGCCAAGAGGACCAAGAGGACCAAGACCACAGACATCACTGCTGACCCTGGACGAC

SEQ ID NO: 4는 SARS-CoV-2 스파이크 당단백질, 아미노산 서열이다

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIY
SKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV
VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAI
PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
NQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA
SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
SGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 5는 SARS-CoV-2 스파이크 당단백질, 염기 서열(Wuhan-Hu-1: Genbank Ref: MN908947)이다.

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaattctttcacacgt
ggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttctttttccaatgttacttggttccatgctata
catgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatggtgtttattttgcttccactgagaagtctaacataataaga
ggctggatttttggtactactttagattcgaagacccagtccctacttattgttaataacgctactaatgttgttattaaagtctgtgaatttcaatttttgt
aatgatccatttttgggtgtttattaccacaaaaacaacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatat
gtctctcagccttttcttatggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttaaaatatat
tctaagcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaacatcactaggttt
caaacttttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcagcttattatgtgggttatcttcaacct
aggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcacttgaccctctctcagaaacaaagtgtacgttgaaatccttc
actgtagaaaaaggaatctatcaaacttctaactttagagtccaaccaacagaatctattgttagatttcctaatattacaaacttgtgccctttttggtgaa
gtttttaacgccaccagatttgcatctgtttatgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattt

tccacttttaagtgttatggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcaga
caaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaattctaacaatcttgat
tctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagagagatatttcaactgaaatctatcaggccggt
agcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatggtttccaacccactaatggtgttggttaccaaccatacagagta
gtagtactttctttgaacttctacatgcaccagcaactgtttgtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaat
ggtttaacaggcacaggtgttcttactgagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgat
ccacagacacttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgttctttat
caggatgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggttctaatgtttttcaaacacgt
gcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtgcaggtatatgcgctagttatcagactcagactaattct
cctcggcgggcacgtagtgtagctagtcaatccatcattgcctacactatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccata
cccacaaattttactattagtgttaccacagaaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaa
tgcagcaatcttttgttgcaatatggcagttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttttt
gcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaaaaccaagcaagaggtca
tttattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtgattgccttggtgatattgctgctagagacctcatt
tgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagatgaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttct
ggttggacctttggtgcaggtgctgcattacaaataccatttgctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgag
aaccaaaaattgattgccaaccaatttaatagtgctattggcaaaattcaagactcactttcttccacagcaagtgcacttggaaaacttcaagatgtggtc
aaccaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatcctttcacgtcttgacaaa
gttgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgactcaacaattaattagagctgcagaaatcagagct
tctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatcaaaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcag
tcagcacctcatggtgtagtcttcttgcatgtgacttatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacac
tttcctcgtgaaggtgtctttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtg
tctggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggagttagataaatatttt
aagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaaaaagaaattgaccgcctcaatgaggttgcc
aagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagcagtatataaaatggccatggtacatttggctaggttttatagctggcttg
attgccatagtaatggtgacaattatgctttgctgtatgaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaa
gacgactctgagccagtgctcaaaggagtcaaattacattacacataa

SEQ ID NO: 6은 SARS-CoV-2 스파이크 당단백질, 염기 서열, 인간 세포에서의 발현에 최적화된 코돈이다.

ATGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGAGCTCCCAGTGCGTGAACCTGACCACAAGGACCCAGCTCCCCCCTGCCTATACCAATTCCTTCACACGG
GGCGTGTACTATCCCGACAAGGTGTTTAGATCTAGCGTGCTGCACTCCACACAGGATCTGTTTCTGCCTTTCTTTTCTAACGTGACCTGGTTCCACGCCATC
CATGTGAGCGGCACCAATGGCACAAAGCGGTTCGACAATCCAGTGCTGCCCTTTAACGATGGCGTGTACTTCGCCTCCACCGAGAAGTCTAACATCATCAGA
GGCTGGATCTTTGGCACCACACTGGACAGCAAGACACAGTCCCTGCTGATCGTGAACAATGCCACCAACGTGGTCATCAAGGTGTGCGAGTTCCAGTTTTGT
AATGATCCATTCCTGGGCGTGTACTATCACAAGAACAATAAGTCTTGGATGGAGAGCGAGTTTCGCGTGTATTCCTCTGCCAACAATTGCACATTTGAGTAC
GTGTCCCAGCCCTTCCTGATGGACCTGGAGGGCAAGCAGGGCAATTTCAAGAACCTGAGGGAGTTCGTGTTTAAGAATATCGATGGCTACTTCAAGATCTAC
TCCAAGCACACCCCAATCAACCTGGTGCGCGACCTGCCACAGGGCTTCTCTGCCCTGGAGCCACTGGTGGATCTGCCCATCGGCATCAACATCACCCGGTTT
CAGACACTGCTGGCCCTGCACAGAAGCTACCTGACACCAGGCGACAGCTCCTCTGGATGGACCGCCGGGGCCGCCGCCTACTATGTGGGCTATCTGCAGCCC
AGGACCTTCCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTGCGCCCTGGACCCCCTGTCTGAGACCAAGTGTACACTGAAGAGCTTT
ACCGTGGAGAAGGGCATCTATCAGACAAGCAATTTCAGGGTGCAGCCTACCGAGTCCATCGTGCGCTTTCCCAATATCACAAACCTGTGCCCTTTTGGCGAG
GTGTTCAACGCAACCAGGTTCGCCAGCGTGTACGCATGGAATAGGAAGCGCATCTCCAACTGCGTGGCCGACTATTCTGTGCTGTACAACAGCGCCTCCTTC
TCTACCTTTAAGTGCTATGGCGTGAGCCCCACAAAGCTGAATGACCTGTGCTTTACCAACGTGTACGCCGATTCCTTCGTGATCAGGGGCGACGAGGTGCGC
CAGATCGCACCAGGACAGACAGGCAAGATCGCAGACTACAATTATAAGCTGCCTGACGATTTCACCGGCTGCGTGATCGCCTGGAACTCTAACAATCTGGAT
AGCAAAGTGGGCGGCAACTACAATTATCTGTACCGGCTGTTTAGAAAGTCTAATCTGAAGCCATTCGAGAGGGACATCTCCACAGAGATCTACCAGGCCGGC
TCTACCCCCTGCAATGGCGTGGAGGGCTTTAACTGTTATTTCCCTCTGCAGAGCTACGGCTTCCAGCCAACAAACGGCGTGGGCTATCAGCCCTACCGCGTG
GTGGTGCTGTCTTTTGAGCTGCTGCACGCACCTGCAACAGTGTGCGGACCAAAGAAGAGCACCAATCTGGTGAAGAACAAGTGCGTGAACTTCAACTTCAAC
GGCCTGACCGGCACAGGCGTGCTGACCGAGTCCAACAAGAAGTTCCTGCCTTTTCAGCAGTTCGGCAGGGACATCGCAGATACCACAGACGCCGTGCGCGAC
CCTCAGACCCTGGAGATCCTGGACATCACACCATGCTCCTTCGGCGGCGTGTCTGTGATCACACCAGGCACCAATACAAGCAACCAGGTGGCCGTGCTGTAT
CAGGACGTGAATTGTACCGAGGTGCCCGTGGCCATCCACGCAGATCAGCTCACCCCTACATGGCGGGTGTACTCTACCGGCAGCAACGTGTTCCAGACAAGA
GCCGGCTGCCTGATCGGAGCCGAGCATGTGAACAATAGCTATGAGTGCGACATCCCTATCGGAGCCGGCATCTGTGCCTCCTACCAGACCCAGACAAACTCC
CCACGGAGAGCCCGGTCTGTGGCCAGCCAGTCCATCATCGCCTATACCATGAGCCTGGGGGCCGAGAACAGCGTGGCCTACTCCAACAATTCTATCGCCATC
CCTACCAACTTCACAATCTCCGTGACCACAGAGATCCTGCCAGTGAGCATGACCAAGACATCCGTGGACTGCACAATGTATATCTGTGGCGATTCCACCGAG

TGCTCTAACCTGCTGCTGCAGTACGGCTCTTTTTGTACCCAGCTCAACAGAGCCCTGACAGGCATCGCCGTGGAGCAGGACAAGAACACACAGGAGGTGTTC
GCCCAGGTGAAGCAGATCTACAAGACCCCACCCATCAAGGACTTTGGCGGCTTCAACTTCAGCCAGATCCTGCCCGATCCTAGCAAGCCATCCAAGCGGTCT
TTTATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGCCGATGCCGGCTTCATCAAGCAGTATGGCGATTGCCTGGGCGACATCGCCGCCAGAGACCTGATC
TGTGCCCAGAAGTTTAATGGCCTGACCGTGCTGCCTCCACTGCTGACAGATGAGATGATCGCCCAGTACACATCTGCCCTGCTGGCCGGCACCATCACAAGC
GGATGGACCTTCGGGGCCGGGGCCGCCCTGCAGATCCCCTTTGCCATGCAGATGGCCTATCGGTTCAACGGCATCGGCGTGACCCAGAATGTGCTGTACGAG
AACCAGAAGCTGATCGCCAATCAGTTTAACTCCGCCATCGGCAAGATCCAGGACTCTCTGAGCTCCACAGCCAGCGCCCTGGGCAAGCTGCAGGATGTGGTG
AATCAGAACGCCCAGGCCCTGAATACCCTGGTGAAGCAGCTCAGCAGCAACTTCGGGGCCATCAGCAGCGTGCTGAACGACATCCTGAGCCGGCTGGACAAG
GTGGAGGCAGAGGTGCAGATCGACCGGCTGATCACAGGCAGACTGCAGTCCCTGCAGACCTACGTGACACAGCAGCTCATCAGGGCCGCCGAGATCAGGGCC
TCTGCCAATCTGGCCGCCACCAAGATGAGCGAGTGCGTGCTGGGCCAGTCCAAGAGAGTGGACTTTTGTGGCAAGGGCTATCACCTGATGAGCTTCCCACAG
TCCGCCCCTCACGGAGTGGTGTTTCTGCATGTGACCTACGTGCCAGCCCAGGAGAAGAACTTCACCACAGCCCCCGCAATCTGCCACGATGGCAAGGCACAC
TTTCCCCGGGAGGGCGTGTTCGTGAGCAACGGCACCCACTGGTTTGTGACACAGCGCAATTTCTACGAGCCACAGATCATCACCACAGACAATACATTCGTG
TCCGGCAACTGTGACGTGGTCATCGGCATCGTGAACAATACCGTGTATGATCCTCTGCAGCCAGAGCTGGACTCTTTTAAGGAGGAGCTGGATAAGTACTTC
AAGAATCACACCAGCCCCGACGTGGATCTGGGCGACATCTCTGGCATCAATGCCAGCGTGGTGAACATCCAGAAGGAGATCGACAGGCTGAACGAGGTGGCC
AAGAATCTGAACGAGTCCCTGATCGATCTGCAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTGGCCCTGGTATATCTGGCTGGGCTTCATCGCCGGCCTG
ATCGCCATCGTGATGGTGACCATCATGCTGTGCTGTATGACAAGCTGCTGTTCCTGCCTGAAGGGCTGCTGTTCTTGTGGCTCCTGCTGTAAGTTTGATGAG
GACGATAGCGAGCCTGTGCTGAAGGGCGTGAAGCTGCACTACACCTGA

SEQ ID NO: 7은 SARS-CoV 스파이크 당단백질, 아미노산 서열(HKU-39849, Genbank Ref: JN854286.1)이다.

MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWV
FGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPID
VVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFR
VVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADY
NYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATV
CGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHA
DQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMA
KTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFM
KQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQE
SLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK
RVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDP
LQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCS
CLKGACSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 8은 SARS-CoV 스파이크 당단백질, 염기 서열이다.

ATGTTTATTTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTGACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTCAACATACT
TCATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTTAGATCAGACACTCTTTATTTAACTCAGGATTTATTTCTTCCATTTTATTCTAATGTTACAGGG
TTTCATACTATTAATCATACGTTTGGCAACCCTGTCATACCTTTTAAGGATGGTATTTATTTTGCTGCCACAGAGAAATCAAATGTTGTCCGTGGTTGGGTT
TTTGGTTCTACCATGAACAACAAGTCACAGTCGGTGATTATTATTAACAATTCTACTAATGTTGTTATACGAGCATGTAACTTTGAATTGTGTGACAACCCT
TTCTTTGCTGTTTCTAAACCCATGGGTACACAGACACATACTATGATATTCGATAATGCATTTAATTGCACTTTCGAGTACATATCTGATGCCTTTTCGCTT
GATGTTTCAGAAAAGTCAGGTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAAGATGGGTTTCTCTATGTTTATAAGGGCTATCAACCTATAGAT
GTAGTTCGTGATCTACCTTCTGGTTTTAACACTTTGAAACCTATTTTTAAGTTGCCTCTTGGTATTAACATTACAAATTTTAGAGCCATTCTTACAGCCTTT
TCACCTGCTCAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTTGGCTATTTAAAGCCAACTACATTTATGCTCAAGTATGATGAAAATGGTACAATC
ACAGATGCTGTTGATTGTTCTCAAAATCCACTTGCTGAACTCAAATGCTCTGTTAAGAGCTTTGAGATTGACAAAGGAATTTACCAGACCTCTAATTTCAGG
GTTGTTCCCTCAGGAGATGTTGTGAGATTCCCTAATATTACAAACTTGTGTCCTTTTGGAGAGGTTTTTAATGCTACTAAATTCCCTTCTGTCTATGCATGG
GAGAGAAAAAAAATTTCTAATTGTGTTGCTGATTACTCTGTGCTCTACAACTCAACATTTTTTTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGTTG
AATGATCTTTGCTTCTCCAATGTCTATGCAGATTCTTTTGTAGTCAAGGGAGATGATGTAAGACAAATAGCGCCAGGACAAACTGGTGTTATTGCTGATTAT
AATTATAAATTGCCAGATGATTTCATGGGTTGTGTCCTTGCTTGGAATACTAGGAACATTGATGCTACTTCAACTGGTAATTATAATTATAAATATAGGTAT
CTTAGACATGGCAAGCTTAGGCCCTTTGAGAGAGACATATCTAATGTGCCTTTCTCCCCTGATGGCAAACCTTGCACCCCACCTGCTCTTAATTGTTATTGG
CCATTAAATGATTATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTGTAGTACTTTCTTTTGAACTTTTAAATGCACCGGCCACGGTT
TGTGGACCAAAATTATCCACTGACCTTATTAAGAACCAGTGTGTCAATTTTAATTTTAATGGACTCACTGGTACTGGTGTGTTAACTCCTTCTTCAAAGAGA
TTTCAACCATTTCAACAATTTGGCCGTGATGTTTCTGATTTCACTGATTCCGTTCGAGATCCTAAAACATCTGAAATATTAGACATTTCACCTTGCTCTTTT
GGGGGTGTAAGTGTAATTACACCTGGAACAAATGCTTCATCTGAAGTTGCTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAGCAATTCATGCA

GATCAACTCACACCAGCTTGGCGCATATATTCTACTGGAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTATAGGAGCTGAGCATGTCGACACTTCTTAT
GAGTGCGACATTCCTATTGGAGCTGGCATTTGTGCTAGTTACCATACAGTTTCTTTATTACGTAGTACTAGCCAAAAATCTATTGTGGCTTATACTATGTCT
TTAGGTGCTGATAGTTCAATTGCTTACTCTAATAACACCATTGCTATACCTACTAACTTTTCAATTAGCATTACTACAGAAGTAATGCCTGTTTCTATGGCT
AAAACCTCCGTAGATTGTAATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTTCTCCAATATGGTAGCTTTTGCACACAACTAAATCGTGCA
CTCTCAGGTATTGCTGCTGAACAGGATCGCAACACACGTGAAGTGTTCGCTCAAGTCAAACAAATGTACAAAACCCCAACTTTGAAATATTTTGGTGGTTTT
AATTTTTCACAAATATTACCTGACCCTCTAAAGCCAACTAAGAGGTCTTTTATTGAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATG
AAGCAATATGGCGAATGCCTAGGTGATATTAATGCTAGAGATCTCATTTGTGCGCAGAAGTTCAATGGACTTACAGTGTTGCCACCTCTGCTCACTGATGAT
ATGATTGCTGCCTACACTGCTGCTCTAGTTAGTGGTACTGCCACTGCTGGATGGACATTTGGTGCTGGCGCTGCTCTTCAAATACCTTTTGCTATGCAAATG
GCATATAGGTTCAATGGCATTGGAGTTACCCAAAATGTTCTCTATGAGAACCAAAAACAAATCGCCAACCAATTTAACAAGGCGATTAGTCAAATTCAAGAA
TCACTTACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTGTTAACCAGAATGCTCAAGCATTAAACACACTTGTTAAACAACTTAGCTCTAATTTT
GGTGCAATTTCAAGTGTGCTAAATGATATCCTTTCGCGACTTGATAAAGTCGAGGCGGAGGTACAAATTGACAGGTTAATTACAGGCAGACTTCAAAGCCTT
CAAACCTATGTAACACAACAACTAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAAATGTCTGAGTGTGTTCTTGGACAATCAAAA
AGAGTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCCACAAGCAGCCCCGCATGGTGTTGTCTTCCTACATGTCACGTATGTGCCATCCCAGGAG
AGGAACTTCACCACAGCGCCAGCAATTTGTCATGAAGGCAAAGCATACTTCCCTCGTGAAGGTGTTTTTGTGTTTAATGGCACTTCTTGGTTTATTACACAG
AGGAACTTCTTTTCTCCACAAATAATTACTACAGACAATACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAACAACACAGTTTATGATCCT
CTGCAACCTGAGCTTGACTCATTCAAAGAAGAGCTGGACAAGTACTTCAAAAATCATACATCACCAGATGTTGATCTTGGCGACATTTCAGGCATTAACGCT
TCTGTCGTCAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTCGCTAAAAATTTAAATGAATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAA
TATATTAAATGGCCTTGGTATGTTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTCATGGTTACAATCTTGCTTTGTTGCATGACTAGTTGTTGCAGT
TGCCTCAAGGGTGCATGCTCTTGTGGTTCTTGCTGCAAGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAA

SEQ ID NO: 9는 SARS-CoV 스파이크 당단백질, 염기 서열, 인간 세포에서 발현을 위해 최적화된 코돈이다.

ATGTTCATCTTTCTGCTGTTCCTGACCCTGACAAGCGGCTCCGACCTGGATAGGTGCACCACATTTGACGATGTGCAGGCCCCCAACTACACACAGCACACC
AGCTCCATGAGGGGCGTGTACTATCCTGATGAGATCTTCCGCTCTGACACACTGTACCTGACCCAGGACCTGTTCCTGCCTTTTTATAGCAATGTGACAGGC
TTCCACACCATCAATCACACATTTGGCAACCCCGTGATCCCTTTCAAGGATGGCATCTACTTTGCCGCCACCGAGAAGTCTAACGTGGTGCGGGGATGGGTG
TTCGGCAGCACAATGAACAATAAGTCTCAGAGCGTGATCATCATCAACAATAGCACCAACGTGGTCATCAGAGCCTGCAATTTTGAGCTGTGCGACAACCCC
TTCTTTGCCGTGTCCAAGCCTATGGGCACCCAGACACACACCATGATCTTTGATAATGCCTTCAACTGTACCTTTGAGTACATCAGCGATGCCTTTTCCCTG
GACGTGTCTGAGAAGTCCGGCAACTTCAAGCACCTGAGGGAGTTCGTGTTTAAGAATAAGGACGGCTTCCTGTACGTGTATAAGGGCTATCAGCCCATCGAT
GTGGTGCGCGACCTGCCTTCCGGCTTCAACACCCTGAAGCCAATCTTTAAGCTGCCCCTGGGCATCAATATCACCAACTTCAGGGCCATCCTGACAGCCTTT
AGCCCAGCACAGGACATCTGGGGCACCAGCGCCGCCGCCTACTTCGTGGGCTATCTGAAGCCCACCACCTTCATGCTGAAGTACGATGAGAACGGCACAATC
ACCGACGCCGTGGATTGCAGCCAGAATCCACTGGCCGAGCTGAAGTGTTCCGTGAAGTCTTTCGAGATCGACAAGGGCATCTATCAGACCTCCAACTTTAGG
GTGGTGCCATCTGGCGATGTGGTGCGCTTCCCAAATATCACCAACCTGTGCCCCTTCGGCGAGGTGTTTAATGCCACAAAGTTCCCCAGCGTGTACGCCTGG
GAGCGCAAGAAGATCAGCAACTGCGTGGCCGACTACTCCGTGCTGTATAATAGCACCTTCTTCAGCACCTTCAAGTGCTACGGCGTGAGCGCCACCAAGCTG
AATGACCTGTGCTTCTCTAACGTGTATGCCGATAGCTTTGTGGTGAAGGGCGACGATGTGAGGCAGATCGCACCTGGACAGACCGGCGTGATCGCAGACTAC
AACTATAAGCTGCCAGACGATTTCATGGGCTGCGTGCTGGCCTGGAATACACGCAACATCGATGCCACATCCACCGGCAACTACAATTATAAGTACCGGTAT
CTGAGACACGGCAAGCTGCGGCCCTTCGAGAGAGACATCTCCAATGTGCCATTTTCTCCAGATGGCAAGCCATGCACCCCACCTGCCCTGAATTGTTACTGG
CCTCTGAACGACTACGGCTTCTATACCACAACCGGCATCGGCTACCAGCCTTATAGGGTGGTGGTGCTGTCCTTTGAGCTGCTGAACGCACCTGCAACCGTG
TGCGGACCAAAGCTGTCTACAGATCTGATCAAGAATCAGTGCGTGAACTTCAACTTCAACGGCCTGACAGGCACCGGCGTGCTGACCCCTTCTAGCAAGCGG
TTCCAGCCATTTCAGCAGTTCGGCAGAGACGTGAGCGATTTCACCGACTCCGTGCGCGACCCAAAGACATCCGAGATCCTGGACATCAGCCCCTGCTCCTTT
GGCGGCGTGTCTGTGATCACACCTGGCACCAACGCCTCCTCTGAGGTGGCCGTGCTGTACCAGGATGTGAATTGTACCGACGTGAGCACAGCAATCCACGCA
GACCAGCTCACCCCAGCATGGCGGATCTATTCCACCGGCAACAACGTGTTCCAGACACAGGCAGGATGCCTGATCGGAGCCGAGCATGTGGATACAAGCTAC
GAGTGCGACATCCCCATCGGAGCCGGCATCTGTGCCTCTTATCACACCGTGAGCCTGCTGAGATCCACATCTCAGAAGTCTATCGTGGCCTACACCATGAGC
CTGGGGGCCGATAGCTCCATCGCCTATTCCAACAATACCATCGCCATCCCAACAAACTTCAGCATCTCCATCACAACCGAAGTGATGCCCGTGTCCATGGCC
AAGACCTCTGTGGACTGCAACATGTACATCTGTGGCGATAGCACAGAGTGCGCCAATCTGCTGCTGCAGTATGGCTCCTTTTGTACCCAGCTCAACCGGGCC
CTGTCTGGAATCGCCGCCGAGCAGGACAGGAATACACGCGAGGTGTTCGCCCAGGTGAAGCAGATGTACAAGACACCTACCCTGAAGTATTTTGGCGGCTTC
AACTTTTCTCAGATCCTGCCTGATCCACTGAAGCCAACCAAGCGGAGCTTCATCGAGGACCTGCTGTTTAATAAGGTGACACTGGCCGATGCCGGCTTCATG
AAGCAGTACGGCGAGTGCCTGGGCGACATCAACGCCAGAGACCTGATCTGTGCCCAGAAGTTTAATGGCCTGACCGTGCTGCCACCCCTGCTGACAGACGAT
ATGATCGCAGCATATACCGCCGCCCTGGTGTCCGGCACAGCCACCGCCGGCTGGACCTTCGGGGCCGGGGCCGCCCTGCAGATCCCTTTCGCCATGCAGATG
GCCTACCGGTTTAACGGCATCGGCGTGACCCAGAATGTGCTGTATGAGAACCAGAAGCAGATCGCCAATCAGTTTAACAAGGCCATCAGCCAGATCCAGGAG
TCCCTGACAACCACATCTACCGCCCTGGGCAAGCTGCAGGACGTGGTGAATCAGAACGCCCAGGCCCTGAATACACTGGTGAAGCAGCTCAGCAGCAACTTC
GGGGCCATCAGCAGCGTGCTGAACGACATCCTGAGCCGGCTGGACAAGGTGGAGGCAGAGGTGCAGATCGATAGGCTGATCACCGGCAGACTGCAGTCTCTG
CAGACATACGTGACCCAGCAGCTCATCAGGGCCGCCGAGATCAGAGCCAGCGCCAACCTGGCCGCCACAAAGATGTCCGAGTGCGTGCTGGGCCAGTCTAAG

AGGGTGGACTTCTGTGGCAAGGGCTACCACCTGATGTCCTTTCCACAGGCCGCCCCTCACGGAGTGGTGTTCCTGCATGTGACCTATGTGCCTTCTCAGGAG
CGCAACTTTACCACAGCCCCAGCAATCTGCCACGAGGGCAAGGCATACTTCCCCCGGGAGGGCGTGTTCGTGTTTAACGGCACCTCCTGGTTTATCACACAG
AGAAATTTCTTTTCCCCTCAGATCATCACCACAGACAATACCTTCGTGAGCGGCAACTGTGACGTGGTCATCGGCATCATCAACAATACAGTGTACGATCCT
CTGCAGCCAGAGCTGGACAGCTTCAAGGAGGAGCTGGATAAGTACTTCAAGAACCACACCTCCCCCGACGTGGATCTGGGCGACATCAGCGGCATCAATGCC
TCCGTGGTGAACATCCAGAAGGAGATCGACAGACTGAATGAGGTGGCCAAGAATCTGAACGAGTCCCTGATCGATCTGCAGGAGCTGGGCAAGTACGAGCAG
TATATCAAGTGGCCATGGTACGTGTGGCTGGGCTTCATCGCCGGCCTGATCGCCATCGTGATGGTGACCATCCTGCTGTGCTGTATGACATCTTGCTGTAGC
TGCCTGAAGGGAGCCTGCTCCTGTGGCTCTTGCTGTAAGTTTGACGAGGACGATAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTATACCTGA

SEQ ID NO: 10은 MERS-CoV 스파이크 당단백질, 아미노산 서열이다.

MIHSVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIYPQGRTYSNITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFV
ANYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSF
ATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWFGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSD
RKAWAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVECDFSPLLSGTPPQVYNFKRLVFTNCN
YNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRLLSDD
RTEVPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVS
GRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCV
LGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPRSVRSVPGEMRLASIAFNHPIQVDQLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQK
CEQLLREYGQFCSKINQALHGANLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADPGYMQGYDDCMQQGPA
SARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFQ
KVQDAVNNNAQALSKLASELSNTFGAISASIGDIIQRLDVLEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQLAKDKVNECVKAQSKRSGFCGQGTH
IVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTNCIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLG
NSTGIDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYIWLGFIAGLVALALCVFFILCCTGCGTN
CMGKLKCNRCCDRYEEYDLEPHKVHVH

SEQ ID NO: 11은 MERS-CoV 스파이크 당단백질, 염기 서열(EMC/2012, Genbank Ref: JX869059.2)이다.

ATGATACACTCAGTGTTTCTACTGATGTTCTTGTTAACACCTACAGAAAGTTACGTTGATGTAGGGCCAGATTCTGTTAAGTCTGCTTGTATTGAGGTTGAT
ATACAACAGACTTTCTTTGATAAAACTTGGCCTAGGCCAATTGATGTTTCTAAGGCTGACGGTATTATATACCCTCAAGGCCGTACATATTCTAACATAACT
ATCACTTATCAAGGTCTTTTTCCCTATCAGGGAGACCATGGTGATATGTATGTTTACTCTGCAGGACATGCTACAGGCACAACTCCACAAAAGTTGTTTGTA
GCTAACTATTCTCAGGACGTCAAACAGTTTGCTAATGGGTTTGTCGTCCGTATAGGAGCAGCTGCCAATTCCACTGGCACTGTTATTATTAGCCCATCTACC
AGCGCTACTATACGAAAAATTTACCCTGCTTTTATGCTGGGTTCTTCAGTTGGTAATTTCTCAGATGGTAAAATGGGCCGCTTCTTCAATCATACTCTAGTT
CTTTTGCCCGATGGATGTGGCACTTTACTTAGAGCTTTTTATTGTATTCTAGAGCCTCGCTCTGGAAATCATTGTCCTGCTGGCAATTCCTATACTTCTTTT
GCCACTTATCACACTCCTGCAACAGATTGTTCTGATGGCAATTACAATCGTAATGCCAGTCTGAACTCTTTTAAGGAGTATTTTAATTTACGTAACTGCACC
TTTATGTACACTTATAACATTACCGAAGATGAGATTTTAGAGTGGTTTGGCATTACACAAACTGCTCAAGGTGTTCACCTCTTCTCATCTCGGTATGTTGAT
TTGTACGGCGGCAATATGTTTCAATTTGCCACCTTGCCTGTTTATGATACTATTAAGTATTATTCTATCATTCCTCACAGTATTCGTTCTATCCAAAGTGAT
AGAAAAGCTTGGGCTGCCTTCTACGTATATAAACTTCAACCGTTAACTTTCCTGTTGGATTTTTCTGTTGATGGTTATATACGCAGAGCTATAGACTGTGGT
TTTAATGATTTGTCACAACTCCACTGCTCATATGAATCCTTCGATGTTGAATCTGGAGTTTATTCAGTTTCGTCTTTCGAAGCAAAACCTTCTGGCTCAGTT
GTGGAACAGGCTGAAGGTGTTGAATGTGATTTTTCACCTCTTCTGTCTGGCACACCTCCTCAGGTTTATAATTTCAAGCGTTTGGTTTTTACCAATTGCAAT
TATAATCTTACCAAATTGCTTTCACTTTTTTCTGTGAATGATTTTACTTGTAGTCAAATATCTCCAGCAGCAATTGCTAGCAACTGTTATTCTTCACTGATT
TTGGATTACTTTTCATACCCACTTAGTATGAAATCCGATCTCAGTGTTAGTTCTGCTGGTCCAATATCCCAGTTTAATTATAAACAGTCCTTTTCTAATCCC
ACATGTTTGATTTTAGCGACTGTTCCTCATAACCTTACTACTATTACTAAGCCTCTTAAGTACAGCTATATTAACAAGTGCTCTCGTCTTCTTTCTGATGAT
CGTACTGAAGTACCTCAGTTAGTGAACGCTAATCAATACTCACCCTGTGTATCCATTGTCCCATCCACTGTGTGGGAAGACGGTGATTATTATAGGAAACAA
CTATCTCCACTTGAAGGTGGTGGCTGGCTTGTTGCTAGTGGCTCAACTGTTGCCATGACTGAGCAATTACAGATGGGCTTTGGTATTACAGTTCAATATGGT
ACAGACACCAATAGTGTTTGCCCCAAGCTTGAATTTGCTAATGACACAAAAATTGCCTCTCAATTAGGCAATTGCGTGGAATATTCCCTCTATGGTGTTTCG
GGCCGTGGTGTTTTTCAGAATTGCACAGCTGTAGGTGTTCGACAGCAGCGCTTTGTTTATGATGCGTACCAGAATTTAGTTGGCTATTATTCTGATGATGGC
AACTACTACTGTTTGCGTGCTTGTGTTAGTGTTCCTGTTTCTGTCATCTATGATAAAGAAACTAAAACCCACGCTACTCTATTTGGTAGTGTTGCATGTGAA
CACATTTCTTCTACCATGTCTCAATACTCCCGTTCTACGCGATCAATGCTTAAACGGCGAGATTCTACATATGGCCCCCTTCAGACACCTGTTGGTTGTGTC
CTAGGACTTGTTAATTCCTCTTTGTTCGTAGAGGACTGCAAGTTGCCTCTTGGTCAATCTCTCTGTGCTCTTCCTGACACACCTAGTACTCTCACACCTCGC
AGTGTGCGCTCTGTTCCAGGTGAAATGCGCTTGGCATCCATTGCTTTTAATCATCCTATTCAGGTTGATCAACTTAATAGTAGTTATTTTAAATTAAGTATA
CCCACTAATTTTTCCTTTGGTGTGACTCAGGAGTACATTCAGACAACCATTCAGAAAGTTACTGTTGATTGTAAACAGTACGTTTGCAATGGTTTCCAGAAG
TGTGAGCAATTACTGCGCGAGTATGGCCAGTTTTGTTCCAAAATAAACCAGGCTCTCCATGGTGCCAATTTACGCCAGGATGATTCTGTACGTAATTTGTTT
GCGAGCGTGAAAAGCTCTCAATCATCTCCTATCATACCAGGTTTTGGAGGTGACTTTAATTTGACACTTCTAGAACCTGTTTCTATATCTACTGGCAGTCGT

AGTGCACGTAGTGCTATTGAGGATTTGCTATTTGACAAAGTCACTATAGCTGATCCTGGTTATATGCAAGGTTACGATGATTGCATGCAGCAAGGTCCAGCA
TCAGCTCGTGATCTTATTTGTGCTCAATATGTGGCTGGTTACAAAGTATTACCTCCTCTTATGGATGTTAATATGGAAGCCGCGTATACTTCATCTTTGCTT
GGCAGCATAGCAGGTGTTGGCTGGACTGCTGGCTTATCCTCCTTTGCTGCTATTCCATTTGCACAGAGTATCTTTTATAGGTTAAACGGTGTTGGCATTACT
CAACAGGTTCTTTCAGAGAACCAAAAGCTTATTGCCAATAAGTTTAATCAGGCTCTGGGAGCTATGCAAACAGGCTTCACTACAACTAATGAAGCTTTTCAG
AAGGTTCAGGATGCTGTGAACAACAATGCACAGGCTCTATCCAAATTAGCTAGCGAGCTATCTAATACTTTTGGTGCTATTTCCGCCTCTATTGGAGACATC
ATACAACGTCTTGATGTTCTCGAACAGGACGCCCAAATAGACAGACTTATTAATGGCCGTTTGACAACACTAAATGCTTTTGTTGCACAGCAGCTTGTTCGT
TCCGAATCAGCTGCTCTTTCCGCTCAATTGGCTAAAGATAAAGTCAATGAGTGTGTCAAGGCACAATCCAAGCGTTCTGGATTTTGCGGTCAAGGCACACAT
ATAGTGTCCTTTGTTGTAAATGCCCCTAATGGCCTTTACTTCATGCATGTTGGTTATTACCCTAGCAACCACATTGAGGTTGTTTCTGCTTATGGTCTTTGC
GATGCAGCTAACCCTACTAATTGTATAGCCCCTGTTAATGGCTACTTTATTAAAACTAATAACACTAGGATTGTTGATGAGTGGTCATATACTGGCTCGTCC
TTCTATGCACCTGAGCCCATTACCTCCCTTAATACTAAGTATGTTGCACCACAGGTGACATACCAAAACATTTCTACTAACCTCCCTCCTCCTCTTCTCGGC
AATTCCACCGGGATTGACTTCCAAGATGAGTTGGATGAGTTTTTCAAAAATGTTAGCACCAGTATACCTAATTTTGGTTCCCTAACACAGATTAATACTACA
TTACTCGATCTTACCTACGAGATGTTGTCTCTTCAACAAGTTGTTAAAGCCCTTAATGAGTCTTACATAGACCTTAAAGAGCTTGGCAATTATACTTATTAC
AACAAATGGCCGTGGTACATTTGGCTTGGTTTCATTGCTGGGCTTGTTGCCTTAGCTCTATGCGTCTTCTTCATACTGTGCTGCACTGGTTGTGGCACAAAC
TGTATGGGAAAACTTAAGTGTAATCGTTGTTGTGATAGATACGAGGAATACGACCTCGAGCCGCATAAGGTTCATGTTCACTAA

SEQ ID NO: 12는 MERS-CoV 스파이크 당단백질, 염기 서열, 인간 세포에서 발현을 위해 최적화된 코돈이다.

ATGATCCACAGCGTGTTCCTGCTGATGTTTCTGCTGACACCTACCGAGTCCTACGTGGATGTGGGCCCAGACTCTGTGAAGAGCGCCTGCATCGAGGTGGAC
ATCCAGCAGACATTCTTTGACAAGACCTGGCCCAGACCCATCGACGTGAGCAAGGCAGACGGAATCATCTACCCACAGGGACGCACATATAGCAACATCACA
ATCACCTACCAGGGCCTGTTCCCTTATCAGGGCGACCACGGCGATATGTACGTGTATAGCGCCGGCCACGCAACCGGCACCACACCACAGAAGCTGTTTGTG
GCCAATTATTCCCAGGACGTGAAGCAGTTCGCCAACGGATTTGTGGTGCGGATCGGGGCCGCCGCCAACAGCACAGGCACCGTGATCATCTCTCCCAGCACA
TCCGCCACCATCAGAAAGATCTACCCTGCCTTTATGCTGGGCAGCTCCGTGGGCAACTTCTCCGATGGCAAGATGGGCAGGTTCTTTAATCACACACTGGTG
CTGCTGCCAGACGGATGCGGCACCCTGCTGAGGGCCTTCTACTGTATCCTGGAGCCCCGCTCTGGAAATCACTGCCCTGCCGGCAACTCCTACACCTCTTTT
GCCACATATCACACCCCTGCCACAGACTGTTCCGATGGCAATTATAACCGGAATGCCAGCCTGAACTCCTTCAAGGAGTACTTTAATCTGAGAAACTGCACC
TTCATGTACACATATAATATCACCGAGGATGAGATCCTGGAGTGGTTCGGCATCACACAGACCGCCCAGGGCGTGCACCTGTTTTCTAGCAGATACGTGGAT
CTGTATGGCGGCAACATGTTCCAGTTTGCCACACTGCCAGTGTATGACACCATCAAGTACTATAGCATCATCCCCCACTCTATCCGGAGCATCCAGTCCGAC
AGAAAGGCCTGGGCCGCCTTCTACGTGTATAAGCTGCAGCCCCTGACCTTCCTGCTGGATTTTTCCGTGGACGGCTACATCCGGAGAGCCATCGATTGCGGC
TTTAACGACCTGTCTCAGCTCCACTGTTCTTATGAGAGCTTCGATGTGGAGTCTGGCGTGTACAGCGTGTCCTCTTTTGAGGCCAAGCCATCTGGCAGCGTG
GTGGAGCAGGCAGAGGGAGTGGAGTGCGACTTCTCCCCACTGCTGTCTGGCACACCACCTCAGGTGTATAATTTCAAGAGGCTGGTGTTTACAAACTGTAAT
TACAACCTGACCAAGCTGCTGTCCCTGTTCTCTGTGAACGACTTTACCTGCAGCCAGATCTCCCCTGCCGCCATCGCCTCCAATTGTTATAGCTCCCTGATC
CTGGATTACTTCTCTTATCCCCTGTCTATGAAGAGCGACCTGTCCGTGTCTAGCGCCGGCCCTATCAGCCAGTTTAATTACAAGCAGTCCTTCTCTAACCCC
ACATGCCTGATCCTGGCCACCGTGCCTCACAACCTGACCACAATCACAAAGCCACTGAAGTACTCCTATATCAATAAGTGCAGCAGGCTGCTGTCCGACGAT
CGCACCGAGGTGCCTCAGCTCGTGAACGCCAACCAGTACTCTCCATGCGTGAGCATCGTGCCATCCACCGTGTGGGAGGACGGCGATTACTATAGAAAGCAG
CTCAGCCCACTGGAGGGAGGAGGATGGCTGGTGGCCAGCGGCTCCACAGTGGCCATGACCGAGCAGCTCCAGATGGGCTTCGGCATCACAGTGCAGTACGGC
ACAGATACCAATAGCGTGTGCCCCAAGCTGGAGTTTGCCAACGACACCAAGATCGCCTCCCAGCTCGGCAATTGCGTGGAGTACTCCCTGTATGGCGTGTCT
GGCAGAGGCGTGTTCCAGAACTGTACAGCCGTGGGCGTGCGGCAGCAGCGGTTCGTGTACGATGCCTATCAGAACCTGGTGGGCTACTATAGCGACGATGGC
AATTACTATTGCCTGAGGGCATGCGTGAGCGTGCCCGTGAGCGTGATCTACGACAAGGAGACAAAGACCCACGCCACCCTGTTCGGCTCCGTGGCCTGCGAG
CACATCTCCTCTACAATGTCTCAGTATTCTAGGAGCACCCGCTCTATGCTGAAGAGGCGCGACAGCACATACGGACCACTGCAGACCCCTGTGGGATGCGTG
CTGGGCCTGGTGAACAGCAGCCTGTTTGTGGAGGATTGCAAGCTGCCACTGGGCCAGTCTCTGTGCGCACTGCCAGACACCCCCAGCACACTGACCCCACGG
TCTGTGAGAAGCGTGCCCGGAGAGATGAGACTGGCCAGCATCGCCTTCAATCACCCTATCCAGGTGGATCAGCTCAACAGCAGCTACTTTAAGCTGAGCATC
CCAACAAACTTCTCCTTTGGCGTGACCCAGGAGTATATCCAGACCACAATCCAGAAGGTGACCGTGGACTGCAAGCAGTACGTGTGCAATGGCTTCCAGAAG
TGCGAGCAGCTCCTGAGGGAGTATGGCCAGTTTTGTTCCAAGATCAATCAGGCCCTGCACGGAGCCAACCTGAGGCAGGACGATTCCGTGAGAAACCTGTTC
GCCTCTGTGAAGTCCTCTCAGAGCTCCCCTATCATCCCAGGCTTCGGCGGCGACTTCAACCTGACCCTGCTGGAGCCCGTGTCCATCTCTACCGGCAGCAGG
TCCGCCCGCAGCGCCATCGAGGATCTGCTGTTTGACAAGGTGACCATCGCCGACCCAGGCTACATGCAGGGCTATGACGATTGCATGCAGCAGGGACCAGCC
TCCGCCCGCGATCTGATCTGTGCCCAGTACGTGGCCGGCTATAAGGTGCTGCCACCCCTGATGGACGTGAACATGGAGGCCGCCTATACATCTAGCCTGCTG
GGCAGCATCGCAGGAGTGGGATGGACCGCCGGCCTGTCCTCTTTCGCCGCAATCCCTTTTGCCCAGTCTATCTTCTACCGGCTGAACGGCGTGGGCATCACA
CAGCAGGTGCTGAGCGAGAATCAGAAGCTGATCGCCAATAAGTTCAACCAGGCCCTGGGGGCCATGCAGACCGGCTTTACCACAACCAACGAGGCCTTCCAG
AAGGTGCAGGATGCCGTGAACAATAACGCACAGGCCCTGTCCAAGCTGGCCTCCGAGCTGTCTAATACCTTCGGGGCCATCAGCGCCAGCATCGGCGACATC
ATCCAGCGCCTGGACGTGCTGGAGCAGGATGCCCAGATCGACAGGCTGATCAATGGCCGCCTGACAACCCTGAACGCCTTTGTGGCACAGCAGCTCGTGCGG
AGCGAGTCTGCCGCCCTGAGCGCCCAGCTCGCCAAGGACAAGGTGAACGAGTGCGTGAAGGCCCAGAGCAAGCGGTCCGGCTTTTGTGGCCAGGGCACCCAC
ATCGTGTCCTTCGTGGTGAATGCCCCTAACGGCCTGTACTTTATGCATGTGGGCTACTATCCAAGCAACCACATCGAGGTGGTGTCCGCCTATGGCCTGTGC
GATGCCGCCAATCCTACAAACTGTATCGCCCCAGTGAATGGCTACTTCATCAAGACCAATAACACACGGATCGTGGACGAGTGGTCCTACACCGGCAGCTCC

TTTTATGCCCCCGAGCCTATCACATCTCTGAACACCAAGTACGTGGCCCCACAGGTGACATATCAGAATATCAGCACCAACCTGCCTCCACCCCTGCTGGGC
AATTCCACCGGCATCGACTTCCAGGATGAGCTGGACGAGTTCTTTAAGAATGTGAGCACATCCATCCCCAACTTTGGCAGCCTGACCCAGATCAACACAACC
CTGCTGGATCTGACATACGAGATGCTGTCTCTGCAGCAGGTGGTGAAGGCCCTGAATGAGAGCTACATCGACCTGAAGGAGCTGGGCAATTATACCTACTAT
AACAAGTGGCCTTGGTACATCTGGCTGGGCTTCATCGCAGGCCTGGTGGCCCTGGCCCTGTGCGTGTTCTTTATCCTGTGCTGTACAGGCTGCGGCACCAAT
TGTATGGGCAAGCTGAAGTGTAACCGGTGCTGTGATAGATACGAGGAGTATGACCTGGAGCCACACAAGGTGCATGTGCACTGA

SEQ ID NO: 13은 SARS-CoV-2 "프롤린 변형된" 스파이크 당단백질, 아미노산 서열이다.

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIY
SKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV
VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAI
PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
NQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA
SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
SGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 14는 SARS-CoV-2 "프롤린 변형된" 스파이크 당단백질, 염기 서열이다.

ATGTTTGTTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAATTCTTTCACACGT
GGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATA
CATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGA
GGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGT
AATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATAT
GTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTTAAGAATATTGATGGTTATTTTAAAATATAT
TCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTT
CAAACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCT
AGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAAATCCTTC
ACTGTAGAAAAAGGAATCTATCAAACTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAA
GTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATTT
TCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGA
CAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTGAT
TCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGT
AGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACCATACAGAGTA
GTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAAT
GGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGAT
CCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTAT
CAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGT
GCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCT
CCTCGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATA
CCCACAAATTTTACTATTAGTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAA
TGCAGCAATCTTTTGTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGAACAAGACAAAAACACCCAAGAAGTTTTT
GCACAAGTCAAACAAATTTACAAAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATT
TGTGCACAAAAGTTTAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCT
GGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAG
AACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACTTCAAGATGTGGTC
AACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATATCCTTTCACGTCTTGACCCT

CCTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCT
TCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAG
TCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGATGGAAAAGCACAC
TTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTG
TCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTT
AAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTGCC
AAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTG
ATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAA
GACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAA

SEQ ID NO: 15는 SARS-CoV-2 "프롤린 변형된" 스파이크 당단백질, 염기 서열, 인간 세포에서 발현을 위해 최적화된 코돈이다.

ATGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGAGCTCCCAGTGCGTGAACCTGACCACAAGGACCCAGCTCCCCCCTGCCTATACCAATTCCTTCACACGG
GGCGTGTACTATCCAGACAAGGTGTTTAGATCTAGCGTGCTGCACTCCACACAGGATCTGTTTCTGCCCTTCTTTTCTAACGTGACCTGGTTCCACGCCATC
CATGTGAGCGGCACCAATGGCACAAAGCGGTTCGACAATCCTGTGCTGCCCTTCAACGATGGCGTGTACTTCGCCTCCACCGAGAAGTCTAACATCATCAGA
GGCTGGATCTTTGGCACCACACTGGACAGCAAGACACAGTCCCTGCTGATCGTGAACAATGCCACCAACGTGGTCATCAAGGTGTGCGAGTTCCAGTTTTGT
AATGATCCTTTCCTGGGCGTGTACTATCACAAGAACAATAAGTCTTGGATGGAGAGCGAGTTTCGCGTGTATTCCTCTGCCAACAATTGCACATTTGAGTAC
GTGTCCCAGCCATTCCTGATGGACCTGGAGGGCAAGCAGGGCAATTTCAAGAACCTGAGGGAGTTCGTGTTTAAGAATATCGATGGCTACTTCAAGATCTAC
TCCAAGCACACCCCTATCAACCTGGTGCGCGACCTGCCACAGGGCTTCTCTGCCCTGGAGCCTCTGGTGGATCTGCCAATCGGCATCAACATCACCCGGTTT
CAGACACTGCTGGCCCTGCACAGAAGCTACCTGACACCTGGCGACAGCTCCTCTGGATGGACCGCCGGGGCCGCCGCCTACTATGTGGGCTATCTGCAGCCA
AGGACCTTCCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTGCGCCCTGGACCCCCTGTCTGAGACCAAGTGTACACTGAAGAGCTTT
ACCGTGGAGAAGGGCATCTATCAGACAAGCAATTTCAGGGTGCAGCCCACCGAGTCCATCGTGCGCTTTCCAAATATCACAAACCTGTGCCCCTTTGGCGAG
GTGTTCAACGCAACCAGGTTCGCCAGCGTGTACGCATGGAATAGGAAGCGCATCTCCAACTGCGTGGCCGACTATTCTGTGCTGTACAACAGCGCCTCCTTC
TCTACCTTTAAGTGCTATGGCGTGAGCCCCACAAAGCTGAATGACCTGTGCTTTACCAACGTGTACGCCGATTCCTTCGTGATCAGGGGCGACGAGGTGCGC
CAGATCGCACCAGGACAGACAGGCAAGATCGCCGACTACAATTATAAGCTGCCCGACGATTTCACCGGCTGCGTGATCGCCTGGAACTCTAACAATCTGGAT
AGCAAAGTGGGCGGCAACTACAATTATCTGTACCGGCTGTTTAGAAAGTCTAATCTGAAGCCTTTCGAGAGGGACATCTCCACAGAGATCTACCAGGCCGGC
TCTACCCCATGCAATGGCGTGGAGGGCTTTAACTGTTATTTCCCCCTGCAGAGCTACGGCTTCCAGCCTACAAACGGCGTGGGCTATCAGCCATACCGCGTG
GTGGTGCTGTCTTTTGAGCTGCTGCACGCACCAGCAACAGTGTGCGGACCTAAGAAGAGCACCAATCTGGTGAAGAACAAGTGCGTGAACTTCAACTTCAAC
GGCCTGACCGGCACAGGCGTGCTGACCGAGTCCAACAAGAAGTTCCTGCCCTTTCAGCAGTTCGGCAGGGACATCGCAGATACCACAGACGCCGTGCGCGAC
CCCCAGACCCTGGAGATCCTGGACATCACACCTTGCTCCTTCGGCGGCGTGTCTGTGATCACACCTGGCACCAATACAAGCAACCAGGTGGCCGTGCTGTAT
CAGGACGTGAATTGTACCGAGGTGCCAGTGGCCATCCACGCCGATCAGCTCACCCCCACATGGCGGGTGTACTCTACCGGCAGCAACGTGTTCCAGACAAGA
GCCGGCTGCCTGATCGGAGCCGAGCATGTGAACAATAGCTATGAGTGCGACATCCCCATCGGAGCCGGCATCTGTGCCTCCTACCAGACCCAGACAAACTCC
CCTCGGAGAGCCCGGTCTGTGGCCAGCCAGTCCATCATCGCCTATACCATGAGCCTGGGGGCCGAGAACAGCGTGGCCTACTCCAACAATTCTATCGCCATC
CCCACCAACTTCACAATCTCCGTGACCACAGAGATCCTGCCTGTGAGCATGACCAAGACATCCGTGGACTGCACAATGTATATCTGTGGCGATTCCACCGAG
TGCTCTAACCTGCTGCTGCAGTACGGCTCTTTTTGTACCCAGCTCAACAGAGCCCTGACAGGCATCGCCGTGGAGCAGGACAAGAACACACAGGAGGTGTTC
GCCCAGGTGAAGCAGATCTACAAGACCCCACCCATCAAGGACTTTGGCGGCTTCAACTTCAGCCAGATCCTGCCAGATCCCAGCAAGCCTTCCAAGCGGTCT
TTTATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGCCGATGCCGGCTTCATCAAGCAGTATGGCGATTGCCTGGGCGACATCGCCGCCAGAGACCTGATC
TGTGCCCAGAAGTTTAATGGCCTGACCGTGCTGCCTCCACTGCTGACAGATGAGATGATCGCCCAGTACACATCTGCCCTGCTGGCCGGCACCATCACAAGC
GGATGGACCTTCGGGGCCGGGGCCGCCCTGCAGATCCCATTTGCCATGCAGATGGCCTATCGGTTCAACGGCATCGGCGTGACCCAGAATGTGCTGTACGAG
AACCAGAAGCTGATCGCCAATCAGTTTAACTCCGCCATCGGCAAGATCCAGGACTCTCTGAGCTCCACAGCCAGCGCCCTGGGCAAGCTGCAGGATGTGGTG
AATCAGAACGCCCAGGCCCTGAATACCCTGGTGAAGCAGCTCAGCAGCAACTTCGGGGCCATCAGCAGCGTGCTGAACGACATCCTGAGCCGGCTGGACCCC
CCTGAGGCAGAGGTGCAGATCGACCGGCTGATCACAGGCAGACTGCAGTCCCTGCAGACCTACGTGACACAGCAGCTCATCAGGGCCGCCGAGATCAGGGCC
TCTGCCAATCTGGCCGCCACCAAGATGAGCGAGTGCGTGCTGGGCCAGTCCAAGAGAGTGGACTTTTGTGGCAAGGGCTATCACCTGATGAGCTTCCCACAG
TCCGCCCCCCACGGAGTGGTGTTTCTGCATGTGACCTACGTGCCTGCCCAGGAGAAGAACTTCACCACAGCCCCAGCCATCTGCCACGATGGCAAGGCCCAC
TTTCCCAGGGAGGGCGTGTTCGTGAGCAACGGCACCCACTGGTTTGTGACACAGCGCAATTTCTACGAGCCTCAGATCATCACCACAGACAATACATTCGTG
TCCGGCAACTGTGACGTGGTCATCGGCATCGTGAACAATACCGTGTATGATCCCCTGCAGCCTGAGCTGGACTCTTTTAAGGAGGAGCTGGATAAGTACTTC
AAGAATCACACCAGCCCCGACGTGGATCTGGGCGACATCTCTGGCATCAATGCCAGCGTGGTGAACATCCAGAAGGAGATCGACAGGCTGAACGAGGTGGCC
AAGAATCTGAACGAGTCCCTGATCGATCTGCAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTGGCCATGGTATATCTGGCTGGGCTTCATCGCCGGCCTG
ATCGCCATCGTGATGGTGACCATCATGCTGTGCTGTATGACAAGCTGCTGTTCCTGCCTGAAGGGCTGCTGTTCTTGTGGCTCCTGCTGTAAGTTTGATGAG
GACGATAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGA

SEQ ID NO: 16은 SARS-CoV-2 "퓨린 절단 변형된" 스파이크 당단백질, 아미노산 서열이다.

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIY
SKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV
VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAI
PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
NQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA
SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
SGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 17은 SARS-CoV-2 "퓨린 절단 변형된" 스파이크 당단백질, 염기 서열이다.

ATGTTTGTTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAATTCTTTCACACGT
GGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATA
CATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGA
GGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGT
AATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATAT
GTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTTAAGAATATTGATGGTTATTTTAAAATATAT
TCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTT
CAAACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCT
AGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAAATCCTTC
ACTGTAGAAAAAGGAATCTATCAAACTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAA
GTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATTT
TCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGA
CAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTGAT
TCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGT
AGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACCATACAGAGTA
GTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAAT
GGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGAT
CCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTAT
CAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGT
GCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCT
CCTGGTAGTGCAAGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATA
CCCACAAATTTTACTATTAGTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAA
TGCAGCAATCTTTTGTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGAACAAGACAAAAACACCCAAGAAGTTTTT
GCACAAGTCAAACAAATTTACAAAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATT
TGTGCACAAAAGTTTAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCT
GGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAG
AACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACTTCAAGATGTGGTC
AACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATATCCTTTCACGTCTTGACAAA
GTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCT
TCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAG
TCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGATGGAAAAGCACAC
TTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTG
TCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTT

AAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTGCC
AAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTG
ATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAA
GACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAA

SEQ ID NO: 18은 SARS-CoV-2 "퓨린 절단 변형된" 스파이크 당단백질, 염기 서열, 인간 세포에서 발현을 위해 최적화된 코돈이다.

ATGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGAGCTCCCAGTGCGTGAACCTGACCACAAGGACCCAGCTCCCCCCTGCCTATACCAATTCCTTCACAAGG
GGCGTGTACTATCCCGACAAGGTGTTTCGCTCTAGCGTGCTGCACAGCACACAGGATCTGTTTCTGCCTTTCTTTTCCAACGTGACCTGGTTCCACGCCATC
CATGTGAGCGGCACCAATGGCACAAAGAGGTTCGACAATCCAGTGCTGCCCTTTAACGATGGCGTGTACTTCGCCTCTACCGAGAAGAGCAACATCATCCGC
GGCTGGATCTTTGGCACCACACTGGACTCCAAGACACAGTCTCTGCTGATCGTGAACAATGCCACCAACGTGGTCATCAAGGTGTGCGAGTTCCAGTTTTGT
AATGATCCATTCCTGGGCGTGTACTATCACAAGAACAATAAGAGCTGGATGGAGTCCGAGTTTCGCGTGTATTCCTCTGCCAACAATTGCACATTTGAGTAC
GTGTCCCAGCCCTTCCTGATGGACCTGGAGGGCAAGCAGGGCAATTTCAAGAACCTGCGGGAGTTCGTGTTTAAGAATATCGATGGCTACTTCAAGATCTAC
AGCAAGCACACCCCAATCAACCTGGTGAGAGACCTGCCACAGGGCTTCTCCGCCCTGGAGCCACTGGTGGATCTGCCCATCGGCATCAACATCACCAGGTTT
CAGACACTGCTGGCCCTGCACCGCAGCTACCTGACACCAGGCGACAGCTCCTCTGGATGGACCGCCGGGGCCGCCGCCTACTATGTGGGCTATCTGCAGCCC
CGGACCTTCCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTGCGCCCTGGACCCCCTGTCCGAGACCAAGTGTACACTGAAGTCTTTT
ACCGTGGAGAAGGGCATCTATCAGACATCTAATTTCCGGGTGCAGCCTACCGAGAGCATCGTGAGATTTCCCAATATCACAAACCTGTGCCCTTTTGGCGAG
GTGTTCAACGCCACCAGATTCGCCAGCGTGTACGCCTGGAATCGGAAGAGAATCAGCAACTGCGTGGCCGACTATTCCGTGCTGTACAACTCTGCCAGCTTC
TCCACCTTTAAGTGCTATGGCGTGTCTCCCACAAAGCTGAATGACCTGTGCTTTACCAACGTGTACGCCGATAGCTTCGTGATCAGGGGCGACGAGGTGAGA
CAGATCGCACCAGGACAGACAGGCAAGATCGCAGACTACAATTATAAGCTGCCTGACGATTTCACCGGCTGCGTGATCGCCTGGAACAGCAACAATCTGGAT
TCCAAAGTGGGCGGCAACTACAATTATCTGTACAGGCTGTTTCGCAAGTCCAATCTGAAGCCATTCGAGCGGGACATCAGCACAGAGATCTACCAGGCAGGC
TCCACCCCATGCAATGGAGTGGAGGGCTTTAACTGTTATTTCCCTCTGCAGTCTTACGGCTTCCAGCCAACAAACGGCGTGGGCTATCAGCCCTACAGAGTG
GTGGTGCTGTCCTTTGAGCTGCTGCACGCACCTGCAACAGTGTGCGGACCAAAGAAGTCTACCAATCTGGTGAAGAACAAGTGCGTGAACTTCAACTTCAAC
GGCCTGACCGGCACAGGCGTGCTGACCGAGTCCAACAAGAAGTTCCTGCCTTTTCAGCAGTTCGGCAGAGACATCGCCGATACCACAGACGCCGTGAGAGAC
CCTCAGACCCTGGAGATCCTGGACATCACACCATGCTCTTTCGGCGGCGTGAGCGTGATCACACCAGGCACCAATACAAGCAACCAGGTGGCCGTGCTGTAT
CAGGACGTGAATTGTACCGAGGTGCCCGTGGCCATCCACGCAGATCAGCTCACCCCTACATGGAGGGTGTACTCCACCGGCTCTAACGTGTTCCAGACACGC
GCCGGATGCCTGATCGGAGCCGAGCATGTGAACAATTCTTATGAGTGCGACATCCCTATCGGAGCCGGCATCTGTGCCAGCTACCAGACCCAGACAAACAGC
CCAGGCTCCGCCAGCTCCGTGGCCTCTCAGAGCATCATCGCCTATACCATGAGCCTGGGGGCCGAGAATAGCGTGGCCTACTCTAACAATAGCATCGCCATC
CCTACCAACTTCACAATCTCCGTGACCACAGAGATCCTGCCAGTGTCCATGACCAAGACATCTGTGGACTGCACAATGTATATCTGTGGCGATTCTACCGAG
TGCAGCAACCTGCTGCTGCAGTACGGCAGCTTTTGTACCCAGCTCAACCGGGCCCTGACAGGAATCGCAGTGGAGCAGGACAAGAACACACAGGAGGTGTTC
GCCCAGGTGAAGCAGATCTACAAGACCCCACCCATCAAGGACTTTGGCGGCTTCAACTTCAGCCAGATCCTGCCCGATCCTTCCAAGCCATCTAAGAGGAGC
TTTATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGCCGATGCCGGCTTCATCAAGCAGTATGGCGATTGCCTGGGCGACATCGCAGCCCGCGACCTGATC
TGTGCCCAGAAGTTTAATGGCCTGACCGTGCTGCCTCCACTGCTGACAGATGAGATGATCGCACAGTACACATCCGCCCTGCTGGCCGGCACCATCACATCT
GGATGGACCTTCGGGGCCGGGGCCGCCCTGCAGATCCCCTTTGCCATGCAGATGGCCTATAGATTCAACGGCATCGGCGTGACCCAGAATGTGCTGTACGAG
AACCAGAAGCTGATCGCCAATCAGTTTAACTCCGCCATCGGCAAGATCCAGGACTCCCTGTCTAGCACAGCCTCTGCCCTGGGCAAGCTGCAGGATGTGGTG
AATCAGAACGCCCAGGCCCTGAATACCCTGGTGAAGCAGCTCAGCAGCAACTTCGGGGCCATCAGCAGCGTGCTGAACGACATCCTGAGCCGGCTGGACAAG
GTGGAGGCAGAGGTGCAGATCGACAGGCTGATCACAGGCCGCCTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTCATCAGGGCCGCCGAGATCAGAGCC
TCCGCCAATCTGGCCGCCACCAAGATGTCTGAGTGCGTGCTGGGCCAGAGCAAGCGCGTGGACTTTTGTGGCAAGGGCTATCACCTGATGTCCTTCCCACAG
TCTGCCCCTCACGGAGTGGTGTTTCTGCATGTGACCTACGTGCCAGCCCAGGAGAAGAACTTCACCACAGCCCCCGCAATCTGCCACGATGGCAAGGCACAC
TTTCCTCGGGAGGGCGTGTTCGTGTCTAACGGCACCCACTGGTTTGTGACACAGAGAAATTTCTACGAGCCACAGATCATCACCACAGACAATACATTCGTG
AGCGGCAACTGTGACGTGGTCATCGGCATCGTGAACAATACCGTGTATGATCCTCTGCAGCCAGAGCTGGACTCCTTTAAGGAGGAGCTGGATAAGTACTTC
AAGAATCACACCTCTCCCGACGTGGATCTGGGCGACATCAGCGGCATCAATGCCTCCGTGGTGAACATCCAGAAGGAGATCGACAGGCTGAACGAGGTGGCC
AAGAATCTGAACGAGTCCCTGATCGATCTGCAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTGGCCCTGGTATATCTGGCTGGGCTTCATCGCCGGCCTG
ATCGCCATCGTGATGGTGACCATCATGCTGTGCTGTATGACATCCTGCTGTTCTTGCCTGAAGGGCTGCTGTAGCTGTGGCTCCTGCTGTAAGTTTGATGAG
GACGATAGCGAGCCTGTGCTGAAGGGCGTGAAGCTGCACTACACCTGA

SEQ ID NO: 19는 SARS-CoV-2 "프롤린 및 퓨린 절단 변형된" 스파이크 당단백질, 아미노산 서열이다.

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIY
SKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR

QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV
VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAI
PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
NQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA
SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
SGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 20은 SARS-CoV-2 "프롤린 및 퓨린 절단 변형된" 스파이크 당단백질, 염기 서열이다.

ATGTTTGTTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAATTCTTTCACACGT
GGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATA
CATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGA
GGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGT
AATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATAT
GTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTTAAGAATATTGATGGTTATTTTAAAATATAT
TCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTT
CAAACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCT
AGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAAATCCTTC
ACTGTAGAAAAAGGAATCTATCAAACTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAA
GTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATTT
TCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGA
CAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTGAT
TCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGT
AGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACCATACAGAGTA
GTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAAT
GGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGAT
CCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTAT
CAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGT
GCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCT
CCTGGTAGTGCAAGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATA
CCCACAAATTTTACTATTAGTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAA
TGCAGCAATCTTTTGTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGAACAAGACAAAAACACCCAAGAAGTTTTT
GCACAAGTCAAACAAATTTACAAAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATT
TGTGCACAAAAGTTTAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCT
GGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAG
AACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACTTCAAGATGTGGTC
AACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATATCCTTTCACGTCTTGACCCT
CCTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCT
TCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAG
TCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGATGGAAAAGCACAC
TTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTG
TCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTT
AAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTGCC
AAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTG
ATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAA
GACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAA

SEQ ID NO: 21은 SARS-Cov-2 "프롤린 및 퓨린 절단 변형" 스파이크 당단백질, 염기 서열, 인간 세포에서의 발현

에 최적화된 코돈이다.

ATGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGAGCTCCCAGTGCGTGAACCTGACCACAAGGACCCAGCTCCCCCCTGCCTATACCAATTCCTTCACAAGG
GGCGTGTACTATCCAGACAAGGTGTTTCGCTCTAGCGTGCTGCACAGCACACAGGATCTGTTTCTGCCCTTCTTTTCCAACGTGACCTGGTTCCACGCCATC
CATGTGAGCGGCACCAATGGCACAAAGAGGTTCGACAATCCTGTGCTGCCCTTCAACGATGGCGTGTACTTCGCCTCTACCGAGAAGAGCAACATCATCCGC
GGCTGGATCTTTGGCACCACACTGGACTCCAAGACACAGTCTCTGCTGATCGTGAACAATGCCACCAACGTGGTCATCAAGGTGTGCGAGTTCCAGTTTTGT
AATGATCCTTTCCTGGGCGTGTACTATCACAAGAACAATAAGAGCTGGATGGAGTCCGAGTTTCGCGTGTATTCCTCTGCCAACAATTGCACATTTGAGTAC
GTGTCCCAGCCATTCCTGATGGACCTGGAGGGCAAGCAGGGCAATTTCAAGAACCTGCGGGAGTTCGTGTTTAAGAATATCGATGGCTACTTCAAGATCTAC
AGCAAGCACACCCCTATCAACCTGGTGAGAGACCTGCCACAGGGCTTCTCCGCCCTGGAGCCTCTGGTGGATCTGCCAATCGGCATCAACATCACCAGGTTT
CAGACACTGCTGGCCCTGCACCGCAGCTACCTGACACCTGGCGACAGCTCCTCTGGATGGACCGCCGGGGCCGCCGCCTACTATGTGGGCTATCTGCAGCCA
CGGACCTTCCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTGCGCCCTGGACCCCCTGTCCGAGACCAAGTGTACACTGAAGTCTTTT
ACCGTGGAGAAGGGCATCTATCAGACATCTAATTTCCGGGTGCAGCCCACCGAGAGCATCGTGAGATTTCCAAATATCACAAACCTGTGCCCCTTTGGCGAG
GTGTTCAACGCCACCAGATTCGCCAGCGTGTACGCCTGGAATCGGAAGAGAATCAGCAACTGCGTGGCCGACTATTCCGTGCTGTACAACTCTGCCAGCTTC
TCCACCTTTAAGTGCTATGGCGTGTCTCCCACAAAGCTGAATGACCTGTGCTTTACCAACGTGTACGCCGATAGCTTCGTGATCAGGGGCGACGAGGTGAGA
CAGATCGCACCAGGCCAGACAGGCAAGATCGCCGACTACAATTATAAGCTGCCCGACGATTTCACCGGCTGCGTGATCGCCTGGAACAGCAACAATCTGGAT
TCCAAAGTGGGCGGCAACTACAATTATCTGTACAGGCTGTTTCGCAAGTCCAATCTGAAGCCTTTCGAGCGGGACATCAGCACAGAGATCTACCAGGCCGGC
TCCACCCCATGCAATGGCGTGGAGGGCTTTAACTGTTATTTCCCCCTGCAGTCTTACGGCTTCCAGCCTACAAACGGCGTGGGCTATCAGCCATACAGAGTG
GTGGTGCTGTCCTTTGAGCTGCTGCACGCACCAGCAACAGTGTGCGGACCTAAGAAGTCTACCAATCTGGTGAAGAACAAGTGCGTGAACTTCAACTTCAAC
GGCCTGACCGGCACAGGCGTGCTGACCGAGTCCAACAAGAAGTTCCTGCCCTTTCAGCAGTTCGGCAGAGACATCGCCGATACCACAGACGCCGTGAGAGAC
CCCCAGACCCTGGAGATCCTGGACATCACACCTTGCTCTTTCGGCGGCGTGAGCGTGATCACACCTGGCACCAATACAAGCAACCAGGTGGCCGTGCTGTAT
CAGGACGTGAATTGTACCGAGGTGCCAGTGGCCATCCACGCCGATCAGCTCACCCCCACATGGAGGGTGTACTCCACCGGCTCTAACGTGTTCCAGACACGC
GCCGGATGCCTGATCGGAGCCGAGCATGTGAACAATTCTTATGAGTGCGACATCCCCATCGGAGCCGGCATCTGTGCCAGCTACCAGACCCAGACAAACAGC
CCTGGCTCCGCCAGCTCCGTGGCCTCTCAGAGCATCATCGCCTATACCATGAGCCTGGGGGCCGAGAATAGCGTGGCCTACTCTAACAATAGCATCGCCATC
CCCACCAACTTCACAATCTCCGTGACCACAGAGATCCTGCCCGTGAGCATGACCAAGACATCTGTGGACTGCACAATGTATATCTGTGGCGATTCTACCGAG
TGCAGCAACCTGCTGCTGCAGTACGGCAGCTTTTGTACCCAGCTCAACCGGGCCCTGACAGGAATCGCAGTGGAGCAGGACAAGAACACACAGGAGGTGTTC
GCCCAGGTGAAGCAGATCTACAAGACCCCACCCATCAAGGACTTTGGCGGCTTCAACTTCAGCCAGATCCTGCCAGATCCCTCCAAGCCTTCTAAGAGGAGC
TTTATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGCCGATGCCGGCTTCATCAAGCAGTATGGCGATTGCCTGGGCGACATCGCAGCCCGCGACCTGATC
TGTGCCCAGAAGTTTAATGGCCTGACCGTGCTGCCTCCACTGCTGACAGATGAGATGATCGCACAGTACACATCCGCCCTGCTGGCCGGCACCATCACATCT
GGATGGACCTTCGGGGCCGGGGCCGCCCTGCAGATCCCATTTGCCATGCAGATGGCCTATAGATTCAACGGCATCGGCGTGACCCAGAATGTGCTGTACGAG
AACCAGAAGCTGATCGCCAATCAGTTTAACTCCGCCATCGGCAAGATCCAGGACTCCCTGTCTAGCACAGCCTCTGCCCTGGGCAAGCTGCAGGATGTGGTG
AATCAGAACGCCCAGGCCCTGAATACCCTGGTGAAGCAGCTCAGCAGCAACTTCGGGGCCATCAGCAGCGTGCTGAACGACATCCTGAGCCGGCTGGACCCC
CCTGAGGCAGAGGTGCAGATCGACAGGCTGATCACAGGCCGCCTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTCATCAGGGCCGCCGAGATCAGAGCC
TCCGCCAATCTGGCCGCCACCAAGATGTCTGAGTGCGTGCTGGGCCAGAGCAAGCGCGTGGACTTTTGTGGCAAGGGCTATCACCTGATGTCCTTCCCACAG
TCTGCCCCCCACGGAGTGGTGTTTCTGCATGTGACCTACGTGCCTGCCCAGGAGAAGAACTTCACCACAGCCCCAGCCATCTGCCACGATGGCAAGGCACAC
TTTCCCCGGGAGGGCGTGTTCGTGTCTAACGGCACCCACTGGTTTGTGACACAGAGAAATTTCTACGAGCCTCAGATCATCACCACAGACAATACATTCGTG
AGCGGCAACTGTGACGTGGTCATCGGCATCGTGAACAATACCGTGTATGATCCCCTGCAGCCTGAGCTGGACTCCTTTAAGGAGGAGCTGGATAAGTACTTC
AAGAATCACACCTCTCCCGACGTGGATCTGGGCGACATCAGCGGCATCAATGCCTCCGTGGTGAACATCCAGAAGGAGATCGACAGGCTGAACGAGGTGGCC
AAGAATCTGAACGAGTCCCTGATCGATCTGCAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTGGCCATGGTATATCTGGCTGGGCTTCATCGCCGGCCTG
ATCGCCATCGTGATGGTGACCATCATGCTGTGCTGTATGACATCCTGCTGTTCTTGCCTGAAGGGCTGCTGTAGCTGTGGCTCCTGCTGTAAGTTTGATGAG
GACGATAGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGA

SEQ ID NO: 22는 SARS-CoV-2 "프롤린 및 퓨린 절단 변형 및 VSV-G TMCyt SWAP" 스파이크 당단백질, 아미노산 서열이다.

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIY
SKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV
VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAI
PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE

NQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA
SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
SGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKFFFIIGLIIGLFL
VLRVGIHLCIKLKHTKKRQIYTDIEMNRLGK

SEQ ID NO: 23은 SARS CoV-2 "프롤린 및 퓨린 절단 변형 및 VSV-G TMCyt SWAP" 스파이크 당단백질, 염기 서열이다.

ATGTTTGTTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCCCCTGCATACACTAATTCTTTCACACGT
GGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATA
CATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCTTCCACTGAGAAGTCTAACATAATAAGA
GGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGT
AATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTTTGAATAT
GTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTTAAGAATATTGATGGTTATTTTAAAATATAT
TCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTT
CAAACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCT
AGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAAATCCTTC
ACTGTAGAAAAAGGAATCTATCAAACTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAA
GTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATTT
TCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGA
CAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTGAT
TCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGT
AGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACCATACAGAGTA
GTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAAT
GGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGAT
CCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTAT
CAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGT
GCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCT
CCTGGTAGTGCAAGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATA
CCCACAAATTTTACTATTAGTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAA
TGCAGCAATCTTTTGTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGAACAAGACAAAAACACCCAAGAAGTTTTT
GCACAAGTCAAACAAATTTACAAAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCA
TTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATT
TGTGCACAAAAGTTTAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCT
GGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAG
AACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACTTCAAGATGTGGTC
AACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTAAATGATATCCTTTCACGTCTTGACCCT
CCTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCT
TCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAG
TCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGATGGAAAAGCACAC
TTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTG
TCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTT
AAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTTGCC
AAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATTTTTCTTTATCATAGGGTTAATCATTGGACTATTCTTG
GTTCTCCGAGTTGGTATCCATCTTTGCATTAAATTAAAGCACACCAAGAAAAGACAGATTTATACAGACATAGAGATGAACCGACTTGGAAAGTAA

SEQ ID NO: 24는 SARS CoV-2 "프롤린 및 퓨린 절단 변형 및 VSV-G TMCyt SWAP" 스파이크 당단백질, 염기 서열, 인간 세포에서의 발현에 최적화된 코돈이다.

ATGTTCGTGTTCCTGGTGCTGCTGCCTCTGGTGAGCTCCCAGTGCGTGAACCTGACCACAAGGACCCAGCTCCCCCCTGCCTATACCAATTCCTTTACAAGG
GGCGTGTACTATCCAGACAAGGTGTTCCGCTCTAGCGTGCTGCACTCTACACAGGATCTGTTCCTGCCCTTCTTTAGCAACGTGACCTGGTTTCACGCCATC
CATGTGAGCGGCACCAATGGCACAAAGCGGTTTGACAATCCTGTGCTGCCATTCAACGATGGCGTGTACTTTGCCTCCACCGAGAAGTCTAACATCATCAGA
GGCTGGATCTTCGGCACCACACTGGACAGCAAGACACAGTCCCTGCTGATCGTGAACAATGCCACCAACGTGGTCATCAAGGTGTGCGAGTTTCAGTTCTGT

AATGATCCTTTTCTGGGCGTGTACTATCACAAGAACAATAAGTCTTGGATGGAGAGCGAGTTCCGCGTGTATTCCTCTGCCAACAATTGTACATTCGAGTAC
GTGTCCCAGCCATTTCTGATGGACCTGGAGGGCAAGCAGGGCAACTTCAAGAACCTGCGGGAGTTCGTGTTCAAGAATATCGATGGCTATTTCAAGATCTAC
TCTAAGCACACCCCTATCAACCTGGTGCGCGACCTGCCACAGGGCTTTAGCGCCCTGGAGCCTCTGGTGGATCTGCCAATCGGCATCAACATCACCAGGTTC
CAGACACTGCTGGCCCTGCACCGCAGCTACCTGACACCTGGCGACAGCTCCTCTGGATGGACCGCCGGGGCCGCCGCCTACTATGTGGGCTATCTGCAGCCA
CGGACCTTTCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTGCGCCCTGGACCCCCTGAGCGAGACCAAGTGTACACTGAAGTCCTTC
ACCGTGGAGAAGGGCATCTATCAGACATCCAATTTTCGGGTGCAGCCCACCGAGTCTATCGTGAGATTCCCAAATATCACAAACCTGTGCCCCTTCGGCGAG
GTGTTTAACGCCACCAGATTCGCCAGCGTGTACGCCTGGAATCGGAAGAGAATCTCTAACTGCGTGGCCGACTATAGCGTGCTGTACAACTCTGCCAGCTTT
TCCACCTTCAAGTGCTATGGCGTGTCCCCCACAAAGCTGAATGACCTGTGCTTCACCAACGTGTACGCCGATTCTTTTGTGATCAGGGGCGACGAGGTGAGA
CAGATCGCACCAGGCCAGACAGGCAAGATCGCCGACTACAATTATAAGCTGCCCGACGATTTCACCGGCTGCGTGATCGCCTGGAACTCTAACAATCTGGAT
AGCAAAGTGGGCGGCAACTACAATTATCTGTACAGGCTGTTCCGCAAGAGCAATCTGAAGCCTTTTGAGCGGGACATCTCTACAGAGATCTACCAGGCCGGC
AGCACCCCATGCAATGGCGTGGAGGGCTTCAACTGTTATTTTCCCCTGCAGTCCTACGGCTTTCAGCCTACCAACGGCGTGGGCTATCAGCCATACAGAGTG
GTGGTGCTGAGCTTCGAGCTGCTGCACGCACCAGCAACAGTGTGCGGACCTAAGAAGTCCACCAATCTGGTGAAGAACAAGTGCGTGAACTTCAACTTCAAC
GGCCTGACCGGCACAGGCGTGCTGACCGAGTCCAATAAGAAGTTTCTGCCCTTCCAGCAGTTTGGCCGGGACATCGCCGATACCACAGACGCCGTGAGAGAC
CCCCAGACCCTGGAGATCCTGGACATCACACCTTGCTCCTTCGGCGGCGTGTCTGTGATCACACCTGGCACCAATACAAGCAACCAGGTGGCCGTGCTGTAT
CAGGACGTGAATTGTACCGAGGTGCCAGTGGCCATCCACGCCGATCAGCTCACCCCCACATGGCGGGTGTACTCCACCGGCTCTAACGTGTTCCAGACAAGA
GCCGGCTGCCTGATCGGAGCCGAGCATGTGAACAATTCCTATGAGTGCGACATCCCCATCGGAGCCGGCATCTGTGCCTCTTACCAGACCCAGACAAACAGC
CCTGGCTCCGCCAGCTCCGTGGCCTCTCAGAGCATCATCGCCTATACCATGAGCCTGGGGGCCGAGAACAGCGTGGCCTACTCTAACAATAGCATCGCCATC
CCCACCAACTTTACAATCTCTGTGACCACAGAGATCCTGCCTGTGAGCATGACCAAGACATCCGTGGACTGCACAATGTATATCTGTGGCGATTCCACCGAG
TGCTCTAACCTGCTGCTGCAGTACGGCAGCTTCTGTACCCAGCTCAACCGGGCCCTGACAGGAATCGCAGTGGAGCAGGACAAGAACACACAGGAGGTGTTT
GCCCAGGTGAAGCAGATCTACAAGACCCCACCCATCAAGGACTTCGGCGGCTTTAATTTCTCCCAGATCCTGCCAGATCCCTCCAAGCCATCTAAGCGGAGC
TTCATCGAGGACCTGCTGTTTAACAAGGTGACCCTGGCCGATGCCGGCTTTATCAAGCAGTATGGCGATTGCCTGGGCGACATCGCCGCCAGAGACCTGATC
TGTGCCCAGAAGTTCAATGGCCTGACCGTGCTGCCTCCACTGCTGACAGATGAGATGATCGCACAGTACACAAGCGCCCTGCTGGCCGGCACCATCACATCC
GGATGGACCTTCGGGGCCGGGGCCGCCCTGCAGATCCCCTTCGCCATGCAGATGGCCTATAGGTTTAACGGCATCGGCGTGACCCAGAATGTGCTGTACGAG
AACCAGAAGCTGATCGCCAATCAGTTCAACTCCGCCATCGGCAAGATCCAGGACAGCCTGTCTAGCACAGCCTCCGCCCTGGGCAAGCTGCAGGATGTGGTG
AATCAGAACGCCCAGGCCCTGAATACCCTGGTGAAGCAGCTCAGCAGCAACTTCGGGGCCATCAGCAGCGTGCTGAACGACATCCTGAGCCGGCTGGACCCC
CCTGAGGCAGAGGTGCAGATCGACAGGCTGATCACAGGCCGCCTGCAGTCTCTGCAGACCTATGTGACACAGCAGCTCATCAGGGCCGCCGAGATCAGAGCC
AGCGCCAATCTGGCCGCCACCAAGATGTCCGAGTGCGTGCTGGGCCAGTCTAAGCGCGTGGACTTCTGTGGCAAGGGCTATCACCTGATGAGCTTTCCACAG
TCCGCCCCCCACGGAGTGGTGTTCCTGCATGTGACCTACGTGCCTGCCCAGGAGAAGAACTTTACCACAGCCCCAGCCATCTGCCACGATGGCAAGGCACAC
TTCCCCAGGGAGGGCGTGTTCGTGAGCAACGGCACCCACTGGTTCGTGACACAGCGCAACTTCTACGAGCCTCAGATCATCACCACAGACAATACATTCGTG
TCTGGCAACTGTGACGTGGTCATCGGCATCGTGAACAATACCGTGTATGATCCCCTGCAGCCTGAGCTGGACAGCTTCAAGGAGGAGCTGGATAAGTACTTT
AAGAATCACACCTCCCCCGACGTGGATCTGGGCGACATCTCTGGCATCAATGCCAGCGTGGTGAACATCCAGAAGGAGATCGACAGGCTGAACGAGGTGGCC
AAGAATCTGAACGAGAGCCTGATCGATCTGCAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTTCTTTTTCATCATCGGCCTGATCATCGGCCTGTTCCTG
GTGCTGCGCGTGGGCATCCACCTGTGCATCAAGCTGAAGCACACCAAGAAGAGGCAGATCTACACAGACATCGAGATGAACCGCCTGGGCAAGTGA

SEQ ID NO: 25는 VSV-G TMCyt SWAP 아미노산 서열을 갖는, SARS-CoV-2 스파이크 당단백질이다.

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIY
SKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV
VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAI
PTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS
FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE
NQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRA
SANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
SGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIK**FFFIIGLIIGLFL**
VLRVGIHLCIKLKHTKKRQIYTDIEMNRLGK

SEQ ID NO: 26은 VSV-G TMCyt SWAP, 염기 서열을 갖는 SARS-CoV-2 스파이크 당단백질이다.

CTCGAGGTTTAAACGAATTCCGCCACCATGTTTGTTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATTACCC

CCTGCATACACTAATTCTTTCACACGTGGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTT
TCCAATGTTACTTGGTTCCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCT
TCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTT
ATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCT
AGTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTTAAG
AATATTGATGGTTATTTTAAAATATATTCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTG
CCAATAGGTATTAACATCACTAGGTTTCAAACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCA
GCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCA
GAAACAAAGTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCCTAAT
ATTACAAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTAT
TCTGTCCTATATAATTCCGCATCATTTTCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCA
TTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTT
ATAGCTTGGAATTCTAACAATCTTGATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGAT
ATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAAT
GGTGTTGGTTACCAACCATACAGAGTAGTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAA
AACAAATGTGTCAATTTCAACTTCAATGGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATT
GCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAAT
ACTTCTAACCAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCT
ACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGC
GCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTT
GCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTAGTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACA
ATGTACATTTGTGGTGATTCAACTGAATGCAGCAATCTTTTGTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGAA
CAAGACAAAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAATATTACCA
GATCCATCAAAACCAAGCAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTT
GGTGATATTGCTGCTAGAGACCTCATTTGTGCACAAAAGTTTAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCT
GCACTGTTAGCGGGTACAATCACTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTTAATGGTATT
GGAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGATTGCCAACCAATTTAATAGTGCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGT
GCACTTGGAAAACTTCAAGATGTGGTCAACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTA
AATGATATCCTTTCACGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAA
TTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAG
GGCTATCATCTTATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACTGCTCCT
GCCATTTGTCATGATGGAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAACACAAAGGAATTTTTATGAACCACAA
ATCATTACTACAGACAACACATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCA
TTCAAGGAGGAGTTAGATAAATATTTTAAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAATGCTTCAGTTGTAAACATTCAAAAA
GAAATTGACCGCCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATTTTTC
TTTATCATAGGGTTAATCATTGGACTATTCTTGGTTCTCCGAGTTGGTATCCATCTTTGCATTAAATTAAAGCACACCAAGAAAAGACAGATTTATACAGAC
ATAGAGATGAACCGACTTGGAAAGTAAGAATTCCACGTGGGATCC

SEQ ID NO: 27은 VSV-G TMCyt SWAP을 갖는 SARS CoV-2 스파이크 당단백질, 염기 서열, 인간 세포에서의 발현에 최적화된 코돈이다.

CTCGAGGTTTAAACGAATTCCGCCACCATGTTCGTGTTCCTGGTGCTGCTGCCCCTGGTGTCCAGCCAGTGCGTGAATCTGACCACCCGGACCCAACTGCCT
CCCGCCTACACAAACTCTTTCACCAGAGGGGTTTATTACCCCGATAAGGTGTTCAGAAGCTCAGTGCTTCATTCTACCCAGGACCTGTTTCTGCCTTTTTTC
AGCAACGTCACATGGTTCCACGCCATCCACGTCAGCGGAACCAACGGCACGAAGCGGTTCGACAATCCTGTGCTGCCTTTTAACGACGGCGTCTACTTTGCC
AGCACGGAAAAGAGCAACATTATCCGGGGATGGATCTTCGGCACCACCCTGGACTCTAAAACCCAGAGCCTGTTGATCGTGAACAACGCAACCAATGTGGTG
ATCAAGGTCTGCGAGTTCCAATTTTGCAACGATCCTTTCCTGGGCGTGTACTACCACAAGAACAACAAGTCTTGGATGGAATCTGAGTTCCGCGTCTACAGC
AGCGCAAACAACTGCACATTTGAGTACGTGTCTCAGCCTTTTCTGATGGACCTGGAAGGAAAGCAGGGAAATTTCAAGAACCTGCGGGAGTTCGTGTTCAAG
AACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCCATCAACCTCGTGAGAGACCTGCCCCAGGGCTTCAGCGCCCTGGAACCCCTGGTGGACCTT
CCCATAGGAATCAACATCACACGGTTCCAGACACTGCTGGCCCTGCATAGAAGCTACCTGACCCCTGGAGATTCTAGCAGCGGCTGGACCGCCGGCGCTGCC
GCTTACTACGTCGGATACCTGCAGCCTAGAACCTTCCTGTTGAAGTACAACGAGAACGGCACCATCACAGATGCCGTGGACTGCGCCCTGGACCCCCTGAGC
GAGACAAAGTGCACCCTGAAGAGCTTCACCGTGGAGAAGGGCATCTACCAGACAAGCAACTTCAGAGTGCAGCCTACCGAGTCAATCGTGAGATTCCCAAAC
ATCACCAACCTTTGTCCTTTCGGCGAGGTATTTAACGCCACCCGGTTCGCCAGCGTGTACGCCTGGAATAGGAAGCGGATCAGCAACTGCGTGGCCGATTAC

AGCGTGCTCTATAACAGCGCCAGTTTTAGCACTTTCAAGTGCTACGGAGTCTCTCCTACAAAGCTGAACGACCTGTGCTTCACCAACGTGTATGCCGACAGC
TTCGTCATCCGGGGCGACGAGGTGCGACAGATCGCTCCTGGCCAGACCGGCAAGATAGCCGACTACAACTACAAGCTGCCTGACGACTTCACAGGCTGCGTG
ATCGCTTGGAACAGCAACAATCTGGATAGCAAAGTGGGCGGCAACTATAACTACCTGTACAGACTGTTCCGGAAGTCCAATCTCAAGCCGTTTGAGAGAGAC
ATCAGCACCGAAATCTACCAGGCTGGATCTACACCCTGCAACGGCGTCGAAGGCTTCAATTGTTACTTCCCTCTGCAATCTTACGGCTTCCAGCCCACCAAC
GGCGTGGGCTACCAGCCCTACAGAGTGGTTGTGCTGAGCTTCGAGCTGCTGCACGCCCCAGCTACAGTGTGCGGCCCTAAGAAATCTACAAACCTGGTCAAG
AACAAGTGTGTGAACTTCAACTTCAATGGCCTGACGGGCACCGGCGTGCTGACAGAGAGCAACAAGAAGTTCCTGCCTTTCCAGCAATTTGGCAGAGACATC
GCCGACACCACCGACGCCGTGCGCGACCCTCAGACCCTGGAAATTCTGGACATCACCCCATGTTCTTTCGGCGGCGTGTCCGTCATTACGCCAGGCACCAAT
ACCAGCAACCAGGTGGCCGTGCTTTATCAGGATGTGAATTGTACCGAAGTTCCTGTTGCAATCCACGCCGACCAACTGACCCCCACATGGAGAGTGTACTCT
ACCGGCAGCAACGTGTTCCAAACGAGAGCCGGATGCCTGATTGGAGCTGAGCATGTGAACAACAGCTACGAGTGCGATATTCCAATCGGAGCCGGCATCTGC
GCCTCCTACCAAACACAAACCAACTCCCCTCGTAGAGCGAGAAGCGTGGCCTCTCAGAGCATCATCGCCTACACCATGAGCCTGGGTGCCGAAAACTCCGTG
GCTTACTCCAACAACAGCATCGCCATCCCTACAAATTTCACCATCAGCGTGACAACCGAGATCCTGCCTGTGTCCATGACCAAGACCAGCGTGGACTGCACG
ATGTACATCTGCGGAGATAGCACCGAGTGCAGCAATCTGCTACTGCAGTATGGCAGCTTCTGCACCCAACTGAACAGAGCACTGACCGGCATTGCTGTGGAA
CAGGACAAGAATACCCAGGAGGTGTTCGCCCAAGTGAAGCAGATTTACAAGACACCCCCTATCAAGGACTTCGGAGGCTTCAACTTCAGCCAGATCCTGCCT
GACCCTAGCAAGCCAAGCAAAAGATCCTTTATCGAAGATCTGCTGTTTAACAAGGTGACACTGGCCGATGCCGGCTTTATCAAGCAGTACGGCGACTGCCTG
GGAGACATCGCCGCCAGAGACCTGATCTGTGCTCAGAAATTTAACGGGCTGACCGTGCTGCCACCTCTGCTGACAGATGAGATGATCGCTCAGTACACCAGC
GCCCTGCTGGCCGGCACAATTACCTCCGGCTGGACCTTCGGAGCCGGAGCCGCCCTGCAGATCCCCTTCGCCATGCAGATGGCCTACCGGTTCAATGGCATC
GGCGTCACCCAAAACGTGCTCTATGAGAACCAGAAGCTGATCGCAAACCAGTTCAACTCCGCCATCGGTAAGATCCAGGACAGTCTGAGCAGCACGGCGTCT
GCCCTGGGCAAGCTCCAGGACGTGGTGAACCAGAACGCCCAGGCCCTTAACACCCTGGTGAAACAACTGAGCAGCAACTTCGGTGCCATTTCCAGCGTTCTC
AATGACATCCTGAGCAGACTGGATAAGGTGGAAGCCGAGGTGCAGATCGACCGGCTGATCACCGGACGGCTGCAGAGCCTGCAGACGTACGTGACCCAGCAA
TTAATCAGAGCTGCCGAGATCAGAGCCAGCGCCAATCTGGCTGCCACCAAAATGAGCGAATGTGTGCTGGGCCAGTCAAAGAGAGTGGATTTTTGTGGCAAA
GGCTACCACCTGATGTCCTTCCCTCAGTCTGCCCCTCACGGCGTGGTGTTCCTCCATGTGACCTATGTGCCTGCTCAGGAGAAGAACTTTACCACAGCCCCT
GCTATCTGCCACGACGGAAAGGCCCACTTCCCCAGAGAGGGCGTGTTTGTGTCCAATGGCACACACTGGTTCGTGACCCAAAGAAACTTCTACGAGCCCCAG
ATCATCACCACAGACAACACCTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATCGTGAACAACACAGTGTACGACCCCCTGCAACCTGAGCTGGACAGC
TTCAAAGAGGAACTGGACAAATACTTCAAGAATCACACCAGCCCTGATGTGGATCTGGGCGACATCAGCGGCATCAACGCCAGCGTCGTGAACATCCAGAAG
GAAATCGACAGACTGAACGAAGTGGCCAAGAACCTGAACGAGAGCCTCATCGATCTGCAGGAGCTGGGCAAGTACGAGCAGTACATCAAATGGCCTTTCTTC
TTCATCATCGGCCTGATTATCGGCCTGTTCCTCGTGCTGAGAGTGGGCATCCACCTGTGCATCAAGCTTAAGCACACAAAAAAGCGGCAGATTTACACCGAC
ATCGAGATGAACCGGCTGGGCAAATGAGAATTCCACGTGGGATCC

발명을 실시하기 위한 구체적인 내용

[0026] SARS-CoV, MERS-CoV 및 SARS-Cov-2와 같은 코로나바이러스는 약 30,000개 염기의 RNA 게놈을 갖는 외피 바이러스이다. 그들은 코로나바이러스의 베타 속에 속한다. 그들은 숙주 세포에서 유래한 지질 이중층으로 둘러싸인 뉴클레오캡시드를 함유한다. 외피에 고정된 스파이크 단백질("S"라고 함)은 숙주 수용체에 결합한 다음 바이러스와 숙주 막(viral and host membranes)을 융합하여 코로나바이러스가 숙주 세포로 들어가는 것을 매개한다. 정의된 수용체 결합 도메인(defined receptor binding domain)은 안지오텐신 전환 효소 2(ACE2)에 대한 수용체이다. (Wan et al., J. Vir. (2020) 94:1). 코로나바이러스 S 단백질은 S1 서브유닛의 3개 복제(copies) 및 S2 서브유닛의 3개 복제를 포함한다. 코로나바이러스 S 단백질은 단백질 생합성 동안 퓨린에 의해 S1 및 S2 서브유닛으로 절단된다. 두 개의 서브유닛은 삼량체화되고 준안정 예비융합 형태(metastable prefusion conformation)로 폴딩된다(fold). S1 서브유닛은 수용체 결합을 담당하고 S2 서브유닛은 막 융합을 매개한다.

[0027] SARS-CoV 및 SARS-CoV-2 스파이크 단백질은 약 76%의 서열 상동성(sequence homology)을 공유하며, 이는 이 두 바이러스가 동일한 수용체인 ACE2를 공유함을 시사한다. SARS-CoV-2와 MERS-CoV 사이에는 서열 유사성이 더 낮다.

[0028] 2019년 12월부터 2020년 3월까지 4개월 동안 환자로부터 분리된 SARS-CoV-2의 게놈에 대한 연구에 따르면 인간 균주의 전반적인 유사성이 4개월 동안 감소한 것으로 나타났다. 이는 바이러스의 돌연변이가 인구 내에서 0.988468로발생했음을 나타내며, 이는 평균 348.33의 뉴클레오티드 차이에 해당한다. 이러한 변화는 이 바이러스의 진화적 변화를 의미하며, 이는 약독화 또는 더 치명적인 균주를 초래할 수 있다(Li et al 2020. Xidan University). 그 후, 2020년 3월 D614 이전에 우세했던 바이러스 변종는 D614가 잘 확립된 영역에서도 스파이크 단백질 G614에 단일 아미노산 변화를 갖는 다른 변종에 의해 추월되었다(Korber et al, (2020) 세포, 4:812-827). 그 후 2020년 후반에 보고된 COVID-19 사례의 예상치 못한 증가는 새로운 변종인 영국의 B.1.1.7과 남아프리카의 501Y.V2의 출현에 기인한다(Fontanet et al, (2021) the Lancet , 397: 952-954) . 두 변종 모두 스파이크 단백질의 수용체 결합 도메인에 돌연변이(N501Y)가 있으며, 이 돌연변이는 전달 증가(increased

transmission)에 기여하는 것으로 보고된 것이다. 전달 증가에 대한 추정치는 40%에서 70% 사이이다. 501Y.V2 변종는 스파이크 단백질에 2개의 추가 돌연변이(E484K 및 K417N)가 있어 항체에 잠재적인 면역 탈출을 부여한다. 우려되는 또 다른 변종인 P.1이 또 다른 돌연변이 세트(N501Y, E484K 및 K417T)와 함께 브라질에서 나타났다.

[0029] 중요한 관심사는 현재 사용 가능한 COVID-19 백신이 SARS-CoV-2 변종의 감염이나 질병으로부터 보호할 수 있는지 여부이다. 예비 연구에 따르면 mRNA COVID-19 백신으로 면역된 개인의 혈청은 영국 변종와 유사한 유사바이러스를 중화했지만 남아프리카 공화국 변종와 유사한 유사바이러스에 대해서는 덜 효과적이었다(Yang et al (2021) Nature, doi.org/10.1038/ s41586-021-03324-6). 또한, 바이러스 벡터 백신을 사용한 예비 연구의 결과는 영국 변종에 대해서는 우수한 효능을 나타내었지만 남아프리카 변종에 대해서는 낮은 효능을 보여주었다(Madhi et al (2021) N.E.J.M. DOI: 10.1056/NEJMoa2102214). 따라서, 다중 돌연변이를 포함하는 코로나바이러스의 변종 균주에 의한 감염으로부터 보호하기 위해 광범위하게 반응성인 항체의 생산을 유도할 수 있는 백신이 필요할 것으로 보인다.

[0030] 본 발명자들은 VLP를 포함하는 베타 코로나바이러스에 대한 백신을 만들었다. VLP는 일반적으로 하나 이상의 바이러스 단백질로 구성된 다중 단백질 구조이다. VLP는 바이러스의 형태를 모방하지만 유전 물질이 없기 때문에 감염성이 아니다. 그들은 적절한 상황에서 바이러스 구조 단백질의 발현에 따라 형성(또는 "자가 조립")할 수 있다. VLP 백신은 약독화 바이러스를 사용하여 제조된 보다 전통적인 백신의 단점 중 일부를 극복한다. 생산 공정 중에 살아있는 바이러스가 없어도 생산할 수 있기 때문이다. 다양한 VLP가 제조되어 왔다. 예를 들어, 외피 단백질 및/또는 표면 당단백질이 있거나 없는 단일 또는 다중 캡시드 단백질을 포함하는 VLP가 제조되었다. 어떤 경우에는 VLP가 외피가 없고(비-enveloped) 단 하나의 주요 캡시드 단백질의 발현에 의해 조립된다. 다른 경우에, VLP는 엔벨로프되고 상응하는 천연 바이러스에서 발견되는 다중 항원성 단백질을 포함할 수 있다. 외피 VLP의 자가 조립은 숙주 세포막(Garrone et al., 2011 Science Trans. Med. 3: 1-8)과의 "출아"와 융합에 필요한 복잡한 반응 때문에 외피가 없는 VLP보다 더 복잡하다. VLP는 완전히 둘러싸인 별도의 입자를 형성한다. 따라서, 외피 VLP의 자가 조립은 성공적이지 않을 수 있고 외피 VLP 입자의 형성 및 안정성을 예측하기 어렵다. 손상되지 않은(intact) VLP의 형성은 전자 현미경을 사용하여 입자를 이미징하여 확인할 수 있다.

[0031] VLP는 일반적으로 해당 고유 바이러스와 유사하며 다가 미립자 구조일 수 있다. 본 발명은 VLP의 맥락에서 표면 당단백질의 제시가 다른 형태의 항원 제시, 예를 들어 VLP와 회합되지 않은 가용성 항원과 비교하여 이러한 폴리펩타이드에 대한 중화 항체의 유도에 유리하다는 인식을 포함한다. 중화 항체는 대부분 3차 또는 4차 구조를 인식한다(recognize). 이것은 종종 외피 당단백질과 같은 항원 단백질을 본래의 바이러스 형태로 제시해야 한다. VLP의 현재 에피토프는 B 세포 수용체의 효율적인 가교를 가능하게 하는 고도로 구조화되고 반복적인 어레이에 있어 고친화성 B 세포의 활성화 및 확장 및 후속 항체 생산을 유도한다(Bachmann, 1993). 실제로, B 세포 항원의 VLP 발현은 동일한 양의 재조합 단백질을 사용한 면역화에 비해 중화 역가를 10배 이상 향상시켰다(Kirchmeier, 2014). 따라서 VLP를 백신 방식으로 사용하면 재조합 단백질 또는 DNA/mRNA 기반 접근 방식에 비해 친화도가 더 높은 B 세포 레퍼토리를 확장할 수 있으며, 후자의 접근 방식은 널리 사용되는 두 가지 COVID-19 백신에서 사용된다.

[0032] 본 발명의 VLP는 레트로바이러스 벡터를 포함한다. 레트로바이러스는 레트로비리다에(Retroviridae) 계통에 속하는 외피 RNA 바이러스이다. 레트로바이러스에 의한 숙주 세포의 감염 후, RNA는 효소 역전사효소를 통해 DNA로 전사된다. 그런 다음 DNA는 인테그라제 효소에 의해 숙주 세포의 게놈에 통합되고 그 후 숙주 세포의 DNA의 일부로 복제된다. 레트로비리다에 계통에는 알파retrovirus, Betaretrovirus, Gammearetrovirus, Deltaretrovirus, Epsilonretrovirus, Lentivirus 및 Spumavirus 속이 포함된다. 이 레트로바이러스 계열의 숙주는 일반적으로 척추동물이다. 레트로바이러스는 숙주 세포막에서 유래한 지질 이중층으로 둘러싸인 구형 뉴클레오캡시드(바이러스 구조 단백질과 복합체를 이루는 바이러스 게놈)를 포함하는 감염성 비리온을 생성한다.

[0033] 레트로바이러스 벡터는 레트로바이러스 유래 게놈이 부족하여 복제되지 않는 VLP를 생성하는 데 사용할 수 있다. 이것은 레트로바이러스의 대부분의 암호화 영역을 유전자 또는 염기 서열로 대체함으로써 달성되며, 벡터가 추가 복제에 필요한 단백질을 만들 수 없도록 한다. 입자 표면에 존재하는 당단백질의 특성에 따라 VLP는 숙주 세포에 결합하여 숙주 세포에 들어가는 능력이 제한되지만 증식할 수는 없다. 따라서 VLP는 면역원성 조성물(예: 백신)로서 안전하게 투여될 수 있다.

[0034] 본 발명은 하나 이상의 레트로바이러스 구조 단백질을 포함하는 VLP를 이용한다. 일부 구현예에서, 본 발명에 따라 사용하기 위한 구조 단백질은 알파레트로바이러스(예를 들어, 조류 백혈병 바이러스), 베타레트로바이러스

(마우스 유방 종양 바이러스), 감미아레트로바이러스(쥐 백혈병 바이러스), 델타레트로바이러스(소 백혈병 바이러스), 엡실론레트로바이러스(Walley)이다. 피부 육종 바이러스), 렌티바이러스(인간 면역결핍 바이러스 1) 또는 스푸마바이러스(침팬지 거품 바이러스) 구조 단백질이다. 특정 구현예에서, 구조적 다중단백질은 뮤린 백혈병 바이러스(MLV) 구조 단백질이고, 본 발명의 한 구현예에서, 몰로니 뮤린 백혈병 바이러스(MMLV)의 구조적 단백질이다. 이러한 레트로바이러스의 게놈은 데이터베이스에서 쉽게 구할 수 있다.

[0035] 일부 구현예에서, 본 발명에 따라 사용하기 위한 레트로바이러스 구조 단백질은 Gag 폴리펩타이드이다. 레트로바이러스의 Gag 단백질은 전체적인 구조적 유사성을 가지며 레트로바이러스의 각 그룹 내에서 아미노산 수준에서 보존된다. 레트로바이러스 Gag 단백질은 주로 바이러스 조립에서 기능한다. 일부 숙주 세포에서 일부 바이러스(예: MMLV와 같은 쥐 백혈병 바이러스)의 Gag 발현은 발현 생성물이 VLP로 자가 조립되는 결과를 초래할 수 있다. 다단백질 형태의 Gag 유전자 발현 산물은 VLP의 핵심 구조 단백질을 생성을 증가시킨다. 기능적으로, Gag 폴리프로테인은 3개의 도메인으로 나뉘며,: Gag 폴리프로테인을 세포막으로 표적화하는 막 결합 도메인; Gag 중합을 촉진하는 상호작용 도메인; 및 숙주 세포로부터 초기 비리온의 방출을 촉진하는 후기 도메인이 그것이다. 일반적으로 바이러스 입자 집합을 매개하는 Gag 단백질의 형태는 다단백질이다. 레트로바이러스는 Gag 폴리단백질로 구성된 미성숙 캡시드를 조립하지만 미성숙 바이러스 입자의 구조 단백질로서 Gag를 갖는 바이러스 프로테아제와 같은 다른 바이러스 요소가 없다.

[0036] 본 발명에 사용하기에 적합한 Gag 폴리펩타이드는 공지된 레트로바이러스 Gag 폴리펩타이드와 실질적으로 상동성이다. MMLV-Gag 유전자는, 성숙한 비리온에서, MLV 프로테아제에 의해 4개의 구조 단백질(매트릭스(MA), p12, 캡시드(CA); 및뉴클레오캡시드(NC))로 분해되는 65kDa 폴리단백질 전구체를 암호화한다. MLV 프로테아제가 없는 경우, 다단백질은 절단되지 않은 채로 남아 있고 생성된 입자는 미성숙한 형태로 남아 있다. 미성숙 입자의 형태는 성숙한 입자의 형태와 다르다. 본 발명의 일부 구현예에서, Gag 서열은 MLV 프로테아제를 암호화하는 유전자를 포함하지 않는다. MMLV 핵산을 암호화하는 유전자는 SEQ ID NO: 2이다. MMLV 핵산의 예시적인 코돈 최적화 서열은 SEQ ID NO: 3으로 제공된다.

[0037] 따라서, 일부 구현예에서, 본 발명에 적합한 Gag 폴리펩타이드는 MMLV-Gag 폴리펩타이드(SEQ ID NO: 1)와 실질적으로 상동성이다. 일부 구현예에서, 본 발명에 적합한 Gag 폴리펩타이드는, SEQ ID NO:1과 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92% 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 아미노산 서열을 갖는다. 일부 구현에에서, 적합한 MMLV-Gag 폴리펩타이드는, SEQ ID NO: 1과 동일하거나 실질적으로 동일한 핵산 서열에 의해 암호화된다.

[0038] 일부 구현예에서, 적합한 MMLV-Gag 폴리펩타이드는, SEQ ID NO:2에 대한 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 서열 동일성을 갖는 핵산 서열에 의해 암호화된다. 일부 구현예에서, 적합한 MMLV-Gag 폴리펩타이드는 SEQ ID NO: 2 또는 그의 코돈 축퇴 버전(SEQ ID NO: 2)을 갖는 핵산 서열에 의해 암호화된다.

[0039] 이 분야의 통상의 기술자에게 잘 알려진 바와 같이, 특정 아미노산(즉, 코돈)을 암호화하는 핵산을, 숙주 유기체에서 더 잘 발현다른 코돈으로 대체함으로써 핵산 서열의 발현을 개선하는 것이 가능하다. 이 효과가 발생하는 한 가지 이유는, 상이한 유기체마다 상이한 코돈을 선호하는 사실에 기인한다. 코돈 선호도에 따라 더 나은 발현을 달성하기 위해 핵산 서열을 변경하는 과정을 코돈 최적화라고 한다. 다양한 유기체에서 코돈 사용 바이아스(codon use bias)을 분석하기 위한 다양한 방법이 당업계에 공지되어 있고, 코돈 최적화된 유전자 서열의 설계에서 이러한 분석을 구현하기 위해 많은 컴퓨터 알고리즘이 개발되어 왔다. 따라서, 일부 구현예에서, 적합한 MMLV-Gag 폴리펩타이드는, MMLV-Gag를 암호화하고 그리고 SEQ ID NO:3에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93% 94%, 95%, 96%, 97%, 98%, 99% 서열 동일성(sequence identity)을 가지는,핵산 서열의 코돈 최적화 버젼(codon optimized version)에 의해 암호화된다. 일부 구현예에서, 적합한 MMLV-Gag 폴리펩타이드는, SEQ ID NO:3과 동일하거나 실질적으로 동일한 핵산 서열에 의해 암호화된다.

[0040] 당업계에 잘 알려진 바와 같이, 아미노산 또는 핵산 서열은, 염기 서열에 대한 BLASTN 및 BLASTP, gapped BLAST 및 PSI-BLAST와 같은 상용 컴퓨터 프로그램에서 이용 가능한 알고리즘을 포함하는 다양한 알고리즘을 사용하여 비교할 수 있다. 이러한 프로그램의 예는, Altschul, et al., 1990, J. Mol. Biol., 215(3): 403-410; Altschul, et al., 1996, Methods in Enzymology 266:460-480; Altschul, et al., 1997 Nucleic Acids Res. 25:3389-3402; Baxevanis, et al., 1998, Bioinformatics: A Practical Guide to Analysis of Genes and 단백질s, Wiley; 및 Misener, et al., (eds.), 1999, Bioinformatics Methods and Prot℃ols (Methods in Molecular Biology, Vol. 132), Humana Press에 설명되어 있다. 상동 서열을 확인하는 것 외에도, 위에서 언급

한 프로그램은 일반적으로 상동 정도의 표시를 제공한다. 일부 구현예에서, 만일 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상의 상응하는 잔기가, 잔기의 관련 스트레치에 대해(over a relevant stretch) 상동이면, 2개의 서열이 실질적으로 상동인 것으로 간주된다. 일부 구현예에서, 그 관련 스트레치는 완전한 서열이다(the relevant stretch is a complete sequence). 일부 구현예에서, 그 관련 스트레치는, 적어도 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500개 이상의 잔기이다.

[0041] 대안적으로, 본 발명에 사용된 Gag 폴리펩타이드는, 실질적인 자가 조립 활성을 유지하면서, 야생형 또는 자연 발생 Gag 폴리펩타이드와 비교하여, 하나 이상의 아미노산 치환, 결실(deletions) 및/또는 삽입을 함유하는 변형된 레트로바이러스 Gag 폴리펩타이드일 수 있다. 일반적으로, 본질적으로, Gag 단백질은 레트로바이러스 프로테아제, 역전사 효소(reverse transcriptase) 및 인테그라제 효소 활성을 함유할 수 있는 큰 C-말단 신장부(C-terminal extension)를 포함한다. 그러나. VLP 조립에는 일반적으로 그러한 구성 요소가 필요하지 않다. 일부 경우에, 레트로바이러스 Gag 단백질은 단독(예: C-말단 신장부 부재; 게놈 RNA, 역전사 효소, 바이러스 프로테아제 또는 외피 단백질 중 하나 이상 부재)으로 자가 조립되어 시험관 내 및 생체 내 모두에서 VLP를 형성할 수 있다(Sharma S et al., 1997, Pr℃. Natl. Acad. Sci. USA 94: 10803-8).

[0042] 본 출원의 발명자들은 대상체에서 면역 반응을 일으킬 수 있는 표면에 베타 코로나바이러스 외피 당단백질을 발현하는 VLP를 만들었다. 체액성 면역 반응은 ,항체 분자에 의해 매개되는 면역 반응이다. 중화 항체라고 하는 특정 항체는, 바이러스에 의해 발현되는 특정 단백질 또는 항원을 인식하고 그에 결합함으로써(by recognizing and binding to a particular 단백질 or 항gen) 바이러스에 의한 감염 및 관련 생물학적 작용(biological effects)으로부터 세포를 방어한다. 코로나바이러스의 외피 단백질은 중화 항체 생산의 중요한 표적이다. 레트로바이러스 Gag-기반 외피 VLP가 중화 항체 반응을 유도할 목적으로 다양한 외피 당단백질을 발현하는데 사용될 수 있다는 것은 이 분야의 통상의 기술자에게 잘 알려져 있다. 보다 구체적으로, HIV-1 gp120, 메타뉴모바이러스 및 인플루엔자 HA와 같은 클래스 I 바이러스 융합 단백질뿐만 아니라 VSV G 단백질 및 CMV gB 단백질과 같은 클래스 III 융합 단백질의 발현에 대한 증거가 존재한다(Mammano et al., 1997, J Virol.71:3341-3345, Levy et al., 2013, Vaccine 31:2778-2785, Lemaitre et al., 2011, Clin. Microbiol. Infect.1:732-737, Garrone et al., Kirch201 al., 2014, CVI 21: 174-180). 그러나, 특히 MLV 유래 Gag에서 코로나바이러스 스파이크 단백질의 발현에 대해서는 알려진 바가 거의 없다. 미국 특허 8,920,812gh에서, 실시예 1은 클래스 II 바이러스 융합 단백질인 RSV F 당단백질을, MLV Gag를 사용하여 생산된 VLP의 표면에서, 발현하지 못하는 것을 기술하고 있다. 그 발명자는 RSV F 당단백질의 존재가 세포막을 통한 Gag 바이러스 입자의 출아(budding)를 방해한다고 가정하였다(컬럼 41, 라인 50 참조). 따라서 레트로바이러스 Gag 기반 외피 바이러스 유사 입자가 코로나바이러스 스파이크 단백질을 성공적으로 발현하는 데 사용될 수 있다는 것은 예측할 수 없었다. 그럼에도 불구하고, 본 발명자들은, VLP의 표면 상에, 하나 이상의 스파이크 폴리펩타이드 항원(예를 들어, SARS CoV-2, SARS CoV 및 MERS-CoV로부터)을 포함하는 베타 코로나바이러스 백신의 여러 상이한 구현예를 만들었다. 일부 구현예에서, 스파이크 폴리펩타이드 항원은 변형된 폴리펩타이드를 포함하여 구성된다. 일부 구현예에서, 스파이크 폴리펩타이드 항원은 하나보다 많은 유전적 변형(genetic modification)을 갖는다.

[0043] 일부 구현예에서, 본 발명의 VLP는, 베타 코로나바이러스로부터의 스파이크 폴리펩타이드(예를 들어, 베타 코로나바이러스 스파이크 폴리펩타이드의 세포외 부분의 전부 또는 일부) 및 자연에서 베타 코로나바이러스 단백질(예: 다른 바이러스)에서 발견되지 않은 막횡단 및/또는 세포질 도메인의 융합 단백질을 포함한다. . 일부 구현예에서, 융합 단백질은, 베타 코로나바이러스로부터의 스파이크 폴리펩타이드(예를 들어, 스파이크 폴리펩타이드의 세포외 부분의 전부 또는 일부) 및 VSV로부터의 당단백질 G(VSV-G라고 함)에서 자연에서 발견되는 막횡단 도메인 및/또는 세포질 도메인을 포함한다. VSV-G 단백질의 뉴클레오티드 및 아미노산 서열은 당업계에 공지되어 있다.

[0044] VSV-G의 막횡단 부위는 바이러스 당단백질을 세포막으로 표적화하는 기능을 할 수 있다(Compton T et al., 1989, Pr℃ Natl Acad Sci USA 86:4112-4116). 다른 단백질의 막횡단 및 세포질 도메인에 대한 VSV-G의 막횡단 및 세포질 도메인의 교환(swapping)은 VLP 제제에서 관심 단백질의 수율을 증가시키고, 그리고 항체 반응 중화에 대한 면역원성을 증가시키는 데 사용되었다(Garrone et al., 2011). 이 변형은, HCV-E1 단백질(Garrone et al, 2011) 및 CMV-gB 단백질(Kirchmeier et al, 2014)을 발현하는 VLP의 수율과 활성을 증가시키는 데 성공적이었다. 그러나, 이 변형은 또한 특정 바이러스 항원과 함께 사용될 때 면역원성의 상당한 손실과 관련이 있다. 또한, 일부 당단백질의 발현은 항원성 당단백질의 막횡단/세포질 도메인을 VSV의 막횡단/세포질 도메인으

로 대체한 후 감소했다. 예를 들어, SARS 바이러스에서 당단백질의 손실이 보고되었다(Broer et al., 2006, J. Vir. 80, 1302-1310). RSV에서, 항원성 표면 단백질이 막횡단 성분을 VSV의 서열로 대체하도록 변형될 때 면역원성의 상당한 손실이 관찰되었다(실시예 6 참조).

[0045] 일부 구현예에서, 본 발명의 면역원성 조성물은 SARS-CoV-2로부터의 야생형 스파이크 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되고, 그의 서열은 SEQ ID NO: 4 또는 SEQ ID NO: 4 의 코돈 축퇴성 버전이다. 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 5로 표시된다. SEQ ID NO: 5의 코돈 최적화된 버전은 SEQ ID NO: 6으로 표시된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 4에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 아미노산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은 SEQ ID NO: 4의 아미노산 서열 또는 SEQ ID NO: 4의 코돈 축퇴성 버전을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서 , 폴리펩타이드는, SEQ ID NO: 5에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%,92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 6에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%인 SEQ ID NO: 5의 서열 , 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 5의 핵산의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는 SEQ ID NO: 6을 갖는 핵산 서열에 의해 암호화된다.

[0046] 일부 구현예에서, 본 발명의 면역원성 조성물은, 막횡단 및 세포질 분절을 VSV의 상응하는 분절로 대체하도록 변형된 SARS-CoV-2로부터의 변형된 스파이크 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되며, 이의 서열은 SEQ ID NO : 26 또는 SEQ ID NO: 26의 코돈 축퇴 버전이다. 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 25로 제시된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 25에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상의 상동성인, 아미노 산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 25 또는 SEQ ID NO: 25의 코돈 축퇴 버전인, 아미노산 서열 을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 26에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 돌연변이된 폴리펩타이드는 SEQ ID NO: 26을 갖는 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 27과에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인,EjQ ID NO: 26의 코돈 최적화된 버전에 의해 암호화된다.

[0047] 일부 구현예에서, 본 발명의 면역원성 조성물은, SARS-CoV로부터의 야생형 스파이크 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되고, 이의 서열은 SEQ ID NO: 7 또는 SEQ ID NO: 7 의 코돈 축퇴성 버전이다. 폴리펩타이드를 암호화하는 핵산은, SEQ ID NO: 8로 표시된다. SEQ ID NO: 8의 코돈 최적화된 버전은 SEQ ID NO: 9로 표시된다. 일부 구현예에서, 본 발명은, SEQ ID NO:7에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 아미노산 서열을 갖는, 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은, EQ ID NO:7의 아미노산 서열 또는 EQ ID NO:7의 코돈 축퇴성 버전을 갖는, 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 8에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 9에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 8의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 9를 갖는 핵산 서열에 의해 암호화된다.

[0048] 일부 구현예에서, 본 발명의 면역원성 조성물은, MERS-CoV로부터의 야생형 스파이크 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되고, 그의 서열은 SEQ ID NO: 10 또는 SEQ ID NO: 10의 코돈 축퇴 버전이다. 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 11로 표시된다. SEQ ID NO:11의 코돈 최적화된 버전은, SEQ ID NO: 12로 표시된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 10에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 아미노산 서열을 가지는, 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은 SEQ ID NO: 10인 아미노산 서열 또는 SEQ ID NO: 10의 코돈 축퇴성 버전을 갖는 폴리펩타이드를 포함하여 구성되는 VLP

를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 11에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는 하기의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 12에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 11의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는 SEQ ID NO: 12를 갖는 핵산 서열에 의해 암호화된다.

[0049] 일부 구현예에서, 본 발명의 면역원성 조성물은 베타 코로나바이러스 스파이크 당단백질의 변종를 포함하여 구성되는 VLP를 포함하여 구성된다. 일부 구현예에서, 변종 스파이크 당단백질은 스파이크 폴리펩타이드로부터 퓨린 절단 부위를 결실시키도록(delete) 변형되었다. 일부 구현예에서, 스파이크 당단백질은, 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하도록 변형되었다. 일부 구현예에서, 스파이크 당단백질은 퓨린 절단 부위를 결실시키고 그리고 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하도록 변형되었다. 이러한 각각의 변형 사항은 아래에서 추가로 설명된다.

[0050] 코로나바이러스 스파이크 단백질은, 프로테아제인 퓨린이 비리온 성숙 과정에서 S 폴리펩타이드를 S1 및 S2 서브유닛으로 절단하는 부위를 포함하는 것으로 알려져 있다. 퓨린 절단 부위를 제거하여 스파이크 폴리펩타이드를 미성숙 형태로 유지하도록 아미노산 서열이 변형된, 변형된 스파이크 단백질 컨스트럭트가 생성되었다, 정상적인 절단과 성숙을 거치지 않는, 스파이크 단백질의 퓨린 절단 부위 돌연변이 버전(furin-cleavage site mutated version)은 향상된 세포 수용체 결합 및 세포 진입(enhanced cell receptor binding and cell entry)을 보일 수 있으며, 이는 이 구조에 대한 면역이 더 큰 중화 활성을 갖는 체액성 면역을 초래할 수 있음을 나타낸다.

[0051] 일부 구현예에서, 본 발명의 면역원성 조성물은, 야생형 또는 자연 발생 SARS-CoV-2 스파이크 폴리펩타이드와 비교하여 돌연변이된 퓨린 절단 부위를 갖는 변형된 SARS-CoV-2 스파이크 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성된다 . 예시적인 변형된 SARS-CoV-2 폴리펩타이드에 대한 서열은, SEQ ID NO: 16으로 제시된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 16에 대하여 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 아미노산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 16인 아미노산 서열 또는 SEQ ID NO: 16의 코돈 축퇴 버전을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO: 17에 대하여 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%,99% 이상 상동성인 핵산 서열에 의해 암호화된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO: 18에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 9 0%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 SEQ ID NO: 17의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 돌연변이된 폴리펩타이드는, SEQ ID NO: 18을 갖는 핵산 서열에 의해 암호화된다.

[0052] SARS-CoV 및 MERS-CoV에 대한 이전 연구에서, 2개의 아미노산 잔기를 프롤린 잔기로 치환하면, S2 서브유닛이 융합 전 형태로 안정화되는 것으로 알려져 있다(Pallesen et al., 2017 PNAS. 114:35; Wrapp et al(2020)Science: 367: 1260-1263). 그러므로, 이러한 돌연변이가 코로나 바이러스에 대한 면역 반응을 향상시킬 수 있다. 따라서, 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하도록 변형된 SARS-CoV-2 폴리펩타이드 컨스트럭트를 제조하였다. 일부 구현예에서, 본 발명의 면역원성 조성물은, 야생형 또는 자연 발생 SARS-CoV-2 폴리펩타이드와 비교하여, 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하도록 변형된 SARS-CoV-2 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성된다. 예시적인 변형된 폴리펩타이드의 서열은 SEQ ID NO: 13에 제시되어 있다. 일부 구현예에서, 본 발명은, SEQ ID NO: 13에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 아미노산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 13인 아미노산 서열 또는 SEQ ID NO: 13의 코돈 축퇴 버전을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO:14에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상 상동성인,핵산 서열에 의해 암호화된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO:15에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인. SEQ ID NO:14의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된

다. 일부 구현예에서, 변형된 폴리펩타이드는 SEQ ID NO:15를 갖는 핵산 서열에 의해 암호화된다.

[0053] 추가 변형에서, 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하도록 변형되고 퓨린 절단 부위를 제거하도록 추가로 변형된 SARS-CoV-2 폴리펩타이드 컨스트럭트가 제조되었다. 일부 구현예에서, 본 발명의 면역원성 조성물은, 야생형 또는 자연 발생 SARS-CoV-2 폴리펩타이드와 비교하여, 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하고 퓨린 절단 부위를 제거하도록 변형된 SARS-CoV-2 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성된다. . 예시적인 변형된 폴리펩타이드의 서열은 SEQ ID NO: 19에 제시되어 있다. 일부 구현예에서, 본 발명은, SEQ ID NO: 19에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 아미노산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 본 발명은 다음을 포함하는 면역원성 조성물을 포함하여 구성된다: 일부 구현예에서, 본 발명은, SEQ ID NO: 19인 아미노산 서열 또는 SEQ ID NO: 19의 코돈 축퇴 버전을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다.. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO: 20에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO: 21에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인 SEQ ID NO: 20의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 변형된 폴리펩타이드는 SEQ ID NO: 21을 갖는 핵산 서열에 의해 암호화된다.

[0054] 일부 구현예에서, 본 명세서에 기술된 VLP는, 코로나바이러스 스파이크 폴리펩타이드의 세포외 도메인(또는 그의 일부), 및 VSV로부터의 외피 단백질로부터 유래한막횡단 부위 및 세포질 꼬리를 포함하여 구성되는 융합 단백질을 포함하여 구성된다. 일부 구현예에서, 본 발명의 면역원성 조성물은, 막횡단 부위 및 세포질 꼬리를 VSV로부터의 막횡단 부위 및 세포질 꼬리로 대체하도록 변형된 코로나바이러스 스파이크 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성된다. 본 명세서에 기술된 임의의 코로나바이러스 스파이크 단백질은 막횡단 부위 및 세포질 꼬리를 VSV로부터의 막횡단 부위 및 세포질 꼬리로 대체하도록 변형될 수 있다.

[0055] 한 특정 구현예에서, 본 발명자들은 막횡단 부위 및 세포질 꼬리를 VSV로부터 유래한막횡단 부위 및 세포질 꼬리로 대체하고; 라이신(986) 및 발린(987) 잔기를 프롤린 잔기로 대체하며; 퓨린 절단 부위를 제거하도록 변형된, SARS-CoV-2 스파이크 단백질을 만들었다(constructed). 이 3중 변형된 SARS-CoV-2 단백질은, 향상된 안정성에 대한 이중 프롤린 돌연변이, 그리고 향상된 수용체 결합과 관련된 돌연변이된 퓨린 절단 부위를 포함한다. 또한, 이는, VLP 외피 상의 개선된 발현과 관련된, VSV로부터의 막횡단 부위 및 세포질 꼬리를 포함한다. 이 삼중 변형된 코로나바이러스 스파이크 폴리펩타이드의 서열은, SEQ ID NO: 22로 표시된다(상기 서열의 끝에 굵은 글씨로 VSV로부터의 부분과 함께 표시됨). 일부 구현예에서, 본 발명은, SEQ ID NO: 22에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 아미노산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO: 23에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인,핵산 서열에 의해 암호화된다. 일부 구현예에서, 변형된 폴리펩타이드는, SEQ ID NO: 23에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 23의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 변형된 폴리펩타이드는 SEQ ID NO: 24를 갖는 핵산 서열에 의해 암호화된다.

[0056] 항원의 혼합물은, 광범위한 반응성 면역을 유도할 수 있으므로 코로나바이러스 항원의 조합은 면역 반응의 폭을 향상시키는 데 사용할 수 있다. VLP는 2개(2 가) 또는 3개(3가) 바이러스 항원을 천연 형태로 발현시켜 강력한 B 세포 반응을 유도하는 데 사용할 수 있다. Zika 바이러스 에피토프를 사용한 이전 연구에서는 단일 2가 VLP에서 2개의 항원 조합이 동일한 항원을 발현하는 2개의 1가 VLP보다 훨씬 더 강력한 면역 반응을 생성하는 것으로 나타났다(미국 특허 8920812).

[0057] 따라서, 본 발명의 VLP는, 2개의 야생형 코로나바이러스 스파이크 단백질, 본 명세서에 기술된 2개의 변형된 코로나바이러스 스파이크 단백질, 또는 본 명세서에 기술된 야생형 및 변형된 코로나바이러스 스파이크 단백질의 임의의 조합을 함유하는 2가 VLP를 포함한다. 본 발명의 VLP는 또한, 3개의 야생형 코로나바이러스 스파이크 단백질, 본 명세서에 기술된 3개의 변형된 코로나바이러스 스파이크 단백질, 또는 본 명세서에 기술된 야생형 및 변형된 코로나바이러스 스파이크 단백질의 임의의 조합을 함유하는, 3가 VLP를 포함한다. 2가 또는 3가 VLP 상에서 발현되는 임의의 야생형 또는 변형된 스파이크 단백질 중 하나 이상은, 막횡단 부위 및 세포질 꼬리를 VSV

로부터의 막횡단 부위 및 세포질 꼬리로 대체하도록 추가로 변형될 수 있다.

[0058] 바람직한 구현예에서, 본 발명의 VLP는, SARS-CoV-2로부터의 스파이크 단백질, SARS-CoV로부터의 스파이크 단백질 및 MERS-CoV로부터의 스파이크 단백질을 포함하여 구성되는, 3가 VLP이다. 스파이크 단백질 중 하나 이상은, 막횡단 부위 및 세포질 꼬리를 VSV로부터의 막횡단 부위 및 세포질 꼬리로 대체하도록 변형될 수 있다.

[0059] 일부 구현예에서, 본 발명의 면역원성 조성물은,그 서열이 SEQ ID NO: 4 또는 SEQ ID NO: 4의 코돈 축퇴성 버전인 SARS-CoV-2로부터의 야생형 스파이크 폴리펩타이드; 그 서열이 SEQ ID NO: 7 또는 SEQ ID NO: 7의 코돈 축퇴 버전인 SARS-CoV의 스파이크 폴리펩타이드; 및 그 서열이 EQ ID NO: 10 또는 SEQ ID NO: 10의 코돈 축퇴 버전인 MERS로부터의 스파이크 폴리펩타이드;를 포함하여 구성되는 3가 VLP를 포함하여 구성된다. SARS-CoV-2 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 5로 표시된다. SEQ ID NO: 5의 코돈 최적화 버전은 SEQ ID NO: 6으로 표시된다. SARS-CoV-2 폴리펩타이드를 암호화하는 핵산은, SEQ ID NO: 8로 표시된다. SEQ ID NO: 8의 코돈 최적화된 버전은 SEQ ID NO: 9로 표시된다. MERS 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 11로 표시된다. SEQ ID NO:11의 최적화된 버전은 SEQ ID NO:12로 표시된다. 일부 구현예에서, 본 발명은, 아미노산 서열이 SEQ ID NO: 4에 대하여 적어도 80%, 85%, 90% , 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성이고, SEQ ID NO: 7에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성이며 그리고 SEQ ID NO: 10에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 아미노산 서열을 갖는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, SARS-CoV-2 폴리펩타이드는, SEQ ID NO: 5에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% , 98%, 99% 또는 그 이상 상동성인 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 6에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 5의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 6을 갖는 핵산 서열에 의해 암호화된다. 일부 구현예에서, SAR-CoV 폴리펩타이드는, SEQ ID NO: 8에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 9에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상상동성인,SEQ ID NO: 8의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 9를 갖는 핵산 서열에 의해 암호화된다. 일부 구현예에서, MERS 폴리펩타이드는, SEQ ID NO: 11에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, 핵산 서열의 코돈 최적화 버전에 의해 암호화됨 SEQ ID NO: 11에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상인, SEQ ID NO: 11의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는 SEQ ID NO: 12를 갖는 핵산 서열에 의해 암호화된다.

[0060] 일부 구현예에서, 본 발명의 면역원성 조성물은, 그 서열이 SEQ ID NO: 22이거나 또는 SEQ ID NO: 22의 코돈 축퇴성 버전인 SARS-CoV-2로부터의 변형된 스파이크 폴리펩타이드; 그 서열이 SEQ ID NO: 7이거나 또는 SEQ ID NO: 7의 코돈 축퇴 버전인 SARS-CoV의 스파이크 폴리펩타이드; 및 그 서열이 SEQ ID NO: 10이거나 또는 SEQ ID NO: 10의 코돈 축퇴 버전인 MERS로부터의 스파이크 폴리펩타이드;를 포함하여 구성되는 3가 VLP를 포함하여 구성된다. SARS-CoV-2 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 5로 제시된다. SEQ ID NO: 5의 코돈 최적화된 버전은 SEQ ID NO: 6으로 표시된다. SARS-CoV 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 8로 표시된다. SEQ ID NO: 8의 코돈 최적화된 버전SEQ ID NO: 9로 표시된다. MERS 폴리펩타이드를 암호화하는 핵산은 SEQ ID NO: 11로 표시된다. SEQ ID NO:11의 코돈 최적화된 버전은 SEQ ID NO: 12로 표시된다. 일부 구현예에서, 본 발명은, SEQ ID NO: 22에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성이고, SEQ ID NO: 7에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%와 98%, 99% 또는 그 이상 상동성이며, 그리고 SEQ ID NO: 10에 대하여 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 과 98%, 99% 또는 그 이상 상동성인, 아미노산 서열을 가지는 폴리펩타이드를 포함하여 구성되는 VLP를 포함하여 구성되는 면역원성 조성물을 포함하여 구성된다. 일부 구현예에서, SARS-CoV-2 폴리펩타이드는, SEQ ID NO: 23에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인,핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 24에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 23의 핵산 서열의 코돈 최적화된 버전에 의

해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 24를 갖는 핵산 서열에 의해 암호화된다. 일부 구현예에서, SAR-CoV 폴리펩타이드는, SEQ ID NO: 8에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 9에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 8의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 9를 갖는 핵산 서열에 의해 암호화된다. 일부 구현예에서, MERS 폴리펩타이드는, SEQ ID NO: 11에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, 핵산 서열에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는, SEQ ID NO: 12에 대하여 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 이상 상동성인, SEQ ID NO: 11의 핵산 서열의 코돈 최적화된 버전에 의해 암호화된다. 일부 구현예에서, 폴리펩타이드는 SEQ ID NO: 12를 갖는 핵산 서열에 의해 암호화된다.

[0061] 실시예에서 볼 수 있는 바와 같이, 본 발명의 VLP는, SARS-CoV-2에 대한 강한 면역 반응을 이끌어낼 수 있었다. 특히, 본 명세서에 기술된 각각의 변형된 스파이크 변종는, 강한 면역 반응을 유도하는데 효과적이었다(실시예 5 참조). 본 발명의 3가 VLP(실시예 6 참조)는 SARS-Cov-2, SARS-CoV 및 MERS에 대한 항체 반응을 유도하였다. 더욱이, 본 발명의 3가 VLP에 의한 면역화는, 백신에 포함되지 않은, 관련된 계절성 인간 코로나바이러스, OC43을 인식하는 항체를 유도하였고, 이는 3가 VLP가 코로나바이러스에 대한 면역을 확장시키는 능력이 있음을 입증한다. 예기치 않게, 1가 VLP를 사용한 면역화에 비해, 3가 VLP는 기능적, 중화 활성을 갖는 항체의 유도를 풍부하게 하였다(enrichment). 중화 항체의 이러한 항체 유도 풍부화를 표 8에 나타내었으며, 백신 접종된 마우스로부터 얻은 혈청을 사용하여, 종말점 항체 바인딩 역가에 대한 종말점 중화 항체 역가의 비율을 보여준다.

[0062] 삼중 변형된 SARS-CoV-2 스파이크 단백질을 발현하는 1가 VLP는 골든 햄스터(실시예 7)에 대한 시험 접종(challenge) 연구에 의해 입증된 바와 같이 SARS-CoV-2에 의한 감염에 대해 상당한 보호를 제공한다. 실시예 7에 나타낸 바와 같이, VLP 백신을 접종한 햄스터는 체중으로 나타낸 바와 같이 바이러스 RNA 수준이 유의하게 낮았고 임상 발현이 개선되었다. 더욱이, 면역된 햄스터는 백신을 접종하지 않은 햄스터보다 더 강한 사이토카인 반응을 나타낼 수 있었다.

[0063] 본 발명의 VLP는, SARS-CoV-2의 변종에 대해 효과적인 광범위한 면역 반응을 입증하였다. 실시예 9에 기술된 바와 같이, 삼중 변형 SARS-CoV-2 스파이크 단백질, 천연 MERS 스파이크 단백질 및 천연 SARS-CoV 단백질을 발현하는 3가 VLP 및 삼중 변형 SARS-CoV-2 스파이크 단백질을 발현하는 1가 VLP를, 마우스에서 SARS-CoV-2의 501Y.V2(남아프리카 공화국) 변종에 대한 항체를 유도하는 능력에 대하여 평가하였다. 놀랍게도, 3가 및 1가 컨스트럭트 모두 501Y.V2 변종에 대한 항체를 유도했다. 더욱 놀라운 것은, 항체 역가가 501Y.V2 균주와 SARS-CoV-2의 원래 L 균주에 대해 유사했다는 사실이었다. 따라서, 삼중 변형된 SARS-CoV-2 스파이크 단백질을 발현하는 3가 및 1가 VLP는 모두 다른 COVID-19 백신으로부터 상당한 회피(escape)를 보여준 SARS-CoV-2 변종에 대한 강력한 항체 반응을 유도하는 데 예기치 않게 효과적이었다.

[0064] 본 발명의 VLP는, 또한 항체 반응의 성질에 영향을 미쳤다. 실시예 10에 나타낸 바와 같이, 야생형 SARS-CoV-2 스파이크 단백질을 발현하는 본 발명의 1가 VLP로 면역화된 마우스는 재조합 스파이크 단백질로 면역화된 마우스보다 더 많은 양의 IgG2b를 생성하였다. 더 높은 IgG2b는 TH1 면역 반응과 연관되어 더 높은 수준의 세포 매개 면역을 유발할 수 있다. 따라서, VLP 상의 스파이크 단백질의 제시(presentation)는 세포-매개 면역과 상관된 반응을 가져왔다.

[0065] VLP를 포함하여 구성되는 조성물은 전형적으로 다양한 크기를 갖는 VLP의 혼합물을 포함할 것이라는 것이 이해될 것이다. 아래 나열된 직경 값은 혼합물 내에서 가장 빈번한 직경에 해당하는 것으로 이해되어야 한다. 일부 구현예에서, 조성물 내 소포의 > 90%는 가장 빈번한 값(예를 들어, 1000 ± 500 nm)의 50% 내에 있는 직경을 가질 것이다. 일부 구현예에서, 분포는 더 좁을 수 있고, 예를 들어 조성물 내 소포의 > 90%는 가장 빈번한 값의 40, 30, 20, 10 또는 5% 내에 있는 직경을 가질 수 있다. 일부 구현예에서, 초음파 처리를 사용하여 VLP 형성을 촉진하고 그리고/또는 VLP 크기를 변경할 수 있다. 일부 구현예에서, 여과, 투석 및/또는 원심분리를 사용하여 VLP 크기 분포를 조정할 수 있다.

[0066] 일반적으로, 본 발명의 방법에 따라 생성된 VLP는 임의의 크기일 수 있다. 특정 구현예에서, 조성물은 약 20 nm 내지 약 300 nm 범위의 직경을 갖는 VLP를 포함할 수 있다. 일부 구현예에서, VLP는, 하한이 20, 30, 40, 50, 60, 70, 80, 90, 또는 100 nm으로 제한되고, 상한이 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200,

190, 180, 또는 170 nm으로 제한되는 범위 내의 직경을 갖는 것을 특징으로 한다. 일부 구현예에서, 집단 내의 (within a population) VLP는, 하한이 20, 30, 40, 50, 60, 70, 80, 90, 또는 100 nm로 제한되고, 상한이 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 또는 170 nm으로 제한되는 범위 내의 평균 직경을 나타낸다. 일부 구현예에서, 집단 내의 VLP는, 0.5 미만(예를 들어, 0.45 미만, 0.4 미만, 또는 0.3 미만)인 다분산 지수(polydispersity index)를 갖는다. 일부 구현예에서, VLP 직경은 나노사이징(nanosizing)에 의해 결정된다. 일부 구현예에서, VLP 직경은 전자 현미경에 의해 결정된다.

[0067] 본 발명에 따른 VLP는, 이 분야의 통상의 기술자에게 공지된 일반적인 방법론에 따라 제조될 수 있다. 예를 들어, 핵산 분자, 재구성된 벡터 또는 플라스미드는 당업계에 공지된 서열을 사용하여 제조될 수 있다. 이러한 서열은 저장소(bank)에서 입수할 수 있고, 원료(material)는 다양한 컬렉션, 공개된 플라스미드(published plasmids) 등에서 얻을 수 있다. 이러한 요소는 이 분야의 통상의 기술자에게 잘 알려진 기술을 사용하여 분리되고 조작될 수 있거나 당업계에서 입수할 수 있다. 다양한 합성 또는 인공 서열(artificial sequences)은 또한 컴퓨터 분석 또는 라이브러리의 (고처리량) 스크리닝을 통해 생성될 수 있다. VLP에 대한 폴리펩타이드의 재조합 발현은 하나 이상의 폴리펩타이드(들)를 암호화하는 폴리뉴클레오타이드를 함유하는 발현 벡터의 구축을 필요로 한다. 하나 이상의 폴리펩타이드를 암호화하는 폴리뉴클레오타이드가 얻어지면, 폴리펩타이드의 생산을 위한 벡터는 당업계에 공지된 기술을 사용하는 재조합 DNA 기술에 의해 생성될 수 있다. 본 발명에 따라 이용될 수 있는 발현 벡터는 특히, 포유동물 및 조류 발현 벡터, 바큘로바이러스 발현 벡터, 식물 발현 벡터(예를 들어, 콜리플라워 모자이크 바이러스(CaMV), 담배 모자이크 바이러스(TMV)), 플라스미드 발현 벡터(예: Ti 플라스미드)를 포함하되 그에 한정되지 않는다..

[0068] 본 발명의 VLP는, 임의의 이용가능한 단백질 발현 시스템에서 생산될 수 있다. 전형적으로, 발현 벡터는 통상적인 기술에 의해 숙주 세포로 전달되고, 이어서 형질감염된 세포는 VLP를 생성하기 위해 통상적인 기술에 의해 배양된다. 일부 구현예에서, VLP는 세포의 일시적 형질감염을 사용하여 생성된다. 일부 구현예에서, VLP는 안정적으로 형질감염된 세포를 사용하여 생산된다. VLP 생산에 사용될 수 있는 전형적인 세포주는, 인간 배아 신장(HEK) 293, WI 38, 차이니즈 햄스터 난소(CHO), 원숭이 신장(COS), HT1080, C10, HeLa, 아기 햄스터 신장(BHK), 3T3, C127, CV-1, HaK, NS/0 및 L-929 세포과 같은 포유동물 세포주를 포함하지만 이에 국한되지 않다. 특정 비제한적 예는, BALB/c 마우스 골수종 계통(line)(NSO/1, ECACC 번호: 85110503); 인간 망막아세포(PER.C6(CruCell, Leiden, The Netherlands)); SV40에 의해 형질전환된 원숭이 신장 CV1 계통(COS-7, ATCC CRL 1651); 인간 배아 신장 계통(현탁 배양에서 성장을 위해 서브클로닝된 293 세포, Graham et al., J. Gen Virol., 36:59 (1977)); 아기 햄스터 신장 세포(BHK, ATCC CCL 10); 차이니즈 햄스터 난소 세포 +/-DHFR(CHO, Urlaub and Chasin, Pr℃. Natl. Acad. Sci. USA, 77:4216 (1980)); 마우스 세르톨리 세포(TM4, Mather, Biol. Reprod., 23:243-251(1980)); 원숭이 신장 세포(CV1 ATCC CCL 70); 아프리카 녹색 원숭이 신장 세포(VERO-76, ATCC CRL-1 587); 인간 자궁경부암 세포(HeLa, ATCC CCL 2); 개 신장 세포(MDCK, ATCC CCL 34); 버팔로 래트 간 세포(BRL 3A, ATCC CRL 1442); 인간 폐 세포(W138, ATCC CCL 75); 인간 간 세포(Hep G2, HB 8065); 마우스 유선 종양(MMT 060562, ATCC CCL51); TRI 세포(Mather et al., Annals N.Y. Acad. Sci., 383:44-68 (1982)); MRC 5 세포; FS4 세포; 및 인간 간종 계통(Hep G2)을 포함하지만, 그에 한정되지 않는다. 일부 구현예에서, VLP 생산을 위해 사용될 수 있는 세포주(cell lines)는, 곤충(예를 들어, Sf-9, Sf-21, Tn-368, Hi5) 또는 식물(예를 들어, 레구미노사, 곡물 또는 담배) 세포를 포함한다. 일부 구현예에서, 특히 글리코실화가 단백질 기능에 중요한 경우, 포유동물 세포가 단백질 발현 및/또는 VLP 생산에 바람직하다는 것이 이해될 것이다(예를 들어, Roldao A et al., 2010 Ext Rev Vaccines 9:1149-76 참조) .

[0069] 삽입된 서열(inserted sequences)의 발현을 조절하거나 특정 방식으로 유전자 생성물을 변형 및 처리하는 세포 균주가 선택될 수 있음을 이해할 것이다. 단백질 제품의 이러한 변형(예: 글리코실화) 및 가공(예: 절단 또는 막으로의 이송)은 VLP 생성 또는 VLP 폴리펩타이드 또는 추가 폴리펩타이드(예: 애주번트 또는 추가 항원)의 기능에 중요할 수 있다. 상이한 세포는, 단백질 및 유전자 산물의 번역 후 처리(post-translational pr℃essing) 및 변형을 위한 특징적이고 특정한 메커니즘을 가지고 있다. 발현된 외래 단백질의 정확한 변형 및 처리를 보장하기 위해 적절한 세포주 또는 숙주 시스템을 선택할 수 있다. 일반적으로, 1차 전사체의 적절한 처리, 유전자 산물의 글리코실화 및 인산화를 위한 적절한 세포 기구(cellular machinery)를 보유하는 진핵 숙주 세포(패키징 세포(예: 293T 인간 배아 신장 세포)라고도 함)가 본 발명에 따라 사용될 수 있다.

[0070] VLP는, 여러 가지 중에서도, 원심분리, 구배법(gradients), 자당밀도-구배 초원심분리, 접선 유동 여과(tangential flow filtration) 및 크로마토그래피(예를 들어, 이온 교환(음이온 및 양이온), 아피니티(affinity) 및 사이징 컬럼 크로마토그래피), 또는 차등 용해도(differential solubility)와 같은 공지된 기술

에 따라 정제될 수 있다. 대안적으로, 또는 추가적으로, 세포 상청액(cell supernatant)을 정제 단계 없이 직접 사용할 수 있다. 추가 항원 또는 애주번트와 같은 추가 엔티티(entities)가 정제된 VLP에 추가될 수 있다.

[0071] 본 발명에 따르면, 세포는 단일 발현 벡터로 형질감염될 수 있다. 일부 구현예에서, 단일 발현 벡터는 VLP의 하나 이상의 요소(예를 들어, 구조적 다단백질, 코로나바이러스 스파이크 단백질 등 중 하나 이상)를 암호화한다. 예를 들어, 일부 구현예에서, 단일 발현 벡터는 VLP의 2개 이상의 요소를 암호화한다. 일부 구현예에서, 단일 발현 벡터는 VLP의 3개 이상의 요소를 암호화한다. 본 발명의 한 구현예에서, 단일 발현 벡터는 Gag 폴리펩타이드 및 코로나바이러스 스파이크 당단백질을 암호화한다.

[0072] 일부 구현예에서, 세포는 2개 이상의 발현 벡터로 형질감염된다. 예를 들어, 일부 구현예에서, 세포는 Gag 폴리펩타이드를 암호화하는 제1 벡터 및 코로나바이러스 스파이크 당단백질을 암호화하는 제2 벡터로 형질감염되고, 코로나바이러스 스파이크 당단백질을 포함하는 "1가" VLP가 생성된다. 일부 구현예에서, 세포는 Gag 폴리펩타이드를 암호화하는 제1 벡터, 코로나바이러스 스파이크 당단백질을 암호화하는 제2 벡터 및 또 다른 코로나바이러스 스파이크 당단백질을 암호화하는 제3 벡터로 형질감염된다. 이러한 구현예에서, 2개의 코로나바이러스 스파이크 당단백질을 포함하는 "2가" VLP가 생성된다. 일부 구현예에서, 세포는 Gag 폴리펩타이드를 암호화하는 제1 벡터, 코로나바이러스 스파이크 당단백질을 암호화하는 제2 벡터, 및 2개의 코로나바이러스 스파이크 당단백질을 암호화하는 제3 벡터로 형질감염된다. 이러한 구현예에서, 3개의 코로나바이러스 스파이크 당단백질을 포함하는 "3가" VLP가 생성된다.

[0073] 실시예에서 추가로 나타내는 바와 같이, SARS-CoV-2 스파이크 단백질의 변형은 VLP의 수율에 상당한 영향을 미쳤다. 표 1을 참조하면, 실시예 3에서 삼중 변형된 SARS-CoV-2 스파이크 단백질을 발현하는 VLP(그룹 3)는 SARS-CoV-2 스파이크 단백질을 발현하는 다른 1가 VLP보다 유의하게 더 높은 스파이크 단백질 수율을 나타냈다. 따라서, VLP의 이 구현예는 더 많은 양으로 제조될 수 있으며, 이는 대유행 상황에서 수요를 해결하는 데 중요한다.

[0074] 일부 구현예에서, 1가, 2가 또는 3가 VLP가 혼합된다. 예를 들어, 일부 구현예에서, 1가 및 2가 VLP가 혼합되어 3가 VLP 혼합물을 형성한다. 일부 구현예에서 2개의 1가 VLP가 혼합되어 2가 VLP 혼합물을 형성한다.

[0075] 본 발명은 본 명세서에 기술된 VLP를 포함하여 구성되고, 임의선택적으로, 본 명세서에 기술된 당단백질 및 당단백질 변종를 추가로 포함하여 구성되는 제약 조성물을 제공한다. 일부 구현예에서, 본 발명은 VLP 및 적어도 하나의 제약상 허용되는 부형제, 애주번트 및/또는 담체를 제공한다. 이러한 약제학적 조성물은 임의선택적으로 하나 이상의 추가 치료 활성 물질을 포함하여 구성되고 그리고 /또는 조합하여 투여될 수 있다. 제공된 약제학적 조성물은 SARS, MERS 또는 COVID-19 감염 또는 SARS, MERS 또는 COVID-19 감염과 관련된 또는 상관된 네거티브 영향들(negative ramifications)의 방지에, 예방제(prophylactic agents)(즉, 백신)로서 유용하다. 제공된 약제학적 조성물은 또한 SARS-CoV-2의 특정 변종에 대한 예방제로서 유용하다. 일부 구현예에서, 제약 조성물은 인간에게 투여하기 위해 제형화된다.

[0076] 본 명세서에서 제공되는 약제학적 조성물은 멸균 주사가능한 형태(예를 들어, 피하 주사 또는 정맥내 주입에 적합한 형태)로 제공될 수 있다. 예를 들어, 일부 구현예에서, 제약 조성물은 주사에 적합한 액체 투여 형태로 제공된다. 일부 구현예에서, 제약 조성물은, 분말(예를 들어, 동결건조 및/또는 멸균)로서 제공되고, 임의선택적으로 진공 하에, 주사 전에 수성 희석제(예를 들어, 물, 완충액, 염 용액 등)로 재구성된다.(reconstituted). 일부 구현예에서, 제약 조성물은 물, 염화나트륨 용액, 아세트산나트륨 용액, 벤질 알코올 용액, 인산염 완충 식염수 등에 희석 및/또는 재구성된다. 일부 구현예에서, 분말은 수성 희석제와 부드럽게 혼합되어야 한다(예를 들어, 쉐이킹 식).

[0077] 일부 구현예에서, 제공된 제약 조성물은 하나 이상의 제약상 허용되는 부형제(예를 들어, 보존제, 불활성 희석제, 분산제, 계면 활성제 및/또는 유화제, 완충제 등)를 포함한다. 적합한 부형제는 예를 들어 물, 식염수, 덱스트로스, 수크로스, 트레할로스, 글리세롤, 에탄올 또는 그 유사물, 및 이들의 조합을 포함한다. Remington's The Science and Practice of Pharmacy, 21st Edition, A. R. Gennaro, (Lippincott, Williams & Wilkins, Baltimore, MD, 2006)은 제약 조성물을 제형화하는데 사용되는 다양한 부형제 및 그의 제조를 위한 공지된 기술을 개시한다. 임의의 통상적인 부형제 매질이 바람직하지 않은 생물학적 효과를 생성하거나 제약 조성물의 다른 성분(들)과 유해한 방식으로 상호작용하는 것과 같이 물질 또는 그의 유도체와 양립할 수 없는 경우를 제외하고, 그의 사용은 본 발명의 범위에 속하는 것으로 고려된다. 일부 구현예에서, 제약 조성물은 보존제를 포함하여 구성되지 않는다.

[0078] 일부 구현예에서, 제약 조성물은 백신으로서 충분히 면역원성이다(예를 들어, 애주번트의 부재 하에). 일부 구현예에서, 약학 조성물의 면역원성은 애주번트를 포함함으로써 향상된다. 임의의 애주번트가 본 발명에 따라 사용될 수 있다. 많은 수의 애주번트가 알려져 있다. 그러한 많은 화합물의 유용한 개요는 미국국립 보건원에서 준비했으며 찾을 수 있다(www.niaid.nih.gov/daids/vaccine/pdf/compendium.pdf). 또한 Allison, 1998, Dev. Biol. Stand., 92:3-11, Unkeless et al., 1998, Annu. Immunol., 6:251-281 및 Phillips et al., 1992, Vaccine, 10:151-158이 참고가 된다. 수백 가지의 상이한 애주번트가 당업계에 공지되어 있고 본 발명의 실시에 사용될 수 있다. 본 발명에 따라 사용될 수 있는 예시적인 애주번트는 사이토카인, 겔형 애주번트(예: 수산화알루미늄, 인산 알루미늄, 인산 칼슘 등), 미생물 애주번트(예: 면역조절 DNA 서열 CpG 모티프, 모노포스포릴 지질 A와 같은 내독소; 콜레라 독소, 대장균 열 불안정 독소 및 백일해 독소와 같은 외독소; 무라밀 디펩티드 등), 오일-에멀젼 및 유화제 기반 애주번트(예: Freund's Adjuvant, MF59) [Novartis], SAF 등), 미립자 애주번트(예: 리포솜, 생분해성 미소구체, 사포닌 등), 합성 애주번트(예: 비이온성 블록 코폴리머, 무라밀 펩티드 유사체, 폴리포스파젠, 합성 폴리뉴클레오티드 등) 및/또는 이들의 조합을 포함하되, 그에 한정되지 않는다. 다른 예시적인 애주번트는 일부 폴리머(예를 들어, 미국 특허 5,500,161호에 기술된 폴리포스파젠; Q57, QS21, 스쿠알렌, 테트라클로로데카옥사이드 등)를 포함한다.

[0079] 일부 구현예에서, 제약 조성물은 냉장 및/또는 냉동될 수 있는 형태로 제공된다. 일부 구현예에서, 제약 조성물은 냉장 및/또는 냉동될 수 없는 형태로 제공된다. 일부 구현예에서, 재구성된 용액 및/또는 액체 투여 형태는 재구성 후 특정 기간(예를 들어, 2시간, 12시간, 24시간, 2일, 5일, 7일, 10일, 2주, 한 달, 두 달 또는 그 이상) 저장될 수 있다. 일부 구현예에서, 지정된 시간보다 더 긴 시간 동안 VLP 제제를 저장하면 VLP 분해가 발생한다.

[0080] 본 발명에 따른 약제학적 조성물은 단일 단위 투여량(unit dose) 및/또는 복수의 단일 단위 투여량으로 제조, 포장 및/또는 판매될 수 있다. 본 명세서에 사용된 "단위 투여량"은 미리 결정된 양의 활성 성분을 포함하여 구성되는 약제학적 조성물의 개별적 양(discrete amount)이다. 활성 성분의 양은 일반적으로 대상체에게 투여되는 용량 및/또는 이러한 용량의 편리한 분획, 예를 들어 이러한 용량의 1/2 또는 1/3과 동일하다.

[0081] 본 발명에 의한 약제학적 조성물의 활성 성분, 약제학적으로 허용되는 부형제, 및/또는 임의의 추가 성분의 상대적인 양은 대상체의 정체(identity), 크기 및/또는 상태에 따라 그리고/또는 조성물이 투여되는 경로에 따라 달라질 수 있다. 예로서, 조성물은 0.1% 내지 100%(w/w) 활성 성분을 포함하여 구성될 수 있다.

[0082] 본 발명의 제공된 조성물 및 방법은 인간 성인 및 어린이를 포함하는 대상체에서 SARS, MERS 또는 COVID-19 감염의 예방 및/또는 치료에 유용하다. 그러나 일반적으로 모든 동물과 함께 사용할 수 있다. 원하는 경우, 본 발명의 방법은 또한 양, 조류, 소, 돼지 및 말 품종과 같은 농장 동물에 사용될 수 있다. 본 발명의 목적을 위해, 백신 접종은 질병 유발 인자에 대한 노출 전, 노출 동안 및/또는 노출 후에, 그리고 특정 구현예에서, 병원체(agent)에 노출되기 전, 노출 동안 및/또는 직후에 투여될 수 있다. 일부 구현예에서, 백신 접종은, 적절한 시간 간격으로, 백신 접종 조성물의 다중 투여를 포함한다.

[0083] 본 명세서에 기술된 조성물은 일반적으로 면역 반응을 유도하는 데 필요하거나 충분한 시간 및 양으로 투여될 것이다. 투여 요법은 단일 투여량 또는 일정 기간에 걸친 복수회 투여량으로 구성될 수 있다. 투여될 면역원성 조성물(예를 들어, VLP)의 정확한 양은 대상체마다 다를 수 있고 여러 요인에 따라 달라질 수 있다. 따라서, 일반적으로 사용되는 정확한 용량은 대상체의 체중 및 투여 경로뿐만 아니라 대상체의 연령에 따라 달라질 것이라는 것을 이해할 것이다. 특정 구현예에서 VLP 조성물의 특정량은 단일 투여량으로 투여된다. 특정 구현예에서, VLP 조성물의 특정량은 1회 초과 투여량(예를 들어, 1-12개월으로 나누어 1-3회 투여량)으로 투여된다.

[0084] 일부 구현예에서, 제공된 조성물은 초기 투여량 및 적어도 1회의 추가 투여량(booster dose)으로 투여된다. 일부 구현예에서, 제공된 조성물은 1차 투여량 및 2, 3 또는 4개의 추가 투여량으로 투여된다. 일부 구현예에서, 제공된 조성물은 1차 투여(initial dose) 및 1차 투여 후 약 1개월, 약 2개월, 약 3개월, 약 4개월, 약 5개월, 또는 약 6개월 후에 적어도 1회의 추가 투여량으로 투여된다. 일부 구현예에서, 제공된 조성물은 1차 투여 후 약 6개월, 약 7개월, 약 8개월, 약 9개월, 약 10개월, 약 11개월, 또는 약 1년 후에 제2 추가 추가 투여량으로 투여된다. 일부 구현예에서, 제공된 조성물은 1년, 2년, 3년, 4년, 5년, 6년, 7년, 8년, 9년 또는 10년마다 추가 투여량으로 투여된다.

[0085] 특정 구현예에서, 제공된 조성물은 예를 들어 주사에 의한 비경구 전달을 위해 제형화될 수 있다. 이러한 구현예에서, 투여는 예를 들어 정맥내, 근육내, 피내 또는 피하, 또는 주입 또는 무바늘 주사 기술을 통한 것일 수 있다. 특정 구현예에서, 조성물은 경구 전달(peroral delivery), 경구 전달(oral delivery), 비강내 전달, 협

측 전달, 설하 전달, 경피 전달, 경피적 전달, 복강내 전달, 질내 전달, 직장 전달 또는 두개내(intracranial) 전달을 위해 제형화될 수 있다.

[0086] 일부 구현예에서, 대상체에 투여시 제공된 VLP는 대상체에서 체액성 면역 반응을 유도한다. 일부 구현예에서, 대상체의 체액성 면역 반응은 약 1개월 이상, 약 2개월 이상, 약 3개월 이상, 약 4개월 이상, 약 5개월 이상, 약 6개월 이상, 적어도 약 7개월, 적어도 약 8개월, 적어도 약 9개월, 적어도 약 10개월, 적어도 약 11개월, 적어도 약 12개월, 적어도 약 13개월, 적어도 약 14개월, 적어도 약 15개월 이상, 적어도 약 16개월, 적어도 약 17개월, 적어도약 18개월, 적어도 약 19개월, 적어도 약 20개월, 적어도 약 21개월, 적어도 약 22개월, 적어도 약 23개월, 적어도 약 24개월, 적어도 약 28개월, 적어도 약 32개월, 적어도 약 36개월, 적어도 약 40개월, 적어도 약 44개월, 적어도약 48개월, 또는 그 이상. 유지된다.

[0087] 일부 구현예에서, 대상체에 투여시 제공된 VLP는 대상체에서 세포 면역 반응을 유도한다. 일부 구현예에서, 대상체의 세포 면역 반응은 약 1개월 이상, 약 2개월 이상, 약 3개월 이상, 약 4개월 이상, 약 5개월 이상, 약 6개월 이상, 약 7개월 이상, 약 8개월 이상, 약 9개월 이상, 약 10개월 이상, 약 11개월 이상, 또는 12개월 이상 유지된다.

[0088] 여기에 언급된 모든 간행물, 특허 출원, 특허 및 기타 참고 문헌은 그 전체가 참고 문헌으로 통합된다. 또한, 재료, 방법 및 예는 예시일 뿐이며 제한하려는 의도가 아니다. 달리 정의되지 않는 한, 여기에서 사용된 모든 기술 및 과학 용어는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 본 명세서에 기술된 것과 유사하거나 동등한 방법 및 재료가 본 발명의 실시 또는 시험에 사용될 수 있지만, 적합한 방법 및 재료가 본 명세서에 기술되어 있다.

[0089] **실시예**

[0090] 하기 실시예는 본 명세서에 기술된 특정 조성물을 제조하고 실시하는 일부 예시적인 모드를 설명한다. 이들 실시예는 단지 예시를 위한 것이며 본 명세서에 기술된 조성물 및 방법의 범위를 제한하려는 것이 아님을 이해해야 한다.

[0091] **실시예 1: DNA 발현 플라스미드의 구축**

[0092] 이 실시예는 재조합 코로나바이러스 스파이크 유전자 서열의 발현을 위한 발현 플라스미드 및 컨스트럭트의 개발을 설명한다. 표준 발현 플라스미드는 일반적으로, 포유류 기원의 프로모터 서열, 인트론 서열(intron sequence), 폴리아데닐화 신호 서열(PolyA), pUC 복제 기원 서열(origin of replication sequence)[pUC - pBR322는 colE1 기원/복제 개시 부위(origin/site of replication initiation)이며, 대장균(E. Coli)(DH5α)과 같은 박테리아의 플라스미드 복제에 사용됨] 및 플라스미드 플라크 선택을 위한 선택 마커로서 항생물질 내성 유전자(antibiotic resistance gene)로 구성된다. 인트론 다음의 플라스미드 내에는 관심 유전자 또는 부분 유전자 서열을 접합하는 데 사용할 수 있는 다양한 제한 효소 부위(restriction enzyme sites)가 있다.

[0093] Propol II 발현 플라스미드는, pHCMV[HCMV에 대한 초기 프로모터(early promoter)], 베타-글로빈 인트론(BGL Intron), 토끼 글로빈 폴리아덴실화 신호 서열(PolyA), pUC 복제 기원 서열[pUC - pBR322는 [pUC - pBR322는 colE1 기원/복제 개시 부위이며, 대장균(DH5α)과 같은 박테리아의 플라스미드 복제에 사용됨], 암피실린 내성 유전자 β-락타마제[Amp R - 플라스미드에 대한 선택 마커는 암피실린(100㎍/ml)에 대한 내성을 부여함]를 함유한다.

[0094] Gag MMLV 발현 컨스트럭트("MLV-Gag")을 개발하기 위해, MMLV(그 C 말단 Pol 서열이 없는 Gag)의 Gag 폴리단백질을 암호화하는 상보적 DNA(cDNA) 서열(SEQ ID NO: 3)을 Propol II 발현 벡터에서 클로닝하였다. 모든 코로나바이러스 발현 컨스트럭트를 개발하기 위해, 각각 다음:

[0095] i) SARS-CoV-2(SEQ ID NO:: 6);

[0096] ii) SARS-CoV(SEQ ID NO:: 9);

[0097] iii) 메르스(SEQ ID NO: 12);

[0098] iv) SARS-CoV-2 프롤린 변형 스파이크 당단백질(SEQ ID NO: 15);

[0099] v) SARS-CoV-2 퓨린 절단 변형(SEQ ID NO: 18);

[0100] vi) SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 21);

[0101] vii) VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO:: 24); 그리고

[0102] viii) VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 변형(SEQ ID NO: 26);의 서열이, Propol II 발현 벡터에 클로닝되었다. SARS-CoV-2 서열은 원래 중국 우한에서 확인된 바이러스의 L 균주에서 유래했다.

[0103] DNA 플라스미드를 적격 대장균(DH5α)에서 증폭하고,그리고 표준 프로토콜에 따라 내독소가 없는 제조 키트로 정제하였다.

[0104] **실시예 2: 바이러스 유사 입자의 생산**

[0105] 이 실시예는 실시예 1에 기술된 다양한 재조합 코로나바이러스 스파이크 항원을 함유하는 바이러스-유사 입자의 생산 방법을 설명한다.

[0106] HEK 293 세포로부터 유래된 293SF-3F6 세포주는, 무혈청 화학적으로 정의된 배지(CA 2,252,972 및 US 6,210,922)에서 성장된 독점적인(proprietary) 현탁 세포 배양물이다. HEK 293SF-3F6 세포를 37℃, 5% CO_2에서 80 rpm의 속도로 진동 플라스크(shaker flask)에서 스케일-업(scaled up)한 다음, 표적 세포 밀도를 얻기 위해 L-글루타민이 보충된 HyQSF4 Transfx293 배지(GE Bioscience)를 사용하여 바이오리액터에 시딩했다(seeded). 0.9 ~ 120만 세포/ml 및 높은 생존율(>90%). 세포를 코로나바이러스 외피 폴리펩타이드를 암호화하는 플라스미드, Gag를 암호화하는 플라스미드 및 형질감염 제제로서 고품질 폴리에틸렌이민(PEIpro™)을 사용하여 약 100만 세포/ml의 세포 밀도로 공동 형질감염시켰다. DNA 플라스미드 및 형질감염제는 OptiPRO SFM 배지(GE Biosciences)에서 제조되었다. 바이오리액터를 매일(트랜스펙션 후 ~24시간 및 48시간) 모니터링하고 세포 밀도, 생존율 및 세포 직경을 기록했다. 생산 브로쓰(production broth)를 형질감염 후 48시간에 거두어들였다(harvested).

[0107] 전체 단백질은 브래드포드 분석 정량 키트(BioRad)에 의해 분취량에서 결정되었다(determined on an aliquot). Bradford 단백질 분석은 Coomassie Brilliant Blue G-250의 산성 용액에 대한 최대 흡광도가, 단백질에 결합할 때, 465 nm에서 595 nm로 이동한다는 관찰을 기반으로 한다. 소수성 및 이온성 상호작용 모두 염료의 음이온성 형태를 안정화시켜, 가시적인 색상 변화를 일으킨다. 분광 광도계를 사용하여, 595 nm에서, 샘플 및 Bradford Protein Reagent 염료의 흡광도를 측정했다.

[0108] **실시예 3: 1가 백신 후보의 생산**

[0109] 실시예 2에 기술된 방법을 사용하여 4개의 상이한 1가 바이러스 유사 입자를 생성하였다. 바이러스 유사 입자를 하기 4개의 SARS-CoV-2 염기 서열 중 하나로 형질감염시켰다:

[0110] 1. 천연 형태의 SARS-CoV-2(SEQ ID NO: 6);

[0111] 2. SARS-CoV-2 프롤린 및 퓨린 절단 변형(SARS-CoV-2 Proline and Furin Cleavage Modified)(SEQ ID NO: 21);

[0112] 3. VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24); 또는

[0113] 4. VSV의 TM/Cyt를 갖는 SARS-CoV-2 변형(SEQ ID NO: 26).

[0114] 생성된 생성물의 총 항원 함량을 측정하여 그 결과를 표 1에 나타내었다.

[0115] 표 1 : 1가 SARS-CoV-2 바이러스 유사 입자 수율

표 1

[0116]

그룹	1가 VLP 설명	Gag (μg)	SARS-CoV-2 스파이크 단백질(μg)
1	천연 SARS-Cov-2 (SEQ ID #6)	23,157	16.2
2	SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 21)	32,220	229.9
3	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 24)	22,916	639.2
4	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 변형(SEQ ID NO: 26)	19,332	49.95

[0117] 표 1의 데이터에서 알 수 있는 바와 같이, 그룹 3, 세포질 및 막횡단 세그먼트를 VSV 형태의 상응하는 세그먼트로 대체함으로써 변형된 SARS-CoV-2 서열을 사용하여 상당히 더 높은 수율을 얻을 수 있었다.

[0118] **실시예 4: 3가 백신 후보의 생산**

[0119] 실시예 2에 기술된 방법을 사용하여 4개의 상이한 3가 바이러스 유사 입자를 생산하였다. 각각의 입자는, Gag 암호화 플라스미드(plasmids encoding Gag), MERS로부터의 항원 서열(SEQ ID NO: 12), SARS-CoV로부터의 항원 서열(서열 ID NO: 9) 및 다음 두 가지 SARS-CoV-2 서열 중 하나로 형질감염시켰다:

[0120] 1. SARS-CoV-2 외피 폴리펩타이드의 천연 형태(SEQ ID NO: 6); 또는

[0121] 2. VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24).

[0122] 생성된 생성물의 항원 함량을 측정하여 그 결과를 표 2에 나타내었다.

[0123] 표 2 : 3가 코로나바이러스 바이러스 유사 입자 수율

표 2

[0124]

그룹	3가 VLP - SARS-CoV-2 스파이크 단백질	Gag (μg)	SARS-CoV-2 스파이크 단백질 (μg)
1	천연 SARS-Cov-2 (SEQ ID #6)	20,358	28.8
2	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24)	17,382	404.4

[0125] 표 2의 데이터에서 알 수 있는 바와 같이, 그룹 2, VSV G 단백질로부터의 TM/Cyt로 추가 변형된 스파이크 단백질의 안정화된 예비융합 형태를 갖는 SARS-CoV-2 서열(SEQ ID NO: 24),을 사용하여 상당히 더 높은 3가 VLP 수율을 얻었다.

[0126] **실시예 5: 1가 SARS-CoV-2 VLP 백신 컨스트럭트의 효능(Potency) 평가**

[0127] 투약 미경험(naive) 6-8 주령 C57/BL6 마우스(n=10)를 아래 표 3에 나타낸 SARS-CoV-2 VLP 백신 제형의 대략 1/20 내지 1/50 인간 투여량(human dose)으로 2회 면역화시켰다. 면역화는 0일 및 21일에 이루어졌다. 그 동물을 면역화 14일 후에 희생시키고, 항-스파이크 단백질 항체 역가 및 중화 항체의 후속 분석을 위해 혈청을 수집하였다.

[0128] SARS-CoV-2 VLP는 표 3에 나타낸 바와 같이 인산 알루미늄 애주번트(Adjuphos®)와 함께 제형화되었다.

[0129] 표 3 : 1가 SARS-CoV-2 VLP 백신 제형

표 3

[0130]

그룹	1가 SARS-CoV-2	μg SARS-CoV-2 스파이크/dose	μg Gag /dose	Al+++ μg/dose	투여량/동물 (μℓ)
1	SARS-CoV-2 천연 스파이크 단백질 VLP	0.07	100.1	125	500
2	SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 21)	0.14	19.6	125	500
3		0.7	98.1	125	500
4	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 24)	0.14	5.0	125	500
5		0.7	25.05	125	500
6	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 변형 (SEQ ID NO: 26)	0.07	27.1	125	500
7		0.14	54.2	125	500

[0131] 항-스파이크 SARS-CoV-2 항체 역가는 다음과 같이 측정되었다: 96 웰 플레이트를 SARS-COV-2 스파이크 단백질 (S1 + S2)(Sinobiological, Cat# 40589-V08B1)(0.1 DPBS의 μg/ml)로 코팅하였다. 다음 날, 그 플레이트를 ELISA 세척 완충액 중 5% 우유로 37℃에서 1시간 동안 차단했다. 그 플레이트를 세척 완충액으로 세척한 다음, 1:10,000 내지 1:1,200,000으로 시작하여 개별 마우스 혈청의 2배 희석액을 첨가했다. 그 플레이트를 37℃에서 1.5시간 동안 인큐베이션한 후, 플레이트 세척 및 ELISA 세척 완충액 중 1% 우유에 1:5,000으로 희석된 2차 항체: 염소 항-마우스(Secondary Antibody: Secondary Antibody: Goat anti-Mouse) IgG1(Bethyl, Cat# A90-

131P)을 첨가하였다. 그 플레이트를 37℃에서 1시간 동안 인큐베이션하였다. 그 플레이트에 TMB 1성분 마이크로웰 기질을 첨가하고, 실온에서 10분 동안 인큐베이션한 다음, 정지액(Stop solution)을 첨가하였다. MAXline 플레이트 판독기를 사용하여 450 nm에서 흡광도를 판독했다. 결과를 아래 표 4에 나타내었다.

[0132] 표 4에 보고된 항-스파이크 총 IgG 바인딩 역가는 재조합 SARS-CoV-2 스파이크 단백질에 대한 ELISA 측정에 의해 여전히 0.1 이상의 광학 밀도를 갖는 혈청의 가장 높은 희석도를 나타낸다. 예상외로, VSV-G 단백질(그룹 4)의 TM/Cyt로 추가 변형된 SARS-CoV-2 스파이크 단백질의 안정화된 예비융합 형태를 발현하는 VLP의 단일 투여량으로 마우스의 면역화는, SARS-CoV-2 스파이크 단백질의 유사한 용량을 발현하지만 다른 프레젠테이션(presentations)(그룹 2, 7)으로 발현하는 VLP로 마우스를 면역화하는 것 보다, 극적으로 더 강력한 항체 반응을 유도했다.

[0133] 각 백신 접종 14일 후 마우스의 항체 역가는 표 4에 나와 있다. P1 및 P2는 1차 및 2차 백신 접종을 나타낸다. 결과는 개별 동물 중에서 풀링되었다(pooled).

[0134] 표 4 : 1가 SARS-CoV-2 VLP 백신 항체 역가

표 4

[0135]

그룹	백신 접종	항-SARS-CoV-2 스파이크 전체 IgG 바인딩 역가
풀링된 그룹(Pooled Group) 1	P1Vd14	9,099
	P2Vd14	310,103
풀링된 그룹 2	P1Vd14	74,612
	P2Vd14	424,883
풀링된 그룹 3	P1Vd14	262,689
	P2Vd14	321,427
풀링된 그룹 4	P1Vd14	341,493
	P2Vd14	670,735
풀링된 그룹 5	P1Vd14	619,766
	P2Vd14	359,528
풀링된 그룹 6	P1Vd14	1,108
	P2Vd14	302,500
풀링된 그룹 7	P1Vd14	4,093
	P2Vd14	221,990

[0136] 표 4에 나타난 바와 같이, 각각의 1가 VLP 백신 제형은 마우스에서 강력한 항체 반응을 유도하였다. 거의 모든 제형에서, 반응은 두 번째 백신 접종에 의해 강력하게 향상되었다. 하나의 그룹, VSV의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24)을 기반으로 하는 백신 제형으로 구성된 그룹 5,는 두 번째 백신 접종 후 감소된 반응을 보였다. 그러나 1차 접종 후 반응이 매우 높아 2차 접종 시 마우스의 면역 반응이 소진되었을 가능성이 높아졌다. 이 반응은 인간과 같은 더 큰 포유류에서 볼 수 없을 수도 있다.

[0137] 중화 항체를 다음과 같이 시험하였다. SARS-CoV-2 바이러스의 캐나다배양균(Canadian isolate)의 100 플라크 형성 단위(plaque forming units, pfu)로 구성되는 일정량의 바이러스를, 테스트할 마우스 혈청 표본의 2 배 희석액(2-fold dilutions)과 혼합했으며, 희석 범위는 40~5120 배였고, 뒤이어 개별 바이러스에 적합한 세포주의 세포에 그 혼합물을 플레이팅하였다. 그 플라크 형성 단위의 농도는 며칠 후에 형성된 플라크의 수에 의해 결정된다. 그런 다음 플라크를 시각화하기 위해 필수 염료(vital dye)(예: 크리스탈 바이올렛 또는 뉴트럴 레드)를 추가하고 테스트 혈청이 포함된 개별 플레이트의 플라크 수를 네거티브 대조군 혈청에 존재하는 플라크 수로 나누어 중화 백분율을 계산했다. 플라크 형성 단위는 현미경 관찰 또는 감염된 세포와 반응하는 특정 염료의 관찰에 의해 측정되었다. 해석은 일반적으로 총 플라크(비리온)의 50%를 억제할 수 있는 혈청의 마지막 희석인 50% 중화를 기반으로 한다. 80 및 90의 플라크 감소 중화 테스트(plaque reduction neutralization test, PRNT) 임계값은 각각 플라크를 80% 또는 90%까지 감소시킬 수 있는 혈청 희석을 나타sos다. 결과는 표 5에 나와 있다.

[0138] 표 5 : 1가 SARS-CoV-2 VLP 백신 중화 항체

표 5

[0139]

시험 그룹	PRNT50	PRNT80	PRNT90
POV (백신 접종 전)	**		
P1VD14 그룹 1	**		
P2VD14 그룹1	2560	1280	640
P1VD14 그룹 2	320	160	80
P2VD14 그룹 2	***	5120	2560
P1VD14 그룹 3	320	285	160
P2VD14 그룹 3	5120	2560	1280
P1VD14 그룹 4	640	320	160
P2VD14 그룹 4	***	5120	2560
P1VD14 그룹 5	2560	640	320
P2VD14 그룹 5	5120	2276	1280
P1VD14 그룹 6	**		
P2VD14 그룹 6	5120	2560	640
P1VD14 그룹 7	160	80	**
P2VD14 그룹 7	2560	1280	320

[0140] ** PRNT의 최저 희석 한계 미만(역가 40),

[0141] *** PRNT의 최고 희석 범위 초과(역가 5120)

[0142] 표 5에 나타낸 바와 같이, 모든 1가 백신 컨스트럭트가 중화 항체 반응을 유도하였다. 이 반응은 엄격한 PRNT 90 임계값의 데이터에서 알 수 있듯이 매우 강력했다.

[0143] **실시예 6: 3가 SARS-CoV-2 VLP 백신 컨스트럭트의 효능 평가**

[0144] 아래 서열 모두를 포함하는 항원 플라스미드를 사용하여 실시예 2의 방법을 사용하여 3가 VLP를 제조하였다:

[0145] i) SARS-CoV-2(SEQ ID NO:: 6);

[0146] ii) SARS-CoV(SEQ ID NO:: 9); 및

[0147] iii) MERS(SEQ ID NO: 12).

[0148] 3가 VLP, 1가 VLP[천연 SARS-CoV-2(SEQ ID NO: 6)를 발현함], 재조합 SARS-CoV-2(SEQ ID NO: 25) 및 Gag 단백질 단독(SEQ ID NO:)을 포함하는 백신 제형 NO:1)을 생쥐의 생체 내에서 시험하였다. 재조합 SARS-CoV-2(SEQ ID NO: 25)는 캐나다 국립 연구 위원회(National Research Council of Canada)에서 제공했다. 백신은 표 6에 나타낸 바와 같이 인산 알루미늄 애주번트(Adjuphos®)로 제형화되었다.

[0149] 40 마리의 투약 미경험 6-8주 령 C57/BL6 마우스(10 마리씩 4개 그룹)를 아래 표 6에 나타낸 백신 제형의 인간 투여량의 대략 1/20 내지 1/50로 3회 면역화시켰다. 면역화는 0일, 21일 및 42일에 하였다. 마지막 면역화 14일 후 동물을 희생시키고, 항-스파이크 단백질 항체 역가 및 중화 항체의 후속 분석을 위해 혈청을 수집하였다.

[0150] 표 6 : 백신 제형

표 6

[0151]

시험 그룹	설명	μg SARS-CoV-2 스파이크/dose	μg Gag /dose	Al+++ μg/dose	투여량/ 동물 (μℓ)
1	1가 SARS-CoV-2 천연 스파이크 단백질 VLP	0.1	50	125	250
2	3가 SARS-CoV-2 (천연 스파이크 단백질); SARS-CoV; MERS-CoV VLP	0.1	50	125	250
3	SARS-CoV-2 스파이크 재조합형	0.1	N/A	125	250
4	(엠프티(Empty) Gag	N/A	50	125	250

[0152] 항-스파이크 SARS-CoV-2, 항-SARS 및 항-MERS 항체 역가를, 실시예 5에 설명된 기술을 사용하여, 각 그룹에 대해 아래 포획(capture) 항원[SARS-COV-2 스파이크 단백질(S1 + S2), Sino Biological, Cat# 40589-V08B1, SARS-COV스파이크 단백질(S1 + S2), MyBioSource, Cat# MBS434077 및 MERS-CoV 스파이크 단백질(S1 + S2), Sino Biological, Cat# 40069-V08B]으로 측정하였다. 그 결과를 표 7에 나타내었다.

[0153] 표 7 : 코로나바이러스 항체 역가

표 7

[0154]

그룹	시점	항-SARS-CoV2 스파이크 바인딩 역가	항-SARS-CoV 스파이크 바인딩 역가	항-MERS 스파이크 바인딩 역가
풀링된 그룹 1 1가 VLP	P1Vd14	2700	300	네거티브
	P2Vd14	72900	8100	100
	P3Vd14	218700	24300	네거티브
풀링된 그룹 2 3가 VLP	P1Vd14	2700	900	>2700
	P2Vd14	24300	24300	>72900
	P3Vd14	>72900 (72900 ~ 218700)	72900	656100
풀링된 그룹 3 안정화된 재조합 SARS-CoV2	P1Vd14	2700	900	네거티브
	P2Vd14	72900	24300	900
	P3Vd14	218700	24300	네거티브
풀링된 그룹 4 엠프티 Gag VLP	P1Vd14	네거티브	네거티브	네거티브
	P2Vd14	네거티브	네거티브	네거티브
	P3Vd14	네거티브	네거티브	네거티브

[0155] 표 7에 나타난 바와 같이 3가 VLP(그룹 2)는 SARS-CoV-2, SARS 및 MERS의 3가지 코로나바이러스 모두에 대한 항체 반응을 유도했다. 이는 3가 백신 후보가 세 가지 주요 코로나바이러스 모두에 대해 면역학적 보호를 다시 제공할 가능성이 있음을 보여준다.

[0156] 3차 백신 접종 후 개별 동물에 대한 항-SARS-CoV-2 바인딩 및 PRNT 80 중화 역가를 아래 표 8에 나타내었다. 중화 항체는 실시예 5에 기술된 방법을 사용하여 측정하였다.

[0157] 표 8 : 개별 마우스에 대한 항-SARS-CoV-2 바인딩 및 중화 역가

표 8

[0158]

그룹 #	설명	마우스 #	항- SARS-CoV2 바인딩 역가	항- SARS-CoV2 바인딩 역가 기하 평균	항 SARS-CoV2 중화 항체 (PRNT 80)	중화 항체 기하 평균	nAb/ 바인딩역가 비율
1	SARS-CoV-2 천연 스파이크 단백질 eVLP	6	1,057,003	355538	2560	1810	0.002
		7	221,691		640		0.003
		8	515,381		1280		0.002
		9	240,199		5120		0.021
		10	124,759		1280		0.01
		11	1,320,203		5120		0.004
		12	258,417		1280		0.005
		13	252,174		640		0.003
		14	454,574		5120		0.011
		15	228,032		1280		0.006

2	SARS-CoV-2+ SARS-CoV-+MERS-CoV 천연 스파이크 단백질 eVLP	21	333,263	247080	>5120	>5120	0.015
		22	299,575		>5120		0.017
		23	208,191		>5120		0.024
		24	142,458		>5120		0.036
		25	294,434		>5120		0.017
		26	129,136		>5120		0.04
		27	322,558		>5120		0.016
		28	390,847		>5120		0.013
		29	278,663		>5120		0.019
		30	214,394		>5120		0.024

[0159] 표 8에 나타낸 데이터에서 알 수 있는 바와 같이, 3가 VLP는, 바인딩 역가가 더 낮았음에도 불구하고, 1가 SARS-CoV-2 VLP보다 더 높은 중화 항체 반응을 유도함을 보여준다. 이는 표 8의 마지막 열에서, 바인딩 항체에 대한 중화 항체의 역가 비율을 관찰할 때 특히 명백하다. 이는 3가 백신 후보가 COVID-19에 대해 더 강력한 면역학적 보호를 제공할 가능성을 보여준다.

[0160] 각 백신 접종 후 14일째에 마우스로부터 얻은 혈청을, 인간을 감염시키고 감기를 유발하는 것으로 알려진 다른 코로나바이러스(HCoV-OC43)와의 교차 반응성에 대해 시험했다. 항체 역가는, 포획 항원으로 인간 코로나바이러스(HCoV-OC43) 스파이크 단백질(S1+S2 ECD, His Tag), Sino, Cat#40607-V08B, 스톡 0.25 mg/mL를 사용하여 위에서 설명한 대로 ELISA를 사용하여 측정되었다. 그 결과를 아래 표 9에 나타내었다.

[0161] 표 9 : HCoV-OC43 스파이크 단백질에 대한 마우스 혈청의 교차 반응성

표 9

[0162]

그룹	백신 접종 넘버	항-HCoV-OC43 스파이크 전체 IgG 바인딩 역가
풀링된 그룹 1 1가 VLPs	P1VD14	네거티브
	P2VD14	100
	P3VD14	300
공동그룹 2 3가 VLPs	P1VD14	네거티브
	P2Vd14	900
	P3VD14	2700
풀링된 그룹 3 NRC 안정화 재조합형 SSARS-CoV2	P1VD14	네거티브
	P2VD14	300
	P3Vd14	300
풀링된 그룹 4 엠프티 eVLP	P1VD14	네거티브
	P2VD14	네거티브
	P3VD14	네거티브

[0163] 위의 표 9에서 볼 수 있듯이 3가 VLP 백신 후보(그룹 2)는 감기를 유발하는 인간 코로나바이러스에 대해 더 높은 교차 반응성을 보였다. 따라서 3가 후보는 1가 VLP 또는 재조합 SARS-CoV-2 스파이크 단백질 단독보다 코로나바이러스에 대한 광범위한 보호 가능성을 입증했다.

[0164] 백신 제형의 효능을 평가하기 위해, 4명의 회복된 COVID-19 환자로부터 수집된 인간 혈청(HS)에서 중화 항체를 또한 측정하였고, 그 결과를 4개의 상이한 시험군에 의해 유도된 중화 항체와 비교하였다. PRNT 50 및 PRNT 90은 실시예 5에 기술된 방법을 사용하여 1차 및 2차 백신 접종 후에 결정되었다. 각 그룹에 대한 풀링된 결과는 아래 표 10에 나와 있다.

[0165] 표 10 - 항 SARS-CoV-2 중화 항체

표 10

[0166]

샘플	PRNT50	PRNT90
HS1	80	**
HS2	160	40

HS3	1280	320
HS4	160	80
POV (백신 접종 전)	**	**
P1VD14 GR1	**	**
P1Vd14 GR2	**	**
P1Vd14 GR3	40	**
P1Vd14 GR4	**	**
P2Vd14 GR1	640	160
P2Vd14 GR2	320	**
P2Vd14 GR3	640	80
P2Vd14 GR4	**	**

[0167] ** PRNT의 최저 희석 한계 미만(역가 40)

[0168] 표 10에서 볼 수 있듯이, 1가 VLP 백신은 PRNT 50과 90으로 측정한 사람 4명 중 3명에서 코로나19 감염보다 더 많은 중화 항체를 유도했다. 3가 VLP 백신은 PRNT 50에 의해 측정된 4명의 인간 환자 중 3명에서 코로나19 감염보다 더 많은 중화 항체를 유도했다. 따라서 백신 컨스트럭트는 SARS-CoV-2에 대한 노출보다 면역 보호를 유도하는 데 적어도 효과적이고, 잠재적으로 더 효과적이다.

[0169] **실시예 7 : 1가 SARS-CoV-2 VLP 백신 컨스트럭트의 보호 효과 평가**

[0170] 시리아 골든 햄스터(수컷, 대략 5-6 주령)를 두 그룹으로 나누고 아래 표 11에 나타낸 제형, 구체적으로 인산 알루미늄 애주번트(Adjuphos®)(그룹 B) 및 식염수 대조군(그룹 A)으로 제형화된 삼중 변형된 SARS-CoV2 VLP(SEQ ID: 24)를 포함하여 구성되는 시험 샘플의 2 투여량으로 면역화했다. 면역화는 근육내 주사를 통해 0일 및 21일에 이루어졌다. 42일째에 모든 동물에게 동물당 1 x 105 TCID50의 시험 접종 바이러스 투여량으로(at a challenge virus dose) 양쪽 콧구멍을 통해 50 ㎕의 SARS-CoV-2를 비강내로 시험 접종하였다(challenged). SST(혈청 분리 튜브) 혈액 샘플(각각 약 0.5 ml)을 1차 면역화 전 0일, 14일 및 35일에 각각 수집하였다. 부검 시 최종 혈액 샘플을 수집했다. 비강 세척액은 35일, 43일, 44일, 45일, 47일, 49일, 51일, 53일 및 56일에 수집되었다. 각 그룹의 동물의 절반은 시험 접종 후 3일에 안락사되었고 나머지 동물은 시험 접종 후 14일에 안락사되었다.

[0171] 표 11 : 1가 SARS-CoV-2 VLP 백신 제형

표 11

[0172]

그룹	시험물	㎍ SARS-CoV-2 스파이크 /dose	Al+++ ㎍/dose	투여량/동물1 (㎕)
A	식염수 대조군	N/A	N/A	100
B	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 24)	1	125	100

[0173] 부검 시, 육안적 폐 병리(gross lung pathology)를 평가하고 병변을 포함하는 폐엽의 비율을 평가했다(estimated). 폐 조직은 qRT-PCR 및 바이러스 세포 배양에 의해 바이러스 양(viral load)에 대해 분석되었다. 유사하게, qRT-PCR 및 바이러스 세포 배양에 의해 바이러스 양에 대해 비갑개를 수집하였다.

[0174] 비강 세척제로부터 RNA의 추출은 Qiagen 시약(QIAamp Viral RNA Mini Kit Cat No./ID: 52906)을 사용하여 수행되었다. 간단히 말해서, 140 ㎕의 비강 세척액을 560 ㎕의 바이러스 용해 완충액(Buffer AVL)에 첨가하였다. 혼합물을 실온에서 10분 동안 인큐베이션하였다. 간단한 원심분리 후, 용액을 560 ㎕의 100% 에탄올을 함유하는 새로운 튜브로 옮기고, 튜브를 실온에서 10분 동안 인큐베이션하였다. 그런 다음 RNA를 정제하고 0.04% 아지드화나트륨(용출 완충액 AVE)을 함유하는 RNase 유리수 60㎕로 용리했다.

[0175] 약 100㎍의 조직을 사용하여 폐엽 및 비갑개로부터 RNA의 추출을 완료하였다. 조직을 30Hz에서 6분 동안 TissueLyserII(Qiagen)에서 멸균 스테인리스 스틸 비드와 함께 600㎕의 용해 완충액(RLT Qiagen)에서 균질화했다. 용액을 5000 x g에서 5분 동안 원심분리했다. 상청액을 600 ㎕의 70% 에탄올을 함유한 새로운 튜브에 옮기고, 그 튜브를 실온에서 10분 동안 인큐베이션하였다. 그런 다음 바이러스 RNA를 Qiagen RNeasy Mini Kit(Cat

No/ID: 74106)를 사용하여 정제하고 50㎕ 용리 완충액으로 용리했다.

[0176] SARS-CoV-2 특이적 프라이머를 사용하여 비강 세척, 폐 조직 및 비갑개 샘플로부터의 RNA에 대해 qRT-PCR 검정을 수행하였다(표 12). 프라이머는 약 60℃의 어닐링 온도를 가졌다. qRT-PCR에는 Qiagen Quantifast RT-PCR Probe 키트를 사용하였고, qRT-PCR 반응은 OneStep Plus(Applied Biosystems) 기계를 이용하여 수행하였다. qRT-PCR 결과는 SARS-CoV-2의 env 유전자를 함유하는 선형화된 플라스미드 DNA 샘플로 표준 곡선을 생성함으로써 반응당 복제 수(copy number)로 표현되었다. 개별 샘플의 Ct 값을 표준 곡선과 함께 사용하여 각 샘플의 복제 수를 결정했다.

[0177] 표 12 : 사용된 프라이머의 서열

표 12

[0178]

프라이머	서열	서열번호
순방향 프라이머 (Fwd)	ACAGGTACGTTAATAGTTAATAGCGT	28
역방향 프라이머 (Rev)	ATATTGCAGCAGTACGCACACA	29
표지 프로브(Labelled Probe)	ACACTAGCCATCCTTACTGCGCTTCG	30

[0179] 감염성 바이러스를 평가하기 위해 바이러스 적정 분석을 수행했다. Vero'76 세포(ATCC CRL-1587)를 사용하여 96웰 플레이트에서 분석을 수행했다. 중간 조직 배양 감염 투여량(TCID50)은 세포의 세포변성 효과(CPE)의 현미경 관찰에 의해 결정되었다. 바이러스를 정량화하고 TCID50/ml 또는 TCID50/그램으로 보고했다. TCID50 값은 Excel의 Spearman & Karber 알고리즘을 사용하여 계산되었다.

[0180] 항-스파이크 SARS-CoV-2 항체 역가는 혈청 샘플에 대해 수행된 ELISA에 의해 측정하였다. 플레이트를 스파이크 S1+S2 Ag(Cat#40589-V08B1, Sino Biological Inc.)로 코팅했다. 코팅 농도는 0.1 ug/mL였다. 플레이트는 0.05% Tween 20을 함유하는 PBS 중 5% 무지방 탈지분유로 차단되었다. 혈청의 4배 희석액을 사용했다. ThermoFisher(PA1-29626)의 염소 항-햄스터(Goat anti-Hamster) IgG HRP를 1:7000에서 이차 항체로 사용했다. 플레이트는 OPD 퍼옥시다제 기재(0.5 mg/ml)(Thermo Scientific Pierce 34006)로 디벨롭되었다(developed). 2.5 M 황산으로 반응을 멈추고 490 nm에서 흡광도를 측정하였다. 검정 전체에 걸쳐 플레이트를 0.05% Tween 20을 함유하는 PBS로 세척하였다. 검정을 이중으로 수행하였다. 역가를 희석(dilutions)의 종말점으로 보고하였다.

[0181] 스파이크 단백질 수용체 결합 도메인(receptor binding domain, "RBD")에 대한 항체를 다음과 같이 측정하였다. 항-SARS-CoV-2 스파이크 S1 RBD IgG 항체 바인딩 역가를 간접 ELISA를 사용하여 혈청 샘플로부터 결정하였다. 재조합 SARS-CoV-2 스파이크 S1 RBD 단백질을 미세역가(microtiter) 플레이트에 밤새 흡착시킨 다음, 그 플레이트를 세척 완충액 중 5% 탈지유 용액으로 1시간 동안 차단했다. 차단 및 세척 후, 샘플을 마이크로플레이트에 첨가하고 1.5시간 동안 인큐베이션하였다. HRP-복합 goat anti-Syrian Hamster IgG-Fc를 검출 항체로 사용하고 마이크로플레이트에서 1시간 동안 인큐베이션하였다. 신호는 테트라메틸벤지딘(TMB) 기재 용액으로 디벨롭되었으며 TMB Microwell 기재용 450㎕ 액체 정지액(Liquid Stop Solution)을 추가하여 반응을 중단했다. 흡광도는 ELISA 마이크로웰 플레이트 판독기를 사용하여 450 nm에서 판독되었다.

[0182] 시험 접종 SARS-CoV-2 바이러스에 대한 바이러스 중화 분석은 세포주 Vero'76을 사용하여 혈청 샘플에 대해 수행되었다. 혈청 샘플을 56℃에서 30분 동안 열 불활성화시켰다(heat-inactivated). 혈청 샘플을 연속 희석했다(2배 연속 희석). 실험은 기술 복제로(in technical duplicates) 수행되었다. 바이러스를 웰 당(per well) 50㎕(접종원 크기 = 25TCID50) 중 $TCID_{50}$의 농도로 배지에서 희석하였다. 그런 다음 60㎕의 바이러스 용액을 60㎕의 연속 희석된 혈청 샘플과 혼합했다. 혼합물을 5% CO2와 함께 37℃에서 1시간 동안 인큐베이션하였다. 사전 배양된 바이러스-혈청 혼합물(100㎕)을 90% 컨플루언트 사전 시드 Vero'76 세포를 포함하는 96웰 평평한 바닥 플레이트의 웰로 옮겼다. 플레이트를 5일 동안 5% CO_2와 함께 37℃에서 인큐베이션했다. 플레이트는 감염 후 1일째(dpi)에 현미경을 사용하여 오염에 대해, 그리고 감염 후 3일 및 5일에 세포변성 효과(cytopathic effect)에 대해 현미경을 사용하여 관찰했다. 5 dpi에서 CPE가 없는 웰에 대한 혈청 희석 인자를 혈청 중화 역가로 정의했다. 초기 혈청 희석 인자는 1:20이었다.

[0183] 중화 항체를 다음과 같이 시험하였다. Vero 세포를 감염 48시간 전에 6-웰 플레이트에 8 x 105세포/웰로 시딩했다. 혈청을 56℃에서 30분 동안 열 불활성화한 다음 얼음으로 옮겼다. 혈청을 바이러스 감염 배지로 1:10 희석

한 다음, 각각의 희석된 혈청을 사용하여 1/20 내지 1:40960(8회 후속 희석)을 제공하기 위해 1/2배 연속 희석을 수행했다. 동일한 부피의 희석된 혈청 및 바이러스(혈청 희석액당 100 pfu)를 혼합하고 37℃에서 1시간 동안 인큐베이션했다. 혈청 및 바이러스 대조군이 포함되지 않았다. 세포를 PBS로 세척하고 각 바이러스/혈청을 세포를 함유하는 각 웰에 옮기고 혼합하고 플레이트를 간격을 두고 흔들면서 37℃에서 1시간 동안 인큐베이션하였다. 1시간 흡착 후, 과잉 접종물을 제거하고 2 ml 바이러스 감염 배지/아가로스 혼합물을 세포에 오버레이(overlay, 덧씌움)했다. 오버레이가 굳도록 하고, 플레이트를 37℃에서 72시간 동안 인큐베이션했다. 감염 후 72시간에 세포를 크리스탈 바이올렛으로 염색하였다. 모든 희석액에 대해 플라크를 정량화하고 PRNT 역가를 계산했다. 무혈청 대조군에 기초한 모든 희석액에 대한 플라크 감소율 %(% plaque reduction)은 PRNT 역가 50, 80 및 90을 결정하기 위해 Reed-Muench 공식을 사용하여 계산되었다.

[0184] 폐 조직은 또한, 부검에서 수집된 사이토카인 유전자 발현에 대해 정량화되었다. IL-4, IL-10, IL-13, TNF-알파 및 IFN-감마의 유전자 발현은 표 13에 나타낸 프라이머를 사용하여 qRT-PCR에 의해 오른쪽 두개골 및 오른쪽 꼬리 폐엽(caudal lung lobe)에서 결정되었다. 베타-액틴 유전자 발현은 참조용으로 사용되었다.

[0185] 표 13 : 프라이머 서열

표 13

[0186]

유전자 표적	순방향 프라이머 5'-> 3'	서열번호	역방향 프라이머 5'->3'	서열번호
IL-4	CCACGGAGAAAGACCTCATCTG	31	GGGTCACCTCATGTTGGAAATAAA	32
IL-10	GTTGCCAAACCTTATCAGAAATGA	33	TTCTGGCCCGTGGTTCTCT	34
IL-13	AAATGGCGGGTTCTGTGC	35	AATATCCTCTGGGTCTTGTAGATGG	36
TNF-알파	GGAGTGGCTGAGCCATCGT	37	AGCTGGTTGTCTTTGAGAGACATG	38
IFN-감마	GGCCATCCAGAGGAGCATAG	39	TTTCTCCATGCTGCTGTTGAA	40
베타-액틴	ACTGCCGCATCCTCTTCCT	41	TCGTTGCCAATGGTGATGAC	42

[0187] 폐 조직을 RNAlater에서 수집하고, RLT 용해 완충액(Qiagen RNeasy Mini Kit, Cat No/ID:74106)을 사용하는 Qiagen RNeasy Mini 추출 키트를 사용하여 RNA를 분리했다. RNA 농도와 순도의 지표인 260/280 비율은 나노드롭(nanodrop) 분광광도계로 측정했다. cDNA는 500 ng의 RNA를 템플레이트(template)로 하여 iScript™ Reverse Transcription Supermix를 사용하여 합성되었다. cDNA는 25℃에서 5분, 46℃에서 20분, 95℃에서 1분의 프로그램에 따라 합성되었다. 마스터 믹스는, 10% 초과량의(at 10% overage)의 하우스 키핑 유전자(house keeping gene); 1.84㎕ 뉴클레아제 없는 H_2O; 순방향 프라이머 0.08 ㎕; 역방향 프라이머 0.08 ㎕; 및 SYBR 10 ㎕ [SYBR® Green PCR Master Mix (SsoAdvanced™ Universal SYBR® Green Supermix #1725275)];뿐만 아니라, 관심 있는 각 유전자에 대해 준비되었다. 각 PCR 반응에 대해 12㎕의 마스터 믹스를 8㎕의 RNA와 조합했다. 로딩 후 플레이트를 1분 동안 1500RPM에서 원심분리하여 모든 액체를 웰의 바닥으로 되돌려 놓았다. qPCR은 Bio-Rad Therm℃ycler(Bio-Rad CX1000)를 사용하여 수행되었다. 데이터는 Bio-Rad CFX Maestro 소프트웨어를 사용하여 분석되었다. 데이터는 Excel의 ΔΔCt 공식에 의한 폴드 변경 계산(fold change calculation)을 위해 Ct 값의 형태로 Excel 스프레드시트로 내보내진다.

[0188] 동물의 임상 관찰에 기초한 결과는 모든 동물이 면역화 단계 전반에 걸쳐 건강했음을 나타내었다. 모든 동물은 정상적인 활동 수준을 보였고 임상 징후는 없었다. 체중 증가는 식염수 대조군(그룹 A)과 비교할 때 시험물로 백신 접종된 그룹(그룹 B)에서 정상이었다.

[0189] SARS-CoV-2 스파이크 단백질에 대한 항체 역가로 측정한 백신 접종에 대한 면역 반응은 1차 백신 접종 후 14일 및 2차 백신 접종 후 14일에 표 14에 나와 있다(P1 및 P2는 1차 및 2차 백신 접종을 나타냄). 표시된 결과는 각 그룹의 동물에 대한 Geo 평균이다.

[0190] 표 14 ; SARS-CoV-2 VLP 백신 항체 역가

표 14

[0191]

	1차 백신 접종 (day 14)		2차 백신 접종 (day 14)	
	항-SARS-CoV-2 S1 RBD	항-SARS-CoV-2 스파이크 전체 IgG	항-SARS-CoV-2 S1 RBD	항-SARS-CoV-2 스파이크 전체 IgG
그룹 A (대조군)	-	3	-	3
그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24)	7,446	1,222	268,399	22,868

[0192] 그룹 B의 동물(3중 변형된 1가 SARS-CoV-2 VLP 백신으로 면역화됨)은 두 번째 백신 접종 2주 후에 높은 수준의 항스파이크 항체를 생성했다. 첫 번째 백신 접종 후 2주에 그룹 B의 동물 12마리 중 10마리가 스파이크 방지 항체를 생성했다(데이터는 표시되지 않음). 그룹 A의 동물(식염수 대조군)은 스파이크 방지 항체 생산이 없었다. 삼중 변형된 1가 SARS-CoV-2 VLP 백신은 또한 첫 번째 면역 접종 후 14일에 검출 가능한 수준의항-SARS-CoV-2-S1 RBD IgG 항체를 유도했다. 항체 역가의 실질적인 증가는 2차 면역화 후 14일째에 관찰되었다. 대조군 A에서는 항-SARS-CoV-2-S1 RBD IgG가 검출되지 않았다.

[0193] 첫 번째 백신 접종 후 14일째에 각 그룹에 대하여 PRNT에 의해 결정된 중화 항체를 표 15에 나타냈다(평균 값 표시). 그 값은 각각 입력 바이러스의 50%(PRNT50), 80%(PRNT80) 또는 90%(PRNT90)의 억제를 나타내는 최고 희석의 역수(reciprocal of highest dilution)를 나타낸다.

[0194] 표 15 : SARS-CoV-2 VLP 백신 중화 항체

표 15

[0195]

시험 그룹 - 1차 백신 접종후 14일	PRNT50	PRNT80	PRNT90
풀링된 그룹 A (대조군)	**	(N/A)	(N/A)
그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24)	441	241	198

[0196] ** PRNT의 최저 희석 한계 미만(역가 40)

[0197] 그룹 B의 모든 동물은 표 15에 나타낸 바와 같이 면역화 후 2주에 바이러스 중화 항체를 생성하였다. 그룹 A의 동물은 표 15에 나타낸 바와 같이 어떠한 중화 항체도 생성하지 않았다.

[0198] 시험 접종 후 3일(dpc)에, 그룹 B의 모든 동물 및 그룹 A의 단 한 마리의 동물이 중화 항체를 생성하였다(데이터는 나타내지 않음). 14 dpc에서 그룹 A와 B의 모든 동물은 중화 항체를 생성했다(데이터는 표시되지 않음).

[0199] 시험 접종 단계 동안, 2 마리를 제외한 모든 동물은 활동적이었고 정상적인 활동 수준을 가졌으며 비정상적인 코 징후를 나타내지 않았다.

[0200] 시험 접종 후 매일 동물의 체중을 측정하였다. 시험 접종 후 그룹 A 동물은 초기 체중의 약 15%를 잃었고 6-8 dpc에서 최고치를 기록했다. 그룹 B 동물의 % 체중 변화의 평균은 약 1-2%에 불과했으며 2 dpc에서 최고조에 달했다. 체중 데이터는 시험 접종 후 0일 및 3일 및 6일에 대해 아래 표 16에 제시되어 있다.

[0201] 표 16 : 바이러스 시험 접종 후 햄스터의 체중

표 16

[0202]

그룹	동물 ID	0 일(Day)		3일째		6일째
		BW	BW 평균	BW	BW 평균	BW

A (대조군)	401	175.1	171.5±14.5	160.9	160.9±13.9	Euth
	403	182.1		167.3		155.4
	408	172.6		160.7		Euth
	412	173.6		161		149.1
	416*	165.9		163.2		137.1
	422	196.3		185.8		170.5
	423	162.8		149.6		Euth
	429	185.2		171.6		159.6
	434	174.8		170.00		161.5
	435	156.4		143.7		Euth
	440	174.3		164.9		Euth
	443	139.4		132.1		Euth
B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 24))	402	159.5	174.2±14.3	155.7	171.6±14.3	Euth
	404	181.2		177.1		179.6
	410	161.4		160.1		Euth
	411	190.0		183.6		182.1
	418	181.3		179.1		185.1
	421*	188.4		189.5		191.6
	426	160.0		159.0		Euth
	428	175.8		173.3		172.8
	431	172.3		169.1		170.3
	436	152.7		147.3		Euth
	442	199.1		195.4		Euth
	446	168.8		169.9		Euth

[0203] 계획대로 3일째에 안락사, 또는 인도적으로 안락사 *(416 및 421)

[0204] 표 16에서 알 수 있는 바와 같이, 식염수 용액을 투여한 동물은 시험 접종 3일 및 6일 후에 상당한 체중이 감소한 반면, 백신을 투여받은 동물은 3일에 체중이 상당히 적게 감소하고 6일에 체중이 증가했다.

[0205] 시험 접종 후 비강 세척액에서 측정된 바이러스 RNA가 표 17에 제시되어 있다. 조사된 모든 날에, 표 17에 도시된 바와 같이 백신 접종된 동물(그룹 B)은 그룹 A 동물(대조군)보다 비강 세척액에서 더 낮은 바이러스 RNA 수준을 가졌다 (시험 접종 후 매일의 복제/Rxn 표시). 시험 접종 후 2일 동안에만 바이러스 RNA 수준이 그룹 A와 비교하여 그룹 B에서 유의하게 더 낮았다(p=0.0206).

[0206] 표 17 : 비강 세척액의 바이러스 RNA

표 17

[0207]

그룹 ID	시험 접종 16일 후 바이러스RNA 평균 (복제/Rxn)							
	0	2	3	5	7	9	11	14
그룹 A	0.1	3708602	340897	94932	12409	6231	130	3146
그룹 B	0.1	205616	105469	31861	93	20	2461	2

[0208] 대조군(그룹 A) 및 백신 접종된(그룹 B) 동물에 대한 다양한 조직에서 시험처리 후 3일째의 바이러스 RNA를 표 18에 나타내었다(복제/그램에 대한 값 표시). 시험 접종 후 3일째에, 바이러스 RNA는 모든 동물에서 폐의 우두엽(right cranial lobe)(RCra) 및 우미엽(right caudal lobe)(RCau) 및 비갑개(nasal turbinates)에서 검출가능하였다. 그룹 A와 비교할 때, 그룹 B의 RCra에서 바이러스 RNA의 수준은 상당히 낮았다. 유사하게, RCau의 RNA 수준은 그룹 A보다 그룹 B에서 유의하게 낮았다. 비갑개에서 그룹 B의 바이러스 RNA 수준은 그룹 A보다 유의하게 낮았다.

[0209] 표 18 : 시험 접종 후 3일째 조직의 바이러스 RNA

표 18

[0210]

그룹 ID	우두엽	우미엽	비갑개
그룹 A(대조군)	2613650533	1608083300	2099767910

그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24))	108006	42465	1205108163

[0211] 대조군(그룹 A) 및 백신 접종된(그룹 B) 동물에 대한 다양한 조직에서 시험 접종 후 14일째의 바이러스 RNA를 표 19에 나타내었다(복제/그램에 대한 값 표시). 챌린지 후 14일째에 RCra, RCau 또는 비갑개에서 모든 그룹 A 동물과 그룹 B의 일부 동물에서 바이러스 RNA가 검출되었다. RCra 및 RCau의 RNA 수준은 그룹 A와 그룹 B에서 유의하게 달랐다.

[0212] 표 19 : 접종 후 14일째 조직의 바이러스 RNA

표 19

[0213]

그룹 ID	우두엽	우미엽	비갑개
그룹 A(대조군)	127855.5	62075.0	3104415.2
그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24))	99.43	105.78	100398.57

[0214] 대조군(그룹 A) 및 백신 접종된(그룹 B) 동물에 대한 시험처리 후 3일째 다양한 조직의 감염성 바이러스를 표 20에 나타내었다(TCID50/그램에 대한 값 표시). 시험처리 후 3일째에, 폐의 우두엽 및 우미엽 및 비갑개에서 그룹 A의 모든 동물에서 감염성 바이러스가 검출 가능하였다. 그룹 B의 감염성 바이러스의 역가는 그룹 A의 것보다 유의하게 낮았다. 시험 접종 후 14일에 감염성 바이러스는 어떤 동물에서도 검출되지 않았다(데이터는 표시되지 않음).

[0215] 표 20 : 접종 후 3일째 조직의 감염성 바이러스(TCID50/그램)

표 20

[0216]

그룹 ID	우두엽		우미엽		비갑개	
	TCID50/그램	평균	TCID50/그램	평균	TCID50/그램	평균
그룹 A (대조군)	1.51E+07	2.63E+07	6.71E+02	1.65E+06	1.15E+07	2.00E+07
	1.03E+07		3.94E+05		2.39E+06	
	3.89E+07		3.45E+05		3.76E+07	
	1.30E+07		1.23E+06		4.66E+05	
	1.32E+06		5.99E+06		2.47E+07	
그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24))	1.79E+01	1.48E+01	8.78E+00	3.45E+00	1.26E+02	3.58E+06
	1.49E+01		2.56E+00		5.43E+05	
	2.39E+01		2.28E+00		5.86E+01	
	9.05E+00		2.42E+00		2.09E+07	
	1.12E+01		2.59E+00		6.85E+00	

[0217] 더 무거운 폐는 더 진행된 질병과 관련이 있다. 따라서, 신체에 대한 폐 무게의 비율은 더 심각한 질병 상태와 상관 관계가 있다. 표 21은 시험 접종 3일 후 대조군(그룹 A) 및 백신 접종된(그룹 B) 동물의 폐 중량 대 체중 비율을 보여준다. 그룹 B 동물의 동물은 체중 대 폐 중량 비율이 상당히 낮았다.

[0218] 표 21 : 체중에 대한 폐 중량의 비율(%)

표 21

[0219]

그룹 ID	폐 중량(g)	체중(g)	폐 중량/체중 비	평균 ±SD

그룹 A (대조군)	1.4	160.9	0.87	0.76±0.15
	1.2	170.1	0.71	
	1.5	160.7	0.93	
	1.1	160.9	0.68	
	0.88	137.1	0.64	
	1.4	183.7	0.76	
	1.6	149.6	1.07	
	1.3	183.6	0.71	
	1	182.3	0.55	
	1.1	143.7	0.77	
	1.3	164.9	0.79	
	0.8	132.1	0.61	
그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형) (SEQ ID NO: 24)	0.9	155.7	0.58	0.56±0.07
	1	188.2	0.53	
	1	160.1	0.62	
	0.9	188.2	0.48	
	0.9	195.2	0.46	
	0.98	191.6	0.51	
	1	159	0.63	
	0.9	180.1	0.50	
	1	183.6	0.54	
	1	147.3	0.68	
	1.2	195.4	0.61	
	NA	169.9	missing	

[0220] 부검 후, 폐 조직을 포르말린에 고정하고, 임베드하고(embedded), 절단하고(sectioned), 헤마톡실린 및 에오신(H&E)으로 염색했다. 슬라이드는 면허 소지(board-certified) 병리학자가 검사하고, 표 22에 표시된 대로 0-4의 척도로 점수를 매겼다.

[0221] 표 22 : 폐 조직학 점수(중앙값)

표 22

[0222]

그룹 ID	접종후 일 수	영향받은 (affected) 유조직의 부분	영향 받은 구역의 염증 침윤의 강도	폐포성 폐세포 (pneumocytes) 이상 발달 정도	폐포간 출혈 (Interalveolar hemorrhage)	폐기종 정도
그룹 A (대조군)	3	3	3	0	2	0
		3	3	0	1	0
		3	3	0	2	0
		1	3	0	3	0
		1	3	0	2	0
		1	2	0	1	0
	14	1	2	3	0	0
		1	2	2	0	0
		4	4	3	4	0
		4	2	2	0	0
		2	3	2	0	0
		1	2	1	0	0

그룹 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형 (SEQ ID NO: 24))	3	0	0	0	0	1
		1	2	0	0	0
		1	1	0	0	0
		1	2	0	0	0
		1	1	0	1	0
		1	1	0	0	0
	14	0	0	0	0	0
		0	0	0	0	0
		0	0	0	0	0
		1	1	0	0	0
		0	0	0	0	0
		0	0	0	0	0

[0223] 표 22에서 알 수 있는 바와 같이, 대조군(그룹 A)의 동물은, 시험 접종후 3일 및 14일에 유의한 질병 병리를 나타냈다. 대조적으로, 백신 접종된 동물(그룹 B)은 3일에 약간의 경미한 병리를 나타냈으나, 14일까지 대부분 회복되었다. 따라서, 백신은 질병 유발 폐 병리학에 대해 상당한 보호를 제공했다.

[0224] 폐 조직, 특히 유조직 및 세기관지/기관지에서 SARS-CoV-2 바이러스를 관찰하기 위해, 면역조직화학 염색을 수행하였다. 염색 작업이 관찰되었고, 두 동물 그룹에 대한 점수가 표 23에 나와 있다.

[0225] 표 23 : SARS-CoV-2에 대한 폐의 면역조직화학적 점수(중앙값)

표 23

[0226]

그룹 ID	접종후 일 수	유조직 IHC		세기관지/기관지	
		개별 점수	중앙 값	개별 점수	중앙 값
그룹 A (대조군)	3	4	4±0.84	2	3±0.52
		4		3	
		4		2	
		4		3	
		3		3	
		2		3	
	14	0	0±1.21	2	2±0.75
		0		2	
		3		1	
		1		1	
		0		3	
		0		2	
대조군 B (VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24))	3	0	0±0	2	1±0.89
		0		1	
		0		2	
		0		0	
		0		1	
		0		0	
	14	0	0±0	0	1±0.63
		0		1	
		0		2	
		0		1	
		0		1	
		0		1	

[0227] 백신을 접종한 동물은 식염수 대조군 동물(그룹 A)보다 유조직과 세기관지/기관지 모두에서 바이러스 얼룩이 훨씬 적었다. 시험 접종 후 14일에 바이러스 스테인(virus stain)은 백신 접종 그룹에서 약간 낮았지만 유조직 또는 세기관지/기관지 그룹 간에 유사했다.

[0228] 우두엽, 우미엽, 및 비갑개에서 사이토카인 IL-4, IL-10, IL-13, TNF-알파 및 IFN-감마의 전사 수준을 qRT-PCR에 의해 결정했다. 시험 접종 후 3일째에, IL-10, IL-13 및 IFN-감마는 그룹 B의 우두엽엽 및 우미엽에서 차별

적인 발현을 나타냈다(표 24 및 25 참조). 비갑개에서 IL-10 및 IFN-감마는 차등 발현을 나타내었다(표 26 참조). 시험 접조 후 14일에 우두엽, 우미엽 및 비갑개에서 IL-4, IL-10, IL-13, TNF-알파 및 IFN-감마의 전사 수준은 그룹 간에 유사했다(표 27-29에 나타냄).

[0229] 표 24 : 접종 3일 후 우두엽에서 사이토카인의 전사 프로파일(폴드 체인지)

표 24

[0230]

그룹 ID	IL-4	IL-10	IL-13	TNF-알파	IFN-감마
그룹 A	1.04	1.09	1.3	1.07알파	1.14
그룹 B	1.8	0.37	7.1	0.61	0.27

[0231] 표 25 : 접종 3일 후 우미엽에서 사이토카인의 전사 프로파일(폴드 체인지)

표 25

[0232]

그룹 ID	IL-4	IL-10	IL-13	TNF-알파	IFN-감마
그룹 A	1.02	1.03	1.33	1.04	1.21
그룹 B	0.88	0.33	2.4	0.55	0.35

[0233] 표 26 : 접종 3일 후 비갑개에서 사이토카인의 전사 프로파일(폴드 체인지)

표 26

[0234]

그룹 ID	IL-4	IL-10	IL-13	TNF-알파	IFN-감마
그룹 A	1.22	1.05	1.26	1.16	1.09
그룹 B	1.44	0.84	2.87	0.45	0.82

[0235] 표 27 : 접종 14일 후 우두엽에서 사이토카인의 전사 프로파일(폴드 체인지)

표 27

[0236]

그룹 ID	IL-4	IL-10	IL-13	TNF-알파	IFN-감마
그룹 A	1.07	1.25	1.57	1.34	1.24
그룹 B	1.37	0.92	2.22	0.77	0.61

[0237] 표 28 : 접종 14일 후 우미엽에서 사이토카인의 전사 프로파일(폴드 체인지)

표 28

[0238]

그룹 ID	IL-4	IL-10	IL-13	TNF-알파	IFN-감마
그룹 A	1.09	1.51	1.13	1.11	1.31
그룹 B	1.38	0.78	0.63	0.94	0.65

[0239] 표 29 접종 14일 후 비갑개에서 사이토카인의 전사 프로파일(폴드 체인지)

표 29

[0240]

그룹 ID	IL-4	IL-10	IL-13	TNF-알파	IFN-감마
그룹 A	1.46	1.45	1.43	1.41	1.46
그룹 B	0.73	0.72	1.40	0.65	0.57

[0241] 실시예 8: 1가 SARS-CoV-2 VLP 백신 컨스트럭트의 단일 투여량의 효능 및 보호 효과 평가

[0242] 시리아 골든 햄스터(수컷, 대략 6-7 주령)를 아래 표 30에 나타낸 1가 삼중 변형된 SARS-CoV-2 VLP 백신 제형으로 면역화시켰다. 면역화는 근육내 주사를 통해 21일째에만 발생하였다. 중화 항체의 후속 분석을 위해 0일 및

35일에 혈청을 수집하였다.

[0243] 표 30 : 1가 SARS-CoV-2 VLP 백신 제형

표 30

[0244]

그룹	시험물	㎍ SARS-CoV-2 스파이크 /dose	A1+++ ㎍/dose	투여량/ 동물 (㎕)
A	식염수 대조군	N/A	N/A	100
B	VSV로부터의 TM/Cyt를 갖는 SARS-CoV-2 프롤린 및 퓨린 절단 변형(SEQ ID NO: 24)	1.4	125	100

[0245] 중화 항체를 그룹 B의 동물에 대해 실시예 7에 기술된 바와 같이 플라크 감소 중화 시험(PRNT)을 사용하여 시험하였다. 결과를 표 31에 나타내었다(평균 값을 나타냄).

[0246] 표 31 : 1가 SARS-CoV-2 VLP 백신 중화 항체

표 31

[0247]

시험 그룹	PRNT50	PRNT80	PRNT90
0 일(day) (백신 접종전)	**	-	-
백신 접종후 14일 그룹 B	190	65	45

[0248] ** PRNT의 최저 희석 한계 미만(역가 40)

[0249] 실시예 7의 그룹 B(동물이 0일 및 21일에 1㎍ SARS-CoV-2 스파이크/투여량의 면역화를 받은 경우)와 비교하여, 이 실시예 8의 그룹 B에 있는 동물(동물이 21일에 1.4㎍ SARS-CoV-2 스파이크/투여량을 받은 경우)이 더 높은 혈청 중화 항체 반응을 나타냈다. 이 데이터는 1가 SARS-CoV-2 VLP 백신의 단 1회 투여만으로 효과적인 예방접종이 됨을 뒷받침한다.

[0250] 실시예 7에 기술된 바와 같이, 42일째에 시험 접종 연구를 수행하였다. 표 32는 시험 접종 전 동물의 평균 체중(그램)을 나타낸다.

[0251] 표 32 : 시험 접종전 동물의 평균 체중

표 32

[0252]

Day	그룹 A	그룹 B
-1	113.8	116.6
7	131.7	134.6
10	136.3	139.9
22	154.5	157.6
28	160.4	164.0
35	168.4	172.0
42	174	178.5

[0253] 표 33은 시험 접종 후 동물의 평균 체중(그램)을 보여준다. 표 33에서 볼 수 있는 바와 같이, 백신 1회 투여량을 받은 동물은 식염수를 투여받은 동물보다 체중 감소가 적었다.

[0254] 표 33 : 시험 접종 후 동물의 평균 체중

표 33

[0255]

일(day)	그룹 A	그룹 B
1	172.1	174.6
2	168.3	172.6
3	165.3	173.4

4	163.7	168.5
5	159.1	168.5
6	155.7	168.5
7	154.7	170.1
8	157.5	171.7
9	162.9	173.8
10	164.4	174.0
11	166.6	175.3
12	169.4	177.1
13	172.0	177.8
14	173.5	178.9

[0256] 표 34는 시험 접종 후 동물의 체중의 평균 변화 %를 나타낸다.

[0257] 표 34 : 접종 후 동물의 평균 체중 변화 %

표 34

[0258]

Day	그룹 A	그룹 B
1	-1.09	-2.24
2	-3.29	-3.30
3	-5.08	-2.90
4	-6.51	-2.10
5	-9.14	-2.15
6	-11.11	-2.11
7	-11.64	-1.22
8	-10.05	-0.31
9	-7.03	0.87
10	-6.22	0.93
11	-4.94	1.63
12	-3.35	2.69
13	-1.89	3.09
14	-1.00	3.74

[0259] 이들 데이터는, 21일째에 1.4 ㎍ SARS-CoV-2 스파이크/투여량의 단일 면역화가 대조군에 비해 바이러스 시험 접종(viral challenge) 후 체중 감소를 예방하는 데 효과적이었다는 것을 보여준다.

[0260] **실시예 9: 남아프리카 공화국 SARS-CoV-2 변종에 대한 항체 역가를 위한 1가 및 3가 SARS-CoV-2 VLP 백신 컨스트럭트의 평가**

[0261] 3중 변형된 SARS-CoV-2 스파이크 단백질을 갖는 1가 및 3가 SARS-CoV-2 VLP 백신 컨스트럭트를 남아프리카 공화국 SARS-CoV-2 변종에 대한 항체 생성에 대해 평가하였다. 마우스를 아래 표 35에 나타낸 SARS-CoV-2 VLP 백신 제형으로 2회(실시예 6에 기술된 바와 같이 0일 및 21일에) IP 면역화시켰다(immunized IP). 면역화 14일 후에 동물을 희생시키고, 안티 스파이크 단백질 항체 역가의 후속 분석을 위해 혈청을 수집하였다.

[0262] SARS-CoV-2 VLP를 표 35에 나타낸 바와 같이 인산 알루미늄 애주번트(Adjuphos®)와 함께 제형화하였다.

[0263] 표 35 : SARS-CoV-2 VLP 백신 제형

표 35

[0264]

그룹	시험물	㎍ SARS-CoV-2 스파이크/dose	Al+++ ㎍/dose	투여량/ 동물 (㎕)
1	VSV로부터의 TM/Cyt를 갖는 1가 VLP - SARS-CoV-2프롤린 및 퓨린 절단 변형(SEQ ID NO: 24)	0.2	125	500
2	3가 VLP (v4) - SARS-CoV-2 스파이크 단백질 (실시예 4에 설명됨)	0.2	125	500

[0265] 표 36 : SARS-CoV-2 VLP 백신 항체 역가

표 36

[0266]

그룹	2차 백신 접종후 일 수	항- SARS-CoV-2 스파이크 전체 IgG 바인딩 역가	항- SARS-CoV-2 아프리카 변종 스파이크 전체 IgG 바인딩 역가
그룹 1	14	128,850	121,511
그룹 2	14	215,232	211,080

[0267] 이 데이터는 1가 및 3가 백신을 주사한 마우스가 SARS-CoV의 스파이크 단백질의 남아프리카 공화국 변종에 바인딩하는 항체를 생성했음을 보여준다.

[0268] 실시예 10: SARS-CoV-2 VLP 백신 컨스트럭트의 아이소타입(isotype) 역가 평가

[0269] 다른 연구에서, 표 37에 나타낸 백신 컨스트럭트로 마우스를 면역화한 후 항체 역가의 아이소타입을 평가하였다.

[0270] 표 37 : SARS-CoV-2 VLP 백신 항체 역가

표 37

[0271]

그룹		㎍ SARS-CoV-2 스파이크 /dose	Al+++ ㎍/dose	투여량/ 동물 (㎕)
1	1가 천연 SARS-Cov-2 (SEQ ID NO: 6) VLP	0.2	125	500
2	안정화 재조합형 SARS-CoV2 스파이크 단백질 (비-VLP)	0.2	125	500

[0272] 마우스를 2회(실시예 6에 기술된 바와 같이 0일 및 21일에) IP 면역화시켰다. 면역화 14일 후에 동물을 희생시키고, 항-스파이크 단백질 항체 역가의 후속 분석을 위해 혈청을 수집하였다.

[0273] 표 38에 나타난 바와 같이, 예상외로 VLP가 재조합 스파이크 단백질과 동일한 양/농도의 명반(alum)으로 제형화되었을 때, 균형 잡힌 항체 반응이 나타났다(IgG1/IgG2b). 증가된 IgG2b 생산은 세포 매개 면역을 나타내는 TH1 면역 반응과 관련이 있고, 이것은 VLP 컨스트럭트가, 더 효과적인 TH1 면역 반응과 상관관계가 있는, 상승된 수준의 IgG2b 발현을 가져왔음을 나타낸다.

[0274] 표 38 : SARS-CoV-2 VLP 백신 항체 역가

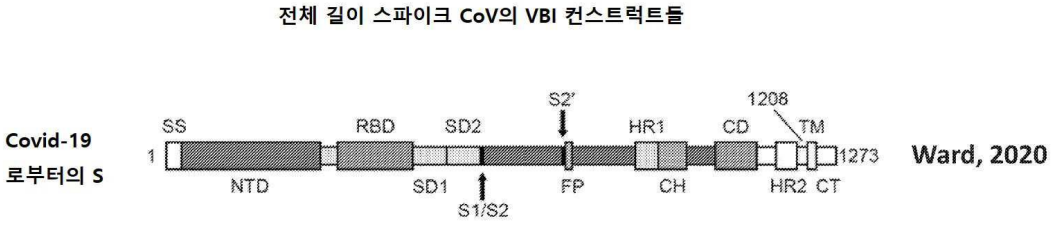
표 38

[0275]

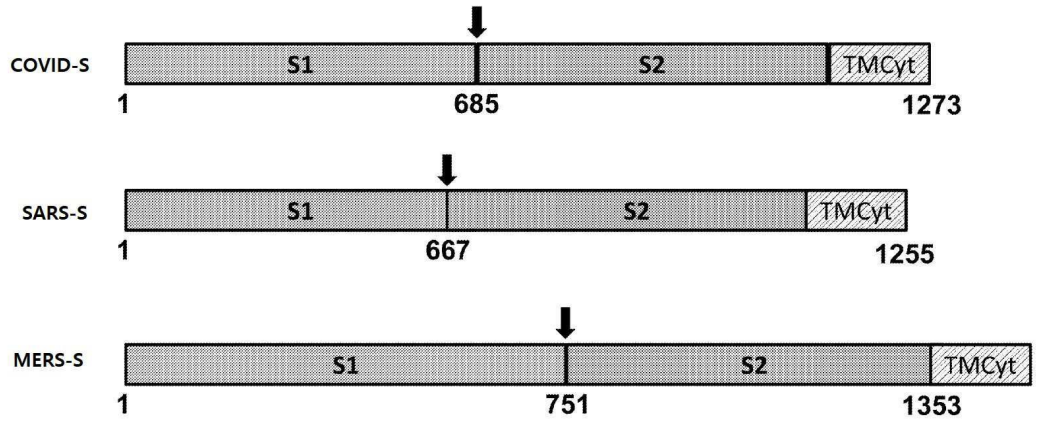
그룹	2차 접종후 일 수	항- SARS-CoV-2 스파이크 IgG1 바인딩 역가	항- SARS-CoV-2 스파이크 IgG2b 바인딩 역가
그룹 1	14	172,105	116,633
그룹 2	14	198,469	9,674

도면

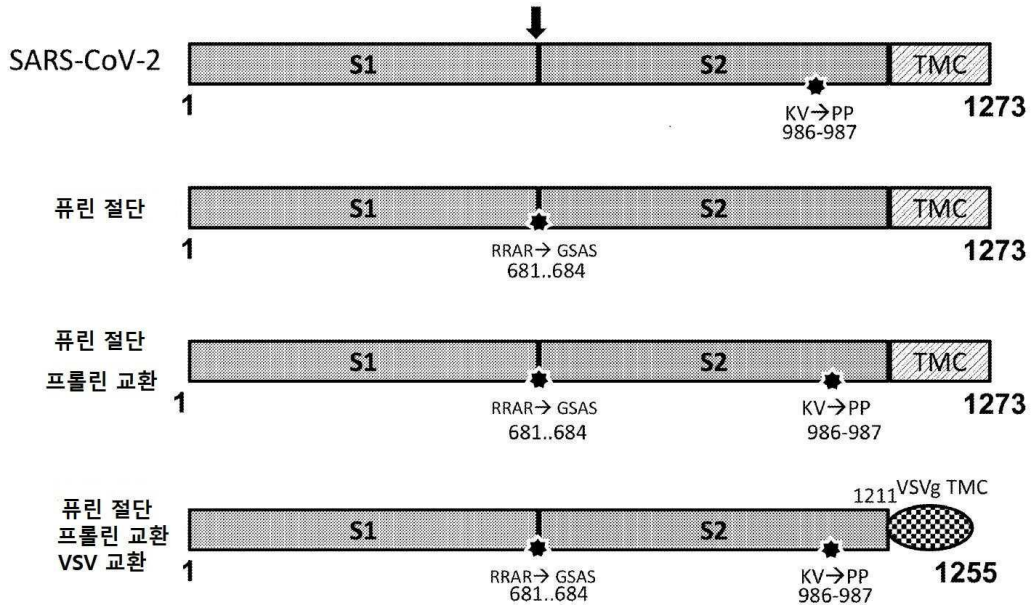
도면1



도면2



도면3



서 열 목 록

<110> VARIATION BIOTECHNOLOGIES INC.

<120> VACCINE COMPOSITIONS FOR THE TREATMENT OF CORONAVIRUS

<130> PAT 110628W-90

<140> PCT/US2021/025020

<141> 2021-03-30

<150> 63/070,150

<151> 2020-08-25

<150> 63/002,237

<151> 2020-03-30

<160> 44

<170> PatentIn version 3.5

<210> 1

<211> 538

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 1

Met Gly Gln Thr Val Thr Thr Pro Leu Ser Leu Thr Leu Gly His Trp

1 5 10 15

Lys Asp Val Glu Arg Ile Ala His Asn Gln Ser Val Asp Val Lys Lys

20 25 30

Arg Arg Trp Val Thr Phe Cys Ser Ala Glu Trp Pro Thr Phe Asn Val

35 40 45

Gly Trp Pro Arg Asp Gly Thr Phe Asn Arg Asp Leu Ile Thr Gln Val

50 55 60

Lys Ile Lys Val Phe Ser Pro Gly Pro His Gly His Pro Asp Gln Val

65 70 75 80

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

85 90 95

Val Lys Pro Phe Val His Pro Lys Pro Pro Pro Pro Leu Pro Pro Ser

100 105 110

Ala Pro Ser Leu Pro Leu Glu Pro Pro Arg Ser Thr Pro Pro Arg Ser

115 120 125

Ser Leu Tyr Pro Ala Leu Thr Pro Ser Leu Gly Ala Lys Pro Lys Pro

130 135 140

Gln Val Leu Ser Asp Ser Gly Gly Pro Leu Ile Asp Leu Leu Thr Glu

145 150 155 160

Asp Pro Pro Pro Tyr Arg Asp Pro Arg Pro Pro Pro Ser Asp Arg Asp

165 170 175

Gly Asn Gly Gly Glu Ala Thr Pro Ala Gly Glu Ala Pro Asp Pro Ser

180 185 190

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

195 200 205

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

210 215 220

Leu Gln Tyr Trp Pro Phe Ser Ser Ser Asp Leu Tyr Asn Trp Lys Asn

225 230 235 240

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

245 250 255

Glu Ser Val Leu Ile Thr His Gln Pro Thr Trp Asp Asp Cys Gln Gln

260 265 270

Leu Leu Gly Thr Leu Leu Thr Gly Glu Glu Lys Gln Arg Val Leu Leu

275 280 285

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

290 295 300

Pro Asn Glu Val Asp Ala Ala Phe Pro Leu Glu Arg Pro Asp Trp Asp

305 310 315 320

Tyr Thr Thr Gln Ala Gly Arg Asn His Leu Val His Tyr Arg Gln Leu

325 330 335

Leu Leu Ala Gly Leu Gln Asn Ala Gly Arg Ser Pro Thr Asn Leu Ala

340 345 350

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

355 360 365

Leu Glu Arg Leu Lys Glu Ala Tyr Arg Arg Tyr Thr Pro Tyr Asp Pro

370 375 380

Glu Asp Pro Gly Gln Glu Thr Asn Val Ser Met Ser Phe Ile Trp Gln
385 390 395 400
Ser Ala Pro Asp Ile Gly Arg Lys Leu Glu Arg Leu Glu Asp Leu Lys
405 410 415
Asn Lys Thr Leu Gly Asp Leu Val Arg Glu Ala Glu Lys Ile Phe Asn
420 425 430
Lys Arg Glu Thr Pro Glu Glu Arg Glu Glu Arg Ile Arg Arg Glu Thr

435 440 445
Glu Glu Lys Glu Glu Arg Arg Arg Thr Glu Asp Glu Gln Lys Glu Lys
450 455 460
Glu Arg Asp Arg Arg Arg His Arg Glu Met Ser Lys Leu Leu Ala Thr
465 470 475 480
Val Val Ser Gly Gln Lys Gln Asp Arg Gln Gly Gly Glu Arg Arg Arg
485 490 495
Ser Gln Leu Asp Arg Asp Gln Cys Ala Tyr Cys Lys Glu Lys Gly His
500 505 510

Trp Ala Lys Asp Cys Pro Lys Lys Pro Arg Gly Pro Arg Gly Pro Arg
515 520 525
Pro Gln Thr Ser Leu Leu Thr Leu Asp Asp
530 535

<210> 2
<211> 1614
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polynucleotide
<400> 2
atgggccaga ctgttaccac tcccttaagt ttgaccttag gtcactggaa agatgtcgag 60
cggatcgctc acaaccagtc ggtagatgtc aagaagagac gttgggttac cttctgctct 120

gcagaatggc caacctttaa cgtcggatgg ccgcgagacg gcacctttaa ccgagacctc 180
atcacccagg ttaagatcaa ggtcttttca cctggcccgc atggacaccc agaccaggtc 240
ccctacatcg tgacctggga agccttggct tttgaccccc ctccctgggt caagcccttt 300
gtacacccta agcctccgcc tcctcttcct ccatccgccc cgtctctccc ccttgaacct 360

cctcgttcga ccccgcctcg atcctccctt tatccagccc tcactccttc tctaggcgcc 420

aaacctaaac ctcaagttct ttctgacagt ggggggccgc tcatcgacct acttacagaa 480

gacccccccgc cttataggga cccaagacca ccccccttccg acagggacgg aaatggtgga 540

gaagcgaccc ctgcgggaga ggcaccggac ccctccccaa tggcatctcg cctacgtggg 600

agacgggagc cccctgtggc cgactccact acctcgcagg cattcccccct ccgcgcagga 660

ggaaacggac agcttcaata ctggccgttc tcctcttctg acctttacaa ctggaaaaat 720

aataaccctt ctttttctga agatccaggt aaactgacag ctctgatcga gtctgttctc 780

atcacccatc agcccacctg ggacgactgt cagcagctgt tggggactct gctgaccgga 840

gaagaaaaac aacgggtgct cttagaggct agaaaggcgg tgcggggcga tgatgggcgc 900

cccactcaac tgcccaatga agtcgatgcc gcttttcccc tcgagcgccc agactgggat 960

tacaccaccc aggcaggtag gaaccaccta gtccactatc gccagttgct cctagcgggt 1020

ctccaaaacg cgggcagaag ccccaccaat ttggccaagg taaaaggaat aacacaaggg 1080

cccaatgagt ctccctcggc cttcctagag agacttaagg aagcctatcg caggtacact 1140

ccttatgacc ctgaggaccc agggcaagaa actaatgtgt ctatgtcttt catttggcag 1200

tctgccccag acattgggag aaagttagag aggttagaag atttaaaaaa caagacgctt 1260

ggagatttgg ttagagaggc agaaaagatc tttaataaac gagaaacccc ggaagaaaga 1320

gaggaacgta tcaggagaga aacagaggaa aaagaagaac gccgtaggac agaggatgag 1380

cagaaagaga aagaaagaga tcgtaggaga catagagaga tgagcaagct attggccact 1440

gtcgttagtg gacagaaaca ggatagacag ggaggagaac gaaggaggtc ccaactcgat 1500

cgcgaccagt gtgcctactg caaagaaaag gggcactggg ctaaagattg tcccaagaaa 1560

ccacgaggac ctcggggacc aagaccccag acctccctcc tgaccctaga tgac 1614

<210> 3

<211> 1614

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 3

atgggacaga ccgtcacaac acccctgagc ctgaccctgg gacattggaa agacgtggag 60

aggatcgcac ataaccagag cgtggacgtg aagaaacgga gatgggtcac attctgcagt 120

gctgagtggc caacttttaa tgtgggatgg ccccgagacg gcactttcaa cagggatctg 180

atcacccagg tgaagatcaa ggtctttagc ccaggacctc acggacatcc agaccaggtg 240

ccttatatcg tcacctggga ggcactggcc ttcgatcccc ctccatgggt gaagccattt 300

gtccacccaa aaccacctcc accactgcct ccaagtgccc cttcactgcc actggaacca 360

ccccggagca caccaccccg cagctccctg tatcctgctc tgactccatc tctgggcgca 420

aagccaaaac cacaggtgct gagcgactcc ggaggaccac tgattgacct gctgacagag 480

gacccccccac cataccgaga tcctcggcct ccaccaagcg accgcgatgg aaatggagga 540

gaggctactc ctgccggcga agcccctgac ccatctccaa tggctagtag gctgcgcggc 600

aggcgcgagc ctccagtggc agatagcacc acatcccagg ccttccctct gagggctggg 660

ggaaatgggc agctccagta ttggccattt tctagttcag acctgtacaa ctggaagaac 720

aataacccct ctttcagtga ggaccccggc aaactgaccg ccctgatcga atccgtgctg 780

attacccatc agcccacatg ggacgattgt cagcagctcc tgggcaccct gctgaccgga 840

gaggaaaagc agcgcgtgct gctggaggct cgcaaagcag tccgagggga cgatggacgg 900

cccacacagc tccctaatga ggtggacgcc gcttttccac tggaaagacc cgactgggat 960

tatactaccc aggcagggag aaaccacctg gtccattaca ggcagctcct gctggcaggc 1020

ctgcagaatg ccgggagatc ccccaccaac ctggccaagg tgaaaggcat cacacagggg 1080

cctaatgagt caccaagcgc ctttctggag aggctgaagg aagcttaccg acggtatacc 1140

ccatacgacc ctgaggaccc cggacaggaa acaaacgtct ccatgtcttt catctggcag 1200

tctgccccag acattgggcg gaagctggag agactggaag acctgaagaa caagaccctg 1260

ggcgacctgg tgcgggaggc tgaaaagatc ttcaacaaac gggagacccc cgaggaaaga 1320

gaggaaagga ttagaaggga aactgaggaa aaggaggaac gccgacggac cgaggacgaa 1380

cagaaggaga aagaacgaga tcggcggcgg caccgggaga tgtcaaagct gctggccacc 1440

gtggtcagcg gacagaaaca ggacagacag ggaggagagc gacggagaag ccagctcgac 1500

agggatcagt gcgcatactg taaggaaaaa ggccattggg ccaaggattg ccccaaaaag 1560

ccaagaggac caagaggacc aagaccacag acatcactgc tgaccctgga cgac 1614

<210> 4

<211> 1273

<212> PRT

<213> Severe acute respiratory syndrome coronavirus 2

<400> 4

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val

1 5 10 15

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

20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu

35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp

50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp

65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu

85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser

100 105 110

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

115 120 125

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

130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr

145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu

165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe

180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr

195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu

210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys

290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val

305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys

325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala

340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu

355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro

370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe

385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly

405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys

420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn

435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe

450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys

465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly

485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val

500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys

515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe

580 585 590
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605
Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
675 680 685
Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
690 695 700
Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile
705 710 715 720
Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val

725 730 735
Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu
740 745 750
Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr
755 760 765
Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln
770 775 780

Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe
785 790 795 800

Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
805 810 815
Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly
820 825 830
Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp
835 840 845
Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu
850 855 860
Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly

865 870 875 880
Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
885 890 895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
900 905 910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn
915 920 925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala
930 935 940

Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn
945 950 955 960
Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val
965 970 975
Leu Asn Asp Ile Leu Ser Arg Leu Asp 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 Leu

1010 1015 1020
Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val

1025 1030 1035 1040

Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala

1045 1050 1055

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu

1060 1065 1070

Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His

1075 1080 1085

Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val

1090 1095 1100

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr

1105 1110 1115 1120

Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr

1125 1130 1135

Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu

1140 1145 1150

Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp

1155 1160 1165

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

1170 1175 1180

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

1185 1190 1195 1200

Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile

1205 1210 1215

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

1220 1225 1230

Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys

1235 1240 1245

Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val

1250 1255 1260

Leu Lys Gly Val Lys Leu His Tyr Thr

1265 1270

<210> 5

<211> 3822

<212> DNA

<213> Severe acute respiratory syndrome coronavirus 2

<400> 5

atgtttgttt ttcttgtttt attgccacta gtctctagtc agtgtgttaa tcttacaacc 60

agaactcaat taccccctgc atacactaat tctttcacac gtggtgttta ttaccctgac 120

aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc tttcttttcc 180

aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat 240

aaccctgtcc taccatttaa tgatggtgtt tattttgctt ccactgagaa gtctaacata 300

ataagaggct ggatttttgg tactacttta gattcgaaga cccagtccct acttattgtt 360

aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420

ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480

tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggaccttgaa 540

ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat tgatggttat 600

tttaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660

tcggctttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720

ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg ttggacagct 780

ggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840

gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaaag 900

tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960

caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc ttttggtgaa 1020

gtttttaacg ccaccagatt tgcatctgtt tatgcttgga acaggaagag aatcagcaac 1080

tgtgttgctg attattctgt cctatataat tccgcatcat tttccacttt taagtgttat 1140

ggagtgtctc ctactaaatt aaatgatctc tgctttacta atgtctatgc agattcattt 1200

gtaattagag gtgatgaagt cagacaaatc gctccagggc aaactggaaa gattgctgat 1260

tataattata aattaccaga tgattttaca ggctgcgtta tagcttggaa ttctaacaat 1320

cttgattcta aggttggtgg taattataat tacctgtata gattgtttag gaagtctaat 1380

ctcaaacctt ttgagagaga tatttcaact gaaatctatc aggccggtag cacaccttgt 1440

aatggtgttg aaggttttaa ttgttacttt cctttacaat catatggttt ccaacccact 1500

aatggtgttg gttaccaacc atacagagta gtagtacttt cttttgaact tctacatgca 1560

ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa atgtgtcaat 1620

ttcaacttca atggtttaac aggcacaggt gttcttactg agtctaacaa aaagtttctg 1680

cctttccaac aatttggcag agacattgct gacactactg atgctgtccg tgatccacag 1740

acacttgaga ttcttgacat tacaccatgt tcttttggtg gtgtcagtgt tataacacca 1800

ggaacaaata cttctaacca ggttgctgtt ctttatcagg atgttaactg cacagaagtc 1860

cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc tacaggttct 1920

aatgtttttc aaacacgtgc aggctgttta ataggggctg aacatgtcaa caactcatat 1980

gagtgtgaca tacccattgg tgcaggtata tgcgctagtt atcagactca gactaattct 2040

cctcggcggg cacgtagtgt agctagtcaa tccatcattg cctacactat gtcacttggt 2100

gcagaaaatt cagttgctta ctctaataac tctattgcca tacccacaaa ttttactatt 2160

agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga ttgtacaatg 2220

tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg cagtttttgt 2280

acacaattaa accgtgcttt aactggaata gctgttgaac aagacaaaaa cacccaagaa 2340

gtttttgcac aagtcaaaca aatttacaaa acaccaccaa ttaaagattt tggtggtttt 2400

aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt tattgaagat 2460

ctacttttca acaaagtgac acttgcagat gctggcttca tcaaacaata tggtgattgc 2520

cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg ccttactgtt 2580

ttgccacctt tgctcacaga tgaaatgatt gctcaataca cttctgcact gttagcgggt 2640

acaatcactt ctggttggac ctttggtgca ggtgctgcat tacaaatacc atttgctatg 2700

caaatggctt ataggtttaa tggtattgga gttacacaga atgttctcta tgagaaccaa 2760

aaattgattg ccaaccaatt taatagtgct attggcaaaa ttcaagactc actttcttcc 2820

acagcaagtg cacttggaaa acttcaagat gtggtcaacc aaaatgcaca agctttaaac 2880

acgcttgtta aacaacttag ctccaatttt ggtgcaattt caagtgtttt aaatgatatc 2940

ctttcacgtc ttgacaaagt tgaggctgaa gtgcaaattg ataggttgat cacaggcaga 3000

cttcaaagtt tgcagacata tgtgactcaa caattaatta gagctgcaga aatcagagct 3060

tctgctaatc ttgctgctac taaaatgtca gagtgtgtac ttggacaatc aaaaagagtt 3120

gatttttgtg gaaagggcta tcatcttatg tccttccctc agtcagcacc tcatggtgta 3180

gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac tgctcctgcc 3240

atttgtcatg atggaaaagc acactttcct cgtgaaggtg tctttgtttc aaatggcaca 3300

cactggtttg taacacaaag gaatttttat gaaccacaaa tcattactac agacaacaca 3360

tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt ttatgatcct 3420

ttgcaacctg aattagactc attcaaggag gagttagata aatattttaa gaatcataca 3480

tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt aaacattcaa 3540
aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct catcgatctc 3600
caagaacttg gaaagtatga gcagtatata aaatggccat ggtacatttg gctaggtttt 3660
atagctggct tgattgccat agtaatggtg acaattatgc tttgctgtat gaccagttgc 3720
tgtagttgtc tcaagggctg ttgttcttgt ggatcctgct gcaaatttga tgaagacgac 3780
tctgagccag tgctcaaagg agtcaaatta cattacacat aa 3822

<210> 6
<211> 3822
<212> DNA

<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polynucleotide
<400> 6

atgttcgtgt ttctggtgct gctgcctctg gtgagctccc agtgcgtgaa cctgaccaca 60
aggacccagc tccccctgc ctataccaat tccttcacac ggggcgtgta ctatcccgac 120
aaggtgttta gatctagcgt gctgcactcc acacaggatc tgtttctgcc tttcttttct 180
aacgtgacct ggttccacgc catccatgtg agcggcacca atggcacaaa gcggttcgac 240
aatccagtgc tgcccttaa cgatggcgtg tacttcgcct ccaccgagaa gtctaacatc 300

atcagaggct ggatctttgg caccacactg gacagcaaga cacagtccct gctgatcgtg 360
aacaatgcca ccaacgtggt catcaaggtg tgcgagttcc agttttgtaa tgatccattc 420
ctgggcgtgt actatcacaa gaacaataag tcttggatgg agagcgagtt tcgcgtgtat 480
tcctctgcca acaattgcac atttgagtac gtgtcccagc ccttcctgat ggacctggag 540
ggcaagcagg gcaatttcaa gaacctgagg gagttcgtgt ttaagaatat cgatggctac 600
ttcaagatct actccaagca caccccaatc aacctggtgc gcgacctgcc acagggcttc 660
tctgccctgg agccactggt ggatctgccc atcggcatca acatcacccg gtttcagaca 720

ctgctggccc tgcacagaag ctacctgaca ccaggcgaca gctcctctgg atggaccgcc 780
ggggccgccg cctactatgt gggctatctg cagcccagga ccttcctgct gaagtacaac 840
gagaatggca ccatcacaga cgcagtggat tgcgccctgg acccccctgtc tgagaccaag 900
tgtacactga agagctttac cgtggagaag ggcatctatc agacaagcaa tttcagggtg 960
cagcctaccg agtccatcgt gcgctttccc aatatcacaa acctgtgccc ttttggcgag 1020
gtgttcaacg caaccaggtt cgccagcgtg tacgcatgga ataggaagcg catctccaac 1080

tgcgtggccg actattctgt gctgtacaac agcgcctcct tctctacctt taagtgctat 1140

ggcgtgagcc ccacaaagct gaatgacctg tgctttacca acgtgtacgc cgattccttc 1200
gtgatcaggg gcgacgaggt gcgccagatc gcaccaggac agacaggcaa gatcgcagac 1260
tacaattata agctgcctga cgatttcacc ggctgcgtga tcgcctggaa ctctaacaat 1320
ctggatagca aagtgggcgg caactacaat tatctgtacc ggctgtttag aaagtctaat 1380
ctgaagccat tcgagaggga catctccaca gagatctacc aggccggctc taccccctgc 1440
aatggcgtgg agggctttaa ctgttatttc cctctgcaga gctacggctt ccagccaaca 1500
aacggcgtgg gctatcagcc ctaccgcgtg gtggtgctgt cttttgagct gctgcacgca 1560

cctgcaacag tgtgcggacc aaagaagagc accaatctgg tgaagaacaa gtgcgtgaac 1620
ttcaacttca acggcctgac cggcacaggc gtgctgaccg agtccaacaa gaagttcctg 1680
ccttttcagc agttcggcag ggacatcgca gataccacag acgccgtgcg cgaccctcag 1740
accctggaga tcctggacat cacaccatgc tccttcggcg gcgtgtctgt gatcacacca 1800
ggcaccaata caagcaacca ggtggccgtg ctgtatcagg acgtgaattg taccgaggtg 1860
cccgtggcca tccacgcaga tcagctcacc cctacatggc gggtgtactc taccggcagc 1920
aacgtgttcc agacaagagc cggctgcctg atcggagccg agcatgtgaa caatagctat 1980

gagtgcgaca tccctatcgg agccggcatc tgtgcctcct accagaccca gacaaactcc 2040
ccacggagag cccggtctgt ggccagccag tccatcatcg cctataccat gagcctgggg 2100
gccgagaaca gcgtggccta ctccaacaat tctatcgcca tccctaccaa cttcacaatc 2160
tccgtgacca cagagatcct gccagtgagc atgaccaaga catccgtgga ctgcacaatg 2220
tatatctgtg gcgattccac cgagtgctct aacctgctgc tgcagtacgg ctctttttgt 2280
acccagctca acagagccct gacaggcatc gccgtggagc aggacaagaa cacacaggag 2340
gtgttcgccc aggtgaagca gatctacaag accccaccca tcaaggactt tggcggcttc 2400

aacttcagcc agatcctgcc cgatcctagc aagccatcca agcggtcttt tatcgaggac 2460
ctgctgttca acaaggtgac cctggccgat gccggcttca tcaagcagta tggcgattgc 2520
ctgggcgaca tcgccgccag agacctgatc tgtgcccaga agtttaatgg cctgaccgtg 2580
ctgcctccac tgctgacaga tgagatgatc gcccagtaca catctgccct gctggccggc 2640
accatcacaa gcggatggac cttcggggcc ggggccgccc tgcagatccc ctttgccatg 2700
cagatggcct atcggttcaa cggcatcggc gtgacccaga atgtgctgta cgagaaccag 2760
aagctgatcg ccaatcagtt taactccgcc atcggcaaga tccaggactc tctgagctcc 2820

acagccagcg ccctgggcaa gctgcaggat gtggtgaatc agaacgccca ggccctgaat 2880

accctggtga agcagctcag cagcaacttc ggggccatca gcagcgtgct gaacgacatc 2940

ctgagccggc tggacaaggt ggaggcagag gtgcagatcg accggctgat cacaggcaga 3000

ctgcagtccc tgcagaccta cgtgacacag cagctcatca gggccgccga gatcagggcc 3060

tctgccaatc tggccgccac caagatgagc gagtgcgtgc tgggccagtc caagagagtg 3120

gacttttgtg gcaagggcta tcacctgatg agcttcccac agtccgcccc tcacggagtg 3180

gtgtttctgc atgtgaccta cgtgccagcc caggagaaga acttcaccac agcccccgca 3240

atctgccacg atggcaaggc acactttccc cgggagggcg tgttcgtgag caacggcacc 3300

cactggtttg tgacacagcg caatttctac gagccacaga tcatcaccac agacaataca 3360

ttcgtgtccg gcaactgtga cgtggtcatc ggcatcgtga acaataccgt gtatgatcct 3420

ctgcagccag agctggactc ttttaaggag gagctggata agtacttcaa gaatcacacc 3480

agccccgacg tggatctggg cgacatctct ggcatcaatg ccagcgtggt gaacatccag 3540

aaggagatcg acaggctgaa cgaggtggcc aagaatctga acgagtccct gatcgatctg 3600

caggagctgg gcaagtatga gcagtacatc aagtggccct ggtatatctg gctgggcttc 3660

atcgccggcc tgatcgccat cgtgatggtg accatcatgc tgtgctgtat gacaagctgc 3720

tgttcctgcc tgaagggctg ctgttcttgt ggctcctgct gtaagtttga tgaggacgat 3780

agcgagcctg tgctgaaggg cgtgaagctg cactacacct ga 3822

<210> 7

<211> 1255

<212> PRT

<213> Severe acute respiratory syndrome-related coronavirus

<400> 7

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln

20 25 30

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

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

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

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

325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr

340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly

355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala

370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly

385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe

405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser

420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu

435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly

450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp

465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val

485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly

500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn

515 520 525

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

530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp

545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys

565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser

580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr

595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr

610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu

625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile

645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys

660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala

675 680 685

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile

690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys

705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu

725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile

740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys

755 760 765

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe

770 775 780

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile

785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met

805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile

820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
850 855 860

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865 870 875 880

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn

885 890 895

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
900 905 910

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
915 920 925

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
930 935 940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945 950 955 960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
965 970 975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
980 985 990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
995 1000 1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
1010 1015 1020

Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His

1025 1030 1035 1040

Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
1045 1050 1055

Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
1060 1065 1070

Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
1075 1080 1085

Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
1090 1095 1100

Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
1105 1110 1115 1120
Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
1125 1130 1135
Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
1140 1145 1150
Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
1155 1160 1165
Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu

1170 1175 1180
Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu
1185 1190 1195 1200
Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu
1205 1210 1215
Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys
1220 1225 1230
Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
1235 1240 1245

Gly Val Lys Leu His Tyr Thr
1250 1255

<210> 8
<211> 3768
<212> DNA
<213> Severe acute respiratory syndrome-related coronavirus
<400> 8
atgtttattt tcttattatt tcttactctc actagtggta gtgaccttga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt 120
tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttatttctt 180
ccattttatt ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc 240

atacctttta aggatggtat ttattttgct gccacagaga aatcaaatgt tgtccgtggt 300

tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat 600
caacctatag atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720

gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac 900
cagacctcta atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca 960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
gagagaaaaa aaatttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt 1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140

aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga 1200
caaactggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc 1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat 1320
aggtatctta gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc 1380
tcccctgatg gcaaaccttg caccccacct gctcttaatt gttattggcc attaaatgat 1440
tatggttttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct 1500
tttgaacttt taaatgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt 1560

aagaaccagt gtgtcaattt taattttaat ggactcactg gtactggtgt gttaactcct 1620
tcttcaaaga gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat 1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc 1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac 1980

catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040

ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100

tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160

aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220

ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tcgcaacaca 2280

cgtgaagtgt tcgctcaagt caaacaaatg tacaaaaccc caactttgaa atattttggt 2340

ggttttaatt tttcacaaat attacctgac cctctaaagc caactaagag gtcttttatt 2400

gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc 2460

gaatgcctag gtgatattaa tgctagagat ctcatttgtg cgcagaagtt caatggactt 2520

acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580

agtggtactg ccactgctgg atggacattt ggtgctggcg ctgctcttca aataccttt 2640

gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700

aaccaaaaac aaatcgccaa ccaatttaac aaggcgatta gtcaaattca agaatcactt 2760

acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820

ttaaacacac ttgttaaaca acttagctct aattttggtg caatttcaag tgtgctaaat 2880

gatatccttt cgcgacttga taaagtcgag gcggaggtac aaattgacag gttaattaca 2940

ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000

agggcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060

agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgcat 3120

ggtgttgtct tcctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180

ccagcaattt gtcatgaagg caaagcatac ttccctcgtg aaggtgtttt tgtgtttaat 3240

ggcacttctt ggtttattac acagaggaac ttcttttctc cacaaataat tactacagac 3300

aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360

gatcctctgc aacctgagct tgactcattc aaagaagagc tggacaagta cttcaaaaat 3420

catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgtcgtcaac 3480

attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaatga atcactcatt 3540

gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggta tgtttggctc 3600

ggcttcattg ctggactaat tgccatcgtc atggttacaa tcttgctttg ttgcatgact 3660

agttgttgca gttgcctcaa gggtgcatgc tcttgtggtt cttgctgcaa gtttgatgag 3720

gatgactctg agccagttct caagggtgtc aaattacatt acacataa 3768

<210> 9

<211> 3768

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 9

atgttcatct ttctgctgtt cctgaccctg acaagcggct ccgacctgga taggtgcacc 60

acatttgacg atgtgcaggc ccccaactac acacagcaca ccagctccat gaggggcgtg 120

tactatcctg atgagatctt ccgctctgac acactgtacc tgacccagga cctgttcctg 180

ccttttttata gcaatgtgac aggcttccac accatcaatc acacatttgg caaccccgtg 240

atcccctttca aggatggcat ctactttgcc gccaccgaga agtctaacgt ggtgcgggga 300

tgggtgttcg gcagcacaat gaacaataag tctcagagcg tgatcatcat caacaatagc 360

accaacgtgg tcatcagagc ctgcaatttt gagctgtgcg acaaccccctt ctttgccgtg 420

tccaagccta tgggcaccca gacacacacc atgatctttg ataatgcctt caactgtacc 480

tttgagtaca tcagcgatgc cttttccctg gacgtgtctg agaagtccgg caacttcaag 540

cacctgaggg agttcgtgtt taagaataag gacggcttcc tgtacgtgta taagggctat 600

cagcccatcg atgtggtgcg cgacctgcct tccggcttca acaccctgaa gccaatcttt 660

aagctgcccc tgggcatcaa tatcaccaac ttcagggcca tcctgacagc ctttagccca 720

gcacaggaca tctggggcac cagcgccgcc gcctacttcg tgggctatct gaagcccacc 780

accttcatgc tgaagtacga tgagaacggc acaatcaccg acgccgtgga ttgcagccag 840

aatccactgg ccgagctgaa gtgttccgtg aagtctttcg agatcgacaa gggcatctat 900

cagacctcca actttagggt ggtgccatct ggcgatgtgg tgcgcttccc aaatatcacc 960

aacctgtgcc ccttcggcga ggtgtttaat gccacaaagt tccccagcgt gtacgcctgg 1020

gagcgcaaga agatcagcaa ctgcgtggcc gactactccg tgctgtataa tagcaccttc 1080

ttcagcacct tcaagtgcta cggcgtgagc gccaccaagc tgaatgacct gtgcttctct 1140

aacgtgtatg ccgatagctt tgtggtgaag ggcgacgatg tgaggcagat cgcacctgga 1200

cagaccggcg tgatcgcaga ctacaactat aagctgccag acgatttcat gggctgcgtg 1260

ctggcctgga atacacgcaa catcgatgcc acatccaccg gcaactacaa ttataagtac 1320

cggtatctga gacacggcaa gctgcggccc ttcgagagag acatctccaa tgtgccattt 1380

tctccagatg gcaagccatg caccccacct gccctgaatt gttactggcc tctgaacgac 1440

tacggcttct ataccacaac cggcatcggc taccagcctt atagggtggt ggtgctgtcc 1500

tttgagctgc tgaacgcacc tgcaaccgtg tgcggaccaa agctgtctac agatctgatc 1560

aagaatcagt gcgtgaactt caacttcaac ggcctgacag gcaccggcgt gctgacccct 1620

tctagcaagc ggttccagcc atttcagcag ttcggcagag acgtgagcga tttcaccgac 1680

tccgtgcgcg acccaaagac atccgagatc ctggacatca gcccctgctc ctttggcggc 1740

gtgtctgtga tcacacctgg caccaacgcc tcctctgagg tggccgtgct gtaccaggat 1800

gtgaattgta ccgacgtgag cacagcaatc cacgcagacc agctcacccc agcatggcgg 1860

atctattcca ccggcaacaa cgtgttccag acacaggcag gatgcctgat cggagccgag 1920

catgtggata caagctacga gtgcgacatc cccatcggag ccggcatctg tgcctcttat 1980

cacaccgtga gcctgctgag atccacatct cagaagtcta tcgtggccta caccatgagc 2040

ctgggggccg atagctccat cgcctattcc aacaatacca tcgccatccc aacaaacttc 2100

agcatctcca tcacaaccga agtgatgccc gtgtccatgg ccaagacctc tgtggactgc 2160

aacatgtaca tctgtggcga tagcacagag tgcgccaatc tgctgctgca gtatggctcc 2220

ttttgtaccc agctcaaccg ggccctgtct ggaatcgccg ccgagcagga caggaataca 2280

cgcgaggtgt tcgcccaggt gaagcagatg tacaagacac ctaccctgaa gtattttggc 2340

ggcttcaact tttctcagat cctgcctgat ccactgaagc caaccaagcg gagcttcatc 2400

gaggacctgc tgtttaataa ggtgacactg gccgatgccg gcttcatgaa gcagtacggc 2460

gagtgcctgg gcgacatcaa cgccagagac ctgatctgtg cccagaagtt taatggcctg 2520

accgtgctgc cacccctgct gacagacgat atgatcgcag catataccgc cgccctggtg 2580

tccggcacag ccaccgccgg ctggaccttc ggggccgggg ccgccctgca gatccctttc 2640

gccatgcaga tggcctaccg gtttaacggc atcggcgtga cccagaatgt gctgtatgag 2700

aaccagaagc agatcgccaa tcagtttaac aaggccatca gccagatcca ggagtccctg 2760

acaaccacat ctaccgccct gggcaagctg caggacgtgg tgaatcagaa cgcccaggcc 2820

ctgaatacac tggtgaagca gctcagcagc aacttcgggg ccatcagcag cgtgctgaac 2880

gacatcctga gccggctgga caaggtggag gcagaggtgc agatcgatag gctgatcacc 2940

ggcagactgc agtctctgca gacatacgtg acccagcagc tcatcagggc cgccgagatc 3000

agagccagcg ccaacctggc cgccacaaag atgtccgagt gcgtgctggg ccagtctaag 3060

agggtggact tctgtggcaa gggctaccac ctgatgtcct ttccacaggc cgcccctcac 3120

ggagtggtgt tcctgcatgt gacctatgtg ccttctcagg agcgcaactt taccacagcc 3180

ccagcaatct gccacgaggg caaggcatac ttcccccggg agggcgtgtt cgtgtttaac 3240

ggcacctcct ggtttatcac acagagaaat ttcttttccc ctcagatcat caccacagac 3300

aataccttcg tgagcggcaa ctgtgacgtg gtcatcggca tcatcaacaa tacagtgtac 3360

gatcctctgc agccagagct ggacagcttc aaggaggagc tggataagta cttcaagaac 3420

cacacctccc ccgacgtgga tctgggcgac atcagcggca tcaatgcctc cgtggtgaac 3480

atccagaagg agatcgacag actgaatgag gtggccaaga atctgaacga gtccctgatc 3540

gatctgcagg agctgggcaa gtacgagcag tatatcaagt ggccatggta cgtgtggctg 3600

ggcttcatcg ccggcctgat cgccatcgtg atggtgacca tcctgctgtg ctgtatgaca 3660

tcttgctgta gctgcctgaa gggagcctgc tcctgtggct cttgctgtaa gtttgacgag 3720

gacgatagcg agcccgtgct gaagggcgtg aagctgcact atacctga 3768

<210> 10

<211> 1353

<212> PRT

<213> Middle East respiratory syndrome-related coronavirus

<400> 10

Met Ile His Ser Val Phe Leu Leu Met Phe Leu Leu Thr Pro Thr Glu
1 5 10 15

Ser Tyr Val Asp Val Gly Pro Asp Ser Val Lys Ser Ala Cys Ile Glu
20 25 30

Val Asp Ile Gln Gln Thr Phe Phe Asp Lys Thr Trp Pro Arg Pro Ile
35 40 45

Asp Val Ser Lys Ala Asp Gly Ile Ile Tyr Pro Gln Gly Arg Thr Tyr
50 55 60

Ser Asn Ile Thr Ile Thr Tyr Gln Gly Leu Phe Pro Tyr Gln Gly Asp
65 70 75 80

His Gly Asp Met Tyr Val Tyr Ser Ala Gly His Ala Thr Gly Thr Thr

85 90 95

Pro Gln Lys Leu Phe Val Ala Asn Tyr Ser Gln Asp Val Lys Gln Phe
100 105 110

Ala Asn Gly Phe Val Val Arg Ile Gly Ala Ala Ala Asn Ser Thr Gly
115 120 125

Thr Val Ile Ile Ser Pro Ser Thr Ser Ala Thr Ile Arg Lys Ile Tyr
130 135 140

Pro Ala Phe Met Leu Gly Ser Ser Val Gly Asn Phe Ser Asp Gly Lys
145 150 155 160

Met Gly Arg Phe Phe Asn His Thr Leu Val Leu Leu Pro Asp Gly Cys
165 170 175
Gly Thr Leu Leu Arg Ala Phe Tyr Cys Ile Leu Glu Pro Arg Ser Gly
180 185 190
Asn His Cys Pro Ala Gly Asn Ser Tyr Thr Ser Phe Ala Thr Tyr His
195 200 205
Thr Pro Ala Thr Asp Cys Ser Asp Gly Asn Tyr Asn Arg Asn Ala Ser
210 215 220
Leu Asn Ser Phe Lys Glu Tyr Phe Asn Leu Arg Asn Cys Thr Phe Met

225 230 235 240
Tyr Thr Tyr Asn Ile Thr Glu Asp Glu Ile Leu Glu Trp Phe Gly Ile
245 250 255
Thr Gln Thr Ala Gln Gly Val His Leu Phe Ser Ser Arg Tyr Val Asp
260 265 270
Leu Tyr Gly Gly Asn Met Phe Gln Phe Ala Thr Leu Pro Val Tyr Asp
275 280 285
Thr Ile Lys Tyr Tyr Ser Ile Ile Pro His Ser Ile Arg Ser Ile Gln
290 295 300

Ser Asp Arg Lys Ala Trp Ala Ala Phe Tyr Val Tyr Lys Leu Gln Pro
305 310 315 320
Leu Thr Phe Leu Leu Asp Phe Ser Val Asp Gly Tyr Ile Arg Arg Ala
325 330 335
Ile Asp Cys Gly Phe Asn Asp Leu Ser Gln Leu His Cys Ser Tyr Glu
340 345 350
Ser Phe Asp Val Glu Ser Gly Val Tyr Ser Val Ser Ser Phe Glu Ala
355 360 365
Lys Pro Ser Gly Ser Val Val Glu Gln Ala Glu Gly Val Glu Cys Asp

370 375 380
Phe Ser Pro Leu Leu Ser Gly Thr Pro Pro Gln Val Tyr Asn Phe Lys
385 390 395 400
Arg Leu Val Phe Thr Asn Cys Asn Tyr Asn Leu Thr Lys Leu Leu Ser
405 410 415

Leu Phe Ser Val Asn Asp Phe Thr Cys Ser Gln Ile Ser Pro Ala Ala
420 425 430
Ile Ala Ser Asn Cys Tyr Ser Ser Leu Ile Leu Asp Tyr Phe Ser Tyr
435 440 445

Pro Leu Ser Met Lys Ser Asp Leu Ser Val Ser Ser Ala Gly Pro Ile
450 455 460
Ser Gln Phe Asn Tyr Lys Gln Ser Phe Ser Asn Pro Thr Cys Leu Ile
465 470 475 480
Leu Ala Thr Val Pro His Asn Leu Thr Thr Ile Thr Lys Pro Leu Lys
485 490 495
Tyr Ser Tyr Ile Asn Lys Cys Ser Arg Leu Leu Ser Asp Asp Arg Thr
500 505 510
Glu Val Pro Gln Leu Val Asn Ala Asn Gln Tyr Ser Pro Cys Val Ser
515 520 525
Ile Val Pro Ser Thr Val Trp Glu Asp Gly Asp Tyr Tyr Arg Lys Gln
530 535 540
Leu Ser Pro Leu Glu Gly Gly Gly Trp Leu Val Ala Ser Gly Ser Thr
545 550 555 560
Val Ala Met Thr Glu Gln Leu Gln Met Gly Phe Gly Ile Thr Val Gln
565 570 575
Tyr Gly Thr Asp Thr Asn Ser Val Cys Pro Lys Leu Glu Phe Ala Asn
580 585 590

Asp Thr Lys Ile Ala Ser Gln Leu Gly Asn Cys Val Glu Tyr Ser Leu
595 600 605
Tyr Gly Val Ser Gly Arg Gly Val Phe Gln Asn Cys Thr Ala Val Gly
610 615 620
Val Arg Gln Gln Arg Phe Val Tyr Asp Ala Tyr Gln Asn Leu Val Gly
625 630 635 640
Tyr Tyr Ser Asp Asp Gly Asn Tyr Tyr Cys Leu Arg Ala Cys Val Ser
645 650 655
Val Pro Val Ser Val Ile Tyr Asp Lys Glu Thr Lys Thr His Ala Thr

660 665 670

Leu Phe Gly Ser Val Ala Cys Glu His Ile Ser Ser Thr Met Ser Gln

675 680 685

Tyr Ser Arg Ser Thr Arg Ser Met Leu Lys Arg Arg Asp Ser Thr Tyr

690 695 700

Gly Pro Leu Gln Thr Pro Val Gly Cys Val Leu Gly Leu Val Asn Ser

705 710 715 720

Ser Leu Phe Val Glu Asp Cys Lys Leu Pro Leu Gly Gln Ser Leu Cys

725 730 735

Ala Leu Pro Asp Thr Pro Ser Thr Leu Thr Pro Arg Ser Val Arg Ser

740 745 750

Val Pro Gly Glu Met Arg Leu Ala Ser Ile Ala Phe Asn His Pro Ile

755 760 765

Gln Val Asp Gln Leu Asn Ser Ser Tyr Phe Lys Leu Ser Ile Pro Thr

770 775 780

Asn Phe Ser Phe Gly Val Thr Gln Glu Tyr Ile Gln Thr Thr Ile Gln

785 790 795 800

Lys Val Thr Val Asp Cys Lys Gln Tyr Val Cys Asn Gly Phe Gln Lys

805 810 815

Cys Glu Gln Leu Leu Arg Glu Tyr Gly Gln Phe Cys Ser Lys Ile Asn

820 825 830

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

835 840 845

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

850 855 860

Phe Gly Gly Asp Phe Asn Leu Thr Leu Leu Glu Pro Val Ser Ile Ser

865 870 875 880

Thr Gly Ser Arg Ser Ala Arg Ser Ala Ile Glu Asp Leu Leu Phe Asp

885 890 895

Lys Val Thr Ile Ala Asp Pro Gly Tyr Met Gln Gly Tyr Asp Asp Cys

900 905 910

Met Gln Gln Gly Pro Ala Ser Ala Arg Asp Leu Ile Cys Ala Gln Tyr

915 920 925

Val Ala Gly Tyr Lys Val Leu Pro Pro Leu Met Asp Val Asn Met Glu

930 935 940

Ala Ala Tyr Thr Ser Ser Leu Leu Gly Ser Ile Ala Gly Val Gly Trp

945 950 955 960

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

965 970 975

Phe Tyr Arg Leu Asn Gly Val Gly Ile Thr Gln Gln Val Leu Ser Glu

980 985 990

Asn Gln Lys Leu Ile Ala Asn Lys Phe Asn Gln Ala Leu Gly Ala Met

995 1000 1005

Gln Thr Gly Phe Thr Thr Thr Asn Glu Ala Phe Gln Lys Val Gln Asp

1010 1015 1020

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

1025 1030 1035 1040

Ser Asn Thr Phe Gly Ala Ile Ser Ala Ser Ile Gly Asp Ile Ile Gln

1045 1050 1055

Arg Leu Asp Val Leu Glu Gln Asp Ala Gln Ile Asp Arg Leu Ile Asn

1060 1065 1070

Gly Arg Leu Thr Thr Leu Asn Ala Phe Val Ala Gln Gln Leu Val Arg

1075 1080 1085

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

1090 1095 1100

Glu Cys Val Lys Ala Gln Ser Lys Arg Ser Gly Phe Cys Gly Gln Gly

1105 1110 1115 1120

Thr His Ile Val Ser Phe Val Val Asn Ala Pro Asn Gly Leu Tyr Phe

1125 1130 1135

Met His Val Gly Tyr Tyr Pro Ser Asn His Ile Glu Val Val Ser Ala

1140 1145 1150

Tyr Gly Leu Cys Asp Ala Ala Asn Pro Thr Asn Cys Ile Ala Pro Val

1155 1160 1165

Asn Gly Tyr Phe Ile Lys Thr Asn Asn Thr Arg Ile Val Asp Glu Trp

1170 1175 1180

Ser Tyr Thr Gly Ser Ser Phe Tyr Ala Pro Glu Pro Ile Thr Ser Leu

1185 1190 1195 1200

Asn Thr Lys Tyr Val Ala Pro Gln Val Thr Tyr Gln Asn Ile Ser Thr

1205 1210 1215

Asn Leu Pro Pro Pro Leu Leu Gly Asn Ser Thr Gly Ile Asp Phe Gln

1220 1225 1230

Asp Glu Leu Asp Glu Phe Phe Lys Asn Val Ser Thr Ser Ile Pro Asn

1235 1240 1245

Phe Gly Ser Leu Thr Gln Ile Asn Thr Thr Leu Leu Asp Leu Thr Tyr

1250 1255 1260

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

1265 1270 1275 1280

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

1285 1290 1295

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Val Ala Leu Ala Leu

1300 1305 1310

Cys Val Phe Phe Ile Leu Cys Cys Thr Gly Cys Gly Thr Asn Cys Met

1315 1320 1325

Gly Lys Leu Lys Cys Asn Arg Cys Cys Asp Arg Tyr Glu Glu Tyr Asp

1330 1335 1340

Leu Glu Pro His Lys Val His Val His

1345 1350

<210> 11

<211> 4062

<212> DNA

<213> Middle East respiratory syndrome-related coronavirus

<400> 11

atgatacact cagtgtttct actgatgttc ttgttaacac ctacagaaag ttacgttgat 60

gtagggccag attctgttaa gtctgcttgt attgaggttg atatacaaca gactttcttt 120

gataaaactt ggcctaggcc aattgatgtt tctaaggctg acggtattat ataccctcaa 180

ggccgtacat attctaacat aactatcact tatcaaggtc tttttcccta tcagggagac 240

catggtgata tgtatgttta ctctgcagga catgctacag gcacaactcc acaaaagttg 300

tttgtagcta actattctca ggacgtcaaa cagtttgcta atgggtttgt cgtccgtata 360

ggagcagctg ccaattccac tggcactgtt attattagcc catctaccag cgctactata 420

cgaaaaattt accctgcttt tatgctgggt tcttcagttg gtaatttctc agatggtaaa 480

atgggccgct tcttcaatca tactctagtt cttttgcccg atggatgtgg cactttactt 540

agagcttttt attgtattct agagcctcgc tctggaaatc attgtcctgc tggcaattcc 600

tatacttctt ttgccactta tcacactcct gcaacagatt gttctgatgg caattacaat 660

cgtaatgcca gtctgaactc ttttaaggag tattttaatt tacgtaactg cacctttatg 720

tacacttata acattaccga agatgagatt ttagagtggt ttggcattac acaaactgct 780

caaggtgttc acctcttctc atctcggtat gttgatttgt acggcggcaa tatgtttcaa 840

tttgccacct tgcctgttta tgatactatt aagtattatt ctatcattcc tcacagtatt 900

cgttctatcc aaagtgatag aaaagcttgg gctgccttct acgtatataa acttcaaccg 960

ttaactttcc tgttggattt ttctgttgat ggttatatac gcagagctat agactgtggt 1020

tttaatgatt tgtcacaact ccactgctca tatgaatcct tcgatgttga atctggagtt 1080

tattcagttt cgtctttcga agcaaaacct tctggctcag ttgtggaaca ggctgaaggt 1140

gttgaatgtg atttttcacc tcttctgtct ggcacacctc ctcaggttta taatttcaag 1200

cgtttggttt ttaccaattg caattataat cttaccaaat tgctttcact tttttctgtg 1260

aatgatttta cttgtagtca aatatctcca gcagcaattg ctagcaactg ttattcttca 1320

ctgattttgg attacttttc atacccactt agtatgaaat ccgatctcag tgttagttct 1380

gctggtccaa tatcccagtt taattataaa cagtcctttt ctaatcccac atgtttgatt 1440

ttagcgactg ttcctcataa ccttactact attactaagc ctcttaagta cagctatatt 1500

aacaagtgct ctcgtcttct ttctgatgat cgtactgaag tacctcagtt agtgaacgct 1560

aatcaatact caccctgtgt atccattgtc ccatccactg tgtgggaaga cggtgattat 1620

tataggaaac aactatctcc acttgaaggt ggtggctggc ttgttgctag tggctcaact 1680

gttgccatga ctgagcaatt acagatgggc tttggtatta cagttcaata tggtacagac 1740

accaatagtg tttgccccaa gcttgaattt gctaatgaca caaaaattgc ctctcaatta 1800

ggcaattgcg tggaatattc cctctatggt gtttcgggcc gtggtgtttt tcagaattgc 1860

acagctgtag gtgttcgaca gcagcgcttt gtttatgatg cgtaccagaa tttagttggc 1920

tattattctg atgatggcaa ctactactgt ttgcgtgctt gtgttagtgt tcctgtttct 1980

gtcatctatg ataaagaaac taaaacccac gctactctat ttggtagtgt tgcatgtgaa 2040

cacatttctt ctaccatgtc tcaatactcc cgttctacgc gatcaatgct taaacggcga 2100

gattctacat atggcccct tcagacacct gttggttgtg tcctaggact tgttaattcc 2160

tctttgttcg tagaggactg caagttgcct cttggtcaat ctctctgtgc tcttcctgac 2220

acacctagta ctctcacacc tcgcagtgtg cgctctgttc caggtgaaat gcgcttggca 2280

tccattgctt ttaatcatcc tattcaggtt gatcaactta atagtagtta ttttaaatta 2340

agtataccca ctaatttttc ctttggtgtg actcaggagt acattcagac aaccattcag 2400

aaagttactg ttgattgtaa acagtacgtt tgcaatggtt tccagaagtg tgagcaatta 2460

ctgcgcgagt atggccagtt ttgttccaaa ataaaccagg ctctccatgg tgccaattta 2520

cgccaggatg attctgtacg taatttgttt gcgagcgtga aaagctctca atcatctcct 2580

atcataccag gttttggagg tgactttaat ttgacacttc tagaacctgt ttctatatct 2640

actggcagtc gtagtgcacg tagtgctatt gaggatttgc tatttgacaa agtcactata 2700

gctgatcctg gttatatgca aggttacgat gattgcatgc agcaaggtcc agcatcagct 2760

cgtgatctta tttgtgctca atatgtggct ggttacaaag tattacctcc tcttatggat 2820

gttaatatgg aagccgcgta tacttcatct ttgcttggca gcatagcagg tgttggctgg 2880

actgctggct tatcctcctt tgctgctatt ccatttgcac agagtatctt ttataggtta 2940

aacggtgttg gcattactca acaggttctt tcagagaacc aaaagcttat tgccaataag 3000

tttaatcagg ctctgggagc tatgcaaaca ggcttcacta caactaatga agcttttcag 3060

aaggttcagg atgctgtgaa caacaatgca caggctctat ccaaattagc tagcgagcta 3120

tctaatactt ttggtgctat ttccgcctct attggagaca tcatacaacg tcttgatgtt 3180

ctcgaacagg acgcccaaat agacagactt attaatggcc gtttgacaac actaaatgct 3240

tttgttgcac agcagcttgt tcgttccgaa tcagctgctc tttccgctca attggctaaa 3300

gataaagtca atgagtgtgt caaggcacaa tccaagcgtt ctggattttg cggtcaaggc 3360

acacatatag tgtcctttgt tgtaaatgcc cctaatggcc tttacttcat gcatgttggt 3420

tattacccta gcaaccacat tgaggttgtt tctgcttatg gtctttgcga tgcagctaac 3480

cctactaatt gtatagcccc tgttaatggc tactttatta aaactaataa cactaggatt 3540

gttgatgagt ggtcatatac tggctcgtcc ttctatgcac ctgagcccat tacctccctt 3600

aatactaagt atgttgcacc acaggtgaca taccaaaaca tttctactaa cctccctcct 3660

cctcttctcg gcaattccac cgggattgac ttccaagatg agttggatga gtttttcaaa 3720

aatgttagca ccagtatacc taattttggt tccctaacac agattaatac tacattactc 3780

gatcttacct acgagatgtt gtctcttcaa caagttgtta aagcccttaa tgagtcttac 3840

atagaccttа aagagcttgg caattatact tattacaaca aatggccgtg gtacatttgg 3900

cttggtttca ttgctgggct tgttgcctta gctctatgcg tcttcttcat actgtgctgc 3960

actggttgtg gcacaaactg tatgggaaaa cttaagtgta atcgttgttg tgatagatac 4020

gaggaatacg acctcgagcc gcataaggtt catgttcact aa 4062

<210> 12

<211> 4062

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 12

atgatccaca gcgtgttcct gctgatgttt ctgctgacac ctaccgagtc ctacgtggat 60

gtgggcccag actctgtgaa gagcgcctgc atcgaggtgg acatccagca gacattcttt 120

gacaagacct ggcccagacc catcgacgtg agcaaggcag acggaatcat ctacccacag 180

ggacgcacat atagcaacat cacaatcacc taccagggcc tgttccctta tcagggcgac 240

cacggcgata tgtacgtgta tagcgccggc cacgcaaccg gcaccacacc acagaagctg 300

tttgtggcca attattccca ggacgtgaag cagttcgcca acggatttgt ggtgcggatc 360

ggggccgccg ccaacagcac aggcaccgtg atcatctctc ccagcacatc cgccaccatc 420

agaaagatct accctgcctt tatgctgggc agctccgtgg gcaacttctc cgatggcaag 480

atgggcaggt tctttaatca cacactggtg ctgctgccag acggatgcgg caccctgctg 540

agggccttct actgtatcct ggagccccgc tctggaaatc actgccctgc cggcaactcc 600

tacacctctt ttgccacata tcacacccct gccacagact gttccgatgg caattataac 660

cggaatgcca gcctgaactc cttcaaggag tactttaatc tgagaaactg caccttcatg 720

tacacatata atatcaccga ggatgagatc ctggagtggt tcggcatcac acagaccgcc 780

cagggcgtgc acctgttttc tagcagatac gtggatctgt atggcggcaa catgttccag 840

tttgccacac tgccagtgta tgacaccatc aagtactata gcatcatccc ccactctatc 900

cggagcatcc agtccgacag aaaggcctgg gccgccttct acgtgtataa gctgcagccc 960

ctgaccttcc tgctggattt ttccgtggac ggctacatcc ggagagccat cgattgcggc 1020

tttaacgacc tgtctcagct ccactgttct tatgagagct tcgatgtgga gtctggcgtg 1080

tacagcgtgt cctcttttga ggccaagcca tctggcagcg tggtggagca ggcagaggga 1140

gtggagtgcg acttctcccc actgctgtct ggcacaccac ctcaggtgta taatttcaag 1200

aggctggtgt ttacaaactg taattacaac ctgaccaagc tgctgtccct gttctctgtg 1260

aacgacttta cctgcagcca gatctcccct gccgccatcg cctccaattg ttatagctcc 1320

ctgatcctgg attacttctc ttatcccctg tctatgaaga gcgacctgtc cgtgtctagc 1380

gccggcccta tcagccagtt taattacaag cagtccttct ctaaccccac atgcctgatc 1440

ctggccaccg tgcctcacaa cctgaccaca atcacaaagc cactgaagta ctcctatatc 1500

aataagtgca gcaggctgct gtccgacgat cgcaccgagg tgcctcagct cgtgaacgcc 1560

aaccagtact ctccatgcgt gagcatcgtg ccatccaccg tgtgggagga cggcgattac 1620

tatagaaagc agctcagccc actggaggga ggaggatggc tggtggccag cggctccaca 1680

gtggccatga ccgagcagct ccagatgggc ttcggcatca cagtgcagta cggcacagat 1740

accaatagcg tgtgccccaa gctggagttt gccaacgaca ccaagatcgc ctcccagctc 1800

ggcaattgcg tggagtactc cctgtatggc gtgtctggca gaggcgtgtt ccagaactgt 1860

acagccgtgg gcgtgcggca gcagcggttc gtgtacgatg cctatcagaa cctggtgggc 1920

tactatagcg acgatggcaa ttactattgc ctgagggcat gcgtgagcgt gcccgtgagc 1980

gtgatctacg acaaggagac aaagacccac gccaccctgt tcggctccgt ggcctgcgag 2040

cacatctcct ctacaatgtc tcagtattct aggagcaccc gctctatgct gaagaggcgc 2100

gacagcacat acggaccact gcagacccct gtgggatgcg tgctgggcct ggtgaacagc 2160

agcctgtttg tggaggattg caagctgcca ctgggccagt ctctgtgcgc actgccagac 2220

acccccagca cactgacccc acggtctgtg agaagcgtgc ccggagagat gagactggcc 2280

agcatcgcct tcaatcaccc tatccaggtg gatcagctca acagcagcta ctttaagctg 2340

agcatcccaa caaacttctc ctttggcgtg acccaggagt atatccagac cacaatccag 2400

aaggtgaccg tggactgcaa gcagtacgtg tgcaatggct tccagaagtg cgagcagctc 2460

ctgagggagt atggccagtt ttgttccaag atcaatcagg ccctgcacgg agccaacctg 2520

aggcaggacg attccgtgag aaacctgttc gcctctgtga agtcctctca gagctccgct 2580

atcatcccag gcttcggcgg cgacttcaac ctgaccctgc tggagcccgt gtccatctct 2640

accggcagca ggtccgcccg cagcgccatc gaggatctgc tgtttgacaa ggtgaccatc 2700

gccgacccag gctacatgca gggctatgac gattgcatgc agcagggacc agcctccgcc 2760

cgcgatctga tctgtgccca gtacgtggcc ggctataagg tgctgccacc cctgatggac 2820

gtgaacatgg aggccgccta tacatctagc ctgctgggca gcatcgcagg agtgggatgg 2880

accgccggcc tgtcctcttt cgccgcaatc ccttttgccc agtctatctt ctaccggctg 2940

aacggcgtgg gcatcacaca gcaggtgctg agcgagaatc agaagctgat cgccaataag 3000

ttcaaccagg ccctgggggc catgcagacc ggctttacca caaccaacga ggccttccag 3060

aaggtgcagg atgccgtgaa caataacgca caggccctgt ccaagctggc ctccgagctg 3120

tctaatacct tcggggccat cagcgccagc atcggcgaca tcatccagcg cctggacgtg 3180

ctggagcagg atgcccagat cgacaggctg atcaatggcc gcctgacaac cctgaacgcc 3240

tttgtggcac agcagctcgt gcggagcgag tctgccgccc tgagcgccca gctcgccaag 3300

gacaaggtga acgagtgcgt gaaggcccag agcaagcggt ccggcttttg tggccagggc 3360

acccacatcg tgtccttcgt ggtgaatgcc cctaacggcc tgtactttat gcatgtgggc 3420

tactatccaa gcaaccacat cgaggtggtg tccgcctatg gcctgtgcga tgccgccaat 3480

cctacaaact gtatcgcccc agtgaatggc tacttcatca agaccaataa cacacggatc 3540

gtggacgagt ggtcctacac cggcagctcc ttttatgccc ccgagcctat cacatctctg 3600

aacaccaagt acgtggcccc acaggtgaca tatcagaata tcagcaccaa cctgcctcca 3660

cccctgctgg gcaattccac cggcatcgac ttccaggatg agctggacga gttctttaag 3720

aatgtgagca catccatccc caactttggc agcctgaccc agatcaacac aaccctgctg 3780

gatctgacat acgagatgct gtctctgcag caggtggtga aggccctgaa tgagagctac 3840

atcgacctga aggagctggg caattatacc tactataaca agtggccttg gtacatctgg 3900

ctgggcttca tcgcaggcct ggtggccctg gccctgtgcg tgttctttat cctgtgctgt 3960

acaggctgcg gcaccaattg tatgggcaag ctgaagtgta accggtgctg tgatagatac 4020

gaggagtatg acctggagcc acacaaggtg catgtgcact ga 4062

<210> 13

<211> 1273

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 13

Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val

1 5 10 15

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

20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu

35 40 45

```
His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp
     50                  55                  60
Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
 65                  70                  75                  80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
                 85                  90                  95
Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
            100                 105                 110
Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
        115                 120                 125
Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
    130                 135                 140
Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr

145                 150                 155                 160
Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
                165                 170                 175
Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
            180                 185                 190
Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
        195                 200                 205
Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
    210                 215                 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225                 230                 235                 240
Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
                245                 250                 255
Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
            260                 265                 270
Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
        275                 280                 285
Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
```

290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val

305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys

325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala

340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu

355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro

370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe

385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly

405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys

420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn

435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe

450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys

465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly

485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val

500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys

515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn

530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu

545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val

565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe

580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val

595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile

610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser

625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val

645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala

660 665 670

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

675 680 685

Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser

690 695 700

Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile

705 710 715 720

Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val

725 730 735

Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu

740 745 750

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

755 760 765

Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln

770 775 780

Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe

785 790 795 800

Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
805 810 815
Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly
820 825 830
Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp
835 840 845
Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu
850 855 860
Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly

865 870 875 880
Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
885 890 895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
900 905 910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn
915 920 925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala
930 935 940

Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn
945 950 955 960
Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val
965 970 975
Leu Asn Asp Ile Leu Ser Arg Leu Asp Pro Pro Glu Ala Glu Val Gln
980 985 990
Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val
995 1000 1005
Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu

1010 1015 1020
Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val
1025 1030 1035 1040
Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala
1045 1050 1055

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu
1060 1065 1070
Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His
1075 1080 1085

Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1090 1095 1100
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr
1105 1110 1115 1120
Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr
1125 1130 1135
Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu
1140 1145 1150
Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp

1155 1160 1165
Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp
1170 1175 1180
Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu
1185 1190 1195 1200
Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile
1205 1210 1215
Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile
1220 1225 1230

Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys
1235 1240 1245
Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val
1250 1255 1260
Leu Lys Gly Val Lys Leu His Tyr Thr
1265 1270

<210> 14
<211> 3822
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 14

atgtttgttt ttcttgtttt attgccacta gtctctagtc agtgtgttaa tcttacaacc 60

agaactcaat taccccctgc atacactaat tctttcacac gtggtgttta ttaccctgac 120

aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc tttcttttcc 180

aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat 240

aaccctgtcc taccatttaa tgatggtgtt tattttgctt ccactgagaa gtctaacata 300

ataagaggct ggatttttgg tactacttta gattcgaaga cccagtccct acttattgtt 360

aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420

ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480

tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggaccttgaa 540

ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat tgatggttat 600

tttaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660

tcggctttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720

ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg ttggacagct 780

ggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840

gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaaag 900

tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960

caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc ttttggtgaa 1020

gtttttaacg ccaccagatt tgcatctgtt tatgcttgga acaggaagag aatcagcaac 1080

tgtgttgctg attattctgt cctatataat tccgcatcat tttccacttt taagtgttat 1140

ggagtgtctc ctactaaatt aaatgatctc tgctttacta atgtctatgc agattcattt 1200

gtaattagag gtgatgaagt cagacaaatc gctccagggc aaactggaaa gattgctgat 1260

tataattata aattaccaga tgattttaca ggctgcgtta tagcttggaa ttctaacaat 1320

cttgattcta aggttggtgg taattataat tacctgtata gattgtttag gaagtctaat 1380

ctcaaacctt ttgagagaga tatttcaact gaaatctatc aggccggtag cacaccttgt 1440

aatggtgttg aaggttttaa ttgttacttt cctttacaat catatggttt ccaacccact 1500

aatggtgttg gttaccaacc atacagagta gtagtacttt cttttgaact tctacatgca 1560

ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa atgtgtcaat 1620

ttcaacttca atggtttaac aggcacaggt gttcttactg agtctaacaa aaagtttctg 1680

cctttccaac aatttggcag agacattgct gacactactg atgctgtccg tgatccacag 1740

acacttgaga ttcttgacat tacaccatgt tcttttggtg gtgtcagtgt tataacacca 1800

ggaacaaata cttctaacca ggttgctgtt ctttatcagg atgttaactg cacagaagtc 1860

cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc tacaggttct 1920

aatgtttttc aaacacgtgc aggctgttta ataggggctg aacatgtcaa caactcatat 1980

gagtgtgaca tacccattgg tgcaggtata tgcgctagtt atcagactca gactaattct 2040

cctcggcggg cacgtagtgt agctagtcaa tccatcattg cctacactat gtcacttggt 2100

gcagaaaatt cagttgctta ctctaataac tctattgcca tacccacaaa ttttactatt 2160

agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga ttgtacaatg 2220

tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg cagtttttgt 2280

acacaattaa accgtgcttt aactggaata gctgttgaac aagacaaaaa cacccaagaa 2340

gtttttgcac aagtcaaaca aatttacaaa acaccaccaa ttaaagattt tggtggtttt 2400

aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt tattgaagat 2460

ctacttttca acaaagtgac acttgcagat gctggcttca tcaaacaata tggtgattgc 2520

cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg ccttactgtt 2580

ttgccacctt tgctcacaga tgaaatgatt gctcaataca cttctgcact gttagcgggt 2640

acaatcactt ctggttggac ctttggtgca ggtgctgcat tacaaatacc atttgctatg 2700

caaatggctt ataggtttaa tggtattgga gttacacaga atgttctcta tgagaaccaa 2760

aaattgattg ccaaccaatt taatagtgct attggcaaaa ttcaagactc actttcttcc 2820

acagcaagtg cacttggaaa acttcaagat gtggtcaacc aaaatgcaca agctttaaac 2880

acgcttgtta aacaacttag ctccaatttt ggtgcaattt caagtgtttt aaatgatatc 2940

ctttcacgtc ttgaccctcc tgaggctgaa gtgcaaattg ataggttgat cacaggcaga 3000

cttcaaagtt tgcagacata tgtgactcaa caattaatta gagctgcaga aatcagagct 3060

tctgctaatc ttgctgctac taaaatgtca gagtgtgtac ttggacaatc aaaaagagtt 3120

gatttttgtg gaaagggcta tcatcttatg tccttccctc agtcagcacc tcatggtgta 3180

gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac tgctcctgcc 3240

atttgtcatg atggaaaagc acactttcct cgtgaaggtg tctttgtttc aaatggcaca 3300

cactggtttg taacacaaag gaatttttat gaaccacaaa tcattactac agacaacaca 3360

tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt ttatgatcct 3420

ttgcaacctg aattagactc attcaaggag gagttagata aatattttaa gaatcataca 3480

tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt aaacattcaa 3540

aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct catcgatctc 3600

caagaacttg gaaagtatga gcagtatata aaatggccat ggtacatttg gctaggtttt 3660

atagctggct tgattgccat agtaatggtg acaattatgc tttgctgtat gaccagttgc 3720

tgtagttgtc tcaagggctg ttgttcttgt ggatcctgct gcaaatttga tgaagacgac 3780

tctgagccag tgctcaaagg agtcaaatta cattacacat aa 3822

<210> 15

<211> 3822

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 15

atgttcgtgt ttctggtgct gctgcctctg gtgagctccc agtgcgtgaa cctgaccaca 60

aggacccagc tccccccctgc ctataccaat tccttcacac ggggcgtgta ctatccagac 120

aaggtgttta gatctagcgt gctgcactcc acacaggatc tgtttctgcc cttcttttct 180

aacgtgacct ggttccacgc catccatgtg agcggcacca atggcacaaa gcggttcgac 240

aatcctgtgc tgcccttcaa cgatggcgtg tacttcgcct ccaccgagaa gtctaacatc 300

atcagaggct ggatctttgg caccacactg gacagcaaga cacagtccct gctgatcgtg 360

aacaatgcca ccaacgtggt catcaaggtg tgcgagttcc agttttgtaa tgatcctttc 420

ctgggcgtgt actatcacaa gaacaataag tcttggatgg agagcgagtt tcgcgtgtat 480

tcctctgcca acaattgcac atttgagtac gtgtcccagc cattcctgat ggacctggag 540

ggcaagcagg gcaatttcaa gaacctgagg gagttcgtgt ttaagaatat cgatggctac 600

ttcaagatct actccaagca cacccctatc aacctggtgc gcgacctgcc acagggcttc 660

tctgccctgg agcctctggt ggatctgcca atcggcatca acatcacccg gtttcagaca 720

ctgctggccc tgcacagaag ctacctgaca cctggcgaca gctcctctgg atggaccgcc 780

ggggccgccg cctactatgt gggctatctg cagccaagga ccttcctgct gaagtacaac 840

gagaatggca ccatcacaga cgcagtggat tgcgccctgg acccccctgtc tgagaccaag 900

tgtacactga agagctttac cgtggagaag ggcatctatc agacaagcaa tttcagggtg 960

cagcccaccg agtccatcgt gcgctttcca aatatcacaa acctgtgccc ctttggcgag 1020

gtgttcaacg caaccaggtt cgccagcgtg tacgcatgga ataggaagcg catctccaac 1080

tgcgtggccg actattctgt gctgtacaac agcgcctcct tctctacctt taagtgctat 1140

ggcgtgagcc ccacaaagct gaatgacctg tgctttacca acgtgtacgc cgattccttc 1200
gtgatcaggg gcgacgaggt gcgccagatc gcaccaggac agacaggcaa gatcgccgac 1260
tacaattata agctgcccga cgatttcacc ggctgcgtga tcgcctggaa ctctaacaat 1320
ctggatagca aagtgggcgg caactacaat tatctgtacc ggctgtttag aaagtctaat 1380
ctgaagcctt tcgagaggga catctccaca gagatctacc aggccggctc taccccatgc 1440
aatggcgtgg agggctttaa ctgttatttc cccctgcaga gctacggctt ccagcctaca 1500
aacggcgtgg gctatcagcc ataccgcgtg gtggtgctgt cttttgagct gctgcacgca 1560

ccagcaacag tgtgcggacc taagaagagc accaatctgg tgaagaacaa gtgcgtgaac 1620
ttcaacttca acggcctgac cggcacaggc gtgctgaccg agtccaacaa gaagttcctg 1680
ccctttcagc agttcggcag ggacatcgca gataccacag acgccgtgcg cgacccccag 1740
accctggaga tcctggacat cacaccttgc tccttcggcg gcgtgtctgt gatcacacct 1800
ggcaccaata caagcaacca ggtggccgtg ctgtatcagg acgtgaattg taccgaggtg 1860
ccagtggcca tccacgccga tcagctcacc cccacatggc gggtgtactc taccggcagc 1920
aacgtgttcc agacaagagc cggctgcctg atcggagccg agcatgtgaa caatagctat 1980

gagtgcgaca tccccatcgg agccggcatc tgtgcctcct accagaccca gacaaactcc 2040
cctcggagag cccggtctgt ggccagccag tccatcatcg cctataccat gagcctgggg 2100
gccgagaaca gcgtggccta ctccaacaat tctatcgcca tccccaccaa cttcacaatc 2160
tccgtgacca cagagatcct gcctgtgagc atgaccaaga catccgtgga ctgcacaatg 2220
tatatctgtg gcgattccac cgagtgctct aacctgctgc tgcagtacgg ctctttttgt 2280
acccagctca acagagccct gacaggcatc gccgtggagc aggacaagaa cacacaggag 2340
gtgttcgccc aggtgaagca gatctacaag accccaccca tcaaggactt tggcggcttc 2400

aacttcagcc agatcctgcc agatcccagc aagccttcca agcggtcttt tatcgaggac 2460
ctgctgttca acaaggtgac cctggccgat gccggcttca tcaagcagta tggcgattgc 2520
ctgggcgaca tcgccgccag agacctgatc tgtgcccaga agtttaatgg cctgaccgtg 2580
ctgcctccac tgctgacaga tgagatgatc gcccagtaca catctgccct gctggccggc 2640
accatcacaa gcggatggac cttcggggcc ggggccgccc tgcagatccc atttgccatg 2700
cagatggcct atcggttcaa cggcatcggc gtgacccaga atgtgctgta cgagaaccag 2760
aagctgatcg ccaatcagtt taactccgcc atcggcaaga tccaggactc tctgagctcc 2820

acagccagcg ccctgggcaa gctgcaggat gtggtgaatc agaacgccca ggccctgaat 2880

accctggtga agcagctcag cagcaacttc ggggccatca gcagcgtgct gaacgacatc 2940

ctgagccggc tggacccccc tgaggcagag gtgcagatcg accggctgat cacaggcaga 3000

ctgcagtccc tgcagaccta cgtgacacag cagctcatca gggccgccga gatcagggcc 3060

tctgccaatc tggccgccac caagatgagc gagtgcgtgc tgggccagtc caagagagtg 3120

gacttttgtg gcaagggcta tcacctgatg agcttcccac agtccgcccc ccacggagtg 3180

gtgtttctgc atgtgaccta cgtgcctgcc caggagaaga acttcaccac agccccagcc 3240

atctgccacg atggcaaggc ccactttccc agggagggcg tgttcgtgag caacggcacc 3300

cactggtttg tgacacagcg caatttctac gagcctcaga tcatcaccac agacaataca 3360

ttcgtgtccg gcaactgtga cgtggtcatc ggcatcgtga acaataccgt gtatgatccc 3420

ctgcagcctg agctggactc ttttaaggag gagctggata agtacttcaa gaatcacacc 3480

agccccgacg tggatctggg cgacatctct ggcatcaatg ccagcgtggt gaacatccag 3540

aaggagatcg acaggctgaa cgaggtggcc aagaatctga acgagtccct gatcgatctg 3600

caggagctgg gcaagtatga gcagtacatc aagtggccat ggtatatctg gctgggcttc 3660

atcgccggcc tgatcgccat cgtgatggtg accatcatgc tgtgctgtat gacaagctgc 3720

tgttcctgcc tgaagggctg ctgttcttgt ggctcctgct gtaagtttga tgaggacgat 3780

agcgagcccg tgctgaaggg cgtgaagctg cactacacct ga 3822

<210> 16

<211> 1273

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 16

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

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

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

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp

65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu

85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser

100 105 110

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

115 120 125

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

130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr

145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu

165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe

180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr

195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu

210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys

290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val

305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335
Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350
Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365
Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro

370 375 380
Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400
Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415
Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430
Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460
Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480
Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495
Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510
Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys

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

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser Pro Gly Ser Ala Ser Ser Val Ala
675 680 685

Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
690 695 700

Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile
705 710 715 720

Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val
725 730 735

Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu
740 745 750

Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr
755 760 765

Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln
770 775 780

Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe
785 790 795 800

Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
805 810 815

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

820 825 830

Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp

835 840 845

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

850 855 860

Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly

865 870 875 880

Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile

885 890 895

Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr

900 905 910

Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn

915 920 925

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

930 935 940

Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn

945 950 955 960

Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val

965 970 975

Leu Asn Asp Ile Leu Ser Arg Leu Asp 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 Leu

1010 1015 1020

Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val

1025 1030 1035 1040

Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala

1045 1050 1055

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu

1060 1065 1070

Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His

1075 1080 1085

Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val

1090 1095 1100

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr

1105 1110 1115 1120

Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr

1125 1130 1135

Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu

1140 1145 1150

Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp

1155 1160 1165

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

1170 1175 1180

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

1185 1190 1195 1200

Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile

1205 1210 1215

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

1220 1225 1230

Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys

1235 1240 1245

Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val

1250 1255 1260

Leu Lys Gly Val Lys Leu His Tyr Thr

1265 1270

<210> 17

<211> 3822

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 17

atgtttgttt ttcttgtttt attgccacta gtctctagtc agtgtgttaa tcttacaacc 60

agaactcaat tacccccctgc atacactaat tctttcacac gtggtgttta ttaccctgac 120

aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc tttcttttcc 180

aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat 240

aaccctgtcc taccatttaa tgatggtgtt tattttgctt ccactgagaa gtctaacata 300

ataagaggct ggatttttgg tactacttta gattcgaaga cccagtccct acttattgtt 360

aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420

ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480

tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggaccttgaa 540

ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat tgatggttat 600

tttaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660

tcggctttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720

ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg ttggacagct 780

ggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840

gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaaag 900

tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960

caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc ttttggtgaa 1020

gtttttaacg ccaccagatt tgcatctgtt tatgcttgga acaggaagag aatcagcaac 1080

tgtgttgctg attattctgt cctatataat tccgcatcat tttccacttt taagtgttat 1140

ggagtgtctc ctactaaatt aaatgatctc tgctttacta atgtctatgc agattcattt 1200

gtaattagag gtgatgaagt cagacaaatc gctccagggc aaactggaaa gattgctgat 1260

tataattata aattaccaga tgattttaca ggctgcgtta tagcttggaa ttctaacaat 1320

cttgattcta aggttggtgg taattataat tacctgtata gattgtttag gaagtctaat 1380

ctcaaacctt ttgagagaga tatttcaact gaaatctatc aggccggtag cacaccttgt 1440

aatggtgttg aaggttttaa ttgttacttt cctttacaat catatggttt ccaacccact 1500

aatggtgttg gttaccaacc atacagagta gtagtacttt cttttgaact tctacatgca 1560

ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa atgtgtcaat 1620

ttcaacttca atggtttaac aggcacaggt gttcttactg agtctaacaa aaagtttctg 1680

cctttccaac aatttggcag agacattgct gacactactg atgctgtccg tgatccacag 1740

acacttgaga ttcttgacat tacaccatgt tcttttggtg gtgtcagtgt tataacacca 1800

ggaacaaata cttctaacca ggttgctgtt ctttatcagg atgttaactg cacagaagtc 1860

cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc tacaggttct 1920

aatgtttttc aaacacgtgc aggctgttta ataggggctg aacatgtcaa caactcatat 1980

gagtgtgaca tacccattgg tgcaggtata tgcgctagtt atcagactca gactaattct 2040

cctggtagtg caagtagtgt agctagtcaa tccatcattg cctacactat gtcacttggt 2100

gcagaaaatt cagttgctta ctctaataac tctattgcca tacccacaaa ttttactatt 2160

agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga ttgtacaatg 2220

tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg cagtttttgt 2280

acacaattaa accgtgcttt aactggaata gctgttgaac aagacaaaaa cacccaagaa 2340

gtttttgcac aagtcaaaca aatttacaaa acaccaccaa ttaaagattt tggtggtttt 2400

aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt tattgaagat 2460

ctacttttca acaaagtgac acttgcagat gctggcttca tcaaacaata tggtgattgc 2520

cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg ccttactgtt 2580

ttgccacctt tgctcacaga tgaaatgatt gctcaataca cttctgcact gttagcgggt 2640

acaatcactt ctggttggac ctttggtgca ggtgctgcat tacaaatacc atttgctatg 2700

caaatggctt ataggtttaa tggtattgga gttacacaga atgttctcta tgagaaccaa 2760

aaattgattg ccaaccaatt taatagtgct attggcaaaa ttcaagactc actttcttcc 2820

acagcaagtg cacttggaaa acttcaagat gtggtcaacc aaaatgcaca agctttaaac 2880

acgcttgtta aacaacttag ctccaatttt ggtgcaattt caagtgtttt aaatgatatc 2940

ctttcacgtc ttgacaaagt tgaggctgaa gtgcaaattg ataggttgat cacaggcaga 3000

cttcaaagtt tgcagacata tgtgactcaa caattaatta gagctgcaga aatcagagct 3060

tctgctaatc ttgctgctac taaaatgtca gagtgtgtac ttggacaatc aaaaagagtt 3120

gatttttgtg gaaagggcta tcatcttatg tccttccctc agtcagcacc tcatggtgta 3180

gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac tgctcctgcc 3240

atttgtcatg atggaaaagc acactttcct cgtgaaggtg tctttgtttc aaatggcaca 3300

cactggtttg taacacaaag gaatttttat gaaccacaaa tcattactac agacaacaca 3360

tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt ttatgatcct 3420

ttgcaacctg aattagactc attcaaggag gagttagata aatattttaa gaatcataca 3480

tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt aaacattcaa 3540

aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct catcgatctc 3600

caagaacttg gaaagtatga gcagtatata aaatggccat ggtacatttg gctaggtttt 3660

atagctggct tgattgccat agtaatggtg acaattatgc tttgctgtat gaccagttgc 3720

tgtagttgtc tcaagggctg ttgttcttgt ggatcctgct gcaaatttga tgaagacgac 3780

tctgagccag tgctcaaagg agtcaaatta cattacacat aa 3822

<210> 18

<211> 3822

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 18

atgttcgtgt ttctggtgct gctgcctctg gtgagctccc agtgcgtgaa cctgaccaca 60

aggacccagc tccccccctgc ctataccaat tccttcacaa ggggcgtgta ctatcccgac 120

aaggtgtttc gctctagcgt gctgcacagc acacaggatc tgtttctgcc tttcttttcc 180

aacgtgacct ggttccacgc catccatgtg agcggcacca atggcacaaa gaggttcgac 240

aatccagtgc tgccctttaa cgatggcgtg tacttcgcct ctaccgagaa gagcaacatc 300

atccgcggct ggatctttgg caccacactg gactccaaga cacagtctct gctgatcgtg 360

aacaatgcca ccaacgtggt catcaaggtg tgcgagttcc agttttgtaa tgatccattc 420

ctgggcgtgt actatcacaa gaacaataag agctggatgg agtccgagtt tcgcgtgtat 480

tcctctgcca acaattgcac atttgagtac gtgtcccagc ccttcctgat ggacctggag 540

ggcaagcagg gcaatttcaa gaacctgcgg gagttcgtgt ttaagaatat cgatggctac 600

ttcaagatct acagcaagca caccccaatc aacctggtga gagacctgcc acagggcttc 660

tccgccctgg agccactggt ggatctgccc atcggcatca acatcaccag gtttcagaca 720

ctgctggccc tgcaccgcag ctacctgaca ccaggcgaca gctcctctgg atggaccgcc 780

ggggccgccg cctactatgt gggctatctg cagccccgga ccttcctgct gaagtacaac 840

gagaatggca ccatcacaga cgcagtggat tgcgccctgg acccccctgtc cgagaccaag 900

tgtacactga agtcttttac cgtggagaag ggcatctatc agacatctaa tttccgggtg 960

cagcctaccg agagcatcgt gagatttccc aatatcacaa acctgtgccc ttttggcgag 1020

gtgttcaacg ccaccagatt cgccagcgtg tacgcctgga atcggaagag aatcagcaac 1080

tgcgtggccg actattccgt gctgtacaac tctgccagct tctccacctt taagtgctat 1140

ggcgtgtctc ccacaaagct gaatgacctg tgctttacca acgtgtacgc cgatagcttc 1200

gtgatcaggg gcgacgaggt gagacagatc gcaccaggac agacaggcaa gatcgcagac 1260

tacaattata agctgcctga cgatttcacc ggctgcgtga tcgcctggaa cagcaacaat 1320

ctggattcca aagtgggcgg caactacaat tatctgtaca ggctgtttcg caagtccaat 1380

ctgaagccat tcgagcggga catcagcaca gagatctacc aggcaggctc caccccatgc 1440

aatggagtgg agggctttaa ctgttatttc cctctgcagt cttacggctt ccagccaaca 1500

aacggcgtgg gctatcagcc ctacagagtg gtggtgctgt cctttgagct gctgcacgca 1560

cctgcaacag tgtgcggacc aaagaagtct accaatctgg tgaagaacaa gtgcgtgaac 1620

ttcaacttca acggcctgac cggcacaggc gtgctgaccg agtccaacaa gaagttcctg 1680

ccttttcagc agttcggcag agacatcgcc gataccacag acgccgtgag agaccctcag 1740

accctggaga tcctggacat cacaccatgc tctttcggcg gcgtgagcgt gatcacacca 1800

ggcaccaata caagcaacca ggtggccgtg ctgtatcagg acgtgaattg taccgaggtg 1860

cccgtggcca tccacgcaga tcagctcacc cctacatgga gggtgtactc caccggctct 1920

aacgtgttcc agacacgcgc cggatgcctg atcggagccg agcatgtgaa caattcttat 1980

gagtgcgaca tccctatcgg agccggcatc tgtgccagct accagaccca gacaaacagc 2040

ccaggctccg ccagctccgt ggcctctcag agcatcatcg cctataccat gagcctgggg 2100

gccgagaata gcgtggccta ctctaacaat agcatcgcca tccctaccaa cttcacaatc 2160

tccgtgacca cagagatcct gccagtgtcc atgaccaaga catctgtgga ctgcacaatg 2220

tatatctgtg gcgattctac cgagtgcagc aacctgctgc tgcagtacgg cagcttttgt 2280

acccagctca accgggccct gacaggaatc gcagtggagc aggacaagaa cacacaggag 2340

gtgttcgccc aggtgaagca gatctacaag accccaccca tcaaggactt tggcggcttc 2400

aacttcagcc agatcctgcc cgatccttcc aagccatcta agaggagctt tatcgaggac 2460

ctgctgttca acaaggtgac cctggccgat gccggcttca tcaagcagta tggcgattgc 2520

ctgggcgaca tcgcagcccg cgacctgatc tgtgcccaga agtttaatgg cctgaccgtg 2580

ctgcctccac tgctgacaga tgagatgatc gcacagtaca catccgccct gctggccggc 2640

accatcacat ctggatggac cttcggggcc ggggccgccc tgcagatccc ctttgccatg 2700

cagatggcct atagattcaa cggcatcggc gtgacccaga atgtgctgta cgagaaccag 2760

aagctgatcg ccaatcagtt taactccgcc atcggcaaga tccaggactc cctgtctagc 2820

acagcctctg ccctgggcaa gctgcaggat gtggtgaatc agaacgccca ggccctgaat 2880

accctggtga agcagctcag cagcaacttc ggggccatca gcagcgtgct gaacgacatc 2940

ctgagccggc tggacaaggt ggaggcagag gtgcagatcg acaggctgat cacaggccgc 3000

ctgcagagcc tgcagaccta cgtgacacag cagctcatca gggccgccga gatcagagcc 3060

tccgccaatc tggccgccac caagatgtct gagtgcgtgc tgggccagag caagcgcgtg 3120

gacttttgtg gcaagggcta tcacctgatg tccttcccac agtctgcccc tcacggagtg 3180

gtgtttctgc atgtgaccta cgtgccagcc caggagaaga acttcaccac agcccccgca 3240

atctgccacg atggcaaggc acactttcct cgggagggcg tgttcgtgtc taacggcacc 3300
cactggtttg tgacacagag aaatttctac gagccacaga tcatcaccac agacaataca 3360
ttcgtgagcg gcaactgtga cgtggtcatc ggcatcgtga acaataccgt gtatgatcct 3420
ctgcagccag agctggactc ctttaaggag gagctggata agtacttcaa gaatcacacc 3480
tctcccgacg tggatctggg cgacatcagc ggcatcaatg cctccgtggt gaacatccag 3540
aaggagatcg acaggctgaa cgaggtggcc aagaatctga acgagtccct gatcgatctg 3600
caggagctgg gcaagtatga gcagtacatc aagtggccct ggtatatctg gctgggcttc 3660

atcgccggcc tgatcgccat cgtgatggtg accatcatgc tgtgctgtat gacatcctgc 3720
tgttcttgcc tgaagggctg ctgtagctgt ggctcctgct gtaagtttga tgaggacgat 3780
agcgagcctg tgctgaaggg cgtgaagctg cactacacct ga 3822

<210> 19
<211> 1273
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<400> 19

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

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

85 90 95
Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

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

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175
Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190
Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205
Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220
Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240
Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255
Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270
Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285
Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320
Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335
Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350
Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro

370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
385 390 395 400
Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
405 410 415
Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
420 425 430
Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460
Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480
Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
485 490 495
Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
500 505 510
Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys

515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val

595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile

610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser

625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val

645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala

660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser Pro Gly Ser Ala Ser Ser Val Ala

675 680 685

Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser

690 695 700

Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile

705 710 715 720

Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val

725 730 735

Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu

740 745 750

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

755 760 765

Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln

770 775 780

Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe

785 790 795 800

Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser

805 810 815

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

820 825 830

Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp

835 840 845

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

850 855 860

Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly

865 870 875 880

Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
885 890 895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
900 905 910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn
915 920 925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala
930 935 940
Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn

945 950 955 960
Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val
965 970 975
Leu Asn Asp Ile Leu Ser Arg Leu Asp Pro Pro Glu Ala Glu Val Gln
980 985 990
Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val
995 1000 1005
Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu
1010 1015 1020

Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val
1025 1030 1035 1040
Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala
1045 1050 1055
Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu
1060 1065 1070
Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His
1075 1080 1085
Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val

1090 1095 1100
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr
1105 1110 1115 1120

Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr

1125 1130 1135

Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu

1140 1145 1150

Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp

1155 1160 1165

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

1170 1175 1180

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

1185 1190 1195 1200

Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile

1205 1210 1215

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

1220 1225 1230

Met Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys

1235 1240 1245

Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val

1250 1255 1260

Leu Lys Gly Val Lys Leu His Tyr Thr

1265 1270

<210> 20

<211> 3822

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 20

atgtttgttt ttcttgtttt attgccacta gtctctagtc agtgtgttaa tcttacaacc 60

agaactcaat taccccctgc atacactaat tctttcacac gtggtgttta ttaccctgac 120

aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc tttcttttcc 180

aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat 240

aaccctgtcc taccatttaa tgatggtgtt tattttgctt ccactgagaa gtctaacata 300

ataagaggct ggatttttgg tactacttta gattcgaaga cccagtccct acttattgtt 360

aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420

ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480

tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggaccttgaa 540

ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat tgatggttat 600

tttaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660

tcggctttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720

ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg ttggacagct 780

ggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840

gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaaag 900

tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960

caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc ttttggtgaa 1020

gtttttaacg ccaccagatt tgcatctgtt tatgcttgga acaggaagag aatcagcaac 1080

tgtgttgctg attattctgt cctatataat tccgcatcat tttccacttt taagtgttat 1140

ggagtgtctc ctactaaatt aaatgatctc tgctttacta atgtctatgc agattcattt 1200

gtaattagag gtgatgaagt cagacaaatc gctccagggc aaactggaaa gattgctgat 1260

tataattata aattaccaga tgattttaca ggctgcgtta tagcttggaa ttctaacaat 1320

cttgattcta aggttggtgg taattataat tacctgtata gattgtttag gaagtctaat 1380

ctcaaacctt ttgagagaga tatttcaact gaaatctatc aggccggtag cacaccttgt 1440

aatggtgttg aaggttttaa ttgttacttt cctttacaat catatggttt ccaacccact 1500

aatggtgttg gttaccaacc atacagagta gtagtacttt cttttgaact tctacatgca 1560

ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa atgtgtcaat 1620

ttcaacttca atggtttaac aggcacaggt gttcttactg agtctaacaa aaagtttctg 1680

cctttccaac aatttggcag agacattgct gacactactg atgctgtccg tgatccacag 1740

acacttgaga ttcttgacat tacaccatgt tcttttggtg gtgtcagtgt tataacacca 1800

ggaacaaata cttctaacca ggttgctgtt ctttatcagg atgttaactg cacagaagtc 1860

cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc tacaggttct 1920

aatgtttttc aaacacgtgc aggctgttta ataggggctg aacatgtcaa caactcatat 1980

gagtgtgaca tacccattgg tgcaggtata tgcgctagtt atcagactca gactaattct 2040

cctggtagtg caagtagtgt agctagtcaa tccatcattg cctacactat gtcacttggt 2100

gcagaaaatt cagttgctta ctctaataac tctattgcca tacccacaaa ttttactatt 2160

agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga ttgtacaatg 2220

tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg cagtttttgt 2280
acacaattaa accgtgcttt aactggaata gctgttgaac aagacaaaaa cacccaagaa 2340
gtttttgcac aagtcaaaca aatttacaaa acaccaccaa ttaaagattt tggtggtttt 2400
aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt tattgaagat 2460
ctacttttca acaaagtgac acttgcagat gctggcttca tcaaacaata tggtgattgc 2520
cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg ccttactgtt 2580
ttgccacctt tgctcacaga tgaaatgatt gctcaataca cttctgcact gttagcgggt 2640

acaatcactt ctggttggac ctttggtgca ggtgctgcat tacaaatacc atttgctatg 2700
caaatggctt ataggtttaa tggtattgga gttacacaga atgttctcta tgagaaccaa 2760
aaattgattg ccaaccaatt taatagtgct attggcaaaa ttcaagactc actttcttcc 2820
acagcaagtg cacttggaaa acttcaagat gtggtcaacc aaaatgcaca agctttaaac 2880
acgcttgtta aacaacttag ctccaatttt ggtgcaattt caagtgtttt aaatgatatc 2940
ctttcacgtc ttgaccctcc tgaggctgaa gtgcaaattg ataggttgat cacaggcaga 3000
cttcaaagtt tgcagacata tgtgactcaa caattaatta gagctgcaga aatcagagct 3060

tctgctaatc ttgctgctac taaaatgtca gagtgtgtac ttggacaatc aaaaagagtt 3120
gatttttgtg gaaagggcta tcatcttatg tccttccctc agtcagcacc tcatggtgta 3180
gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac tgctcctgcc 3240
atttgtcatg atggaaaagc acactttcct cgtgaaggtg tctttgtttc aaatggcaca 3300
cactggtttg taacacaaag gaatttttat gaaccacaaa tcattactac agacaacaca 3360
tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt ttatgatcct 3420
ttgcaacctg aattagactc attcaaggag gagttagata aatattttaa gaatcataca 3480

tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt aaacattcaa 3540
aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct catcgatctc 3600
caagaacttg gaaagtatga gcagtatata aaatggccat ggtacatttg gctaggtttt 3660
atagctggct tgattgccat agtaatggtg acaattatgc tttgctgtat gaccagttgc 3720
tgtagttgtc tcaagggctg ttgttcttgt ggatcctgct gcaaatttga tgaagacgac 3780
tctgagccag tgctcaaagg agtcaaatta cattacacat aa 3822

<210> 21

<211> 3822

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 21

atgttcgtgt ttctggtgct gctgcctctg gtgagctccc agtgcgtgaa cctgaccaca 60

aggacccagc tccccccctgc ctataccaat tccttcacaa ggggcgtgta ctatccagac 120

aaggtgtttc gctctagcgt gctgcacagc acacaggatc tgtttctgcc cttcttttcc 180

aacgtgacct ggttccacgc catccatgtg agcggcacca atggcacaaa gaggttcgac 240

aatcctgtgc tgcccttcaa cgatggcgtg tacttcgcct ctaccgagaa gagcaacatc 300

atccgcggct ggatctttgg caccacactg gactccaaga cacagtctct gctgatcgtg 360

aacaatgcca ccaacgtggt catcaaggtg tgcgagttcc agttttgtaa tgatcctttc 420

ctgggcgtgt actatcacaa gaacaataag agctggatgg agtccgagtt tcgcgtgtat 480

tcctctgcca acaattgcac atttgagtac gtgtcccagc cattcctgat ggacctggag 540

ggcaagcagg gcaatttcaa gaacctgcgg gagttcgtgt ttaagaatat cgatggctac 600

ttcaagatct acagcaagca cacccctatc aacctggtga gagacctgcc acagggcttc 660

tccgccctgg agcctctggt ggatctgcca atcggcatca acatcaccag gtttcagaca 720

ctgctggccc tgcaccgcag ctacctgaca cctggcgaca gctcctctgg atggaccgcc 780

ggggccgccg cctactatgt gggctatctg cagccacgga ccttcctgct gaagtacaac 840

gagaatggca ccatcacaga cgcagtggat tgcgccctgg acccccctgtc cgagaccaag 900

tgtacactga agtcttttac cgtggagaag ggcatctatc agacatctaa tttccgggtg 960

cagcccaccg agagcatcgt gagatttcca aatatcacaa acctgtgccc ctttggcgag 1020

gtgttcaacg ccaccagatt cgccagcgtg tacgcctgga atcggaagag aatcagcaac 1080

tgcgtggccg actattccgt gctgtacaac tctgccagct tctccacctt taagtgctat 1140

ggcgtgtctc ccacaaagct gaatgacctg tgctttacca acgtgtacgc cgatagcttc 1200

gtgatcaggg gcgacgaggt gagacagatc gcaccaggcc agacaggcaa gatcgccgac 1260

tacaattata agctgcccga cgatttcacc ggctgcgtga tcgcctggaa cagcaacaat 1320

ctggattcca aagtgggcgg caactacaat tatctgtaca ggctgtttcg caagtccaat 1380

ctgaagcctt tcgagcggga catcagcaca gagatctacc aggccggctc caccccatgc 1440

aatggcgtgg agggctttaa ctgttatttc cccctgcagt cttacggctt ccagcctaca 1500

aacggcgtgg gctatcagcc atacagagtg gtggtgctgt cctttgagct gctgcacgca 1560

ccagcaacag tgtgcggacc taagaagtct accaatctgg tgaagaacaa gtgcgtgaac 1620

ttcaacttca acggcctgac cggcacaggc gtgctgaccg agtccaacaa gaagttcctg 1680

ccctttcagc agttcggcag agacatcgcc gataccacag acgccgtgag agacccccag 1740

accctggaga tcctggacat cacaccttgc tctttcggcg gcgtgagcgt gatcacacct 1800

ggcaccaata caagcaacca ggtggccgtg ctgtatcagg acgtgaattg taccgaggtg 1860

ccagtggcca tccacgccga tcagctcacc cccacatgga gggtgtactc caccggctct 1920

aacgtgttcc agacacgcgc cggatgcctg atcggagccg agcatgtgaa caattcttat 1980

gagtgcgaca tccccatcgg agccggcatc tgtgccagct accagaccca gacaaacagc 2040

cctggctccg ccagctccgt ggcctctcag agcatcatcg cctataccat gagcctgggg 2100

gccgagaata gcgtggccta ctctaacaat agcatcgcca tccccaccaa cttcacaatc 2160

tccgtgacca cagagatcct gcccgtgagc atgaccaaga catctgtgga ctgcacaatg 2220

tatatctgtg gcgattctac cgagtgcagc aacctgctgc tgcagtacgg cagcttttgt 2280

acccagctca accgggccct gacaggaatc gcagtggagc aggacaagaa cacacaggag 2340

gtgttcgccc aggtgaagca gatctacaag accccaccca tcaaggactt tggcggcttc 2400

aacttcagcc agatcctgcc agatccctcc aagccttcta agaggagctt tatcgaggac 2460

ctgctgttca acaaggtgac cctggccgat gccggcttca tcaagcagta tggcgattgc 2520

ctgggcgaca tcgcagcccg cgacctgatc tgtgcccaga agtttaatgg cctgaccgtg 2580

ctgcctccac tgctgacaga tgagatgatc gcacagtaca catccgccct gctggccggc 2640

accatcacat ctggatggac cttcggggcc ggggccgccc tgcagatccc atttgccatg 2700

cagatggcct atagattcaa cggcatcggc gtgacccaga atgtgctgta cgagaaccag 2760

aagctgatcg ccaatcagtt taactccgcc atcggcaaga tccaggactc cctgtctagc 2820

acagcctctg ccctgggcaa gctgcaggat gtggtgaatc agaacgccca ggccctgaat 2880

accctggtga agcagctcag cagcaacttc ggggccatca gcagcgtgct gaacgacatc 2940

ctgagccggc tggacccccc tgaggcagag gtgcagatcg acaggctgat cacaggccgc 3000

ctgcagagcc tgcagaccta cgtgacacag cagctcatca gggccgccga gatcagagcc 3060

tccgccaatc tggccgccac caagatgtct gagtgcgtgc tgggccagag caagcgcgtg 3120

gacttttgtg gcaagggcta tcacctgatg tccttcccac agtctgcccc ccacggagtg 3180

gtgtttctgc atgtgaccta cgtgcctgcc caggagaaga acttcaccac agccccagcc 3240

atctgccacg atggcaaggc acactttccc cgggagggcg tgttcgtgtc taacggcacc 3300

cactggtttg tgacacagag aaatttctac gagcctcaga tcatcaccac agacaataca 3360

ttcgtgagcg gcaactgtga cgtggtcatc ggcatcgtga acaataccgt gtatgatccc 3420

ctgcagcctg agctggactc ctttaaggag gagctggata agtacttcaa gaatcacacc 3480

tctcccgacg tggatctggg cgacatcagc ggcatcaatg cctccgtggt gaacatccag 3540

aaggagatcg acaggctgaa cgaggtggcc aagaatctga acgagtccct gatcgatctg 3600

caggagctgg gcaagtatga gcagtacatc aagtggccat ggtatatctg gctgggcttc 3660

atcgccggcc tgatcgccat cgtgatggtg accatcatgc tgtgctgtat gacatcctgc 3720

tgttcttgcc tgaagggctg ctgtagctgt ggctcctgct gtaagtttga tgaggacgat 3780

agcgagcccg tgctgaaggg cgtgaagctg cactacacct ga 3822

<210> 22

<211> 1255

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 22

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

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

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

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu

85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140
Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175
Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190
Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205
Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220
Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240
Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255
Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
260 265 270
Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
275 280 285
Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
305 310 315 320
Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
325 330 335
Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
340 345 350
Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
355 360 365
Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro

370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe

385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly

405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys

420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn

435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe

450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys

465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly

485 490 495

Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val

500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys

515 520 525

Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn

530 535 540

Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu

545 550 555 560

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val

565 570 575

Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe

580 585 590

Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val

595 600 605

Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile

610 615 620

His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser

625 630 635 640

Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val

645 650 655

Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala

660 665 670

Ser Tyr Gln Thr Gln Thr Asn Ser Pro Gly Ser Ala Ser Ser Val Ala

675 680 685

Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser

690 695 700

Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile

705 710 715 720

Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val

725 730 735

Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu

740 745 750

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

755 760 765

Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln

770 775 780

Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe

785 790 795 800

Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser

805 810 815

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

820 825 830

Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp

835 840 845

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

850 855 860

Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly

865 870 875 880

Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
885 890 895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
900 905 910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn
915 920 925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala
930 935 940
Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn

945 950 955 960
Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val
965 970 975
Leu Asn Asp Ile Leu Ser Arg Leu Asp Pro Pro Glu Ala Glu Val Gln
980 985 990
Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val
995 1000 1005
Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu
1010 1015 1020

Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val
1025 1030 1035 1040
Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala
1045 1050 1055
Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu
1060 1065 1070
Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His
1075 1080 1085
Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val

1090 1095 1100
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr
1105 1110 1115 1120
Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr
1125 1130 1135

Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu
1140 1145 1150
Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp
1155 1160 1165

Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp
1170 1175 1180
Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu
1185 1190 1195 1200
Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Phe Phe Phe Ile Ile
1205 1210 1215
Gly Leu Ile Ile Gly Leu Phe Leu Val Leu Arg Val Gly Ile His Leu
1220 1225 1230
Cys Ile Lys Leu Lys His Thr Lys Lys Arg Gln Ile Tyr Thr Asp Ile

1235 1240 1245
Glu Met Asn Arg Leu Gly Lys
1250 1255

<210> 23
<211> 3768
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polynucleotide
<400> 23
atgtttgttt ttcttgtttt attgccacta gtctctagtc agtgtgttaa tcttacaacc 60
agaactcaat taccccctgc atacactaat tctttcacac gtggtgttta ttaccctgac 120
aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc tttcttttcc 180

aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat 240
aaccctgtcc taccatttaa tgatggtgtt tattttgctt ccactgagaa gtctaacata 300
ataagaggct ggatttttgg tactacttta gattcgaaga cccagtccct acttattgtt 360
aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420
ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480
tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggaccttgaa 540

ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat tgatggttat 600

tttaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660

tcggctttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720

ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg ttggacagct 780

ggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840

gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaaag 900

tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960

caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc ttttggtgaa 1020

gtttttaacg ccaccagatt tgcatctgtt tatgcttgga acaggaagag aatcagcaac 1080

tgtgttgctg attattctgt cctatataat tccgcatcat tttccacttt taagtgttat 1140

ggagtgtctc ctactaaatt aaatgatctc tgctttacta atgtctatgc agattcattt 1200

gtaattagag gtgatgaagt cagacaaatc gctccagggc aaactggaaa gattgctgat 1260

tataattata aattaccaga tgattttaca ggctgcgtta tagcttggaa ttctaacaat 1320

cttgattcta aggttggtgg taattataat tacctgtata gattgtttag gaagtctaat 1380

ctcaaacctt ttgagagaga tatttcaact gaaatctatc aggccggtag cacaccttgt 1440

aatggtgttg aaggttttaa ttgttacttt cctttacaat catatggttt ccaacccact 1500

aatggtgttg gttaccaacc atacagagta gtagtacttt cttttgaact tctacatgca 1560

ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa atgtgtcaat 1620

ttcaacttca atggtttaac aggcacaggt gttcttactg agtctaacaa aaagtttctg 1680

cctttccaac aatttggcag agacattgct gacactactg atgctgtccg tgatccacag 1740

acacttgaga ttcttgacat tacaccatgt tcttttggtg gtgtcagtgt tataacacca 1800

ggaacaaata cttctaacca ggttgctgtt ctttatcagg atgttaactg cacagaagtc 1860

cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc tacaggttct 1920

aatgtttttc aaacacgtgc aggctgttta ataggggctg aacatgtcaa caactcatat 1980

gagtgtgaca tacccattgg tgcaggtata tgcgctagtt atcagactca gactaattct 2040

cctggtagtg caagtagtgt agctagtcaa tccatcattg cctacactat gtcacttggt 2100

gcagaaaatt cagttgctta ctctaataac tctattgcca tacccacaaa ttttactatt 2160

agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga ttgtacaatg 2220

tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg cagtttttgt 2280

acacaattaa accgtgcttt aactggaata gctgttgaac aagacaaaaa cacccaagaa 2340

gtttttgcac aagtcaaaca aatttacaaa acaccaccaa ttaaagattt tggtggtttt 2400

aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt tattgaagat 2460

ctacttttca acaaagtgac acttgcagat gctggcttca tcaaacaata tggtgattgc 2520

cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg ccttactgtt 2580

ttgccacctt tgctcacaga tgaaatgatt gctcaataca cttctgcact gttagcgggt 2640

acaatcactt ctggttggac ctttggtgca ggtgctgcat tacaaatacc atttgctatg 2700

caaatggctt ataggtttaa tggtattgga gttacacaga atgttctcta tgagaaccaa 2760

aaattgattg ccaaccaatt taatagtgct attggcaaaa ttcaagactc actttcttcc 2820

acagcaagtg cacttggaaa acttcaagat gtggtcaacc aaaatgcaca agctttaaac 2880

acgcttgtta aacaacttag ctccaatttt ggtgcaattt caagtgtttt aaatgatatc 2940

ctttcacgtc ttgaccctcc tgaggctgaa gtgcaaattg ataggttgat cacaggcaga 3000

cttcaaagtt tgcagacata tgtgactcaa caattaatta gagctgcaga aatcagagct 3060

tctgctaatc ttgctgctac taaaatgtca gagtgtgtac ttggacaatc aaaaagagtt 3120

gatttttgtg gaaagggcta tcatcttatg tccttccctc agtcagcacc tcatggtgta 3180

gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac tgctcctgcc 3240

atttgtcatg atggaaaagc acactttcct cgtgaaggtg tctttgtttc aaatggcaca 3300

cactggtttg taacacaaag gaatttttat gaaccacaaa tcattactac agacaacaca 3360

tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt ttatgatcct 3420

ttgcaacctg aattagactc attcaaggag gagttagata aatattttaa gaatcataca 3480

tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt aaacattcaa 3540

aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct catcgatctc 3600

caagaacttg gaaagtatga gcagtatata aaatttttct ttatcatagg gttaatcatt 3660

ggactattct tggttctccg agttggtatc catctttgca ttaaattaaa gcacaccaag 3720

aaaagacaga tttatacaga catagagatg aaccgacttg gaaagtaa 3768

<210> 24

<211> 3768

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 24

atgttcgtgt tcctggtgct gctgcctctg gtgagctccc agtgcgtgaa cctgaccaca 60

aggacccagc tccccccctgc ctataccaat tcctttacaa ggggcgtgta ctatccagac 120

aaggtgttcc gctctagcgt gctgcactct acacaggatc tgttcctgcc cttctttagc 180

aacgtgacct ggtttcacgc catccatgtg agcggcacca atggcacaaa gcggtttgac 240

aatcctgtgc tgccattcaa cgatggcgtg tactttgcct ccaccgagaa gtctaacatc 300

atcagaggct ggatcttcgg caccacactg gacagcaaga cacagtccct gctgatcgtg 360

aacaatgcca ccaacgtggt catcaaggtg tgcgagtttc agttctgtaa tgatcctttt 420

ctgggcgtgt actatcacaa gaacaataag tcttggatgg agagcgagtt ccgcgtgtat 480

tcctctgcca acaattgtac attcgagtac gtgtcccagc catttctgat ggacctggag 540

ggcaagcagg gcaacttcaa gaacctgcgg gagttcgtgt tcaagaatat cgatggctat 600

ttcaagatct actctaagca cacccctatc aacctggtgc gcgacctgcc acagggcttt 660

agcgccctgg agcctctggt ggatctgcca atcggcatca acatcaccag gttccagaca 720

ctgctggccc tgcaccgcag ctacctgaca cctggcgaca gctcctctgg atggaccgcc 780

ggggccgccg cctactatgt gggctatctg cagccacgga cctttctgct gaagtacaac 840

gagaatggca ccatcacaga cgcagtggat tgcgccctgg acccccctgag cgagaccaag 900

tgtacactga agtccttcac cgtggagaag ggcatctatc agacatccaa ttttcgggtg 960

cagcccaccg agtctatcgt gagattccca aatatcacaa acctgtgccc cttcggcgag 1020

gtgtttaacg ccaccagatt cgccagcgtg tacgcctgga atcggaagag aatctctaac 1080

tgcgtggccg actatagcgt gctgtacaac tctgccagct tttccacctt caagtgctat 1140

ggcgtgtccc ccacaaagct gaatgacctg tgcttcacca acgtgtacgc cgattctttt 1200

gtgatcaggg gcgacgaggt gagacagatc gcaccaggcc agacaggcaa gatcgccgac 1260

tacaattata agctgcccga cgatttcacc ggctgcgtga tcgcctggaa ctctaacaat 1320

ctggatagca aagtgggcgg caactacaat tatctgtaca ggctgttccg caagagcaat 1380

ctgaagcctt ttgagcggga catctctaca gagatctacc aggccggcag caccccatgc 1440

aatggcgtgg agggcttcaa ctgttatttt cccctgcagt cctacggctt tcagcctacc 1500

aacggcgtgg gctatcagcc atacagagtg gtggtgctga gcttcgagct gctgcacgca 1560

ccagcaacag tgtgcggacc taagaagtcc accaatctgg tgaagaacaa gtgcgtgaac 1620

ttcaacttca acggcctgac cggcacaggc gtgctgaccg agtccaataa gaagtttctg 1680

cccttccagc agtttggccg ggacatcgcc gataccacag acgccgtgag agacccccag 1740

accctggaga tcctggacat cacaccttgc tccttcggcg gcgtgtctgt gatcacacct 1800

ggcaccaata caagcaacca ggtggccgtg ctgtatcagg acgtgaattg taccgaggtg 1860

ccagtggcca tccacgccga tcagctcacc cccacatggc gggtgtactc caccggctct 1920

aacgtgttcc agacaagagc cggctgcctg atcggagccg agcatgtgaa caattcctat 1980

gagtgcgaca tccccatcgg agccggcatc tgtgcctctt accagaccca gacaaacagc 2040

cctggctccg ccagctccgt ggcctctcag agcatcatcg cctataccat gagcctgggg 2100

gccgagaaca gcgtggccta ctctaacaat agcatcgcca tccccaccaa ctttacaatc 2160

tctgtgacca cagagatcct gcctgtgagc atgaccaaga catccgtgga ctgcacaatg 2220

tatatctgtg gcgattccac cgagtgctct aacctgctgc tgcagtacgg cagcttctgt 2280

acccagctca accgggccct gacaggaatc gcagtggagc aggacaagaa cacacaggag 2340

gtgtttgccc aggtgaagca gatctacaag accccaccca tcaaggactt cggcggcttt 2400

aatttctccc agatcctgcc agatccctcc aagccatcta agcggagctt catcgaggac 2460

ctgctgttta acaaggtgac cctggccgat gccggcttta tcaagcagta tggcgattgc 2520

ctgggcgaca tcgccgccag agacctgatc tgtgcccaga agttcaatgg cctgaccgtg 2580

ctgcctccac tgctgacaga tgagatgatc gcacagtaca caagcgccct gctggccggc 2640

accatcacat ccggatggac cttcggggcc ggggccgccc tgcagatccc cttcgccatg 2700

cagatggcct ataggtttaa cggcatcggc gtgacccaga atgtgctgta cgagaaccag 2760

aagctgatcg ccaatcagtt caactccgcc atcggcaaga tccaggacag cctgtctagc 2820

acagcctccg ccctgggcaa gctgcaggat gtggtgaatc agaacgccca ggccctgaat 2880

accctggtga agcagctcag cagcaacttc ggggccatca gcagcgtgct gaacgacatc 2940

ctgagccggc tggacccccc tgaggcagag gtgcagatcg acaggctgat cacaggccgc 3000

ctgcagtctc tgcagaccta tgtgacacag cagctcatca gggccgccga gatcagagcc 3060

agcgccaatc tggccgccac caagatgtcc gagtgcgtgc tgggccagtc taagcgcgtg 3120

gacttctgtg gcaagggcta tcacctgatg agctttccac agtccgcccc ccacggagtg 3180

gtgttcctgc atgtgaccta cgtgcctgcc caggagaaga actttaccac agccccagcc 3240

atctgccacg atggcaaggc acacttcccc agggagggcg tgttcgtgag caacggcacc 3300

cactggttcg tgacacagcg caacttctac gagcctcaga tcatcaccac agacaataca 3360

ttcgtgtctg gcaactgtga cgtggtcatc ggcatcgtga acaataccgt gtatgatccc 3420

ctgcagcctg agctggacag cttcaaggag gagctggata agtactttaa gaatcacacc 3480

tcccccgacg tggatctggg cgacatctct ggcatcaatg ccagcgtggt gaacatccag 3540

aaggagatcg acaggctgaa cgaggtggcc aagaatctga acgagagcct gatcgatctg 3600

caggagctgg gcaagtatga gcagtacatc aagttctttt tcatcatcgg cctgatcatc 3660

ggcctgttcc tggtgctgcg cgtgggcatc cacctgtgca tcaagctgaa gcacaccaag 3720

aagaggcaga tctacacaga catcgagatg aaccgcctgg gcaagtga 3768

<210> 25

<211> 1255

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 25

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

Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe
20 25 30

Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu
35 40 45

His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp

50 55 60

Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp
65 70 75 80

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
100 105 110

Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
115 120 125

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

130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe

180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr

195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu

210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys

290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val

305 310 315 320

Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys

325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala

340 345 350

Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu

355 360 365

Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro

370 375 380

Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe

385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly

405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys

420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
435 440 445
Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
450 455 460
Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
465 470 475 480
Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly

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

Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605
Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser

625 630 635 640
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655
Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
675 680 685

Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
690 695 700

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

770 775 780
Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe
785 790 795 800
Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
805 810 815
Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly
820 825 830
Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp
835 840 845

Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu
850 855 860
Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly
865 870 875 880
Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile
885 890 895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr
900 905 910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn

915 920 925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala

930 935 940

Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn

945 950 955 960

Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val

965 970 975

Leu Asn Asp Ile Leu Ser Arg Leu Asp 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 Leu

1010 1015 1020

Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val

1025 1030 1035 1040

Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala

1045 1050 1055

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu

1060 1065 1070

Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His

1075 1080 1085

Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val

1090 1095 1100

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn Thr

1105 1110 1115 1120

Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn Asn Thr

1125 1130 1135

Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu

1140 1145 1150

Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp

1155 1160 1165

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

1170 1175 1180

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

1185 1190 1195 1200

Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Phe Phe Phe Ile Ile

1205 1210 1215

Gly Leu Ile Ile Gly Leu Phe Leu Val Leu Arg Val Gly Ile His Leu

1220 1225 1230

Cys Ile Lys Leu Lys His Thr Lys Lys Arg Gln Ile Tyr Thr Asp Ile

1235 1240 1245

Glu Met Asn Arg Leu Gly Lys

1250 1255

<210> 26

<211> 3819

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 26

ctcgaggttt aaacgaattc cgccaccatg tttgtttttc ttgttttatt gccactagtc 60

tctagtcagt gtgttaatct tacaaccaga actcaattac cccctgcata cactaattct 120

ttcacacgtg gtgtttatta ccctgacaaa gttttcagat cctcagtttt acattcaact 180

caggacttgt tcttaccttt cttttccaat gttacttggt tccatgctat acatgtctct 240

gggaccaatg gtactaagag gtttgataac cctgtcctac catttaatga tggtgtttat 300

tttgcttcca ctgagaagtc taacataata agaggctgga tttttggtac tactttagat 360

tcgaagaccc agtccctact tattgttaat aacgctacta atgttgttat taaagtctgt 420

gaatttcaat tttgtaatga tccatttttg ggtgtttatt accacaaaaa caacaaaagt 480

tggatggaaa gtgagttcag agtttattct agtgcgaata attgcacttt tgaatatgtc 540

tctcagcctt ttcttatgga ccttgaagga aaacagggta atttcaaaaa tcttagggaa 600

tttgtgttta agaatattga tggttatttt aaaatatatt ctaagcacac gcctattaat 660

ttagtgcgtg atctccctca gggtttttcg gctttagaac cattggtaga tttgccaata 720

ggtattaaca tcactaggtt tcaaacttta cttgctttac atagaagtta tttgactcct 780

ggtgattctt cttcaggttg gacagctggt gctgcagctt attatgtggg ttatcttcaa 840

cctaggactt ttctattaaa atataatgaa aatggaacca ttacagatgc tgtagactgt 900

gcacttgacc ctctctcaga aacaaagtgt acgttgaaat ccttcactgt agaaaaagga 960

atctatcaaa cttctaactt tagagtccaa ccaacagaat ctattgttag atttcctaat 1020

attacaaact tgtgcccttt tggtgaagtt tttaacgcca ccagatttgc atctgtttat 1080

gcttggaaca ggaagagaat cagcaactgt gttgctgatt attctgtcct atataattcc 1140

gcatcatttt ccacttttaa gtgttatgga gtgtctccta ctaaattaaa tgatctctgc 1200

tttactaatg tctatgcaga ttcatttgta attagaggtg atgaagtcag acaaatcgct 1260

ccagggcaaa ctggaaagat tgctgattat aattataaat taccagatga ttttacaggc 1320

tgcgttatag cttggaattc taacaatctt gattctaagg ttggtggtaa ttataattac 1380

ctgtatagat tgtttaggaa gtctaatctc aaacctttg agagagatat ttcaactgaa 1440

atctatcagg ccggtagcac accttgtaat ggtgttgaag gttttaattg ttactttcct 1500

ttacaatcat atggtttcca acccactaat ggtgttggtt accaaccata cagagtagta 1560

gtactttctt ttgaacttct acatgcacca gcaactgttt gtggacctaa aaagtctact 1620

aatttggtta aaaacaaatg tgtcaatttc aacttcaatg gtttaacagg cacaggtgtt 1680

cttactgagt ctaacaaaaa gtttctgcct ttccaacaat ttggcagaga cattgctgac 1740

actactgatg ctgtccgtga tccacagaca cttgagattc ttgacattac accatgttct 1800

tttggtggtg tcagtgttat aacaccagga acaaatactt ctaaccaggt tgctgttctt 1860

tatcaggatg ttaactgcac agaagtccct gttgctattc atgcagatca acttactcct 1920

acttggcgtg tttattctac aggttctaat gtttttcaaa cacgtgcagg ctgtttaata 1980

ggggctgaac atgtcaacaa ctcatatgag tgtgacatac ccattggtgc aggtatatgc 2040

gctagttatc agactcagac taattctcct cggcgggcac gtagtgtagc tagtcaatcc 2100

atcattgcct acactatgtc acttggtgca gaaaattcag ttgcttactc taataactct 2160

attgccatac ccacaaattt tactattagt gttaccacag aaattctacc agtgtctatg 2220

accaagacat cagtagattg tacaatgtac atttgtggtg attcaactga atgcagcaat 2280

cttttgttgc aatatggcag tttttgtaca caattaaacc gtgctttaac tggaatagct 2340

gttgaacaag acaaaaacac ccaagaagtt tttgcacaag tcaaacaaat ttacaaaaca 2400

ccaccaatta aagattttgg tggttttaat ttttcacaaa tattaccaga tccatcaaaa 2460

ccaagcaaga ggtcatttat tgaagatcta cttttcaaca aagtgacact tgcagatgct 2520

ggcttcatca aacaatatgg tgattgcctt ggtgatattg ctgctagaga cctcatttgt 2580

gcacaaaagt ttaacggcct tactgttttg ccacctttgc tcacagatga aatgattgct 2640

caatacactt ctgcactgtt agcgggtaca atcacttctg gttggacctt tggtgcaggt 2700

gctgcattac aaataccatt tgctatgcaa atggcttata ggtttaatgg tattggagtt 2760

acacagaatg ttctctatga gaaccaaaaa ttgattgcca accaatttaa tagtgctatt 2820

ggcaaaattc aagactcact ttcttccaca gcaagtgcac ttggaaaact tcaagatgtg 2880

gtcaaccaaa atgcacaagc tttaaacacg cttgttaaac aacttagctc caattttggt 2940

gcaatttcaa gtgttttaaa tgatatcctt tcacgtcttg acaaagttga ggctgaagtg 3000

caaattgata ggttgatcac aggcagactt caaagtttgc agacatatgt gactcaacaa 3060

ttaattagag ctgcagaaat cagagcttct gctaatcttg ctgctactaa aatgtcagag 3120

tgtgtacttg gacaatcaaa aagagttgat ttttgtggaa agggctatca tcttatgtcc 3180

ttccctcagt cagcacctca tggtgtagtc ttcttgcatg tgacttatgt ccctgcacaa 3240

gaaaagaact tcacaactgc tcctgccatt tgtcatgatg gaaaagcaca ctttcctcgt 3300

gaaggtgtct ttgtttcaaa tggcacacac tggtttgtaa cacaaaggaa tttttatgaa 3360

ccacaaatca ttactacaga caacacattt gtgtctggta actgtgatgt tgtaatagga 3420

attgtcaaca acacagttta tgatcctttg caacctgaat tagactcatt caaggaggag 3480

ttagataaat attttaagaa tcatacatca ccagatgttg atttaggtga catctctggc 3540

attaatgctt cagttgtaaa cattcaaaaa gaaattgacc gcctcaatga ggttgccaag 3600

aatttaaatg aatctctcat cgatctccaa gaacttggaa agtatgagca gtatataaaa 3660

tggccatttt tctttatcat agggttaatc attggactat tcttggttct ccgagttggt 3720

atccatcttt gcattaaatt aaagcacacc aagaaaagac agatttatac agacatagag 3780

atgaaccgac ttggaaagta agaattccac gtgggatcc 3819

<210> 27

<211> 3819

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 27

ctcgaggttt aaacgaattc cgccaccatg ttcgtgttcc tggtgctgct gcccctggtg 60

tccagccagt gcgtgaatct gaccacccgg acccaactgc ctcccgccta cacaaactct 120

ttcaccagag gggtttatta ccccgataag gtgttcagaa gctcagtgct tcattctacc 180

caggacctgt ttctgccttt tttcagcaac gtcacatggt tccacgccat ccacgtcagc 240

ggaaccaacg gcacgaagcg gttcgacaat cctgtgctgc cttttaacga cggcgtctac 300

tttgccagca cggaaaagag caacattatc cggggatgga tcttcggcac caccctggac 360

tctaaaaccc agagcctgtt gatcgtgaac aacgcaacca atgtggtgat caaggtctgc 420

gagttccaat tttgcaacga tcctttcctg ggcgtgtact accacaagaa caacaagtct 480

tggatggaat ctgagttccg cgtctacagc agcgcaaaca actgcacatt tgagtacgtg 540

tctcagcctt ttctgatgga cctggaagga aagcagggaa atttcaagaa cctgcgggag 600

ttcgtgttca agaacatcga cggctacttc aagatctaca gcaagcacac ccccatcaac 660

ctcgtgagag acctgcccca gggcttcagc gccctggaac ccctggtgga ccttcccata 720

ggaatcaaca tcacacggtt ccagacactg ctggccctgc atagaagcta cctgaccccct 780

ggagattcta gcagcggctg gaccgccggc gctgccgctt actacgtcgg atacctgcag 840

cctagaacct tcctgttgaa gtacaacgag aacggcacca tcacagatgc cgtggactgc 900

gccctggacc ccctgagcga gacaaagtgc accctgaaga gcttcaccgt ggagaagggc 960

atctaccaga caagcaactt cagagtgcag cctaccgagt caatcgtgag attcccaaac 1020

atcaccaacc tttgtccttt cggcgaggta tttaacgcca cccggttcgc cagcgtgtac 1080

gcctggaata ggaagcggat cagcaactgc gtggccgatt acagcgtgct ctataacagc 1140

gccagtttta gcactttcaa gtgctacgga gtctctccta caaagctgaa cgacctgtgc 1200

ttcaccaacg tgtatgccga cagcttcgtc atccggggcg acgaggtgcg acagatcgct 1260

cctggccaga ccggcaagat agccgactac aactacaagc tgcctgacga cttcacaggc 1320

tgcgtgatcg cttggaacag caacaatctg gatagcaaag tgggcggcaa ctataactac 1380

ctgtacagac tgttccggaa gtccaatctc aagccgtttg agagagacat cagcaccgaa 1440

atctaccagg ctggatctac accctgcaac ggcgtcgaag gcttcaattg ttacttccct 1500

ctgcaatctt acggcttcca gcccaccaac ggcgtgggct accagcccta cagagtggtt 1560

gtgctgagct tcgagctgct gcacgcccca gctacagtgt gcggccctaa gaaatctaca 1620

aacctggtca agaacaagtg tgtgaacttc aacttcaatg gcctgacggg caccggcgtg 1680

ctgacagaga gcaacaagaa gttcctgcct ttccagcaat ttggcagaga catcgccgac 1740

accaccgacg ccgtgcgcga ccctcagacc ctggaaattc tggacatcac cccatgttct 1800

ttcggcggcg tgtccgtcat tacgccaggc accaatacca gcaaccaggt ggccgtgctt 1860

tatcaggatg tgaattgtac cgaagttcct gttgcaatcc acgccgacca actgaccccc 1920

acatggagag tgtactctac cggcagcaac gtgttccaaa cgagagccgg atgcctgatt 1980

ggagctgagc atgtgaacaa cagctacgag tgcgatattc caatcggagc cggcatctgc 2040

gcctcctacc aaacacaaac caactcccct cgtagagcga gaagcgtggc ctctcagagc 2100

atcatcgcct acaccatgag cctgggtgcc gaaaactccg tggcttactc caacaacagc 2160

atcgccatcc ctacaaattt caccatcagc gtgacaaccg agatcctgcc tgtgtccatg 2220

accaagacca gcgtggactg cacgatgtac atctgcggag atagcaccga gtgcagcaat 2280

ctgctactgc agtatggcag cttctgcacc caactgaaca gagcactgac cggcattgct 2340

gtggaacagg acaagaatac ccaggaggtg ttcgcccaag tgaagcagat ttacaagaca 2400

ccccctatca aggacttcgg aggcttcaac ttcagccaga tcctgcctga ccctagcaag 2460

ccaagcaaaa gatcctttat cgaagatctg ctgtttaaca aggtgacact ggccgatgcc 2520

ggctttatca agcagtacgg cgactgcctg ggagacatcg ccgccagaga cctgatctgt 2580

gctcagaaat ttaacgggct gaccgtgctg ccacctctgc tgacagatga gatgatcgct 2640

cagtacacca gcgccctgct ggccggcaca attacctccg gctggacctt cggagccgga 2700

gccgccctgc agatcccctt cgccatgcag atggcctacc ggttcaatgg catcggcgtc 2760

acccaaaacg tgctctatga gaaccagaag ctgatcgcaa accagttcaa ctccgccatc 2820

ggtaagatcc aggacagtct gagcagcacg gcgtctgccc tgggcaagct ccaggacgtg 2880

gtgaaccaga acgcccaggc ccttaacacc ctggtgaaac aactgagcag caacttcggt 2940

gccatttcca gcgttctcaa tgacatcctg agcagactgg ataaggtgga agccgaggtg 3000

cagatcgacc ggctgatcac cggacggctg cagagcctgc agacgtacgt gacccagcaa 3060

ttaatcagag ctgccgagat cagagccagc gccaatctgg ctgccaccaa aatgagcgaa 3120

tgtgtgctgg gccagtcaaa gagagtggat ttttgtggca aaggctacca cctgatgtcc 3180

ttccctcagt ctgcccctca cggcgtggtg ttcctccatg tgacctatgt gcctgctcag 3240

gagaagaact ttaccacagc ccctgctatc tgccacgacg gaaaggccca cttccccaga 3300

gagggcgtgt ttgtgtccaa tggcacacac tggttcgtga cccaaagaaa cttctacgag 3360

ccccagatca tcaccacaga caacaccttc gtgagcggca actgcgacgt ggtgatcggc 3420

atcgtgaaca acacagtgta cgacccccctg caacctgagc tggacagctt caaagaggaa 3480

ctggacaaat acttcaagaa tcacaccagc cctgatgtgg atctgggcga catcagcggc 3540

atcaacgcca gcgtcgtgaa catccagaag gaaatcgaca gactgaacga agtggccaag 3600

aacctgaacg agagcctcat cgatctgcag gagctgggca agtacgagca gtacatcaaa 3660

tggcctttct tcttcatcat cggcctgatt atcggcctgt tcctcgtgct gagagtgggc 3720

atccacctgt gcatcaagct taagcacaca aaaaagcggc agatttacac cgacatcgag 3780

atgaaccggc tgggcaaatg agaattccac gtgggatcc 3819

<210> 28

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 28

acaggtacgt taatagttaa tagcgt 26

<210> 29

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 29

atattgcagc agtacgcaca ca 22

<210> 30

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic probe

<400> 30

acactagcca tccttactgc gcttcg 26

<210> 31

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 31

ccacggagaa agacctcatc tg 22

<210> 32

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 32

gggtcacctc atgttggaaa taaa 24

<210> 33

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223>

Description of Artificial Sequence: Synthetic primer

<400> 33

gttgccaaac cttatcagaa atga 24

<210> 34

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 34

ttctggcccg tggttctct 19

<210> 35

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 35

aaatggcggg ttctgtgc 18

<210> 36

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 36

aatatcctct gggtcttgta gatgg 25

<210> 37

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 37

ggagtggctg agccatcgt 19

<210> 38

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 38

agctggttgt ctttgagaga catg 24

<210> 39

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 39

ggccatccag aggagcatag 20

<210> 40

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 40

tttctccatg ctgctgttga a 21

<210> 41

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 41

actgccgcat cctcttcct 19

<210> 42

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223>

Description of Artificial Sequence: Synthetic primer

<400> 42

tcgttgccaa tggtgatgac 20

<210> 43

<211> 4

<212> PRT

<213> Severe acute respiratory syndrome coronavirus 2

<400> 43

Arg Arg Ala Arg

1

<210> 44

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 44

Gly Ser Ala Ser

1