(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0079057
(43) 공개일자 2023년06월05일

(51) 국제특허분류(Int. Cl.)
C07K 14/005 (2006.01) *A61K 39/00* (2006.01)
A61K 39/215 (2006.01) *A61P 31/14* (2006.01)
C12N 15/62 (2006.01) *C12N 15/82* (2006.01)
C12N 7/00 (2006.01)
(52) CPC특허분류
C07K 14/005 (2013.01)
A61K 39/215 (2013.01)
(21) 출원번호 10-2023-7010167
(22) 출원일자(국제) 2021년08월31일
심사청구일자 없음
(85) 번역문제출일자 2023년03월24일
(86) 국제출원번호 PCT/CA2021/051201
(87) 국제공개번호 WO 2022/047575
국제공개일자 2022년03월10일
(30) 우선권주장
63/073,327 2020년09월01일 미국(US)
63/211,716 2021년06월17일 미국(US)

(71) 출원인
메디카고 인코포레이티드
캐나다, 퀘벡 쥐1브이 3브이9, 세인트-포이, 슈트 600, 루트드 레그리즈, 1020
(72) 발명자
라보이에 피에르-올리비에르
캐나다 쥐1엔 2엘7 퀘벡주 퀘벡 프랭클린 621
다우스트 마르크-안드레
캐나다 쥐1엑스 4엔4 퀘벡주 퀘벡 뤼 루이스-피셋 3925
(74) 대리인
김진희, 김태홍

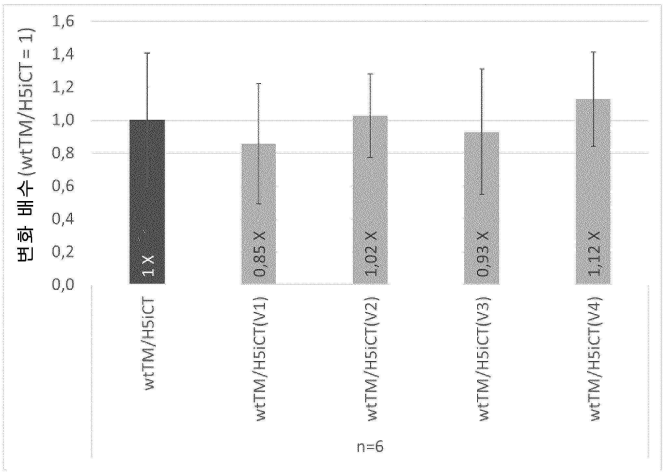
전체 청구항 수 : 총 60 항

(54) 발명의 명칭 **변형된 코로나바이러스 구조 단백질**

(57) 요 약

변형된 코로나바이러스 스파이크(S) 단백질, 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP), 및 변형된 S 단백질을 코딩하는 핵산을 제공한다. 숙주 또는 숙주 세포에서의 변형된 S 단백질 및 VLP 생산 방법 또한 기술한다. 변형된 S 단백질은 막횡단 도메인(TM) 또는 TM의 일부, 및 사이토졸 테일(CT) 또는 CT의 일부를 포함할 수 있고, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것이고, 여기서, TM 또는 TM의 일부는 CT 또는 CT의 일부와 이종성이다. 추가로, 변형된 코로나바이러스(S)-단백질 또는 VLP를 포함하는 조성물을 피험체에게 투여하는 단계를 포함하는, 피험체에서 코로나바이러스 감염에 대한 면역을 유도하는 방법을 기술한다.

대 표 도 - 도5b



(52) CPC특허분류
A61P 31/14 (2018.01)
C12N 15/62 (2013.01)
C12N 15/8258 (2013.01)
C12N 7/00 (2013.01)
C07K 2319/00 (2013.01)
C12N 2760/16122 (2013.01)
C12N 2770/20022 (2013.01)
C12N 2770/20023 (2013.01)
C12N 2770/20034 (2013.01)

명 세 서

청구범위

청구항 1

연속적으로,

- 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

- 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는

- 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

- 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT

를 포함하는 키메라 TMCT인 TMCT

를 포함하는, 변형된 코로나바이러스 S 단백질.

청구항 2

제1항에 있어서, TM이 CT에 직접 융합된 것인 변형된 코로나바이러스 S 단백질.

청구항 3

제1항에 있어서, TM이 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열 및 인플루엔자 HA 단백질 TM으로부터 유래된 C 말단 서열을 포함하는 키메라 TM인 변형된 코로나바이러스 S 단백질.

청구항 4

제3항에 있어서, 키메라 TM가, 서열 번호 18 또는 서열 번호 169의 아미노산 1-20에 상응하는 적어도 20개의 아미노산, 또는 서열 번호 118 또는 164의 아미노산 1-21에 상응하는 적어도 21개의 아미노산, 또는 서열 번호 123의 아미노산 1-22에 상응하는 적어도 22개의 아미노산을 포함하는 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산을 포함하는 것인 변형된 코로나바이러스 S 단백질.

청구항 5

제4항에 있어서, 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산이 AGL 또는 AGL의 보존 치환, MAGL 또는 MAGL의 보존 치환으로부터 선택되는 것인 변형된 코로나바이러스 S 단백질.

청구항 6

제1항에 있어서, CT가 코로나바이러스 S 단백질 CT로부터 유래된 N 말단 서열 및 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열을 포함하는 키메라 CT인 변형된 코로나바이러스 S 단백질.

청구항 7

제6항에 있어서, 키메라 CT가 서열 번호 18, 126, 128, 129, 130 또는 131의 아미노산 27-37에 상응하는 아미노산; 또는 서열 번호 127의 아미노산 27-36에 상응하는 아미노산을 포함하는 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열, 및 코로나바이러스 S 단백질 CT의 N 말단 단부로부터의 하나 이상의 아미노산을 포함하는 것인 변형된 코로나바이러스 S 단백질.

청구항 8

제7항에 있어서, 코로나바이러스 S 단백질 CT의 N 말단 단부로부터의 하나 이상의 아미노산이 C 또는 C의 보존 치환, CC 또는 CC의 보존 치환, 또는 CCM 또는 CCM의 보존 치환으로부터 선택되는 것인 변형된 코로나바이러스

S 단백질.

청구항 9

제3항에 있어서, 키메라 TM이 서열 번호 18 또는 서열 번호 169의 아미노산 1-20에, 또는 서열 번호 118 또는 서열 번호 164의 아미노산 1-21에, 또는 서열 번호 123의 아미노산 1-22에 상응하는 아미노산을 포함하는 것인 변형된 코로나바이러스 S 단백질.

청구항 10

제4항에 있어서, 키메라 CT가 서열 번호 18, 126, 128, 129, 130 또는 131의 아미노산 27-37; 또는 서열 번호 127의 아미노산 27-36에 상응하는 아미노산을 포함하는 것인 변형된 코로나바이러스 S 단백질.

청구항 11

제1항에 있어서, 키메라 TMCT가 서열 번호 18 또는 서열 번호 169의 아미노산 1-20에, 또는 서열 번호 118 또는 서열 번호 164의 아미노산 1-21에, 또는 서열 번호 123의 아미노산 1-22에 상응하는 아미노산을 포함하는 키메라 TM, 및 서열 번호 18, 126, 128, 129, 130 또는 131의 아미노산 27-37에; 서열 번호 127의 아미노산 27-36에 상응하는 아미노산 또는 그의 조합을 포함하는 키메라 CT를 포함하는 것인 변형된 코로나바이러스 S 단백질.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 상기 S 단백질이 야생형 코로나바이러스 S 단백질 아미노산 서열과 비교하였을 때, 하나 이상의 아미노산 치환을 포함하는 것인 변형된 S 단백질.

청구항 13

제12항에 있어서, 하나 이상의 아미노산 치환이

i) S1과 S2 서브유닛 사이의 절단 부위에서의 프로세싱을 제한하는 치환,

ii) 하나 이상의 아미노산의 하나 이상의 프롤린으로의 치환, 또는

iii) S1과 S2 서브유닛 사이의 절단 부위에서의 프로세싱을 제한하는 치환 및 하나 이상의 아미노산의 하나 이상의 프롤린으로의 치환

을 포함하는 것인 변형된 S 단백질.

청구항 14

제12항 또는 제13항에 있어서, 하나 이상의 아미노산 치환이 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 971 및 972의 아미노산에 상응하는 것인 변형된 S 단백질.

청구항 15

제12항 또는 제13항에 있어서, 하나 이상의 아미노산 치환이 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 802, 927, 971 및 972의 아미노산에 상응하는 것인 변형된 S 단백질.

청구항 16

제14항 또는 제15항에 있어서, 변형된 S 단백질이 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 또는 그의 조합의 아미노산에 상응하는 하나 이상의 아미노산 치환을 추가로 포함하는 것인 변형된 S 단백질.

청구항 17

제16항에 있어서, 서열 번호 2의 위치 667의 아미노산에 상응하는 아미노산의 치환이 글리신, 또는 글리신의 보존 치환으로의 치환이고, 서열 번호 2의 위치 668에 상응하는 아미노산의 치환이 세린, 또는 세린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 670에 상응하는 아미노산의 치환이 세린, 또는 세린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 971 및 972의 아미노산에 상응하는 아미노산의 치환이 프롤린, 또는 프롤린의 보

존 치환으로의 치환인 변형된 S 단백질.

청구항 18

제16항에 있어서, 서열 번호 2의 위치 667의 아미노산에 상응하는 아미노산의 치환이 글리신, 또는 글리신의 보존 치환으로의 치환이고, 서열 번호 2의 위치 668에 상응하는 아미노산의 치환이 세린, 또는 세린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 670에 상응하는 아미노산의 치환이 세린, 또는 세린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 802, 927, 971 및 972의 아미노산에 상응하는 아미노산의 치환이 프롤린, 또는 프롤린의 보존 치환으로의 치환인 변형된 S 단백질.

청구항 19

제14항 내지 제18항 중 어느 한 항에 있어서, 변형된 S 단백질이 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 923의 아미노산에 상응하는 아미노산 치환을 추가로 포함하는 것인 변형된 S 단백질.

청구항 20

제19항에 있어서, 서열 번호 2의 위치 667의 아미노산에 상응하는 아미노산의 치환이 글리신, 또는 글리신의 보존 치환으로의 치환이고, 서열 번호 2의 위치 668에 상응하는 아미노산의 치환이 세린, 또는 세린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 670에 상응하는 아미노산의 치환이 세린, 또는 세린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 802, 927, 971 및 972의 아미노산에 상응하는 아미노산의 치환이 프롤린, 또는 프롤린의 보존 치환으로의 치환이고, 서열 번호 2의 위치 923에 상응하는 아미노산의 치환이 페닐알라닌, 또는 페닐알라닌의 보존 치환으로의 치환인 변형된 S 단백질.

청구항 21

제12항 내지 제20항 중 어느 한 항에 있어서, 하나 이상의 치환이 융합 전 상태로 S 단백질을 유지시키거나, 또는 숙주 또는 숙주 세포에서 발현되었을 때, 숙주 또는 숙주 세포에서 발현된, 하나 이상의 치환을 포함하지 않는 상응하는 S 단백질의 수율과 비교하였을 때, 더 높은 수율로 변형된 S 단백질을 생산하는 것인 변형된 S 단백질.

청구항 22

제1항 내지 제21항 중 어느 한 항에 있어서, S 단백질이 S1 서브유닛 및 S2 서브유닛을 포함하고, 여기서, S2 서브유닛이 키메라 TMCT를 포함하는 것인 변형된 S 단백질.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 인플루엔자 HA 단백질이 인플루엔자 타입 B 또는 인플루엔자 서브타입 H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, 또는 H16으로부터 유래된 것인 변형된 S 단백질.

청구항 24

제1항 내지 제23항 중 어느 한 항에 있어서, 코로나바이러스 S 단백질이 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래된 것인 변형된 S 단백질.

청구항 25

제1항 내지 제23항 중 어느 한 항에 있어서, 엑토도메인이 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래되고, TM 또는 TM의 일부가 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래되거나, 또는 엑토도메인, 및 TM 또는 TM의 일부가 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래된 것인 변형된 S 단백질.

청구항 26

제1항 내지 제25항 중 어느 한 항에 있어서, TMCT가 서열 번호 18, 19, 37, 38, 39, 64, 126, 127, 128, 129, 130, 131, 118, 119, 120, 123, 124, 125, 134, 135, 164, 165, 166, 169, 170, 171, 172 또는 173의 서열과

약 80% 내지 약 100%의 동일성을 가진 서열을 포함하는 것인 변형된 S 단백질.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 변형된 S 단백질이 변형된 S 단백질 및 신호 펩티드를 포함하는 전구체 단백질로서 생산되는 것인 변형된 S 단백질.

청구항 28

제27항에 있어서, 신호 펩티드가 변형된 S 단백질에 대해 천연 또는 비천연인 변형된 S 단백질.

청구항 29

제27항 또는 제28항에 있어서, 신호 펩티드가 단백질 디술피드 이소머라제(PDI)의 신호 펩티드로부터 유래된 것인 변형된 S 단백질.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 변형된 S 단백질이 식물 특이적 N-글리칸을 포함하는 것인 변형된 S 단백질.

청구항 31

제1항에 있어서, S 단백질의 아미노산 서열이 서열 번호 5, 21, 30, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 95, 96, 97, 108, 109, 110, 144, 145, 146, 155, 156 또는 157의 아미노산과, 또는 서열 번호 47의 아미노산 24-1259, 서열 번호 48의 아미노산 25-1259, 서열 번호 49의 아미노산 25-1259, 서열 번호 50의 아미노산 25-1259, 서열 번호 51의 아미노산 25-1259, 서열 번호 52의 아미노산 25-1259, 서열 번호 53의 아미노산 25-1259, 서열 번호 54의 아미노산 25-1259, 서열 번호 55의 아미노산 25-1259, 서열 번호 56의 아미노산 25-1259, 서열 번호 57의 아미노산 25-1259, 서열 번호 58의 아미노산 25-1259, 서열 번호 59의 아미노산 25-1262, 서열 번호 60의 아미노산 25-1261, 서열 번호 61의 아미노산 25-1258, 또는 서열 번호 62의 아미노산 25-1256, 서열 번호 95의 아미노산 25-1243, 서열 번호 96의 아미노산 25-1240, 서열 번호 97의 아미노산 25-1243, 서열 번호 108의 아미노산 25-1341, 서열 번호 109의 아미노산 25-1338, 서열 번호 110의 아미노산 25-1341, 또는 서열 번호 144의 아미노산 25-1351, 서열 번호 145의 아미노산 25-1348, 서열 번호 146의 아미노산 25-1351, 서열 번호 155의 아미노산 25-1159, 서열 번호 156의 아미노산 25-1156, 또는 서열 번호 157의 아미노산 25-1159와 80% 내지 100%의 동일성을 포함하는 것인 변형된 S 단백질.

청구항 32

제1항 내지 제30항 중 어느 한 항에 있어서, CT 또는 CT의 일부가 서열 번호 15의 서열과, 또는 서열 번호 6, 8, 7, 9, 10, 12, 13 또는 14의 아미노산 35-50과, 또는 서열 번호 11의 아미노산 34-49와, 또는 서열 번호 3의 아미노산 553-568과, 또는 서열 번호 18의 아미노산 22-37과, 또는 서열 번호 19의 아미노산 21-40과, 또는 서열 번호 37의 아미노산 21-39와, 또는 서열 번호 38의 아미노산 25-36과, 또는 서열 번호 39의 아미노산 24-34와, 또는 서열 번호 126, 128, 129, 130 또는 131의 아미노산 22-37과, 또는 서열 번호 127의 아미노산 22-36과 80% 내지 100%의 동일성을 포함하는 것인 변형된 S 단백질.

청구항 33

제27항에 있어서, 전구체 단백질이 서열 번호 1의 아미노산 1-1234와, 또는 서열 번호 5의 아미노산 1-1234와, 서열 번호 30의 아미노산 1-1243과, 서열 번호 95의 아미노산 1-1227과, 서열 번호 108의 아미노산 1-1325와, 서열 번호 112의 아미노산 1-1216과, 서열 번호 113의 아미노산 1-1318과, 서열 번호 144의 아미노산 1-1335와, 서열 번호 155의 아미노산 1-1143과, 서열 번호 158의 아미노산 1-1325와, 서열 번호 159의 아미노산 1-1135와 80% 내지 100%의 동일성을 포함하고, 여기서, CT의 아미노산 서열이 서열 번호 15의 서열과, 또는 서열 번호 6, 8, 7, 9, 10, 12, 13 또는 14의 아미노산 35-50과, 또는 서열 번호 11의 아미노산 34-49와, 또는 서열 번호 3의 아미노산 553-568과 80% 내지 100%의 동일성을 포함하는 것인 변형된 S 단백질.

청구항 34

제1항 내지 제33항 중 어느 한 항의 변형된 S 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산.

청구항 35

제1항 내지 제33항 중 어느 한 항의 변형된 S 단백질을 포함하는 삼량체.

청구항 36

제1항 내지 제33항 중 어느 한 항의 변형된 S 단백질, 또는 제35항의 삼량체를 포함하는 바이러스 유사 입자(VLP).

청구항 37

제36항에 있어서, VLP가 식물 지질을 추가로 포함하는 것인 VLP.

청구항 38

유효 용량의, 제1항 내지 제33항 중 어느 한 항의 변형된 S 단백질, 제35항의 삼량체, 또는 제36항 또는 제37항의 VLP, 및 약학적으로 허용되는 담체, 애주번트, 비히클 또는 부형제를 포함하는 조성물.

청구항 39

유효 용량의, 제1항 내지 제33항 중 어느 한 항의 변형된 S 단백질, 제35항의 삼량체, 제36항 또는 제37항의 VLP, 또는 제38항의 조성물을 포함하는, 면역 반응을 유도하기 위한 백신.

청구항 40

제39항에 있어서, 백신이 VLP의 혼합물을 포함하는 다가 백신인 백신.

청구항 41

피험체에게 제38항의 조성물, 또는 제39항 또는 제40항의 백신을 투여하는 단계를 포함하는, 피험체에서 코로나바이러스 감염에 대한 면역을 유도하는 방법.

청구항 42

제41항에 있어서, 조성물 또는 백신이 피험체에게 1회 투여되는 것인 방법.

청구항 43

제41항에 있어서, 조성물 또는 백신이 피험체에게 다회 투여되는 것인 방법.

청구항 44

제41항에 있어서, 조성물 또는 백신이 초기 용량으로 투여되고, 초기 용량 투여로부터 1일 내지 6개월 사이에 1회 이상의 후속 용량이 투여되는 것인 방법.

청구항 45

제38항의 조성물, 또는 제39항 또는 제40항의 백신을 사용하여 제조된 항체 또는 항체 단편.

청구항 46

제1항 내지 제33항 중 어느 한 항의 변형된 S 단백질, 제35항의 삼량체, 제34항의 핵산, 또는 제36항 또는 제37항의 VLP를 포함하는 비인간 숙주 또는 숙주 세포.

청구항 47

비인간 숙주 또는 숙주 세포에서 변형된 S 단백질을 생산하는 방법으로서,

a) 제34항의 핵산을 비인간 숙주 또는 숙주 세포 내로 도입하거나, 또는 제34항의 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계, 및

b) 핵산의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 변형된 S 단백질을 생산

하는 단계

를 포함하는, 비인간 숙주 또는 숙주 세포에서 변형된 S 단백질을 생산하는 방법.

청구항 48

비인간 숙주 또는 숙주 세포에서 바이러스 유사 입자(VLP)를 생산하는 방법으로서,

a) 제34항의 핵산을 비인간 숙주 또는 숙주 세포 내로 도입하거나, 또는 제34항의 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계, 및

b) 핵산의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 VLP를 생산하는 단계

를 포함하는, 비인간 숙주 또는 숙주 세포에서 VLP를 생산하는 방법.

청구항 49

제47항 또는 제48항에 있어서, 방법이 비인간 숙주 또는 숙주 세포를 수확하는 단계인 단계 c)를 추가로 포함하는 것인 방법.

청구항 50

비인간 숙주 또는 숙주 세포에서 변형된 코로나바이러스 S 단백질의 생산 수율을 증가시키는 방법으로서,

a) 제34항의 핵산을 숙주 또는 숙주 세포 내로 도입하거나, 또는 제34항의 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계; 및

b) 핵산에 의해 코딩되는 변형된 S 단백질의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜, 비변형된 S 단백질을 발현하는 숙주 또는 숙주 세포와 비교하여 더 높은 수율로 변형된 S 단백질을 생산하는 단계

를 포함하는, 비인간 숙주 또는 숙주 세포에서 변형된 코로나바이러스 S 단백질의 생산 수율을 증가시키는 방법.

청구항 51

비인간 숙주 또는 숙주 세포에서 바이러스 유사 입자(VLP)의 생산 수율을 증가시키는 방법으로서,

a) 연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 코딩하는 핵산을 비인간 숙주 또는 숙주 세포 내로 도입하는 단계로서, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 단계, 또는

연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 코딩하는 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계로서, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 단계; 및

b) 핵산에 의해 코딩되는 변형된 S 단백질의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜, 비변형된 S 단백질을 발현하는 숙주 또는 숙주 세포에서의 VLP의 수율과 비교하여 더 높은 수율로 변형된 S 단백질을 포함하는 VLP를 생산하는 단계

를 포함하는, 비인간 숙주 또는 숙주 세포에서 VLP의 생산 수율을 증가시키는 방법.

청구항 52

비인간 숙주 또는 숙주 세포에서 바이러스 유사 입자(VLP)를 생산하는 방법으로서,

a) 연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 코딩하는 핵산을 비인간 숙주 또는 숙주 세포 내로 도입하는 단계

로서, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 단계; 또는

연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 코딩하는 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계로서, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 단계; 및

b) 핵산의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 VLP를 생산하는 단계

를 포함하는, 비인간 숙주 또는 숙주 세포에서 VLP를 생산하는 방법.

청구항 53

제48항, 제51항 또는 제52항 중 어느 한 항에 있어서, VLP가 추가로 비인간 숙주 또는 숙주 세포로부터 추출 및 정제되는 것인 방법.

청구항 54

제48항, 제51항 또는 제52항 중 어느 한 항의 방법에 의해 생산된 VLP.

청구항 55

제47항 또는 제50항의 방법에 의해 생산된 변형된 S 단백질.

청구항 56

제46항 내지 제55항 중 어느 한 항에 있어서, 비인간 숙주 또는 숙주 세포가 식물, 식물의 일부, 식물 세포, 진균, 진균 세포, 곤충, 곤충 세포, 동물 또는 동물 세포를 포함하는 것인 방법, 또는 비인간 숙주 또는 숙주 세포.

청구항 57

바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 포함하고, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래되고, 상기 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 971 및 972에 치환을 포함하는 것인 조성물.

청구항 58

바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 포함하고, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래되고, 상기 S 단백질은 위치 667에 글리신 치환, 위치 668에 세린 치환, 위치 670에 세린 치환, 위치 971에 프롤린 치환, 및 위치 972에 프롤린 치환을 포함하고, 위치는 서열 번호 2의 참조 아미노산 서열에 상응하는 것인 조성물.

청구항 59

바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 변형된 코로나바이러스 S 단백질을 포함하고, 변형된 S 단백질은 서열 번호 21의 서열, 또는 서열 번호 51의 아미노산 25-1259를 포함하는 것인 조성물.

청구항 60

바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 연속적으로, 코로나바이러스 S 단백질로부터 유래된 엑

토도메인, 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 코로나바이러스 S 단백질을 포함하고, 여기서, TMCT는 막횡단 도메인(TM) 및 사이토졸 테일(CT)을 포함하는 키메라 TMCT이고, 상기 TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래되고, 상기 CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래되고, 상기 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 802, 923, 927, 971 및 972에 치환을 포함하는 것인 조성물.

발명의 설명

기 술 분 야

[0001] **본 발명의 분야**

[0002] 본 개시내용은 변형된 바이러스 구조 단백질에 관한 것이다. 본 발명은 또한 변형된 바이러스 구조 단백질을 포함하는 바이러스 유사 입자(VLP: virus-like particle), 및 숙주 또는 숙주 세포에서 VLP를 생산하는 방법에 관한 것이다.

배 경 기 술

[0003] **배경**

[0004] 코로나바이러스(CoV)는 코로나비리다에(Coronaviridae) 과, 아테리비리다에(*Arteriviridae*) 과, 메소니비리다에(*Mesoniviridae*) 과, 및 로니비리다에(*Roniviridae*) 과를 포함하는 니도바이랄레스(*Nidovirales*) 목에 속하는 가장 큰 바이러스 그룹이다. 코로나비리나에(*Coronavirinae*)는 코로나비리다에 과 중의 두 아과 중 하나를 구성하고, 여기서, 나머지 다른 한 아과는 토로비리나에(*Torovirinae*)이다. 코로나비리나에는 알파, 베타, 감마 및 델타 코로나바이러스인 4개 속으로 추가로 세분된다. 알파 코로나바이러스 및 베타 코로나바이러스의 구성원은 포유동물에서만 발견된다. 알파코로나바이러스 속은 두 가지 인간 바이러스 종인 HCoV-229E와 HCoV-NL63을 포함한다. 중요한 동물 알파코로나바이러스는 돼지의 전염성 위장염 바이러스 및 고양이 전염성 복막염 바이러스이다.

[0005] 중증 급성 호흡기 증후군 코로나바이러스 2(SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2, 이는 또한 2019-nCoV 및 HCoV-19로도 공지)는 신규 계통 B 베타코로나바이러스(베타-CoV)로서, 사망률과 이환율이 높아 전 세계적으로 공중 보건에 중대한 영향을 미치는 호흡기 질환인 코로나바이러스 질환 2019(COVID-19)를 유발한다. 예컨대, 2020년 시작된 팬데믹과 같은 SARS-CoV-2의 발생은 바이러스의 인큐베이션 기간 및 전염성으로 인해 의료 시스템에 가장 중요한 도전과제이다. COVID-19에 대한 치료가 시급히 필요하지만, SARS-CoV-2 발생을 장기적으로 관리하려면 효과적인 백신이 필요할 것이다.

[0006] 코로나바이러스 비리온은 저온 전자 단층 촬영 및 저온 전자 현미경법에 의한 최근 연구에서 나타난 바와 같이 직경이 대략 118-140 nm인 구형이다.

[0007] 코로나바이러스의 가장 두드러진 구조적 특징은 비리온 표면에서 나오는 곤봉 형상의 스파이크 돌기이다. 코로나바이러스 입자는, 스파이크(S), 막(M), 및 외피(E) 단백질, 및 일부 코로나바이러스의 경우에만 헤마글루티닌-에스테라제(HE: hemagglutinin-esterase) 단백질인 3 또는 4가지 유형의 구조 단백질이 삽입되어 있는 지질 이중층에 의해 둘러싸여 있는, 뉴클레오캡시드(N) 인단백질과 바이러스 게놈 RNA 사이의 회합에 의해 형성된 나선형 뉴클레오캡시드 구조로 구성된다(문헌 [Masters PS. The molecular biology of coronaviruses. Adv Virus Res. 2006; 66:193-292.]).

[0008] 막(M) 단백질은 비리온에서 가장 풍부한 구조 단백질이다. 이는 3개의 막횡단 도메인을 가진 작은(~25-30 kDa) 단백질이고, 그의 형상이 비리온에 부여되는 것으로 간주된다. 외피(E) 단백질은 크기가 8.4 내지 12 kDa 범위인, 76-109개의 아미노산으로 구성된 짧은 내재성 막 단백질이다. 1차 및 2차 구조는 E가 7-12개의 아미노산으로 구성된 짧은 친수성 아미노 말단을 가지고 있고, 그 뒤에 25개의 아미노산으로 구성된 큰 소수성 막횡단 도메인(TMD: transmembrane domain)이 있으며, 상기 단백질의 대부분을 차지하는 긴 친수성 카복실 말단으로 끝난다는 것을 보여준다. E 단백질은 예컨대, 어셈블리, 출아, 외피 형성 및 병인과 같은 바이러스 수명 주기의 여러 측면에 관여한다.

[0009] 스파이크(S) 단백질은 많은 코로나바이러스에 대한 숙주 수용체의 인식 뿐만 아니라, 바이러스가 세포로 진입하기 위한 바이러스 및 숙주 세포막의 융합에 필요한 당단백질이다(문헌 [Belouzard et al., Viruses 2012 Jun;4(6):1011-33]). 바이러스 외피 표면의 1차 당단백질로서 코로나비리다에의 S 단백질은 SARS-CoV-2 감염을

비롯한 자연 감염에 의해 유도된 중화 항체의 주요 표적이며, 실험 백신 후보에서 표적화되는 중요한 항원이다.

[0010] SARS-CoV-2 S 단백질은 다른 코로나바이러스의 S 단백질과 마찬가지로 초기에 전구체 단백질로 합성된다. 개별 전구체 S 단백질은 동종삼량체를 형성하고, 골지 구획 내에서 글리코실화될 뿐만 아니라, 신호 펩티드를 제거하기 위해 프로세싱된다. S 단백질은 막 융합을 촉진하기 위해 2단계 프로테아제 매개 활성화를 필요로 한다. SARS-CoV-2 S 단백질은 S1/S2 연접부에 있는 RRAR 퓨린 절단 부위로 구별되는데, 이는 골지 구획에서 프로세싱되어 융합 전 입체형태의 동종삼량체 내에서 S1/S2 프로토머로 비공유적으로 결합된 상태 그대로 유지되는 S1 및 S2 폴리펩티드(또는 서브유닛)라는 별개의 두 폴리펩티드가 생성되는 것으로 추정된다(문헌 [Walls et al. Cell 2020 181(2) p281-292]; [Li et al. eLife 2019; 8: e51230.]). S1/S2 연접부에서의 퓨린 절단 및 융합 펩티드의 상류의 S2' 부위에서의 절단은 바이러스가 세포 표면 또는 엔도솜에 진입하는 동안 발생하며, 여러 프로테아제에 의해 매개될 수 있다.

[0011] 이 삼량체는 수용체 결합 도메인(RBD) 에피토프를 통해 숙주 세포 상의 표적 수용체에 결합하기 전에 융합 전 입체형태로 유지된다. 수용체 결합은 융합 전 삼량체를 불안정화시켜, 그 결과로 S1 서브유닛은 배출되고, S2 서브유닛은 바이러스의 세포막에의 융합을 통해 안정한 융합 후 형태로 전이된다(문헌 [Wrapp et al. Science, 13 Mar 2020, Vol. 367, Issue 6483, pp. 1260-1263]). SARS-CoV-2에 감염된 개체의 중화 항체는 S 단백질의 S1 서브유닛의 RBD를 표적화하는 것으로 나타났다(문헌 [Premkumar, L., 2020 *Science Immunology* 11 Jun 2020: Vol. 5, Issue 48]).

[0012] 융합 전 형태에서의 S 단백질 엑토도메인의 안정화는 가능하게는 더욱 안정한 융합 후 구조를 취하고자 하는 경향으로 인해 발생하는 트리거링 또는 미스폴딩을 방지함으로써 재조합 발현 수율을 증가시키는 경향이 있다(문헌 [Hsieh et al. Science 2020, 369 p.1501-1505]).

[0013] S 단백질 엑토도메인에 대한 돌연변이가 융합 전 입체형태의 안정화를 촉진하는 것으로 나타났다. (Pallesen, J. et al.)의 WO 2018/081318 및 그의 동반 간행물(문헌 [PNAS August 29, 2017 114 (35)])에는 융합 전 입체형태의 MERS-CoV 스파이크 단백질의 S 엑토도메인 삼량체를 안정화시키는 헵타드 반복부 1(HR1: heptad repeat 1)과 중심 나선 사이의 연접부에서의, 또는 그 인근에서의 이중 프롤린 치환, 및 S 엑토도메인의 S1/S2 절단 부위 및 S2' 절단 부위에서의 프로테아제 절단을 막는 치환이 개시되어 있다. 상동성 아미노산 잔기에서 이중 프롤린 치환으로 안정화된 SARS-CoV-2 S 단백질은 저온-EM에 의해 고해상도 구조를 결정하는 데 사용되었다(문헌 [Wrapp et al Science 2020 367, 1260-1263]; [Walls et al. Cell 2020, 181, 281-292]). 추가로, 퓨린 인식 부위의 파괴는 S 단백질을 융합 전 입체형태로 유지시키는 것으로 간주된다(문헌 [Wrapp et al Science 2020 367, 1260-1263]). 그러나, 이러한 치환에도 불구하고, SARS-CoV-2 S 단백질 엑토도메인은 여전히 불안정하고, 포유동물 세포에서 신뢰할 수 있게 생산하기가 어렵기 때문에 효과적인 고수율의 서브유닛 백신 개발을 방해한다(문헌 [Hsieh et al. Science 2020, 369 p.1501-1505]).

[0014] (Hsieh et al.)(문헌 [Science 2020, 369 p.1501-1505])은 이전에 결정된 저온-EM 구조에 기초하여 100개 이상의 구조 유도 SARS-CoV-2 스파이크 단백질 변이체를 디자인하고, 포유동물 세포에서 발현시켰다. 변이체를 생화학적, 생물물리학적 및 구조적으로 특징화하여 수율 및 안정성을 증가시키는 치환을 확인하였다.(Hsieh et al.) 등은 이중 프롤린 치환에 비해 스파이크의 발현 및/또는 안정성을 증가시키는 다중 프롤린, 디술피드 결합, 염 브릿지 및 공동 충전 치환을 보고하고 있다. 가장 우수한 것으로 확인된 변이체인 HexaPro는 6개의 유익한 프롤린 치환을 가지며, 이를 통해 그의 모체 구성물보다 10배 더 높은 발현을 보이고, 열 스트레스, 실온 보관 및 다회에 걸친 동결-해동을 견딜 수 있다.

[0015] S2 서브유닛은 3개의 도메인: 큰 엑토도메인, 막횡단 도메인(TM) 및 세포질 테일(CT)로 나눌 수 있다. S 단백질의 세포질 테일은 이전에 어셈블리에 필요한 것으로 밝혀졌다. 코로나비리다에의 CT에서 두 가지 고유한 유지 신호: i) 소포체 회수 신호(ERRS: endoplasmic reticulum retrieval signal) 및/또는 ii) 티로신 의존성 국재화 신호(YxxI 또는 YxxF 모티프)가 발견될 수 있다. ERRS는 코토머 복합체 I(COPI: coatomer complex I)에 결합하는 이염기성 KxHxx 모티프를 포함한다. 모티프는 SARS 막(M) 단백질과 함께 공동 발현될 때, SARS S 단백질을 ERGIC/골지 영역으로 국재화하는 데 필요하며, 국재화는 KxHxx 모티프를 돌연변이시킴으로써 중단될 수 있다(문헌 [McBride et al. J. Virol.Feb 2007, 81 (5) 2418-2428]). ERRS를 포함하는 S 단백질은 COPI 소포체로 동원되고, 골지에서 역행하여 소포체(ER)로 회수된다. ER과 골지 사이의 S 단백질의 반복된 순환을 통해 S 단백질의 세포내 유지가 이루어진다. 알파코로나바이러스 및 베타코로나바이러스의 S 단백질, 둘 모두 ERRS를 포함한다(문헌 [Ujike et al. Journal of General Virology (2016), 97, 1853-1864]).

[0016] 베타코로나바이러스의 S 단백질, 예컨대, MERS-CoV, SARS-CoV 및 SARS-CoV 2의 S 단백질은 ERRS만을 보유하고,

세포내에서 유지될 수 없고, 그로 인해 S 단백질은 원형질막으로 방출된다. ERRS가 결여된 돌연변이체 SARS-CoV S 단백질은 원형질막으로 수송되는 반면, 천연 S 단백질은 M 단백질과 함께 공동 발현될 때 출아 부위 근처에서 M 단백질과 상호작용하여 S 단백질의 세포내 유지를 유도하고, 이는 SARS-CoV의 ERRS가 M 단백질과의 상호작용에 의해 특히 내측 골지 구획 뒤에 S 단백질이 축적되는 것의 원인이 되며, 이를 통해 S 단백질이 VLP 내로 도입된다는 것을 시사하는 것이다(문헌 [Ujike et al. Journal of General Virology (2016), 97, 1853-1864]). 최근, ERRS 제거가 SARS-CoV-2 S 단백질을 렌티바이러스 슈도비리온 내로의 도입을 촉진하는 것으로 밝혀졌다(문헌 [Ou et al., 2020 Nature Communications volume 11, Article number: 1620]).

[0017] (Yu et al.)(문헌 [2020 Science])은 세포질 테일, 및 막횡단 도메인이 다양하게 결실된 SARS-CoV-2 S 단백질의 6가지 변이체를 발현하는 프로토타입 DNA 백신 세트를 구성하였고, 레수스 마카크(rhesus macaque)에서 SARS-CoV-2 바이러스 챌린지에 대한 그의 면역원성 및 방어 효능에 대하여 평가하였다. SARS-CoV-2 S 단백질 엑토도메인의 가용성 단편이 (바이러스 복제를 나타내는) sgmRNA의 수준 감소를 유도한 반면, 전장의 S 단백질 면역원으로 최적의 방어가 달성되었다.

[0018] (Broer et al.)(문헌 [2006 J. Virol. p. 1302-1310])은 SARS-CoV S-슈도타이핑된 레트로바이러스(SARSpp)를 이용하여 SARS-CoV의 감염성 및 막 융합 활성에서의 S 단백질의 막횡단 및 세포질 도메인의 역할을 연구하였다. S의 세포질 도메인이 수포성 구내염 바이러스 G 단백질(VSV-G: vesicular stomatitis virus G protein)로부터 유래된 세포질 도메인으로 대체된 SARSpp는 야생형의 최대 40%까지 감염성이 있었다. 그에 반해, TMD 및 VSV-G의 세포질 도메인, 둘 모두 포함하는 SARSpp는 감염성이 심각하게 손상되었다 (<5%). 이는 S의 TMD가 SARS-CoV의 진입 프로세스에 관여할 수 있다는 것을 보여주는 것이다.

[0019] 백신접종은 감염 전에 가능성이 있는 작용제에 대한 면역 반응을 일으키도록 피험체를 유도함으로써 질환에 대한 방어를 제공한다. 통상적으로, 이는 면역원으로서 생 약독화 또는 전체 불활성화된 형태의 감염원 사용을 통해 달성되었다. 전체 바이러스(예컨대, 사멸 또는 약독화된 바이러스)를 백신으로 사용하는 것의 위험을 피하기 위해, 바이러스 단백질 또는 서브유닛 또는 그의 재조합 버전이 백신으로 추구되었다. 천연 또는 재조합 바이러스 단백질을 백신 제제로 사용하는 데 있어 주요 장애물은 단백질의 입체형태가 그의 자연 환경에서 항원을 모방하도록 하는 것이다. 적합한 애주번트, 및 펩티드의 경우, 담체 단백질을 사용하여 면역 반응을 부스팅시킬 수 있다. 추가로, 백신으로서의 바이러스 단백질 또는 서브유닛은 주로 체액성 반응을 유도할 수 있는 바, 따라서, 지속적인 면역을 유발하지 못할 수 있다. 서브유닛 백신은 전체 불활성화된 바이러스가 우수한 방어를 제공하는 것으로 입증될 수 있는 질환에는 효과가 없을 수 있다.

[0020] 바이러스 유사 입자(VLP)는 면역계에 개선된 항원 제시와 함께 바람직한 입체형태로 바이러스 단백질을 발현하기 위한 면역원성 조성물에 사용될 수 있다. VLP는 성숙한 비리온과 매우 유사하지만, 바이러스 게놈 물질을 포함하지 않으며, 비복제성이기 때문에 백신으로 투여하기에 안전하다. 추가로, VLP는 VLP의 표면에 바이러스 당단백질을 발현하도록 조작될 수 있으며, 이는 그의 고유의 생리학적 구성이다. VLP는 무손상 비리온과 유사하고, 다가 미립자 구조이기 때문에, VLP는 가용성 외피 단백질 항원보다 당단백질에 대한 중화 항체를 유도하는 데 더 효과적일 수 있다.

[0021] 포유동물 세포주, 박테리아, 곤충 세포주, 효모 및 식물 세포를 비롯한 다양한 발현 시스템이 VLP를 생산하는 데 사용되어 왔다. 30개 이상의 상이한 바이러스에 대한 VLP가 백신용으로 곤충 및 포유동물 시스템에서 생성되었다(문헌 [Noad, R. and Roy, P., 2003, Trends Microbiol 11: 438-44]). VLP는 또한 식물에서도 생산되었다(WO2009/076778; WO2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235; WO2010/006452; WO2011/03522; WO 2010/148511; WO2014153674, 및 WO2012/083445 참조).

[0022] VLP는 곤충 및 포유동물 세포의 S 단백질, M 단백질, E 단백질을 비롯한, 중증 급성 호흡기 증후군 코로나바이러스(SARS-CoV 또는 SARS-CoV-1)의 천연 표면 단백질로 생산되었다(문헌 [Liu et al., 2008, J Virol., p. 11318-11330]). SARS-CoV-2 바이러스 유사 입자(VLP)는 또한 포유동물 세포에서 바이러스 표면 단백질 S, M 및 E를 공동 발현함으로써 어셈블리되었다(문헌 [Xu et al. Front. Bioeng. Biotechnol., 30 July 2020]). 연구에 따르면 M 단백질은 바이러스 유사 입자(VLP) 형성에 필수불가결하다는 것이 추가로 나타났다(문헌 [Siu et al. Journal of Virology (2008) 82:11318-11330], [Huang et al. Journal of Virology (2004) 78:12557-12565]). 포유동물 세포에서 막 단백질(M) 및 작은 외피 단백질(E)의 발현은 SARS-CoV-2 VLP의 효율적인 형성 및 방출에 필수적이다(문헌 [Xu et al. Front. Bioeng. Biotechnol., 30 July 2020]). 그럼에도 불구하고, SARS-CoV VLP 어셈블리에 필요한 최소 요건은 여전히 논란의 여지가 있다. (Y. Huang et al.)(문헌 [Journal of Virology (2004) 78:12557-12565])은 형질감염된 인간 세포에서 M 및 N 바이러스 단백질의 공동 발현만을 필요로 하는

VLP의 형성을 기술한 반면, (Siu et al.)(문헌 [Journal of Virology (2008)82:11318-11330])은 형질감염된 포유동물 세포에서 SARS-CoV VLP의 효율적인 생산 및 방출을 위해 E 및 N 단백질, 둘모두 M 단백질과 공동 발현되어야 한다는 것을 보여주었다.

[0023] WO2012/083445에는 S 단백질의 막횡단 도메인 및 사이토졸 테일 도메인(TM/CT)이 인플루엔자 HA 단백질로부터의 TM/CT로 대체된 것인, 식물에서의 SARS CoV S 단백질 생산이 개시되어 있다.

[0024] 몇몇 그룹은 효과적인 백신 전략법으로서 SARS-CoV VLP를 사용한 면역화를 제안하였다. 곤충 세포에서 생산된 VLP 또는 포유동물 세포에서 생산된 키메라 MHV/SARS-CoV VLP가 이들 연구에서 사용되었다(문헌 [Lokugamage et al. Vaccine 2008 Feb 6;26(6):797-808], [Lu et al. 2007 Immunology 122496-5024]).

[0025] 그러나, SARS-CoV-2 VLP의 효과적인 규모 확장 및 전 세계 인구의 광범위한 백신 접종 요구를 충족시키는 데 필요한 양으로의 제조를 위해서는 효율적인 바이러스 구조 단백질 및 VLP 생산이 요구된다.

발명의 내용

[0026] **본 발명의 요약**

[0027] 본 발명은 변형된 바이러스 구조 단백질에 관한 것이다. 본 발명은 또한 변형된 바이러스 구조 단백질을 포함하는 바이러스 유사 입자(VLP) 및 숙주 또는 숙주 세포에서 VLP를 생산하는 방법에 관한 것이다. 더욱 구체적으로, 본 발명은 변형된 코로나바이러스 S 단백질에 관한 것이다. 본 발명은 또한 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP) 및 숙주 또는 숙주 세포에서 VLP를 생산하는 방법에 관한 것이다.

[0028] 한 측면에서, 연속적으로,

[0029] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0030] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는

[0031] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0032] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT

[0033] 를 포함하는 키메라 TMCT인 TMCT

[0034] 를 포함하는 변형된 코로나바이러스 S 단백질을 제공한다.

[0035] 본원에 기술된 바와 같은 변형된 S 단백질은 삼량체를 형성할 수 있다. 따라서, 본원에 기술된 바와 같은 변형된 코로나바이러스 S 단백질을 포함하는 삼량체를 제공한다.

[0036] 추가 측면에서, 상기 기술된 바와 같은 변형된 S 단백질, 또는 변형된 S 단백질을 포함하는 삼량체를 포함하는 바이러스 유사 입자(VLP)를 제공한다. 따라서, VLP는 변형된 코로나바이러스 S 단백질, 또는 변형된 S 단백질을 포함하는 삼량체를 포함하고, 변형된 S 단백질은

[0037] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0038] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0039] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함한다.

[0040] VLP는 식물 지질을 추가로 포함할 수 있다.

[0041] TM은 CT에 직접 융합될 수 있다. TM은 코로나바이러스 S 단백질 TM으로부터 유래된 것일 수 있고, CT는 인플루엔자 HA 단백질 CT로부터 유래된 것일 수 있다. 추가로, TM은 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열 및 인플루엔자 HA 단백질 TM으로부터 유래된 C 말단 서열을 포함하는 키메라 TM일 수 있다. 키메라 TM은 서열 번호 18 또는 서열 번호 169의 아미노산 1-20에 상응하는 적어도 20개의 아미노산, 또는 서열 번호 118 또는 164의 아미노산 1-21에 상응하는 적어도 21개의 아미노산, 또는 서열 번호 123의 아미노산 1-22에 상응하는 적어도 22개의 아미노산을 포함하는, 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산을 포함할 수 있다. 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산은 AGL 또는 AGL의 보존 치환, MAGL 또는 MAGL의 보존 치환으로부터 선

택될 수 있다. 키메라 TM은 서열 번호 18의 아미노산 1-20에 상응하는 아미노산을 포함할 수 있다.

[0042] CT는 코로나바이러스 S 단백질 CT로부터 유래된 N 말단 서열 및 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열을 포함하는 키메라 CT일 수 있다. 키메라 CT는 서열 번호 18, 126, 127, 128, 129, 130 또는 131의 아미노산 27-37에 상응하는 적어도 11개의 아미노산을 포함하는 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열, 및 코로나바이러스 S 단백질 CT의 N 말단 단부로부터의 하나 이상의 아미노산을 포함할 수 있다. 코로나바이러스 S 단백질 CT의 N 말단 단부로부터의 하나 이상의 아미노산은 C 또는 C의 보존 치환, CC 또는 CC의 보존 치환, 또는 CCM 또는 CCM의 보존 치환으로부터 선택될 수 있다. 키메라 CT는 서열 번호 18, 126, 128, 129, 130 또는 131의 아미노산 27-37; 또는 서열 번호 127의 아미노산 27-36에 상응하는 아미노산을 포함할 수 있다. 한 측면에서, 키메라 TMCT는 서열 번호 18 또는 서열 번호 169의 아미노산 1-20, 또는 서열 번호 118 또는 서열 번호 164의 아미노산 1-21, 또는 서열 번호 123의 아미노산 1-22에 상응하는 아미노산을 포함하는 키메라 TM, 서열 번호 18, 126, 127, 128, 129, 130 또는 131의 아미노산 27-37에 상응하는 아미노산, 또는 그의 조합을 포함할 수 있다.

[0043] CT 또는 CT의 일부는 서열 번호 15의 서열과, 또는 서열 번호 6, 8, 7, 9, 10, 12, 13 또는 14의 아미노산 35-50과, 또는 서열 번호 11의 아미노산 34-49와, 또는 서열 번호 3의 아미노산 553-568과, 또는 서열 번호 18의 아미노산 22-37과, 또는 서열 번호 19의 아미노산 21-40과, 또는 서열 번호 37의 아미노산 21-39와, 또는 서열 번호 38의 아미노산 25-36과, 또는 서열 번호 39의 아미노산 24-34와, 또는 서열 번호 126, 128, 129, 130 또는 131의 아미노산 22-37과; 또는 서열 번호 127의 아미노산 22-36과 80% 내지 100%의 동일성을 포함할 수 있다. TM 또는 TM의 일부는 서열 번호 132 또는 133의 서열과 80% 내지 100%의 동일성을 포함할 수 있다.

[0044] TMCT는 서열 번호 18, 19, 37, 38, 39, 64, 126, 127, 128, 129, 130, 131, 118, 119, 120, 123, 124, 125, 134, 135 164, 165, 166, 169, 170, 171, 172 또는 173의 서열과 약 80% 내지 약 100%의 동일성을 가진 서열을 포함할 수 있다.

[0045] 변형된 S 단백질은 S1 서브유닛 및 S2 서브유닛을 포함할 수 있고, 여기서, S2 서브유닛은 키메라 TMCT를 포함한다.

[0046] 변형된 S 단백질은 변형된 S 단백질 및 신호 펩티드를 포함하는 전구체 단백질로서 생산될 수 있다. 변형된 S 단백질 및 신호 펩티드를 포함하는 전구체 단백질은 서열 번호 1의 아미노산 1-1234와, 또는 서열 번호 5의 아미노산 1-1234, 서열 번호 21의 아미노산 1-1219와, 또는 서열 번호 30의 아미노산 1-1243과 80% 내지 100%의 동일성을 포함할 수 있고, 여기서, CT의 아미노산 서열은 서열 번호 15의 서열과, 또는 서열 번호 6, 8, 7, 9, 10, 12, 13 또는 14의 아미노산 35-50과, 또는 서열 번호 11의 아미노산 34-49와, 또는 서열 번호 3의 아미노산 553-568과 80% 내지 100%의 동일성을 포함한다.

[0047] 신호 펩티드는 S 단백질에 대해 천연 또는 비천연인 것일 수 있다. 비천연 신호 펩티드는 단백질 디술피드 이소머라제(PDI: protein disulfide-isomerase)의 신호 펩티드로부터 유래된 것일 수 있다. 변형된 S 단백질은 식물 특이적 N-글리칸을 추가로 포함할 수 있다.

[0048] 변형된 S 단백질 중의 CT 또는 CT의 일부는 인플루엔자 타입 B 또는 인플루엔자 서브타입 H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, 또는 H16으로부터 유래된 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것일 수 있다. 인플루엔자 헤마글루티닌(HA) 단백질은 인플루엔자 타입 B 또는 인플루엔자 서브타입 H1, H3, H5, H6, H7 또는 H9로부터 유래된 것일 수 있다.

[0049] 변형된 S 단백질의 엑토도메인은 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래된 것일 수 있고, TM 또는 TM의 일부는 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래된 것일 수 있거나, 또는 엑토도메인 및 TM 또는 TM의 일부, 둘 모두 SARS-CoV-2, SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV로부터 유래된 것일 수 있다.

[0050] 추가 측면에서, 변형된 S 단백질은 야생형 코로나바이러스 아미노산 서열과 비교하였을 때, 하나 이상의 아미노산 치환을 포함할 수 있다. 하나 이상의 치환은 융합 전 상태로 S 단백질을 유지할 수 있다.

[0051] 하나 이상의 아미노산 치환은 i) S1과 S2 서브유닛 사이의 절단 부위에서의 프로세싱을 제한하는 치환, ii) 하나 이상의 아미노산의 하나 이상의 프롤린으로의 치환, 또는 iii) S1과 S2 서브유닛 사이의 절단 부위에서의 프로세싱을 제한하는 치환 및 하나 이상의 아미노산의 하나 이상의 프롤린으로의 치환을 포함할 수 있다.

[0052] 하나 이상의 치환은 융합 전 상태로 S 단백질을 유지할 수 있거나, 또는 숙주 또는 숙주 세포에서 발현되었을

때 하나 이상의 치환을 포함하지 않는 상응하는 S 단백질의 수율과 비교하여, 숙주 또는 숙주 세포에서 발현되었을 때 더 높은 수율로 변형된 S 단백질을 생산한다.

[0053] 하나 이상의 아미노산 치환은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 802, 923, 927, 971, 972 또는 그의 조합의 아미노산에 상응할 수 있다.

[0054] 한 측면에서, 하나 이상의 아미노산 치환은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 971 및 972의 아미노산에 상응한다. 또 다른 측면에서, 하나 이상의 아미노산 치환은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 802, 927, 971 및 972의 아미노산에 상응한다. 추가로, 변형된 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 또는 그의 조합의 아미노산에 상응하는 하나 이상의 아미노산 치환을 포함할 수 있다. 따라서, 변형된 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668 및 670의 아미노산에 상응하는 하나 이상의 아미노산 치환을 포함할 수 있다.

[0055] 한 측면에서, 하나 이상의 치환은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 971 및 972의 아미노산에 상응할 수 있다. 서열 번호 2의 위치 667의 아미노산에 상응하는 아미노산의 치환은 글리신, 또는 글리신의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 668의 아미노산에 상응하는 아미노산의 치환은 세린, 또는 세린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 670의 아미노산에 상응하는 아미노산의 치환은 세린, 또는 세린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 971의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 972의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있다. 상기 기술된 바와 같은 변형된 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 923의 아미노산에 상응하는 아미노산의 치환을 추가로 포함할 수 있다.

[0056] 또 다른 측면에서, 하나 이상의 아미노산 치환은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 802, 927, 971 및 972의 아미노산에 상응할 수 있다. 서열 번호 2의 위치 667의 아미노산에 상응하는 아미노산의 치환은 글리신, 또는 글리신의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 668의 아미노산에 상응하는 아미노산의 치환은 세린, 또는 세린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 670의 아미노산에 상응하는 아미노산의 치환은 세린, 또는 세린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 802의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 927의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 971의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 972의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있다.

[0057] 또 다른 측면에서, 상기 기술된 바와 같은 변형된 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 923의 아미노산에 상응하는 아미노산의 치환을 추가로 포함할 수 있다. 서열 번호 2의 위치 667의 아미노산에 상응하는 아미노산의 치환은 글리신, 또는 글리신의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 668의 아미노산에 상응하는 아미노산의 치환은 세린, 또는 세린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 670의 아미노산에 상응하는 아미노산의 치환은 세린, 또는 세린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 802, 927, 971 및 972의 아미노산에 상응하는 아미노산의 치환은 프롤린, 또는 프롤린의 보존 치환으로의 치환일 수 있고, 서열 번호 2의 위치 923의 아미노산에 상응하는 아미노산의 치환은 페닐알라닌, 또는 페닐알라닌의 보존 치환으로의 치환일 수 있다.

[0058] 변형된 S 단백질은 서열 번호 5, 21, 30, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 95, 96, 97, 108, 109, 110, 144, 145, 146, 155, 156 또는 157의 아미노산과, 또는 서열 번호 47의 아미노산 24-1259, 서열 번호 48의 아미노산 25-1259, 서열 번호 49의 아미노산 25-1259, 서열 번호 50의 아미노산 25-1259, 서열 번호 51의 아미노산 25-1259, 서열 번호 52의 아미노산 25-1259, 서열 번호 53의 아미노산 25-1259, 서열 번호 54의 아미노산 25-1259, 서열 번호 55의 아미노산 25-1259, 서열 번호 56의 아미노산 25-1259, 서열 번호 57의 아미노산 25-1259, 서열 번호 58의 아미노산 25-1259, 서열 번호 59의 아미노산 25-1262, 서열 번호 60의 아미노산 25-1261, 서열 번호 61의 아미노산 25-1258, 또는 서열 번호 62의 아미노산 25-1256, 서열 번호 95의 아미노산 25-1243, 서열 번호 96의 아미노산 25-1240, 서열 번호 97의 아미노산 25-1243, 서열 번호 108의 아미노산 25-1341, 서열 번호 109의 아미노산 25-1338, 서열 번호 110의 아미노산 25-1341, 서열 번호 144의 아미노산 25-1351, 서열 번호 145의 아미노산 25-1348, 서열 번호 146의 아미노산 25-1351, 서열 번호 155의 아미노산 25-1159, 서열 번호 156의 아미노산 25-1156, 또는 서열 번호 157의 아미노산

25-1159와 80% 내지 100%의 동일성을 포함할 수 있다.

[0059] 또 다른 측면에서, 상기 기술된 바와 같은 변형된 S 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산을 제공한다.

[0060] 추가 측면에서, 유효 용량의, 상기 기술된 바와 같은 변형된 S 단백질, 변형된 S 단백질을 포함하는 삼량체, 또는 변형된 S 단백질을 포함하는 VLP, 및 약학적으로 허용되는 담체, 애주번트, 비히클 또는 부형제를 포함하는 조성물을 제공한다. 추가의 또 다른 측면에서, 면역 반응을 유도하기 위한 백신을 제공한다. 백신은 유효 용량의, 상기 기술된 바와 같은, 변형된 S 단백질, 상기 기술된 바와 같은 변형된 S 단백질을 포함하는 삼량체 또는 변형된 S 단백질을 포함하는 VLP를 포함한다.

[0061] 조성물은 약학적으로 허용되는 담체, 애주번트, 비히클 또는 부형제를 추가로 포함한다. 추가 측면에서, 면역 반응을 유도하기 위한 백신을 제공한다. 백신은 유효 용량의, 상기 기술된 바와 같은 변형된 코로나바이러스를 포함하는 VLP를 포함한다. 백신은 VLP의 혼합물을 포함하는 다가 백신일 수 있다.

[0062] 추가의 또 다른 측면에서, 상기 기술된 바와 같은 변형된 S 단백질, 삼량체 또는 VLP를 포함하는 (비인간) 숙주 또는 숙주 세포를 제공한다. 추가의 또 다른 측면에서, 상기 기술된 바와 같은 VLP를 포함하는 숙주 또는 숙주 세포를 제공한다. 또 다른 측면에서, 유효량의, 상기 기술된 바와 같은 변형된 S 단백질을 포함하는 VLP를 포함하는 조성물을 제공한다.

[0063] 추가의 또 다른 측면에서, S 단백질 삼량체를 제공한다. 삼량체는 변형된 S 단백질이

[0064] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0065] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0066] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0067] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하는 것인, 변형된 코로나바이러스 S 단백질을 포함한다. 삼량체 중 변형된 S 단백질은 상기 기술된 바와 같이, 야생형 코로나바이러스 아미노산 서열과 비교하여 하나 이상의 아미노산 치환을 포함할 수 있다. 추가 측면에서, 유효 용량의 상기 기술된 바와 같은 삼량체 및 약학적으로 허용되는 담체, 애주번트, 비히클 또는 부형제를 포함하는 조성물을 제공한다. 또 다른 측면에서, 상기 기술된 바와 같은 삼량체를 포함하는 바이러스 유사 입자(VLP) 또한 제공한다. VLP는 식물 지질을 추가로 포함할 수 있다. 또 다른 측면에서, 유효 용량의 상기 기술된 바와 같은 삼량체 및 약학적으로 허용되는 담체, 애주번트, 비히클 또는 부형제를 포함하는 조성물을 제공한다. 추가 측면에서, 면역 반응을 유도하기 위한 백신을 제공한다. 백신은 유효 용량의 상기 기술된 바와 같은 삼량체를 포함한다. 추가 측면에서, 면역 반응을 유도하기 위한 백신을 제공한다. 백신은 유효 용량의 상기 기술된 바와 같은 삼량체를 포함하는 VLP를 포함한다. 백신은 VLP의 혼합물을 포함하는 다가 백신일 수 있다. 추가의 또 다른 측면에서, 상기 기술된 바와 같은 삼량체, 또는 삼량체를 포함하는 VLP를 포함하는 비인간 숙주 또는 숙주 세포를 제공한다.

[0068] 또 다른 측면에서, 상기 기술된 바와 같은 조성물 또는 백신을 투여하는 단계를 포함하는, 피험체에서 코로나바이러스 감염에 대한 면역을 유도하는 방법을 제공한다. 조성물 또는 백신은 피험체에게 1회 투여될 수 있거나, 또는 조성물 또는 백신은 다회에 걸쳐 피험체에게 투여될 수 있다. 조성물 또는 백신은 초기 용량으로 투여될 수 있고, 초기 용량 투여로부터 1일 내지 6개월 사이에 1회 이상의 후속 용량이 투여될 수 있다. 후속 투여는 초기 용량 투여 후 21일 후에 투여될 수 있다.

[0069] 또 다른 측면에서, 상기 기술된 바와 같은 조성물 또는 백신을 사용하여 제조된 항체 또는 항체 단편을 제공한다.

[0070] 추가의 또 다른 측면에서,

[0071] A) a) 비인간 숙주 또는 숙주 세포 내로 상기 기술된 바와 같은 변형된 S 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산을 도입하거나; 또는 상기 기술된 바와 같은 변형된 S 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계, 및

[0072] b) 핵산의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 VLP를 생산하는 단계를 포함하는, (비인간) 숙주 또는 숙주 세포에서 바이러스 유사 입자(VLP)를 생산하는 방법을 제공한다.

[0073] 추가 단계 c)에서, 비인간 숙주 또는 숙주 세포를 수확할 수 있다.

[0074] 추가 측면에서,

[0075] B) a) 비인간 숙주 또는 숙주 세포 내로 상기 기술된 바와 같은 변형된 S 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산을 도입하거나; 또는 상기 기술된 바와 같은 변형된 S 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계; 및

[0076] b) 핵산에 의해 코딩되는 변형된 S 단백질의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 유사하거나, 또는 동일한 조건하에서 비변형된 S 단백질을 발현하는 숙주 또는 숙주 세포와 비교하여 더 높은 수율로 변형된 S 단백질을 생산하는 단계를 포함하는,

[0077] (비인간) 숙주 또는 숙주 세포에서 코로나바이러스 S 단백질의 생산 수율을 증가시키는 방법을 제공한다.

[0078] 추가 단계 c)에서, 비인간 숙주 또는 숙주 세포를 수확할 수 있다.

[0079] 추가의 또 다른 측면에서,

[0080] C) a) 비인간 숙주 또는 숙주 세포 내로,

[0081] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0082] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0083] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0084] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하는 변형된 코로나바이러스 S 단백질을 코딩하는 핵산을 도입;

[0085] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0086] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0087] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0088] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하는 변형된 S 단백질을 코딩하는 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계; 및

[0089] b) 핵산에 의해 코딩되는 변형된 S 단백질의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 유사하거나, 또는 동일한 조건하에서 비변형된 S 단백질을 발현하는 숙주 또는 숙주 세포에서의 VLP의 수율과 비교하여 더 높은 수율로 변형된 S 단백질을 포함하는 VLP를 생산하는 단계를 포함하는,

[0090] (비인간) 숙주 또는 숙주 세포에서 바이러스 유사 입자(VLP)의 생산 수율을 증가시키는 방법을 제공한다.

[0091] 추가 측면에서, D) a) 비인간 숙주 또는 숙주 세포 내로,

[0092] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0093] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0094] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0095] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하는 변형된 코로나바이러스 S 단백질을 코딩하는 핵산을 도입하거나; 또는

[0096] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0097] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0098] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0099] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하는 변형된 S 단백질을 코딩하는 핵산을 포함하는 비인간 숙주 또는 숙주 세포를 제공하는 단계; 및

[0100] b) 핵산의 발현을 허용하는 조건하에서 비인간 숙주 또는 숙주 세포를 인큐베이션시켜 바이러스 유사 입자(VLP)를 생산하는 단계를 포함하는,

[0101] (비인간) 숙주 또는 숙주 세포에서 VLP를 생산하는 방법을 제공한다.

[0102] 추가 측면에서, 방법 A), B), C) 또는 D)의 VLP는 추가로 숙주 또는 숙주 세포로부터 추출 및 정제될 수 있다. 숙주 또는 숙주 세포는 식물, 식물 세포, 진균, 진균 세포, 곤충, 곤충 세포, 동물 또는 동물 세포를 포함할 수 있다. 방법 A), B), C) 또는 D)의 숙주 또는 숙주 세포는 식물, 식물의 일부 또는 식물 세포일 수 있다.

[0103] 또 다른 측면에서, 방법 A), B), C) 또는 D)에 의해 생산된 VLP를 제공한다.

[0104] 추가로, 추가의 또 다른 측면에서, 애주번트 및 바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 변형된 코로나바이러스 S 단백질을 포함하고, 변형된 S 단백질은

[0105] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0106] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0107] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0108] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하고, 여기서, 변형된 S 단백질은 서열 번호 2의 참조 아미노산 서열과 비교하였을 때, 위치 667, 668, 670, 971 및 972에 치환을 추가로 포함하는 것인, 조성물을 제공한다.

[0109] 추가의 또 다른 측면에서, 바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 변형된 코로나바이러스 S 단백질을 포함하고, 변형된 S 단백질은

[0110] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0111] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는 키메라 TMCT이며,

[0112] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및

[0113] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함하는 것인 TMCT를 포함하고; 여기서, 변형된 S 단백질은 위치 667에 글리신 치환, 위치 668에 세린 치환, 위치 670에 세린 치환, 위치 971에 프롤린 치환 및 위치 972에 프롤린 치환을 포함하고, 위치는 서열 번호 2의 참조 아미노산 서열에 상응하는 것인, 조성물을 제공한다. 인플루엔자 헤마글루티닌(HA) 단백질은 인플루엔자 타입 B 또는 인플루엔자 서브타입 H1, H3, H5, H6, H7 또는 H9로부터 유래된 것일 수 있다. 조성물은 애주번트를 추가로 포함할 수 있다.

[0114] 추가 측면에서, 바이러스 유사 입자(VLP)를 포함하는 조성물로서, VLP는 변형된 코로나바이러스 S 단백질을 포함하고, 변형된 S 단백질은 서열 번호 21의 서열을 포함하는 것인, 조성물을 제공한다. 조성물은 애주번트를 추가로 포함할 수 있다.

[0115] 본 발명의 요약은 반드시 본 발명의 모든 특징을 기술하는 것은 아니다.

도면의 간단한 설명

[0116] **도면의 간단한 설명**

본 발명의 이들 및 다른 특징은 첨부된 도면이 참조되는 하기 설명으로부터 더 자명해 질 것이다:

도 1은 코로나바이러스 S 단백질 및 S1/S2(aa 685/686) 및 S2' 절단 부위의 위치의 개략도를 보여주는 것이다. SP: 신호 펩티드(aa 1-15); NTD: N 말단 도메인(aa 16-306); RBD: 수용체 결합 도메인(aa 335-527); FP: 융합 펩티드(aa 816-833); HR1: 헵타드 반복부 1(aa 908-991); HR2: 헵타드 반복부 2(aa 1166-1207); TM: 막횡단 도메인(aa 1214-1234); CT: 세포질 테일(aa 1235-1273). 각 영역의 잔기 번호(aa)는 SARS-CoV-2(2019-nCoV)의 S 단백질 중의 그의 위치에 상응한다.

도 2는 예시적인 인플루엔자 균주로부터의 아미노산 서열의 정렬을 보여주는 것이다. H1 A/California/07/2009 (서열 번호 6), H2 A/Singapore/1/1957 HA(서열 번호 7), H3 A/Minnesota/41/2019 HA(서열 번호 8), H5 A/Indonesia/5/05 HA(서열 번호 9), H6 A/Teal/Hong Kong/W312/97 HA(서열 번호 10), H7 A/Guangdong/17SF003/2016 HA(서열 번호 11), H9 A/Hong Kong/1073/99 HA(서열 번호 12), 및

B/Washington/02/2019 HA(서열 번호 13)에 대한 헤마글루티닌(HA)의 엑토도메인, 막횡단 도메인(TM), 및 세포질 테일 도메인(CT)의 C 말단 영역이 제시되어 있다. 상기 서열에 대한 컨센서스 서열 또한 제시되어 있다(서열 번호 14).

도 3a는 하기 5' UTR: nbMT78(구성물 8586), nbCSY65(구성물 8589) 및 nbHEL40(구성물 8591)의 제어하에 천연(야생형) 막횡단 도메인 및 세포질 테일(wtTMCT)을 포함하는 SARS-CoV-2 S 단백질; nbMT78(구성물 8592), nbCSY65(구성물 8595) 및 nbHEL40(구성물 8597)의 제어하에, 천연(야생형) 막횡단 도메인 및 세포질 테일(wtTMCT)이 인플루엔자 H5 A/Indonesia/5/05의 인플루엔자 헤마글루티닌(HA)의 TMCT(H5iTMCT)로 대체된 변형된 SARS-CoV-2 단백질; 및 nbMT78(구성물 8610), nbCSY65(구성물 8611) 및 nbHEL40(구성물 8671)의 제어하에, 천연(야생형) 세포질 테일(wtCT)이 인플루엔자 H5 A/Indonesia/5/05의 인플루엔자 헤마글루티닌(HA)의 CT(H5iCT)로 대체된 변형된 SARS-CoV-2 단백질을 발현하는 식물에서의 SARS-CoV-2 S 단백질 축적 차이의 정량화된 변화 배수를 보여주는 것이다. (nCOV S(GSAS-2P)로 지칭되는) SARS-CoV-2 S 단백질 서열은 서열 번호 2의 참조 서열 기준으로 하기 치환: R667G, R668S, R670S, K971P 및 V972P를 가진다. 결과는 1로 설정된, 구성물 8591로부터의 SARS-CoV-2 S 단백질 축적으로 정규화되었다. **도 3b**는 비환원 SDS-PAGE 겔 상의 정화된 미정제 추출물의 단백질 분리를 보여주는 것이다. 하기 변형된 S 단백질이 식물에서 발현되었다: 레인 1: nbMT78 제어하의 야생형 막횡단 및 사이토졸 테일 도메인(wt TMCT)을 포함하는 S 단백질; 레인 2: nbCSY65 제어하의 야생형 막횡단 및 사이토졸 테일 도메인(wt TMCT)을 포함하는 S 단백질; 레인 3: nbHEL40 제어하의 야생형 막횡단 및 사이토졸 테일 도메인(wt TMCT)을 포함하는 S 단백질; 레인 4: nbMT78 제어하의 SARS-CoV-2 엑토도메인, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)로부터의 막횡단 및 사이토졸 테일 도메인(H5i TMCT)을 포함하는 변형된 S 단백질; 레인 5: nbCSY65 제어하의 SARS-CoV-2 엑토도메인, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)로부터의 막횡단 및 사이토졸 테일 도메인(H5i TMCT)을 포함하는 변형된 S 단백질; 레인 6: nbHEL40 제어하의 SARS-CoV-2 엑토도메인, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)로부터의 막횡단 및 사이토졸 테일 도메인(H5i TMCT)을 포함하는 변형된 S 단백질; 레인 7: nbMT78 제어하의 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H5 A/Indonesia/5/05의 헤마글루티닌(HA)로부터의 사이토졸 테일 도메인(CT)(H5i CT)을 포함하는 변형된 S 단백질; 레인 8: nbCSY65 제어하의 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H5 A/Indonesia/5/05의 헤마글루티닌(HA)로부터의 사이토졸 테일 도메인(CT)(H5i CT)을 포함하는 변형된 S 단백질; 레인 9: nbHEL40 제어하의 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H5 A/Indonesia/5/05의 헤마글루티닌(HA)로부터의 사이토졸 테일 도메인(CT)(H5i CT)을 포함하는 변형된 S 단백질. 변형된 S 단백질의 분자량은 약 150 kDa이고, 이는 화살표로 표시되어 있다. **도 3c**는 도 3b에 도시된 동일한 일련의 용해물의 웨스턴 블롯 분석을 보여주는 것이며, 레인은 도 3b에 도시된 바와 같은 레인에 상응한다. 상단 패널은 항SARS-CoV-2 S1 항체(40150-R007)를 이용한 검출을 보여주는 것이다. 하단 패널은 항SARS-CoV-2 S2 항체(NB100-56578)를 이용한 검출을 보여주는 것이다. (S1 및 S2 서브유닛을 포함하는) SARS-CoV-2 단백질의 단량체의 분자량은 약 150 kDa이다(비환원).

도 4a는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H5 Indo); SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H1 A/California/7/2009의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H1 California); SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H3 A/Minnesota/41/2019의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H3 Minnesota); SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H6 A/Teal/Hong Kong/W312/97의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H6 Hong Kong); SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H7 A/Guangdong/17SF003/2016의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H7 Guangdong); SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H9 A/Hong Kong/1073/99의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H9 Hong Kong); 및 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 B/Washington/02/2019의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(B Washington)을 발현하는 식물에서의 SARS-CoV-2 S 단백질 축적 차이의 정량화된 변화 배수를 보여주는 것이다. 결과는 1로 설정된, 천연(야생형) 세포질 테일(wtCT)이 nbHEL40의 제어하에 인플루엔자 H5 A/Indonesia/5/05의 인플루엔자 헤마글루티닌(HA)의 CT(H5iCT)로 대체된 변형된 SARS-CoV-2(H5 Indo)를 코딩하는 구성물 8671로부터의 SARS-CoV-2 S 단백질 축적으로 정규화되었다. **도 4b**는 하기 구성물로부터 변형된 S 단백질을 발현하는 식물로부터의 미정제 용해물의 웨스턴 블롯 분석을 보여주는 것이다: 레인 2, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H5 Indo); 레인 3, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H1 A/California/7/2009의 헤마글루티

닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H1 California); 레인 4, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H3 A/Minnesota/41/2019의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H3 Minn); 레인 5, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H6 A/Teal/Hong Kong/W312/97의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H6 HK); 레인 6, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H7 A/Guangdong/17SF003/2016의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H7 Guan); 레인 7, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 H9 A/Hong Kong/1073/99의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(H9 HK); 레인 8, SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단, 및 B/Washington/02/2019의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질(B Wash). 레인 1은 모의 아그로침투(Agroinfiltration)로 처리된 식물로부터의 미정제 용해물이다. Sino 400150-R007 항체는 SARS-COV-2 S 단백질의 S1 서브유닛 검출을 위해 사용되었다. (S1 및 S2 서브유닛을 포함하는) SARS-CoV-2 단백질의 단량체의 분자량은 약 150 kDa이다(비환원). **도 4c**는 SARS-COV-2 S 단백질의 S2 서브유닛을 검출하는 NB100-56578 항체를 제외한, 도 4b에 도시된 동일한 일련의 용해물의 웨스턴 블롯 분석을 보여주는 것이다.

도 5a는 천연 SARS-CoV-2 S 단백질의 C 말단 영역(wtTM/wtCT), 인플루엔자 H5 헤마글루티닌(HA)의 C 말단 영역(H5iTM/H5iCT), 야생형 막횡단 도메인(TM) 및 인플루엔자 H5 HA 사이토졸 테일(CT) 도메인을 포함하는 변형된 SARS-CoV-2 S 단백질의 C 말단 영역(wtTM/H5iCT), 및 SARS-CoV-2 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는 4개의 대안적인 버전의 변형된 S 단백질의 C 말단 영역(wtTM/H5iCT V1-V4)의 아미노산 서열을 보여주는 것이다. 코로나바이러스 S 단백질로부터의 TM 도메인은 밑줄체로 표시되어 있고, 인플루엔자 HA로부터 유래된 CT 도메인은 굵은체로 제시되어 있다. **도 5b**는 1로 설정된 야생형 막횡단 도메인(TM) 및 인플루엔자 H5 HA 사이토졸 테일(CT) 도메인을 포함하는 키메라 TMCT(wtTM/H5iCT)를 가진 변형된 SARS-CoV-2 S 단백질을 발현하는 식물에서의 변형된 SARS-CoV-2 S 단백질 축적 대비, 도 5a에 도시된 바와 같은, 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 4개의 각 변이체 변형된 S 단백질(wtTM/H5iCT, V1-V4)을 발현하는 식물에서의 SARS-CoV-2 S 단백질 축적 차이의 정량화된 변화 배수를 보여주는 것이다.

도 6a는 야생형 막횡단 및 사이토졸 테일 도메인을 포함하는 SARS-COV-2 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(wtTMCT; 구성물 8591). **도 6b**는 SARS-CoV-2 엑토도메인 및 인플루엔자 H5 헤마글루티닌 막횡단 도메인 및 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5i TMCT; 구성물 8597). **도 6c**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H5 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5i CT; 구성물 8671). **도 6d**는 SARS-CoV-2 엑토도메인 및 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 가진 대안적인 버전의 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5i CT V1; 구성물 8980). **도 6e**는 SARS-CoV-2 엑토도메인 및 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 가진 대안적인 버전의 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5i CT V2; 구성물 8981). **도 6f**는 SARS-CoV-2 엑토도메인 및 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 SARS-CoV-2 엑토도메인 및 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 가진 SARS-CoV-2 엑토도메인을 포함하는 대안적인 버전의 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5i CT V3; 구성물 8982). **도 6g**는 SARS-CoV-2 엑토도메인 및 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 가진 대안적인 버전의 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5i CT V4; 구성물 8983). **도 6h**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H1 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H1 CT; 구성물 7390). **도 6i**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H3 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H3 CT; 구성물 7391). **도 6j**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H6 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H6 CT; 구성물 7392). **도 6k**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H7 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H7 CT; 구성물 7393). **도 6l**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 H9 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것

이다(H9 CT; 구성물 7394). **도 6m**는 SARS-CoV-2 엑토도메인, SARS-CoV-2 막횡단 도메인 및 인플루엔자 HA B 헤마글루티닌 사이토졸 테일 도메인을 포함하는 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(HA B CT; 구성물 7395).

도 7a는 수용자 벡터 8501의 개략도를 보여주는 것이다. **도 7b**는 수용자 벡터 8500의 개략도를 보여주는 것이다. **도 7c**는 수용자 벡터 8716의 개략도를 보여주는 것이다.

도 8a는 벡터 8586의 개략도를 보여주는 것이다. **도 8b**는 벡터 8589의 개략도를 보여주는 것이다. **도 8c**는 벡터 8591의 개략도를 보여주는 것이다.

도 9a는 벡터 8592의 개략도를 보여주는 것이다. **도 9b**는 벡터 8595의 개략도를 보여주는 것이다. **도 9c**는 벡터 8597의 개략도를 보여주는 것이다.

도 10a는 벡터 8610의 개략도를 보여주는 것이다. **도 10b**는 벡터 8611의 개략도를 보여주는 것이다. **도 10c**는 벡터 8671의 개략도를 보여주는 것이다.

도 11a는 추가 치환을 포함하는 변형된 SARS-CoV-2 S 단백질(wtTM/H5iCT)을 발현하는 식물에서의 축적의 정량화된 변화 배수를 보여주는 것이다. 변형된 SARS-CoV-2 S 단백질은 하기 치환을 가진다: "GSAS-2P": R667G, R668S, R670S, K971P 및 V972P; "GSAS-4P": R667G, R668S, R670S, K971P, V972P, F802P 및 A927P; 및 "GSAS-6P": R667G, R668S, R670S, K971P, V972P, F802P, A877P, A884P 및 A927P (서열 번호 2의 참조 서열 기준). 결과는 1로 설정된, 변형된 SARS-CoV-2 (wtTM/H5iCT) 및 GSAS+2P 치환의 축적으로 정규화되었다. **도 11b**는 각각의 변형된 SARS-CoV-2 S 단백질이 L923F 치환을 추가로 포함하는 것인 축적의 정량화된 변화 배수와 비교하여도 11a에 대해 기술된 바와 같은 GSAS-2P, GSAS-4P, 및 GSAS-6P 치환을 각각 포함하는 변형된 SARS-CoV-2 S 단백질(wtTM/H5iCT)을 발현하는 식물에서의 축적의 정량화된 변화 배수를 보여주는 것이다. 결과는 1로 설정된, 변형된 SARS-CoV-2(wtTM/H5iCT) 및 GSAS+2P 치환의 축적으로 정규화되었다.

도 12a는 벡터 8980의 개략도를 보여주는 것이다. **도 12b**는 벡터 8981의 개략도를 보여주는 것이다. **도 12c**는 벡터 8982의 개략도를 보여주는 것이다. **도 12d**는 벡터 8983의 개략도를 보여주는 것이다.

도 13a는 벡터 7390의 개략도를 보여주는 것이다. **도 13b**는 벡터 7391의 개략도를 보여주는 것이다. **도 13c**는 벡터 7392의 개략도를 보여주는 것이다. **도 13d**는 벡터 7393의 개략도를 보여주는 것이다. **도 13e**는 벡터 7394의 개략도를 보여주는 것이다. **도 13f**는 벡터 7395의 개략도를 보여주는 것이다.

도 14a는 벡터 8953의 개략도를 보여주는 것이다. **도 14b**는 벡터 8940의 개략도를 보여주는 것이다.

도 15a는 벡터 8933의 개략도를 보여주는 것이다. **도 15b**는 벡터 8960의 개략도를 보여주는 것이다. **도 15c**는 벡터 8947의 개략도를 보여주는 것이다.

도 16a는 하기 구성물로부터 변형된 S 단백질을 발현하는 식물로부터의 미정제 용해물의 웨스턴 블롯 분석을 보여주는 것이다: 레인 1, SARS-CoV-1 엑토도메인, 막횡단, 및 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("wtTMCT," 구성물 9231); 레인 2, SARS-CoV-1로부터의 엑토도메인, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)으로부터의 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 S 단백질("H5iTMCT," 구성물 9232); 레인 3, SARS-CoV-1로부터의 엑토도메인 및 막횡단 도메인, 및 H5 A/Indonesia/5/05(H5 Indo)의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H5iCT," 구성물 9233); 레인 4, SARS-CoV-1 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는, SARS-CoV-1로부터의 엑토도메인 및 막횡단 도메인, 및 H5 A/Indonesia/5/05(H5 Indo)의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H5iCT(V4)," 구성물 9234); 레인 5, SARS-CoV-1로부터의 엑토도메인 및 막횡단 도메인, 및 H1 A/California/7/2009의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H1cCT," 구성물 9235). 검출에 사용된 1차 항체는 시노 바이오로지칼즈(Sino Biologicals)로부터의 SARS-CoV 스파이크 S1 서브유닛 항체(SARS-CoV Spike S1 Subunit Antibody)(40150-MM08, 1/5000)였다. 검출에 사용된 2차 항체는 JIR로부터의 염소 항마우스(115-035-146, 1/10 000)였다. 변형된 S 단백질의 분자량은 약 150 kDa이다.

도 16b, 16c 및 **16d**는 불연속 이오딕사놀 밀도 구배로부터의 분획 F5(30%), F6(30%), F7(25%), F8(25%), F9(25%), F10(15%) 및 F11(15%)의 웨스턴 블롯 분석을 보여주는 것이다. 상기 분획에서 단백질의 축적은 더 큰 분자량의 구조체 형성, 즉, VLP 형성을 나타낸다. 미정제 용해물의 분획으로부터의 웨스턴 블롯의 경우, 검출에 사용된 1차 항체는 시노 바이오로지칼즈로부터의 SARS-CoV 스파이크 S1 서브유닛

항체(40150-MM08(1/5000))였고, 검출에 사용된 2차 항체는 염소 항마우스(JIR, 115-035-146(1/10000))였다. **도 16b**: 천연 TMCT 도메인을 포함하는 SARS-CoV-1 S 단백질(2P+R667A 치환 포함)을 발현하는 식물로부터의 미정제 용해물(wtTMCT, 구성물 9231)을 분석하였다. **도 16c**: H5 A/Indonesia/5/05 HA로부터의 TMCT를 가진 변형된 SARS-CoV-1 S 단백질(2P+R667A 치환 포함)을 발현하는 식물로부터의 미정제 용해물(H5iTMCT, 구성물 9232). **도 16d**: H5 A/Indonesia/5/05 HA로부터의 세포질 테일을 가진 변형된 SARS-CoV-1 S 단백질(2P+R667A 치환 포함)을 발현하는 식물로부터의 미정제 용해물(H5iCT, 구성물 9233).

도 17a는 천연 TMCT 도메인을 포함하는 SARS-COV-1 S 단백질(2P+R667A 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(wtTMCT, 구성물 9231). **도 17b**는 H5 A/Indonesia/5/05 HA로부터의 TMCT를 가진 변형된 SARS-CoV-1 S 단백질(2P+R667A 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iTMCT, 구성물 9232). **도 17c**는 H5 A/Indonesia/5/05 HA로부터의 세포질 테일을 가진 변형된 SARS-CoV-1 S 단백질(2P+R667A 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iCT, 구성물 9233).

도 18a는 벡터 9231의 개략도를 보여주는 것이다. **도 18b**는 벡터 9232의 개략도를 보여주는 것이다. **도 18c**는 벡터 9233의 개략도를 보여주는 것이다. **도 18d**는 벡터 9234의 개략도를 보여주는 것이다. **도 18e**는 벡터 9235의 개략도를 보여주는 것이다.

도 19a는 하기 구성물로부터 변형된 S 단백질을 발현하는 식물로부터의 미정제 용해물의 웨스턴 블롯 분석을 보여주는 것이다: 레인 1, MERS-CoV 엑토도메인, 막횡단, 및 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("wtTMCT," 구성물 9246); 레인 2, MERS-CoV로부터의 엑토도메인, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)으로부터의 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 S 단백질("H5iTMCT," 구성물 9247); 레인 3, MERS-CoV로부터의 엑토도메인 및 막횡단 도메인, 및 H5 A/Indonesia/5/05(H5 Indo)의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H5iCT," 구성물 9249); 레인 4, MERS-CoV 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는, MERS-CoV로부터의 엑토도메인 및 막횡단 도메인, 및 H5 A/Indonesia/5/05(H5 Indo)의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H5iCT(V4)," 구성물 9250); 레인 5, MERS-CoV로부터의 엑토도메인 및 막횡단 도메인, 및 H1 A/California/7/2009의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H1cCT," 구성물 9251). 검출에 사용된 1차 항체는 시노 바이오로지칼즈로부터의 MERS-CoV 스파이크 단백질 S1 항체(N 말단)(100208-RP02, 1/5000)이었다. 검출에 사용된 2차 항체는 JIR로부터의 염소 항마우스(115-035-144, 1/10 000)였다. 변형된 S 단백질의 분자량은 약 175 kDa이다. **도 19b**는 천연 TMCT 도메인을 포함하는 MERS-COV S 단백질(ASVG+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(wtTMCT, 구성물 9246). **도 19c**는 H5 A/Indonesia/5/05 HA로부터의 TMCT를 가진 변형된 MERS-CoV S 단백질(ASVG+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iTMCT, 구성물 9232). **도 19d**는 H5 A/Indonesia/5/05 HA로부터의 세포질 테일을 가진 변형된 MERS-CoV S 단백질(ASVG+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iCT, 구성물 9249). **도 19e**는 MERS-CoV 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는, H5 A/Indonesia/5/05 HA로부터의 사이토졸 테일을 가진 변형된 MERS-CoV S 단백질(ASVG+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다("H5iCT(V4)," 구성물 9250). **도 19f**는 H1 A/California/7/2009 HA로부터의 세포질 테일을 가진 변형된 MERS-CoV S 단백질(ASVG+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H1cCT, 구성물 9251).

도 20a는 벡터 9246의 개략도를 보여주는 것이다. **도 20b**는 벡터 9247의 개략도를 보여주는 것이다. **도 20c**는 벡터 9249의 개략도를 보여주는 것이다. **도 20d**는 벡터 9250의 개략도를 보여주는 것이다. **도 20e**는 벡터 9251의 개략도를 보여주는 것이다.

도 21은 수용자 벡터 7147의 개략도를 보여주는 것이다.

도 22는 천연 신호 펩티드가 제거된 천연 SARS-CoV-2, SARS-CoV-1, 및 MERS-CoV S 단백질 서열(서열 번호 2, 114, 및 115)의 정렬을 보여주는 것이다. 신호 펩티드가 없는 천연 SARS-CoV-2 S 단백질(서열 번호 2) 중의 RRAR 퓨린 절단 부위(667-670) 및 F802P, A877P, A884P, A927P, K971P, V972P에 상응하는 잔기는 각각 신호 펩티드가 없는 천연 SARS-CoV-1 S 단백질(서열 번호 114) 및 천연 MERS S 단백질(서열 번호 115)로부터의 상동성 잔기와 함께 박스로 표시되어 있다.

도 23a는 하기 구성물로부터 변형된 S 단백질을 발현하는 식물로부터의 미정제 용해물의 웨스턴 블롯 분석을 보여주는 것이다: 레인 3, OC43-CoV 엑토도메인, 막횡단, 및 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("wtTMCT," 구성물 9269); 레인 4, OC43-CoV로부터의 엑토도메인, 및 H5 A/Indonesia/5/05의 헤마글루티닌(HA)으로부터의 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함하는 변형된 S 단백질("H5iTMCT," 구성물 9270); 레인 5, OC43-CoV로부터의 엑토도메인 및 막횡단 도메인, 및 H5 A/Indonesia/5/05(H5 Indo)의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H5iCT," 구성물 9272); 레인 6, OC43-CoV 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는, OC43-CoV로부터의 엑토도메인 및 막횡단 도메인, 및 H5 A/Indonesia/5/05(H5 Indo)의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H5iCT (V4)," 구성물 9273); 레인 7, OC43-CoV로부터의 엑토도메인 및 막횡단 도메인, 및 H1 A/California/7/2009의 헤마글루티닌(HA)으로부터의 사이토졸 테일 도메인을 포함하는 변형된 S 단백질("H1cCT," 구성물 9274). 검출에 사용된 1차 항체는 안티바이즈-온라인(Antibodies-online)으로부터의 항코로나바이러스 OC43 스파이크 단백질(ABIN2754654, 1/1000)이었다. 검출에 사용된 2차 항체는 JIR로부터의 염소 항토끼(111-035-144, 1/10 000)였다. 변형된 S 단백질의 분자량은 약 150 kDa이다. **도 23b**는 H5 A/Indonesia/5/05 HA로부터의 TMCT를 가진 변형된 OC43-CoV S 단백질(GGSGS+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iTMCT, 구성물 9270). **도 23c**는 H5 A/Indonesia/5/05 HA로부터의 세포질 테일을 가진 변형된 OC43-CoV S 단백질(GGSGS+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iCT, 구성물 9272). **도 23d**는 OC43-CoV 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는, H5 A/Indonesia/5/05 HA로부터의 사이토졸 테일을 가진 변형된 OC43-CoV S 단백질(GGSGS+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iCT (V4), 구성물 9273). **도 23e**는 H1 A/California/7/2009 HA로부터의 세포질 테일을 가진 변형된 OC43-CoV S 단백질(GGSGS+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H1cCT, 구성물 9274).

도 24a는 벡터 9269의 개략도를 보여주는 것이다. **도 24b**는 벡터 9270의 개략도를 보여주는 것이다. **도 24c**는 벡터 9272의 개략도를 보여주는 것이다. **도 24d**는 벡터 9273의 개략도를 보여주는 것이다. **도 24e**는 벡터 9274의 개략도를 보여주는 것이다.

도 25a는 천연 TMCT 도메인을 포함하는 229E-CoV S 단백질(R567A+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(wtTMCT, 구성물 9310). **도 25b**는 H5 A/Indonesia/5/05 HA로부터의 TMCT를 가진 변형된 229E-CoV S 단백질(R567A+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iTMCT, 구성물 9311). **도 25c**는 H5 A/Indonesia/5/05 HA로부터의 세포질 테일을 가진 변형된 229E-CoV S 단백질(R567A+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iCT, 구성물 9312). **도 25d**는 229E-CoV 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진이 있는, H5 A/Indonesia/5/05 HA로부터의 사이토졸 테일을 가진 변형된 229E-CoV S 단백질(R567A+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H5iCT (V4), 구성물 9313). **도 25e**는 H1 A/California/7/2009 HA로부터의 세포질 테일을 가진 변형된 229E-CoV S 단백질(R567A+2P 치환 포함)을 포함하는 바이러스 유사 입자(VLP)의 전자 현미경 사진을 보여주는 것이다(H1cCT, 구성물 9314).

도 26a는 벡터 9310의 개략도를 보여주는 것이다. **도 26b**는 벡터 9311의 개략도를 보여주는 것이다. **도 26c**는 벡터 9312의 개략도를 보여주는 것이다. **도 26d**는 벡터 9313의 개략도를 보여주는 것이다. **도 26e**는 벡터 9314의 개략도를 보여주는 것이다.

발명을 실시하기 위한 구체적인 내용

[0117] **상세한 설명**

[0118] 하기 설명은 바람직한 실시양태의 것이다.

[0119] 본원에서 사용되는 바, "포함하는," "가진," "비롯한" 및 "함유하는"이라는 용어 및 그의 문법적 파생어는 포괄적이거나, 개방적이며, 추가의 언급되지 않은 요소 및/또는 방법 단계를 배제하지 않는다. 용도 또는 방법과 관련하여 본원에서 사용될 때 "본질적으로 ~로 이루어진"이라는 용어는 추가 요소 및/또는 방법 단계가 존재할 수 있지만, 이러한 추가가 언급된 방법 또는 용도가 작용하는 방식에 실질적으로 영향을 미치지 않는다는 것을 의미한다. 용도 또는 방법과 관련하여 본원에서 사용될 때, "~로 이루어진"이라는 용어는 추가 요소 및/또는 방법

단계의 존재를 배제한다. 특정 요소 및/또는 단계를 포함하는 것으로 본원에 기술된 용도 또는 방법은 또한 특정 실시양태에서, 본질적으로 상기 요소 및/또는 단계로 이루어질 수 있고, 다른 실시양태에서는 이들 실시양태가 구체적으로 언급되었는지 여부와는 상관없이, 상기 요소 및/또는 단계로 이루어질 수 있다. 추가로, 단수형의 사용에는 복수형도 포함되고, "또는"은 달리 언급되지 않는 한, "및/또는"을 의미한다. 본원에서 사용되는 바, "복수"라는 용어는 1 초과, 예를 들어, 2 이상, 3 이상, 4 이상 등을 의미한다. 본원에서 달리 정의하지 않는 한, 본원에서 사용되는 모든 기술 및 과학 용어는 당업계이 숙련가가 일반적으로 이해하는 것과 동일한 의미를 가진다. 본원에서 사용되는 바, "약"이라는 용어는 주어진 값으로부터 대략 +/- 10% 변동을 의미한다. 그러한 변동은 그가 구체적으로 언급되었는지 여부와는 상관없이, 본원에 제공된 임의의 주어진 값에 항상 포함된다는 것을 이해하여야 한다. "포함하는"이라는 용어와 함께 사용될 때, "하나"("a" 또는 "an")라는 단어의 사용은 "하나"를 의미할 수 있지만, 이는 또한 "하나 이상," "적어도 하나" 및 "하나 또는 그 초과"의 의미와도 일치한다.

[0120] 본 설명은 변형된 바이러스 구조 단백질, 및 숙주 또는 숙주 세포에서의 그이 생산에 관한 것이다. 변형된 바이러스 구조 단백질은 연속적으로, 엑토도메인, 막횡단 도메인(TM) 또는 TM의 일부, 및 사이토졸 테일(CT) 도메인 또는 CT의 일부를 포함하고, 여기서, 엑토도메인 및 TM 또는 TM의 일부는 코로나비리다에로부터 유래된 것이고, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것이다.

[0121] 변형된 바이러스 구조 단백질은 변형된 코로나바이러스 구조 단백질일 수 있고, 여기서, 사이토졸 테일 도메인 또는 사이토졸 테일 도메인의 일부는 인플루엔자 헤마글루티닌(HA) 단백질의 사이토졸 테일 도메인 또는 사이토졸 테일 도메인의 일부로 대체된 것이다. 예를 들어, 변형된 바이러스 구조 단백질은 변형된 코로나바이러스 스파이크 또는 표면(S) 단백질일 수 있고, 여기서, S 단백질의 사이토졸 테일 도메인 또는 사이토졸 테일 도메인의 일부는 인플루엔자 헤마글루티닌(HA) 단백질의 사이토졸 테일 도메인 또는 사이토졸 테일 도메인의 일부로 대체된 것이다.

[0122] 본 개시내용은 변형된 바이러스 구조 단백질로서, 여기서, 변형된 바이러스 구조 단백질의 엑토도메인 및 막횡단 도메인은 코로나바이러스 S 단백질의 엑토도메인 및 막횡단 도메인으로부터 유래된 것일 수 있고, 여기서, 사이토졸 테일 도메인은 인플루엔자 헤마글루티닌(HA) 단백질의 사이토졸 테일 도메인으로부터 유래된 것인, 변형된 바이러스 구조 단백질을 제공한다.

[0123] 변형된 S 단백질은 키메라 변형된 S 단백질 또는 키메라 S 단백질일 수 있다. "키메라 S 단백질"이란, 단일 폴리펩티드로 융합되는 2개 이상의 공급원으로부터의 아미노산 서열 및/또는 단백질 도메인 또는 단백질 도메인의 일부를 포함하는 단백질 또는 폴리펩티드를 의미한다. 예를 들어, 제한하는 것은 아니지만, 키메라 S 단백질의 엑토도메인 및 막횡단 도메인(TM) 또는 TM의 일부는 제1 바이러스 구조 단백질, 예를 들어, 코로나바이러스 S 단백질로부터 유래된 것일 수 있고, 세포질 테일(CT) 또는 CT의 일부는 제2 바이러스 구조 단백질로부터 유래된 것일 수 있고, 예를 들어, CT는 인플루엔자 HA로부터 유래된 것일 수 있다. 추가로, 엑토도메인은 제1 바이러스 구조 단백질, 예를 들어, 제1 코로나바이러스 S 단백질로부터 유래된 것일 수 있, TM 또는 TM의 일부는 제2 바이러스 구조 단백질, 예를 들어, 제2 코로나바이러스 S 단백질로부터 유래된 것일 수 있고, CT 또는 CT의 일부는 제3 바이러스 구조 단백질로부터 유래된 것일 수 있고, 예를 들어, CT는 인플루엔자 HA로부터 유래된 것일 수 있다. 따라서, 변형된 S 단백질 또는 키메라 S 단백질은 키메라 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함할 수 있다.

[0124] 변형된 코로나바이러스 S 단백질은 연속적으로,

[0125] - 코로나바이러스 S 단백질로부터 유래된 엑토도메인,

[0126] - 막횡단 및 사이토졸 테일 도메인(TMCT)으로서, 여기서, TMCT는

[0127] - 막횡단 도메인(TM)으로서, 여기서, TM 또는 TM의 일부는 코로나바이러스 S 단백질로부터 유래된 것일 수 있는 TM, 및

[0128] - 사이토졸 테일(CT)로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것일 수 있는 CT를 포함할 수 있는 키메라 TMCT일 수 있는 TMCT를 포함할 수 있다.

[0129] TM 또는 TM의 일부는 CT 또는 CT의 일부에 직접 융합 또는 연결될 수 있거나, 또는 TM 또는 TM의 일부는 개재 펩티드 서열에 의해 CT 또는 CT의 일부에 융합 또는 연결될 수 있다.

[0130] 추가로, TM은 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질 TM으로부터

유래된 C 말단 서열을 포함할 수 있는 키메라 TM일 수 있다. CT는 코로나바이러스 S 단백질 CT로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열을 포함할 수 있는 키메라 CT일 수 있다.

[0131] 따라서, 키메라 TMCT는 천연 코로나바이러스 S 단백질 TM, 키메라 코로나바이러스 S 단백질/인플루엔자 HA TM, 천연 인플루엔자 HA CT, 키메라 인플루엔자 HA/코로나바이러스 S 단백질 CT 또는 그의 조합을 포함할 수 있다. 키메라 코로나바이러스 S 단백질/인플루엔자 HA TM은 코로나바이러스 S 단백질의 TM으로부터의 서열 및 인플루엔자 HA의 TM으로부터의 서열을 포함한다. 유사하게, 키메라 인플루엔자 HA/코로나바이러스 S 단백질 CT는 인플루엔자 HA의 CT로부터의 서열 및 코로나바이러스 S 단백질의 CT로부터의 서열을 포함한다.

[0132] "키메라 막횡단 및 사이토졸 테일 도메인" 또는 "키메라 TMCT"는 코로나바이러스 S 단백질 TMCT에 대해 천연의 것이 아닌 TMCT를 지칭한다. 키메라 TMCT는 자연상에서 발견되지 않는 서열을 포함한다. 따라서, TMCT는 코로나바이러스 S 단백질의 엑토도메인과 이종성인 서열을 포함할 수 있다. "이종성"이라는 용어는 상이한 생물학적 또는 합성 공급원으로부터 기원하는 서열 또는 도메인을 지칭한다. 예를 들어, 키메라 TMCT는 엑토도메인과 동일한 코로나바이러스 S 단백질로부터 유래된 TM 또는 TM의 일부를 포함할 수 있고, 즉, TM은 S 단백질의 엑토도메인과 상동성일 수 있거나, 또는 TM 또는 TM의 일부는 상이한 바이러스 TM, 예를 들어, 엑토도메인과 상이한 코로나바이러스 S 단백질로부터의 TM으로부터 유래될 수 있고, 즉, TM은 S 단백질의 엑토도메인과 이종성일 수 있다. CT 또는 CT의 일부는 변형된 S 단백질의 엑토도메인, TM, 또는 엑토도메인 및 TM, 둘 모두와 이종성인 CT로부터 유래될 수 있다.

[0133] 코로나바이러스 S 단백질, 변형된 S 단백질 또는 변형된 코로나바이러스 S 단백질의 엑토도메인 및 막횡단 도메인 또는 막횡단 도메인의 일부는 코로나비리다에 과 바이러스의 임의의 구성원으로부터 유래된 것일 수 있다. 예를 들어, 코로나바이러스 S 단백질, 변형된 S 단백질 또는 변형된 코로나바이러스 S 단백질의 엑토도메인 및 막횡단 도메인은 예를 들어, 코로나바이러스, 예컨대, 알파코로나바이러스(알파-CoV), 베타코로나바이러스(베타-CoV), 감마코로나바이러스(감마-CoV) 또는 델타코로나바이러스(델타-CoV)으로부터 유래된 것일 수 있다. 예를 들어, 코로나바이러스는 알파코로나바이러스(알파-CoV) 또는 베타코로나바이러스(베타-CoV)일 수 있다. 알파코로나바이러스는 뒤비나코바이러스(Duvinacovirus), 예컨대, 예를 들어, HCoV-229E(229E-CoV)일 수 있거나, 또는 세트라코바이러스(Setracovirus), 예컨대, 예를 들어, HCoV-NL63일 수 있다. 바람직한 실시양태에서, 코로나바이러스는 베타코로나바이러스(베타-CoV)이다. 베타코로나바이러스는 계통 A 베타코로나바이러스, 예컨대, 예를 들어, HCoV-OC43(OC43-CoV) 또는 HCoV-HKU1(HKU1-CoV), 계통 B 베타코로나바이러스, 예컨대, 예를 들어, SARS-CoV(이는 또한 SARS-CoV-1로도 지칭) 또는 SARS-CoV-2 및 그의 변이체, 또는 계통 C 베타코로나바이러스, 예컨대, 예를 들어, MERS-CoV일 수 있다.

[0134] 코로나바이러스 S 단백질, 변형된 S 단백질 또는 변형된 코로나바이러스 S 단백질의 엑토도메인 및 막횡단 도메인 또는 막횡단 도메인의 일부는 B.1.1.7 계통("알파" 변이체)(20I/501Y.V1, MW531680.1), B.1.351 계통("베타" 변이체)(20H/501Y.V2), P.1 계통("감마" 변이체)(20J/501Y.V3), B 1.617.2 계통("델타" 변이체), B.1.525 균주, B.1.429 계통("ETA" 변이체) 또는 코로나바이러스 S 단백질에서 자연적으로 발생하는 돌연변이를 포함하는 균주의 다른 변이체, 또는 그의 자연적으로 발생된 재조합 균주를 포함하나, 이에 제한되지 않는, SARS-CoV-2 계통의 변이체로부터 유래된 것일 수 있다.

[0135] 한 실시양태에서, 변형된 바이러스 구조 단백질의 엑토도메인 및 막횡단 도메인 또는 막횡단 도메인의 일부는 SARS-CoV-2 계통(이는 또한 SARS-CoV-2 변이체로도 지칭)의 코로나바이러스의 스파이크 단백질(S)로부터 유래된 것이다. 다른 실시양태에서, 변형된 바이러스 구조 단백질의 엑토도메인 및 막횡단 도메인 또는 막횡단 도메인의 일부는 SARS-CoV-1, MERS-CoV, OC43-CoV 또는 229E-CoV 또는 그의 변이체의 스파이크 단백질(S)로부터 유래된 것이다.

[0136] 변형된 바이러스 구조 단백질과 관련하여, 본원에서 사용되는 바, "변형된"이라는 용어는 코로나비리다에로부터의 구조 단백질 중의 세포질 테일 도메인(CT) 또는 CT의 일부의 이종성 바이러스의 CT 또는 CT의 일부로의 대체를 지칭할 수 있다. 예를 들어, 변형된 바이러스 구조 단백질은 S 단백질의 CT 또는 CT의 일부가 인플루엔자 헤마글루티닌(HA)의 CT 또는 CT의 일부로 대체된 코로나바이러스 S 단백질일 수 있다.

[0137] 따라서, 변형된 바이러스 구조 단백질은 막횡단 도메인(TM) 또는 TM의 일부, 및 사이토졸 테일(CT) 또는 CT의 일부를 포함하는 변형된 코로나바이러스 스파이크(S) 단백질로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것일 수 있고, 여기서, TM 또는 TM의 일부는 CT 또는 CT의 일부와 이종성인 것인, 변형된 코로나바이러스 스파이크(S) 단백질일 수 있다. 추가로, 변형된 S 단백질은 막횡단 도메인(TM) 또는 TM의 일부, 및 사이토졸 테일(CT) 또는 CT의 일부를 포함하고, 여기서, CT 또는 CT의 일부는 인플루엔자

헤마글루티닌(HA) 단백질로부터 유래된 것일 수 있고, 여기서, CT 또는 CT의 일부는 TM 또는 TM의 일부와 이종성인 것이다.

[0138] 따라서, 한 측면에서, 막횡단 도메인(TM) 또는 TM의 일부, 및 사이토졸 테일(CT) 또는 CT의 일부를 포함하는 변형된 코로나바이러스 스파이크(S) 단백질로서, 여기서, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것이고, 여기서, TM 또는 TM의 일부는 CT 또는 CT의 일부와 이종성인 것인, 변형된 코로나바이러스 스파이크(S) 단백질을 제공한다. 변형된 코로나바이러스 스파이크(S) 단백질은 또한 변형된 S 단백질로도 지칭된다.

[0139] 세포질 테일 도메인은 또한 "세포질 테일," "사이토졸 테일," "사이토졸 테일 도메인," "CT, "CTD," "세포질(cytoplasmic) 도메인," "세포질(cytoplasm) 도메인," "CP, "CPD" 또는 "C 말단 도메인" 및 유사한 표현으로도 지칭될 수 있다. 세포질 테일 도메인은 또한 세포질 테일 도메인의 일부를 포함할 수 있다.

[0140] 변형된 바이러스 구조 단백질, 예컨대, 본원에 개시된 바와 같은 변형된 S 단백질은 야생형 또는 비변형된 바이러스 구조 단백질(예를 들어, S 단백질)과 비교하여 개선된 특징을 가진 것으로 밝혀졌다. 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질의 개선된 특징의 예로는 숙주 또는 숙주 세포에서 발현되었을 때, 야생형 또는 비변형된 바이러스 구조 단백질과 비교하여 변형된 바이러스 구조 단백질 수율 증가; 숙주 또는 숙주 세포에서 발현되었을 때, 야생형 또는 비변형된 바이러스 구조 단백질과 비교하여 바이러스 구조 단백질의 무결성, 안정성, 또는 무결성 및 안정성, 둘 모두의 개선; 본원에 기술된 바와 같은 변형을 포함하지 않는 바이러스 구조 단백질을 포함하는 VLP의 무결성, 안정성 또는 무결성 및 안정성과 비교하여, 변형된 바이러스 구조 단백질로 구성된 바이러스 유사 입자(VLP)의 무결성, 안정성, 또는 무결성 및 안정성, 둘 모두의 개선; 숙주 세포에서 발현되었을 때, 동일하거나, 또는 실질적으로 유사한 숙주 세포에서 발현되는 변형된 바이러스 구조 단백질을 포함하지 않는 VLP의 수율과 비교하여 변형된 바이러스 구조 단백질을 포함하는 VLP의 수율 증가를 포함하나, 이에 제한되지 않는다.

[0141] 추가로, 숙주 또는 숙주 세포에서 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질을 포함하는 바이러스 유사 입자(VLP)를 생산하는 방법 또한 기술한다. 천연 또는 야생형 CT가 본원에 기술된 바와 같은 인플루엔자 HA의 CT로 대체된 변형된 S 단백질과 같은 변형된 바이러스 구조 단백질을 포함하는 VLP가 생산되었을 때, 숙주에서의 VLP 생산 수율은 i) 천연 CT를 포함하거나, 또는 ii) 막횡단 도메인(TM) 및 CT가 인플루엔자 HA의 TM 및 CT로 대체된 변형된 바이러스 구조 단백질을 포함하는 바이러스 구조 단백질을 포함하는 VLP의 수율과 비교하여 증가된 것으로 관찰되었다.

[0142] 막횡단 도메인은 또한 "TM" 또는 "TMD"로도 지칭될 수 있다. 막횡단 및 세포질 테일 도메인은 TMCT 또는 TM/CT로 지칭될 수 있다.

[0143] 도 3a는 변형된 S 단백질(예컨대, 변형된 SARS-CoV-2 S 단백질)이 식물에서 발현되었을 때, 변형된 S 단백질의 수율 또는 단백질 축적(변화 배수로 표시)은 천연 막횡단 및 세포질 테일(TMCT)이 인플루엔자 HA로부터의 TMCT로 대체된 경우(구성물 8592, 8595, 및 8597), 천연 TMCT(구성물 8586, 8589, 및 8591)를 포함하는 S 단백질의 수율 또는 단백질 축적과 비교하여 대략 2배 증가되었다는 것을 보여주는 것이다. 추가로, 오직 세포질 테일(CT)만이 인플루엔자 HA의 CT로 대체된 경우(구성물 8610, 8611, 및 8671), 변형된 S 단백질(예컨대, 변형된 SARS-CoV-2 S 단백질)이 식물에서 발현되었을 때, 인플루엔자 HA의 CT를 포함하는 변형된 S 단백질의 단백질 축적(변화 배수로 표시)은 TMCT가 인플루엔자 HA의 TMCT로 대체된 변형된 S 단백질의 축적과 비교하여 대략 1.74 내지 2.14배로 추가로 증가하였다. 상응하여, 인플루엔자 HA의 CT를 포함하는 변형된 S 단백질의 단백질 축적은 천연 막횡단 및 세포질 테일(wtTMCT)을 포함하는 S 단백질의 축적과 비교하여 대략 3.57 내지 4.40배 증가하였다.

[0144] 도 3b는 미정제 식물 추출물로부터의 야생형 TMCT(wt TMCT)를 포함하는 S 단백질, 또는 인플루엔자 HA의 TMCT(H5i TMCT)를 포함하는 변형된 S 단백질의 단백질 축적과 비교하여 인플루엔자 HA로부터의 세포질 테일(H5i CT)을 포함하는 변형된 S 단백질(변형된 SARS-CoV-2 S 단백질)에 대해 단백질 축적이 더 높은 것으로 관찰되었다는 것을 보여주는 것이다. 인플루엔자 HA로부터의 세포질 테일(H5i CT)을 포함하는 변형된 S 단백질 쿠마시 염색 단독으로 시각화된다. 인플루엔자 HA로부터의 세포질 테일(H5i CT)을 포함하는 변형된 S 단백질에 대한 밴드가 야생형 TMCT(wt TMCT)를 포함하는 S 단백질 또는 인플루엔자 HA의 TMCT(H5i TMCT)를 포함하는 변형된 S 단백질의 밴드와 비교하여 더욱 뚜렷하고, 더 두껍다 - S 단백질로 마킹된 약 150 kDa의 밴드를 참조한다. 밴드 두께는 존재하는 단백질 양에 상응하며, H5i CT S 단백질에 대하여 더 많은 단백질이 축적되었다는 것을 나타내는 것이다. 사용된 발현 인핸서와는 상관없이 더 높은 단백질 축적인 관찰되었다.

[0145] 하기에서 더 상세하게 추가로 논의되는 바와 같이, 변형된 S 단백질이 인플루엔자 HA로부터의 세포질 테일을 포함하는 SARS-CoV-1 S 단백질(도 16a 참조), 또는 인플루엔자 HA로부터의 세포질 테일을 포함하는 MERS CoV S 단백질(도 19a 참조)을 포함하는 경우에 유사한 결과을 수득하였다.

[0146] 도 3c는 미정제 식물 추출물의 웨스턴 블롯 분석에 의한 S 단백질(SARS-CoV-2 S 단백질) 축적을 보여주는 것이다. 인플루엔자 HA로부터의 세포질 테일(H5i CT)을 포함하는 변형된 S 단백질이 식물에서 발현되었을 때, 변형된 S 단백질의 축적은, 야생형 TMCT(wt TMCT), 및 막횡단 도메인 및 세포질 테일(TMCT) 도메인, 둘 모두 인플루엔자 HA로부터의 TMCT(H5i TMCT)로 대체된 변형된 S 단백질과 비교하여 더 높은 것으로 관찰되었다. 도 3c의 웨스턴 블롯 분석은 SARS CoV-2 S 단백질이 S1 도메인/서브유닛(상단 패널, 항SARS-CoV-2 S1 항체로 검출) 및 S2 도메인/서브유닛(하단 패널, 항SARS-CoV-2 S2 항체로 검출), 둘 모두를 포함하고, 분자량이 약 150 kDa이라는 것도 추가로 보여준다.

[0147] 본 기술내용은 변형된 코로나바이러스 스파이크 또는 표면 단백질(S 단백질)일 수 있는 변형된 바이러스 구조 단백질을 제공한다. 변형된 S 단백질은 연속적으로, 엑토도메인, 막횡단 도메인(TM) 또는 TM의 일부, 및 사이토졸 테일(CT) 도메인 또는 CT의 일부를 포함하고, 여기서, 엑토도메인 및 막횡단 도메인은 코로나바이러스로부터 유래된 것이고, CT 또는 CT의 일부는 인플루엔자 헤마글루티닌(HA) 단백질의 CT로부터 유래된 것이다. 엑토도메인 및 막횡단 도메인 또는 TM의 일부는 동일한 코로나바이러스로부터 유래된 것일 수 있다. 따라서, 변형된 구조 단백질의 엑토도메인 및 막횡단 도메인 또는 TM의 일부는 서로 동종성인 반면(즉, 이종성이 아닌 반면), CT 또는 CT의 일부는 엑토도메인 및 막횡단 도메인과 이종성이다.

[0148] 추가로, 변형된 S 단백질의 막횡단 도메인(TM) 또는 TM의 일부는 엑토도메인과 상이한 코로나바이러스로부터 유래된 것일 수 있다. 따라서, 변형된 S 단백질 중의 TM 또는 TM의 일부는 변형된 S 단백질의 엑토도메인 및 CT 도메인 또는 CT의 일부, 둘 모두와 이종성일 수 있다(동종성이 아닐 수 있다). 유사하게, 엑토도메인은 변형된 S 단백질의 TM 또는 TM의 일부 및 CT 도메인 또는 CT의 일부와 이종성일 수 있다(동종성이 아닐 수 있다). 예를 들어, 변형된 S 단백질의 엑토도메인은 제1 코로나바이러스로부터 유래된 것일 수 있고, TM 또는 TM의 일부는 제2 코로나바이러스로부터 유래된 것일 수 있고, CT 또는 CT의 일부는 인플루엔자 HA로부터 유래된 것일 수 있다. 제1 코로나바이러스 및 제2 코로나바이러스는 상이한 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주에 속하는 것일 수 있다. 따라서, 제1 코로나바이러스 및 제2 코로나바이러스는 서로 이종성일 수 있고, 이는 또한 각각 CT 또는 CT의 일부의 유래 기점이 되는 바이러스 과와 이종성일 수 있다.

[0149] 예를 들어, S 단백질 엑토도메인의 유래 기점이 되는 제1 코로나바이러스는 임의의 코로나바이러스, 예를 들어, 알파코로나바이러스(알파-CoV) 또는 베타코로나바이러스(베타-CoV)로부터의 것일 수 있다. S 단백질의 엑토도메인의 유래 기점이 될 수 있는 제1 코로나바이러스의 비제한적인 예는 뒤비나코바이러스, 예를 들어, HCoV-229E, 세트라코바이러스, 예를 들어, HCoV-NL63, 계통 A 베타코로나바이러스, 예를 들어, HCoV-OC43 또는 HCoV-HKU1, 계통 B 베타코로나바이러스, 예를 들어, SARS-CoV 또는 SARS-CoV 2 또는 계통 C 베타코로나바이러스, 예를 들어, MERS-CoV이다. TM의 유래 기점이 되는 제2 코로나바이러스는 엑토도메인의 유래 기점이 되는 제1 코로나바이러스와 상이한 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주에 속하는 것일 수 있다. 예를 들어, S 단백질 TM의 유래 기점이 되는 제2 코로나바이러스는 제2 코로나바이러스가 제1 코로나바이러스와 이종성인 한, 임의의 코로나바이러스 예를 들어, 알파코로나바이러스(알파-CoV) 또는 베타코로나바이러스(베타-CoV)로부터 유래된 것일 수 있다. S 단백질의 TM의 유래 기점이 될 수 있는 제2 코로나바이러스의 비제한적인 예는 뒤비나코바이러스, 예를 들어, HCoV-229E(이는 또한 229E-CoV로도 지칭), 세트라코바이러스, 예를 들어, HCoV-NL63 (NL63-CoV), 계통 A 베타코로나바이러스, 예를 들어, HCoV-OC43(이는 또한 OC43-CoV로도 지칭) 또는 HCoV-HKU1(HKU1-CoV), 계통 B 베타코로나바이러스, 예를 들어, SARS-CoV(이는 또한 SARS-CoV 1로도 지칭) 또는 SARS-CoV 2 또는 계통 C 베타코로나바이러스, 예를 들어, MERS-CoV(이는 또한 "MERS"로도 지칭)이다.

[0150] 코로나바이러스 S 단백질, 예컨대, SARS-CoV-1 S 단백질, SARS-CoV-2 S 단백질, MERS CoV S 단백질, OC43-CoV S 단백질, 또는 229E-CoV S 단백질 중의 도메인은 당업계에 공지되어 있는 방법에 의해 쉽게 확인될 수 있다. 예를 들어, 도메인 예컨대, 막횡단 도메인은 예를 들어, 막횡단 예측 프로그램(예컨대, Expert 단백질 분석 System; ExPASy.org, 스위스 생물정보학 연구소(Swiss Institute of Bioinformatics)에 의해 운용; 또는 문헌 [Dense Alignment Surface Method, Cserzo M., et al. 1997, Prot. Eng. vol. 10, no. 6, 673-676]; [Lolkema J.S. 1998, FEMS Microbiol Rev. 22, no 4, 305-322])을 사용하여 단백질의 아미노산 서열의 소수성 정도를 측정함으로써, 단백질의 아미노산 서열의 하이드로패시 프로파일(예컨대, Kyte-Doolittle Hydropathy Profile)을 측정함으로써, 3차원 단백질 구조를 측정하고, 막에서 열역학적으로 안정적인 구조(예컨대, 단일 알파 나선, 수개의 막횡단 알파 나선으로 이루어진 안정적인 복합체, 막횡단 베타 배럴, 베타-나선, 또는 막에서 열역학적으

로 안정적인 임의의 다른 구조)를 확인함으로써 확인될 수 있다.

[0151] 추가로, 코로나바이러스 S 단백질 내의 도메인은 예를 들어, 서열 정렬에 의해 알려진 단백질 서열과 비교함으로써 결정될 수 있다. 비교를 위한 서열 정렬 방법은 당업계에 널리 공지되어 있으며, 하기에 추가로 기술되는 바와 같다.

[0152] 코로나바이러스 S 단백질의 도메인 및 도메인 조직은 널리 공지되어 있고, 기술되어 있다. 모든 코로나바이러스 스파이크 단백질(S 단백질)은 2개의 서브유닛 또는 도메인에서 동일한 조직을 공유한다: 수용체 결합을 담당하는, S1로 명명되는 N 말단 서브유닛(또는 도메인) 및 바이러스 부착, 막 융합 및 바이러스 진입을 담당하는 C 말단 S2 서브유닛(또는 도메인).

[0153] 도 1a는 코로나바이러스 S 단백질의 서브유닛 및 도메인 및 S1/S2 S2' 절단 부위의 위치와 함께 코로나바이러스 S 단백질의 개략도를 보여주는 것이다. S1 서브유닛은 바이러스 막에 대해 원위부에 있고, 그의 숙주 수용체에 대한 바이러스 부착을 매개하는 수용체 결합 도메인(RBD)을 포함한다. S2 서브유닛은 융합 단백질 기구, 예컨대, 융합 펩티드, 2개의 헵타드 반복 서열(HR1 및 HR2), 융합 당단백질 및 막횡단 도메인에 전형적인 중심 나선, 및 사이토졸 테일 도메인을 포함한다(예를 들어, 문헌 [Kirchdoerfer et al. Nature 2016 Mar 3;531(7592):118-2](상기 문헌은 본원에서 참조로 포함된다) 참조).

[0154] 막횡단 도메인(TM) 및 세포질 테일 도메인(CT)은 S2 서브유닛의 C 말단 단부에 위치한다. 이들 도메인은 모든 코로나바이러스에서 보존되지만(도 1a 및 문헌 [Corver et al. 2009, Virol J. 2009; 6: 230](상기 문헌은 본원에서 참조로 포함된다) 참조), 상이한 참조, 그룹 및 저자는 이들 도메인과 관련하여 상이한 아미노산 넘버링을 언급하였다.

[0155] 예를 들어, SARS-CoV-2의 S 단백질에서 아미노산(aa: amino acid): 1214-1234가 TM에 할당될 수 있고, aa 1235-1273이 CT에 할당될 수 있다(예를 들어, UniProtKB - P0DTC2 (SPIKE_SARS2) 참조). SARS-CoV-2(서열 번호. 1)의 서열을 문헌 [Kirchdoerfer et al. (Nature 2016 Mar 3;531(7592):118-2)]의 SARS-CoV-1의 서열과 정렬할 때, SARS-CoV-2 TM은 아미노산 1214-1236에 상응하고, SARS-CoV-2 CT는 아미노산: 1237-1273에 상응한다.

[0156] 본 개시내용의 목적을 위해, 천연(비변형된) S 단백질의 TM 및 CT는 코로나바이러스 S 단백질 참조 서열(서열 번호 1)과 정렬되었을 때 하기 아미노산에 상응한다: TM: 아미노산 1214-1234 및 CT: 아미노산: 1235-1273.

[0157] 신호 펩티드(SP)를 포함하는 15개의 아미노산이 S 단백질로부터 제거될 때, TM은 참조 서열 서열 번호 2의 아미노산 1199-1219에 상응하고, CT는 서열 번호 2의 아미노산 1220-1258에 상응한다(또한 참조 서열 및 넘버링에 대해서는 표 1 참조).

[0158] 코로나바이러스 S 단백질의 TM은 소수성 서열이 뒤따르는 고도로 보존된 N 말단의 방향족이 풍부한 스트레치를 가진다(도 22 및 문헌 [Corver et al. Virology Journal volume 6, 230 (2009)] 참조). 코로나바이러스 S 단백질 TM 도메인의 컨센서스 서열은 하기와 같다:

[0159] WYXWLGFIAGLXAXXX{X}VXXXL(서열 번호 132), (여기서, {X}는 부재할 수 있다).

[0160] 예를 들어, 코로나바이러스 S 단백질 TM 도메인 컨센서스 서열은

[0161] WY[I/V]WLGFIAGL[V/I]A[L/I][A/V][L/M]{X}V[F/T][F/I]XL(서열 번호 133), (여기서, {X}는 C일 수 있거나, 또는 부재할 수 있다)일 수 있다.

표 1. 변형된 S 단백질 중의 TM 및 CT 도메인의 위치 및 참조 서열 중의 상응하는 아미노산 위치의 비제한적인 예.

S 단백질	막횡단 도메인 (TM)	세포질 테일 도메인(CT)
변형된 S 단백질[1] **[SARS-CoV-2 H5iCT]** (서열 번호 21)	1199-1219	1220-1235
SARS-CoV-2[2] (서열 번호 1)	1214-1234	1235-1273
SARS-CoV-2[3] (서열 번호 2)	1199-1219	1220-1258
SARS-CoV (SARS-CoV-1)[2] (서열 번호 112)	1196-1216	1217-1255
SARS-CoV (SARS-CoV-1)[3] (서열 번호 114)	1183-1203	1204-1242
MERS (MERS-CoV)[2] (서열 번호 113)	1297-1318	1319-1353
MERS (MERS-CoV)[3] (서열 번호 115)	1280-1301	1302-1336
OC43-CoV[2] (서열 번호 158)	1233-1325	1326-1360
OC43-CoV[3] (서열 번호 160)	1291-1311	1312-1346
229E-CoV[2] (서열 번호 159)	1116-1135	1136-1173
229E-CoV[3] (서열 번호 161)	1100-1119	1120-1157

[1] 넘버링은 신호 펩티드를 배제하고, CT는 인플루엔자 HA로부터 유래된 것이다.
[2] 넘버링은 신호 펩티드를 포함하고, CT는 천연 CT이다.
[3] 넘버링은 신호 펩티드를 배제하고, CT는 천연 CT이다.

[0162] 키메라 TMCT이며,

[0163] TM 및 CT 도메인에 할당된 잔기의 번호 넘버링에 있어 차이가 있지만, 당업자는 예를 들어, 하기 기술되는 바와 같이, 공지된 방법을 사용하여 코로나바이러스 S 단백질 중의 성가 도메인의 경계 또는 바운더리를 결정할 수 있을 것이다.

[0164] 변형된 코로나바이러스 S 단백질에서, 인플루엔자 HA로부터 유래될 수 있는 이종성 CT 또는 CT의 일부는 코로나바이러스 S 단백질의 TM 또는 TM의 일부의 C 말단 단부에 직접 융합될 수 있거나, 이종성 CT 또는 CT의 일부는 개재 펩티드 서열(이는 또한 링커 또는 링커 서열로도 지칭)을 사용하여 코로나바이러스 S 단백질의 TM 또는 TM의 일부의 C 말단 단부에 융합될 수 있다. 따라서, 변형된 S 단백질은 개재 펩티드을 포함할 수 있고, 여기서, 개재 펩티드 서열은 CT 또는 CT의 일부를 TM 또는 TM의 일부의 C 말단 단부에 융합에 융합시킨다.

[0165] 이종성 CT, CT의 일부 또는 이종성 CT와 함께 개재 펩티드 서열은 (예를 들어, 서열 번호 1, 2, 21, 114, 115, 160 또는 161과 관련하여 표 1에 정의된 바와 같이, TM 도메인의 C 말단의 4개의 아미노산 내에서) TM 도메인의 C 말단부, 또는 (예를 들어, 서열 번호 1, 2, 21, 114, 115, 160 또는 161과 관련하여 표 1에 정의된 바와 같이, CT 도메인의 N 말단의 4개의 아미노산 내에서) CT 도메인의 N 말단부의 아미노산에 융합될 수 있다.

[0166] 예를 들어, 코로나바이러스 TM은 서열 번호 1의 아미노산 1230-1238 중 어느 하나에 상응하는 아미노산 잔기에서 종료될 수 있다. 따라서, 코로나바이러스 TM의 C 말단 단부는 서열 번호 1의 아미노산 1230-1238 중 어느 하나에 상응하는 아미노산일 수 있다. 한 예에서, 코로나바이러스 TM은 서열 번호 1의 아미노산 1230에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM은 서열 번호 1의 아미노산 1231에 상응하는 아미노산 잔기에서 종료될 수 있다. 추가 예에서, TM은 서열 번호 1의 아미노산 1232에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM은 서열 번호 1의 아미노산 1233에 상응하는 아미노산 잔기에서 종료될 수 있다. 추가 예에서, TM은서열 번호 1의 아미노산 1234에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM은 서열 번호 1의 아미노산 1235에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM은 서열 번호 1의 아미노산 1236에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM은 서열 번호 1의 아미노산 1237에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM은 서열 번호 1의 아미노산 1238에 상응하는 아미노산 잔기에서 종료될 수 있다. 바람직한 실시양태에서, TM은 서열 번호 1의 아미노산 1234에 상응하는 아미노산 잔기에서 종료될 수 있다.

[0167] 또 다른 예에서, 코로나바이러스 TM 또는 TM의 일부는 서열 번호 2 또는 21의 아미노산 1215-1219 중 어느 하나

에 상응하는 아미노산 잔기에서 종료될 수 있다. 따라서, 코로나바이러스 TM 또는 TM의 일부의 C 말단 단부는 서열 번호 2 또는 21의 아미노산 1215-1224 중 어느 하나에 상응하는 아미노산일 수 있다. 한 예에서, 코로나바이러스 TM 또는 TM의 일부는 서열 번호 2 또는 21 중 아미노산 1215에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM 또는 TM의 일부는 서열 번호 2 또는 21 중 아미노산 1216에 상응하는 아미노산 잔기에서 종료될 수 있다. 추가 예에서, TM 또는 TM의 일부는 서열 번호 2 또는 21 중 아미노산 1217에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM 또는 TM의 일부는 서열 번호 2 또는 21 중 아미노산 1218에 상응하는 아미노산 잔기에서 종료될 수 있다. 또 다른 예에서, TM 또는 TM의 일부는 서열 번호 2 또는 21 중 아미노산 1219에 상응하는 아미노산 잔기에서 종료될 수 있다.

[0168] 코로나바이러스 S 단백질로부터의 TM 또는 TM의 일부의 C 말단 단부에 이종성 CT를 융합시킬 수 있는 개재 펩티드 서열의 길이는 0-10개의 아미노산 길이일 수 있다. 따라서, 개재 펩티드 서열의 길이는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 아미노산 길이일 수 있다. 개재 펩티드 서열은 코로나바이러스 단백질로부터 유래될 수 있고, 예를 들어, 개재 펩티드 서열은 코로나바이러스 S 단백질로부터의 TM의 C 말단 단부로부터, 또는 코로나바이러스 S 단백질의 CT의 N 말단 단부로부터, 또는 그 둘로부터 유래된 것일 수 있다. 개재 펩티드 서열은 추가로 인플루엔자 HA 단백질로부터 유래된 것일 수 있고, 예를 들어, 개재 펩티드 서열은 인플루엔자 HA 단백질로부터의 TM의 C 말단 단부로부터, 또는 인플루엔자 HA 단백질의 CT의 N 말단 단부로부터, 또는 그 둘로부터 유래된 것일 수 있다. 추가로, 개재 펩티드 서열은 코로나바이러스 및/또는 변형된 S 단백질의 HA 부분과 이종성일 수 있거나, 또는 개재 펩티드 서열은 인공 서열일 수 있다.

[0169] 변형된 S 단백질의 (키메라 TMCT로도 또한 지칭되는) TM/CT 도메인의 서열의 비제한적인 예는 하기 제시되어 있다. 코로나바이러스 S 단백질로부터의 TM 도메인의 서열은 밑줄체로 표시되어 있고, 인플루엔자 HA 유래 CT 도메인은 굵은체로 표시되어 있다. 굵은 이탤릭체로 표시된 서열은 인플루엔자 HA의 TM으로부터 유래된 서열이다. 밑줄체의 이탤릭체로 표시된 서열은 코로나바이러스 S 단백질의 CT로부터 유래된 서열이다.

SARS-CoV-2

- WYIWLGFIAGLIAIVMVTIML**SLWMCSNGSLQCRICI** (서열 번호 18) (wtTM/H5iCT)
- WYIWLGFIAGLIAIVMVTIM***MAGL*SLWMCSNGSLQCRICI** (서열 번호 19)(wtTM/H5iCT V1)
- WYIWLGFIAGLIAIVMVTIM***AGL*SLWMCSNGSLQCRICI** (서열 번호 37) (wtTM/H5iCT V2)
- WYIWLGFIAGLIAIVMVTIML*CCM***CSNGSLQCRICI** (서열 번호 38) (wtTM/H5iCT V3)
- WYIWLGFIAGLIAIVMVTIML*CC***SNGSLQCRICI** (서열 번호 39) (wtTM/H5iCT V4)
- WYIWLGFIAGLIAIVMVTIML**SFWMCSNGSLQCRICI** (서열 번호 126)(wtTM/H1iCT)
- WYIWLGFIAGLIAIVMVTIML**MWACQKGNIRCNICI** (서열 번호 127) (wtTM/H3iCT)
- WYIWLGFIAGLIAIVMVTIML**GLWMCSNGSMQCRICI** (서열 번호 128) (wtTM/H6iCT)
- WYIWLGFIAGLIAIVMVTIML**VFICVKNGNMRCTICI** (서열 번호 129) (wtTM/H7iCT)
- WYIWLGFIAGLIAIVMVTIML**LFWAMSNGSCRCNICI** (서열 번호 130) (wtTM/H9iCT)
- WYIWLGFIAGLIAIVMVTIML**VVYMVSRDNVSCSICL** (서열 번호 131) (wtTM/BiCT)

[0170]

SARS-CoV-1

- WYVWLGFIAGLIAIVMVTILL**SLWM CSNGSLQCRICI** (서열 번호 118) (wtTM/H5iCT)
- WYVWLGFIAGLIAIVMVTILL*CC***SNGSLQCRICI** (서열 번호 119) (wtTM/H5iCT V4)
- WYVWLGFIAGLIAIVMVTILL**SFWM CSNGSLQCRICI** (서열 번호 120) (wtTM/H1cCT)

MERS-CoV

- WYIWLGFIAGLVALALCVFFIL**SLWMCSNGSLQCRICI** (서열 번호 123) (wtTM/H5iCT)
- WYIWLGFIAGLVALALCVFFIL*CC***SNGSLQCRICI** (서열 번호 124) (wtTM/H5iCT V4)
- WYIWLGFIAGLVALALCVFFIL**SFWMCSNGSLQCRICI** (서열 번호 125) (wtTM/ H1cCT)

OC43-CoV

- WYVWLLICLAGVAMLVLLFFI**SLWMCSNGSLQCRICI** (서열 번호 164) (wtTM/H5iCT)
- WYVWLLICLAGVAMLVLLFFI*CC***SNGSLQCRICI** (서열 번호 165) (wtTM/H5iCT V4)
- WYVWLLICLAGVAMLVLLFFI**SFWMCSNGSLQCRICI** (서열 번호 166) (wtTM/ H1cCT)

229E-CoV

- WWVWLCISVVLIFVVSMLLL**SLWMCSNGSLQCRICI** (서열 번호 169) (wtTM/H5iCT)
- WWVWLCISVVLIFVVSMLLL*CC***SNGSLQCRICI** (서열 번호 170) (wtTM/H5iCT V4)
- WWVWLCISVVLIFVVSMLLL**SFWMCSNGSLQCRICI** (서열 번호 171) (wtTM/ H1cCT)

[0171]

[0172] 변형된 코로나바이러스 S 단백질은 키메라 TMCT를 포함할 수 있다. 예를 들어, 키메라 TMCT는 표 1b에 제시된 바와 같이 코로나바이러스 S 단백질로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질로부터 유래된 C 말단 서열을 포함할 수 있다. TM은 "S 단백질 TM 서열"로 라벨링된 열에 제시된 바와 같은 서열을 포함할 수 있고, CT는 "HA 단백질 CT 서열"로 라벨링된 열에 제시된 바와 같은 서열을 포함할 수 있다. CT 및 TM은 "S 단백질 CT 서열" 및/또는 "HA 단백질 TM 서열"로 라벨링된 열에 제시된 바와 같은 서열(이는 또한 하기 추가로 기술되는 바와 같이, 개재 서열 또는 링커로도 지칭)에 의해 연결될 수 있다.

표 1B: 변형된 S 단백질 중의 TM 및 CT 서열의 비제한적인 예. 참조 서열 내의 아미노산 위치가 명시되어 있다.

서열 번호	S 단백질 TM 서열	S 단백질 CT 서열	HA 단백질 TM 서열	HA 단백질 CT 서열
18	1-21	-	-	22-37
19	1-20	-	21-24	25-40
37	1-20	-	21-23	24-39
38	1-21	22-24	-	25-36
39	1-21	22-23	-	24-34
126	1-21			22-37
127	1-21			22-36
128	1-21			22-37
129	1-21			22-37
130	1-21			22-37
131	1-21			22-37
118	1-21			22-37
119	1-21	22-23	-	24-34
120	1-21			22-37
123	1-22			23-38
124	1-22	23-24	-	25-35
125	1-22			23-38
164	1-21			22-37
165	1-21	22-23	-	24-34
166	1-21			22-37
169	1-20			21-36
170	1-20	21-22	-	23-33
171	1-20			21-36

[0173]

[0174] 예를 들어, 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열은 적어도 하기를 포함할 수 있다:

[0175] - 서열 번호 18, 19, 37, 38, 39, 118, 119, 123, 124, 164, 165, 169 또는 170의 아미노산 1-19에 상응하는 적어도 19개의 아미노산;

[0176] 서열 번호 18, 19, 37, 38, 39, 118, 119, 123, 124, 164, 165, 169 또는 170의 아미노산 1-20에 상응하는 적어도 20개의 아미노산;

[0177] 서열 번호 18, 19, 37, 38, 39, 118, 119, 123, 124, 164, 165, 169 또는 170의 아미노산 1-21에 상응하는 적어도 21개의 아미노산;

[0178] 서열 번호 18, 19, 37, 38, 39, 118, 119, 123, 124, 164, 165, 169 또는 170의 아미노산 1-22에 상응하는 적어도 22개의 아미노산;

[0179] 서열 번호 18, 19, 37, 38, 39, 118, 119, 123, 124, 164, 165, 169 또는 170의 아미노산 1-23에 상응하는 적어도 23개의 아미노산;

[0180] 서열 번호 18, 19, 37, 38, 39, 118, 119, 123, 124, 164, 165, 169 또는 170의 아미노산 1-24에 상응하는 적어도 24개의 아미노산.

[0181] 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열은 서열 번호 18 또는 169의 아미노산 1-20에 상응하는 적어도 20개의 아미노산, 또는 서열 번호 118 또는 164의 아미노산 1-21에 상응하는 적어도 21개의 아미노산, 또는 서열 번호 123의 아미노산 1-22에 상응하는 적어도 22개의 아미노산 및 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산을 포함할 수 있다. 코로나바이러스 S 단백질 TM으로부터 유래된 N 말단 서열은 서열 번호 18 또는 169의 아미노산 1-20에 상응하는 적어도 20개의 아미노산, 또는 서열 번호 118 또는 164의 아미노산 1-21에 상응하는 적어도 21개의 아미노산, 또는 서열 번호 123의 아미노산 1-22에 상응하는 적어도 22개의 아미노산 및 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산을 포함할 수 있다. 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산은 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 아미노산을 포함할 수 있다. 예를 들어, 하나 이상의 아미노산은 2, 3 또는 4개의 아미노산일 수 있다. 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산은 A, C, G, L, S, M, W, 또는 A, C, G, L, S, M, W의 보존 치환, 또는 그의 조합일 수 있다. 한 예에서, 인플루엔자 HA 단백질 TM의 C 말단 단부로부터의 하나 이상의 아미노산은 AG 또는 AG의 보존 치환, AGL 또는 AGL의 보존 치환, MAGL 또는 MAGL의 보존

치환으로부터 선택될 수 있다.

[0182] 변형된 코로나바이러스 S 단백질은 또한 코로나바이러스 S 단백질 CT로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열을 포함하는 키메라 CT를 포함할 수 있다.

[0183] 코로나바이러스 S 단백질 CT로부터 유래된 N 말단 서열은 하나 이상의 아미노산을 포함할 수 있다. 코로나바이러스 S 단백질 CT의 N 말단 단부로부터의 하나 이상의 아미노산은 0, 1, 2, 3, 4 또는 5개의 아미노산을 포함할 수 있다. 예를 들어, 하나 이상의 아미노산은 1, 2 또는 3개의 아미노산일 수 있다. 코로나바이러스 S 단백질 CT의 N 말단 단부로부터의 하나 이상의 아미노산은 C 또는 M, 또는 C 또는 M의 보존 치환일 수 있다. 한 예에서, 코로나바이러스 S 단백질의 N 말단 단부로부터의 하나 이상의 아미노산은 C 또는 C의 보존 치환, CC 또는 CC의 보존 치환, 또는 CCM 또는 CCM의 보존 치환으로부터 선택될 수 있다.

[0184] 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열은 서열 번호 18의 아미노산 27-37에 상응하는 적어도 11개의 아미노산을 포함할 수 있다. 인플루엔자 HA 단백질 CT로부터 유래된 N 말단 서열은 서열 번호 18의 아미노산 26-37에 상응하는 적어도 12개의 아미노산, 서열 번호 18의 아미노산 25-37에 상응하는 적어도 13개의 아미노산, 서열 번호 18의 아미노산 24-37에 상응하는 적어도 14개의 아미노산, 서열 번호 18의 아미노산 23-37에 상응하는 적어도 15개의 아미노산, 또는 서열 번호 18의 아미노산 22-37에 상응하는 적어도 16개의 아미노산을 추가로 포함할 수 있다.

[0185] 또 다른 예에서, 인플루엔자 HA 단백질 CT로부터 유래된 C 말단 서열은 하기를 포함할 수 있다:

[0186] - 서열 번호 126, 127, 128, 129, 130 또는 131의 아미노산 27-37에 상응하는 적어도 11개의 아미노산;

[0187] - 서열 번호 126, 127, 128, 129, 130 또는 131의 아미노산 26-37에 상응하는 적어도 12개의 아미노산;

[0188] - 서열 번호 126, 127, 128, 129, 130 또는 131의 아미노산 25-37에 상응하는 적어도 13개의 아미노산;

[0189] - 서열 번호 126, 127, 128, 129, 130 또는 131의 아미노산 24-37에 상응하는 적어도 14개의 아미노산;

[0190] - 서열 번호 126, 127, 128, 129, 130 또는 131의 아미노산 23-37에 상응하는 적어도 15개의 아미노산; 또는

[0191] - 서열 번호 126, 127, 128, 129, 130 또는 131의 아미노산 22-37에 상응하는 적어도 16개의 아미노산.

[0192] 예를 들어, CT는 표 1B(HA 단백질 CT 서열)에 제시된 바와 같은 서열을 포함할 수 있다. 예를 들어, CT는 서열 번호 18, 126, 128, 129, 130, 131, 118, 120, 164 또는 166의 아미노산 22-37; 또는 서열 번호 19의 아미노산 25-40; 또는 서열 번호 37의 아미노산 24-39; 또는 서열 번호 38의 아미노산 25-36; 또는 서열 번호 39 또는 119의 아미노산 24-34; 또는 서열 번호 127의 아미노산 22-36; 또는 서열 번호 118 또는 164의 아미노산 22-37; 또는 서열 번호 123 또는 125의 아미노산 23-38; 또는 서열 번호 124의 아미노산 25-35; 또는 서열 번호 165의 아미노산 24-34; 또는 서열 번호 169의 아미노산 21-36 또는 서열 번호 170의 아미노산 23-33 또는 서열 번호 21-36의 아미노산을 포함할 수 있다.

[0193] 인플루엔자 CT 또는 CT의 일부는 개재 펩티드 서열로 S 단백질의 TM 또는 TM의 일부에 융합 또는 연결될 수 있다. 예를 들어, 개재 펩티드 서열은 인플루엔자 CT, S 단백질 TM 또는 그의 조합으로부터 유래될 수 있거나, 또는 개재 펩티드 서열은 인공 서열일 수 있다. 개재 펩티드 서열의 길이는 다양할 수 있다. 예를 들어, 개재 펩티드 서열의 길이는 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 아미노산 길이일 수 있고, 바람직하게, 개재 펩티드 서열의 길이는 2 내지 8개의 아미노산 길이이다. 한 예에서, 개재 펩티드 서열의 길이는 2개의 아미노산 길이이고, 예를 들어, 개재 펩티드 서열은 서열 LC를 포함할 수 있다. 또 다른 예에서, 개재 펩티드 서열의 길이는 4개의 아미노산 길이이고, 예를 들어, 개재 펩티드 서열은 서열 LCCM을 포함할 수 있다. 또 다른 예에서, 개재 펩티드 서열의 길이는 5개의 아미노산 길이일 수 있고, 예를 들어, 개재 펩티드 서열은 서열 LSLWM을 포함할 수 있다. 또 다른 예에서, 개재 펩티드 서열의 길이는 7개의 아미노산 길이일 수 있고, 예를 들어, 개재 펩티드 서열은 서열 AGLSLWM을 포함할 수 있다. 추가 예에서, 개재 펩티드 서열의 길이는 8개의 아미노산 길이일 수 있고, 예를 들어, 개재 펩티드 서열은 서열 MAGLSLWM을 포함할 수 있다.

[0194] 예를 들어, 변형된 S 단백질의 TMCT는 하기 서열, 또는 그와 90-100%, 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함할 수 있다:

-WYIWLGFIAGLIAIVMVTIM - (X)n - CSNGSXXCXICI (서열 번호 64),

-WYVWLGFIAGLIAIVMVTIL - (X)n - CSNGSXXCXICI (서열 번호 134), 또는

-WYIWLGFIAGLVALALCVFFIL -(X)n - CSNGSXXCXICI (서열 번호 135),

-WYVWLLICLAGVAMLVLLFFI -(X)n - CSNGSXXCXICI (서열 번호 172),

-WWVWLCISVVLIFVVSMLLL -(X)n - CSNGSXXCXICI (서열 번호 173)

[0195]

[0196] 여기서, $(X)_n$은 개재 펩티드 서열이고, 여기서, 개재 펩티드 서열의 길이는 0 내지 n개의 아미노산 잔기 길이일 수 있고, 여기서, n은 0-10개, 예를 들어, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 아미노산인 임의의 길이일 수 있고, 여기서, X는 임의의 아미노산을 포함할 수 있고, 예를 들어, X는 A, C, G, F, L, S, M, W, 또는 A, C, G, F, L, S, M, W의 보존 치환, 또는 그의 조합일 수 있다.

[0197] 비제한적인 개재 펩티드 서열 $(X)_n$은 하기를 포함할 수 있다:

[0198] - $(X)_0$(개재 펩티드 서열은 부재한다);

[0199] - 예를 들어, 서열 L 또는 C, 또는 L 또는 C의 보존 치환을 포함하는 $(X)_1$;

[0200] - 예를 들어, 서열 LC, 또는 LC의 보존 치환을 포함하는 $(X)_2$;

[0201] - 예를 들어, 서열 LCC, 또는 LCC의 보존 치환을 포함하는 $(X)_3$;

[0202] - 예를 들어, 서열 LCCM 또는 LCCM의 보존 치환, SLWM 또는 SLWM의 보존 치환, SFWM 또는 SFWM의 보존 치환을 포함하는 $(X)_4$;

[0203] - 예를 들어, 서열 LSLWM, 또는 LSLWM의 보존 치환, LSLWM 또는 LSLWM의 보존 치환, LSFWM 또는 LSFWM의 보존 치환을 포함하는 $(X)_5$;

[0204] - 예를 들어, 서열 GLSLWM, 또는 GLSLWM의 보존 치환을 포함하는 $(X)_6$;

[0205] - 예를 들어, 서열 AGLSLWM, 또는 AGLSLWM의 보존 치환을 포함하는 $(X)_7$; 또는

[0206] - 예를 들어, 서열 MAGLSLWM, 또는 MAGLSLWM의 보존 치환을 포함하는 $(X)_8$.

[0207] 도 5b는 식물에서 발현되었을 때, 코로나바이러스 S 단백질의 세포질 테일이 H5 A/Indonesia/5/05 HA의 코로나바이러스 S 단백질로 대체된 참조 변형된 S 단백질(wtTM/H5iCT)(서열 번호 18, 구성물 8671의 생성물)의 단백질 축적 대비, SARS-COV-2 막횡단(TM) 도메인과 H5 A/Indonesia/5/05 HA 사이토졸 테일(CT) 도메인 사이에 가변 마진(개재 펩티드 서열)이 있는 대안적인 버전의 C 말단 영역을 포함하는 변형된 S 단백질인 wtTM/H5iCT V1(서열 번호 19, 구성물 8980의 생성물), wtTM/H5iCT V2(서열 번호 37, 구성물 8981의 생성물), wtTM/H5iCT V3(서열 번호 38, 구성물 #8982의 생성물) 및 wtTM/H5iCT V4(서열 번호 39, 구성물 8983의 생성물)의 단백질 축적의 변화 배수를 보여주는 것이다. 시험된 변형된 S 단백질 모두 단백질 축적을 보였고, 여기서, 대안적인 버전과 wtTM/H5iCT 참조 대조군 사이에 통계적으로 유의적인 차이는 없었다.

[0208] 유사하게, wtTM/H5iCT V4 버전의 TMCT와 함께 SARS-CoV-1 S 단백질(도 16a), 또는 wtTM/H5iCT V4 버전의 TMCT와 함께 MERS S 단백질(도 19a)을 포함하는 변형된 S 단백질은 식물에서 발현되었을 때, 야생형 S 단백질(wtTMCT) 또는 TMCT가 H5 A/Indonesia/5/05 HA의 TMCT로 대체된 S 단백질(H5iTMCT)의 단백질 축적과 비교하여 증가된 단백질 축적을 보였다. 추가로, wtTM/H5iCT V4 버전의 TMCT를 포함하는 OC43 CoV S 단백질은 식물에서 발현되었을 때, 야생형 TMCT를 포함하는 OC43 CoV S 단백질(wtTMCT)의 단백질 축적과 비교하여 증가된 단백질 축적을 보였다(도 23a).

[0209] 따라서, 변형된 S 단백질은 TM 및 CT 도메인(TM/CT)을 포함할 수 있고, 여기서, CT 또는 CT의 일부는 개재 펩티

드 서열을 통해 TM 또는 TM의 일부의 C 말단 단부에 융합되고, 여기서, 개재 펩티드 서열은 서열 X_n을 포함한다.

[0210] 추가로, 변형된 S 단백질은 서열 번호 18, 19, 37, 38, 39, 64, 126, 127, 128, 129, 130, 131, 118, 119, 120, 123, 124, 125, 134, 135, 164, 165, 166, 169, 170, 171, 172 또는 173의 서열과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함하는 TM 및 CT 도메인(TM/CT)을 포함할 수 있다.

[0211] 변형된 S 단백질은 서열 번호 18의 아미노산 22-37, 서열 번호 19의 아미노산 21-40, 서열 번호 37의 아미노산 21-39, 서열 번호 38의 아미노산 25-36 또는 서열 번호 39의 아미노산 24-34, 서열 번호 126의 아미노산 22-37, 서열 번호 127의 아미노산 22-36, 서열 번호 128의 아미노산 22-37, 서열 번호 129의 아미노산 22-37, 서열 번호 130의 아미노산 22-37, 또는 서열 번호 131의 아미노산 22-37과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함하는 CT 또는 CT의 일부를 포함할 수 있다.

[0212] 변형된 S 단백질은 서열 번호 18의 아미노산 1-20, 서열 번호 19의 아미노산 1-20, 서열 번호 37의 아미노산 1-20, 서열 번호 38의 아미노산 1-24, 서열 번호 39의 아미노산 1-23, 서열 번호 118의 아미노산 1-21, 서열 번호 119의 아미노산 1-23, 서열 번호 123의 아미노산 1-22, 서열 번호 124의 아미노산 1-24, 서열 번호 164의 아미노산 1-21, 서열 번호 165의 아미노산 1-23, 서열 번호 169의 아미노산 1-20, 또는 서열 번호 170의 아미노산 1-22와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함하는 TM 또는 TM의 일부를 포함할 수 있다. 추가로, 본원에 기술된 바와 같은 변형된 S 단백질은 서열 번호 132 또는 133의 서열과 80% 내지 100%의 동일성을 포함하는 TM 또는 TM의 일부를 포함할 수 있다.

[0213] 추가로, 변형된 S 단백질은 서열 번호 5, 59, 60, 61, 62, 95, 96, 97, 108, 109 또는 110의 서열과 70% 내지 100%의 서열 동일성, 또는 서열 유사성을 포함할 수 있고, 예를 들어, 변형된 S 단백질은 서열 번호 5, 59, 60, 61, 62, 95, 96, 97, 108, 109 또는 110의 서열과 약 70, 75, 80, 85, 87, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함할 수 있다.

[0214] 바이러스 구조 단백질, 예컨대, 예를 들어, 코로나바이러스 S 단백질의 세포질 테일 도메인(CT) 또는 CT의 일부는 하기 기술되는 바와 같은 인플루엔자 헤마글루티닌(HA)으로부터의 CT 또는 CT의 일부로 대체될 수 있고, 생성된 단백질은 변형된 바이러스 구조 단백질로 지칭된다. 따라서, 천연 CT 또는 천연 CT의 일부가 HA로부터의 CT 또는 CT의 일부로 대체된 코로나바이러스 S 단백질은 변형된 코로나바이러스 S 단백질 또는 변형된 S 단백질로 지칭될 수 있다. 상기에 추가로 기술된 바와 같이, HA CT 또는 HA CT의 일부는 코로나바이러스 TM 도메인의 N 말단 단부에 직접 융합될 수 있거나, 또는 개재 펩티드 서열을 통해 코로나바이러스 TM 또는 TM의 일부의 N 말단 단부에 융합될 수 있다. 따라서, HA CT 또는 HA CT의 일부는 개재 펩티드 서열을 통해 S 단백질 TM 또는 S 단백질 TM의 일부의 C 말단 단부에 융합될 수 있다.

[0215] 인플루엔자 "헤마글루티닌" 또는 "HA"는 일반적으로 신호 펩티드, HA1 도메인, 및 C 말단에 막 관통 앵커 부위 및 작은 세포질 테일을 포함하는 HA2 도메인을 포함하는 동종삼량체 막 타입 I 당단백질이다(예를 들어, 도 1c 및 도 2 참조). 다양한 인플루엔자 균주로부터의 HA의 아미노산 서열이 당업계에 널리 공지되어 있다. 추가로, HA를 코딩하는 아미노산 서열 및 뉴클레오티드 서열은 널리 공지되어 있고, 이용가능하다 - 예를 들어, 생물방어 공중 보건 베이스(BioDefence Public Health base)(Influenza Virus; URL: biohealthbase.org 참조) 또는 미국 국립 생물공학 정보 센터(National Center for Biotechnology Information)(URL: ncbi.nlm.nih.gov 참조)(상기 둘 모두 본원에서 참조로 포함된다)을 참조한다. 상이한 인플루엔자 균주로부터의 HA 세포질 테일 도메인의 예시적인 아미노산 서열이 도 2에 제시되어 있다.

[0216] 상이한 참조 및 그룹이 HA의 CT에 상이한 길이를 할당하지만, CT의 N 말단 서열은 상이한 인플루엔자 서브타입 및 균주로부터의 HA 사이에서 보존되고, 적어도 5개의 잔기가 13개의 HA 서브타입 중 적어도 10개에 대해 서열 동일성을 가진 것으로 나타났다(문헌 [Simpson and Lamb 1992, Journal of Virology, 790-803]). 도 2는 예시적인 인플루엔자 균주로부터의 아미노산 서열 및 HA 단백질의 N 말단부의 보존 서열의 정렬을 보여주는 것이다. 인플루엔자 세포질 테일(CT) 도메인의 컨센서스 서열은

[0217] XXWMCSNGSXXCXICI(서열 번호 15)이다(도 2, 의 서열 번호 14의 C 말단 단부 또한 참조).

[0218] HA 세포질 테일 도메인 컨센서스 서열에 상응하는 CT 서열은 상기 논의된 바와 같이 직접 또는 개재 펩티드 서

열(링커 서열)을 통해 코로나바이러스 S 단백질의 TM의 C 말단 단부에 융합될 수 있다.

[0219] 추가로, 천연 인플루엔자 HA TM/CT 경계로부터 N 말단 또는 C 말단에 위치하는 아미노산 잔기는 또한 변형된 코로나바이러스 S 단백질의 TM 또는 TM의 일부에 직접 또는 개재 펩티드 서열을 통해 융합되는 CT 서열에 포함될 수 있다.

[0220] 따라서, CT 또는 CT의 일부의 서열은 예를 들어, 서열 번호 14의 아미노산 30-40 중 어느 하나에 상응하는 아미노산 잔기에서 시작될 수 있다. 따라서, CT 서열의 N 말단 단부는 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14의 아미노산 30-40 중 어느 하나에 상응하는 아미노산일 수 있다. 한 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 30에 상응하는 아미노산 잔기에서 시작될 수 있다. 또 다른 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 31에 상응하는 아미노산 잔기에서 시작될 수 있다. 추가 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 32에 상응하는 아미노산 잔기에서 시작될 수 있다. 또 다른 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 33에 상응하는 아미노산 잔기에서 시작될 수 있다. 추가 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 34에 상응하는 아미노산 잔기에서 시작될 수 있다. 또 다른 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 35에 상응하는 아미노산 잔기에서 시작될 수 있다. 추가 예에서, CT 서열은 서열 번호 6-13 또는 14 중의 아미노산 36에 상응하는 아미노산 잔기에서 시작될 수 있다. 또 다른 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 37에 상응하는 아미노산 잔기에서 시작될 수 있다. 추가 예에서, CT 서열은 서열 번호 6-13 또는 14 중의 아미노산 38에 상응하는 아미노산 잔기에서 시작될 수 있다. 또 다른 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 39에 상응하는 아미노산 잔기에서 시작될 수 있다. 추가 예에서, CT 서열은 서열 번호 6, 7, 8, 9, 10, 11, 12, 13 또는 14 중의 아미노산 40에 상응하는 아미노산 잔기에서 시작될 수 있다.

[0221] 변형된 S 단백질의 세포질 테일(CT) 또는 CT의 일부는 어느 하나의 인플루엔자 타입, 서브타입 또는 균주의 헤마글루티닌(HA)의 CT 또는 CT의 일부로부터 유래된 것일 수 있다. 예를 들어, CT는 인플루엔자 타입 A 또는 인플루엔자 타입 B로부터의 HA로부터 유래된 것일 수 있다. 예를 들어, CT는 인플루엔자 서브타입 H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, 또는 H16의 HA로부터 유래된 것일 수 있다. CT는 예를 들어, 서브타입 H1, H2, H3, H5, H6, H7 또는 H9의 HA로부터 유래된 것일 수 있다. 추가로, CT 또는 CT의 일부는 인플루엔자 타입 B의 HA로부터 유래된 것일 수 있다. 타입 B 인플루엔자는 계통 B/Yamagata 또는 B/Victoria로부터의 것일 수 있다.

[0222] 예를 들어, 변형된 S 단백질의 CT 또는 CT의 일부는 헤마글루티닌(HA) 인플루엔자 H1, H3, H5, H6, H7, H9 또는 B 균주의 CT로부터 유래된 것일 수 있다. HA CT의 유래 기점이 될 수 있는 인플루엔자 균주의 비제한적인 예는 인플루엔자 H1 California/7/2009, H2 A/Singapore/1/1957, H3 A/Minnesota/41/2019, H5 A/Indonesia/5/05, H6 A/Teal/Hong Kong/W312/97, H7 A/Guangdong/17SF003/2016, H9 A/Hong Kong/1073/99 또는 B/Washington/02/2019이다. HA CT의 아미노산 서열의 비제한적인 예는 도 2에 제시되어 있다.

[0223] 도 4a에 제시된 바와 같이, SARS-CoV-2 S 단백질의 천연 세포질 테일(CT)이 인플루엔자 HA H1 California/7/2009(H1 Calif), H3 A/Minnesota/41/2019(H3 Minn), H6 A/Teal/Hong Kong/W312/97(H6 HK), H7 A/Guangdong/17SF003/2016(H7 Guan), H9 A/Hong Kong/1073/99(H9 HK) 또는 B/Washington/02/2019(B Wash)로부터의 CT로 대체되었을 때, H5 A/Indonesia/5/05로부터의 CT를 포함하는 SARS-CoV-2 S(H5 Indo)와 비교하여 상기 변형된 SARS-CoV-2 S 단백질에 대하여 단백질 축적 변화 배수가 유사한 것으로 관찰되었다. 웨스턴 블롯 분석을 통해 상기 관찰결과를 확인하였다(도 4b 및 4c 참조).

[0224] SARS-CoV-1 S 단백질의 천연 세포질 테일(CT), MERS S 단백질의 천연 CT, 또는 OC43 CoV S 단백질의 천연 CT가 인플루엔자 HA H1 California/7/2009로부터의 CT(H1cCT)로 대체되었을 때, 유사한 결과를 수득하였다(도 16a, 19a, 및 23a 참조).

[0225] 따라서, 세포질 테일 도메인(CT) 또는 CT의 일부는 서열 번호 15의 서열과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 30-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 31-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 32-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 33-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 34-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 35-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 36-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 37-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아

미노산 38-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 39-50과, 또는 서열 번호 6, 7, 8, 9, 10, 12, 13, 14의 아미노산 40-50과, 또는 서열 번호 11의 아미노산 31-49와, 또는 서열 번호 11의 아미노산 32-49와, 또는 서열 번호 11의 아미노산 33-49와, 또는 서열 번호 11의 아미노산 34-49와, 또는 서열 번호 11의 아미노산 35-49와, 또는 서열 번호 11의 아미노산 36-49와, 또는 서열 번호 11의 아미노산 37-49와, 또는 서열 번호 11의 아미노산 38-49와, 또는 서열 번호 11의 아미노산 39-49와, 또는 서열 번호 3의 아미노산 548-568과, 또는 서열 번호 3의 아미노산 549-568과, 또는 서열 번호 3의 아미노산 550-568과, 또는 서열 번호 3의 아미노산 551-568과, 또는 서열 번호 3의 아미노산 552-568과, 또는 서열 번호 3의 아미노산 553-568과, 또는 서열 번호 3의 아미노산 554-568과, 또는 서열 번호 3의 아미노산 555-568과, 또는 서열 번호 3의 아미노산 556-568과, 또는 서열 번호 3의 아미노산 557-568과, 또는 서열 번호 3의 아미노산 558-568과 약 70, 75, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가질 수 있다.

[0226] 추가로, 변형된 S 단백질은 서열 번호 5, 53, 54, 55, 56, 57 또는 58의 서열과 70% 내지 100%의 서열 동일성, 또는 서열 유사성을 포함할 수 있고, 예를 들어, 변형된 S 단백질은 서열 번호 5, 53, 54, 55, 56, 57 또는 58의 서열과, 또는 서열 번호 53의 아미노산 25-1259, 서열 번호 54의 아미노산 25-1259, 서열 번호 55의 아미노산 25-1259, 서열 번호 56의 아미노산 25-1259, 서열 번호 57의 아미노산 25-1259, 또는 서열 번호 58의 아미노산 25-1259와 약 70, 75, 80, 85, 87, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함할 수 있다.

[0227] 추가 실시양태에서, 변형된 S 단백질 엑토도메인 및/또는 막횡단 도메인은 SARS-CoV-2 S 단백질 이외의 다른 코로나바이러스 S 단백질로부터, 예를 들어, SARS-CoV-1 S 단백질, MERS-CoV S 단백질, OC43-CoV S 단백질, 229E-CoV S 단백질 등으로부터 수득될 수 있다.

[0228] 도 16a에 제시된 바와 같이, 미정제 식물 추출물로부터의 야생형 TMCT(wt TMCT)를 포함하는 S 단백질, 또는 인플루엔자 H5 HA의 TMCT(H5i TMCT)를 포함하는 변형된 S 단백질의 단백질 축적과 비교하여, 인플루엔자 H5 HA로부터의 CT(H5iCT), 인플루엔자 H1 HA로부터의 CT(H1cCT) 또는 SARS-CoV-1 TM과 H5 HA CT 사이에 가변 마진이 있는, 인플루엔자 H5 HA로부터의 CT("H5iCT(V4))를 포함하는 변형된 SARS-CoV-1 S 단백질에 대해 단백질 축적이 더 높은 것으로 관찰되었다. 변형된 S 단백질은 고분자량의 구조체로 어셈블리되었고(도 16b-16d 참조), 이는 VLP인 것으로 확인되었다(도 17a-17c 참조). 변형된 TMCT 및/또는 CT를 포함하는 SARS-CoV-1 S 단백질과 함께 동일한 겔 상에 제시되었을 때, 천연 TMCT를 포함하는 SARS-CoV-1 단백질에 대한 단백질 축적량은 웨스턴 블롯 분석의 검출 한계 미만이었지만(도 16a 참조), 천연 TMCT를 포함하는 SARS-CoV-1 단백질만 단독으로 겔 상에 존재할 때에는 신호는 웨스턴 블롯 분석에 의해 검출될 수 있었고(도 16b 참조), VLP는 전자 현미경법에 의해 관찰되었다(도 17b 참조).

[0229] 변형된 MERS-CoV S 단백질에 대해 유사한 결과를 수득하였다(도 19a 참조). 야생형 TMCT(wt TMCT)를 포함하는 S 단백질, 또는 인플루엔자 H5 HA의 TMCT(H5i TMCT)를 포함하는 변형된 S 단백질의 단백질 축적과 비교하여, 인플루엔자 H5 HA로부터의 CT(H5iCT), 인플루엔자 H1 HA로부터의 CT(H1cCT) 또는 MERS-CoV TM과 H5 HA CT 사이에 가변 마진이 있는, 인플루엔자 H5 HA로부터의 CT("H5iCT(V4))를 포함하는 변형된 MERS-CoV S 단백질에 대해 단백질 축적이 더 높은 것으로 관찰되었다(대략 175 kDa의 단백질 밴드 참조). 인플루엔자 H1 HA CT(H1cCT)를 포함하는 변형된 MERS-CoV의 경우, 축적이 가장 높은 것으로 관찰되었다. 대략 100 kDa에서 관찰된 더 작은 밴드는 S 단백질의 단백질분해 절단 단편일 가능성이 가장 높다. 이론에 얽매이지 않고, 천연 CT의 인플루엔자 HA CT로의 대체가 MERS S 단백질을 안정화시키고, S 단백질의 절단을 감소시키는 것으로 간주된다.

[0230] 도 23a에 추가로 제시된 바와 같아, 천연 OC43 CoV S 단백질 TMCT를 포함하는 OC43 CoV S 단백질을 발현하는 식물에서 낮은 단백질 수율이 관찰되었다. 그러나, 천연 OC43 CoV S 단백질 TMCT가 인플루엔자 H5 HA로부터의 TMCT(H5iTMCT), 인플루엔자 H5 HA로부터의 CT(H5iCT), 인플루엔자 H1 HA로부터의 CT(H1cCT) 또는 OC43-CoV H5 HA CT 사이에 가변 마진이 있는, 인플루엔자 H5 HA로부터의 CT("H5iCT(V4))로 대체되었을 때에는 천연 TMCT를 포함하는 OC43 CoV S 단백질과 비교하여 단백질 축적이 더 높은 것으로 관찰되었다(약 150 kDa의 밴드 참조. 겔에 제시된 더 큰 밴드는 단백질 삼량체인 것으로 간주된다). 변형된 229E-CoV S 단백질에서도 유사한 결과가 관찰되었다(데이터는 제시되지 않음).

[0231] 추가로, 인플루엔자 H5 HA로부터의 TMCT(H5iTMCT), 인플루엔자 H5 HA로부터의 CT(H5iCT), 또는 인플루엔자 H1 HA로부터의 CT를 포함하는 MERS-CoV S 단백질, OC43-CoV S 단백질, 및 229E-CoV S 단백질은 도 19b-19f, 23b-23e, 및 25a-25e에 제시된 바와 같이 VLP를 형성하는 것으로 관찰되었다.

[0232] 따라서, 본 개시내용은 바이러스 구조 단백질의 엑토도메인 및 막횡단 도메인(TM) 또는 TM의 일부가 코로나바이러스로부터 유래되고, 사이토졸 테일(CT) 또는 CT의 일부가 인플루엔자 단백질로부터 유래된 것인, "변형된 바이러스 구조 단백질," "바이러스 구조 융합 단백질" 또는 "키메라 바이러스 구조 단백질"을 제공한다. 예를 들어, 엑토도메인 및 막횡단 도메인은 코로나바이러스 스파이크(S) 단백질로부터 유래된 것일 수 있고, 사이토졸 테일(CT) 또는 CT의 일부는 인플루엔자 HA 단백질로부터 유래된 것일 수 있다. 변형된 S 단백질은 연속적으로 i) (S1 서브유닛, 및 S2 서브유닛의 FP, HR1 및 HR2 도메인을 포함하는) 코로나바이러스 S 단백질로부터 유래된 엑토도메인, ii) 코로나바이러스 막횡단 도메인(TM) 또는 코로나바이러스 TM의 일부, 및 iii) 인플루엔자 HA 세포질 테일 도메인(CT) 또는 HA CT의 일부를 포함할 수 있다. 따라서, 변형된 S 단백질에서, CT 또는 CT의 일부는 TM 및 엑토도메인과 이종성이다. 유사하게, 변형된 S 단백질의 TM(및 엑토도메인)은 CT와 이종성이다. 엑토도메인 및 막횡단 도메인(TM)은 동일한 코로나바이러스로부터 유래된 것일 수 있거나(즉, 엑토도메인 및 TM은 서로 동종성일 수 있거나), 또는 엑토도메인은 제1 코로나바이러스로부터 유래된 것일 수 있고, TM은 제2 코로나바이러스로부터 유래된 것일 수 있다(즉, 엑토도메인 및 TM은 서로 이종성이다).

[0233] "융합 단백질"로도 지칭되는, "키메라 단백질," 또는 "키메라 폴리펩티드"란, 단일 폴리펩티드로서 융합된, 2개 이상의 공급원으로부터의 아미노산 서열, 예를 들어, 제한하는 것은 아니지만, 제1 바이러스 구조 단백질로부터 유래된, 예를 들어, 코로나바이러스 S 단백질로부터 유래된 엑토도메인 및 막횡단 도메인, 및 제2 바이러스 구조 단백질로부터 유래된 세포질 테일(CT), 예를 들어, 인플루엔자 HA로부터의 CT를 포함하는 단백질 또는 폴리펩티드를 의미한다.

[0234] 변형된 코로나바이러스 S 단백질은 막횡단 및 사이토졸 테일 도메인(TMCT)을 포함할 수 있고, 여기서, TMCT는 키메라 TMCT이다. 키메라 TMCT는 막횡단 도메인(TM) 또는 TM의 일부가 코로나바이러스 S 단백질로부터 유래된 것인 TM, 및 사이토졸 테일(CT) 또는 CT의 일부가 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래된 것인 CT를 포함할 수 있다. 추가로, 키메라 TMCT는 천연 코로나바이러스 S 단백질 TM, 키메라 코로나바이러스 S 단백질/인플루엔자 HA TM, 천연 인플루엔자 HA CT, 키메라 인플루엔자 HA/코로나바이러스 S 단백질 CT 또는 그의 조합을 포함할 수 있다. 예를 들어, 변형된 코로나바이러스 S 단백질은 천연 인플루엔자 HA CT 및 키메라 TM을 포함하는 키메라 TMCT를 포함할 수 있고, 여기서, 키메라 TM은 코로나바이러스 S 단백질의 TM으로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질의 TM으로부터 유래된 C 말단 서열을 포함한다. 또 다른 예에서, 변형된 코로나바이러스 S 단백질은 천연 코로나바이러스 S 단백질 TM 및 키메라 CT를 포함하는 키메라 TMCT를 포함할 수 있고, 여기서, 키메라 CT는 코로나바이러스 S 단백질로부터 유래된 N 말단 서열, 및 인플루엔자 HA 단백질로부터 유래된 C 말단 서열을 포함한다. 추가 예에서, 변형된 코로나바이러스 S 단백질은, 코로나바이러스 S 단백질의 TM으로부터 유래된 N 말단 서열 및 인플루엔자 HA 단백질의 TM으로부터 유래된 C 말단 서열을 포함하는 키메라 TM, 및 코로나바이러스 S 단백질로부터 유래된 N 말단 서열 및 인플루엔자 HA 단백질로부터 유래된 C 말단 서열을 포함하는 키메라 CT를 포함하는 키메라 TMCT를 포함할 수 있다.

[0235] 본 개시내용에서 변형된 S 단백질 또는 변형된 코로나바이러스 스파이크(S) 단백질을 언급할 때, 막횡단 도메인(TM) 또는 S 단백질 TM의 일부 및 사이토졸 테일(CT) 또는 CT의 일부를 포함하는 변형된 코로나바이러스 스파이크(S) 단백질로서, 여기서, CT는 인플루엔자 헤마글루티닌(HA) 단백질로부터 유래되고, 여기서, TM은 CT와 이종성인 것을 의미한다.

[0236] 변형된 S 단백질은 서열 번호 5, 21, 30, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 95, 96, 97, 108, 109, 110, 144, 145, 146, 155, 156 또는 157의 서열과 70% 내지 100%의 서열 동일성, 또는 서열 유사성을 포함할 수 있고, 예를 들어, 변형된 S 단백질은 서열 번호 5, 21, 30, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 95, 96, 97, 108, 109, 110, 144, 145, 146, 155, 156 또는 157의 서열과, 또는 서열 번호 47의 아미노산 25-1259, 서열 번호 48의 아미노산 25-1259, 서열 번호 49의 아미노산 25-1259, 서열 번호 50의 아미노산 25-1259, 서열 번호 51의 아미노산 25-1259, 서열 번호 52의 아미노산 25-1259, 서열 번호 53의 아미노산 25-1259, 서열 번호 54의 아미노산 25-1259, 서열 번호 55의 아미노산 25-1259, 서열 번호 56의 아미노산 25-1259, 서열 번호 57의 아미노산 25-1259, 서열 번호 58의 아미노산 25-1259, 서열 번호 59의 아미노산 25-1262, 서열 번호 60의 아미노산 25-1261, 서열 번호 61의 아미노산 25-1258, 서열 번호 62의 아미노산 25-1256, 서열 번호 95의 아미노산 25-1243, 서열 번호 96의 아미노산 25-1240, 서열 번호 97의 아미노산 25-1243, 서열 번호 108의 아미노산 25-1341, 서열 번호 109의 아미노산 25-1338, 서열 번호 110의 아미노산 25-1341, 서열 번호 144의 아미노산 25-1351, 서열 번호 145의 아미노산 25-1348, 서열 번호 146의 아미노산 25-1351, 서열 번호 155의 아미노산 25-1159, 서열 번호 156의 아미노산 25-1156, 또는 서열 번호 157의 아미노산 25-1159와 약 70, 75, 80, 85, 87, 90, 91, 92, 93, 94, 95, 96, 97,

98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 서열을 포함할 수 있다.

[0237] 변형된 S 단백질은 추가로 변형된 S 단백질 전구체(이는 또한 전구체 S 단백질로도 지칭)로 생산되거나, 또는 합성될 수 있고, 여기서, S 단백질 전구체는 변형된 S 단백질 및 신호 펩티드를 포함하고, 여기서, 신호 펩티드는 코로나바이러스에 대해 천연인 것이거나(즉, 엑토도메인과 동종성이거나), 또는 신호 펩티드는 엑토도메인에 대해 비천연이거나, 또는 이종성일 수 있다. 비제한적인 예에서, 천연 신호 펩티드는 단백질 디술피드 이소머라제(PDI)로부터의 신호 펩티드로 대체될 수 있다.

[0238] 변형된 S 단백질 전구체는 엑토도메인에 대해 비천연이거나, 또는 이종성인 신호 펩티드를 포함할 수 있다. 비천연 신호 펩티드는 전체 천연 신호 펩티드를 대체할 수 있거나, 또는 코로나바이러스 S 단백질의 천연 신호 펩티드의 일부를 대체할 수 있다. 추가로, 비천연 또는 이종성 신호 펩티드는 변형된 S 단백질의 N 말단에 직접 융합될 수 있거나, 또는 비천연 또는 이종성 신호 펩티드는 개재 펩티드 서열을 사용하여 변형된 S 단백질의 N 말단에 융합될 수 있다.

[0239] 신호 펩티드(또한 신호 서열, 표적화 신호, 국재화 신호, 국재화 서열, 전이 펩티드, 리더 서열 또는 리더 펩티드로도 지칭)는 분비 경로로 향하게 되는 새로 합성된 대부분의 단백질의 N 말단에 존재하는 짧은 펩티드이다. 신호 펩티드는 단백질을 소포체 및 골지체를 포함하는 내막 시스템으로 표적화하는 역할을 하며, 여기서, ER 내강 내에 위치하는 신호 펩티다제에 의해 공동 번역에 의해 제거되고, 성숙 단백질이 생성된다. 표적화 서열을 확인하기 위한 실험적 방법은 시간이 많이 소요되고, 힘들기 때문에, 표적화 신호를 예측하는 다양한 계산 접근법이 개발되었으며, 당업계에 널리 공지되어 있다. 신호 펩티드는 일반적으로 서열 유사성이 낮지만, 몇 가지 특징적인 특성을 공유한다. 신호 서열 및 그의 절단 부위를 예측하기 위해, 이러한 특징적인 특성을 고려하는 많은 예측 방법, 예를 들어, 시그널P(SignalP)(문헌 [Bendtsen et al., J Mol Biol. 2004 Jul 16;340(4):783-95.]; [Petersen et al., Nature Methods volume 8, pages785-786(2011)]), 시그널-CF(Signal-CF)(문헌 [Chou and Shen, Biochem Biophys Res Commun. 2007 Jun 8;357(3):633-40]), 및 시그널-BLAST(Signal-BLAST)(문헌 [Frank and Sippl, Bioinformatics, 2008 Oct 1;24(19):2172-6])(상기 문헌들은 본원에서 참조로 포함된다)가 개발되었다.

[0240] 예측 프로그램을 사용하여, SARS-CoV-2 S 단백질에 대한 신호 펩티드 절단 부위는 서열 번호 1의 서열에 상응하는 서열의 위치 15와 16 사이인 것으로 예측된다. 그러나, SARS-CoV-2 S 단백질에 대한 신호 펩티드 절단 부위는 서열 번호 1의 서열에 상응하는 서열의 다른 연속 위치 사이인 것으로 예측되거나, 또는 존재할 수 있다. 예를 들어, SARS-CoV-2 S 단백질에 대한 신호 펩티드 절단 부위 또한 서열 번호 1의 서열에 상응하는 서열의 위치 13과 14 사이인 것으로 예측되거나, 또는 존재할 수 있다.

[0241] 천연 SARS-CoV-2 S 단백질(천연 신호 펩티드 서열 포함)의 N 말단 영역이 하기 제시되어 있다:

[0242] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNS(서열 번호 63).

[0243] 예측된 신호 펩티드 서열(SP)은 밑줄체로 표시되어 있다. 회색으로 음영 처리된 서열은 표 2에 제시된 서열에 상응한다. 성숙한 SARS-CoV-2 S 단백질의 제1 아미노산 잔기는 발린(V)일 수 있고, 여기서, 그의 위치는 전구체 S 단백질(천연 신호 펩티드를 포함하는 천연 SARS-CoV-2 S 단백질)의 V16에 상응하는 1(+1)로 지정될 수 있다. 성숙 SARS-CoV-2 S 단백질의 제1 아미노산 잔기는 표 2에 제시된 바와 같이 서열 번호 1 또는 서열 번호 63의 다른 잔기에 존재할 수 있다. 예를 들어, 성숙 SARS-CoV-2 S 단백질의 제1 아미노산 잔기는 글루타민(Q)일 수 있고, 여기서, 그의 위치는 14(-2)로 지정될 수 있다.

표 2: 신호 펩티드(SP) 절단 부위를 둘러싼 SARS-CoV-2 S 단백질 서열의 일부. 잔기의 넘버링은 천연 신호 펩티드를 포함하는 N 말단으로부터(SP 포함하의 #) 또는 위치 V1의 예측된 절단 부위로부터의(SP 부재하의 #) 것(이는 신호 펩티드를 포함하는 서열(전구체 S 단백질)에서의 V16과 같다)이다.

	S	**S**	**Q**	**C**	**V**	**N**	**L**	**T**	**T**	**R**	**T**
SP 포함하의 #	12	13	14	15	16	17	18	19	20	21	22
SP 부재하의 #*	-4	-3	-2	-1	1	2	3	4	5	6	7

*성숙 단백질

[0244]

[0245] 발현된 단백질 또는 폴리펩티드의 아포플라스트로의 국재화를 지시하기 위한 신호 펩티드 또는 펩티드 서열은

천연(단백질에 대해) 신호 또는 리더 서열, 또는 이종성 신호 서열, 예를 들어, 제한하는 것은 아니지만, 쌀 아밀라제 신호 펩티드(문헌 [McCormick 1999, Proc Natl Acad Sci USA 96:703-708]) 또는 단백질 디술피드 이소머라제 신호 펩티드(PDI)를 포함하나, 이에 제한되지 않는다. 따라서, 본원에 기술된 바, 변형된 S 단백질은 변형된 S 단백질 및 이종성 아미노산 신호 펩티드 서열을 포함하는 전구체 단백질로서 생산될 수 있다. 예를 들어, 변형된 S 단백질 전구체는 단백질 디술피드 이소머라제로부터의 신호 펩티드(PDI SP; 수탁 번호 Z11499의 뉴클레오티드 32-103)를 포함할 수 있다.

[0246] 따라서, 본 개시내용은 변형된 S 단백질 및 천연, 또는 비천연 신호 펩티드를 포함하는 변형된 S 단백질 전구체, 및 상기 단백질을 코딩하는 핵산을 제공한다.

[0247] 변형된 바이러스 구조 단백질은 변형된 S 단백질일 수 있고, 여기서, 변형된 S 단백질은 단량체 또는 단일 쇄 변형된 S 단백질이다. 단량체 또는 단일 쇄 변형된 S 단백질은 S1 도메인(서브유닛) 및 S2 도메인(서브유닛)을 포함할 수 있고, 여기서, S2 도메인(서브유닛)은 S 단백질의 천연 CT를 인플루엔자 HA 단백질의 CT로 대체하도록 변형되고, 여기서, 변형된 S 단백질은 단일 인접한 폴리펩티드 쇄이다. 단량체 또는 단일 쇄 변형된 S 단백질은 삼량체화되어, 삼량체 변형된 S 단백질로 지칭되는 삼량체를 형성할 수 있다. 삼량체는 일반적으로 비공유 결합으로 결합된 3개의 단백질로 형성된 거대분자 복합체이다.

[0248] S 단백질은 보존된 활성화 절단 부위에서 2개의 폴리펩티드 쇄, S1 서브유닛 및 S2 서브유닛으로 절단되며, 이는 동종삼량체 내에서 S1/S2 프로토머로서 회합된 상태로 남아 있다. 이론에 의해 얽매이기를 바라지 않으면서, S 단백질의 서브유닛으로의 절단은 바이러스 감염성에 중요할 수 있지만, 단백질의 삼량체화에 필수적이지는 않을 수 있다.

[0249] 변형된 S 단백질은 하나 이상의 치환, 대체 또는 돌연변이를 추가로 포함할 수 있다. 예를 들어, 변형된 S 단백질은 발현, 수율, 안정성을 증가시키거나, 또는 적합한 발현 시스템에서 변형된 S 단백질의 발현, 수율 및 안정성을 증가시키기 위해 엑토도메인에 하나 이상의 치환, 대체 또는 돌연변이를 포함할 수 있다.

[0250] 예를 들어, 변형된 S 단백질은 S1/S2 및/또는 S2' 프로테아제 절단 부위에 대한 치환 또는 돌연변이를 포함하여 이들 부위에서 프로테아제 절단을 방지할 수 있다. 따라서, 숙주 또는 숙주 세포에서 생산될 때, 변형된 S 단백질은 별개의 S1 및 S2 서브유닛 또는 폴리펩티드 쇄로 절단되지 않는다.

[0251] 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질은 변형된 바이러스 구조 단백질의 삼량체로 추가로 어셈블리될 수 있다. 따라서, 본원에 기술된 바와 같은 변형된 S 단백질을 포함하는 코로나바이러스 단백질 삼량체를 추가로 제공한다. 삼량체는 단일 쇄 변형된 S 단백질을 포함할 수 있고, 여기서, 단일 쇄 변형된 S 단백질은 S1 서브유닛 및 S2 서브유닛을 포함하고, 여기서, S2 서브유닛의 CT는 인플루엔자 헤마글루티닌(HA)의 CT로 대체된다.

[0252] 삼량체는 추가로 융합 전 입체형태로 안정화될 수 있다. 따라서, 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질은 융합 전 입체형태에서 융합 후 입체형태로의 S 단백질의 입체형태 변화를 억제시키기 위해 하나 이상의 치환, 대체 또는 돌연변이를 추가로 포함할 수 있고, 이로써, S 단백질 또는 S 단백질 삼량체를 융합 전 입체형태로 안정화시킬 수 있다.

[0253] "아미노산 치환" 또는 "치환"이란, 단백질의 아미노산 서열 중의 아미노산을 상이한 아미노산으로 대체하는 것을 의미한다. 용어 아미노산, 아미노산 잔기 또는 잔기는 본 개시내용에서 상호교환적으로 사용된다. 하나 이상의 아미노산은 단백질의 아미노산 서열의 전체 길이를 변경하지 않으면서, 이 위치에서 원래 또는 야생형 아미노산과 상이한 하나 이상의 아미노산으로 대체되거나 치환될 수 있다.

[0254] 예를 들어, 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질은 프롤린 치환, 디술피드 결합 및 염 브릿지의 형성을 허용하는 치환, 및/또는 공동 충전 치환에 의해 안정화될 수 있다.

[0255] (Hsieh et al.)(문헌 [Science 2020, 369 p.1501-1505](상기 문헌은 본원에서 참조로 포함된다)은 포유동물 세포에서 다양한 SARS-CoV-2 스파이크 단백질 변이체를 디자인하고, 발현시켰다. HexaPro로 지칭되는 6개의 프롤린 치환을 포함하는 S 단백질 변이체는 오직 이중 프롤린 치환만을 가진 변이체와 비교하여 S 단백질보다 9.8X 더 높은 발현을 보였고, Tm은 ~5℃ 증가하였고, 포유동물 세포주에서 삼량체 융합 전 입체형태를 유지하였다. HexaPro 변이체는 (Hsieh et al.)에 의하면 가장 우수한 변이체인 것으로 간주된다.

[0256] 본 개시내용에서, 서열 번호 2의 위치 802, 927, 971 및 972에 상응하는 4개의 프롤린 치환("4P") 및 위치 923에의 추가의 단일 아미노산 치환의 조합이 이루어진 경우, 최고 수율이 관찰되었다. 추가로, 위치 802, 877,

884, 927, 971 및 972에 상응하는 6개의 프롤린 치환("6P") 및 위치 923에의 추가의 단일 아미노산 치환의 조합이 이루어진 경우에도 또한 수율이 더 높은 것으로 관찰되었다.

[0257] 본원에서 제공되는 바와 같이, 변형된 S 단백질은 숙주 또는 숙주 세포, 예를 들어, 식물 또는 식물 세포에서 변형된 단백질의 안정성, 수율 또는 안정성 및 수율을 증가시키기 위해 하나 이상의 치환, 대체 또는 돌연변이를 추가로 포함할 수 있다.

[0258] 본원에 기술된 바와 같은 변형된 S 단백질은 그의 아미노산 서열 중 하기 기술되는 바와 같은 참조 서열 내의 아미노산에 상응하는 임의의 하나 이상의 아미노산에서 하나 이상의 돌연변이, 변형, 또는 치환을 포함할 수 있다.

[0259] "아미노산에 상응한다," "아미노산에 상응하는," 또는 서열에 상응하는" 등은 아미노산(또는 뉴클레오티드)이 하기 기술되는 바와 같은 참조 코로나바이러스 서열과의 서열 정렬에서 아미노산(또는 뉴클레오티드)에 상응한다는 것을 의미한다. 코로나바이러스 서열 중의 상응하는 아미노산 위치는 코로나바이러스 S 단백질의 공지된 서열에의 정렬에 의해 결정될 수 있다. 비교를 위한 서열 정렬 방법은 당업계에 널리 공지되어 있고, 이는 하기에 추가로 기술된다. 상응하는 아미노산의 예는 표 3에 제시되어 있다.

표 3: 코로나바이러스 S 단백질 중 상응하는 아미노산/잔기 위치의 위치.
(참조 서열은 명시되어 있다).

	S 단백질 중 아미노산 위치(명시된 서열 기준)								
SARS-CoV-2[1] (서열 번호 2)	667	668	670	802	877	884	927	971	972
SARS-CoV-2[2] (서열 번호 1)	682	683	685	817	892	899	942	986	987
SARS-CoV-1[1] (서열 번호 114)	651	652	654	786	861	868	911	955	956
SARS-CoV-1[2] (서열 번호 112)	664	665	667	799	874	881	924	968	969
MERS-CoV[1]	730	731	733	872	949	956	999	1043	1044
(서열 번호 115)									
MERS-CoV[2] (서열 번호 113)	747	748	750	889	966	973	1016	1060	1061
OC43-CoV[1] (서열 번호 160)	748	749	751	898	N/A	976	1019	1063	1064
OC43-CoV[2] (서열 번호 158)	762	763	765	912	N/A	990	1033	1077	1078
229E-CoV[1] (서열 번호 161)	551	552	554	675	N/A	754	811	855	856
229E-CoV[2] (서열 번호 159)	567	568	570	691	N/A	770	827	871	872

1 넘버링은 신호 펩티드(SP)를 배제한다.
2 넘버링은 신호 펩티드(SP)를 포함한다.

[0260]

[0261] 예를 들어, 변형된 S 단백질은 예를 들어, WO 2018/081318(상기 문헌은 본원에서 참조로 포함된다)에 기술된 바와 같이, HR1 도메인과 중심 나선 도메인 사이의 바운더리에, 또는 그 인근에, 융합 전 입체형태에서 S 엑토도메인 삼량체를 안정화시키는 하나 이상의(예를 들어, 2개의 연속) 프롤린 치환을 가질 수 있다. 추가로, 하나 이상의 치환은 S1 서브유닛과 S2 서브유닛 사이의 절단 부위에서의 프로세싱 또는 절단을 제한할 수 있고/거나, 방지할 수 있다.

[0262] 변형된 S 단백질은 표 3에 제시된 바와 같은 위치에 하나 이상의 치환을 포함할 수 있다. 예를 들어, 변형된 S 단백질은 서열 번호 2의 참조 서열(SARS-CoV-2) 중 위치 667, 668, 670, 802, 877, 884, 923, 927, 971, 972, 또는 그의 조합에 상응하는 위치에 하나 이상의 치환을 포함할 수 있다. SARS-CoV-1, MERS-CoV, OC43-CoV 및 229E-CoV의 S 단백질 중의 상응하는 위치는 표 3에 제시되어 있다. 다른 코로나바이러스로부터의 S 단백질 중의 상응하는 아미노산 위치는 당업계에 공지된 방법에 의해 결정될 수 있다.

[0263] *GSAS - 2P(971 및 972)*

[0264] 예를 들어, 변형된 S 단백질은 서열 번호 2의 아미노산 서열의 아미노산 at 위치 667, 668, 670, 971 또는 972에 상응하는 하나 이상의 아미노산에 하나 이상의 치환을 가질 수 있다.

[0265] 한 측면에서, 변형된 S 단백질은 위치 667, 668, 670 또는 그의 조합(서열 번호 2에 따른 넘버링)에 상응하는 치환, 변형, 또는 돌연변이를 포함할 수 있다. 예를 들어, 위치 667에 상응하는 아미노산은 글리신(G) 또는 글

리신(G)의 보존 치환으로 치환될 수 있고, 위치 668에 상응하는 아미노산은 세린(S) 또는 세린(S)의 보존 치환으로 치환될 수 있고, 위치 670에 상응하는 아미노산은 세린(S) 또는 세린(S)의 보존 치환으로 치환될 수 있다.

[0266] 변형된 S 단백질은 위치 971, 972에 상응 또는 위치 971 및 972에(서열 번호 2에 따른 넘버링) 치환, 변형, 또는 돌연변이를 추가로 포함할 수 있다. 예를 들어, 위치 971 및/또는 972에 상응하는 아미노산은 프롤린(P) 또는 프롤린(P)의 보존 치환으로 치환될 수 있다.

[0267] 변형된 S 단백질은 서열 번호 2의 위치 667, 668, 670, 971, 972에 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환을 포함할 수 있다. 하나 이상의 치환을 포함하는 변형된 S 단백질은 융합 전 입체형태로 안정화될 수 있다. 추가로, 변형된 S 단백질은 융합 전 입체형태로 안정화된 삼량체를 형성할 수 있다.

[0268] 예를 들어, 변형된 S 단백질은 하기 치환을 포함할 수 있다(서열 번호 2에 따른 넘버링): R667G, R668S, R670S(본원에서 "GSAS"로 지칭). 변형된 S 단백질은 또한 하기 치환을 포함할 수 있다(서열 번호 2에 따른 넘버링): K971P 및 V972P(본원에서 "2P"로 지칭). 추가로 변형된 S 단백질은 하기 치환을 가질 수 있다(서열 번호 2에 따른 넘버링): R667G, R668S, R670S, K971P 및 V972P(본원에서 "GSAS-2P"로 지칭).

[0269] 예를 들어, 변형된 S 단백질은 서열 번호 47 서열의 아미노산 서열과, 또는 서열 번호 47의 아미노산 25-1259와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 가질 수 있고, 여기서, 아미노산 서열은 위치 667에 글리신(G) 또는 글리신(G)의 보존 치환, 위치 668에 세린(S) 또는 세린(S)의 보존 치환, 위치 670에 세린(S) 또는 세린(S)의 보존 치환, 위치 971 및 972에 프롤린(P) 또는 프롤린(P)의 보존 치환을 갖고, 여기서, 변형된 S 단백질은 발현되었을 때, VLP를 형성한다.

[0270] 또 다른 예에서, 변형된 S 단백질은 서열 번호 114의 아미노산 서열의 위치 654, 955 또는 956에, 또는 서열 번호 115의 아미노산 서열의 위치 730, 733, 1043 또는 1044에의 아미노산에 상응하는 하나 이상의 아미노산에 하나 이상의 치환을 가질 수 있다.

[0271] 예를 들어, 변형된 S 단백질은 하기 치환을 포함할 수 있다: R654A(서열 번호 114에 따른 넘버링) 또는 R730A 및/또는 R733G(서열 번호 115에 따른 넘버링). 변형된 S 단백질은 또한 하기 치환을 가질 수 있다: K955P 및/또는 V956P(서열 번호 114에 따른 넘버링) 또는 V1043P 및/또는 L1044P(서열 번호 115에 따른 넘버링). 추가로 변형된 S 단백질은 하기 치환을 가질 수 있다: R654A, K955P 및 V956P(서열 번호 114에 따른 넘버링) 또는 R730A, R733G, V1043P, L1044P(서열 번호 115에 따른 넘버링).

[0272] *GSAS-4P(802, 927, 971 및 972)*

[0273] 변형된 S 단백질은 위치 667, 668, 및 670에의 아미노산에 상응하는 아미노산에 치환, 및 위치 802, 927, 971 및 972에 상응하는 하나 이상의 잔기에 추가로 하나 이상의 치환을 추가로 가질 수 있다(서열 번호 2에 따른 넘버링). 예를 들어, 위치 802, 927, 971 및 972에 상응하는 아미노산은 프롤린(P) 또는 프롤린(P)의 보존 치환으로 치환될 수 있다.

[0274] 도 11a에 제시된 바와 같이, "GSAS" 변형 및 하기 변형: F802P, A927P, K971P, V972P를 가진 변형된 S 단백질("GSAS-4P"로 지칭, 구성물 8953으로부터 발현)은 "GSAS-2P" S 단백질(구성물 8671로부터 발현)의 수율과 비교하여 변형된 S 단백질 수율에 있어 2.47배 증가된 증가를 보였다.

[0275] 따라서, 변형된 S 단백질은 서열 번호 2의 위치 667, 668, 670, 802, 927, 971 및 972에의 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환을 포함할 수 있다.

[0276] 예를 들어, 변형된 S 단백질은 서열 번호 48의 아미노산 서열과, 또는 서열 번호 48의 아미노산 25-1259와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 가질 수 있고, 여기서, 아미노산 서열은 위치 667에 글리신(G) 또는 글리신(G)의 보존 치환, 위치 668에 세린(S) 또는 세린(S)의 보존 치환, 위치 670에 세린(S) 또는 세린(S)의 보존 치환, 위치 802, 927, 971 및 972에 프롤린(P) 또는 프롤린(P)의 보존 치환을 갖고, 여기서, 변형된 S 단백질은 발현되었을 때, VLP를 형성한다.

[0277] 또 다른 예에서, 변형된 S 단백질은 서열 번호 114의 아미노산 서열의 위치 654, 786, 911, 955 또는 956에서의, 또는 서열 번호 115의 아미노산 서열의 위치 730, 733, 872, 999, 1043 또는 1044에의 아미노산에

상응하는 하나 이상의 아미노산에 하나 이상의 치환을 가질 수 있다.

[0278] 예를 들어, 변형된 S 단백질은 하기 치환을 포함할 수 있다: R654A(서열 번호 114에 따른 넘버링) 또는 R730A 및/또는 R733G(서열 번호 115에 따른 넘버링). 변형된 S 단백질은 또한 하기 치환을 가질 수 있다: F786P, S911P, K955P 및/또는 V956P(서열 번호 114에 따른 넘버링) 또는 A872P, N999P, V1043P 및/또는 L1044P(서열 번호 115에 따른 넘버링). 추가로 변형된 S 단백질은 하기 치환을 가질 수 있다: R654A, F786P, S911P, K955P 및 V956P(서열 번호 114에 따른 넘버링) 또는 R730A, R733G, A872P, N999P, V1043P, L1044P(서열 번호 115에 따른 넘버링).

[0279] *GSAS-6P(802, 877, 884, 927, 971 및 972)*

[0280] 변형된 S 단백질은 위치 667, 668, 및 670에의 아미노산에 상응하는 아미노산에 치환, 및 위치 802, 877, 884, 927, 971, 및 972에 상응하는 하나 이상의 잔기에 추가로 하나 이상의 치환을 추가로 가질 수 있다(서열 번호 2에 따른 넘버링). 예를 들어, 위치 802, 877, 884, 927, 971, 및 972에 상응하는 아미노산은 프롤린(P) 또는 프롤린(P)의 보존 치환으로 치환될 수 있다(서열 번호 2에 따른 넘버링).

[0281] 도 11a에 제시된 바와 같이, "GSAS" 변형 및 하기 변형: F802P, A877P, A884P, A927P, K971P, V972P를 가진 변형된 S 단백질("GSAS-6P"로 지칭, 구성물 8940으로부터 발현)은 "GSAS-2P" S 단백질(구성물 8671로부터 발현)의 수율과 비교하여 S 단백질 수율에 있어 2.11배 증가된 증가를 보였다.

[0282] 따라서, 변형된 S 단백질은 서열 번호 2의 위치 667, 668, 670, 802, 877, 884, 927, 971 및 972에의 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환을 포함할 수 있다.

[0283] 예를 들어, 변형된 S 단백질은 서열 번호 49의 아미노산 서열과, 또는 서열 번호 48의 아미노산 25-1259과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 가질 수 있고, 여기서, 아미노산 서열은 위치 667에 글리신(G) 또는 글리신(G)의 보존 치환, 위치 668에 세린(S) 또는 세린(S)의 보존 치환, 위치 670에 세린(S) 또는 세린(S)의 보존 치환, 위치 802, 802, 877, 884, 927, 971 및 972에 프롤린(P) 또는 프롤린(P)의 보존 치환을 갖고, 여기서, 변형된 S 단백질은 발현되었을 때, VLP를 형성한다.

[0284] 또 다른 예에서, 변형된 S 단백질은 서열 번호 114의 아미노산 서열의 위치 654, 786, 861, 868, 911, 955 또는 956에의, 또는 서열 번호 115의 아미노산 서열의 위치 730, 733, 872, 949, 956, 999, 1043 또는 1044에의 아미노산에 상응하는 하나 이상의 아미노산에 하나 이상의 치환을 가질 수 있다.

[0285] 예를 들어, 변형된 S 단백질은 은 하기 치환을 포함할 수 있다: R654A(서열 번호 114에 따른 넘버링) 또는 R730A 및/또는 R733G(서열 번호 115에 따른 넘버링). 변형된 S 단백질은 또한 하기 치환을 가질 수 있다: F786P, A861P, A868P, S911P, K955P 및/또는 V956P(서열 번호 114에 따른 넘버링) 또는 A872P, S949P, A956P, N999P, V1043P 및/또는 L1044P(서열 번호 115에 따른 넘버링). 추가로 변형된 S 단백질은 하기 치환을 가질 수 있다: R654A, F786P, A861P, A868P, S911P, K955P 및 V956P(서열 번호 114에 따른 넘버링) 또는 R730A, R733G, A872P, S949P, A956P, N999P, V1043P 및 L1044P(서열 번호 115에 따른 넘버링).

[0286] *위치 923에서의 치환*

[0287] 본원에 기술된 바와 같은 변형된 S 단백질은 위치 923(서열 번호 2에 따른 넘버링)에 상응하는 치환, 변형 또는 돌연변이를 추가로 포함할 수 있다. 예를 들어, 위치 923에 상응하는 아미노산은 페닐알라닌(F) 또는 페닐알라닌(F)의 보존된 치환으로 치환될 수 있다.

[0288] 도 11b에 제시된 바와 같이, "GSAS-2P" 변형 및 L923F 치환을 가진 변형된 S 단백질(구성물 8933으로부터 발현)은 L923F 치환을 포함하지 않는 "GSAS-2P" S 단백질(구성물 8671로부터 발현)의 수율과 비교하여 S 단백질 수율에 있어 1.36배의 증가를 보였다. "GSAS-4P" 변형 및 L923F 치환을 가진 변형된 S 단백질(구성물 8960으로부터 발현)은 L923F 치환을 포함하지 않는 "GSAS-2P" S 단백질(구성물 8671로부터 발현)의 수율과 비교하여 변형된 S 단백질 수율에 있어 2.88배의 증가를 보였다. "GSAS-6P" 변형 및 L923F 치환을 가진 변형된 S 단백질(구성물 8947로부터 발현)은 L923F 치환을 포함하지 않는 "GSAS-2P" S 단백질(구성물 8671로부터 발현)의 수율과 비교하여 수율에 있어 2.47배의 증가를 보였다.

[0289] 따라서, 변형된 S 단백질은 서열 번호 2의 위치 667, 668, 670, 927, 971, 972, 802, 877, 884, 923 또는 그의 조합에의 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환

을 포함할 수 있다. 예를 들어, 변형된 S 단백질은 서열 번호 2의 위치 667, 668, 670, 971, 972, 923, 또는 그의 조합(GSAS-2P-923), 서열 번호 2의 667, 668, 670, 927, 971, 972, 802 923, 또는 그의 조합(GSAS-4P-923), 또는 서열 번호 2의 서열 번호 2의 667, 668, 670, 927, 971, 972, 802, 877, 884, 923 또는 그의 조합(GSAS-6P-923)에의 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환을 포함할 수 있다.

[0290] 예를 들어, 변형된 S 단백질은

[0291] - 서열 번호 2의 위치 667, 668, 670, 971, 972 및 923(GSAS-2P-923),

[0292] - 서열 번호 2의 667, 668, 670, 927, 971, 972, 802 및 923(GSAS-4P-923), 또는

[0293] - 서열 번호 2의 667, 668, 670, 927, 971, 972, 802, 877, 884 및 923(GSAS-6P-923)에의 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환을 포함할 수 있다.

[0294] 예를 들어, 변형된 S 단백질은 서열 번호 50의 아미노산 서열, 또는 서열 번호 50의 아미노산 25-1259와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 가질 수 있고(여기서, 아미노산 서열은 위치 667에 글리신(G) 글리신(G)의 보존 치환, 위치 668 및 670에 세린(S) 또는 세린(S)의 보존 치환, 위치 971 및 972에 프롤린(P) 또는 프롤린(P)의 보존 치환, 및 위치 923에 페닐알라닌(F) 또는 페닐알라닌(F)의 보존 치환을 가진다); 서열 번호 51 또는 서열 번호 51의 아미노산 25-1259와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 가질 수 있고(여기서, 아미노산 서열은 위치 667에 글리신(G) 또는 글리신(G)의 보존 치환, 위치 668 및 670에 세린(S) 또는 세린(S)의 보존 치환, 위치 927, 971, 972 및 802에 프롤린(P) 또는 프롤린(P)의 보존 치환, 및 위치 923에 페닐알라닌(F) 또는 페닐알라닌(F)의 보존 치환을 가진다); 또는 서열 번호 52 또는 서열 번호 52의 아미노산 25-1259와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 가질 수 있고(여기서, 아미노산 서열은 위치 667에 글리신(G) 또는 글리신(G)의 보존 치환, 위치 668 및 670에 세린(S) 또는 세린(S)의 보존 치환, 위치 927, 971, 972, 802, 877 및 884에 프롤린(P) 또는 프롤린(P)의 보존 치환, 및 위치 923에 페닐알라닌(F) 또는 페닐알라닌(F)의 보존 치환을 가진다), 여기서, 변형된 S 단백질은 발현되었을 때, VLP를 형성한다.

[0295] 따라서, 하기:

[0296] 1. 키메라 막횡단 및 세포질 테일 도메인(TMCT);

[0297] 2. 상응하는 야생형 코로나바이러스 S 단백질과 비교하여 위치 667, 668, 및/또는 670(서열 번호 2에 따른 넘버링)에의 아미노산에 상응하는 하나 이상의 치환;

[0298] 3. 상응하는 야생형 코로나바이러스 S 단백질과 비교하여 위치 802, 877, 884, 927, 971, 및/또는 972(서열 번호 2에 따른 넘버링)에의 아미노산에 상응하는 하나 이상의 치환;

[0299] 4. 상응하는 야생형 코로나바이러스 S 단백질과 비교하여 위치 923(서열 번호 2에 따른 넘버링)에 상응하는 치환;

[0300] 5. 또는 1-4하에 기술된 바와 같은 변형 및/또는 치환의 조합을 포함할 수 있는 변형된 코로나바이러스 S 단백질을 제공한다.

[0301] 본원에서 사용되는 바, 용어 "보존 치환" 또는 "보존적 치환" 및 그의 문법적 파생어는 참조 아미노산과 상이하지만(치환), 기술한 치환 또는 기술한 잔기와 동일한 부류의 아미노산에 있는 아미노산을 지칭한다(즉, 비극성 잔기를 대체하는 비극성 잔기, 방향족 잔기를 대체하는 방향족 잔기, 극성-비하전된 잔기를 대체하는 극성-비하전된 잔기, 하전된 잔기를 대체하는 하전된 잔기). 보존적 치환에 대한 추가 정보는 예를 들어, 문헌 [Sahin-Toth et al. (Protein ScL, 3:240-247, 1994)], [Hochuli et al (Bio/Technology, 6:1321-1325, 1988)], [Henikoff S, and Henikoff JG (Proc. Natl. Acad. Sci. USA 89: 10915-10919, 1992)]]에서, 및 널리 사용되는 유전학 및 분자 생물학 교재에서 살펴볼 수 있다.

[0302] 변형된 바이러스 구조 단백질은 추가로 글리코실화될 수 있다. 코로나바이러스 S 단백질, 코로나바이러스 M 단백질 및 코로나바이러스 E 단백질은 글리코실화되고, N 연결된 글리코실화 및 O 연결된 글리코실화, 둘 모두가 이루어진다.

[0303] 변형된 바이러스 구조 단백질은 변형된 바이러스 구조 단백질이 발현되는 숙주 또는 숙주 세포에 고유한 글리코실화 패턴을 포함할 수 있다. 예를 들어, 식물 또는 식물 세포에서 발현될 때, 변형된 바이러스 구조 단백질은 식물 특이적 N-글리칸을 포함할 수 있다. 따라서, 식물 특이적 N-글리칸을 가진 변형된 바이러스 구조 단백질 또한 제공한다.

[0304] 본원에 기술된 바, 변형된 바이러스 구조 단백질의 사이토졸 테일 도메인(CT)은 인플루엔자 헤마글루티닌(HA)으로부터의 CT로 대체될 수 있다. 상기 기술된 바와 같은 바이러스 구조 단백질의 엑토도메인 및 막횡단 도메인(TM)은 인플루엔자 HA 사이토졸 테일 도메인(CT)에 융합되고, 이로써, CT는 바이러스 구조 단백질, 예컨대, S 단백질의 엑토도메인 및 막횡단 도메인과 관련하여 이종성이다. 변형된 S 단백질은 바이러스 유사 입자(VLP)로 자기 어셈블리될 수 있다.

[0305] 따라서, 본 기술내용은 추가로 바이러스 유사 입자(VLP)에 관한 것이다. 더욱 구체적으로, 본 기술내용은 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질을 포함하는 VLP, 및 숙주 또는 숙주 세포에서 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질을 이용하여 VLP를 생산하는 방법에 관한 것이다. VLP는 본원에 기술된 바와 같은, 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질을 포함한다.

[0306] 도 6c, 17a, 17b, 및 17c에 제시된 바와 같이, 천연 또는 야생형 CT가 인플루엔자 HA 단백질로부터의 CT에 의해 대체된, 변형된 S 단백질(변형된 SARS-CoV-2 또는 변형된 SARS-CoV-1 S 단백질)로 예시되는 변형된 바이러스 구조 단백질은 식물에서 발현될 때, VLP로 자기 어셈블리된다. VLP는 동일한 식물 발현 시스템에서 천연 TM/CT 서열을 포함하는 S 단백질(도 6a 및 17a 참조), 또는 H5 인플루엔자 TM/CT 서열을 포함하는 변형된 S 단백질(도 6b 및 17b 참조)로 생산된 VLP와 유사하다.

[0307] 추가로, 도 6d, 6e, 6f 및 6g에 제시된 바와 같이, TM과 인플루엔자 CT 도메인 사이에 가변 마진 또는 바운더리(개재 펩티드 서열)가 있는 변형된 S 단백질 또한 식물에서 발현될 때, VLP로 자기 어셈블리된다.

[0308] 추가로, 도 6h, 6i, 6j, 6k, 6l 및 6m에 제시된 바와 같이, 천연 또는 야생형 CT가 각각 H1, H3, H6, H7, H9 및 B 인플루엔자로부터의 인플루엔자 HA 단백질로부터의 CT에 의해 대체된 변형된 S 단백질 또한 식물에서 발현될 때, VLP로 자기 어셈블리된다.

[0309] 추가로, 도 19b-19f, 23b-23e, 및 25a-25e에 제시된 바와 같이, 변형된 S 단백질이 인플루엔자 H5 HA로부터의 TMCT(H5iTMCT), 인플루엔자 H5 HA로부터의 CT(H5iCT), 또는 인플루엔자 H1 HA로부터의 CT를 가진, MERS-CoV, OC43-CoV, 및 229E-CoV로부터 유래된 변형된 S 단백질 또한 VLP를 형성하였다.

[0310] 용어 바이러스-유사 입자"(VLP) 또는 "바이러스 유사 입자들" 또는 "VLP들"은 일반적으로 감염에서 생성된 비리온과 형태학적 및 항원적으로 유사하지만, 복제에 충분한 유전 정보가 없으며, 따라서, 비감염성인 바이러스 유사 구조를 지칭한다. VLP는 자기 어셈블리되고, 하나 이상의 구조 단백질, 예컨대, 예를 들어, 변형된 바이러스 구조 단백질, 예를 들어, 제한하는 것은 아니지만, 변형된 S 단백질을 포함하는 구조이다. 따라서, VLP는 변형된 S 단백질을 포함할 수 있다. VLP는 변형된 S 단백질로 구성된 바이러스 구조 단백질을 추가로 포함할 수 있다. 따라서, 일부 실시양태에서, VLP에는 코로나바이러스 M 단백질 및/또는 코로나바이러스 E 단백질이 결여되어 있거나, 또는 없을 수 있다. 따라서, 일부 실시양태에서, 본원에 기술된 바와 같은 변형된 바이러스 구조 단백질로부터 생성된 VLP는 코로나바이러스 M 단백질, 코로나바이러스 E 단백질 또는 코로나바이러스 M 단백질 및 코로나바이러스 E 단백질을 포함하지 않는다. 추가로, 일부 실시양태에서, VLP는 코로나비리다에 또는 인플루엔자 바이러스와 이종성인 바이러스로부터의 구조 또는 비구조 단백질을 포함하지 않고, 예를 들어, VLP는 코로나비리다에로부터 유래된 것인 아닌 바이러스로부터의 구조 또는 비구조 단백질을 포함하지 않는다.

[0311] 또 다른 실시양태에서, VLP는 코로나바이러스 E 단백질, 코로나바이러스 M 단백질 및 변형된 코로나바이러스 S 단백질을 포함할 수 있다. 또 다른 실시양태에서, VLP는 코로나바이러스 E 단백질 및 변형된 코로나바이러스 S 단백질을 포함할 수 있다. 또 다른 실시양태에서, VLP는 코로나바이러스 M 단백질 및 변형된 코로나바이러스 S 단백질을 포함할 수 있다. 추가로, VLP는 코로나바이러스 E 단백질, 변형된 코로나바이러스 M 단백질 및 변형된 코로나바이러스 S 단백질을 포함할 수 있다. VLP는 변형된 코로나바이러스 E 단백질, 변형된 코로나바이러스 M 단백질 및 변형된 코로나바이러스 S 단백질을 추가로 포함할 수 있다. 또 다른 실시양태에서, VLP는 변형된 코로나바이러스 E 단백질 및 변형된 코로나바이러스 S 단백질을 포함할 수 있다. 또 다른 실시양태에서, VLP는 변형된 코로나바이러스 M 단백질 및 변형된 코로나바이러스 S 단백질을 포함할 수 있다.

[0312] VLP는 식물 및 식물 세포를 비롯한, 적합한 숙주 또는 식물 및 식물 세포에서 생성될 수 있다. 적합한 조건 하에서의 숙주 또는 숙주 세포로부터의 추출 후, 단리 및 적합한 조건 하에서의 추가 정제시, VLP는 무손상 구조

로 회수될 수 있다

[0313] VLP는 임의의 적합한 방법, 예를 들어, 화학적 또는 생화학적 추출을 사용하여 정제 또는 추출될 수 있다. VLP는 건조, 열, pH, 계면활성제 및 세제에 상대적으로 민감하다. 따라서, 수율을 최대화하고, 세포 단백질로 VLP 분획의 오염을 최소화하고, 단백질 또는 VLP, 및 필요할 경우, 연관된 지질 외피 또는 막의 무결성을 유지하는 방법, 세포벽을 느슨하게 하여 단백질, 또는 VLP를 방출하는 방법을 사용하는 것이 유용할 수 있다. 예를 들어, SDS 또는 트리톤(Triton)™ X-100과 같은 세제 또는 계면활성제의 사용을 최소화하거나, 제거하면, VLP 추출 수율을 개선시키는 데 유익할 수 있다. 이어서, VLP를 예를 들어, 전자 현미경법(도 4b 참조)에 의해, 또는 크기 배제 크로마토그래피에 의해 구조 및 크기에 대해 평가할 수 있다.

[0314] 외피 바이러스, 예컨대, 코로나바이러스의 경우, 지질층 또는 막이 바이러스에 의해 유지되는 것이 유리할 수 있다. 지질의 조성은 시스템에 따라 다를 수 있으며(예컨대, 식물 생산 외피 바이러스는 외피에 식물 지질 또는 피토스테롤을 포함할 수 있다), 개선된 면역 반응에 기여할 수 있다.

[0315] 따라서, 숙주 또는 숙주 세포에서 생산된 VLP는 숙주 또는 숙주 세포의 원형질막으로부터의 지질을 포함할 수 있다. 예를 들어, 식물에서 생산된 VLP는 식물 기원의 지질("식물 지질")을 함유할 수 있고, 곤충 세포에서 생산된 VLP는 곤충 세포의 원형질막으로부터의 지질(일반적으로 "곤충 지질"로 지칭)을 포함할 수 있고, 포유동물 세포에서 생산된 VLP는 포유동물 세포의 원형질막으로부터의 지질(일반적으로 "포유동물 지질"로 지칭)을 포함할 수 있다.

[0316] 식물 지질 또는 식물 유래 지질은 지질 이중층 형태일 수 있고, VLP를 둘러싸는 외피를 추가로 포함할 수 있다. 식물 유래 지질은 인지질, 트리, 디, 및 모노글리세리드 뿐만 아니라, 지용성 스테롤 또는 스테롤을 포함하는 대사산물을 비롯한, VLP가 생산된 식물의 원형질막의 지질 성분을 포함할 수 있다. 예로는 포스파티딜콜린(PC), 포스파티딜에탄올아민(PE), 포스파티딜이노시톨, 포스파티딜세린, 글리코스핑고지질, 피토스테롤 또는 그의 조합을 포함한다. 피토스테롤의 예로는 캄페스테롤, 스티그마스테롤, 에르고스테롤, 브라시카스테롤, 델타-7-스티그마스테롤, 델타-7-아베나스테롤, 다우노스테롤, 시토스테롤, 24-메틸콜레스테롤, 콜레스테롤 또는 베타-시토스테롤을 포함한다. 당업자가 이해할 수 있는 바와 같이, 세포 원형질막의 지질 조성은 세포의 수득 기점이 되는 세포 또는 유기체 또는 종의 배양 또는 성장 조건에 따라 달라질 수 있다. 일반적으로, 베타-시토스테롤은 가장 풍부한 피토스테롤이다.

[0317] 이론에 의해 얽매이기를 바라지 않으면서, 식물 유래 지질을 포함하는 식물 생산 VLP는 다른 제조 시스템에서 제조된 VLP보다 더 강한 면역 반응을 유도할 수 있고, 상기 식물 생산 VLP에 의해 유도된 면역 반응은 생 백신 또는 약독화된 전체 바이러스 백신에 의해 유도된 면역 반응과 비교하여 더 강력할 수 있다.

[0318] 추가로, 식물 지질의 존재의 잠재적 애주번트 효과 이외에도, 항원 제시 세포에 의한 당단백질 항원의 포획을 용이하게 하는 식물 N-글리칸의 능력은 식물에서 VLP의 생산에 유리할 수 있다.

[0319] 식물내에서 생산된 VLP는 식물 특이적 N-글리칸을 포함하는 변형된 바이러스 구조 단백질을 포함할 수 있다. 따라서, 본 개시내용은 또한 식물 특이적 N-글리칸을 가진 변형된 바이러스 구조 단백질을 포함하는 VLP를 제공한다. 추가로, 식물 지질 및 식물 특이적 N-글리칸을 가진 변형된 바이러스 구조 단백질을 포함하는 VLP를 제공한다.

[0320] 숙주 또는 숙주 세포에서 변형된 구조 단백질을 포함하는 바이러스 유사 입자(VLP)를 생산하는 방법 또한 제공한다. 추가로, 숙주 또는 숙주 세포에서 변형된 구조 단백질을 포함하는 바이러스 유사 입자(VLP)의 생산 수율을 증가시키는 방법 또한 제공한다. 본 방법은 변형된 구조 단백질을 코딩하는 서열을 포함하는 핵산을 숙주 또는 숙주 세포로 도입하는 단계 및 핵산의 발현을 허용하는 조건하에서 숙주 또는 숙주 세포를 인큐베이션시켜 VLP를 생성하는 단계를 포함한다. 변형된 바이러스 구조 단백질은 비변형된 바이러스 구조 단백질을 발현하는 숙주 또는 숙주 세포와 비교하여 더 높은 수율로 생산될 수 있다

[0321] 예를 들어, 도 3a에 제시된 바와 같이, 바이러스 구조 단백질의 세포질 테일(CT)이 인플루엔자 HA의 CT로 대체되어 변형된 바이러스 구조 단백질, 예를 들어, 변형된 S 단백질을 생산할 때, 식물에서 발현된 VLP의 수율은 증가될 수 있다. 도 11a 및 11b에 추가로 제시된 바와 같이, 변형된 S 단백질이 서열 번호 2의 위치 667, 668, 670, 802, 923, 927, 971 및/또는 972의 아미노산에 상응하는 아미노산의 하나 이상의 치환을 포함하거나, 또는 그로 구성된 하나 이상의 치환을 추가로 포함할 때, 식물에서 발현되었을 때, 변형된 S 단백질을 포함하는 VLP의 수율은 추가로 증가될 수 있다.

[0322] 변형된 바이러스 구조 단백질(예컨대, 변형된 S 단백질)의 수율, 또는 숙주 또는 숙주 세포, 예를 들어, 식물

또는 식물 세포에서 생산된 변형된 바이러스 구조 단백질을 포함하는 VLP의 수율은 상응하는 비변형된 바이러스 구조 단백질의 수율, 또는 상응하는 비변형된 바이러스 구조 단백질을 포함하는 VLP의 수율과 비교하여 1.1-10배, 또는 그 사이의 임의의 양만큼 증가될 수 있다. 예를 들어, 변형된 바이러스 구조 단백질(예컨대, 변형된 S 단백질)의 수율, 또는 숙주 또는 숙주 세포에서의 (변형된 바이러스 구조 단백질을 포함하는) VLP의 수율은 동일한 조건하에서 숙주 또는 숙주 세포에서 생산되었을 때, 상응하는 비변형된 바이러스 구조 단백질의 수율, 또는 상응하는 비변형된 바이러스 구조 단백질을 포함하는 VLP의 수율과 비교하여 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.6, 5.7, 5.8, 5.9, 6, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10배 또는 그 사이의 임의의 양만큼 증가될 수 있다.

[0323] 본원에 기술된 변형된 바이러스 구조 단백질은 서열 번호 1, 2, 5, 21, 30, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 95, 96, 97, 108, 109, 110, 112, 113, 114, 115, 144, 145, 146, 155, 156, 157, 158, 159, 160, 또는 161의 아미노산 서열과, 서열 번호 47의 아미노산 25-1259, 서열 번호 48의 아미노산 25-1259, 서열 번호 49의 아미노산 25-1259, 서열 번호 50의 아미노산 25-1259, 서열 번호 51의 아미노산 25-1259, 서열 번호 52의 아미노산 25-1259, 서열 번호 53의 아미노산 25-1259, 서열 번호 54의 아미노산 25-1259, 서열 번호 55의 아미노산 25-1259, 서열 번호 56의 아미노산 25-1259, 서열 번호 57의 아미노산 25-1259, 서열 번호 58의 아미노산 25-1259, 서열 번호 59의 아미노산 25-1262, 서열 번호 60의 아미노산 25-1261, 서열 번호 61의 아미노산 25-1258, 서열 번호 62의 아미노산 25-1256, 서열 번호 95의 아미노산 25-1243, 서열 번호 96의 아미노산 25-1240, 서열 번호 97의 아미노산 25-1243, 서열 번호 108의 아미노산 25-1341, 서열 번호 109의 아미노산 25-1338, 서열 번호 110의 아미노산 25-1341, 서열 번호 144의 아미노산 25-1351, 서열 번호 145의 아미노산 25-1348, 서열 번호 146의 아미노산 25-1351, 서열 번호 155의 아미노산 25-1159, 서열 번호 156의 아미노산 25-1156, 또는 서열 번호 157의 아미노산 25-1159와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 포함하는 변형된 S 단백질을 포함하고, 여기서, 변형된 S 단백질은 숙주 또는 숙주 세포에서 발현되었을 때, VLP를 형성한다. 변형된 S 단백질의 엑토도메인 및 막횡단 도메인의 아미노산 서열은 서열 번호 1의 아미노산 1-1234와, 서열 번호 2와 아미노산 1-1219와, 서열 번호 5의 아미노산 1-1234와, 서열 번호 21의 아미노산 1-1219와, 서열 번호 30의 아미노산 1-1243과, 서열 번호 47의 아미노산 25-1243과, 서열 번호 48의 아미노산 25-1243과, 서열 번호 49의 아미노산 25-1243과, 서열 번호 50의 아미노산 25-1243과, 서열 번호 51의 아미노산 25-1243과, 서열 번호 52의 아미노산 25-1243과, 서열 번호 53의 아미노산 25-1243과, 서열 번호 54의 아미노산 25-1243과, 서열 번호 55의 아미노산 25-1243과, 서열 번호 56의 아미노산 25-1243과, 서열 번호 57의 아미노산 25-1243과, 서열 번호 58의 아미노산 25-1243과, 서열 번호 59의 아미노산 25-1242와, 서열 번호 60의 아미노산 25-1242와, 서열 번호 61의 아미노산 25-1246과, 또는 서열 번호 62의 아미노산 25-1245, 서열 번호 95의 아미노산 25-1227, 서열 번호 96의 아미노산 25-1227, 서열 번호 97의 아미노산 25-1227, 서열 번호 108의 아미노산 25-1325, 서열 번호 109의 아미노산 25-1325, 서열 번호 110의 아미노산 25-1325, 서열 번호 144의 아미노산 25-1335, 서열 번호 145의 아미노산 25-1335, 서열 번호 146의 아미노산 25-1335, 서열 번호 155의 아미노산 25-1143, 서열 번호 156의 아미노산 25-1143, 또는 서열 번호 157의 아미노산 25-1143과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 갖고, 변형된 S 단백질의 세포질 테일 도메인(CT)의 아미노산 서열은 서열 번호 15의 서열과, 또는 서열 번호 6, 8, 7, 9, 10, 12, 13, 14의 아미노산 35-50과, 또는 서열 번호 11의 아미노산 34-49와 함께, 또는 서열 번호 3의 아미노산 553-568과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 갖고, 여기서, 변형된 S 단백질은 숙주 또는 숙주 세포에서 발현되었을 때, VLP를 형성한다.

[0324] 추가로, 변형된 바이러스 구조 단백질은 서열 번호 22, 26, 29, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 90, 91, 92, 95, 96, 97, 103, 104, 105, 139, 140, 141, 150, 151, 또는 152에 따른 뉴클레오티드 서열과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 뉴클레오티드 서열에 의해 코딩될 수 있고, 여기서, 뉴클레오티드 서열은 숙주 또는 숙주 세포에서 발현되었을 때, VLP를 형성하는 변형된 S 단백질을 코딩한다.

[0325] 서열 번호 1, 2, 5, 21, 30, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 95, 96, 97, 108, 109, 110, 144, 145, 146, 155, 156 또는 157의 아미노산 서열과 약 70, 75, 80, 85, 87, 90, 91, 92, 93

94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 포함하는 변형된 S 단백질을 코딩하는 뉴클레오티드 서열로서, 여기서, 변형된 S 단백질은 숙주 또는 숙주 세포에서 발현되었을 때, VLP를 형성하는 것인, 뉴클레오티드 서열을 추가로 제공한다. 뉴클레오티드 서열은 서열 번호 1의 아미노산 1-1234와, 서열 번호 2의 아미노산 1-1219와, 서열 번호 5의 아미노산 1-1234와, 서열 번호 21의 아미노산 1-1219와, 또는 서열 번호 30의 아미노산 1-1243과, 서열 번호 47의 아미노산 25-1243과, 서열 번호 48의 아미노산 25-1243과, 서열 번호 49의 아미노산 25-1243과, 서열 번호 50의 아미노산 25-1243과, 서열 번호 51의 아미노산 25-1243과, 서열 번호 52의 아미노산 25-1243과, 서열 번호 53의 아미노산 25-1243과, 서열 번호 54의 아미노산 25-1243과, 서열 번호 55의 아미노산 25-1243과, 서열 번호 56의 아미노산 25-1243과, 서열 번호 57의 아미노산 25-1243과, 서열 번호 58의 아미노산 25-1243과, 서열 번호 59의 아미노산 25-1242와, 서열 번호 60의 아미노산 25-1242와, 서열 번호 61의 아미노산 25-1246과, 서열 번호 62의 아미노산 25-1245, 서열 번호 95의 아미노산 25-1227, 서열 번호 96의 아미노산 25-1227, 서열 번호 97의 아미노산 25-1227, 서열 번호 108의 아미노산 25-1325, 서열 번호 109의 아미노산 25-1325, 서열 번호 110의 아미노산 25-1325, 서열 번호 144의 아미노산 25-1335, 서열 번호 145의 아미노산 25-1335, 서열 번호 146의 아미노산 25-1335, 서열 번호 155의 아미노산 25-1143, 서열 번호 156의 아미노산 25-1143, 또는 서열 번호 157의 아미노산 25-1143과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 변형된 S 단백질의 엑토도메인 및 막횡단 도메인의 아미노산 서열을 코딩할 수 있고, 변형된 S 단백질의 세포질 테일 도메인의 아미노산 서열은 서열 번호 15의 서열과, 또는 서열 번호 6, 8, 7, 9, 10, 12, 13, 14의 아미노산 35-50과, 또는 서열 번호 11의 아미노산 34-49와, 또는 서열 번호 3의 아미노산 553-568과 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 갖고, 여기서, 변형된 S 단백질은 숙주 또는 숙주 세포에서 발현되었을 때, VLP를 형성한다.

[0326] 서열 번호 5, 21, 30, 또는 47-62의 아미노산 서열과, 또는 서열 번호 47의 아미노산 24-1259, 서열 번호 48의 아미노산 25-1259, 서열 번호 49의 아미노산 25-1259, 서열 번호 50의 아미노산 25-1259, 서열 번호 51의 아미노산 25-1259, 서열 번호 52의 아미노산 25-1259, 서열 번호 53의 아미노산 25-1259, 서열 번호 54의 아미노산 25-1259, 서열 번호 55의 아미노산 25-1259, 서열 번호 56의 아미노산 25-1259, 서열 번호 57의 아미노산 25-1259, 서열 번호 58의 아미노산 25-1259, 서열 번호 59의 아미노산 25-1262, 서열 번호 60의 아미노산 25-1261, 서열 번호 61의 아미노산 25-1258, 또는 서열 번호 62의 아미노산 25-1256, 서열 번호 95의 아미노산 25-1243, 서열 번호 96의 아미노산 25-1240, 서열 번호 97의 아미노산 25-1243, 서열 번호 108의 아미노산 25-1341, 서열 번호 109의 아미노산 25-1338, 서열 번호 110의 아미노산 25-1341, 서열 번호 144의 아미노산 25-1351, 서열 번호 145의 아미노산 25-1348, 서열 번호 146의 아미노산 25-1351, 서열 번호 1455의 아미노산 25-1159, 서열 번호 156의 아미노산 25-1156 또는 서열 번호 157의 아미노산 25-1159와 약 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% 또는 그 사이의 임의의 양의 서열 동일성, 또는 서열 유사성을 가진 아미노산 서열을 포함하는 변형된 S 단백질을 코딩하는 뉴클레오티드 서열로서, 여기서, 변형된 S 단백질은 숙주 또는 숙주 세포에서 발현되었을 때, VLP를 형성하는 것인 뉴클레오티드 서열을 추가로 제공한다.

[0327] 특정 서열을 언급할 때, 용어 "유사성(%)," "서열 유사성," "동일성(%)," 또는 "서열 동일성"은 예를 들어, 위스콘신 대학교(University of Wisconsin) GCG 소프트웨어 프로그램에 제시된 바와 같이, 또는 수동 정렬 및 시각적 검사에 의해 사용된다(예컨대, 문헌 [Current Protocols in Molecular Biology, Ausubel et al., eds. 1995 supplement] 참조). 비교를 위한 서열 정렬 방법은 당업계에 널리 공지되어 있다. 비교를 위한 최적의 서열 정렬은 예를 들어, (Smith & Waterman) (문헌 [1981, Adv. Appl. Math. 2:482])의 알고리즘을 사용하여, (Needleman & Wunsch) (문헌 [1970, J. Mol. Biol. 48:443]48:443)의 정렬 알고리즘에 의해, (Pearson & Lipman) (문헌 문헌 [1988, Proc Natl Acad Sci USA 85:2444])의 유사성 방법 검색에 의해, 이러한 알고리즘의 전산화 구현에 의해(예를 들어, GAP, BESTFIT, FASTA, 및 TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr, Madison, Wis) 수행될 수 있다.

[0328] 서열 동일성(%) 및 서열 유사성(%)을 결정하는 데 적합한 알고리즘의 예는 BLAST 및 BLAST 2.0 알고리즘이고, 이는 각각 문헌 [Altschul et al., (1977, Nuc. Acids Res. 25:3389-3402)] 및 [Altschul et al., (1990, J. Mol. Biol. 215:403-410)]에 기술되어 있다. BLAST 및 BLAST 2.0은 본원에 기술된 파라미터를 이용하여 본 개시내용의 핵산 및 단백질에 대한 서열 동일성(%)을 결정하는 데 사용된다. 예를 들어, BLASTN 프로그램 (뉴클레오티드 서열인 경우)은 디폴트로 워드길이(W) 11, 예상치(E) 10, M=5, N=-4 및 두 가닥의 비교를 사용할 수 있다. 아미노산 서열인 경우, BLASTP 프로그램은 디폴트로 워드 길이 3, 및 예상치(E) 10, 및 BLOSUM62 스코어링 매트릭스(문헌 [Henikoff & Henikoff, 1989, Proc. Natl. Acad. Sci. USA 89:10915] 참조) 정렬값(B) 50, 예상

치(E) 10, M=5, N=-4, 및 두 가닥의 비교를 사용할 수 있다. BLAST 분석을 수행하기 위한 소프트웨어는 미국 생물공학 정보센터 (National Center for Biotechnology Information)(URL: ncbi.nlm.nih.gov/ 참조))를 통해 공개적으로 이용가능하다.

[0329] 본 개시내용에서 언급된 핵산 서열 또는 뉴클레오티드 서열은 핵산 서열 또는 뉴클레오티드 서열이 엄격한 하이브리드화 조건하에서 본원에서 정의된 바와 같은 하나 이상의 뉴클레오티드 서열 또는 핵산 서열 또는 뉴클레오티드 서열의 보체와 하이브리드화한다면, 서열, 또는 서열의 보체와 "실질적으로 상동성," "실질적으로 유사" 또는 "실질적으로 동일한" 것일 수 있다. 상기 상동성 서열이 본원에 기술된 서열, 또는 코딩된 생성물의 하나 이상의 특성을 보인다면, 적어도 약 70%, 또는 70 내지 100%, 또는 그 사이의 임의의 양, 예를 들어, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100%, 또는 그 사이의 임의의 양의 뉴클레오티드가 정의된 길이의 뉴클레오티드 서열에 걸쳐 매칭될 때, 서열은 "실질적으로 상동성," "실질적으로 유사," "실질적으로 동일한" 것이다.

[0330] 많은 유기체는 성장하는 펩티드 쇄에서 특정 아미노산의 삽입을 코딩하기 위해 특정 코돈 사용에 대한 편향을 나타낸다. 코돈 선호도 또는 코돈 편향, 유기체간 코돈 사용빈도 차이는 유전자 코드의 축퇴성에 의해 제공되며, 많은 유기체에서 잘 문서화되어 있다. 코돈 편향은 종종 메신저 RNA(mRNA: messenger RNA)의 번역 효율성과 상관성이 있고, 이는 결국 특히 번역되는 코돈의 특성 및 특정 전달 RNA(tRNA: transfer RNA) 분자의 이용가능성에 의존하는 것으로 간주된다. 세포에서 선택된 tRNA의 우세는 일반적으로 펩티드 합성에서 가장 빈번하게 사용되는 코돈을 반영한다. 따라서, 유전자는 코돈 최적화에 기초하여 주어진 유기체에서 최적의 유전자 발현을 위해 맞춤화될 수 있다. 이종성 발현 단백질을 코딩하는 뉴클레오티드 서열을 최적화하는 프로세스는 발현 수율을 증진시키기 위한 중요한 단계일 수 있다. 최적화 요건에는 외래 단백질을 생산하는 숙주의 능력을 개선시키는 단계를 포함할 수 있다.

[0331] 비효율적인 단백질 코딩 영역의 번역 동역학적 성질을 개선시키기 위한 상이한 코돈 최적화 기술이 당업계에 공지되어 있다. 이러한 기술은 주로 특정 숙주 유기체에 대한 코돈 사용빈도를 확인하는 데 의존한다. 특정 유전자 또는 서열이 상기 유기체에서 발현되어야 하는 경우, 이때 상기 유전자 및 서열의 코딩 서열은 숙주 유기체의 보다 빈번하게 사용되는 코돈에 의해 관심 서열의 코돈을 대체하도록 변형될 것이다.

[0332] "코돈 최적화"는 천연 서열의 코돈 중 적어도 하나, 1개 초과 또는 상당한 수를 또 다른 유기체 또는 종의 유전자에서 더 빈번하게 또는 가장 빈번하게 사용될 수 있는 코돈으로 대체함으로써 진행되는, 관심 숙주 또는 숙주 세포에서의 발현 증진을 위해 핵산 서열을 변형시키는 것으로 정의된다. 다양한 종이 특정 아미노산의 특정 코돈에 대해 특정 편향을 나타낸다.

[0333] 본 개시내용은 코돈 최적화된 합성 폴리뉴클레오티드 서열, 예를 들어, 서열이 인간 코돈 사용빈도 또는 식물 코돈 사용빈도에 대해 최적화된 합성 폴리뉴클레오티드 서열을 포함한다. 이어서, 코돈 최적화된 폴리뉴클레오티드 서열은 숙주, 예를 들어, 식물에서 발현될 수 있다. 더욱 구체적으로, 인간 코돈 사용빈도 또는 식물 코돈 사용빈도에 대해 최적화된 서열은 식물에서 발현될 수 있다. 이론에 의해 얽매이기를 바라지 않으면서, 인간 코돈에 대해 최적화된 서열은 서열의 구아닌-시토신 함량(GC 함량)을 증가시키고, 식물이 숙주로서 사용되는 경우, 발현 수율을 증진시키는 것으로 간주된다.

[0334] 본원에서 사용되는 바, 용어 "구성물," "벡터" 또는 "발현 벡터"는 외인성 뉴클레오티드 서열(예를 들어, 본원에 기술된 바와 같은 변형된 바이러스 구조 단백질을 코딩하는 뉴클레오티드 서열)을 숙주 세포(예컨대, 식물 세포) 내로 전달하고, 숙주 세포에서 외인성 핵산 서열의 발현을 지시하는 재조합 핵산을 지칭한다. "발현 카세트"는 숙주 세포에서의 관심 핵산의 전사를 위한 적절한 프로모터 또는 다른 조절 요소의 제어하의, 및 그에 작동가능하게(또는 작동적으로) 연결된 관심 뉴클레오티드 서열을 포함하는 핵산을 지칭한다. 당업자가 이해하는 바와 같이, 발현 카세트는 숙주 세포(예컨대, 식물 숙주)에서 활성인 임의의 서열인 종결(종결자) 서열을 포함할 수 있다. 예를 들어, 식물에서 종결 서열은 이분 RNA 바이러스, 예컨대, 코모바이러스의 RNA-2 게놈 세그먼트로부터 유래될 수 있고, 종결 서열은 NOS 종결인자일 수 있거나, 또는 종결인자 서열은 알팔파 플라스토시아닌 유전자의 3'UTR로부터 수득될 수 있다.

[0335] 본원에 기술된 바와 같이, 변형된 바이러스 구조 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산은 숙주, 숙주의 일부 또는 숙주 세포에서 바이러스 구조 단백질의 발현을 증진시키는 서열을 추가로 포함할 수 있다. 발현을 증진시키는 서열은 변형된 바이러스 구조 단백질을 코딩하는 핵산과 작동적으로 회합된 5' UTR 인핸서 요소 또는 식물 유래 발현 인핸서를 포함할 수 있다. 변형된 바이러스 구조 단백질을 코딩하는 서열은 또한 예를 들어, 인간 코돈 사용빈도 최적화, GC 함량 증가, 또는 그의 조합에 의해 발현을 증가시키기 위해 최적화

될 수 있다.

[0336] "조절 영역" "조절 요소" 또는 "프로모터"는 항상 그러한 것은 아니지만, 전형적으로는 DNA 또는 RNA, 또는 DNA 및 RNA, 둘 모두로 구성될 수 있는 유전자의 단백질 코딩 영역 상류에 위치하는 핵산의 일부를 의미한다. 조절 영역이 활성 상태이고, 관심 뉴클레오티드 서열과 작동적으로 회합되어 있거나, 또는 작동적으로 연결되어 있을 때, 이를 통해 관심 뉴클레오티드 서열은 발현될 수 있다. 조절 요소는 기관 특이성을 매개하거나, 또는 발생 또는 시간적 유전자 활성화를 제어할 수 있다. "조절 영역"은 프로모터 요소, 기본 프로모터 활성을 나타내는 핵심 프로모터 요소, 외부 자극에 반응하여 유도가능한 요소, 예컨대, 음성적 조절 요소와 같은 프로모터 활성을 매개하는 요소 또는 전사 인핸서를 포함한다. 본원에서 사용되는 바, "조절 영역"은 또한 전사 이후 활성인 요소, 예를 들어, 유전자 발현을 조정하는 조절 요소, 예컨대, 번역 및 전사 인핸서, 번역 및 전사 리프레서, 상류 활성화 서열, 및 mRNA 불안정성 결정요인을 포함한다. 이들 후자의 요소들 중 일부는 코딩 영역 가까이에 위치할 수 있다.

[0337] 본 개시내용의 맥락에서, 용어 "조절 요소" 또는 "조절 영역"은 전형적으로, 항상 그런 것은 아니지만, 일반적으로는 RNA 폴리머라제 및/또는 특정 부위에서 전사가 시작하는 데 필요한 다른 요소들에 대한 인식을 제공하여 코딩 영역의 발현을 제어하는, 구조 유전자의 코딩 서열에 대해 상류(5')에 위치하는 DNA 서열을 지칭한다. 그러나, 인트론에 또는 서열의 3'에 위치하는 다른 뉴클레오티드 서열 또한 관심 코딩 영역의 발현 조절에 기여할 수 있다는 것으로 이해되어야 한다. RNA 폴리머라제 또는 특정 부위에서 개시를 보장하기 위한 다른 전사 인자에 대한 인식을 제공하는 조절 요소의 예로는 프로모터 요소가 있다. 전부는 아니지만, 대부분의 진핵생물 프로모터 요소는 일반적으로 전사 시작 부위의 대략 25개의 염기쌍 상류에 위치하는 아데노신 및 티미닌 뉴클레오티드 염기쌍으로 구성된 보존된 핵산 서열인 TATA 박스를 포함한다. 프로모터 요소는 전사 개시를 담당하는 기본 프로모터 요소 뿐만 아니라. 유전자 발현을 변형시키는 다른 조절 요소도 포함할 수 있다.

[0338] 발생 조절성, 유도성 또는 구성적 조절 영역을 비롯한, 여러 타입의 조절 영역이 존재한다. 발생적으로 조절되거나, 제어하에서 유전자의 차등적 발현을 제어하는 조절 영역은 특정 기관 또는 조직의 발생 중의 특정 시간에 상기 기관 또는 기관의 조직 내에서 활성화된다. 그러나, 발생적으로 조절되는 일부 조절 영역은 특정 발생 단계에서 특정 기관 또는 조직 내에서 우선적으로 활성화될 수 있으며, 또한 발생 조절 방식으로 활성화되거나, 식물 내의 다른 기관 또는 조직에서도 기저 수준에서 활성화될 수 있다. 조직 특이적 조절 영역, 예를 들어, 종자 특이적 조절 영역의 예로는 나핀 프로모터 및 크루시페린 프로모터를 포함한다(문헌 [Rask et al., 1998, J. Plant Physiol. 152: 595-599]; [Bilodeau et al., 1994, Plant Cell 14: 125-130]). 잎 특이적 프로모터의 예로는 플라스토시아닌 프로모터를 포함한다(US 7,125,978(이는 본원에서 참조로 포함된다) 참조).

[0339] 유도성 조절 영역은 유도물질에 반응하여 하나 이상의 DNA 서열 또는 유전자의 전사를 직접 또는 간접적으로 활성화시킬 수 있는 영역이다. 유도물질의 부재하에서 DNA 서열 또는 유전자는 전사되지 않을 것이다. 전형적으로, 전사를 활성화시키기 위해 유도성 조절 영역에 특이적으로 결합하는 단백질 인자는 비활성 형태로 존재할 수 있으며, 이어서, 유도물질에 의해 직접 또는 간접적으로 활성 형태로 전환된다. 그러나 단백질 인자 또한 부재할 수 있다. 유도물질은 예컨대, 단백질, 대사물질, 성장 조절제, 제초제 또는 페놀 화합물과 같은 화학 작용제 또는 열, 냉기, 염분 또는 독성 요소에 의해 직접적으로 또는 예컨대, 바이러스 같은 병원체 또는 병인 물질의 작용을 통해 간접적으로 가해지는 생리적 스트레스일 수 있다. 유도성 조절 영역을 포함하는 식물 세포는 예컨대, 분무, 물주기, 가열 또는 유사한 방법과 같이, 세포 또는 식물에 유도물질을 외부적으로 적용함으로써 유도물질에 노출될 수 있다. 유도성 조절 요소는 식물 또는 비식물 유전자로부터 유래될 수 있다(예컨대, 문헌 [Gatz, C. and Lenk, I.R.P., 1998, Trends Plant Sci. 3, 352-358]). 잠재적인 유도성 프로모터의 예로는 테트라사이클린 유도성 프로모터(문헌 [Gatz, C.,1997, Ann. Rev. Plant Physiol. Plant Mol. Biol. 48, 89-108]), 스테로이드 유도성 프로모터(문헌 [Aoyama, T. and Chua, N.H.,1997, Plant J. 2, 397-404]) 및 에탄올 유도성 프로모터(문헌 [Salter, M.G., et al, 1998, Plant Journal 16, 127-132]; [Caddick, M.X., et al, 1998, Nature Biotech. 16, 177-180]) 사이토키닌 유도성 IB6 및 CKI1 유전자(문헌 [Brandstatter, I. and Kieber, J.J.,1998, Plant Cell 10, 1009-1019]; [Kakimoto, T., 1996, Science 274, 982-985]) 및 옥신 유도성 요소, DR5(문헌 [Ulmasov, T., et al., 1997, Plant Cell 9, 1963-1971])을 포함하나, 이에 제한되지 않는다.

[0340] 구성적 조절 영역은 식물의 다양한 부분에 걸쳐 및 식물 발생 내내 지속적으로 유전자의 발현을 지시한다. 공지된 구성적 조절 요소의 예로는 CaMV 35S 전사체(p35S; 문헌 [Odell et al., 1985, Nature, 313: 810-812](상기 문헌은 본원에서 참조로 포함된다)), 쌀 액틴 1(문헌 [Zhang et al, 1991, Plant Cell, 3: 1155-1165]), 액틴 2(문헌 [An *et al*., 1996, *Plant J*., 10: 107-121]), 또는 tms 2(U.S. 5,428,147), 및 트리오스포스페이트 이

소머라제 1(문헌 [Xu et. al., 1994, Plant Physiol. 106: 459-467]) 유전자, 옥수수 유비퀴틴 1 유전자(문헌 [Cornejo et al, 1993, Plant Mol. Biol. 29: 637-646]), 아라비돕시스(*Arabidopsis*) 유비퀴틴 1 및 6 유전자(문헌 [Holtorf et al, 1995, Plant Mol. Biol. 29: 637-646]), 담배 번역 개시 인자 4A 유전자(문헌 [Mandel et al, 1995 Plant Mol. Biol. 29: 995-1004]), 카사바 엽맥 모자이크 바이러스 프로모터, pCAS, (문헌 [Verdaguer et al., 1996]); 리불로스 비포스페이트 카복실라제의 작은 서브유닛의 프로모터, pRbcS: (문헌 [Outchkourov et al., 2003]), pUbi(외떡잎식물 및 쌍떡잎식물에 대해)를 포함한다.

[0341] 본원에서 사용되는 바, 용어 "구성적"은 구성적 조절 영역의 제어하에서 뉴클레오티드 서열이 모든 세포 타입에서 동일한 수준으로 발현되지만, 비록 풍부함에서의 차이가 자주 관찰되더라도, 서열이 광범위한 세포 타입에서 발현된다는 것을 반드시 나타내는 것은 아니다.

[0342] 본 개시내용의 하나 이상의 유전적 구성물은 또한 필요할 수 있는 추가 인핸서, 번역 또는 전사 인핸서를 포함할 수 있다. 인핸서는 전사되는 서열의 5' 또는 3'에 위치할 수 있다. 인핸서 영역은 당업자에게 널리 공지되어 있고, ATG 개시 코돈, 인접 서열 등을 포함할 수 있다. 개시 코돈이 존재하는 경우, 전사된 서열의 정확한 번역을 제공하기 위해 코딩 서열의 리딩 프레임("프레임내")과 동위상일 수 있다.

[0343] 용어 "5'UTR" 또는 "5' 비번역 영역," "5' 리더 서열" 또는 "5' UTR 인핸서 요소"는 번역되지 않는 mRNA의 영역을 지칭한다. 5'UTR은 전형적으로 전사 시작 부위에서 시작하여 코딩 영역의 번역 개시 부위 또는 시작 코돈 바로 앞에서 끝난다. 5' UTR은 mRNA 전사체의 안정성 및/또는 번역을 조정할 수 있다.

[0344] 본원에서 사용되는 바, 용어 "식물 유래 발현 인핸서"는 식물로부터 수득된 뉴클레오티드 서열로서, 5'UTR을 코딩하는 뉴클레오티드 서열을 지칭한다. 식물 유래 발현 인핸서의 예는 미국 특허 가출원 제62/643,053호(2018년 3월 14일 출원) 및 국제 출원 번호 PCT/CA2019/050319(2019년 3월 14일 출원)(이는 본원에서 참조로 포함된다)에, 또는 문헌 [Diamos A.G. et al. (2016, Front Plt Sci. 7:1-15](상기 문헌은 본원에서 참조로 포함된다)에 기술되어 있다. 식물 유래 발현 인핸서는 US 62/643,053 및 PCT/CA2019/050319에 기술된 바와 같이, nbEPI42, nbSNS46, nbCSY65, nbHEL40, nbSEP44, nbMT78, nbATL75, nbDJ46, nbCHP79, nbEN42, atHSP69, atGRP62, atPK65, atRP46, nb30S72, nbGT61, nbPV55, nbPPI43, nbPM64 및 nbH2A86로부터 선택될 수 있다. 식물 유래 발현 인핸서는 식물 유래 발현 인핸서 서열 및 관심 뉴클레오티드 서열, 예를 들어, 변형된 S 단백질을 코딩하는 뉴클레오티드 서열과 작동적으로 연결된 조절 영역을 포함하는 식물 발현 시스템 내에서 사용될 수 있다.

[0345] RNA의 안정성 및/또는 번역 효율은 3' 비번역 영역(3'UTR)을 포함함으로써 추가로 개선될 수 있다. 따라서, 본 기술내용의 하나 이상의 유전적 구성물은 3' UTR을 추가로 포함할 수 있다.

[0346] 3' 비번역 영역은 폴리아데닐화 신호 및 mRNA 프로세싱 또는 유전자 발현에 영향을 줄 수 있는 임의의 다른 조절 신호를 포함할 수 있다. 폴리아데닐화 신호는 일반적으로 mRNA 전구체의 3' 단부에 폴리아데닐산 트랙을 부가하는 것을 특징으로 한다. 폴리아데닐화 신호는 일반적으로 표준 형태 5' AATAAA-3'에 대한 상동성의 존재에 의해 인식되지만, 변형은 드물지 않다. 적합한 3' 영역의 비제한적 예는 아그로박테리움(*Agrobacterium*) 종양 유도(Ti: tumor inducing) 플라스미드 유전자, 예컨대, 노팔린 신타제(Nos 유전자) 및 식물 유전자, 예컨대, 대두 저장 단백질 유전자, 리불로스-1,5-비스포스페이트 카복실라제 유전자의 작은 서브유닛(ssRUBISCO; US 4,962,028(이는 본원에서 참조로 포함된다))의 폴리아데닐화 신호를 함유하는 3' 전사된 비번역 영역, US 7,125,978(이는 본원에서 참조로 포함된다)에 기술된, 플라스토시아닌 발현을 조절하는 데 사용되는 프로모터, 아라카차 바이러스 B 분리주 유전자(AvB: Arracacha virus B)(서열 번호 40)로부터 유래된 3' UTR, 비트 괴저성 엽맥 황화 바이러스(trBNYVV: Beet necrotic yellow vein virus)(서열 번호 41)로부터 유래된 3' UTR, 써던 빈 모자이크 바이러스(SBMV: Southern bean mosaic virus)(서열 번호 42)로부터 유래된 3' UTR, 순무 링스팟 바이러스(TuRSV: Turnip ringspot virus)(서열 번호 43)로부터 유래된 3' UTR, 동부 모자이크 바이러스(CPMV: Cowpea Mosaic Virus)(서열 번호 44)로부터 유래된 3' UTR, 브로드 빈 트루 모자이크 바이러스(BBTMV: Broad bean true mosaic virus)(서열 번호 45)로부터 유래된 3' UTR, 또는 우르미아 멜론 바이러스 (trOUMV: Ourmia melon virus)(서열 번호 46)로부터 유래된 3' UTR이 있다. 3'UTR은 발현 수준을 조정하기 위해 이종성 서열로부터 유래된 5'UTR과 함께 사용될 수 있다.

[0347] 따라서, 3'UTR의 제어하에, 및 그에 작동가능하게 (또는 작동적으로) 연결된 관심 (예컨대, 변형된 바이러스 구조 단백질) 뉴클레오티드 서열을 포함하는 핵산을 포함하는 "구성물," "벡터," "발현 벡터" 또는 "발현 카세트"를 제공한다. 추가로, 핵산은 관심 (예컨대, 변형된 바이러스 구조 단백질) 뉴클레오티드 서열에 작동가능하게 (또는 작동적으로) 연결된 3'UTR을 포함할 수 있다.

[0348] 변형된 바이러스 구조 단백질은 필요에 따라 예컨대, 식물 또는 식물 세포와 같은 숙주 또는 숙주 세포의 임의의 세포내 또는 세포외 공간, 소기관 또는 조직으로 표적화될 수 있다. 발현된 단백질을 특정 위치에 국재화시키기 위해, 단백질을 코딩하는 핵산은 신호 펩티드 또는 리더 서열을 코딩하는 핵산 서열에 연결될 수 있다. 대안적으로, 신호 펩티드는 전이 펩티드, 신호 서열, 리더 서열, 표적화 신호, 국재화 신호, 국재화 서열, 전이 펩티드 또는 리더 펩티드로 언급될 수 있다.

[0349] 본 기술내용의 하나 이상의 변형된 유전적 구성물은 본 개시내용의 핵산, 또는 뉴클레오티드 서열, 또는 구성물, 또는 벡터에 의해 형질전환되는 임의의 적합한 숙주 또는 숙주 세포에서 발현될 수 있다. 숙주 또는 숙주 세포는 식물, 진균, 박테리아, 곤충 및 동물, 예를 들어, 포유동물을 포함하는 임의의 공급원으로부터 유래될 수 있다. 따라서, 숙주 또는 숙주 세포는 식물 또는 식물 세포, 진균 또는 진균 세포, 박테리아 또는 박테리아 세포, 곤충 또는 곤충 세포, 및 동물 또는 동물 세포로부터 선택될 수 있다. 포유동물 또는 동물은 인간이 아닐 수 있다. 바람직한 실시양태에서, 숙주 또는 숙주 세포는 식물, 식물의 일부 또는 식물 세포이다.

[0350] 본원에서 사용되는 바, 용어 "식물," "식물의 일부," "식물 일부," "식물성 물질(matter)," "식물 바이오매스," "식물성 물질(material)," "식물 추출물" 또는 "식물 잎"은 본원에서 기술된 하나 이상의 핵산 발현을 위한 전사, 번역, 및 번역 후 기구를 제공할 수 있고/거나, 발현된 단백질 또는 VLP의 추출 및 정제 기점이 될 수 있는, 전체 식물, 조직, 세포, 또는 그의 임의의 분획, 세포내 식물 성분, 세포외 식물 성분, 식물의 지질 또는 고체 추출물, 또는 그의 조합을 포함할 수 있다. 식물은 초본 식물을 포함할 수 있지만, 이에 제한되지 않는다. 초본 식물은 1년생, 2년생 또는 다년생 식물일 수 있다. 식물은 추가로 제한하는 것은 아니지만, 예를 들어, 카놀라, 브라시카 종(*Brassica* spp.), 옥수수, 니코티아나 종(*Nicotiana* spp.), (담배) 예를 들어, 니코티아나 벤타미아나(*Nicotiana benthamiana*), 니코티아나 러스티카(*Nicotiana rustica*), 니코티아나 타바쿰(*Nicotiana tabacum*), 니코티아나 알라타(*Nicotiana alata*), 아라비돕시스 탈리아나(*Arabidopsis thaliana*), 알팔파, 감자, 고구마(이포뫼아 바타투스(*Ipomoea batatus*)), 인삼, 완두콩, 귀리, 쌀, 대두, 밀, 보리, 해바라기, 목화, 옥수수, 호밀 (세칼레 세레알레(*Secale cereale*)), 수수(소루굼 비콜로르(*Sorghum bicolor*), 소루굼 불가레(*Sorghum vulgare*)), 잇꽃(카르타무스 틴크토리우스(*Carthamus tinctorius*))을 비롯한, 농작물을 포함할 수 있다.

[0351] 본원에서 사용되는 바, 용어 "식물 일부"는 잎, 줄기, 뿌리, 꽃, 과실, 잎, 줄기, 뿌리, 꽃, 과실로부터 수득되는 식물 세포, 잎, 줄기, 뿌리, 꽃, 과실로부터 수득되는 식물 추출물, 또는 그의 조합을 포함하나, 이에 제한되지 않는, 식물의 임의의 부분을 지칭한다. 한 실시양태에서, 식물 일부는 식물, 예컨대, 예를 들어, 잎, 줄기, 꽃 및 과실의 지상부를 지칭한다. 본원에서 사용되는 바, 용어 "식물 추출물"은 식물, 식물의 일부, 식물 세포, 또는 그의 조합을 물리적으로(예를 들어, 동결시킨 후, 이어서, 적합한 완충제 중에서 추출하여), 기계적으로(예를 들어, 식물 또는 식물의 일부를 그라인딩 또는 균질화한 후, 적합한 완충제 중에서 추출하여), 효소적으로(예를 들어, 세포벽 분해 효소를 사용하여), 화학적으로(예를 들어, 하나 이상의 킬레이터 또는 완충제를 사용하여), 또는 그의 조합으로 처리한 후 수득되는, 식물 유래 생성물을 지칭한다. 식물 추출물은 원치않는 식물 성분, 예를 들어, 세포벽 파편을 제거하기 위하여 추가적으로 프로세싱될 수 있다. 식물 추출물은 식물, 식물의 일부 또는 식물 세포로부터의 하나 이상의 성분, 예를 들어, 식물, 식물의 일부, 또는 식물 세포로부터의 단백질(단백질 복합체, 단백질 상부구조 및/또는 VLP 포함), 핵산, 지질, 탄수화물, 또는 그의 조합의 회수를 돕기 위해 수득될 수 있다. 식물 추출물이 단백질을 포함하는 경우, 이는 단백질 추출물로 불릴 수도 있다. 단백질 추출물은 식물 조직으로부터 하나 이상의 단백질, 단백질 복합체, 예를 들어, 단백질 삼량체, 단백질 상부구조 및/또는 VLP를 포함하는 미정제 식물 추출물, 부분적으로 정제된 식물 또는 단백질 추출물, 또는 정제된 생성물일 수 있다. 원하는 경우, 단백질 추출물 또는 식물 추출물은 당업자에게 공지된 기술을 사용하여 부분적으로 정제될 수 있고, 예를 들어, 추출물은 염 또는 pH 침전, 원심분리, 구배 밀도 원심분리, 여과, 크로마토그래피, 예를 들어, 크기 배제 크로마토그래피, 이온 교환 크로마토그래피, 친화성 크로마토그래피 또는 그의 조합이 수행될 수 있다. 단백질 추출물은 또한 당업자에게 공지된 기술을 사용하여 정제될 수 있다.

[0352] 본 개시내용의 구성물은 Ti 플라스미드, Ri 플라스미드, 식물 바이러스 벡터, 직접적인 DNA 형질 전환, 미세주입, 전기천공 등을 사용하여 식물 세포에 도입될 수 있다. 상기 기술에 대한 리뷰를 위해, 예를 들어, 문헌 [Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academy Press, New York VIII, pp. 421-463 (1988)]; [Geierson and Corey, *Plant Molecular Biology*, 2d Ed. (1988)]; 및 [Miki and Iyer, Fundamentals of Gene Transfer in Plants. In *Plant Metabolism*, 2d Ed. DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison Wesly, Langmans Ltd. London, pp. 561-579 (1997)]을 참조한다. 다른 방법으로는 직접적인 DNA 흡수, 리포솜 사용, 전기천공, 예를 들어, 원형질체, 미세주입, 미세발사체 또는 위스

커를 사용하는 것, 및 진공 침투를 포함한다. 예를 들어, 문헌 [Bilang, et al. (Gene 100: 247-250 (1991)], [Scheid et al. (Mol. Gen. Genet. 228: 104-112, 1991)], [Guerche et al. (Plant Science 52: 111-116, 1987)], [Neuhause et al. (Theor. Appl Genet. 75: 30-36, 1987)], [Klein et al., Nature 327: 70-73 (1987)]; [Howell et al. (Science 208: 1265, 1980)], [Horsch et al. (Science 227: 1229-1231, 1985)], [DeBlock et al., Plant Physiology 91: 694-701, 1989)], [Methods for Plant Molecular Biology (Weissbach and Weissbach, eds., Academic Press Inc., 1988)], [Methods in Plant Molecular Biology (Schuler and Zielinski, eds., Academic Press Inc., 1989)], [Liu and Lomonossoff (J Virol Meth, 105:343-348, 2002,)], 미국 특허 제4,945,050호; 제5,036,006호; 및 제5,100,792호, 미국 특허 출원 시리얼 제08/438,666호 (1995년 5월 10일 출원), 및 제07/951,715호(1992년 9월 25일 출원)(상기 모두 본원에서 참조로 포함된다)를 참조한다.

[0353] 하기 기술되는 바와 같이, 일시적 발현 방법을 사용하여 본 개시내용의 구성물을 발현할 수 있다(문헌 [Liu and Lomonossoff, 2002, Journal of Virological Methods, 105:343-348] (상기 문헌은 본원에서 참조로 포함된다) 참조). 대안적으로, 문헌 [Kapila et al., 1997](상기 문헌은 본원에서 참조로 포함된다)에 기술된 바와 같이, 진공 기반 일시적 발현 방법이 사용될 수 있다. 상기 방법은 제한하는 것은 아니지만, 예를 들어, 아그로접종 (Agro-inoculation) 또는 아그로침투, 시린지 침투의 방법을 포함할 수 있지만, 상기 언급된 바와 같이 다른 일시적 방법도 사용될 수 있다. 아그로접종, 아그로침투, 시린지 침투의 경우, 원하는 핵산을 포함하는 아그로박테리아의 혼합물이 조직, 예를 들어, 잎, 식물의 지상부 (줄기, 잎 및 꽃 포함), 식물의 다른 일부 (줄기, 뿌리, 꽃) 또는 전체 식물의 세포사이 공간으로 들어간다. 표피를 통과한 후, 아그로박테리아는 t-DNA 카피를 감염시키고, 그를 세포로 전달한다. t-DNA는 에피솜에 의해 전사되고, mRNA는 번역되어, 감염된 세포에서 관심 단백질의 생산으로 이어지지만, 핵 내부로 t-DNA의 통과는 일시적이다.

[0354] 형질전환된 식물 세포의 확인을 돕기 위해, 본 개시내용의 구성물은 식물 선별가능한 마커를 포함하도록 추가로 조작될 수도 있다. 유용한 선별가능한 마커는 화학물질, 예컨대, 항생제, 예를 들어, 젠타마이신, 히그로마이신, 카나마이신, 또는 제초제, 포스피노트리신, 글리포세이트, 클로로술푸론 등에 대한 내성을 제공하는 효소를 포함한다. 유사하게, 변색에 의해 확인가능한 화합물의 생산을 제공하는 효소, 예컨대, GUS(베타-글루쿠로니다제), 또는 발광, 예컨대, 루시페라제 또는 GFP가 사용될 수도 있다.

[0355] 또한 본 개시내용의 일부로 간주되는 것은 본원에 기술된 일시적 단백질 발현에 적합한 플랫폼 식물로서 사용될 수 있는 본 개시내용의 유전적 구성물을 함유하는 트랜스제닉 식물, 식물 세포 또는 종자이다. 식물 세포로부터 전체 식물을 재생하는 방법은 또한 당업계에 공지되어 있다(예를 들어, 문헌 [Guerineau and Mullineaux (1993, Plant transformation and expression vectors. In: Plant Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, pp 121-148] 참조). 일반적으로, 형질전환된 식물 세포는 예컨대, 항생제와 같은 선별제를 함유할 수 있는 적절한 배지에서 배양되며, 여기서, 형질전환된 식물 세포의 확인을 용이하게 하기 위해 선별가능한 마커가 사용된다. 일단 캘러스가 형성되면, 공지된 방법에 따라 적절한 식물 호르몬을 사용하고, 식물 재생을 위해 싹을 발근 배지로 옮겨 싹 형성을 촉진할 수 있다. 이어서, 식물은 종자에서 또는 영양 번식 기술을 사용하여 반복적인 세대를 확립하는 데 사용될 수 있다. 트랜스제닉 식물은 또한 조직 배양을 사용하지 않고도 생성될 수 있다. 이들 유기체의 안정적인 형질전환 및 재생은 당업계에 확립되어 있으며 당업자에게 공지되어 있다. 이용가능한 기술은 문헌 [Vasil et al. (Cell Culture and Somatic Cell Genetics of Plants, VoI I, II and III, Laboratory Procedures and Their Applications, Academic Press, 1984)], 및 [Weissbach and Weissbach (Methods for Plant Molecular Biology, Academic Press, 1989)]에서 리뷰된다. 형질전환 및 재생된 식물을 수득하는 방법은 본 개시내용에서 중요하지 않다.

[0356] 식물, 식물 일부 또는 식물 세포가 둘 이상의 핵산 구성물로 형질전환되거나, 또는 공동 형질전환되어야 하는 경우, 핵산 구성물은 감염 이벤트로 아그로박테리움 내로 도입될 수 있고, 그 결과 핵산이 풀링되고, 박테리아 세포가 형질감염 된다. 대안적으로, 구성물은 연속적으로 도입될 수 있다. 이러한 경우, 제1 구성물은 기술된 바와 같이 아그로박테리움 내로 도입되고, 세포는 단독으로 형질전환된 박테리아만이 성장할 수 있는 선별 조건 (예컨대, 항생제 존재시) 하에서 성장한다. 제1 선택 단계 이후, 제2 핵산 구성물이 기술된 바와 같이, 아그로박테리움 내로 도입되고, 세포는 이중 선별 조건하에서 성장하며, 이중 형질전환된 박테리아만이 성장할 수 있다. 이어서, 이중 형질전환된 박테리아는 본원에 기술된 바와 같이 식물, 식물의 일부 또는 식물 세포를 형질전환시키는 데 사용될 수 있거나, 또는 제3 핵산 구성물을 수용하도록 추가의 형질전환 단계를 거칠 수 있다.

[0357] 대안적으로, 식물, 식물 일부, 또는 식물 세포가 둘 이상의 핵산 구성물로 형질전환되거나, 또는 공동 형질전환되어야 하는 경우, 핵산 구성물은 아그로박테리움 세포의 혼합물을 식물, 식물 일부 또는 식물 세포로 동시-침

투시킴으로써, 식물 내로 도입될 수 있고, 각각의 아그로박테리움 세포는 식물 내에 도입되는 하나 이상의 구성물을 포함할 수 있다. 침투 단계 중에 식물, 식물 일부 또는 식물 세포 내에서의 구성물내 관심 뉴클레오티드 서열의 상대적인 발현 수준을 변화시키기 위하여, 원하는 구성물을 포함하는 다양한 아그로박테리아 집단의 농도가 달라질 수 있다.

[0358] 본원에 기술된 바와 같은, 변형된 바이러스 표면 단백질 또는 변형된 바이러스 표면 단백질을 포함하는 VLP는 피험체에서 면역 반응을 유도하는 데 사용될 수 있다.

[0359] "면역 반응"은 일반적으로 피험체의 적성 면역계의 반응을 지칭한다. 적응 면역계는 일반적으로 체액성 반응, 및 세포 매개 반응을 포함한다. 체액성 반응은 B 림프구 계통의 세포(B 세포)에서 생산된, 분비된 항체에 의해 매개되는 면역의 한 측면이다. 분비된 항체는 침입한 미생물(예컨대, 바이러스 또는 박테리)의 표면에 있는 항원에 결합하며, 이는 파괴하기 위해서 이들을 표시한다. 체액성 면역은 일반적으로 항체 생산 및 그에 수반되는 프로세스 뿐만 아니라, Th2 세포 활성화를 포함하는 항체의 이펙터 기능 및 사이토카인 생산, 기억 세포 생성, 식균 작용의 옵소닌 촉진, 병원체 제거 등을 지칭한다. 용어 "조정하다" 또는 "조정" 등은 일반적으로 공지되거나, 사용되는, 일부는 본 명세서에 예시된 여러 검정법 중 어느 하나에 의해 결정된 특정 반응 또는 파라미터의 증가 또는 감소를 지칭한다. 그 중 일부는 본 명세서에 예시된다.

[0360] 세포 매개 반응은 항체를 포함하지 않으며, 대신에 대식세포, 자연 살해 세포(NK: natural killer cell), 항원 특이적 세포독성 T 림프구의 활성화 및 항원에 반응하는 다양한 사이토카인의 방출을 수반하는 면역 반응이다. 세포 매개 면역은 일반적으로 일부 Th 세포 활성화, Tc 세포 활성화 및 T 세포 매개 반응을 지칭하는데 사용된다. 세포 매개 면역은 바이러스 감염에 반응하는 데 특히 중요할 수 있다.

[0361] 예를 들어, 항원 특이적 CD8 양성 T 림프구의 유도는 ELISPOT 검정법을 사용하여 측정될 수 있고; CD4 양성 T 림프구의 자극은 증식 검정법을 사용하여 측정될 수 있다. 항코로나바이러스 항체 역가는 ELISA 검정법을 사용하여 정량화될 수 있고; 항원 특이적 또는 교차 반응성 항체의 이소타입은 또한 항이소타입 항체(예컨대, 항 IgG, IgA, IgE 또는 IgM)를 사용하여 측정될 수 있다. 상기 검정법을 수행하기 위한 방법 및 기술은 당업계에 널리 공지되어 있다.

[0362] 사이토카인 존재 또는 수준도 또한 정량화될 수 있다. 예를 들어, T 헬퍼 세포 반응(Th1/Th2)은 ELISA(예컨대, BD Biosciences OptEIA 키트)를 사용한 IFN-γ 및 IL-4 분비 세포의 측정을 특징으로 할 수 있다. 피험체로부터 수득된 말초 혈액 단핵 세포(PBMC: peripheral blood mononuclear cell) 또는 비장세포를 배양하고, 상청액을 분석할 수 있다. T 림프구는 또한 당업계에 공지된 마커 특이적 형광 라벨 및 방법을 사용하여 형광 활성화 세포 분류(FACS: fluorescence-activated cell sorting)에 의해 정량화될 수 있다.

[0363] 피험체에서 면역 반응을 특징화하기 위해 미세중화 검정법 또한 수행될 수 있으며, 예를 들어, 문헌 [Rowe *et al.*, 1973]의 방법을 참조한다. 바이러스 중화 역가는 세포의 크리스탈 바이올렛 고정/착색 후 용해 플라크의 측정(플라크 검정법); 시험관내 배양에서 세포 용해의 현미경 관찰, 및 2) 코로나바이러스의 ELISA 및 분광광도 검출을 비롯한, 다양한 방법으로 정량화할 수 있다.

[0364] 본원에서 사용되는 바, 용어 "에피토프" 또는 "에피토프들"은 항체가 특이적으로 결합하는 항원의 구조적 부분을 지칭한다.

[0365] 본원에 기술된 바와 같은 변형된 바이러스 구조 단백질, 삼량체 또는 삼량체 변형된 바이러스 구조 단백질 또는 변형된 바이러스 구조 단백질을 포함하는 VLP를 피험체, 또는 숙주 동물에게 투여하여 항체 또는 항체 단편을 생산하는 단계를 포함하는, 항체 또는 항체 단편을 생산하는 방법을 제공한다. 본 방법에 의해 생산된 항체 또는 항체 단편 또한 제공한다.

[0366] 따라서, 본 개시내용은 또한 피험체에서 코로나바이러스 감염에 대한 면역성을 유도하기 위한, 본원에 기술된 바와 같은 바이러스 구조 단백질 또는 변형된 바이러스 구조 단백질을 포함하는 VLP의 용도를 제공한다. 또한 변형된 바이러스 구조 단백질 또는 변형된 바이러스 구조 단백질을 포함하는 VLP를 피험체 또는 숙주 동물에게 투여함으로써 제조된 항체 또는 항체 단편 또한 본원에서 개시된다.

[0367] 피험체에서 면역 반응을 유도하기 위한, 유효 용량의, 본원에 기술된 바와 같은 변형된 바이러스 구조 단백질 또는 변형된 바이러스 구조 단백질을 포함하는 VLP, 및 약학적으로 허용되는 담체, 애주번트, 비히클, 또는 부형제를 포함하는 조성물을 추가로 제공한다. 유효 용량의 변형된 바이러스 구조 단백질 또는 변형된 바이러스 구조 단백질을 포함하는 VLP를 포함하는, 피험체에서 코로나바이러스에 대한 면역 반응을 유도하기 위한 백신

또한 제공한다.

[0368] 조성물 내의 VLP 중 적어도 하나가 본원에 기술된 바와 같은 변형된 코로나바이러스 S 단백질을 포함한다면, VLP의 혼합물을 포함할 수 있는 조성물도 추가로 제공한다. 예를 들어, 하나 이상의 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주 각각으로부터의 하나 이상의 변형된 S 단백질을 포함하는 각 코로나바이러스 S 단백질이 발현될 수 있고, 상응하는 VLP는 정제될 수 있다. VLP 혼합물 중 하나 이상의 VLP가 본원에 기술된 바와 같은 변형된 S 단백질을 포함한다면, 2개 이상의 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주)로부터 수득된 바이러스 유사 입자는 원하는 대로 조합되어 VLP 혼합물을 생산할 수 있다. VLP는 원하는 비로, 예를 들어, 거의 등가인 비로 조합 또는 생산될 수 있거나, 또는 한 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주가 조성물 중 다수의 VLP를 구성하도록 하는 방식으로 조합될 수 있다. 하나 이상의 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주 각각으로부터 유래된 엑토도메인 및/또는 TM 또는 TM의 일부를 포함하는 하나 이상의 변형된 S 단백질을 포함하는 VLP의 조성물로서, 이로써, 본 개시내용에서 제공되는 바와 같은 상이한 변형된 S 단백질의 혼합물이 임의의 개별 VLP 조성물에 존재할 수 있는 것인 VLP의 조성물을 추가로 제공한다.

[0369] 조성물 또는 백신은 한 타입의 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터의 변형된 바이러스 구조 단백질, 예컨대, 변형된 S 단백질을 포함하는 VLP를 포함할 수 있거나, 또는 조성물 또는 백신은 다중 VLP 타입을 포함할 수 있고, 여기서, 각 VLP 타입은 변형된 S 단백질을 포함하고, 여기서, 동일한 VLP 중의 변형된 S 단백질은 한 타입의 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터 유래된 것이고, 즉, 조성물 또는 백신은 상이한 코로나바이러스 VLP의 혼합물을 포함할 수 있고, 여기서, 각 VLP는 동일한 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터의 변형된 S 단백질을 포함할 수 있다. 예를 들어, 조성물 또는 백신은 제1 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터의 제1 변형된 S 단백질을 포함하는 제1 VLP, 및 제2 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터의 제2 변형된 S 단백질을 포함하는 제2 VLP를 포함할 수 있다. 추가로, 조성물은 또한 제3 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터의 제3 변형된 S 단백질을 포함하는 제3 VLP를 포함할 수 있고/거나, 조성물 또는 백신은 제4 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주로부터의 제4 변형된 S 단백질을 포함하는 제4 VLP를 포함할 수 있다. 따라서, 본 기술내용은 또한 1가(단가), 또는 다가(monovalent)(다가(polyvalent))인 조성물 또는 백신을 제공한다. 1가 조성물 또는 백신은 피험체를 단일 타입의 코로나바이러스 균주에 대해 면역화시킬 수 있는 반면, 다가 조성물 또는 백신은 피험체를 1 초과의 코로나바이러스 균주에 대해 면역화시킬 수 있다. 예를 들어, 조성물 또는 백신은 투여시 피험체를 2개의 상이한 코로나바이러스 과, 서브군, 타입, 서브타입, 계통 또는 균주 타입에 대해 면역화시킬 수 있는 2가 조성물 또는 백신일 수 있다. 추가로, 조성물 또는 백신은 3가 조성물일 수 있거나, 또는 조성물은 4가(tetravalent 또는 quadrivalent) 조성물 또는 백신일 수 있다.

[0370] 추가로, 다가 조성물은 상이한 HA 세포질 테일을 포함하는 하나 이상의 변형된 S 단백질을 포함하는 VLP를 포함할 수 있다. 예를 들어, 다가 조성물은, 각각이 S 단백질 엑토도메인, S 단백질 막횡단 도메인, 및 인플루엔자 H1, H3, H5, H6, H7, H9 또는 B 균주로부터의 HA로부터 유래된 세포질 테일을 포함하는 것인, 2개 이상의 변형된 S 단백질을 포함하는 한 VLP 또는 복수의 VLP를 포함할 수 있다. 인플루엔자 균주의 비제한적인 예로는 예를 들어, H1 California/7/2009, H3 A/Minnesota/41/2019, H5 A/Indonesia/5/05, H6 A/Teal/Hong Kong/W312/97, H7 A/Guangdong/17SF003/2016, H9 A/Hong Kong/1073/99 또는 B/Washington/02/2019가 있다.

[0371] 다중 타입 VLP를 포함하는 다가 조성물 또는 백신은 피험체에서 면역 반응을 유도하기 위한 것으로, 약학적으로 허용되는 담체, 애주번트, 비히클, 또는 부형제를 추가로 포함할 수 있다.

[0372] 백신 항원에 대한 피험체의 면역 반응을 증진시키기 위한 애주번 시스템은 널리 공지되어 있고, 본 명세서에 기술된 바와 같은 백신 또는 약학적 조성물과 함께 사용될 수 있다. 사용될 수 있는 애주번트에는 많은 타입이 있다. 인간용으로 일반적인 애주번트는 수산화알루미늄, 인산알루미늄 및 인산칼슘이 있다. 또한, 오일 에멀젼(프로인트 불완전 애주번트(FIA: Freund's incomplete adjuvant), Montanide™, Adjuvant 65 및 Lipovant™과 같은 수중유 또는 유중수 에멀젼), 박테리아(또는 그의 합성 유도체)로부터의 생성물, 내독소, 지방산, 파라핀 또는 식물성 오일, 콜레스테롤 및 지방족 아민 또는 예를 들어, 스쿠알렌과 같은 천연 유기 화합물 기반의 다수의 애주번트가 있다. 사용될 수 있는 비제한적인 애주번트로는 예를 들어, 스쿠알렌 오일의 수중유 에멀젼(예를 들어, MF-59 또는 AS03), 스쿠알렌 오일의 수중유 에멀젼(예를 들어, MF-59 또는 AS03), 안정한 에멀젼(SE: stable emulsion)에 통합된 합성 TLR4 효능제 글루코피라노실 지질 A(GLA: glucopyranosyl 지질 A)로 구성된

애주번트(GLA-SE) 또는 CpG 1018 톨-유사 수용체(TLR9) 효능제 애주번트를 포함한다.

[0373] 따라서, 백신 또는 약학적 조성물은 하나 이상의 애주번트를 포함할 수 있다. 예를 들어, 백신 또는 약학적 조성물은 수산화알루미늄, 인산알루미늄, 인산칼슘, 수중유 또는 유중수 에멀젼, 스쿠알렌을 포함하는 에멀젼(예를 들어, MF-59 또는 AS03), GLA-SE를 포함하는 에멀젼, 또는 CpG 1018 애주번트를 포함할 수 있다.

[0374] 본 기술내용의 약학적 조성물, 백신 또는 제제는 그 자체가 공지된 방식으로, 예컨대, 통상적인 혼합, 용해, 과립화, 당의정 제조, 분말화, 유화, 캡슐화, 포획 또는 정제화 프로세스에 의해 제조될 수 있다.

[0375] 약학적 조성물, 백신 또는 제제는 투여 전에 임의의 구성 성분의 혼합 또는 사전 혼합에 의해, 예를 들어, 최종 제제, 백신 또는 약학적 조성물을 투여하기 전에 수행되는 단계로서, 2개 이상의 백신 현탁액, 약학적으로 허용되는 담체, 애주번트, 비히클 또는 부형제의 수동 또는 기계 보조 혼합에 의해 생산될 수 있다.

[0376] 약학적 조성물, 백신 또는 제제는 피험체에게 경구적으로, 피내로, 비내로, 근육내로, 복강내로, 정맥내로, 또는 피하로 투여될 수 있다

[0377] 주사제는 액체 용액 또는 현탁액, 주사 전에 액체에 용액 또는 현탁액에 적합한 고체 형태, 또는 에멀젼으로서 통상적인 형태로 제조될 수 있다. 적합한 부형제는 예를 들어, 물, 식염수, 덱스트로스, 만닛톨, 락토스, 레시틴, 알부민, 글루타민산나트륨, 시스테인 하이드로클로라이드 등이다. 추가로, 원하는 경우, 주사 가능한 약약학적 조성물은 예컨대, 습윤화제, pH 완충화제 등과 같은 비독성 보조 물질을 소량 함유할 수 있다. 생리학적으로 적합한 완충제로는 행크스 용액(Hanks's solution), 링거액(Ringer's solution) 또는 생리식염수 완충제를 포함하나, 이에 제한되지 않는다. 원하는 경우, 흡수 증진 제제(예를 들어, 리포솜)가 사용될 수 있다.

[0378] 조성물 또는 백신은 피험체에게 1회 투여될 수 있다(단일 용량). 추가로, 백신 또는 조성물은 피험체에게 다회 투여될 수 있다(다회 용량). 따라서 조성물, 제제, 또는 백신은 면역 반응을 유도하기 위해 단일 용량으로 피험체에게 투여될 수 있거나, 또는 조성물, 제제, 또는 백신은 다회 투여될 수 있다(다회 투여량). 예를 들어, 조성물 또는 백신 용량은 2, 3, 4 또는 5회 투여될 수 있다.따라서, 조성물 또는 백신은 피험체에게 초기 용량으로 투여될 수 있고, 이어서, 1 이상의 용량이 피험체에게 투여될 수 있다. 용량 투여는 서로 시간상 거리를 두고 이루어질 수 있다. 예를 들어, 초기 용량 투여 후, 1 이상의 후속 용량은 초기 용량 투여로부터 1일, 2일, 3일, 4일, 5일, 6일, 1주, 2주, 3주, 1개월, 2개월, 3개월, 4개월, 5개월 또는 6개월 또는 그 사이의 임의의 시점에 투여될 수 있다. 추가로, 조성물 또는 백신은 매년 투여될 수 있다. 예를 들어, 조성물 또는 백신은 계절 백신으로서 투여될 수 있다.

[0379] 본 개시내용은 추가로 하기 서열을 제공한다.

표 4: 서열의 서열 번호 및 설명

서열 번호:	서열 설명
1	천연 SARS-CoV-2 S 단백질 wtTM/CT AA (P0DTC2)
2	신호 펩티드(SP)를 포함하지 않는 천연 SARS-CoV-2 S 단백질 wtTM/CT AA (P0DTC2)
3	H5 A/Indonesia/5/05 헤마글루티닌 (HA) AA (A5A5L7)
4	H5 A/Indonesia/5/05 헤마글루티닌 (HA) 바이러스 cDNA (EF541394.1)
5	H5 A/Indonesia/5/05 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 AA
6	H1 California/7/2009 헤마글루티닌 TM/CT 영역 AA
7	H2 A/Singapore/1/1957 HA 헤마글루티닌 TM/CT 영역 AA
8	H3 A/Minnesota/41/2019 헤마글루티닌 TM/CT 영역 AA
9	H5 A/Indonesia/5/05 헤마글루티닌 TM/CT 영역 AA
10	H6 A/Teal/Hong Kong/W312/97 헤마글루티닌 TM/CT 영역 AA
11	H7 A/Guangdong/17SF003/2016 헤마글루티닌 TM/CT 영역 AA
12	H9 A/Hong Kong/1073/99 헤마글루티닌 TM/CT 영역 AA
13	B/Washington/02/2019 헤마글루티닌 TM/CT 영역 AA
14	인플루엔자 헤마글루티닌 TM/CT 영역의 C 말단 영역의 컨센서스 서열
15	인플루엔자 헤마글루티닌의 CT 도메인의 컨센서스 서열
16	천연 SARS-CoV-2 S 단백질의 TM/CT 영역 (wtTM/CT)
17	H5 A/Indonesia/5/05 헤마글루티닌의 TM/CT (H5iTM/CT)
18	H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질의 TM/CT 영역
19	H5i 헤마글루티닌 CT를 포함하는 변형된 S 단백질의 TM/CT 영역 (변이 1)
20	(SP 부재) SARS-CoV-2 S 단백질 GSAS+PP wtTM/CT AA
21	(SP 부재) H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
22	PDI-SARS-CoV-2 S단백질 GSAS+PP wtTM/CT-DNA
23	PDI-SARS-CoV-2 S단백질 GSAS+PP wtTM/CT-AA (구성물 8591의 생성물)
24	IF(PDI)-CoV(opt2).c
25	IF(AVB)-CoV(opt2).r
26	PDI 변형된 SARS-CoV-2 S 단백질 GSAS+PP H5iTM/CT-DNA
27	PDI 변형된 SARS-CoV-2 S 단백질 GSAS+PP H5iTM/CT-AA (구성물 8597의 생성물)
28	IF(Avb)-H5I.r
29	PDI 변형된 SARS-CoV-2 S 단백질 GSAS+PP wtTM/H5iCT-DNA
30	PDI 변형된 SARS-CoV-2 S 단백질 GSAS+PP wtTM/H5iCT-AA (구성물 8671의 생성물)
31	클로닝 벡터 8501 (좌측 T-DNA에서 우측 T-DNA)

[0380]

서열 번호:	서열 설명
32	구성물 8586(2X35S(+C) prom에서 NOS term)
33	클로닝 벡터 8500(좌측 T-DNA에서 우측 T-DNA]
34	구성물 8589(2X35S(+C) prom에서 NOS term)
35	클로닝 벡터 8716(좌측 T-DNA에서 우측 T-DNA)
36	구성물 8591(2X35S(+C) prom에서 NOS term)
37	H5i 헤마글루티닌 CT를 포함하는 변형된 S 단백질의 C 말단 영역 (대체물 2)
38	H5i 헤마글루티닌 CT를 포함하는 변형된 S 단백질의 C 말단 영역 (대체물 3)
39	H5i 헤마글루티닌 CT를 포함하는 변형된 S 단백질의 C 말단 영역 (대체물 4)
40	3'UTR AvB (아라카차 바이러스 B 분리주)
41	3'UTR trBNYVV (비트 괴저성 엽맥 황화 바이러스)
42	3' UTR SBMV (써던 빈 모자이크 바이러스)
43	3' UTR TuRSV (순무 링스팟 바이러스)
44	3' UTR CPMV (동부 모자이크 바이러스)
45	3' BBTMV (브로드 빈 트루 모자이크 바이러스)
46	3' UTR trOUMV (우르미아 멜론 바이러스)
	치환을 포함하는 변형된 S 단백질
47	PDI-S 단백질 GSAS-2P (구성물 8671의 생성물)
48	PDI-S 단백질 GSAS-4P (구성물 8953의 생성물)
49	PDI-S 단백질 GSAS-6P (구성물 8940의 생성물)
50	PDI-S 단백질 GSAS-2P-923 (구성물 8933의 생성물)
51	PDI-S 단백질 GSAS-4P-923 (구성물 8960의 생성물)
52	PDI-S 단백질 GSAS-6P-923(구성물 8947의 생성물)
	상이한 HA 균주로부터의 CT를 포함하는 변형된 S
53	PDI-S 단백질+H1 Cal (구성물 7390의 생성물)
54	PDI-S 단백질+H3 Minn (구성물 7391의 생성물)
55	PDI-S 단백질+H6 HK (구성물 7392의 생성물)
56	PDI-S 단백질+H7 Guangdong (구성물 7393의 생성물)
57	PDI-S 단백질+H9 HK (구성물 7394의 생성물)
58	PDI-S 단백질+ B/ Wash (구성물 7395의 생성물)
	변형된 S 단백질 C 말단 변이
59	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V1) (구성물 8980의 생성물)
60	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V2) (구성물 8981의 생성물)
61	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V3) (구성물 8982의 생성물)
62	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V4) (구성물 8983의 생성물)
	N 말단 영역
63	천연 SARS-CoV-2 단백질(천연 신호 펩티드 포함)의 N-말단 영역
64	개재 펩티드 서열 X_n을 포함하는 변형된 SARS-CoV-2 S 단백질의 TM/CT 영역
	치환을 포함하는 변형된 S 단백질(DNA)
65	PDI-S 단백질 GSAS+4P-DNA
66	PDI-S 단백질 GSAS+6P-DNA
67	PDI-S 단백질 GSAS+2P+L923F-DNA
68	PDI-S 단백질 GSAS+4P+L923F-DNA
69	PDI-S 단백질 GSAS+6P+L923F-DNA
	변형된 S 단백질 C 말단 변이 (DNA)
70	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V1) DNA
71	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V2) DNA
72	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V3) DNA

[0381]

서열 번호:	서열 설명
73	H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V4) DNA
	상이한 HA 균주로부터의 CT를 포함하는 변형된 S (DNA) 및 프라이머
74	PDI-S 단백질 +H1 Cal DNA
75	PDI-S 단백질 +H3 Minn DNA
76	PDI-S 단백질 +H6 HK DNA
77	PDI-S 단백질 +H7 Guangdong DNA
78	PDI-S 단백질 +H9 HK DNA
79	PDI-S 단백질 + B/ Wash DNA
80	IF-H1HawaiiCT.r
81	IF-H3MinnesotaCT.r
82	IF-HongKongCT.r
83	IF-GuangdongCT.r
84	IF-H9HKCT.r
85	IF-BWashCT.r
	변형된 **SARS-COV-1 S** 단백질 **(DNA)** 및 프라이머
86	IF(nbHEL40)-PDI.c
87	IF(AvB+wtCT).r
88	PDI-SARS-COV-1 wtTMCT-DNA
89	PDI-SARS-COV-1 H5iTMCT-DNA
90	PDI-SARS-COV-1 H5iCT-DNA
91	PDI-SARS-COV-1 H5iCT(V4)-DNA
92	PDI-SARS-COV-1 H1cCT-DNA
	변형된 **SARS-COV-1 S** 단백질 **(AA)**
93	PDI-SARS-COV-1 wtTMCT-AA
94	PDI-SARS-COV-1 H5iTMCT-AA
95	PDI-SARS-COV-1 H5iCT-AA
96	PDI-SARS-COV-1 H5iCT(V4)-AA
97	PDI-SARS-COV-1 H1cCT-AA
	변형된 **MERS S** 단백질 **(DNA)** 및 프라이머
98	IF(AvB+wtCT-MERS).r
99	IF(H1cCT-wtTM).r
100	IF(H5ITMCT).r
101	PDI-MERS-wtTMCT-DNA
102	PDI-MERS-H5iTMCT-DNA
103	PDI-MERS-H5iCT-DNA
104	PDI-MERS-H5iCT(V4)-DNA
105	PDI-MERS-H1cCT-DNA
	변형된 **SARS-COV-1 S** 단백질 **(AA)**
106	PDI-MERS-wtTMCT-AA
107	PDI-MERS-H5iTMCT-AA
108	PDI-MERS-H5iCT-AA
109	PDI-MERS-H5iCT(V4)-AA
110	PDI-MERS-H1cCT-AA
	다른 서열
111	클로닝 벡터 7147 (좌측 T-DNA에서 우측 T-DNA)
112	천연 SARS-CoV-1 S 단백질 wtTM/CT AA (P59594)
113	천연 MERS S 단백질 wtTM/CT AA (AFY13307)

[0382]

서열 번호:	서열 설명
114	신호 펩티드를 포함하지 않는 천연 SARS-CoV-1 S 단백질 wtTM/CT AA (P59594)
115	신호 펩티드를 포함하지 않는 천연 MERS S 단백질 wtTM/CT AA (AFY13307)
116	변형된 PDI-SARS-COV-1 wtTMCT의 TMCT 영역-AA
117	변형된 PDI-SARS-COV-1 H5iTMCT의 TMCT 영역-AA
118	변형된 PDI-SARS-COV-1 H5iCT의 TMCT 영역-AA
119	변형된 PDI-SARS-COV-1 H5iCT(V4)의 TMCT 영역-AA
120	변형된 PDI-SARS-COV-1 H1cCT의 TMCT 영역-AA
121	변형된 PDI-MERS-wtTMCT의 TMCT 영역-AA
122	변형된 PDI-MERS-H5iTMCT의 TMCT 영역-AA
123	변형된 PDI-MERS-H5iCT의 TMCT 영역-AA
124	변형된 PDI-MERS-H5iCT(V4)의 TMCT 영역-AA
125	변형된 PDI-MERS-H1cCT의 TMCT 영역-AA
126	변형된 PDI-S-단백질+H1 Cal의 TMCT 영역
127	변형된 PDI-S-단백질+H3 Minn의 TMCT 영역
128	변형된 PDI-S-단백질+H6 HK의 TMCT 영역
129	변형된 PDI-S-단백질+H7 Guangdong의 TMCT 영역
130	변형된 PDI-S-단백질+H9 HK의 TMCT 영역
131	변형된 PDI-S-단백질+ B/ Wash의 TMCT 영역
132	코로나바이러스 S 단백질의 TM 도메인의 컨센서스 서열
133	코로나바이러스 S 단백질의 TM 도메인의 컨센서스 서열
134	개재 펩티드 서열 X_n을 포함하는 변형된 SARS-CoV-1 S 단백질의 TM/CT 영역
135	개재 펩티드 서열 X_n을 포함하는 변형된 MERS S 단백질의 TM/CT 영역
	변형된 OC43-CoV S 단백질 (DNA) 및 프라이머
136	IF(AvB+wtCT-OC43).r
137	PDI-OC43-wtTMCT-DNA
138	PDI-OC43-H5iTMCT-DNA
139	PDI-OC43-H5iCT-DNA
140	PDI-OC43-H5iCT(V4)-DNA
141	PDI-OC43-H1cCT-DNA
	변형된 OC43-CoV S 단백질 (AA)
142	PDI-OC43-wtTMCT-AA
143	PDI-OC43-H5iTMCT-AA
144	PDI-OC43-H5iCT-AA
145	PDI-OC43-H5iCT(V4)-AA
146	PDI-OC43-H1cCT-AA
	변형된 229E-CoV S 단백질 (DNA) 및 프라이머
147	IF(CoV229EwtCT).r
148	PDI-229E -wtTMCT-DNA
149	PDI-229E -H5iTMCT-DNA
150	PDI-229E -H5iCT-DNA
151	PDI-229E -H5iCT(V4)-DNA
152	PDI-229E -H1cCT-DNA
	변형된 229E-CoV S 단백질 (AA)
153	PDI-229E -wtTMCT-AA
154	PDI-229E -H5iTMCT-AA
155	PDI-229E -H5iCT-AA
156	PDI-229E -H5iCT(V4)-AA

[0383]

서열 번호:	서열 설명
157	PDI-229E-H1cCT-AA
	다른 서열
158	천연 OC43-CoV S 단백질 wtTM/CT AA (AVR40344)
159	천연 229E S 단백질 wtTM/CT AA (P15423)
160	신호 펩티드를 포함하지 않는 천연 OC43-CoV S 단백질 wtTM/CT AA (AVR40344)
161	신호 펩티드를 포함하지 않는 천연 229E S 단백질 wtTM/CT AA (P15423)
162	변형된 PDI-OC43-COV wtTMCT의 TMCT 영역-AA
163	변형된 PDI-OC43-COV H5iTMCT의 TMCT 영역-AA
164	변형된 PDI-OC43-COV H5iCT의 TMCT 영역-AA
165	변형된 PDI-OC43-COV H5iCT(V4)의 TMCT 영역-AA
166	변형된 PDI-OC43-COV H1cCT의 TMCT 영역-AA
167	변형된 PDI-229E-wtTMCT의 TMCT 영역-AA
168	변형된 PDI-229E-H5iTMCT의 TMCT 영역-AA
169	변형된 PDI-229E-H5iCT-AA의 TMCT 영역
170	변형된 PDI-229E-H5iCT(V4)의 TMCT 영역-AA
171	변형된 PDI-229E-H1cCT의 TMCT 영역-AA
172	개재 펩티드 서열 X_n을 포함하는 변형된 OC43-CoV S 단백질의 TM/CT 영역
173	개재 펩티드 서열 X_n을 포함하는 변형된 OC43-CoV S 단백질의 TM/CT 영역

[0384]

[0385] 본 발명은 하기 실시예에서 추가로 예시될 것이다.

[0386] **실시예**

[0387] 실시예 1: 변형된 구조 단백질 구성물의 제조

[0388] 당업계에 널리 공지된 기술을 이용하여 SARS-CoV-2 S 단백질 구성물을 제조하였다. 예를 들어, wtTMCT를 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 8586, 8589, 8591, 도 8a-8c 참조)을 하기 기술되는 바와 같이 클로닝하였다. 유사한 기술을 이용하여 H5iTMCT를 포함하는 SARS-COV-2 스파이크 단백질에 대한 구성물(구성물 번호 8592, 8595, 8597, 도 9a-9c 참조) 및 H5iCT를 포함하는 SARS-COV-2 스파이크 단백질에 대한 구성물(구성물 번호 8610, 8611, 8671, 도 10a-10c 참조)을 수득하였고, 서열 프라이머, 주형 및 생성물은 표 5에 기술되어 있다.

[0389] *wtTMCT를 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 8586, 8589, 8591)*

[0390] 알팔파 PDI 분비 신호 펩티드(PDISP)에 융합된, GSAS+K971P+V972P 엑토도메인 돌연변이를 포함하고, SARS-CoV-2로부터의 천연 막횡단 도메인 및 천연 세포질 테일(wtTMCT)을 포함하는 성숙 SARS-CoV-2 스파이크(S) 단백질 2를 코딩하는 서열(서열 번호 23)을 하기 PCR 기반 방법을 사용하여 3개의 상이한 발현 시스템으로 클로닝하였다. 주형으로서 wtTMCT를 포함하는 PDISP-SARS-COV-2 스파이크 단백질 유전자 서열(서열 번호 22)을 이용하여, 프라이머 IF(PDI)-CoV(opt2).c(서열 번호 24) 및 IF(AVB)-CoV(opt2).r(서열 번호 25)을 사용하여 SARS-COV-2 스파이크 단백질(wtTMCT) 코딩 서열을 함유하는 단편을 증폭시켰다. 인-퓨전(In-Fusion) 클로닝 시스템(Clontech: 미국 캘리포니아주 마운틴 뷰 소재)을 이용하여 PCR 생성물을 3개의 상이한 발현 시스템으로 클로닝하였다.

[0391] 제1 발현 시스템의 경우, 구성물 번호 8501(도 7a)을 AatII 및 StuI 제한 효소로 분해하고, 제1 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 8501은 2X35S(+C)/nbMT78/PDI/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 31에 제시되어 있다. 생성된 구성물은 번호 8586(서열 번호 32)으로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 SARS-COV-2로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 23에 제시되어 있다. 플라스미드 8586의 대표도는 도 8a에 제시되어 있다.

[0392] 제2 발현 시스템의 경우, 구성물 번호 8500(도 7b) 또한 AatII 및 StuI 제한 효소로 분해하고, 제2 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 8500은 2X35S(+C)/nbCSY65/PDI/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 33에 제시되어 있다. 생성된 구성물은 번호 8589(서열 번호 34)로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 SARS-COV-2로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 23에 제시되어 있다. 플라스미드 8589의 대표도는 도 8b에 제시되어 있다.

[0393] 제3 발현 시스템의 경우, 구성물 번호 8716(도 7c) 또한 AatII 및 StuI 제한 효소로 분해하고, 제3 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 8716은 2X35S(+C)/nbHEL40/PDI/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 35에 제시되어 있다. 생성된 구성물은 번호 8591(서열 번호 36)로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 SARS-COV-2로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 23에 제시되어 있다. 플라스미드 8591의 대표도는 도 8c에 제시되어 있다.

[0394] *H5iTMCT를 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 8592, 8595, 8597)*

[0395] GSAS+K971P+V972P 엑토도메인 돌연변이를 포함하고, H5 A/Indonesia/5/05 HA로부터의 막횡단 도메인 및 세포질 테일(H5iTMCT)을 포함하는, SARS-CoV-2로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열(서열 번호 27)을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 기술된 동일한 3개의 발현

시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 상기 기술된 것과 유사한 기술을 이용하여, 수용자 구성물 8501로부터 구성물 번호 8592(도 9a)를 유도하였고, 수용자 구성물 8500으로부터 구성물 번호 8595(도 9b)를 유도하였고, 수용자 구성물 8716으로부터 구성물 번호 8597(도 9c)을 유도하였고, 프라이머, 주형 및 생성물은 하기 표 5에 제공되어 있다.

[0396] *H5iCT를 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 8610, 8611, 8671)*

[0397] GSAS+K971P+V972P 엑토도메인 돌연변이를 포함하고, H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT)을 포함하는, SARS-CoV-2로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열(서열 번호 30)을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 기술된 동일한 3개의 발현 시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 상기 기술된 것과 유사한 기술을 이용하여, 수용자 구성물 8501로부터 구성물 번호 8610(도 10a)를 유도하였고, 수용자 구성물 8500으로부터 구성물 번호 8611(도 10b)를 유도하였고, 수용자 구성물 8716으로부터 구성물 번호 8671(도 10c)을 유도하였고, 프라이머, 주형 및 생성물은 하기 표 5에 제공되어 있다.

[0398] *대안적 TM CT 융합 서열을 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 8980, 8981, 8982, 8983)*

[0399] 서열 번호 19에 제시된 바와 같은, GSAS+K971P+V972P 엑토도메인 돌연변이를 포함하고, H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT)을 포함하는, SARS-CoV-2로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 구성물 8671에 대해 기술된 바와 같이 동일한 발현 시스템으로 클로닝하여 구성물 8980을 수득하였다(도 12a). 서열 번호 37(구성물 8981, 도 12b), 서열 번호 38(구성물 8982, 도 12c), 및 서열 번호 39(구성물 8983, 도 12d)에 제시된 바와 같은, GSAS+K971P+V972P 엑토도메인 돌연변이를 포함하고, H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT)을 포함하는, SARS-CoV-2로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열에 대하여 유사한 구성물을 생성하였다.

[0400] *다른 HA 균주로부터의 CT를 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 7390, 7391, 7392, 7393, 7394, 및 7395)*

[0401] GSAS+K971P+V972P 엑토도메인 돌연변이을 포함하고, H1 A/California/7/2009 HA로부터의 세포질 테일(H1CT)을 포함하는, SARS-CoV-2로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 구성물 8671에 대해 기술된 바와 같이 동일한 발현 시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 따라서, 생성된 구성물 7390은 H1 A/California/7/2009 HA 세포질 테일(H1CT)을 포함하는 변형된 S 단백질을 코딩한다(도 13a). H3 A/Minnesota/41/2019(구성물 7391, H3 CT)(도 13b), H6 A/Teal/Hong Kong/W312/97(구성물 7392, H6 CT)(도 13c), H7 A/Guangdong/17SF003/2016(구성물 7393, H7 CT)(도 13d), H9 A/Hong Kong/1073/99(구성물 7394, H9h CT)(도 13e) 또는 B/Washington/02/2019(구성물 7395, HA B CT)(도 13f)에 대하여 유사한 구성물을 생성하였다.

[0402] *치환을 포함하는 SARS-COV-2 스파이크 단백질(구성물 번호 8933, 8960, 8947)*

[0403] 예컨대, R667G, R668S, R670S, F802P, A877P, A884P, A927P, K971P, V972P, 및 L923F와 같은, S 단백질 중의 돌연변이의 조합을 포함하는 변형된 SARS-CoV-2 S 단백질 구성물을 당업계에 널리 공지된 기술을 이용하여, 및 기본적으로 상기 기술된 바와 같이 제조하였다. 구성물은 하기 치환을 가진다: 구성물 8933: R667G, R668S, R670S, K971P, V972P 및 L923F("GSAS-2P-923"); 구성물 8960: R667G, R668S, R670S, F802P, A927P, K971P, V972P 및 L923F("GSAS-4P-923") 및 구성물 8947: R667G, R668S, R670S, F802P, A877P, A884P, A927P, K971P, V972P 및 L923F("GSAS-6P-923").

[0404] *wtTMCT 및 변형된 TMCT를 포함하는 SARS-COV-1 스파이크 단백질(구성물 번호 9231, 9232, 9233, 9234, 9235)*

[0405] 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된, R654A+K955P+V956P 엑토도메인 돌연변이를 포함하고, SARS-CoV-1로부터의 천연 막횡단 도메인 및 천연 세포질 테일(wtTMCT)을 포함하는 성숙 SARS-CoV-1 스파이크(S) 단백질을 코딩하는 서열(서열 번호 88)을 PCR 기반 방법에 의해 하기 발현 시스템으로 클로닝하였다. 주형으로서 wtTMCT를 포함하는 PDISP-SARS-COV-1 스파이크 단백질 유전자 서열(서열 번호 88)을 이용하여, 프라이머 IF(nbHEL40)-PDI.c(서열 번호 86) 및 IF(AvB+wtCT).r(서열 번호 87)을 사용하여 PDISP-SARS-COV-1 스파이크 단백질(wtTMCT) 코딩 서열을 함유하는 단편을 증폭시켰다. 인-퓨전 클로닝 시스템(Clontech: 미국 캘리포니아주 마운틴 뷰 소재)을 이용하여 PCR 생성물을 하기 발현 시스템으로 클로닝하였다.

[0406] 구성물 번호 7147(도 21)을 AatII 및 StuI 제한 효소로 분해하고, 제1 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 7147은 2X35S(+C)/nbHEL40/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 111에 제시되어 있다. 생성된 구성물은 번호 9231로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 SARS-COV-1로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 93에 제시되어 있다. 플라스미드 9231의 대표도는 도 18a에 제시되어 있다.

[0407] R654A+ K955P+V956P 엑토도메인 돌연변이, 및 i) H5 A/Indonesia/5/05 HA로부터의 막횡단 도메인 및 세포질 테일(H5iTMCT), ii) H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT 및 변이 H5iCT(V4)), 또는 iii) H1 A/California/7/2009 HA로부터의 세포질 테일(H1cCT)을 포함하는, SARS-CoV-1로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 구성물 9231에 대해 기술된 바와 같이 동일한 발현 시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 따라서, 생성된 구성물 9232, 9233, 9234, 9235는 H5 A/Indonesia/5/05 TMCT(H5iTMCT)를 포함하는 변형된 S 단백질(도 18b, 서열 번호 94), H5 A/Indonesia/5/05 CT(H5iCT)를 포함하는 변형된 SARS-COV-1 S 단백질(도 18c, 서열 번호 95), H5 A/Indonesia/5/05 CT 변이체(H5iCT(V4))를 포함하는 변형된 S 단백질(도 18d, 서열 번호 96), 또는 H1 A/California/7/2009 CT(H1cCT)를 포함하는 변형된 S 단백질(도 18e, 서열 번호 97)을 코딩한다.

[0408] *wtTMCT 및 변형된 TMCT를 포함하는 MERS-CoV 스파이크 단백질(구성물 번호 9246, 9247, 9249, 9250, 9251)*

[0409] 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된, R730A+R733G+ V1043P+L1044P 엑토도메인 돌연변이를 포함하고, MERS-CoV로부터의 천연 막횡단 도메인 및 천연 세포질 테일(wtTMCT)을 포함하는 성숙 SARS-CoV-1 스파이크(S) 단백질을 코딩하는 서열(서열 번호 101)을 PCR 기반 방법에 의해 하기 발현 시스템으로 클로닝하였다. 주형으로서 wtTMCT를 포함하는 PDISP-SARS-COV-1 스파이크 단백질 유전자 서열(서열 번호 101)을 이용하여, 프라이머 IF(nbHEL40)-PDI.c(서열 번호 86) 및 IF(AvB+wtCT-MERS).r(서열 번호 98)을 사용하여 PDISP-MERS-COV 스파이크 단백질(wtTMCT) 코딩 서열을 함유하는 단편을 증폭시켰다. 인-퓨전 클로닝 시스템(Clontech: 미국 캘리포니아주 마운틴 뷰 소재)을 이용하여 PCR 생성물을 하기 발현 시스템으로 클로닝하였다.

[0410] 구성물 번호 7147(도 21)을 AatII 및 StuI 제한 효소로 분해하고, 제1 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 7147은 2X35S(+C)/nbHEL40/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 111에 제시되어 있다. 생성된 구성물은 번호 9246으로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 MERS-COV로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 106에 제시되어 있다. 플라스미드 9246의 대표도는 도 20a에 제시되어 있다.

[0411] R730A+R733G+ V1043P+L1044P 엑토도메인 돌연변이, 및 i) H5 A/Indonesia/5/05 HA로부터의 막횡단 도메인 및 세포질 테일(H5iTMCT), ii) H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT 및 변이 H5iCT(V4)), 또는 iii) H1 A/California/7/2009 HA로부터의 세포질 테일(H1cCT)을 포함하는, MERS-CoV로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 구성물 9246에 대해 기술된 바와 같이 동일한 발현 시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 따라서, 생성된 구성물 9247, 9249, 9250, 9251은 H5 A/Indonesia/5/05 TMCT(H5iTMCT)를 포함하는 변형된 MERS-COV S 단백질(도 20b, 서열 번호 107), H5 A/Indonesia/5/05 CT(H5iCT)를 포함하는 변형된 S 단백질(도 20c, 서열 번호 108), H5 A/Indonesia/5/05 CT 변이체(H5iCT(V4))를 포함하는 변형된 S 단백질(도 20d, 서열 번호 109), 또는 H1 A/California/7/2009 CT(H1cCT)를 포함하는 변형된 S 단백질(도 20e, 서열 번호 110)을 코딩한다.

[0412] *wtTMCT 및 변형된 TMCT를 포함하는 OC43-CoV 스파이크 단백질(구성물 번호 9269, 9270, 9272, 9273 및 9274)*

[0413] 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된, R761G+R762G+R764G+R765S+A1077P+L1078P 엑토도메인 돌연변이를 포함하고, OC43-CoV로부터의 천연 막횡단 도메인 및 천연 세포질 테일(wtTMCT)을 포함하는 성숙 OC43-CoV 스파이크(S) 단백질을 코딩하는 서열(서열 번호 137)을 PCR 기반 방법에 의해 하기 발현 시스템으로 클로닝하였다.

주형으로서 wtTMCT를 포함하는 PDISP-OC43-COV 스파이크 단백질 유전자 서열(서열 번호 137)을 이용하여, 프라이머 IF(nbHEL40)-PDI.c(서열 번호 86) 및 IF(AvB+wtCT-OC43).r(서열 번호 136)을 사용하여 PDISP-OC43-COV 스파이크 단백질(wtTMCT) 코딩 서열을 함유하는 단편을 증폭시켰다. 인-퓨전 클로닝 시스템(Clontech: 미국 캘리포니아주 마운틴 뷰 소재)을 이용하여 PCR 생성물을 하기 발현 시스템으로 클로닝하였다.

[0414] 구성물 번호 7147(도 21)을 AatII 및 StuI 제한 효소로 분해하고, 제1 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 7147은 2X35S(+C)/nbHEL40/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 111에 제시되어 있다. 생성된 구성물은 번호 9269로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 OC43-COV로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 142에 제시되어 있다. 플라스미드 9269의 대표도는 도 24a에 제시되어 있다.

[0415] R761G+R762G+R764G+R765S+A1077P+L1078P 엑토도메인 돌연변이, 및 i) H5 A/Indonesia/5/05 HA로부터의 막횡단 도메인 및 세포질 테일(H5iTMCT), ii) H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT 및 변이 H5iCT(V4)), 또는 iii) H1 A/California/7/2009 HA로부터의 세포질 테일(H1cCT)을 포함하는, OC43-CoV로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 구성물 9269에 대해 기술된 바와 같이 동일한 발현 시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 따라서, 생성된 구성물 9270, 9272, 9273 및 9274는 H5 A/Indonesia/5/05 TMCT(H5iTMCT)를 포함하는 변형된 OC43-COV S 단백질(도 24b, 서열 번호 143), H5 A/Indonesia/5/05 CT(H5iCT)를 포함하는 변형된 S 단백질(도 24c, 서열 번호 144), H5 A/Indonesia/5/05 CT 변이체(H5iCT(V4))를 포함하는 변형된 S 단백질(도 24d, 서열 번호 145), 또는 H1 A/California/7/2009 CT(H1cCT)를 포함하는 변형된 S 단백질(도 24e, 서열 번호 146)을 코딩한다.

[0416] *wtTMCT 및 변형된 TMCT를 포함하는 229E-CoV 스파이크 단백질(구성물 번호 9310, 9311, 9312, 9313 및 9314)*

[0417] 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된, R567A+T871P+I872P 엑토도메인 돌연변이를 포함하고, 229E-CoV로부터의 천연 막횡단 도메인 및 천연 세포질 테일(wtTMCT)을 포함하는 성숙 229E-CoV 스파이크(S) 단백질을 코딩하는 서열(서열 번호 148)을 PCR 기반 방법에 의해 하기 발현 시스템으로 클로닝하였다. 주형으로서 wtTMCT를 포함하는 PDISP-OC43-COV 스파이크 단백질 유전자 서열(서열 번호 148)을 이용하여, 프라이머 IF(nbHEL40)-PDI.c(서열 번호 86) 및 IF(CoV229EwtCT).r(서열 번호 147)을 사용하여 PDISP-229E-COV 스파이크 단백질(wtTMCT) 코딩 서열을 함유하는 단편을 증폭시켰다. 인-퓨전 클로닝 시스템(Clontech: 미국 캘리포니아주 마운틴 뷰 소재)을 이용하여 PCR 생성물을 하기 발현 시스템으로 클로닝하였다.

[0418] 구성물 번호 7147(도 21)을 AatII 및 StuI 제한 효소로 분해하고, 제1 인-퓨전 어셈블리 반응을 위해 선형화된 플라스미드를 이용하였다. 구성물 번호 7147은 2X35S(+C)/nbHEL40/AvB/NOS 기반 발현 카세트에서 관심 유전자의 "인 퓨전" 클로닝을 위해 의도된 수용자 플라스미드이다. 상기 수용자 플라스미드는 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결인자하에 TBSV P19 침묵 억제인자의 공동 발현을 위한 유전적 구성물을 포함한다. 백본은 pCAMBIA 바이너리 플라스미드이고, 좌측에서 우측으로 t-DNA 경계의 서열이 서열 번호 111에 제시되어 있다. 생성된 구성물은 번호 9310으로 제공된다. 알팔파 PDI 분비신호 펩티드(PDISP)에 융합된 229E-COV로부터의 성숙 스파이크 단백질의 아미노산 서열은 서열 번호 153에 제시되어 있다. 플라스미드 9310의 대표도는 도 26a에 제시되어 있다.

[0419] R567A+T871P+I872P 엑토도메인 돌연변이, 및 i) H5 A/Indonesia/5/05 HA로부터의 막횡단 도메인 및 세포질 테일(H5iTMCT), ii) H5 A/Indonesia/5/05 HA로부터의 세포질 테일(H5iCT 및 변이 H5iCT(V4)), 또는 iii) H1 A/California/7/2009 HA로부터의 세포질 테일(H1cCT)을 포함하는, 229E-CoV로부터의 성숙 스파이크(S) 단백질을 코딩하는 서열을 알팔파 PDI 분비신호 펩티드(PDISP)에 융합시키고, 유사한 PCR 기반 방법에 의해 상기 구성물 9310에 대해 기술된 바와 같이 동일한 발현 시스템으로 클로닝하였다(프라이머에 대해 표 5 및 사용된 서열에 대해 실시예 3 참조). 따라서, 생성된 구성물 9311, 9312, 9313 및 9314는 H5 A/Indonesia/5/05 TMCT(H5iTMCT)를 포함하는 변형된 229E-COV S 단백질(도 26b, 서열 번호 154), H5 A/Indonesia/5/05 CT(H5iCT)를 포함하는 변형된 S 단백질(도 26c, 서열 번호 155), H5 A/Indonesia/5/05 CT 변이체(H5iCT(V4))를 포함하는 변형된 S 단백질(도 26d, 서열 번호 156), 또는 H1 A/California/7/2009 CT(H1cCT)를 포함하는 변형된 S 단백질(도 26e, 서열 번호 157)을 코딩한다.

[0420]

표 5. 구성물 번호, 프라이머 및 주형 목록

구성물 #	5' UTR	관심 유전자 (GOI)	TMCT 영역	프라이머 1 (단편 1의 정방향 프라이머)	프라이머 2 (단편 1의 역방향 프라이머)	제1 PCR용 주형/ 생성된 유전자	생성된 단백질	수용자 플라스미드
8586	nbMT78	SARS-CoV-2 S	wt TMCT	서열 번호 24	서열 번호 25	서열 번호 22	서열 번호 23	서열 번호 31
8589	nbCSY65	SARS-CoV-2 S	wt TMCT	서열 번호 24	서열 번호 25	서열 번호 22	서열 번호 23	서열 번호 33
8591	nbHEL40	SARS-CoV-2 S	wt TMCT	서열 번호 24	서열 번호 25	서열 번호 22	서열 번호 23	서열 번호 35
8592	nbMT78	SARS-CoV-2 S	H5iTMCT	서열 번호 24	서열 번호 28	서열 번호 26	서열 번호 27	서열 번호 31
8595	nbCSY65	SARS-CoV-2 S	H5iTMCT	서열 번호 24	서열 번호 28	서열 번호 26	서열 번호 27	서열 번호 33
8597	nbHEL40	SARS-CoV-2 S	H5iTMCT	서열 번호 24	서열 번호 28	서열 번호 26	서열 번호 27	서열 번호 35
8610	nbMT78	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 29	서열 번호 30	서열 번호 31
8611	nbCSY65	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 29	서열 번호 30	서열 번호 33
8671	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 29	서열 번호 30	서열 번호 35
8953	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 65	서열 번호 48	서열 번호 35
8940	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 66	서열 번호 49	서열 번호 35
8933	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 67	서열 번호 50	서열 번호 35
8960	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 68	서열 번호 51	서열 번호 35
8947	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT	서열 번호 24	서열 번호 28	서열 번호 69	서열 번호 52	서열 번호 35
8980	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT (V1)	서열 번호 24	서열 번호 28	서열 번호 70	서열 번호 59	서열 번호 35
8981	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT (V2)	서열 번호 24	서열 번호 28	서열 번호 71	서열 번호 60	서열 번호 35
8982	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT (V3)	서열 번호 24	서열 번호 28	서열 번호 72	서열 번호 61	서열 번호 35
8983	nbHEL40	SARS-CoV-2 S	wtTM/H5iCT (V4)	서열 번호 24	서열 번호 28	서열 번호 73	서열 번호 62	서열 번호 35
7390	nbHEL40	SARS-CoV-2 S	wtTM/H1cCT	서열 번호 24	서열 번호 80	서열 번호 74	서열 번호 53	서열 번호 35
7391	nbHEL40	SARS-CoV-2 S	wtTM/H3mCT	서열 번호 24	서열 번호 81	서열 번호 75	서열 번호 54	서열 번호 35
7392	nbHEL40	SARS-CoV-2 S	wtTM/H6hkCT	서열 번호 24	서열 번호 82	서열 번호 76	서열 번호 55	서열 번호 35
7393	nbHEL40	SARS-CoV-2 S	wtTM/H7gCT	서열 번호 24	서열 번호 83	서열 번호 77	서열 번호 56	서열 번호 35
7394	nbHEL40	SARS-CoV-2 S	wtTM/H9hkCT	서열 번호 24	서열 번호 84	서열 번호 78	서열 번호 57	서열 번호 35
7395	nbHEL40	SARS-CoV-2 S	wtTM/BwCT	서열 번호 24	서열 번호 85	서열 번호 79	서열 번호 58	서열 번호 35
9231	nbHEL40	SARS-COV-1 S	wtTMCT	서열 번호 86	서열 번호 87	서열 번호 88	서열 번호 93	서열 번호 111
9232	nbHEL40	SARS-COV-1 S	H5iTMCT	서열 번호 86	서열 번호 28	서열 번호 89	서열 번호 94	서열 번호 111
9233	nbHEL40	SARS-COV-1 S	wtTM/H5iCT	서열 번호 86	서열 번호 28	서열 번호 90	서열 번호 95	서열 번호 111

구성물 #	5' UTR	관심 유전자 (GOI)	TMCT 영역	프라이머 1 (단편 1의 정방향 프라이머)	프라이머 2 (단편 1의 역방향 프라이머)	제1 PCR용 주형/ 생성된 유전자	생성된 단백질	수용자 플라스미드
9234	nbHEL40	SARS-COV-1 S	wtTM/H5iCT (V4)	서열 번호 86	서열 번호 28	서열 번호 91	서열 번호 96	서열 번호 111
9235	nbHEL40	SARS-COV-1 S	wtTM/H1cCT	서열 번호 86	서열 번호 80	서열 번호 92	서열 번호 97	서열 번호 111
9246	nbHEL40	MERS S	wtTMCT	서열 번호 86	서열 번호 98	서열 번호 101	SEQ ID NO: 106	서열 번호 111
9247	nbHEL40	MERS S	H5iTMCT	서열 번호 86	서열 번호 100	서열 번호 102	서열 번호 107	서열 번호 111
9249	nbHEL40	MERS S	wtTM/H5iCT	서열 번호 86	서열 번호 100	서열 번호 103	서열 번호 108	서열 번호 111
9250	nbHEL40	MERS S	wtTM/H5iCT (V4)	서열 번호 86	서열 번호 100	서열 번호 104	서열 번호 109	서열 번호 111
9251	nbHEL40	MERS S	wtTM/H1cCT	서열 번호 86	서열 번호 99	서열 번호 105	서열 번호 110	서열 번호 111
9269	nbHEL40	OC43 S	wtTMCT	서열 번호 86	서열 번호 136	서열 번호 137	서열 번호 142	서열 번호 111
9270	nbHEL40	OC43 S	H5iTMCT	서열 번호 86	서열 번호 100	서열 번호 138	서열 번호 143	서열 번호 111
9272	nbHEL40	OC43 S	wtTM/H5iCT	서열 번호 86	서열 번호 100	서열 번호 139	서열 번호 144	서열 번호 111
9273	nbHEL40	OC43 S	wtTM/H5iCT (V4)	서열 번호 86	서열 번호 100	서열 번호 140	서열 번호 145	서열 번호 111
9274	nbHEL40	OC43 S	wtTM/H1cCT	서열 번호 86	서열 번호 99	서열 번호 141	서열 번호 146	서열 번호 111
9310	nbHEL40	229E S	wtTMCT	서열 번호 86	서열 번호 147	서열 번호 148	서열 번호 153	서열 번호 111
9311	nbHEL40	229E S	H5iTMCT	서열 번호 86	서열 번호 100	서열 번호 149	서열 번호 154	서열 번호 111
9312	nbHEL40	229E S	wtTM/H5iCT	서열 번호 86	서열 번호 100	서열 번호 150	서열 번호 155	서열 번호 111
9313	nbHEL40	229E S	wtTM/H5iCT (V4)	서열 번호 86	서열 번호 100	서열 번호 151	서열 번호 156	서열 번호 111
9314	nbHEL40	229E S	wtTM/H1cCT	서열 번호 86	서열 번호 99	서열 번호 152	서열 번호 157	서열 번호 111

[0421]

[0422] 실시예 2: 방법

[0423] 아그로박테리움 투메파시엔스(*Agrobacterium tumefaciens*) 형질감염

[0424] 아그로박테리움 투메파시엔스 균주 AGL1을 문헌 [D'Aoust *et al.*, 2008 (*Plant Biotech. J.* 6:930-40)]에 기술된 방법을 이용하여 SARS-CoV-2 변형된 S 단백질 발현 벡터로 전기천공하여 형질감염시켰다. 형질감염된 아그로박테리움을 10 mM 2-(N-모르폴리노)에탄술폰산(MES), 20 μM 아세토시린곤, 50 μg/ml 카나마이신 및 25 μg/ml 카르베니실린(pH 5.6)으로 보충된 YEB 배지에서 OD_{600} 0.6 내지 1.6까지 성장시켰다. 아그로박테리움 현탁액을 사용 전에 원심분리하고, 침투 배지(10 mM $MgCl_2$ 및 10 mM MES pH 5.6) 중에 재현탁시켰다.

[0425] 식물 바이오매스, 접종물 및 아그로침투 준비

[0426] N. 벤타미아나(*N. benthamiana*) 식물을 시판용 피트 모스 배지로 채워진 평지에서 종자로부터 재배하였다. 식물을 16/8 광주기 및 25℃ 주/20℃ 야인 온도상에서 온실에서 성장시켰다. 시딩 3주 후, 개별 묘목을 골라 화분에 이식하고, 동일한 환경 조건하에서 추가로 3주 동안 온실에서 자라도록 두었다.

[0427] 각 발현 벡터로 형질감염된 아그로박테리아를 OD_{600}이 0.6 내지 1.6에 도달할 때까지 10 mM 2-(N-모르폴리노)에탄술폰산 (MES), 20 μM 아세토시린곤, 50 μg/ml 카나마이신 및 25 μg/ml의 카르베니실린(pH 5.6)으로 보충된 YEB 배지 중에서 성장시켰다. 아그로박테리움 현탁액을 사용 전에 원심분리하고, 침투 배지(10 mM $MgCl_2$ 및 10

mM MES pH 5.6) 중에 재현탁시키고, 4℃에서 밤새도록 보관하였다. 침투 당일, 배양 배치를 2.5 배양 부피로 희석하고, 사용 전에 가온시켰다. N. 벤타미아나 전체 식물을 2 min 동안 20-40 Torr의 진공하에 밀폐된 스테인리스 스틸 탱크의 박테리아 현탁액에 뒤집어서 거꾸로 두었다. 식물을 수확할 때까지 6일 또는 9일의 인큐베이션 기간 동안 온실로 돌려보냈다.

[0428] 잎 수확 및 총 단백질 및 VLP 추출

[0429] 인큐베이션 후, 식물의 지상부를 수확하고, -80℃에서 동결시키고, 조각으로 분쇄하였다. 전체 가용성 단백질을 pH 8.0 + 500 mM NaCl, 0.4 μg/ml 메타비술파이트 및 1 mM 페닐메탄술포닐 플루오라이드에서 2 부피의 냉 50 mM 트리스 완충제 중에서 냉동 분쇄된 식물 물질의 각 샘플을 기계적으로 균질화(Polytron)하여 추출하였다. 균질화 후, 슬러리를 4℃에서 10 min 동안 10,000 g에서 원심분리하고, 상기 정화된 미정제 추출물(상청액)을 분석을 위해 보관하였다.

[0430] 참조 표준으로서 소 혈청 알부민을 사용하여 브래드포드 검정법(Bradford assay)(Bio-Rad: 미국 캘리포니아주 허큘리스 소재)에 의해 정화된 미정제 추출물의 전체 단백질 함량을 측정하였다. 크리테리온™ TGX 스테인-프리™(Criterion™ TGX Stain-Free™) 프리캐스트 겔(Bio-Rad Laboratories: 미국 캘리포니아주 허큘리스 소재)을 이용하여 환원 조건하에서 SDS-PAGE에 의해 단백질을 분리하였다. 쿠마시 브릴리언트 블루(Coomassie Brilliant Blue)로 겔을 염색하여 단백질을 시각화하였다. 대안적으로, 단백질을 Gel Doc™ EZ 이미징 시스템(Bio-Rad Laboratories: 미국 캘리포니아주 허큘리스 소재)으로 시각화하고, 면역검출을 위해 폴리비닐렌 디플루오라이드(PVDF) 막(Roche Diagnostics Corporation: 미국 인디애나주 인디애나폴리스 소재) 상에 전기전달하였다. 면역블롯 전에, 막을 4℃에서 16-18 h 동안 5% 탈지유 및 트리스-완충처리된 염수 중 0.1% 트윈-20(TBS-T)으로 차단시켰다.

[0431] VLP 정제를 위해, 2 부피의 추출 완충제(50 mM 트리스 완충제(pH 7.0) + 500 mM NaCl)를 포함하는 블렌더를 사용하여 기계적 추출에 의해 동결된 바이오매스로부터 단백질을 추출하고, 0.5 M 시트르산을 사용하여 pH를 6.1로 강하시켰다. 슬러리를 큰 공극 나일론 필터를 통해 여과하여 큰 파편을 제거하고, 4℃에서 5 min 동안 5000 g로 원심분리하였다. 상청액을 수집하고, 30 min 동안 (4℃) 5000 g로 다시 원심분리하여 추가 파편을 제거하고, 정화 필터를 통과시켰다. 이어서, 상청액을 불연속 이오딕사놀 밀도 구배에 로딩하였다. 분석 밀도 구배 원심분리를 하기와 같이 수행하였다: 트리스 완충제 중 불연속 이오딕사놀 밀도 구배(35%의 이오딕사놀 3 ml, 30%의 이오딕사놀 3 ml, 25%의 이오딕사놀 3 ml, 15%의 이오딕사놀 3 ml 및 10%의 이오딕사놀 5 ml)를 포함하는 38 mL 튜브를 준비하고, 바이러스 유사 입자를 함유하는 22 ml의 추출물과 오버레이시켰다. 구배를 2시간 동안 (4℃) 동안 120 000 g로 원심분리하였다. 원심분리 후, 1 mL 분획을 아래에서 위로 수집하고, 단백질 염색 또는 웨스턴 블롯과 함께 조합된 SDS-PAGE에 의해 분획을 분석하였다. 분획 6 내지 9를 풀링하고, 아미콘(Amicon) 원심분리 장치를 사용하여 완충제를 교환하였다. 단백질 함량은 브래드포드 검정법으로 측정한다.

[0432] 단백질 분석 및 면역블롯팅

[0433] 면역블롯팅은 TBS-트윈 20 0.1% 중 2% 탈지유에 희석된 1차 mAb(항S1, Sino Biological, 카탈로그 번호 40150-R007 또는 항S2, Novus biological, 카탈로그 번호 NB100-56578)를 이용하여 제1 인큐베이션하는 동안 수행하였다. TBS-트윈 20 0.1% 중 2% 탈지유 중에서 화학발광 검출을 위한 2차 항체로서 퍼옥시다제 접합된 염소 항토끼(Jackson Immunoresearch, 카탈로그 번호 115-035-144)를 이용하였다. 기질로서 루미놀(Roche Diagnostics Corporation)을 이용하여 화학발광에 의해 면역반응성 복합체를 검출하였다. EZ-링크 플러스® 활성화 퍼옥시다제(EZ-Link Plus® Activated Peroxidase) 접합 키트(Pierce: 미국 일리노이주 록퍼드 소재)를 이용하여 인간 IgG 항체의 호스래디쉬 퍼옥시다제 효소 접합을 수행하였다.

[0434] 초목에서, 수율을 정화된 미정제 추출물에서 평가하고, 모세관 기반 전기영동 방법(Protein Simple, BioTechne) 기술 및 WES 분석 시스템을 이용하여 분석하였다. 간략하면, 미정제 추출물로부터의 가용성 단백질을 모세관에서 분자량별로 분리하고, 매트릭스에 고정시켰다. 정제된 VLP를 사용하는 표준 곡선을 사용하여 S 단백질 양을 측정하고, 항S2 항체(Novus biological, 카탈로그 번호 NB100-56578)를 제조업체 지침에 따라 검출에 사용하였다. 이어서, 1로 설정된 비교물질 구성물을 이용하여 수율을 정규화한다.

[0435] SARS-CoV S 단백질 검출을 위한 1차 항체는 시노 바이오로지칼즈로부터의 SARS-CoV 스파이크 S1 서브유닛 항체(40150-MM08(1/5000))였고, 검출을 위해 사용된 2차 항체는 염소 항마우스(JIR, 115-035-146(1/10000))였다. MERS CoV S 단백질 검출을 위한 1차 항체는 시노 바이오로지칼로부터의 MERS-CoV 스파이크 단백질 S1 항체(N 말단)(100208-RP02, 1/5000)였다. 검출을 위해 사용된 2차 항체는 JIR로부터의 염소 항마우스(115-035-144, 1/10

000)였다. 검출을 위해 사용된 1차 항체는 안티바이즈-온라인으로부터의 항코로나바이러스 OC43 스파이크 단백질(ABIN2754654, 1/1000)이었다. 검출을 위해 사용된 2차 항체는 JIR로부터의 염소 항토끼(111-035-144, 1/10000)였다.

[0436] 전자 현미경법

[0437] 발현된 S 단백질이 VLP로 어셈블리되었는지 여부를 결정하기 위해, 면역 포획된 입자의 투과 전자 현미경(TEM: transmission electron microscopy)을 정제된 VLP에 대해 수행하였다. 글로우 방전 탄소/구리 그리드(10s, 0.3 mbar)를 5 min 동안 20 ㎕의 정제된 VLP(100 ㎍/ml)에 놓은 후, 멸균 증류수로 4회 세척하였다. 그리드를 20 ㎕의 2% 우라닐 아세테이트에 1 min 동안 플로팅한 후, 습한 여과지를 터치하여 과량의 용액을 제거하고, TEM(Tecnai Microscope)에서 관측하기 전에 여과지에서 24h 동안 건조시킨다.

[0438] 실시예 3: 서열

[0439] 하기 서열을 상기 기술된 실시예에서 사용하였다.

천연 SARS-CoV-2 S 단백질 wtTM/CT AA (P0DTC2) (서열 번호 1)

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWF
HAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV
CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFK
NIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTA
GAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTES
IVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDL
CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRL
FRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHA
PATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL
DITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCL
IGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIP
TNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE
VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIA
ARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISS
VLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRV
DFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVT
QRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDIS
GINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLC
CMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

신호 펩티드(SP)를 포함하지 않는 천연 SARS-CoV-2 S 단백질 wtTM/CT AA (P0DTC2)

(서열 번호 2)

VNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFD
NPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYY

[0440]

HKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTF
LLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGE
VFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGD
EVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDIST
EIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVK
NKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIP
IGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVS
MTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTV
LPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIAN
QFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQS
APHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDR
LNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC
GSCCKFDEDDSEPVLKGVKLHYT

H5 A/Indonesia/5/05 헤마글루티닌 (HA) AA (A5A5L7) (서열 번호 3)

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDLDGVKP
LILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELKHLLSRINHFEKI
QIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNTNQEDLLVLWGIH
HPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEFFWTILKPNDAINFESN
GNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKSN
RLVLATGLRNSPQRESRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKA
IDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLD
FHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEIS
GVKLESIGTYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

H5 A/Indonesia/5/05 헤마글루티닌 (HA) 바이러스 cDNA (EF541394.1) (서열 번호 4)

CTGTAAAAATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATT
TGCATTGGTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTAC
TGTTACACATGCCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAG
TGAAGCCTCTAATTTTAAGAGATTGTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGAC
GAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTG
TTACCCAGGGAGTTTCAACGACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTG
AGAAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTGAGCTCA
GCATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAAGAACAG
TACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAAGATCTTTTGGTACTGTGGG
GAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAAAACCCAACCACCTATATT
TCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAATAGCTACTAGATCCAAAGTAAA
CGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAATTTTAAAACCTAATGATGCAATCAACTTCG
AGAGTAATGGAAATTTCATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGCA
ATTATGAAAAGTGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGAT
AAACTCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGA
AATCAAACAGATTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAGAAAA
AAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGG
TTGGTATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTC
AAAAGGCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTT

[0441]

GAGGCCGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGA
AGACGGGTTTCTAGATGTCTGGACTTATAATGCCGAACTTCTGGTTCTCATGGAAAATGAGAGAA
CTCTAGACTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTAGGGAT
AATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACAAATGTGATAATGAATGTATGGA
AAGTATAAGAAACGGAACGTACAACTATCCGCAGTATTCAGAAGAAGCAAGATTAAAAAGAGAGG
AAATAAGTGGGGTAAAATTGGAATCAATAGGAACTTACCAAATACTGTCAATTTATTCAACAGTG
GCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATC
GTTACAATGCAGAATTTGCATTTAAATTTGTGAGTTCAG

H5 A/Indonesia/5/05 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 AA (서열 번호 5)

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWF
HAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKV
CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFK
NIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTA
GAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTES
IVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDL
CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRL
FRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHA
PATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL
DITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCL
IGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIP
TNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQE
VFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIA
ARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISS
VLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRV
DFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVT
QRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDIS
GINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLS
LWMCSNGSLQCRICI

H1 A/California/7/2009 헤마글루티닌 TM/CT AA (서열 번호 6)

IDGVKLESTRIYQILAIYSTVASSLVLVVSLGAISFWMCSNGSLQCRICI

H2 A/Singapore/1/1957 헤마글루티닌 TM/CT AA (서열 번호 7)

IKGVKLSSMGVYQILAIYATVAGSLSLAIMMAGISFWMCSNGSLQCRICI

H3 A/Minnesota/41/2019 헤마글루티닌 TM/CT AA (서열 번호 8)

IKGVELKSGYKDWILWISFAISCFLLCVALLGFIMWACQKGNIRCNICI

H5 A/Indonesia/5/05 헤마글루티닌 TM/CT AA (서열 번호 9)

ISGVKLESIGTYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

H6 A/Teal/Hong Kong/W312/97 헤마글루티닌 TM/CT AA (서열 번호 10)

IESVKLENLGVYQILAIYSTVSSSLVLVGLIMAMGLWMCSNGSMQCRICI

[0442]

H7 A/Guangdong/17SF003/2016 헤마글루티닌 TM/CT AA (서열 번호 11)

IDPVKLSSGYKDVILWFSFGASCFILLAIVMGLVFICVKNGNMRCTICI

H9 A/Hong Kong/1073/99 헤마글루티닌 TM/CT AA (서열 번호 12)

IEGVKLESEGTYKILTIYSTVASSLVLAMGFAAFLFWAMSNGSCRCNICI

B/Washington/02/2019 헤마글루티닌 TM/CT AA (서열 번호 13)

AASLNDDGLDNHTILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSICL

인플루엔자 헤마글루티닌의 C 말단 영역의 컨센서스 서열 (서열 번호 14)

IXGVKLXSXGXYXILXIYSTVASSLXLXXXXXXXXXWMCSNGSXXCXICI

인플루엔자 헤마글루티닌의 CT 도메인의 컨센서스 서열 (서열 번호 15)

XXWMCSNGSXXCXICI

천연 SARS-CoV-2 S 단백질 wtTM/CT의 C 말단 영역 (서열 번호 16)

WYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVK
LHYT

H5 A/Indonesia/5/05 헤마글루티닌의 C 말단 영역 (서열 번호 17)

ISGVKLESIGTYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질의 C 말단 영역 (서열 번호 18)

WYIWLGFIAGLIAIVMVTIMLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질의 C 말단 영역,
변이 1 (서열 번호 19)

WYIWLGFIAGLIAIVMVTIMMAGLSLWMCSNGSLQCRICI

(SP 부재) SARS-CoV-2 S 단백질 GSAS+PP wtTM/CT AA (서열 번호 20)

VNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFD
NPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYY
HKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTF
LLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGE
VFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGD

[0443]

EVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDIST
EIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVK
NKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIP
IGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVS
MTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTV
LPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIAN
QFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQS
APHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDR
LNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSC
GSCCKFDEDDSEPVLKGVKLHYT

(SP 부재) H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA (서열 번호 21)

VNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFD
NPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYY
HKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTF
LLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGE
VFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGD
EVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDIST
EIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVK
NKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIP
IGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVS
MTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTV
LPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIAN
QFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQS
APHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDR
LNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLSLWMCSNGSLQCRICI

PDI-SARS-CoV-2 S 단백질 GSAS+PP wtTM/CT-DNA (서열 번호 22)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGAT
CTTCGCGGTGAATCTTACGACGCGAACACAGTTACCACCCGCATATACAAATAGCTTCACTCGGG
GTGTTTATTACCCCGACAAAGTGTTCAGGTCCTCCGTGCTCCACTCAACACAGGACCTCTTTCTT
CCTTTCTTTTCTAACGTGACATGGTTTCATGCCATTCATGTATCCGGCACTAACGGTACTAAGAG
GTTCGATAATCCTGTGCTCCCTTTCAATGACGGCGTTTACTTTGCAAGCACAGAGAAGAGTAACA
TCATCCGAGGTTGGATCTTTGGCACTACCCTCGATTCAAAGACGCAGAGCCTCCTCATTGTGAAC
AATGCCACTAACGTGGTGATCAAAGTTTGCGAGTTTCAGTTCTGCAATGACCCTTTCTTGGGGGT
GTACTATCATAAGAACAACAAGTCTTGGATGGAATCTGAATTCCGCGTCTATAGCAGCGCCAACA
ACTGCACCTTTGAATACGTGTCCCAGCCCTTCCTTATGGACCTGGAGGGAAAGCAGGGAAACTTT
AAGAATCTGAGAGAGTTCGTGTTTAAAAATATCGACGGCTATTTTAAGATCTATTCTAAGCACAC
GCCTATTAATCTCGTGCGCGATCTTCCACAAGGCTTCAGCGCCCTGGAACCACTCGTGGACCTCC
CAATTGGTATCAACATCACTAGATTTCAGACTCTGCTTGCCCTCCACCGATCCTATCTGACACCC

[0444]

GGAGACTCCTCTAGCGGCTGGACTGCCGGCGCTGCCGCTTATTACGTTGGTTATCTTCAGCCACG
CACGTTCCTGCTGAAGTATAACGAGAATGGTACTATTACCGATGCCGTGGATTGTGCCCTTGACC
CCCTGTCCGAAACTAAGTGCACACTCAAGTCATTCACTGTGGAAAAAGGAATCTACCAGACAAGC
AATTTTCGGGTCCAGCCTACTGAGAGCATTGTGCGCTTTCCTAACATCACAAATCTTTGCCCCTT
CGGAGAGGTTTTCAATGCTACACGGTTTGCCTCCGTGTATGCCTGGAACCGCAAGAGAATTTCCA
ATTGCGTGGCCGATTACTCCGTGCTCTACAATAGTGCAAGCTTTAGCACCTTTAAGTGCTATGGC
GTATCCCCTACTAAGCTTAACGACTTGTGTTTCACAAACGTGTATGCCGACTCCTTTGTGATACG
GGGCGACGAAGTTAGACAGATAGCACCAGGACAGACGGGAAAGATAGCTGACTACAACTATAAGC
TTCCTGATGACTTCACTGGCTGCGTTATCGCGTGGAATTCTAACAACCTGGACTCAAAAGTCGGC
GGCAACTATAACTATCTCTATCGGCTGTTCCGCAAGAGTAACCTTAAGCCCTTTGAGAGAGATAT
AAGCACTGAAATCTACCAGGCTGGCAGTACGCCCTGTAATGGCGTGGAAGGCTTTAATTGTTATT
TTCCACTGCAATCCTATGGTTTTCAGCCAACCAATGGCGTGGGCTACCAACCATACCGCGTCGTG
GTGCTCTCCTTTGAACTGCTCCACGCTCCCGCGACTGTCTGCGGCCCCAAGAAGTCCACGAACCT
TGTGAAGAATAAGTGCGTTAATTTTAATTTCAACGGCCTCACTGGAACAGGAGTGCTCACTGAGA
GTAACAAGAAGTTCCTGCCATTTCAACAATTTGGCAGAGACATAGCCGATACTACTGACGCCGTT
AGGGACCCCCAGACCCTCGAGATTCTCGATATAACGCCCTGCTCCTTCGGTGGAGTTTCCGTGAT
CACGCCAGGCACCAATACCAGTAACCAGGTCGCCGTGCTGTATCAGGATGTCAACTGTACTGAGG
TGCCCGTAGCCATCCATGCGGATCAGCTCACACCAACTTGGAGGGTGTACAGCACCGGCTCCAAT
GTATTCCAGACTCGGGCCGGATGCCTTATTGGCGCCGAACACGTGAACAATAGTTACGAATGCGA
TATTCCAATTGGCGCCGGAATCTGTGCTAGCTACCAGACTCAGACGAACTCCCCAGGCAGCGCCA
GCAGCGTTGCCAGCCAGTCAATCATCGCTTATACAATGTCACTTGGAGCCGAAAACTCCGTGGCT
TACTCAAACAACAGCATCGCCATCCCCACAAACTTCACCATATCCGTGACAACTGAGATTCTGCC
AGTGTCCATGACTAAGACGTCCGTAGATTGCACTATGTACATATGCGGCGACAGCACAGAATGTT
CTAATCTGCTGCTGCAATATGGAAGCTTCTGCACTCAACTGAACAGAGCGCTCACAGGCATCGCC
GTGGAGCAGGATAAGAATACCCAGGAGGTGTTCGCCCAAGTTAAGCAGATCTACAAGACCCCACC
CATAAAGGATTTCGGTGGATTCAATTTTAGTCAGATACTCCCAGACCCATCTAAGCCATCCAAGA
GGAGCTTTATCGAGGATCTTTTGTTTAACAAAGTTACTCTGGCCGACGCCGGTTTCATCAAGCAG
TACGGAGATTGCCTCGGCGACATCGCTGCTCGTGACCTCATCTGTGCGCAAAAGTTTAACGGTCT
GACGGTGCTGCCTCCCCTCCTTACTGATGAAATGATCGCCCAGTATACCAGCGCACTCCTCGCTG
GCACCATAACATCCGGTTGGACATTCGGCGCTGGTGCAGCACTGCAGATACCATTCGCCATGCAA
ATGGCATATCGTTTCAACGGTATCGGTGTCACACAGAATGTCCTATATGAGAACCAGAAGCTGAT
CGCAAATCAGTTCAATAGTGCCATCGGAAAAATCCAGGATAGCCTTAGCAGCACAGCCTCAGCCC
TTGGCAAACTCCAGGATGTCGTGAACCAGAATGCCCAGGCTCTCAATACCCTCGTGAAGCAGCTC
TCATCTAATTTCGGCGCAATTTCCAGTGTCCTCAACGACATCCTCAGCCGCCTCGACCCCCCCGA
GGCCGAAGTGCAGATTGACAGACTGATTACAGGTCGACTCCAGAGCCTCCAGACTTACGTGACTC
AGCAGCTGATAAGAGCCGCCGAGATAAGGGCCAGCGCTAACCTGGCTGCCACAAAGATGTCTGAG
TGCGTGCTGGGCCAGTCCAAGAGAGTAGACTTCTGTGGCAAAGGCTACCATCTGATGAGCTTCCC
ACAATCCGCACCTCACGGCGTAGTGTTCCTCCACGTGACATATGTACCGGCTCAGGAGAAGAATT
TCACTACCGCTCCTGCTATATGCCATGATGGAAAGGCTCACTTCCCCCGGGAGGGGGTGTTCGTG
TCCAACGGCACCCATTGGTTTGTGACTCAGCGGAATTTCTACGAACCCCAGATCATAACCACTGA
CAACACATTTGTGTCCGGAAATTGTGACGTGGTCATTGGAATAGTGAACAACACTGTTTATGATC
CACTGCAGCCAGAACTTGACAGCTTTAAGGAGGAGCTCGACAAGTACTTCAAGAATCATACGTCA
CCAGATGTGGACCTCGGAGATATTAGCGGTATCAATGCCAGTGTTGTCAATATTCAGAAGGAAAT
AGACCGCCTTAATGAGGTCGCCAAAAATCTGAACGAGAGCCTCATCGATCTTCAGGAGCTGGGCA
AATATGAGCAGTACATCAAGTGGCCTTGGTATATTTGGCTTGGCTTCATCGCCGGCCTGATCGCC
ATAGTAATGGTCACAATTATGCTCTGCTGCATGACCTCTTGCTGCTCCTGTCTGAAAGGCTGCTG
CTCTTGCGGATCCTGCTGCAAATTTGATGAGGATGACAGTGAACCAGTCCTGAAGGGCGTGAAGC
TGCACTATACTTAG

[0445]

PDI-SARS-CoV-2 S 단백질 GSAS+PP wtTM/CT-AA (서열 번호 23)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

IF(PDI)-CoV(opt2).c (서열 번호 24)

TCTCAGATCTTCGCGGTGAATCTTACGACGCGAACACAGTTACCACCCGCAT

IF(AVB)-CoV(opt2).r (서열 번호 25)

ACGACACGACTAAGGCCTCTAAGTATAGTGCAGCTTCACGCCCTTCAGGAC

PDI- 변형된 SARS-CoV-2 S 단백질 GSAS+PP H5iTM/CT-DNA (서열 번호 26)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGAT
CTTCGCGGTGAATCTTACGACGCGAACACAGTTACCACCCGCATATACAAATAGCTTCACTCGGG
GTGTTTATTACCCCGACAAAGTGTTCAGGTCCTCCGTGCTCCACTCAACACAGGACCTCTTTCTT
CCTTTCTTTTCTAACGTGACATGGTTTCATGCCATTCATGTATCCGGCACTAACGGTACTAAGAG
GTTCGATAATCCTGTGCTCCCTTTCAATGACGGCGTTTACTTTGCAAGCACAGAGAAGAGTAACA
TCATCCGAGGTTGGATCTTTGGCACTACCCTCGATTCAAAGACGCAGAGCCTCCTCATTGTGAAC
AATGCCACTAACGTGGTGATCAAAGTTTGCGAGTTTCAGTTCTGCAATGACCCTTTCTTGGGGGT
GTACTATCATAAGAACAACAAGTCTTGGATGGAATCTGAATTCCGCGTCTATAGCAGCGCCAACA
ACTGCACCTTTGAATACGTGTCCCAGCCCTTCCTTATGGACCTGGAGGGAAAGCAGGGAAACTTT
AAGAATCTGAGAGAGTTCGTGTTTAAAAATATCGACGGCTATTTTAAGATCTATTCTAAGCACAC
GCCTATTAATCTCGTGCGCGATCTTCCACAAGGCTTCAGCGCCCTGGAACCACTCGTGGACCTCC
CAATTGGTATCAACATCACTAGATTTCAGACTCTGCTTGCCCTCCACCGATCCTATCTGACACCC
GGAGACTCCTCTAGCGGCTGGACTGCCGGCGCTGCCGCTTATTACGTTGGTTATCTTCAGCCACG
CACGTTCCTGCTGAAGTATAACGAGAATGGTACTATTACCGATGCCGTGGATTGTGCCCTTGACC
CCCTGTCCGAAACTAAGTGCACACTCAAGTCATTCACTGTGGAAAAAGGAATCTACCAGACAAGC
AATTTTCGGGTCCAGCCTACTGAGAGCATTGTGCGCTTTCCTAACATCACAAATCTTTGCCCCTT
CGGAGAGGTTTTCAATGCTACACGGTTTGCCTCCGTGTATGCCTGGAACCGCAAGAGAATTTCCA

[0446]

ATTGCGTGGCCGATTACTCCGTGCTCTACAATAGTGCAAGCTTTAGCACCTTTAAGTGCTATGGC
GTATCCCCTACTAAGCTTAACGACTTGTGTTTCACAAACGTGTATGCCGACTCCTTTGTGATACG
GGGCGACGAAGTTAGACAGATAGCACCAGGACAGACGGGAAAGATAGCTGACTACAACTATAAGC
TTCCTGATGACTTCACTGGCTGCGTTATCGCGTGGAATTCTAACAACCTGGACTCAAAAGTCGGC
GGCAACTATAACTATCTCTATCGGCTGTTCCGCAAGAGTAACCTTAAGCCCTTTGAGAGAGATAT
AAGCACTGAAATCTACCAGGCTGGCAGTACGCCCTGTAATGGCGTGGAAGGCTTTAATTGTTATT
TTCCACTGCAATCCTATGGTTTTCAGCCAACCAATGGCGTGGGCTACCAACCATACCGCGTCGTG
GTGCTCTCCTTTGAACTGCTCCACGCTCCCGCGACTGTCTGCGGCCCCAAGAAGTCCACGAACCT
TGTGAAGAATAAGTGCGTTAATTTTAATTTCAACGGCCTCACTGGAACAGGAGTGCTCACTGAGA
GTAACAAGAAGTTCCTGCCATTTCAACAATTTGGCAGAGACATAGCCGATACTACTGACGCCGTT
AGGGACCCCCAGACCCTCGAGATTCTCGATATAACGCCCTGCTCCTTCGGTGGAGTTTCCGTGAT
CACGCCAGGCACCAATACCAGTAACCAGGTCGCCGTGCTGTATCAGGATGTCAACTGTACTGAGG
TGCCCGTAGCCATCCATGCGGATCAGCTCACACCAACTTGGAGGGTGTACAGCACCGGCTCCAAT
GTATTCCAGACTCGGGCCGGATGCCTTATTGGCGCCGAACACGTGAACAATAGTTACGAATGCGA
TATTCCAATTGGCGCCGGAATCTGTGCTAGCTACCAGACTCAGACGAACTCCCCAGGCAGCGCCA
GCAGCGTTGCCAGCCAGTCAATCATCGCTTATACAATGTCACTTGGAGCCGAAAACTCCGTGGCT
TACTCAAACAACAGCATCGCCATCCCCACAAACTTCACCATATCCGTGACAACTGAGATTCTGCC
AGTGTCCATGACTAAGACGTCCGTAGATTGCACTATGTACATATGCGGCGACAGCACAGAATGTT
CTAATCTGCTGCTGCAATATGGAAGCTTCTGCACTCAACTGAACAGAGCGCTCACAGGCATCGCC
GTGGAGCAGGATAAGAATACCCAGGAGGTGTTCGCCCAAGTTAAGCAGATCTACAAGACCCCACC
CATAAAGGATTTCGGTGGATTCAATTTTAGTCAGATACTCCCAGACCCATCTAAGCCATCCAAGA
GGAGCTTTATCGAGGATCTTTTGTTTAACAAAGTTACTCTGGCCGACGCCGGTTTCATCAAGCAG
TACGGAGATTGCCTCGGCGACATCGCTGCTCGTGACCTCATCTGTGCGCAAAAGTTTAACGGTCT
GACGGTGCTGCCTCCCCTCCTTACTGATGAAATGATCGCCCAGTATACCAGCGCACTCCTCGCTG
GCACCATAACATCCGGTTGGACATTCGGCGCTGGTGCAGCACTGCAGATACCATTCGCCATGCAA
ATGGCATATCGTTTCAACGGTATCGGTGTCACACAGAATGTCCTATATGAGAACCAGAAGCTGAT
CGCAAATCAGTTCAATAGTGCCATCGGAAAAATCCAGGATAGCCTTAGCAGCACAGCCTCAGCCC
TTGGCAAACTCCAGGATGTCGTGAACCAGAATGCCCAGGCTCTCAATACCCTCGTGAAGCAGCTC
TCATCTAATTTCGGCGCAATTTCCAGTGTCCTCAACGACATCCTCAGCCGCCTCGACCCCCCCGA
GGCCGAAGTGCAGATTGACAGACTGATTACAGGTCGACTCCAGAGCCTCCAGACTTACGTGACTC
AGCAGCTGATAAGAGCCGCCGAGATAAGGGCCAGCGCTAACCTGGCTGCCACAAAGATGTCTGAG
TGCGTGCTGGGCCAGTCCAAGAGAGTAGACTTCTGTGGCAAAGGCTACCATCTGATGAGCTTCCC
ACAATCCGCACCTCACGGCGTAGTGTTCCTCCACGTGACATATGTACCGGCTCAGGAGAAGAATT
TCACTACCGCTCCTGCTATATGCCATGATGGAAAGGCTCACTTCCCCCGGGAGGGGGTGTTCGTG
TCCAACGGCACCCATTGGTTTGTGACTCAGCGGAATTTCTACGAACCCCAGATCATAACCACTGA
CAACACATTTGTGTCCGGAAATTGTGACGTGGTCATTGGAATAGTGAACAACACTGTTTATGATC
CACTGCAGCCAGAACTTGACAGCTTTAAGGAGGAGCTCGACAAGTACTTCAAGAATCATACGTCA
CCAGATGTGGACCTCGGAGATATTAGCGGTATCAATGCCAGTGTTGTCAATATTCAGAAGGAAAT
AGACCGCCTTAATGAGGTCGCCAAAAATCTGAACGAGAGCCTCATCGATCTTCAGGAGCTGGGCA
AATATGAGCAGTACATCAAGTGGCCTTGGTATCAAATACTGTCAATTTATTCAACAGTGGCGAGT
TCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACA
ATGCAGAATTTGCATTTAA

PDI- 변형된 SARS-CoV-2 S 단백질 GSAS+PP H5iTM/CT-AA (서열 번호 27)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG

[0447]

VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYQILSIYSTVAS
SLALAIMMAGLSLWMCSNGSLQCRICI

IF(Avb)-H5I.r (서열 번호 28)

ACGACACGACTAAGGCCTTTAAATGCAAATTCTGCATTGTAACGATCC

PDI- 변형된 SARS-CoV-2 S 단백질 GSAS+PP wtTM/H5iCT-DNA (서열 번호 29)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGAT
CTTCGCGGTGAATCTTACGACGCGAACACAGTTACCACCCGCATATACAAATAGCTTCACTCGGG
GTGTTTATTACCCCGACAAAGTGTTCAGGTCCTCCGTGCTCCACTCAACACAGGACCTCTTTCTT
CCTTTCTTTTCTAACGTGACATGGTTTCATGCCATTCATGTATCCGGCACTAACGGTACTAAGAG
GTTCGATAATCCTGTGCTCCCTTTCAATGACGGCGTTTACTTTGCAAGCACAGAGAAGAGTAACA
TCATCCGAGGTTGGATCTTTGGCACTACCCTCGATTCAAAGACGCAGAGCCTCCTCATTGTGAAC
AATGCCACTAACGTGGTGATCAAAGTTTGCGAGTTTCAGTTCTGCAATGACCCTTTCTTGGGGGT
GTACTATCATAAGAACAACAAGTCTTGGATGGAATCTGAATTCCGCGTCTATAGCAGCGCCAACA
ACTGCACCTTTGAATACGTGTCCCAGCCCTTCCTTATGGACCTGGAGGGAAAGCAGGGAAACTTT
AAGAATCTGAGAGAGTTCGTGTTTAAAAATATCGACGGCTATTTTAAGATCTATTCTAAGCACAC
GCCTATTAATCTCGTGCGCGATCTTCCACAAGGCTTCAGCGCCCTGGAACCACTCGTGGACCTCC
CAATTGGTATCAACATCACTAGATTTCAGACTCTGCTTGCCCTCCACCGATCCTATCTGACACCC
GGAGACTCCTCTAGCGGCTGGACTGCCGGCGCTGCCGCTTATTACGTTGGTTATCTTCAGCCACG
CACGTTCCTGCTGAAGTATAACGAGAATGGTACTATTACCGATGCCGTGGATTGTGCCCTTGACC
CCCTGTCCGAAACTAAGTGCACACTCAAGTCATTCACTGTGGAAAAAGGAATCTACCAGACAAGC
AATTTTCGGGTCCAGCCTACTGAGAGCATTGTGCGCTTTCCTAACATCACAAATCTTTGCCCCTT
CGGAGAGGTTTTCAATGCTACACGGTTTGCCTCCGTGTATGCCTGGAACCGCAAGAGAATTTCCA
ATTGCGTGGCCGATTACTCCGTGCTCTACAATAGTGCAAGCTTTAGCACCTTTAAGTGCTATGGC
GTATCCCCTACTAAGCTTAACGACTTGTGTTTCACAAACGTGTATGCCGACTCCTTTGTGATACG
GGGCGACGAAGTTAGACAGATAGCACCAGGACAGACGGGAAAGATAGCTGACTACAACTATAAGC
TTCCTGATGACTTCACTGGCTGCGTTATCGCGTGGAATTCTAACAACCTGGACTCAAAAGTCGGC
GGCAACTATAACTATCTCTATCGGCTGTTCCGCAAGAGTAACCTTAAGCCCTTTGAGAGAGATAT
AAGCACTGAAATCTACCAGGCTGGCAGTACGCCCTGTAATGGCGTGGAAGGCTTTAATTGTTATT
TTCCACTGCAATCCTATGGTTTTCAGCCAACCAATGGCGTGGGCTACCAACCATACCGCGTCGTG
GTGCTCTCCTTTGAACTGCTCCACGCTCCCGCGACTGTCTGCGGCCCCAAGAAGTCCACGAACCT
TGTGAAGAATAAGTGCGTTAATTTTAATTTCAACGGCCTCACTGGAACAGGAGTGCTCACTGAGA
GTAACAAGAAGTTCCTGCCATTTCAACAATTTGGCAGAGACATAGCCGATACTACTGACGCCGTT
AGGGACCCCCAGACCCTCGAGATTCTCGATATAACGCCCTGCTCCTTCGGTGGAGTTTCCGTGAT

[0448]

CACGCCAGGCACCAATACCAGTAACCAGGTCGCCGTGCTGTATCAGGATGTCAACTGTACTGAGG
TGCCCGTAGCCATCCATGCGGATCAGCTCACACCAACTTGGAGGGTGTACAGCACCGGCTCCAAT
GTATTCCAGACTCGGGCCGGATGCCTTATTGGCGCCGAACACGTGAACAATAGTTACGAATGCGA
TATTCCAATTGGCGCCGGAATCTGTGCTAGCTACCAGACTCAGACGAACTCCCCAGGCAGCGCCA
GCAGCGTTGCCAGCCAGTCAATCATCGCTTATACAATGTCACTTGGAGCCGAAAACTCCGTGGCT
TACTCAAACAACAGCATCGCCATCCCCACAAACTTCACCATATCCGTGACAACTGAGATTCTGCC
AGTGTCCATGACTAAGACGTCCGTAGATTGCACTATGTACATATGCGGCGACAGCACAGAATGTT
CTAATCTGCTGCTGCAATATGGAAGCTTCTGCACTCAACTGAACAGAGCGCTCACAGGCATCGCC
GTGGAGCAGGATAAGAATACCCAGGAGGTGTTCGCCCAAGTTAAGCAGATCTACAAGACCCCACC
CATAAAGGATTTCGGTGGATTCAATTTTAGTCAGATACTCCCAGACCCATCTAAGCCATCCAAGA
GGAGCTTTATCGAGGATCTTTTGTTTAACAAAGTTACTCTGGCCGACGCCGGTTTCATCAAGCAG
TACGGAGATTGCCTCGGCGACATCGCTGCTCGTGACCTCATCTGTGCGCAAAAGTTTAACGGTCT
GACGGTGCTGCCTCCCCTCCTTACTGATGAAATGATCGCCCAGTATACCAGCGCACTCCTCGCTG
GCACCATAACATCCGGTTGGACATTCGGCGCTGGTGCAGCACTGCAGATACCATTCGCCATGCAA
ATGGCATATCGTTTCAACGGTATCGGTGTCACACAGAATGTCCTATATGAGAACCAGAAGCTGAT
CGCAAATCAGTTCAATAGTGCCATCGGAAAAATCCAGGATAGCCTTAGCAGCACAGCCTCAGCCC
TTGGCAAACTCCAGGATGTCGTGAACCAGAATGCCCAGGCTCTCAATACCCTCGTGAAGCAGCTC
TCATCTAATTTCGGCGCAATTTCCAGTGTCCTCAACGACATCCTCAGCCGCCTCGACCCCCCCGA
GGCCGAAGTGCAGATTGACAGACTGATTACAGGTCGACTCCAGAGCCTCCAGACTTACGTGACTC
AGCAGCTGATAAGAGCCGCCGAGATAAGGGCCAGCGCTAACCTGGCTGCCACAAAGATGTCTGAG
TGCGTGCTGGGCCAGTCCAAGAGAGTAGACTTCTGTGGCAAAGGCTACCATCTGATGAGCTTCCC
ACAATCCGCACCTCACGGCGTAGTGTTCCTCCACGTGACATATGTACCGGCTCAGGAGAAGAATT
TCACTACCGCTCCTGCTATATGCCATGATGGAAAGGCTCACTTCCCCCGGGAGGGGGTGTTCGTG
TCCAACGGCACCCATTGGTTTGTGACTCAGCGGAATTTCTACGAACCCCAGATCATAACCACTGA
CAACACATTTGTGTCCGGAAATTGTGACGTGGTCATTGGAATAGTGAACAACACTGTTTATGATC
CACTGCAGCCAGAACTTGACAGCTTTAAGGAGGAGCTCGACAAGTACTTCAAGAATCATACGTCA
CCAGATGTGGACCTCGGAGATATTAGCGGTATCAATGCCAGTGTTGTCAATATTCAGAAGGAAAT
AGACCGCCTTAATGAGGTCGCCAAAAATCTGAACGAGAGCCTCATCGATCTTCAGGAGCTGGGCA
AATATGAGCAGTACATCAAGTGGCCTTGGTATATTTGGCTTGGCTTCATCGCCGGCCTGATCGCC
ATAGTAATGGTCACAATTATGCTCTCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAAT
TTGCATTTAA

PDI-변형된 SARS-CoV-2 S 단백질 GSAS+PP wtTM/H5iCT-AA (서열 번호 30)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS

[0449]

PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

클로닝 벡터 8501 (좌측 T-DNA에서 우측 T-DNA) (서열 번호 31)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGT
TTTTAATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAA
AAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTA
CTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGAT
ATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTGGTCAA
AAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGT
TGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAA
ATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTT
TTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCATTTGATTAA
ACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTG
GTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCCATAGAGTCAGTTAACTCAT
TTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAAAAAAACG
GTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATC
CAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACA
TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTT
TATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATAC
AAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAA
ACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCC
TTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTC
GAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGAT
ATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTTGGATCTTGGACGGGAGATTCG
GTTAACTATGCAGCATCTCGATTTTTCGGTTTCGACCAGATCGGATGTACCTATAGTATTCGGTT
TCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCTTCAGCATCTCTGTGAGATGGCAA
TTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGGAAAGTAATGTATCAAGAGGA
TGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAATGCTTCTTCGTCTCCT
ATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGTAGAAGAGCTTAATTAATCGTTGTTGT
TATGAAATACTATTTGTATGAGATGAACTGGTGTAATGTAATTCATTTACATAAGTGGAGTCAGA
ATCAGAATGTTTCCTCCATAACTAACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTT
GAACAACTAAAATTGAACATCTTTTGCCACAACTTTATAAGTGGTTAATATAGCTCAAATATATG
GTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACG
TTATTAACTACTAATTTTATATCATCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCA
TGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAAT
ACGCAAACCGCCTCTCCCCGCGCGTTGGGAATTACTAGCGCGTGTCGACAAGCTTGCATGCCGGT
CAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACC
AAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCA
GCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTG
CGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCAC
CCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGT
GATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCC
CAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCC
ACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT
GTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCT
ATATAAGGAAGTTCATTTCATTTGGAGAGGCACACAATTTGCTTTAGTGATTAAACTTTCTTTTA

[0450]

CAACAAATTAAAGGTCTATTATCTCCCAACAACATAAGAAAACAATGGCGAAAAACGTTGCGATT
TTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGACGTCACTCCTCAG
CCAAAACGACACCCCCATCTGTCTATCCACTGGCCCCTGGATCTGCTGCCCAAACTAACTCCATG
GTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACTCTGG
ATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCCTGCAGTCTGACCTCTACACTCTGAGCA
GCTCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCCACCCG
GCCAGCAGCACCAAGGTGGACAAGAAAATTGTGCCCAGGGATTGTGGTTGTAAGCCTTGCATATG
TACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCCCAAAGCCCAAGGATGTGCTCACCATTA
CTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTC
AGCTGGTTTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAA
CAGCACTTTCCGCTCAGTCAGTGAACTTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGA
CGTCCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGG
CAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAAAGGCCTTAG
TCGTGTCGTTTTTCAAATAATATAATCCTTTTAGGGTTTTAGTTAGTTTAAATTTTCTGTTGCTC
CTGTTTAGCAGGTCGTGCCTTCAGCAAGCACACAAAAACAGAGTGTTTATTTTAAGTTGTTTGTT
TAGTGATTCAAAAAAAAAATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTT
GCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACAT
GTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAAT
ACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATG
TTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCCACGTGACTAGTGGCACTGGCCGTCGTTTT
ACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTG
AATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACT
ATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTA

구성물 8586 (2X35S prom 에서 NOS term) (서열 번호 32)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCC
CAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCC
ACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT
GTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAA
GACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCC
CCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTG
ATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCT
CTATATAAGGAAGTTCATTTCATTTGGAGAGGCACACAATTTGCTTTAGTGATTAAACTTTCTTT
TACAACAAATTAAAGGTCTATTATCTCCCAACAACATAAGAAAACAATGGCGAAAAACGTTGCGA
TTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGTGAATCTTACG
ACGCGAACACAGTTACCACCCGCATATACAAATAGCTTCACTCGGGGTGTTTATTACCCCGACAA
AGTGTTCAGGTCCTCCGTGCTCCACTCAACACAGGACCTCTTTCTTCCTTTCTTTTCTAACGTGA
CATGGTTTCATGCCATTCATGTATCCGGCACTAACGGTACTAAGAGGTTCGATAATCCTGTGCTC
CCTTTCAATGACGGCGTTTACTTTGCAAGCACAGAGAAGAGTAACATCATCCGAGGTTGGATCTT
TGGCACTACCCTCGATTCAAAGACGCAGAGCCTCCTCATTGTGAACAATGCCACTAACGTGGTGA
TCAAAGTTTGCGAGTTTCAGTTCTGCAATGACCCTTTCTTGGGGGTGTACTATCATAAGAACAAC
AAGTCTTGGATGGAATCTGAATTCCGCGTCTATAGCAGCGCCAACAACTGCACCTTTGAATACGT
GTCCCAGCCCTTCCTTATGGACCTGGAGGGAAAGCAGGGAAACTTTAAGAATCTGAGAGAGTTCG
TGTTTAAAAATATCGACGGCTATTTTAAGATCTATTCTAAGCACACGCCTATTAATCTCGTGCGC

[0451]

GATCTTCCACAAGGCTTCAGCGCCCTGGAACCACTCGTGGACCTCCCAATTGGTATCAACATCAC
TAGATTTCAGACTCTGCTTGCCCTCCACCGATCCTATCTGACACCCGGAGACTCCTCTAGCGGCT
GGACTGCCGGCGCTGCCGCTTATTACGTTGGTTATCTTCAGCCACGCACGTTCCTGCTGAAGTAT
AACGAGAATGGTACTATTACCGATGCCGTGGATTGTGCCCTTGACCCCCTGTCCGAAACTAAGTG
CACACTCAAGTCATTCACTGTGGAAAAAGGAATCTACCAGACAAGCAATTTTCGGGTCCAGCCTA
CTGAGAGCATTGTGCGCTTTCCTAACATCACAAATCTTTGCCCCTTCGGAGAGGTTTTCAATGCT
ACACGGTTTGCCTCCGTGTATGCCTGGAACCGCAAGAGAATTTCCAATTGCGTGGCCGATTACTC
CGTGCTCTACAATAGTGCAAGCTTTAGCACCTTTAAGTGCTATGGCGTATCCCCTACTAAGCTTA
ACGACTTGTGTTTCACAAACGTGTATGCCGACTCCTTTGTGATACGGGGCGACGAAGTTAGACAG
ATAGCACCAGGACAGACGGGAAAGATAGCTGACTACAACTATAAGCTTCCTGATGACTTCACTGG
CTGCGTTATCGCGTGGAATTCTAACAACCTGGACTCAAAAGTCGGCGGCAACTATAACTATCTCT
ATCGGCTGTTCCGCAAGAGTAACCTTAAGCCCTTTGAGAGAGATATAAGCACTGAAATCTACCAG
GCTGGCAGTACGCCCTGTAATGGCGTGGAAGGCTTTAATTGTTATTTTCCACTGCAATCCTATGG
TTTTCAGCCAACCAATGGCGTGGGCTACCAACCATACCGCGTCGTGGTGCTCTCCTTTGAACTGC
TCCACGCTCCCGCGACTGTCTGCGGCCCCAAGAAGTCCACGAACCTTGTGAAGAATAAGTGCGTT
AATTTTAATTTCAACGGCCTCACTGGAACAGGAGTGCTCACTGAGAGTAACAAGAAGTTCCTGCC
ATTTCAACAATTTGGCAGAGACATAGCCGATACTACTGACGCCGTTAGGGACCCCCAGACCCTCG
AGATTCTCGATATAACGCCCTGCTCCTTCGGTGGAGTTTCCGTGATCACGCCAGGCACCAATACC
AGTAACCAGGTCGCCGTGCTGTATCAGGATGTCAACTGTACTGAGGTGCCCGTAGCCATCCATGC
GGATCAGCTCACACCAACTTGGAGGGTGTACAGCACCGGCTCCAATGTATTCCAGACTCGGGCCG
GATGCCTTATTGGCGCCGAACACGTGAACAATAGTTACGAATGCGATATTCCAATTGGCGCCGGA
ATCTGTGCTAGCTACCAGACTCAGACGAACTCCCCAGGCAGCGCCAGCAGCGTTGCCAGCCAGTC
AATCATCGCTTATACAATGTCACTTGGAGCCGAAAACTCCGTGGCTTACTCAAACAACAGCATCG
CCATCCCCACAAACTTCACCATATCCGTGACAACTGAGATTCTGCCAGTGTCCATGACTAAGACG
TCCGTAGATTGCACTATGTACATATGCGGCGACAGCACAGAATGTTCTAATCTGCTGCTGCAATA
TGGAAGCTTCTGCACTCAACTGAACAGAGCGCTCACAGGCATCGCCGTGGAGCAGGATAAGAATA
CCCAGGAGGTGTTCGCCCAAGTTAAGCAGATCTACAAGACCCCACCCATAAAGGATTTCGGTGGA
TTCAATTTTAGTCAGATACTCCCAGACCCATCTAAGCCATCCAAGAGGAGCTTTATCGAGGATCT
TTTGTTTAACAAAGTTACTCTGGCCGACGCCGGTTTCATCAAGCAGTACGGAGATTGCCTCGGCG
ACATCGCTGCTCGTGACCTCATCTGTGCGCAAAAGTTTAACGGTCTGACGGTGCTGCCTCCCCTC
CTTACTGATGAAATGATCGCCCAGTATACCAGCGCACTCCTCGCTGGCACCATAACATCCGGTTG
GACATTCGGCGCTGGTGCAGCACTGCAGATACCATTCGCCATGCAAATGGCATATCGTTTCAACG
GTATCGGTGTCACACAGAATGTCCTATATGAGAACCAGAAGCTGATCGCAAATCAGTTCAATAGT
GCCATCGGAAAAATCCAGGATAGCCTTAGCAGCACAGCCTCAGCCCTTGGCAAACTCCAGGATGT
CGTGAACCAGAATGCCCAGGCTCTCAATACCCTCGTGAAGCAGCTCTCATCTAATTTCGGCGCAA
TTTCCAGTGTCCTCAACGACATCCTCAGCCGCCTCGACCCCCCCGAGGCCGAAGTGCAGATTGAC
AGACTGATTACAGGTCGACTCCAGAGCCTCCAGACTTACGTGACTCAGCAGCTGATAAGAGCCGC
CGAGATAAGGGCCAGCGCTAACCTGGCTGCCACAAAGATGTCTGAGTGCGTGCTGGGCCAGTCCA
AGAGAGTAGACTTCTGTGGCAAAGGCTACCATCTGATGAGCTTCCCACAATCCGCACCTCACGGC
GTAGTGTTCCTCCACGTGACATATGTACCGGCTCAGGAGAAGAATTTCACTACCGCTCCTGCTAT
ATGCCATGATGGAAAGGCTCACTTCCCCCGGGAGGGGGTGTTCGTGTCCAACGGCACCCATTGGT
TTGTGACTCAGCGGAATTTCTACGAACCCCAGATCATAACCACTGACAACACATTTGTGTCCGGA
AATTGTGACGTGGTCATTGGAATAGTGAACAACACTGTTTATGATCCACTGCAGCCAGAACTTGA
CAGCTTTAAGGAGGAGCTCGACAAGTACTTCAAGAATCATACGTCACCAGATGTGGACCTCGGAG
ATATTAGCGGTATCAATGCCAGTGTTGTCAATATTCAGAAGGAAATAGACCGCCTTAATGAGGTC
GCCAAAAATCTGAACGAGAGCCTCATCGATCTTCAGGAGCTGGGCAAATATGAGCAGTACATCAA
GTGGCCTTGGTATATTTGGCTTGGCTTCATCGCCGGCCTGATCGCCATAGTAATGGTCACAATTA
TGCTCTGCTGCATGACCTCTTGCTGCTCCTGTCTGAAAGGCTGCTGCTCTTGCGGATCCTGCTGC
AAATTTGATGAGGATGACAGTGAACCAGTCCTGAAGGGCGTGAAGCTGCACTATACTTAGAGGCC
TTAGTCGTGTCGTTTTTCAAATAATATAATCCTTTTAGGGTTTTAGTTAGTTTAAATTTTCTGTT
GCTCCTGTTTAGCAGGTCGTGCCTTCAGCAAGCACACAAAAACAGAGTGTTTATTTTAAGTTGTT
[0452] TGTTTAGTGATTCAAAAAAAAAATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCC

TGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTA
ACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATT
TAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATC
TATGTTACTAGAT

클로닝 벡터 8500 (좌측 T-DNA에서 우측 T-DNA) (서열 번호 33)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGT
TTTTAATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAA
AAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTA
CTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGAT
ATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTGGTCAA
AAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGT
TGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAA
ATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTT
TTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCATTTGATTAA
ACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTG
GTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCCATAGAGTCAGTTAACTCAT
TTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAAAAAAACG
GTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATC
CAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACA
TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTT
TATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATAC
AAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAA
ACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCC
TTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTC
GAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGAT
ATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTTGGATCTTGGACGGGAGATTCG
GTTAACTATGCAGCATCTCGATTTTTCGGTTTCGACCAGATCGGATGTACCTATAGTATTCGGTT
TCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCTTCAGCATCTCTGTGAGATGGCAA
TTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGGAAAGTAATGTATCAAGAGGA
TGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAATGCTTCTTCGTCTCCT
ATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGTAGAAGAGCTTAATTAATCGTTGTTGT
TATGAAATACTATTTGTATGAGATGAACTGGTGTAATGTAATTCATTTACATAAGTGGAGTCAGA
ATCAGAATGTTTCCTCCATAACTAACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTT
GAACAACTAAAATTGAACATCTTTTGCCACAACTTTATAAGTGGTTAATATAGCTCAAATATATG
GTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACG
TTATTAACTACTAATTTTATATCATCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCA
TGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAAT
ACGCAAACCGCCTCTCCCCGCGCGTTGGGAATTACTAGCGCGTGTCGACAAGCTTGCATGCCGGT
CAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACC
AAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCA
GCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTG
CGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCAC
CCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGT
GATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCC
CAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCC
ACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT
GTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCT

[0453]

ATATAAGGAAGTTCATTTCATTTGGAGAGGCACTTTTCTAATCAATCATCAAACAGAACGCAGAA
AATTTCCTAAAAACAAAAAAAAGGCATACAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGT
TTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGACGTCACTCCTCAGCCAAAACGACACC
CCCATCTGTCTATCCACTGGCCCCTGGATCTGCTGCCCAAACTAACTCCATGGTGACCCTGGGAT
GCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGC
GGTGTGCACACCTTCCCAGCTGTCCTGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGT
CCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCA
AGGTGGACAAGAAAATTGTGCCCAGGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAA
GTATCATCTGTCTTCATCTTCCCCCCAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAA
GGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAG
ATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGC
TCAGTCAGTGAACTTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGACGTCCAGATTTTG
GCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTG
GATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAAAGGCCTTAGTCGTGTCGTTTTT
CAAATAATATAATCCTTTTAGGGTTTTAGTTAGTTTAAATTTTCTGTTGCTCCTGTTTAGCAGGT
CGTGCCTTCAGCAAGCACACAAAAACAGAGTGTTTATTTTAAGTTGTTTGTTTAGTGATTCAAAA
AAAAAATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGA
TGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACG
TTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAA
CAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCT
AGAGTCTCAAGCTTGGCGCGCCCACGTGACTAGTGGCACTGGCCGTCGTTTTACAACGTCGTGAC
TGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG
TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCT
AGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGAC
AGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTA

구성물 8589 (2X35S prom 에서 NOS term) (서열 번호 34)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCC
CAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCC
ACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT
GTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAA
GACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCC
CCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTG
ATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCT
CTATATAAGGAAGTTCATTTCATTTGGAGAGGCACTTTTCTAATCAATCATCAAACAGAACGCAG
AAAATTTCCTAAAAACAAAAAAAAGGCATACAAATGGCGAAAAACGTTGCGATTTTCGGCTTATT
GTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGTGAATCTTACGACGCGAACACAGT
TACCACCCGCATATACAAATAGCTTCACTCGGGGTGTTTATTACCCCGACAAAGTGTTCAGGTCC
TCCGTGCTCCACTCAACACAGGACCTCTTTCTTCCTTTCTTTTCTAACGTGACATGGTTTCATGC
CATTCATGTATCCGGCACTAACGGTACTAAGAGGTTCGATAATCCTGTGCTCCCTTTCAATGACG
GCGTTTACTTTGCAAGCACAGAGAAGAGTAACATCATCCGAGGTTGGATCTTTGGCACTACCCTC
GATTCAAAGACGCAGAGCCTCCTCATTGTGAACAATGCCACTAACGTGGTGATCAAAGTTTGCGA
GTTTCAGTTCTGCAATGACCCTTTCTTGGGGGTGTACTATCATAAGAACAACAAGTCTTGGATGG
AATCTGAATTCCGCGTCTATAGCAGCGCCAACAACTGCACCTTTGAATACGTGTCCCAGCCCTTC
CTTATGGACCTGGAGGGAAAGCAGGGAAACTTTAAGAATCTGAGAGAGTTCGTGTTTAAAAATAT

[0454]

CGACGGCTATTTTAAGATCTATTCTAAGCACACGCCTATTAATCTCGTGCGCGATCTTCCACAAG
GCTTCAGCGCCCTGGAACCACTCGTGGACCTCCCAATTGGTATCAACATCACTAGATTTCAGACT
CTGCTTGCCCTCCACCGATCCTATCTGACACCCGGAGACTCCTCTAGCGGCTGGACTGCCGGCGC
TGCCGCTTATTACGTTGGTTATCTTCAGCCACGCACGTTCCTGCTGAAGTATAACGAGAATGGTA
CTATTACCGATGCCGTGGATTGTGCCCTTGACCCCCTGTCCGAAACTAAGTGCACACTCAAGTCA
TTCACTGTGGAAAAAGGAATCTACCAGACAAGCAATTTTCGGGTCCAGCCTACTGAGAGCATTGT
GCGCTTTCCTAACATCACAAATCTTTGCCCCTTCGGAGAGGTTTTCAATGCTACACGGTTTGCCT
CCGTGTATGCCTGGAACCGCAAGAGAATTTCCAATTGCGTGGCCGATTACTCCGTGCTCTACAAT
AGTGCAAGCTTTAGCACCTTTAAGTGCTATGGCGTATCCCCTACTAAGCTTAACGACTTGTGTTT
CACAAACGTGTATGCCGACTCCTTTGTGATACGGGGCGACGAAGTTAGACAGATAGCACCAGGAC
AGACGGGAAAGATAGCTGACTACAACTATAAGCTTCCTGATGACTTCACTGGCTGCGTTATCGCG
TGGAATTCTAACAACCTGGACTCAAAAGTCGGCGGCAACTATAACTATCTCTATCGGCTGTTCCG
CAAGAGTAACCTTAAGCCCTTTGAGAGAGATATAAGCACTGAAATCTACCAGGCTGGCAGTACGC
CCTGTAATGGCGTGGAAGGCTTTAATTGTTATTTTCCACTGCAATCCTATGGTTTTCAGCCAACC
AATGGCGTGGGCTACCAACCATACCGCGTCGTGGTGCTCTCCTTTGAACTGCTCCACGCTCCCGC
GACTGTCTGCGGCCCCAAGAAGTCCACGAACCTTGTGAAGAATAAGTGCGTTAATTTTAATTTCA
ACGGCCTCACTGGAACAGGAGTGCTCACTGAGAGTAACAAGAAGTTCCTGCCATTTCAACAATTT
GGCAGAGACATAGCCGATACTACTGACGCCGTTAGGGACCCCCAGACCCTCGAGATTCTCGATAT
AACGCCCTGCTCCTTCGGTGGAGTTTCCGTGATCACGCCAGGCACCAATACCAGTAACCAGGTCG
CCGTGCTGTATCAGGATGTCAACTGTACTGAGGTGCCCGTAGCCATCCATGCGGATCAGCTCACA
CCAACTTGGAGGGTGTACAGCACCGGCTCCAATGTATTCCAGACTCGGGCCGGATGCCTTATTGG
CGCCGAACACGTGAACAATAGTTACGAATGCGATATTCCAATTGGCGCCGGAATCTGTGCTAGCT
ACCAGACTCAGACGAACTCCCCAGGCAGCGCCAGCAGCGTTGCCAGCCAGTCAATCATCGCTTAT
ACAATGTCACTTGGAGCCGAAAACTCCGTGGCTTACTCAAACAACAGCATCGCCATCCCCACAAA
CTTCACCATATCCGTGACAACTGAGATTCTGCCAGTGTCCATGACTAAGACGTCCGTAGATTGCA
CTATGTACATATGCGGCGACAGCACAGAATGTTCTAATCTGCTGCTGCAATATGGAAGCTTCTGC
ACTCAACTGAACAGAGCGCTCACAGGCATCGCCGTGGAGCAGGATAAGAATACCCAGGAGGTGTT
CGCCCAAGTTAAGCAGATCTACAAGACCCCACCCATAAAGGATTTCGGTGGATTCAATTTTAGTC
AGATACTCCCAGACCCATCTAAGCCATCCAAGAGGAGCTTTATCGAGGATCTTTTGTTTAACAAA
GTTACTCTGGCCGACGCCGGTTTCATCAAGCAGTACGGAGATTGCCTCGGCGACATCGCTGCTCG
TGACCTCATCTGTGCGCAAAAGTTTAACGGTCTGACGGTGCTGCCTCCCCTCCTTACTGATGAAA
TGATCGCCCAGTATACCAGCGCACTCCTCGCTGGCACCATAACATCCGGTTGGACATTCGGCGCT
GGTGCAGCACTGCAGATACCATTCGCCATGCAAATGGCATATCGTTTCAACGGTATCGGTGTCAC
ACAGAATGTCCTATATGAGAACCAGAAGCTGATCGCAAATCAGTTCAATAGTGCCATCGGAAAAA
TCCAGGATAGCCTTAGCAGCACAGCCTCAGCCCTTGGCAAACTCCAGGATGTCGTGAACCAGAAT
GCCCAGGCTCTCAATACCCTCGTGAAGCAGCTCTCATCTAATTTCGGCGCAATTTCCAGTGTCCT
CAACGACATCCTCAGCCGCCTCGACCCCCCCGAGGCCGAAGTGCAGATTGACAGACTGATTACAG
GTCGACTCCAGAGCCTCCAGACTTACGTGACTCAGCAGCTGATAAGAGCCGCCGAGATAAGGGCC
AGCGCTAACCTGGCTGCCACAAAGATGTCTGAGTGCGTGCTGGGCCAGTCCAAGAGAGTAGACTT
CTGTGGCAAAGGCTACCATCTGATGAGCTTCCCACAATCCGCACCTCACGGCGTAGTGTTCCTCC
ACGTGACATATGTACCGGCTCAGGAGAAGAATTTCACTACCGCTCCTGCTATATGCCATGATGGA
AAGGCTCACTTCCCCCGGGAGGGGGTGTTCGTGTCCAACGGCACCCATTGGTTTGTGACTCAGCG
GAATTTCTACGAACCCCAGATCATAACCACTGACAACACATTTGTGTCCGGAAATTGTGACGTGG
TCATTGGAATAGTGAACAACACTGTTTATGATCCACTGCAGCCAGAACTTGACAGCTTTAAGGAG
GAGCTCGACAAGTACTTCAAGAATCATACGTCACCAGATGTGGACCTCGGAGATATTAGCGGTAT
CAATGCCAGTGTTGTCAATATTCAGAAGGAAATAGACCGCCTTAATGAGGTCGCCAAAAATCTGA
ACGAGAGCCTCATCGATCTTCAGGAGCTGGGCAAATATGAGCAGTACATCAAGTGGCCTTGGTAT
ATTTGGCTTGGCTTCATCGCCGGCCTGATCGCCATAGTAATGGTCACAATTATGCTCTGCTGCAT
GACCTCTTGCTGCTCCTGTCTGAAAGGCTGCTGCTCTTGCGGATCCTGCTGCAAATTTGATGAGG
ATGACAGTGAACCAGTCCTGAAGGGCGTGAAGCTGCACTATACTTAGAGGCCTTAGTCGTGTCGT
TTTTCAAATAATATAATCCTTTTAGGGTTTTAGTTAGTTTAAATTTTCTGTTGCTCCTGTTTAGC
[0455] AGGTCGTGCCTTCAGCAAGCACACAAAAACAGAGTGTTTATTTTAAGTTGTTTGTTTAGTGATTC

AAAAAAAAAATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTT
GCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCAT
GACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAG
AAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

클로닝 벡터 8761 (좌측 T-DNA에서 우측 T-DNA) (서열 번호 35)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGT
TTTTAATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAA
AAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTA
CTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGAT
ATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTGGTCAA
AAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGT
TGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAA
ATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTT
TTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCATTTGATTAA
ACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTG
GTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCCATAGAGTCAGTTAACTCAT
TTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAAAAAAACG
GTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATC
CAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACA
TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTT
TATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATAC
AAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAA
ACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCC
TTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTC
GAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGAT
ATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTTGGATCTTGGACGGGAGATTCG
GTTAACTATGCAGCATCTCGATTTTTCGGTTTCGACCAGATCGGATGTACCTATAGTATTCGGTT
TCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCTTCAGCATCTCTGTGAGATGGCAA
TTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGGAAAGTAATGTATCAAGAGGA
TGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAATGCTTCTTCGTCTCCT
ATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGTAGAAGAGCTTAATTAATCGTTGTTGT
TATGAAATACTATTTGTATGAGATGAACTGGTGTAATGTAATTCATTTACATAAGTGGAGTCAGA
ATCAGAATGTTTCCTCCATAACTAACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTT
GAACAACTAAAATTGAACATCTTTTGCCACAACTTTATAAGTGGTTAATATAGCTCAAATATATG
GTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACG
TTATTAACTACTAATTTTATATCATCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCA
TGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAAT
ACGCAAACCGCCTCTCCCCGCGCGTTGGGAATTACTAGCGCGTGTCGACAAGCTTGCATGCCGGT
CAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACC
AAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCA
GCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTG
CGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCAC
CCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGT
GATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCC
CAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCC
ACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT

[0456]

GTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCT
ATATAAGGAAGTTCATTTCATTTGGAGAGGCACTCCATTTGAATCTATCAAACCAAAACACATTG
AGCAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTC
TCAGATCTTCGCGACGTCACTCCTCAGCCAAAACGACACCCCCATCTGTCTATCCACTGGCCCCT
GGATCTGCTGCCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGA
GCCAGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCC
TGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCGAG
ACCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCAAGGTGGACAAGAAAATTGTGCCCAG
GGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCC
CAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATC
AGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACACAGCTCA
GACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACTTCCCATCATGC
ACCAGGACTGGCTCAATGGCAAGGAGACGTCCAGATTTTGGCGATCTATTCAACTGTCGCCAGTT
CATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAG
TGTAGAATATGTATTTAAAGGCCTTAGTCGTGTCGTTTTTCAAATAATATAATCCTTTTAGGGTT
TTAGTTAGTTTAAATTTTCTGTTGCTCCTGTTTAGCAGGTCGTGCCTTCAGCAAGCACACAAAAA
CAGAGTGTTTATTTTAAGTTGTTTGTTTAGTGATTCAAAAAAAAAATCGTTCAAACATTTGGCAA
TAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAA
TTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGA
TTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGAT
AAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCCAC
GTGACTAGTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGAT
CGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGAT
TGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAA
GAGAAAAGAGCGTTTA

구성물 8591 (2X35S prom 에서 NOS term) (서열 번호 36)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCC
CAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCC
ACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT
GTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAA
GACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCC
CCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTG
ATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCT
CTATATAAGGAAGTTCATTTCATTTGGAGAGGCACTCCATTTGAATCTATCAAACCAAAACACAT
TGAGCAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCT
TCTCAGATCTTCGCGGTGAATCTTACGACGCGAACACAGTTACCACCCGCATATACAAATAGCTT
CACTCGGGGTGTTTATTACCCCGACAAAGTGTTCAGGTCCTCCGTGCTCCACTCAACACAGGACC
TCTTTCTTCCTTTCTTTTCTAACGTGACATGGTTTCATGCCATTCATGTATCCGGCACTAACGGT
ACTAAGAGGTTCGATAATCCTGTGCTCCCTTTCAATGACGGCGTTTACTTTGCAAGCACAGAGAA
GAGTAACATCATCCGAGGTTGGATCTTTGGCACTACCCTCGATTCAAAGACGCAGAGCCTCCTCA
TTGTGAACAATGCCACTAACGTGGTGATCAAAGTTTGCGAGTTTCAGTTCTGCAATGACCCTTTC
TTGGGGGTGTACTATCATAAGAACAACAAGTCTTGGATGGAATCTGAATTCCGCGTCTATAGCAG
CGCCAACAACTGCACCTTTGAATACGTGTCCCAGCCCTTCCTTATGGACCTGGAGGGAAAGCAGG
GAAACTTTAAGAATCTGAGAGAGTTCGTGTTTAAAAATATCGACGGCTATTTTAAGATCTATTCT

[0457]

AAGCACACGCCTATTAATCTCGTGCGCGATCTTCCACAAGGCTTCAGCGCCCTGGAACCACTCGT
GGACCTCCCAATTGGTATCAACATCACTAGATTTCAGACTCTGCTTGCCCTCCACCGATCCTATC
TGACACCCGGAGACTCCTCTAGCGGCTGGACTGCCGGCGCTGCCGCTTATTACGTTGGTTATCTT
CAGCCACGCACGTTCCTGCTGAAGTATAACGAGAATGGTACTATTACCGATGCCGTGGATTGTGC
CCTTGACCCCCTGTCCGAAACTAAGTGCACACTCAAGTCATTCACTGTGGAAAAAGGAATCTACC
AGACAAGCAATTTTCGGGTCCAGCCTACTGAGAGCATTGTGCGCTTTCCTAACATCACAAATCTT
TGCCCCTTCGGAGAGGTTTTCAATGCTACACGGTTTGCCTCCGTGTATGCCTGGAACCGCAAGAG
AATTTCCAATTGCGTGGCCGATTACTCCGTGCTCTACAATAGTGCAAGCTTTAGCACCTTTAAGT
GCTATGGCGTATCCCCTACTAAGCTTAACGACTTGTGTTTCACAAACGTGTATGCCGACTCCTTT
GTGATACGGGGCGACGAAGTTAGACAGATAGCACCAGGACAGACGGGAAAGATAGCTGACTACAA
CTATAAGCTTCCTGATGACTTCACTGGCTGCGTTATCGCGTGGAATTCTAACAACCTGGACTCAA
AAGTCGGCGGCAACTATAACTATCTCTATCGGCTGTTCCGCAAGAGTAACCTTAAGCCCTTTGAG
AGAGATATAAGCACTGAAATCTACCAGGCTGGCAGTACGCCCTGTAATGGCGTGGAAGGCTTTAA
TTGTTATTTTCCACTGCAATCCTATGGTTTTCAGCCAACCAATGGCGTGGGCTACCAACCATACC
GCGTCGTGGTGCTCTCCTTTGAACTGCTCCACGCTCCCGCGACTGTCTGCGGCCCCAAGAAGTCC
ACGAACCTTGTGAAGAATAAGTGCGTTAATTTTAATTTCAACGGCCTCACTGGAACAGGAGTGCT
CACTGAGAGTAACAAGAAGTTCCTGCCATTTCAACAATTTGGCAGAGACATAGCCGATACTACTG
ACGCCGTTAGGGACCCCCAGACCCTCGAGATTCTCGATATAACGCCCTGCTCCTTCGGTGGAGTT
TCCGTGATCACGCCAGGCACCAATACCAGTAACCAGGTCGCCGTGCTGTATCAGGATGTCAACTG
TACTGAGGTGCCCGTAGCCATCCATGCGGATCAGCTCACACCAACTTGGAGGGTGTACAGCACCG
GCTCCAATGTATTCCAGACTCGGGCCGGATGCCTTATTGGCGCCGAACACGTGAACAATAGTTAC
GAATGCGATATTCCAATTGGCGCCGGAATCTGTGCTAGCTACCAGACTCAGACGAACTCCCCAGG
CAGCGCCAGCAGCGTTGCCAGCCAGTCAATCATCGCTTATACAATGTCACTTGGAGCCGAAAACT
CCGTGGCTTACTCAAACAACAGCATCGCCATCCCCACAAACTTCACCATATCCGTGACAACTGAG
ATTCTGCCAGTGTCCATGACTAAGACGTCCGTAGATTGCACTATGTACATATGCGGCGACAGCAC
AGAATGTTCTAATCTGCTGCTGCAATATGGAAGCTTCTGCACTCAACTGAACAGAGCGCTCACAG
GCATCGCCGTGGAGCAGGATAAGAATACCCAGGAGGTGTTCGCCCAAGTTAAGCAGATCTACAAG
ACCCCACCCATAAAGGATTTCGGTGGATTCAATTTTAGTCAGATACTCCCAGACCCATCTAAGCC
ATCCAAGAGGAGCTTTATCGAGGATCTTTTGTTTAACAAAGTTACTCTGGCCGACGCCGGTTTCA
TCAAGCAGTACGGAGATTGCCTCGGCGACATCGCTGCTCGTGACCTCATCTGTGCGCAAAAGTTT
AACGGTCTGACGGTGCTGCCTCCCCTCCTTACTGATGAAATGATCGCCCAGTATACCAGCGCACT
CCTCGCTGGCACCATAACATCCGGTTGGACATTCGGCGCTGGTGCAGCACTGCAGATACCATTCG
CCATGCAAATGGCATATCGTTTCAACGGTATCGGTGTCACACAGAATGTCCTATATGAGAACCAG
AAGCTGATCGCAAATCAGTTCAATAGTGCCATCGGAAAAATCCAGGATAGCCTTAGCAGCACAGC
CTCAGCCCTTGGCAAACTCCAGGATGTCGTGAACCAGAATGCCCAGGCTCTCAATACCCTCGTGA
AGCAGCTCTCATCTAATTTCGGCGCAATTTCCAGTGTCCTCAACGACATCCTCAGCCGCCTCGAC
CCCCCCGAGGCCGAAGTGCAGATTGACAGACTGATTACAGGTCGACTCCAGAGCCTCCAGACTTA
CGTGACTCAGCAGCTGATAAGAGCCGCCGAGATAAGGGCCAGCGCTAACCTGGCTGCCACAAAGA
TGTCTGAGTGCGTGCTGGGCCAGTCCAAGAGAGTAGACTTCTGTGGCAAAGGCTACCATCTGATG
AGCTTCCCACAATCCGCACCTCACGGCGTAGTGTTCCTCCACGTGACATATGTACCGGCTCAGGA
GAAGAATTTCACTACCGCTCCTGCTATATGCCATGATGGAAAGGCTCACTTCCCCCGGGAGGGGG
TGTTCGTGTCCAACGGCACCCATTGGTTTGTGACTCAGCGGAATTTCTACGAACCCCAGATCATA
ACCACTGACAACACATTTGTGTCCGGAAATTGTGACGTGGTCATTGGAATAGTGAACAACACTGT
TTATGATCCACTGCAGCCAGAACTTGACAGCTTTAAGGAGGAGCTCGACAAGTACTTCAAGAATC
ATACGTCACCAGATGTGGACCTCGGAGATATTAGCGGTATCAATGCCAGTGTTGTCAATATTCAG
AAGGAAATAGACCGCCTTAATGAGGTCGCCAAAAATCTGAACGAGAGCCTCATCGATCTTCAGGA
GCTGGGCAAATATGAGCAGTACATCAAGTGGCCTTGGTATATTTGGCTTGGCTTCATCGCCGGCC
TGATCGCCATAGTAATGGTCACAATTATGCTCTGCTGCATGACCTCTTGCTGCTCCTGTCTGAAA
GGCTGCTGCTCTTGCGGATCCTGCTGCAAATTTGATGAGGATGACAGTGAACCAGTCCTGAAGGG
CGTGAAGCTGCACTATACTTAGAGGCCTTAGTCGTGTCGTTTTTCAAATAATATAATCCTTTTAG
GGTTTTAGTTAGTTTAAATTTTCTGTTGCTCCTGTTTAGCAGGTCGTGCCTTCAGCAAGCACACA
[0458] AAAACAGAGTGTTTATTTTAAGTTGTTTGTTTAGTGATTCAAAAAAAAAATCGTTCAAACATTTG

GCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGT
TGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTT
ATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTA
GGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질의 C 말단 영역,

변이 2 (서열 번호 37)

WYIWLGFIAGLIAIVMVTIMAGLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질의 C 말단 영역,

변이 3 (서열 번호 38)

WYIWLGFIAGLIAIVMVTIMLCCMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 변형된 SARS-CoV-2 S 단백질의 C 말단 영역,

변이 4 (서열 번호 39)

WYIWLGFIAGLIAIVMVTIMLCCSNGSLQCRICI

3'UTR AvB (아라카차 바이러스 B) (서열 번호 40)

TAGTCGTGTCGTTTTTCAAATAATATAATCCTTTTAGGGTTTTAGTTAGTTTAAATTTTCTGTTG
CTCCTGTTTAGCAGGTCGTGCCTTCAGCAAGCACACAAAAACAGAGTGTTTATTTTAAGTTGTTT
GTTTAGTGATTCAAAAAAAAA

3'UTR trBNYVV (비트 괴저성 엽맥 황화 바이러스) (서열 번호 41)

CCTATCTTGATGAAGGTTGTTGTGGTTTTCTCATTACTGTTTTATTATTGTTTGAGTTGCTTATG
TCGGTTCTTGATTATGTGGTGCATAATTATTGAACTAATTGTTTGTTGGGTTGTAATGTACTGAC
TGGGTGTGAATTGTACCAGTCGTTAAAGGGTTTACTATCAGTATATTGATAT

3' UTR SBMV (써던 빈 모자이크 바이러스) (서열 번호 42)

TGAGGAGTTGTATAATAATACCTGCACCCTTCTCTTTGGCAGGGAGGGTGTTTCGTTTTCACAAT
GCCACGCGCTTGAGGGAGAATGCACGTTAATCATCCCTCCGCTAGTGATGGAGCGTAATCCAAAA
GT

3' UTR TuRSV (순무 링스팟 바이러스) (서열 번호 43)

TGATTTATAATAGCCATAGATTAAGTTTAAATGTATTACGTTTGTATTTTATTCTCTTTTTTAAG
TTTCCTATGTTGTTTTAAATTAAATATCTGTATAATTAGTAGATGTAAATCTGCTTTGTGCGTTT
GACAGTCTGTGGAAACGCACTGGTTCATGAGATAGGACCACCTAGGAGGTAGGACTCTGGGTTTT
AATTATCTCATTTCTTAGTTATACCGTATTATATATATGATTTAGTAGTAATTGTTTTCTCTTGA
TATGTATTATTACTTTTTTATT

3' UTR CPMV (동부 모자이크 바이러스) (서열 번호 44)

ATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTG
TGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCC
CTTCAGCAAGGACACAAAAAGATTTTAATTTTATT

[0459]

3' UTR BBTMV (브로드 빈 트루 모자이크 바이러스) (서열 번호 45)

TAGTTTTCTTCCGCTTTTCTTTTGTAGTGTGTGGTTTTCTTTGTTTCTTCTTTTCTTTTCTCTTT
CCTTTTCTCTTACTCCTGCCTGGCAGGTCGTGCCTTCAGTAAGCACAACAAAAATATGCATTTAT
TAGAGTATTTCTTTCTTCTTTAGCATAAAGGTATTGAAGACCTATAAACTTCGTCCGGGTTGGGG
AAAGTACCAGCTTAGCATATCTTTAGAAAACTATATAGAGCTCTTTACCTTGAGTTGTTTCCTAA
AGTTTATGCAAAAAA

3' UTR trOUMV (우르미아 멜론 바이러스) (서열 번호 46)

CTCACGTCTGGGGTGAGCCCTAGCCAAATAGGAAAACGATAAGCGCTTTGCATGCAAAATGAGTT
GGGCCACAAGTGCCACTCGCAGCGAAGGCGGTCTGAGGTTTCCCCCTGGCGGTTACTTCCATATC
TTTGGGAGATAACTGGG

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 47)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+4P AA
(서열 번호 48)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQ

[0460]

YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+6P AA
(서열 번호 49)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP+923 AA
(서열 번호 50)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSFSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS

[0461]

PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+4P+923 AA
(서열 번호 51)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSFSSTPSALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+6P+923 AA
(서열 번호 52)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSFSSTPSALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSLWMCSNGSLQCRICI

[0462]

H1 Cal 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 53)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLSFWMCSNGSLQCRICI

H3 Minn 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 54)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLMWACQKGNIRCNICI

H6 HK 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 55)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS

[0463]

NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLGLWMCSNGSMQCRICI

H7 Guangdong 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 56)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLVFICVKNGNMRCTICI

H9 HK 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 57)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN

[0464]

VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLLFWAMSNGSCRCNICI

B/Wash 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP AA
(서열 번호 58)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP (대체물 1) AA
(서열 번호 59)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL

[0465]

SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMMAGLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP (대체물 2) AA (서열 번호 60)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMAGLSLWMCSNGSLQCRICI

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP (대체물 3) AA (서열 번호 61)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLCCMCSNGSLQCRICI

[0466]

H5i 헤마글루티닌 CT를 포함하는 (PDI) 변형된 SARS-CoV-2 S 단백질 GSAS+PP (대체물 4) AA (서열 번호 62)

MAKNVAIFGLLFSLLVLVPSQIFAVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFL
PFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVN
NATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF
KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTP
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTS
NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVV
VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVASQSIIAYTMSLGAENSVA
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQ
YGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQL
SSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSE
CVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFV
SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
PDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIA
IVMVTIMLCCSNGSLQCRICI

천연 SARS-CoV-2 S 단백질의 N 말단 영역(천연 신호 펩티드 서열 밑줄체로 표시) (서열 번호 63)

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNS

개재 펩티드 서열 $(X)_n$을 포함하는 변형된 TMCT (서열 번호 64)

WYIWLGFIAGLIAIVMVTIM(X)$_n$CSNGSXXCXICI

PDI-S 단백질 GSAS+4P-DNA (서열 번호 65)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca

[0467]

attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactccttttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggtttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaatttttagtcagatactcccagacccatctaagccatccaaga
ggagccccatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacaccctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctctttatggatgtgctccaatggatcgttacaatgcagaat
ttgcatttaa

PDI-S 단백질 GSAS+6P-DNA (서열 번호 66)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttctttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt

[0468]

gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagccccatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtcccgcactgcagataccattccccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacaccctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctctttatggatgtgctccaatggatcgttacaatgcagaat
[0469] ttgcatttaa

PDI-S 단백질 GSAS+2P+L923F-DNA (서열 번호 67)

[0470]

```
atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaatttttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagcttcagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
```

tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctctttatggatgtgctccaatggatcgttacaatgcagaat
ttgcatttaa

PDI-S 단백질 GSAS+4P+L923F-DNA (서열 번호 68)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactccttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagccccatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct

[0471]

gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagcttcagcagcacaccctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgaccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctctttatggatgtgctccaatggatcgttacaatgcagaat
ttgcatttaa

PDI-S 단백질 GSAS+6P+L923F-DNA (서열 번호 69)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
ccttttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat

[0472]

gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaatttttagtcagatactcccagacccatctaagccatccaaga
ggagccccatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtcccgcactgcagataccattccccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagcttcagcagcacaccctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctctttatggatgtgctccaatggatcgttacaatgcagaat
ttgcatttaa

H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V1) DNA (서열 번호 70)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
ccttttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc

[0473]

ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgatggccggcctctctttatggatgtgctccaatggatcgttaca
atgcagaatttgcatttaa

H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V2) DNA (서열 번호 71)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
ccttttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac

[0474]

[0475]

gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaatttttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgaccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatggccggcctctctttatggatgtgctccaatggatcgttacaatg
cagaatttgcatttaa

H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V3) DNA (서열 번호 72)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttctttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga

[0476]

caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctgctgcatgtgctccaatggatcgttacaatgcagaatttg
catttaa

H5i 헤마글루티닌 CT를 포함하는 PDI 변형된 S 단백질 (V4) DNA (서열 번호 73)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgcccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactccttttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggcccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa

[0477]

atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctctgctgctccaatggatcgttacaatgcagaatttgcattta
a

PDI-S- 단백질+H1 Cal DNA (서열 번호 74)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca

[0478]

gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaatttttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctcagcttctggatgtgctctaatgggtctctacagtgtagaat
atgtatttaa

PDI-S- 단백질+H3 Minn DNA (서열 번호 75)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgaccctttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat

[0479]

aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctcatgtgggcctgtcagaagggcaacatcagatgcaacatctg
catctaa

PDI-S- 단백질+H6 HK DNA (서열 번호 76)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttctttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc

[0480]

ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaacctataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggaccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaatttttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgaccccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctcggtctttggatgtgttcaaatggttcaatgcagtgcaggat
atgtatataa

PDI-S- 단백질+H7 Guangdong DNA (서열 번호 77)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg

[0481]

gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
ccttttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctccttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctccctcctcttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgaccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca

[0482]

ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctcgtcttcatatgtgtgaagaatggaaacatgcggtgcactat
ttgtatataa

PDI-S- 단백질+H9 HK DNA (서열 번호 78)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
ccttttttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgcccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactcctttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctccttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc
agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc

[0483]

ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctcctgttctgggccatgtccaatggatcttgcagatgcaacat
ttgtatataa

PDI-S-단백질+ B/ Wash DNA (서열 번호 79)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcggtgaatcttacgacgcgaacacagttaccacccgcatatacaaatagcttcactcggg
gtgtttattaccccgacaaagtgttcaggtcctccgtgctccactcaacacaggacctctttctt
cctttcttttctaacgtgacatggtttcatgccattcatgtatccggcactaacggtactaagag
gttcgataatcctgtgctccctttcaatgacggcgtttactttgcaagcacagagaagagtaaca
tcatccgaggttggatctttggcactaccctcgattcaaagacgcagagcctcctcattgtgaac
aatgccactaacgtggtgatcaaagtttgcgagtttcagttctgcaatgacccttttcttgggggt
gtactatcataagaacaacaagtcttggatggaatctgaattccgcgtctatagcagcgccaaca
actgcacctttgaatacgtgtcccagcccttccttatggacctggagggaaagcagggaaacttt
aagaatctgagagagttcgtgtttaaaaatatcgacggctattttaagatctattctaagcacac
gcctattaatctcgtgcgcgatcttccacaaggcttcagcgccctggaaccactcgtggacctcc
caattggtatcaacatcactagatttcagactctgcttgccctccaccgatcctatctgacaccc
ggagactcctctagcggctggactgccggcgctgccgcttattacgttggttatcttcagccacg
cacgttcctgctgaagtataacgagaatggtactattaccgatgccgtggattgtgcccttgacc
ccctgtccgaaactaagtgcacactcaagtcattcactgtggaaaaaggaatctaccagacaagc
aattttcgggtccagcctactgagagcattgtgcgctttcctaacatcacaaatctttgccccctt
cggagaggttttcaatgctacacggtttgcctccgtgtatgcctggaaccgcaagagaatttcca
attgcgtggccgattactccgtgctctacaatagtgcaagctttagcacctttaagtgctatggc
gtatcccctactaagcttaacgacttgtgtttcacaaacgtgtatgccgactccttgtgatacg
gggcgacgaagttagacagatagcaccaggacagacgggaaagatagctgactacaactataagc
ttcctgatgacttcactggctgcgttatcgcgtggaattctaacaacctggactcaaaagtcggc
ggcaactataactatctctatcggctgttccgcaagagtaaccttaagccctttgagagagatat
aagcactgaaatctaccaggctggcagtacgccctgtaatggcgtggaaggctttaattgttatt
ttccactgcaatcctatggttttcagccaaccaatggcgtgggctaccaaccataccgcgtcgtg
gtgctctcctttgaactgctccacgctcccgcgactgtctgcggccccaagaagtccacgaacct
tgtgaagaataagtgcgttaattttaatttcaacggcctcactggaacaggagtgctcactgaga
gtaacaagaagttcctgccatttcaacaatttggcagagacatagccgatactactgacgccgtt
agggacccccagaccctcgagattctcgatataacgccctgctccttcggtggagtttccgtgat
cacgccaggcaccaataccagtaaccaggtcgccgtgctgtatcaggatgtcaactgtactgagg
tgcccgtagccatccatgcggatcagctcacaccaacttggagggtgtacagcaccggctccaat
gtattccagactcgggccggatgccttattggcgccgaacacgtgaacaatagttacgaatgcga
tattccaattggcgccggaatctgtgctagctaccagactcagacgaactccccaggcagcgcca
gcagcgttgccagccagtcaatcatcgcttatacaatgtcacttggagccgaaaactccgtggct
tactcaaacaacagcatcgccatccccacaaacttcaccatatccgtgacaactgagattctgcc

[0484]

agtgtccatgactaagacgtccgtagattgcactatgtacatatgcggcgacagcacagaatgtt
ctaatctgctgctgcaatatggaagcttctgcactcaactgaacagagcgctcacaggcatcgcc
gtggagcaggataagaatacccaggaggtgttcgcccaagttaagcagatctacaagaccccacc
cataaaggatttcggtggattcaattttagtcagatactcccagacccatctaagccatccaaga
ggagctttatcgaggatcttttgtttaacaaagttactctggccgacgccggtttcatcaagcag
tacggagattgcctcggcgacatcgctgctcgtgacctcatctgtgcgcaaaagtttaacggtct
gacggtgctgcctcccctccttactgatgaaatgatcgcccagtataccagcgcactcctcgctg
gcaccataacatccggttggacattcggcgctggtgcagcactgcagataccattcgccatgcaa
atggcatatcgtttcaacggtatcggtgtcacacagaatgtcctatatgagaaccagaagctgat
cgcaaatcagttcaatagtgccatcggaaaaatccaggatagccttagcagcacagcctcagccc
ttggcaaactccaggatgtcgtgaaccagaatgcccaggctctcaataccctcgtgaagcagctc
tcatctaatttcggcgcaatttccagtgtcctcaacgacatcctcagccgcctcgacccccccga
ggccgaagtgcagattgacagactgattacaggtcgactccagagcctccagacttacgtgactc
agcagctgataagagccgccgagataagggccagcgctaacctggctgccacaaagatgtctgag
tgcgtgctgggccagtccaagagagtagacttctgtggcaaaggctaccatctgatgagcttccc
acaatccgcacctcacggcgtagtgttcctccacgtgacatatgtaccggctcaggagaagaatt
tcactaccgctcctgctatatgccatgatggaaaggctcacttcccccgggaggggtgttcgtg
tccaacggcacccattggtttgtgactcagcggaatttctacgaaccccagatcataaccactga
caacacatttgtgtccggaaattgtgacgtggtcattggaatagtgaacaacactgtttatgatc
cactgcagccagaacttgacagctttaaggaggagctcgacaagtacttcaagaatcatacgtca
ccagatgtggacctcggagatattagcggtatcaatgccagtgttgtcaatattcagaaggaaat
agaccgccttaatgaggtcgccaaaaatctgaacgagagcctcatcgatcttcaggagctgggca
aatatgagcagtacatcaagtggccttggtatatttggcttggcttcatcgccggcctgatcgcc
atagtaatggtcacaattatgctcgttgtttatatggtctccagagacaatgtttcttgctccat
ttgtctataa

IF-H1HawaiiCT.r (서열 번호 80)

acgacacgactaaggcctttaaatacatattctacactgtagagaccc

IF-H3MinnesotaCT.r (서열 번호 81)

acgacacgactaaggcctttagatgcagatgttgcatctgatgttgcccttctg

IF-HongKongCT.r (서열 번호 82)

acgacacgactaaggcctttatatacatatcctgcactgcattgaaccattt

IF-GuangdongCT.r (서열 번호 83)

acgacacgactaaggcctttatatacaaatagtgcaccgcatgtttcca

IF-H9HKCT.r (서열 번호 84)

acgacacgactaaggcctttatatacaaatgttgcatctgcaagatccat

IF-BWashCT.r (서열 번호 85)

acgacacgactaaggcctttatagacaaatggagcaagaaacattgtctc

[0485]

IF(nbHEL40)-PDI.c (서열 번호 86)

ccaaaacacattgagcaaaatggcgaaaaacgttgcgattttcggcttat

IF(AvB+wtCT).r (서열 번호 87)

ACGACACGACTAAGGCCTTTAGGTATAATGGAGTTTCACCCCCTTCAGAA

PDI-SARS-COV-1 wtTMCT-DNA (서열 번호 88)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTCCGATCTGGATCGGTGCACTACATTTGACGATGTGCAGGCACCTAATTATACTCAGC
ACACCTCTTCCATGCGGGGCGTGTACTACCCCGACGAGATTTTTAGAAGTGACACACTGTACCTG
ACCCAAGACCTTTTTCTCCCATTTTATAGCAATGTCACGGGATTCCACACTATCAATCACACATT
CGGAAACCCTGTTATCCCCTTTAAAGACGGTATCTACTTTGCTGCTACTGAGAAATCGAATGTTG
TTCGCGGTTGGGTGTTCGGCTCAACCATGAACAACAAGAGTCAGTCAGTAATAATTATAAACAAC
TCAACCAATGTCGTCATCAGGGCTTGCAACTTCGAGCTCTGCGATAACCCCTTCTTTGCGGTGTC
CAAACCGATGGGCACTCAGACCCATACCATGATTTTTGACAATGCCTTTAATTGTACTTTCGAAT
ACATCAGCGATGCCTTCTCGCTGGATGTGTCCGAGAAGAGCGGTAACTTCAAGCATTTGCGGGAG
TTCGTTTTCAAAAACAAAGACGGGTTTTTGTATGTCTACAAAGGCTATCAGCCCATTGATGTGGT
CAGGGATCTGCCCAGTGGTTTCAACACACTGAAGCCCATCTTCAAGTTGCCACTCGGCATCAACA
TCACTAATTTCCGCGCCATCCTAACTGCTTTTTCGCCAGCGCAGGATATTTGGGGGACATCCGCC
GCAGCTTACTTCGTGGGATATCTGAAGCCCACCACTTTTATGCTAAAGTACGACGAGAATGGCAC
CATCACCGATGCCGTGGACTGCTCACAGAATCCTCTCGCAGAGCTCAAGTGTTCAGTGAAGTCAT
TTGAGATCGACAAAGGGATCTACCAAACATCTAACTTTCGCGTCGTGCCTTCCGGTGACGTAGTC
CGTTTCCCAAATATCACTAACTTGTGCCCTTTTGGTGAAGTATTCAACGCAACCAAATTCCCCAG
TGTGTATGCGTGGGAGCGCAAAAAGATCAGTAACTGTGTGGCTGACTATAGTGTTCTGTACAATA
GCACCTTCTTCAGCACCTTCAAGTGTTATGGAGTGAGCGCTACAAAACTGAACGATCTTTGTTTC
TCAAACGTGTACGCCGATTCATTTGTCGTTAAAGGTGACGATGTGAGGCAGATCGCTCCAGGCCA
GACAGGTGTGATTGCTGACTATAATTACAAACTGCCAGACGACTTCATGGGGTGCGTGCTAGCTT
GGAATACAAGAAACATTGACGCCACCTCCACGGGAAATTACAATTACAAGTATCGTTACCTTCGC
CATGGAAAGTTGAGACCCTTCGAGCGTGATATAAGTAACGTGCCCTTTAGTCCAGATGGAAAACC
CTGCACACCCCCTGCTCTCAATTGCTATTGGCCTCTCAATGACTACGGCTTTTACACAACTACTG
GCATCGGATACCAGCCTTACCGGGTCGTGGTGCTCAGTTTTGAGTTGCTTAACGCACCCGCCACC
GTGTGTGGTCCTAAACTTTCTACTGACCTGATTAAAAACCAATGCGTCAACTTCAATTTTAACGG
GCTGACCGGCACCGGTGTCCTGACCCCTAGCTCTAAGAGATTCCAGCCTTTTCAGCAGTTCGGGA
GGGATGTGAGCGACTTTACCGACTCTGTCAGGGATCCAAAGACCAGCGAGATACTGGATATCTCG
CCCTGCAGTTTCGGTGGCGTGTCCGTTATTACACCTGGCACCAACGCCTCCTCAGAGGTGGCGGT
GCTCTATCAAGATGTCAACTGCACTGATGTGTCAACTGCCATCCATGCCGATCAGCTGACCCCCG
CCTGGCGCATCTACAGTACCGGGAACAACGTTTTTCAGACCCAGGCCGGCTGTCTAATCGGCGCA
GAGCACGTTGACACATCCTACGAATGTGACATACCTATCGGGGCAGGCATTTGCGCTAGCTACCA
TACCGTGTCACTGTTGGCTTCCACGTCACAAAAGTCAATCGTTGCCTACACGATGAGTCTGGGGG
CTGACTCATCTATCGCCTACAGCAACAATACCATTGCAATTCCCACAAACTTCAGTATCTCCATC
ACAACAGAGGTGATGCCCGTTTCTATGGCTAAAACATCAGTCGATTGCAATATGTATATATGCGG
CGATAGTACTGAGTGCGCCAATCTCTTGTTACAGTACGGCTCCTTTTGTACCCAGCTGAACCGAG
CACTGTCTGGAATCGCCGCAGAACAGGATCGCAATACCCGGGAAGTCTTCGCCCAGGTGAAGCAG
ATGTACAAAACGCCCACTCTCAAGTATTTCGGCGGATTCAACTTTTCTCAGATTTTGCCTGACCC
GCTCAAGCCAACAAAACGATCTTTTATCGAAGACCTTCTGTTTAACAAGGTCACACTGGCGGATG
CTGGGTTCATGAAACAGTACGGTGAATGCCTGGGGGACATCAATGCCAGAGATCTGATCTGCGCC
CAGAAATTCAATGGCTTAACAGTCCTCCCACCTCTCTTGACCGACGATATGATCGCTGCGTACAC
CGCTGCTCTGGTATCGGGCACCGCGACTGCTGGCTGGACCTTTGGTGCCGGAGCCGCACTCCAGA
TCCCATTCGCCATGCAGATGGCCTACCGCTTCAACGGAATCGGGGTCACCCAGAACGTGCTGTAT

[0486]

GAGAACCAGAAACAGATCGCCAATCAGTTCAATAAGGCAATTAGTCAGATTCAGGAGAGTCTTAC
CACTACCAGCACCGCCCTGGGCAAGCTGCAAGATGTTGTGAACCAGAATGCGCAGGCATTAAACA
CTCTGGTTAAACAGCTGAGCTCAAATTTTGGTGCAATCTCTTCAGTTCTGAACGATATCCTGAGT
CGGCTGGATCCGCCAGAGGCTGAAGTGCAAATTGATCGTTTGATCACCGGGAGGCTACAATCTCT
GCAGACGTACGTGACCCAGCAGCTCATCCGGGCAGCCGAAATTCGCGCATCAGCCAACCTCGCTG
CAACTAAGATGTCTGAGTGCGTGCTGGGCCAGAGTAAGAGGGTGGACTTTTGTGGTAAGGGATAC
CACCTCATGTCCTTTCCGCAAGCGGCTCCCCACGGCGTGGTTTTCTTACACGTTACCTATGTGCC
ATCCCAAGAACGCAATTTCACCACCGCTCCAGCTATCTGTCATGAGGGCAAAGCATATTTCCCCA
GGGAAGGAGTATTTGTGTTTAATGGCACGTCCTGGTTTATAACCCAACGTAACTTTTTCTCCCCA
CAGATTATCACAACCGACAACACATTCGTGTCTGGGAATTGTGACGTCGTGATCGGGATCATTAA
CAATACCGTTTACGATCCCTTGCAGCCCGAGCTTGACTCCTTTAAAGAGGAACTAGACAAATACT
TTAAGAATCACACCTCACCGGACGTAGATTTGGGAGACATCTCTGGAATTAATGCCTCTGTGGTG
AATATCCAGAAGGAGATCGACCGCCTGAATGAAGTCGCCAAGAACCTCAACGAGTCCCTGATAGA
TCTGCAAGAACTGGGCAAATATGAACAGTACATCAAATGGCCGTGGTACGTGTGGTTGGGCTTTA
TCGCTGGACTTATTGCAATCGTGATGGTGACGATTCTGCTCTGCTGTATGACTTCCTGCTGCTCT
TGTCTGAAGGGCGCCTGTAGCTGTGGTTCCTGCTGCAAGTTCGACGAAGACGACTCCGAACCAGT
TCTGAAGGGGGTGAAACTCCATTATACCTAA

PDI-SARS-COV-1 H5iTMCT-DNA (서열 번호 89)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTCCGATCTGGATCGGTGCACTACATTTGACGATGTGCAGGCACCTAATTATACTCAGC
ACACCTCTTCCATGCGGGGCGTGTACTACCCCGACGAGATTTTTAGAAGTGACACACTGTACCTG
ACCCAAGACCTTTTTCTCCCATTTTATAGCAATGTCACGGGATTCCACACTATCAATCACACATT
CGGAAACCCTGTTATCCCCTTTAAAGACGGTATCTACTTTGCTGCTACTGAGAAATCGAATGTTG
TTCGCGGTTGGGTGTTCGGCTCAACCATGAACAACAAGAGTCAGTCAGTAATAATTATAAACAAC
TCAACCAATGTCGTCATCAGGGCTTGCAACTTCGAGCTCTGCGATAACCCCTTCTTTGCGGTGTC
CAAACCGATGGGCACTCAGACCCATACCATGATTTTTGACAATGCCTTTAATTGTACTTTCGAAT
ACATCAGCGATGCCTTCTCGCTGGATGTGTCCGAGAAGAGCGGTAACTTCAAGCATTTGCGGGAG
TTCGTTTTCAAAAACAAAGACGGGTTTTTGTATGTCTACAAAGGCTATCAGCCCATTGATGTGGT
CAGGGATCTGCCCAGTGGTTTCAACACACTGAAGCCCATCTTCAAGTTGCCACTCGGCATCAACA
TCACTAATTTCCGCGCCATCCTAACTGCTTTTTCGCCAGCGCAGGATATTTGGGGGACATCCGCC
GCAGCTTACTTCGTGGGATATCTGAAGCCCACCACTTTTATGCTAAAGTACGACGAGAATGGCAC
CATCACCGATGCCGTGGACTGCTCACAGAATCCTCTCGCAGAGCTCAAGTGTTCAGTGAAGTCAT
TTGAGATCGACAAAGGGATCTACCAAACATCTAACTTTCGCGTCGTGCCTTCCGGTGACGTAGTC
CGTTTCCCAAATATCACTAACTTGTGCCCTTTTGGTGAAGTATTCAACGCAACCAAATTCCCCAG
TGTGTATGCGTGGGAGCGCAAAAAGATCAGTAACTGTGTGGCTGACTATAGTGTTCTGTACAATA
GCACCTTCTTCAGCACCTTCAAGTGTTATGGAGTGAGCGCTACAAAACTGAACGATCTTTGTTTC
TCAAACGTGTACGCCGATTCATTTGTCGTTAAAGGTGACGATGTGAGGCAGATCGCTCCAGGCCA
GACAGGTGTGATTGCTGACTATAATTACAAACTGCCAGACGACTTCATGGGGTGCGTGCTAGCTT
GGAATACAAGAAACATTGACGCCACCTCCACGGGAAATTACAATTACAAGTATCGTTACCTTCGC
CATGGAAAGTTGAGACCCTTCGAGCGTGATATAAGTAACGTGCCCTTTAGTCCAGATGGAAAACC
CTGCACACCCCCTGCTCTCAATTGCTATTGGCCTCTCAATGACTACGGCTTTTACACAACTACTG
GCATCGGATACCAGCCTTACCGGGTCGTGGTGCTCAGTTTTGAGTTGCTTAACGCACCCGCCACC
GTGTGTGGTCCTAAACTTTCTACTGACCTGATTAAAAACCAATGCGTCAACTTCAATTTTAACGG
GCTGACCGGCACCGGTGTCCTGACCCCTAGCTCTAAGAGATTCCAGCCTTTTCAGCAGTTCGGGA
GGGATGTGAGCGACTTTACCGACTCTGTCAGGGATCCAAAGACCAGCGAGATACTGGATATCTCG
CCCTGCAGTTTCGGTGGCGTGTCCGTTATTACACCTGGCACCAACGCCTCCTCAGAGGTGGCGGT
GCTCTATCAAGATGTCAACTGCACTGATGTGTCAACTGCCATCCATGCCGATCAGCTGACCCCCG
CCTGGCGCATCTACAGTACCGGGAACAACGTTTTTCAGACCCAGGCCGGCTGTCTAATCGGCGCA
GAGCACGTTGACACATCCTACGAATGTGACATACCTATCGGGGCAGGCATTTGCGCTAGCTACCA

[0487]

TACCGTGTCACTGTTGGCTTCCACGTCACAAAAGTCAATCGTTGCCTACACGATGAGTCTGGGGG
CTGACTCATCTATCGCCTACAGCAACAATACCATTGCAATTCCCACAAACTTCAGTATCTCCATC
ACAACAGAGGTGATGCCCGTTTCTATGGCTAAAACATCAGTCGATTGCAATATGTATATATGCGG
CGATAGTACTGAGTGCGCCAATCTCTTGTTACAGTACGGCTCCTTTTGTACCCAGCTGAACCGAG
CACTGTCTGGAATCGCCGCAGAACAGGATCGCAATACCCGGGAAGTCTTCGCCCAGGTGAAGCAG
ATGTACAAAACGCCCACTCTCAAGTATTTCGGCGGATTCAACTTTTCTCAGATTTTGCCTGACCC
GCTCAAGCCAACAAAACGATCTTTTATCGAAGACCTTCTGTTTAACAAGGTCACACTGGCGGATG
CTGGGTTCATGAAACAGTACGGTGAATGCCTGGGGGACATCAATGCCAGAGATCTGATCTGCGCC
CAGAAATTCAATGGCTTAACAGTCCTCCCACCTCTCTTGACCGACGATATGATCGCTGCGTACAC
CGCTGCTCTGGTATCGGGCACCGCGACTGCTGGCTGGACCTTTGGTGCCGGAGCCGCACTCCAGA
TCCCATTCGCCATGCAGATGGCCTACCGCTTCAACGGAATCGGGGTCACCCAGAACGTGCTGTAT
GAGAACCAGAAACAGATCGCCAATCAGTTCAATAAGGCAATTAGTCAGATTCAGGAGAGTCTTAC
CACTACCAGCACCGCCCTGGGCAAGCTGCAAGATGTTGTGAACCAGAATGCGCAGGCATTAAACA
CTCTGGTTAAACAGCTGAGCTCAAATTTTGGTGCAATCTCTTCAGTTCTGAACGATATCCTGAGT
CGGCTGGATCCGCCAGAGGCTGAAGTGCAAATTGATCGTTTGATCACCGGGAGGCTACAATCTCT
GCAGACGTACGTGACCCAGCAGCTCATCCGGGCAGCCGAAATTCGCGCATCAGCCAACCTCGCTG
CAACTAAGATGTCTGAGTGCGTGCTGGGCCAGAGTAAGAGGGTGGACTTTTGTGGTAAGGGATAC
CACCTCATGTCCTTTCCGCAAGCGGCTCCCCACGGCGTGGTTTTCTTACACGTTACCTATGTGCC
ATCCCAAGAACGCAATTTCACCACCGCTCCAGCTATCTGTCATGAGGGCAAAGCATATTTCCCCA
GGGAAGGAGTATTTGTGTTTAATGGCACGTCCTGGTTTATAACCCAACGTAACTTTTTCTCCCCA
CAGATTATCACAACCGACAACACATTCGTGTCTGGGAATTGTGACGTCGTGATCGGGATCATTAA
CAATACCGTTTACGATCCCTTGCAGCCCGAGCTTGACTCCTTTAAAGAGGAACTAGACAAATACT
TTAAGAATCACACCTCACCGGACGTAGATTTGGGAGACATCTCTGGAATTAATGCCTCTGTGGTG
AATATCCAGAAGGAGATCGACCGCCTGAATGAAGTCGCCAAGAACCTCAACGAGTCCCTGATAGA
TCTGCAAGAACTGGGCAAATATGAACAGTACATCAAATGGCCGTGGTACcaaatactgtcaattt
attcaacagtggcgagttccctagcactggcaatcatgatggctggtctatctttatggatgtgc
tccaatggatcgttacaatgcagaatttgcattTAA

PDI-SARS-COV-1 H5iCT-DNA (서열 번호 90)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTCCGATCTGGATCGGTGCACTACATTTGACGATGTGCAGGCACCTAATTATACTCAGC
ACACCTCTTCCATGCGGGGCGTGTACTACCCCGACGAGATTTTTAGAAGTGACACACTGTACCTG
ACCCAAGACCTTTTTCTCCCATTTTATAGCAATGTCACGGGATTCCACACTATCAATCACACATT
CGGAAACCCTGTTATCCCCTTTAAAGACGGTATCTACTTTGCTGCTACTGAGAAATCGAATGTTG
TTCGCGGTTGGGTGTTCGGCTCAACCATGAACAACAAGAGTCAGTCAGTAATAATTATAAACAAC
TCAACCAATGTCGTCATCAGGGCTTGCAACTTCGAGCTCTGCGATAACCCCTTCTTTGCGGTGTC
CAAACCGATGGGCACTCAGACCCATACCATGATTTTTGACAATGCCTTTAATTGTACTTTCGAAT
ACATCAGCGATGCCTTCTCGCTGGATGTGTCCGAGAAGAGCGGTAACTTCAAGCATTTGCGGGAG
TTCGTTTTCAAAAACAAAGACGGGTTTTTGTATGTCTACAAAGGCTATCAGCCCATTGATGTGGT
CAGGGATCTGCCCAGTGGTTTCAACACACTGAAGCCCATCTTCAAGTTGCCACTCGGCATCAACA
TCACTAATTTCCGCGCCATCCTAACTGCTTTTTCGCCAGCGCAGGATATTTGGGGGACATCCGCC
GCAGCTTACTTCGTGGGATATCTGAAGCCCACCACTTTTATGCTAAAGTACGACGAGAATGGCAC
CATCACCGATGCCGTGGACTGCTCACAGAATCCTCTCGCAGAGCTCAAGTGTTCAGTGAAGTCAT
TTGAGATCGACAAAGGGATCTACCAAACATCTAACTTTCGCGTCGTGCCTTCCGGTGACGTAGTC
CGTTTCCCAAATATCACTAACTTGTGCCCTTTTGGTGAAGTATTCAACGCAACCAAATTCCCCAG
TGTGTATGCGTGGGAGCGCAAAAAGATCAGTAACTGTGTGGCTGACTATAGTGTTCTGTACAATA
GCACCTTCTTCAGCACCTTCAAGTGTTATGGAGTGAGCGCTACAAAACTGAACGATCTTTGTTTC
TCAAACGTGTACGCCGATTCATTTGTCGTTAAAGGTGACGATGTGAGGCAGATCGCTCCAGGCCA
GACAGGTGTGATTGCTGACTATAATTACAAACTGCCAGACGACTTCATGGGGTGCGTGCTAGCTT
GGAATACAAGAAACATTGACGCCACCTCCACGGGAAATTACAATTACAAGTATCGTTACCTTCGC

[0488]

CATGGAAAGTTGAGACCCTTCGAGCGTGATATAAGTAACGTGCCCTTTAGTCCAGATGGAAAACC
CTGCACACCCCCTGCTCTCAATTGCTATTGGCCTCTCAATGACTACGGCTTTTACACAACTACTG
GCATCGGATACCAGCCTTACCGGGTCGTGGTGCTCAGTTTTGAGTTGCTTAACGCACCCGCCACC
GTGTGTGGTCCTAAACTTTCTACTGACCTGATTAAAAACCAATGCGTCAACTTCAATTTTAACGG
GCTGACCGGCACCGGTGTCCTGACCCCTAGCTCTAAGAGATTCCAGCCTTTTCAGCAGTTCGGGA
GGGATGTGAGCGACTTTACCGACTCTGTCAGGGATCCAAAGACCAGCGAGATACTGGATATCTCG
CCCTGCAGTTTCGGTGGCGTGTCCGTTATTACACCTGGCACCAACGCCTCCTCAGAGGTGGCGGT
GCTCTATCAAGATGTCAACTGCACTGATGTGTCAACTGCCATCCATGCCGATCAGCTGACCCCCG
CCTGGCGCATCTACAGTACCGGGAACAACGTTTTTCAGACCCAGGCCGGCTGTCTAATCGGCGCA
GAGCACGTTGACACATCCTACGAATGTGACATACCTATCGGGGCAGGCATTTGCGCTAGCTACCA
TACCGTGTCACTGTTGGCTTCCACGTCACAAAAGTCAATCGTTGCCTACACGATGAGTCTGGGGG
CTGACTCATCTATCGCCTACAGCAACAATACCATTGCAATTCCCACAAACTTCAGTATCTCCATC
ACAACAGAGGTGATGCCCGTTTCTATGGCTAAAACATCAGTCGATTGCAATATGTATATATGCGG
CGATAGTACTGAGTGCGCCAATCTCTTGTTACAGTACGGCTCCTTTTGTACCCAGCTGAACCGAG
CACTGTCTGGAATCGCCGCAGAACAGGATCGCAATACCCGGGAAGTCTTCGCCCAGGTGAAGCAG
ATGTACAAAACGCCCACTCTCAAGTATTTCGGCGGATTCAACTTTTCTCAGATTTTGCCTGACCC
GCTCAAGCCAACAAAACGATCTTTTATCGAAGACCTTCTGTTTAACAAGGTCACACTGGCGGATG
CTGGGTTCATGAAACAGTACGGTGAATGCCTGGGGGACATCAATGCCAGAGATCTGATCTGCGCC
CAGAAATTCAATGGCTTAACAGTCCTCCCACCTCTCTTGACCGACGATATGATCGCTGCGTACAC
CGCTGCTCTGGTATCGGGCACCGCGACTGCTGGCTGGACCTTTGGTGCCGGAGCCGCACTCCAGA
TCCCATTCGCCATGCAGATGGCCTACCGCTTCAACGGAATCGGGGTCACCCAGAACGTGCTGTAT
GAGAACCAGAAACAGATCGCCAATCAGTTCAATAAGGCAATTAGTCAGATTCAGGAGAGTCTTAC
CACTACCAGCACCGCCCTGGGCAAGCTGCAAGATGTTGTGAACCAGAATGCGCAGGCATTAAACA
CTCTGGTTAAACAGCTGAGCTCAAATTTTGGTGCAATCTCTTCAGTTCTGAACGATATCCTGAGT
CGGCTGGATCCGCCAGAGGCTGAAGTGCAAATTGATCGTTTGATCACCGGGAGGCTACAATCTCT
GCAGACGTACGTGACCCAGCAGCTCATCCGGGCAGCCGAAATTCGCGCATCAGCCAACCTCGCTG
CAACTAAGATGTCTGAGTGCGTGCTGGGCCAGAGTAAGAGGGTGGACTTTTGTGGTAAGGGATAC
CACCTCATGTCCTTTCCGCAAGCGGCTCCCCACGGCGTGGTTTTCTTACACGTTACCTATGTGCC
ATCCCAAGAACGCAATTTCACCACCGCTCCAGCTATCTGTCATGAGGGCAAAGCATATTTCCCCA
GGGAAGGAGTATTTGTGTTTAATGGCACGTCCTGGTTTATAACCCAACGTAACTTTTTCTCCCCA
CAGATTATCACAACCGACAACACATTCGTGTCTGGGAATTGTGACGTCGTGATCGGGATCATTAA
CAATACCGTTTACGATCCCTTGCAGCCCGAGCTTGACTCCTTTAAAGAGGAACTAGACAAATACT
TTAAGAATCACACCTCACCGGACGTAGATTTGGGAGACATCTCTGGAATTAATGCCTCTGTGGTG
AATATCCAGAAGGAGATCGACCGCCTGAATGAAGTCGCCAAGAACCTCAACGAGTCCCTGATAGA
TCTGCAAGAACTGGGCAAATATGAACAGTACATCAAATGGCCGTGGTACGTGTGGTTGGGCTTTA
TCGCTGGACTTATTGCAATCGTGATGGTGACGATTCTGCTCtctttatggatgtgctccaatgga
tcgttacaatgcagaatttgcattTAA

PDI-SARS-COV-1 H5iCT(V4)-DNA (서열 번호 91)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTCCGATCTGGATCGGTGCACTACATTTGACGATGTGCAGGCACCTAATTATACTCAGC
ACACCTCTTCCATGCGGGGCGTGTACTACCCCGACGAGATTTTTAGAAGTGACACACTGTACCTG
ACCCAAGACCTTTTTCTCCCATTTTATAGCAATGTCACGGGATTCCACACTATCAATCACACATT
CGGAAACCCTGTTATCCCCTTTAAAGACGGTATCTACTTTGCTGCTACTGAGAAATCGAATGTTG
TTCGCGGTTGGGTGTTCGGCTCAACCATGAACAACAAGAGTCAGTCAGTAATAATTATAAACAAC
TCAACCAATGTCGTCATCAGGGCTTGCAACTTCGAGCTCTGCGATAACCCCTTCTTTGCGGTGTC
CAAACCGATGGGCACTCAGACCCATACCATGATTTTTGACAATGCCTTTAATTGTACTTTCGAAT
ACATCAGCGATGCCTTCTCGCTGGATGTGTCCGAGAAGAGCGGTAACTTCAAGCATTTGCGGGAG
TTCGTTTTCAAAAACAAAGACGGGTTTTTGTATGTCTACAAAGGCTATCAGCCCATTGATGTGGT
CAGGGATCTGCCCAGTGGTTTCAACACACTGAAGCCCATCTTCAAGTTGCCACTCGGCATCAACA

[0489]

TCACTAATTTCCGCGCCATCCTAACTGCTTTTTCGCCAGCGCAGGATATTTGGGGGACATCCGCC
GCAGCTTACTTCGTGGGATATCTGAAGCCCACCACTTTTATGCTAAAGTACGACGAGAATGGCAC
CATCACCGATGCCGTGGACTGCTCACAGAATCCTCTCGCAGAGCTCAAGTGTTCAGTGAAGTCAT
TTGAGATCGACAAAGGGATCTACCAAACATCTAACTTTCGCGTCGTGCCTTCCGGTGACGTAGTC
CGTTTCCCAAATATCACTAACTTGTGCCCTTTTGGTGAAGTATTCAACGCAACCAAATTCCCCAG
TGTGTATGCGTGGGAGCGCAAAAAGATCAGTAACTGTGTGGCTGACTATAGTGTTCTGTACAATA
GCACCTTCTTCAGCACCTTCAAGTGTTATGGAGTGAGCGCTACAAAACTGAACGATCTTTGTTTC
TCAAACGTGTACGCCGATTCATTTGTCGTTAAAGGTGACGATGTGAGGCAGATCGCTCCAGGCCA
GACAGGTGTGATTGCTGACTATAATTACAAACTGCCAGACGACTTCATGGGGTGCGTGCTAGCTT
GGAATACAAGAAACATTGACGCCACCTCCACGGGAAATTACAATTACAAGTATCGTTACCTTCGC
CATGGAAAGTTGAGACCCTTCGAGCGTGATATAAGTAACGTGCCCTTTAGTCCAGATGGAAAACC
CTGCACACCCCCTGCTCTCAATTGCTATTGGCCTCTCAATGACTACGGCTTTTACACAACTACTG
GCATCGGATACCAGCCTTACCGGGTCGTGGTGCTCAGTTTTGAGTTGCTTAACGCACCCGCCACC
GTGTGTGGTCCTAAACTTTCTACTGACCTGATTAAAAACCAATGCGTCAACTTCAATTTTAACGG
GCTGACCGGCACCGGTGTCCTGACCCCTAGCTCTAAGAGATTCCAGCCTTTTCAGCAGTTCGGGA
GGGATGTGAGCGACTTTACCGACTCTGTCAGGGATCCAAAGACCAGCGAGATACTGGATATCTCG
CCCTGCAGTTTCGGTGGCGTGTCCGTTATTACACCTGGCACCAACGCCTCCTCAGAGGTGGCGGT
GCTCTATCAAGATGTCAACTGCACTGATGTGTCAACTGCCATCCATGCCGATCAGCTGACCCCCG
CCTGGCGCATCTACAGTACCGGGAACAACGTTTTTCAGACCCAGGCCGGCTGTCTAATCGGCGCA
GAGCACGTTGACACATCCTACGAATGTGACATACCTATCGGGGCAGGCATTTGCGCTAGCTACCA
TACCGTGTCACTGTTGGCTTCCACGTCACAAAAGTCAATCGTTGCCTACACGATGAGTCTGGGGG
CTGACTCATCTATCGCCTACAGCAACAATACCATTGCAATTCCCACAAACTTCAGTATCTCCATC
ACAACAGAGGTGATGCCCGTTTCTATGGCTAAAACATCAGTCGATTGCAATATGTATATATGCGG
CGATAGTACTGAGTGCGCCAATCTCTTGTTACAGTACGGCTCCTTTTGTACCCAGCTGAACCGAG
CACTGTCTGGAATCGCCGCAGAACAGGATCGCAATACCCGGGAAGTCTTCGCCCAGGTGAAGCAG
ATGTACAAAACGCCCACTCTCAAGTATTTCGGCGGATTCAACTTTTCTCAGATTTTGCCTGACCC
GCTCAAGCCAACAAAACGATCTTTTATCGAAGACCTTCTGTTTAACAAGGTCACACTGGCGGATG
CTGGGTTCATGAAACAGTACGGTGAATGCCTGGGGGACATCAATGCCAGAGATCTGATCTGCGCC
CAGAAATTCAATGGCTTAACAGTCCTCCCACCTCTCTTGACCGACGATATGATCGCTGCGTACAC
CGCTGCTCTGGTATCGGGCACCGCGACTGCTGGCTGGACCTTTGGTGCCGGAGCCGCACTCCAGA
TCCCATTCGCCATGCAGATGGCCTACCGCTTCAACGGAATCGGGGTCACCCAGAACGTGCTGTAT
GAGAACCAGAAACAGATCGCCAATCAGTTCAATAAGGCAATTAGTCAGATTCAGGAGAGTCTTAC
CACTACCAGCACCGCCCTGGGCAAGCTGCAAGATGTTGTGAACCAGAATGCGCAGGCATTAAACA
CTCTGGTTAAACAGCTGAGCTCAAATTTTGGTGCAATCTCTTCAGTTCTGAACGATATCCTGAGT
CGGCTGGATCCGCCAGAGGCTGAAGTGCAAATTGATCGTTTGATCACCGGGAGGCTACAATCTCT
GCAGACGTACGTGACCCAGCAGCTCATCCGGGCAGCCGAAATTCGCGCATCAGCCAACCTCGCTG
CAACTAAGATGTCTGAGTGCGTGCTGGGCCAGAGTAAGAGGGTGGACTTTTGTGGTAAGGGATAC
CACCTCATGTCCTTTCCGCAAGCGGCTCCCCACGGCGTGGTTTTCTTACACGTTACCTATGTGCC
ATCCCAAGAACGCAATTTCACCACCGCTCCAGCTATCTGTCATGAGGGCAAAGCATATTTCCCCA
GGGAAGGAGTATTTGTGTTTAATGGCACGTCCTGGTTTATAACCCAACGTAACTTTTTCTCCCCA
CAGATTATCACAACCGACAACACATTCGTGTCTGGGAATTGTGACGTCGTGATCGGGATCATTAA
CAATACCGTTTACGATCCCTTGCAGCCCGAGCTTGACTCCTTTAAAGAGGAACTAGACAAATACT
TTAAGAATCACACCTCACCGGACGTAGATTTGGGAGACATCTCTGGAATTAATGCCTCTGTGGTG
AATATCCAGAAGGAGATCGACCGCCTGAATGAAGTCGCCAAGAACCTCAACGAGTCCCTGATAGA
TCTGCAAGAACTGGGCAAATATGAACAGTACATCAAATGGCCGTGGTACGTGTGGTTGGGCTTTA
TCGCTGGACTTATTGCAATCGTGATGGTGACGATTCTGCTCtgctgctccaatggatcgttacaa
tgcagaatttgcattTAA

[0490]

PDI-SARS-COV-1 H1cCT-DNA (서열 번호 92)

[0491]

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTCCGATCTGGATCGGTGCACTACATTTGACGATGTGCAGGCACCTAATTATACTCAGC
ACACCTCTTCCATGCGGGGGCGTGTACTACCCCGACGAGATTTTTAGAAGTGACACACTGTACCTG
ACCCAAGACCTTTTTCTCCCATTTTATAGCAATGTCACGGGATTCCACACTATCAATCACACATT
CGGAAACCCTGTTATCCCCTTTAAAGACGGTATCTACTTTGCTGCTACTGAGAAATCGAATGTTG
TTCGCGGTTGGGTGTTCGGCTCAACCATGAACAACAAGAGTCAGTCAGTAATAATTATAAACAAC
TCAACCAATGTCGTCATCAGGGCTTGCAACTTCGAGCTCTGCGATAACCCCTTCTTTGCGGTGTC
CAAACCGATGGGCACTCAGACCCATACCATGATTTTTGACAATGCCTTTAATTGTACTTTCGAAT
ACATCAGCGATGCCTTCTCGCTGGATGTGTCCGAGAAGAGCGGTAACTTCAAGCATTTGCGGGAG
TTCGTTTTCAAAAACAAAGACGGGTTTTTGTATGTCTACAAAGGCTATCAGCCCATTGATGTGGT
CAGGGATCTGCCCAGTGGTTTCAACACACTGAAGCCCATCTTCAAGTTGCCACTCGGCATCAACA
TCACTAATTTCCGCGCCATCCTAACTGCTTTTTCGCCAGCGCAGGATATTTGGGGGACATCCGCC
GCAGCTTACTTCGTGGGATATCTGAAGCCCACCACTTTTATGCTAAAGTACGACGAGAATGGCAC
CATCACCGATGCCGTGGACTGCTCACAGAATCCTCTCGCAGAGCTCAAGTGTTCAGTGAAGTCAT
TTGAGATCGACAAAGGGATCTACCAAACATCTAACTTTCGCGTCGTGCCTTCCGGTGACGTAGTC
CGTTTCCCAAATATCACTAACTTGTGCCCTTTTGGTGAAGTATTCAACGCAACCAAATTCCCCAG
TGTGTATGCGTGGGAGCGCAAAAAGATCAGTAACTGTGTGGCTGACTATAGTGTTCTGTACAATA
GCACCTTCTTCAGCACCTTCAAGTGTTATGGAGTGAGCGCTACAAAACTGAACGATCTTTGTTTC
TCAAACGTGTACGCCGATTCATTTGTCGTTAAAGGTGACGATGTGAGGCAGATCGCTCCAGGCCA
GACAGGTGTGATTGCTGACTATAATTACAAACTGCCAGACGACTTCATGGGGTGCGTGCTAGCTT
GGAATACAAGAAACATTGACGCCACCTCCACGGGAAATTACAATTACAAGTATCGTTACCTTCGC
CATGGAAAGTTGAGACCCTTCGAGCGTGATATAAGTAACGTGCCCTTTAGTCCAGATGGAAAACC
CTGCACACCCCCTGCTCTCAATTGCTATTGGCCTCTCAATGACTACGGCTTTTACACAACTACTG
GCATCGGATACCAGCCTTACCGGGTCGTGGTGCTCAGTTTTGAGTTGCTTAACGCACCCGCCACC
GTGTGTGGTCCTAAACTTTCTACTGACCTGATTAAAAACCAATGCGTCAACTTCAATTTTAACGG
GCTGACCGGCACCGGTGTCCTGACCCCTAGCTCTAAGAGATTCCAGCCTTTTCAGCAGTTCGGGA
GGGATGTGAGCGACTTTACCGACTCTGTCAGGGATCCAAAGACCAGCGAGATACTGGATATCTCG
CCCTGCAGTTTCGGTGGCGTGTCCGTTATTACACCTGGCACCAACGCCTCCTCAGAGGTGGCGGT
GCTCTATCAAGATGTCAACTGCACTGATGTGTCAACTGCCATCCATGCCGATCAGCTGACCCCCG
CCTGGCGCATCTACAGTACCGGGAACAACGTTTTTCAGACCCAGGCCGGCTGTCTAATCGGCGCA
GAGCACGTTGACACATCCTACGAATGTGACATACCTATCGGGGCAGGCATTTGCGCTAGCTACCA
TACCGTGTCACTGTTGGCTTCCACGTCACAAAAGTCAATCGTTGCCTACACGATGAGTCTGGGGG
CTGACTCATCTATCGCCTACAGCAACAATACCATTGCAATTCCCACAAACTTCAGTATCTCCATC
ACAACAGAGGTGATGCCCGTTTCTATGGCTAAAACATCAGTCGATTGCAATATGTATATATGCGG
CGATAGTACTGAGTGCGCCAATCTCTTGTTACAGTACGGCTCCTTTTGTACCCAGCTGAACCGAG
CACTGTCTGGAATCGCCGCAGAACAGGATCGCAATACCCGGGAAGTCTTCGCCCAGGTGAAGCAG
ATGTACAAAACGCCCACTCTCAAGTATTTCGGCGGATTCAACTTTTCTCAGATTTTGCCTGACCC
GCTCAAGCCAACAAAACGATCTTTTATCGAAGACCTTCTGTTTAACAAGGTCACACTGGCGGATG
CTGGGTTCATGAAACAGTACGGTGAATGCCTGGGGGACATCAATGCCAGAGATCTGATCTGCGCC
CAGAAATTCAATGGCTTAACAGTCCTCCCACCTCTCTTGACCGACGATATGATCGCTGCGTACAC
CGCTGCTCTGGTATCGGGCACCGCGACTGCTGGCTGGACCTTTGGTGCCGGAGCCGCACTCCAGA
TCCCATTCGCCATGCAGATGGCCTACCGCTTCAACGGAATCGGGGTCACCCAGAACGTGCTGTAT
GAGAACCAGAAACAGATCGCCAATCAGTTCAATAAGGCAATTAGTCAGATTCAGGAGAGTCTTAC
CACTACCAGCACCGCCCTGGGCAAGCTGCAAGATGTTGTGAACCAGAATGCGCAGGCATTAAACA
CTCTGGTTAAACAGCTGAGCTCAAATTTTGGTGCAATCTCTTCAGTTCTGAACGATATCCTGAGT
CGGCTGGATCCGCCAGAGGCTGAAGTGCAAATTGATCGTTTGATCACCGGGAGGCTACAATCTCT
GCAGACGTACGTGACCCAGCAGCTCATCCGGGCAGCCGAAATTCGCGCATCAGCCAACCTCGCTG
CAACTAAGATGTCTGAGTGCGTGCTGGGCCAGAGTAAGAGGGTGGACTTTTGTGGTAAGGGATAC
CACCTCATGTCCTTTCCGCAAGCGGCTCCCCACGGCGTGGTTTTCTTACACGTTACCTATGTGCC
ATCCCAAGAACGCAATTTCACCACCGCTCCAGCTATCTGTCATGAGGGCAAAGCATATTTCCCCA
GGGAAGGAGTATTTGTGTTTAATGGCACGTCCTGGTTTATAACCCAACGTAACTTTTTCTCCCCA
CAGATTATCACAACCGACAACACATTCGTGTCTGGGAATTGTGACGTCGTGATCGGGATCATTAA

CAATACCGTTTACGATCCCTTGCAGCCCGAGCTTGACTCCTTTAAAGAGGAACTAGACAAATACT
TTAAGAATCACACCTCACCGGACGTAGATTTGGGAGACATCTCTGGAATTAATGCCTCTGTGGTG
AATATCCAGAAGGAGATCGACCGCCTGAATGAAGTCGCCAAGAACCTCAACGAGTCCCTGATAGA
TCTGCAAGAACTGGGCAAATATGAACAGTACATCAAATGGCCGTGGTACGTGTGGTTGGGCTTTA
TCGCTGGACTTATTGCAATCGTGATGGTGACGATTCTGCTCagcttctggatgtgctctaatggg
tctctacagtgtagaatatgtattTAA

PDI-SARS-COV-1 wtTMCT-AA (서열 번호 93)

MAKNVAIFGLLFSLLVLVPSQIFASDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYL
TQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINN
STNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLRE
FVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSA
AAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVV
RFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCF
SNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLR
HGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPAT
VCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS
PCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGA
EHVDTSYECDIPIGAGICASYHTVSLLASTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISI
TTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQ
MYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICA
QKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILS
RLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY
HLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSP
QIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV
NIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCS
CLKGACSCGSCCKFDEDDSEPVLKGVKLHYT

PDI-SARS-COV-1 H5iTMCT-AA (서열 번호 94)

MAKNVAIFGLLFSLLVLVPSQIFASDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYL
TQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINN
STNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLRE
FVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSA
AAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVV
RFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCF
SNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLR
HGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPAT
VCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS
PCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGA
EHVDTSYECDIPIGAGICASYHTVSLLASTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISI
TTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQ
MYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICA
QKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILS
RLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY
HLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSP
QIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV

[0492]

NIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYQILSIYSTVASSLALAIMMAGLSLWMC
SNGSLQCRICI

PDI-SARS-COV-1 H5iCT-AA (서열 번호 95)

MAKNVAIFGLLFSLLVLVPSQIFASDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYL
TQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINN
STNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLRE
FVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSA
AAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVV
RFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCF
SNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLR
HGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPAT
VCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS
PCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGA
EHVDTSYECDIPIGAGICASYHTVSLLASTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISI
TTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQ
MYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICA
QKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILS
RLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY
HLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSP
QIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV
NIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLSLWMCSNG
SLQCRICI

PDI-SARS-COV-1 H5iCT(V4)-AA (서열 번호 96)

MAKNVAIFGLLFSLLVLVPSQIFASDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYL
TQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINN
STNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLRE
FVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSA
AAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVV
RFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCF
SNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLR
HGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPAT
VCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS
PCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGA
EHVDTSYECDIPIGAGICASYHTVSLLASTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISI
TTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQ
MYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICA
QKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILS
RLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY
HLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSP
QIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV
NIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCSNGSLQ
CRICI

[0493]

PDI-SARS-COV-1 H1cCT-AA (서열 번호 97)

MAKNVAIFGLLFSLLVLVPSQIFASDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYL
TQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINN
STNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLRE
FVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSA
AAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVV
RFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCF
SNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLR
HGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPAT
VCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS
PCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGA
EHVDTSYECDIPIGAGICASYHTVSLLASTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISI
TTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQ
MYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICA
QKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILS
RLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY
HLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSP
QIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV
NIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLSFWMCSNG
SLQCRICI

IF(AvB+wtCT-MERS).r (서열 번호 98)

ACGACACGACTAAGGCCTTCAGTGAACGTGGACCTTGTGAGGCTCAAGGTCATACTCCTC

IF(H1cCT-wtTM).r (서열 번호 99)

ACGACACGACTAAGGCCTTCAAATACATATTCTACACTGTAGAGACCCA

IF(H5ITMCT).r (서열 번호 100)

ACGACACGACTAAGGCCTTCAAATGCAAATTCTGCATTGTAACGATCC

PDI-MERS-wtTMCT-DNA (서열 번호 101)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTATGTCGATGTGGGTCCCGATAGTGTTAAGTCCGCCTGCATCGAAGTGGACATTCAGC
AGACCTTCTTCGATAAGACTTGGCCTCGGCCAATTGATGTGTCCAAGGCCGACGGCATTATCTAC
CCCCAAGGTCGGACATATTCCAACATAACTATCACCTATCAGGGGCTATTCCCTTATCAGGGCGA
CCATGGGGACATGTACGTTTACAGCGCTGGTCACGCTACAGGGACGACCCCCCAGAAGCTCTTCG
TGGCGAACTATAGTCAGGACGTGAAACAGTTTGCCAACGGTTTTGTAGTGCGCATCGGGGCAGCC
GCTAACTCCACTGGTACTGTTATTATCAGCCCTTCCACGAGTGCCACAATTCGAAAGATCTATCC
GGCCTTCATGCTAGGATCCTCTGTGGGCAATTTTAGCGACGGTAAGATGGGTCGGTTCTTCAACC
ACACGCTTGTGCTGCTTCCCGATGGGTGCGGTACTTTGCTGAGGGCCTTTTACTGTATCCTAGAG
CCCCGATCCGGCAACCACTGCCCCGCCGGGAACTCGTATACTTCCTTTGCCACTTATCATACTCC
AGCCACGGATTGTAGCGATGGGAACTACAATAGGAACGCCAGTTTGAATTCCTTTAAAGAGTACT
TCAACTTGCGGAATTGTACCTTCATGTATACATATAACATTACTGAGGACGAAATTCTCGAATGG
TTCGGAATCACTCAAACAGCCCAGGGAGTGCACCTCTTTAGTTCTCGCTATGTGGACTTATATGG
AGGCAATATGTTTCAATTCGCCACCTTACCCGTCTACGATACGATCAAGTATTACTCGATCATAC

[0494]

CCCACTCCATTAGGTCCATTCAGAGCGATCGCAAGGCATGGGCCGCATTCTATGTGTATAAGCTC
CAGCCCCTGACCTTCCTCTTGGATTTCTCCGTGGACGGCTACATCAGAAGGGCTATCGATTGCGG
GTTCAACGACCTCAGCCAGCTGCATTGTTCTTATGAGAGCTTTGACGTGGAAAGCGGAGTTTACT
CAGTCTCTTCCTTTGAGGCTAAACCTTCAGGTAGCGTCGTAGAGCAAGCAGAGGGTGTGGAGTGC
GATTTCTCACCACTGCTCAGCGGAACCCCACCCCAGGTCTACAACTTTAAGCGGCTCGTGTTCAC
AAACTGTAACTATAACTTGACTAAGTTGCTGTCACTCTTTTCCGTGAATGATTTTACATGCTCCC
AAATCAGCCCAGCCGCTATTGCGTCTAATTGCTATTCCTCATTGATCCTGGATTACTTCAGTTAC
CCCCTCTCTATGAAGAGCGATCTCTCGGTTAGTAGCGCTGGGCCTATTTCCCAGTTTAACTACAA
ACAATCCTTTTCCAATCCAACATGCCTGATCTTAGCTACTGTACCCCACAACCTGACTACTATTA
CGAAGCCACTCAAGTACTCATACATTAATAAGTGCAGCCGATTCCTCAGTGATGATCGCACCGAA
GTGCCGCAGCTTGTAAACGCGAACCAGTACTCCCCATGCGTCTCTATTGTGCCTTCTACAGTGTG
GGAAGACGGCGATTATTATAGAAAGCAGCTGTCGCCACTGGAAGGTGGCGGGTGGCTAGTTGCCA
GTGGGTCCACAGTTGCCATGACCGAGCAACTTCAGATGGGGTTTGGCATAACAGTGCAGTATGGT
ACCGATACGAACAGCGTGTGTCCAAAATTGGAATTTGCTAACGACACCAAGATCGCCTCCCAGTT
GGGAAATTGTGTTGAATATTCCCTGTACGGAGTGTCAGGCCGGGGGGTGTTCCAAAATTGCACCG
CCGTGGGAGTGAGGCAGCAAAGATTCGTGTACGACGCATACCAGAATCTAGTCGGATACTATTCT
GACGATGGAAACTACTACTGTCTGCGCGCTTGCGTCTCAGTGCCCGTGAGTGTCATATATGATAA
GGAGACCAAGACTCACGCTACTCTCTTTGGTTCTGTCGCGTGCGAACACATTTCCTCTACAATGT
CCCAGTATAGTCGCTCCACTCGGTCTATGTTAAAGCGCAGAGACAGTACCTACGGCCCTCTACAG
ACACCTGTGGGGTGCGTTCTCGGCCTTGTCAATTCTAGCCTGTTTGTGGAGGATTGTAAGCTGCC
CCTTGGTCAAAGCTTATGCGCACTGCCCGATACGCCCAGCACACTTACACCAGCTTCAGTGGGGT
CCGTCCCCGGGGAAATGAGATTGGCCTCGATCGCTTTCAACCACCCCATACAGGTGGATCAGCTC
AACTCGTCATACTTCAAGCTAAGCATCCCTACTAATTTCTCCTTTGGTGTGACTCAGGAGTACAT
TCAGACCACAATTCAAAAGGTGACCGTTGACTGCAAGCAGTATGTGTGCAACGGGTTCCAGAAAT
GTGAACAGCTGCTCCGGGAGTATGGCCAGTTCTGTTCTAAAATCAACCAGGCCCTCCACGGAGCA
AACCTTAGGCAGGACGATTCTGTCAGAAACCTCTTTGCCAGCGTCAAGAGTTCTCAGAGTTCCCC
TATTATACCTGGCTTCGGCGGGGATTTCAACCTGACACTACTTGAACCTGTAAGCATATCAACCG
GAAGTCGCAGTGCCCGTTCCGCCATCGAGGATCTGCTCTTCGACAAAGTAACTATTGCAGATCCC
GGATACATGCAGGGGTATGACGACTGCATGCAGCAGGGTCCAGCCTCTGCAAGGGATCTGATATG
CGCACAGTATGTCGCTGGGTACAAAGTGTTGCCTCCTCTCATGGACGTGAACATGGAAGCGGCCT
ATACCTCCTCACTTCTAGGCTCCATAGCGGGCGTGGGATGGACCGCAGGGCTTTCAAGCTTCGCC
GCAATTCCCTTTGCTCAATCTATCTTCTACAGGCTTAATGGCGTTGGAATCACCCAGCAGGTGTT
AAGCGAAAACCAGAAATTGATTGCCAATAAGTTTAACCAAGCTTTGGGGGCCATGCAGACAGGCT
TTACAACCACAAACGAGGCTTTCCATAAAGTACAGGATGCGGTAAACAATAACGCACAAGCCCTG
TCAAAGCTGGCTTCAGAGCTCTCAAATACATTTGGCGCTATATCCGCGTCTATCGGCGATATCAT
ACAACGGTTGGACCCACCCGAACAGGACGCACAGATTGATCGTTTGATCAACGGGAGGCTTACCA
CCTTAAACGCTTTTGTGGCCCAGCAACTGGTGCGGTCTGAGAGCGCCGCCTTGAGCGCTCAGCTG
GCAAAGGATAAAGTGAATGAATGCGTGAAAGCTCAATCAAAGAGAAGTGGGTTTTGTGGGCAGGG
TACTCATATTGTTTCCTTTGTGGTGAACGCCCCAAATGGACTCTACTTTATGCATGTTGGATACT
ACCCGAGCAACCACATCGAGGTCGTTTCCGCCTATGGGCTTTGTGACGCAGCAAACCCTACTAAC
TGTATCGCGCCAGTTAATGGCTACTTTATTAAAACAAATAACACACGCATTGTGGATGAATGGAG
TTACACAGGGTCCAGCTTCTACGCTCCAGAGCCTATCACCTCTCTGAACACAAAGTATGTGGCAC
CTCAGGTCACATATCAGAACATCTCGACAAACCTGCCCCCCCCACTCTTGGGCAACTCCACAGGG
ATCGACTTCCAGGACGAGCTTGACGAATTCTTCAAGAACGTGTCCACCAGTATCCCTAATTTTGG
TTCGCTGACCCAAATTAACACAACCCTGCTCGATCTGACATATGAAATGCTTTCACTACAGCAGG
TGGTCAAAGCGTTGAACGAGTCGTATATCGACCTGAAAGAGTTAGGGAATTACACATACTATAAC
AAATGGCCCTGGTATATTTGGTTAGGATTCATTGCCGGGCTGGTGGCCCTTGCCTTGTGCGTATT
TTTCATCTTGTGCTGTACCGGTTGCGGTACGAATTGCATGGGAAAACTGAAATGTAATCGGTGCT
GCGATCGCTATGAGGAGTATGACCTTGAGCCTCACAAGGTCCACGTTCACTGA

[0495]

PDI-MERS-H5iTMCT-DNA (서열 번호 102)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTATGTCGATGTGGGTCCCGATAGTGTTAAGTCCGCCTGCATCGAAGTGGACATTCAGC
AGACCTTCTTCGATAAGACTTGGCCTCGGCCAATTGATGTGTCCAAGGCCGACGGCATTATCTAC
CCCCAAGGTCGGACATATTCCAACATAACTATCACCTATCAGGGGCTATTCCCTTATCAGGGCGA
CCATGGGGACATGTACGTTTACAGCGCTGGTCACGCTACAGGGACGACCCCCCAGAAGCTCTTCG
TGGCGAACTATAGTCAGGACGTGAAACAGTTTGCCAACGGTTTTGTAGTGCGCATCGGGGCAGCC
GCTAACTCCACTGGTACTGTTATTATCAGCCCTTCCACGAGTGCCACAATTCGAAAGATCTATCC
GGCCTTCATGCTAGGATCCTCTGTGGGCAATTTTAGCGACGGTAAGATGGGTCGGTTCTTCAACC
ACACGCTTGTGCTGCTTCCCGATGGGTGCGGTACTTTGCTGAGGGCCTTTTACTGTATCCTAGAG
CCCCGATCCGGCAACCACTGCCCCGCCGGGAACTCGTATACTTCCTTTGCCACTTATCATACTCC
AGCCACGGATTGTAGCGATGGGAACTACAATAGGAACGCCAGTTTGAATTCCTTTAAAGAGTACT
TCAACTTGCGGAATTGTACCTTCATGTATACATATAACATTACTGAGGACGAAATTCTCGAATGG
TTCGGAATCACTCAAACAGCCCAGGGAGTGCACCTCTTTAGTTCTCGCTATGTGGACTTATATGG
AGGCAATATGTTTCAATTCGCCACCTTACCCGTCTACGATACGATCAAGTATTACTCGATCATAC
CCCACTCCATTAGGTCCATTCAGAGCGATCGCAAGGCATGGGCCGCATTCTATGTGTATAAGCTC
CAGCCCCTGACCTTCCTCTTGGATTTCTCCGTGGACGGCTACATCAGAAGGGCTATCGATTGCGG
GTTCAACGACCTCAGCCAGCTGCATTGTTCTTATGAGAGCTTTGACGTGGAAAGCGGAGTTTACT
CAGTCTCTTCCTTTGAGGCTAAACCTTCAGGTAGCGTCGTAGAGCAAGCAGAGGGTGTGGAGTGC
GATTTCTCACCACTGCTCAGCGGAACCCCACCCCAGGTCTACAACTTTAAGCGGCTCGTGTTCAC
AAACTGTAACTATAACTTGACTAAGTTGCTGTCACTCTTTTCCGTGAATGATTTTACATGCTCCC
AAATCAGCCCAGCCGCTATTGCGTCTAATTGCTATTCCTCATTGATCCTGGATTACTTCAGTTAC
CCCCTCTCTATGAAGAGCGATCTCTCGGTTAGTAGCGCTGGGCCTATTTCCCAGTTTAACTACAA
ACAATCCTTTTCCAATCCAACATGCCTGATCTTAGCTACTGTACCCCACAACCTGACTACTATTA
CGAAGCCACTCAAGTACTCATACATTAATAAGTGCAGCCGATTCCTCAGTGATGATCGCACCGAA
GTGCCGCAGCTTGTAAACGCGAACCAGTACTCCCCATGCGTCTCTATTGTGCCTTCTACAGTGTG
GGAAGACGGCGATTATTATAGAAAGCAGCTGTCGCCACTGGAAGGTGGCGGGTGGCTAGTTGCCA
GTGGGTCCACAGTTGCCATGACCGAGCAACTTCAGATGGGGTTTGGCATAACAGTGCAGTATGGT
ACCGATACGAACAGCGTGTGTCCAAAATTGGAATTTGCTAACGACACCAAGATCGCCTCCCAGTT
GGGAAATTGTGTTGAATATTCCCTGTACGGAGTGTCAGGCCGGGGGGTGTTCCAAAATTGCACCG
CCGTGGGAGTGAGGCAGCAAAGATTCGTGTACGACGCATACCAGAATCTAGTCGGATACTATTCT
GACGATGGAAACTACTACTGTCTGCGCGCTTGCGTCTCAGTGCCCGTGAGTGTCATATATGATAA
GGAGACCAAGACTCACGCTACTCTCTTTGGTTCTGTCGCGTGCGAACACATTTCCTCTACAATGT
CCCAGTATAGTCGCTCCACTCGGTCTATGTTAAAGCGCAGAGACAGTACCTACGGCCCTCTACAG
ACACCTGTGGGGTGCGTTCTCGGCCTTGTCAATTCTAGCCTGTTTGTGGAGGATTGTAAGCTGCC
CCTTGGTCAAAGCTTATGCGCACTGCCCGATACGCCCAGCACACTTACACCAGCTTCAGTGGGGT
CCGTCCCCGGGGAAATGAGATTGGCCTCGATCGCTTTCAACCACCCCATACAGGTGGATCAGCTC
AACTCGTCATACTTCAAGCTAAGCATCCCTACTAATTTCTCCTTTGGTGTGACTCAGGAGTACAT
TCAGACCACAATTCAAAAGGTGACCGTTGACTGCAAGCAGTATGTGTGCAACGGGTTCCAGAAAT
GTGAACAGCTGCTCCGGGAGTATGGCCAGTTCTGTTCTAAAATCAACCAGGCCCTCCACGGAGCA
AACCTTAGGCAGGACGATTCTGTCAGAAACCTCTTTGCCAGCGTCAAGAGTTCTCAGAGTTCCCC
TATTATACCTGGCTTCGGCGGGGATTTCAACCTGACACTACTTGAACCTGTAAGCATATCAACCG
GAAGTCGCAGTGCCCGTTCCGCCATCGAGGATCTGCTCTTCGACAAAGTAACTATTGCAGATCCC
GGATACATGCAGGGGTATGACGACTGCATGCAGCAGGGTCCAGCCTCTGCAAGGGATCTGATATG
CGCACAGTATGTCGCTGGGTACAAAGTGTTGCCTCCTCTCATGGACGTGAACATGGAAGCGGCCT
ATACCTCCTCACTTCTAGGCTCCATAGCGGGCGTGGGATGGACCGCAGGGCTTTCAAGCTTCGCC
GCAATTCCCTTTGCTCAATCTATCTTCTACAGGCTTAATGGCGTTGGAATCACCCAGCAGGTGTT
AAGCGAAAACCAGAAATTGATTGCCAATAAGTTTAACCAAGCTTTGGGGGCCATGCAGACAGGCT
TTACAACCACAAACGAGGCTTTCCATAAAGTACAGGATGCGGTAAACAATAACGCACAAGCCCTG
TCAAAGCTGGCTTCAGAGCTCTCAAATACATTTGGCGCTATATCCGCGTCTATCGGCGATATCAT
ACAACGGTTGGACCCACCCGAACAGGACGCACAGATTGATCGTTTGATCAACGGGAGGCTTACCA
CCTTAAACGCTTTTGTGGCCCAGCAACTGGTGCGGTCTGAGAGCGCCGCCTTGAGCGCTCAGCTG
GCAAAGGATAAAGTGAATGAATGCGTGAAAGCTCAATCAAAGAGAAGTGGGTTTTGTGGGCAGGG

[0496]

TACTCATATTGTTTCCTTTGTGGTGAACGCCCCAAATGGACTCTACTTTATGCATGTTGGATACT
ACCCGAGCAACCACATCGAGGTCGTTTCCGCCTATGGGCTTTGTGACGCAGCAAACCCTACTAAC
TGTATCGCGCCAGTTAATGGCTACTTTATTAAAACAAATAACACACGCATTGTGGATGAATGGAG
TTACACAGGGTCCAGCTTCTACGCTCCAGAGCCTATCACCTCTCTGAACACAAAGTATGTGGCAC
CTCAGGTCACATATCAGAACATCTCGACAAACCTGCCCCCCCCACTCTTGGGCAACTCCACAGGG
ATCGACTTCCAGGACGAGCTTGACGAATTCTTCAAGAACGTGTCCACCAGTATCCCTAATTTTGG
TTCGCTGACCCAAATTAACACAACCCTGCTCGATCTGACATATGAAATGCTTTCACTACAGCAGG
TGGTCAAAGCGTTGAACGAGTCGTATATCGACCTGAAAGAGTTAGGGAATTACACATACTATAAC
AAATGGCCCTGGTATcaaatactgtcaatttattcaacagtggcgagttccctagcactggcaat
catgatggctggtctatctttatggatgtgctccaatggatcgttacaatgcagaatttgcattT
GA

PDI-MERS-H5iCT-DNA (서열 번호 103)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTATGTCGATGTGGGTCCCGATAGTGTTAAGTCCGCCTGCATCGAAGTGGACATTCAGC
AGACCTTCTTCGATAAGACTTGGCCTCGGCCAATTGATGTGTCCAAGGCCGACGGCATTATCTAC
CCCCAAGGTCGGACATATTCCAACATAACTATCACCTATCAGGGGCTATTCCCTTATCAGGGCGA
CCATGGGGACATGTACGTTTACAGCGCTGGTCACGCTACAGGGACGACCCCCCAGAAGCTCTTCG
TGGCGAACTATAGTCAGGACGTGAAACAGTTTGCCAACGGTTTTGTAGTGCGCATCGGGGCAGCC
GCTAACTCCACTGGTACTGTTATTATCAGCCCTTCCACGAGTGCCACAATTCGAAAGATCTATCC
GGCCTTCATGCTAGGATCCTCTGTGGGCAATTTTAGCGACGGTAAGATGGGTCGGTTCTTCAACC
ACACGCTTGTGCTGCTTCCCGATGGGTGCGGTACTTTGCTGAGGGCCTTTTACTGTATCCTAGAG
CCCCGATCCGGCAACCACTGCCCCGCCGGGAACTCGTATACTTCCTTTGCCACTTATCATACTCC
AGCCACGGATTGTAGCGATGGGAACTACAATAGGAACGCCAGTTTGAATTCCTTTAAAGAGTACT
TCAACTTGCGGAATTGTACCTTCATGTATACATATAACATTACTGAGGACGAAATTCTCGAATGG
TTCGGAATCACTCAAACAGCCCAGGGAGTGCACCTCTTTAGTTCTCGCTATGTGGACTTATATGG
AGGCAATATGTTTCAATTCGCCACCTTACCCGTCTACGATACGATCAAGTATTACTCGATCATAC
CCCACTCCATTAGGTCCATTCAGAGCGATCGCAAGGCATGGGCCGCATTCTATGTGTATAAGCTC
CAGCCCCTGACCTTCCTCTTGGATTTCTCCGTGGACGGCTACATCAGAAGGGCTATCGATTGCGG
GTTCAACGACCTCAGCCAGCTGCATTGTTCTTATGAGAGCTTTGACGTGGAAAGCGGAGTTTACT
CAGTCTCTTCCTTTGAGGCTAAACCTTCAGGTAGCGTCGTAGAGCAAGCAGAGGGTGTGGAGTGC
GATTTCTCACCACTGCTCAGCGGAACCCCACCCCAGGTCTACAACTTTAAGCGGCTCGTGTTCAC
AAACTGTAACTATAACTTGACTAAGTTGCTGTCACTCTTTTCCGTGAATGATTTTACATGCTCCC
AAATCAGCCCAGCCGCTATTGCGTCTAATTGCTATTCCTCATTGATCCTGGATTACTTCAGTTAC
CCCCTCTCTATGAAGAGCGATCTCTCGGTTAGTAGCGCTGGGCCTATTTCCCAGTTTAACTACAA
ACAATCCTTTTCCAATCCAACATGCCTGATCTTAGCTACTGTACCCCACAACCTGACTACTATTA
CGAAGCCACTCAAGTACTCATACATTAATAAGTGCAGCCGATTCCTCAGTGATGATCGCACCGAA
GTGCCGCAGCTTGTAAACGCGAACCAGTACTCCCCATGCGTCTCTATTGTGCCTTCTACAGTGTG
GGAAGACGGCGATTATTATAGAAAGCAGCTGTCGCCACTGGAAGGTGGCGGGTGGCTAGTTGCCA
GTGGGTCCACAGTTGCCATGACCGAGCAACTTCAGATGGGGTTTGGCATAACAGTGCAGTATGGT
ACCGATACGAACAGCGTGTGTCCAAAATTGGAATTTGCTAACGACACCAAGATCGCCTCCCAGTT
GGGAAATTGTGTTGAATATTCCCTGTACGGAGTGTCAGGCCGGGGGGTGTTCCAAAATTGCACCG
CCGTGGGAGTGAGGCAGCAAAGATTCGTGTACGACGCATACCAGAATCTAGTCGGATACTATTCT
GACGATGGAAACTACTACTGTCTGCGCGCTTGCGTCTCAGTGCCCGTGAGTGTCATATATGATAA
GGAGACCAAGACTCACGCTACTCTCTTTGGTTCTGTCGCGTGCGAACACATTTCCTCTACAATGT
CCCAGTATAGTCGCTCCACTCGGTCTATGTTAAAGCGCAGAGACAGTACCTACGGCCCTCTACAG
ACACCTGTGGGGTGCGTTCTCGGCCTTGTCAATTCTAGCCTGTTTGTGGAGGATTGTAAGCTGCC
CCTTGGTCAAAGCTTATGCGCACTGCCCGATACGCCCAGCACACTTACACCAGCTTCAGTGGGGT
CCGTCCCCGGGGAAATGAGATTGGCCTCGATCGCTTTCAACCACCCCATACAGGTGGATCAGCTC
AACTCGTCATACTTCAAGCTAAGCATCCCTACTAATTTCTCCTTTGGTGTGACTCAGGAGTACAT
TCAGACCACAATTCAAAAGGTGACCGTTGACTGCAAGCAGTATGTGTGCAACGGGTTCCAGAAAT

[0497]

GTGAACAGCTGCTCCGGGAGTATGGCCAGTTCTGTTCTAAAATCAACCAGGCCCTCCACGGAGCA
AACCTTAGGCAGGACGATTCTGTCAGAAACCTCTTTGCCAGCGTCAAGAGTTCTCAGAGTTCCCC
TATTATACCTGGCTTCGGCGGGGATTTCAACCTGACACTACTTGAACCTGTAAGCATATCAACCG
GAAGTCGCAGTGCCCGTTCCGCCATCGAGGATCTGCTCTTCGACAAAGTAACTATTGCAGATCCC
GGATACATGCAGGGGTATGACGACTGCATGCAGCAGGGTCCAGCCTCTGCAAGGGATCTGATATG
CGCACAGTATGTCGCTGGGTACAAAGTGTTGCCTCCTCTCATGGACGTGAACATGGAAGCGGCCT
ATACCTCCTCACTTCTAGGCTCCATAGCGGGCGTGGGATGGACCGCAGGGCTTTCAAGCTTCGCC
GCAATTCCCTTTGCTCAATCTATCTTCTACAGGCTTAATGGCGTTGGAATCACCCAGCAGGTGTT
AAGCGAAAACCAGAAATTGATTGCCAATAAGTTTAACCAAGCTTTGGGGGCCATGCAGACAGGCT
TTACAACCACAAACGAGGCTTTCCATAAAGTACAGGATGCGGTAAACAATAACGCACAAGCCCTG
TCAAAGCTGGCTTCAGAGCTCTCAAATACATTTGGCGCTATATCCGCGTCTATCGGCGATATCAT
ACAACGGTTGGACCCACCCGAACAGGACGCACAGATTGATCGTTTGATCAACGGGAGGCTTACCA
CCTTAAACGCTTTTGTGGCCCAGCAACTGGTGCGGTCTGAGAGCGCCGCCTTGAGCGCTCAGCTG
GCAAAGGATAAAGTGAATGAATGCGTGAAAGCTCAATCAAAGAGAAGTGGGTTTTGTGGGCAGGG
TACTCATATTGTTTCCTTTGTGGTGAACGCCCCAAATGGACTCTACTTTATGCATGTTGGATACT
ACCCGAGCAACCACATCGAGGTCGTTTCCGCCTATGGGCTTTGTGACGCAGCAAACCCTACTAAC
TGTATCGCGCCAGTTAATGGCTACTTTATTAAAACAAATAACACACGCATTGTGGATGAATGGAG
TTACACAGGGTCCAGCTTCTACGCTCCAGAGCCTATCACCTCTCTGAACACAAAGTATGTGGCAC
CTCAGGTCACATATCAGAACATCTCGACAAACCTGCCCCCCCCCACTCTTGGGCAACTCCACAGGG
ATCGACTTCCAGGACGAGCTTGACGAATTCTTCAAGAACGTGTCCACCAGTATCCCTAATTTTGG
TTCGCTGACCCAAATTAACACAACCCTGCTCGATCTGACATATGAAATGCTTTCACTACAGCAGG
TGGTCAAAGCGTTGAACGAGTCGTATATCGACCTGAAAGAGTTAGGGAATTACACATACTATAAC
AAATGGCCCTGGTATATTTGGTTAGGATTCATTGCCGGGCTGGTGGCCCTTGCCTTGTGCGTATT
TTTCATCTTGtctttatggatgtgctccaatggatcgttacaatgcagaatttgcattTGA

PDI-MERS-H5iCT(V4)-DNA (서열 번호 104)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTATGTCGATGTGGGTCCCGATAGTGTTAAGTCCGCCTGCATCGAAGTGGACATTCAGC
AGACCTTCTTCGATAAGACTTGGCCTCGGCCAATTGATGTGTCCAAGGCCGACGGCATTATCTAC
CCCCAAGGTCGGACATATTCCAACATAACTATCACCTATCAGGGGCTATTCCCTTATCAGGGCGA
CCATGGGGACATGTACGTTTACAGCGCTGGTCACGCTACAGGGACGACCCCCCAGAAGCTCTTCG
TGGCGAACTATAGTCAGGACGTGAAACAGTTTGCCAACGGTTTTGTAGTGCGCATCGGGGCAGCC
GCTAACTCCACTGGTACTGTTATTATCAGCCCTTCCACGAGTGCCACAATTCGAAAGATCTATCC
GGCCTTCATGCTAGGATCCTCTGTGGGCAATTTTAGCGACGGTAAGATGGGTCGGTTCTTCAACC
ACACGCTTGTGCTGCTTCCCGATGGGTGCGGTACTTTGCTGAGGGCCTTTTACTGTATCCTAGAG
CCCCGATCCGGCAACCACTGCCCCGCCGGGAACTCGTATACTTCCTTTGCCACTTATCATACTCC
AGCCACGGATTGTAGCGATGGGAACTACAATAGGAACGCCAGTTTGAATTCCTTTAAAGAGTACT
TCAACTTGCGGAATTGTACCTTCATGTATACATATAACATTACTGAGGACGAAATTCTCGAATGG
TTCGGAATCACTCAAACAGCCCAGGGAGTGCACCTCTTTAGTTCTCGCTATGTGGACTTATATGG
AGGCAATATGTTTCAATTCGCCACCTTACCCGTCTACGATACGATCAAGTATTACTCGATCATAC
CCCACTCCATTAGGTCCATTCAGAGCGATCGCAAGGCATGGGCCGCATTCTATGTGTATAAGCTC
CAGCCCCTGACCTTCCTCTTGGATTTCTCCGTGGACGGCTACATCAGAAGGGCTATCGATTGCGG
GTTCAACGACCTCAGCCAGCTGCATTGTTCTTATGAGAGCTTTGACGTGGAAAGCGGAGTTTACT
CAGTCTCTTCCTTTGAGGCTAAACCTTCAGGTAGCGTCGTAGAGCAAGCAGAGGGTGTGGAGTGC
GATTTCTCACCACTGCTCAGCGGAACCCCACCCCAGGTCTACAACTTTAAGCGGCTCGTGTTCAC
AAACTGTAACTATAACTTGACTAAGTTGCTGTCACTCTTTTCCGTGAATGATTTTACATGCTCCC
AAATCAGCCCAGCCGCTATTGCGTCTAATTGCTATTCCTCATTGATCCTGGATTACTTCAGTTAC
CCCCTCTCTATGAAGAGCGATCTCTCGGTTAGTAGCGCTGGGCCTATTTCCCAGTTTAACTACAA
ACAATCCTTTTCCAATCCAACATGCCTGATCTTAGCTACTGTACCCCACAACCTGACTACTATTA
CGAAGCCACTCAAGTACTCATACATTAATAAGTGCAGCCGATTCCTCAGTGATGATCGCACCGAA

[0498]

GTGCCGCAGCTTGTAAACGCGAACCAGTACTCCCCATGCGTCTCTATTGTGCCTTCTACAGTGTG
GGAAGACGGCGATTATTATAGAAAGCAGCTGTCGCCACTGGAAGGTGGCGGGTGGCTAGTTGCCA
GTGGGTCCACAGTTGCCATGACCGAGCAACTTCAGATGGGGTTTGGCATAACAGTGCAGTATGGT
ACCGATACGAACAGCGTGTGTCCAAAATTGGAATTTGCTAACGACACCAAGATCGCCTCCCAGTT
GGGAAATTGTGTTGAATATTCCCTGTACGGAGTGTCAGGCCGGGGGGTGTTCCAAAATTGCACCG
CCGTGGGAGTGAGGCAGCAAAGATTCGTGTACGACGCATACCAGAATCTAGTCGGATACTATTCT
GACGATGGAAACTACTACTGTCTGCGCGCTTGCGTCTCAGTGCCCGTGAGTGTCATATATGATAA
GGAGACCAAGACTCACGCTACTCTCTTTGGTTCTGTCGCGTGCGAACACATTTCCTCTACAATGT
CCCAGTATAGTCGCTCCACTCGGTCTATGTTAAAGCGCAGAGACAGTACCTACGGCCCTCTACAG
ACACCTGTGGGGTGCGTTCTCGGCCTTGTCAATTCTAGCCTGTTTGTGGAGGATTGTAAGCTGCC
CCTTGGTCAAAGCTTATGCGCACTGCCCGATACGCCCAGCACACTTACACCAGCTTCAGTGGGGT
CCGTCCCCGGGGAAATGAGATTGGCCTCGATCGCTTTCAACCACCCCATACAGGTGGATCAGCTC
AACTCGTCATACTTCAAGCTAAGCATCCCTACTAATTTCTCCTTTGGTGTGACTCAGGAGTACAT
TCAGACCACAATTCAAAAGGTGACCGTTGACTGCAAGCAGTATGTGTGCAACGGGTTCCAGAAAT
GTGAACAGCTGCTCCGGGAGTATGGCCAGTTCTGTTCTAAAATCAACCAGGCCCTCCACGGAGCA
AACCTTAGGCAGGACGATTCTGTCAGAAACCTCTTTGCCAGCGTCAAGAGTTCTCAGAGTTCCCC
TATTATACCTGGCTTCGGCGGGGATTTCAACCTGACACTACTTGAACCTGTAAGCATATCAACCG
GAAGTCGCAGTGCCCGTTCCGCCATCGAGGATCTGCTCTTCGACAAAGTAACTATTGCAGATCCC
GGATACATGCAGGGGTATGACGACTGCATGCAGCAGGGTCCAGCCTCTGCAAGGGATCTGATATG
CGCACAGTATGTCGCTGGGTACAAAGTGTTGCCTCCTCTCATGGACGTGAACATGGAAGCGGCCT
ATACCTCCTCACTTCTAGGCTCCATAGCGGGCGTGGGATGGACCGCAGGGCTTTCAAGCTTCGCC
GCAATTCCCTTTGCTCAATCTATCTTCTACAGGCTTAATGGCGTTGGAATCACCCAGCAGGTGTT
AAGCGAAAACCAGAAATTGATTGCCAATAAGTTTAACCAAGCTTTGGGGGCCATGCAGACAGGCT
TTACAACCACAAACGAGGCTTTCCATAAAGTACAGGATGCGGTAAACAATAACGCACAAGCCCTG
TCAAAGCTGGCTTCAGAGCTCTCAAATACATTTGGCGCTATATCCGCGTCTATCGGCGATATCAT
ACAACGGTTGGACCCACCCGAACAGGACGCACAGATTGATCGTTTGATCAACGGGAGGCTTACCA
CCTTAAACGCTTTTGTGGCCCAGCAACTGGTGCGGTCTGAGAGCGCCGCCTTGAGCGCTCAGCTG
GCAAAGGATAAAGTGAATGAATGCGTGAAAGCTCAATCAAAGAGAAGTGGGTTTTGTGGGCAGGG
TACTCATATTGTTTCCTTTGTGGTGAACGCCCCAAATGGACTCTACTTTATGCATGTTGGATACT
ACCCGAGCAACCACATCGAGGTCGTTTCCGCCTATGGGCTTTGTGACGCAGCAAACCCTACTAAC
TGTATCGCGCCAGTTAATGGCTACTTTATTAAAACAAATAACACACGCATTGTGGATGAATGGAG
TTACACAGGGTCCAGCTTCTACGCTCCAGAGCCTATCACCTCTCTGAACACAAAGTATGTGGCAC
CTCAGGTCACATATCAGAACATCTCGACAAACCTGCCCCCCCCCACTCTTGGGCAACTCCACAGGG
ATCGACTTCCAGGACGAGCTTGACGAATTCTTCAAGAACGTGTCCACCAGTATCCCTAATTTTGG
TTCGCTGACCCAAATTAACACAACCCTGCTCGATCTGACATATGAAATGCTTTCACTACAGCAGG
TGGTCAAAGCGTTGAACGAGTCGTATATCGACCTGAAAGAGTTAGGGAATTACACATACTATAAC
AAATGGCCCTGGTATATTTGGTTAGGATTCATTGCCGGGCTGGTGGCCCTTGCCTTGTGCGTATT
TTTCATCTTGtgctgctccaatggatcgttacaatgcagaatttgcattTGA

PDI-MERS-H1cCT-DNA (서열 번호 105)

atggcgaaaaacgttgcgatttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgTATGTCGATGTGGGTCCCGATAGTGTTAAGTCCGCCTGCATCGAAGTGGACATTCAGC
AGACCTTCTTCGATAAGACTTGGCCTCGGCCAATTGATGTGTCCAAGGCCGACGGCATTATCTAC
CCCCAAGGTCGGACATATTCCAACATAACTATCACCTATCAGGGGCTATTCCCTTATCAGGGCGA
CCATGGGGACATGTACGTTTACAGCGCTGGTCACGCTACAGGGACGACCCCCCAGAAGCTCTTCG
TGGCGAACTATAGTCAGGACGTGAAACAGTTTGCCAACGGTTTTGTAGTGCGCATCGGGGCAGCC
GCTAACTCCACTGGTACTGTTATTATCAGCCCTTCCACGAGTGCCACAATTCGAAAGATCTATCC
GGCCTTCATGCTAGGATCCTCTGTGGGCAATTTTAGCGACGGTAAGATGGGTCGGTTCTTCAACC
ACACGCTTGTGCTGCTTCCCGATGGGTGCGGTACTTTGCTGAGGGCCTTTTACTGTATCCTAGAG
CCCCGATCCGGCAACCACTGCCCCGCCGGGAACTCGTATACTTCCTTTGCCACTTATCATACTCC

[0499]

AGCCACGGATTGTAGCGATGGGAACTACAATAGGAACGCCAGTTTGAATTCCTTTAAAGAGTACT
TCAACTTGCGGAATTGTACCTTCATGTATACATATAACATTACTGAGGACGAAATTCTCGAATGG
TTCGGAATCACTCAAACAGCCCAGGGAGTGCACCTCTTTAGTTCTCGCTATGTGGACTTATATGG
AGGCAATATGTTTCAATTCGCCACCTTACCCGTCTACGATACGATCAAGTATTACTCGATCATAC
CCCACTCCATTAGGTCCATTCAGAGCGATCGCAAGGCATGGGCCGCATTCTATGTGTATAAGCTC
CAGCCCCTGACCTTCCTCTTGGATTTCTCCGTGGACGGCTACATCAGAAGGGCTATCGATTGCGG
GTTCAACGACCTCAGCCAGCTGCATTGTTCTTATGAGAGCTTTGACGTGGAAAGCGGAGTTTACT
CAGTCTCTTCCTTTGAGGCTAAACCTTCAGGTAGCGTCGTAGAGCAAGCAGAGGGTGTGGAGTGC
GATTTCTCACCACTGCTCAGCGGAACCCCACCCCAGGTCTACAACTTTAAGCGGCTCGTGTTCAC
AAACTGTAACTATAACTTGACTAAGTTGCTGTCACTCTTTTCCGTGAATGATTTTACATGCTCCC
AAATCAGCCCAGCCGCTATTGCGTCTAATTGCTATTCCTCATTGATCCTGGATTACTTCAGTTAC
CCCCTCTCTATGAAGAGCGATCTCTCGGTTAGTAGCGCTGGGCCTATTTCCCAGTTTAACTACAA
ACAATCCTTTTCCAATCCAACATGCCTGATCTTAGCTACTGTACCCCACAACCTGACTACTATTA
CGAAGCCACTCAAGTACTCATACATTAATAAGTGCAGCCGATTCCTCAGTGATGATCGCACCGAA
GTGCCGCAGCTTGTAAACGCGAACCAGTACTCCCCATGCGTCTCTATTGTGCCTTCTACAGTGTG
GGAAGACGGCGATTATTATAGAAAGCAGCTGTCGCCACTGGAAGGTGGCGGGTGGCTAGTTGCCA
GTGGGTCCACAGTTGCCATGACCGAGCAACTTCAGATGGGGTTTGGCATAACAGTGCAGTATGGT
ACCGATACGAACAGCGTGTGTCCAAAATTGGAATTTGCTAACGACACCAAGATCGCCTCCCAGTT
GGGAAATTGTGTTGAATATTCCCTGTACGGAGTGTCAGGCCGGGGGGTGTTCCAAAATTGCACCG
CCGTGGGAGTGAGGCAGCAAAGATTCGTGTACGACGCATACCAGAATCTAGTCGGATACTATTCT
GACGATGGAAACTACTACTGTCTGCGCGCTTGCGTCTCAGTGCCCGTGAGTGTCATATATGATAA
GGAGACCAAGACTCACGCTACTCTCTTTGGTTCTGTCGCGTGCGAACACATTTCCTCTACAATGT
CCCAGTATAGTCGCTCCACTCGGTCTATGTTAAAGCGCAGAGACAGTACCTACGGCCCTCTACAG
ACACCTGTGGGGTGCGTTCTCGGCCTTGTCAATTCTAGCCTGTTTGTGGAGGATTGTAAGCTGCC
CCTTGGTCAAAGCTTATGCGCACTGCCCGATACGCCCAGCACACTTACACCAGCTTCAGTGGGGT
CCGTCCCCGGGGAAATGAGATTGGCCTCGATCGCTTTCAACCACCCCATACAGGTGGATCAGCTC
AACTCGTCATACTTCAAGCTAAGCATCCCTACTAATTTCTCCTTTGGTGTGACTCAGGAGTACAT
TCAGACCACAATTCAAAAGGTGACCGTTGACTGCAAGCAGTATGTGTGCAACGGGTTCCAGAAAT
GTGAACAGCTGCTCCGGGAGTATGGCCAGTTCTGTTCTAAAATCAACCAGGCCCTCCACGGAGCA
AACCTTAGGCAGGACGATTCTGTCAGAAACCTCTTTGCCAGCGTCAAGAGTTCTCAGAGTTCCCC
TATTATACCTGGCTTCGGCGGGGATTTCAACCTGACACTACTTGAACCTGTAAGCATATCAACCG
GAAGTCGCAGTGCCCGTTCCGCCATCGAGGATCTGCTCTTCGACAAAGTAACTATTGCAGATCCC
GGATACATGCAGGGGTATGACGACTGCATGCAGCAGGGTCCAGCCTCTGCAAGGGATCTGATATG
CGCACAGTATGTCGCTGGGTACAAAGTGTTGCCTCCTCTCATGGACGTGAACATGGAAGCGGCCT
ATACCTCCTCACTTCTAGGCTCCATAGCGGGCGTGGGATGGACCGCAGGGCTTTCAAGCTTCGCC
GCAATTCCCTTTGCTCAATCTATCTTCTACAGGCTTAATGGCGTTGGAATCACCCAGCAGGTGTT
AAGCGAAAACCAGAAATTGATTGCCAATAAGTTTAACCAAGCTTTGGGGGCCATGCAGACAGGCT
TTACAACCACAAACGAGGCTTTCCATAAAGTACAGGATGCGGTAAACAATAACGCACAAGCCCTG
TCAAAGCTGGCTTCAGAGCTCTCAAATACATTTGGCGCTATATCCGCGTCTATCGGCGATATCAT
ACAACGGTTGGACCCACCCGAACAGGACGCACAGATTGATCGTTTGATCAACGGGAGGCTTACCA
CCTTAAACGCTTTTGTGGCCCAGCAACTGGTGCGGTCTGAGAGCGCCGCCTTGAGCGCTCAGCTG
GCAAAGGATAAAGTGAATGAATGCGTGAAAGCTCAATCAAAGAGAAGTGGGTTTTGTGGGCAGGG
TACTCATATTGTTTCCTTTGTGGTGAACGCCCCAAATGGACTCTACTTTATGCATGTTGGATACT
ACCCGAGCAACCACATCGAGGTCGTTTCCGCCTATGGGCTTTGTGACGCAGCAAACCCTACTAAC
TGTATCGCGCCAGTTAATGGCTACTTTATTAAAACAAATAACACACGCATTGTGGATGAATGGAG
TTACACAGGGTCCAGCTTCTACGCTCCAGAGCCTATCACCTCTCTGAACACAAAGTATGTGGCAC
CTCAGGTCACATATCAGAACATCTCGACAAACCTGCCCCCCCCCACTCTTGGGCAACTCCACAGGG
ATCGACTTCCAGGACGAGCTTGACGAATTCTTCAAGAACGTGTCCACCAGTATCCCTAATTTTGG
TTCGCTGACCCAAATTAACACAACCCTGCTCGATCTGACATATGAAATGCTTTCACTACAGCAGG
TGGTCAAAGCGTTGAACGAGTCGTATATCGACCTGAAAGAGTTAGGGAATTACACATACTATAAC
AAATGGCCCTGGTATATTTGGTTAGGATTCATTGCCGGGCTGGTGGCCCTTGCCTTGTGCGTATT
TTTCATCTTGagcttctggatgtgctctaatgggtctctacagtgtagaatatgtattTGA

[0500]

PDI-MERS-wtTMCT-AA (서열 번호 106)

MAKNVAIFGLLFSLLVLVPSQIFAYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIY
PQGRTYSNITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAA
ANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILE
PRSGNHCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEW
FGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKL
QPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVEC
DFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSY
PLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTE
VPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYG
TDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYS
DDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQ
TPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVDQL
NSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGA
NLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADP
GYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFA
AIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQAL
SKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQL
AKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTN
CIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTG
IDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYN
KWPWYIWLGFIAGLVALALCVFFILCCTGCGTNCMGKLKCNRCCDRYEEYDLEPHKVHVH

PDI-MERS-H5iTMCT-AA (서열 번호 107)

MAKNVAIFGLLFSLLVLVPSQIFAYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIY
PQGRTYSNITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAA
ANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILE
PRSGNHCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEW
FGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKL
QPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVEC
DFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSY
PLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTE
VPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYG
TDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYS
DDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQ
TPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVDQL
NSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGA
NLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADP
GYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFA
AIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQAL
SKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQL
AKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTN
CIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTG
IDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYN
KWPWYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

[0501]

PDI-MERS-H5iCT-AA (서열 번호 108)

MAKNVAIFGLLFSLLVLVPSQIFAYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIY
PQGRTYSNITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAA
ANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILE
PRSGNHCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEW
FGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKL
QPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVEC
DFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSY
PLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTE
VPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYG
TDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYS
DDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQ
TPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVDQL
NSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGA
NLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADP
GYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFA
AIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQAL
SKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQL
AKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTN
CIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTG
IDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYN
KWPWYIWLGFIAGLVALALCVFFILSLWMCSNGSLQCRICI

PDI-MERS-H5iCT(V4)-AA (서열 번호 109)

MAKNVAIFGLLFSLLVLVPSQIFAYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIY
PQGRTYSNITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAA
ANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILE
PRSGNHCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEW
FGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKL
QPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVEC
DFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSY
PLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTE
VPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYG
TDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYS
DDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQ
TPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVDQL
NSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGA
NLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADP
GYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFA
AIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQAL
SKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQL
AKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTN
CIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTG
IDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYN
KWPWYIWLGFIAGLVALALCVFFILCCSNGSLQCRICI

[0502]

PDI-MERS-H1cCT-AA (서열 번호 110)

MAKNVAIFGLLFSLLVLVPSQIFAYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIY
PQGRTYSNITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAA
ANSTGTVIISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILE
PRSGNHCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEW
FGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKL
QPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVEC
DFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSY
PLSMKSDLSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTE
VPQLVNANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYG
TDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYS
DDGNYYCLRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQ
TPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVDQL
NSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGA
NLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADP
GYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFA
AIPFAQSIFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQAL
SKLASELSNTFGAISASIGDIIQRLDPPEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQL
AKDKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTN
CIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTG
IDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYN
KWPWYIWLGFIAGLVALALCVFFILSFWMCSNGSLQCRICI

클로닝 벡터 7147 (좌측 T-DNA에서 우측 T-DNA) (서열 번호 111)

[0503]

tggcaggatatattgtggtgtaaacaaattgacgcttagacaacttaataacacattgcggacgt
ttttaatgtactgaattaacgccgaatcccgggctggtatatttatatgttgtcaaataactcaa
aaaccataaaagtttaagttagcaagtgtgtacatttttacttgaacaaaaatattcacctacta
ctgttataaatcattattaaacattagagtaaagaaatatggatgataagaacaagagtagtgat
attttgacaacaattttgttgcaacatttgagaaaatttttgttgttctctctttttcattggtcaa
aaacaatagagagagaaaaaggaagagggagaataaaaacataatgtgagtatgagagagaaagt
tgtacaaaagttgtaccaaaatagttgtacaaatatcattgaggaatttgacaaaagctacacaa
ataagggttaattgctgtaaataaataaggatgacgcattagagagatgtaccattagagaattt
ttggcaagtcattaaaaagaaagaataaattatttttaaaattaaaagttgagtcatttgattaa
acatgtgattatttaatgaattgatgaaagagttggattaaagttgtattagtaattagaatttg
gtgtcaaatttaatttgacatttgatcttttcctatatattgccccatagagtcagttaactcat
ttttatatttcatagatcaaataagagaaataacggtatattaatccctccaaaaaaaaaaaacg
gtatatttactaaaaaatctaagccacgtaggaggataacaggatccccgtaggaggataacatc
caatccaaccaatcacaacaatcctgatgagataacccactttaagcccacgcatctgtggcaca
tctacattatctaaatcacacattcttccacacatctgagccacacaaaaaccaatccacatctt
tatcacccattctataaaaaatcacactttgtgagtctacactttgattcccttcaaacacatac
aaagagaagagactaattaattaattaatcatcttgagagaaaatggaacgagctatacaaggaa
acgacgctagggaacaagctaacagtgaacgttgggatggaggatcaggaggtaccacttctccc
ttcaaacttcctgacgaaagtccgagttggactgagtggcggctacataacgatgagacgaattc
gaatcaagataatccccttggtttcaaggaaagctggggtttcgggaaagttgtatttaagagat
atctcagatacgacaggacggaagcttcactgcacagagtccttggatcttggacgggagattcg
gttaactatgcagcatctcgatttttcggtttcgaccagatcggatgtacctatagtattcggtt
tcgaggagttagtatcaccgtttctggagggtcgcgaactcttcagcatctctgtgagatggcaa
ttcggtctaagcaagaactgctacagcttgccccaatcgaagtggaaagtaatgtatcaagagga
tgccctgaaggtactcaaaccttcgaaaaagaaagcgagtaagttaaaatgcttcttcgtctcct
atttataatatggtttgttattgttaattttgttcttgtagaagagcttaattaatcgttgttgt
tatgaaatactatttgtatgagatgaactggtgtaatgtaattcatttacataagtggagtcaga

atcagaatgtttcctccataactaactagacatgaagacctgccgcgtacaattgtcttatattt
gaacaactaaaattgaacatcttttgccacaactttataagtggttaatatagctcaaatatatg
gtcaagttcaatagattaataatggaaatatcagttatcgaaattcattaacaatcaacttaacg
ttattaactactaattttatatcatccccctttgataaatgatagtacaccaattaggaaggagca
tgctcgcctaggagattgtcgtttcccgccttcagtttgcaagctgctctagccgtgtagccaat
acgcaaaccgcctctccccgcgcgttgggaattactagcgcgtgtcgacaagcttgcatgccggt
caacatggtggagcacgacacacttgtctactccaaaaatatcaaagatacagtctcagaagacc
aaagggcaattgagacttttcaacaaagggtaatatccggaaacctcctcggattccattgccca
gctatctgtcactttattgtgaagatagtggaaaaggaaggtggctcctacaaatgccatcattg
cgataaaggaaaggccatcgttgaagatgcctctgccgacagtggtcccaaagatggacccccac
ccacgaggagcatcgtggaaaaagaagacgttccaaccacgtcttcaaagcaagtggattgatgt
gataacatggtggagcacgacacacttgtctactccaaaaatatcaaagatacagtctcagaaga
ccaaagggcaattgagacttttcaacaaagggtaatatccggaaacctcctcggattccattgcc
cagctatctgtcactttattgtgaagatagtggaaaaggaaggtggctcctacaaatgccatcat
tgcgataaaggaaaggccatcgttgaagatgcctctgccgacagtggtcccaaagatggacccccc
acccacgaggagcatcgtggaaaaagaagacgttccaaccacgtcttcaaagcaagtggattgat
gtgatatctccactgacgtaagggatgacgcacaatcccactatccttcgcaagacccttcctct
atataaggaagttcatttcatttggagaggcactccatttgaatctatcaaaccaaaacacattg
agacgtcacgtactcctcagccaaaacgacacccccatctgtctatccactggcccctggatctg
ctgcccaaactaactccatggtgaccctgggatgcctggtcaagggctatttccctgagccagtg
acagtgacctggaactctggatccctgtccagcggtgtgcacaccttcccagctgtcctgcagtc
tgacctctacactctgagcagctcagtgactgtcccctccagcacctggcccagcgagaccgtca
cctgcaacgttgcccacccggccagcagcaccaaggtggacaagaaaattgtgcccagggattgt
ggttgtaagccttgcatatgtacagtcccagaagtatcatctgtcttcatcttccccccaaagcc
caaggatgtgctcaccattactctgactcctaaggtcacgtgtgttgtggtagacatcagcaagg
atgatcccgaggtccagttcagctggtttgtagatgatgtggaggtgcacacagctcagacgcaa
ccccgggaggagcagttcaacagcactttccgctcagtcagtgaacttcccatcatgcaccagga
ctggctcaatggcaaggagacgtccagattttggcgatctattcaactgtcgccagttcattggt
actggtagtctccctgggggcaatcagtttctggatgtgctctaatgggtctctacagtgtagaa
tatgtatttaaaggccttagtcgtgtcgtttttcaaataatataatcctttagggttttagttta
gtttaaattttctgttgctcctgtttagcaggtcgtgccttcagcaagcacacaaaaacagagtg
tttattttaagttgtttgtttagtgattcaaaaaaaaaatcgttcaaacatttggcaataaagtt
tcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttgaattacgtt
aagcatgtaataattaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagt
cccgcaattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaattat
cgcgcgcggtgtcatctatgttactagatctctagagtctcaagcttggcgcgcccacgtgacta
gtggcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcg
ccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttc
ccaacagttgcgcagcctgaatggcgaatgctagagcagcttgagcttggatcagattgtcgtt
tcccgccttcagtttaaactatcagtgtttgacaggatatattggcgggtaaacctaagagaaaa
gagcgttta

천연 SARS-CoV-1 S 단백질 wtTM/CT AA P59594 (서열 번호 112)

MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSN
VTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNF
ELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLY
VYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPT
TFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPF
GEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVK
GDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDI
SNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLI

[0504]

KNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVSVIT
PGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDI
PIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAK
TSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFG
GFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPP
LLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFN
KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQI
DRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPH
GVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVS
GNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNE
VAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSC
CKFDEDDSEPVLKGVKLHYT

천연 MERS S 단백질 wtTM/CT AA AFY13307 (서열 번호 113)

MIHSVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIYPQGRTYS
NITITYQGLFPYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTV
IISPSTSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHC
PAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWFGITQTA
QGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKLQPLTFLL
DFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVECDFSPLLS
GTPPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSD
LSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTEVPQLVNA
NQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVC
PKLEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDDGNYYC
LRACVSVPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCVL
GLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPRSVRSVPGEMRLASIAFNHPIQVDQLNSSYFKL
SIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGANLRQDDS
VRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADPGYMQGYD
DCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQS
IFYRLNGVGITQQVLSENQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQALSKLASEL
SNTFGAISASIGDIIQRLDVLEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQLAKDKVNE
CVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTNCIAPVNG
YFIKTNNTRIVDEWSYTGSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTGIDFQDEL
DEFFKNVSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYIW
LGFIAGLVALALCVFFILCCTGCGTNCMGKLKCNRCCDRYEEYDLEPHKVHVH

신호 펩티드를 포함하지 않는 천연 SARS-CoV-1 S 단백질 wtTM/CT AA P59594
(서열 번호 114)

SDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGN
PVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKP
MGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRD
LPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTIT
DAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVY
AWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTG
VIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCT
PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLT
GTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLY
QDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTV
SLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDS
TECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLK
PTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAA

[0505]

LVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTT
STALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQT
YVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQ
ERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNT
VYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQ
ELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLK
GVKLHYT

신호 펩티드를 포함하지 않는 천연 MERS S 단백질 wtTM/CT AA AFY13307 (서열 번호 115)

YVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKADGIIYPQGRTYSNITITYQGLFPYQGDHG
DMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKIYPAF
MLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSFATYHTPAT
DCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWFGITQTAQGVHLFSSRYVDLYGGN
MFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAWAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFN
DLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVEQAEGVECDFSPLLSGTPPQVYNFKRLVFTNC
NYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQS
FSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTEVPQLVNANQYSPCVSIVPSTVWED
GDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVCPKLEFANDTKIASQLGN
CVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDDGNYYCLRACVSVPVSVIYDKET
KTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDSTYGPLQTPVGCVLGLVNSSLFVEDCKLPLG
QSLCALPDTPSTLTPRSVRSVPGEMRLASIAFNHPIQVDQLNSSYFKLSIPTNFSFGVTQEYIQT
TIQKVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGANLRQDDSVRNLFASVKSSQSSPII
PGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTIADPGYMQGYDDCMQQGPASARDLICAQ
YVAGYKVLPPLMDVNMEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSE
NQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQALSKLASELSNTFGAISASIGDIIQR
LDVLEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQLAKDKVNECVKAQSKRSGFCGQGTH
IVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTNCIAPVNGYFIKTNNTRIVDEWSYT
GSSFYAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNSTGIDFQDELDEFFKNVSTSIPNFGSL
TQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYIWLGFIAGLVALALCVFFI
LCCTGCGTNCMGKLKCNRCCDRYEEYDLEPHKVHVH

변형된 PDI-SARS-COV-1 wtTMCT의 TMCT 영역-AA (서열 번호 116)

WYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT

변형된 PDI-SARS-COV-1 H5iTMCT의 TMCT 영역-AA (서열 번호 117)

WYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

변형된 PDI-SARS-COV-1 H5iCT의 TMCT 영역-AA (서열 번호 118)

WYVWLGFIAGLIAIVMVTILLSLWMCSNGSLQCRICI

변형된 PDI-SARS-COV-1 H5iCT(V4)의 TMCT 영역-AA (서열 번호 119)

WYVWLGFIAGLIAIVMVTILLCCSNGSLQCRICI

변형된 PDI-SARS-COV-1 H1cCT의 TMCT 영역-AA (서열 번호 120)

WYVWLGFIAGLIAIVMVTILLSFWMCSNGSLQCRICI

[0506]

변형된 PDI-MERS-wtTMCT의 TMCT 영역-AA (서열 번호 121)

WYIWLGFIAGLVALALCVFFILCCTGCGTNCMGKLKCNRCCDRYEEYDLEPHKVHVH

변형된 PDI-MERS-H5iTMCT의 TMCT 영역-AA (서열 번호 122)

WYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

변형된 PDI-MERS-H5iCT의 TMCT 영역-AA (서열 번호 123)

WYIWLGFIAGLVALALCVFFILSLWMCSNGSLQCRICI

변형된 PDI-MERS-H5iCT(V4)의 TMCT 영역-AA (서열 번호 124)

WYIWLGFIAGLVALALCVFFILCCSNGSLQCRICI

변형된 PDI-MERS-H1cCT의 TMCT 영역-AA (서열 번호 125)

WYIWLGFIAGLVALALCVFFILSFWMCSNGSLQCRICI

변형된 PDI-S-단백질+H1 Cal의 TMCT 영역 (서열 번호 126)

WYIWLGFIAGLIAIVMVTIMLSFWMCSNGSLQCRICI

변형된 PDI-S-단백질+H3 Minn의 TMCT 영역 (서열 번호 127)

WYIWLGFIAGLIAIVMVTIMLMWACQKGNIRCNICI

변형된 PDI-S-단백질+H6 HK의 TMCT 영역 (서열 번호 128)

WYIWLGFIAGLIAIVMVTIMLGLWMCSNGSMQCRICI

변형된 PDI-S-단백질+H7 Guangdong의 TMCT 영역 (서열 번호 129)

WYIWLGFIAGLIAIVMVTIMLVFICVKNGNMRCTICI

변형된 PDI-S-단백질+H9 HK의 TMCT 영역 (서열 번호 130)

WYIWLGFIAGLIAIVMVTIMLLFWAMSNGSCRCNICI

변형된 PDI-S-단백질+ B/Wash의 TMCT 영역 (서열 번호 131)

WYIWLGFIAGLIAIVMVTIMLVVYMVSRDNVSCSICL

코로나바이러스 S 단백질의 TM 도메인의 컨센서스 서열 (서열 번호 132)

WYXWLGFIAGLXAXXX{X}VXXXL ({X}는 부재할 수 있다)

코로나바이러스 S 단백질의 TM 도메인의 컨센서스 서열 (서열 번호 133)

WY[I/V]WLGFIAGL[V/I]A[L/I][A/V][L/M]{X}V[F/T][F/I]XL

[0507] (여기서, {X}는 C일 수 있거나, 또는 부재할 수 있다)

개재 펩티드 서열 X_n을 포함하는 변형된 SARS-CoV-1 S 단백질의 TM/CT 영역 (서열 번호 134)

WYVWLGFIAGLIAIVMVTIL - (X)n - CSNGSXXCXICI

개재 펩티드 서열 X_n을 포함하는 변형된 MERS S 단백질의 TM/CT 영역 (서열 번호 135)

WYIWLGFIAGLVALALCVFFIL -(X)n - CSNGSXXCXICI

IF(AvB+wtCT-OC43).r ((서열 번호 136)

ACGACACGACTAAGGCCTTCAGTCGTCATGCGAGGTCTTAATGACAAGC

PDI-OC43-wtTMCT-DNA (서열 번호 137)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgGTGATCGGCGATCTGAATTGTACCCTGGATCCCCGCCTGAAAGGGAGCTTTAACAACC
GAGATACAGGACCCCCGTCTATATCCATAGATACAGTGGATGTTACGAACGGGCTCGGCACCTAC
TATGTGCTAGACCGAGTTTATTTGAACACCACCTTATTCCTCAATGGATACTACCCAACTTCAGG
TAGTACTTACAGAAACATGGCGCTGAAGGGTACGGATCTGCTGAGCACCCTATGGTTTAAACCTC
CCTTCCTCTCGGACTTTATTAATGGCATCTTCGCTAAGGTGAAAAACACGAAGGTTTTCAAAGAT
GGAGTGATGTATTCAGAGTTCCCTGCGATCACCATTGGAAGTACCTTCGTGAATACTTCCTATAG
CGTGGTGGTTCAACCACGGACAATCAACTCCACCCAGGACGGCGTCAACAAGCTCCAGGGATTGC
TGGAGGTGTCAGTCTGTCAATATAACATGTGTGAGTACCCACACACTATCTGTCACCCTAATCTA
GGCAACCACTTTAAGGAACTGTGGCACTACGATACGGGGGTGGTAAGTTGCTTATATAAGAGAAA
TTTCACCTATGATGTTAATGCAACGTACCTGTACTTTCACTTCTATCAAGAAGGAGGAACTTTCT
ACGCATATTTCACAGATACCGGCTTTGTGACGAAATTCTTATTCAACGTTTACCTCGGAATGGCA
TTAAGCCATTATTACGTGATGCCTCTCACTTGCATCAGACGCCCTAAGGATGGTTTTTCTCTGGA
GTACTGGGTCACTCCCCTGACACCACGGCAGTACCTGCTTGCTTTTAACCAGGACGGTATCATTT
TTAATGCCGTCGATTGTATGAGCGATTTTATGAGCGAGATAAAGTGCAAGACCCAATCTATTGCT
CCGCCCACGGGGGTGTACGAACTGAATGGTTACACCGTCCAGCCCGTTGCCGATGTATATAGACG
GAAACCAGACCTGCCCAATTGCAACATCGAAGCTTGGTTAAACGATAAGTCAGTGCCCTCCCCCC
TCAATTGGGAGAGGAAGACTTTCTCCAACTGTAATTTCAACATGTCAAGCCTGATGTCTTTCATT
CAAGCCGATTCGTTCACTTGTAATAATATAGATGCAGCAAAGATCTATGGTATGTGCTTCAGTTC
CATCACAATAGATAAGTTTGCAATACCAAACCGTCGCAAGGTGGACCTTCAGCTCGGCAACCTGG
GCTATCTGCAGTCCAGCAATTATAGAATAGACACCACCGCCACATCATGTCAGCTGTACTATAAC
CTCCCAGCAGCGAACGTCAGTGTTAGTAGGTTCAATCCTTCTACCTGGAATAAAAGGTTTGGATT
CATCGAAGATAGTGTGTTCGTACCTCAGCCAACAGGAGTGTTCACCAATCACAGCGTGGTCTACG
CCCAACATTGCTTCAAGGCACCCAAAAATTTCTGCCCATGTAGCAGTTGCTCCTGCCCGGGTAAG
AACAATGGGATCGGCACCTGCCCAGCAGGCACCAATTCACTTACATGCGATAATCTGTGTACACT
GGATCCTATTACACTTAAGGCCCCTGATACCTACAAATGCCCCCAGAGCAAGAGCCTGGTCGGTA
TCGGAGAACACTGTTCCGGACTTGCAGTAAAAAGCGACTATTGTGGAAATAACTCTTGCACTTGT
CAGCCACAAGCCTTCCTCGGTTGGTCCGCTGACTCTTGTTTACAAGGGGATAAGTGTAACATCTT
CGCAAATTTCATCTTACACGATGTGAATAACGGCTTAACATGCAGCACAGATCTCCAGAAGGCAA
ACACAGAGATCGAATTAGGAGTCTGCGTTAATTACGATCTCTACGGGATCTCTGGCCAGGGCATC
TTCGTGGAGGTTAATGCTACCTACTACAATAGTTGGCAAAATCTGCTCTACGATAGCAATGGCAA
CCTCTATGGATTCAGAGACTATATTACTAACAGGACGTTCATGATTCACTCGTGCTATTCCGGGC
GGGTGTCAGCAGCTTATCACGCAAATTCTTCAGAGCCAGCTCTGCTATTCCGAAACATAAAATGT
AATTACGTGTTCAATAATTCACTGACTCGGCAGCTGCAGCCGATTAATTACAGCTTCGACAGCTA
CCTTGGTTGCGTTGTTAACGCCTACAACTCCACTGCCATATCAGTTCAGACCTGCGACCTTACTG
TGGGCTCTGGCTATTGTGTCGATTATTCAAAGAACGGGGGGAGCGGGTCCGCAATAACAACTGGC
TATAGGTTCACCAATTTTGAGCCTTTCACCGTGAATAGTGTCAACGATAGCCTGGAGCCTGTCGG

[0508]

AGGTCTTTATGAGATACAAATCCCCTCCGAGTTCACAATTGGCAACATGGAAGAGTTCATCCAGA
CGAGTTCCCCAAAGGTGACGATCGATTGCGCGGCTTTCGTCTGCGGCGACTACGCCGCATGCAAG
TTACAACTCGTTGAGTATGGAAGTTTTTGCGATAATATAAACGCAATTCTGACTGAAGTGAACGA
ACTGCTGGACACCACTCAGTTGCAGGTGGCAAATTCGCTCATGAACGGCGTGACACTGTCAACCA
AACTGAAGGACGGTGTCAATTTCAATGTGGATGACATTAACTTCAGCCCCGTACTGGGCTGTTTG
GGTAGTGAGTGTTCTAAGGCTAGCAGCCGCTCCGCCATTGAGGACTTGTTGTTTGATAAAGTTAA
GCTGAGTGACGTTGGATTTGTTGAGGCGTATAATAACTGTACCGGTGGTGCAGAGATAAGGGATC
TGATCTGTGTCCAGAGTTATAAGGGGATTAAGGTTCTCCCCCCGCTACTCTCGGAGAATCAGATA
TCAGGATACACCCTGGCCGCTACCTCAGCCTCGCTGTTTCCCCCTTGGACCGCTGCCGCCGGTGT
CCCATTTTATTTGAATGTGCAGTATCGGATCAACGGTCTGGGAGTGACAATGGACGTGCTGTCTC
AGAACCAGAAACTGATCGCCAATGCATTCAACAATGCTCTGCACGCCATCCAGCAAGGGTTTGAC
GCTACAAATTCTGCCCTCGTAAAAATCCAGGCCGTGGTGAATGCTAACGCCGAAGCCCTTAATAA
TCTGCTCCAGCAGCTTTCTAACCGCTTTGGAGCTATTTCTGCCTCACTGCAGGAAATTCTATCCA
GACTGGATCCCCCTGAGGCAGAAGCCCAAATCGACCGTCTCATAAACGGCAGACTCACTGCTCTT
AACGCCTACGTTAGTCAACAATTGAGCGATTCGACCTTGGTGAAATTCAGCGCAGCTCAGGCTAT
GGAGAAGGTGAACGAGTGCGTGAAGTCACAGAGCTCCAGAATCAATTTCTGTGGCAATGGGAACC
ATATCATCTCCTTGGTTCAGAATGCTCCCTACGGCCTGTATTTCATCCACTTCAACTACGTGCCC
ACGAAGTACGTTACAGCCAAAGTGTCCCCCGGACTGTGCATCGCTGGTAACAGGGGCATTGCACC
AAAATCCGGCTACTTCGTCAATGTCAACAACACATGGATGTATACTGGGAGTGGTTATTATTACC
CTGAACCTATAACAGAGAACAATGTAGTAGTCATGTCCACATGCGCCGTCAATTATACTAAGGCC
CCCTATGTTATGCTCAACACTTCAATTCCCAATCTCCCGGATTTCAAAGAAGAGCTGGATCAGTG
GTTTAAGAATCAGACATCCGTGGCCCCTGACTTAAGCTTGGATTATATCAATGTGACTTTTTTAG
ACTTACAGGTCGAGATGAACCGACTCCAGGAAGCTATAAAAGTACTGAACCACTCCTATATCAAT
CTGAAAGATATCGGTACATACGAATATTACGTAAAATGGCCTTGGTATGTGTGGCTACTAATTTG
CCTTGCGGGCGTGGCTATGCTGGTCCTGCTGTTCTTCATTTGCTGCTGTACTGGGTGCGGTACTT
CCTGCTTCAAAAAGTGCGGTGGTTGCTGCGACGACTACACTGGGTACCAGGAGCTTGTCATTAAG
ACCTCGCATGACGACTGA

PDI-OC43-H5iTMCT-DNA (서열 번호 138)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgGTGATCGGCGATCTGAATTGTACCCTGGATCCCCGCCTGAAAGGGAGCTTTAACAACC
GAGATACAGGACCCCCGTCTATATCCATAGATACAGTGGATGTTACGAACGGGCTCGGCACCTAC
TATGTGCTAGACCGAGTTTATTTGAACACCACCTTATTCCTCAATGGATACTACCCAACTTCAGG
TAGTACTTACAGAAACATGGCGCTGAAGGGTACGGATCTGCTGAGCACCCTATGGTTTAAACCTC
CCTTCCTCTCGGACTTTATTAATGGCATCTTCGCTAAGGTGAAAAACACGAAGGTTTTCAAAGAT
GGAGTGATGTATTCAGAGTTCCCTGCGATCACCATTGGAAGTACCTTCGTGAATACTTCCTATAG
CGTGGTGGTTCAACCACGGACAATCAACTCCACCCAGGACGGCGTCAACAAGCTCCAGGGATTGC
TGGAGGTGTCAGTCTGTCAATATAACATGTGTGAGTACCCACACACTATCTGTCACCCTAATCTA
GGCAACCACTTTAAGGAACTGTGGCACTACGATACGGGGGTGGTAAGTTGCTTATATAAGAGAAA
TTTCACCTATGATGTTAATGCAACGTACCTGTACTTTCACTTCTATCAAGAAGGAGGAACTTTCT
ACGCATATTTCACAGATACCGGCTTTGTGACGAAATTCTTATTCAACGTTTACCTCGGAATGGCA
TTAAGCCATTATTACGTGATGCCTCTCACTTGCATCAGACGCCCTAAGGATGGTTTTTCTCTGGA
GTACTGGGTCACTCCCCTGACACCACGGCAGTACCTGCTTGCTTTTAACCAGGACGGTATCATTT
TTAATGCCGTCGATTGTATGAGCGATTTTATGAGCGAGATAAAGTGCAAGACCCAATCTATTGCT
CCGCCCACGGGGGTGTACGAACTGAATGGTTACACCGTCCAGCCCGTTGCCGATGTATATAGACG
GAAACCAGACCTGCCCAATTGCAACATCGAAGCTTGGTTAAACGATAAGTCAGTGCCCTCCCCCC
TCAATTGGGAGAGGAAGACTTTCTCCAACTGTAATTTCAACATGTCAAGCCTGATGTCTTTCATT
CAAGCCGATTCGTTCACTTGTAATAATATAGATGCAGCAAAGATCTATGGTATGTGCTTCAGTTC
CATCACAATAGATAAGTTTGCAATACCAAACCGTCGCAAGGTGGACCTTCAGCTCGGCAACCTGG
GCTATCTGCAGTCCAGCAATTATAGAATAGACACCACCGCCACATCATGTCAGCTGTACTATAAC

[0509]

CTCCCAGCAGCGAACGTCAGTGTTAGTAGGTTCAATCCTTCTACCTGGAATAAAAGGTTTGGATT
CATCGAAGATAGTGTGTTCGTACCTCAGCCAACAGGAGTGTTCACCAATCACAGCGTGGTCTACG
CCCAACATTGCTTCAAGGCACCCAAAAATTTCTGCCCATGTAGCAGTTGCTCCTGCCCGGGTAAG
AACAATGGGATCGGCACCTGCCCAGCAGGCACCAATTCACTTACATGCGATAATCTGTGTACACT
GGATCCTATTACACTTAAGGCCCCTGATACCTACAAATGCCCCCAGAGCAAGAGCCTGGTCGGTA
TCGGAGAACACTGTTCCGGACTTGCAGTAAAAAGCGACTATTGTGGAAATAACTCTTGCACTTGT
CAGCCACAAGCCTTCCTCGGTTGGTCCGCTGACTCTTGTTTACAAGGGGATAAGTGTAACATCTT
CGCAAATTTCATCTTACACGATGTGAATAACGGCTTAACATGCAGCACAGATCTCCAGAAGGCAA
ACACAGAGATCGAATTAGGAGTCTGCGTTAATTACGATCTCTACGGGATCTCTGGCCAGGGCATC
TTCGTGGAGGTTAATGCTACCTACTACAATAGTTGGCAAAATCTGCTCTACGATAGCAATGGCAA
CCTCTATGGATTCAGAGACTATATTACTAACAGGACGTTCATGATTCACTCGTGCTATTCCGGGC
GGGTGTCAGCAGCTTATCACGCAAATTCTTCAGAGCCAGCTCTGCTATTCCGAAACATAAAATGT
AATTACGTGTTCAATAATTCACTGACTCGGCAGCTGCAGCCGATTAATTACAGCTTCGACAGCTA
CCTTGGTTGCGTTGTTAACGCCTACAACTCCACTGCCATATCAGTTCAGACCTGCGACCTTACTG
TGGGCTCTGGCTATTGTGTCGATTATTCAAAGAACGGGGGGAGCGGGTCCGCAATAACAACTGGC
TATAGGTTCACCAATTTTGAGCCTTTCACCGTGAATAGTGTCAACGATAGCCTGGAGCCTGTCGG
AGGTCTTTATGAGATACAAATCCCCTCCGAGTTCACAATTGGCAACATGGAAGAGTTCATCCAGA
CGAGTTCCCCAAAGGTGACGATCGATTGCGCGGCTTTCGTCTGCGGCGACTACGCCGCATGCAAG
TTACAACTCGTTGAGTATGGAAGTTTTTGCGATAATATAAACGCAATTCTGACTGAAGTGAACGA
ACTGCTGGACACCACTCAGTTGCAGGTGGCAAATTCGCTCATGAACGGCGTGACACTGTCAACCA
AACTGAAGGACGGTGTCAATTTCAATGTGGATGACATTAACTTCAGCCCCGTACTGGGCTGTTTG
GGTAGTGAGTGTTCTAAGGCTAGCAGCCGCTCCGCCATTGAGGACTTGTTGTTTGATAAAGTTAA
GCTGAGTGACGTTGGATTTGTTGAGGCGTATAATAACTGTACCGGTGGTGCAGAGATAAGGGATC
TGATCTGTGTCCAGAGTTATAAGGGGATTAAGGTTCTCCCCCCCGCTACTCTCGGAGAATCAGATA
TCAGGATACACCCTGGCCGCTACCTCAGCCTCGCTGTTTCCCCCTTGGACCGCTGCCGCCGGTGT
CCCATTTTATTTGAATGTGCAGTATCGGATCAACGGTCTGGGAGTGACAATGGACGTGCTGTCTC
AGAACCAGAAACTGATCGCCAATGCATTCAACAATGCTCTGCACGCCATCCAGCAAGGGTTTGAC
GCTACAAATTCTGCCCTCGTAAAAATCCAGGCCGTGGTGAATGCTAACGCCGAAGCCCTTAATAA
TCTGCTCCAGCAGCTTTCTAACCGCTTTGGAGCTATTTCTGCCTCACTGCAGGAAATTCTATCCA
GACTGGATCCCCCTGAGGCAGAAGCCCAAATCGACCGTCTCATAAACGGCAGACTCACTGCTCTT
AACGCCTACGTTAGTCAACAATTGAGCGATTCGACCTTGGTGAAATTCAGCGCAGCTCAGGCTAT
GGAGAAGGTGAACGAGTGCGTGAAGTCACAGAGCTCCAGAATCAATTTCTGTGGCAATGGGAACC
ATATCATCTCCTTGGTTCAGAATGCTCCCTACGGCCTGTATTTCATCCACTTCAACTACGTGCCC
ACGAAGTACGTTACAGCCAAAGTGTCCCCCGGACTGTGCATCGCTGGTAACAGGGGCATTGCACC
AAAATCCGGCTACTTCGTCAATGTCAACAACACATGGATGTATACTGGGAGTGGTTATTATTACC
CTGAACCTATAACAGAGAACAATGTAGTAGTCATGTCCACATGCGCCGTCAATTATACTAAGGCC
CCCTATGTTATGCTCAACACTTCAATTCCCAATCTCCCGGATTTCAAAGAAGAGCTGGATCAGTG
GTTTAAGAATCAGACATCCGTGGCCCCTGACTTAAGCTTGGATTATATCAATGTGACTTTTTTAG
ACTTACAGGTCGAGATGAACCGACTCCAGGAAGCTATAAAAGTACTGAACCACTCCTATATCAAT
CTGAAAGATATCGGTACATACGAATATTACGTAAAATGGCCTTGGTATcaaatactgtcaattta
ttcaacagtggcgagttccctagcactggcaatcatgatggctggtctatctttatggatgtgct
ccaatggatcgttacaatgcagaatttgcattTGA

PDI-OC43-H5iCT-DNA (서열 번호 139)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgGTGATCGGCGATCTGAATTGTACCCTGGATCCCCGCCTGAAAGGGAGCTTTAACAACC
GAGATACAGGACCCCCGTCTATATCCATAGATACAGTGGATGTTACGAACGGGCTCGGCACCTAC
TATGTGCTAGACCGAGTTTATTTGAACACCACCTTATTCCTCAATGGATACTACCCAACTTCAGG
TAGTACTTACAGAAACATGGCGCTGAAGGGTACGGATCTGCTGAGCACCCTATGGTTTAAACCTC
CCTTCCTCTCGGACTTTATTAATGGCATCTTCGCTAAGGTGAAAAACACGAAGGTTTTCAAAGAT

[0510]

GGAGTGATGTATTCAGAGTTCCCTGCGATCACCATTGGAAGTACCTTCGTGAATACTTCCTATAG
CGTGGTGGTTCAACCACGGACAATCAACTCCACCCAGGACGGCGTCAACAAGCTCCAGGGATTGC
TGGAGGTGTCAGTCTGTCAATATAACATGTGTGAGTACCCACACACTATCTGTCACCCTAATCTA
GGCAACCACTTTAAGGAACTGTGGCACTACGATACGGGGGTGGTAAGTTGCTTATATAAGAGAAA
TTTCACCTATGATGTTAATGCAACGTACCTGTACTTTCACTTCTATCAAGAAGGAGGAACTTTCT
ACGCATATTTCACAGATACCGGCTTTGTGACGAAATTCTTATTCAACGTTTACCTCGGAATGGCA
TTAAGCCATTATTACGTGATGCCTCTCACTTGCATCAGACGCCCTAAGGATGGTTTTTCTCTGGA
GTACTGGGTCACTCCCCTGACACCACGGCAGTACCTGCTTGCTTTTAACCAGGACGGTATCATTT
TTAATGCCGTCGATTGTATGAGCGATTTTATGAGCGAGATAAAGTGCAAGACCCAATCTATTGCT
CCGCCCACGGGGGTGTACGAACTGAATGGTTACACCGTCCAGCCCGTTGCCGATGTATATAGACG
GAAACCAGACCTGCCCAATTGCAACATCGAAGCTTGGTTAAACGATAAGTCAGTGCCCTCCCCCC
TCAATTGGGAGAGGAAGACTTTCTCCAACTGTAATTTCAACATGTCAAGCCTGATGTCTTTCATT
CAAGCCGATTCGTTCACTTGTAATAATATAGATGCAGCAAAGATCTATGGTATGTGCTTCAGTTC
CATCACAATAGATAAGTTTGCAATACCAAACCGTCGCAAGGTGGACCTTCAGCTCGGCAACCTGG
GCTATCTGCAGTCCAGCAATTATAGAATAGACACCACCGCCACATCATGTCAGCTGTACTATAAC
CTCCCAGCAGCGAACGTCAGTGTTAGTAGGTTCAATCCTTCTACCTGGAATAAAAGGTTTGGATT
CATCGAAGATAGTGTGTTCGTACCTCAGCCAACAGGAGTGTTCACCAATCACAGCGTGGTCTACG
CCCAACATTGCTTCAAGGCACCCAAAAATTTCTGCCCATGTAGCAGTTGCTCCTGCCCGGGTAAG
AACAATGGGATCGGCACCTGCCCAGCAGGCACCAATTCACTTACATGCGATAATCTGTGTACACT
GGATCCTATTACACTTAAGGCCCCTGATACCTACAAATGCCCCCAGAGCAAGAGCCTGGTCGGTA
TCGGAGAACACTGTTCCGGACTTGCAGTAAAAAGCGACTATTGTGGAAATAACTCTTGCACTTGT
CAGCCACAAGCCTTCCTCGGTTGGTCCGCTGACTCTTGTTTACAAGGGGATAAGTGTAACATCTT
CGCAAATTTCATCTTACACGATGTGAATAACGGCTTAACATGCAGCACAGATCTCCAGAAGGCAA
ACACAGAGATCGAATTAGGAGTCTGCGTTAATTACGATCTCTACGGGATCTCTGGCCAGGGCATC
TTCGTGGAGGTTAATGCTACCTACTACAATAGTTGGCAAAATCTGCTCTACGATAGCAATGGCAA
CCTCTATGGATTCAGAGACTATATTACTAACAGGACGTTCATGATTCACTCGTGCTATTCCGGGC
GGGTGTCAGCAGCTTATCACGCAAATTCTTCAGAGCCAGCTCTGCTATTCCGAAACATAAAATGT
AATTACGTGTTCAATAATTCACTGACTCGGCAGCTGCAGCCGATTAATTACAGCTTCGACAGCTA
CCTTGGTTGCGTTGTTAACGCCTACAACTCCACTGCCATATCAGTTCAGACCTGCGACCTTACTG
TGGGCTCTGGCTATTGTGTCGATTATTCAAAGAACGGGGGGAGCGGGTCCGCAATAACAACTGGC
TATAGGTTCACCAATTTTGAGCCTTTCACCGTGAATAGTGTCAACGATAGCCTGGAGCCTGTCGG
AGGTCTTTATGAGATACAAATCCCCTCCGAGTTCACAATTGGCAACATGGAAGAGTTCATCCAGA
CGAGTTCCCCAAAGGTGACGATCGATTGCGCGGCTTTCGTCTGCGGCGACTACGCCGCATGCAAG
TTACAACTCGTTGAGTATGGAAGTTTTTGCGATAATATAAACGCAATTCTGACTGAAGTGAACGA
ACTGCTGGACACCACTCAGTTGCAGGTGGCAAATTCGCTCATGAACGGCGTGACACTGTCAACCA
AACTGAAGGACGGTGTCAATTTCAATGTGGATGACATTAACTTCAGCCCCGTACTGGGCTGTTTG
GGTAGTGAGTGTTCTAAGGCTAGCAGCCGCTCCGCCATTGAGGACTTGTTGTTTGATAAAGTTAA
GCTGAGTGACGTTGGATTTGTTGAGGCGTATAATAACTGTACCGGTGGTGCAGAGATAAGGGATC
TGATCTGTGTCCAGAGTTATAAGGGGATTAAGGTTCTCCCCCCCGCTACTCTCGGAGAATCAGATA
TCAGGATACACCCTGGCCGCTACCTCAGCCTCGCTGTTTCCCCCTTGGACCGCTGCCGCCGGTGT
CCCATTTTATTTGAATGTGCAGTATCGGATCAACGGTCTGGGAGTGACAATGGACGTGCTGTCTC
AGAACCAGAAACTGATCGCCAATGCATTCAACAATGCTCTGCACGCCATCCAGCAAGGGTTTGAC
GCTACAAATTCTGCCCTCGTAAAAATCCAGGCCGTGGTGAATGCTAACGCCGAAGCCCTTAATAA
TCTGCTCCAGCAGCTTTCTAACCGCTTTGGAGCTATTTCTGCCTCACTGCAGGAAATTCTATCCA
GACTGGATCCCCCTGAGGCAGAAGCCCAAATCGACCGTCTCATAAACGGCAGACTCACTGCTCTT
AACGCCTACGTTAGTCAACAATTGAGCGATTCGACCTTGGTGAAATTCAGCGCAGCTCAGGCTAT
GGAGAAGGTGAACGAGTGCGTGAAGTCACAGAGCTCCAGAATCAATTTCTGTGGCAATGGGAACC
ATATCATCTCCTTGGTTCAGAATGCTCCCTACGGCCTGTATTTCATCCACTTCAACTACGTGCCC
ACGAAGTACGTTACAGCCAAAGTGTCCCCCGGACTGTGCATCGCTGGTAACAGGGGCATTGCACC
AAAATCCGGCTACTTCGTCAATGTCAACAACACATGGATGTATACTGGGAGTGGTTATTATTACC
CTGAACCTATAACAGAGAACAATGTAGTAGTCATGTCCACATGCGCCGTCAATTATACTAAGGCC
CCCTATGTTATGCTCAACACTTCAATTCCCAATCTCCCGGATTTCAAAGAAGAGCTGGATCAGTG

[0511]

GTTTAAGAATCAGACATCCGTGGCCCCTGACTTAAGCTTGGATTATATCAATGTGACTTTTTTAG
ACTTACAGGTCGAGATGAACCGACTCCAGGAAGCTATAAAAGTACTGAACCACTCCTATATCAAT
CTGAAAGATATCGGTACATACGAATATTACGTAAAATGGCCTTGGTATGTGTGGCTACTAATTTG
CCTTGCGGGCGTGGCTATGCTGGTCCTGCTGTTCTTCATTtctttatggatgtgctccaatggat
cgttacaatgcagaatttgcattTGA

PDI-OC43-H5iCT(V4)-DNA (서열 번호 140)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgGTGATCGGCGATCTGAATTGTACCCTGGATCCCCGCCTGAAAGGGAGCTTTAACAACC
GAGATACAGGACCCCCGTCTATATCCATAGATACAGTGGATGTTACGAACGGGCTCGGCACCTAC
TATGTGCTAGACCGAGTTTATTTGAACACCACCTTATTCCTCAATGGATACTACCCAACTTCAGG
TAGTACTTACAGAAACATGGCGCTGAAGGGTACGGATCTGCTGAGCACCCTATGGTTTAAACCTC
CCTTCCTCTCGGACTTTATTAATGGCATCTTCGCTAAGGTGAAAAACACGAAGGTTTTCAAAGAT
GGAGTGATGTATTCAGAGTTCCCTGCGATCACCATTGGAAGTACCTTCGTGAATACTTCCTATAG
CGTGGTGGTTCAACCACGGACAATCAACTCCACCCAGGACGGCGTCAACAAGCTCCAGGGATTGC
TGGAGGTGTCAGTCTGTCAATATAACATGTGTGAGTACCCACACACTATCTGTCACCCTAATCTA
GGCAACCACTTTAAGGAACTGTGGCACTACGATACGGGGGTGGTAAGTTGCTTATATAAGAGAAA
TTTCACCTATGATGTTAATGCAACGTACCTGTACTTTCACTTCTATCAAGAAGGAGGAACTTTCT
ACGCATATTTCACAGATACCGGCTTTGTGACGAAATTCTTATTCAACGTTTACCTCGGAATGGCA
TTAAGCCATTATTACGTGATGCCTCTCACTTGCATCAGACGCCCTAAGGATGGTTTTTCTCTGGA
GTACTGGGTCACTCCCCTGACACCACGGCAGTACCTGCTTGCTTTTAACCAGGACGGTATCATTT
TTAATGCCGTCGATTGTATGAGCGATTTTATGAGCGAGATAAAGTGCAAGACCCAATCTATTGCT
CCGCCCACGGGGGTGTACGAACTGAATGGTTACACCGTCCAGCCCGTTGCCGATGTATATAGACG
GAAACCAGACCTGCCCAATTGCAACATCGAAGCTTGGTTAAACGATAAGTCAGTGCCCTCCCCCC
TCAATTGGGAGAGGAAGACTTTCTCCAACTGTAATTTCAACATGTCAAGCCTGATGTCTTTCATT
CAAGCCGATTCGTTCACTTGTAATAATATAGATGCAGCAAAGATCTATGGTATGTGCTTCAGTTC
CATCACAATAGATAAGTTTGCAATACCAAACCGTCGCAAGGTGGACCTTCAGCTCGGCAACCTGG
GCTATCTGCAGTCCAGCAATTATAGAATAGACACCACCGCCACATCATGTCAGCTGTACTATAAC
CTCCCAGCAGCGAACGTCAGTGTTAGTAGGTTCAATCCTTCTACCTGGAATAAAAGGTTTGGATT
CATCGAAGATAGTGTGTTCGTACCTCAGCCAACAGGAGTGTTCACCAATCACAGCGTGGTCTACG
CCCAACATTGCTTCAAGGCACCCAAAAATTTCTGCCCATGTAGCAGTTGCTCCTGCCCGGGTAAG
AACAATGGGATCGGCACCTGCCCAGCAGGCACCAATTCACTTACATGCGATAATCTGTGTACACT
GGATCCTATTACACTTAAGGCCCCTGATACCTACAAATGCCCCCAGAGCAAGAGCCTGGTCGGTA
TCGGAGAACACTGTTCCGGACTTGCAGTAAAAAGCGACTATTGTGGAAATAACTCTTGCACTTGT
CAGCCACAAGCCTTCCTCGGTTGGTCCGCTGACTCTTGTTTACAAGGGGATAAGTGTAACATCTT
CGCAAATTTCATCTTACACGATGTGAATAACGGCTTAACATGCAGCACAGATCTCCAGAAGGCAA
ACACAGAGATCGAATTAGGAGTCTGCGTTAATTACGATCTCTACGGGATCTCTGGCCAGGGCATC
TTCGTGGAGGTTAATGCTACCTACTACAATAGTTGGCAAAATCTGCTCTACGATAGCAATGGCAA
CCTCTATGGATTCAGAGACTATATTACTAACAGGACGTTCATGATTCACTCGTGCTATTCCGGGC
GGGTGTCAGCAGCTTATCACGCAAATTCTTCAGAGCCAGCTCTGCTATTCCGAAACATAAAATGT
AATTACGTGTTCAATAATTCACTGACTCGGCAGCTGCAGCCGATTAATTACAGCTTCGACAGCTA
CCTTGGTTGCGTTGTTAACGCCTACAACTCCACTGCCATATCAGTTCAGACCTGCGACCTTACTG
TGGGCTCTGGCTATTGTGTCGATTATTCAAAGAACGGGGGGAGCGGGTCCGCAATAACAACTGGC
TATAGGTTCACCAATTTTGAGCCTTTCACCGTGAATAGTGTCAACGATAGCCTGGAGCCTGTCGG
AGGTCTTTATGAGATACAAATCCCCTCCGAGTTCACAATTGGCAACATGGAAGAGTTCATCCAGA
CGAGTTCCCCAAAGGTGACGATCGATTGCGCGGCTTTCGTCTGCGGCGACTACGCCGCATGCAAG
TTACAACTCGTTGAGTATGGAAGTTTTTGCGATAATATAAACGCAATTCTGACTGAAGTGAACGA
ACTGCTGGACACCACTCAGTTGCAGGTGGCAAATTCGCTCATGAACGGCGTGACACTGTCAACCA
AACTGAAGGACGGTGTCAATTTCAATGTGGATGACATTAACTTCAGCCCCGTACTGGGCTGTTTG
[0512] GGTAGTGAGTGTTCTAAGGCTAGCAGCCGCTCCGCCATTGAGGACTTGTTGTTTGATAAAGTTAA

GCTGAGTGACGTTGGATTTGTTGAGGCGTATAATAACTGTACCGGTGGTGCAGAGATAAGGGATC
TGATCTGTGTCCAGAGTTATAAGGGGATTAAGGTTCTCCCCCCGCTACTCTCGGAGAATCAGATA
TCAGGATACACCCTGGCCGCTACCTCAGCCTCGCTGTTTCCCCCTTGGACCGCTGCCGCCGGTGT
CCCATTTTATTTGAATGTGCAGTATCGGATCAACGGTCTGGGAGTGACAATGGACGTGCTGTCTC
AGAACCAGAAACTGATCGCCAATGCATTCAACAATGCTCTGCACGCCATCCAGCAAGGGTTTGAC
GCTACAAATTCTGCCCTCGTAAAAATCCAGGCCGTGGTGAATGCTAACGCCGAAGCCCTTAATAA
TCTGCTCCAGCAGCTTTCTAACCGCTTTGGAGCTATTTCTGCCTCACTGCAGGAAATTCTATCCA
GACTGGATCCCCCTGAGGCAGAAGCCCAAATCGACCGTCTCATAAACGGCAGACTCACTGCTCTT
AACGCCTACGTTAGTCAACAATTGAGCGATTCGACCTTGGTGAAATTCAGCGCAGCTCAGGCTAT
GGAGAAGGTGAACGAGTGCGTGAAGTCACAGAGCTCCAGAATCAATTTCTGTGGCAATGGGAACC
ATATCATCTCCTTGGTTCAGAATGCTCCCTACGGCCTGTATTTCATCCACTTCAACTACGTGCCC
ACGAAGTACGTTACAGCCAAAGTGTCCCCCGGACTGTGCATCGCTGGTAACAGGGGCATTGCACC
AAAATCCGGCTACTTCGTCAATGTCAACAACACATGGATGTATACTGGGAGTGGTTATTATTACC
CTGAACCTATAACAGAGAACAATGTAGTAGTCATGTCCACATGCGCCGTCAATTATACTAAGGCC
CCCTATGTTATGCTCAACACTTCAATTCCCAATCTCCCGGATTTCAAAGAAGAGCTGGATCAGTG
GTTTAAGAATCAGACATCCGTGGCCCCTGACTTAAGCTTGGATTATATCAATGTGACTTTTTTAG
ACTTACAGGTCGAGATGAACCGACTCCAGGAAGCTATAAAAGTACTGAACCACTCCTATATCAAT
CTGAAAGATATCGGTACATACGAATATTACGTAAAATGGCCTTGGTATGTGTGGCTACTAATTTG
CCTTGCGGGCGTGGCTATGCTGGTCCTGCTGTTCTTCATTtgctgctccaatggatcgttacaat
gcagaatttgcattTGA

PDI-OC43-H1cCT-DNA (서열 번호 141)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgGTGATCGGCGATCTGAATTGTACCCTGGATCCCCGCCTGAAAGGGAGCTTTAACAACC
GAGATACAGGACCCCCGTCTATATCCATAGATACAGTGGATGTTACGAACGGGCTCGGCACCTAC
TATGTGCTAGACCGAGTTTATTTGAACACCACCTTATTCCTCAATGGATACTACCCAACTTCAGG
TAGTACTTACAGAAACATGGCGCTGAAGGGTACGGATCTGCTGAGCACCCTATGGTTTAAACCTC
CCTTCCTCTCGGACTTTATTAATGGCATCTTCGCTAAGGTGAAAAACACGAAGGTTTTCAAAGAT
GGAGTGATGTATTCAGAGTTCCCTGCGATCACCATTGGAAGTACCTTCGTGAATACTTCCTATAG
CGTGGTGGTTCAACCACGGACAATCAACTCCACCCAGGACGGCGTCAACAAGCTCCAGGGATTGC
TGGAGGTGTCAGTCTGTCAATATAACATGTGTGAGTACCCACACACTATCTGTCACCCTAATCTA
GGCAACCACTTTAAGGAACTGTGGCACTACGATACGGGGGTGGTAAGTTGCTTATATAAGAGAAA
TTTCACCTATGATGTTAATGCAACGTACCTGTACTTTCACTTCTATCAAGAAGGAGGAACTTTCT
ACGCATATTTCACAGATACCGGCTTTGTGACGAAATTCTTATTCAACGTTTACCTCGGAATGGCA
TTAAGCCATTATTACGTGATGCCTCTCACTTGCATCAGACGCCCTAAGGATGGTTTTTCTCTGGA
GTACTGGGTCACTCCCCTGACACCACGGCAGTACCTGCTTGCTTTTAACCAGGACGGTATCATTT
TTAATGCCGTCGATTGTATGAGCGATTTTATGAGCGAGATAAAGTGCAAGACCCAATCTATTGCT
CCGCCCACGGGGGTGTACGAACTGAATGGTTACACCGTCCAGCCCGTTGCCGATGTATATAGACG
GAAACCAGACCTGCCCAATTGCAACATCGAAGCTTGGTTAAACGATAAGTCAGTGCCCTCCCCCC
TCAATTGGGAGAGGAAGACTTTCTCCAACTGTAATTTCAACATGTCAAGCCTGATGTCTTTCATT
CAAGCCGATTCGTTCACTTGTAATAATATAGATGCAGCAAAGATCTATGGTATGTGCTTCAGTTC
CATCACAATAGATAAGTTTGCAATACCAAACCGTCGCAAGGTGGACCTTCAGCTCGGCAACCTGG
GCTATCTGCAGTCCAGCAATTATAGAATAGACACCACCGCCACATCATGTCAGCTGTACTATAAC
CTCCCAGCAGCGAACGTCAGTGTTAGTAGGTTCAATCCTTCTACCTGGAATAAAAGGTTTGGATT
CATCGAAGATAGTGTGTTCGTACCTCAGCCAACAGGAGTGTTCACCAATCACAGCGTGGTCTACG
CCCAACATTGCTTCAAGGCACCCAAAAATTTCTGCCCATGTAGCAGTTGCTCCTGCCCGGGTAAG
AACAATGGGATCGGCACCTGCCCAGCAGGCACCAATTCACTTACATGCGATAATCTGTGTACACT
GGATCCTATTACACTTAAGGCCCCTGATACCTACAAATGCCCCCAGAGCAAGAGCCTGGTCGGTA
TCGGAGAACACTGTTCCGGACTTGCAGTAAAAAGCGACTATTGTGGAAATAACTCTTGCACTTGT
CAGCCACAAGCCTTCCTCGGTTGGTCCGCTGACTCTTGTTTACAAGGGGATAAGTGTAACATCTT

[0513]

CGCAAATTTCATCTTACACGATGTGAATAACGGCTTAACATGCAGCACAGATCTCCAGAAGGCAA
ACACAGAGATCGAATTAGGAGTCTGCGTTAATTACGATCTCTACGGGATCTCTGGCCAGGGCATC
TTCGTGGAGGTTAATGCTACCTACTACAATAGTTGGCAAAATCTGCTCTACGATAGCAATGGCAA
CCTCTATGGATTCAGAGACTATATTACTAACAGGACGTTCATGATTCACTCGTGCTATTCCGGGC
GGGTGTCAGCAGCTTATCACGCAAATTCTTCAGAGCCAGCTCTGCTATTCCGAAACATAAAATGT
AATTACGTGTTCAATAATTCACTGACTCGGCAGCTGCAGCCGATTAATTACAGCTTCGACAGCTA
CCTTGGTTGCGTTGTTAACGCCTACAACTCCACTGCCATATCAGTTCAGACCTGCGACCTTACTG
TGGGCTCTGGCTATTGTGTCGATTATTCAAAGAACGGGGGGAGCGGGTCCGCAATAACAACTGGC
TATAGGTTCACCAATTTTGAGCCTTTCACCGTGAATAGTGTCAACGATAGCCTGGAGCCTGTCGG
AGGTCTTTATGAGATACAAATCCCCTCCGAGTTCACAATTGGCAACATGGAAGAGTTCATCCAGA
CGAGTTCCCCAAAGGTGACGATCGATTGCGCGGCTTTCGTCTGCGGCGACTACGCCGCATGCAAG
TTACAACTCGTTGAGTATGGAAGTTTTTGCGATAATATAAACGCAATTCTGACTGAAGTGAACGA
ACTGCTGGACACCACTCAGTTGCAGGTGGCAAATTCGCTCATGAACGGCGTGACACTGTCAACCA
AACTGAAGGACGGTGTCAATTTCAATGTGGATGACATTAACTTCAGCCCCGTACTGGGCTGTTTG
GGTAGTGAGTGTTCTAAGGCTAGCAGCCGCTCCGCCATTGAGGACTTGTTGTTTGATAAAGTTAA
GCTGAGTGACGTTGGATTTGTTGAGGCGTATAATAACTGTACCGGTGGTGCAGAGATAAGGGATC
TGATCTGTGTCCAGAGTTATAAGGGGATTAAGGTTCTCCCCCCGCTACTCTCGGAGAATCAGATA
TCAGGATACACCCTGGCCGCTACCTCAGCCTCGCTGTTTCCCCCTTGGACCGCTGCCGCCGGTGT
CCCATTTTATTTGAATGTGCAGTATCGGATCAACGGTCTGGGAGTGACAATGGACGTGCTGTCTC
AGAACCAGAAACTGATCGCCAATGCATTCAACAATGCTCTGCACGCCATCCAGCAAGGGTTTGAC
GCTACAAATTCTGCCCTCGTAAAAATCCAGGCCGTGGTGAATGCTAACGCCGAAGCCCTTAATAA
TCTGCTCCAGCAGCTTTCTAACCGCTTTGGAGCTATTTCTGCCTCACTGCAGGAAATTCTATCCA
GACTGGATCCCCCTGAGGCAGAAGCCCAAATCGACCGTCTCATAAACGGCAGACTCACTGCTCTT
AACGCCTACGTTAGTCAACAATTGAGCGATTCGACCTTGGTGAAATTCAGCGCAGCTCAGGCTAT
GGAGAAGGTGAACGAGTGCGTGAAGTCACAGAGCTCCAGAATCAATTTCTGTGGCAATGGGAACC
ATATCATCTCCTTGGTTCAGAATGCTCCCTACGGCCTGTATTTCATCCACTTCAACTACGTGCCC
ACGAAGTACGTTACAGCCAAAGTGTCCCCCGGACTGTGCATCGCTGGTAACAGGGGCATTGCACC
AAAATCCGGCTACTTCGTCAATGTCAACAACACATGGATGTATACTGGGAGTGGTTATTATTACC
CTGAACCTATAACAGAGAACAATGTAGTAGTCATGTCCACATGCGCCGTCAATTATACTAAGGCC
CCCTATGTTATGCTCAACACTTCAATTCCCAATCTCCCGGATTTCAAAGAAGAGCTGGATCAGTG
GTTTAAGAATCAGACATCCGTGGCCCCTGACTTAAGCTTGGATTATATCAATGTGACTTTTTTAG
ACTTACAGGTCGAGATGAACCGACTCCAGGAAGCTATAAAAGTACTGAACCACTCCTATATCAAT
CTGAAAGATATCGGTACATACGAATATTACGTAAAATGGCCTTGGTATGTGTGGCTACTAATTTG
CCTTGCGGGCGTGGCTATGCTGGTCCTGCTGTTCTTCATTagcttctggatgtgctctaatgggt
ctctacagtgtagaatatgtattTGA

PDI-OC43-wtTMCT-AA (서열 번호 142)

MAKNVAIFGLLFSLLVLVPSQIFAVIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTY
YVLDRVYLNTTLFLNGYYPTSGSTYRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKD
GVMYSEFPAITIGSTFVNTSYSVVVQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNL
GNHFKELWHYDTGVVSCLYKRNFTYDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMA
LSHYYVMPLTCIRRPKDGFSLEYWVTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIA
PPTGVYELNGYTVQPVADVYRRKPDLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFI
QADSFTCNNIDAAKIYGMCFSSITIDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYN
LPAANVSVSRFNPSTWNKRFGFIEDSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGK
NNGIGTCPAGTNSLTCDNLCTLDPITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTC
QPQAFLGWSADSCLQGDKCNIFANFILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGI
FVEVNATYYNSWQNLLYDSNGNLYGFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKC
NYVFNNSLTRQLQPINYSFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNGGSGSAITTG
YRFTNFEPFTVNSVNDSLEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACK

[0514]

LQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCL
GSECSKASSRSAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQI
SGYTLAATSASLFPPWTAAAGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFD
ATNSALVKIQAVVNANAEALNNLLQQLSNRFGAISASLQEILSRLDPPEAEAQIDRLINGRLTAL
NAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVP
TKYVTAKVSPGLCIAGNRGIAPKSGYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKA
PYVMLNTSIPNLPDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYIN
LKDIGTYEYYVKWPWYVWLLICLAGVAMLVLLFFICCCTGCGTSCFKKCGGCCDDYTGYQELVIK
TSHDD

PDI-OC43-H5iTMCT-AA (서열 번호 143)

MAKNVAIFGLLFSLLVLVPSQIFAVIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTY
YVLDRVYLNTTLFLNGYYPTSGSTYRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKD
GVMYSEFPAITIGSTFVNTSYSVVVQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNL
GNHFKELWHYDTGVVSCLYKRNFTYDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMA
LSHYYVMPLTCIRRPKDGFSLEYWVTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIA
PPTGVYELNGYTVQPVADVYRRKPDLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFI
QADSFTCNNIDAAKIYGMCFSSITIDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYN
LPAANVSVSRFNPSTWNKRFGFIEDSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGK
NNGIGTCPAGTNSLTCDNLCTLDPITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTC
QPQAFLGWSADSCLQGDKCNIFANFILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGI
FVEVNATYYNSWQNLLYDSNGNLYGFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKC
NYVFNNSLTRQLQPINYSFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNGGSGSAITTG
YRFTNFEPFTVNSVNDSLEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACK
LQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCL
GSECSKASSRSAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQI
SGYTLAATSASLFPPWTAAAGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFD
ATNSALVKIQAVVNANAEALNNLLQQLSNRFGAISASLQEILSRLDPPEAEAQIDRLINGRLTAL
NAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVP
TKYVTAKVSPGLCIAGNRGIAPKSGYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKA
PYVMLNTSIPNLPDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYIN
LKDIGTYEYYVKWPWYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

PDI-OC43-H5iCT-AA (서열 번호 144)

MAKNVAIFGLLFSLLVLVPSQIFAVIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTY
YVLDRVYLNTTLFLNGYYPTSGSTYRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKD
GVMYSEFPAITIGSTFVNTSYSVVVQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNL
GNHFKELWHYDTGVVSCLYKRNFTYDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMA
LSHYYVMPLTCIRRPKDGFSLEYWVTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIA
PPTGVYELNGYTVQPVADVYRRKPDLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFI
QADSFTCNNIDAAKIYGMCFSSITIDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYN
LPAANVSVSRFNPSTWNKRFGFIEDSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGK
NNGIGTCPAGTNSLTCDNLCTLDPITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTC
QPQAFLGWSADSCLQGDKCNIFANFILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGI
FVEVNATYYNSWQNLLYDSNGNLYGFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKC
NYVFNNSLTRQLQPINYSFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNGGSGSAITTG
YRFTNFEPFTVNSVNDSLEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACK
LQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCL
GSECSKASSRSAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQI

[0515]

SGYTLAATSASLFPPWTAAAGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFD
ATNSALVKIQAVVNANAEALNNLLQQLSNRFGAISASLQEILSRLDPPEAEAQIDRLINGRLTAL
NAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVP
TKYVTAKVSPGLCIAGNRGIAPKSGYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKA
PYVMLNTSIPNLPDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYIN
LKDIGTYEYYVKWPWYVWLLICLAGVAMLVLLFFISLWMCSNGSLQCRICI

PDI-OC43-H5iCT(V4)-AA (서열 번호 145)

MAKNVAIFGLLFSLLVLVPSQIFAVIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTY
YVLDRVYLNTTLFLNGYYPTSGSTYRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKD
GVMYSEFPAITIGSTFVNTSYSVVVQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNL
GNHFKELWHYDTGVVSCLYKRNFTYDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMA
LSHYYVMPLTCIRRPKDGFSLEYWVTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIA
PPTGVYELNGYTVQPVADVYRRKPDLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFI
QADSFTCNNIDAAKIYGMCFSSITIDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYN
LPAANVSVSRFNPSTWNKRFGFIEDSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGK
NNGIGTCPAGTNSLTCDNLCTLDPITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTC
QPQAFLGWSADSCLQGDKCNIFANFILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGI
FVEVNATYYNSWQNLLYDSNGNLYGFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKC
NYVFNNSLTRQLQPINYSFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNGGSGSAITTG
YRFTNFEPFTVNSVNDSLEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACK
LQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCL
GSECSKASSRSAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQI
SGYTLAATSASLFPPWTAAAGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFD
ATNSALVKIQAVVNANAEALNNLLQQLSNRFGAISASLQEILSRLDPPEAEAQIDRLINGRLTAL
NAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVP
TKYVTAKVSPGLCIAGNRGIAPKSGYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKA
PYVMLNTSIPNLPDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYIN
LKDIGTYEYYVKWPWYVWLLICLAGVAMLVLLFFICCSNGSLQCRICI

PDI-OC43-H1cCT-AA (서열 번호 146)

MAKNVAIFGLLFSLLVLVPSQIFAVIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTY
YVLDRVYLNTTLFLNGYYPTSGSTYRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKD
GVMYSEFPAITIGSTFVNTSYSVVVQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNL
GNHFKELWHYDTGVVSCLYKRNFTYDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMA
LSHYYVMPLTCIRRPKDGFSLEYWVTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIA
PPTGVYELNGYTVQPVADVYRRKPDLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFI
QADSFTCNNIDAAKIYGMCFSSITIDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYN
LPAANVSVSRFNPSTWNKRFGFIEDSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGK
NNGIGTCPAGTNSLTCDNLCTLDPITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTC
QPQAFLGWSADSCLQGDKCNIFANFILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGI
FVEVNATYYNSWQNLLYDSNGNLYGFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKC
NYVFNNSLTRQLQPINYSFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNGGSGSAITTG
YRFTNFEPFTVNSVNDSLEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACK
LQLVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCL
GSECSKASSRSAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQI
SGYTLAATSASLFPPWTAAAGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFD
ATNSALVKIQAVVNANAEALNNLLQQLSNRFGAISASLQEILSRLDPPEAEAQIDRLINGRLTAL

[0516]

NAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVP
TKYVTAKVSPGLCIAGNRGIAPKSGYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKA
PYVMLNTSIPNLPDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYIN
LKDIGTYEYYVKWPWYVWLLICLAGVAMLVLLFFISFWMCSNGSLQCRICI

IF(CoV229EwtCT).r (서열 번호 147)

ACGACACGACTAAGGCCTTCACTGTATGTGGATCTTTTCGACATCGTA

PDI-229E -wtTMCT-DNA (서열 번호 148)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgCAAACGACTAATGGGCTGAACACCAGTTACAGCGTCTGTAACGGCTGCGTCGGATATA
GCGAGAACGTGTTCGCAGTGGAAAGTGGGGGGTACATTCCTTCCGACTTCGCTTTCAATAACTGG
TTTCTCCTGACTAACACAAGCTCCGTCGTGGATGGCGTGGTCAGGTCCTTTCAGCCTCTTCTCCT
GAATTGCCTGTGGTCTGTGTCCGGGTTAAGATTCACTACAGGCTTCGTATACTTCAACGGGACGG
GCCGGGGGGATTGCAAGGGCTTCTCCTCCGACGTGCTGTCAGATGTGATCCGTTACAATCTGAAC
TTCGAAGAGAACTTACGGCGGGGGACAATCCTGTTCAAAACATCATATGGCGTAGTCGTATTTTA
CTGCACCAATAATACCCTGGTGAGTGGGGACGCCCATATTCCCTTCGGAACAGTGCTGGGTAACT
TTTACTGTTTTGTCAACACTACGATCGGAAACGAAACCACTAGCGCCTTTGTCGGAGCTCTGCCA
AAAACAGTTAGGGAGTTCGTGATCTCTCGGACCGGTCACTTCTATATCAACGGCTACCGTTATTT
TACTTTGGGCAACGTCGAAGCCGTCAATTTTAATGTGACAACTGCAGAGACAACTGACTTTTGCA
CTGTGGCTCTCGCCAGTTATGCCGATGTGCTGGTGAATGTAAGTCAAACGTCAATTGCCAACATC
ATCTATTGTAACTCAGTAATCAACCGGCTCCGCTGTGACCAACTCTCATTCGACGTCCCCGACGG
ATTCTATTCCACGAGCCCGATTCAGAGCGTGGAACTGCCAGTTTCCATCGTATCCCTCCCAGTTT
ACCACAAGCACACTTTTATCGTTCTCTACGTAGATTTTAAACCCCAGTCAGGAGGAGGGAAATGC
TTCAACTGCTACCCGGCTGGCGTGAACATCACCTTGGCCAATTTTAATGAAACTAAAGGGCCCCT
TTGCGTGGATACGTCACACTTTACCACAAAGTATGTTGCAGTCTATGCTAACGTCGGCAGGTGGT
CAGCGTCCATTAACACAGGCAATTGCCCGTTCTCTTTCGGGAAAGTGAACAACTTCGTGAAGTTT
GGAAGTGTGTGCTTCAGTTTGAAAGACATTCCGGGCGGCTGCGCCATGCCTATTGTGGCTAATTG
GGCTTATTCCAAGTACTACACCATTGGCTCTCTCTACGTTAGCTGGAGCGACGGTGACGGTATAA
CGGGCGTACCACAACCGGTGGAAGGGGTCAGCTCTTTCATGAATGTCACTCTGGACAAGTGTACC
AAATATAATATATACGATGTGAGTGGAGTGGGCGTTATACGCGTGTCTAACGACACCTTTCTAAA
CGGCATAACCTACACAAGCACGTCAGGCAATCTGTTAGGTTTTAAAGACGTCACTAAAGGCACTA
TATATAGCATCACCCCATGCAACCCACCTGATCAATTAGTCGTATATCAGCAAGCTGTTGTGGGT
GCTATGCTGTCAGAAAACTTCACCAGCTACGGGTTCTCCAATGTGGTGGAACTGCCCAAATTCTT
TTACGCTAGCAATGGCACATATAACTGTACTGACGCCGTCTTGACTTACAGTTCATTCGGAGTGT
GCGCGGACGGCAGCATTATCGCCGTGCAGCCGGCCAATGTCAGCTATGATTCCGTTTCCGCCATC
GTGACAGCCAACTTGTCGATTCCCTCTAACTGGACAACGTCTGTCCAAGTCGAATATCTGCAGAT
CACCTCAACCCCCATAGTAGTCGATTGCTCAACCTACGTCTGCAACGGTAATGTCAGATGTGTCG
AGCTGCTCAAGCAGTACACCTCCGCCTGTAAGACTATTGAGGATGCATTAAGAAATAGTGCAAGA
TTGGAAAGCGCCGATGTGTCGGAAATGCTAACCTTCGATAAGAAGGCATTCACACTGGCGAACGT
AAGCTCTTTCGGCGATTACAACCTGTCTTCGGTAATCCCTAGCTTGCCCACATCCGGCTCTCGGG
TGGCGGGGCGGAGCGCTATCGAGGACATTTTATTCTCGAAACTGGTTACATCTGGGCTCGGAACT
GTGGACGCCGATTACAAGAAGTGCACCAAGGGCCTAAGCATCGCCGACCTCGCCTGTGCTCAGTA
CTACAACGGAATTATGGTGCTGCCAGGTGTCGCTGACGCAGAGCGGATGGCTATGTATACCGGCA
GTCTCATTGGCGGGATTGCGTTGGGCGGCCTGACGTCCGCTGTCTCCATCCCTTTCTCTCTGGCT
ATACAAGCCCGACTGAATTATGTGGCCCTGCAGACTGATGTCCTGCAAGAAAATCAGAAGATTCT
TGCCGCCAGCTTCAACAAGGCCATGACTAATATTGTGGATGCGTTTACCGGAGTGAATGACGCCA
TCACCCAAACGTCCCAAGCCCTGCAGACAGTCGCCACGGCGTTAAACAAAATCCAGGATGTAGTG

[0517]

AATCAGCAAGGGAACAGCTTGAATCACCTGACGTCCCAGTTAAGACAGAACTTTCAGGCAATCAG
TAGCTCAATCCAGGCTATCTACGATCGATTAGATCCTCCTCAGGCAGATCAGCAGGTGGATCGGC
TCATCACCGGCCGCCTCGCGGCATTGAATGTTTTCGTAAGTCATACCTTGACCAAGTACACGGAG
GTGAGGGCCAGTCGCCAGCTGGCTCAGCAAAAAGTGAATGAGTGTGTGAAATCACAGAGCAAACG
GTACGGGTTTTGTGGAAATGGGACGCACATCTTTAGCATCGTTAATGCTGCCCCCGAAGGGTTAG
TCTTCCTGCACACTGTGCTCCTTCCTACCCAGTATAAAGATGTCGAAGCATGGTCTGGGCTCTGT
GTCGATGGAACTAACGGTTATGTCCTTCGACAGCCAAACCTCGCTCTCTATAAAGAAGGGAATTA
CTATAGGATCACCTCAAGAATCATGTTCGAGCCCAGGATACCAACAATGGCCGATTTTGTGCAGA
TTGAAAATTGTAACGTGACCTTTGTGAATATCAGTCGATCCGAGCTTCAAACGATTGTTCCTGAG
TACATCGACGTGAATAAAACTCTACAAGAGCTGTCCTATAAACTGCCTAATTATACCGTGCCTGA
CCTTGTAGTCGAGCAATACAACCAGACTATTCTGAACCTGACATCGGAAATCTCTACATTGGAGA
ATAAAAGCGCCGAGCTCAATTACACAGTGCAGAAGCTGCAGACCCTGATCGACAATATTAACAGC
ACTCTTGTGGACTTAAAGTGGCTGAACCGTGTGGAGACTTACATCAAGTGGCCCTGGTGGGTGTG
GCTCTGTATTTCCGTGGTCCTTATATTTGTTGTAAGTATGCTGCTCCTGTGCTGTTGCTCAACCG
GGTGCTGCGGTTTTTTCTCCTGTTTCGCCTCATCCATCCGTGGCTGTTGTGAGAGCACTAAACTG
CCATATTACGATGTCGAAAAGATCCACATACAGTGA

PDI-229E -H5iTMCT-DNA (서열 번호 : 149)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgCAAACGACTAATGGGCTGAACACCAGTTACAGCGTCTGTAACGGCTGCGTCGGATATA
GCGAGAACGTGTTCGCAGTGGAAAGTGGGGGGTACATTCCTTCCGACTTCGCTTTCAATAACTGG
TTTCTCCTGACTAACACAAGCTCCGTCGTGGATGGCGTGGTCAGGTCCTTTCAGCCTCTTCTCCT
GAATTGCCTGTGGTCTGTGTCCGGGTTAAGATTCACTACAGGCTTCGTATACTTCAACGGGACGG
GCCGGGGGGGATTGCAAGGGCTTCTCCTCCGACGTGCTGTCAGATGTGATCCGTTACAATCTGAAC
TTCGAAGAGAACTTACGGCGGGGGACAATCCTGTTCAAAACATCATATGGCGTAGTCGTATTTTA
CTGCACCAATAATACCCTGGTGAGTGGGGACGCCCATATTCCCTTCGGAACAGTGCTGGGTAACT
TTTACTGTTTTGTCAACACTACGATCGGAAACGAAACCACTAGCGCCTTTGTCGGAGCTCTGCCA
AAAACAGTTAGGGAGTTCGTGATCTCTCGGACCGGTCACTTCTATATCAACGGCTACCGTTATTT
TACTTTGGGCAACGTCGAAGCCGTCAATTTTAATGTGACAACTGCAGAGACAACTGACTTTTGCA
CTGTGGCTCTCGCCAGTTATGCCGATGTGCTGGTGAATGTAAGTCAAACGTCAATTGCCAACATC
ATCTATTGTAACTCAGTAATCAACCGGCTCCGCTGTGACCAACTCTCATTCGACGTCCCCGACGG
ATTCTATTCCACGAGCCCGATTCAGAGCGTGGAACTGCCAGTTTCCATCGTATCCCTCCCAGTTT
ACCACAAGCACACTTTTATCGTTCTCTACGTAGATTTTAAACCCCAGTCAGGAGGAGGGAAATGC
TTCAACTGCTACCCGGCTGGCGTGAACATCACCTTGGCCAATTTTAATGAAACTAAAGGGCCCCT
TTGCGTGGATACGTCACACTTTACCACAAAGTATGTTGCAGTCTATGCTAACGTCGGCAGGTGGT
CAGCGTCCATTAACACAGGCAATTGCCCGTTCTCTTTCGGGAAAGTGAACAACTTCGTGAAGTTT
GGAAGTGTGTGCTTCAGTTTGAAAGACATTCCGGGCGGCTGCGCCATGCCTATTGTGGCTAATTG
GGCTTATTCCAAGTACTACACCATTGGCTCTCTCTACGTTAGCTGGAGCGACGGTGACGGTATAA
CGGGCGTACCACAACCGGTGGAAGGGGTCAGCTCTTTCATGAATGTCACTCTGGACAAGTGTACC
AAATATAATATATACGATGTGAGTGGAGTGGGCGTTATACGCGTGTCTAACGACACCTTTCTAAA
CGGCATAACCTACACAAGCACGTCAGGCAATCTGTTAGGTTTTAAAGACGTCACTAAAGGCACTA
TATATAGCATCACCCCATGCAACCCACCTGATCAATTAGTCGTATATCAGCAAGCTGTTGTGGGT
GCTATGCTGTCAGAAAACTTCACCAGCTACGGGTTCTCCAATGTGGTGGAACTGCCCAAATTCTT
TTACGCTAGCAATGGCACATATAACTGTACTGACGCCGTCTTGACTTACAGTTCATTCGGAGTGT
GCGCGGACGGCAGCATTATCGCCGTGCAGCCGGCCAATGTCAGCTATGATTCCGTTTCCGCCATC
GTGACAGCCAACTTGTCGATTCCCTCTAACTGGACAACGTCTGTCCAAGTCGAATATCTGCAGAT
CACCTCAACCCCCATAGTAGTCGATTGCTCAACCTACGTCTGCAACGGTAATGTCAGATGTGTCG
AGCTGCTCAAGCAGTACACCTCCGCCTGTAAGACTATTGAGGATGCATTAAGAAATAGTGCAAGA
TTGGAAAGCGCCGATGTGTCGGAAATGCTAACCTTCGATAAGAAGGCATTCACACTGGCGAACGT
AAGCTCTTTCGGCGATTACAACCTGTCTTCGGTAATCCCTAGCTTGCCCACATCCGGCTCTCGGG

[0518]

TGGCGGGGCGGAGCGCTATCGAGGACATTTTATTCTCGAAACTGGTTACATCTGGGCTCGGAACT
GTGGACGCCGATTACAAGAAGTGCACCAAGGGCCTAAGCATCGCCGACCTCGCCTGTGCTCAGTA
CTACAACGGAATTATGGTGCTGCCAGGTGTCGCTGACGCAGAGCGGATGGCTATGTATACCGGCA
GTCTCATTGGCGGGATTGCGTTGGGCGGCCTGACGTCCGCTGTCTCCATCCCTTTCTCTCTGGCT
ATACAAGCCCGACTGAATTATGTGGCCCTGCAGACTGATGTCCTGCAAGAAAATCAGAAGATTCT
TGCCGCCAGCTTCAACAAGGCCATGACTAATATTGTGGATGCGTTTACCGGAGTGAATGACGCCA
TCACCCAAACGTCCCAAGCCCTGCAGACAGTCGCCACGGCGTTAAACAAAATCCAGGATGTAGTG
AATCAGCAAGGGAACAGCTTGAATCACCTGACGTCCCAGTTAAGACAGAACTTTCAGGCAATCAG
TAGCTCAATCCAGGCTATCTACGATCGATTAGATCCTCCTCAGGCAGATCAGCAGGTGGATCGGC
TCATCACCGGCCGCCTCGCGGCATTGAATGTTTTCGTAAGTCATACCTTGACCAAGTACACGGAG
GTGAGGGCCAGTCGCCAGCTGGCTCAGCAAAAAGTGAATGAGTGTGTGAAATCACAGAGCAAACG
GTACGGGTTTTGTGGAAATGGGACGCACATCTTTAGCATCGTTAATGCTGCCCCCGAAGGGTTAG
TCTTCCTGCACACTGTGCTCCTTCCTACCCAGTATAAAGATGTCGAAGCATGGTCTGGGCTCTGT
GTCGATGGAACTAACGGTTATGTCCTTCGACAGCCAAACCTCGCTCTCTATAAAGAAGGGAATTA
CTATAGGATCACCTCAAGAATCATGTTCGAGCCCAGGATACCAACAATGGCCGATTTTGTGCAGA
TTGAAAATTGTAACGTGACCTTTGTGAATATCAGTCGATCCGAGCTTCAAACGATTGTTCCTGAG
TACATCGACGTGAATAAAACTCTACAAGAGCTGTCCTATAAACTGCCTAATTATACCGTGCCTGA
CCTTGTAGTCGAGCAATACAACCAGACTATTCTGAACCTGACATCGGAAATCTCTACATTGGAGA
ATAAAAGCGCCGAGCTCAATTACACAGTGCAGAAGCTGCAGACCCTGATCGACAATATTAACAGC
ACTCTTGTGGACTTAAAGTGGCTGAACCGTGTGGAGACTTACATCAAGTGGCCCTGGTGGcaaat
actgtcaatttattcaacagtggcgagttccctagcactggcaatcatgatggctggtctatctt
tatggatgtgctccaatggatcgttacaatgcagaatttgcattTGA

PDI-229E -H5iCT-DNA (서열 번호 150)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgCAAACGACTAATGGGCTGAACACCAGTTACAGCGTCTGTAACGGCTGCGTCGGATATA
GCGAGAACGTGTTCGCAGTGGAAAGTGGGGGGTACATTCCTTCCGACTTCGCTTTCAATAACTGG
TTTCTCCTGACTAACACAAGCTCCGTCGTGGATGGCGTGGTCAGGTCCTTTCAGCCTCTTCTCCT
GAATTGCCTGTGGTCTGTGTCCGGGTTAAGATTCACTACAGGCTTCGTATACTTCAACGGGACGG
GCCGGGGGGATTGCAAGGGCTTCTCCTCCGACGTGCTGTCAGATGTGATCCGTTACAATCTGAAC
TTCGAAGAGAACTTACGGCGGGGGACAATCCTGTTCAAAACATCATATGGCGTAGTCGTATTTTA
CTGCACCAATAATACCCTGGTGAGTGGGGACGCCCATATTCCCTTCGGAACAGTGCTGGGTAACT
TTTACTGTTTTGTCAACACTACGATCGGAAACGAAACCACTAGCGCCTTTGTCGGAGCTCTGCCA
AAAACAGTTAGGGAGTTCGTGATCTCTCGGACCGGTCACTTCTATATCAACGGCTACCGTTATTT
TACTTTGGGCAACGTCGAAGCCGTCAATTTTAATGTGACAACTGCAGAGACAACTGACTTTTGCA
CTGTGGCTCTCGCCAGTTATGCCGATGTGCTGGTGAATGTAAGTCAAACGTCAATTGCCAACATC
ATCTATTGTAACTCAGTAATCAACCGGCTCCGCTGTGACCAACTCTCATTCGACGTCCCCGACGG
ATTCTATTCCACGAGCCCGATTCAGAGCGTGGAACTGCCAGTTTCCATCGTATCCCTCCCAGTTT
ACCACAAGCACACTTTTATCGTTCTCTACGTAGATTTTAAACCCCAGTCAGGAGGAGGGAAATGC
TTCAACTGCTACCCGGCTGGCGTGAACATCACCTTGGCCAATTTTAATGAAACTAAAGGGCCCCT
TTGCGTGGATACGTCACACTTTACCACAAAGTATGTTGCAGTCTATGCTAACGTCGGCAGGTGGT
CAGCGTCCATTAACACAGGCAATTGCCCGTTCTCTTTCGGGAAAGTGAACAACTTCGTGAAGTTT
GGAAGTGTGTGCTTCAGTTTGAAAGACATTCCGGGCGGCTGCGCCATGCCTATTGTGGCTAATTG
GGCTTATTCCAAGTACTACACCATTGGCTCTCTCTACGTTAGCTGGAGCGACGGTGACGGTATAA
CGGGCGTACCACAACCGGTGGAAGGGGTCAGCTCTTTCATGAATGTCACTCTGGACAAGTGTACC
AAATATAATATATACGATGTGAGTGGAGTGGGCGTTATACGCGTGTCTAACGACACCTTTCTAAA
CGGCATAACCTACACAAGCACGTCAGGCAATCTGTTAGGTTTTAAAGACGTCACTAAAGGCACTA
TATATAGCATCACCCCATGCAACCCACCTGATCAATTAGTCGTATATCAGCAAGCTGTTGTGGGT
GCTATGCTGTCAGAAAACTTCACCAGCTACGGGTTCTCCAATGTGGTGGAACTGCCCAAATTCTT
TTACGCTAGCAATGGCACATATAACTGTACTGACGCCGTCTTGACTTACAGTTCATTCGGAGTGT

[0519]

GCGCGGACGGCAGCATTATCGCCGTGCAGCCGGCCAATGTCAGCTATGATTCCGTTTCCGCCATC
GTGACAGCCAACTTGTCGATTCCCTCTAACTGGACAACGTCTGTCCAAGTCGAATATCTGCAGAT
CACCTCAACCCCCATAGTAGTCGATTGCTCAACCTACGTCTGCAACGGTAATGTCAGATGTGTCG
AGCTGCTCAAGCAGTACACCTCCGCCTGTAAGACTATTGAGGATGCATTAAGAAATAGTGCAAGA
TTGGAAAGCGCCGATGTGTCGGAAATGCTAACCTTCGATAAGAAGGCATTCACACTGGCGAACGT
AAGCTCTTTCGGCGATTACAACCTGTCTTCGGTAATCCCTAGCTTGCCCACATCCGGCTCTCGGG
TGGCGGGGCGGAGCGCTATCGAGGACATTTTATTCTCGAAACTGGTTACATCTGGGCTCGGAACT
GTGGACGCCGATTACAAGAAGTGCACCAAGGGCCTAAGCATCGCCGACCTCGCCTGTGCTCAGTA
CTACAACGGAATTATGGTGCTGCCAGGTGTCGCTGACGCAGAGCGGATGGCTATGTATACCGGCA
GTCTCATTGGCGGGATTGCGTTGGGCGGCCTGACGTCCGCTGTCTCCATCCCTTTCTCTCTGGCT
ATACAAGCCCGACTGAATTATGTGGCCCTGCAGACTGATGTCCTGCAAGAAAATCAGAAGATTCT
TGCCGCCAGCTTCAACAAGGCCATGACTAATATTGTGGATGCGTTTACCGGAGTGAATGACGCCA
TCACCCAAACGTCCCAAGCCCTGCAGACAGTCGCCACGGCGTTAAACAAAATCCAGGATGTAGTG
AATCAGCAAGGGAACAGCTTGAATCACCTGACGTCCCAGTTAAGACAGAACTTTCAGGCAATCAG
TAGCTCAATCCAGGCTATCTACGATCGATTAGATCCTCCTCAGGCAGATCAGCAGGTGGATCGGC
TCATCACCGGCCGCCTCGCGGCATTGAATGTTTTCGTAAGTCATACCTTGACCAAGTACACGGAG
GTGAGGGCCAGTCGCCAGCTGGCTCAGCAAAAAGTGAATGAGTGTGTGAAATCACAGAGCAAACG
GTACGGGTTTTGTGGAAATGGGACGCACATCTTTAGCATCGTTAATGCTGCCCCCGAAGGGTTAG
TCTTCCTGCACACTGTGCTCCTTCCTACCCAGTATAAAGATGTCGAAGCATGGTCTGGGCTCTGT
GTCGATGGAACTAACGGTTATGTCCTTCGACAGCCAAACCTCGCTCTCTATAAAGAAGGGAATTA
CTATAGGATCACCTCAAGAATCATGTTCGAGCCCAGGATACCAACAATGGCCGATTTTGTGCAGA
TTGAAAATTGTAACGTGACCTTTGTGAATATCAGTCGATCCGAGCTTCAAACGATTGTTCCTGAG
TACATCGACGTGAATAAAACTCTACAAGAGCTGTCCTATAAACTGCCTAATTATACCGTGCCTGA
CCTTGTAGTCGAGCAATACAACCAGACTATTCTGAACCTGACATCGGAAATCTCTACATTGGAGA
ATAAAAGCGCCGAGCTCAATTACACAGTGCAGAAGCTGCAGACCCTGATCGACAATATTAACAGC
ACTCTTGTGGACTTAAAGTGGCTGAACCGTGTGGAGACTTACATCAAGTGGCCCTGGTGGGTGTG
GCTCTGTATTTCCGTGGTCCTTATATTTGTTGTAAGTATGCTGCTCCTGtctttatggatgtgct
ccaatggatcgttacaatgcagaatttgcattTGA

PDI-229E -H5iCT(V4)-DNA (서열 번호 151)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgCAAACGACTAATGGGCTGAACACCAGTTACAGCGTCTGTAACGGCTGCGTCGGATATA
GCGAGAACGTGTTCGCAGTGGAAAGTGGGGGGTACATTCCTTCCGACTTCGCTTTCAATAACTGG
TTTCTCCTGACTAACACAAGCTCCGTCGTGGATGGCGTGGTCAGGTCCTTTCAGCCTCTTCTCCT
GAATTGCCTGTGGTCTGTGTCCGGGTTAAGATTCACTACAGGCTTCGTATACTTCAACGGGACGG
GCCGGGGGGATTGCAAGGGCTTCTCCTCCGACGTGCTGTCAGATGTGATCCGTTACAATCTGAAC
TTCGAAGAGAACTTACGGCGGGGGACAATCCTGTTCAAAACATCATATGGCGTAGTCGTATTTTA
CTGCACCAATAATACCCTGGTGAGTGGGGACGCCCATATTCCCTTCGGAACAGTGCTGGGTAACT
TTTACTGTTTTGTCAACACTACGATCGGAAACGAAACCACTAGCGCCTTTGTCGGAGCTCTGCCA
AAAACAGTTAGGGAGTTCGTGATCTCTCGGACCGGTCACTTCTATATCAACGGCTACCGTTATTT
TACTTTGGGCAACGTCGAAGCCGTCAATTTTAATGTGACAACTGCAGAGACAACTGACTTTTGCA
CTGTGGCTCTCGCCAGTTATGCCGATGTGCTGGTGAATGTAAGTCAAACGTCAATTGCCAACATC
ATCTATTGTAACTCAGTAATCAACCGGCTCCGCTGTGACCAACTCTCATTCGACGTCCCCGACGG
ATTCTATTCCACGAGCCCGATTCAGAGCGTGGAACTGCCAGTTTCCATCGTATCCCTCCCAGTTT
ACCACAAGCACACTTTTATCGTTCTCTACGTAGATTTTAAACCCCAGTCAGGAGGAGGGAAATGC
TTCAACTGCTACCCGGCTGGCGTGAACATCACCTTGGCCAATTTTAATGAAACTAAAGGGCCCCT
TTGCGTGGATACGTCACACTTTACCACAAAGTATGTTGCAGTCTATGCTAACGTCGGCAGGTGGT
CAGCGTCCATTAACACAGGCAATTGCCCGTTCTCTTTCGGGAAAGTGAACAACTTCGTGAAGTTT
GGAAGTGTGTGCTTCAGTTTGAAAGACATTCCGGGCGGCTGCGCCATGCCTATTGTGGCTAATTG
GGCTTATTCCAAGTACTACACCATTGGCTCTCTCTACGTTAGCTGGAGCGACGGTGACGGTATAA

[0520]

CGGGCGTACCACAACCGGTGGAAGGGGTCAGCTCTTTCATGAATGTCACTCTGGACAAGTGTACC
AAATATAATATATACGATGTGAGTGGAGTGGGCGTTATACGCGTGTCTAACGACACCTTTCTAAA
CGGCATAACCTACACAAGCACGTCAGGCAATCTGTTAGGTTTTAAAGACGTCACTAAAGGCACTA
TATATAGCATCACCCCATGCAACCCACCTGATCAATTAGTCGTATATCAGCAAGCTGTTGTGGGT
GCTATGCTGTCAGAAAACTTCACCAGCTACGGGTTCTCCAATGTGGTGGAACTGCCCAAATTCTT
TTACGCTAGCAATGGCACATATAACTGTACTGACGCCGTCTTGACTTACAGTTCATTCGGAGTGT
GCGCGGACGGCAGCATTATCGCCGTGCAGCCGGCCAATGTCAGCTATGATTCCGTTTCCGCCATC
GTGACAGCCAACTTGTCGATTCCCTCTAACTGGACAACGTCTGTCCAAGTCGAATATCTGCAGAT
CACCTCAACCCCCATAGTAGTCGATTGCTCAACCTACGTCTGCAACGGTAATGTCAGATGTGTCG
AGCTGCTCAAGCAGTACACCTCCGCCTGTAAGACTATTGAGGATGCATTAAGAAATAGTGCAAGA
TTGGAAAGCGCCGATGTGTCGGAAATGCTAACCTTCGATAAGAAGGCATTCACACTGGCGAACGT
AAGCTCTTTCGGCGATTACAACCTGTCTTCGGTAATCCCTAGCTTGCCCACATCCGGCTCTCGGG
TGGCGGGGCGGAGCGCTATCGAGGACATTTTATTCTCGAAACTGGTTACATCTGGGCTCGGAACT
GTGGACGCCGATTACAAGAAGTGCACCAAGGGCCTAAGCATCGCCGACCTCGCCTGTGCTCAGTA
CTACAACGGAATTATGGTGCTGCCAGGTGTCGCTGACGCAGAGCGGATGGCTATGTATACCGGCA
GTCTCATTGGCGGGATTGCGTTGGGCGGCCTGACGTCCGCTGTCTCCATCCCTTTCTCTCTGGCT
ATACAAGCCCGACTGAATTATGTGGCCCTGCAGACTGATGTCCTGCAAGAAAATCAGAAGATTCT
TGCCGCCAGCTTCAACAAGGCCATGACTAATATTGTGGATGCGTTTACCGGAGTGAATGACGCCA
TCACCCAAACGTCCCAAGCCCTGCAGACAGTCGCCACGGCGTTAAACAAAATCCAGGATGTAGTG
AATCAGCAAGGGAACAGCTTGAATCACCTGACGTCCCAGTTAAGACAGAACTTTCAGGCAATCAG
TAGCTCAATCCAGGCTATCTACGATCGATTAGATCCTCCTCAGGCAGATCAGCAGGTGGATCGGC
TCATCACCGGCCGCCTCGCGGCATTGAATGTTTTCGTAAGTCATACCTTGACCAAGTACACGGAG
GTGAGGGCCAGTCGCCAGCTGGCTCAGCAAAAAGTGAATGAGTGTGTGAAATCACAGAGCAAACG
GTACGGGTTTTGTGGAAATGGGACGCACATCTTTAGCATCGTTAATGCTGCCCCCGAAGGGTTAG
TCTTCCTGCACACTGTGCTCCTTCCTACCCAGTATAAAGATGTCGAAGCATGGTCTGGGCTCTGT
GTCGATGGAACTAACGGTTATGTCCTTCGACAGCCAAACCTCGCTCTCTATAAAGAAGGGAATTA
CTATAGGATCACCTCAAGAATCATGTTCGAGCCCAGGATACCAACAATGGCCGATTTTGTGCAGA
TTGAAAATTGTAACGTGACCTTTGTGAATATCAGTCGATCCGAGCTTCAAACGATTGTTCCTGAG
TACATCGACGTGAATAAAACTCTACAAGAGCTGTCCTATAAACTGCCTAATTATACCGTGCCTGA
CCTTGTAGTCGAGCAATACAACCAGACTATTCTGAACCTGACATCGGAAATCTCTACATTGGAGA
ATAAAAGCGCCGAGCTCAATTACACAGTGCAGAAGCTGCAGACCCTGATCGACAATATTAACAGC
ACTCTTGTGGACTTAAAGTGGCTGAACCGTGTGGAGACTTACATCAAGTGGCCCTGGTGGGTGTG
GCTCTGTATTTCCGTGGTCCTTATATTTGTTGTAAGTATGCTGCTCCTGtgctgctccaatggat
cgttacaatgcagaatttgcattTGA

PDI-229E -H1cCT-DNA (서열 번호 152)

atggcgaaaaacgttgcgattttcggcttattgttttctcttcttgtgttggttccttctcagat
cttcgcgCAAACGACTAATGGGCTGAACACCAGTTACAGCGTCTGTAACGGCTGCGTCGGATATA
GCGAGAACGTGTTCGCAGTGGAAAGTGGGGGGTACATTCCTTCCGACTTCGCTTTCAATAACTGG
TTTCTCCTGACTAACACAAGCTCCGTCGTGGATGGCGTGGTCAGGTCCTTTCAGCCTCTTCTCCT
GAATTGCCTGTGGTCTGTGTCCGGGTTAAGATTCACTACAGGCTTCGTATACTTCAACGGGACGG
GCCGGGGGGATTGCAAGGGCTTCTCCTCCGACGTGCTGTCAGATGTGATCCGTTACAATCTGAAC
TTCGAAGAGAACTTACGGCGGGGGACAATCCTGTTCAAAACATCATATGGCGTAGTCGTATTTTA
CTGCACCAATAATACCCTGGTGAGTGGGGACGCCCATATTCCCTTCGGAACAGTGCTGGGTAACT
TTTACTGTTTTGTCAACACTACGATCGGAAACGAAACCACTAGCGCCTTTGTCGGAGCTCTGCCA
AAAACAGTTAGGGAGTTCGTGATCTCTCGGACCGGTCACTTCTATATCAACGGCTACCGTTATTT
TACTTTGGGCAACGTCGAAGCCGTCAATTTTAATGTGACAACTGCAGAGACAACTGACTTTTGCA
CTGTGGCTCTCGCCAGTTATGCCGATGTGCTGGTGAATGTAAGTCAAACGTCAATTGCCAACATC
ATCTATTGTAACTCAGTAATCAACCGGCTCCGCTGTGACCAACTCTCATTCGACGTCCCCGACGG
ATTCTATTCCACGAGCCCGATTCAGAGCGTGGAACTGCCAGTTTCCATCGTATCCCTCCCAGTTT

[0521]

ACCACAAGCACACTTTTATCGTTCTCTACGTAGATTTTAAACCCCAGTCAGGAGGAGGGAAATGC
TTCAACTGCTACCCGGCTGGCGTGAACATCACCTTGGCCAATTTTAATGAAACTAAAGGGCCCCT
TTGCGTGGATACGTCACACTTTACCACAAAGTATGTTGCAGTCTATGCTAACGTCGGCAGGTGGT
CAGCGTCCATTAACACAGGCAATTGCCCGTTCTCTTTCGGGAAAGTGAACAACTTCGTGAAGTTT
GGAAGTGTGTGCTTCAGTTTGAAAGACATTCCGGGCGGCTGCGCCATGCCTATTGTGGCTAATTG
GGCTTATTCCAAGTACTACACCATTGGCTCTCTCTACGTTAGCTGGAGCGACGGTGACGGTATAA
CGGGCGTACCACAACCGGTGGAAGGGGTCAGCTCTTTCATGAATGTCACTCTGGACAAGTGTACC
AAATATAATATATACGATGTGAGTGGAGTGGGCGTTATACGCGTGTCTAACGACACCTTTCTAAA
CGGCATAACCTACACAAGCACGTCAGGCAATCTGTTAGGTTTTAAAGACGTCACTAAAGGCACTA
TATATAGCATCACCCCATGCAACCCACCTGATCAATTAGTCGTATATCAGCAAGCTGTTGTGGGT
GCTATGCTGTCAGAAAACTTCACCAGCTACGGGTTCTCCAATGTGGTGGAACTGCCCAAATTCTT
TTACGCTAGCAATGGCACATATAACTGTACTGACGCCGTCTTGACTTACAGTTCATTCGGAGTGT
GCGCGGACGGCAGCATTATCGCCGTGCAGCCGGCCAATGTCAGCTATGATTCCGTTTCCGCCATC
GTGACAGCCAACTTGTCGATTCCCTCTAACTGGACAACGTCTGTCCAAGTCGAATATCTGCAGAT
CACCTCAACCCCCATAGTAGTCGATTGCTCAACCTACGTCTGCAACGGTAATGTCAGATGTGTCG
AGCTGCTCAAGCAGTACACCTCCGCCTGTAAGACTATTGAGGATGCATTAAGAAATAGTGCAAGA
TTGGAAAGCGCCGATGTGTCGGAAATGCTAACCTTCGATAAGAAGGCATTCACACTGGCGAACGT
AAGCTCTTTCGGCGATTACAACCTGTCTTCGGTAATCCCTAGCTTGCCCACATCCGGCTCTCGGG
TGGCGGGGCGGAGCGCTATCGAGGACATTTTATTCTCGAAACTGGTTACATCTGGGCTCGGAACT
GTGGACGCCGATTACAAGAAGTGCACCAAGGGCCTAAGCATCGCCGACCTCGCCTGTGCTCAGTA
CTACAACGGAATTATGGTGCTGCCAGGTGTCGCTGACGCAGAGCGGATGGCTATGTATACCGGCA
GTCTCATTGGCGGGATTGCGTTGGGCGGCCTGACGTCCGCTGTCTCCATCCCTTTCTCTCTGGCT
ATACAAGCCCGACTGAATTATGTGGCCCTGCAGACTGATGTCCTGCAAGAAAATCAGAAGATTCT
TGCCGCCAGCTTCAACAAGGCCATGACTAATATTGTGGATGCGTTTACCGGAGTGAATGACGCCA
TCACCCAAACGTCCCAAGCCCTGCAGACAGTCGCCACGGCGTTAAACAAAATCCAGGATGTAGTG
AATCAGCAAGGGAACAGCTTGAATCACCTGACGTCCCAGTTAAGACAGAACTTTCAGGCAATCAG
TAGCTCAATCCAGGCTATCTACGATCGATTAGATCCTCCTCAGGCAGATCAGCAGGTGGATCGGC
TCATCACCGGCCGCCTCGCGGCATTGAATGTTTTCGTAAGTCATACCTTGACCAAGTACACGGAG
GTGAGGGCCAGTCGCCAGCTGGCTCAGCAAAAAGTGAATGAGTGTGTGAAATCACAGAGCAAACG
GTACGGGTTTTGTGGAAATGGGACGCACATCTTTAGCATCGTTAATGCTGCCCCCGAAGGGTTAG
TCTTCCTGCACACTGTGCTCCTTCCTACCCAGTATAAAGATGTCGAAGCATGGTCTGGGCTCTGT
GTCGATGGAACTAACGGTTATGTCCTTCGACAGCCAAACCTCGCTCTCTATAAAGAAGGGAATTA
CTATAGGATCACCTCAAGAATCATGTTCGAGCCCAGGATACCAACAATGGCCGATTTTGTGCAGA
TTGAAAATTGTAACGTGACCTTTGTGAATATCAGTCGATCCGAGCTTCAAACGATTGTTCCTGAG
TACATCGACGTGAATAAAACTCTACAAGAGCTGTCCTATAAACTGCCTAATTATACCGTGCCTGA
CCTTGTAGTCGAGCAATACAACCAGACTATTCTGAACCTGACATCGGAAATCTCTACATTGGAGA
ATAAAAGCGCCGAGCTCAATTACACAGTGCAGAAGCTGCAGACCCTGATCGACAATATTAACAGC
ACTCTTGTGGACTTAAAGTGGCTGAACCGTGTGGAGACTTACATCAAGTGGCCCTGGTGGGTGTG
GCTCTGTATTTCCGTGGTCCTTATATTTGTTGTAAGTATGCTGCTCCTGagcttctggatgtgct
ctaatgggtctctacagtgtagaatatgtattTGA

PDI-229E -wtTMCT-AA (서열 번호 153)

MAKNVAIFGLLFSLLVLVPSQIFAQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNW
FLLTNTSSVVDGVVRSFQPLLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLN
FEENLRRGTILFKTSYGVVVFYCTNNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALP
KTVREFVISRTGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANI
IYCNSVINRLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKC
FNCYPAGVNITLANFNETKGPLCVDTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKF
GSVCFSLKDIPGGCAMPIVANWAYSKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCT
KYNIYDVSGVGVIRVSNDTFLNGITYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVG

[0522]

AMLSENFTSYGFSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQPANVSYDSVSAI
VTANLSIPSNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSAR
LESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGT
VDADYKKCTKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLA
IQARLNYVALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVV
NQQGNSLNHLTSQLRQNFQAISSSIQAIYDRLDPPQADQQVDRLITGRLAALNVFVSHTLTKYTE
VRASRQLAQQKVNECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLC
VDGTNGYVLRQPNLALYKEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPE
YIDVNKTLQELSYKLPNYTVPDLVVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINS
TLVDLKWLNRVETYIKWPWWVWLCISVVLIFVVSMLLLCCCSTGCCGFFSCFASSIRGCCESTKL
PYYDVEKIHIQ

PDI-229E -H5iTMCT-AA (서열 번호 154)

MAKNVAIFGLLFSLLVLVPSQIFAQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNW
FLLTNTSSVVDGVVRSFQPLLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLN
FEENLRRGTILFKTSYGVVVFYCTNNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALP
KTVREFVISRTGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANI
IYCNSVINRLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKC
FNCYPAGVNITLANFNETKGPLCVDTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKF
GSVCFSLKDIPGGCAMPIVANWAYSKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCT
KYNIYDVSGVGVIRVSNDTFLNGITYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVG
AMLSENFTSYGFSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQPANVSYDSVSAI
VTANLSIPSNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSAR
LESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGT
VDADYKKCTKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLA
IQARLNYVALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVV
NQQGNSLNHLTSQLRQNFQAISSSIQAIYDRLDPPQADQQVDRLITGRLAALNVFVSHTLTKYTE
VRASRQLAQQKVNECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLC
VDGTNGYVLRQPNLALYKEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPE
YIDVNKTLQELSYKLPNYTVPDLVVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINS
TLVDLKWLNRVETYIKWPWWVWLCISVVLIFVVSMLLLSLWMCSNGSLQCRICI

PDI-229E -H5iCT-AA (서열 번호 155)

MAKNVAIFGLLFSLLVLVPSQIFAQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNW
FLLTNTSSVVDGVVRSFQPLLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLN
FEENLRRGTILFKTSYGVVVFYCTNNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALP
KTVREFVISRTGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANI
IYCNSVINRLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKC
FNCYPAGVNITLANFNETKGPLCVDTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKF
GSVCFSLKDIPGGCAMPIVANWAYSKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCT
KYNIYDVSGVGVIRVSNDTFLNGITYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVG
AMLSENFTSYGFSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQPANVSYDSVSAI
VTANLSIPSNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSAR
LESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGT
VDADYKKCTKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLA
IQARLNYVALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVV
NQQGNSLNHLTSQLRQNFQAISSSIQAIYDRLDPPQADQQVDRLITGRLAALNVFVSHTLTKYTE
VRASRQLAQQKVNECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLC
VDGTNGYVLRQPNLALYKEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPE

YIDVNKTLQELSYKLPNYTVPDLVVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINS
TLVDLKWLNRVETYIKWPWWVWLCISVVLIFVVSMLLLSLWMCSNGSLQCRICI

[0523]

PDI-229E -H5iCT(V4)-AA (서열 번호 156)

MAKNVAIFGLLFSLLVLVPSQIFAQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNW
FLLTNTSSVVDGVVRSFQPLLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLN
FEENLRRGTILFKTSYGVVVFYCTNNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALP
KTVREFVISRTGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANI
IYCNSVINRLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKC
FNCYPAGVNITLANFNETKGPLCVDTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKF
GSVCFSLKDIPGGCAMPIVANWAYSKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCT
KYNIYDVSGVGVIRVSNDTFLNGITYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVG
AMLSENFTSYGFSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQPANVSYDSVSAI
VTANLSIPSNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSAR
LESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGT
VDADYKKCTKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLA
IQARLNYVALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVV
NQQGNSLNHLTSQLRQNFQAISSSIQAIYDRLDPPQADQQVDRLITGRLAALNVFVSHTLTKYTE
VRASRQLAQQKVNECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLC
VDGTNGYVLRQPNLALYKEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPE
YIDVNKTLQELSYKLPNYTVPDLVVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINS
TLVDLKWLNRVETYIKWPWWVWLCISVVLIFVVSMLLLCCSNGSLQCRICI

PDI-229E-H1cCT-AA (서열 번호 157)

MAKNVAIFGLLFSLLVLVPSQIFAQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNW
FLLTNTSSVVDGVVRSFQPLLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLN
FEENLRRGTILFKTSYGVVVFYCTNNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALP
KTVREFVISRTGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANI
IYCNSVINRLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKC
FNCYPAGVNITLANFNETKGPLCVDTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKF
GSVCFSLKDIPGGCAMPIVANWAYSKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCT
KYNIYDVSGVGVIRVSNDTFLNGITYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVG
AMLSENFTSYGFSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQPANVSYDSVSAI
VTANLSIPSNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSAR
LESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGT
VDADYKKCTKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLA
IQARLNYVALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVV
NQQGNSLNHLTSQLRQNFQAISSSIQAIYDRLDPPQADQQVDRLITGRLAALNVFVSHTLTKYTE
VRASRQLAQQKVNECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLC
VDGTNGYVLRQPNLALYKEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPE
YIDVNKTLQELSYKLPNYTVPDLVVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINS
TLVDLKWLNRVETYIKWPWWVWLCISVVLIFVVSMLLLSFWMCSNGSLQCRICI

천연 OC43-CoV S 단백질 wtTM/CT AA (AVR40344) (서열 번호 158)

MFLILLISLPTAFAVIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTYYVLDRVYLNT
TLFLNGYYPTSGSTYRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKDGVMYSEFPAI

[0524]

TIGSTFVNTSYSVVVQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNLGNHFKELWHY
DTGVVSCLYKRNFTYDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMALSHYYVMPLT
CIRRPKDGFSLEYWVTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIAPPTGVYELNG
YTVQPVADVYRRKPDLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFIQADSFTCNNI
DAAKIYGMCFSSITIDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYNLPAANVSVSR
FNPSTWNKRFGFIEDSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGKNNGIGTCPAG
TNSLTCDNLCTLDPITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTCQPQAFLGWSA
DSCLQGDKCNIFANFILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGIFVEVNATYYN
SWQNLLYDSNGNLYGFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKCNYVFNNSLTR
QLQPINYSFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNRRSRRAITTGYRFTNFEPFT
VNSVNDSLEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACKLQLVEYGSFC
DNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCLGSECSKASSR
SAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQISGYTLAATSA
SLFPPWTAAAGVPFYLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFDATNSALVKIQ
AVVNANAEALNNLLQQLSNRFGAISASLQEILSRLDALEAEAQIDRLINGRLTALNAYVSQQLSD
STLVKFSAAQAMEKVNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVPTKYVTAKVSP
GLCIAGNRGIAPKSGYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKAPYVMLNTSIP
NLPDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYINLKDIGTYEYY
VKWPWYVWLLICLAGVAMLVLLFFICCCTGCGTSCFKKCGGCCDDYTGYQELVIKTSHDD

천연 229E S 단백질 wtTM/CT AA (P15423) (서열 번호 159)

MFVLLVAYALLHIAGCQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNWFLLTNTSS
VVDGVVRSFQPLLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLNFEENLRRG
TILFKTSYGVVVFYCTNNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALPKTVREFVI
SRTGHFYINGYRYFTLGNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANIIYCNSVIN
RLRCDQLSFDVPDGFYSTSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKCFNCYPAGV
NITLANFNETKGPLCVDTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKFGSVCFSLK
DIPGGCAMPIVANWAYSKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCTKYNIYDVS
GVGVIRVSNDTFLNGITYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVGAMLSENFT
SYGFSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQPRNVSYDSVSAIVTANLSIP
SNWTTSVQVEYLQITSTPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSARLESADVSE
MLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGTVDADYKKC
TKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLAIQARLNYV
ALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVVNQQGNSLN
HLTSQLRQNFQAISSSIQAIYDRLDTIQADQQVDRLITGRLAALNVFVSHTLTKYTEVRASRQLA
QQKVNECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLCVDGTNGYV
LRQPNLALYKEGNYYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPEYIDVNKTL
QELSYKLPNYTVPDLVVEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINSTLVDLKWL
NRVETYIKWPWWVWLCISVVLIFVVSMLLLCCCSTGCCGFFSCFASSIRGCCESTKLPYYDVEKI
HIQ

신호 펩티드를 포함하지 않는 천연 OC43-CoV S 단백질 wtTM/CT AA (AVR40344)
(서열 번호 160)

VIGDLNCTLDPRLKGSFNNRDTGPPSISIDTVDVTNGLGTYYVLDRVYLNTTLFLNGYYPTSGST
YRNMALKGTDLLSTLWFKPPFLSDFINGIFAKVKNTKVFKDGVMYSEFPAITIGSTFVNTSYSVV
VQPRTINSTQDGVNKLQGLLEVSVCQYNMCEYPHTICHPNLGNHFKELWHYDTGVVSCLYKRNFT
YDVNATYLYFHFYQEGGTFYAYFTDTGFVTKFLFNVYLGMALSHYYVMPLTCIRRPKDGFSLEYW
VTPLTPRQYLLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIAPPTGVYELNGYTVQPVADVYRRKP

[0525]

DLPNCNIEAWLNDKSVPSPLNWERKTFSNCNFNMSSLMSFIQADSFTCNNIDAAKIYGMCFSSIT
IDKFAIPNRRKVDLQLGNLGYLQSSNYRIDTTATSCQLYYNLPAANVSVSRFNPSTWNKRFGFIE
DSVFVPQPTGVFTNHSVVYAQHCFKAPKNFCPCSSCSCPGKNNGIGTCPAGTNSLTCDNLCTLDP
ITLKAPDTYKCPQSKSLVGIGEHCSGLAVKSDYCGNNSCTCQPQAFLGWSADSCLQGDKCNIFAN
FILHDVNNGLTCSTDLQKANTEIELGVCVNYDLYGISGQGIFVEVNATYYNSWQNLLYDSNGNLY
GFRDYITNRTFMIHSCYSGRVSAAYHANSSEPALLFRNIKCNYVFNNSLTRQLQPINYSFDSYLG
CVVNAYNSTAISVQTCDLTVGSGYCVDYSKNRRSRRAITTGYRFTNFEPFTVNSVNDSLEPVGGL
YEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACKLQLVEYGSFCDNINAILTEVNELL
DTTQLQVANSLMNGVTLSTKLKDGVNFNVDDINFSPVLGCLGSECSKASSRSAIEDLLFDKVKLS
DVGFVEAYNNCTGGAEIRDLICVQSYKGIKVLPPLLSENQISGYTLAATSASLFPPWTAAAGVPF
YLNVQYRINGLGVTMDVLSQNQKLIANAFNNALHAIQQGFDATNSALVKIQAVVNANAEALNNLL
QQLSNRFGAISASLQEILSRLDALEAEAQIDRLINGRLTALNAYVSQQLSDSTLVKFSAAQAMEK
VNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFNYVPTKYVTAKVSPGLCIAGNRGIAPKS
GYFVNVNNTWMYTGSGYYYPEPITENNVVVMSTCAVNYTKAPYVMLNTSIPNLPDFKEELDQWFK
NQTSVAPDLSLDYINVTFLDLQVEMNRLQEAIKVLNHSYINLKDIGTYEYYVKWPWYVWLLICLA
GVAMLVLLFFICCCTGCGTSCFKKCGGCCDDYTGYQELVIKTSHDD

신호 펩티드를 포함하지 않는 천연 229E S 단백질 wtTM/CT AA (P15423) (서열 번호 161)

QTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNWFLLTNTSSVVDGVVRSFQPLLLNC
LWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLNFEENLRRGTILFKTSYGVVVFYCT
NNTLVSGDAHIPFGTVLGNFYCFVNTTIGNETTSAFVGALPKTVREFVISRTGHFYINGYRYFTL
GNVEAVNFNVTTAETTDFCTVALASYADVLVNVSQTSIANIIYCNSVINRLRCDQLSFDVPDGFY
STSPIQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKCFNCYPAGVNITLANFNETKGPLCV
DTSHFTTKYVAVYANVGRWSASINTGNCPFSFGKVNNFVKFGSVCFSLKDIPGGCAMPIVANWAY
SKYYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCTKYNIYDVSGVGVIRVSNDTFLNGI
TYTSTSGNLLGFKDVTKGTIYSITPCNPPDQLVVYQQAVVGAMLSENFTSYGFSNVVELPKFFYA
SNGTYNCTDAVLTYSSFGVCADGSIIAVQPRNVSYDSVSAIVTANLSIPSNWTTSVQVEYLQITS
TPIVVDCSTYVCNGNVRCVELLKQYTSACKTIEDALRNSARLESADVSEMLTFDKKAFTLANVSS
FGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKLVTSGLGTVDADYKKCTKGLSIADLACAQYYN
GIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFSLAIQARLNYVALQTDVLQENQKILAA
SFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVVNQQGNSLNHLTSQLRQNFQAISSS
IQAIYDRLDTIQADQQVDRLITGRLAALNVFVSHTLTKYTEVRASRQLAQQKVNECVKSQSKRYG
FCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLCVDGTNGYVLRQPNLALYKEGNYYR
ITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIVPEYIDVNKTLQELSYKLPNYTVPDLV
VEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINSTLVDLKWLNRVETYIKWPWWVWLC
ISVVLIFVVSMLLLCCCSTGCCGFFSCFASSIRGCCESTKLPYYDVEKIHIQ

변형된 PDI-OC43-COV wtTMCT의 TMCT 영역-AA (서열 번호 162)

WYVWLLICLAGVAMLVLLFFICCCTGCGTSCFKKCGGCCDDYTGYQELVIKTSHDD

변형된 PDI-OC43-COV H5iTMCT의 TMCT 영역-AA (서열 번호 163)

WYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

변형된 PDI-OC43-COV H5iCT의 TMCT 영역-AA (서열 번호 164)

WYVWLLICLAGVAMLVLLFFISLWMCSNGSLQCRICI

[0526]

변형된 PDI-OC43-COV H5iCT(V4)의 TMCT 영역-AA (서열 번호 165)

WYVWLLICLAGVAMLVLLFFICCSNGSLQCRICI

변형된 PDI-OC43-COV H1cCT의 TMCT 영역-AA (서열 번호 166)

WYVWLLICLAGVAMLVLLFFISFWMCSNGSLQCRICI

변형된 PDI-229E-wtTMCT의 TMCT 영역-AA (서열 번호 167)

WWVWLCISVVLIFVVSMLLLCCCSTGCCGFFSCFASSIRGCCESTKLPYYDVEKIHIQ

변형된 PDI-229E-H5iTMCT의 TMCT 영역-AA (서열 번호 168)

WWQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

변형된 PDI-229E-H5iCT의 TMCT 영역-AA(서열 번호 169)

WWVWLCISVVLIFVVSMLLLSLWMCSNGSLQCRICI

변형된 PDI-229E-H5iCT(V4)의 TMCT 영역-AA (서열 번호 170)

WWVWLCISVVLIFVVSMLLLCCSNGSLQCRICI

변형된 PDI-229E-H1cCT의 TMCT 영역-AA (서열 번호 171)

WWVWLCISVVLIFVVSMLLLSFWMCSNGSLQCRICI

개재 펩티드 서열 X_n을 포함하는 변형된 OC43-CoV S 단백질의 TM/CT 영역
(서열 번호 172)

WYVWLLICLAGVAMLVLLFFI -(X)n - CSNGSXXCXICI

개재 펩티드 서열 X_n을 포함하는 변형된 OC43-CoV S 단백질의 TM/CT 영역 (서열 번호 173)

WWVWLCISVVLIFVVSMLLL -(X)n - CSNGSXXCXICI

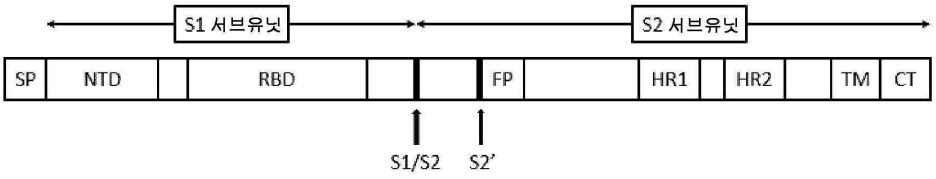
[0527]

[0528] 모든 인용은 본원에서 참조로 포함된다.

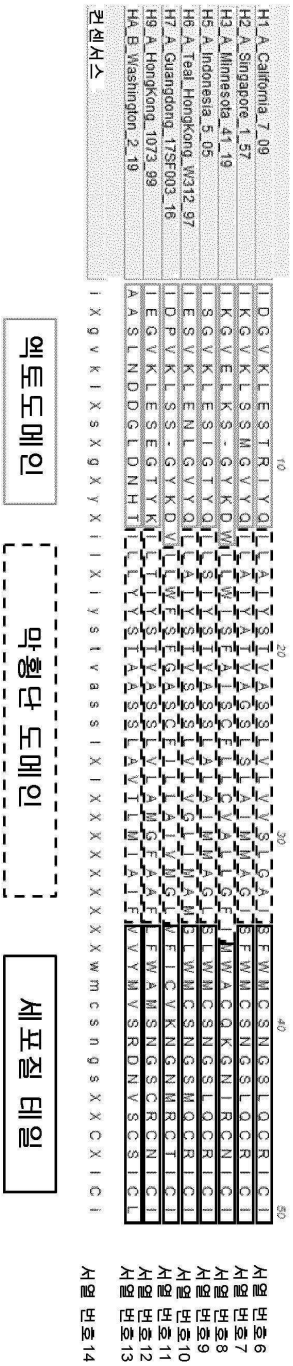
[0529] 본 발명은 하나 이상의 실시예와 관련하여 기술되었다. 그러나, 청구범위에 정의된 바와 같은 본 발명의 범주를 벗어나지 않으면서, 다수의 변형 및 수정이 이루어질 수 있다는 것이 당업자에게 자명할 것이다.

도면

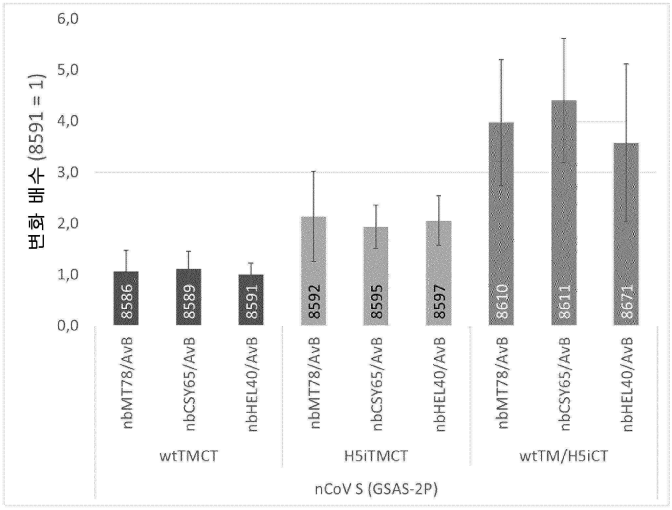
도면1



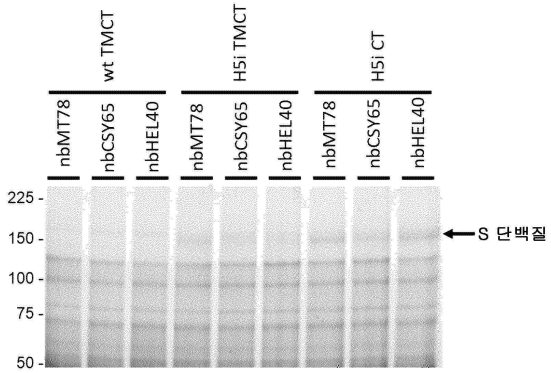
도면2


헤드도메인
막힐단 도메인
세포질 테일
서열 번호 6
서열 번호 7
서열 번호 8
서열 번호 9
서열 번호 10
서열 번호 11
서열 번호 12
서열 번호 13
서열 번호 14

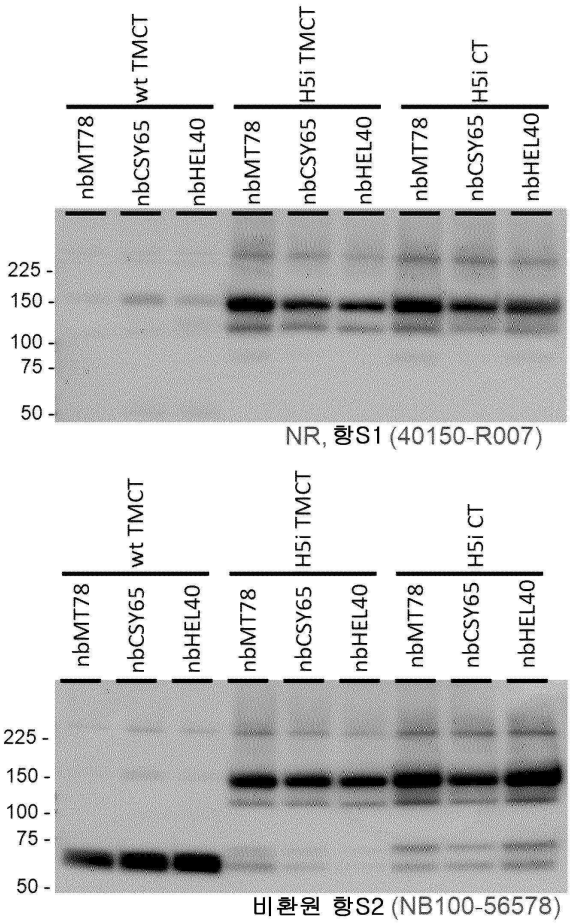
도면3a


변화 배수 (8591 = 1)
6,0
5,0
4,0
3,0
2,0
1,0
0,0
8586
8589
8591
8592
8595
8597
8610
8611
8671
nbMT78/AvB
nbCSY65/AvB
nbHEL40/AvB
wtTMCT
H5iTMCT
wtTM/H5iCT
nCoV S (GSAS-2P)

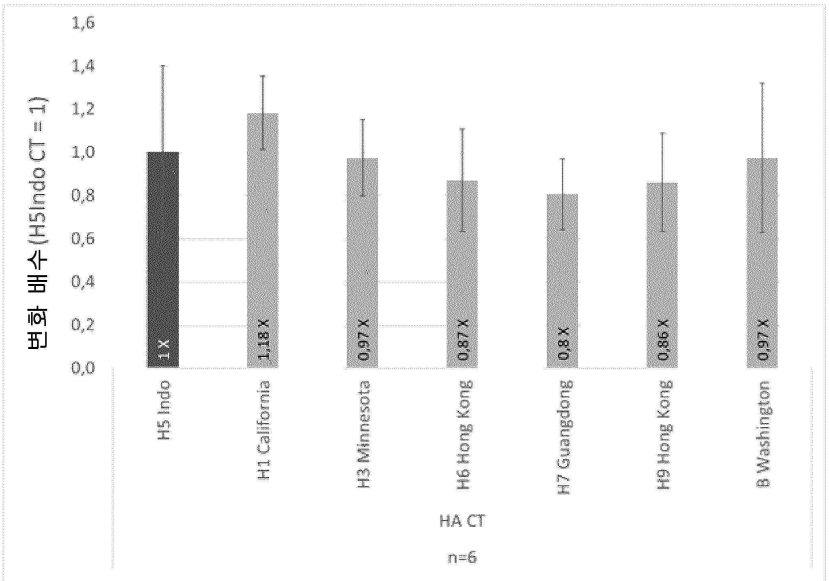
도면3b


wt TMCT
H5i TMCT
H5i CT
nbMT78
nbCSY65
nbHEL40
225
150
100
75
50
S 단백질

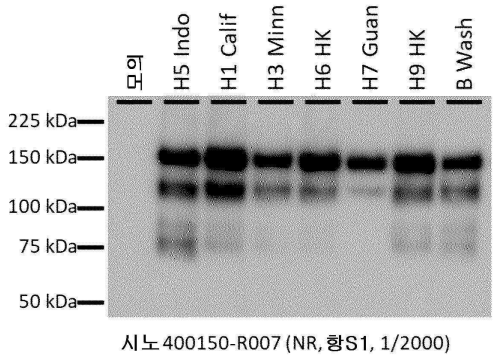
도면3c


wt TMCT
H5i TMCT
H5i CT
nbMT78
nbCSY65
nbHEL40
225
150
100
75
50
NR, 항S1 (40150-R007)
비환원 항S2 (NB100-56578)

도면4a

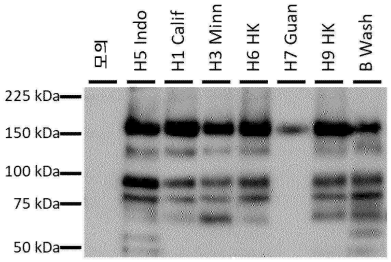

변화 배수 (H5Indo CT = 1)
1,6
1,4
1,2
1,0
0,8
0,6
0,4
0,2
0,0
1 X
1,18 X
0,97 X
0,87 X
0,8 X
0,86 X
0,97 X
H5 Indo
H1 California
H3 Minnesota
H6 Hong Kong
H7 Guangdong
H9 Hong Kong
B Washington
HA CT
n=6

도면4b



시노400150-R007 (NR, 항S1, 1/2000)

도면4c



NB100-56578 (NR, 항S2 , 1/5000)

도면5a

wtTM/wtCT를 포함하는 SARS-CoV-2 S 단백질

…YEQYIKWPWYIWLGFIAGLIAIVMVTIML**CCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT**

H5iTM/H5iCT를 포함하는 H5 A/Indonesia/5/05의 헤마글루티닌

…EISGVKLESIGTYQILSIYSTVASSLALAIMMAGL**SLWMCSNGSLQCRICI**

wtTM/H5iCT를 포함하는 변형된 SARS-CoV-2 S 단백질

…YEQYIKWPWYIWLGFIAGLIAIVMVTIML**SLWMCSNGSLQCRICI**

wtTM/H5iCT를 포함하는 변형된 SARS-CoV-2 S 단백질(V1)

…YEQYIKWPWYIWLGFIAGLIAIVMVTIM**MAGLSLWMCSNGSLQCRICI**

wtTM/H5iCT를 포함하는 변형된 SARS-CoV-2 S 단백질(V2)

…YEQYIKWPWYIWLGFIAGLIAIVMVTIM**AGLSLWMCSNGSLQCRICI**

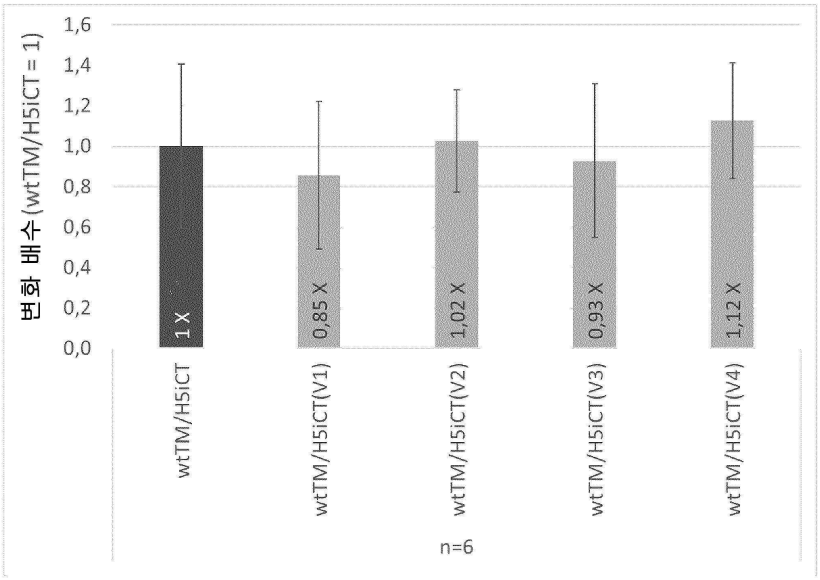
wtTM/H5iCT를 포함하는 변형된 SARS-CoV-2 S 단백질(V3)

…YEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCM**CSNGSLQCRICI**

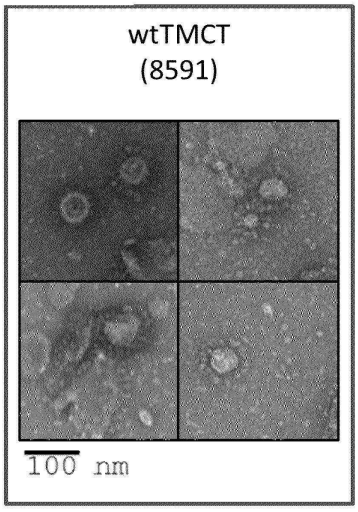
wtTM/H5iCT를 포함하는 변형된 SARS-CoV-2 S 단백질(V4)

…YEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC**SNGSLQCRIC**

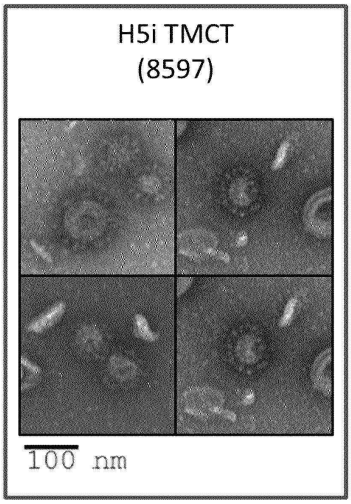
도면5b


변화 배수(wtTM/H5iCT = 1)
1,6
1,4
1,2
1,0
0,8
0,6
0,4
0,2
0,0
1 X
0,85 X
1,02 X
0,93 X
1,12 X
wtTM/H5iCT
wtTM/H5iCT(V1)
wtTM/H5iCT(V2)
wtTM/H5iCT(V3)
wtTM/H5iCT(V4)
n=6

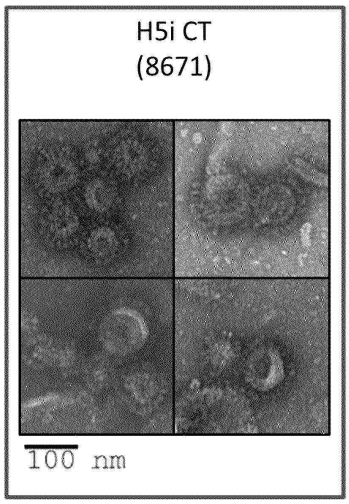
도면6a


wtTMCT
(8591)
100 nm

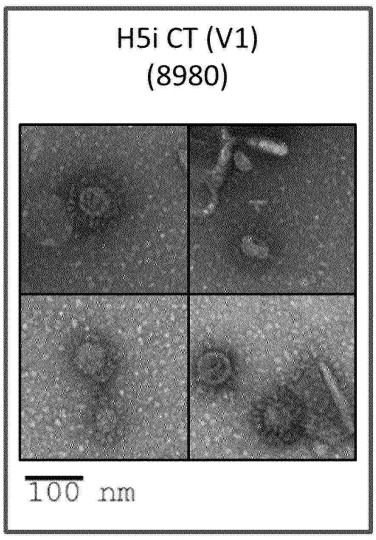
도면6b



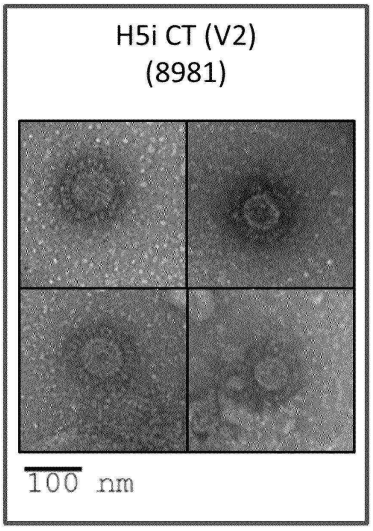
도면6c



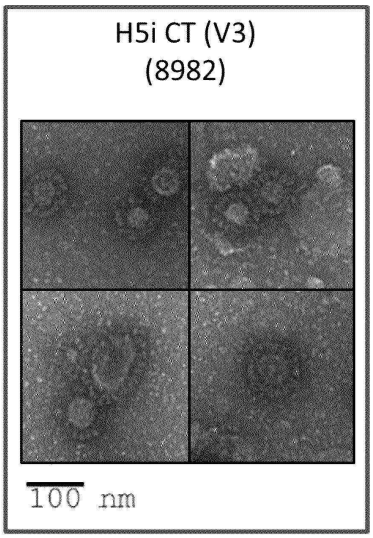
도면6d



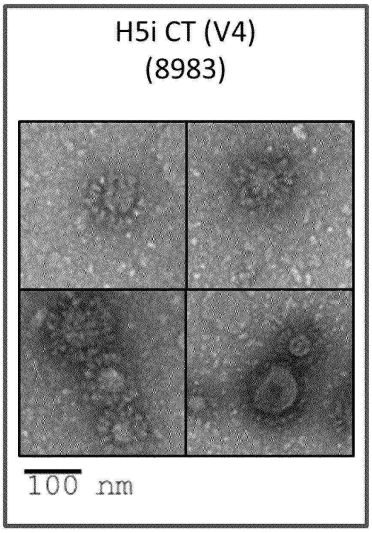
도면6e



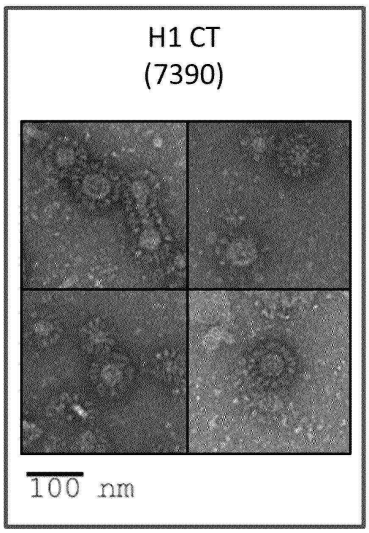
도면6f



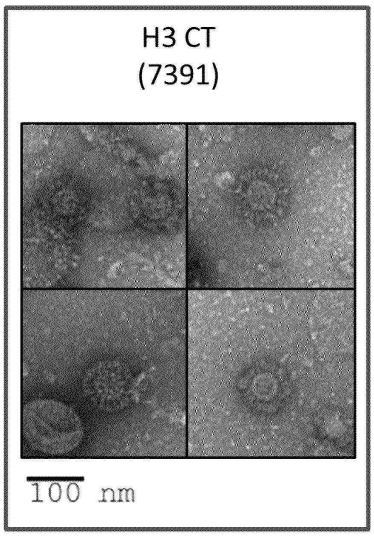
도면6g



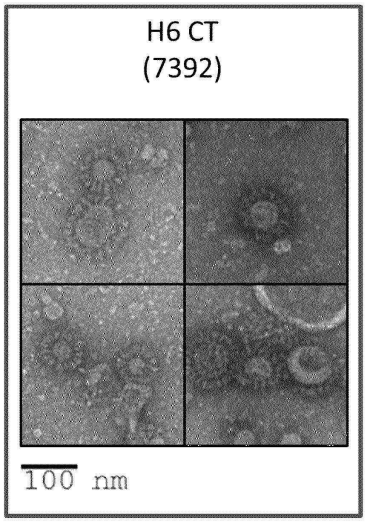
도면6h



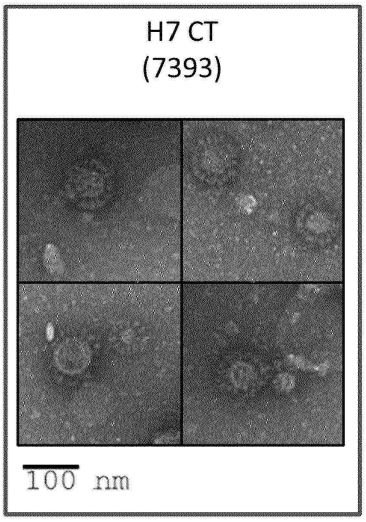
도면6i



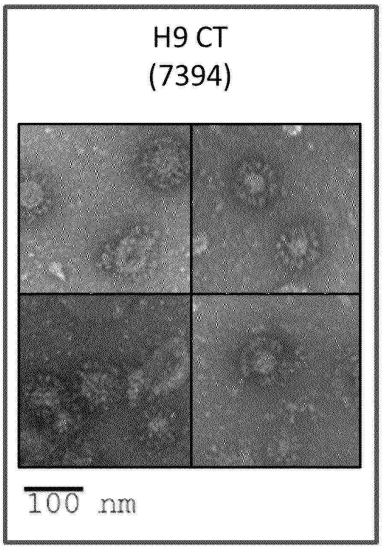
도면6j



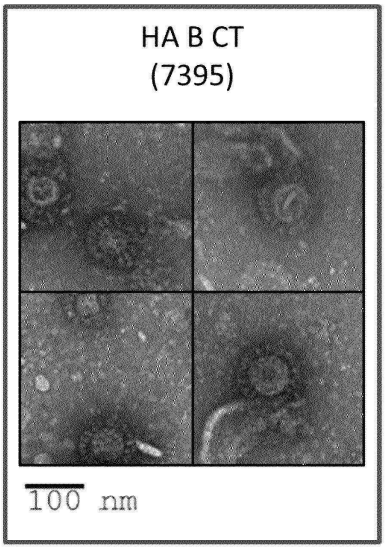
도면6k



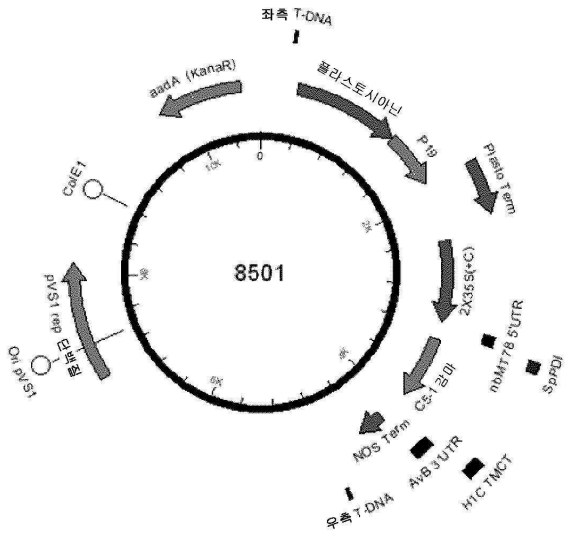
도면6l



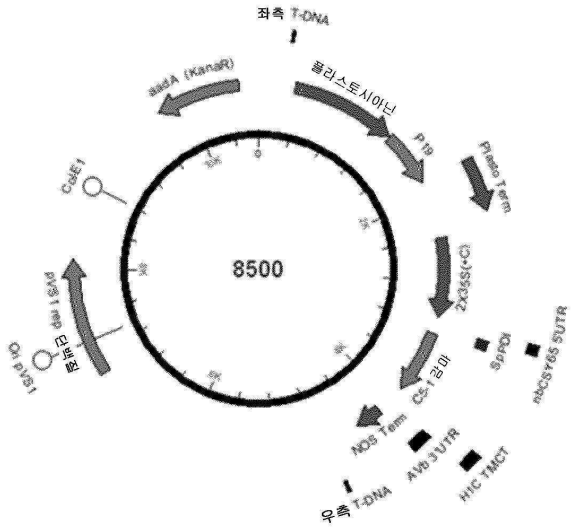
도면6m


HA B CT
(7395)
100 nm

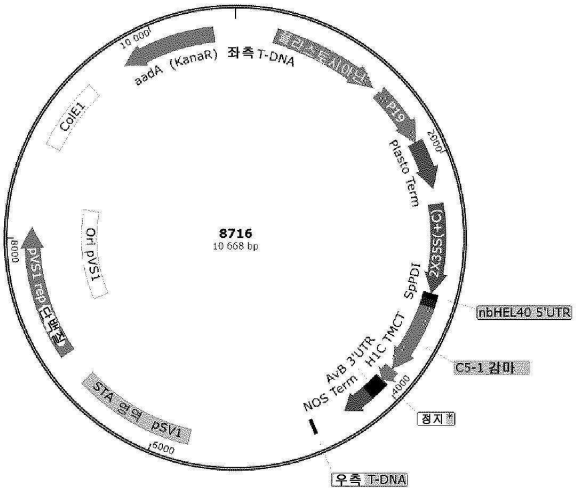
도면7a


좌측 T-DNA
aadA (KanaR)
플라스토시아닌
P19
Plasto Term
2X35S(+C)
CS-1 리더
nOM178 5'UTR
SpPDI
NOS Term
AvB 3'UTR
H1C TMCT
우측 T-DNA
ColE1
pVS1 rep 단백질
OriV pVS1
8501

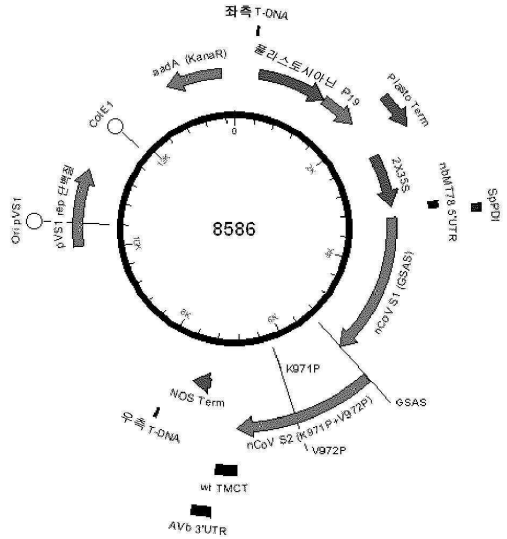
도면7b



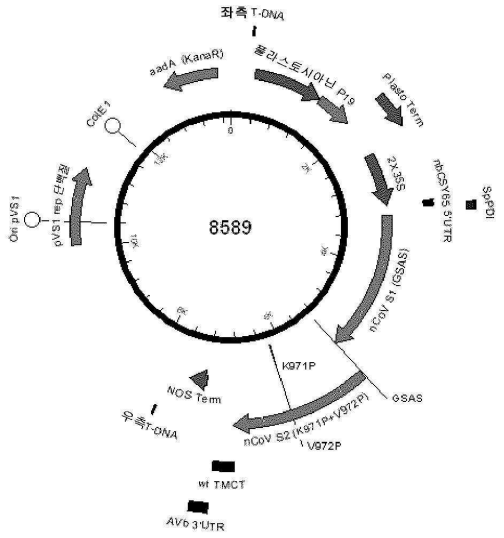
도면7c



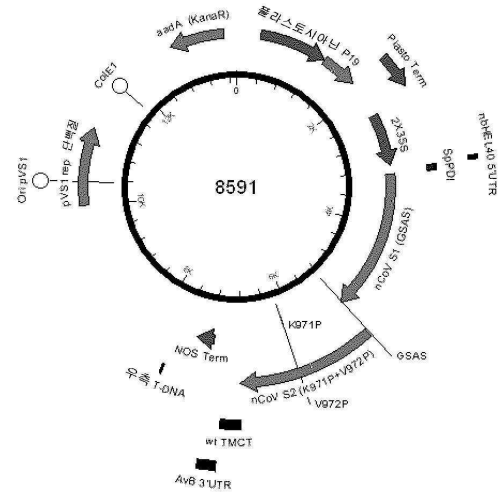
도면8a



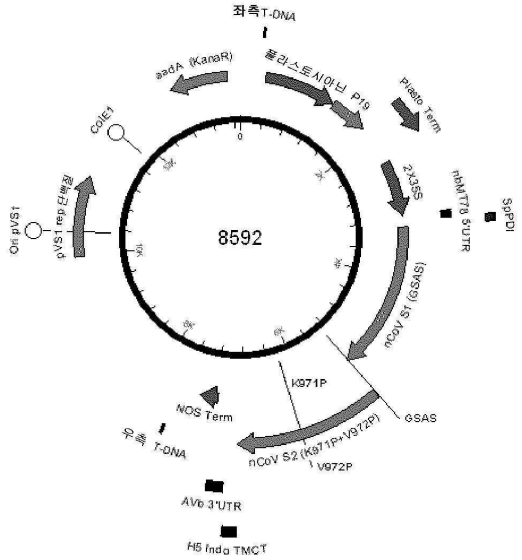
도면8b


좌측 T-DNA
aadA (KanaR)
폴리스토시아닌 P19
Plasto Term
2X 35S
ntGSV65 5'UTR
SpPDI
nCoV S1 (GSAS)
CdE1
Ori pVS1
pVS1 rep 단백질
8589
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)
V972P
wt TMCT
AVb 3'UTR

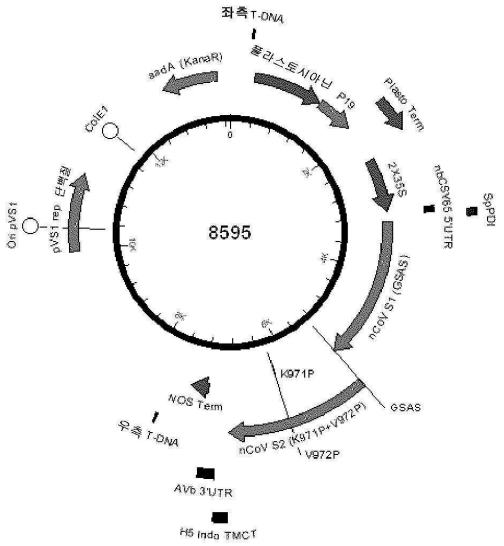
도면8c


aadA (KanaR)
폴리스토시아닌 P19
Plasto Term
2X 35S
nbHE40 5'UTR
SpPDI
nCoV S1 (GSAS)
CdE1
Ori pVS1
pVS1 rep 단백질
8591
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)
V972P
wt TMCT
Avb 3'UTR

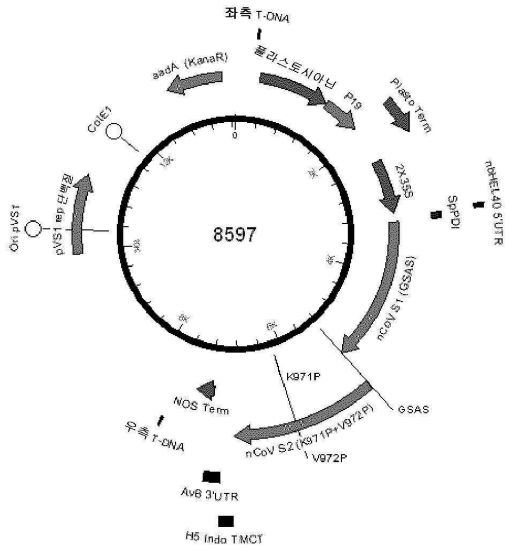
도면9a


좌측T-DNA
aadA (KanaR)
플라스토시아닌 P19
Pisato Term
CoIE1
2X35S
nbMT78 5'UTR
SpPDI
pVS1 rep 단백질
Ori pVS1
8592
nCoV S1 (GSAS)
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)
V972P
AVb 3'UTR
H5 hdo TMCT

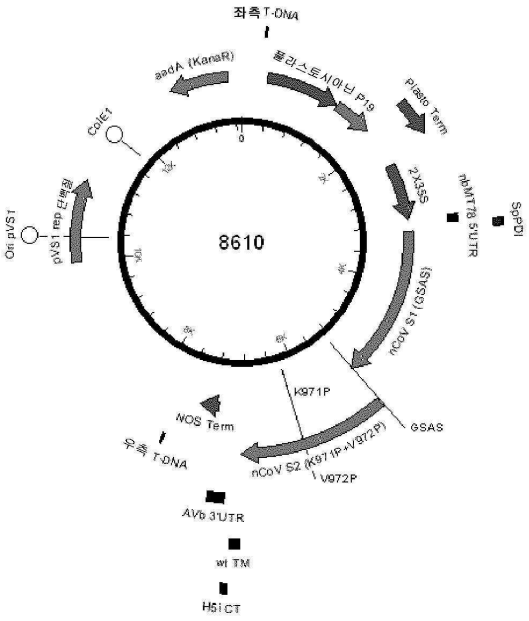
도면9b


좌측T-DNA
aadA (KanaR)
플라스토시아닌 P19
Pisato Term
CoIE1
2X35S
nbCSV65 5'UTR
SpPDI
pVS1 rep 단백질
Ori pVS1
8595
nCoV S1 (GSAS)
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)
V972P
AVb 3'UTR
H5 hdo TMCT

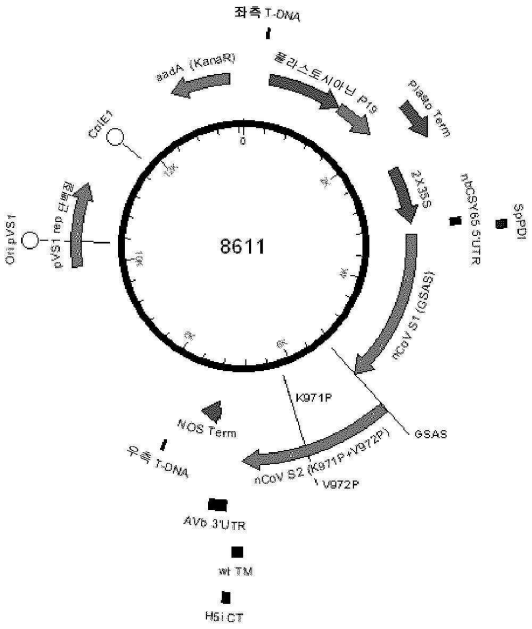
도면9c


좌측 T-DNA
aadA (KanaR)
폴리스토시아닌 P19
Pisato Term
2X35S
SpPDI
nbHEL40 5'UTR
nCoV S1 (GSAS)
GSAS
K971P
V972P
nCoV S2 (K971P+V972P)
NOS Term
우측 T-DNA
AvB 3'UTR
H5 Indo TMCT
ColE1
Ori pVS1
pVS1 rep 단백질
8597

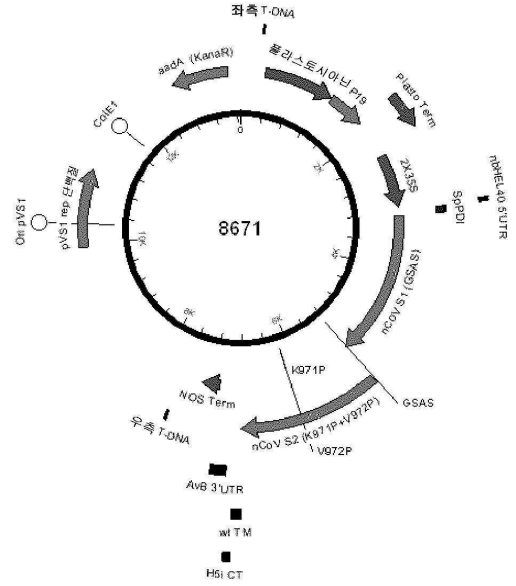
도면10a


좌측 T-DNA
aadA (KanaR)
폴리스토시아닌 P19
Pisato Term
2X35S
SoPDI
nbMT78 5'UTR
nCoV S1 (GSAS)
GSAS
K971P
V972P
nCoV S2 (K971P+V972P)
NOS Term
우측 T-DNA
AVb 3'UTR
wt TM
H5 CT
ColE1
Ori pVS1
pVS1 rep 단백질
8610

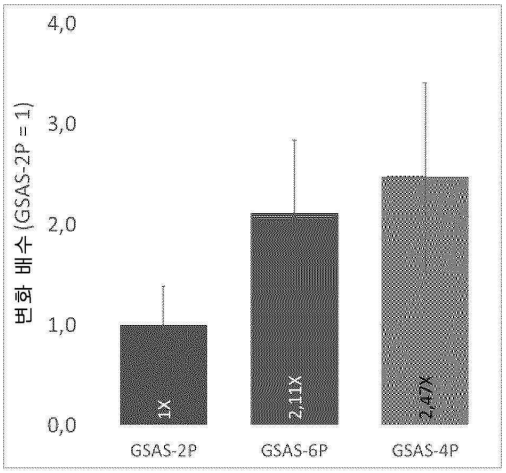
도면10b


좌측 T-DNA
aadA (KanaR)
폴리스토시아닌 P19
Pigato Term
ColE1
2X35S
nbCSV65 5'UTR
SoPDI
Ori pVS1
pVS1 rep 단백질
8611
nCoV S1 (GSAS)
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)
V972P
AVb 3'UTR
wt TM
H5i CT

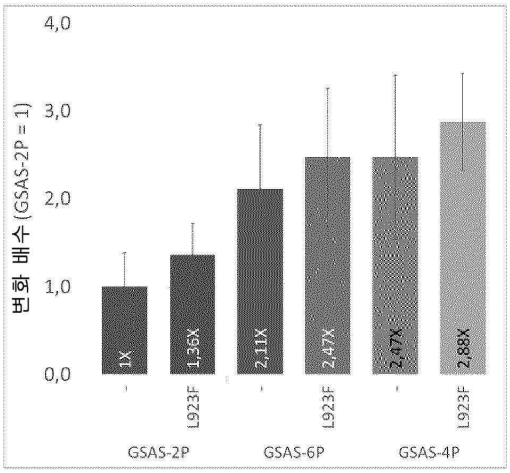
도면10c


좌측 T-DNA
aadA (KanaR)
폴리스토시아닌 P19
Pigato Term
ColE1
2X35S
nbHEL40 5'UTR
SoPDI
Ori pVS1
pVS1 rep 단백질
8671
nCoV S1 (GSAS)
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)
V972P
AVb 3'UTR
wt TM
H5i CT

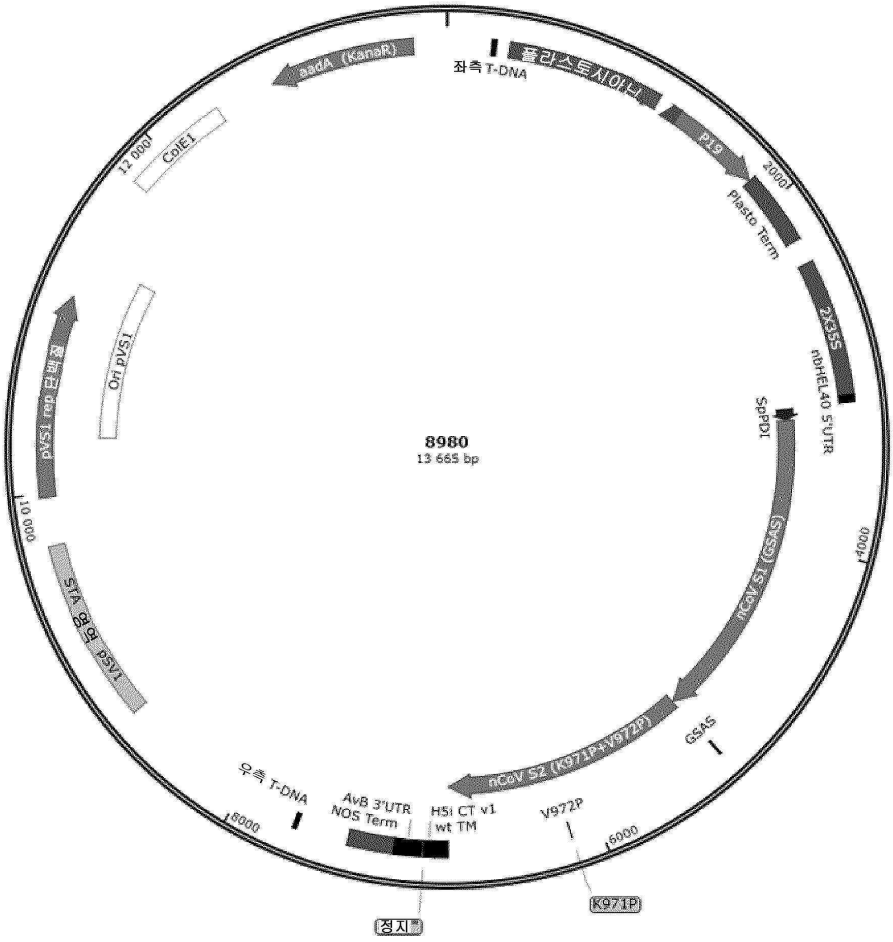
도면11a


4,0
3,0
2,0
1,0
0,0
변화 배수 (GSAS-2P = 1)
1X
2,11X
2,47X
GSAS-2P
GSAS-6P
GSAS-4P

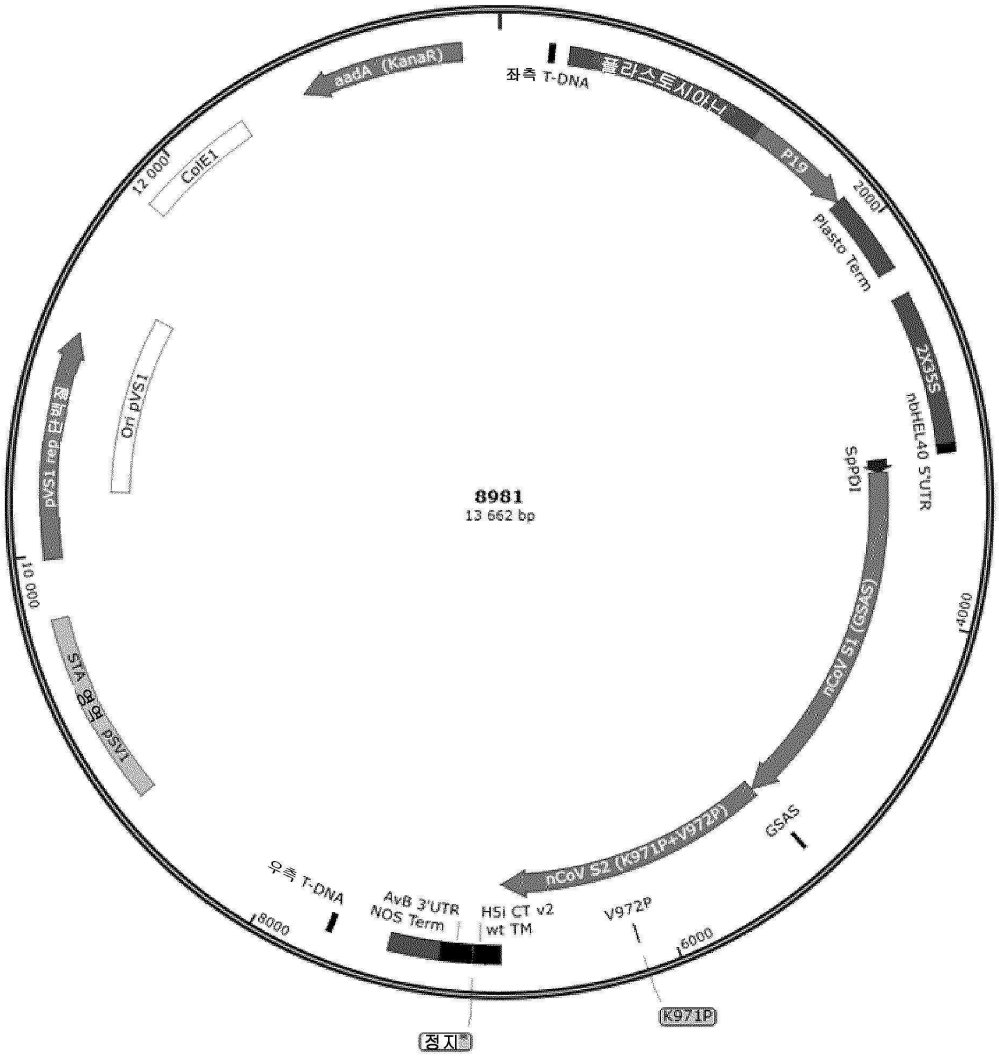
도면11b


4,0
3,0
2,0
1,0
0,0
변화 배수 (GSAS-2P = 1)
1X
1,36X
2,11X
2,47X
2,47X
2,88X
-
L923F
-
L923F
-
L923F
GSAS-2P
GSAS-6P
GSAS-4P

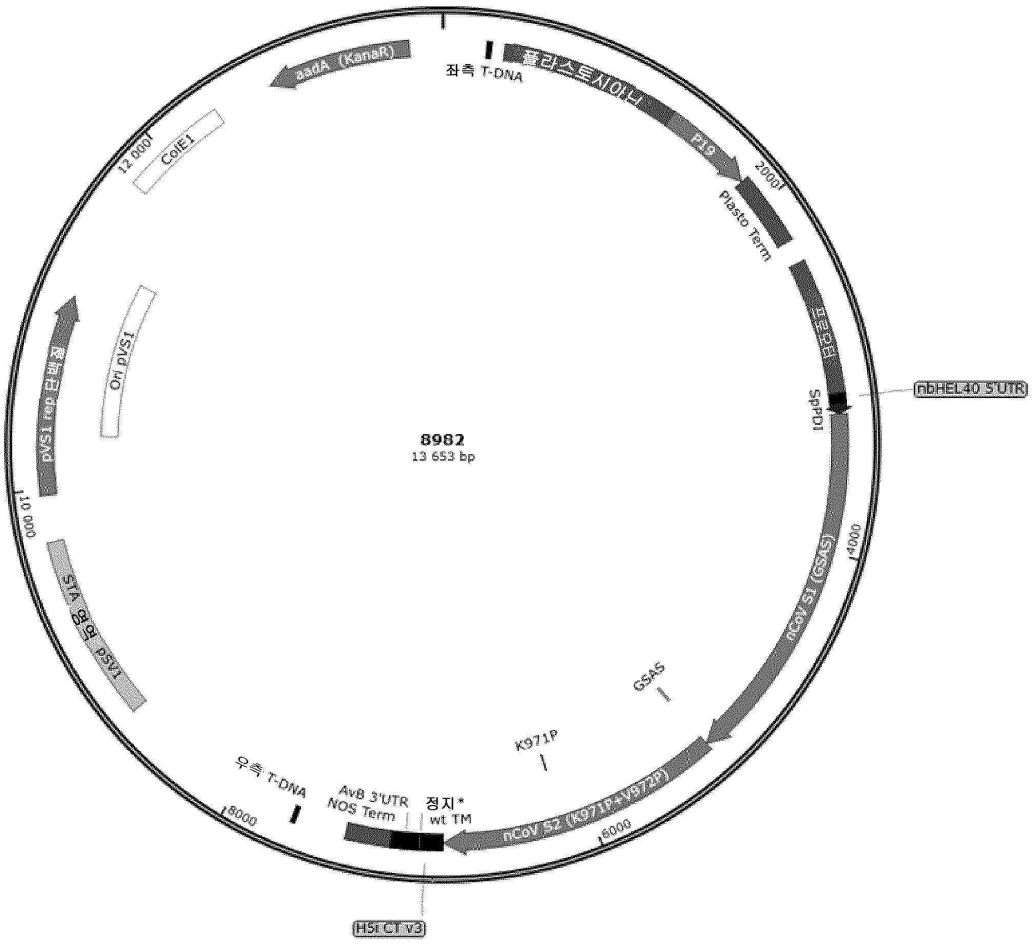
도면12a


8980
13 665 bp
aadA (KanaR)
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
2X35S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
GSAS
nCoV S2 (K971P+V972P)
V972P
K971P
H5i CT v1
wt TM
AvB 3'UTR
NOS Term
우측 T-DNA
SfA 융합 pSV1
pVS1 rep 단백질
Ori pVS1
ColE1
2000
4000
6000
8000
10 000
12 000
정지

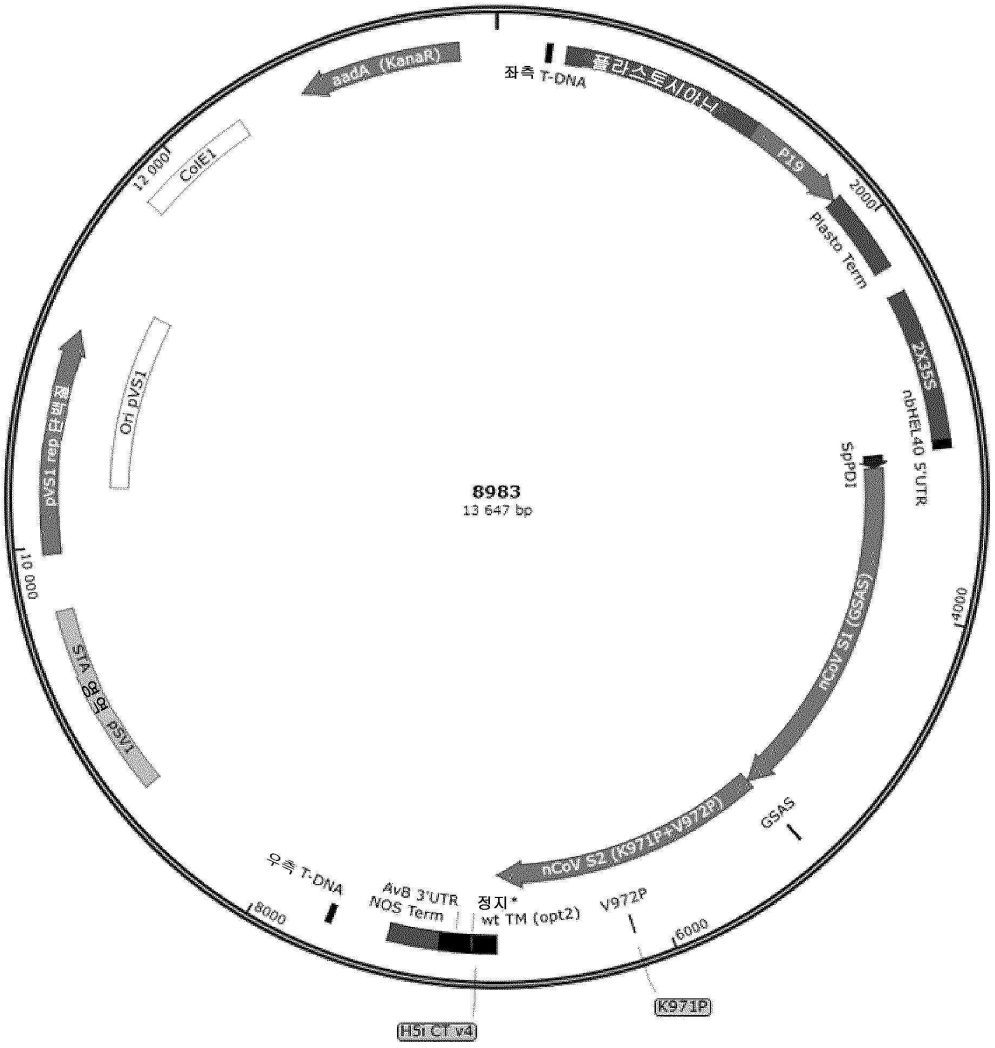
도면12b


8981
13 662 bp
aadA (KanaR)
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
2X35S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
GSAS
nCoV S2 (K971P+V972P)
우측 T-DNA
AvB 3'UTR
NOS Term
H5i CT v2
wt TM
V972P
K971P
정지
S1A 영역 pSV1
pVS1 rep 단백질
Ori pVS1
ColE1
2000
4000
6000
8000
10 000
12 000

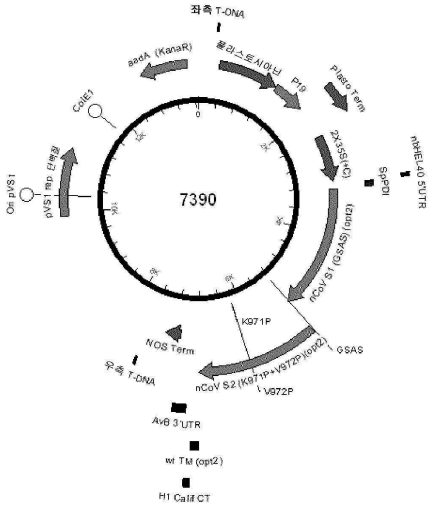
도면12c


8982
13 653 bp
aadA (KanaR)
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
SpPDI
nbHEL40 5'UTR
nCoV S1 (GSAS)
GSAS
K971P
nCoV S2 (K971P+V972P)
AvB 3'UTR
NOS Term
정지*
wt TM
우측 T-DNA
H5t CT v3
STA 영역 pSV1
pVS1 rep 단백질
Ori pVS1
CoiE1
2000
4000
6000
8000
10 000
12 000

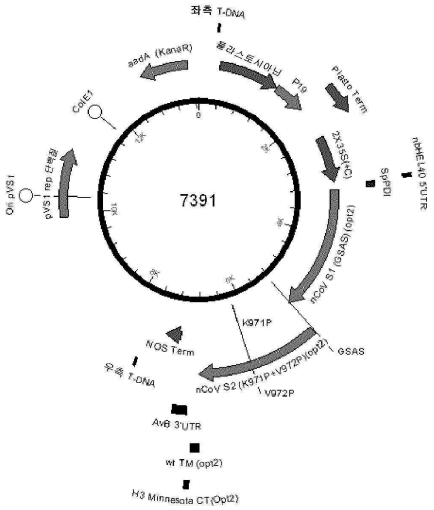
도면12d


badA (KanaR)
좌측 T-DNA
폴리스토시아닌
P19
Plasto Term
2X35S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
GSAS
nCoV S2 (K971P+V972P)
V972P
K971P
H5i CT v4
우측 T-DNA
Avβ 3'UTR
NOS Term
정지*
wt TM (opt2)
ColE1
Ori pVS1
pVS1 rep 단백질
STA 융합 pSV1
8983
13 647 bp
2000
4000
6000
8000
10 000
12 000

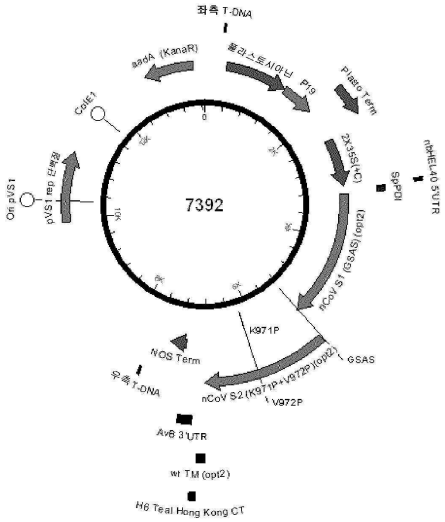
도면13a


좌측 T-DNA
badA (KanaR)
폴리스토시아닌
P19
Plasto Term
2X35S(+C)
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS) (opt2)
GSAS
K971P
nCoV S2 (K971P+V972P)(opt2)
V972P
NOS Term
우측 T-DNA
Avβ 3'UTR
wt TM (opt2)
H1 Calif CT
ColE1
Ori pVS1
pVS1 rep 단백질
7390

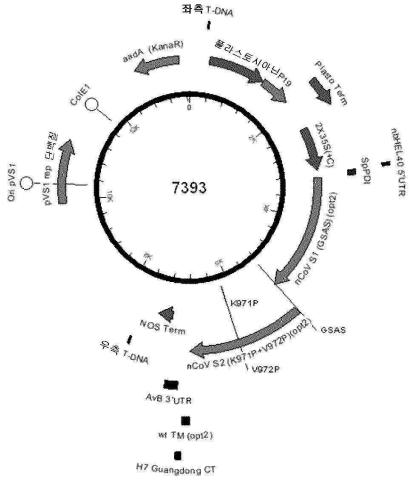
도면13b



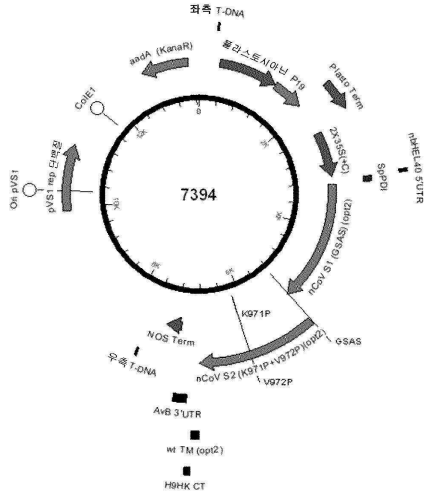
도면13c



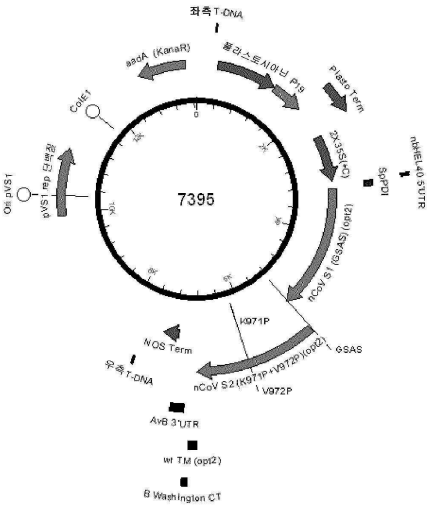
도면13d



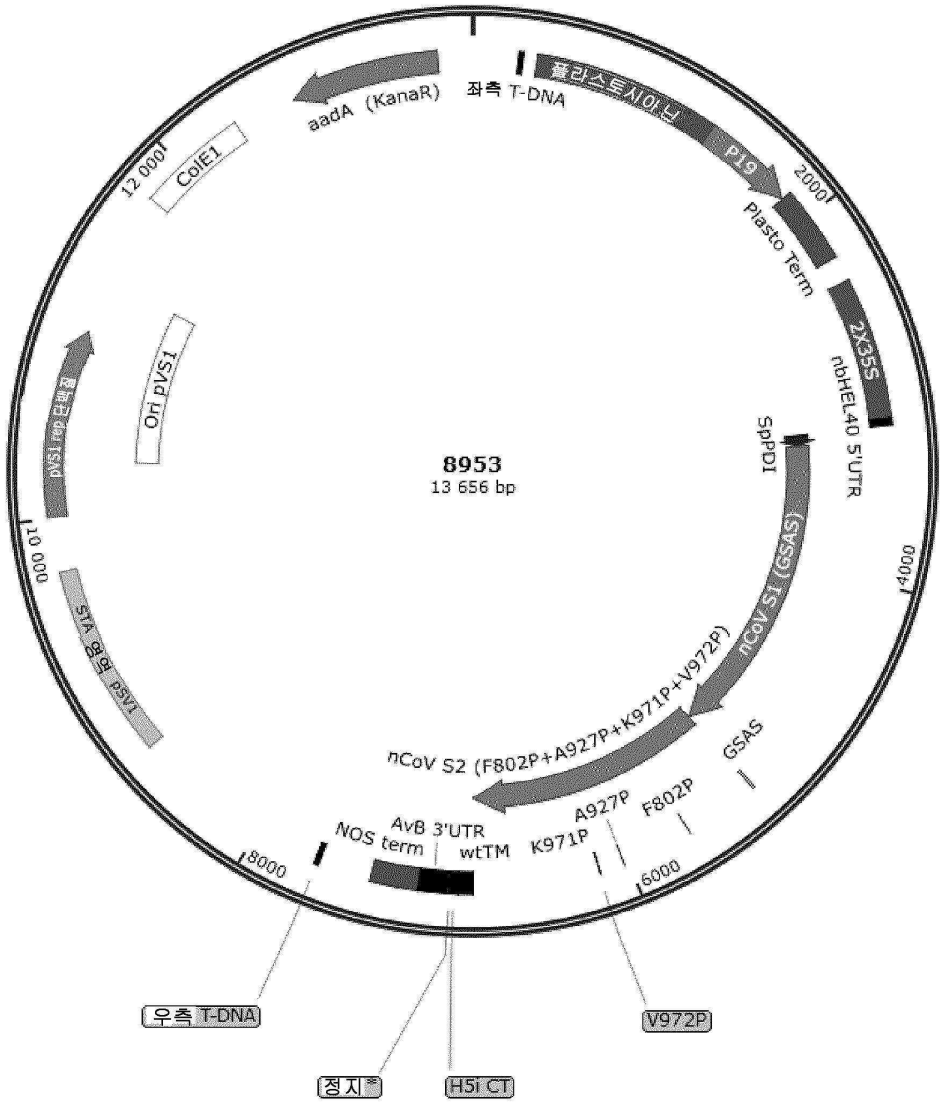
도면13e


좌측 T-DNA
aadA (KanaR)
불리스토시아누 P19
Plasto Term
2x35S(+Ω)
SoPDI
nbHEL40 5'UTR
nCoV S1 (GSAS)(opt2)
ColE1
Ole pVS1
pVS1 rep 단백질
7394
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)(opt2)
V972P
AvB 3'UTR
wt TM (opt2)
H9HK CT

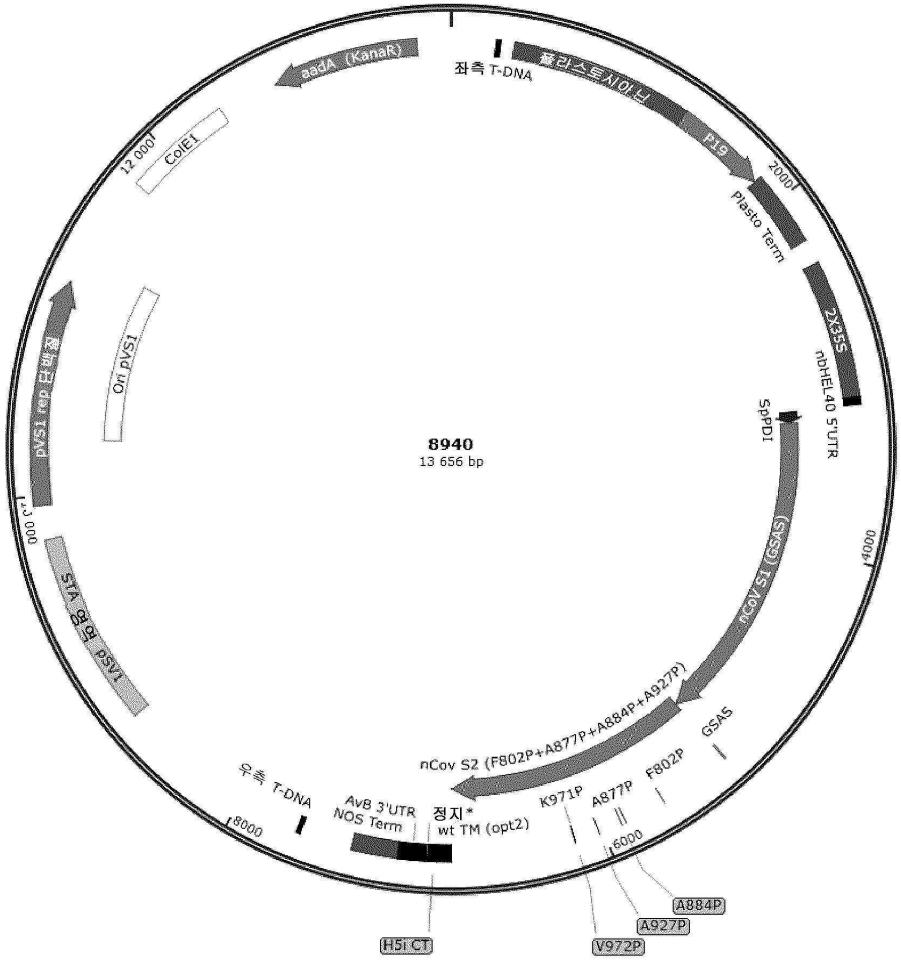
도면13f


좌측 T-DNA
aadA (KanaR)
불리스토시아누 P19
Plasto Term
2x35S(+Ω)
SoPDI
nbHEL40 5'UTR
nCoV S1 (GSAS)(opt2)
ColE1
Ole pVS1
pVS1 rep 단백질
7395
K971P
GSAS
NOS Term
우측 T-DNA
nCoV S2 (K971P+V972P)(opt2)
V972P
AvB 3'UTR
wt TM (opt2)
B Washington CT

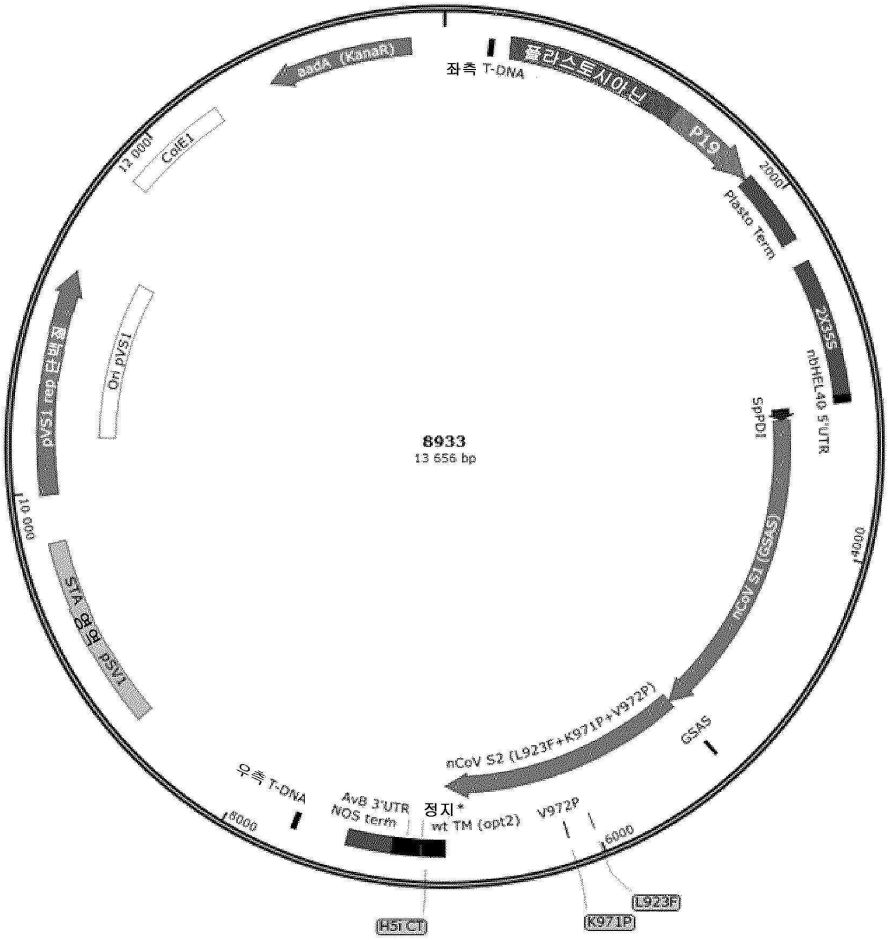
도면14a


8953
13 656 bp
aadA (KanaR)
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
2X35S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
nCoV S2 (F802P+A927P+K971P+V972P)
GSAS
F802P
A927P
K971P
AvB 3'UTR
wtTM
NOS term
S1A 융합 pSV1
pVS1 rep단백질
Ori pVS1
ColE1
2000
4000
6000
8000
10 000
12 000
우측 T-DNA
정지
H5I CT
V972P

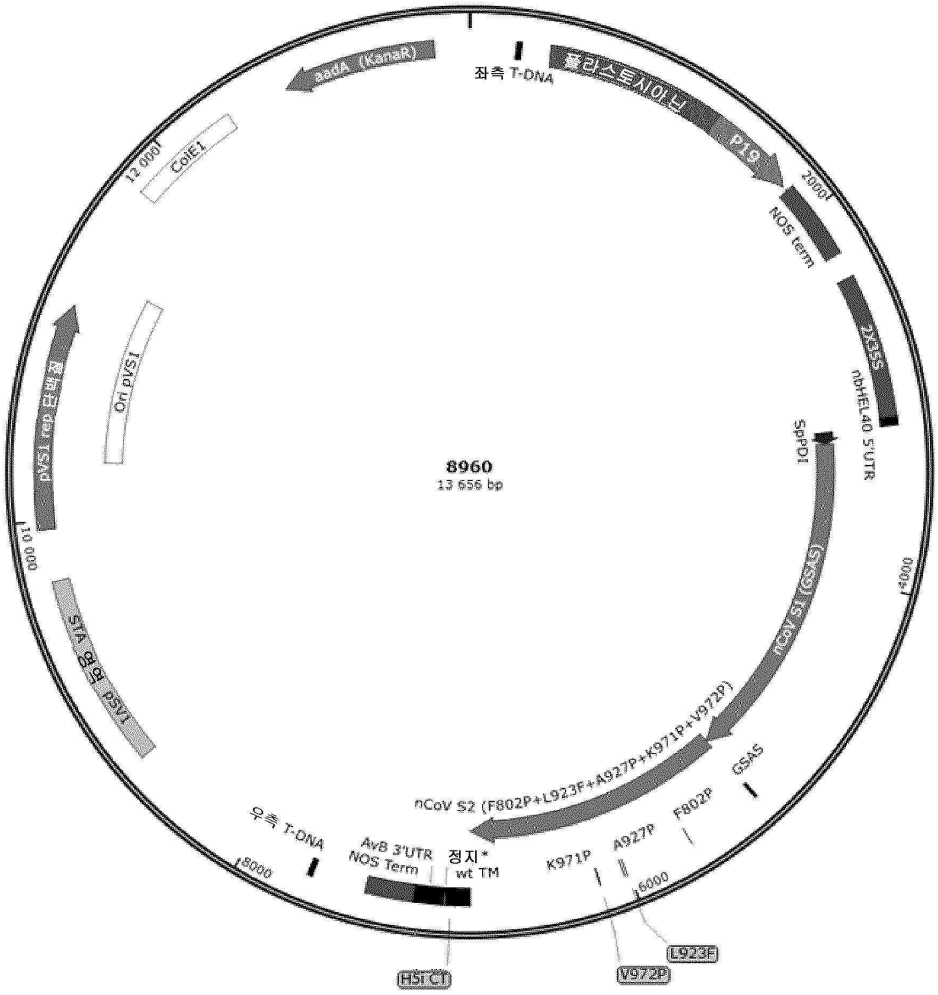
도면14b


8940
13 656 bp
좌측 T-DNA
aadA (KanaR)
ColE1
12 000
Ori pVS1
pVS1 rep 단백질
pVS1 STA 영역
8000
우측 T-DNA
Avb 3'UTR
NOS Term
정지*
wt TM (opt2)
H5i CT
nCov S2 (F802P+A877P+A884P+A927P)
K971P
A877P
F802P
GSAS
A884P
A927P
V972P
6000
nCov S1 (GSAS)
SpPDI
nbHEL40 5'UTR
2X35S
Plasto Term
P19
4000

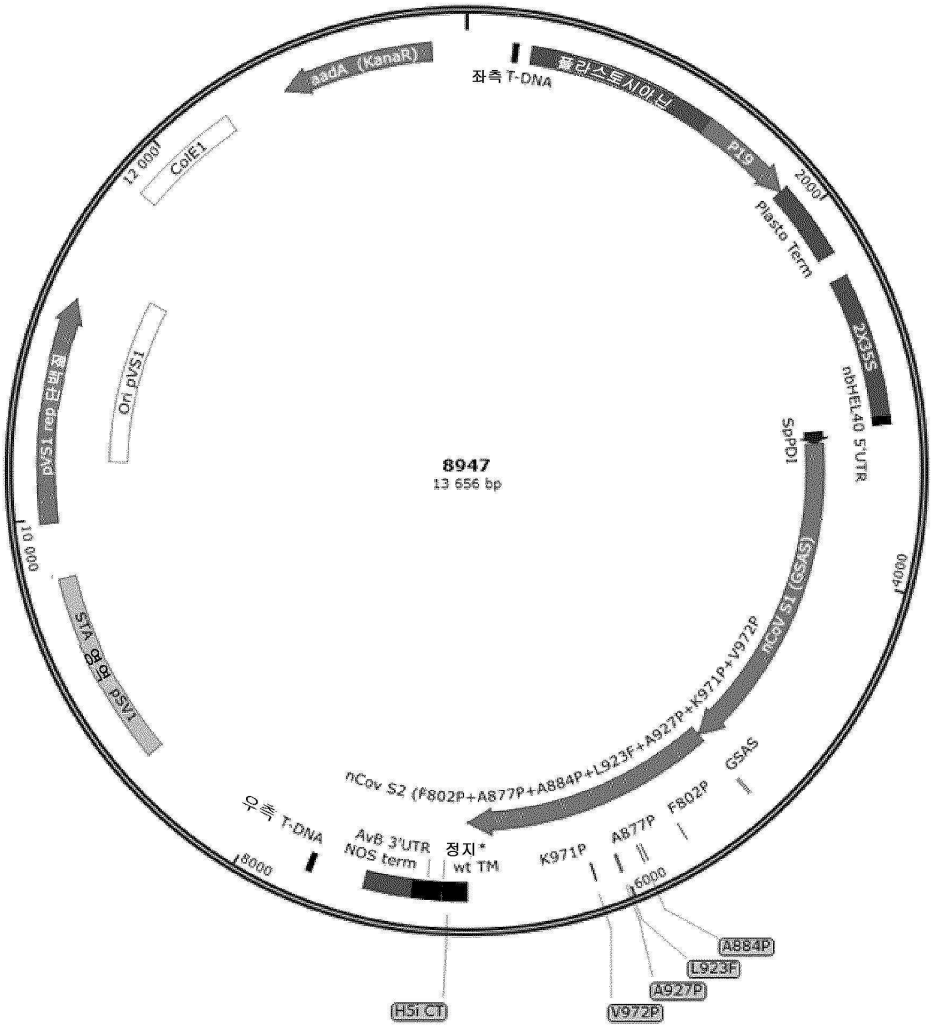
도면15a


8933
13 656 bp
좌측 T-DNA
플라스틱시아닌
P19
Plasto Term
2X35S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
GSAS
nCoV S2 (L923F+K971P+V972P)
V972P
L923F
K971P
정지*
wt TM (opt2)
AvB 3'UTR
NOS term
H5t CT
우측 T-DNA
STA 융합 pSV1
융합 pSV1
Ori pVS1
ColE1
aadA (KanaR)
2000
4000
6000
8000
10 000
12 000

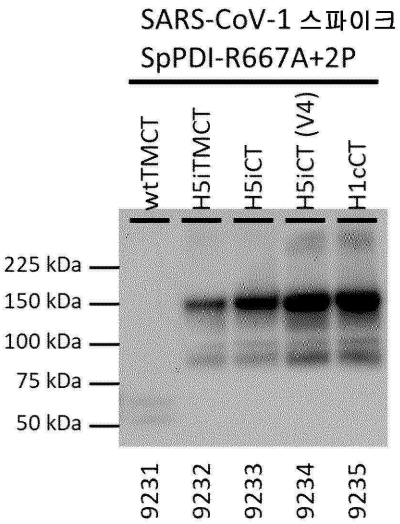
도면15b


8960
13 656 bp
aadA (KSmaR)
좌측 T-DNA
폴리스틱시아닌
P19
2000
NOS term
2xCaMV 35S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
GSAS
F802P
A927P
K971P
V972P
L923F
nCoV S2 (F802P+L923F+A927P+K971P+V972P)
우측 T-DNA
AvB 3'UTR
NOS Term
정지*
wt TM
H5t CT
pVS1 rep 단백질
On pVS1
CoiE1
STA 융합 pSV1
12 000
10 000
8000
6000
4000

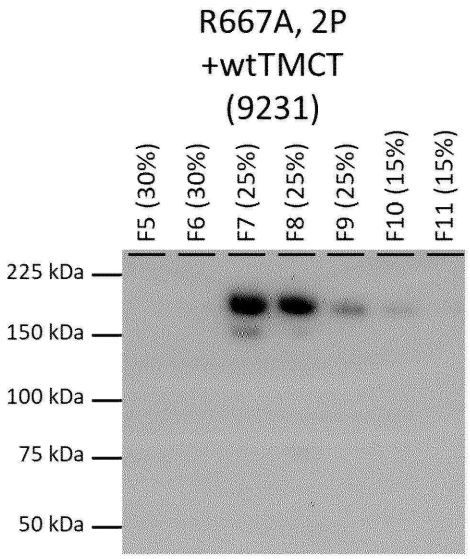
도면15c


8947
13 656 bp
aadA (KanaR)
좌측 T-DNA
퓨리스토시이닌
P19
Plasto Term
2000
2435S
nbHEL40 5'UTR
SpPDI
nCoV S1 (GSAS)
4000
GSAS
F802P
A877P
K971P
nCov S2 (F802P+A877P+A884P+L923F+A927P+K971P+V972P
AvB 3'UTR
NOS term
정지*
wt TM
우측 T-DNA
8000
6000
10 000
STA 융합 pSV1
pVS1 rep 단백질
Ori pVS1
CoIE1
12 000
A884P
L923F
A927P
V972P
H5i CT

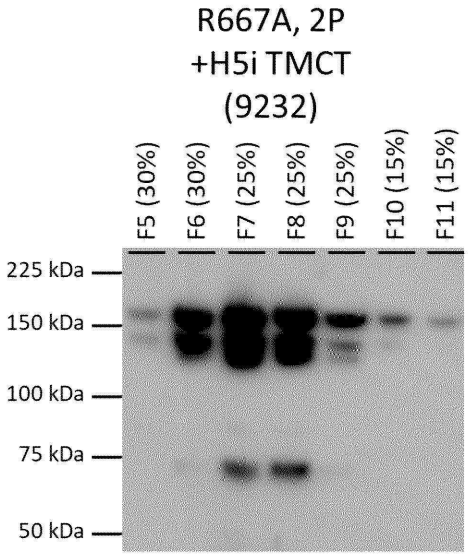
도면16a


SARS-CoV-1 스파이크
SpPDI-R667A+2P
wtTMCT
H5iTMCT
H5iCT
H5iCT (V4)
H1cCT
225 kDa
150 kDa
100 kDa
75 kDa
50 kDa
9231
9232
9233
9234
9235

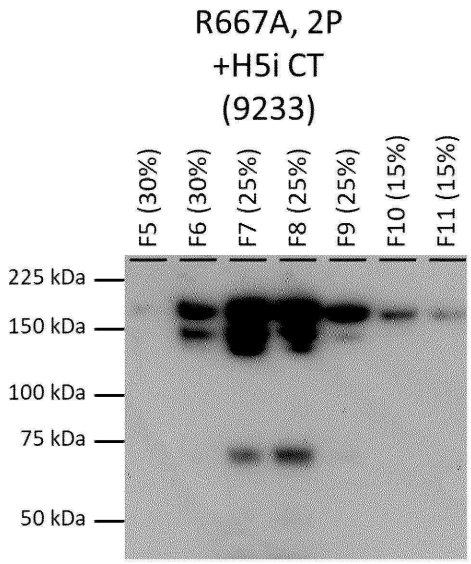
도면16b


R667A, 2P
+wtTMCT
(9231)
F5 (30%)
F6 (30%)
F7 (25%)
F8 (25%)
F9 (25%)
F10 (15%)
F11 (15%)
225 kDa
150 kDa
100 kDa
75 kDa
50 kDa

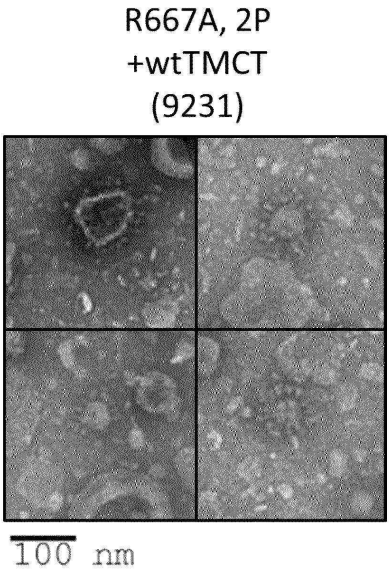
도면16c


R667A, 2P
+H5i TMCT
(9232)
F5 (30%)
F6 (30%)
F7 (25%)
F8 (25%)
F9 (25%)
F10 (15%)
F11 (15%)
225 kDa
150 kDa
100 kDa
75 kDa
50 kDa

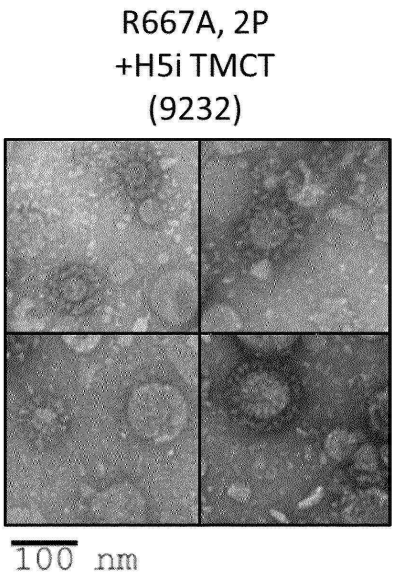
도면16d


R667A, 2P
+H5i CT
(9233)
F5 (30%)
F6 (30%)
F7 (25%)
F8 (25%)
F9 (25%)
F10 (15%)
F11 (15%)
225 kDa
150 kDa
100 kDa
75 kDa
50 kDa

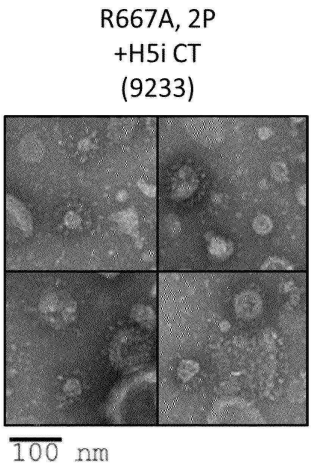
도면17a


R667A, 2P
+wtTMCT
(9231)
100 nm

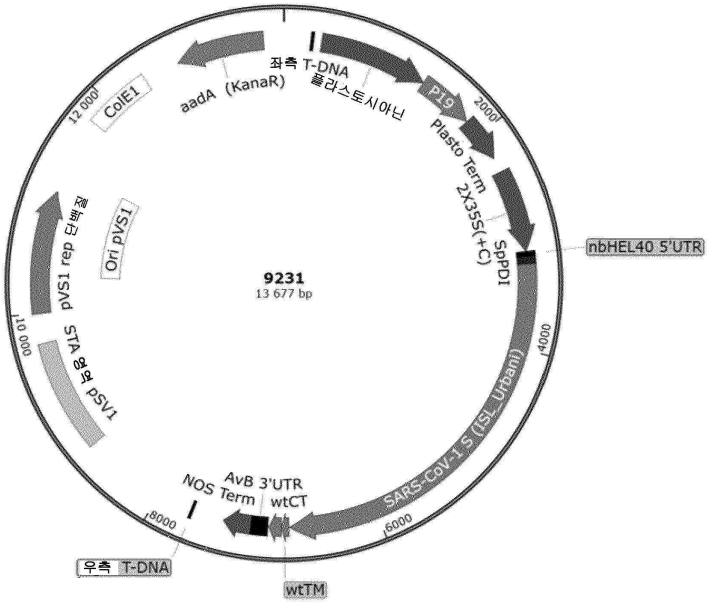
도면17b



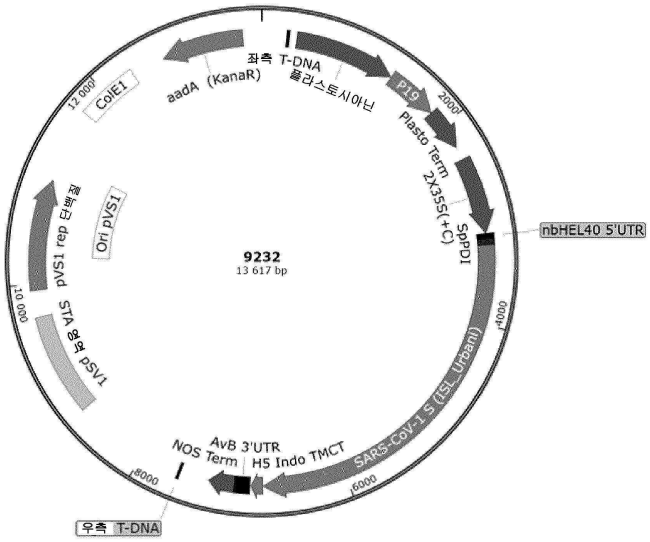
도면17c



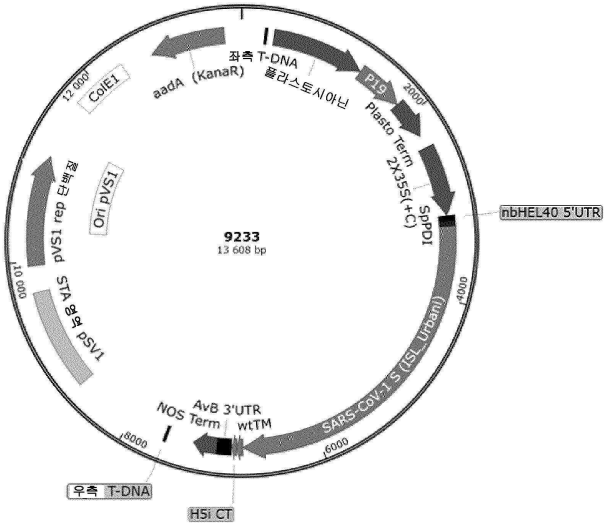
도면18a


9231
13 677 bp
좌측 T-DNA
aadA (KanaR)
ColE1
12 000
2000
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
4000
SARS-CoV-1 S (ISL_Urbani)
6000
AvB 3'UTR
wtCT
NOS Term
8000
우측 T-DNA
wtTM
10 000
STA 영역 pSV1
pVS1 rep 단백질
Ori pVS1

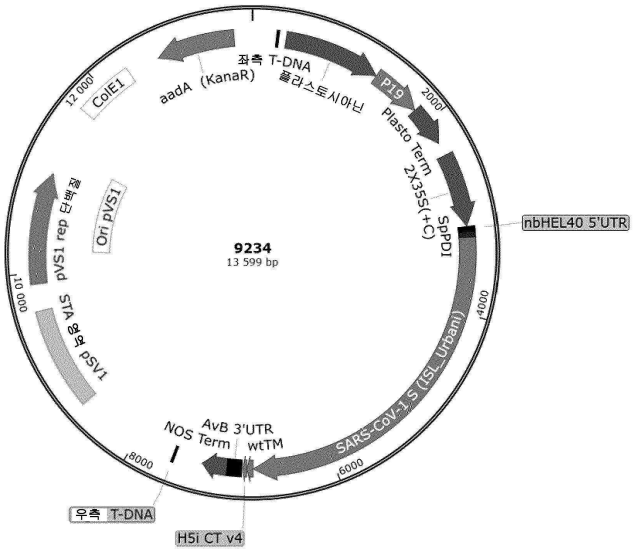
도면18b


9232
13 617 bp
좌측 T-DNA
aadA (KanaR)
ColE1
12 000
2000
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
4000
SARS-CoV-1 S (ISL_Urbani)
6000
AvB 3'UTR
H5 Indo TMCT
NOS Term
8000
우측 T-DNA
10 000
STA 영역 pSV1
pVS1 rep 단백질
Ori pVS1

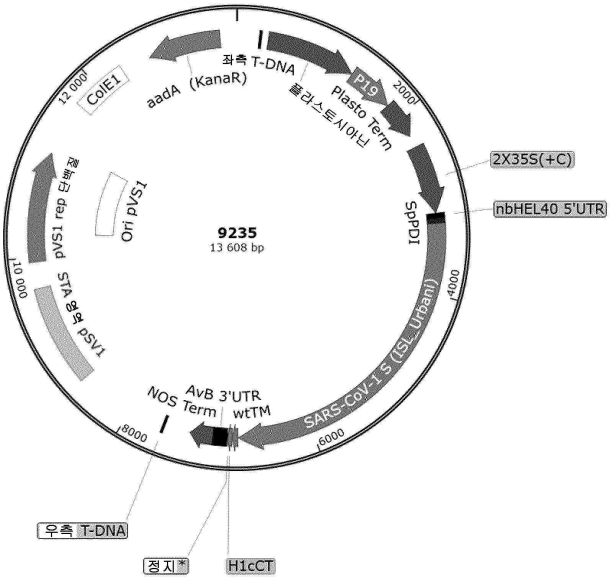
도면18c


9233
13 608 bp
좌측 T-DNA
aadA (KanaR)
플라스토시아닌
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
SARS-CoV-1 S (ISL_Urbani)
AvB 3'UTR
wtTM
NOS Term
우측 T-DNA
H5i CT
ColE1
Ori pVS1
pVS1 rep 단백질
STA 영역 pSV1
2000
4000
6000
8000
10 000
12 000

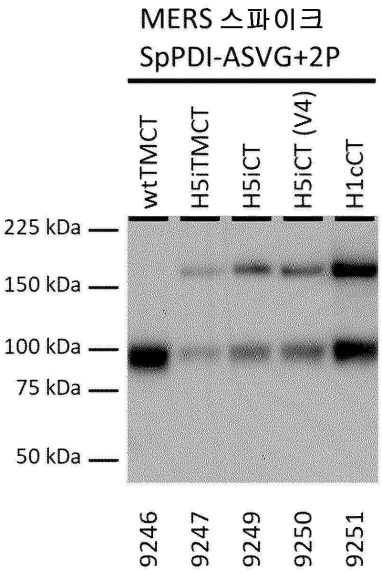
도면18d


9234
13 599 bp
좌측 T-DNA
aadA (KanaR)
플라스토시아닌
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
SARS-CoV-1 S (ISL_Urbani)
AvB 3'UTR
wtTM
NOS Term
우측 T-DNA
H5i CT v4
ColE1
Ori pVS1
pVS1 rep 단백질
STA 영역 pSV1
2000
4000
6000
8000
10 000
12 000

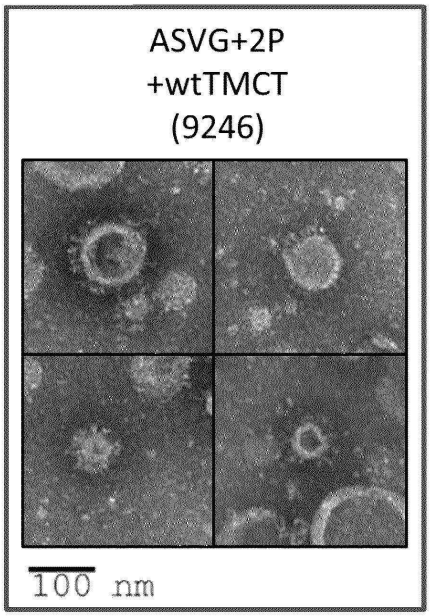
도면18e


9235
13 608 bp
좌측 T-DNA
aadA (KanaR)
ColE1
Ori pVS1
pVS1 rep 단백질
STA 영역 pSV1
NOS Term
AvB 3'UTR
wtTM
SARS-CoV-1 S (ISL_Urbani)
SpPDI
Plasto Term
P19
2X35S(+C)
nbHEL40 5'UTR
우측 T-DNA
정지*
H1cCT

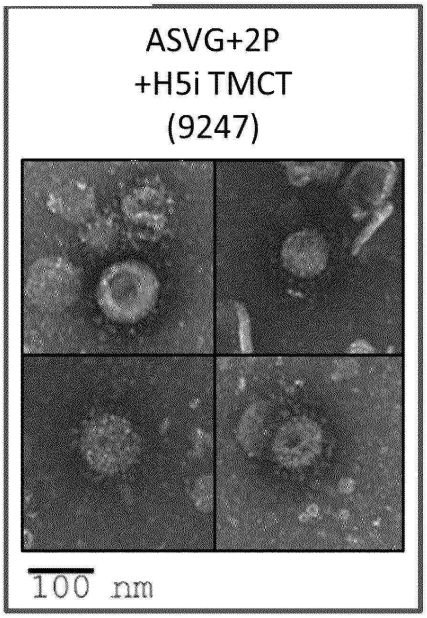
도면19a


MERS 스파이크
SpPDI-ASVG+2P
wtTMCT
H5iTMCT
H5iCT
H5iCT (V4)
H1cCT
225 kDa
150 kDa
100 kDa
75 kDa
50 kDa
9246
9247
9249
9250
9251

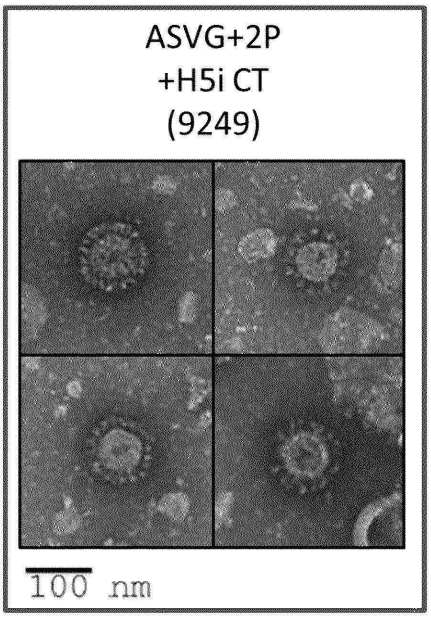
도면19b



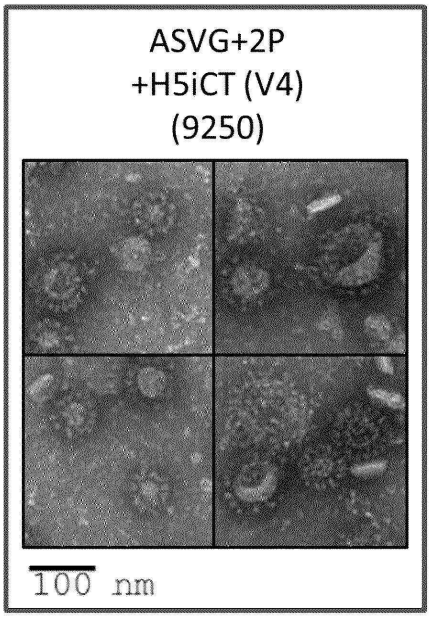
도면19c



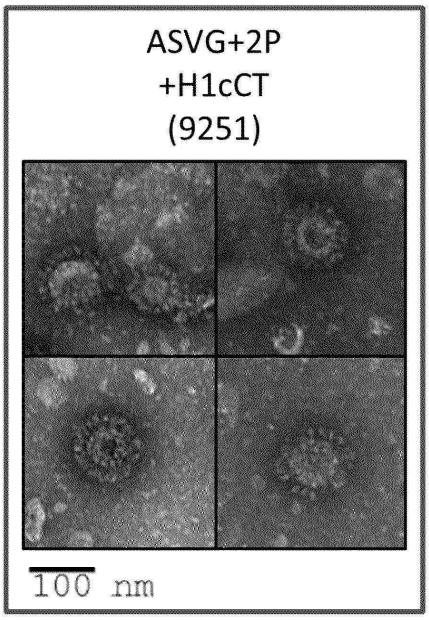
도면19d



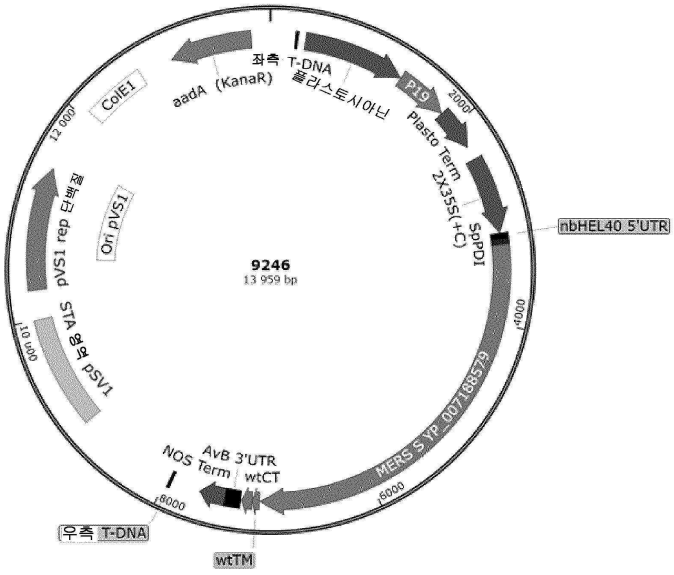
도면19e



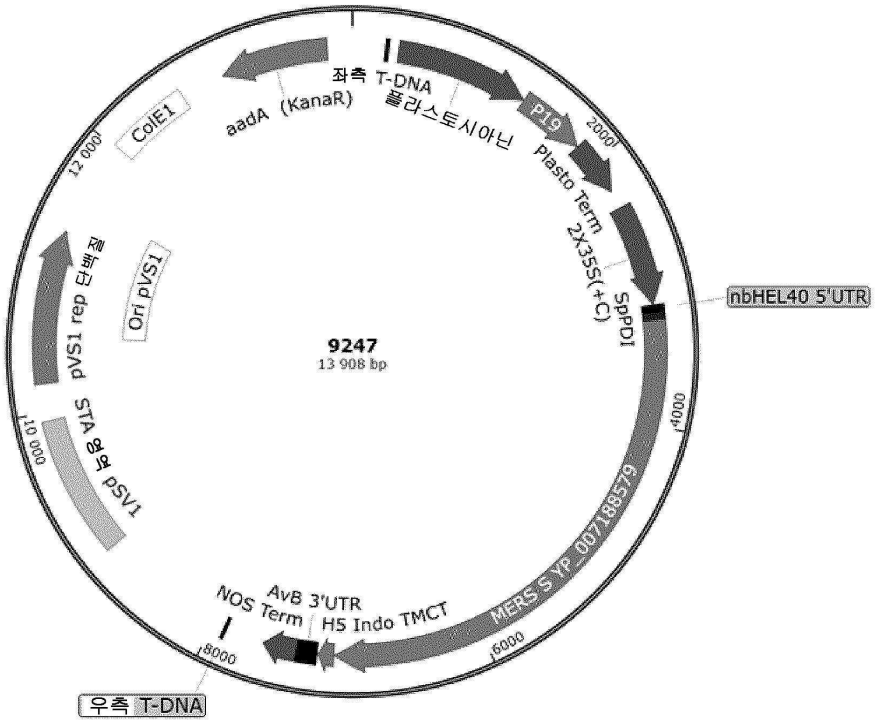
도면19f


ASVG+2P
+H1cCT
(9251)
100 nm

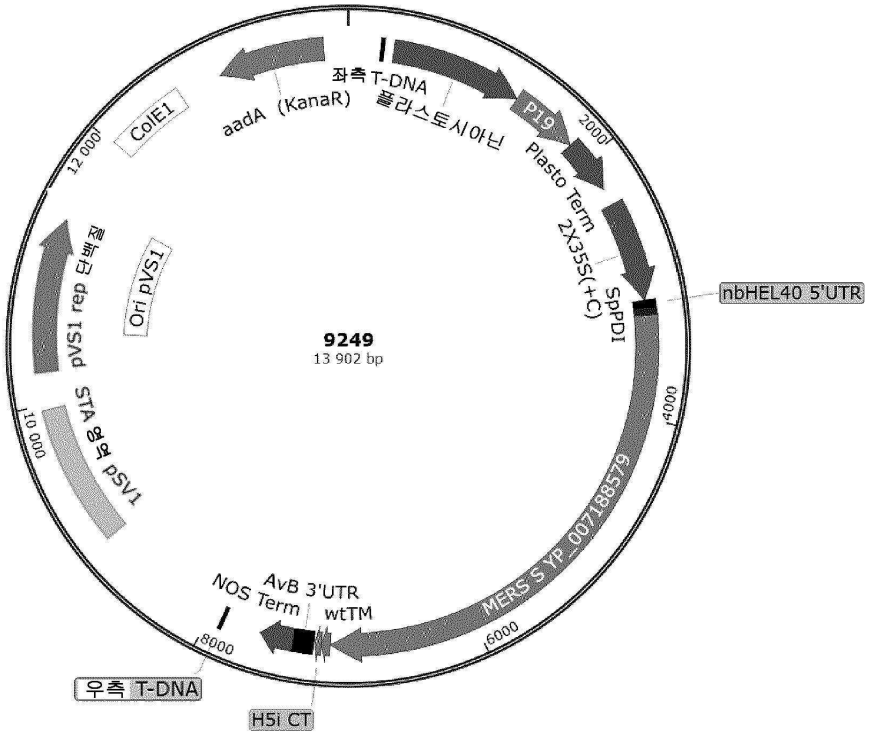
도면20a


9246
13 959 bp
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
2x35S(+C)
SpPDI
nbHEL40 5'UTR
MERS S YP_009047831.1
AvB 3'UTR
wtCT
NOS Term
우측 T-DNA
wtTM
STA 영역 pSV1
pVS1 rep 단백질
On pVS1
ColE1
aadA (KanaR)
2000
4000
6000
8000
10 000
12 000

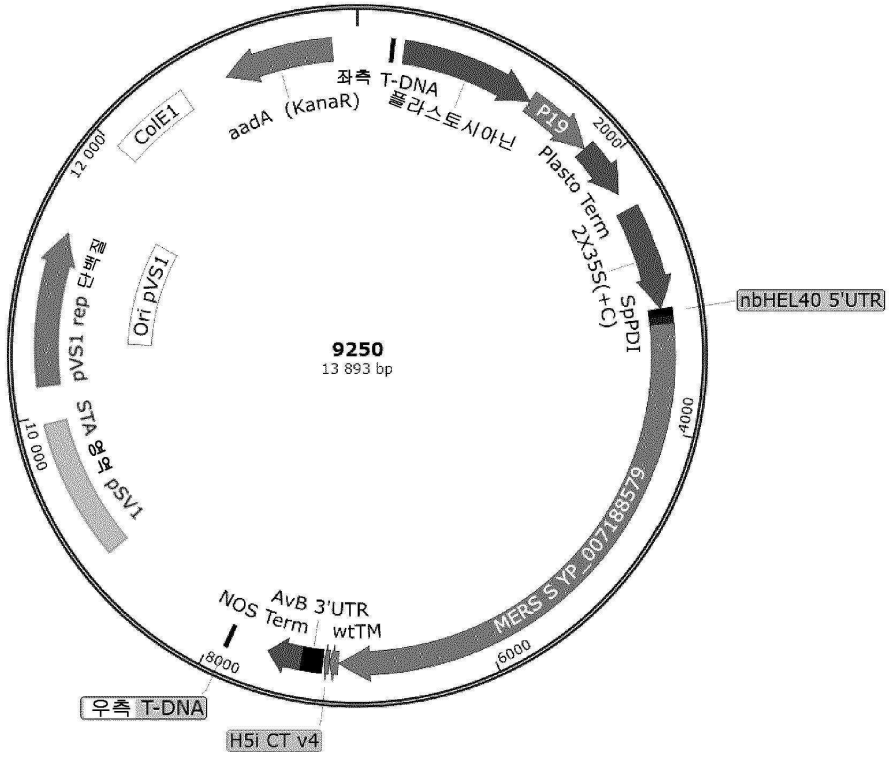
도면20b


9247
13 908 bp
좌측 T-DNA
aadA (KanaR)
플라스토시아닌
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
MERS S YP_007188579
H5 Indo TMCT
AvB 3'UTR
NOS Term
우측 T-DNA
STA 영역 pSV1
pVS1 rep 단백질
Ori pVS1
ColE1
2000
4000
6000
8000
10 000
12 000

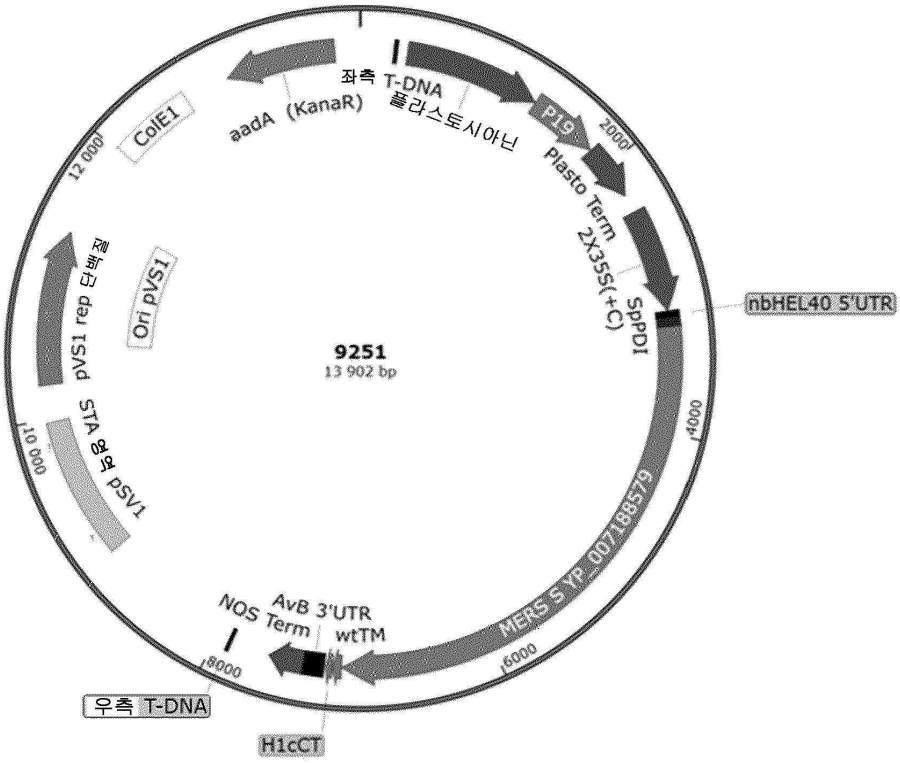
도면20c


9249
13 902 bp
좌측 T-DNA
aadA (KanaR)
플라스토시아닌
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
MERS S YP_007188579
AvB 3'UTR
wtTM
NOS Term
우측 T-DNA
H5i CT
STA 영역 pSV1
pVS1 rep 단백질
Ori pVS1
ColE1
2000
4000
6000
8000
10 000
12 000

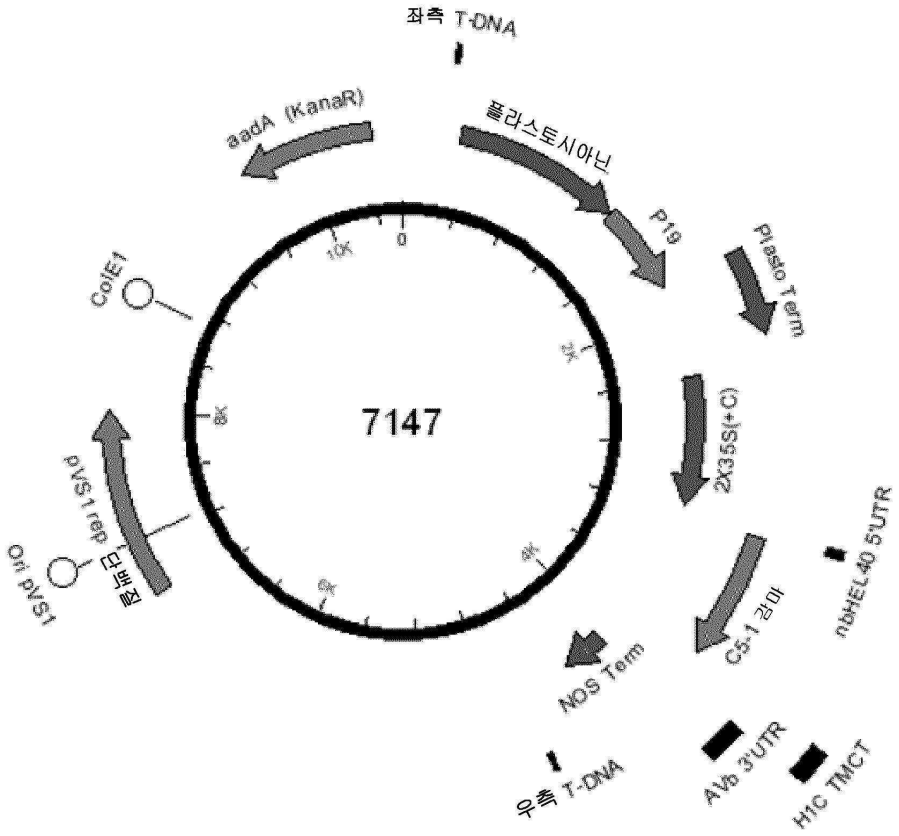
도면20d


9250
13 893 bp
좌측 T-DNA
aadA (KanaR)
ColE1
Ori pVS1
pVS1 rep 단백질
STA 융모 pSV1
NOS Term
AvB 3'UTR
wtTM
MERS S YP_007188579
SpPDI
2X35S(+C)
Plasto Term
P19
플라스토시아닌
nbHEL40 5'UTR
우측 T-DNA
H5i CT v4
2000
4000
6000
8000
10 000
12 000

도면20e


9251
13 902 bp
좌측 T-DNA
aadA (KanaR)
ColE1
Ori pVS1
pVS1 rep 단백질
STA 융모 pSV1
NOS Term
AvB 3'UTR
wtTM
MERS S YP_007188579
SpPDI
2X35S(+C)
Plasto Term
P19
플라스토시아닌
nbHEL40 5'UTR
우측 T-DNA
H1cCT
2000
4000
6000
8000
10 000
12 000

도면21


좌측 T-DNA
aadA (KanaR)
플라스토시아닌
P19
plastocyanin Term
2X35S(+C)
C5-1 리더
nbHEL40 5'UTR
NOS Term
우측 T-DNA
AVb 3'UTR
H1C TMCT
pVS1 rep
단백질
Ori pVS1
ColE1
7147
0
2K
4K
6K
8K
10K

도면22a

```
Native_MERS_S_no_signal_peptide_SEQ_ID_115      YVDVGPDSVKSACIEVDIQQTFFDKTWPRP-IDVSKADGIIYPQGRTYSNITITYQGLF-   58
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 ----------SDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFL   47
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  ----------VNL----TTRTQLPPAY--TNSFTRGVYYPDKVFRSSVLHSTQDLFL   41
                                                          :: .  *  .  * *: **:  *. *.**

Native_MERS_S_no_signal_peptide_SEQ_ID_115      PYQGDHGDMYVYSAGHATGTTPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTVIISPS  118
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 PFYSNVT---GFHTIN------HT----PGNPVIPFKDGIYFAAT-----------E  80
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  PFFSNVT---WFHAIHVSGTNGTKR----FDNPVLPFNDGVYFAST-----------E  81
                                                *: .:     : : :    :.: * * :*. .                .

Native_MERS_S_no_signal_peptide_SEQ_ID_115      TSATIRKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAF--YCILEPRSGN  176
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 ----KSNVVRGWVFGSTMNNKSQ---------SVIIINNSTNVVIRACNFELCDNPFFAV  127
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  ----KSNIIRGWIFGTTLDSKTQ---------SLLIVNNATNVVIKVCEFQFCNDPFLGV  128
                                                  ::  .:::*::::. ::          ::::: :. ..::::.          :* .

Native_MERS_S_no_signal_peptide_SEQ_ID_115      HCPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEW  236
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 SKPMGT-----------------------------QTHTMIFDNAFNCTFEYISDAFSLDVSEK  162
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  YYHKNN-----------------------KSWMESEFRVYSSANNCTFEYVSQPFLMDLEGK  167
                                                ..                       .  : .  **** * :  ::

Native_MERS_S_no_signal_peptide_SEQ_ID_115      FGIYQTAQG-VHLFSSRYVDLYGGN-----------------MFQFATLPVYDTIKYYSIP  280
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 SGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAIL  222
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  QGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLL  227
                                                 * :  :  *.  . :. :*          :  :  **: .*. : :

Native_MERS_S_no_signal_peptide_SEQ_ID_115      HSIR---SIQSDRKAW----AAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCS  333
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 TAFS------PAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCS  276
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  ALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCT  287
                                                          .  *      **:;*  *:* **:*.:. :* *  *:**. : *:: :*:

Native_MERS_S_no_signal_peptide_SEQ_ID_115      YESFDVESGVYSVSSFEAKPSGSVVEQAEG-VECDFSPLLSG-TPPQVYNFKRLVFTNCN  391
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 VKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCV  336
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  LKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCV  347
                                                 :** ::.*:*..*.*.. *: .:*.  :      * *. ::..    .** ::*  ::**

Native_MERS_S_no_signal_peptide_SEQ_ID_115      YNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFN  451
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 ADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYN  396
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  ADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYN  407
                                                 : : * .   .. *.*  :* : : .  *::::  * *     .  ::: ...* *::::*

Native_MERS_S_no_signal_peptide_SEQ_ID_115      YKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRFLSDDRTEVPQLVNANQYSPCVS  511
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 YKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTP  456
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNG  467
                                                **   .   *:: .  : :.      :* *       *    . ::      .**
```

도면22b

```
Native_MERS_S_no_signal_peptide_SEQ_ID_115      IVPST-VWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQYGTDTNSVCPK 570
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 -PALNCYW-----------PLNDYGFYTTTGIGYQPYRVVVLSFELL----NAPATVCGP 500
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  VEGFNCYF-----------PLQSYGFQPTNGVGYQPYRVVVLSFELL----HAPATVCGP 512
                                                              .  :            **:. *:  :.*  .  : :.* :        :**

Native_MERS_S_no_signal_peptide_SEQ_ID_115      LEFANDTKIASQLGNCVEYSLYGVSGRGVFQNCTAVGVRQQRFVYDAYQNLVGYYSD--D 628
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 -----KLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFT-DSVRDPKT 554
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  -----KKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTT-DAVRDPQT 566
                                                     . .      .:***:::. *:;* ***:  ..          *:*  *  :  .    *

Native_MERS_S_no_signal_peptide_SEQ_ID_115      GNYYCLRACVSVPVSVIYD--KETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRRDST 686
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 SEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLT--PAWRIYSTG 612
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  LEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT--PTWRVYSTG 624
                                                  :   :  *     ****     : :.  *.*: .* * .:  ::    . : :  : ::

Native_MERS_S_no_signal_peptide_SEQ_ID_115      YGPLQTPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPSTLTPRSVRSVPGEMRLASIA 746
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 NNVFQTQAGCLIGAEHVD-TSYECDIPIGAGICASYHTVS-L--LRSTSQKSI---VA 663
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  SNVFQTRAGCLIGAEHVN-NSYECDIPIGAGICASYQTQT-NSPRRARSVASQSI---IA 679
                                                 . :** .**:;*  : .    :*.:*:* .:** .* :      *.  :      :*

Native_MERS_S_no_signal_peptide_SEQ_ID_115      FNHPIQV-DQLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCKQYVCNGFQKCEQLLR 805
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 YTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLL 723
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  YTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLL 739
                                                :  . : ...:  *  .:*******:.:* * : .:: *.:***. *:*.. :* :**

Native_MERS_S_no_signal_peptide_SEQ_ID_115      EYGQFCSKINQALHGANLRQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTG 865
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 QYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGF-NFSQILPDP---L 779
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  QYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF-NFSQILPDP---S 795
                                                :**.**:::*:** *   .** .:::**.**.  ::* :  *** *::: * *

Native_MERS_S_no_signal_peptide_SEQ_ID_115      SRSARSAIEDLLFDKVTIADPGYMQGYDDCMQGPASARDLICAQYVAGYKVLPPLMDVN 925
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 KPTKRSFIEDLLFNKVTLADAGFMKQYGECL--GDINARDLICAQKFNGLTVLPPLLTDD 837
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  KPSKRSFIEDLLFNKVTLADAGFIKQYGDCL--GDIAARDLICAQKFNGLTVLPPLLTDE 853
                                                . : **  ********:***:** *:::  *.:*:   *  ******** . * .*****:  :

Native_MERS_S_no_signal_peptide_SEQ_ID_115      MEAAYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANKFN 985
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 MIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFN 897
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  MIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN 913
                                                * * **::*:..      *** * .:  ***** .: **:**:*:**:** **** ***:**

Native_MERS_S_no_signal_peptide_SEQ_ID_115      QALGAMQTGFTTTNEAFHKVQDAVNNNAQALSKLASELSNTFGAISASIGDIIQRLDVLE 1045
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE 957
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  SAIGKIQDSLSSTPSALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE 973
                                                .*:. :* .:::* *: *:** **:***** ..*..:** ..***** :.**:.***.:*
```

도면22c

```
Native_MERS_S_no_signal_peptide_SEQ_ID_115      QDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQLAKDKVNECVKAQSKRSGFCGQGTH 1105
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 AEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH 1017
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  AEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH 1033
                                                 :.******.*** :*:::*****:* .    **:**  *:.*** .**** .***:* *

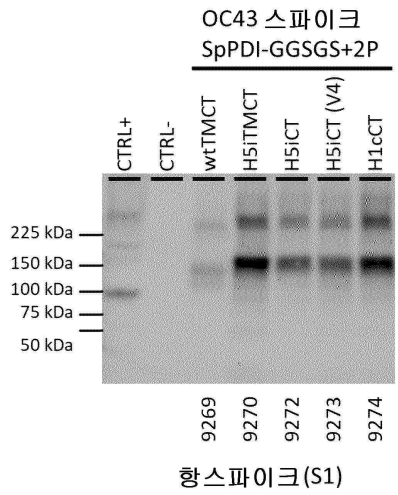
Native_MERS_S_no_signal_peptide_SEQ_ID_115      IVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLCDAANPTNCIAPVNGYFIKTNNTRIVD 1165
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKA---YFPREGVFVFN----GT 1069
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  LMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKA---HFPREGVFVSN----GT 1085
                                                ::***   **:*: *:*** * *::. :..:* .:*. .:     * :* *: .

Native_MERS_S_no_signal_peptide_SEQ_ID_115      EWSYTGSSFYAPEPITSLNTKYVAPQVTYQN-ISTNLPPELLGNSTGIDFQDELDEFFKN 1224
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 SWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQ--PELDSFKEELDKYFKN 1127
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  HWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQ--PELDSFKEELDKYFKN 1143
                                                 * *  .*: *: **: **    .  . :...:   **        .*:;***:;***

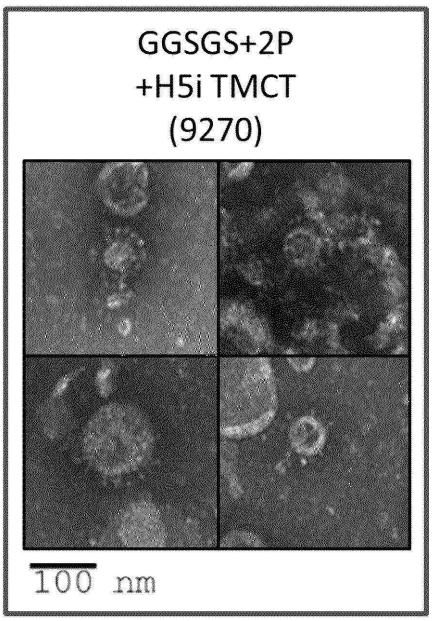
Native_MERS_S_no_signal_peptide_SEQ_ID_115      VSTSIPNFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYIWL 1284
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 HTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL 1187
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  HTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWL 1203
                                                 ::    ::*.:: **:::::::  *:  *:::*.* **** ***:***:*  * *****:**

Native_MERS_S_no_signal_peptide_SEQ_ID_115      GFIAGLVALALCVFFILCCTGCGTNCMGKLKCNRCCDRYEEYDLEPHKVHVH----1336
Native_SARS-CoV-1_S_no_signal_peptide_SEQ_ID_114 GFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKF-DEDDSEPVLKGVKLHYT 1242
Native_SARS-CoV-2_S_no_signal_peptide_SEQ_ID_2  GFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKF-DEDDSEPVLKGVKLHYT 1258
                                                *******:*:: .::: * *.* :   *  .*. **.  :* * **    *:
```

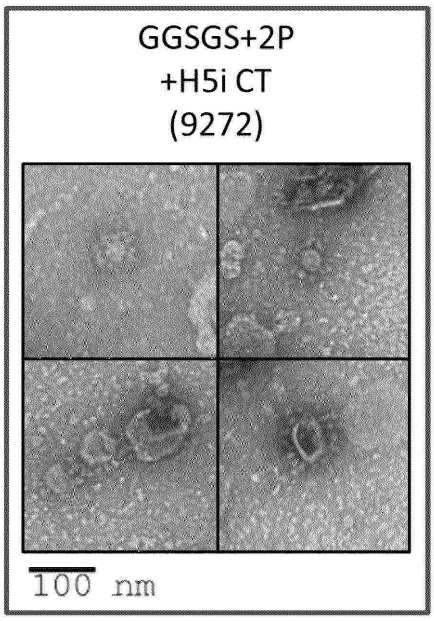
도면23a


OC43 스파이크
SpPDI-GGSGS+2P
CTRL+
CTRL-
wtTMCT
H5iTMCT
H5iCT
H5iCT (V4)
H1cCT
225 kDa
150 kDa
100 kDa
75 kDa
50 kDa
9269
9270
9272
9273
9274
항스파이크(S1)

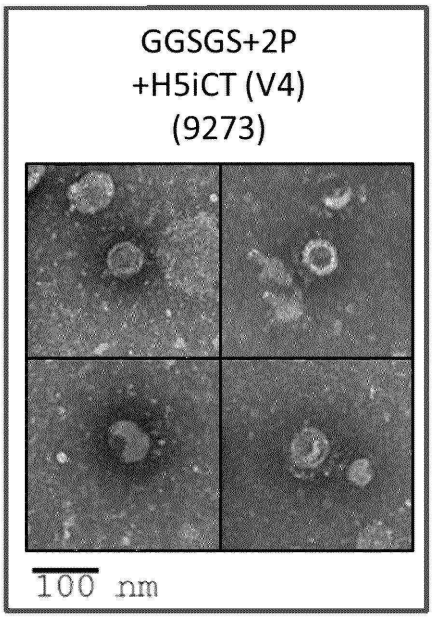
도면23b


GGSGS+2P
+H5i TMCT
(9270)
100 nm

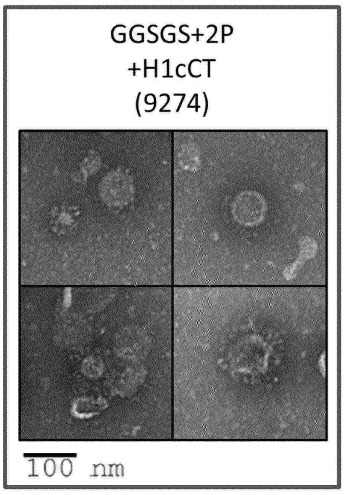
도면23c



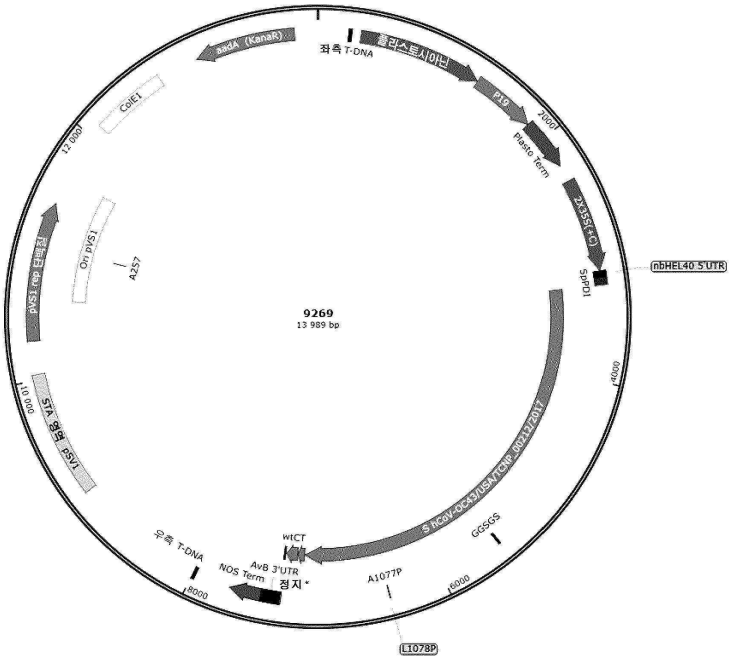
도면23d



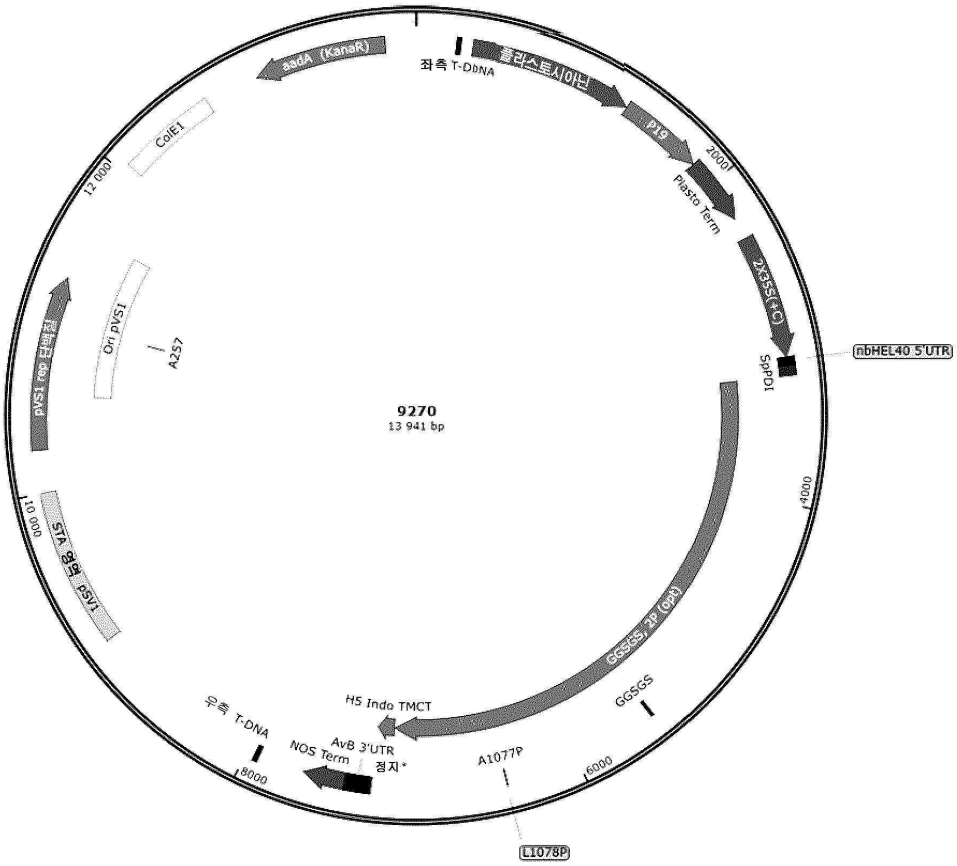
도면23e


GGSGS+2P
+H1cCT
(9274)
100 nm

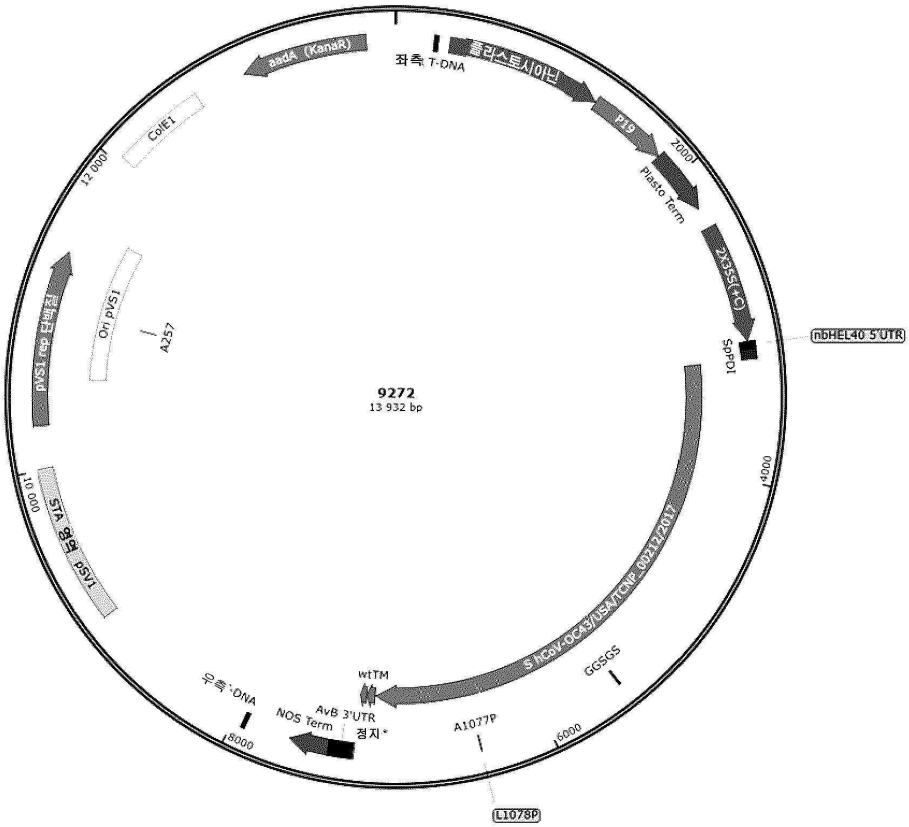
도면24a


aadA (KanaR)
좌측 T-DNA
CoIE1
2000
Nos Term
2x35S(+Ω)
SpPDI
nbHEL40 5'UTR
9269
13 989 bp
Ori pVS1
A257
pVS1 rep
STA 영역 pVS1
wtCT
GGSGS
AvB 3'UTR
NOS Term
정지
우측 T-DNA
A1077P
L1078P

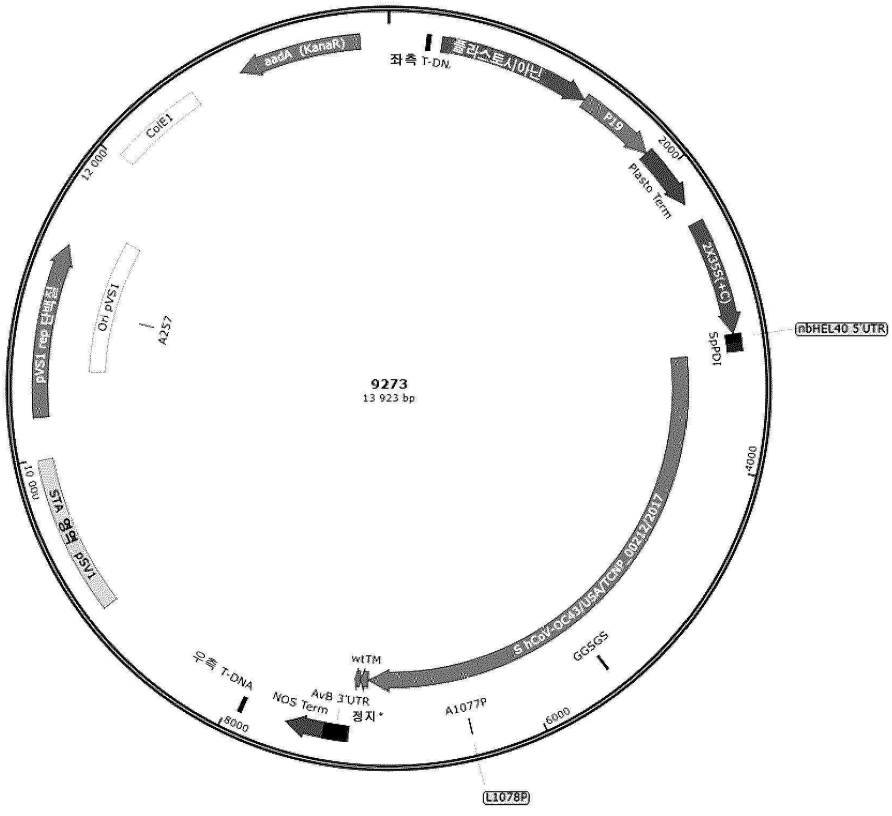
도면24b


9270
13 941 bp
aadA (KanaR)
좌측 T-DbNA
플라스토시아닌
P1P
Plasto Term
2A35S(+C)
SpPDI
nbHEL40 5'UTR
GGSGS 2P (GPI)
GGSGS
H5 Indo TMCT
AvB 3'UTR
정지*
NOS Term
우측 T-DNA
A1077P
L1078P
ColE1
Ori pVS1
A2S7
pVS1 Rep ori
STA 융합 pSV1
12 000
2000
4000
6000
8000
10 000

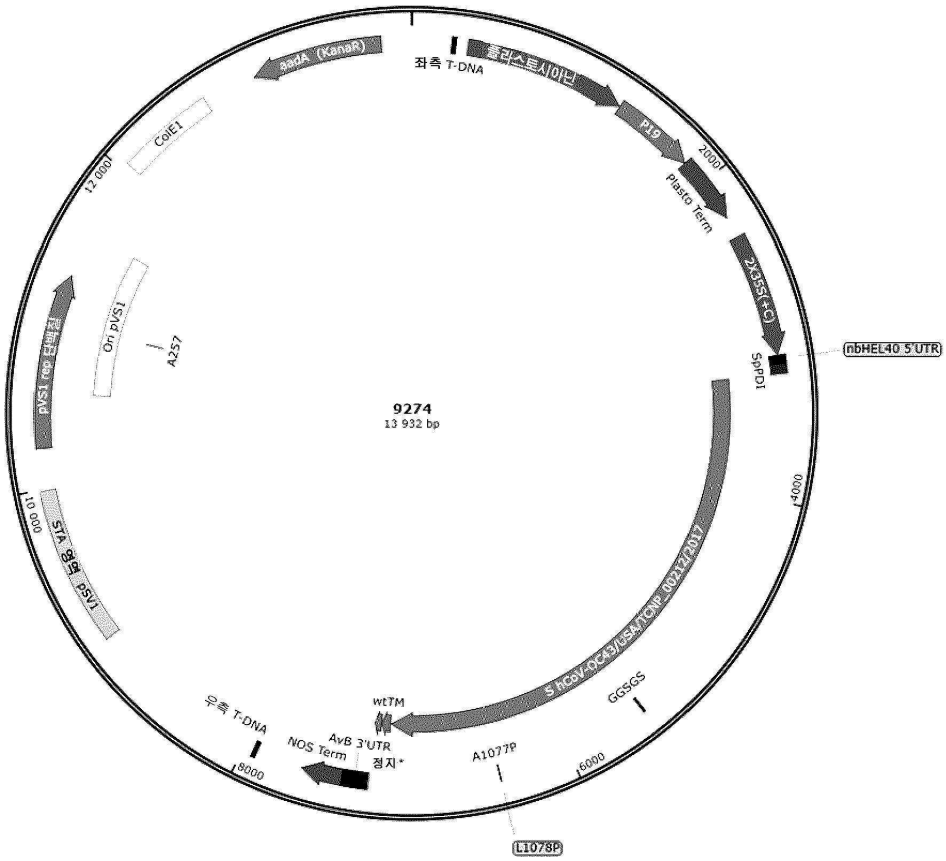
도면24c


9272
13 932 bp
aadA (KanaR)
좌측 T-DNA
폴리스토시아닌
P19
Plasto Term
2X35S(+Ω)
SpPDI
nbHEL40 5'UTR
ColE1
Ori PVS1
A257
레플리카제 repA PSV1
STA 영역 PSV1
우측 T-DNA
NOS Term
AvB 3'UTR
정지*
wtTM
GGSGS
A1077P
L1078P

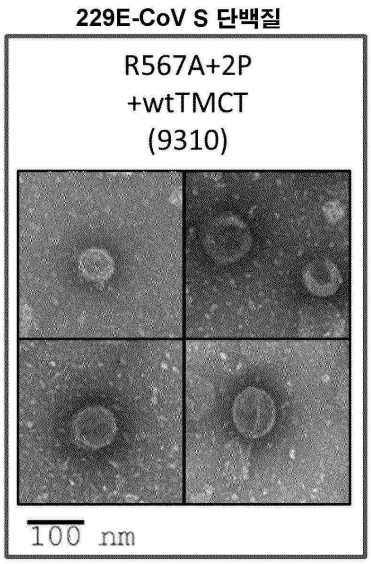
도면24d


9273
13 923 bp
aadA (KanaR)
좌측 T-DNA
폴리스토시아닌
P19
Plasto Term
2X35S(+Ω)
SpPDI
nbHEL40 5'UTR
ColE1
Ori PVS1
A257
레플리카제 repA PSV1
STA 영역 PSV1
우측 T-DNA
NOS Term
AvB 3'UTR
정지*
wtTM
GGSGS
A1077P
L1078P

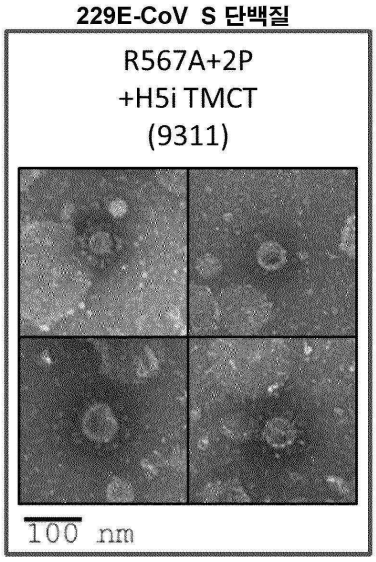
도면24e


9274
13 932 bp
좌측 T-DNA
PJP
Plasto Term
2x35S(+Ω)
SpPDI
nbHEL40 5'UTR
4000
6000
A1077P
L1078P
GGSGS
wtTM
AvB 3'UTR
정지*
NOS Term
우측 T-DNA
8000
10 000
STA 영역 pSV1
pVS1 rep 단백질
On pVS1
A257
12 000
ColE1
2000
aadA (KanaR)

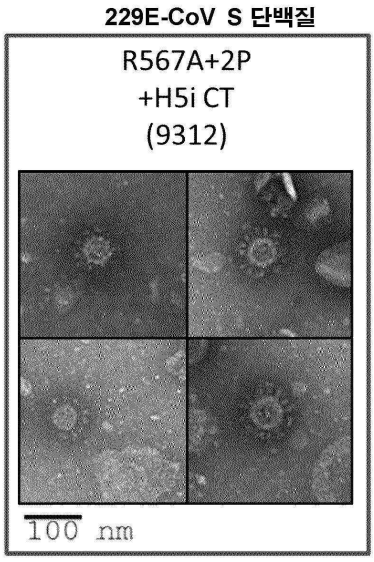
도면25a


229E-CoV S 단백질
R567A+2P
+wtTMCT
(9310)
100 nm

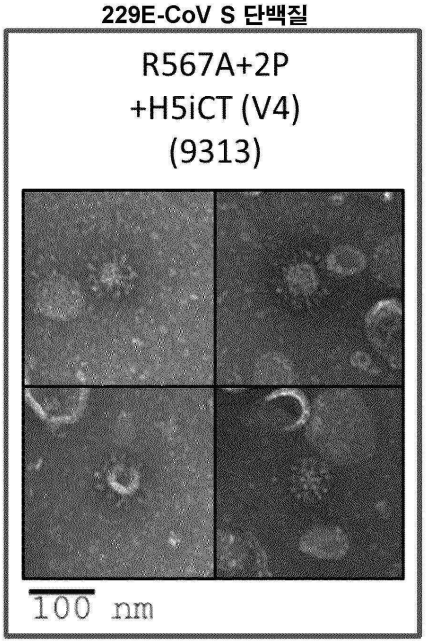
도면25b


229E-CoV S 단백질
R567A+2P
+H5i TMCT
(9311)
100 nm

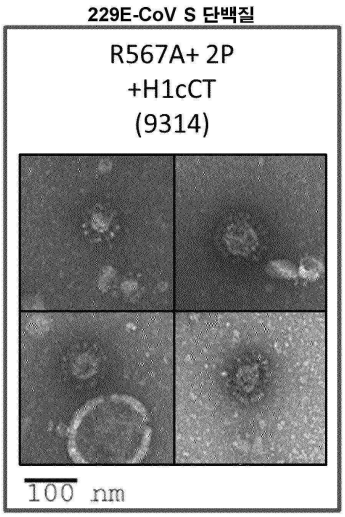
도면25c


229E-CoV S 단백질
R567A+2P
+H5i CT
(9312)
100 nm

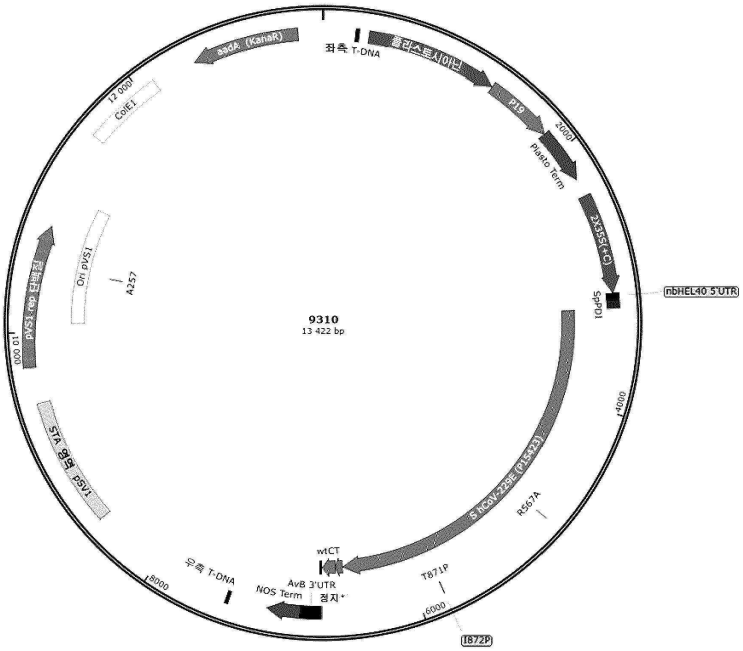
도면25d


229E-CoV S 단백질
R567A+2P
+H5iCT (V4)
(9313)
100 nm

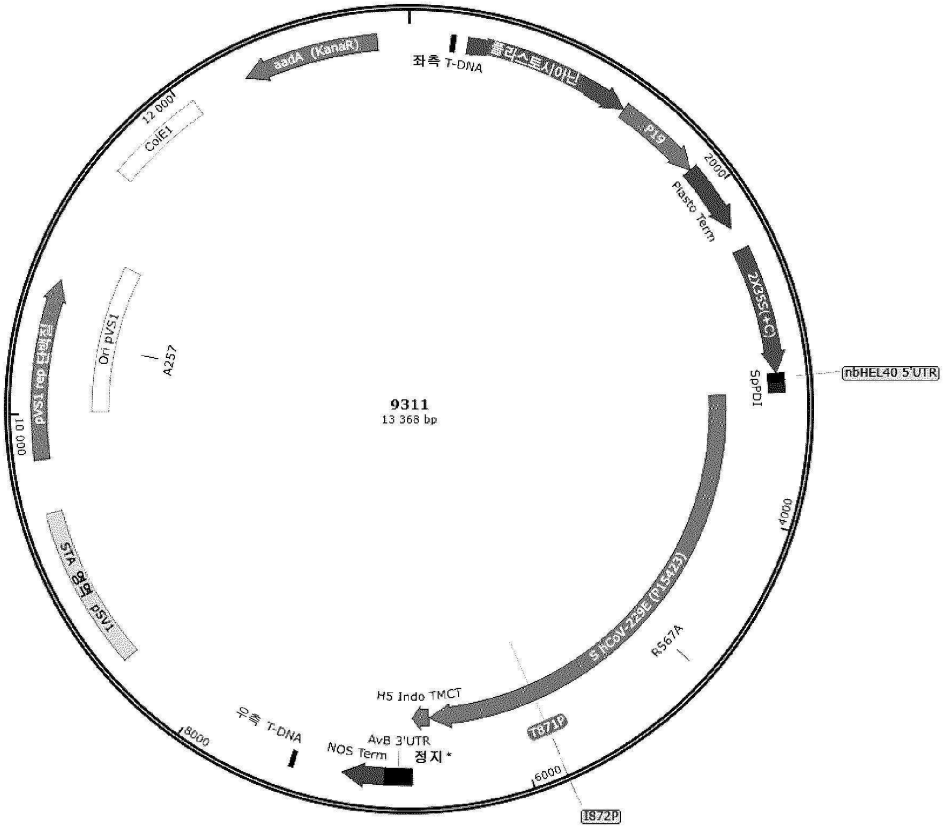
도면25e


229E-CoV S 단백질
R567A+ 2P
+H1cCT
(9314)
100 nm

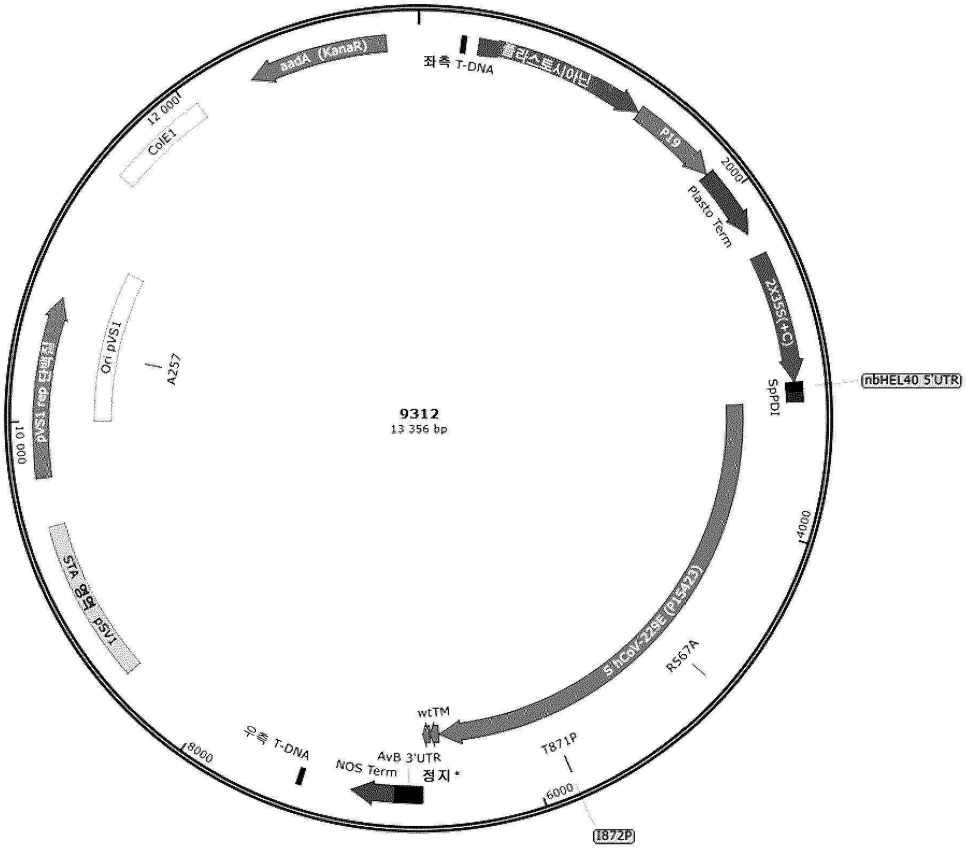
도면26a


aadA (KanaR)
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
R567A
T871P
I872P
wtCT
AvB 3'UTR
NOS Term
정지 *
우측 T-DNA
STA 융합 PSV1
PVS1 rep 단백질
Ori PVS1
A257
CoiEI
9310
13 422 bp

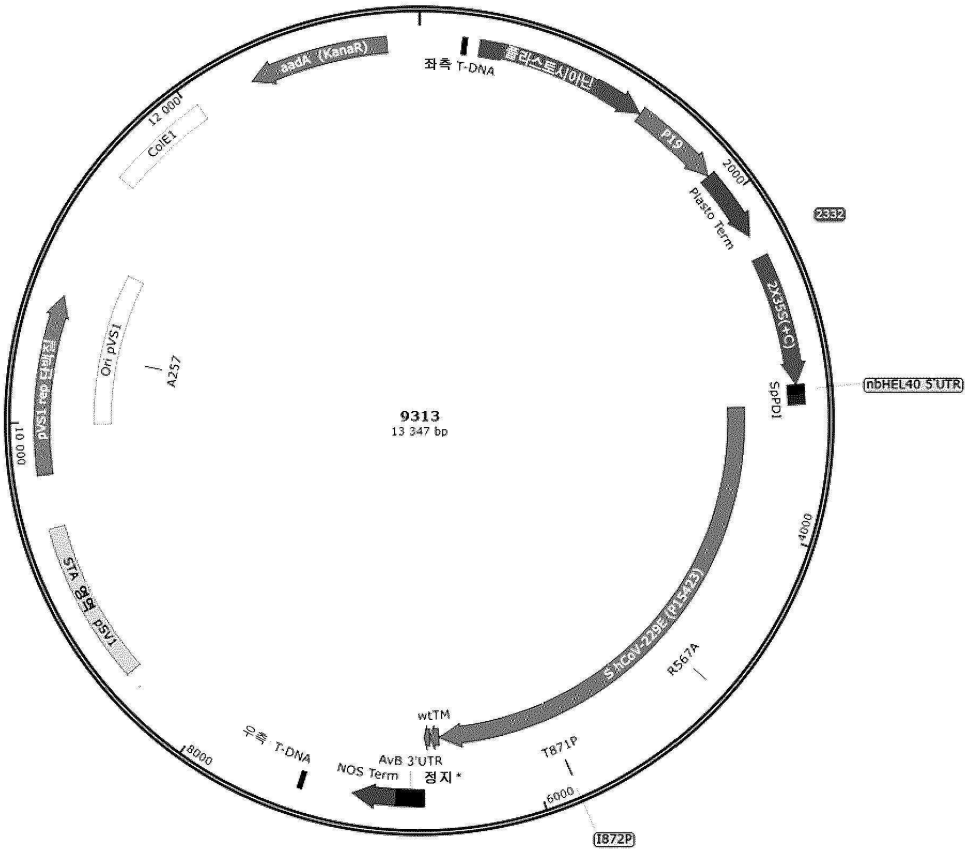
도면26b


aadA (KanaR)
좌측 T-DNA
플라스토시아닌
P19
Plasto Term
2X35S(+C)
SpPDI
nbHEL40 5'UTR
R567A
T871P
I872P
H5 Indo TMCT
AvB 3'UTR
NOS Term
정지 *
우측 T-DNA
STA 융합 PSV1
PVS1 rep 단백질
Ori PVS1
A257
CoiEI
9311
13 368 bp

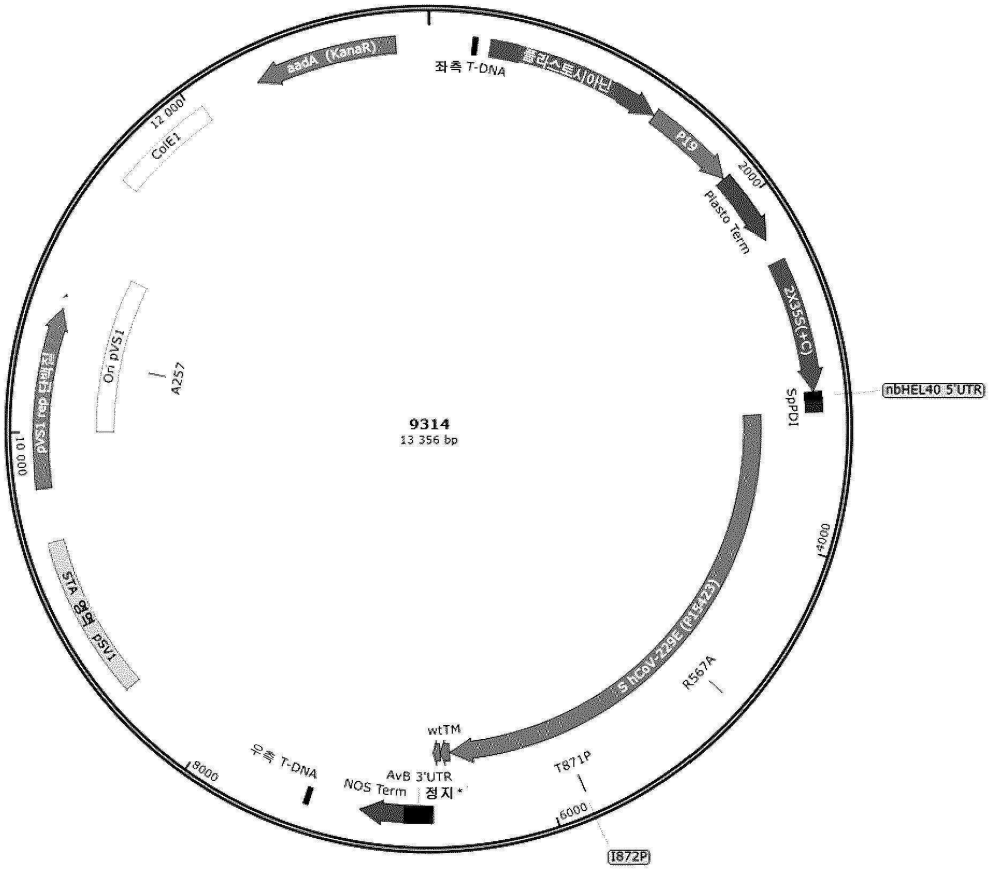
도면26c


9312
13 356 bp
aadA (KanaR)
좌측 T-DNA
플라스토시아닌
PIP
PsbTo Term
2X35S(+Ω)
SpPDI
nbHEL40 5'UTR
14000
S1HC51-2SE (R567A)
R567A
T871P
I872P
6000
wtTM
AvB 3'UTR
NOS Term
정지 *
우측 T-DNA
18000
STA 융합 PSV1
PVS1 rep 단백질
10 000
Ori PVS1
A257
CoIE1
12 000
2000

도면26d


9313
13 347 bp
aadA (KanaR)
좌측 T-DNA
클라스트시아닌
P19
Plasto Term
2X35S(+C)
SoPDI
nbHEL40 5'UTR
2332
4000
S.hCoV-229E (RTD25)
R567A
T871P
I872P
6000
wtTM
AvB 3'UTR
NOS Term
정지 *
우측 T-DNA
8000
STA 영역 pSV1
pVS1 Rep-Trfa
10 000
Ori pVS1
A257
ColE1
12 000
2000

도면26e



서 열 목 록

SEQUENCE LISTING

<110> Medicago Inc.

<120> Modified Viral Structural Protein

<130> V815385WO

<150> US 63/073,327

<151> 2020-09-01

<150> US 63/211,716

<151> 2021-06-17

<160> 173

<170> PatentIn version 3.5

<210> 1

<211> 1273

<212> PRT

<213> SARS-CoV-2

<400> 1

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

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

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

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

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

Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu
85 90 95

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

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

Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
130 135 140

Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
145 150 155 160

Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
165 170 175

Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
180 185 190

Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
245 250 255

Gly Trp Thr Ala Gly Ala Ala 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
1010 1015 1020

Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys
1025 1030 1035

Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro
1040 1045 1050

Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val
1055 1060 1065

Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His
1070 1075 1080

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn
1085 1090 1095

Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln
1100 1105 1110

Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val
1115 1120 1125

Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro
1130 1135 1140

Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1145 1150 1155

His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn
1160 1165 1170

Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu
1175 1180 1185

Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu
1190 1195 1200

Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu
1205 1210 1215

Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met
1220 1225 1230

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

<210> 2
<211> 1258
<212> PRT
<213> SARS-CoV-2
<400> 2

Val Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser
1 5 10 15

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

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

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

Glu Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln
210 215 220
Thr Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser
225 230 235 240
Ser Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln
245 250 255
Pro Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp
260 265 270

Ala Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu
275 280 285
Lys Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg
290 295 300
Val Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu
305 310 315 320
Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr
325 330 335

Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val
340 345 350
Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser
355 360 365
Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser
370 375 380
Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr

385 390 395 400

Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly
405 410 415
Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly
420 425 430
Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro
435 440 445
Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro
450 455 460

Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr
465 470 475 480
Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val
485 490 495
Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro
500 505 510
Lys Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe
515 520 525

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

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

Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys
645 650 655

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

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

Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg
785 790 795 800
Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala
805 810 815
Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg
820 825 830
Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro
835 840 845

Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala
850 855 860
Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln
865 870 875 880
Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

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

980 985 990

Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn

995 1000 1005

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

1010 1015 1020

Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro

1025 1030 1035

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

1040 1045 1050

Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His

1055 1060 1065

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn

1070 1075 1080

Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln

1085 1090 1095

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

1100 1105 1110

Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro

1115 1120 1125

Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1130 1135 1140
His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn
1145 1150 1155

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

Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys
1220 1225 1230
Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro
1235 1240 1245
Val Leu Lys Gly Val Lys Leu His Tyr Thr
1250 1255

<210> 3
<211> 568
<212> PRT
<213> Influenza
<400> 3
Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30
Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45
Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60
Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn

65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
100 105 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
115 120 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
130 135 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145 150 155 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
165 170 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
180 185 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
195 200 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
210 215 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
260 265 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305 310 315 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
325 330 335

Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
340 345 350
Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
355 360 365
Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
370 375 380
Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
385 390 395 400

Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
405 410 415
Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
420 425 430
Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
435 440 445
Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
450 455 460

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465 470 475 480
Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
485 490 495
Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
500 505 510
Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
515 520 525

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540
Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545 550 555 560
Ser Leu Gln Cys Arg Ile Cys Ile

565

<210> 4

<211> 1729

<212> DNA

<213> Influenza

<400> 4

ctgtaaaaat ggagaaaata gtgcttcttc ttgcaatagt cagtcttgtt aaaagtgatc 60

agatttgcat tggttaccat gcaaacaatt caacagagca ggttgacaca atcatggaaa 120

agaacgttac tgttacacat gcccaagaca tactggaaaa gacacacaac gggaagctct 180

gcgatctaga tggagtgaag cctctaattt taagagattg tagtgtagct ggatggctcc 240

tcgggaaccc aatgtgtgac gaattcatca atgtaccgga atggtcttac atagtggaga 300

aggccaatcc aaccaatgac ctctgttacc cagggagttt caacgactat gaagaactga 360

aacacctatt gagcagaata aaccattttg agaaaattca aatcatcccc aaaagttctt 420

ggtccgatca tgaagcctca tcaggagtga gctcagcatg tccatacctg ggaagtccct 480

ccttttttag aaatgtggta tggcttatca aaaagaacag tacataccca acaataaaga 540

aaagctacaa taataccaac caagaagatc ttttggtact gtggggaatt caccatccta 600

atgatgcggc agagcagaca aggctatatc aaaacccaac cacctatatt tccattggga 660

catcaacact aaaccagaga ttggtaccaa aaatagctac tagatccaaa gtaaacgggc 720

aaagtggaag gatggagttc ttctggacaa ttttaaaacc taatgatgca atcaacttcg 780

agagtaatgg aaatttcatt gctccagaat atgcatacaa aattgtcaag aaaggggact 840

cagcaattat gaaaagtgaa ttggaatatg gtaactgcaa caccaagtgt caaactccaa 900

tgggggcgat aaactctagt atgccattcc acaacataca ccctctcacc atcggggaat 960

gccccaaata tgtgaaatca aacagattag tccttgcaac agggctcaga aatagccctc 1020

aaagagagag cagaagaaaa aagagaggac tatttggagc tatagcaggt tttatagagg 1080

gaggatggca gggaatggta gatggttggt atgggtacca ccatagcaat gagcagggga 1140

gtgggtacgc tgcagacaaa gaatccactc aaaaggcaat agatggagtc accaataagg 1200

tcaactcaat cattgacaaa atgaacactc agtttgaggc cgttggaagg gaatttaata 1260

acttagaaag gagaatagag aatttaaaca agaagatgga agacgggttt ctagatgtct 1320

ggacttataa tgccgaactt ctggttctca tggaaaatga gagaactcta gactttcatg 1380

actcaaatgt taagaacctc tacgacaagg tccgactaca gcttagggat aatgcaaagg 1440

agctgggtaa cggttgtttc gagttctatc acaaatgtga taatgaatgt atggaaagta 1500

taagaaacgg aacgtacaac tatccgcagt attcagaaga agcaagatta aaaagagagg 1560

aaataagtgg ggtaaaattg gaatcaatag gaacttacca aatactgtca atttattcaa 1620

cagtggcgag ttccctagca ctggcaatca tgatggctgg tctatcttta tggatgtgct 1680

ccaatggatc gttacaatgc agaatttgca tttaaatttg tgagttcag 1729

<210> 5

<211> 1250

<212> PRT

<213> Artificial Sequence

<220><223> Modified SARS-CoV-2 with H5 A/Indonesia/5/05 Hemagglutinin CT AA

<400> 5

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

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

100 105 110

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

115 120 125

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

1010 1015 1020

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

1025 1030 1035

Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro

1040 1045 1050

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

1055 1060 1065

Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His

1070 1075 1080

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn

1085 1090 1095

Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln

1100 1105 1110

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

1115 1120 1125

Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro

1130 1135 1140

Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn

1145 1150 1155

His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

1235 1240 1245

Cys Ile

1250

<210> 6

<211> 50

<212> PRT

<213> Influenza

<400> 6

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

1 5 10 15

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

20 25 30

Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

35 40 45

Cys Ile

50

<210> 7

<211> 50

<212> PRT

<213> Influenza

<400> 7

Ile Lys Gly Val Lys Leu Ser Ser Met Gly Val Tyr Gln Ile Leu Ala

1 5 10 15

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

20 25 30

Gly Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

35 40 45

Cys Ile

50

<210> 8

<211> 49

<212> PRT

<213> Influenza

<400> 8

Ile Lys Gly Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp

1 5 10 15

Ile Ser Phe Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly

20 25 30

Phe Ile Met Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys

35 40 45

Ile

<210> 9

<211> 50

<212> PRT

<213> Influenza

<400> 9

Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile Leu Ser

1 5 10 15

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

20 25 30

Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

35 40 45

Cys Ile

50

<210> 10

<211> 50

<212> PRT

<213> Influenza

<400> 10

Ile Glu Ser Val Lys Leu Glu Asn Leu Gly Val Tyr Gln Ile Leu Ala
1 5 10 15

Ile Tyr Ser Thr Val Ser Ser Ser Leu Val Leu Val Gly Leu Ile Met
20 25 30

Ala Met Gly Leu Trp Met Cys Ser Asn Gly Ser Met Gln Cys Arg Ile
35 40 45

Cys Ile

50

<210> 11

<211> 49

<212> PRT

<213> Influenza

<400> 11

Ile Asp Pro Val Lys Leu Ser Ser Gly Tyr Lys Asp Val Ile Leu Trp
1 5 10 15

Phe Ser Phe Gly Ala Ser Cys Phe Ile Leu Leu Ala Ile Val Met Gly
20 25 30

Leu Val Phe Ile Cys Val Lys Asn Gly Asn Met Arg Cys Thr Ile Cys
35 40 45

Ile

<210> 12

<211> 50

<212> PRT

<213> Influenza

<400> 12

Ile Glu Gly Val Lys Leu Glu Ser Glu Gly Thr Tyr Lys Ile Leu Thr

1 5 10 15

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

20 25 30

Ala Phe Leu Phe Trp Ala Met Ser Asn Gly Ser Cys Arg Cys Asn Ile

35 40 45

Cys Ile

50

<210> 13

<211> 50

<212> PRT

<213> Influenza

<400> 13

Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu

1 5 10 15

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

20 25 30

Ile Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile

35 40 45

Cys Leu

50

<210> 14

<211> 50

<212> PRT

<213> Artificial Sequence

<220><223> Consensus Sequence of C-Terminal Region of Influenza Hemagglutinin

<220><221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (7)..(7)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (9)..(9)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (11)..(11)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (13)..(13)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (16)..(16)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (26)..(26)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (28)..(36)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (44)..(45)

<223

> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (47)..(47)

<223> Xaa can be any naturally occurring amino acid

<400> 14

Ile Xaa Gly Val Lys Leu Xaa Ser Xaa Gly Xaa Tyr Xaa Ile Leu Xaa
1 5 10 15

Ile Tyr Ser Thr Val Ala Ser Ser Leu Xaa Leu Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Trp Met Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa Ile
35 40 45

Cys Ile
50

<210> 15

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Consensus Sequence of CT Domain of Influenza Hemagglutinin

<220><221> misc_feature

<222> (1)..(2)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (10)..(11)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (13)..(13)

<223> Xaa can be any naturally occurring amino acid

<400> 15

Xaa Xaa Trp Met Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa Ile Cys Ile

1 5 10 15

<210> 16

<211> 69

<212> PRT

<213> SARS-CoV-2

<400> 16

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

1 5 10 15

Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys

20 25 30

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

35 40 45

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

50 55 60

Lys Leu His Tyr Thr

65

<210> 17

<211> 50

<212> PRT

<213> Influenza

<400> 17

Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile Leu Ser

1 5 10 15

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

20 25 30

Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

35 40 45

Cys Ile

50

<210> 18

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> C-Terminal Region of Modified SARS-CoV-2 S protein with H5i Hemagglutinin CT

<400> 18

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

1 5 10 15

Val Thr Ile Met Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln

20 25 30

Cys Arg Ile Cys Ile

35

<210> 19

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> C-Terminal Region of Modified SARS-CoV-2 S protein with H5i Hemagglutinin CT, Variation 1

<400> 19

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

1 5 10 15

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

20 25 30

Ser Leu Gln Cys Arg Ile Cys Ile

35 40

<210> 20

<211> 1258

<212> PRT

<213> Artificial Sequence

<220><223> (no SP) SARS-CoV-2 S protein GSAS+PP wtTM/CT AA

<400> 20

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

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

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

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

165 170 175

Phe Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His
180 185 190
Thr Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu
195 200 205
Glu Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln
210 215 220
Thr Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser
225 230 235 240

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

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

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

Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly
420 425 430

Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro
435 440 445

Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro
450 455 460

Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr
465 470 475 480

Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val
485 490 495

Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro
500 505 510

Lys Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe
515 520 525

Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe
530 535 540

Leu Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala
545 550 555 560

Val Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser
565 570 575

Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln
580 585 590

Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala
595 600 605

Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly
610 615 620

Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His
625 630 635 640

Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys
645 650 655

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

660 665 670

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

675 680 685

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

690 695 700

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

705 710 715 720

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

725 730 735

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

740 745 750

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

755 760 765

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

770 775 780

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

785 790 795 800

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

805 810 815

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

820 825 830

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

835 840 845

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

850 855 860

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

865 870 875 880

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

885 890 895

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

900 905 910

Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser
915 920 925
Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu
930 935 940

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

Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys
1010 1015 1020
Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro
1025 1030 1035
Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val
1040 1045 1050
Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His
1055 1060 1065

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn
1070 1075 1080
Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln
1085 1090 1095
Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val
1100 1105 1110
Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro
1115 1120 1125

Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1130 1135 1140
His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

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

1235 1240 1245

Val Leu Lys Gly Val Lys Leu His Tyr Thr

1250 1255

<210> 21

<211> 1235

<212> PRT

<213> Artificial Sequence

<220><223> (no SP) Modified SARS-CoV-2 S protein GSAS+PP with H5i Hemagglutinin CT AA

<400> 21

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

Glu Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp
85 90 95

Ser Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val
100 105 110

Ile Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val
115 120 125

Tyr Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val
130 135 140

Tyr Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe
145 150 155 160

Leu Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu
165 170 175

Phe Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His
180 185 190

Thr Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu
195 200 205

Glu Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln
210 215 220

Thr Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser
225 230 235 240

Ser Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln
245 250 255

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

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

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

Val Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu
305 310 315 320

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

325 330 335

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

340 345 350

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

355 360 365

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

370 375 380

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

385 390 395 400

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

405 410 415

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

420 425 430

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

435 440 445

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

450 455 460

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

465 470 475 480

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

485 490 495

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

500 505 510

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

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln
580 585 590
Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala
595 600 605

Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly
610 615 620
Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His
625 630 635 640
Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys
645 650 655
Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Gly Ser Ala Ser Ser Val
660 665 670

Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn
675 680 685
Ser Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr
690 695 700
Ile Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser
705 710 715 720
Val Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn
725 730 735

Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu
740 745 750
Thr Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala
755 760 765
Gln Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly
770 775 780
Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg
785 790 795 800

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala
805 810 815
Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg

820 825 830

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

835 840 845

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

850 855 860

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

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

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

980 985 990

Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn

995 1000 1005

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

1010 1015 1020

Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro

1025 1030 1035

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

1040 1045 1050

Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His

1055 1060 1065

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn
1070 1075 1080
Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln
1085 1090 1095
Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val
1100 1105 1110

Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro
1115 1120 1125
Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1130 1135 1140
His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn
1145 1150 1155
Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu
1160 1165 1170

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

Cys Ile
1235

<210> 22
<211> 3849
<212> DNA
<213> Artificial Sequence
<220><223> PDI-SARS-CoV-2 S protein GSAS+PP wtTM/CT-DNA
<400> 22
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280
ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340
gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460
ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760
acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880
gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940
gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300
gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360
ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420
atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480
ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540
atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600
aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct gctgcatgac ctcttgctgc tcctgtctga aaggctgctg ctcttgcgga 3780

tcctgctgca aatttgatga ggatgacagt gaaccagtcc tgaagggcgt gaagctgcac 3840

tatacttag 3849

<210> 23

<211> 1282

<212> PRT

<213> Artificial Sequence

<220><223> PDI-SARS-CoV-2 S protein GSAS+PP wtTM/CT-AA

<400> 23

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140

Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190
Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220
Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240
Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255
Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285
Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320
Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350
Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380
Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415
Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685
Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser
705 710 715 720

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

Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
805 810 815
Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
820 825 830
Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845

Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880
Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895
Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met Thr Ser
1235 1240 1245
Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys
1250 1255 1260
Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly Val Lys
1265 1270 1275

Leu His Tyr Thr
1280

<210> 24
<211> 52
<212> DNA
<213> Artificial Sequence
<220><223> IF(PDI)-CoV(opt2).c
<400> 24
tctcagatct tcgcggtgaa tcttacgacg cgaacacagt taccacccgc at 52

<210> 25
<211> 51
<212> DNA
<213> Artificial Sequence
<220><223> IF(AVB)-CoV(opt2).r
<400> 25
acgacacgac taaggcctct aagtatagtg cagcttcacg cccttcagga c 51

<210> 26

<211> 3789

<212> DNA

<213> Artificial Sequence

<220><223> PDI-Modified SARS-CoV-2 S protein GSAS+PP H5iTM/CT-DNA

<400> 26

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atcaaatact gtcaatttat tcaacagtgg cgagttccct agcactggca 3720

atcatgatgg ctggtctatc tttatggatg tgctccaatg gatcgttaca atgcagaatt 3780

tgcatttaa 3789

<210> 27

<211> 1262

<212> PRT

<213> Artificial Sequence

<220><223> PDI-Modified SARS-CoV-2 S protein GSAS+PP H5iTM/CT-AA

<400> 27

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

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

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

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

290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320
Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335
Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350
Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser

355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser

370 375 380

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

385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val

405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640
Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655
Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp
660 665 670
Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685
Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser
705 710 715 720
Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu
725 730 735
Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750
Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe
755 760 765
Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp
770 775 780
Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro

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

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val

1220 1225 1230

Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly Leu Ser Leu

1235 1240 1245

Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255 1260

<210> 28

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> IF(Avb)-H5I.r

<400> 28

acgacacgac taaggccttt aaatgcaaat tctgcattgt aacgatcc 48

<210> 29

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-Modified SARS-CoV-2 S protein GSAS+PP wtTM/H5iCT-DNA

<400

> 29

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa 3780

<210> 30

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> PDI-Modified SARS-CoV-2 S protein GSAS+PP wtTM/H5iCT-AA

<400> 30

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140

Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr

165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190
Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205
Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220
Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240

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

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

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

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr
420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460

Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn
485 490 495

Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525

Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540

Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560

Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575

Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr
595 600 605

Pro Gly Thr Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val
610 615 620

Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

900 905 910

Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925
Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
930 935 940

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

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035
Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050
Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080
Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125

Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys

1235 1240 1245

Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 31

<211> 4474

<212> DNA

<213> Artificial Sequence

<220><223> Cloning vector 8501 from left to right T-DNA

<400> 31

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg cacacaattt gctttagtga ttaaactttc ttttacaaca aattaaaggt 2940

ctattatctc ccaacaacat aagaaaacaa tggcgaaaaa cgttgcgatt ttcggcttat 3000

tgttttctct tcttgtgttg gttccttctc agatcttcgc gacgtcactc ctcagccaaa 3060

acgacacccc catctgtcta tccactggcc cctggatctg ctgcccaaac taactccatg 3120

gtgaccctgg gatgcctggt caagggctat ttccctgagc cagtgacagt gacctggaac 3180

tctggatccc tgtccagcgg tgtgcacacc ttcccagctg tcctgcagtc tgacctctac 3240

actctgagca gctcagtgac tgtcccctcc agcacctggc ccagcgagac cgtcacctgc 3300

aacgttgccc acccggccag cagcaccaag gtggacaaga aaattgtgcc cagggattgt 3360

ggttgtaagc cttgcatatg tacagtccca gaagtatcat ctgtcttcat cttccccca 3420

aagcccaagg atgtgctcac cattactctg actcctaagg tcacgtgtgt tgtggtagac 3480

atcagcaagg atgatcccga ggtccagttc agctggtttg tagatgatgt ggaggtgcac 3540

acagctcaga cgcaaccccg ggaggagcag ttcaacagca ctttccgctc agtcagtgaa 3600

cttcccatca tgcaccagga ctggctcaat ggcaaggaga cgtccagatt ttggcgatct 3660

attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc agtttctgga 3720

tgtgctctaa tgggtctcta cagtgtagaa tatgtattta aaggccttag tcgtgtcgtt 3780

tttcaaataa tataatcctt ttagggtttt agttagttta aattttctgt tgctcctgtt 3840

tagcaggtcg tgccttcagc aagcacacaa aaacagagtg tttattttaa gttgtttgtt 3900

tagtgattca aaaaaaaaat cgttcaaaca tttggcaata aagtttctta agattgaatc 3960

ctgttgccgg tcttgcgatg attatcatat aatttctgtt gaattacgtt aagcatgtaa 4020

taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt agagtcccgc 4080

aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat 4140

cgcgcgcggt gtcatctatg ttactagatc tctagagtct caagcttggc gcgcccacgt 4200

gactagtggc actggccgtc gttttacaac gtcgtgactg ggaaaaccct ggcgttaccc 4260

aacttaatcg ccttgcagca catccccctt tcgccagctg gcgtaatagc gaagaggccc 4320

gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg cgaatgctag agcagcttga 4380

gcttggatca gattgtcgtt tcccgccttc agtttaaact atcagtgttt gacaggatat 4440

attggcgggt aaacctaaga gaaaagagcg ttta 4474

<210> 32

<211> 5083

<212> DNA

<213> Artificial Sequence

<220><223> Construct 8586 from 2X35S prom to NOS term

<400> 32

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggcac acaatttgct ttagtgatta aactttcttt 780

tacaacaaat taaaggtcta ttatctccca acaacataag aaaacaatgg cgaaaaacgt 840

tgcgattttc ggcttattgt tttctcttct tgtgttggtt ccttctcaga tcttcgcggt 900

gaatcttacg acgcgaacac agttaccacc cgcatataca aatagcttca ctcggggtgt 960

ttattacccc gacaaagtgt tcaggtcctc cgtgctccac tcaacacagg acctctttct 1020

tcctttcttt tctaacgtga catggtttca tgccattcat gtatccggca ctaacggtac 1080

taagaggttc gataatcctg tgctcccttt caatgacggc gtttactttg caagcacaga 1140

gaagagtaac atcatccgag gttggatctt tggcactacc ctcgattcaa agacgcagag 1200

cctcctcatt gtgaacaatg ccactaacgt ggtgatcaaa gtttgcgagt ttcagttctg 1260

caatgaccct ttcttggggg tgtactatca taagaacaac aagtcttgga tggaatctga 1320

attccgcgtc tatagcagcg ccaacaactg cacctttgaa tacgtgtccc agcccttcct 1380

tatggacctg gagggaaagc agggaaactt taagaatctg agagagttcg tgtttaaaaa 1440

tatcgacggc tattttaaga tctattctaa gcacacgcct attaatctcg tgcgcgatct 1500

tccacaaggc ttcagcgccc tggaaccact cgtggacctc ccaattggta tcaacatcac 1560

tagatttcag actctgcttg ccctccaccg atcctatctg acacccggag actcctctag 1620

cggctggact gccggcgctg ccgcttatta cgttggttat cttcagccac gcacgttcct 1680

gctgaagtat aacgagaatg gtactattac cgatgccgtg gattgtgccc ttgacccccc 1740

gtccgaaact aagtgcacac tcaagtcatt cactgtggaa aaaggaatct accagacaag 1800

caattttcgg gtccagccta ctgagagcat tgtgcgcttt cctaacatca caaatctttg 1860

ccccttcgga gaggttttca atgctacacg gtttgcctcc gtgtatgcct ggaaccgcaa 1920

gagaatttcc aattgcgtgg ccgattactc cgtgctctac aatagtgcaa gctttagcac 1980

ctttaagtgc tatggcgtat cccctactaa gcttaacgac ttgtgtttca caaacgtgta 2040

tgccgactcc tttgtgatac ggggcgacga agttagacag atagcaccag gacagacggg 2100

aaagatagct gactacaact ataagcttcc tgatgacttc actggctgcg ttatcgcgtg 2160

gaattctaac aacctggact caaaagtcgg cggcaactat aactatctct atcggctgtt 2220

ccgcaagagt aaccttaagc cctttgagag agatataagc actgaaatct accaggctgg 2280

cagtacgccc tgtaatggcg tggaaggctt taattgttat tttccactgc aatcctatgg 2340

ttttcagcca accaatggcg tgggctacca accataccgc gtcgtggtgc tctcctttga 2400

actgctccac gctcccgcga ctgtctgcgg ccccaagaag tccacgaacc ttgtgaagaa 2460

taagtgcgtt aattttaatt tcaacggcct cactggaaca ggagtgctca ctgagagtaa 2520

caagaagttc ctgccatttc aacaatttgg cagagacata gccgatacta ctgacgccgt 2580

tagggacccc cagaccctcg agattctcga tataacgccc tgctccttcg gtggagtttc 2640

cgtgatcacg ccaggcacca ataccagtaa ccaggtcgcc gtgctgtatc aggatgtcaa 2700

ctgtactgag gtgcccgtag ccatccatgc ggatcagctc acaccaactt ggagggtgta 2760

cagcaccggc tccaatgtat tccagactcg ggccggatgc cttattggcg ccgaacacgt 2820

gaacaatagt tacgaatgcg atattccaat tggcgccgga atctgtgcta gctaccagac 2880

tcagacgaac tccccaggca gcgccagcag cgttgccagc cagtcaatca tcgcttatac 2940

aatgtcactt ggagccgaaa actccgtggc ttactcaaac aacagcatcg ccatccccac 3000

aaacttcacc atatccgtga caactgagat tctgccagtg tccatgacta agacgtccgt 3060

agattgcact atgtacatat gcggcgacag cacagaatgt tctaatctgc tgctgcaata 3120

tggaagcttc tgcactcaac tgaacagagc gctcacaggc atcgccgtgg agcaggataa 3180

gaatacccag gaggtgttcg cccaagttaa gcagatctac aagaccccac ccataaagga 3240

tttcggtgga ttcaatttta gtcagatact cccagaccca tctaagccat ccaagaggag 3300

ctttatcgag gatcttttgt ttaacaaagt tactctggcc gacgccggtt tcatcaagca 3360

gtacggagat tgcctcggcg acatcgctgc tcgtgacctc atctgtgcgc aaaagtttaa 3420

cggtctgacg gtgctgcctc ccctccttac tgatgaaatg atcgcccagt ataccagcgc 3480

actcctcgct ggcaccataa catccggttg gacattcggc gctggtgcag cactgcagat 3540

accattcgcc atgcaaatgg catatcgttt caacggtatc ggtgtcacac agaatgtcct 3600

atatgagaac cagaagctga tcgcaaatca gttcaatagt gccatcggaa aaatccagga 3660

tagccttagc agcacagcct cagcccttgg caaactccag gatgtcgtga accagaatgc 3720

ccaggctctc aataccctcg tgaagcagct ctcatctaat ttcggcgcaa tttccagtgt 3780

cctcaacgac atcctcagcc gcctcgaccc ccccgaggcc gaagtgcaga ttgacagact 3840

gattacaggt cgactccaga gcctccagac ttacgtgact cagcagctga taagagccgc 3900

cgagataagg gccagcgcta acctggctgc cacaaagatg tctgagtgcg tgctgggcca 3960

gtccaagaga gtagacttct gtggcaaagg ctaccatctg atgagcttcc cacaatccgc 4020

acctcacggc gtagtgttcc tccacgtgac atatgtaccg gctcaggaga agaatttcac 4080

taccgctcct gctatatgcc atgatggaaa ggctcacttc ccccgggagg gggtgttcgt 4140

gtccaacggc acccattggt ttgtgactca gcggaatttc tacgaacccc agatcataac 4200

cactgacaac acatttgtgt ccggaaattg tgacgtggtc attggaatag tgaacaacac 4260

tgtttatgat ccactgcagc cagaacttga cagctttaag gaggagctcg acaagtactt 4320

caagaatcat acgtcaccag atgtggacct cggagatatt agcggtatca atgccagtgt 4380

tgtcaatatt cagaaggaaa tagaccgcct taatgaggtc gccaaaaatc tgaacgagag 4440

cctcatcgat cttcaggagc tgggcaaata tgagcagtac atcaagtggc cttggtatat 4500

ttggcttggc ttcatcgccg gcctgatcgc catagtaatg gtcacaatta tgctctgctg 4560

catgacctct tgctgctcct gtctgaaagg ctgctgctct tgcggatcct gctgcaaatt 4620

tgatgaggat gacagtgaac cagtcctgaa gggcgtgaag ctgcactata cttagaggcc 4680

ttagtcgtgt cgtttttcaa ataatataat ccttttaggg ttttagttag tttaaatttt 4740

ctgttgctcc tgtttagcag gtcgtgcctt cagcaagcac acaaaaacag agtgtttatt 4800

ttaagttgtt tgtttagtga ttcaaaaaaa aaatcgttca aacatttggc aataaagttt 4860

cttaagattg aatcctgttg ccggtcttgc gatgattatc atataatttc tgttgaatta 4920
cgttaagcat gtaataatta acatgtaatg catgacgtta tttatgagat gggtttttat 4980
gattagagtc ccgcaattat acatttaata cgcgatagaa aacaaaatat agcgcgcaaa 5040
ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta gat 5083

<210> 33
<211> 4461
<212> DNA
<213> Artificial Sequence
<220><223> Cloning vector 8500 from left to right T-DNA
<400> 33

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120
aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300
gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360
aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420
aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540
aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600
atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660
taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720
tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780
gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840
acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960
accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080
aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg cacttttcta atcaatcatc aaacagaacg cagaaaattt cctaaaaaca 2940

aaaaaaaggc atacaaatgg cgaaaaacgt tgcgattttc ggcttattgt tttctcttct 3000

tgtgttggtt ccttctcaga tcttcgcgac gtcactcctc agccaaaacg acacccccat 3060
ctgtctatcc actggcccct ggatctgctg cccaaactaa ctccatggtg accctgggat 3120
gcctggtcaa gggctatttc cctgagccag tgacagtgac ctggaactct ggatccctgt 3180
ccagcggtgt gcacaccttc ccagctgtcc tgcagtctga cctctacact ctgagcagct 3240
cagtgactgt cccctccagc acctggccca gcgagaccgt cacctgcaac gttgcccacc 3300
cggccagcag caccaaggtg gacaagaaaa ttgtgcccag ggattgtggt tgtaagcctt 3360
gcatatgtac agtcccagaa gtatcatctg tcttcatctt cccccccaaag cccaaggatg 3420

tgctcaccat tactctgact cctaaggtca cgtgtgttgt ggtagacatc agcaaggatg 3480
atcccgaggt ccagttcagc tggtttgtag atgatgtgga ggtgcacaca gctcagacgc 3540
aaccccggga ggagcagttc aacagcactt tccgctcagt cagtgaactt cccatcatgc 3600
accaggactg gctcaatggc aaggagacgt ccagattttg gcgatctatt caactgtcgc 3660
cagttcattg gtactggtag tctccctggg ggcaatcagt ttctggatgt gctctaatgg 3720
gtctctacag tgtagaatat gtatttaaag gccttagtcg tgtcgttttt caaataatat 3780
aatcctttta gggttttagt tagtttaaat tttctgttgc tcctgtttag caggtcgtgc 3840

cttcagcaag cacacaaaaa cagagtgttt attttaagtt gtttgtttag tgattcaaaa 3900
aaaaaatcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct 3960
tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta 4020
atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta 4080
atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc 4140
atctatgtta ctagatctct agagtctcaa gcttggcgcg cccacgtgac tagtggcact 4200
ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct 4260

tgcagcacat ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc 4320
ttcccaacag ttgcgcagcc tgaatggcga atgctagagc agcttgagct tggatcagat 4380
tgtcgtttcc cgccttcagt ttaaactatc agtgtttgac aggatatatt ggcgggtaaa 4440
cctaagagaa aagagcgttt a 4461

<210> 34
<211> 5070
<212> DNA
<213> Artificial Sequence
<220><223> Construct 8589 from 2X35S prom to NOS term

<400> 34

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggcac ttttctaatc aatcatcaaa cagaacgcag 780

aaaatttcct aaaaacaaaa aaaaggcata caaatggcga aaaacgttgc gattttcggc 840

ttattgtttt ctcttcttgt gttggttcct tctcagatct tcgcggtgaa tcttacgacg 900

cgaacacagt taccacccgc atatacaaat agcttcactc ggggtgttta ttaccccgac 960

aaagtgttca ggtcctccgt gctccactca acacaggacc tctttcttcc tttcttttct 1020

aacgtgacat ggtttcatgc cattcatgta tccggcacta acggtactaa gaggttcgat 1080

aatcctgtgc tccctttcaa tgacggcgtt tactttgcaa gcacagagaa gagtaacatc 1140

atccgaggtt ggatctttgg cactaccctc gattcaaaga cgcagagcct cctcattgtg 1200

aacaatgcca ctaacgtggt gatcaaagtt tgcgagtttc agttctgcaa tgacccttc 1260

ttgggggtgt actatcataa gaacaacaag tcttggatgg aatctgaatt ccgcgtctat 1320

agcagcgcca acaactgcac ctttgaatac gtgtcccagc ccttccttat ggacctggag 1380

ggaaagcagg gaaactttaa gaatctgaga gagttcgtgt ttaaaaatat cgacggctat 1440

tttaagatct attctaagca cacgcctatt aatctcgtgc gcgatcttcc acaaggcttc 1500

agcgccctgg aaccactcgt ggacctccca attggtatca acatcactag atttcagact 1560

ctgcttgccc tccaccgatc ctatctgaca cccggagact cctctagcgg ctggactgcc 1620

ggcgctgccg cttattacgt tggttatctt cagccacgca cgttcctgct gaagtataac 1680

gagaatggta ctattaccga tgccgtggat tgtgcccttg acccccctgtc cgaaactaag 1740

tgcacactca agtcattcac tgtggaaaaa ggaatctacc agacaagcaa ttttcgggtc 1800
cagcctactg agagcattgt gcgctttcct aacatcacaa atctttgccc cttcggagag 1860
gttttcaatg ctacacggtt tgcctccgtg tatgcctgga accgcaagag aatttccaat 1920
tgcgtggccg attactccgt gctctacaat agtgcaagct ttagcacctt taagtgctat 1980
ggcgtatccc ctactaagct taacgacttg tgtttcacaa acgtgtatgc cgactccttt 2040
gtgatacggg gcgacgaagt tagacagata gcaccaggac agacgggaaa gatagctgac 2100
tacaactata agcttcctga tgacttcact ggctgcgtta tcgcgtggaa ttctaacaac 2160

ctggactcaa aagtcggcgg caactataac tatctctatc ggctgttccg caagagtaac 2220
cttaagccct ttgagagaga tataagcact gaaatctacc aggctggcag tacgccctgt 2280
aatggcgtgg aaggctttaa ttgttatttt ccactgcaat cctatggttt tcagccaacc 2340
aatggcgtgg gctaccaacc ataccgcgtc gtggtgctct cctttgaact gctccacgct 2400
cccgcgactg tctgcggccc caagaagtcc acgaaccttg tgaagaataa gtgcgttaat 2460
tttaatttca acggcctcac tggaacagga gtgctcactg agagtaacaa gaagttcctg 2520
ccatttcaac aatttggcag agacatagcc gatactactg acgccgttag ggacccccag 2580

accctcgaga ttctcgatat aacgccctgc tccttcggtg gagtttccgt gatcacgcca 2640
ggcaccaata ccagtaacca ggtcgccgtg ctgtatcagg atgtcaactg tactgaggtg 2700
cccgtagcca tccatgcgga tcagctcaca ccaacttgga gggtgtacag caccggctcc 2760
aatgtattcc agactcgggc cggatgcctt attggcgccg aacacgtgaa caatagttac 2820
gaatgcgata ttccaattgg cgccggaatc tgtgctagct accagactca gacgaactcc 2880
ccaggcagcg ccagcagcgt tgccagccag tcaatcatcg cttatacaat gtcacttgga 2940
gccgaaaact ccgtggctta ctcaaacaac agcatcgcca tccccacaaa cttcaccata 3000

tccgtgacaa ctgagattct gccagtgtcc atgactaaga cgtccgtaga ttgcactatg 3060
tacatatgcg gcgacagcac agaatgttct aatctgctgc tgcaatatgg aagcttctgc 3120
actcaactga acagagcgct cacaggcatc gccgtggagc aggataagaa tacccaggag 3180
gtgttcgccc aagttaagca gatctacaag accccaccca taaaggattt cggtggattc 3240
aattttagtc agatactccc agacccatct aagccatcca agaggagctt tatcgaggat 3300
cttttgttta acaaagttac tctggccgac gccggtttca tcaagcagta cggagattgc 3360
ctcggcgaca tcgctgctcg tgacctcatc tgtgcgcaaa agtttaacgg tctgacggtg 3420

ctgcctcccc tccttactga tgaaatgatc gcccagtata ccagcgcact cctcgctggc 3480

accataacat ccggttggac attcggcgct ggtgcagcac tgcagatacc attcgccatg 3540

caaatggcat atcgtttcaa cggtatcggt gtcacacaga atgtcctata tgagaaccag 3600

aagctgatcg caaatcagtt caatagtgcc atcggaaaaa tccaggatag ccttagcagc 3660

acagcctcag cccttggcaa actccaggat gtcgtgaacc agaatgccca ggctctcaat 3720

accctcgtga agcagctctc atctaatttc ggcgcaattt ccagtgtcct caacgacatc 3780

ctcagccgcc tcgacccccc cgaggccgaa gtgcagattg acagactgat tacaggtcga 3840

ctccagagcc tccagactta cgtgactcag cagctgataa gagccgccga gataagggcc 3900

agcgctaacc tggctgccac aaagatgtct gagtgcgtgc tgggccagtc caagagagta 3960

gacttctgtg gcaaaggcta ccatctgatg agcttcccac aatccgcacc tcacggcgta 4020

gtgttcctcc acgtgacata tgtaccggct caggagaaga atttcactac cgctcctgct 4080

atatgccatg atggaaaggc tcacttcccc cgggaggggg tgttcgtgtc caacggcacc 4140

cattggtttg tgactcagcg gaatttctac gaaccccaga tcataaccac tgacaacaca 4200

tttgtgtccg gaaattgtga cgtggtcatt ggaatagtga acaacactgt ttatgatcca 4260

ctgcagccag aacttgacag ctttaaggag gagctcgaca agtacttcaa gaatcatacg 4320

tcaccagatg tggacctcgg agatattagc ggtatcaatg ccagtgttgt caatattcag 4380

aaggaaatag accgccttaa tgaggtcgcc aaaaatctga acgagagcct catcgatctt 4440

caggagctgg gcaaatatga gcagtacatc aagtggcctt ggtatatttg gcttggcttc 4500

atcgccggcc tgatcgccat agtaatggtc acaattatgc tctgctgcat gacctcttgc 4560

tgctcctgtc tgaaaggctg ctgctcttgc ggatcctgct gcaaatttga tgaggatgac 4620

agtgaaccag tcctgaaggg cgtgaagctg cactatactt agaggcctta gtcgtgtcgt 4680

ttttcaaata atataatcct tttagggttt tagttagttt aaattttctg ttgctcctgt 4740

ttagcaggtc gtgccttcag caagcacaca aaaacagagt gtttatttta agttgtttgt 4800

ttagtgattc aaaaaaaaaa tcgttcaaac atttggcaat aaagtttctt aagattgaat 4860

cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta 4920

ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg 4980

caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta 5040

tcgcgcgcgg tgtcatctat gttactagat 5070

<210> 35

<211> 4436

<212> DNA

<213> Artificial Sequence

<220><223> Cloning vector 8716 from left to right T-DNA

<400> 35

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg cactccattt gaatctatca aaccaaaaca cattgagcaa aatggcgaaa 2940

aacgttgcga ttttcggctt attgttttct cttcttgtgt tggttccttc tcagatcttc 3000

gcgacgtcac tcctcagcca aaacgacacc cccatctgtc tatccactgg cccctggatc 3060

tgctgcccaa actaactcca tggtgaccct gggatgcctg gtcaagggct atttccctga 3120

gccagtgaca gtgacctgga actctggatc cctgtccagc ggtgtgcaca ccttcccagc 3180

tgtcctgcag tctgacctct acactctgag cagctcagtg actgtcccct ccagcacctg 3240

gcccagcgag accgtcacct gcaacgttgc ccacccggcc agcagcacca aggtggacaa 3300

gaaaattgtg cccagggatt gtggttgtaa gccttgcata tgtacagtcc cagaagtatc 3360

atctgtcttc atcttccccc caaagcccaa ggatgtgctc accattactc tgactcctaa 3420

ggtcacgtgt gttgtggtag acatcagcaa ggatgatccc gaggtccagt tcagctggtt 3480

tgtagatgat gtggaggtgc acacagctca gacgcaaccc cgggaggagc agttcaacag 3540

cactttccgc tcagtcagtg aacttcccat catgcaccag gactggctca atggcaagga 3600

gacgtccaga ttttggcgat ctattcaact gtcgccagtt cattggtact ggtagtctcc 3660

ctgggggcaa tcagtttctg gatgtgctct aatgggtctc tacagtgtag aatatgtatt 3720

taaaggcctt agtcgtgtcg tttttcaaat aatataatcc ttttagggtt ttagttagtt 3780

taaatttttct gttgctcctg tttagcaggt cgtgccttca gcaagcacac aaaaacagag 3840

tgtttatttt aagttgtttg tttagtgatt caaaaaaaaa atcgttcaaa catttggcaa 3900

taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg 3960

ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg 4020

gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag 4080

cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tctctagagt 4140

ctcaagcttg gcgcgcccac gtgactagtg gcactggccg tcgttttaca acgtcgtgac 4200

tgggaaaacc ctggcgttac ccaacttaat cgccttgcag cacatccccc tttcgccagc 4260

tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg cagcctgaat 4320

ggcgaatgct agagcagctt gagcttggat cagattgtcg tttcccgcct tcagtttaaa 4380

ctatcagtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag cgttta 4436

<210> 36

<211> 5045

<212> DNA

<213> Artificial Sequence

<220><223> Construct 8591 from 2X35S prom to NOS term

<400> 36

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggcac tccatttgaa tctatcaaac caaaacacat 780

tgagcaaaat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt cttgtgttgg 840

ttccttctca gatcttcgcg gtgaatctta cgacgcgaac acagttacca cccgcatata 900

caaatagctt cactcggggt gtttattacc ccgacaaagt gttcaggtcc tccgtgctcc 960

actcaacaca ggacctcttt cttcctttct tttctaacgt gacatggttt catgccattc 1020

atgtatccgg cactaacggt actaagaggt tcgataatcc tgtgctccct ttcaatgacg 1080

gcgtttactt tgcaagcaca gagaagagta acatcatccg aggttggatc tttggcacta 1140

ccctcgattc aaagacgcag agcctcctca ttgtgaacaa tgccactaac gtggtgatca 1200

aagtttgcga gtttcagttc tgcaatgacc ctttcttggg ggtgtactat cataagaaca 1260

acaagtcttg gatggaatct gaattccgcg tctatagcag cgccaacaac tgcacctttg 1320

aatacgtgtc ccagcccttc cttatggacc tggagggaaa gcagggaaac tttaagaatc 1380

tgagagagtt cgtgtttaaa aatatcgacg gctattttaa gatctattct aagcacacgc 1440

ctattaatct cgtgcgcgat cttccacaag gcttcagcgc cctggaacca ctcgtggacc 1500

tcccaattgg tatcaacatc actagatttc agactctgct tgccctccac cgatcctatc 1560

tgacacccgg agactcctct agcggctgga ctgccggcgc tgccgcttat tacgttggtt 1620

atcttcagcc acgcacgttc ctgctgaagt ataacgagaa tggtactatt accgatgccg 1680

tggattgtgc ccttgacccc ctgtccgaaa ctaagtgcac actcaagtca ttcactgtgg 1740

aaaaaggaat ctaccagaca agcaattttc gggtccagcc tactgagagc attgtgcgct 1800

ttcctaacat cacaaatctt tgccccttcg gagaggtttt caatgctaca cggtttgcct 1860

ccgtgtatgc ctggaaccgc aagagaattt ccaattgcgt ggccgattac tccgtgctct 1920

acaatagtgc aagctttagc acctttaagt gctatggcgt atcccctact aagcttaacg 1980

acttgtgttt cacaaacgtg tatgccgact cctttgtgat acggggcgac gaagttagac 2040

agatagcacc aggacagacg ggaaagatag ctgactacaa ctataagctt cctgatgact 2100

tcactggctg cgttatcgcg tggaattcta acaacctgga ctcaaaagtc ggcggcaact 2160

ataactatct ctatcggctg ttccgcaaga gtaaccttaa gcccctttgag agagatataa 2220

gcactgaaat ctaccaggct ggcagtacgc cctgtaatgg cgtggaaggc tttaattgtt 2280

attttccact gcaatcctat ggttttcagc caaccaatgg cgtgggctac caaccatacc 2340

gcgtcgtggt gctctccttt gaactgctcc acgctcccgc gactgtctgc ggccccaaga 2400

agtccacgaa ccttgtgaag aataagtgcg ttaattttaa tttcaacggc ctcactggaa 2460

caggagtgct cactgagagt aacaagaagt tcctgccatt tcaacaattt ggcagagaca 2520
tagccgatac tactgacgcc gttagggacc cccagaccct cgagattctc gatataacgc 2580
cctgctcctt cggtggagtt tccgtgatca cgccaggcac caataccagt aaccaggtcg 2640
ccgtgctgta tcaggatgtc aactgtactg aggtgcccgt agccatccat gcggatcagc 2700
tcacaccaac ttggagggtg tacagcaccg gctccaatgt attccagact cgggccggat 2760
gccttattgg cgccgaacac gtgaacaata gttacgaatg cgatattcca attggcgccg 2820
gaatctgtgc tagctaccag actcagacga actccccagg cagcgccagc agcgttgcca 2880

gccagtcaat catcgcttat acaatgtcac ttggagccga aaactccgtg gcttactcaa 2940
acaacagcat cgccatcccc acaaacttca ccatatccgt gacaactgag attctgccag 3000
tgtccatgac taagacgtcc gtagattgca ctatgtacat atgcggcgac agcacagaat 3060
gttctaatct gctgctgcaa tatggaagct tctgcactca actgaacaga gcgctcacag 3120
gcatcgccgt ggagcaggat aagaataccc aggaggtgtt cgcccaagtt aagcagatct 3180
acaagacccc acccataaag gatttcggtg gattcaattt tagtcagata ctcccagacc 3240
catctaagcc atccaagagg agctttatcg aggatctttt gtttaacaaa gttactctgg 3300

ccgacgccgg tttcatcaag cagtacggag attgcctcgg cgacatcgct gctcgtgacc 3360
tcatctgtgc gcaaaagttt aacggtctga cggtgctgcc tcccctcctt actgatgaaa 3420
tgatcgccca gtataccagc gcactcctcg ctggcaccat aacatccggt tggacattcg 3480
gcgctggtgc agcactgcag ataccattcg ccatgcaaat ggcatatcgt ttcaacggta 3540
tcggtgtcac acagaatgtc ctatatgaga accagaagct gatcgcaaat cagttcaata 3600
gtgccatcgg aaaaatccag gatagcctta gcagcacagc ctcagccctt ggcaaactcc 3660
aggatgtcgt gaaccagaat gcccaggctc tcaataccct cgtgaagcag ctctcatcta 3720

atttcggcgc aatttccagt gtcctcaacg acatcctcag ccgcctcgac ccccccgagg 3780
ccgaagtgca gattgacaga ctgattacag gtcgactcca gagcctccag acttacgtga 3840
ctcagcagct gataagagcc gccgagataa gggccagcgc taacctggct gccacaaaga 3900
tgtctgagtg cgtgctgggc cagtccaaga gagtagactt ctgtggcaaa ggctaccatc 3960
tgatgagctt cccacaatcc gcacctcacg gcgtagtgtt cctccacgtg acatatgtac 4020
cggctcagga gaagaatttc actaccgctc ctgctatatg ccatgatgga aaggctcact 4080
tcccccggga gggggtgttc gtgtccaacg gcacccattg gtttgtgact cagcggaatt 4140

tctacgaacc ccagatcata accactgaca acacatttgt gtccggaaat tgtgacgtgg 4200

tcattggaat agtgaacaac actgtttatg atccactgca gccagaactt gacagcttta 4260

aggaggagct cgacaagtac ttcaagaatc atacgtcacc agatgtggac ctcggagata 4320

ttagcggtat caatgccagt gttgtcaata ttcagaagga aatagaccgc cttaatgagg 4380

tcgccaaaaa tctgaacgag agcctcatcg atcttcagga gctgggcaaa tatgagcagt 4440

acatcaagtg gccttggtat atttggcttg gcttcatcgc cggcctgatc gccatagtaa 4500

tggtcacaat tatgctctgc tgcatgacct cttgctgctc ctgtctgaaa ggctgctgct 4560

cttgcggatc ctgctgcaaa tttgatgagg atgacagtga accagtcctg aagggcgtga 4620

agctgcacta tacttagagg ccttagtcgt gtcgtttttc aaataatata atccttttag 4680

ggttttagtt agtttaaatt ttctgttgct cctgtttagc aggtcgtgcc ttcagcaagc 4740

acacaaaaac agagtgttta ttttaagttg tttgtttagt gattcaaaaa aaaaatcgtt 4800

caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt gcgatgatta 4860

tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcatgacgt 4920

tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag 4980

aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac 5040

tagat 5045

<210> 37

<211> 39

<212> PRT

<213> Artificial Sequence

<220><223> C-Terminal Region of Modified SARS-CoV-2 S protein with H5i Hemagglutinin CT, Variation 2

<400> 37

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

1 5 10 15

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

20 25 30

Leu Gln Cys Arg Ile Cys Ile

35

<210> 38

<211> 36

<212> PRT

<213> Artificial Sequence

<220><223> C-Terminal Region of Modified SARS-CoV-2 S protein with H5i Hemagglutinin CT, Variation 3

<400> 38

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met
1 5 10 15

Val Thr Ile Met Leu Cys Cys Met Cys Ser Asn Gly Ser Leu Gln Cys

20 25 30

Arg Ile Cys Ile

35

<210> 39

<211> 34

<212> PRT

<213> Artificial Sequence

<220><223> C-Terminal Region of Modified SARS-CoV-2 S protein with H5i Hemagglutinin CT, Variation 4

<400> 39

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met
1 5 10 15

Val Thr Ile Met Leu Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

20 25 30

Cys Ile

<210> 40

<211> 151

<212> DNA

<213> Arracacha Virus B

<400> 40

tagtcgtgtc gtttttcaaa taatataatc cttttagggt tttagttagt ttaaattttc 60

tgttgctcct gtttagcagg tcgtgccttc agcaagcaca caaaaacaga gtgtttattt 120

taagttgttt gtttagtgat tcaaaaaaaa a 151

<210> 41

<211> 182

<212> DNA

<213> Beet necrotic yellow vein virus

<400> 41

cctatcttga tgaaggttgt tgtggttttc tcattactgt tttattattg tttgagttgc 60

ttatgtcggt tcttgattat gtggtgcata attattgaac taattgtttg ttgggttgta 120

atgtactgac tgggtgtgaa ttgtaccagt cgttaaaggg tttactatca gtatattgat 180

at 182

<210> 42

<211> 132

<212> DNA

<213> Southern bean mosaic virus

<400> 42

tgaggagttg tataataata cctgcaccct tctctttggc agggagggtg tttcgttttc 60

acaatgccac gcgcttgagg gagaatgcac gttaatcatc cctccgctag tgatggagcg 120

taatccaaaa gt 132

<210> 43

<211> 282

<212> DNA

<213

> Turnip ringspot virus

<400> 43

tgatttataa tagccataga ttaagtttaa atgtattacg tttgtatttt attctctttt 60

ttaagtttcc tatgttgttt taaattaaat atctgtataa ttagtagatg taaatctgct 120

ttgtgcgttt gacagtctgt ggaaacgcac tggttcatga gataggacca cctaggaggt 180

aggactctgg gttttaatta tctcatttct tagttatacc gtattatata tatgatttag 240

tagtaattgt tttctcttga tatgtattat tacttttta tt 282

<210> 44

<211> 165

<212> DNA

<213> Cowpea Mosaic virus

<400> 44

attttcttta gtttgaattt actgttattc ggtgtgcatt tctatgtttg gtgagcggtt 60

ttctgtgctc agagtgtgtt tattttatgt aatttaattt ctttgtgagc tcctgtttag 120

caggtcgtcc cttcagcaag gacacaaaaa gattttaatt ttatt 165

<210> 45

<211> 275

<212> DNA

<213> Broad bean true mosaic virus

<400> 45

tagttttctt ccgcttttct tttgtagtgt gtggttttct ttgtttcttc ttttcttttc 60

tctttccttt tctcttactc ctgcctggca ggtcgtgcct tcagtaagca caacaaaaat 120

atgcatttat tagagtattt ctttcttctt tagcataaag gtattgaaga cctataaact 180

tcgtccgggt tggggaaagt accagcttag catatcttta gaaaactata tagagctctt 240

taccttgagt tgtttcctaa agtttatgca aaaaa 275

<210> 46

<211> 147

<212> DNA

<213> Ourmia melon virus

<400> 46

ctcacgtctg gggtgagccc tagccaaata ggaaaacgat aagcgctttg catgcaaaat 60

gagttgggcc acaagtgcca ctcgcagcga aggcggtctg aggtttcccc ctggcggtta 120

cttccatatc tttgggagat aactggg 147

<210> 47

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H5i

Hemagglutinin CT AA

<400> 47

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

20 25 30

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

35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe

50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80
Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95
Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110
Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140
Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160
Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175
Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205
Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220
Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240
Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270
Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285
Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300

Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335
Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350
Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380

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

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

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540
Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
805 810 815
Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
820 825 830

Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845
Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880
Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925
Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
930 935 940
Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960

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

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035
Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys

1235 1240 1245

Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 48

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+4P with H5i Hemagglutinin CT AA

<400> 48

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140

Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp

210 215 220

Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285

Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300

Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365

Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380

Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr
420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460

Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480
Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn
485 490 495

Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510
Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540
Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560

Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590
Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr
595 600 605
Pro Gly Thr Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val
610 615 620

Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640
Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655
Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp
660 665 670
Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685

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

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Pro Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035
Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050
Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080
Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125

Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1145 1150 1155
Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1160 1165 1170
Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys

1235 1240 1245

Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 49

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+6P with H5i Hemagglutinin CT AA

<400> 49

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

20 25 30

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

35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe

50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser

65 70 75 80

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

85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile

115 120 125

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

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

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

Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320
Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335
Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser

370 375 380

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

385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val

405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640
Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655
Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp
660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685
Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser
705 710 715 720
Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu
725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys
740 745 750
Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe
755 760 765
Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp
770 775 780
Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800

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

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Pro Ala Leu Gln Ile Pro Phe Pro Met Gln Met Ala

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125
Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1145 1150 1155
Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1160 1165 1170

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

Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys
1235 1240 1245
Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1250 1255

<210> 50
<211> 1259
<212> PRT
<213> Artificial Sequence
<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP+923 with H5i Hemagglutinin CT AA
<400> 50
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe

50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser

65 70 75 80

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

85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile

115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe

130 135 140

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

145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr

165 170 175

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

180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly

195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp

210 215 220

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

225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser

245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala

260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr

275 280 285

Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320
Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350
Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380
Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415
Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr
420 425 430
Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445
Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460

Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480
Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn
485 490 495
Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510
Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
805 810 815
Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
820 825 830
Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845

Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880
Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895
Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910

Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925
Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
930 935 940
Asp Ser Phe Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
980 985 990
Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly
995 1000 1005
Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050
Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065
Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080
Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095

Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125
Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1145 1150 1155

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

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys
1235 1240 1245
Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1250 1255

<210> 51

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+4P+923 with H5i Hemagglutinin CT AA

<400> 51

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140

Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220

Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285

Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300

Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365

Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380

Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr
420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

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

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

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

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655
Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp
660 665 670
Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700

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

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp
770 775 780
Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
805 810 815
Asp Pro Ser Lys Pro Ser Lys Arg Ser Pro Ile Glu Asp Leu Leu Phe
820 825 830

Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845
Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880
Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925
Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
1175 1180 1185
Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr
1205 1210 1215
Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys
1235 1240 1245
Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1250 1255

<210> 52
<211> 1259

<212> PRT
<213> Artificial Sequence
<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+6P+923 with H5i Hemagglutinin CT AA
<400> 52
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15
Leu Val Pro Ser Gln Ile Phe Ala Val Asn Leu Thr Thr Arg Thr Gln
20 25 30
Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60
Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80
Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95
Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125
Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140
Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160
Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190
Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205
Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220
Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240

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

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

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365

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

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445
Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460
Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480
Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn
485 490 495

Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510
Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540
Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560

Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590
Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Pro Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880

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

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

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035
Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050
Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Leu Trp Met Cys

1235 1240 1245

Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 53

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H1 Cal Hemagglutinin CT AA

<400> 53

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45
Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60
Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80
Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110
Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125
Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140
Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

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

Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240
Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255
Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala

260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr

275 280 285

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

290 295 300

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

305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val

325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn

340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser

355 360 365

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

370 375 380

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

385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val

405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540

Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560
Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590
Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr
595 600 605

Pro Gly Thr Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val
610 615 620
Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640
Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655
Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp
660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685
Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser
705 710 715 720
Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu
725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys
740 745 750
Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

```
        755                 760                 765
Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp
    770                 775                 780
Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785                 790                 795                 800

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

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

Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
    930                 935                 940
Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945                 950                 955                 960
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
                965                 970                 975
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
            980                 985                 990

Leu Asp Pro Pro Glu Ala Glu Val  Gln Ile Asp Arg Leu  Ile Thr Gly
        995                 1000                1005
```

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

Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065
Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080
Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125
Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1145 1150 1155
Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1160 1165 1170

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

Ile Ala Ile Val Met Val Thr Ile Met Leu Ser Phe Trp Met Cys

1235 1240 1245

Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 54

<211> 1258

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H3 Minn Hemagglutinin CT AA

<400> 54

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

20 25 30

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

35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe

50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser

65 70 75 80

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

85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile

115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe

130 135 140

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

145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr

165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190
Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220
Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240
Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255
Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285
Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320
Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350
Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380
Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415
Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685
Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser
705 710 715 720

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

Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
805 810 815
Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
820 825 830
Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845

Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880
Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895
Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Leu Met Trp Ala Cys Gln
1235 1240 1245
Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
1250 1255

<210> 55
<211> 1259
<212> PRT
<213> Artificial Sequence
<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H6 HK Hemagglutinin CT AA

<400> 55
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15
Leu Val Pro Ser Gln Ile Phe Ala Val Asn Leu Thr Thr Arg Thr Gln
20 25 30
Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45
Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

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

85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile

115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe

130 135 140

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

145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr

165 170 175

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

180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly

195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp

210 215 220

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

225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser

245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala

260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr

275 280 285

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

290 295 300

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

305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val

325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365

Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380

Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr
420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460

Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn
485 490 495

Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525

Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540

Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560

Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845

Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880

Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910

Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925

Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
930 935 940

Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960

Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
980 985 990

Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly
995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035

Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050

Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125
Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140

Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1145 1150 1155
Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1160 1165 1170
Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
1175 1180 1185
Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr
1205 1210 1215
Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Leu Gly Leu Trp Met Cys
1235 1240 1245
Ser Asn Gly Ser Met Gln Cys Arg Ile Cys Ile
1250 1255

<210> 56

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H7 Guangdong Hemagglutinin CT AA

<400> 56

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

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

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140

Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220

Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240

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

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

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

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445
Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460
Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480
Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510
Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540
Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560

Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590
Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr
595 600 605
Pro Gly Thr Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val
610 615 620

Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640
Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala
645 650 655
Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp
660 665 670
Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn
675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr
690 695 700
Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser
705 710 715 720
Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu
725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys
740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe
755 760 765
Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp
770 775 780
Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
820 825 830
Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845
Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880

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

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

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Leu Val Phe Ile Cys Val
1235 1240 1245

Lys Asn Gly Asn Met Arg Cys Thr Ile Cys Ile
1250 1255

<210> 57
<211> 1259
<212> PRT
<213> Artificial Sequence
<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H9 HK Hemagglutinin CT AA
<400> 57
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15
Leu Val Pro Ser Gln Ile Phe Ala Val Asn Leu Thr Thr Arg Thr Gln
20 25 30

Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro
35 40 45
Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60
Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80
Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

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

145 150 155 160

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

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

Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320
Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335
Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380
Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr
420 425 430
Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445
Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460
Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn
485 490 495
Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510
Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540

Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
545 550 555 560
Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590
Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr
595 600 605

Pro Gly Thr Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val
610 615 620
Asn Cys Thr Glu Val Pro Val Ala Ile His Ala Asp Gln Leu Thr Pro
625 630 635 640
Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925

Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
930 935 940
Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
965 970 975
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
980 985 990

Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly
995 1000 1005
Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035
Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050

Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065
Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080
Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125
Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

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

1235 1240 1245

Ser Asn Gly Ser Cys Arg Cys Asn Ile Cys Ile

1250 1255

<210> 58

<211> 1259

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with B/Wash Hemagglutinin CT AA

<400> 58

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

20 25 30

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

35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe

50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80

Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140

Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175

Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220

Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285

Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300

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

305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val

325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn

340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser

355 360 365

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

370 375 380

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

385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val

405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590

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

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

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

Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800
Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly
1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080

Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095

Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125

Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140

Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1145 1150 1155

Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1160 1165 1170

Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
1175 1180 1185

Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr
1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Val Val Tyr Met Val
1235 1240 1245

Ser Arg Asp Asn Val Ser Cys Ser Ile Cys Leu
1250 1255

<210> 59

<211> 1262

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H5i Hemagglutinin CT (alternative 1) AA

<400> 59

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

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80
Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95
Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly
100 105 110
Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe
130 135 140
Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser
145 150 155 160
Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr
165 170 175
Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln
180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly
195 200 205
Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp
210 215 220

Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240
Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270
Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr
275 280 285
Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro
290 295 300
Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys Gly
305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val
325 330 335
Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn
340 345 350
Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser
355 360 365
Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser
370 375 380

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

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460
Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu
725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys
740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe
755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp
770 775 780

Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr Lys Thr
785 790 795 800

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

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
820 825 830

Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp
835 840 845

Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala
865 870 875 880

Gln Tyr Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr
885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
900 905 910

Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
915 920 925

Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln
930 935 940

Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr
1205 1210 1215
Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1220 1225 1230
Ile Ala Ile Val Met Val Thr Ile Met Met Ala Gly Leu Ser Leu
1235 1240 1245
Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1250 1255 1260

<210> 60
<211> 1261
<212> PRT
<213> Artificial Sequence
<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H5i Hemagglutinin CT (alternative 2) AA
<400> 60
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15
Leu Val Pro Ser Gln Ile Phe Ala Val Asn Leu Thr Thr Arg Thr Gln
20 25 30
Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro

35 40 45
Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe
50 55 60
Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser
65 70 75 80
Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn
85 90 95
Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110
Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile
115 120 125
Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe

130 135 140

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

145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr

165 170 175

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

180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly

195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp

210 215 220

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

225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser

245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala

260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr

275 280 285

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

290 295 300

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

305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val

325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn

340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser

355 360 365

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

370 375 380

Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys
385 390 395 400
Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val
405 410 415
Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430
Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn
435 440 445
Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu
450 455 460
Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu
465 470 475 480
Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495
Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val
500 505 510
Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His
515 520 525
Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys
530 535 540
Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val

545 550 555 560
Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln Phe Gly Arg
565 570 575
Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro Gln Thr Leu Glu
580 585 590
Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr
595 600 605
Pro Gly Thr Asn Thr Ser Asn Gln Val Ala Val Leu Tyr Gln Asp Val

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

930 935 940
Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
965 970 975
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
980 985 990
Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly

995 1000 1005
Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
1010 1015 1020
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met
1025 1030 1035
Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly
1040 1045 1050
Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly

1055 1060 1065
Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn
1070 1075 1080
Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe
1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

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

1235 1240 1245

Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255 1260

<210> 61

<211> 1258

<212> PRT

<213> Artificial Sequence

<220><223> (PDI) Modified SARS-CoV-2 S protein GSAS+PP with H5i Hemagglutinin CT (alternative 3) AA

<400> 61

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

20 25 30

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

35 40 45

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

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

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

210 215 220
Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile
225 230 235 240
Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser
245 250 255
Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala
260 265 270
Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr

275 280 285
Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp Pro

290 295 300

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

305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val

325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn

340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser

355 360 365

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

370 375 380

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

385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val

405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

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

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

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg

980 985 990

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

995 1000 1005

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1010 1015 1020

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met Cys Ser

1235 1240 1245

Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 62

<211> 1256

<212> PRT

<213> Artificial Sequence

<220><223> PDI) Modified SARS-CoV-2 S protein GSAS+PP with H5i Hemagglutinin CT (alternative 4) AA

<400> 62

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

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

20 25 30

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

35 40 45

Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe

50 55 60

Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser

65 70 75 80

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

85 90 95

Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly

100 105 110

Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile

115 120 125

Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe

130 135 140

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

145 150 155 160

Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr

165 170 175

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

180 185 190

Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly

195 200 205

Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp

210 215 220

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

225 230 235 240

Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser

245 250 255

Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala

260 265 270

Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr

275 280 285

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

290 295 300

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

305 310 315 320

Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile Val

325 330 335

Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn

340 345 350

Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser

355 360 365

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

370 375 380

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

385 390 395 400

Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val

405 410 415

Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr

420 425 430

Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn

435 440 445

Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu

450 455 460

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

465 470 475 480

Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn

485 490 495

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

500 505 510

Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly Val Ser Val Ile Thr

595 600 605

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

610 615 620

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

625 630 635 640

Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn Val Phe Gln Thr Arg Ala

645 650 655

Gly Cys Leu Ile Gly Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp

660 665 670

Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn

675 680 685

Ser Pro Gly Ser Ala Ser Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr

690 695 700

Thr Met Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser

705 710 715 720

Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu

725 730 735

Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys

740 745 750

Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe

755 760 765

Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln Asp

770 775 780

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

785 790 795 800

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

805 810 815

Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe

820 825 830

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

835 840 845

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

850 855 860

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala

865 870 875 880

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

885 890 895

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

900 905 910

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

915 920 925

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

930 935 940

Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val
945 950 955 960
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser

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

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

1085 1090 1095
Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val
1100 1105 1110
Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn
1115 1120 1125
Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn
1130 1135 1140
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys

1145 1150 1155
Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1160 1165 1170
Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile

1175 1180 1185

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

1190 1195 1200

Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr

1205 1210 1215

Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1220 1225 1230

Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Ser Asn Gly

1235 1240 1245

Ser Leu Gln Cys Arg Ile Cys Ile

1250 1255

<210> 63

<211> 31

<212> PRT

<213> SARS-CoV-2

<400> 63

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

20 25 30

<210> 64

<211> 33

<212> PRT

<213> Artificial Sequence

<220><223> Modified TMCT with intervening peptide sequence (X)n

<220><221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa may be any combination of 0 to 10 amino acids

<220><221> misc_feature

<222> (27)..(28)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (30)..(30)

<223> Xaa can be any naturally occurring amino acid

<400> 64

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met
1 5 10 15

Val Thr Ile Met Xaa Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa Ile Cys
20 25 30

Ile

<210> 65

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S Protein GSAS+4P-DNA

<400> 65

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagccccat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca ccctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240
gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300
gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420
atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480
ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540
atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600
aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660
tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720
attatgctct ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa 3780

<210> 66
<211> 3780
<212> DNA
<213> Artificial Sequence
<220><223> PDI-S Protein GSAS+6P-DNA
<400> 66
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120
ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180
caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240
ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300
tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420
gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480
tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540
tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagccccat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
cccgcactgc agataccatt ccccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760
acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820
ggaaaaatcc aggatagcct tagcagcaca ccctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940
gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240
gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gagggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360
ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420
atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480
ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540
atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600
aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660
tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa 3780

<210> 67

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S Protein GSAS+2P+L923F-DNA

<400> 67

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280
ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340
gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460
ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760
acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagctt cagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880
gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940
gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300
gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360
ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420
atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480
ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540
atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600
aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa 3780

<210> 68

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S Protein GSAS+4P+L923F-DNA

<400> 68

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560
gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620
aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680
ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740
actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800
ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860
tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980
ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280
ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400
ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460
ccatccaaga ggagccccat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820
ggaaaaatcc aggatagctt cagcagcaca ccctcagccc ttggcaaact ccaggatgtc 2880
gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940
gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa 3780

<210> 69

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S Protein GSAS+6P+L923F-DNA

<400> 69

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080
gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140
gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200
ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260
ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320
tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380
ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500
ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560
gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620
aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680
ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740
actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800
ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920
acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980
ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340
gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400
ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460
ccatccaaga ggagccccat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

cccgcactgc agataccatt ccccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagctt cagcagcaca ccctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa 3780

<210> 70

<211> 3789

<212> DNA

<213> Artificial Sequence

<220><223> PDI-Modified S protein with H5i Hemagglutinin CT (V1) DNA

<400> 70

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgatgg ccggcctctc tttatggatg tgctccaatg gatcgttaca atgcagaatt 3780

tgcatttaa 3789

<210> 71

<211> 3786

<212> DNA

<213> Artificial Sequence

<220><223> PDI-Modified S protein with H5i Hemagglutinin CT (V2) DNA

<400> 71

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatggccg gcctctcttt atggatgtgc tccaatggat cgttacaatg cagaatttgc 3780

atttaa 3786

<210> 72

<211> 3777

<212> DNA

<213> Artificial Sequence

<220><223> PDI-Modified S protein with H5i Hemagglutinin CT (V3) DNA

<400> 72

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctct gctgcatgtg ctccaatgga tcgttacaat gcagaatttg catttaa 3777

<210> 73

<211> 3771

<212> DNA

<213> Artificial Sequence

<220><223> PDI-Modified S protein with H5i Hemagglutinin CT (V4) DNA

<400> 73

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760
acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820
ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880
gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240
gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300
gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420
atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480
ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540
atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600
aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660
tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720
attatgctct gctgctccaa tggatcgtta caatgcagaa tttgcattta a 3771

<210> 74
<211> 3780
<212> DNA
<213> Artificial Sequence
<220><223> PDI-S-protein+H1 Cal DNA
<400> 74
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120
ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280
ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340
gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400
ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760
acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820
ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940
gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240
gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360
ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420
atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480
ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540
atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600
aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660
tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctca gcttctggat gtgctctaat gggtctctac agtgtagaat atgtatttaa 3780

<210> 75

<211> 3777

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S-protein+H3 Minn DNA

<400> 75

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620
aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680
ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740
actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800
ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860
tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920
acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280
ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340
gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460
ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760
acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880
gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940
gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000
cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060
ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120
tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180
ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gagggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctca tgtgggcctg tcagaagggc aacatcagat gcaacatctg catctaa 3777

<210> 76

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S-protein+H6 HK DNA

<400> 76

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140
gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200
ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260
ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320
tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380
ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440
atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560
gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620
aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680
ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740
actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800
ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860
tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980
ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280
ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400
ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460
ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520
ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580
gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640
cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700
gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctcg gtctttggat gtgttcaaat ggttcaatgc agtgcaggat atgtatataa 3780

<210> 77

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S-protein+H7 Guangdong DNA

<400> 77

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660
ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720
ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780
ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840
ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900
gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960
atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080
gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140
gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200
ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260
ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320
tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380
ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500
ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560
gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620
aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680
ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740
actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800
ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920
acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980
ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040
gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100
atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160
atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220
actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctcg tcttcatatg tgtgaagaat ggaaacatgc ggtgcactat ttgtatataa 3780

<210> 78

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S-protein+H9 HK DNA

<400> 78

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180
caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240
ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300
tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360
tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420
gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480
tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600
ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660
ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720
ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780
ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840
ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900
gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020
atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080
gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140
gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200
ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260
ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320
tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440
atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500
ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560
gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620
aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680
ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740
actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gaggggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctcc tgttctgggc catgtccaat ggatcttgca gatgcaacat ttgtatataa 3780

<210> 79

<211> 3780

<212> DNA

<213> Artificial Sequence

<220><223> PDI-S-protein+ B/ Wash DNA

<400> 79

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgaatct tacgacgcga acacagttac cacccgcata tacaaatagc 120

ttcactcggg gtgtttatta ccccgacaaa gtgttcaggt cctccgtgct ccactcaaca 180

caggacctct ttcttccttt cttttctaac gtgacatggt ttcatgccat tcatgtatcc 240

ggcactaacg gtactaagag gttcgataat cctgtgctcc ctttcaatga cggcgtttac 300

tttgcaagca cagagaagag taacatcatc cgaggttgga tctttggcac taccctcgat 360

tcaaagacgc agagcctcct cattgtgaac aatgccacta acgtggtgat caaagtttgc 420

gagtttcagt tctgcaatga ccctttcttg ggggtgtact atcataagaa caacaagtct 480

tggatggaat ctgaattccg cgtctatagc agcgccaaca actgcacctt tgaatacgtg 540

tcccagccct tccttatgga cctggaggga aagcagggaa actttaagaa tctgagagag 600

ttcgtgttta aaaatatcga cggctatttt aagatctatt ctaagcacac gcctattaat 660

ctcgtgcgcg atcttccaca aggcttcagc gccctggaac cactcgtgga cctcccaatt 720

ggtatcaaca tcactagatt tcagactctg cttgccctcc accgatccta tctgacaccc 780

ggagactcct ctagcggctg gactgccggc gctgccgctt attacgttgg ttatcttcag 840

ccacgcacgt tcctgctgaa gtataacgag aatggtacta ttaccgatgc cgtggattgt 900

gcccttgacc ccctgtccga aactaagtgc acactcaagt cattcactgt ggaaaaagga 960

atctaccaga caagcaattt tcgggtccag cctactgaga gcattgtgcg ctttcctaac 1020

atcacaaatc tttgcccctt cggagaggtt ttcaatgcta cacggtttgc ctccgtgtat 1080

gcctggaacc gcaagagaat ttccaattgc gtggccgatt actccgtgct ctacaatagt 1140

gcaagcttta gcacctttaa gtgctatggc gtatccccta ctaagcttaa cgacttgtgt 1200

ttcacaaacg tgtatgccga ctcctttgtg atacggggcg acgaagttag acagatagca 1260

ccaggacaga cgggaaagat agctgactac aactataagc ttcctgatga cttcactggc 1320

tgcgttatcg cgtggaattc taacaacctg gactcaaaag tcggcggcaa ctataactat 1380

ctctatcggc tgttccgcaa gagtaacctt aagccctttg agagagatat aagcactgaa 1440

atctaccagg ctggcagtac gccctgtaat ggcgtggaag gctttaattg ttattttcca 1500

ctgcaatcct atggttttca gccaaccaat ggcgtgggct accaaccata ccgcgtcgtg 1560

gtgctctcct ttgaactgct ccacgctccc gcgactgtct gcggccccaa gaagtccacg 1620

aaccttgtga agaataagtg cgttaatttt aatttcaacg gcctcactgg aacaggagtg 1680

ctcactgaga gtaacaagaa gttcctgcca tttcaacaat ttggcagaga catagccgat 1740

actactgacg ccgttaggga cccccagacc ctcgagattc tcgatataac gccctgctcc 1800

ttcggtggag tttccgtgat cacgccaggc accaatacca gtaaccaggt cgccgtgctg 1860

tatcaggatg tcaactgtac tgaggtgccc gtagccatcc atgcggatca gctcacacca 1920

acttggaggg tgtacagcac cggctccaat gtattccaga ctcgggccgg atgccttatt 1980

ggcgccgaac acgtgaacaa tagttacgaa tgcgatattc caattggcgc cggaatctgt 2040

gctagctacc agactcagac gaactcccca ggcagcgcca gcagcgttgc cagccagtca 2100

atcatcgctt atacaatgtc acttggagcc gaaaactccg tggcttactc aaacaacagc 2160

atcgccatcc ccacaaactt caccatatcc gtgacaactg agattctgcc agtgtccatg 2220

actaagacgt ccgtagattg cactatgtac atatgcggcg acagcacaga atgttctaat 2280

ctgctgctgc aatatggaag cttctgcact caactgaaca gagcgctcac aggcatcgcc 2340

gtggagcagg ataagaatac ccaggaggtg ttcgcccaag ttaagcagat ctacaagacc 2400

ccacccataa aggatttcgg tggattcaat tttagtcaga tactcccaga cccatctaag 2460

ccatccaaga ggagctttat cgaggatctt ttgtttaaca aagttactct ggccgacgcc 2520

ggtttcatca agcagtacgg agattgcctc ggcgacatcg ctgctcgtga cctcatctgt 2580

gcgcaaaagt ttaacggtct gacggtgctg cctcccctcc ttactgatga aatgatcgcc 2640

cagtatacca gcgcactcct cgctggcacc ataacatccg gttggacatt cggcgctggt 2700

gcagcactgc agataccatt cgccatgcaa atggcatatc gtttcaacgg tatcggtgtc 2760

acacagaatg tcctatatga gaaccagaag ctgatcgcaa atcagttcaa tagtgccatc 2820

ggaaaaatcc aggatagcct tagcagcaca gcctcagccc ttggcaaact ccaggatgtc 2880

gtgaaccaga atgcccaggc tctcaatacc ctcgtgaagc agctctcatc taatttcggc 2940

gcaatttcca gtgtcctcaa cgacatcctc agccgcctcg acccccccga ggccgaagtg 3000

cagattgaca gactgattac aggtcgactc cagagcctcc agacttacgt gactcagcag 3060

ctgataagag ccgccgagat aagggccagc gctaacctgg ctgccacaaa gatgtctgag 3120

tgcgtgctgg gccagtccaa gagagtagac ttctgtggca aaggctacca tctgatgagc 3180

ttcccacaat ccgcacctca cggcgtagtg ttcctccacg tgacatatgt accggctcag 3240

gagaagaatt tcactaccgc tcctgctata tgccatgatg gaaaggctca cttcccccgg 3300

gagggggtgt tcgtgtccaa cggcacccat tggtttgtga ctcagcggaa tttctacgaa 3360

ccccagatca taaccactga caacacattt gtgtccggaa attgtgacgt ggtcattgga 3420

atagtgaaca acactgttta tgatccactg cagccagaac ttgacagctt taaggaggag 3480

ctcgacaagt acttcaagaa tcatacgtca ccagatgtgg acctcggaga tattagcggt 3540

atcaatgcca gtgttgtcaa tattcagaag gaaatagacc gccttaatga ggtcgccaaa 3600

aatctgaacg agagcctcat cgatcttcag gagctgggca aatatgagca gtacatcaag 3660

tggccttggt atatttggct tggcttcatc gccggcctga tcgccatagt aatggtcaca 3720

attatgctcg ttgtttatat ggtctccaga gacaatgttt cttgctccat ttgtctataa 3780

<210> 80

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> IF-H1HawaiiCT.r

<400> 80

acgacacgac taaggccttt aaatacatat tctacactgt agagaccc 48

<210> 81

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> IF-H3MinnesotaCT.r

<400> 81

acgacacgac taaggccttt agatgcagat gttgcatctg atgttgccct tctg 54

<210> 82

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> IF-HongKongCT.r

<400> 82

acgacacgac taaggccttt atatacatat cctgcactgc attgaaccat tt 52

<210> 83

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> IF-GuangdongCT.r

<400> 83

acgacacgac taaggccttt atatacaaat agtgcaccgc atgtttcca 49

<210> 84

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> IF-H9HKCT.r

<400> 84

acgacacgac taaggccttt atatacaaat gttgcatctg caagatccat 50

<210> 85

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> IF-BWashCT.r

<400> 85

acgacacgac taaggccttt atagacaaat ggagcaagaa acattgtctc 50

<210> 86

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> IF(nbHEL40)-PDI.c

<400> 86

ccaaaacaca ttgagcaaaa tggcgaaaaa cgttgcgatt ttcggcttat 50

<210> 87

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> IF(AvB+wtCT).r

<400> 87

acgacacgac taaggccttt aggtataatg gagtttcacc cccttcagaa 50

<210> 88

<211> 3801

<212> DNA

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 wtTMCT-DNA

<400> 88

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtccgatct ggatcggtgc actacatttg acgatgtgca ggcacctaat 120

tatactcagc acacctcttc catgcggggc gtgtactacc ccgacgagat ttttagaagt 180

gacacactgt acctgaccca agacctttt ctcccatttt atagcaatgt cacgggattc 240

cacactatca atcacacatt cggaaaccct gttatcccct ttaaagacgg tatctacttt 300

gctgctactg agaaatcgaa tgttgttcgc ggttgggtgt tcggctcaac catgaacaac 360

aagagtcagt cagtaataat tataaacaac tcaaccaatg tcgtcatcag ggcttgcaac 420

ttcgagctct gcgataaccc cttctttgcg gtgtccaaac cgatgggcac tcagacccat 480

accatgattt ttgacaatgc ctttaattgt actttcgaat acatcagcga tgccttctcg 540

ctggatgtgt ccgagaagag cggtaacttc aagcatttgc gggagttcgt tttcaaaaac 600

aaagacgggt ttttgtatgt ctacaaaggc tatcagccca ttgatgtggt cagggatctg 660

cccagtggtt tcaacacact gaagcccatc ttcaagttgc cactcggcat caacatcact 720

aatttccgcg ccatcctaac tgctttttcg ccagcgcagg atatttgggg gacatccgcc 780

gcagcttact tcgtgggata tctgaagccc accactttta tgctaaagta cgacgagaat 840

ggcaccatca ccgatgccgt ggactgctca cagaatcctc tcgcagagct caagtgttca 900

gtgaagtcat ttgagatcga caaagggatc taccaaacat ctaactttcg cgtcgtgcct 960

tccggtgacg tagtccgttt cccaaatatc actaacttgt gccctttttgg tgaagtattc 1020

aacgcaacca aattccccag tgtgtatgcg tgggagcgca aaaagatcag taactgtgtg 1080

gctgactata gtgttctgta caatagcacc ttcttcagca ccttcaagtg ttatggagtg 1140

agcgctacaa aactgaacga tctttgtttc tcaaacgtgt acgccgattc atttgtcgtt 1200

aaaggtgacg atgtgaggca gatcgctcca ggccagacag gtgtgattgc tgactataat 1260

tacaaactgc cagacgactt catggggtgc gtgctagctt ggaatacaag aaacattgac 1320

gccacctcca cgggaaatta caattacaag tatcgttacc ttcgccatgg aaagttgaga 1380

cccttcgagc gtgatataag taacgtgccc tttagtccag atggaaaacc ctgcacaccc 1440

cctgctctca attgctattg gcctctcaat gactacggct tttacacaac tactggcatc 1500

ggataccagc cttaccgggt cgtggtgctc agttttgagt tgcttaacgc acccgccacc 1560

gtgtgtggtc ctaaactttc tactgacctg attaaaaacc aatgcgtcaa cttcaatttt 1620

aacgggctga ccggcaccgg tgtcctgacc cctagctcta agagattcca gccttttcag 1680

cagttcggga gggatgtgag cgactttacc gactctgtca gggatccaaa gaccagcgag 1740

atactggata tctcgccctg cagtttcggt ggcgtgtccg ttattacacc tggcaccaac 1800

gcctcctcag aggtggcggt gctctatcaa gatgtcaact gcactgatgt gtcaactgcc 1860

atccatgccg atcagctgac ccccgcctgg cgcatctaca gtaccgggaa caacgttttt 1920

cagacccagg ccggctgtct aatcggcgca gagcacgttg acacatccta cgaatgtgac 1980

atacctatcg gggcaggcat ttgcgctagc taccataccg tgtcactgtt ggcttccacg 2040

tcacaaaagt caatcgttgc ctacacgatg agtctggggg ctgactcatc tatcgcctac 2100

agcaacaata ccattgcaat tcccacaaac ttcagtatct ccatcacaac agaggtgatg 2160

cccgtttcta tggctaaaac atcagtcgat tgcaatatgt atatatgcgg cgatagtact 2220

gagtgcgcca atctcttgtt acagtacggc tccttttgta cccagctgaa ccgagcactg 2280

tctggaatcg ccgcagaaca ggatcgcaat acccgggaag tcttcgccca ggtgaagcag 2340

atgtacaaaa cgcccactct caagtatttc ggcggattca acttttctca gattttgcct 2400

gacccgctca agccaacaaa acgatctttt atcgaagacc ttctgtttaa caaggtcaca 2460

ctggcggatg ctgggttcat gaaacagtac ggtgaatgcc tgggggacat caatgccaga 2520

gatctgatct gcgcccagaa attcaatggc ttaacagtcc tcccacctct cttgaccgac 2580

gatatgatcg ctgcgtacac cgctgctctg gtatcgggca ccgcgactgc tggctggacc 2640

tttggtgccg gagccgcact ccagatccca ttcgccatgc agatggccta ccgcttcaac 2700

ggaatcgggg tcacccagaa cgtgctgtat gagaaccaga aacagatcgc caatcagttc 2760

aataaggcaa ttagtcagat tcaggagagt cttaccacta ccagcaccgc cctgggcaag 2820

ctgcaagatg ttgtgaacca gaatgcgcag gcattaaaca ctctggttaa acagctgagc 2880

tcaaattttg gtgcaatctc ttcagttctg aacgatatcc tgagtcggct ggatccgcca 2940

gaggctgaag tgcaaattga tcgtttgatc accgggaggc tacaatctct gcagacgtac 3000

gtgacccagc agctcatccg ggcagccgaa attcgcgcat cagccaacct cgctgcaact 3060

aagatgtctg agtgcgtgct gggccagagt aagagggtgg acttttgtgg taagggatac 3120

cacctcatgt cctttccgca agcggctccc cacggcgtgg ttttcttaca cgttacctat 3180

gtgccatccc aagaacgcaa tttcaccacc gctccagcta tctgtcatga gggcaaagca 3240

tatttcccca gggaaggagt atttgtgttt aatggcacgt cctggtttat aacccaacgt 3300

aacttttttct ccccacagat tatcacaacc gacaacacat tcgtgtctgg gaattgtgac 3360

gtcgtgatcg ggatcattaa caataccgtt tacgatccct tgcagcccga gcttgactcc 3420

tttaaagagg aactagacaa atactttaag aatcacacct caccggacgt agatttggga 3480
gacatctctg gaattaatgc ctctgtggtg aatatccaga aggagatcga ccgcctgaat 3540
gaagtcgcca agaacctcaa cgagtccctg atagatctgc aagaactggg caaatatgaa 3600
cagtacatca aatggccgtg gtacgtgtgg ttgggcttta tcgctggact tattgcaatc 3660
gtgatggtga cgattctgct ctgctgtatg acttcctgct gctcttgtct gaagggcgcc 3720
tgtagctgtg gttcctgctg caagttcgac gaagacgact ccgaaccagt tctgaagggg 3780
gtgaaactcc attataccta a 3801

<210> 89
<211> 3741
<212> DNA
<213> Artificial Sequence
<220><223> PDI-SARS-COV-1 H5iTMCT-DNA
<400> 89
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cgtccgatct ggatcggtgc actacatttg acgatgtgca ggcacctaat 120
tatactcagc acacctcttc catgcggggc gtgtactacc ccgacgagat ttttagaagt 180
gacacactgt acctgaccca agacctttt ctcccatttt atagcaatgt cacgggattc 240
cacactatca atcacacatt cggaaaccct gttatcccct ttaaagacgg tatctacttt 300
gctgctactg agaaatcgaa tgttgttcgc ggttgggtgt tcggctcaac catgaacaac 360

aagagtcagt cagtaataat tataaacaac tcaaccaatg tcgtcatcag ggcttgcaac 420
ttcgagctct gcgataaccc cttctttgcg gtgtccaaac cgatgggcac tcagacccat 480
accatgattt ttgacaatgc ctttaattgt actttcgaat acatcagcga tgccttctcg 540
ctggatgtgt ccgagaagag cggtaacttc aagcatttgc gggagttcgt tttcaaaaac 600
aaagacgggt ttttgtatgt ctacaaaggc tatcagccca ttgatgtggt cagggatctg 660
cccagtggtt tcaacacact gaagcccatc ttcaagttgc cactcggcat caacatcact 720
aatttccgcg ccatcctaac tgctttttcg ccagcgcagg atatttgggg gacatccgcc 780

gcagcttact tcgtgggata tctgaagccc accactttta tgctaaagta cgacgagaat 840
ggcaccatca ccgatgccgt ggactgctca cagaatcctc tcgcagagct caagtgttca 900
gtgaagtcat ttgagatcga caaagggatc taccaaacat ctaactttcg cgtcgtgcct 960
tccggtgacg tagtccgttt cccaaatatc actaacttgt gccctttttgg tgaagtattc 1020

aacgcaacca aattccccag tgtgtatgcg tgggagcgca aaaagatcag taactgtgtg 1080

gctgactata gtgttctgta caatagcacc ttcttcagca ccttcaagtg ttatggagtg 1140

agcgctacaa aactgaacga tctttgtttc tcaaacgtgt acgccgattc atttgtcgtt 1200

aaaggtgacg atgtgaggca gatcgctcca ggccagacag gtgtgattgc tgactataat 1260

tacaaactgc cagacgactt catggggtgc gtgctagctt ggaatacaag aaacattgac 1320

gccacctcca cgggaaatta caattacaag tatcgttacc ttcgccatgg aaagttgaga 1380

cccttcgagc gtgatataag taacgtgccc tttagtccag atggaaaacc ctgcacaccc 1440

cctgctctca attgctattg gcctctcaat gactacggct tttacacaac tactggcatc 1500

ggataccagc cttaccgggt cgtggtgctc agttttgagt tgcttaacgc acccgccacc 1560

gtgtgtggtc ctaaactttc tactgacctg attaaaaacc aatgcgtcaa cttcaatttt 1620

aacgggctga ccggcaccgg tgtcctgacc cctagctcta agagattcca gccttttcag 1680

cagttcggga gggatgtgag cgactttacc gactctgtca gggatccaaa gaccagcgag 1740

atactggata tctcgccctg cagtttcggt ggcgtgtccg ttattacacc tggcaccaac 1800

gcctcctcag aggtggcggt gctctatcaa gatgtcaact gcactgatgt gtcaactgcc 1860

atccatgccg atcagctgac ccccgcctgg cgcatctaca gtaccgggaa caacgttttt 1920

cagacccagg ccggctgtct aatcggcgca gagcacgttg acacatccta cgaatgtgac 1980

atacctatcg gggcaggcat ttgcgctagc taccataccg tgtcactgtt ggcttccacg 2040

tcacaaaagt caatcgttgc ctacacgatg agtctggggg ctgactcatc tatcgcctac 2100

agcaacaata ccattgcaat tcccacaaac ttcagtatct ccatcacaac agaggtgatg 2160

cccgtttcta tggctaaaac atcagtcgat tgcaatatgt atatatgcgg cgatagtact 2220

gagtgcgcca atctcttgtt acagtacggc tccttttgta cccagctgaa ccgagcactg 2280

tctggaatcg ccgcagaaca ggatcgcaat acccgggaag tcttcgccca ggtgaagcag 2340

atgtacaaaa cgcccactct caagtatttc ggcggattca acttttctca gattttgcct 2400

gacccgctca agccaacaaa acgatctttt atcgaagacc ttctgtttaa caaggtcaca 2460

ctggcggatg ctgggttcat gaaacagtac ggtgaatgcc tgggggacat caatgccaga 2520

gatctgatct gcgcccagaa attcaatggc ttaacagtcc tcccacctct cttgaccgac 2580

gatatgatcg ctgcgtacac cgctgctctg gtatcgggca ccgcgactgc tggctggacc 2640

tttggtgccg gagccgcact ccagatccca ttcgccatgc agatggccta ccgcttcaac 2700

ggaatcgggg tcacccagaa cgtgctgtat gagaaccaga aacagatcgc caatcagttc 2760

aataaggcaa ttagtcagat tcaggagagt cttaccacta ccagcaccgc cctgggcaag 2820

ctgcaagatg ttgtgaacca gaatgcgcag gcattaaaca ctctggttaa acagctgagc 2880

tcaaattttg gtgcaatctc ttcagttctg aacgatatcc tgagtcggct ggatccgcca 2940

gaggctgaag tgcaaattga tcgtttgatc accgggaggc tacaatctct gcagacgtac 3000

gtgacccagc agctcatccg ggcagccgaa attcgcgcat cagccaacct cgctgcaact 3060

aagatgtctg agtgcgtgct gggccagagt aagagggtgg acttttgtgg taagggatac 3120

cacctcatgt cctttccgca agcggctccc cacggcgtgg ttttcttaca cgttacctat 3180

gtgccatccc aagaacgcaa tttcaccacc gctccagcta tctgtcatga gggcaaagca 3240

tatttcccca gggaaggagt atttgtgttt aatggcacgt cctggtttat aacccaacgt 3300

aacttttct ccccacagat tatcacaacc gacaacacat tcgtgtctgg gaattgtgac 3360

gtcgtgatcg ggatcattaa caataccgtt tacgatccct tgcagcccga gcttgactcc 3420

tttaaagagg aactagacaa atactttaag aatcacacct caccggacgt agatttggga 3480

gacatctctg gaattaatgc ctctgtggtg aatatccaga aggagatcga ccgcctgaat 3540

gaagtcgcca agaacctcaa cgagtccctg atagatctgc aagaactggg caaatatgaa 3600

cagtacatca aatggccgtg gtaccaaata ctgtcaattt attcaacagt ggcgagttcc 3660

ctagcactgg caatcatgat ggctggtcta tctttatgga tgtgctccaa tggatcgtta 3720

caatgcagaa tttgcattta a 3741

<210> 90

<211> 3732

<212> DNA

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 H5iCT-DNA

<400> 90

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtccgatct ggatcggtgc actacatttg acgatgtgca ggcacctaat 120

tatactcagc acacctcttc catgcggggc gtgtactacc ccgacgagat ttttagaagt 180

gacacactgt acctgaccca agacctttt ctcccatttt atagcaatgt cacgggattc 240

cacactatca atcacacatt cggaaaccct gttatcccct ttaaagacgg tatctacttt 300

gctgctactg agaaatcgaa tgttgttcgc ggttgggtgt tcggctcaac catgaacaac 360

aagagtcagt cagtaataat tataaacaac tcaaccaatg tcgtcatcag ggcttgcaac 420

ttcgagctct gcgataaccc cttctttgcg gtgtccaaac cgatgggcac tcagacccat 480

accatgattt ttgacaatgc ctttaattgt actttcgaat acatcagcga tgccttctcg 540

ctggatgtgt ccgagaagag cggtaacttc aagcatttgc gggagttcgt tttcaaaaac 600

aaagacgggt ttttgtatgt ctacaaaggc tatcagccca ttgatgtggt cagggatctg 660

cccagtggtt tcaacacact gaagcccatc ttcaagttgc cactcggcat caacatcact 720

aatttccgcg ccatcctaac tgctttttcg ccagcgcagg atatttgggg gacatccgcc 780

gcagcttact tcgtgggata tctgaagccc accactttta tgctaaagta cgacgagaat 840

ggcaccatca ccgatgccgt ggactgctca cagaatcctc tcgcagagct caagtgttca 900

gtgaagtcat ttgagatcga caaagggatc taccaaacat ctaactttcg cgtcgtgcct 960

tccggtgacg tagtccgttt cccaaatatc actaacttgt gccctttgg tgaagtattc 1020

aacgcaacca aattccccag tgtgtatgcg tgggagcgca aaaagatcag taactgtgtg 1080

gctgactata gtgttctgta caatagcacc ttcttcagca ccttcaagtg ttatggagtg 1140

agcgctacaa aactgaacga tctttgtttc tcaaacgtgt acgccgattc atttgtcgtt 1200

aaaggtgacg atgtgaggca gatcgctcca ggccagacag gtgtgattgc tgactataat 1260

tacaaactgc cagacgactt catggggtgc gtgctagctt ggaatacaag aaacattgac 1320

gccacctcca cgggaaatta caattacaag tatcgttacc ttcgccatgg aaagttgaga 1380

cccttcgagc gtgatataag taacgtgccc tttagtccag atggaaaacc ctgcacaccc 1440

cctgctctca attgctattg gcctctcaat gactacggct tttacacaac tactggcatc 1500

ggataccagc cttaccgggt cgtggtgctc agttttgagt tgcttaacgc acccgccacc 1560

gtgtgtggtc ctaaactttc tactgacctg attaaaaacc aatgcgtcaa cttcaatttt 1620

aacgggctga ccggcaccgg tgtcctgacc cctagctcta agagattcca gccttttcag 1680

cagttcggga gggatgtgag cgactttacc gactctgtca gggatccaaa gaccagcgag 1740

atactggata tctcgccctg cagtttcggt ggcgtgtccg ttattacacc tggcaccaac 1800

gcctcctcag aggtggcggt gctctatcaa gatgtcaact gcactgatgt gtcaactgcc 1860

atccatgccg atcagctgac ccccgcctgg cgcatctaca gtaccgggaa caacgttttt 1920

cagacccagg ccggctgtct aatcggcgca gagcacgttg acacatccta cgaatgtgac 1980

atacctatcg gggcaggcat ttgcgctagc taccataccg tgtcactgtt ggcttccacg 2040

tcacaaaagt caatcgttgc ctacacgatg agtctggggg ctgactcatc tatcgcctac 2100

agcaacaata ccattgcaat tcccacaaac ttcagtatct ccatcacaac agaggtgatg 2160

cccgtttcta tggctaaaac atcagtcgat tgcaatatgt atatatgcgg cgatagtact 2220

gagtgcgcca atctcttgtt acagtacggc tccttttgta cccagctgaa ccgagcactg 2280

tctggaatcg ccgcagaaca ggatcgcaat acccgggaag tcttcgccca ggtgaagcag 2340

atgtacaaaa cgcccactct caagtatttc ggcggattca acttttctca gattttgcct 2400

gacccgctca agccaacaaa acgatctttt atcgaagacc ttctgtttaa caaggtcaca 2460
ctggcggatg ctgggttcat gaaacagtac ggtgaatgcc tgggggacat caatgccaga 2520
gatctgatct gcgcccagaa attcaatggc ttaacagtcc tcccacctct cttgaccgac 2580
gatatgatcg ctgcgtacac cgctgctctg gtatcgggca ccgcgactgc tggctggacc 2640
tttggtgccg gagccgcact ccagatccca ttcgccatgc agatggccta ccgcttcaac 2700
ggaatcgggg tcacccagaa cgtgctgtat gagaaccaga aacagatcgc caatcagttc 2760
aataaggcaa ttagtcagat tcaggagagt cttaccacta ccagcaccgc cctgggcaag 2820

ctgcaagatg ttgtgaacca gaatgcgcag gcattaaaca ctctggttaa acagctgagc 2880
tcaaattttg gtgcaatctc ttcagttctg aacgatatcc tgagtcggct ggatccgcca 2940
gaggctgaag tgcaaattga tcgtttgatc accgggaggc tacaatctct gcagacgtac 3000
gtgacccagc agctcatccg ggcagccgaa attcgcgcat cagccaacct cgctgcaact 3060
aagatgtctg agtgcgtgct gggccagagt aagagggtgg acttttgtgg taagggatac 3120
cacctcatgt cctttccgca agcggctccc cacggcgtgg ttttcttaca cgttacctat 3180
gtgccatccc aagaacgcaa tttcaccacc gctccagcta tctgtcatga gggcaaagca 3240

tatttcccca gggaaggagt atttgtgttt aatggcacgt cctggtttat aacccaacgt 3300
aacttttttct ccccacagat tatcacaacc gacaacacat tcgtgtctgg gaattgtgac 3360
gtcgtgatcg ggatcattaa caataccgtt tacgatccct tgcagcccga gcttgactcc 3420
tttaaagagg aactagacaa atactttaag aatcacacct caccggacgt agatttggga 3480
gacatctctg gaattaatgc ctctgtggtg aatatccaga aggagatcga ccgcctgaat 3540
gaagtcgcca agaacctcaa cgagtccctg atagatctgc aagaactggg caaatatgaa 3600
cagtacatca aatggccgtg gtacgtgtgg ttgggcttta tcgctggact tattgcaatc 3660

gtgatggtga cgattctgct ctctttatgg atgtgctcca atggatcgtt acaatgcaga 3720
atttgcattt aa 3732

<210> 91
<211> 3723
<212> DNA
<213> Artificial Sequence
<220><223> PDI-SARS-COV-1 H5iCT(V4)-DNA
<400> 91
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtccgatct ggatcggtgc actacatttg acgatgtgca ggcacctaat 120

tatactcagc acacctcttc catgcggggc gtgtactacc ccgacgagat ttttagaagt 180

gacacactgt acctgaccca agacctttt ctcccatttt atagcaatgt cacgggattc 240

cacactatca atcacacatt cggaaaccct gttatcccct ttaaagacgg tatctacttt 300

gctgctactg agaaatcgaa tgttgttcgc ggttgggtgt tcggctcaac catgaacaac 360

aagagtcagt cagtaataat tataaacaac tcaaccaatg tcgtcatcag ggcttgcaac 420

ttcgagctct gcgataaccc cttctttgcg gtgtccaaac cgatgggcac tcagacccat 480

accatgattt ttgacaatgc ctttaattgt actttcgaat acatcagcga tgccttctcg 540

ctggatgtgt ccgagaagag cggtaacttc aagcatttgc gggagttcgt tttcaaaaac 600

aaagacgggt ttttgtatgt ctacaaaggc tatcagccca ttgatgtggt cagggatctg 660

cccagtggtt tcaacacact gaagcccatc ttcaagttgc cactcggcat caacatcact 720

aatttccgcg ccatcctaac tgctttttcg ccagcgcagg atatttgggg gacatccgcc 780

gcagcttact tcgtgggata tctgaagccc accactttta tgctaaagta cgacgagaat 840

ggcaccatca ccgatgccgt ggactgctca cagaatcctc tcgcagagct caagtgttca 900

gtgaagtcat ttgagatcga caaagggatc taccaaacat ctaactttcg cgtcgtgcct 960

tccggtgacg tagtccgttt cccaaatatc actaacttgt gccctttttgg tgaagtattc 1020

aacgcaacca aattccccag tgtgtatgcg tgggagcgca aaaagatcag taactgtgtg 1080

gctgactata gtgttctgta caatagcacc ttcttcagca ccttcaagtg ttatggagtg 1140

agcgctacaa aactgaacga tctttgtttc tcaaacgtgt acgccgattc atttgtcgtt 1200

aaaggtgacg atgtgaggca gatcgctcca ggccagacag gtgtgattgc tgactataat 1260

tacaaactgc cagacgactt catggggtgc gtgctagctt ggaatacaag aaacattgac 1320

gccacctcca cgggaaatta caattacaag tatcgttacc ttcgccatgg aaagttgaga 1380

cccttcgagc gtgatataag taacgtgccc tttagtccag atggaaaacc ctgcacaccc 1440

cctgctctca attgctattg gcctctcaat gactacggct tttacacaac tactggcatc 1500

ggataccagc cttaccgggt cgtggtgctc agttttgagt tgcttaacgc acccgccacc 1560

gtgtgtggtc ctaaactttc tactgacctg attaaaaacc aatgcgtcaa cttcaatttt 1620

aacgggctga ccggcaccgg tgtcctgacc cctagctcta agagattcca gccttttcag 1680

cagttcggga gggatgtgag cgactttacc gactctgtca gggatccaaa gaccagcgag 1740

atactggata tctcgccctg cagtttcggt ggcgtgtccg ttattacacc tggcaccaac 1800

gcctcctcag aggtggcggt gctctatcaa gatgtcaact gcactgatgt gtcaactgcc 1860

atccatgccg atcagctgac ccccgcctgg cgcatctaca gtaccgggaa caacgttttt 1920

cagacccagg ccggctgtct aatcggcgca gagcacgttg acacatccta cgaatgtgac 1980
atacctatcg gggcaggcat ttgcgctagc taccataccg tgtcactgtt ggcttccacg 2040
tcacaaaagt caatcgttgc ctacacgatg agtctggggg ctgactcatc tatcgcctac 2100
agcaacaata ccattgcaat tcccacaaac ttcagtatct ccatcacaac agaggtgatg 2160
cccgtttcta tggctaaaac atcagtcgat tgcaatatgt atatatgcgg cgatagtact 2220
gagtgcgcca atctcttgtt acagtacggc tccttttgta cccagctgaa ccgagcactg 2280
tctggaatcg ccgcagaaca ggatcgcaat acccgggaag tcttcgccca ggtgaagcag 2340

atgtacaaaa cgcccactct caagtatttc ggcggattca acttttctca gattttgcct 2400
gacccgctca agccaacaaa acgatctttt atcgaagacc ttctgtttaa caaggtcaca 2460
ctggcggatg ctgggttcat gaaacagtac ggtgaatgcc tgggggacat caatgccaga 2520
gatctgatct gcgcccagaa attcaatggc ttaacagtcc tcccacctct cttgaccgac 2580
gatatgatcg ctgcgtacac cgctgctctg gtatcgggca ccgcgactgc tggctggacc 2640
tttggtgccg gagccgcact ccagatccca ttcgccatgc agatggccta ccgcttcaac 2700
ggaatcgggg tcacccagaa cgtgctgtat gagaaccaga aacagatcgc caatcagttc 2760

aataaggcaa ttagtcagat tcaggagagt cttaccacta ccagcaccgc cctgggcaag 2820
ctgcaagatg ttgtgaacca gaatgcgcag gcattaaaca ctctggttaa acagctgagc 2880
tcaaattttg gtgcaatctc ttcagttctg aacgatatcc tgagtcggct ggatccgcca 2940
gaggctgaag tgcaaattga tcgtttgatc accgggaggc tacaatctct gcagacgtac 3000
gtgacccagc agctcatccg ggcagccgaa attcgcgcat cagccaacct cgctgcaact 3060
aagatgtctg agtgcgtgct gggccagagt aagagggtgg acttttgtgg taagggatac 3120
cacctcatgt cctttccgca agcggctccc cacggcgtgg ttttcttaca cgttacctat 3180

gtgccatccc aagaacgcaa tttcaccacc gctccagcta tctgtcatga gggcaaagca 3240
tatttcccca gggaaggagt atttgtgttt aatggcacgt cctggtttat aacccaacgt 3300
aacttttttct ccccacagat tatcacaacc gacaacacat tcgtgtctgg gaattgtgac 3360
gtcgtgatcg ggatcattaa caataccgtt tacgatccct tgcagcccga gcttgactcc 3420
tttaaagagg aactagacaa atactttaag aatcacacct caccggacgt agatttggga 3480
gacatctctg gaattaatgc ctctgtggtg aatatccaga aggagatcga ccgcctgaat 3540
gaagtcgcca agaacctcaa cgagtccctg atagatctgc aagaactggg caaatatgaa 3600

cagtacatca aatggccgtg gtacgtgtgg ttgggcttta tcgctggact tattgcaatc 3660

gtgatggtga cgattctgct ctgctgctcc aatggatcgt tacaatgcag aatttgcatt 3720

taa 3723

<210> 92

<211> 3732

<212> DNA

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 H1cCT-DNA

<400> 92

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtccgatct ggatcggtgc actacatttg acgatgtgca ggcacctaat 120

tatactcagc acacctcttc catgcggggc gtgtactacc ccgacgagat ttttagaagt 180

gacacactgt acctgaccca agacctttt ctcccatttt atagcaatgt cacgggattc 240

cacactatca atcacacatt cggaaaccct gttatcccct ttaaagacgg tatctacttt 300

gctgctactg agaaatcgaa tgttgttcgc ggttgggtgt tcggctcaac catgaacaac 360

aagagtcagt cagtaataat tataaacaac tcaaccaatg tcgtcatcag ggcttgcaac 420

ttcgagctct gcgataaccc cttctttgcg gtgtccaaac cgatgggcac tcagacccat 480

accatgattt ttgacaatgc ctttaattgt actttcgaat acatcagcga tgccttctcg 540

ctggatgtgt ccgagaagag cggtaacttc aagcatttgc gggagttcgt tttcaaaaac 600

aaagacgggt ttttgtatgt ctacaaaggc tatcagccca ttgatgtggt cagggatctg 660

cccagtggtt tcaacacact gaagcccatc ttcaagttgc cactcggcat caacatcact 720

aatttccgcg ccatcctaac tgctttttcg ccagcgcagg atatttgggg gacatccgcc 780

gcagcttact tcgtgggata tctgaagccc accactttta tgctaaagta cgacgagaat 840

ggcaccatca ccgatgccgt ggactgctca cagaatcctc tcgcagagct caagtgttca 900

gtgaagtcat ttgagatcga caaagggatc taccaaacat ctaactttcg cgtcgtgcct 960

tccggtgacg tagtccgttt cccaaatatc actaacttgt gccctttgg tgaagtattc 1020

aacgcaacca aattccccag tgtgtatgcg tgggagcgca aaaagatcag taactgtgtg 1080

gctgactata gtgttctgta caatagcacc ttcttcagca ccttcaagtg ttatggagtg 1140

agcgctacaa aactgaacga tctttgtttc tcaaacgtgt acgccgattc atttgtcgtt 1200

aaaggtgacg atgtgaggca gatcgctcca ggccagacag gtgtgattgc tgactataat 1260

tacaaactgc cagacgactt catggggtgc gtgctagctt ggaatacaag aaacattgac 1320

gccacctcca cgggaaatta caattacaag tatcgttacc ttcgccatgg aaagttgaga 1380

ccctttcgagc gtgatataag taacgtgccc tttagtccag atggaaaacc ctgcacaccc 1440

cctgctctca attgctattg gcctctcaat gactacggct tttacacaac tactggcatc 1500
ggataccagc cttaccgggt cgtggtgctc agttttgagt tgcttaacgc acccgccacc 1560
gtgtgtggtc ctaaactttc tactgacctg attaaaaacc aatgcgtcaa cttcaatttt 1620
aacgggctga ccggcaccgg tgtcctgacc cctagctcta agagattcca gccttttcag 1680
cagttcggga gggatgtgag cgactttacc gactctgtca gggatccaaa gaccagcgag 1740
atactggata tctcgccctg cagtttcggt ggcgtgtccg ttattacacc tggcaccaac 1800
gcctcctcag aggtggcggt gctctatcaa gatgtcaact gcactgatgt gtcaactgcc 1860

atccatgccg atcagctgac ccccgcctgg cgcatctaca gtaccgggaa caacgttttt 1920
cagacccagg ccggctgtct aatcggcgca gagcacgttg acacatccta cgaatgtgac 1980
atacctatcg gggcaggcat ttgcgctagc taccataccg tgtcactgtt ggcttccacg 2040
tcacaaaagt caatcgttgc ctacacgatg agtctggggg ctgactcatc tatcgcctac 2100
agcaacaata ccattgcaat tcccacaaac ttcagtatct ccatcacaac agaggtgatg 2160
cccgtttcta tggctaaaac atcagtcgat tgcaatatgt atatatgcgg cgatagtact 2220
gagtgcgcca atctcttgtt acagtacggc tccttttgta cccagctgaa ccgagcactg 2280

tctggaatcg ccgcagaaca ggatcgcaat acccgggaag tcttcgccca ggtgaagcag 2340
atgtacaaaa cgcccactct caagtatttc ggcggattca acttttctca gattttgcct 2400
gacccgctca agccaacaaa acgatctttt atcgaagacc ttctgtttaa caaggtcaca 2460
ctggcggatg ctgggttcat gaaacagtac ggtgaatgcc tgggggacat caatgccaga 2520
gatctgatct gcgcccagaa attcaatggc ttaacagtcc tcccacctct cttgaccgac 2580
gatatgatcg ctgcgtacac cgctgctctg gtatcgggca ccgcgactgc tggctggacc 2640
tttggtgccg gagccgcact ccagatccca ttcgccatgc agatggccta ccgcttcaac 2700

ggaatcgggg tcacccagaa cgtgctgtat gagaaccaga aacagatcgc caatcagttc 2760
aataaggcaa ttagtcagat tcaggagagt cttaccacta ccagcaccgc cctgggcaag 2820
ctgcaagatg ttgtgaacca gaatgcgcag gcattaaaca ctctggttaa acagctgagc 2880
tcaaattttg gtgcaatctc ttcagttctg aacgatatcc tgagtcggct ggatccgcca 2940
gaggctgaag tgcaaattga tcgtttgatc accgggaggc tacaatctct gcagacgtac 3000
gtgacccagc agctcatccg ggcagccgaa attcgcgcat cagccaacct cgctgcaact 3060
aagatgtctg agtgcgtgct gggccagagt aagagggtgg acttttgtgg taagggatac 3120

cacctcatgt cctttccgca agcggctccc cacggcgtgg ttttcttaca cgttacctat 3180

gtgccatccc aagaacgcaa tttcaccacc gctccagcta tctgtcatga gggcaaagca 3240

tatttcccca gggaaggagt atttgtgttt aatggcacgt cctggtttat aacccaacgt 3300

aacttttct ccccacagat tatcacaacc gacaacacat tcgtgtctgg gaattgtgac 3360

gtcgtgatcg ggatcattaa caataccgtt tacgatccct tgcagcccga gcttgactcc 3420

tttaaagagg aactagacaa atactttaag aatcacacct caccggacgt agatttggga 3480

gacatctctg gaattaatgc ctctgtggtg aatatccaga aggagatcga ccgcctgaat 3540

gaagtcgcca agaacctcaa cgagtccctg atagatctgc aagaactggg caaatatgaa 3600

cagtacatca aatggccgtg gtacgtgtgg ttgggcttta tcgctggact tattgcaatc 3660

gtgatggtga cgattctgct cagcttctgg atgtgctcta atgggtctct acagtgtaga 3720

atatgtattt aa 3732

<210> 93

<211> 1266

<212> PRT

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 wtTMCT-AA

<400> 93

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Ser Asp Leu Asp Arg Cys Thr Thr
20 25 30

Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met
35 40 45

Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr
50 55 60

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

His Thr Ile Asn His Thr Phe Gly Asn Pro Val Ile Pro Phe Lys Asp
85 90 95

Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp
100 105 110

Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile

115 120 125

Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys

130 135 140

Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr His

145 150 155 160

Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser

165 170 175

Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys His

180 185 190

Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr

195 200 205

Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe

210 215 220

Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr

225 230 235 240

Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp

245 250 255

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

260 265 270

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

275 280 285

Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe

290 295 300

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

305 310 315 320

Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe

325 330 335

Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu

340 345 350

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

355 360 365

Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys
370 375 380
Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala Asp Ser Phe Val Val
385 390 395 400

Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Val Ile
405 410 415
Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met Gly Cys Val Leu
420 425 430
Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn
435 440 445
Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu Arg
450 455 460

Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro
465 470 475 480
Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr
485 490 495
Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe
500 505 510
Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser Thr
515 520 525

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

Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu
595 600 605
Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp

610 615 620

Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe
625 630 635 640

Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser
645 650 655

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

Thr Val Ser Leu Leu Ala Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr
675 680 685

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

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

Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys
725 730 735

Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe
740 745 750

Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp
755 760 765

Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr
770 775 780

Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
785 790 795 800

Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
805 810 815

Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly Glu
820 825 830

Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
835 840 845

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala
850 855 860

Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp Thr
865 870 875 880
Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
885 890 895
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
900 905 910

Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln
915 920 925
Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp Val
930 935 940
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
945 950 955 960
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
965 970 975

Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly
980 985 990
Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala
995 1000 1005
Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser
1010 1015 1020
Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys
1025 1030 1035

Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His Gly Val
1040 1045 1050
Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn Phe
1055 1060 1065
Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
1070 1075 1080
Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr
1085 1090 1095

Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln

1160 1165 1170

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

1175 1180 1185

Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile

1190 1195 1200

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

1205 1210 1215

Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys Met Thr Ser Cys

1220 1225 1230

Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys Gly Ser Cys Cys Lys

1235 1240 1245

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

1250 1255 1260

His Tyr Thr

1265

<210> 94

<211> 1246

<212> PRT

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 H5iTMCT-AA

<400> 94

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Ser Asp Leu Asp Arg Cys Thr Thr

20 25 30

Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met

35 40 45

Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr

50 55 60

Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe

65 70 75 80

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

85 90 95

Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp

100 105 110

Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile

115 120 125

Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys

130 135 140

Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr His

145 150 155 160

Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser

165 170 175

Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys His

180 185 190

Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr

195 200 205

Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe

210 215 220

Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr

225 230 235 240

Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp

245 250 255

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

260 265 270

Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile Thr Asp Ala Val Asp
275 280 285
Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe
290 295 300
Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val Val Pro
305 310 315 320

Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe
325 330 335
Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu
340 345 350
Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn
355 360 365
Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys
370 375 380

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

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

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

515 520 525

Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu Thr

530 535 540

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

545 550 555 560

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

565 570 575

Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ser Phe Gly Gly Val

580 585 590

Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu

595 600 605

Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp

610 615 620

Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe

625 630 635 640

Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser

645 650 655

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

660 665 670

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

675 680 685

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

690 695 700

Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met

705 710 715 720

Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys

725 730 735

Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe

740 745 750

Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp

755 760 765

Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr
770 775 780
Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro
785 790 795 800
Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe
805 810 815
Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly Glu
820 825 830

Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
835 840 845
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala
850 855 860
Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp Thr
865 870 875 880
Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
885 890 895

Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
900 905 910
Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln
915 920 925
Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp Val
930 935 940
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
945 950 955 960

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

1010 1015 1020

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

Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr
1085 1090 1095
Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr
1100 1105 1110
Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn
1115 1120 1125
Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu
1130 1135 1140

Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp
1145 1150 1155
Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln
1160 1165 1170
Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu
1175 1180 1185
Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile
1190 1195 1200

Lys Trp Pro Trp Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala
1205 1210 1215
Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp
1220 1225 1230
Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1235 1240 1245

<210> 95

<211> 1243

<212> PRT

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 H5iCT-AA

<400> 95

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Ser Asp Leu Asp Arg Cys Thr Thr

20 25 30

Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met

35 40 45

Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr

50 55 60

Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe

65 70 75 80

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

85 90 95

Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp

100 105 110

Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile

115 120 125

Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys

130 135 140

Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr His

145 150 155 160

Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser

165 170 175

Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys His

180 185 190

Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr

195 200 205

Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe

210 215 220

Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr

225 230 235 240

Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp

245 250 255

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

260 265 270

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

275 280 285

Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe

290 295 300

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

305 310 315 320

Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe

325 330 335

Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu

340 345 350

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

355 360 365

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

370 375 380

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

385 390 395 400

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

405 410 415

Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met Gly Cys Val Leu

420 425 430

Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn

435 440 445

Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu Arg

450 455 460

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

465 470 475 480

Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr

485 490 495

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

500 505 510

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

515 520 525

Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu Thr

530 535 540

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

545 550 555 560

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

565 570 575

Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ser Phe Gly Gly Val

580 585 590

Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu

595 600 605

Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp

610 615 620

Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe

625 630 635 640

Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser

645 650 655

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

660 665 670

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

675 680 685

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

690 695 700

Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met

705 710 715 720

Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys

725 730 735

Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe

740 745 750

Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp

755 760 765

Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr

770 775 780

Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro

785 790 795 800

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

805 810 815

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

820 825 830

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

835 840 845

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala

850 855 860

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

865 870 875 880

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala

885 890 895

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

900 905 910

Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln

915 920 925

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

930 935 940

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

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

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

1085 1090 1095
Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr
1100 1105 1110
Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn
1115 1120 1125
Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu
1130 1135 1140
Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp

1145 1150 1155
Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln
1160 1165 1170
Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu

1175 1180 1185

Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile

1190 1195 1200

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

1205 1210 1215

Ala Ile Val Met Val Thr Ile Leu Leu Ser Leu Trp Met Cys Ser

1220 1225 1230

Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1235 1240

<210> 96

<211> 1240

<212> PRT

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 H5iCT(V4)-AA

<400> 96

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Ser Asp Leu Asp Arg Cys Thr Thr

20 25 30

Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met

35 40 45

Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr

50 55 60

Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe

65 70 75 80

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

85 90 95

Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp

100 105 110

Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile

115 120 125

Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys

130 135 140

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

Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe
210 215 220
Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr
225 230 235 240
Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp
245 250 255
Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr
260 265 270

Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile Thr Asp Ala Val Asp
275 280 285
Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe
290 295 300
Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val Val Pro
305 310 315 320
Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe
325 330 335

Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu
340 345 350
Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn
355 360 365
Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys
370 375 380

Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala Asp Ser Phe Val Val
385 390 395 400

Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Val Ile
405 410 415
Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met Gly Cys Val Leu
420 425 430
Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn
435 440 445
Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu Arg
450 455 460

Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro
465 470 475 480
Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr
485 490 495
Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe
500 505 510
Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser Thr
515 520 525

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

Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu
595 600 605
Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp
610 615 620
Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe

625 630 635 640

Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser

645 650 655

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

660 665 670

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

675 680 685

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

690 695 700

Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met

705 710 715 720

Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys

725 730 735

Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe

740 745 750

Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp

755 760 765

Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr

770 775 780

Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro

785 790 795 800

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

805 810 815

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

820 825 830

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

835 840 845

Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala

850 855 860

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

865 870 875 880

Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
885 890 895
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
900 905 910

Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln
915 920 925
Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp Val
930 935 940
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
945 950 955 960
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
965 970 975

Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly
980 985 990
Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala
995 1000 1005
Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser
1010 1015 1020
Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys
1025 1030 1035

Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His Gly Val
1040 1045 1050
Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn Phe
1055 1060 1065
Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
1070 1075 1080
Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr
1085 1090 1095

Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr
1100 1105 1110
Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln

1160 1165 1170

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

1175 1180 1185

Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile

1190 1195 1200

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

1205 1210 1215

Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys Ser Asn Gly Ser

1220 1225 1230

Leu Gln Cys Arg Ile Cys Ile

1235 1240

<210> 97

<211> 1243

<212> PRT

<213> Artificial Sequence

<220><223> PDI-SARS-COV-1 H1cCT-AA

<400> 97

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Ser Asp Leu Asp Arg Cys Thr Thr

20 25 30

Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met

35 40 45

Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr

50 55 60

Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe

65 70 75 80

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

85 90 95

Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp

100 105 110

Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile

115 120 125

Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys

130 135 140

Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr His

145 150 155 160

Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser

165 170 175

Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys His

180 185 190

Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr

195 200 205

Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe

210 215 220

Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr

225 230 235 240

Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp

245 250 255

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

260 265 270

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

275 280 285

Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe

290 295 300

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

305 310 315 320

Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe
325 330 335
Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu

340 345 350
Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn
355 360 365
Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys
370 375 380
Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala Asp Ser Phe Val Val
385 390 395 400
Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Val Ile

405 410 415
Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met Gly Cys Val Leu
420 425 430
Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn
435 440 445
Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu Arg
450 455 460
Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro

465 470 475 480
Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr
485 490 495
Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe
500 505 510
Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser Thr
515 520 525
Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu Thr

530 535 540
Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg Phe Gln Pro Phe Gln
545 550 555 560
Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp Ser Val Arg Asp Pro

565 570 575

Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ser Phe Gly Gly Val

580 585 590

Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu

595 600 605

Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp

610 615 620

Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe

625 630 635 640

Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser

645 650 655

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

660 665 670

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

675 680 685

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

690 695 700

Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met

705 710 715 720

Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys

725 730 735

Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe

740 745 750

Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp

755 760 765

Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr

770 775 780

Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro

785 790 795 800

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

805 810 815

Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly Glu
820 825 830
Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe
835 840 845
Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala

850 855 860
Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp Thr
865 870 875 880
Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
885 890 895
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn
900 905 910
Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln

915 920 925
Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp Val
930 935 940
Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
945 950 955 960
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg
965 970 975
Leu Asp Pro Pro Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly

980 985 990
Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala
995 1000 1005
Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser
1010 1015 1020
Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys
1025 1030 1035
Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His Gly Val

1040 1045 1050
Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn Phe

1055 1060 1065

Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro

1070 1075 1080

Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr

1085 1090 1095

Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln

1160 1165 1170

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

1175 1180 1185

Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile

1190 1195 1200

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

1205 1210 1215

Ala Ile Val Met Val Thr Ile Leu Leu Ser Phe Trp Met Cys Ser

1220 1225 1230

Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1235 1240

<210> 98

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> IF(AvB+wtCT-MERS).r

<400> 98

acgacacgac taaggccttc agtgaacgtg gaccttgtga ggctcaaggt catactcctc 60

<210> 99

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> IF(H1cCT-wtTM).r

<400> 99

acgacacgac taaggccttc aaatacatat tctacactgt agagaccca 49

<210> 100

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> IF(H5ITMCT).r

<400> 100

acgacacgac taaggccttc aaatgcaaat tctgcattgt aacgatcc 48

<210> 101

<211> 4083

<212> DNA

<213> Artificial Sequence

<220><223> PDI-MERS-wtTMCT-DNA

<400> 101

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtatgtcga tgtgggtccc gatagtgtta agtccgcctg catcgaagtg 120

gacattcagc agaccttctt cgataagact tggcctcggc caattgatgt gtccaaggcc 180

gacggcatta tctacccccа aggtcggaca tattccaaca taactatcac ctatcagggg 240

ctattccctt atcagggcga ccatggggac atgtacgttt acagcgctgg tcacgctaca 300

gggacgaccc cccagaagct cttcgtggcg aactatagtc aggacgtgaa acagtttgcc 360

aacggttttg tagtgcgcat cggggcagcc gctaactcca ctggtactgt tattatcagc 420

ccttccacga gtgccacaat tcgaaagatc tatccggcct tcatgctagg atcctctgtg 480

ggcaatttta gcgacggtaa gatgggtcgg ttcttcaacc acacgcttgt gctgcttccc 540

gatgggtgcg gtactttgct gagggccttt tactgtatcc tagagccccg atccggcaac 600

cactgccccg ccgggaactc gtatacttcc tttgccactt atcatactcc agccacggat 660

tgtagcgatg ggaactacaa taggaacgcc agtttgaatt cctttaaaga gtacttcaac 720

ttgcggaatt gtaccttcat gtatacatat aacattactg aggacgaaat tctcgaatgg 780

ttcggaatca ctcaaacagc ccagggagtg cacctcttta gttctcgcta tgtggactta 840

tatggaggca atatgtttca attcgccacc ttacccgtct acgatacgat caagtattac 900

tcgatcatac cccactccat taggtccatt cagagcgatc gcaaggcatg ggccgcattc 960

tatgtgtata agctccagcc cctgaccttc ctcttggatt tctccgtgga cggctacatc 1020

agaagggcta tcgattgcgg gttcaacgac ctcagccagc tgcattgttc ttatgagagc 1080

tttgacgtgg aaagcggagt ttactcagtc tcttcctttg aggctaaacc ttcaggtagc 1140

gtcgtagagc aagcagaggg tgtggagtgc gatttctcac cactgctcag cggaacccca 1200

ccccaggtct acaactttaa gcggctcgtg ttcacaaact gtaactataa cttgactaag 1260

ttgctgtcac tcttttccgt gaatgatttt acatgctccc aaatcagccc agccgctatt 1320

gcgtctaatt gctattcctc attgatcctg gattacttca gttacccccc ctctatgaag 1380

agcgatctct cggttagtag cgctgggcct atttcccagt ttaactacaa acaatccttt 1440

tccaatccaa catgcctgat cttagctact gtaccccaca acctgactac tattacgaag 1500

ccactcaagt actcatacat taataagtgc agccgattcc tcagtgatga tcgcaccgaa 1560

gtgccgcagc ttgtaaacgc gaaccagtac tccccatgcg tctctattgt gccttctaca 1620

gtgtgggaag acggcgatta ttatagaaag cagctgtcgc cactggaagg tggcgggtgg 1680

ctagttgcca gtgggtccac agttgccatg accgagcaac ttcagatggg gtttggcata 1740

acagtgcagt atggtaccga tacgaacagc gtgtgtccaa aattggaatt tgctaacgac 1800

accaagatcg cctcccagtt gggaaattgt gttgaatatt ccctgtacgg agtgtcaggc 1860

cgggggggtgt tccaaaattg caccgccgtg ggagtgaggc agcaaagatt cgtgtacgac 1920

gcataccaga atctagtcgg atactattct gacgatggaa actactactg tctgcgcgct 1980

tgcgtctcag tgcccgtgag tgtcatatat gataaggaga ccaagactca cgctactctc 2040

tttggttctg tcgcgtgcga acacatttcc tctacaatgt cccagtatag tcgctccact 2100

cggtctatgt taaagcgcag agacagtacc tacggccctc tacagacacc tgtggggtgc 2160

gttctcggcc ttgtcaattc tagcctgttt gtggaggatt gtaagctgcc ccttggtcaa 2220

agcttatgcg cactgcccga tacgcccagc acacttacac cagcttcagt ggggtccgtc 2280

cccggggaaa tgagattggc ctcgatcgct ttcaaccacc ccatacaggt ggatcagctc 2340

aactcgtcat acttcaagct aagcatccct actaatttct cctttggtgt gactcaggag 2400

tacattcaga ccacaattca aaaggtgacc gttgactgca agcagtatgt gtgcaacggg 2460

ttccagaaat gtgaacagct gctccgggag tatggccagt tctgttctaa aatcaaccag 2520

gccctccacg gagcaaacct taggcaggac gattctgtca gaaacctctt tgccagcgtc 2580

aagagttctc agagttcccc tattatacct ggcttcggcg gggatttcaa cctgacacta 2640

cttgaacctg taagcatatc aaccggaagt cgcagtgccc gttccgccat cgaggatctg 2700

ctcttcgaca aagtaactat tgcagatccc ggatacatgc aggggtatga cgactgcatg 2760

cagcagggtc cagcctctgc aagggatctg atatgcgcac agtatgtcgc tgggtacaaa 2820

gtgttgcctc ctctcatgga cgtgaacatg gaagcggcct atacctcctc acttctaggc 2880

tccatagcgg gcgtgggatg gaccgcaggg ctttcaagct tcgccgcaat tccctttgct 2940

caatctatct tctacaggct taatggcgtt ggaatcaccc agcaggtgtt aagcgaaaac 3000

cagaaattga ttgccaataa gtttaaccaa gctttggggg ccatgcagac aggctttaca 3060

accacaaacg aggctttcca taaagtacag gatgcggtaa acaataacgc acaagccctg 3120

tcaaagctgg cttcagagct ctcaaataca tttggcgcta tatccgcgtc tatcggcgat 3180

atcatacaac ggttggaccc acccgaacag gacgcacaga ttgatcgttt gatcaacggg 3240

aggcttacca ccttaaacgc ttttgtggcc cagcaactgg tgcggtctga gagcgccgcc 3300

ttgagcgctc agctggcaaa ggataaagtg aatgaatgcg tgaaagctca atcaaagaga 3360

agtgggtttt gtgggcaggg tactcatatt gtttcctttg tggtgaacgc cccaaatgga 3420

ctctacttta tgcatgttgg atactacccg agcaaccaca tcgaggtcgt ttccgcctat 3480

gggctttgtg acgcagcaaa ccctactaac tgtatcgcgc cagttaatgg ctactttatt 3540

aaaacaaata acacacgcat tgtggatgaa tggagttaca cagggtccag cttctacgct 3600

ccagagccta tcacctctct gaacacaaag tatgtggcac ctcaggtcac atatcagaac 3660

atctcgacaa acctgccccc cccactcttg ggcaactcca cagggatcga cttccaggac 3720

gagcttgacg aattcttcaa gaacgtgtcc accagtatcc ctaattttgg ttcgctgacc 3780

caaattaaca caaccctgct cgatctgaca tatgaaatgc tttcactaca gcaggtggtc 3840

aaagcgttga acgagtcgta tatcgacctg aaagagttag ggaattacac atactataac 3900

aaatggccct ggtatatttg gttaggattc attgccgggc tggtggccct tgccttgtgc 3960

gtatttttca tcttgtgctg taccggttgc ggtacgaatt gcatgggaaa actgaaatgt 4020

aatcggtgct gcgatcgcta tgaggagtat gaccttgagc ctcacaaggt ccacgttcac 4080

tga 4083

<210> 102

<211> 4032

<212> DNA

<213> Artificial Sequence

<220><223> PDI-MERS-H5iTMCT-DNA

<400> 102

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtatgtcga tgtgggtccc gatagtgtta agtccgcctg catcgaagtg 120

gacattcagc agaccttctt cgataagact tggcctcggc caattgatgt gtccaaggcc 180

gacggcatta tctaccccca aggtcggaca tattccaaca taactatcac ctatcagggg 240

ctattccctt atcagggcga ccatggggac atgtacgttt acagcgctgg tcacgctaca 300

gggacgaccc cccagaagct cttcgtggcg aactatagtc aggacgtgaa acagtttgcc 360

aacggttttg tagtgcgcat cggggcagcc gctaactcca ctggtactgt tattatcagc 420

ccttccacga gtgccacaat tcgaaagatc tatccggcct tcatgctagg atcctctgtg 480

ggcaatttta gcgacggtaa gatgggtcgg ttcttcaacc acacgcttgt gctgcttccc 540

gatgggtgcg gtactttgct gagggccttt tactgtatcc tagagccccg atccggcaac 600

cactgccccg ccgggaactc gtatacttcc tttgccactt atcatactcc agccacggat 660

tgtagcgatg ggaactacaa taggaacgcc agtttgaatt cctttaaaga gtacttcaac 720

ttgcggaatt gtaccttcat gtatacatat aacattactg aggacgaaat tctcgaatgg 780

ttcggaatca ctcaaacagc ccagggagtg cacctcttta gttctcgcta tgtggactta 840

tatggaggca atatgtttca attcgccacc ttacccgtct acgatacgat caagtattac 900

tcgatcatac cccactccat taggtccatt cagagcgatc gcaaggcatg ggccgcattc 960

tatgtgtata agctccagcc cctgaccttc ctcttggatt tctccgtgga cggctacatc 1020

agaagggcta tcgattgcgg gttcaacgac ctcagccagc tgcattgttc ttatgagagc 1080

tttgacgtgg aaagcggagt ttactcagtc tcttcctttg aggctaaacc ttcaggtagc 1140

gtcgtagagc aagcagaggg tgtggagtgc gatttctcac cactgctcag cggaacccca 1200

ccccaggtct acaactttaa gcggctcgtg ttcacaaact gtaactataa cttgactaag 1260

ttgctgtcac tcttttccgt gaatgatttt acatgctccc aaatcagccc agccgctatt 1320

gcgtctaatt gctattcctc attgatcctg gattacttca gttacccccc ctctatgaag 1380

agcgatctct cggttagtag cgctgggcct atttcccagt ttaactacaa acaatccttt 1440

tccaatccaa catgcctgat cttagctact gtaccccaca acctgactac tattacgaag 1500

ccactcaagt actcatacat taataagtgc agccgattcc tcagtgatga tcgcaccgaa 1560

gtgccgcagc ttgtaaacgc gaaccagtac tccccatgcg tctctattgt gccttctaca 1620

gtgtgggaag acggcgatta ttatagaaag cagctgtcgc cactggaagg tggcgggtgg 1680

ctagttgcca gtgggtccac agttgccatg accgagcaac ttcagatggg gtttggcata 1740

acagtgcagt atggtaccga tacgaacagc gtgtgtccaa aattggaatt tgctaacgac 1800

accaagatcg cctcccagtt gggaaattgt gttgaatatt ccctgtacgg agtgtcaggc 1860

cgggggggtgt tccaaaattg caccgccgtg ggagtgaggc agcaaagatt cgtgtacgac 1920
gcataccaga atctagtcgg atactattct gacgatggaa actactactg tctgcgcgct 1980
tgcgtctcag tgcccgtgag tgtcatatat gataaggaga ccaagactca cgctactctc 2040
tttggttctg tcgcgtgcga acacatttcc tctacaatgt cccagtatag tcgctccact 2100
cggtctatgt taaagcgcag agacagtacc tacggccctc tacagacacc tgtggggtgc 2160
gttctcggcc ttgtcaattc tagcctgttt gtggaggatt gtaagctgcc ccttggtcaa 2220
agcttatgcg cactgcccga tacgcccagc acacttacac cagcttcagt ggggtccgtc 2280

cccggggaaa tgagattggc ctcgatcgct ttcaaccacc ccatacaggt ggatcagctc 2340
aactcgtcat acttcaagct aagcatccct actaatttct cctttggtgt gactcaggag 2400
tacattcaga ccacaattca aaaggtgacc gttgactgca agcagtatgt gtgcaacggg 2460
ttccagaaat gtgaacagct gctccgggag tatggccagt tctgttctaa aatcaaccag 2520
gccctccacg gagcaaacct taggcaggac gattctgtca gaaacctctt tgccagcgtc 2580
aagagttctc agagttcccc tattatacct ggcttcggcg gggatttcaa cctgacacta 2640
cttgaacctg taagcatatc aaccggaagt cgcagtgccc gttccgccat cgaggatctg 2700

ctcttcgaca aagtaactat tgcagatccc ggatacatgc aggggtatga cgactgcatg 2760
cagcagggtc cagcctctgc aagggatctg atatgcgcac agtatgtcgc tgggtacaaa 2820
gtgttgcctc ctctcatgga cgtgaacatg gaagcggcct atacctcctc acttctaggc 2880
tccatagcgg gcgtgggatg gaccgcaggg ctttcaagct tcgccgcaat tccctttgct 2940
caatctatct tctacaggct taatggcgtt ggaatcaccc agcaggtgtt aagcgaaaac 3000
cagaaattga ttgccaataa gtttaaccaa gctttggggg ccatgcagac aggctttaca 3060
accacaaacg aggctttcca taaagtacag gatgcggtaa acaataacgc acaagccctg 3120

tcaaagctgg cttcagagct ctcaaataca tttggcgcta tatccgcgtc tatcggcgat 3180
atcatacaac ggttggaccc acccgaacag gacgcacaga ttgatcgttt gatcaacggg 3240
aggcttacca ccttaaacgc ttttgtggcc cagcaactgg tgcggtctga gagcgccgcc 3300
ttgagcgctc agctggcaaa ggataaagtg aatgaatgcg tgaaagctca atcaaagaga 3360
agtgggtttt gtgggcaggg tactcatatt gtttcctttg tggtgaacgc cccaaatgga 3420
ctctacttta tgcatgttgg atactacccg agcaaccaca tcgaggtcgt ttccgcctat 3480
gggctttgtg acgcagcaaa ccctactaac tgtatcgcgc cagttaatgg ctactttatt 3540

aaaacaaata acacacgcat tgtggatgaa tggagttaca cagggtccag cttctacgct 3600

ccagagccta tcacctctct gaacacaaag tatgtggcac ctcaggtcac atatcagaac 3660

atctcgacaa acctgccccc cccactcttg ggcaactcca cagggatcga cttccaggac 3720

gagcttgacg aattcttcaa gaacgtgtcc accagtatcc ctaattttgg ttcgctgacc 3780

caaattaaca caaccctgct cgatctgaca tatgaaatgc tttcactaca gcaggtggtc 3840

aaagcgttga acgagtcgta tatcgacctg aaagagttag ggaattacac atactataac 3900

aaatggccct ggtatcaaat actgtcaatt tattcaacag tggcgagttc cctagcactg 3960

gcaatcatga tggctggtct atctttatgg atgtgctcca atggatcgtt acaatgcaga 4020

atttgcattt ga 4032

<210> 103

<211> 4026

<212> DNA

<213> Artificial Sequence

<220><223> PDI-MERS-H5iCT-DNA

<400> 103

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtatgtcga tgtgggtccc gatagtgtta agtccgcctg catcgaagtg 120

gacattcagc agaccttctt cgataagact tggcctcggc caattgatgt gtccaaggcc 180

gacggcatta tctacccccа aggtcggaca tattccaaca taactatcac ctatcagggg 240

ctattccctt atcagggcga ccatggggac atgtacgttt acagcgctgg tcacgctaca 300

gggacgaccc cccagaagct cttcgtggcg aactatagtc aggacgtgaa acagtttgcc 360

aacggttttg tagtgcgcat cggggcagcc gctaactcca ctggtactgt tattatcagc 420

ccttccacga gtgccacaat tcgaaagatc tatccggcct tcatgctagg atcctctgtg 480

ggcaatttta gcgacggtaa gatgggtcgg ttcttcaacc acacgcttgt gctgcttccc 540

gatgggtgcg gtactttgct gagggccttt tactgtatcc tagagccccg atccggcaac 600

cactgccccg ccgggaactc gtatacttcc tttgccactt atcatactcc agccacggat 660

tgtagcgatg ggaactacaa taggaacgcc agtttgaatt cctttaaaga gtacttcaac 720

ttgcggaatt gtaccttcat gtatacatat aacattactg aggacgaaat tctcgaatgg 780

ttcggaatca ctcaaacagc ccagggagtg cacctcttta gttctcgcta tgtggactta 840

tatggaggca atatgtttca attcgccacc ttacccgtct acgatacgat caagtattac 900

tcgatcatac cccactccat taggtccatt cagagcgatc gcaaggcatg ggccgcattc 960

tatgtgtata agctccagcc cctgaccttc ctcttggatt tctccgtgga cggctacatc 1020

agaagggcta tcgattgcgg gttcaacgac ctcagccagc tgcattgttc ttatgagagc 1080

tttgacgtgg aaagcggagt ttactcagtc tcttcctttg aggctaaacc ttcaggtagc 1140
gtcgtagagc aagcagaggg tgtggagtgc gatttctcac cactgctcag cggaacccca 1200
ccccaggtct acaactttaa gcggctcgtg ttcacaaact gtaactataa cttgactaag 1260
ttgctgtcac tcttttccgt gaatgatttt acatgctccc aaatcagccc agccgctatt 1320
gcgtctaatt gctattcctc attgatcctg gattacttca gttacccccc ctctatgaag 1380
agcgatctct cggttagtag cgctgggcct atttcccagt ttaactacaa acaatccttt 1440
tccaatccaa catgcctgat cttagctact gtaccccaca acctgactac tattacgaag 1500

ccactcaagt actcatacat taataagtgc agccgattcc tcagtgatga tcgcaccgaa 1560
gtgccgcagc ttgtaaacgc gaaccagtac tccccatgcg tctctattgt gccttctaca 1620
gtgtgggaag acggcgatta ttatagaaag cagctgtcgc cactggaagg tggcgggtgg 1680
ctagttgcca gtgggtccac agttgccatg accgagcaac ttcagatggg gtttggcata 1740
acagtgcagt atggtaccga tacgaacagc gtgtgtccaa aattggaatt tgctaacgac 1800
accaagatcg cctcccagtt gggaaattgt gttgaatatt ccctgtacgg agtgtcaggc 1860
cgggggggtgt tccaaaattg caccgccgtg ggagtgaggc agcaaagatt cgtgtacgac 1920

gcataccaga atctagtcgg atactattct gacgatggaa actactactg tctgcgcgct 1980
tgcgtctcag tgcccgtgag tgtcatatat gataaggaga ccaagactca cgctactctc 2040
tttggttctg tcgcgtgcga acacatttcc tctacaatgt cccagtatag tcgctccact 2100
cggtctatgt taaagcgcag agacagtacc tacggccctc tacagacacc tgtggggtgc 2160
gttctcggcc ttgtcaattc tagcctgttt gtggaggatt gtaagctgcc ccttggtcaa 2220
agcttatgcg cactgcccga tacgcccagc acacttacac cagcttcagt ggggtccgtc 2280
cccggggaaa tgagattggc ctcgatcgct ttcaaccacc ccatacaggt ggatcagctc 2340

aactcgtcat acttcaagct aagcatccct actaatttct cctttggtgt gactcaggag 2400
tacattcaga ccacaattca aaaggtgacc gttgactgca agcagtatgt gtgcaacggg 2460
ttccagaaat gtgaacagct gctccgggag tatggccagt tctgttctaa aatcaaccag 2520
gccctccacg gagcaaacct taggcaggac gattctgtca gaaacctctt tgccagcgtc 2580
aagagttctc agagttcccc tattatacct ggcttcggcg gggatttcaa cctgacacta 2640
cttgaacctg taagcatatc aaccggaagt cgcagtgccc gttccgccat cgaggatctg 2700
ctcttcgaca aagtaactat tgcagatccc ggatacatgc aggggtatga cgactgcatg 2760

cagcagggtc cagcctctgc aagggatctg atatgcgcac agtatgtcgc tgggtacaaa 2820

gtgttgcctc ctctcatgga cgtgaacatg gaagcggcct atacctcctc acttctaggc 2880

tccatagcgg gcgtgggatg gaccgcaggg ctttcaagct tcgccgcaat tccctttgct 2940

caatctatct tctacaggct taatggcgtt ggaatcaccc agcaggtgtt aagcgaaaac 3000

cagaaattga ttgccaataa gtttaaccaa gctttggggg ccatgcagac aggctttaca 3060

accacaaacg aggctttcca taaagtacag gatgcggtaa acaataacgc acaagccctg 3120

tcaaagctgg cttcagagct ctcaaataca tttggcgcta tatccgcgtc tatcggcgat 3180

atcatacaac ggttggaccc acccgaacag gacgcacaga ttgatcgttt gatcaacggg 3240

aggcttacca ccttaaacgc ttttgtggcc cagcaactgg tgcggtctga gagcgccgcc 3300

ttgagcgctc agctggcaaa ggataaagtg aatgaatgcg tgaaagctca atcaaagaga 3360

agtgggtttt gtgggcaggg tactcatatt gtttcctttg tggtgaacgc cccaaatgga 3420

ctctacttta tgcatgttgg atactacccg agcaaccaca tcgaggtcgt ttccgcctat 3480

gggctttgtg acgcagcaaa ccctactaac tgtatcgcgc cagttaatgg ctactttatt 3540

aaaacaaata acacacgcat tgtggatgaa tggagttaca cagggtccag cttctacgct 3600

ccagagccta tcacctctct gaacacaaag tatgtggcac ctcaggtcac atatcagaac 3660

atctcgacaa acctgccccc cccactcttg ggcaactcca cagggatcga cttccaggac 3720

gagcttgacg aattcttcaa gaacgtgtcc accagtatcc ctaattttgg ttcgctgacc 3780

caaattaaca caaccctgct cgatctgaca tatgaaatgc tttcactaca gcaggtggtc 3840

aaagcgttga acgagtcgta tatcgacctg aaagagttag ggaattacac atactataac 3900

aaatggccct ggtatatttg gttaggattc attgccgggc tggtggccct tgccttgtgc 3960

gtatttttca tcttgtcttt atggatgtgc tccaatggat cgttacaatg cagaatttgc 4020

atttga 4026

<210> 104

<211> 4017

<212> DNA

<213> Artificial Sequence

<220><223> PDI-MERS-H5iCT(V4)-DNA

<400> 104

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtatgtcga tgtgggtccc gatagtgtta agtccgcctg catcgaagtg 120

gacattcagc agaccttctt cgataagact tggcctcggc caattgatgt gtccaaggcc 180

gacggcatta tctaccccca aggtcggaca tattccaaca taactatcac ctatcagggg 240

ctattccctt atcagggcga ccatggggac atgtacgttt acagcgctgg tcacgctaca 300

gggacgaccc cccagaagct cttcgtggcg aactatagtc aggacgtgaa acagtttgcc 360
aacggttttg tagtgcgcat cggggcagcc gctaactcca ctggtactgt tattatcagc 420
ccttccacga gtgccacaat tcgaaagatc tatccggcct tcatgctagg atcctctgtg 480
ggcaatttta gcgacggtaa gatgggtcgg ttcttcaacc acacgcttgt gctgcttccc 540
gatgggtgcg gtactttgct gagggccttt tactgtatcc tagagccccg atccggcaac 600
cactgccccg ccgggaactc gtatacttcc tttgccactt atcatactcc agccacggat 660
tgtagcgatg ggaactacaa taggaacgcc agtttgaatt cctttaaaga gtacttcaac 720

ttgcggaatt gtaccttcat gtatacatat aacattactg aggacgaaat tctcgaatgg 780
ttcggaatca ctcaaacagc ccagggagtg cacctcttta gttctcgcta tgtggactta 840
tatggaggca atatgtttca attcgccacc ttacccgtct acgatacgat caagtattac 900
tcgatcatac cccactccat taggtccatt cagagcgatc gcaaggcatg ggccgcattc 960
tatgtgtata agctccagcc cctgaccttc ctcttggatt tctccgtgga cggctacatc 1020
agaagggcta tcgattgcgg gttcaacgac ctcagccagc tgcattgttc ttatgagagc 1080
tttgacgtgg aaagcggagt ttactcagtc tcttcctttg aggctaaacc ttcaggtagc 1140

gtcgtagagc aagcagaggg tgtggagtgc gatttctcac cactgctcag cggaacccca 1200
ccccaggtct acaactttaa gcggctcgtg ttcacaaact gtaactataa cttgactaag 1260
ttgctgtcac tcttttccgt gaatgatttt acatgctccc aaatcagccc agccgctatt 1320
gcgtctaatt gctattcctc attgatcctg gattacttca gttacccccc ctctatgaag 1380
agcgatctct cggttagtag cgctgggcct atttcccagt ttaactacaa acaatccttt 1440
tccaatccaa catgcctgat cttagctact gtaccccaca acctgactac tattacgaag 1500
ccactcaagt actcatacat taataagtgc agccgattcc tcagtgatga tcgcaccgaa 1560

gtgccgcagc ttgtaaacgc gaaccagtac tccccatgcg tctctattgt gccttctaca 1620
gtgtgggaag acggcgatta ttatagaaag cagctgtcgc cactggaagg tggcgggtgg 1680
ctagttgcca gtgggtccac agttgccatg accgagcaac ttcagatggg gtttggcata 1740
acagtgcagt atggtaccga tacgaacagc gtgtgtccaa aattggaatt tgctaacgac 1800
accaagatcg cctcccagtt gggaaattgt gttgaatatt ccctgtacgg agtgtcaggc 1860
cgggggggtgt tccaaaattg caccgccgtg ggagtgaggc agcaaagatt cgtgtacgac 1920
gcataccaga atctagtcgg atactattct gacgatggaa actactactg tctgcgcgct 1980

tgcgtctcag tgcccgtgag tgtcatatat gataaggaga ccaagactca cgctactctc 2040

tttggttctg tcgcgtgcga acacatttcc tctacaatgt cccagtatag tcgctccact 2100

cggtctatgt taaagcgcag agacagtacc tacggccctc tacagacacc tgtggggtgc 2160

gttctcggcc ttgtcaattc tagcctgttt gtggaggatt gtaagctgcc ccttggtcaa 2220

agcttatgcg cactgcccga tacgcccagc acacttacac cagcttcagt ggggtccgtc 2280

cccggggaaa tgagattggc ctcgatcgct ttcaaccacc ccatacaggt ggatcagctc 2340

aactcgtcat acttcaagct aagcatccct actaatttct cctttggtgt gactcaggag 2400

tacattcaga ccacaattca aaaggtgacc gttgactgca agcagtatgt gtgcaacggg 2460

ttccagaaat gtgaacagct gctccgggag tatggccagt tctgttctaa aatcaaccag 2520

gccctccacg gagcaaacct taggcaggac gattctgtca gaaacctctt tgccagcgtc 2580

aagagttctc agagttcccc tattatacct ggcttcggcg gggatttcaa cctgacacta 2640

cttgaacctg taagcatatc aaccggaagt cgcagtgccc gttccgccat cgaggatctg 2700

ctcttcgaca aagtaactat tgcagatccc ggatacatgc aggggtatga cgactgcatg 2760

cagcagggtc cagcctctgc aagggatctg atatgcgcac agtatgtcgc tgggtacaaa 2820

gtgttgcctc ctctcatgga cgtgaacatg gaagcggcct atacctcctc acttctaggc 2880

tccatagcgg gcgtgggatg gaccgcaggg ctttcaagct tcgccgcaat tccctttgct 2940

caatctatct tctacaggct taatggcgtt ggaatcaccc agcaggtgtt aagcgaaaac 3000

cagaaattga ttgccaataa gtttaaccaa gctttggggg ccatgcagac aggctttaca 3060

accacaaacg aggctttcca taaagtacag gatgcggtaa acaataacgc acaagccctg 3120

tcaaagctgg cttcagagct ctcaaataca tttggcgcta tatccgcgtc tatcggcgat 3180

atcatacaac ggttggaccc acccgaacag gacgcacaga ttgatcgttt gatcaacggg 3240

aggcttacca ccttaaacgc ttttgtggcc cagcaactgg tgcggtctga gagcgccgcc 3300

ttgagcgctc agctggcaaa ggataaagtg aatgaatgcg tgaaagctca atcaaagaga 3360

agtgggtttt gtgggcaggg tactcatatt gtttcctttg tggtgaacgc cccaaatgga 3420

ctctacttta tgcatgttgg atactacccg agcaaccaca tcgaggtcgt ttccgcctat 3480

gggctttgtg acgcagcaaa ccctactaac tgtatcgcgc cagttaatgg ctactttatt 3540

aaaacaaata acacacgcat tgtggatgaa tggagttaca cagggtccag cttctacgct 3600

ccagagccta tcacctctct gaacacaaag tatgtggcac ctcaggtcac atatcagaac 3660

atctcgacaa acctgccccc cccactcttg ggcaactcca cagggatcga cttccaggac 3720

gagcttgacg aattcttcaa gaacgtgtcc accagtatcc ctaattttgg ttcgctgacc 3780

caaattaaca caaccctgct cgatctgaca tatgaaatgc tttcactaca gcaggtggtc 3840

aaagcgttga acgagtcgta tatcgacctg aaagagttag ggaattacac atactataac 3900

aaatggccct ggtatatttg gttaggattc attgccgggc tggtggccct tgccttgtgc 3960

gtatttttca tcttgtgctg ctccaatgga tcgttacaat gcagaatttg catttga 4017

<210> 105

<211> 4026

<212> DNA

<213> Artificial Sequence

<220><223> PDI-MERS-H1cCT-DNA

<400> 105

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgtatgtcga tgtgggtccc gatagtgtta agtccgcctg catcgaagtg 120

gacattcagc agaccttctt cgataagact tggcctcggc caattgatgt gtccaaggcc 180

gacggcatta tctaccccca aggtcggaca tattccaaca taactatcac ctatcagggg 240

ctattccctt atcagggcga ccatggggac atgtacgttt acagcgctgg tcacgctaca 300

gggacgaccc cccagaagct cttcgtggcg aactatagtc aggacgtgaa acagtttgcc 360

aacggttttg tagtgcgcat cggggcagcc gctaactcca ctggtactgt tattatcagc 420

ccttccacga gtgccacaat tcgaaagatc tatccggcct tcatgctagg atcctctgtg 480

ggcaatttta gcgacggtaa gatgggtcgg ttcttcaacc acacgcttgt gctgcttccc 540

gatgggtgcg gtactttgct gagggccttt tactgtatcc tagagccccg atccggcaac 600

cactgccccg ccgggaactc gtatacttcc tttgccactt atcatactcc agccacggat 660

tgtagcgatg ggaactacaa taggaacgcc agtttgaatt cctttaaaga gtacttcaac 720

ttgcggaatt gtaccttcat gtatacatat aacattactg aggacgaaat tctcgaatgg 780

ttcggaatca ctcaaacagc ccagggagtg cacctcttta gttctcgcta tgtggactta 840

tatggaggca atatgtttca attcgccacc ttacccgtct acgatacgat caagtattac 900

tcgatcatac cccactccat taggtccatt cagagcgatc gcaaggcatg ggccgcattc 960

tatgtgtata agctccagcc cctgaccttc ctcttggatt tctccgtgga cggctacatc 1020

agaagggcta tcgattgcgg gttcaacgac ctcagccagc tgcattgttc ttatgagagc 1080

tttgacgtgg aaagcggagt ttactcagtc tcttcctttg aggctaaacc ttcaggtagc 1140

gtcgtagagc aagcagaggg tgtggagtgc gatttctcac cactgctcag cggaacccca 1200

ccccaggtct acaactttaa gcggctcgtg ttcacaaact gtaactataa cttgactaag 1260

ttgctgtcac tcttttccgt gaatgatttt acatgctccc aaatcagccc agccgctatt 1320

gcgtctaatt gctattcctc attgatcctg gattacttca gttacccccct ctctatgaag 1380

agcgatctct cggttagtag cgctgggcct atttcccagt ttaactacaa acaatccttt 1440

tccaatccaa catgcctgat cttagctact gtaccccaca acctgactac tattacgaag 1500

ccactcaagt actcatacat taataagtgc agccgattcc tcagtgatga tcgcaccgaa 1560

gtgccgcagc ttgtaaacgc gaaccagtac tccccatgcg tctctattgt gccttctaca 1620

gtgtgggaag acggcgatta ttatagaaag cagctgtcgc cactggaagg tggcgggtgg 1680

ctagttgcca gtgggtccac agttgccatg accgagcaac ttcagatggg gtttggcata 1740

acagtgcagt atggtaccga tacgaacagc gtgtgtccaa aattggaatt tgctaacgac 1800

accaagatcg cctcccagtt gggaaattgt gttgaatatt ccctgtacgg agtgtcaggc 1860

cgggggggtgt tccaaaattg caccgccgtg ggagtgaggc agcaaagatt cgtgtacgac 1920

gcataccaga atctagtcgg atactattct gacgatggaa actactactg tctgcgcgct 1980

tgcgtctcag tgcccgtgag tgtcatatat gataaggaga ccaagactca cgctactctc 2040

tttggttctg tcgcgtgcga acacatttcc tctacaatgt cccagtatag tcgctccact 2100

cggtctatgt taaagcgcag agacagtacc tacggccctc tacagacacc tgtggggtgc 2160

gttctcggcc ttgtcaattc tagcctgttt gtggaggatt gtaagctgcc ccttggtcaa 2220

agcttatgcg cactgcccga tacgcccagc acacttacac cagcttcagt ggggtccgtc 2280

cccggggaaa tgagattggc ctcgatcgct ttcaaccacc ccatacaggt ggatcagctc 2340

aactcgtcat acttcaagct aagcatccct actaatttct cctttggtgt gactcaggag 2400

tacattcaga ccacaattca aaaggtgacc gttgactgca agcagtatgt gtgcaacggg 2460

ttccagaaat gtgaacagct gctccgggag tatggccagt tctgttctaa aatcaaccag 2520

gccctccacg gagcaaacct taggcaggac gattctgtca gaaacctctt tgccagcgtc 2580

aagagttctc agagttcccc tattatacct ggcttcggcg gggatttcaa cctgacacta 2640

cttgaacctg taagcatatc aaccggaagt cgcagtgccc gttccgccat cgaggatctg 2700

ctcttcgaca aagtaactat tgcagatccc ggatacatgc aggggtatga cgactgcatg 2760

cagcagggtc cagcctctgc aagggatctg atatgcgcac agtatgtcgc tgggtacaaa 2820

gtgttgcctc ctctcatgga cgtgaacatg gaagcggcct atacctcctc acttctaggc 2880

tccatagcgg gcgtgggatg gaccgcaggg ctttcaagct tcgccgcaat tccctttgct 2940

caatctatct tctacaggct taatggcgtt ggaatcaccc agcaggtgtt aagcgaaaac 3000

cagaaattga ttgccaataa gtttaaccaa gctttggggg ccatgcagac aggctttaca 3060

accacaaacg aggctttcca taaagtacag gatgcggtaa acaataacgc acaagccctg 3120

tcaaagctgg cttcagagct ctcaaataca tttggcgcta tatccgcgtc tatcggcgat 3180

atcatacaac ggttggaccc acccgaacag gacgcacaga ttgatcgttt gatcaacggg 3240

aggcttacca ccttaaacgc ttttgtggcc cagcaactgg tgcggtctga gagcgccgcc 3300

ttgagcgctc agctggcaaa ggataaagtg aatgaatgcg tgaaagctca atcaaagaga 3360

agtgggtttt gtgggcaggg tactcatatt gtttcctttg tggtgaacgc cccaaatgga 3420

ctctacttta tgcatgttgg atactacccg agcaaccaca tcgaggtcgt ttccgcctat 3480

gggctttgtg acgcagcaaa ccctactaac tgtatcgcgc cagttaatgg ctactttatt 3540

aaaacaaata acacacgcat tgtggatgaa tggagttaca cagggtccag cttctacgct 3600

ccagagccta tcacctctct gaacacaaag tatgtggcac ctcaggtcac atatcagaac 3660

atctcgacaa acctgccccc cccactcttg ggcaactcca cagggatcga cttccaggac 3720

gagcttgacg aattcttcaa gaacgtgtcc accagtatcc ctaattttgg ttcgctgacc 3780

caaattaaca caaccctgct cgatctgaca tatgaaatgc tttcactaca gcaggtggtc 3840

aaagcgttga acgagtcgta tatcgacctg aaagagttag ggaattacac atactataac 3900

aaatggccct ggtatatttg gttaggattc attgccgggc tggtggccct tgccttgtgc 3960

gtatttttca tcttgagctt ctggatgtgc tctaatgggt ctctacagtg tagaatatgt 4020

atttga 4026

<210> 106

<211> 1360

<212> PRT

<213> Artificial Sequence

<220><223> PDI-MERS-wtTMCT-AA

<400> 106

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Tyr Val Asp Val Gly Pro Asp Ser

20 25 30

Val Lys Ser Ala Cys Ile Glu Val Asp Ile Gln Gln Thr Phe Phe Asp

35 40 45

Lys Thr Trp Pro Arg Pro Ile Asp Val Ser Lys Ala Asp Gly Ile Ile

50 55 60

Tyr Pro Gln Gly Arg Thr Tyr Ser Asn Ile Thr Ile Thr Tyr Gln Gly

65 70 75 80

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

130 135 140
Ala Thr Ile Arg Lys Ile Tyr Pro Ala Phe Met Leu Gly Ser Ser Val
145 150 155 160
Gly Asn Phe Ser Asp Gly Lys Met Gly Arg Phe Phe Asn His Thr Leu
165 170 175
Val Leu Leu Pro Asp Gly Cys Gly Thr Leu Leu Arg Ala Phe Tyr Cys
180 185 190
Ile Leu Glu Pro Arg Ser Gly Asn His Cys Pro Ala Gly Asn Ser Tyr

195 200 205
Thr Ser Phe Ala Thr Tyr His Thr Pro Ala Thr Asp Cys Ser Asp Gly
210 215 220
Asn Tyr Asn Arg Asn Ala Ser Leu Asn Ser Phe Lys Glu Tyr Phe Asn
225 230 235 240
Leu Arg Asn Cys Thr Phe Met Tyr Thr Tyr Asn Ile Thr Glu Asp Glu
245 250 255
Ile Leu Glu Trp Phe Gly Ile Thr Gln Thr Ala Gln Gly Val His Leu

260 265 270
Phe Ser Ser Arg Tyr Val Asp Leu Tyr Gly Gly Asn Met Phe Gln Phe
275 280 285
Ala Thr Leu Pro Val Tyr Asp Thr Ile Lys Tyr Tyr Ser Ile Ile Pro
290 295 300
His Ser Ile Arg Ser Ile Gln Ser Asp Arg Lys Ala Trp Ala Ala Phe
305 310 315 320
Tyr Val Tyr Lys Leu Gln Pro Leu Thr Phe Leu Leu Asp Phe Ser Val

325 330 335

Asp Gly Tyr Ile Arg Arg Ala Ile Asp Cys Gly Phe Asn Asp Leu Ser

340 345 350

Gln Leu His Cys Ser Tyr Glu Ser Phe Asp Val Glu Ser Gly Val Tyr

355 360 365

Ser Val Ser Ser Phe Glu Ala Lys Pro Ser Gly Ser Val Val Glu Gln

370 375 380

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

385 390 395 400

Pro Gln Val Tyr Asn Phe Lys Arg Leu Val Phe Thr Asn Cys Asn Tyr

405 410 415

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

420 425 430

Ser Gln Ile Ser Pro Ala Ala Ile Ala Ser Asn Cys Tyr Ser Ser Leu

435 440 445

Ile Leu Asp Tyr Phe Ser Tyr Pro Leu Ser Met Lys Ser Asp Leu Ser

450 455 460

Val Ser Ser Ala Gly Pro Ile Ser Gln Phe Asn Tyr Lys Gln Ser Phe

465 470 475 480

Ser Asn Pro Thr Cys Leu Ile Leu Ala Thr Val Pro His Asn Leu Thr

485 490 495

Thr Ile Thr Lys Pro Leu Lys Tyr Ser Tyr Ile Asn Lys Cys Ser Arg

500 505 510

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

515 520 525

Gln Tyr Ser Pro Cys Val Ser Ile Val Pro Ser Thr Val Trp Glu Asp

530 535 540

Gly Asp Tyr Tyr Arg Lys Gln Leu Ser Pro Leu Glu Gly Gly Gly Trp

545 550 555 560

Leu Val Ala Ser Gly Ser Thr Val Ala Met Thr Glu Gln Leu Gln Met

565 570 575

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

580 585 590

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

595 600 605

Asn Cys Val Glu Tyr Ser Leu Tyr Gly Val Ser Gly Arg Gly Val Phe

610 615 620

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

625 630 635 640

Ala Tyr Gln Asn Leu Val Gly Tyr Tyr Ser Asp Asp Gly Asn Tyr Tyr

645 650 655

Cys Leu Arg Ala Cys Val Ser Val Pro Val Ser Val Ile Tyr Asp Lys

660 665 670

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

675 680 685

Ile Ser Ser Thr Met Ser Gln Tyr Ser Arg Ser Thr Arg Ser Met Leu

690 695 700

Lys Arg Arg Asp Ser Thr Tyr Gly Pro Leu Gln Thr Pro Val Gly Cys

705 710 715 720

Val Leu Gly Leu Val Asn Ser Ser Leu Phe Val Glu Asp Cys Lys Leu

725 730 735

Pro Leu Gly Gln Ser Leu Cys Ala Leu Pro Asp Thr Pro Ser Thr Leu

740 745 750

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

755 760 765

Ile Ala Phe Asn His Pro Ile Gln Val Asp Gln Leu Asn Ser Ser Tyr

770 775 780

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

785 790 795 800

Tyr Ile Gln Thr Thr Ile Gln Lys Val Thr Val Asp Cys Lys Gln Tyr

805 810 815

Val Cys Asn Gly Phe Gln Lys Cys Glu Gln Leu Leu Arg Glu Tyr Gly

820 825 830

Gln Phe Cys Ser Lys Ile Asn Gln Ala Leu His Gly Ala Asn Leu Arg

835 840 845

Gln Asp Asp Ser Val Arg Asn Leu Phe Ala Ser Val Lys Ser Ser Gln

850 855 860

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

865 870 875 880

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

885 890 895

Ile Glu Asp Leu Leu Phe Asp Lys Val Thr Ile Ala Asp Pro Gly Tyr

900 905 910

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

915 920 925

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

930 935 940

Leu Met Asp Val Asn Met Glu Ala Ala Tyr Thr Ser Ser Leu Leu Gly

945 950 955 960

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

965 970 975

Ile Pro Phe Ala Gln Ser Ile Phe Tyr Arg Leu Asn Gly Val Gly Ile

980 985 990

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

995 1000 1005

Asn Gln Ala Leu Gly Ala Met Gln Thr Gly Phe Thr Thr Thr Asn

1010 1015 1020

Glu Ala Phe His Lys Val Gln Asp Ala Val Asn Asn Asn Ala Gln

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Glu Gln Asp Ala Gln Ile Asp Arg Leu Ile Asn Gly Arg Leu Thr

1070 1075 1080

Thr Leu Asn Ala Phe Val Ala Gln Gln Leu Val Arg Ser Glu Ser

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

Pro Val Asn Gly Tyr Phe Ile Lys Thr Asn Asn Thr Arg Ile Val

1175 1180 1185

Asp Glu Trp Ser Tyr Thr Gly Ser Ser Phe Tyr Ala Pro Glu Pro

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Thr Gly Ile Asp Phe Gln Asp Glu Leu Asp Glu Phe Phe Lys Asn

1235 1240 1245

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

1250 1255 1260

Thr Thr Leu Leu Asp Leu Thr Tyr Glu Met Leu Ser Leu Gln Gln

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

Gly Phe Ile Ala Gly Leu Val Ala Leu Ala Leu Cys Val Phe Phe

1310 1315 1320

Ile Leu Cys Cys Thr Gly Cys Gly Thr Asn Cys Met Gly Lys Leu

1325 1330 1335

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

1340 1345 1350

Pro His Lys Val His Val His

1355 1360

<210> 107

<211> 1343

<212> PRT

<213> Artificial Sequence

<220><223> PDI-MERS-H5iTMCT-AA

<400> 107

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Tyr Val Asp Val Gly Pro Asp Ser

20 25 30

Val Lys Ser Ala Cys Ile Glu Val Asp Ile Gln Gln Thr Phe Phe Asp

35 40 45

Lys Thr Trp Pro Arg Pro Ile Asp Val Ser Lys Ala Asp Gly Ile Ile

50 55 60

Tyr Pro Gln Gly Arg Thr Tyr Ser Asn Ile Thr Ile Thr Tyr Gln Gly

65 70 75 80

Leu Phe Pro Tyr Gln Gly Asp His Gly Asp Met Tyr Val Tyr Ser Ala

85 90 95

Gly His Ala Thr Gly Thr Thr Pro Gln Lys Leu Phe Val Ala Asn Tyr

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

Gly Asn Phe Ser Asp Gly Lys Met Gly Arg Phe Phe Asn His Thr Leu

165 170 175

Val Leu Leu Pro Asp Gly Cys Gly Thr Leu Leu Arg Ala Phe Tyr Cys

180 185 190

Ile Leu Glu Pro Arg Ser Gly Asn His Cys Pro Ala Gly Asn Ser Tyr

195 200 205

Thr Ser Phe Ala Thr Tyr His Thr Pro Ala Thr Asp Cys Ser Asp Gly

210 215 220

Asn Tyr Asn Arg Asn Ala Ser Leu Asn Ser Phe Lys Glu Tyr Phe Asn

225 230 235 240

Leu Arg Asn Cys Thr Phe Met Tyr Thr Tyr Asn Ile Thr Glu Asp Glu

245 250 255

Ile Leu Glu Trp Phe Gly Ile Thr Gln Thr Ala Gln Gly Val His Leu

260 265 270

Phe Ser Ser Arg Tyr Val Asp Leu Tyr Gly Gly Asn Met Phe Gln Phe

275 280 285

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

290 295 300

His Ser Ile Arg Ser Ile Gln Ser Asp Arg Lys Ala Trp Ala Ala Phe

305 310 315 320

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

325 330 335

Asp Gly Tyr Ile Arg Arg Ala Ile Asp Cys Gly Phe Asn Asp Leu Ser

340 345 350

Gln Leu His Cys Ser Tyr Glu Ser Phe Asp Val Glu Ser Gly Val Tyr

355 360 365

Ser Val Ser Ser Phe Glu Ala Lys Pro Ser Gly Ser Val Val Glu Gln

370 375 380

Ala Glu Gly Val Glu Cys Asp Phe Ser Pro Leu Leu Ser Gly Thr Pro
385 390 395 400
Pro Gln Val Tyr Asn Phe Lys Arg Leu Val Phe Thr Asn Cys Asn Tyr

405 410 415
Asn Leu Thr Lys Leu Leu Ser Leu Phe Ser Val Asn Asp Phe Thr Cys
420 425 430
Ser Gln Ile Ser Pro Ala Ala Ile Ala Ser Asn Cys Tyr Ser Ser Leu
435 440 445
Ile Leu Asp Tyr Phe Ser Tyr Pro Leu Ser Met Lys Ser Asp Leu Ser
450 455 460
Val Ser Ser Ala Gly Pro Ile Ser Gln Phe Asn Tyr Lys Gln Ser Phe

465 470 475 480
Ser Asn Pro Thr Cys Leu Ile Leu Ala Thr Val Pro His Asn Leu Thr
485 490 495
Thr Ile Thr Lys Pro Leu Lys Tyr Ser Tyr Ile Asn Lys Cys Ser Arg
500 505 510
Phe Leu Ser Asp Asp Arg Thr Glu Val Pro Gln Leu Val Asn Ala Asn
515 520 525
Gln Tyr Ser Pro Cys Val Ser Ile Val Pro Ser Thr Val Trp Glu Asp

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

595 600 605
Asn Cys Val Glu Tyr Ser Leu Tyr Gly Val Ser Gly Arg Gly Val Phe
610 615 620
Gln Asn Cys Thr Ala Val Gly Val Arg Gln Gln Arg Phe Val Tyr Asp

625 630 635 640

Ala Tyr Gln Asn Leu Val Gly Tyr Tyr Ser Asp Asp Gly Asn Tyr Tyr

645 650 655

Cys Leu Arg Ala Cys Val Ser Val Pro Val Ser Val Ile Tyr Asp Lys

660 665 670

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

675 680 685

Ile Ser Ser Thr Met Ser Gln Tyr Ser Arg Ser Thr Arg Ser Met Leu

690 695 700

Lys Arg Arg Asp Ser Thr Tyr Gly Pro Leu Gln Thr Pro Val Gly Cys

705 710 715 720

Val Leu Gly Leu Val Asn Ser Ser Leu Phe Val Glu Asp Cys Lys Leu

725 730 735

Pro Leu Gly Gln Ser Leu Cys Ala Leu Pro Asp Thr Pro Ser Thr Leu

740 745 750

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

755 760 765

Ile Ala Phe Asn His Pro Ile Gln Val Asp Gln Leu Asn Ser Ser Tyr

770 775 780

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

785 790 795 800

Tyr Ile Gln Thr Thr Ile Gln Lys Val Thr Val Asp Cys Lys Gln Tyr

805 810 815

Val Cys Asn Gly Phe Gln Lys Cys Glu Gln Leu Leu Arg Glu Tyr Gly

820 825 830

Gln Phe Cys Ser Lys Ile Asn Gln Ala Leu His Gly Ala Asn Leu Arg

835 840 845

Gln Asp Asp Ser Val Arg Asn Leu Phe Ala Ser Val Lys Ser Ser Gln

850 855 860

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

865 870 875 880

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

885 890 895

Ile Glu Asp Leu Leu Phe Asp Lys Val Thr Ile Ala Asp Pro Gly Tyr

900 905 910

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

915 920 925

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

930 935 940

Leu Met Asp Val Asn Met Glu Ala Ala Tyr Thr Ser Ser Leu Leu Gly

945 950 955 960

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

965 970 975

Ile Pro Phe Ala Gln Ser Ile Phe Tyr Arg Leu Asn Gly Val Gly Ile

980 985 990

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

995 1000 1005

Asn Gln Ala Leu Gly Ala Met Gln Thr Gly Phe Thr Thr Thr Asn

1010 1015 1020

Glu Ala Phe His Lys Val Gln Asp Ala Val Asn Asn Asn Ala Gln

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Glu Gln Asp Ala Gln Ile Asp Arg Leu Ile Asn Gly Arg Leu Thr

1070 1075 1080

Thr Leu Asn Ala Phe Val Ala Gln Gln Leu Val Arg Ser Glu Ser

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

Pro Val Asn Gly Tyr Phe Ile Lys Thr Asn Asn Thr Arg Ile Val

1175 1180 1185

Asp Glu Trp Ser Tyr Thr Gly Ser Ser Phe Tyr Ala Pro Glu Pro

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Thr Gly Ile Asp Phe Gln Asp Glu Leu Asp Glu Phe Phe Lys Asn

1235 1240 1245

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

1250 1255 1260

Thr Thr Leu Leu Asp Leu Thr Tyr Glu Met Leu Ser Leu Gln Gln

1265 1270 1275

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

1280 1285 1290

Gly Asn Tyr Thr Tyr Tyr Asn Lys Trp Pro Trp Tyr Gln Ile Leu

1295 1300 1305

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

1310 1315 1320

Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln

1325 1330 1335

Cys Arg Ile Cys Ile

1340

<210> 108

<211> 1341

<212> PRT

<213> Artificial Sequence

<220><223> PDI-MERS-H5iCT-AA

<400> 108

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Tyr Val Asp Val Gly Pro Asp Ser
20 25 30

Val Lys Ser Ala Cys Ile Glu Val Asp Ile Gln Gln Thr Phe Phe Asp
35 40 45

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

Tyr Pro Gln Gly Arg Thr Tyr Ser Asn Ile Thr Ile Thr Tyr Gln Gly
65 70 75 80

Leu Phe Pro Tyr Gln Gly Asp His Gly Asp Met Tyr Val Tyr Ser Ala
85 90 95

Gly His Ala Thr Gly Thr Thr Pro Gln Lys Leu Phe Val Ala Asn Tyr
100 105 110

Ser Gln Asp Val Lys Gln Phe Ala Asn Gly Phe Val Val Arg Ile Gly
115 120 125

Ala Ala Ala Asn Ser Thr Gly Thr Val Ile Ile Ser Pro Ser Thr Ser
130 135 140

Ala Thr Ile Arg Lys Ile Tyr Pro Ala Phe Met Leu Gly Ser Ser Val
145 150 155 160

Gly Asn Phe Ser Asp Gly Lys Met Gly Arg Phe Phe Asn His Thr Leu
165 170 175

Val Leu Leu Pro Asp Gly Cys Gly Thr Leu Leu Arg Ala Phe Tyr Cys
180 185 190

Ile Leu Glu Pro Arg Ser Gly Asn His Cys Pro Ala Gly Asn Ser Tyr

195 200 205

Thr Ser Phe Ala Thr Tyr His Thr Pro Ala Thr Asp Cys Ser Asp Gly

210 215 220

Asn Tyr Asn Arg Asn Ala Ser Leu Asn Ser Phe Lys Glu Tyr Phe Asn

225 230 235 240

Leu Arg Asn Cys Thr Phe Met Tyr Thr Tyr Asn Ile Thr Glu Asp Glu

245 250 255

Ile Leu Glu Trp Phe Gly Ile Thr Gln Thr Ala Gln Gly Val His Leu

260 265 270

Phe Ser Ser Arg Tyr Val Asp Leu Tyr Gly Gly Asn Met Phe Gln Phe

275 280 285

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

290 295 300

His Ser Ile Arg Ser Ile Gln Ser Asp Arg Lys Ala Trp Ala Ala Phe

305 310 315 320

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

325 330 335

Asp Gly Tyr Ile Arg Arg Ala Ile Asp Cys Gly Phe Asn Asp Leu Ser

340 345 350

Gln Leu His Cys Ser Tyr Glu Ser Phe Asp Val Glu Ser Gly Val Tyr

355 360 365

Ser Val Ser Ser Phe Glu Ala Lys Pro Ser Gly Ser Val Val Glu Gln

370 375 380

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

385 390 395 400

Pro Gln Val Tyr Asn Phe Lys Arg Leu Val Phe Thr Asn Cys Asn Tyr

405 410 415

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

420 425 430

Ser Gln Ile Ser Pro Ala Ala Ile Ala Ser Asn Cys Tyr Ser Ser Leu

435 440 445

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

500 505 510
Phe Leu Ser Asp Asp Arg Thr Glu Val Pro Gln Leu Val Asn Ala Asn
515 520 525
Gln Tyr Ser Pro Cys Val Ser Ile Val Pro Ser Thr Val Trp Glu Asp
530 535 540
Gly Asp Tyr Tyr Arg Lys Gln Leu Ser Pro Leu Glu Gly Gly Gly Trp
545 550 555 560
Leu Val Ala Ser Gly Ser Thr Val Ala Met Thr Glu Gln Leu Gln Met

565 570 575
Gly Phe Gly Ile Thr Val Gln Tyr Gly Thr Asp Thr Asn Ser Val Cys
580 585 590
Pro Lys Leu Glu Phe Ala Asn Asp Thr Lys Ile Ala Ser Gln Leu Gly
595 600 605
Asn Cys Val Glu Tyr Ser Leu Tyr Gly Val Ser Gly Arg Gly Val Phe
610 615 620
Gln Asn Cys Thr Ala Val Gly Val Arg Gln Gln Arg Phe Val Tyr Asp

625 630 635 640
Ala Tyr Gln Asn Leu Val Gly Tyr Tyr Ser Asp Asp Gly Asn Tyr Tyr
645 650 655
Cys Leu Arg Ala Cys Val Ser Val Pro Val Ser Val Ile Tyr Asp Lys
660 665 670
Glu Thr Lys Thr His Ala Thr Leu Phe Gly Ser Val Ala Cys Glu His
675 680 685
Ile Ser Ser Thr Met Ser Gln Tyr Ser Arg Ser Thr Arg Ser Met Leu

690 695 700

Lys Arg Arg Asp Ser Thr Tyr Gly Pro Leu Gln Thr Pro Val Gly Cys

705 710 715 720

Val Leu Gly Leu Val Asn Ser Ser Leu Phe Val Glu Asp Cys Lys Leu

725 730 735

Pro Leu Gly Gln Ser Leu Cys Ala Leu Pro Asp Thr Pro Ser Thr Leu

740 745 750

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

755 760 765

Ile Ala Phe Asn His Pro Ile Gln Val Asp Gln Leu Asn Ser Ser Tyr

770 775 780

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

785 790 795 800

Tyr Ile Gln Thr Thr Ile Gln Lys Val Thr Val Asp Cys Lys Gln Tyr

805 810 815

Val Cys Asn Gly Phe Gln Lys Cys Glu Gln Leu Leu Arg Glu Tyr Gly

820 825 830

Gln Phe Cys Ser Lys Ile Asn Gln Ala Leu His Gly Ala Asn Leu Arg

835 840 845

Gln Asp Asp Ser Val Arg Asn Leu Phe Ala Ser Val Lys Ser Ser Gln

850 855 860

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

865 870 875 880

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

885 890 895

Ile Glu Asp Leu Leu Phe Asp Lys Val Thr Ile Ala Asp Pro Gly Tyr

900 905 910

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

915 920 925

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

930 935 940

Leu Met Asp Val Asn Met Glu Ala Ala Tyr Thr Ser Ser Leu Leu Gly

945 950 955 960

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

965 970 975

Ile Pro Phe Ala Gln Ser Ile Phe Tyr Arg Leu Asn Gly Val Gly Ile

980 985 990

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

995 1000 1005

Asn Gln Ala Leu Gly Ala Met Gln Thr Gly Phe Thr Thr Thr Asn

1010 1015 1020

Glu Ala Phe His Lys Val Gln Asp Ala Val Asn Asn Asn Ala Gln

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Glu Gln Asp Ala Gln Ile Asp Arg Leu Ile Asn Gly Arg Leu Thr

1070 1075 1080

Thr Leu Asn Ala Phe Val Ala Gln Gln Leu Val Arg Ser Glu Ser

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

Pro Val Asn Gly Tyr Phe Ile Lys Thr Asn Asn Thr Arg Ile Val

1175 1180 1185

Asp Glu Trp Ser Tyr Thr Gly Ser Ser Phe Tyr Ala Pro Glu Pro

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Thr Gly Ile Asp Phe Gln Asp Glu Leu Asp Glu Phe Phe Lys Asn

1235 1240 1245

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

1250 1255 1260

Thr Thr Leu Leu Asp Leu Thr Tyr Glu Met Leu Ser Leu Gln Gln

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

Gly Phe Ile Ala Gly Leu Val Ala Leu Ala Leu Cys Val Phe Phe

1310 1315 1320

Ile Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg

1325 1330 1335

Ile Cys Ile

1340

<210> 109

<211> 1338

<212> PRT

<213> Artificial Sequence

<220><223> PDI-MERS-H5iCT(V4)-AA

<400> 109

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Tyr Val Asp Val Gly Pro Asp Ser

20 25 30

Val Lys Ser Ala Cys Ile Glu Val Asp Ile Gln Gln Thr Phe Phe Asp

35 40 45

Lys Thr Trp Pro Arg Pro Ile Asp Val Ser Lys Ala Asp Gly Ile Ile

50 55 60

Tyr Pro Gln Gly Arg Thr Tyr Ser Asn Ile Thr Ile Thr Tyr Gln Gly

65 70 75 80

Leu Phe Pro Tyr Gln Gly Asp His Gly Asp Met Tyr Val Tyr Ser Ala

85 90 95

Gly His Ala Thr Gly Thr Thr Pro Gln Lys Leu Phe Val Ala Asn Tyr

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

Gly Asn Phe Ser Asp Gly Lys Met Gly Arg Phe Phe Asn His Thr Leu

165 170 175

Val Leu Leu Pro Asp Gly Cys Gly Thr Leu Leu Arg Ala Phe Tyr Cys

180 185 190

Ile Leu Glu Pro Arg Ser Gly Asn His Cys Pro Ala Gly Asn Ser Tyr

195 200 205

Thr Ser Phe Ala Thr Tyr His Thr Pro Ala Thr Asp Cys Ser Asp Gly

210 215 220

Asn Tyr Asn Arg Asn Ala Ser Leu Asn Ser Phe Lys Glu Tyr Phe Asn

225 230 235 240

Leu Arg Asn Cys Thr Phe Met Tyr Thr Tyr Asn Ile Thr Glu Asp Glu

245 250 255

Ile Leu Glu Trp Phe Gly Ile Thr Gln Thr Ala Gln Gly Val His Leu

260 265 270

Phe Ser Ser Arg Tyr Val Asp Leu Tyr Gly Gly Asn Met Phe Gln Phe

275 280 285

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

290 295 300

His Ser Ile Arg Ser Ile Gln Ser Asp Arg Lys Ala Trp Ala Ala Phe

305 310 315 320

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

325 330 335

Asp Gly Tyr Ile Arg Arg Ala Ile Asp Cys Gly Phe Asn Asp Leu Ser

340 345 350

Gln Leu His Cys Ser Tyr Glu Ser Phe Asp Val Glu Ser Gly Val Tyr

355 360 365

Ser Val Ser Ser Phe Glu Ala Lys Pro Ser Gly Ser Val Val Glu Gln

370 375 380

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

385 390 395 400

Pro Gln Val Tyr Asn Phe Lys Arg Leu Val Phe Thr Asn Cys Asn Tyr

405 410 415

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

420 425 430

Ser Gln Ile Ser Pro Ala Ala Ile Ala Ser Asn Cys Tyr Ser Ser Leu

435 440 445

Ile Leu Asp Tyr Phe Ser Tyr Pro Leu Ser Met Lys Ser Asp Leu Ser

450 455 460

Val Ser Ser Ala Gly Pro Ile Ser Gln Phe Asn Tyr Lys Gln Ser Phe

465 470 475 480

Ser Asn Pro Thr Cys Leu Ile Leu Ala Thr Val Pro His Asn Leu Thr

485 490 495

Thr Ile Thr Lys Pro Leu Lys Tyr Ser Tyr Ile Asn Lys Cys Ser Arg

500 505 510

Phe Leu Ser Asp Asp Arg Thr Glu Val Pro Gln Leu Val Asn Ala Asn
515 520 525
Gln Tyr Ser Pro Cys Val Ser Ile Val Pro Ser Thr Val Trp Glu Asp

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

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

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

725 730 735
Pro Leu Gly Gln Ser Leu Cys Ala Leu Pro Asp Thr Pro Ser Thr Leu
740 745 750
Thr Pro Ala Ser Val Gly Ser Val Pro Gly Glu Met Arg Leu Ala Ser

```
        755                 760                 765
Ile Ala Phe Asn His Pro Ile Gln Val Asp Gln Leu Asn Ser Ser Tyr
    770                 775                 780
Phe Lys Leu Ser Ile Pro Thr Asn Phe Ser Phe Gly Val Thr Gln Glu

785                 790                 795                 800
Tyr Ile Gln Thr Thr Ile Gln Lys Val Thr Val Asp Cys Lys Gln Tyr
                805                 810                 815
Val Cys Asn Gly Phe Gln Lys Cys Glu Gln Leu Leu Arg Glu Tyr Gly
            820                 825                 830
Gln Phe Cys Ser Lys Ile Asn Gln Ala Leu His Gly Ala Asn Leu Arg
        835                 840                 845
Gln Asp Asp Ser Val Arg Asn Leu Phe Ala Ser Val Lys Ser Ser Gln

    850                 855                 860
Ser Ser Pro Ile Ile Pro Gly Phe Gly Gly Asp Phe Asn Leu Thr Leu
865                 870                 875                 880
Leu Glu Pro Val Ser Ile Ser Thr Gly Ser Arg Ser Ala Arg Ser Ala
                885                 890                 895
Ile Glu Asp Leu Leu Phe Asp Lys Val Thr Ile Ala Asp Pro Gly Tyr
            900                 905                 910
Met Gln Gly Tyr Asp Asp Cys Met Gln Gln Gly Pro Ala Ser Ala Arg

        915                 920                 925
Asp Leu Ile Cys Ala Gln Tyr Val Ala Gly Tyr Lys Val Leu Pro Pro
    930                 935                 940
Leu Met Asp Val Asn Met Glu Ala Ala Tyr Thr Ser Ser Leu Leu Gly
945                 950                 955                 960
Ser Ile Ala Gly Val Gly Trp Thr Ala Gly Leu Ser Ser Phe Ala Ala
                965                 970                 975
Ile Pro Phe Ala Gln Ser Ile Phe Tyr Arg Leu Asn Gly Val Gly Ile

            980                 985                 990
Thr Gln Gln Val Leu Ser Glu Asn  Gln Lys Leu Ile Ala  Asn Lys Phe
        995                 1000                1005
```

```
Asn Gln  Ala Leu Gly Ala Met  Gln Thr Gly Phe Thr  Thr Thr Asn
    1010                1015                1020
Glu Ala  Phe His Lys Val Gln  Asp Ala Val Asn Asn  Asn Ala Gln
    1025                1030                1035
Ala Leu  Ser Lys Leu Ala Ser  Glu Leu Ser Asn Thr  Phe Gly Ala

    1040                1045                1050
Ile Ser  Ala Ser Ile Gly Asp  Ile Ile Gln Arg Leu  Asp Pro Pro
    1055                1060                1065
Glu Gln  Asp Ala Gln Ile Asp  Arg Leu Ile Asn Gly  Arg Leu Thr
    1070                1075                1080
Thr Leu  Asn Ala Phe Val Ala  Gln Gln Leu Val Arg  Ser Glu Ser
    1085                1090                1095
Ala Ala  Leu Ser Ala Gln Leu  Ala Lys Asp Lys Val  Asn Glu Cys

    1100                1105                1110
Val Lys  Ala Gln Ser Lys Arg  Ser Gly Phe Cys Gly  Gln Gly Thr
    1115                1120                1125
His Ile  Val Ser Phe Val Val  Asn Ala Pro Asn Gly  Leu Tyr Phe
    1130                1135                1140
Met His  Val Gly Tyr Tyr Pro  Ser Asn His Ile Glu  Val Val Ser
    1145                1150                1155
Ala Tyr  Gly Leu Cys Asp Ala  Ala Asn Pro Thr Asn  Cys Ile Ala

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

    1220                1225                1230
Thr Gly  Ile Asp Phe Gln Asp  Glu Leu Asp Glu Phe  Phe Lys Asn
```

1235 1240 1245

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

1250 1255 1260

Thr Thr Leu Leu Asp Leu Thr Tyr Glu Met Leu Ser Leu Gln Gln

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

Gly Phe Ile Ala Gly Leu Val Ala Leu Ala Leu Cys Val Phe Phe

1310 1315 1320

Ile Leu Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1325 1330 1335

<210> 110

<211> 1341

<212> PRT

<213> Artificial Sequence

<220><223> PDI-MERS-H1cCT-AA

<400> 110

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Tyr Val Asp Val Gly Pro Asp Ser

20 25 30

Val Lys Ser Ala Cys Ile Glu Val Asp Ile Gln Gln Thr Phe Phe Asp

35 40 45

Lys Thr Trp Pro Arg Pro Ile Asp Val Ser Lys Ala Asp Gly Ile Ile

50 55 60

Tyr Pro Gln Gly Arg Thr Tyr Ser Asn Ile Thr Ile Thr Tyr Gln Gly

65 70 75 80

Leu Phe Pro Tyr Gln Gly Asp His Gly Asp Met Tyr Val Tyr Ser Ala

85 90 95

Gly His Ala Thr Gly Thr Thr Pro Gln Lys Leu Phe Val Ala Asn Tyr

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

Gly Asn Phe Ser Asp Gly Lys Met Gly Arg Phe Phe Asn His Thr Leu

165 170 175

Val Leu Leu Pro Asp Gly Cys Gly Thr Leu Leu Arg Ala Phe Tyr Cys

180 185 190

Ile Leu Glu Pro Arg Ser Gly Asn His Cys Pro Ala Gly Asn Ser Tyr

195 200 205

Thr Ser Phe Ala Thr Tyr His Thr Pro Ala Thr Asp Cys Ser Asp Gly

210 215 220

Asn Tyr Asn Arg Asn Ala Ser Leu Asn Ser Phe Lys Glu Tyr Phe Asn

225 230 235 240

Leu Arg Asn Cys Thr Phe Met Tyr Thr Tyr Asn Ile Thr Glu Asp Glu

245 250 255

Ile Leu Glu Trp Phe Gly Ile Thr Gln Thr Ala Gln Gly Val His Leu

260 265 270

Phe Ser Ser Arg Tyr Val Asp Leu Tyr Gly Gly Asn Met Phe Gln Phe

275 280 285

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

290 295 300

His Ser Ile Arg Ser Ile Gln Ser Asp Arg Lys Ala Trp Ala Ala Phe

305 310 315 320

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

325 330 335

Asp Gly Tyr Ile Arg Arg Ala Ile Asp Cys Gly Phe Asn Asp Leu Ser

340 345 350

Gln Leu His Cys Ser Tyr Glu Ser Phe Asp Val Glu Ser Gly Val Tyr
355 360 365
Ser Val Ser Ser Phe Glu Ala Lys Pro Ser Gly Ser Val Val Glu Gln
370 375 380

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

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

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

Gly Phe Gly Ile Thr Val Gln Tyr Gly Thr Asp Thr Asn Ser Val Cys
580 585 590
Pro Lys Leu Glu Phe Ala Asn Asp Thr Lys Ile Ala Ser Gln Leu Gly

595 600 605

Asn Cys Val Glu Tyr Ser Leu Tyr Gly Val Ser Gly Arg Gly Val Phe

610 615 620

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

625 630 635 640

Ala Tyr Gln Asn Leu Val Gly Tyr Tyr Ser Asp Asp Gly Asn Tyr Tyr

645 650 655

Cys Leu Arg Ala Cys Val Ser Val Pro Val Ser Val Ile Tyr Asp Lys

660 665 670

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

675 680 685

Ile Ser Ser Thr Met Ser Gln Tyr Ser Arg Ser Thr Arg Ser Met Leu

690 695 700

Lys Arg Arg Asp Ser Thr Tyr Gly Pro Leu Gln Thr Pro Val Gly Cys

705 710 715 720

Val Leu Gly Leu Val Asn Ser Ser Leu Phe Val Glu Asp Cys Lys Leu

725 730 735

Pro Leu Gly Gln Ser Leu Cys Ala Leu Pro Asp Thr Pro Ser Thr Leu

740 745 750

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

755 760 765

Ile Ala Phe Asn His Pro Ile Gln Val Asp Gln Leu Asn Ser Ser Tyr

770 775 780

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

785 790 795 800

Tyr Ile Gln Thr Thr Ile Gln Lys Val Thr Val Asp Cys Lys Gln Tyr

805 810 815

Val Cys Asn Gly Phe Gln Lys Cys Glu Gln Leu Leu Arg Glu Tyr Gly

820 825 830

Gln Phe Cys Ser Lys Ile Asn Gln Ala Leu His Gly Ala Asn Leu Arg

835 840 845

Gln Asp Asp Ser Val Arg Asn Leu Phe Ala Ser Val Lys Ser Ser Gln
850 855 860
Ser Ser Pro Ile Ile Pro Gly Phe Gly Gly Asp Phe Asn Leu Thr Leu
865 870 875 880
Leu Glu Pro Val Ser Ile Ser Thr Gly Ser Arg Ser Ala Arg Ser Ala
885 890 895

Ile Glu Asp Leu Leu Phe Asp Lys Val Thr Ile Ala Asp Pro Gly Tyr
900 905 910
Met Gln Gly Tyr Asp Asp Cys Met Gln Gln Gly Pro Ala Ser Ala Arg
915 920 925
Asp Leu Ile Cys Ala Gln Tyr Val Ala Gly Tyr Lys Val Leu Pro Pro
930 935 940
Leu Met Asp Val Asn Met Glu Ala Ala Tyr Thr Ser Ser Leu Leu Gly
945 950 955 960

Ser Ile Ala Gly Val Gly Trp Thr Ala Gly Leu Ser Ser Phe Ala Ala
965 970 975
Ile Pro Phe Ala Gln Ser Ile Phe Tyr Arg Leu Asn Gly Val Gly Ile
980 985 990
Thr Gln Gln Val Leu Ser Glu Asn Gln Lys Leu Ile Ala Asn Lys Phe
995 1000 1005
Asn Gln Ala Leu Gly Ala Met Gln Thr Gly Phe Thr Thr Thr Asn
1010 1015 1020

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

Thr Leu Asn Ala Phe Val Ala Gln Gln Leu Val Arg Ser Glu Ser

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

Pro Val Asn Gly Tyr Phe Ile Lys Thr Asn Asn Thr Arg Ile Val

1175 1180 1185

Asp Glu Trp Ser Tyr Thr Gly Ser Ser Phe Tyr Ala Pro Glu Pro

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Thr Gly Ile Asp Phe Gln Asp Glu Leu Asp Glu Phe Phe Lys Asn

1235 1240 1245

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

1250 1255 1260

Thr Thr Leu Leu Asp Leu Thr Tyr Glu Met Leu Ser Leu Gln Gln

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

Gly Phe Ile Ala Gly Leu Val Ala Leu Ala Leu Cys Val Phe Phe

1310 1315 1320

Ile Leu Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg

1325 1330 1335

Ile Cys Ile

1340

<210> 111

<211> 4364

<212> DNA

<213> Artificial Sequence

<220><223> Cloning vector 7147 from left to right T-DNA

<400> 111

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg cactccattt gaatctatca aaccaaaaca cattgagacg tcacgtactc 2940

ctcagccaaa acgacacccc catctgtcta tccactggcc cctggatctg ctgcccaaac 3000

taactccatg gtgaccctgg gatgcctggt caagggctat ttccctgagc cagtgacagt 3060

gacctggaac tctggatccc tgtccagcgg tgtgcacacc ttcccagctg tcctgcagtc 3120

tgacctctac actctgagca gctcagtgac tgtcccctcc agcacctggc ccagcgagac 3180
cgtcacctgc aacgttgccc acccggccag cagcaccaag gtggacaaga aaattgtgcc 3240
cagggattgt ggttgtaagc cttgcatatg tacagtccca gaagtatcat ctgtcttcat 3300
cttcccccca aagcccaagg atgtgctcac cattactctg actcctaagg tcacgtgtgt 3360
tgtggtagac atcagcaagg atgatcccga ggtccagttc agctggtttg tagatgatgt 3420
ggaggtgcac acagctcaga cgcaaccccg ggaggagcag ttcaacagca ctttccgctc 3480
agtcagtgaa cttcccatca tgcaccagga ctggctcaat ggcaaggaga cgtccagatt 3540

ttggcgatct attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc 3600
agtttctgga tgtgctctaa tgggtctcta cagtgtagaa tatgtattta aaggccttag 3660
tcgtgtcgtt tttcaaataa tataatcctt ttagggtttt agttagttta aattttctgt 3720
tgctcctgtt tagcaggtcg tgccttcagc aagcacacaa aaacagagtg tttattttaa 3780
gttgtttgtt tagtgattca aaaaaaaaat cgttcaaaca tttggcaata aagtttctta 3840
agattgaatc ctgttgccgg tcttgcgatg attatcatat aatttctgtt gaattacgtt 3900
aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt 3960

agagtcccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag 4020
gataaattat cgcgcgcggt gtcatctatg ttactagatc tctagagtct caagcttggc 4080
gcgcccacgt gactagtggc actggccgtc gttttacaac gtcgtgactg ggaaaaccct 4140
ggcgttaccc aacttaatcg ccttgcagca catccccctt tcgccagctg gcgtaatagc 4200
gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg cgaatgctag 4260
agcagcttga gcttggatca gattgtcgtt tcccgccttc agtttaaact atcagtgttt 4320
gacaggatat attggcgggt aaacctaaga gaaaagagcg ttta 4364

<210> 112
<211> 1255
<212> PRT
<213> SARS-CoV-1
<400> 112

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

1010 1015 1020

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

1070 1075 1080
Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
1085 1090 1095
Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
1100 1105 1110
Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
1115 1120 1125
Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser

1130 1135 1140
Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val
1145 1150 1155
Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
1160 1165 1170
Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
1175 1180 1185
Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu Gly Phe Ile

1190 1195 1200
Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys
1205 1210 1215
Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys Gly
1220 1225 1230
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> 113

<211> 1353

<212> PRT

<213> MERS

<400> 113

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 Phe 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 His Lys Val Gln

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

Gln Gln Leu Val Arg Ser Glu Ser Ala Ala Leu Ser Ala Gln Leu

1085 1090 1095

Ala Lys Asp Lys Val Asn Glu Cys Val Lys Ala Gln Ser Lys Arg

1100 1105 1110

Ser Gly Phe Cys Gly Gln Gly Thr His Ile Val Ser Phe Val Val

1115 1120 1125

Asn Ala Pro Asn Gly Leu Tyr Phe Met His Val Gly Tyr Tyr Pro

1130 1135 1140

Ser Asn His Ile Glu Val Val Ser Ala Tyr Gly Leu Cys Asp Ala

1145 1150 1155

Ala Asn Pro Thr Asn Cys Ile Ala Pro Val Asn Gly Tyr Phe Ile

1160 1165 1170

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

1175 1180 1185

Ser Ser Phe Tyr Ala Pro Glu Pro Ile Thr Ser Leu Asn Thr Lys

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

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

1250 1255 1260

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

1265 1270 1275

Ser Tyr Ile Asp Leu Lys Glu Leu Gly Asn Tyr Thr Tyr Tyr Asn

1280 1285 1290

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

1295 1300 1305

Ala Leu Ala Leu Cys Val Phe Phe Ile Leu Cys Cys Thr Gly Cys

1310 1315 1320

Gly Thr Asn Cys Met Gly Lys Leu Lys Cys Asn Arg Cys Cys Asp

1325 1330 1335

Arg Tyr Glu Glu Tyr Asp Leu Glu Pro His Lys Val His Val His

1340 1345 1350

<210> 114

<211> 1242

<212> PRT

<213> SARS-CoV-1

<400> 114

Ser Asp Leu Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn

1 5 10 15

Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu

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

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

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

Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala
    210                 215                 220
Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe
225                 230                 235                 240
Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn
                245                 250                 255
Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu
            260                 265                 270
```

Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln
275 280 285

Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro
290 295 300

Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys
305 310 315 320

Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val
325 330 335

Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys
340 345 350

Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn
355 360 365

Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile
370 375 380

Ala Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro
385 390 395 400

Asp Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp
405 410 415

Ala Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His
420 425 430

Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser
435 440 445

Pro Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro
450 455 460

Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro
465 470 475 480

Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr
485 490 495

Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val
500 505 510

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

515 520 525

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

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

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

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

Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg
770 775 780

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala
785 790 795 800
Gly Phe Met Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg
805 810 815
Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro
820 825 830
Leu Leu Thr Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser
835 840 845

Gly Thr Ala Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln
850 855 860
Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val
865 870 875 880
Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe
885 890 895
Asn Lys Ala Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr
900 905 910

Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu
915 920 925
Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser
930 935 940
Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val
945 950 955 960
Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr
965 970 975

Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn
980 985 990
Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg
995 1000 1005
Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln

1010 1015 1020

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

1025 1030 1035

Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu

1040 1045 1050

Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly

1055 1060 1065

Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile

1070 1075 1080

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

1085 1090 1095

Ile Gly Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu

1100 1105 1110

Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His

1115 1120 1125

Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Leu Lys Gly Val Lys Leu His Tyr Thr

1235 1240

<210> 115

<211> 1336

<212> PRT

<213> MERS

<400> 115

Tyr Val Asp Val Gly Pro Asp Ser Val Lys Ser Ala Cys Ile Glu Val
1 5 10 15

Asp Ile Gln Gln Thr Phe Phe Asp Lys Thr Trp Pro Arg Pro Ile Asp
20 25 30

Val Ser Lys Ala Asp Gly Ile Ile Tyr Pro Gln Gly Arg Thr Tyr Ser
35 40 45

Asn Ile Thr Ile Thr Tyr Gln Gly Leu Phe Pro Tyr Gln Gly Asp His
50 55 60

Gly Asp Met Tyr Val Tyr Ser Ala Gly His Ala Thr Gly Thr Thr Pro
65 70 75 80

Gln Lys Leu Phe Val Ala Asn Tyr Ser Gln Asp Val Lys Gln Phe Ala
85 90 95

Asn Gly Phe Val Val Arg Ile Gly Ala Ala Ala Asn Ser Thr Gly Thr
100 105 110

Val Ile Ile Ser Pro Ser Thr Ser Ala Thr Ile Arg Lys Ile Tyr Pro
115 120 125

Ala Phe Met Leu Gly Ser Ser Val Gly Asn Phe Ser Asp Gly Lys Met
130 135 140

Gly Arg Phe Phe Asn His Thr Leu Val Leu Leu Pro Asp Gly Cys Gly
145 150 155 160

Thr Leu Leu Arg Ala Phe Tyr Cys Ile Leu Glu Pro Arg Ser Gly Asn
165 170 175

His Cys Pro Ala Gly Asn Ser Tyr Thr Ser Phe Ala Thr Tyr His Thr
180 185 190

Pro Ala Thr Asp Cys Ser Asp Gly Asn Tyr Asn Arg Asn Ala Ser Leu
195 200 205

Asn Ser Phe Lys Glu Tyr Phe Asn Leu Arg Asn Cys Thr Phe Met Tyr

210 215 220

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

Asp Arg Lys Ala Trp Ala Ala Phe Tyr Val Tyr Lys Leu Gln Pro Leu
290 295 300
Thr Phe Leu Leu Asp Phe Ser Val Asp Gly Tyr Ile Arg Arg Ala Ile
305 310 315 320
Asp Cys Gly Phe Asn Asp Leu Ser Gln Leu His Cys Ser Tyr Glu Ser
325 330 335
Phe Asp Val Glu Ser Gly Val Tyr Ser Val Ser Ser Phe Glu Ala Lys
340 345 350

Pro Ser Gly Ser Val Val Glu Gln Ala Glu Gly Val Glu Cys Asp Phe
355 360 365
Ser Pro Leu Leu Ser Gly Thr Pro Pro Gln Val Tyr Asn Phe Lys Arg
370 375 380
Leu Val Phe Thr Asn Cys Asn Tyr Asn Leu Thr Lys Leu Leu Ser Leu
385 390 395 400
Phe Ser Val Asn Asp Phe Thr Cys Ser Gln Ile Ser Pro Ala Ala Ile
405 410 415

Ala Ser Asn Cys Tyr Ser Ser Leu Ile Leu Asp Tyr Phe Ser Tyr Pro
420 425 430
Leu Ser Met Lys Ser Asp Leu Ser Val Ser Ser Ala Gly Pro Ile Ser
435 440 445
Gln Phe Asn Tyr Lys Gln Ser Phe Ser Asn Pro Thr Cys Leu Ile Leu
450 455 460

Ala Thr Val Pro His Asn Leu Thr Thr Ile Thr Lys Pro Leu Lys Tyr
465 470 475 480

Ser Tyr Ile Asn Lys Cys Ser Arg Phe Leu Ser Asp Asp Arg Thr Glu
485 490 495
Val Pro Gln Leu Val Asn Ala Asn Gln Tyr Ser Pro Cys Val Ser Ile
500 505 510
Val Pro Ser Thr Val Trp Glu Asp Gly Asp Tyr Tyr Arg Lys Gln Leu
515 520 525
Ser Pro Leu Glu Gly Gly Gly Trp Leu Val Ala Ser Gly Ser Thr Val
530 535 540

Ala Met Thr Glu Gln Leu Gln Met Gly Phe Gly Ile Thr Val Gln Tyr
545 550 555 560
Gly Thr Asp Thr Asn Ser Val Cys Pro Lys Leu Glu Phe Ala Asn Asp
565 570 575
Thr Lys Ile Ala Ser Gln Leu Gly Asn Cys Val Glu Tyr Ser Leu Tyr
580 585 590
Gly Val Ser Gly Arg Gly Val Phe Gln Asn Cys Thr Ala Val Gly Val
595 600 605

Arg Gln Gln Arg Phe Val Tyr Asp Ala Tyr Gln Asn Leu Val Gly Tyr
610 615 620
Tyr Ser Asp Asp Gly Asn Tyr Tyr Cys Leu Arg Ala Cys Val Ser Val
625 630 635 640
Pro Val Ser Val Ile Tyr Asp Lys Glu Thr Lys Thr His Ala Thr Leu
645 650 655
Phe Gly Ser Val Ala Cys Glu His Ile Ser Ser Thr Met Ser Gln Tyr
660 665 670

Ser Arg Ser Thr Arg Ser Met Leu Lys Arg Arg Asp Ser Thr Tyr Gly
675 680 685
Pro Leu Gln Thr Pro Val Gly Cys Val Leu Gly Leu Val Asn Ser Ser
690 695 700
Leu Phe Val Glu Asp Cys Lys Leu Pro Leu Gly Gln Ser Leu Cys Ala

705 710 715 720

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

725 730 735

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

740 745 750

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

755 760 765

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

770 775 780

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

785 790 795 800

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

805 810 815

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

820 825 830

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

835 840 845

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

850 855 860

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

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Tyr Arg Leu Asn Gly Val Gly Ile Thr Gln Gln Val Leu Ser Glu Asn
965 970 975
Gln Lys Leu Ile Ala Asn Lys Phe Asn Gln Ala Leu Gly Ala Met Gln
980 985 990

Thr Gly Phe Thr Thr Thr Asn Glu Ala Phe His Lys Val Gln Asp Ala
995 1000 1005
Val Asn Asn Asn Ala Gln Ala Leu Ser Lys Leu Ala Ser Glu Leu
1010 1015 1020
Ser Asn Thr Phe Gly Ala Ile Ser Ala Ser Ile Gly Asp Ile Ile
1025 1030 1035
Gln Arg Leu Asp Val Leu Glu Gln Asp Ala Gln Ile Asp Arg Leu
1040 1045 1050

Ile Asn Gly Arg Leu Thr Thr Leu Asn Ala Phe Val Ala Gln Gln
1055 1060 1065
Leu Val Arg Ser Glu Ser Ala Ala Leu Ser Ala Gln Leu Ala Lys
1070 1075 1080
Asp Lys Val Asn Glu Cys Val Lys Ala Gln Ser Lys Arg Ser Gly
1085 1090 1095
Phe Cys Gly Gln Gly Thr His Ile Val Ser Phe Val Val Asn Ala
1100 1105 1110

Pro Asn Gly Leu Tyr Phe Met His Val Gly Tyr Tyr Pro Ser Asn
1115 1120 1125
His Ile Glu Val Val Ser Ala Tyr Gly Leu Cys Asp Ala Ala Asn
1130 1135 1140
Pro Thr Asn Cys Ile Ala Pro Val Asn Gly Tyr Phe Ile Lys Thr
1145 1150 1155
Asn Asn Thr Arg Ile Val Asp Glu Trp Ser Tyr Thr Gly Ser Ser
1160 1165 1170

Phe Tyr Ala Pro Glu Pro Ile Thr Ser Leu Asn Thr Lys Tyr Val
1175 1180 1185
Ala Pro Gln Val Thr Tyr Gln Asn Ile Ser Thr Asn Leu Pro Pro

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Ser Leu Thr Gln Ile Asn Thr Thr Leu Leu Asp Leu Thr Tyr Glu

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

Ala Leu Cys Val Phe Phe Ile Leu Cys Cys Thr Gly Cys Gly Thr

1295 1300 1305

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

1310 1315 1320

Glu Glu Tyr Asp Leu Glu Pro His Lys Val His Val His

1325 1330 1335

<210> 116

<211> 60

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-SARS-COV-1 wtTMCT-AA

<400> 116

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

1 5 10 15

Val Thr Ile Leu Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys

20 25 30

Gly Ala Cys Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser

35 40 45

Glu Pro Val Leu Lys Gly Val Lys Leu His Tyr Thr

50 55 60

<210> 117

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> YQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

<400> 117

Trp Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
1 5 10 15
Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
20 25 30
Ser Leu Gln Cys Arg Ile Cys Ile
35 40

<210> 118

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-SARS-COV-1 H5iCT-AA

<400> 118

Trp Tyr Val Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met
1 5 10 15
Val Thr Ile Leu Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln
20 25 30
Cys Arg Ile Cys Ile
35

<210> 119

<211> 34

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-SARS-COV-1 H5iCT(V4)-AA

<400> 119

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

1 5 10 15

Val Thr Ile Leu Leu Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

20 25 30

Cys Ile

<210> 120

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-SARS-COV-1 H1cCT-AA

<400> 120

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

1 5 10 15

Val Thr Ile Leu Leu Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln

20 25 30

Cys Arg Ile Cys Ile

35

<210> 121

<211> 57

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-MERS-wtTMCT-AA

<400> 121

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

1 5 10 15

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

20 25 30

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

35 40 45

Leu Glu Pro His Lys Val His Val His

50 55

<210> 122

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-MERS-H5iTMCT-AA

<400> 122

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

1 5 10 15

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly

20 25 30

Ser Leu Gln Cys Arg Ile Cys Ile

35 40

<210> 123

<211> 38

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-MERS-H5iCT-AA

<400> 123

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

1 5 10 15

Cys Val Phe Phe Ile Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu

20 25 30

Gln Cys Arg Ile Cys Ile

35

<210> 124

<211> 35

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-MERS-H5iCT(V4)-AA

<400> 124

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

1 5 10 15

Cys Val Phe Phe Ile Leu Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg

20 25 30

Ile Cys Ile

35

<210> 125

<211> 38

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-MERS-H1cCT-AA

<400> 125

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

1 5 10 15

Cys Val Phe Phe Ile Leu Ser Phe Trp Met Cys Ser Asn Gly Ser Leu

20 25 30

Gln Cys Arg Ile Cys Ile

35

<210> 126

<211> 37

<212

> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-S-protein+H1 Cal

<400> 126

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

1 5 10 15

Val Thr Ile Met Leu Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln

20 25 30

Cys Arg Ile Cys Ile

35

<210> 127

<211> 36

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-S-protein+H3 Minn

<400> 127

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

1 5 10 15

Val Thr Ile Met Leu Met Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys

20 25 30

Asn Ile Cys Ile

35

<210> 128

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-S-protein+H6 HK

<400> 128

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

1 5 10 15

Val Thr Ile Met Leu Gly Leu Trp Met Cys Ser Asn Gly Ser Met Gln

20 25 30

Cys Arg Ile Cys Ile

35

<210> 129

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-S-protein+H7 Guangdong

<400> 129

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

1 5 10 15

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

20 25 30

Cys Thr Ile Cys Ile

35

<210> 130

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-S-protein+H9 HK

<400> 130

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met
1 5 10 15

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

20 25 30

Cys Asn Ile Cys Ile
35

<210> 131

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-S-protein+ B/ Wash

<400> 131

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met
1 5 10 15

Val Thr Ile Met Leu Val Val Tyr Met Val Ser Arg Asp Asn Val Ser
20 25 30

Cys Ser Ile Cys Leu
35

<210> 132

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Consensus Sequence of TM Domain of Coronavirus S-protein

<220><221> misc_feature

<222> (3)..(3)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (12)..(12)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (14)..(16)

<223> Xaa can be any naturally occurring amino acid

<220><221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa may be present or absent

<220><221> misc_feature

<222> (19)..(21)

<223> Xaa can be any naturally occurring amino acid

<400> 132

Trp Tyr Xaa Trp Leu Gly Phe Ile Ala Gly Leu Xaa Ala Xaa Xaa Xaa

1 5 10 15

Xaa Val Xaa Xaa Xaa Leu

20

<210> 133

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Consensus Sequence of TM Domain of Coronavirus S-protein

<220><221> MISC_FEATURE

<222> (3)..(3)

<223> Xaa may be Ile or Val

<220><221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa may be Val or Ile

<220><221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa may be Leu or Ile

<220><221> MISC_FEATURE

<222> (15)..(15)

<223> Xaa may be Ala or Val

<220><221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa may be Leu or Met

<220><221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa may be Cys or absent

<220><221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa may be Phe or Thr

<220><221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa may be Phe or Ile

<220><221> misc_feature

<222> (21)..(21)

<223> Xaa can be any naturally occurring amino acid

<400> 133

Trp Tyr Xaa Trp Leu Gly Phe Ile Ala Gly Leu Xaa Ala Xaa Xaa Xaa

1 5 10 15

Xaa Val Xaa Xaa Xaa Leu

20

<210> 134

<211> 33

<212> PRT

<213> Artificial Sequence

<220><223> TM/CT Region of Modified SARS-CoV-1 S protein with intervening

peptide sequence

<220><221> MISC_FEATURE

<222> (21)..(21)

<223> X may be any combination of 0 to 10 amino acids.

<220><221> misc_feature

<222> (27)..(28)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (30)..(30)

<223> Xaa can be any naturally occurring amino acid

<400> 134

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

1 5 10 15

Val Thr Ile Leu Xaa Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa Ile Cys

20 25 30

Ile

<210> 135

<211> 35

<212> PRT

<213> Artificial Sequence

<220><223> TM/CT Region of Modified MERS S protein with intervening peptide sequence

<220><221> MISC_FEATURE

<222> (23)..(23)

<223> X may be any combination of 0 to 10 amino acids

<220><221> misc_feature

<222> (29)..(30)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (32)..(32)

<223> Xaa can be any naturally occurring amino acid

<400> 135

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

1 5 10 15

Cys Val Phe Phe Ile Leu Xaa Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa

20 25 30

Ile Cys Ile

35

<210> 136

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> IF(AvB+wtCT-OC43).r

<400> 136

acgacacgac taaggccttc agtcgtcatg cgaggtctta atgacaagc 49

<210> 137

<

211> 4113

<212> DNA

<213> Artificial Sequence

<220><223> PDI-OC43-wtTMCT-DNA

<400> 137

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgatcgg cgatctgaat tgtaccctgg atccccgcct gaaagggagc 120

tttaacaacc gagatacagg acccccgtct atatccatag atacagtgga tgttacgaac 180

gggctcggca cctactatgt gctagaccga gtttatttga acaccacctt attcctcaat 240

ggatactacc caacttcagg tagtacttac agaaacatgg cgctgaaggg tacggatctg 300

ctgagcaccc tatggtttaa acctcccttc ctctcggact ttattaatgg catcttcgct 360

aaggtgaaaa acacgaaggt tttcaaagat ggagtgatgt attcagagtt ccctgcgatc 420

accattggaa gtaccttcgt gaatacttcc tatagcgtgg tggttcaacc acggacaatc 480

aactccaccc aggacggcgt caacaagctc cagggattgc tggaggtgtc agtctgtcaa 540

tataacatgt gtgagtaccc acacactatc tgtcacccta atctaggcaa ccactttaag 600

gaactgtggc actacgatac gggggtggta agttgcttat ataagagaaa tttcacctat 660

gatgttaatg caacgtacct gtactttcac ttctatcaag aaggaggaac tttctacgca 720

tatttcacag ataccggctt tgtgacgaaa ttcttattca acgtttacct cggaatggca 780

ttaagccatt attacgtgat gcctctcact tgcatcagac gccctaagga tggtttttct 840

ctggagtact gggtcactcc cctgacacca cggcagtacc tgcttgcttt taaccaggac 900

ggtatcattt ttaatgccgt cgattgtatg agcgatttta tgagcgagat aaagtgcaag 960

acccaatcta ttgctccgcc cacgggggtg tacgaactga atggttacac cgtccagccc 1020

gttgccgatg tatatagacg gaaaccagac ctgcccaatt gcaacatcga agcttggtta 1080

aacgataagt cagtgccctc ccccctcaat tgggagagga agactttctc caactgtaat 1140

ttcaacatgt caagcctgat gtctttcatt caagccgatt cgttcacttg taataatata 1200

gatgcagcaa agatctatgg tatgtgcttc agttccatca caatagataa gtttgcaata 1260

ccaaaccgtc gcaaggtgga ccttcagctc ggcaacctgg gctatctgca gtccagcaat 1320

tatagaatag acaccaccgc cacatcatgt cagctgtact ataacctccc agcagcgaac 1380

gtcagtgtta gtaggttcaa tccttctacc tggaataaaa ggtttggatt catcgaagat 1440

agtgtgttcg tacctcagcc aacaggagtg ttcaccaatc acagcgtggt ctacgcccaa 1500

cattgcttca aggcacccaa aaatttctgc ccatgtagca gttgctcctg cccgggtaag 1560

aacaatggga tcggcacctg cccagcaggc accaattcac ttacatgcga taatctgtgt 1620

acactggatc ctattacact taaggcccct gatacctaca aatgccccca gagcaagagc 1680

ctggtcggta tcggagaaca ctgttccgga cttgcagtaa aaagcgacta ttgtggaaat 1740

aactcttgca cttgtcagcc acaagccttc ctcggttggt ccgctgactc ttgtttacaa 1800

ggggataagt gtaacatctt cgcaaatttc atcttacacg atgtgaataa cggcttaaca 1860

tgcagcacag atctccagaa ggcaaacaca gagatcgaat taggagtctg cgttaattac 1920

gatctctacg ggatctctgg ccagggcatc ttcgtggagg ttaatgctac ctactacaat 1980

agttggcaaa atctgctcta cgatagcaat ggcaacctct atggattcag agactatatt 2040

actaacagga cgttcatgat tcactcgtgc tattccgggc gggtgtcagc agcttatcac 2100

gcaaattctt cagagccagc tctgctattc cgaaacataa aatgtaatta cgtgttcaat 2160

aattcactga ctcggcagct gcagccgatt aattacagct tcgacagcta ccttggttgc 2220

gttgttaacg cctacaactc cactgccata tcagttcaga cctgcgacct tactgtgggc 2280

tctggctatt gtgtcgatta ttcaaagaac ggggggagcg ggtccgcaat aacaactggc 2340

tataggttca ccaattttga gcctttcacc gtgaatagtg tcaacgatag cctggagcct 2400

gtcggaggtc tttatgagat acaaatcccc tccgagttca caattggcaa catggaagag 2460

ttcatccaga cgagttcccc aaaggtgacg atcgattgcg cggctttcgt ctgcggcgac 2520

tacgccgcat gcaagttaca actcgttgag tatggaagtt tttgcgataa tataaacgca 2580

attctgactg aagtgaacga actgctggac accactcagt tgcaggtggc aaattcgctc 2640

atgaacggcg tgacactgtc aaccaaactg aaggacggtg tcaatttcaa tgtggatgac 2700

attaacttca gccccgtact gggctgtttg ggtagtgagt gttctaaggc tagcagccgc 2760

tccgccattg aggacttgtt gtttgataaa gttaagctga gtgacgttgg atttgttgag 2820

gcgtataata actgtaccgg tggtgcagag ataagggatc tgatctgtgt ccagagttat 2880

aaggggatta aggttctccc cccgctactc tcggagaatc agatatcagg atacaccctg 2940

gccgctacct cagcctcgct gtttcccccc tggaccgctg ccgccggtgt cccattttat 3000

ttgaatgtgc agtatcggat caacggtctg ggagtgacaa tggacgtgct gtctcagaac 3060

cagaaactga tcgccaatgc attcaacaat gctctgcacg ccatccagca agggtttgac 3120

gctacaaatt ctgccctcgt aaaaatccag gccgtggtga atgctaacgc cgaagccctt 3180

aataatctgc tccagcagct ttctaaccgc tttggagcta tttctgcctc actgcaggaa 3240

attctatcca gactggatcc ccctgaggca gaagcccaaa tcgaccgtct cataaacggc 3300

agactcactg ctcttaacgc ctacgttagt caacaattga gcgattcgac cttggtgaaa 3360
ttcagcgcag ctcaggctat ggagaaggtg aacgagtgcg tgaagtcaca gagctccaga 3420
atcaatttct gtggcaatgg gaaccatatc atctccttgg ttcagaatgc tccctacggc 3480
ctgtatttca tccacttcaa ctacgtgccc acgaagtacg ttacagccaa agtgtccccc 3540
ggactgtgca tcgctggtaa caggggcatt gcaccaaaat ccggctactt cgtcaatgtc 3600
aacaacacat ggatgtatac tgggagtggt tattattacc ctgaacctat aacagagaac 3660
aatgtagtag tcatgtccac atgcgccgtc aattatacta aggcccccta tgttatgctc 3720

aacacttcaa ttcccaatct cccggatttc aaagaagagc tggatcagtg gtttaagaat 3780
cagacatccg tggcccctga cttaagcttg gattatatca atgtgacttt tttagactta 3840
caggtcgaga tgaaccgact ccaggaagct ataaaagtac tgaaccactc ctatatcaat 3900
ctgaaagata tcggtacata cgaatattac gtaaaatggc cttggtatgt gtggctacta 3960
atttgccttg cgggcgtggc tatgctggtc ctgctgttct tcatttgctg ctgtactggg 4020
tgcggtactt cctgcttcaa aaagtgcggt ggttgctgcg acgactacac tgggtaccag 4080
gagcttgtca ttaagacctc gcatgacgac tga 4113

<210> 138
<211> 4065
<212> DNA
<213> Artificial Sequence
<220><223> PDI-OC43-H5iTMCT-DNA
<400> 138
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cggtgatcgg cgatctgaat tgtaccctgg atccccgcct gaaagggagc 120
tttaacaacc gagatacagg acccccgtct atatccatag atacagtgga tgttacgaac 180
gggctcggca cctactatgt gctagaccga gtttatttga acaccacctt attcctcaat 240
ggatactacc caacttcagg tagtacttac agaaacatgg cgctgaaggg tacggatctg 300
ctgagcaccc tatggtttaa acctcccttc ctctcggact ttattaatgg catcttcgct 360

aaggtgaaaa acacgaaggt tttcaaagat ggagtgatgt attcagagtt ccctgcgatc 420
accattggaa gtaccttcgt gaatacttcc tatagcgtgg tggttcaacc acggacaatc 480
aactccaccc aggacggcgt caacaagctc cagggattgc tggaggtgtc agtctgtcaa 540
tataacatgt gtgagtaccc acacactatc tgtcacccta atctaggcaa ccactttaag 600

gaactgtggc actacgatac gggggtggta agttgcttat ataagagaaa tttcacctat 660

gatgttaatg caacgtacct gtactttcac ttctatcaag aaggaggaac tttctacgca 720

tatttcacag ataccggctt tgtgacgaaa ttcttattca acgtttacct cggaatggca 780

ttaagccatt attacgtgat gcctctcact tgcatcagac gccctaagga tggtttttct 840

ctggagtact gggtcactcc cctgacacca cggcagtacc tgcttgcttt taaccaggac 900

ggtatcattt ttaatgccgt cgattgtatg agcgatttta tgagcgagat aaagtgcaag 960

acccaatcta ttgctccgcc cacgggggtg tacgaactga atggttacac cgtccagccc 1020

gttgccgatg tatatagacg gaaaccagac ctgcccaatt gcaacatcga agcttggtta 1080

aacgataagt cagtgccctc ccccctcaat tgggagagga agactttctc caactgtaat 1140

ttcaacatgt caagcctgat gtctttcatt caagccgatt cgttcacttg taataatata 1200

gatgcagcaa agatctatgg tatgtgcttc agttccatca caatagataa gtttgcaata 1260

ccaaaccgtc gcaaggtgga ccttcagctc ggcaacctgg gctatctgca gtccagcaat 1320

tatagaatag acaccaccgc cacatcatgt cagctgtact ataacctccc agcagcgaac 1380

gtcagtgtta gtaggttcaa tccttctacc tggaataaaa ggtttggatt catcgaagat 1440

agtgtgttcg tacctcagcc aacaggagtg ttcaccaatc acagcgtggt ctacgcccaa 1500

cattgcttca aggcacccaa aaatttctgc ccatgtagca gttgctcctg cccgggtaag 1560

aacaatggga tcggcacctg cccagcaggc accaattcac ttacatgcga taatctgtgt 1620

acactggatc ctattacact taaggcccct gatacctaca aatgccccca gagcaagagc 1680

ctggtcggta tcggagaaca ctgttccgga cttgcagtaa aaagcgacta ttgtggaaat 1740

aactcttgca cttgtcagcc acaagccttc ctcggttggt ccgctgactc ttgtttacaa 1800

ggggataagt gtaacatctt cgcaaatttc atcttacacg atgtgaataa cggcttaaca 1860

tgcagcacag atctccagaa ggcaaacaca gagatcgaat taggagtctg cgttaattac 1920

gatctctacg ggatctctgg ccagggcatc ttcgtggagg ttaatgctac ctactacaat 1980

agttggcaaa atctgctcta cgatagcaat ggcaacctct atggattcag agactatatt 2040

actaacagga cgttcatgat tcactcgtgc tattccgggc gggtgtcagc agcttatcac 2100

gcaaattctt cagagccagc tctgctattc cgaaacataa aatgtaatta cgtgttcaat 2160

aattcactga ctcggcagct gcagccgatt aattacagct tcgacagcta ccttggttgc 2220

gttgttaacg cctacaactc cactgccata tcagttcaga cctgcgacct tactgtgggc 2280

tctggctatt gtgtcgatta ttcaaagaac gggggggagcg ggtccgcaat aacaactggc 2340

tataggttca ccaattttga gcctttcacc gtgaatagtg tcaacgatag cctggagcct 2400

gtcggaggtc tttatgagat acaaatcccc tccgagttca caattggcaa catggaagag 2460

ttcatccaga cgagttcccc aaaggtgacg atcgattgcg cggctttcgt ctgcggcgac 2520
tacgccgcat gcaagttaca actcgttgag tatggaagtt tttgcgataa tataaacgca 2580
attctgactg aagtgaacga actgctggac accactcagt tgcaggtggc aaattcgctc 2640
atgaacggcg tgacactgtc aaccaaactg aaggacggtg tcaatttcaa tgtggatgac 2700
attaacttca gccccgtact gggctgtttg ggtagtgagt gttctaaggc tagcagccgc 2760
tccgccattg aggacttgtt gtttgataaa gttaagctga gtgacgttgg atttgttgag 2820
gcgtataata actgtaccgg tggtgcagag ataagggatc tgatctgtgt ccagagttat 2880

aaggggatta aggttctccc cccgctactc tcggagaatc agatatcagg atacaccctg 2940
gccgctacct cagcctcgct gtttcccccc tggaccgctg ccgccggtgt cccattttat 3000
ttgaatgtgc agtatcggat caacggtctg ggagtgacaa tggacgtgct gtctcagaac 3060
cagaaactga tcgccaatgc attcaacaat gctctgcacg ccatccagca agggtttgac 3120
gctacaaatt ctgccctcgt aaaaatccag gccgtggtga atgctaacgc cgaagccctt 3180
aataatctgc tccagcagct ttctaaccgc tttggagcta tttctgcctc actgcaggaa 3240
attctatcca gactggatcc ccctgaggca gaagcccaaa tcgaccgtct cataaacggc 3300

agactcactg ctcttaacgc ctacgttagt caacaattga gcgattcgac cttggtgaaa 3360
ttcagcgcag ctcaggctat ggagaaggtg aacgagtgcg tgaagtcaca gagctccaga 3420
atcaatttct gtggcaatgg gaaccatatc atctccttgg ttcagaatgc tccctacggc 3480
ctgtatttca tccacttcaa ctacgtgccc acgaagtacg ttacagccaa agtgtccccc 3540
ggactgtgca tcgctggtaa caggggcatt gcaccaaaat ccggctactt cgtcaatgtc 3600
aacaacacat ggatgtatac tgggagtggt tattattacc ctgaacctat aacagagaac 3660
aatgtagtag tcatgtccac atgcgccgtc aattatacta aggcccccta tgttatgctc 3720

aacacttcaa ttcccaatct cccggatttc aaagaagagc tggatcagtg gtttaagaat 3780
cagacatccg tggcccctga cttaagcttg gattatatca atgtgacttt tttagactta 3840
caggtcgaga tgaaccgact ccaggaagct ataaaagtac tgaaccactc ctatatcaat 3900
ctgaaagata tcggtacata cgaatattac gtaaaatggc cttggtatca aatactgtca 3960
atttattcaa cagtggcgag ttccctagca ctggcaatca tgatggctgg tctatcttta 4020
tggatgtgct ccaatggatc gttacaatgc agaatttgca tttga 4065

<210> 139
<211> 4056
<212> DNA

<213> Artificial Sequence

<220><223> PDI-OC43-H5iCT-DNA

<400> 139

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgatcgg cgatctgaat tgtaccctgg atccccgcct gaaagggagc 120

tttaacaacc gagatacagg acccccgtct atatccatag atacagtgga tgttacgaac 180

gggctcggca cctactatgt gctagaccga gtttatttga acaccacctt attcctcaat 240

ggatactacc caacttcagg tagtacttac agaaacatgg cgctgaaggg tacggatctg 300

ctgagcaccc tatggtttaa acctcccttc ctctcggact ttattaatgg catcttcgct 360

aaggtgaaaa acacgaaggt tttcaaagat ggagtgatgt attcagagtt ccctgcgatc 420

accattggaa gtaccttcgt gaatacttcc tatagcgtgg tggttcaacc acggacaatc 480

aactccaccc aggacggcgt caacaagctc cagggattgc tggaggtgtc agtctgtcaa 540

tataacatgt gtgagtaccc acacactatc tgtcacccta atctaggcaa ccactttaag 600

gaactgtggc actacgatac gggggtggta agttgcttat ataagagaaa tttcacctat 660

gatgttaatg caacgtacct gtactttcac ttctatcaag aaggaggaac tttctacgca 720

tatttcacag ataccggctt tgtgacgaaa ttcttattca acgtttacct cggaatggca 780

ttaagccatt attacgtgat gcctctcact tgcatcagac gccctaagga tggtttttct 840

ctggagtact gggtcactcc cctgacacca cggcagtacc tgcttgcttt taaccaggac 900

ggtatcattt ttaatgccgt cgattgtatg agcgatttta tgagcgagat aaagtgcaag 960

acccaatcta ttgctccgcc cacgggggtg tacgaactga atggttacac cgtccagccc 1020

gttgccgatg tatatagacg gaaaccagac ctgcccaatt gcaacatcga agcttggtta 1080

aacgataagt cagtgccctc ccccctcaat tgggagagga agactttctc caactgtaat 1140

ttcaacatgt caagcctgat gtctttcatt caagccgatt cgttcacttg taataatata 1200

gatgcagcaa agatctatgg tatgtgcttc agttccatca caatagataa gtttgcaata 1260

ccaaaccgtc gcaaggtgga ccttcagctc ggcaacctgg gctatctgca gtccagcaat 1320

tatagaatag acaccaccgc cacatcatgt cagctgtact ataacctccc agcagcgaac 1380

gtcagtgtta gtaggttcaa tccttctacc tggaataaaa ggtttggatt catcgaagat 1440

agtgtgttcg tacctcagcc aacaggagtg ttcaccaatc acagcgtggt ctacgcccaa 1500

cattgcttca aggcacccaa aaatttctgc ccatgtagca gttgctcctg cccgggtaag 1560

aacaatggga tcggcacctg cccagcaggc accaattcac ttacatgcga taatctgtgt 1620

acactggatc ctattacact taaggcccct gatacctaca aatgccccca gagcaagagc 1680

ctggtcggta tcggagaaca ctgttccgga cttgcagtaa aaagcgacta ttgtggaaat 1740
aactcttgca cttgtcagcc acaagccttc ctcggttggt ccgctgactc ttgtttacaa 1800
ggggataagt gtaacatctt cgcaaatttc atcttacacg atgtgaataa cggcttaaca 1860
tgcagcacag atctccagaa ggcaaacaca gagatcgaat taggagtctg cgttaattac 1920
gatctctacg ggatctctgg ccagggcatc ttcgtggagg ttaatgctac ctactacaat 1980
agttggcaaa atctgctcta cgatagcaat ggcaacctct atggattcag agactatatt 2040
actaacagga cgttcatgat tcactcgtgc tattccgggc gggtgtcagc agcttatcac 2100

gcaaattctt cagagccagc tctgctattc cgaaacataa aatgtaatta cgtgttcaat 2160
aattcactga ctcggcagct gcagccgatt aattacagct tcgacagcta ccttggttgc 2220
gttgttaacg cctacaactc cactgccata tcagttcaga cctgcgacct tactgtgggc 2280
tctggctatt gtgtcgatta ttcaaagaac gggggggagcg ggtccgcaat aacaactggc 2340
tataggttca ccaattttga gcctttcacc gtgaatagtg tcaacgatag cctggagcct 2400
gtcggaggtc tttatgagat acaaatcccc tccgagttca caattggcaa catggaagag 2460
ttcatccaga cgagttcccc aaaggtgacg atcgattgcg cggctttcgt ctgcggcgac 2520

tacgccgcat gcaagttaca actcgttgag tatggaagtt tttgcgataa tataaacgca 2580
attctgactg aagtgaacga actgctggac accactcagt tgcaggtggc aaattcgctc 2640
atgaacggcg tgacactgtc aaccaaactg aaggacggtg tcaatttcaa tgtggatgac 2700
attaacttca gccccgtact gggctgtttg ggtagtgagt gttctaaggc tagcagccgc 2760
tccgccattg aggacttgtt gtttgataaa gttaagctga gtgacgttgg atttgttgag 2820
gcgtataata actgtaccgg tggtgcagag ataagggatc tgatctgtgt ccagagttat 2880
aaggggatta aggttctccc cccgctactc tcggagaatc agatatcagg atacaccctg 2940

gccgctacct cagcctcgct gtttcccccct tggaccgctg ccgccggtgt cccattttat 3000
ttgaatgtgc agtatcggat caacggtctg ggagtgacaa tggacgtgct gtctcagaac 3060
cagaaactga tcgccaatgc attcaacaat gctctgcacg ccatccagca agggtttgac 3120
gctacaaatt ctgccctcgt aaaaatccag gccgtggtga atgctaacgc cgaagcccctt 3180
aataatctgc tccagcagct ttctaaccgc tttggagcta tttctgcctc actgcaggaa 3240
attctatcca gactggatcc ccctgaggca gaagcccaaa tcgaccgtct cataaacggc 3300
agactcactg ctcttaacgc ctacgttagt caacaattga gcgattcgac cttggtgaaa 3360

ttcagcgcag ctcaggctat ggagaaggtg aacgagtgcg tgaagtcaca gagctccaga 3420

atcaatttct gtggcaatgg gaaccatatc atctccttgg ttcagaatgc tccctacggc 3480

ctgtatttca tccacttcaa ctacgtgccc acgaagtacg ttacagccaa agtgtccccc 3540

ggactgtgca tcgctggtaa caggggcatt gcaccaaaat ccggctactt cgtcaatgtc 3600

aacaacacat ggatgtatac tgggagtggt tattattacc ctgaacctat aacagagaac 3660

aatgtagtag tcatgtccac atgcgccgtc aattatacta aggcccccta tgttatgctc 3720

aacacttcaa ttcccaatct cccggatttc aaagaagagc tggatcagtg gtttaagaat 3780

cagacatccg tggcccctga cttaagcttg gattatatca atgtgacttt tttagactta 3840

caggtcgaga tgaaccgact ccaggaagct ataaaagtac tgaaccactc ctatatcaat 3900

ctgaaagata tcggtacata cgaatattac gtaaaatggc cttggtatgt gtggctacta 3960

atttgccttg cgggcgtggc tatgctggtc ctgctgttct tcatttcttt atggatgtgc 4020

tccaatggat cgttacaatg cagaatttgc atttga 4056

<210> 140

<211> 4047

<212> DNA

<213> Artificial Sequence

<220><223> PDI-OC43-H5iCT(V4)-DNA

<400> 140

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgatcgg cgatctgaat tgtaccctgg atccccgcct gaaagggagc 120

tttaacaacc gagatacagg acccccgtct atatccatag atacagtgga tgttacgaac 180

gggctcggca cctactatgt gctagaccga gtttatttga acaccacctt attcctcaat 240

ggatactacc caacttcagg tagtacttac agaaacatgg cgctgaaggg tacggatctg 300

ctgagcaccc tatggtttaa acctcccttc ctctcggact ttattaatgg catcttcgct 360

aaggtgaaaa acacgaaggt tttcaaagat ggagtgatgt attcagagtt ccctgcgatc 420

accattggaa gtaccttcgt gaatacttcc tatagcgtgg tggttcaacc acggacaatc 480

aactccaccc aggacggcgt caacaagctc cagggattgc tggaggtgtc agtctgtcaa 540

tataacatgt gtgagtaccc acacactatc tgtcacccta atctaggcaa ccactttaag 600

gaactgtggc actacgatac gggggtggta agttgcttat ataagagaaa tttcacctat 660

gatgttaatg caacgtacct gtactttcac ttctatcaag aaggaggaac tttctacgca 720

tatttcacag ataccggctt tgtgacgaaa ttcttattca acgtttacct cggaatggca 780

ttaagccatt attacgtgat gcctctcact tgcatcagac gccctaagga tggtttttct 840

ctggagtact gggtcactcc cctgacacca cggcagtacc tgcttgcttt taaccaggac 900

ggtatcattt ttaatgccgt cgattgtatg agcgatttta tgagcgagat aaagtgcaag 960
acccaatcta ttgctccgcc cacgggggtg tacgaactga atggttacac cgtccagccc 1020
gttgccgatg tatatagacg gaaaccagac ctgcccaatt gcaacatcga agcttggtta 1080
aacgataagt cagtgccctc ccccctcaat tgggagagga agactttctc caactgtaat 1140
ttcaacatgt caagcctgat gtctttcatt caagccgatt cgttcacttg taataatata 1200
gatgcagcaa agatctatgg tatgtgcttc agttccatca caatagataa gtttgcaata 1260
ccaaaccgtc gcaaggtgga ccttcagctc ggcaacctgg gctatctgca gtccagcaat 1320

tatagaatag acaccaccgc cacatcatgt cagctgtact ataacctccc agcagcgaac 1380
gtcagtgtta gtaggttcaa tccttctacc tggaataaaa ggtttggatt catcgaagat 1440
agtgtgttcg tacctcagcc aacaggagtg ttcaccaatc acagcgtggt ctacgcccaa 1500
cattgcttca aggcacccaa aaatttctgc ccatgtagca gttgctcctg cccgggtaag 1560
aacaatggga tcggcacctg cccagcaggc accaattcac ttacatgcga taatctgtgt 1620
acactggatc ctattacact taaggcccct gatacctaca aatgccccca gagcaagagc 1680
ctggtcggta tcggagaaca ctgttccgga cttgcagtaa aaagcgacta ttgtggaaat 1740

aactcttgca cttgtcagcc acaagccttc ctcggttggt ccgctgactc ttgtttacaa 1800
ggggataagt gtaacatctt cgcaaatttc atcttacacg atgtgaataa cggcttaaca 1860
tgcagcacag atctccagaa ggcaaacaca gagatcgaat taggagtctg cgttaattac 1920
gatctctacg ggatctctgg ccagggcatc ttcgtggagg ttaatgctac ctactacaat 1980
agttggcaaa atctgctcta cgatagcaat ggcaacctct atggattcag agactatatt 2040
actaacagga cgttcatgat tcactcgtgc tattccgggc gggtgtcagc agcttatcac 2100
gcaaattctt cagagccagc tctgctattc cgaaacataa aatgtaatta cgtgttcaat 2160

aattcactga ctcggcagct gcagccgatt aattacagct tcgacagcta ccttggttgc 2220
gttgttaacg cctacaactc cactgccata tcagttcaga cctgcgacct tactgtgggc 2280
tctggctatt gtgtcgatta ttcaaagaac gggggggagcg ggtccgcaat aacaactggc 2340
tataggttca ccaattttga gcctttcacc gtgaatagtg tcaacgatag cctggagcct 2400
gtcggaggtc tttatgagat acaaatcccc tccgagttca caattggcaa catggaagag 2460
ttcatccaga cgagttcccc aaaggtgacg atcgattgcg cggctttcgt ctgcggcgac 2520
tacgccgcat gcaagttaca actcgttgag tatggaagtt tttgcgataa tataaacgca 2580

attctgactg aagtgaacga actgctggac accactcagt tgcaggtggc aaattcgctc 2640

atgaacggcg tgacactgtc aaccaaactg aaggacggtg tcaatttcaa tgtggatgac 2700

attaacttca gccccgtact gggctgtttg ggtagtgagt gttctaaggc tagcagccgc 2760

tccgccattg aggacttgtt gtttgataaa gttaagctga gtgacgttgg atttgttgag 2820

gcgtataata actgtaccgg tggtgcagag ataagggatc tgatctgtgt ccagagttat 2880

aaggggatta aggttctccc cccgctactc tcggagaatc agatatcagg atacaccctg 2940

gccgctacct cagcctcgct gtttcccccc tggaccgctg ccgccggtgt cccattttat 3000

ttgaatgtgc agtatcggat caacggtctg ggagtgacaa tggacgtgct gtctcagaac 3060

cagaaactga tcgccaatgc attcaacaat gctctgcacg ccatccagca agggtttgac 3120

gctacaaatt ctgccctcgt aaaaatccag gccgtggtga atgctaacgc cgaagccctt 3180

aataatctgc tccagcagct ttctaaccgc tttggagcta tttctgcctc actgcaggaa 3240

attctatcca gactggatcc ccctgaggca gaagcccaaa tcgaccgtct cataaacggc 3300

agactcactg ctcttaacgc ctacgttagt caacaattga gcgattcgac cttggtgaaa 3360

ttcagcgcag ctcaggctat ggagaaggtg aacgagtgcg tgaagtcaca gagctccaga 3420

atcaatttct gtggcaatgg gaaccatatc atctccttgg ttcagaatgc tccctacggc 3480

ctgtatttca tccacttcaa ctacgtgccc acgaagtacg ttacagccaa agtgtccccc 3540

ggactgtgca tcgctggtaa caggggcatt gcaccaaaat ccggctactt cgtcaatgtc 3600

aacaacacat ggatgtatac tgggagtggt tattattacc ctgaacctat aacagagaac 3660

aatgtagtag tcatgtccac atgcgccgtc aattatacta aggcccccta tgttatgctc 3720

aacacttcaa ttcccaatct cccggatttc aaagaagagc tggatcagtg gtttaagaat 3780

cagacatccg tggcccctga cttaagcttg gattatatca atgtgacttt tttagactta 3840

caggtcgaga tgaaccgact ccaggaagct ataaaagtac tgaaccactc ctatatcaat 3900

ctgaaagata tcggtacata cgaatattac gtaaaatggc cttggtatgt gtggctacta 3960

atttgccttg cgggcgtggc tatgctggtc ctgctgttct tcatttgctg ctccaatgga 4020

tcgttacaat gcagaatttg catttga 4047

<210> 141

<211> 4056

<212> DNA

<213> Artificial Sequence

<220><223> PDI-OC43-H1cCT-DNA

<400> 141

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggtgatcgg cgatctgaat tgtaccctgg atccccgcct gaaagggagc 120

tttaacaacc gagatacagg acccccgtct atatccatag atacagtgga tgttacgaac 180

gggctcggca cctactatgt gctagaccga gtttatttga acaccacctt attcctcaat 240

ggatactacc caacttcagg tagtacttac agaaacatgg cgctgaaggg tacggatctg 300

ctgagcaccc tatggtttaa acctcccttc ctctcggact ttattaatgg catcttcgct 360

aaggtgaaaa acacgaaggt tttcaaagat ggagtgatgt attcagagtt ccctgcgatc 420

accattggaa gtaccttcgt gaatacttcc tatagcgtgg tggttcaacc acggacaatc 480

aactccaccc aggacggcgt caacaagctc cagggattgc tggaggtgtc agtctgtcaa 540

tataacatgt gtgagtaccc acacactatc tgtcacccta atctaggcaa ccactttaag 600

gaactgtggc actacgatac gggggtggta agttgcttat ataagagaaa tttcacctat 660

gatgttaatg caacgtacct gtactttcac ttctatcaag aaggaggaac tttctacgca 720

tatttcacag ataccggctt tgtgacgaaa ttcttattca acgtttacct cggaatggca 780

ttaagccatt attacgtgat gcctctcact tgcatcagac gccctaagga tggtttttct 840

ctggagtact gggtcactcc cctgacacca cggcagtacc tgcttgcttt taaccaggac 900

ggtatcattt ttaatgccgt cgattgtatg agcgatttta tgagcgagat aaagtgcaag 960

acccaatcta ttgctccgcc cacgggggtg tacgaactga atggttacac cgtccagccc 1020

gttgccgatg tatatagacg gaaaccagac ctgcccaatt gcaacatcga agcttggtta 1080

aacgataagt cagtgccctc ccccctcaat tgggagagga agactttctc caactgtaat 1140

ttcaacatgt caagcctgat gtctttcatt caagccgatt cgttcacttg taataatata 1200

gatgcagcaa agatctatgg tatgtgcttc agttccatca caatagataa gtttgcaata 1260

ccaaaccgtc gcaaggtgga ccttcagctc ggcaacctgg gctatctgca gtccagcaat 1320

tatagaatag acaccaccgc cacatcatgt cagctgtact ataacctccc agcagcgaac 1380

gtcagtgtta gtaggttcaa tccttctacc tggaataaaa ggtttggatt catcgaagat 1440

agtgtgttcg tacctcagcc aacaggagtg ttcaccaatc acagcgtggt ctacgcccaa 1500

cattgcttca aggcacccaa aaatttctgc ccatgtagca gttgctcctg cccgggtaag 1560

aacaatggga tcggcacctg cccagcaggc accaattcac ttacatgcga taatctgtgt 1620

acactggatc ctattacact taaggcccct gatacctaca aatgccccca gagcaagagc 1680

ctggtcggta tcggagaaca ctgttccgga cttgcagtaa aaagcgacta ttgtggaaat 1740

aactcttgca cttgtcagcc acaagccttc ctcggttggt ccgctgactc ttgtttacaa 1800

ggggataagt gtaacatctt cgcaaatttc atcttacacg atgtgaataa cggcttaaca 1860

tgcagcacag atctccagaa ggcaaacaca gagatcgaat taggagtctg cgttaattac 1920

gatctctacg ggatctctgg ccagggcatc ttcgtggagg ttaatgctac ctactacaat 1980

agttggcaaa atctgctcta cgatagcaat ggcaacctct atggattcag agactatatt 2040

actaacagga cgttcatgat tcactcgtgc tattccgggc gggtgtcagc agcttatcac 2100

gcaaattctt cagagccagc tctgctattc cgaaacataa aatgtaatta cgtgttcaat 2160

aattcactga ctcggcagct gcagccgatt aattacagct tcgacagcta ccttggttgc 2220

gttgttaacg cctacaactc cactgccata tcagttcaga cctgcgacct tactgtgggc 2280

tctggctatt gtgtcgatta ttcaaagaac ggggggagcg ggtccgcaat aacaactggc 2340

tataggttca ccaattttga gcctttcacc gtgaatagtg tcaacgatag cctggagcct 2400

gtcggaggtc tttatgagat acaaatcccc tccgagttca caattggcaa catggaagag 2460

ttcatccaga cgagttcccc aaaggtgacg atcgattgcg cggctttcgt ctgcggcgac 2520

tacgccgcat gcaagttaca actcgttgag tatggaagtt tttgcgataa tataaacgca 2580

attctgactg aagtgaacga actgctggac accactcagt tgcaggtggc aaattcgctc 2640

atgaacggcg tgacactgtc aaccaaactg aaggacggtg tcaatttcaa tgtggatgac 2700

attaacttca gccccgtact gggctgtttg ggtagtgagt gttctaaggc tagcagccgc 2760

tccgccattg aggacttgtt gtttgataaa gttaagctga gtgacgttgg atttgttgag 2820

gcgtataata actgtaccgg tggtgcagag ataagggatc tgatctgtgt ccagagttat 2880

aaggggatta aggttctccc cccgctactc tcggagaatc agatatcagg atacaccctg 2940

gccgctacct cagcctcgct gtttcccccc tggaccgctg ccgccggtgt cccattttat 3000

ttgaatgtgc agtatcggat caacggtctg ggagtgacaa tggacgtgct gtctcagaac 3060

cagaaactga tcgccaatgc attcaacaat gctctgcacg ccatccagca agggtttgac 3120

gctacaaatt ctgccctcgt aaaaatccag gccgtggtga atgctaacgc cgaagcccct 3180

aataatctgc tccagcagct ttctaaccgc tttggagcta tttctgcctc actgcaggaa 3240

attctatcca gactggatcc ccctgaggca gaagcccaaa tcgaccgtct cataaacggc 3300

agactcactg ctcttaacgc ctacgttagt caacaattga gcgattcgac cttggtgaaa 3360

ttcagcgcag ctcaggctat ggagaaggtg aacgagtgcg tgaagtcaca gagctccaga 3420

atcaatttct gtggcaatgg gaaccatatc atctccttgg ttcagaatgc tccctacggc 3480

ctgtatttca tccacttcaa ctacgtgccc acgaagtacg ttacagccaa agtgtccccc 3540

ggactgtgca tcgctggtaa caggggcatt gcaccaaaat ccggctactt cgtcaatgtc 3600

aacaacacat ggatgtatac tgggagtggt tattattacc ctgaacctat aacagagaac 3660

aatgtagtag tcatgtccac atgcgccgtc aattatacta aggcccccta tgttatgctc 3720

aacacttcaa ttcccaatct cccggatttc aaagaagagc tggatcagtg gtttaagaat 3780

cagacatccg tggcccctga cttaagcttg gattatatca atgtgacttt tttagactta 3840

caggtcgaga tgaaccgact ccaggaagct ataaaagtac tgaaccactc ctatatcaat 3900

ctgaaagata tcggtacata cgaatattac gtaaaatggc cttggtatgt gtggctacta 3960

atttgccttg cgggcgtggc tatgctggtc ctgctgttct tcattagctt ctggatgtgc 4020

tctaatgggt ctctacagtg tagaatatgt atttga 4056

<210> 142

<211> 1370

<212> PRT

<213> Artificial Sequence

<220><223> PDI-OC43-wtTMCT-AA

<400> 142

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Val Ile Gly Asp Leu Asn Cys Thr
20 25 30

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

Pro Ser Ile Ser Ile Asp Thr Val Asp Val Thr Asn Gly Leu Gly Thr
50 55 60

Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn Thr Thr Leu Phe Leu Asn
65 70 75 80

Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr Arg Asn Met Ala Leu Lys
85 90 95

Gly Thr Asp Leu Leu Ser Thr Leu Trp Phe Lys Pro Pro Phe Leu Ser
100 105 110

Asp Phe Ile Asn Gly Ile Phe Ala Lys Val Lys Asn Thr Lys Val Phe
115 120 125

Lys Asp Gly Val Met Tyr Ser Glu Phe Pro Ala Ile Thr Ile Gly Ser
130 135 140

Thr Phe Val Asn Thr Ser Tyr Ser Val Val Val Gln Pro Arg Thr Ile

145 150 155 160

Asn Ser Thr Gln Asp Gly Val Asn Lys Leu Gln Gly Leu Leu Glu Val

165 170 175

Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr Pro His Thr Ile Cys His

180 185 190

Pro Asn Leu Gly Asn His Phe Lys Glu Leu Trp His Tyr Asp Thr Gly

195 200 205

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

210 215 220

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

225 230 235 240

Tyr Phe Thr Asp Thr Gly Phe Val Thr Lys Phe Leu Phe Asn Val Tyr

245 250 255

Leu Gly Met Ala Leu Ser His Tyr Tyr Val Met Pro Leu Thr Cys Ile

260 265 270

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

275 280 285

Thr Pro Arg Gln Tyr Leu Leu Ala Phe Asn Gln Asp Gly Ile Ile Phe

290 295 300

Asn Ala Val Asp Cys Met Ser Asp Phe Met Ser Glu Ile Lys Cys Lys

305 310 315 320

Thr Gln Ser Ile Ala Pro Pro Thr Gly Val Tyr Glu Leu Asn Gly Tyr

325 330 335

Thr Val Gln Pro Val Ala Asp Val Tyr Arg Arg Lys Pro Asp Leu Pro

340 345 350

Asn Cys Asn Ile Glu Ala Trp Leu Asn Asp Lys Ser Val Pro Ser Pro

355 360 365

Leu Asn Trp Glu Arg Lys Thr Phe Ser Asn Cys Asn Phe Asn Met Ser

370 375 380

Ser Leu Met Ser Phe Ile Gln Ala Asp Ser Phe Thr Cys Asn Asn Ile

385 390 395 400

Asp Ala Ala Lys Ile Tyr Gly Met Cys Phe Ser Ser Ile Thr Ile Asp

405 410 415

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

420 425 430

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

435 440 445

Ser Cys Gln Leu Tyr Tyr Asn Leu Pro Ala Ala Asn Val Ser Val Ser

450 455 460

Arg Phe Asn Pro Ser Thr Trp Asn Lys Arg Phe Gly Phe Ile Glu Asp

465 470 475 480

Ser Val Phe Val Pro Gln Pro Thr Gly Val Phe Thr Asn His Ser Val

485 490 495

Val Tyr Ala Gln His Cys Phe Lys Ala Pro Lys Asn Phe Cys Pro Cys

500 505 510

Ser Ser Cys Ser Cys Pro Gly Lys Asn Asn Gly Ile Gly Thr Cys Pro

515 520 525

Ala Gly Thr Asn Ser Leu Thr Cys Asp Asn Leu Cys Thr Leu Asp Pro

530 535 540

Ile Thr Leu Lys Ala Pro Asp Thr Tyr Lys Cys Pro Gln Ser Lys Ser

545 550 555 560

Leu Val Gly Ile Gly Glu His Cys Ser Gly Leu Ala Val Lys Ser Asp

565 570 575

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

580 585 590

Trp Ser Ala Asp Ser Cys Leu Gln Gly Asp Lys Cys Asn Ile Phe Ala

595 600 605

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

610 615 620

Leu Gln Lys Ala Asn Thr Glu Ile Glu Leu Gly Val Cys Val Asn Tyr

625 630 635 640

Asp Leu Tyr Gly Ile Ser Gly Gln Gly Ile Phe Val Glu Val Asn Ala
645 650 655
Thr Tyr Tyr Asn Ser Trp Gln Asn Leu Leu Tyr Asp Ser Asn Gly Asn

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

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

785 790 795 800
Val Gly Gly Leu Tyr Glu Ile Gln Ile Pro Ser Glu Phe Thr Ile Gly
805 810 815
Asn Met Glu Glu Phe Ile Gln Thr Ser Ser Pro Lys Val Thr Ile Asp
820 825 830
Cys Ala Ala Phe Val Cys Gly Asp Tyr Ala Ala Cys Lys Leu Gln Leu
835 840 845
Val Glu Tyr Gly Ser Phe Cys Asp Asn Ile Asn Ala Ile Leu Thr Glu

850 855 860
Val Asn Glu Leu Leu Asp Thr Thr Gln Leu Gln Val Ala Asn Ser Leu
865 870 875 880
Met Asn Gly Val Thr Leu Ser Thr Lys Leu Lys Asp Gly Val Asn Phe

885 890 895

Asn Val Asp Asp Ile Asn Phe Ser Pro Val Leu Gly Cys Leu Gly Ser

900 905 910

Glu Cys Ser Lys Ala Ser Ser Arg Ser Ala Ile Glu Asp Leu Leu Phe

915 920 925

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

930 935 940

Cys Thr Gly Gly Ala Glu Ile Arg Asp Leu Ile Cys Val Gln Ser Tyr

945 950 955 960

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

965 970 975

Gly Tyr Thr Leu Ala Ala Thr Ser Ala Ser Leu Phe Pro Pro Trp Thr

980 985 990

Ala Ala Ala Gly Val Pro Phe Tyr Leu Asn Val Gln Tyr Arg Ile Asn

995 1000 1005

Gly Leu Gly Val Thr Met Asp Val Leu Ser Gln Asn Gln Lys Leu

1010 1015 1020

Ile Ala Asn Ala Phe Asn Asn Ala Leu His Ala Ile Gln Gln Gly

1025 1030 1035

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

1040 1045 1050

Asn Ala Asn Ala Glu Ala Leu Asn Asn Leu Leu Gln Gln Leu Ser

1055 1060 1065

Asn Arg Phe Gly Ala Ile Ser Ala Ser Leu Gln Glu Ile Leu Ser

1070 1075 1080

Arg Leu Asp Pro Pro Glu Ala Glu Ala Gln Ile Asp Arg Leu Ile

1085 1090 1095

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

1100 1105 1110

Ser Asp Ser Thr Leu Val Lys Phe Ser Ala Ala Gln Ala Met Glu

1115 1120 1125

Lys Val Asn Glu Cys Val Lys Ser Gln Ser Ser Arg Ile Asn Phe
1130 1135 1140
Cys Gly Asn Gly Asn His Ile Ile Ser Leu Val Gln Asn Ala Pro
1145 1150 1155
Tyr Gly Leu Tyr Phe Ile His Phe Asn Tyr Val Pro Thr Lys Tyr

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

1220 1225 1230
Lys Ala Pro Tyr Val Met Leu Asn Thr Ser Ile Pro Asn Leu Pro
1235 1240 1245
Asp Phe Lys Glu Glu Leu Asp Gln Trp Phe Lys Asn Gln Thr Ser
1250 1255 1260
Val Ala Pro Asp Leu Ser Leu Asp Tyr Ile Asn Val Thr Phe Leu
1265 1270 1275
Asp Leu Gln Val Glu Met Asn Arg Leu Gln Glu Ala Ile Lys Val

1280 1285 1290
Leu Asn His Ser Tyr Ile Asn Leu Lys Asp Ile Gly Thr Tyr Glu
1295 1300 1305
Tyr Tyr Val Lys Trp Pro Trp Tyr Val Trp Leu Leu Ile Cys Leu
1310 1315 1320
Ala Gly Val Ala Met Leu Val Leu Leu Phe Phe Ile Cys Cys Cys
1325 1330 1335
Thr Gly Cys Gly Thr Ser Cys Phe Lys Lys Cys Gly Gly Cys Cys

1340 1345 1350
Asp Asp Tyr Thr Gly Tyr Gln Glu Leu Val Ile Lys Thr Ser His

1355 1360 1365

Asp Asp

1370

<210> 143

<211> 1354

<212> PRT

<213> Artificial Sequence

<220><223> PDI-OC43-H5iTMCT-AA

<400> 143

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Val Ile Gly Asp Leu Asn Cys Thr
20 25 30

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

Pro Ser Ile Ser Ile Asp Thr Val Asp Val Thr Asn Gly Leu Gly Thr
50 55 60

Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn Thr Thr Leu Phe Leu Asn
65 70 75 80

Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr Arg Asn Met Ala Leu Lys
85 90 95

Gly Thr Asp Leu Leu Ser Thr Leu Trp Phe Lys Pro Pro Phe Leu Ser
100 105 110

Asp Phe Ile Asn Gly Ile Phe Ala Lys Val Lys Asn Thr Lys Val Phe
115 120 125

Lys Asp Gly Val Met Tyr Ser Glu Phe Pro Ala Ile Thr Ile Gly Ser
130 135 140

Thr Phe Val Asn Thr Ser Tyr Ser Val Val Val Gln Pro Arg Thr Ile
145 150 155 160

Asn Ser Thr Gln Asp Gly Val Asn Lys Leu Gln Gly Leu Leu Glu Val
165 170 175

Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr Pro His Thr Ile Cys His

180 185 190

Pro Asn Leu Gly Asn His Phe Lys Glu Leu Trp His Tyr Asp Thr Gly

195 200 205

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

210 215 220

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

225 230 235 240

Tyr Phe Thr Asp Thr Gly Phe Val Thr Lys Phe Leu Phe Asn Val Tyr

245 250 255

Leu Gly Met Ala Leu Ser His Tyr Tyr Val Met Pro Leu Thr Cys Ile

260 265 270

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

275 280 285

Thr Pro Arg Gln Tyr Leu Leu Ala Phe Asn Gln Asp Gly Ile Ile Phe

290 295 300

Asn Ala Val Asp Cys Met Ser Asp Phe Met Ser Glu Ile Lys Cys Lys

305 310 315 320

Thr Gln Ser Ile Ala Pro Pro Thr Gly Val Tyr Glu Leu Asn Gly Tyr

325 330 335

Thr Val Gln Pro Val Ala Asp Val Tyr Arg Arg Lys Pro Asp Leu Pro

340 345 350

Asn Cys Asn Ile Glu Ala Trp Leu Asn Asp Lys Ser Val Pro Ser Pro

355 360 365

Leu Asn Trp Glu Arg Lys Thr Phe Ser Asn Cys Asn Phe Asn Met Ser

370 375 380

Ser Leu Met Ser Phe Ile Gln Ala Asp Ser Phe Thr Cys Asn Asn Ile

385 390 395 400

Asp Ala Ala Lys Ile Tyr Gly Met Cys Phe Ser Ser Ile Thr Ile Asp

405 410 415

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

420 425 430

Leu Gly Tyr Leu Gln Ser Ser Asn Tyr Arg Ile Asp Thr Thr Ala Thr
435 440 445
Ser Cys Gln Leu Tyr Tyr Asn Leu Pro Ala Ala Asn Val Ser Val Ser
450 455 460
Arg Phe Asn Pro Ser Thr Trp Asn Lys Arg Phe Gly Phe Ile Glu Asp

465 470 475 480
Ser Val Phe Val Pro Gln Pro Thr Gly Val Phe Thr Asn His Ser Val
485 490 495
Val Tyr Ala Gln His Cys Phe Lys Ala Pro Lys Asn Phe Cys Pro Cys
500 505 510
Ser Ser Cys Ser Cys Pro Gly Lys Asn Asn Gly Ile Gly Thr Cys Pro
515 520 525
Ala Gly Thr Asn Ser Leu Thr Cys Asp Asn Leu Cys Thr Leu Asp Pro

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

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

660 665 670
Leu Tyr Gly Phe Arg Asp Tyr Ile Thr Asn Arg Thr Phe Met Ile His

675 680 685

Ser Cys Tyr Ser Gly Arg Val Ser Ala Ala Tyr His Ala Asn Ser Ser

690 695 700

Glu Pro Ala Leu Leu Phe Arg Asn Ile Lys Cys Asn Tyr Val Phe Asn

705 710 715 720

Asn Ser Leu Thr Arg Gln Leu Gln Pro Ile Asn Tyr Ser Phe Asp Ser

725 730 735

Tyr Leu Gly Cys Val Val Asn Ala Tyr Asn Ser Thr Ala Ile Ser Val

740 745 750

Gln Thr Cys Asp Leu Thr Val Gly Ser Gly Tyr Cys Val Asp Tyr Ser

755 760 765

Lys Asn Gly Gly Ser Gly Ser Ala Ile Thr Thr Gly Tyr Arg Phe Thr

770 775 780

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

785 790 795 800

Val Gly Gly Leu Tyr Glu Ile Gln Ile Pro Ser Glu Phe Thr Ile Gly

805 810 815

Asn Met Glu Glu Phe Ile Gln Thr Ser Ser Pro Lys Val Thr Ile Asp

820 825 830

Cys Ala Ala Phe Val Cys Gly Asp Tyr Ala Ala Cys Lys Leu Gln Leu

835 840 845

Val Glu Tyr Gly Ser Phe Cys Asp Asn Ile Asn Ala Ile Leu Thr Glu

850 855 860

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

865 870 875 880

Met Asn Gly Val Thr Leu Ser Thr Lys Leu Lys Asp Gly Val Asn Phe

885 890 895

Asn Val Asp Asp Ile Asn Phe Ser Pro Val Leu Gly Cys Leu Gly Ser

900 905 910

Glu Cys Ser Lys Ala Ser Ser Arg Ser Ala Ile Glu Asp Leu Leu Phe

915 920 925

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

930 935 940

Cys Thr Gly Gly Ala Glu Ile Arg Asp Leu Ile Cys Val Gln Ser Tyr

945 950 955 960

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

965 970 975

Gly Tyr Thr Leu Ala Ala Thr Ser Ala Ser Leu Phe Pro Pro Trp Thr

980 985 990

Ala Ala Ala Gly Val Pro Phe Tyr Leu Asn Val Gln Tyr Arg Ile Asn

995 1000 1005

Gly Leu Gly Val Thr Met Asp Val Leu Ser Gln Asn Gln Lys Leu

1010 1015 1020

Ile Ala Asn Ala Phe Asn Asn Ala Leu His Ala Ile Gln Gln Gly

1025 1030 1035

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

1040 1045 1050

Asn Ala Asn Ala Glu Ala Leu Asn Asn Leu Leu Gln Gln Leu Ser

1055 1060 1065

Asn Arg Phe Gly Ala Ile Ser Ala Ser Leu Gln Glu Ile Leu Ser

1070 1075 1080

Arg Leu Asp Pro Pro Glu Ala Glu Ala Gln Ile Asp Arg Leu Ile

1085 1090 1095

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

1100 1105 1110

Ser Asp Ser Thr Leu Val Lys Phe Ser Ala Ala Gln Ala Met Glu

1115 1120 1125

Lys Val Asn Glu Cys Val Lys Ser Gln Ser Ser Arg Ile Asn Phe

1130 1135 1140

Cys Gly Asn Gly Asn His Ile Ile Ser Leu Val Gln Asn Ala Pro

1145 1150 1155

Tyr Gly Leu Tyr Phe Ile His Phe Asn Tyr Val Pro Thr Lys Tyr

1160 1165 1170

Val Thr Ala Lys Val Ser Pro Gly Leu Cys Ile Ala Gly Asn Arg

1175 1180 1185

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

1190 1195 1200

Trp Met Tyr Thr Gly Ser Gly Tyr Tyr Tyr Pro Glu Pro Ile Thr

1205 1210 1215

Glu Asn Asn Val Val Val Met Ser Thr Cys Ala Val Asn Tyr Thr

1220 1225 1230

Lys Ala Pro Tyr Val Met Leu Asn Thr Ser Ile Pro Asn Leu Pro

1235 1240 1245

Asp Phe Lys Glu Glu Leu Asp Gln Trp Phe Lys Asn Gln Thr Ser

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

Leu Asn His Ser Tyr Ile Asn Leu Lys Asp Ile Gly Thr Tyr Glu

1295 1300 1305

Tyr Tyr Val Lys Trp Pro Trp Tyr Gln Ile Leu Ser Ile Tyr Ser

1310 1315 1320

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

1325 1330 1335

Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys

1340 1345 1350

Ile

<210> 144

<211> 1351

<212> PRT

<213> Artificial Sequence

<220><223> PDI-OC43-H5iCT-AA

<400> 144

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Val Ile Gly Asp Leu Asn Cys Thr
20 25 30

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

Pro Ser Ile Ser Ile Asp Thr Val Asp Val Thr Asn Gly Leu Gly Thr
50 55 60

Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn Thr Thr Leu Phe Leu Asn
65 70 75 80

Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr Arg Asn Met Ala Leu Lys
85 90 95

Gly Thr Asp Leu Leu Ser Thr Leu Trp Phe Lys Pro Pro Phe Leu Ser
100 105 110

Asp Phe Ile Asn Gly Ile Phe Ala Lys Val Lys Asn Thr Lys Val Phe
115 120 125

Lys Asp Gly Val Met Tyr Ser Glu Phe Pro Ala Ile Thr Ile Gly Ser
130 135 140

Thr Phe Val Asn Thr Ser Tyr Ser Val Val Val Gln Pro Arg Thr Ile
145 150 155 160

Asn Ser Thr Gln Asp Gly Val Asn Lys Leu Gln Gly Leu Leu Glu Val
165 170 175

Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr Pro His Thr Ile Cys His
180 185 190

Pro Asn Leu Gly Asn His Phe Lys Glu Leu Trp His Tyr Asp Thr Gly
195 200 205

Val Val Ser Cys Leu Tyr Lys Arg Asn Phe Thr Tyr Asp Val Asn Ala
210 215 220

Thr Tyr Leu Tyr Phe His Phe Tyr Gln Glu Gly Gly Thr Phe Tyr Ala
225 230 235 240

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

290 295 300
Asn Ala Val Asp Cys Met Ser Asp Phe Met Ser Glu Ile Lys Cys Lys
305 310 315 320
Thr Gln Ser Ile Ala Pro Pro Thr Gly Val Tyr Glu Leu Asn Gly Tyr
325 330 335
Thr Val Gln Pro Val Ala Asp Val Tyr Arg Arg Lys Pro Asp Leu Pro
340 345 350
Asn Cys Asn Ile Glu Ala Trp Leu Asn Asp Lys Ser Val Pro Ser Pro

355 360 365
Leu Asn Trp Glu Arg Lys Thr Phe Ser Asn Cys Asn Phe Asn Met Ser
370 375 380
Ser Leu Met Ser Phe Ile Gln Ala Asp Ser Phe Thr Cys Asn Asn Ile
385 390 395 400
Asp Ala Ala Lys Ile Tyr Gly Met Cys Phe Ser Ser Ile Thr Ile Asp
405 410 415
Lys Phe Ala Ile Pro Asn Arg Arg Lys Val Asp Leu Gln Leu Gly Asn

420 425 430
Leu Gly Tyr Leu Gln Ser Ser Asn Tyr Arg Ile Asp Thr Thr Ala Thr
435 440 445
Ser Cys Gln Leu Tyr Tyr Asn Leu Pro Ala Ala Asn Val Ser Val Ser
450 455 460
Arg Phe Asn Pro Ser Thr Trp Asn Lys Arg Phe Gly Phe Ile Glu Asp
465 470 475 480
Ser Val Phe Val Pro Gln Pro Thr Gly Val Phe Thr Asn His Ser Val

485 490 495

Val Tyr Ala Gln His Cys Phe Lys Ala Pro Lys Asn Phe Cys Pro Cys

500 505 510

Ser Ser Cys Ser Cys Pro Gly Lys Asn Asn Gly Ile Gly Thr Cys Pro

515 520 525

Ala Gly Thr Asn Ser Leu Thr Cys Asp Asn Leu Cys Thr Leu Asp Pro

530 535 540

Ile Thr Leu Lys Ala Pro Asp Thr Tyr Lys Cys Pro Gln Ser Lys Ser

545 550 555 560

Leu Val Gly Ile Gly Glu His Cys Ser Gly Leu Ala Val Lys Ser Asp

565 570 575

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

580 585 590

Trp Ser Ala Asp Ser Cys Leu Gln Gly Asp Lys Cys Asn Ile Phe Ala

595 600 605

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

610 615 620

Leu Gln Lys Ala Asn Thr Glu Ile Glu Leu Gly Val Cys Val Asn Tyr

625 630 635 640

Asp Leu Tyr Gly Ile Ser Gly Gln Gly Ile Phe Val Glu Val Asn Ala

645 650 655

Thr Tyr Tyr Asn Ser Trp Gln Asn Leu Leu Tyr Asp Ser Asn Gly Asn

660 665 670

Leu Tyr Gly Phe Arg Asp Tyr Ile Thr Asn Arg Thr Phe Met Ile His

675 680 685

Ser Cys Tyr Ser Gly Arg Val Ser Ala Ala Tyr His Ala Asn Ser Ser

690 695 700

Glu Pro Ala Leu Leu Phe Arg Asn Ile Lys Cys Asn Tyr Val Phe Asn

705 710 715 720

Asn Ser Leu Thr Arg Gln Leu Gln Pro Ile Asn Tyr Ser Phe Asp Ser

725 730 735

Tyr Leu Gly Cys Val Val Asn Ala Tyr Asn Ser Thr Ala Ile Ser Val

740 745 750

Gln Thr Cys Asp Leu Thr Val Gly Ser Gly Tyr Cys Val Asp Tyr Ser

755 760 765

Lys Asn Gly Gly Ser Gly Ser Ala Ile Thr Thr Gly Tyr Arg Phe Thr

770 775 780

Asn Phe Glu Pro Phe Thr Val Asn Ser Val Asn Asp Ser Leu Glu Pro

785 790 795 800

Val Gly Gly Leu Tyr Glu Ile Gln Ile Pro Ser Glu Phe Thr Ile Gly

805 810 815

Asn Met Glu Glu Phe Ile Gln Thr Ser Ser Pro Lys Val Thr Ile Asp

820 825 830

Cys Ala Ala Phe Val Cys Gly Asp Tyr Ala Ala Cys Lys Leu Gln Leu

835 840 845

Val Glu Tyr Gly Ser Phe Cys Asp Asn Ile Asn Ala Ile Leu Thr Glu

850 855 860

Val Asn Glu Leu Leu Asp Thr Thr Gln Leu Gln Val Ala Asn Ser Leu

865 870 875 880

Met Asn Gly Val Thr Leu Ser Thr Lys Leu Lys Asp Gly Val Asn Phe

885 890 895

Asn Val Asp Asp Ile Asn Phe Ser Pro Val Leu Gly Cys Leu Gly Ser

900 905 910

Glu Cys Ser Lys Ala Ser Ser Arg Ser Ala Ile Glu Asp Leu Leu Phe

915 920 925

Asp Lys Val Lys Leu Ser Asp Val Gly Phe Val Glu Ala Tyr Asn Asn

930 935 940

Cys Thr Gly Gly Ala Glu Ile Arg Asp Leu Ile Cys Val Gln Ser Tyr

945 950 955 960

Lys Gly Ile Lys Val Leu Pro Pro Leu Leu Ser Glu Asn Gln Ile Ser

965 970 975

Gly Tyr Thr Leu Ala Ala Thr Ser Ala Ser Leu Phe Pro Pro Trp Thr

980 985 990

Ala Ala Ala Gly Val Pro Phe Tyr Leu Asn Val Gln Tyr Arg Ile Asn

995 1000 1005

Gly Leu Gly Val Thr Met Asp Val Leu Ser Gln Asn Gln Lys Leu

1010 1015 1020

Ile Ala Asn Ala Phe Asn Asn Ala Leu His Ala Ile Gln Gln Gly

1025 1030 1035

Phe Asp Ala Thr Asn Ser Ala Leu Val Lys Ile Gln Ala Val Val

1040 1045 1050

Asn Ala Asn Ala Glu Ala Leu Asn Asn Leu Leu Gln Gln Leu Ser

1055 1060 1065

Asn Arg Phe Gly Ala Ile Ser Ala Ser Leu Gln Glu Ile Leu Ser

1070 1075 1080

Arg Leu Asp Pro Pro Glu Ala Glu Ala Gln Ile Asp Arg Leu Ile

1085 1090 1095

Asn Gly Arg Leu Thr Ala Leu Asn Ala Tyr Val Ser Gln Gln Leu

1100 1105 1110

Ser Asp Ser Thr Leu Val Lys Phe Ser Ala Ala Gln Ala Met Glu

1115 1120 1125

Lys Val Asn Glu Cys Val Lys Ser Gln Ser Ser Arg Ile Asn Phe

1130 1135 1140

Cys Gly Asn Gly Asn His Ile Ile Ser Leu Val Gln Asn Ala Pro

1145 1150 1155

Tyr Gly Leu Tyr Phe Ile His Phe Asn Tyr Val Pro Thr Lys Tyr

1160 1165 1170

Val Thr Ala Lys Val Ser Pro Gly Leu Cys Ile Ala Gly Asn Arg

1175 1180 1185

Gly Ile Ala Pro Lys Ser Gly Tyr Phe Val Asn Val Asn Asn Thr

1190 1195 1200

Trp Met Tyr Thr Gly Ser Gly Tyr Tyr Tyr Pro Glu Pro Ile Thr

1205 1210 1215

Glu Asn Asn Val Val Val Met Ser Thr Cys Ala Val Asn Tyr Thr

1220 1225 1230

Lys Ala Pro Tyr Val Met Leu Asn Thr Ser Ile Pro Asn Leu Pro

1235 1240 1245

Asp Phe Lys Glu Glu Leu Asp Gln Trp Phe Lys Asn Gln Thr Ser

1250 1255 1260

Val Ala Pro Asp Leu Ser Leu Asp Tyr Ile Asn Val Thr Phe Leu

1265 1270 1275

Asp Leu Gln Val Glu Met Asn Arg Leu Gln Glu Ala Ile Lys Val

1280 1285 1290

Leu Asn His Ser Tyr Ile Asn Leu Lys Asp Ile Gly Thr Tyr Glu

1295 1300 1305

Tyr Tyr Val Lys Trp Pro Trp Tyr Val Trp Leu Leu Ile Cys Leu

1310 1315 1320

Ala Gly Val Ala Met Leu Val Leu Leu Phe Phe Ile Ser Leu Trp

1325 1330 1335

Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1340 1345 1350

<210> 145

<211> 1348

<212> PRT

<213> Artificial Sequence

<

220><223> PDI-OC43-H5iCT(V4)-AA

<400> 145

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Val Ile Gly Asp Leu Asn Cys Thr

20 25 30

Leu Asp Pro Arg Leu Lys Gly Ser Phe Asn Asn Arg Asp Thr Gly Pro

35 40 45

Pro Ser Ile Ser Ile Asp Thr Val Asp Val Thr Asn Gly Leu Gly Thr

50 55 60

Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn Thr Thr Leu Phe Leu Asn

65 70 75 80

Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr Arg Asn Met Ala Leu Lys

85 90 95

Gly Thr Asp Leu Leu Ser Thr Leu Trp Phe Lys Pro Pro Phe Leu Ser

100 105 110

Asp Phe Ile Asn Gly Ile Phe Ala Lys Val Lys Asn Thr Lys Val Phe

115 120 125

Lys Asp Gly Val Met Tyr Ser Glu Phe Pro Ala Ile Thr Ile Gly Ser

130 135 140

Thr Phe Val Asn Thr Ser Tyr Ser Val Val Val Gln Pro Arg Thr Ile

145 150 155 160

Asn Ser Thr Gln Asp Gly Val Asn Lys Leu Gln Gly Leu Leu Glu Val

165 170 175

Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr Pro His Thr Ile Cys His

180 185 190

Pro Asn Leu Gly Asn His Phe Lys Glu Leu Trp His Tyr Asp Thr Gly

195 200 205

Val Val Ser Cys Leu Tyr Lys Arg Asn Phe Thr Tyr Asp Val Asn Ala

210 215 220

Thr Tyr Leu Tyr Phe His Phe Tyr Gln Glu Gly Gly Thr Phe Tyr Ala

225 230 235 240

Tyr Phe Thr Asp Thr Gly Phe Val Thr Lys Phe Leu Phe Asn Val Tyr

245 250 255

Leu Gly Met Ala Leu Ser His Tyr Tyr Val Met Pro Leu Thr Cys Ile

260 265 270

Arg Arg Pro Lys Asp Gly Phe Ser Leu Glu Tyr Trp Val Thr Pro Leu

275 280 285

Thr Pro Arg Gln Tyr Leu Leu Ala Phe Asn Gln Asp Gly Ile Ile Phe

290 295 300

Asn Ala Val Asp Cys Met Ser Asp Phe Met Ser Glu Ile Lys Cys Lys

305 310 315 320

Thr Gln Ser Ile Ala Pro Pro Thr Gly Val Tyr Glu Leu Asn Gly Tyr

325 330 335

Thr Val Gln Pro Val Ala Asp Val Tyr Arg Arg Lys Pro Asp Leu Pro

340 345 350

Asn Cys Asn Ile Glu Ala Trp Leu Asn Asp Lys Ser Val Pro Ser Pro

355 360 365

Leu Asn Trp Glu Arg Lys Thr Phe Ser Asn Cys Asn Phe Asn Met Ser

370 375 380

Ser Leu Met Ser Phe Ile Gln Ala Asp Ser Phe Thr Cys Asn Asn Ile

385 390 395 400

Asp Ala Ala Lys Ile Tyr Gly Met Cys Phe Ser Ser Ile Thr Ile Asp

405 410 415

Lys Phe Ala Ile Pro Asn Arg Arg Lys Val Asp Leu Gln Leu Gly Asn

420 425 430

Leu Gly Tyr Leu Gln Ser Ser Asn Tyr Arg Ile Asp Thr Thr Ala Thr

435 440 445

Ser Cys Gln Leu Tyr Tyr Asn Leu Pro Ala Ala Asn Val Ser Val Ser

450 455 460

Arg Phe Asn Pro Ser Thr Trp Asn Lys Arg Phe Gly Phe Ile Glu Asp

465 470 475 480

Ser Val Phe Val Pro Gln Pro Thr Gly Val Phe Thr Asn His Ser Val

485 490 495

Val Tyr Ala Gln His Cys Phe Lys Ala Pro Lys Asn Phe Cys Pro Cys

500 505 510

Ser Ser Cys Ser Cys Pro Gly Lys Asn Asn Gly Ile Gly Thr Cys Pro

515 520 525

Ala Gly Thr Asn Ser Leu Thr Cys Asp Asn Leu Cys Thr Leu Asp Pro

530 535 540

Ile Thr Leu Lys Ala Pro Asp Thr Tyr Lys Cys Pro Gln Ser Lys Ser

545 550 555 560

Leu Val Gly Ile Gly Glu His Cys Ser Gly Leu Ala Val Lys Ser Asp

565 570 575

Tyr Cys Gly Asn Asn Ser Cys Thr Cys Gln Pro Gln Ala Phe Leu Gly

580 585 590

Trp Ser Ala Asp Ser Cys Leu Gln Gly Asp Lys Cys Asn Ile Phe Ala

595 600 605

Asn Phe Ile Leu His Asp Val Asn Asn Gly Leu Thr Cys Ser Thr Asp

610 615 620

Leu Gln Lys Ala Asn Thr Glu Ile Glu Leu Gly Val Cys Val Asn Tyr

625 630 635 640

Asp Leu Tyr Gly Ile Ser Gly Gln Gly Ile Phe Val Glu Val Asn Ala

645 650 655

Thr Tyr Tyr Asn Ser Trp Gln Asn Leu Leu Tyr Asp Ser Asn Gly Asn

660 665 670

Leu Tyr Gly Phe Arg Asp Tyr Ile Thr Asn Arg Thr Phe Met Ile His

675 680 685

Ser Cys Tyr Ser Gly Arg Val Ser Ala Ala Tyr His Ala Asn Ser Ser

690 695 700

Glu Pro Ala Leu Leu Phe Arg Asn Ile Lys Cys Asn Tyr Val Phe Asn

705 710 715 720

Asn Ser Leu Thr Arg Gln Leu Gln Pro Ile Asn Tyr Ser Phe Asp Ser

725 730 735

Tyr Leu Gly Cys Val Val Asn Ala Tyr Asn Ser Thr Ala Ile Ser Val

740 745 750

Gln Thr Cys Asp Leu Thr Val Gly Ser Gly Tyr Cys Val Asp Tyr Ser

755 760 765

Lys Asn Gly Gly Ser Gly Ser Ala Ile Thr Thr Gly Tyr Arg Phe Thr

770 775 780

Asn Phe Glu Pro Phe Thr Val Asn Ser Val Asn Asp Ser Leu Glu Pro

785 790 795 800

Val Gly Gly Leu Tyr Glu Ile Gln Ile Pro Ser Glu Phe Thr Ile Gly
805 810 815
Asn Met Glu Glu Phe Ile Gln Thr Ser Ser Pro Lys Val Thr Ile Asp

820 825 830
Cys Ala Ala Phe Val Cys Gly Asp Tyr Ala Ala Cys Lys Leu Gln Leu
835 840 845
Val Glu Tyr Gly Ser Phe Cys Asp Asn Ile Asn Ala Ile Leu Thr Glu
850 855 860
Val Asn Glu Leu Leu Asp Thr Thr Gln Leu Gln Val Ala Asn Ser Leu
865 870 875 880
Met Asn Gly Val Thr Leu Ser Thr Lys Leu Lys Asp Gly Val Asn Phe

885 890 895
Asn Val Asp Asp Ile Asn Phe Ser Pro Val Leu Gly Cys Leu Gly Ser
900 905 910
Glu Cys Ser Lys Ala Ser Ser Arg Ser Ala Ile Glu Asp Leu Leu Phe
915 920 925
Asp Lys Val Lys Leu Ser Asp Val Gly Phe Val Glu Ala Tyr Asn Asn
930 935 940
Cys Thr Gly Gly Ala Glu Ile Arg Asp Leu Ile Cys Val Gln Ser Tyr

945 950 955 960
Lys Gly Ile Lys Val Leu Pro Pro Leu Leu Ser Glu Asn Gln Ile Ser
965 970 975
Gly Tyr Thr Leu Ala Ala Thr Ser Ala Ser Leu Phe Pro Pro Trp Thr
980 985 990
Ala Ala Ala Gly Val Pro Phe Tyr Leu Asn Val Gln Tyr Arg Ile Asn
995 1000 1005
Gly Leu Gly Val Thr Met Asp Val Leu Ser Gln Asn Gln Lys Leu

1010 1015 1020
Ile Ala Asn Ala Phe Asn Asn Ala Leu His Ala Ile Gln Gln Gly
1025 1030 1035
Phe Asp Ala Thr Asn Ser Ala Leu Val Lys Ile Gln Ala Val Val

1040 1045 1050

Asn Ala Asn Ala Glu Ala Leu Asn Asn Leu Leu Gln Gln Leu Ser

1055 1060 1065

Asn Arg Phe Gly Ala Ile Ser Ala Ser Leu Gln Glu Ile Leu Ser

1070 1075 1080

Arg Leu Asp Pro Pro Glu Ala Glu Ala Gln Ile Asp Arg Leu Ile

1085 1090 1095

Asn Gly Arg Leu Thr Ala Leu Asn Ala Tyr Val Ser Gln Gln Leu

1100 1105 1110

Ser Asp Ser Thr Leu Val Lys Phe Ser Ala Ala Gln Ala Met Glu

1115 1120 1125

Lys Val Asn Glu Cys Val Lys Ser Gln Ser Ser Arg Ile Asn Phe

1130 1135 1140

Cys Gly Asn Gly Asn His Ile Ile Ser Leu Val Gln Asn Ala Pro

1145 1150 1155

Tyr Gly Leu Tyr Phe Ile His Phe Asn Tyr Val Pro Thr Lys Tyr

1160 1165 1170

Val Thr Ala Lys Val Ser Pro Gly Leu Cys Ile Ala Gly Asn Arg

1175 1180 1185

Gly Ile Ala Pro Lys Ser Gly Tyr Phe Val Asn Val Asn Asn Thr

1190 1195 1200

Trp Met Tyr Thr Gly Ser Gly Tyr Tyr Tyr Pro Glu Pro Ile Thr

1205 1210 1215

Glu Asn Asn Val Val Val Met Ser Thr Cys Ala Val Asn Tyr Thr

1220 1225 1230

Lys Ala Pro Tyr Val Met Leu Asn Thr Ser Ile Pro Asn Leu Pro

1235 1240 1245

Asp Phe Lys Glu Glu Leu Asp Gln Trp Phe Lys Asn Gln Thr Ser

1250 1255 1260

Val Ala Pro Asp Leu Ser Leu Asp Tyr Ile Asn Val Thr Phe Leu

1265 1270 1275

Asp Leu Gln Val Glu Met Asn Arg Leu Gln Glu Ala Ile Lys Val
1280 1285 1290
Leu Asn His Ser Tyr Ile Asn Leu Lys Asp Ile Gly Thr Tyr Glu
1295 1300 1305
Tyr Tyr Val Lys Trp Pro Trp Tyr Val Trp Leu Leu Ile Cys Leu

1310 1315 1320
Ala Gly Val Ala Met Leu Val Leu Leu Phe Phe Ile Cys Cys Ser
1325 1330 1335
Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1340 1345

<210> 146

<211> 1351

<212> PRT

<213> Artificial Sequence

<220><223> PDI-OC43-H1cCT-AA

<400> 146

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Val Ile Gly Asp Leu Asn Cys Thr
20 25 30
Leu Asp Pro Arg Leu Lys Gly Ser Phe Asn Asn Arg Asp Thr Gly Pro
35 40 45
Pro Ser Ile Ser Ile Asp Thr Val Asp Val Thr Asn Gly Leu Gly Thr
50 55 60
Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn Thr Thr Leu Phe Leu Asn
65 70 75 80

Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr Arg Asn Met Ala Leu Lys
85 90 95
Gly Thr Asp Leu Leu Ser Thr Leu Trp Phe Lys Pro Pro Phe Leu Ser
100 105 110
Asp Phe Ile Asn Gly Ile Phe Ala Lys Val Lys Asn Thr Lys Val Phe
115 120 125

Lys Asp Gly Val Met Tyr Ser Glu Phe Pro Ala Ile Thr Ile Gly Ser
130 135 140

Thr Phe Val Asn Thr Ser Tyr Ser Val Val Val Gln Pro Arg Thr Ile
145 150 155 160
Asn Ser Thr Gln Asp Gly Val Asn Lys Leu Gln Gly Leu Leu Glu Val
165 170 175
Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr Pro His Thr Ile Cys His
180 185 190
Pro Asn Leu Gly Asn His Phe Lys Glu Leu Trp His Tyr Asp Thr Gly
195 200 205

Val Val Ser Cys Leu Tyr Lys Arg Asn Phe Thr Tyr Asp Val Asn Ala
210 215 220
Thr Tyr Leu Tyr Phe His Phe Tyr Gln Glu Gly Gly Thr Phe Tyr Ala
225 230 235 240
Tyr Phe Thr Asp Thr Gly Phe Val Thr Lys Phe Leu Phe Asn Val Tyr
245 250 255
Leu Gly Met Ala Leu Ser His Tyr Tyr Val Met Pro Leu Thr Cys Ile
260 265 270

Arg Arg Pro Lys Asp Gly Phe Ser Leu Glu Tyr Trp Val Thr Pro Leu
275 280 285
Thr Pro Arg Gln Tyr Leu Leu Ala Phe Asn Gln Asp Gly Ile Ile Phe
290 295 300
Asn Ala Val Asp Cys Met Ser Asp Phe Met Ser Glu Ile Lys Cys Lys
305 310 315 320
Thr Gln Ser Ile Ala Pro Pro Thr Gly Val Tyr Glu Leu Asn Gly Tyr
325 330 335

Thr Val Gln Pro Val Ala Asp Val Tyr Arg Arg Lys Pro Asp Leu Pro
340 345 350
Asn Cys Asn Ile Glu Ala Trp Leu Asn Asp Lys Ser Val Pro Ser Pro
355 360 365
Leu Asn Trp Glu Arg Lys Thr Phe Ser Asn Cys Asn Phe Asn Met Ser

370 375 380

Ser Leu Met Ser Phe Ile Gln Ala Asp Ser Phe Thr Cys Asn Asn Ile

385 390 395 400

Asp Ala Ala Lys Ile Tyr Gly Met Cys Phe Ser Ser Ile Thr Ile Asp

405 410 415

Lys Phe Ala Ile Pro Asn Arg Arg Lys Val Asp Leu Gln Leu Gly Asn

420 425 430

Leu Gly Tyr Leu Gln Ser Ser Asn Tyr Arg Ile Asp Thr Thr Ala Thr

435 440 445

Ser Cys Gln Leu Tyr Tyr Asn Leu Pro Ala Ala Asn Val Ser Val Ser

450 455 460

Arg Phe Asn Pro Ser Thr Trp Asn Lys Arg Phe Gly Phe Ile Glu Asp

465 470 475 480

Ser Val Phe Val Pro Gln Pro Thr Gly Val Phe Thr Asn His Ser Val

485 490 495

Val Tyr Ala Gln His Cys Phe Lys Ala Pro Lys Asn Phe Cys Pro Cys

500 505 510

Ser Ser Cys Ser Cys Pro Gly Lys Asn Asn Gly Ile Gly Thr Cys Pro

515 520 525

Ala Gly Thr Asn Ser Leu Thr Cys Asp Asn Leu Cys Thr Leu Asp Pro

530 535 540

Ile Thr Leu Lys Ala Pro Asp Thr Tyr Lys Cys Pro Gln Ser Lys Ser

545 550 555 560

Leu Val Gly Ile Gly Glu His Cys Ser Gly Leu Ala Val Lys Ser Asp

565 570 575

Tyr Cys Gly Asn Asn Ser Cys Thr Cys Gln Pro Gln Ala Phe Leu Gly

580 585 590

Trp Ser Ala Asp Ser Cys Leu Gln Gly Asp Lys Cys Asn Ile Phe Ala

595 600 605

Asn Phe Ile Leu His Asp Val Asn Asn Gly Leu Thr Cys Ser Thr Asp

610 615 620

Leu Gln Lys Ala Asn Thr Glu Ile Glu Leu Gly Val Cys Val Asn Tyr
625 630 635 640
Asp Leu Tyr Gly Ile Ser Gly Gln Gly Ile Phe Val Glu Val Asn Ala
645 650 655

Thr Tyr Tyr Asn Ser Trp Gln Asn Leu Leu Tyr Asp Ser Asn Gly Asn
660 665 670
Leu Tyr Gly Phe Arg Asp Tyr Ile Thr Asn Arg Thr Phe Met Ile His
675 680 685
Ser Cys Tyr Ser Gly Arg Val Ser Ala Ala Tyr His Ala Asn Ser Ser
690 695 700
Glu Pro Ala Leu Leu Phe Arg Asn Ile Lys Cys Asn Tyr Val Phe Asn
705 710 715 720

Asn Ser Leu Thr Arg Gln Leu Gln Pro Ile Asn Tyr Ser Phe Asp Ser
725 730 735
Tyr Leu Gly Cys Val Val Asn Ala Tyr Asn Ser Thr Ala Ile Ser Val
740 745 750
Gln Thr Cys Asp Leu Thr Val Gly Ser Gly Tyr Cys Val Asp Tyr Ser
755 760 765
Lys Asn Gly Gly Ser Gly Ser Ala Ile Thr Thr Gly Tyr Arg Phe Thr
770 775 780

Asn Phe Glu Pro Phe Thr Val Asn Ser Val Asn Asp Ser Leu Glu Pro
785 790 795 800
Val Gly Gly Leu Tyr Glu Ile Gln Ile Pro Ser Glu Phe Thr Ile Gly
805 810 815
Asn Met Glu Glu Phe Ile Gln Thr Ser Ser Pro Lys Val Thr Ile Asp
820 825 830
Cys Ala Ala Phe Val Cys Gly Asp Tyr Ala Ala Cys Lys Leu Gln Leu
835 840 845

Val Glu Tyr Gly Ser Phe Cys Asp Asn Ile Asn Ala Ile Leu Thr Glu
850 855 860
Val Asn Glu Leu Leu Asp Thr Thr Gln Leu Gln Val Ala Asn Ser Leu

865 870 875 880

Met Asn Gly Val Thr Leu Ser Thr Lys Leu Lys Asp Gly Val Asn Phe

885 890 895

Asn Val Asp Asp Ile Asn Phe Ser Pro Val Leu Gly Cys Leu Gly Ser

900 905 910

Glu Cys Ser Lys Ala Ser Ser Arg Ser Ala Ile Glu Asp Leu Leu Phe

915 920 925

Asp Lys Val Lys Leu Ser Asp Val Gly Phe Val Glu Ala Tyr Asn Asn

930 935 940

Cys Thr Gly Gly Ala Glu Ile Arg Asp Leu Ile Cys Val Gln Ser Tyr

945 950 955 960

Lys Gly Ile Lys Val Leu Pro Pro Leu Leu Ser Glu Asn Gln Ile Ser

965 970 975

Gly Tyr Thr Leu Ala Ala Thr Ser Ala Ser Leu Phe Pro Pro Trp Thr

980 985 990

Ala Ala Ala Gly Val Pro Phe Tyr Leu Asn Val Gln Tyr Arg Ile Asn

995 1000 1005

Gly Leu Gly Val Thr Met Asp Val Leu Ser Gln Asn Gln Lys Leu

1010 1015 1020

Ile Ala Asn Ala Phe Asn Asn Ala Leu His Ala Ile Gln Gln Gly

1025 1030 1035

Phe Asp Ala Thr Asn Ser Ala Leu Val Lys Ile Gln Ala Val Val

1040 1045 1050

Asn Ala Asn Ala Glu Ala Leu Asn Asn Leu Leu Gln Gln Leu Ser

1055 1060 1065

Asn Arg Phe Gly Ala Ile Ser Ala Ser Leu Gln Glu Ile Leu Ser

1070 1075 1080

Arg Leu Asp Pro Pro Glu Ala Glu Ala Gln Ile Asp Arg Leu Ile

1085 1090 1095

Asn Gly Arg Leu Thr Ala Leu Asn Ala Tyr Val Ser Gln Gln Leu

1100 1105 1110

```
Ser Asp  Ser Thr Leu Val Lys  Phe Ser Ala Ala Gln  Ala Met Glu
    1115                 1120                 1125
Lys Val  Asn Glu Cys Val Lys  Ser Gln Ser Ser Arg  Ile Asn Phe
    1130                 1135                 1140
Cys Gly  Asn Gly Asn His Ile  Ile Ser Leu Val Gln  Asn Ala Pro
    1145                 1150                 1155

Tyr Gly  Leu Tyr Phe Ile His  Phe Asn Tyr Val Pro  Thr Lys Tyr
    1160                 1165                 1170
Val Thr  Ala Lys Val Ser Pro  Gly Leu Cys Ile Ala  Gly Asn Arg
    1175                 1180                 1185
Gly Ile  Ala Pro Lys Ser Gly  Tyr Phe Val Asn Val  Asn Asn Thr
    1190                 1195                 1200
Trp Met  Tyr Thr Gly Ser Gly  Tyr Tyr Tyr Pro Glu  Pro Ile Thr
    1205                 1210                 1215

Glu Asn  Asn Val Val Val Met  Ser Thr Cys Ala Val  Asn Tyr Thr
    1220                 1225                 1230
Lys Ala  Pro Tyr Val Met Leu  Asn Thr Ser Ile Pro  Asn Leu Pro
    1235                 1240                 1245
Asp Phe  Lys Glu Glu Leu Asp  Gln Trp Phe Lys Asn  Gln Thr Ser
    1250                 1255                 1260
Val Ala  Pro Asp Leu Ser Leu  Asp Tyr Ile Asn Val  Thr Phe Leu
    1265                 1270                 1275

Asp Leu  Gln Val Glu Met Asn  Arg Leu Gln Glu Ala  Ile Lys Val
    1280                 1285                 1290
Leu Asn  His Ser Tyr Ile Asn  Leu Lys Asp Ile Gly  Thr Tyr Glu
    1295                 1300                 1305
Tyr Tyr  Val Lys Trp Pro Trp  Tyr Val Trp Leu Leu  Ile Cys Leu
    1310                 1315                 1320
Ala Gly  Val Ala Met Leu Val  Leu Leu Phe Phe Ile  Ser Phe Trp
    1325                 1330                 1335

Met Cys  Ser Asn Gly Ser Leu  Gln Cys Arg Ile Cys  Ile
```

1340 1345 1350

<210> 147

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> IF(CoV229EwtCT).r

<400> 147

acgacacgac taaggccttc actgtatgtg gatcttttcg acatcgta 48

<210> 148

<211> 3546

<212> DNA

<213> Artificial Sequence

<220><223> PDI-229E -wtTMCT-DNA

<400> 148

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgcaaacgac taatgggctg aacaccagtt acagcgtctg taacggctgc 120

gtcggatata gcgagaacgt gttcgcagtg gaaagtgggg ggtacattcc ttccgacttc 180

gctttcaata actggtttct cctgactaac acaagctccg tcgtggatgg cgtggtcagg 240

tcctttcagc ctcttctcct gaattgcctg tggtctgtgt ccgggttaag attcactaca 300

ggcttcgtat acttcaacgg gacgggccgg ggggattgca agggcttctc ctccgacgtg 360

ctgtcagatg tgatccgtta caatctgaac ttcgaagaga acttacggcg ggggacaatc 420

ctgttcaaaa catcatatgg cgtagtcgta ttttactgca ccaataatac cctggtgagt 480

ggggacgccc atattccctt cggaacagtg ctgggtaact tttactgttt tgtcaacact 540

acgatcggaa acgaaaccac tagcgccttt gtcggagctc tgccaaaaac agttagggag 600

ttcgtgatct ctcggaccgg tcacttctat atcaacggct accgttattt tactttgggc 660

aacgtcgaag ccgtcaattt taatgtgaca actgcagaga caactgactt ttgcactgtg 720

gctctcgcca gttatgccga tgtgctggtg aatgtaagtc aaacgtcaat tgccaacatc 780

atctattgta actcagtaat caaccggctc cgctgtgacc aactctcatt cgacgtcccc 840

gacggattct attccacgag cccgattcag agcgtggaac tgccagtttc catcgtatcc 900

ctcccagttt accacaagca cacttttatc gttctctacg tagattttaa accccagtca 960

ggaggaggga aatgcttcaa ctgctacccg gctggcgtga acatcacctt ggccaatttt 1020

aatgaaacta aagggcccct ttgcgtggat acgtcacact ttaccacaaa gtatgttgca 1080

gtctatgcta acgtcggcag gtggtcagcg tccattaaca caggcaattg cccgttctct 1140

ttcgggaaag tgaacaactt cgtgaagttt ggaagtgtgt gcttcagttt gaaagacatt 1200

ccgggcggct gcgccatgcc tattgtggct aattgggctt attccaagta ctacaccatt 1260

ggctctctct acgttagctg gagcgacggt gacggtataa cgggcgtacc acaaccggtg 1320

gaaggggtca gctctttcat gaatgtcact ctggacaagt gtaccaaata taatatatac 1380

gatgtgagtg gagtgggcgt tatacgcgtg tctaacgaca cctttctaaa cggcataacc 1440

tacacaagca cgtcaggcaa tctgttaggt tttaaagacg tcactaaagg cactatatat 1500

agcatcaccc catgcaaccc acctgatcaa ttagtcgtat atcagcaagc tgttgtgggt 1560

gctatgctgt cagaaaactt caccagctac gggttctcca atgtggtgga actgcccaaa 1620

ttcttttacg ctagcaatgg cacatataac tgtactgacg ccgtcttgac ttacagttca 1680

ttcggagtgt gcgcggacgg cagcattatc gccgtgcagc cggccaatgt cagctatgat 1740

tccgtttccg ccatcgtgac agccaacttg tcgattccct ctaactggac aacgtctgtc 1800

caagtcgaat atctgcagat cacctcaacc cccatagtag tcgattgctc aacctacgtc 1860

tgcaacggta atgtcagatg tgtcgagctg ctcaagcagt acacctccgc ctgtaagact 1920

attgaggatg cattaagaaa tagtgcaaga ttggaaagcg ccgatgtgtc ggaaatgcta 1980

accttcgata agaaggcatt cacactggcg aacgtaagct ctttcggcga ttacaacctg 2040

tcttcggtaa tccctagctt gcccacatcc ggctctcggg tggcggggcg gagcgctatc 2100

gaggacattt tattctcgaa actggttaca tctgggctcg gaactgtgga cgccgattac 2160

aagaagtgca ccaagggcct aagcatcgcc gacctcgcct gtgctcagta ctacaacgga 2220

attatggtgc tgccaggtgt cgctgacgca gagcggatgg ctatgtatac cggcagtctc 2280

attggcggga ttgcgttggg cggcctgacg tccgctgtct ccatcccttt ctctctggct 2340

atacaagccc gactgaatta tgtggccctg cagactgatg tcctgcaaga aaatcagaag 2400

attcttgccg ccagcttcaa caaggccatg actaatattg tggatgcgtt taccggagtg 2460

aatgacgcca tcacccaaac gtcccaagcc ctgcagacag tcgccacggc gttaaacaaa 2520

atccaggatg tagtgaatca gcaagggaac agcttgaatc acctgacgtc ccagttaaga 2580

cagaactttc aggcaatcag tagctcaatc caggctatct acgatcgatt agatcctcct 2640

caggcagatc agcaggtgga tcggctcatc accggccgcc tcgcggcatt gaatgttttc 2700

gtaagtcata ccttgaccaa gtacacggag gtgagggcca gtcgccagct ggctcagcaa 2760

aaagtgaatg agtgtgtgaa atcacagagc aaacggtacg ggttttgtgg aaatgggacg 2820

cacatcttta gcatcgttaa tgctgccccc gaagggttag tcttcctgca cactgtgctc 2880

cttcctaccc agtataaaga tgtcgaagca tggtctgggc tctgtgtcga tggaactaac 2940

ggttatgtcc ttcgacagcc aaacctcgct ctctataaag aagggaatta ctataggatc 3000

acctcaagaa tcatgttcga gcccaggata ccaacaatgg ccgattttgt gcagattgaa 3060

aattgtaacg tgacctttgt gaatatcagt cgatccgagc ttcaaacgat tgttcctgag 3120

tacatcgacg tgaataaaac tctacaagag ctgtcctata aactgcctaa ttataccgtg 3180

cctgaccttg tagtcgagca atacaaccag actattctga acctgacatc ggaaatctct 3240

acattggaga ataaaagcgc cgagctcaat tacacagtgc agaagctgca gaccctgatc 3300

gacaatatta acagcactct tgtggactta aagtggctga accgtgtgga gacttacatc 3360

aagtggccct ggtgggtgtg gctctgtatt tccgtggtcc ttatatttgt tgtaagtatg 3420

ctgctcctgt gctgttgctc aaccgggtgc tgcggttttt tctcctgttt cgcctcatcc 3480

atccgtggct gttgtgagag cactaaactg ccatattacg atgtcgaaaa gatccacata 3540

cagtga 3546

<210> 149

<211> 3492

<212> DNA

<213> Artificial Sequence

<220><223> PDI-229E -H5iTMCT-DNA

<400> 149

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgcaaacgac taatgggctg aacaccagtt acagcgtctg taacggctgc 120

gtcggatata gcgagaacgt gttcgcagtg gaaagtgggg ggtacattcc ttccgacttc 180

gctttcaata actggtttct cctgactaac acaagctccg tcgtggatgg cgtggtcagg 240

tcctttcagc ctcttctcct gaattgcctg tggtctgtgt ccgggttaag attcactaca 300

ggcttcgtat acttcaacgg gacgggccgg ggggattgca agggcttctc ctccgacgtg 360

ctgtcagatg tgatccgtta caatctgaac ttcgaagaga acttacggcg ggggacaatc 420

ctgttcaaaa catcatatgg cgtagtcgta ttttactgca ccaataatac cctggtgagt 480

ggggacgccc atattccctt cggaacagtg ctgggtaact tttactgttt tgtcaacact 540

acgatcggaa acgaaaccac tagcgccttt gtcggagctc tgccaaaaac agttagggag 600

ttcgtgatct ctcggaccgg tcacttctat atcaacggct accgttattt tactttgggc 660

aacgtcgaag ccgtcaattt taatgtgaca actgcagaga caactgactt ttgcactgtg 720

gctctcgcca gttatgccga tgtgctggtg aatgtaagtc aaacgtcaat tgccaacatc 780

atctattgta actcagtaat caaccggctc cgctgtgacc aactctcatt cgacgtcccc 840

gacggattct attccacgag cccgattcag agcgtggaac tgccagtttc catcgtatcc 900

ctcccagttt accacaagca cacttttatc gttctctacg tagattttaa accccagtca 960

ggaggaggga aatgcttcaa ctgctacccg gctggcgtga acatcacctt ggccaatttt 1020

aatgaaacta aagggcccct ttgcgtggat acgtcacact ttaccacaaa gtatgttgca 1080

gtctatgcta acgtcggcag gtggtcagcg tccattaaca caggcaattg cccgttctct 1140

ttcgggaaag tgaacaactt cgtgaagttt ggaagtgtgt gcttcagttt gaaagacatt 1200

ccgggcggct gcgccatgcc tattgtggct aattgggctt attccaagta ctacaccatt 1260

ggctctctct acgttagctg gagcgacggt gacggtataa cgggcgtacc acaaccggtg 1320

gaaggggtca gctctttcat gaatgtcact ctggacaagt gtaccaaata taatatatac 1380

gatgtgagtg gagtgggcgt tatacgcgtg tctaacgaca cctttctaaa cggcataacc 1440

tacacaagca cgtcaggcaa tctgttaggt tttaaagacg tcactaaagg cactatatat 1500

agcatcaccc catgcaaccc acctgatcaa ttagtcgtat atcagcaagc tgttgtgggt 1560

gctatgctgt cagaaaactt caccagctac gggttctcca atgtggtgga actgcccaaa 1620

ttcttttacg ctagcaatgg cacatataac tgtactgacg ccgtcttgac ttacagttca 1680

ttcggagtgt gcgcggacgg cagcattatc gccgtgcagc cggccaatgt cagctatgat 1740

tccgtttccg ccatcgtgac agccaacttg tcgattccct ctaactggac aacgtctgtc 1800

caagtcgaat atctgcagat cacctcaacc cccatagtag tcgattgctc aacctacgtc 1860

tgcaacggta atgtcagatg tgtcgagctg ctcaagcagt acacctccgc ctgtaagact 1920

attgaggatg cattaagaaa tagtgcaaga ttggaaagcg ccgatgtgtc ggaaatgcta 1980

accttcgata agaaggcatt cacactggcg aacgtaagct ctttcggcga ttacaacctg 2040

tcttcggtaa tccctagctt gcccacatcc ggctctcggg tggcggggcg gagcgctatc 2100

gaggacattt tattctcgaa actggttaca tctgggctcg gaactgtgga cgccgattac 2160

aagaagtgca ccaagggcct aagcatcgcc gacctcgcct gtgctcagta ctacaacgga 2220

attatggtgc tgccaggtgt cgctgacgca gagcggatgg ctatgtatac cggcagtctc 2280

attggcggga ttgcgttggg cggcctgacg tccgctgtct ccatcccttt ctctctggct 2340

atacaagccc gactgaatta tgtggccctg cagactgatg tcctgcaaga aaatcagaag 2400

attcttgccg ccagcttcaa caaggccatg actaatattg tggatgcgtt taccggagtg 2460

aatgacgcca tcacccaaac gtcccaagcc ctgcagacag tcgccacggc gttaaacaaa 2520

atccaggatg tagtgaatca gcaagggaac agcttgaatc acctgacgtc ccagttaaga 2580

cagaactttc aggcaatcag tagctcaatc caggctatct acgatcgatt agatcctcct 2640

caggcagatc agcaggtgga tcggctcatc accggccgcc tcgcggcatt gaatgttttc 2700

gtaagtcata ccttgaccaa gtacacggag gtgagggcca gtcgccagct ggctcagcaa 2760

aaagtgaatg agtgtgtgaa atcacagagc aaacggtacg ggttttgtgg aaatgggacg 2820
cacatcttta gcatcgttaa tgctgccccc gaagggttag tcttcctgca cactgtgctc 2880
cttcctaccc agtataaaga tgtcgaagca tggtctgggc tctgtgtcga tggaactaac 2940
ggttatgtcc ttcgacagcc aaacctcgct ctctataaag aagggaatta ctataggatc 3000
acctcaagaa tcatgttcga gcccaggata ccaacaatgg ccgattttgt gcagattgaa 3060
aattgtaacg tgacctttgt gaatatcagt cgatccgagc ttcaaacgat tgttcctgag 3120
tacatcgacg tgaataaaac tctacaagag ctgtcctata aactgcctaa ttataccgtg 3180

cctgaccttg tagtcgagca atacaaccag actattctga acctgacatc ggaaatctct 3240
acattggaga ataaaagcgc cgagctcaat tacacagtgc agaagctgca gaccctgatc 3300
gacaatatta acagcactct tgtggactta aagtggctga accgtgtgga gacttacatc 3360
aagtggccct ggtggcaaat actgtcaatt tattcaacag tggcgagttc cctagcactg 3420
gcaatcatga tggctggtct atctttatgg atgtgctcca atggatcgtt acaatgcaga 3480
atttgcattt ga 3492

<210> 150
<211> 3480
<212> DNA
<213> Artificial Sequence

<220><223> PDI-229E -H5iCT-DNA
<400> 150

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cgcaaacgac taatgggctg aacaccagtt acagcgtctg taacggctgc 120
gtcggatata gcgagaacgt gttcgcagtg gaaagtgggg ggtacattcc ttccgacttc 180
gctttcaata actggtttct cctgactaac acaagctccg tcgtggatgg cgtggtcagg 240
tcctttcagc ctcttctcct gaattgcctg tggtctgtgt ccgggttaag attcactaca 300
ggcttcgtat acttcaacgg gacgggccgg ggggattgca agggcttctc ctccgacgtg 360
ctgtcagatg tgatccgtta caatctgaac ttcgaagaga acttacggcg ggggacaatc 420

ctgttcaaaa catcatatgg cgtagtcgta ttttactgca ccaataatac cctggtgagt 480
ggggacgccc atattccctt cggaacagtg ctgggtaact tttactgttt tgtcaacact 540
acgatcggaa acgaaaccac tagcgccttt gtcggagctc tgccaaaaac agttagggag 600
ttcgtgatct ctcggaccgg tcacttctat atcaacggct accgttattt tactttgggc 660

aacgtcgaag ccgtcaattt taatgtgaca actgcagaga caactgactt ttgcactgtg 720

gctctcgcca gttatgccga tgtgctggtg aatgtaagtc aaacgtcaat tgccaacatc 780

atctattgta actcagtaat caaccggctc cgctgtgacc aactctcatt cgacgtcccc 840

gacggattct attccacgag cccgattcag agcgtggaac tgccagtttc catcgtatcc 900

ctcccagttt accacaagca cacttttatc gttctctacg tagattttaa accccagtca 960

ggaggaggga aatgcttcaa ctgctacccg gctggcgtga acatcacctt ggccaatttt 1020

aatgaaacta aagggcccct ttgcgtggat acgtcacact ttaccacaaa gtatgttgca 1080

gtctatgcta acgtcggcag gtggtcagcg tccattaaca caggcaattg cccgttctct 1140

ttcgggaaag tgaacaactt cgtgaagttt ggaagtgtgt gcttcagttt gaaagacatt 1200

ccgggcggct gcgccatgcc tattgtggct aattgggctt attccaagta ctacaccatt 1260

ggctctctct acgttagctg gagcgacggt gacggtataa cgggcgtacc acaaccggtg 1320

gaaggggtca gctctttcat gaatgtcact ctggacaagt gtaccaaata taatatatac 1380

gatgtgagtg gagtgggcgt tatacgcgtg tctaacgaca cctttctaaa cggcataacc 1440

tacacaagca cgtcaggcaa tctgttaggt tttaaagacg tcactaaagg cactatatat 1500

agcatcaccc catgcaaccc acctgatcaa ttagtcgtat atcagcaagc tgttgtgggt 1560

gctatgctgt cagaaaactt caccagctac gggttctcca atgtggtgga actgcccaaa 1620

ttcttttacg ctagcaatgg cacatataac tgtactgacg ccgtcttgac ttacagttca 1680

ttcggagtgt gcgcggacgg cagcattatc gccgtgcagc cggccaatgt cagctatgat 1740

tccgtttccg ccatcgtgac agccaacttg tcgattccct ctaactggac aacgtctgtc 1800

caagtcgaat atctgcagat cacctcaacc cccatagtag tcgattgctc aacctacgtc 1860

tgcaacggta atgtcagatg tgtcgagctg ctcaagcagt acacctccgc ctgtaagact 1920

attgaggatg cattaagaaa tagtgcaaga ttggaaagcg ccgatgtgtc ggaaatgcta 1980

accttcgata agaaggcatt cacactggcg aacgtaagct ctttcggcga ttacaacctg 2040

tcttcggtaa tccctagctt gcccacatcc ggctctcggg tggcggggcg gagcgctatc 2100

gaggacattt tattctcgaa actggttaca tctgggctcg gaactgtgga cgccgattac 2160

aagaagtgca ccaagggcct aagcatcgcc gacctcgcct gtgctcagta ctacaacgga 2220

attatggtgc tgccaggtgt cgctgacgca gagcggatgg ctatgtatac cggcagtctc 2280

attggcggga ttgcgttggg cggcctgacg tccgctgtct ccatcccttt ctctctggct 2340

atacaagccc gactgaatta tgtggccctg cagactgatg tcctgcaaga aaatcagaag 2400

attcttgccg ccagcttcaa caaggccatg actaatattg tggatgcgtt taccggagtg 2460

aatgacgcca tcacccaaac gtcccaagcc ctgcagacag tcgccacggc gttaaacaaa 2520

atccaggatg tagtgaatca gcaagggaac agcttgaatc acctgacgtc ccagttaaga 2580
cagaactttc aggcaatcag tagctcaatc caggctatct acgatcgatt agatcctcct 2640
caggcagatc agcaggtgga tcggctcatc accggccgcc tcgcggcatt gaatgttttc 2700
gtaagtcata ccttgaccaa gtacacggag gtgagggcca gtcgccagct ggctcagcaa 2760
aaagtgaatg agtgtgtgaa atcacagagc aaacggtacg ggttttgtgg aaatgggacg 2820
cacatcttta gcatcgttaa tgctgccccc gaagggttag tcttcctgca cactgtgctc 2880
cttcctaccc agtataaaga tgtcgaagca tggtctgggc tctgtgtcga tggaactaac 2940

ggttatgtcc ttcgacagcc aaacctcgct ctctataaag aagggaatta ctataggatc 3000
acctcaagaa tcatgttcga gcccaggata ccaacaatgg ccgattttgt gcagattgaa 3060
aattgtaacg tgacctttgt gaatatcagt cgatccgagc ttcaaacgat tgttcctgag 3120
tacatcgacg tgaataaaac tctacaagag ctgtcctata aactgcctaa ttataccgtg 3180
cctgaccttg tagtcgagca atacaaccag actattctga acctgacatc ggaaatctct 3240
acattggaga ataaaagcgc cgagctcaat tacacagtgc agaagctgca gaccctgatc 3300
gacaatatta acagcactct tgtggactta aagtggctga accgtgtgga gacttacatc 3360

aagtggccct ggtgggtgtg gctctgtatt tccgtggtcc ttatatttgt tgtaagtatg 3420
ctgctcctgt ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttga 3480

<210> 151

<211> 3471

<212> DNA

<213> Artificial Sequence

<220><223> PDI-229E -H5iCT(V4)-DNA

<400> 151

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60
cagatcttcg cgcaaacgac taatgggctg aacaccagtt acagcgtctg taacggctgc 120
gtcggatata gcgagaacgt gttcgcagtg gaaagtgggg ggtacattcc ttccgacttc 180
gctttcaata actggtttct cctgactaac acaagctccg tcgtggatgg cgtggtcagg 240

tcctttcagc ctcttctcct gaattgcctg tggtctgtgt ccgggttaag attcactaca 300
ggcttcgtat acttcaacgg gacgggccgg ggggattgca agggcttctc ctccgacgtg 360
ctgtcagatg tgatccgtta caatctgaac ttcgaagaga acttacggcg ggggacaatc 420
ctgttcaaaa catcatatgg cgtagtcgta ttttactgca ccaataatac cctggtgagt 480

ggggacgccc atattccctt cggaacagtg ctgggtaact tttactgttt tgtcaacact 540

acgatcggaa acgaaaccac tagcgccttt gtcggagctc tgccaaaaac agttagggag 600

ttcgtgatct ctcggaccgg tcacttctat atcaacggct accgttattt tactttgggc 660

aacgtcgaag ccgtcaattt taatgtgaca actgcagaga caactgactt ttgcactgtg 720

gctctcgcca gttatgccga tgtgctggtg aatgtaagtc aaacgtcaat tgccaacatc 780

atctattgta actcagtaat caaccggctc cgctgtgacc aactctcatt cgacgtcccc 840

gacggattct attccacgag cccgattcag agcgtggaac tgccagtttc catcgtatcc 900

ctcccagttt accacaagca cacttttatc gttctctacg tagattttaa accccagtca 960

ggaggaggga aatgcttcaa ctgctacccg gctggcgtga acatcacctt ggccaatttt 1020

aatgaaacta aagggcccct ttgcgtggat acgtcacact ttaccacaaa gtatgttgca 1080

gtctatgcta acgtcggcag gtggtcagcg tccattaaca caggcaattg cccgttctct 1140

ttcgggaaag tgaacaactt cgtgaagttt ggaagtgtgt gcttcagttt gaaagacatt 1200

ccgggcggct gcgccatgcc tattgtggct aattgggctt attccaagta ctacaccatt 1260

ggctctctct acgttagctg gagcgacggt gacggtataa cgggcgtacc acaaccggtg 1320

gaaggggtca gctctttcat gaatgtcact ctggacaagt gtaccaaata taatatatac 1380

gatgtgagtg gagtgggcgt tatacgcgtg tctaacgaca cctttctaaa cggcataacc 1440

tacacaagca cgtcaggcaa tctgttaggt tttaaagacg tcactaaagg cactatatat 1500

agcatcaccc catgcaaccc acctgatcaa ttagtcgtat atcagcaagc tgttgtgggt 1560

gctatgctgt cagaaaactt caccagctac gggttctcca atgtggtgga actgcccaaa 1620

ttcttttacg ctagcaatgg cacatataac tgtactgacg ccgtcttgac ttacagttca 1680

ttcggagtgt gcgcggacgg cagcattatc gccgtgcagc cggccaatgt cagctatgat 1740

tccgtttccg ccatcgtgac agccaacttg tcgattccct ctaactggac aacgtctgtc 1800

caagtcgaat atctgcagat cacctcaacc cccatagtag tcgattgctc aacctacgtc 1860

tgcaacggta atgtcagatg tgtcgagctg ctcaagcagt acacctccgc ctgtaagact 1920

attgaggatg cattaagaaa tagtgcaaga ttggaaagcg ccgatgtgtc ggaaatgcta 1980

accttcgata agaaggcatt cacactggcg aacgtaagct ctttcggcga ttacaacctg 2040

tcttcggtaa tccctagctt gcccacatcc ggctctcggg tggcggggcg gagcgctatc 2100

gaggacattt tattctcgaa actggttaca tctgggctcg gaactgtgga cgccgattac 2160

aagaagtgca ccaagggcct aagcatcgcc gacctcgcct gtgctcagta ctacaacgga 2220

attatggtgc tgccaggtgt cgctgacgca gagcggatgg ctatgtatac cggcagtctc 2280

attggcggga ttgcgttggg cggcctgacg tccgctgtct ccatcccttt ctctctggct 2340

atacaagccc gactgaatta tgtggccctg cagactgatg tcctgcaaga aaatcagaag 2400
attcttgccg ccagcttcaa caaggccatg actaatattg tggatgcgtt taccggagtg 2460
aatgacgcca tcacccaaac gtcccaagcc ctgcagacag tcgccacggc gttaaacaaa 2520
atccaggatg tagtgaatca gcaagggaac agcttgaatc acctgacgtc ccagttaaga 2580
cagaactttc aggcaatcag tagctcaatc caggctatct acgatcgatt agatcctcct 2640
caggcagatc agcaggtgga tcggctcatc accggccgcc tcgcggcatt gaatgttttc 2700
gtaagtcata ccttgaccaa gtacacggag gtgagggcca gtcgccagct ggctcagcaa 2760

aaagtgaatg agtgtgtgaa atcacagagc aaacggtacg ggttttgtgg aaatgggacg 2820
cacatcttta gcatcgttaa tgctgccccc gaagggttag tcttcctgca cactgtgctc 2880
cttcctaccc agtataaaga tgtcgaagca tggtctgggc tctgtgtcga tggaactaac 2940
ggttatgtcc ttcgacagcc aaacctcgct ctctataaag aagggaatta ctataggatc 3000
acctcaagaa tcatgttcga gcccaggata ccaacaatgg ccgattttgt gcagattgaa 3060
aattgtaacg tgacctttgt gaatatcagt cgatccgagc ttcaaacgat tgttcctgag 3120
tacatcgacg tgaataaaac tctacaagag ctgtcctata aactgcctaa ttataccgtg 3180

cctgaccttg tagtcgagca atacaaccag actattctga acctgacatc ggaaatctct 3240
acattggaga ataaaagcgc cgagctcaat tacacagtgc agaagctgca gaccctgatc 3300
gacaatatta acagcactct tgtggactta aagtggctga accgtgtgga gacttacatc 3360
aagtggccct ggtgggtgtg gctctgtatt tccgtggtcc ttatatttgt tgtaagtatg 3420
ctgctcctgt gctgctccaa tggatcgtta caatgcagaa tttgcatttg a 3471

<210> 152
<211> 3480
<212> DNA
<213> Artificial Sequence
<220><223> PDI-229E -H1cCT-DNA
<400> 152

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgcaaacgac taatgggctg aacaccagtt acagcgtctg taacggctgc 120
gtcggatata gcgagaacgt gttcgcagtg gaaagtgggg ggtacattcc ttccgacttc 180
gctttcaata actggtttct cctgactaac acaagctccg tcgtggatgg cgtggtcagg 240
tcctttcagc ctcttctcct gaattgcctg tggtctgtgt ccgggttaag attcactaca 300

ggcttcgtat acttcaacgg gacgggccgg ggggattgca agggcttctc ctccgacgtg 360

ctgtcagatg tgatccgtta caatctgaac ttcgaagaga acttacggcg ggggacaatc 420

ctgttcaaaa catcatatgg cgtagtcgta ttttactgca ccaataatac cctggtgagt 480

ggggacgccc atattccctt cggaacagtg ctgggtaact tttactgttt tgtcaacact 540

acgatcggaa acgaaaccac tagcgccttt gtcggagctc tgccaaaaac agttagggag 600

ttcgtgatct ctcggaccgg tcacttctat atcaacggct accgttattt tactttgggc 660

aacgtcgaag ccgtcaattt taatgtgaca actgcagaga caactgactt ttgcactgtg 720

gctctcgcca gttatgccga tgtgctggtg aatgtaagtc aaacgtcaat tgccaacatc 780

atctattgta actcagtaat caaccggctc cgctgtgacc aactctcatt cgacgtcccc 840

gacggattct attccacgag cccgattcag agcgtggaac tgccagtttc catcgtatcc 900

ctcccagttt accacaagca cacttttatc gttctctacg tagattttaa accccagtca 960

ggaggaggga aatgcttcaa ctgctacccg gctggcgtga acatcacctt ggccaatttt 1020

aatgaaacta aagggcccct ttgcgtggat acgtcacact ttaccacaaa gtatgttgca 1080

gtctatgcta acgtcggcag gtggtcagcg tccattaaca caggcaattg cccgttctct 1140

ttcgggaaag tgaacaactt cgtgaagttt ggaagtgtgt gcttcagttt gaaagacatt 1200

ccgggcggct gcgccatgcc tattgtggct aattgggctt attccaagta ctacaccatt 1260

ggctctctct acgttagctg gagcgacggt gacggtataa cgggcgtacc acaaccggtg 1320

gaaggggtca gctctttcat gaatgtcact ctggacaagt gtaccaaata taatatatac 1380

gatgtgagtg gagtgggcgt tatacgcgtg tctaacgaca cctttctaaa cggcataacc 1440

tacacaagca cgtcaggcaa tctgttaggt tttaaagacg tcactaaagg cactatatat 1500

agcatcaccc catgcaaccc acctgatcaa ttagtcgtat atcagcaagc tgttgtgggt 1560

gctatgctgt cagaaaactt caccagctac gggttctcca atgtggtgga actgcccaaa 1620

ttcttttacg ctagcaatgg cacatataac tgtactgacg ccgtcttgac ttacagttca 1680

ttcggagtgt gcgcggacgg cagcattatc gccgtgcagc cggccaatgt cagctatgat 1740

tccgtttccg ccatcgtgac agccaacttg tcgattccct ctaactggac aacgtctgtc 1800

caagtcgaat atctgcagat cacctcaacc cccatagtag tcgattgctc aacctacgtc 1860

tgcaacggta atgtcagatg tgtcgagctg ctcaagcagt acacctccgc ctgtaagact 1920

attgaggatg cattaagaaa tagtgcaaga ttggaaagcg ccgatgtgtc ggaaatgcta 1980

accttcgata agaaggcatt cacactggcg aacgtaagct ctttcggcga ttacaacctg 2040

tcttcggtaa tccctagctt gcccacatcc ggctctcggg tggcggggcg gagcgctatc 2100

gaggacattt tattctcgaa actggttaca tctgggctcg gaactgtgga cgccgattac 2160

aagaagtgca ccaagggcct aagcatcgcc gacctcgcct gtgctcagta ctacaacgga 2220
attatggtgc tgccaggtgt cgctgacgca gagcggatgg ctatgtatac cggcagtctc 2280
attggcggga ttgcgttggg cggcctgacg tccgctgtct ccatcccttt ctctctggct 2340
atacaagccc gactgaatta tgtggccctg cagactgatg tcctgcaaga aaatcagaag 2400
attcttgccg ccagcttcaa caaggccatg actaatattg tggatgcgtt taccggagtg 2460
aatgacgcca tcacccaaac gtcccaagcc ctgcagacag tcgccacggc gttaaacaaa 2520
atccaggatg tagtgaatca gcaagggaac agcttgaatc acctgacgtc ccagttaaga 2580

cagaactttc aggcaatcag tagctcaatc caggctatct acgatcgatt agatcctcct 2640
caggcagatc agcaggtgga tcggctcatc accggccgcc tcgcggcatt gaatgttttc 2700
gtaagtcata ccttgaccaa gtacacggag gtgagggcca gtcgccagct ggctcagcaa 2760
aaagtgaatg agtgtgtgaa atcacagagc aaacggtacg ggttttgtgg aaatgggacg 2820
cacatcttta gcatcgttaa tgctgccccc gaagggttag tcttcctgca cactgtgctc 2880
cttcctaccc agtataaaga tgtcgaagca tggtctgggc tctgtgtcga tggaactaac 2940
ggttatgtcc ttcgacagcc aaacctcgct ctctataaag aagggaatta ctataggatc 3000

acctcaagaa tcatgttcga gcccaggata ccaacaatgg ccgattttgt gcagattgaa 3060
aattgtaacg tgacctttgt gaatatcagt cgatccgagc ttcaaacgat tgttcctgag 3120
tacatcgacg tgaataaaac tctacaagag ctgtcctata aactgcctaa ttataccgtg 3180
cctgaccttg tagtcgagca atacaaccag actattctga acctgacatc ggaaatctct 3240
acattggaga ataaaagcgc cgagctcaat tacacagtgc agaagctgca gaccctgatc 3300
gacaatatta acagcactct tgtggactta aagtggctga accgtgtgga gacttacatc 3360
aagtggccct ggtgggtgtg gctctgtatt tccgtggtcc ttatatttgt tgtaagtatg 3420

ctgctcctga gcttctggat gtgctctaat gggtctctac agtgtagaat atgtatttga 3480

<210> 153

<211> 1181

<212> PRT

<213> Artificial Sequence

<220><223> PDI-229E -wtTMCT-AA

<400> 153

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Thr Thr Asn Gly Leu Asn Thr
20 25 30
Ser Tyr Ser Val Cys Asn Gly Cys Val Gly Tyr Ser Glu Asn Val Phe

35 40 45
Ala Val Glu Ser Gly Gly Tyr Ile Pro Ser Asp Phe Ala Phe Asn Asn
50 55 60
Trp Phe Leu Leu Thr Asn Thr Ser Ser Val Val Asp Gly Val Val Arg
65 70 75 80
Ser Phe Gln Pro Leu Leu Leu Asn Cys Leu Trp Ser Val Ser Gly Leu
85 90 95
Arg Phe Thr Thr Gly Phe Val Tyr Phe Asn Gly Thr Gly Arg Gly Asp

100 105 110
Cys Lys Gly Phe Ser Ser Asp Val Leu Ser Asp Val Ile Arg Tyr Asn
115 120 125
Leu Asn Phe Glu Glu Asn Leu Arg Arg Gly Thr Ile Leu Phe Lys Thr
130 135 140
Ser Tyr Gly Val Val Val Phe Tyr Cys Thr Asn Asn Thr Leu Val Ser
145 150 155 160
Gly Asp Ala His Ile Pro Phe Gly Thr Val Leu Gly Asn Phe Tyr Cys

165 170 175
Phe Val Asn Thr Thr Ile Gly Asn Glu Thr Thr Ser Ala Phe Val Gly
180 185 190
Ala Leu Pro Lys Thr Val Arg Glu Phe Val Ile Ser Arg Thr Gly His
195 200 205
Phe Tyr Ile Asn Gly Tyr Arg Tyr Phe Thr Leu Gly Asn Val Glu Ala
210 215 220
Val Asn Phe Asn Val Thr Thr Ala Glu Thr Thr Asp Phe Cys Thr Val

225 230 235 240
Ala Leu Ala Ser Tyr Ala Asp Val Leu Val Asn Val Ser Gln Thr Ser
245 250 255
Ile Ala Asn Ile Ile Tyr Cys Asn Ser Val Ile Asn Arg Leu Arg Cys

260 265 270

Asp Gln Leu Ser Phe Asp Val Pro Asp Gly Phe Tyr Ser Thr Ser Pro

275 280 285

Ile Gln Ser Val Glu Leu Pro Val Ser Ile Val Ser Leu Pro Val Tyr

290 295 300

His Lys His Thr Phe Ile Val Leu Tyr Val Asp Phe Lys Pro Gln Ser

305 310 315 320

Gly Gly Gly Lys Cys Phe Asn Cys Tyr Pro Ala Gly Val Asn Ile Thr

325 330 335

Leu Ala Asn Phe Asn Glu Thr Lys Gly Pro Leu Cys Val Asp Thr Ser

340 345 350

His Phe Thr Thr Lys Tyr Val Ala Val Tyr Ala Asn Val Gly Arg Trp

355 360 365

Ser Ala Ser Ile Asn Thr Gly Asn Cys Pro Phe Ser Phe Gly Lys Val

370 375 380

Asn Asn Phe Val Lys Phe Gly Ser Val Cys Phe Ser Leu Lys Asp Ile

385 390 395 400

Pro Gly Gly Cys Ala Met Pro Ile Val Ala Asn Trp Ala Tyr Ser Lys

405 410 415

Tyr Tyr Thr Ile Gly Ser Leu Tyr Val Ser Trp Ser Asp Gly Asp Gly

420 425 430

Ile Thr Gly Val Pro Gln Pro Val Glu Gly Val Ser Ser Phe Met Asn

435 440 445

Val Thr Leu Asp Lys Cys Thr Lys Tyr Asn Ile Tyr Asp Val Ser Gly

450 455 460

Val Gly Val Ile Arg Val Ser Asn Asp Thr Phe Leu Asn Gly Ile Thr

465 470 475 480

Tyr Thr Ser Thr Ser Gly Asn Leu Leu Gly Phe Lys Asp Val Thr Lys

485 490 495

Gly Thr Ile Tyr Ser Ile Thr Pro Cys Asn Pro Pro Asp Gln Leu Val

500 505 510

Val Tyr Gln Gln Ala Val Val Gly Ala Met Leu Ser Glu Asn Phe Thr
515 520 525
Ser Tyr Gly Phe Ser Asn Val Val Glu Leu Pro Lys Phe Phe Tyr Ala
530 535 540
Ser Asn Gly Thr Tyr Asn Cys Thr Asp Ala Val Leu Thr Tyr Ser Ser

545 550 555 560
Phe Gly Val Cys Ala Asp Gly Ser Ile Ile Ala Val Gln Pro Ala Asn
565 570 575
Val Ser Tyr Asp Ser Val Ser Ala Ile Val Thr Ala Asn Leu Ser Ile
580 585 590
Pro Ser Asn Trp Thr Thr Ser Val Gln Val Glu Tyr Leu Gln Ile Thr
595 600 605
Ser Thr Pro Ile Val Val Asp Cys Ser Thr Tyr Val Cys Asn Gly Asn

610 615 620
Val Arg Cys Val Glu Leu Leu Lys Gln Tyr Thr Ser Ala Cys Lys Thr
625 630 635 640
Ile Glu Asp Ala Leu Arg Asn Ser Ala Arg Leu Glu Ser Ala Asp Val
645 650 655
Ser Glu Met Leu Thr Phe Asp Lys Lys Ala Phe Thr Leu Ala Asn Val
660 665 670
Ser Ser Phe Gly Asp Tyr Asn Leu Ser Ser Val Ile Pro Ser Leu Pro

675 680 685
Thr Ser Gly Ser Arg Val Ala Gly Arg Ser Ala Ile Glu Asp Ile Leu
690 695 700
Phe Ser Lys Leu Val Thr Ser Gly Leu Gly Thr Val Asp Ala Asp Tyr
705 710 715 720
Lys Lys Cys Thr Lys Gly Leu Ser Ile Ala Asp Leu Ala Cys Ala Gln
725 730 735
Tyr Tyr Asn Gly Ile Met Val Leu Pro Gly Val Ala Asp Ala Glu Arg

740 745 750
Met Ala Met Tyr Thr Gly Ser Leu Ile Gly Gly Ile Ala Leu Gly Gly

```
        755                 760                 765
Leu Thr Ser Ala Val Ser Ile Pro Phe Ser Leu Ala Ile Gln Ala Arg
    770                 775                 780
Leu Asn Tyr Val Ala Leu Gln Thr Asp Val Leu Gln Glu Asn Gln Lys
785                 790                 795                 800
Ile Leu Ala Ala Ser Phe Asn Lys Ala Met Thr Asn Ile Val Asp Ala

                805                 810                 815
Phe Thr Gly Val Asn Asp Ala Ile Thr Gln Thr Ser Gln Ala Leu Gln
            820                 825                 830
Thr Val Ala Thr Ala Leu Asn Lys Ile Gln Asp Val Val Asn Gln Gln
        835                 840                 845
Gly Asn Ser Leu Asn His Leu Thr Ser Gln Leu Arg Gln Asn Phe Gln
    850                 855                 860
Ala Ile Ser Ser Ser Ile Gln Ala Ile Tyr Asp Arg Leu Asp Pro Pro

865                 870                 875                 880
Gln Ala Asp Gln Gln Val Asp Arg Leu Ile Thr Gly Arg Leu Ala Ala
                885                 890                 895
Leu Asn Val Phe Val Ser His Thr Leu Thr Lys Tyr Thr Glu Val Arg
            900                 905                 910
Ala Ser Arg Gln Leu Ala Gln Gln Lys Val Asn Glu Cys Val Lys Ser
        915                 920                 925
Gln Ser Lys Arg Tyr Gly Phe Cys Gly Asn Gly Thr His Ile Phe Ser

    930                 935                 940
Ile Val Asn Ala Ala Pro Glu Gly Leu Val Phe Leu His Thr Val Leu
945                 950                 955                 960
Leu Pro Thr Gln Tyr Lys Asp Val Glu Ala Trp Ser Gly Leu Cys Val
                965                 970                 975
Asp Gly Thr Asn Gly Tyr Val Leu Arg Gln Pro Asn Leu Ala Leu Tyr
            980                 985                 990
Lys Glu Gly Asn Tyr Tyr Arg Ile  Thr Ser Arg Ile Met  Phe Glu Pro

        995                 1000                1005
```

Arg Ile Pro Thr Met Ala Asp Phe Val Gln Ile Glu Asn Cys Asn
1010 1015 1020
Val Thr Phe Val Asn Ile Ser Arg Ser Glu Leu Gln Thr Ile Val
1025 1030 1035
Pro Glu Tyr Ile Asp Val Asn Lys Thr Leu Gln Glu Leu Ser Tyr
1040 1045 1050
Lys Leu Pro Asn Tyr Thr Val Pro Asp Leu Val Val Glu Gln Tyr

1055 1060 1065
Asn Gln Thr Ile Leu Asn Leu Thr Ser Glu Ile Ser Thr Leu Glu
1070 1075 1080
Asn Lys Ser Ala Glu Leu Asn Tyr Thr Val Gln Lys Leu Gln Thr
1085 1090 1095
Leu Ile Asp Asn Ile Asn Ser Thr Leu Val Asp Leu Lys Trp Leu
1100 1105 1110
Asn Arg Val Glu Thr Tyr Ile Lys Trp Pro Trp Trp Val Trp Leu

1115 1120 1125
Cys Ile Ser Val Val Leu Ile Phe Val Val Ser Met Leu Leu Leu
1130 1135 1140
Cys Cys Cys Ser Thr Gly Cys Cys Gly Phe Phe Ser Cys Phe Ala
1145 1150 1155
Ser Ser Ile Arg Gly Cys Cys Glu Ser Thr Lys Leu Pro Tyr Tyr
1160 1165 1170
Asp Val Glu Lys Ile His Ile Gln
1175 1180

<210> 154
<211> 1163
<212> PRT
<213> Artificial Sequence
<220><223> PDI-229E -H5iTMCT-AA
<400> 154
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Thr Thr Asn Gly Leu Asn Thr
20 25 30
Ser Tyr Ser Val Cys Asn Gly Cys Val Gly Tyr Ser Glu Asn Val Phe
35 40 45

Ala Val Glu Ser Gly Gly Tyr Ile Pro Ser Asp Phe Ala Phe Asn Asn
50 55 60
Trp Phe Leu Leu Thr Asn Thr Ser Ser Val Val Asp Gly Val Val Arg
65 70 75 80
Ser Phe Gln Pro Leu Leu Leu Asn Cys Leu Trp Ser Val Ser Gly Leu
85 90 95
Arg Phe Thr Thr Gly Phe Val Tyr Phe Asn Gly Thr Gly Arg Gly Asp
100 105 110

Cys Lys Gly Phe Ser Ser Asp Val Leu Ser Asp Val Ile Arg Tyr Asn
115 120 125
Leu Asn Phe Glu Glu Asn Leu Arg Arg Gly Thr Ile Leu Phe Lys Thr
130 135 140
Ser Tyr Gly Val Val Val Phe Tyr Cys Thr Asn Asn Thr Leu Val Ser
145 150 155 160
Gly Asp Ala His Ile Pro Phe Gly Thr Val Leu Gly Asn Phe Tyr Cys
165 170 175

Phe Val Asn Thr Thr Ile Gly Asn Glu Thr Thr Ser Ala Phe Val Gly
180 185 190
Ala Leu Pro Lys Thr Val Arg Glu Phe Val Ile Ser Arg Thr Gly His
195 200 205
Phe Tyr Ile Asn Gly Tyr Arg Tyr Phe Thr Leu Gly Asn Val Glu Ala
210 215 220
Val Asn Phe Asn Val Thr Thr Ala Glu Thr Thr Asp Phe Cys Thr Val
225 230 235 240

Ala Leu Ala Ser Tyr Ala Asp Val Leu Val Asn Val Ser Gln Thr Ser
245 250 255
Ile Ala Asn Ile Ile Tyr Cys Asn Ser Val Ile Asn Arg Leu Arg Cys

260 265 270

Asp Gln Leu Ser Phe Asp Val Pro Asp Gly Phe Tyr Ser Thr Ser Pro

275 280 285

Ile Gln Ser Val Glu Leu Pro Val Ser Ile Val Ser Leu Pro Val Tyr

290 295 300

His Lys His Thr Phe Ile Val Leu Tyr Val Asp Phe Lys Pro Gln Ser

305 310 315 320

Gly Gly Gly Lys Cys Phe Asn Cys Tyr Pro Ala Gly Val Asn Ile Thr

325 330 335

Leu Ala Asn Phe Asn Glu Thr Lys Gly Pro Leu Cys Val Asp Thr Ser

340 345 350

His Phe Thr Thr Lys Tyr Val Ala Val Tyr Ala Asn Val Gly Arg Trp

355 360 365

Ser Ala Ser Ile Asn Thr Gly Asn Cys Pro Phe Ser Phe Gly Lys Val

370 375 380

Asn Asn Phe Val Lys Phe Gly Ser Val Cys Phe Ser Leu Lys Asp Ile

385 390 395 400

Pro Gly Gly Cys Ala Met Pro Ile Val Ala Asn Trp Ala Tyr Ser Lys

405 410 415

Tyr Tyr Thr Ile Gly Ser Leu Tyr Val Ser Trp Ser Asp Gly Asp Gly

420 425 430

Ile Thr Gly Val Pro Gln Pro Val Glu Gly Val Ser Ser Phe Met Asn

435 440 445

Val Thr Leu Asp Lys Cys Thr Lys Tyr Asn Ile Tyr Asp Val Ser Gly

450 455 460

Val Gly Val Ile Arg Val Ser Asn Asp Thr Phe Leu Asn Gly Ile Thr

465 470 475 480

Tyr Thr Ser Thr Ser Gly Asn Leu Leu Gly Phe Lys Asp Val Thr Lys

485 490 495

Gly Thr Ile Tyr Ser Ile Thr Pro Cys Asn Pro Pro Asp Gln Leu Val

500 505 510

Val Tyr Gln Gln Ala Val Val Gly Ala Met Leu Ser Glu Asn Phe Thr
515 520 525
Ser Tyr Gly Phe Ser Asn Val Val Glu Leu Pro Lys Phe Phe Tyr Ala
530 535 540
Ser Asn Gly Thr Tyr Asn Cys Thr Asp Ala Val Leu Thr Tyr Ser Ser
545 550 555 560

Phe Gly Val Cys Ala Asp Gly Ser Ile Ile Ala Val Gln Pro Ala Asn
565 570 575
Val Ser Tyr Asp Ser Val Ser Ala Ile Val Thr Ala Asn Leu Ser Ile
580 585 590
Pro Ser Asn Trp Thr Thr Ser Val Gln Val Glu Tyr Leu Gln Ile Thr
595 600 605
Ser Thr Pro Ile Val Val Asp Cys Ser Thr Tyr Val Cys Asn Gly Asn
610 615 620

Val Arg Cys Val Glu Leu Leu Lys Gln Tyr Thr Ser Ala Cys Lys Thr
625 630 635 640
Ile Glu Asp Ala Leu Arg Asn Ser Ala Arg Leu Glu Ser Ala Asp Val
645 650 655
Ser Glu Met Leu Thr Phe Asp Lys Lys Ala Phe Thr Leu Ala Asn Val
660 665 670
Ser Ser Phe Gly Asp Tyr Asn Leu Ser Ser Val Ile Pro Ser Leu Pro
675 680 685

Thr Ser Gly Ser Arg Val Ala Gly Arg Ser Ala Ile Glu Asp Ile Leu
690 695 700
Phe Ser Lys Leu Val Thr Ser Gly Leu Gly Thr Val Asp Ala Asp Tyr
705 710 715 720
Lys Lys Cys Thr Lys Gly Leu Ser Ile Ala Asp Leu Ala Cys Ala Gln
725 730 735
Tyr Tyr Asn Gly Ile Met Val Leu Pro Gly Val Ala Asp Ala Glu Arg
740 745 750

Met Ala Met Tyr Thr Gly Ser Leu Ile Gly Gly Ile Ala Leu Gly Gly

```
        755                 760                 765
Leu Thr Ser Ala Val Ser Ile Pro Phe Ser Leu Ala Ile Gln Ala Arg
    770                 775                 780
Leu Asn Tyr Val Ala Leu Gln Thr Asp Val Leu Gln Glu Asn Gln Lys
785                 790                 795                 800
Ile Leu Ala Ala Ser Phe Asn Lys Ala Met Thr Asn Ile Val Asp Ala
                805                 810                 815

Phe Thr Gly Val Asn Asp Ala Ile Thr Gln Thr Ser Gln Ala Leu Gln
            820                 825                 830
Thr Val Ala Thr Ala Leu Asn Lys Ile Gln Asp Val Val Asn Gln Gln
        835                 840                 845
Gly Asn Ser Leu Asn His Leu Thr Ser Gln Leu Arg Gln Asn Phe Gln
    850                 855                 860
Ala Ile Ser Ser Ser Ile Gln Ala Ile Tyr Asp Arg Leu Asp Pro Pro
865                 870                 875                 880

Gln Ala Asp Gln Gln Val Asp Arg Leu Ile Thr Gly Arg Leu Ala Ala
                885                 890                 895
Leu Asn Val Phe Val Ser His Thr Leu Thr Lys Tyr Thr Glu Val Arg
            900                 905                 910
Ala Ser Arg Gln Leu Ala Gln Gln Lys Val Asn Glu Cys Val Lys Ser
        915                 920                 925
Gln Ser Lys Arg Tyr Gly Phe Cys Gly Asn Gly Thr His Ile Phe Ser
    930                 935                 940

Ile Val Asn Ala Ala Pro Glu Gly Leu Val Phe Leu His Thr Val Leu
945                 950                 955                 960
Leu Pro Thr Gln Tyr Lys Asp Val Glu Ala Trp Ser Gly Leu Cys Val
                965                 970                 975
Asp Gly Thr Asn Gly Tyr Val Leu Arg Gln Pro Asn Leu Ala Leu Tyr
            980                 985                 990
Lys Glu Gly Asn Tyr Tyr Arg Ile  Thr Ser Arg Ile Met  Phe Glu Pro
        995                 1000                1005
```

Arg Ile Pro Thr Met Ala Asp Phe Val Gln Ile Glu Asn Cys Asn
1010 1015 1020

Val Thr Phe Val Asn Ile Ser Arg Ser Glu Leu Gln Thr Ile Val
1025 1030 1035

Pro Glu Tyr Ile Asp Val Asn Lys Thr Leu Gln Glu Leu Ser Tyr
1040 1045 1050

Lys Leu Pro Asn Tyr Thr Val Pro Asp Leu Val Val Glu Gln Tyr
1055 1060 1065

Asn Gln Thr Ile Leu Asn Leu Thr Ser Glu Ile Ser Thr Leu Glu
1070 1075 1080

Asn Lys Ser Ala Glu Leu Asn Tyr Thr Val Gln Lys Leu Gln Thr
1085 1090 1095

Leu Ile Asp Asn Ile Asn Ser Thr Leu Val Asp Leu Lys Trp Leu
1100 1105 1110

Asn Arg Val Glu Thr Tyr Ile Lys Trp Pro Trp Trp Gln Ile Leu
1115 1120 1125

Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met
1130 1135 1140

Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln
1145 1150 1155

Cys Arg Ile Cys Ile
1160

<210> 155

<211> 1159

<212> PRT

<213> Artificial Sequence

<220><223> PDI-229E -H5iCT-AA

<400> 155

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Thr Thr Asn Gly Leu Asn Thr
20 25 30

Ser Tyr Ser Val Cys Asn Gly Cys Val Gly Tyr Ser Glu Asn Val Phe
35 40 45
Ala Val Glu Ser Gly Gly Tyr Ile Pro Ser Asp Phe Ala Phe Asn Asn
50 55 60
Trp Phe Leu Leu Thr Asn Thr Ser Ser Val Val Asp Gly Val Val Arg
65 70 75 80

Ser Phe Gln Pro Leu Leu Leu Asn Cys Leu Trp Ser Val Ser Gly Leu
85 90 95
Arg Phe Thr Thr Gly Phe Val Tyr Phe Asn Gly Thr Gly Arg Gly Asp
100 105 110
Cys Lys Gly Phe Ser Ser Asp Val Leu Ser Asp Val Ile Arg Tyr Asn
115 120 125
Leu Asn Phe Glu Glu Asn Leu Arg Arg Gly Thr Ile Leu Phe Lys Thr
130 135 140

Ser Tyr Gly Val Val Val Phe Tyr Cys Thr Asn Asn Thr Leu Val Ser
145 150 155 160
Gly Asp Ala His Ile Pro Phe Gly Thr Val Leu Gly Asn Phe Tyr Cys
165 170 175
Phe Val Asn Thr Thr Ile Gly Asn Glu Thr Thr Ser Ala Phe Val Gly
180 185 190
Ala Leu Pro Lys Thr Val Arg Glu Phe Val Ile Ser Arg Thr Gly His
195 200 205

Phe Tyr Ile Asn Gly Tyr Arg Tyr Phe Thr Leu Gly Asn Val Glu Ala
210 215 220
Val Asn Phe Asn Val Thr Thr Ala Glu Thr Thr Asp Phe Cys Thr Val
225 230 235 240
Ala Leu Ala Ser Tyr Ala Asp Val Leu Val Asn Val Ser Gln Thr Ser
245 250 255
Ile Ala Asn Ile Ile Tyr Cys Asn Ser Val Ile Asn Arg Leu Arg Cys
260 265 270

Asp Gln Leu Ser Phe Asp Val Pro Asp Gly Phe Tyr Ser Thr Ser Pro

275 280 285

Ile Gln Ser Val Glu Leu Pro Val Ser Ile Val Ser Leu Pro Val Tyr

290 295 300

His Lys His Thr Phe Ile Val Leu Tyr Val Asp Phe Lys Pro Gln Ser

305 310 315 320

Gly Gly Gly Lys Cys Phe Asn Cys Tyr Pro Ala Gly Val Asn Ile Thr

325 330 335

Leu Ala Asn Phe Asn Glu Thr Lys Gly Pro Leu Cys Val Asp Thr Ser

340 345 350

His Phe Thr Thr Lys Tyr Val Ala Val Tyr Ala Asn Val Gly Arg Trp

355 360 365

Ser Ala Ser Ile Asn Thr Gly Asn Cys Pro Phe Ser Phe Gly Lys Val

370 375 380

Asn Asn Phe Val Lys Phe Gly Ser Val Cys Phe Ser Leu Lys Asp Ile

385 390 395 400

Pro Gly Gly Cys Ala Met Pro Ile Val Ala Asn Trp Ala Tyr Ser Lys

405 410 415

Tyr Tyr Thr Ile Gly Ser Leu Tyr Val Ser Trp Ser Asp Gly Asp Gly

420 425 430

Ile Thr Gly Val Pro Gln Pro Val Glu Gly Val Ser Ser Phe Met Asn

435 440 445

Val Thr Leu Asp Lys Cys Thr Lys Tyr Asn Ile Tyr Asp Val Ser Gly

450 455 460

Val Gly Val Ile Arg Val Ser Asn Asp Thr Phe Leu Asn Gly Ile Thr

465 470 475 480

Tyr Thr Ser Thr Ser Gly Asn Leu Leu Gly Phe Lys Asp Val Thr Lys

485 490 495

Gly Thr Ile Tyr Ser Ile Thr Pro Cys Asn Pro Pro Asp Gln Leu Val

500 505 510

Val Tyr Gln Gln Ala Val Val Gly Ala Met Leu Ser Glu Asn Phe Thr

515 520 525

Ser Tyr Gly Phe Ser Asn Val Val Glu Leu Pro Lys Phe Phe Tyr Ala
530 535 540
Ser Asn Gly Thr Tyr Asn Cys Thr Asp Ala Val Leu Thr Tyr Ser Ser
545 550 555 560
Phe Gly Val Cys Ala Asp Gly Ser Ile Ile Ala Val Gln Pro Ala Asn
565 570 575
Val Ser Tyr Asp Ser Val Ser Ala Ile Val Thr Ala Asn Leu Ser Ile
580 585 590

Pro Ser Asn Trp Thr Thr Ser Val Gln Val Glu Tyr Leu Gln Ile Thr
595 600 605
Ser Thr Pro Ile Val Val Asp Cys Ser Thr Tyr Val Cys Asn Gly Asn
610 615 620
Val Arg Cys Val Glu Leu Leu Lys Gln Tyr Thr Ser Ala Cys Lys Thr
625 630 635 640
Ile Glu Asp Ala Leu Arg Asn Ser Ala Arg Leu Glu Ser Ala Asp Val
645 650 655

Ser Glu Met Leu Thr Phe Asp Lys Lys Ala Phe Thr Leu Ala Asn Val
660 665 670
Ser Ser Phe Gly Asp Tyr Asn Leu Ser Ser Val Ile Pro Ser Leu Pro
675 680 685
Thr Ser Gly Ser Arg Val Ala Gly Arg Ser Ala Ile Glu Asp Ile Leu
690 695 700
Phe Ser Lys Leu Val Thr Ser Gly Leu Gly Thr Val Asp Ala Asp Tyr
705 710 715 720

Lys Lys Cys Thr Lys Gly Leu Ser Ile Ala Asp Leu Ala Cys Ala Gln
725 730 735
Tyr Tyr Asn Gly Ile Met Val Leu Pro Gly Val Ala Asp Ala Glu Arg
740 745 750
Met Ala Met Tyr Thr Gly Ser Leu Ile Gly Gly Ile Ala Leu Gly Gly
755 760 765
Leu Thr Ser Ala Val Ser Ile Pro Phe Ser Leu Ala Ile Gln Ala Arg

770 775 780

Leu Asn Tyr Val Ala Leu Gln Thr Asp Val Leu Gln Glu Asn Gln Lys
785 790 795 800
Ile Leu Ala Ala Ser Phe Asn Lys Ala Met Thr Asn Ile Val Asp Ala
805 810 815
Phe Thr Gly Val Asn Asp Ala Ile Thr Gln Thr Ser Gln Ala Leu Gln
820 825 830
Thr Val Ala Thr Ala Leu Asn Lys Ile Gln Asp Val Val Asn Gln Gln
835 840 845

Gly Asn Ser Leu Asn His Leu Thr Ser Gln Leu Arg Gln Asn Phe Gln
850 855 860
Ala Ile Ser Ser Ser Ile Gln Ala Ile Tyr Asp Arg Leu Asp Pro Pro
865 870 875 880
Gln Ala Asp Gln Gln Val Asp Arg Leu Ile Thr Gly Arg Leu Ala Ala
885 890 895
Leu Asn Val Phe Val Ser His Thr Leu Thr Lys Tyr Thr Glu Val Arg
900 905 910

Ala Ser Arg Gln Leu Ala Gln Gln Lys Val Asn Glu Cys Val Lys Ser
915 920 925
Gln Ser Lys Arg Tyr Gly Phe Cys Gly Asn Gly Thr His Ile Phe Ser
930 935 940
Ile Val Asn Ala Ala Pro Glu Gly Leu Val Phe Leu His Thr Val Leu
945 950 955 960
Leu Pro Thr Gln Tyr Lys Asp Val Glu Ala Trp Ser Gly Leu Cys Val
965 970 975

Asp Gly Thr Asn Gly Tyr Val Leu Arg Gln Pro Asn Leu Ala Leu Tyr
980 985 990
Lys Glu Gly Asn Tyr Tyr Arg Ile Thr Ser Arg Ile Met Phe Glu Pro
995 1000 1005
Arg Ile Pro Thr Met Ala Asp Phe Val Gln Ile Glu Asn Cys Asn
1010 1015 1020

Val Thr Phe Val Asn Ile Ser Arg Ser Glu Leu Gln Thr Ile Val
1025 1030 1035

Pro Glu Tyr Ile Asp Val Asn Lys Thr Leu Gln Glu Leu Ser Tyr
1040 1045 1050
Lys Leu Pro Asn Tyr Thr Val Pro Asp Leu Val Val Glu Gln Tyr
1055 1060 1065
Asn Gln Thr Ile Leu Asn Leu Thr Ser Glu Ile Ser Thr Leu Glu
1070 1075 1080
Asn Lys Ser Ala Glu Leu Asn Tyr Thr Val Gln Lys Leu Gln Thr
1085 1090 1095

Leu Ile Asp Asn Ile Asn Ser Thr Leu Val Asp Leu Lys Trp Leu
1100 1105 1110
Asn Arg Val Glu Thr Tyr Ile Lys Trp Pro Trp Trp Val Trp Leu
1115 1120 1125
Cys Ile Ser Val Val Leu Ile Phe Val Val Ser Met Leu Leu Leu
1130 1135 1140
Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys
1145 1150 1155

Ile

<210> 156
<211> 1156
<212> PRT
<213> Artificial Sequence
<220><223> PDI-229E -H5iCT(V4)-AA
<400> 156
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15
Leu Val Pro Ser Gln Ile Phe Ala Gln Thr Thr Asn Gly Leu Asn Thr
20 25 30
Ser Tyr Ser Val Cys Asn Gly Cys Val Gly Tyr Ser Glu Asn Val Phe

35 40 45

Ala Val Glu Ser Gly Gly Tyr Ile Pro Ser Asp Phe Ala Phe Asn Asn
50 55 60
Trp Phe Leu Leu Thr Asn Thr Ser Ser Val Val Asp Gly Val Val Arg
65 70 75 80
Ser Phe Gln Pro Leu Leu Leu Asn Cys Leu Trp Ser Val Ser Gly Leu
85 90 95
Arg Phe Thr Thr Gly Phe Val Tyr Phe Asn Gly Thr Gly Arg Gly Asp
100 105 110

Cys Lys Gly Phe Ser Ser Asp Val Leu Ser Asp Val Ile Arg Tyr Asn
115 120 125
Leu Asn Phe Glu Glu Asn Leu Arg Arg Gly Thr Ile Leu Phe Lys Thr
130 135 140
Ser Tyr Gly Val Val Val Phe Tyr Cys Thr Asn Asn Thr Leu Val Ser
145 150 155 160
Gly Asp Ala His Ile Pro Phe Gly Thr Val Leu Gly Asn Phe Tyr Cys
165 170 175

Phe Val Asn Thr Thr Ile Gly Asn Glu Thr Thr Ser Ala Phe Val Gly
180 185 190
Ala Leu Pro Lys Thr Val Arg Glu Phe Val Ile Ser Arg Thr Gly His
195 200 205
Phe Tyr Ile Asn Gly Tyr Arg Tyr Phe Thr Leu Gly Asn Val Glu Ala
210 215 220
Val Asn Phe Asn Val Thr Thr Ala Glu Thr Thr Asp Phe Cys Thr Val
225 230 235 240

Ala Leu Ala Ser Tyr Ala Asp Val Leu Val Asn Val Ser Gln Thr Ser
245 250 255
Ile Ala Asn Ile Ile Tyr Cys Asn Ser Val Ile Asn Arg Leu Arg Cys
260 265 270
Asp Gln Leu Ser Phe Asp Val Pro Asp Gly Phe Tyr Ser Thr Ser Pro
275 280 285

Ile Gln Ser Val Glu Leu Pro Val Ser Ile Val Ser Leu Pro Val Tyr
290 295 300

His Lys His Thr Phe Ile Val Leu Tyr Val Asp Phe Lys Pro Gln Ser
305 310 315 320
Gly Gly Gly Lys Cys Phe Asn Cys Tyr Pro Ala Gly Val Asn Ile Thr
325 330 335
Leu Ala Asn Phe Asn Glu Thr Lys Gly Pro Leu Cys Val Asp Thr Ser
340 345 350
His Phe Thr Thr Lys Tyr Val Ala Val Tyr Ala Asn Val Gly Arg Trp
355 360 365

Ser Ala Ser Ile Asn Thr Gly Asn Cys Pro Phe Ser Phe Gly Lys Val
370 375 380
Asn Asn Phe Val Lys Phe Gly Ser Val Cys Phe Ser Leu Lys Asp Ile
385 390 395 400
Pro Gly Gly Cys Ala Met Pro Ile Val Ala Asn Trp Ala Tyr Ser Lys
405 410 415
Tyr Tyr Thr Ile Gly Ser Leu Tyr Val Ser Trp Ser Asp Gly Asp Gly
420 425 430

Ile Thr Gly Val Pro Gln Pro Val Glu Gly Val Ser Ser Phe Met Asn
435 440 445
Val Thr Leu Asp Lys Cys Thr Lys Tyr Asn Ile Tyr Asp Val Ser Gly
450 455 460
Val Gly Val Ile Arg Val Ser Asn Asp Thr Phe Leu Asn Gly Ile Thr
465 470 475 480
Tyr Thr Ser Thr Ser Gly Asn Leu Leu Gly Phe Lys Asp Val Thr Lys
485 490 495

Gly Thr Ile Tyr Ser Ile Thr Pro Cys Asn Pro Pro Asp Gln Leu Val
500 505 510
Val Tyr Gln Gln Ala Val Val Gly Ala Met Leu Ser Glu Asn Phe Thr
515 520 525
Ser Tyr Gly Phe Ser Asn Val Val Glu Leu Pro Lys Phe Phe Tyr Ala

530 535 540

Ser Asn Gly Thr Tyr Asn Cys Thr Asp Ala Val Leu Thr Tyr Ser Ser

545 550 555 560

Phe Gly Val Cys Ala Asp Gly Ser Ile Ile Ala Val Gln Pro Ala Asn

565 570 575

Val Ser Tyr Asp Ser Val Ser Ala Ile Val Thr Ala Asn Leu Ser Ile

580 585 590

Pro Ser Asn Trp Thr Thr Ser Val Gln Val Glu Tyr Leu Gln Ile Thr

595 600 605

Ser Thr Pro Ile Val Val Asp Cys Ser Thr Tyr Val Cys Asn Gly Asn

610 615 620

Val Arg Cys Val Glu Leu Leu Lys Gln Tyr Thr Ser Ala Cys Lys Thr

625 630 635 640

Ile Glu Asp Ala Leu Arg Asn Ser Ala Arg Leu Glu Ser Ala Asp Val

645 650 655

Ser Glu Met Leu Thr Phe Asp Lys Lys Ala Phe Thr Leu Ala Asn Val

660 665 670

Ser Ser Phe Gly Asp Tyr Asn Leu Ser Ser Val Ile Pro Ser Leu Pro

675 680 685

Thr Ser Gly Ser Arg Val Ala Gly Arg Ser Ala Ile Glu Asp Ile Leu

690 695 700

Phe Ser Lys Leu Val Thr Ser Gly Leu Gly Thr Val Asp Ala Asp Tyr

705 710 715 720

Lys Lys Cys Thr Lys Gly Leu Ser Ile Ala Asp Leu Ala Cys Ala Gln

725 730 735

Tyr Tyr Asn Gly Ile Met Val Leu Pro Gly Val Ala Asp Ala Glu Arg

740 745 750

Met Ala Met Tyr Thr Gly Ser Leu Ile Gly Gly Ile Ala Leu Gly Gly

755 760 765

Leu Thr Ser Ala Val Ser Ile Pro Phe Ser Leu Ala Ile Gln Ala Arg

770 775 780

Leu Asn Tyr Val Ala Leu Gln Thr Asp Val Leu Gln Glu Asn Gln Lys
785 790 795 800
Ile Leu Ala Ala Ser Phe Asn Lys Ala Met Thr Asn Ile Val Asp Ala
805 810 815

Phe Thr Gly Val Asn Asp Ala Ile Thr Gln Thr Ser Gln Ala Leu Gln
820 825 830
Thr Val Ala Thr Ala Leu Asn Lys Ile Gln Asp Val Val Asn Gln Gln
835 840 845
Gly Asn Ser Leu Asn His Leu Thr Ser Gln Leu Arg Gln Asn Phe Gln
850 855 860
Ala Ile Ser Ser Ser Ile Gln Ala Ile Tyr Asp Arg Leu Asp Pro Pro
865 870 875 880

Gln Ala Asp Gln Gln Val Asp Arg Leu Ile Thr Gly Arg Leu Ala Ala
885 890 895
Leu Asn Val Phe Val Ser His Thr Leu Thr Lys Tyr Thr Glu Val Arg
900 905 910
Ala Ser Arg Gln Leu Ala Gln Gln Lys Val Asn Glu Cys Val Lys Ser
915 920 925
Gln Ser Lys Arg Tyr Gly Phe Cys Gly Asn Gly Thr His Ile Phe Ser
930 935 940

Ile Val Asn Ala Ala Pro Glu Gly Leu Val Phe Leu His Thr Val Leu
945 950 955 960
Leu Pro Thr Gln Tyr Lys Asp Val Glu Ala Trp Ser Gly Leu Cys Val
965 970 975
Asp Gly Thr Asn Gly Tyr Val Leu Arg Gln Pro Asn Leu Ala Leu Tyr
980 985 990
Lys Glu Gly Asn Tyr Tyr Arg Ile Thr Ser Arg Ile Met Phe Glu Pro
995 1000 1005

Arg Ile Pro Thr Met Ala Asp Phe Val Gln Ile Glu Asn Cys Asn
1010 1015 1020
Val Thr Phe Val Asn Ile Ser Arg Ser Glu Leu Gln Thr Ile Val

1025 1030 1035

Pro Glu Tyr Ile Asp Val Asn Lys Thr Leu Gln Glu Leu Ser Tyr

1040 1045 1050

Lys Leu Pro Asn Tyr Thr Val Pro Asp Leu Val Val Glu Gln Tyr

1055 1060 1065

Asn Gln Thr Ile Leu Asn Leu Thr Ser Glu Ile Ser Thr Leu Glu

1070 1075 1080

Asn Lys Ser Ala Glu Leu Asn Tyr Thr Val Gln Lys Leu Gln Thr

1085 1090 1095

Leu Ile Asp Asn Ile Asn Ser Thr Leu Val Asp Leu Lys Trp Leu

1100 1105 1110

Asn Arg Val Glu Thr Tyr Ile Lys Trp Pro Trp Trp Val Trp Leu

1115 1120 1125

Cys Ile Ser Val Val Leu Ile Phe Val Val Ser Met Leu Leu Leu

1130 1135 1140

Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

1145 1150 1155

<210> 157

<211> 1159

<212> PRT

<213> Artificial Sequence

<220><223> PDI-229E-H1cCT-AA

<400> 157

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Thr Thr Asn Gly Leu Asn Thr

20 25 30

Ser Tyr Ser Val Cys Asn Gly Cys Val Gly Tyr Ser Glu Asn Val Phe

35 40 45

Ala Val Glu Ser Gly Gly Tyr Ile Pro Ser Asp Phe Ala Phe Asn Asn

50 55 60

Trp Phe Leu Leu Thr Asn Thr Ser Ser Val Val Asp Gly Val Val Arg

65 70 75 80

Ser Phe Gln Pro Leu Leu Leu Asn Cys Leu Trp Ser Val Ser Gly Leu

85 90 95

Arg Phe Thr Thr Gly Phe Val Tyr Phe Asn Gly Thr Gly Arg Gly Asp

100 105 110

Cys Lys Gly Phe Ser Ser Asp Val Leu Ser Asp Val Ile Arg Tyr Asn

115 120 125

Leu Asn Phe Glu Glu Asn Leu Arg Arg Gly Thr Ile Leu Phe Lys Thr

130 135 140

Ser Tyr Gly Val Val Val Phe Tyr Cys Thr Asn Asn Thr Leu Val Ser

145 150 155 160

Gly Asp Ala His Ile Pro Phe Gly Thr Val Leu Gly Asn Phe Tyr Cys

165 170 175

Phe Val Asn Thr Thr Ile Gly Asn Glu Thr Thr Ser Ala Phe Val Gly

180 185 190

Ala Leu Pro Lys Thr Val Arg Glu Phe Val Ile Ser Arg Thr Gly His

195 200 205

Phe Tyr Ile Asn Gly Tyr Arg Tyr Phe Thr Leu Gly Asn Val Glu Ala

210 215 220

Val Asn Phe Asn Val Thr Thr Ala Glu Thr Thr Asp Phe Cys Thr Val

225 230 235 240

Ala Leu Ala Ser Tyr Ala Asp Val Leu Val Asn Val Ser Gln Thr Ser

245 250 255

Ile Ala Asn Ile Ile Tyr Cys Asn Ser Val Ile Asn Arg Leu Arg Cys

260 265 270

Asp Gln Leu Ser Phe Asp Val Pro Asp Gly Phe Tyr Ser Thr Ser Pro

275 280 285

Ile Gln Ser Val Glu Leu Pro Val Ser Ile Val Ser Leu Pro Val Tyr

290 295 300

His Lys His Thr Phe Ile Val Leu Tyr Val Asp Phe Lys Pro Gln Ser

305 310 315 320

Gly Gly Gly Lys Cys Phe Asn Cys Tyr Pro Ala Gly Val Asn Ile Thr
325 330 335
Leu Ala Asn Phe Asn Glu Thr Lys Gly Pro Leu Cys Val Asp Thr Ser

340 345 350
His Phe Thr Thr Lys Tyr Val Ala Val Tyr Ala Asn Val Gly Arg Trp
355 360 365
Ser Ala Ser Ile Asn Thr Gly Asn Cys Pro Phe Ser Phe Gly Lys Val
370 375 380
Asn Asn Phe Val Lys Phe Gly Ser Val Cys Phe Ser Leu Lys Asp Ile
385 390 395 400
Pro Gly Gly Cys Ala Met Pro Ile Val Ala Asn Trp Ala Tyr Ser Lys

405 410 415
Tyr Tyr Thr Ile Gly Ser Leu Tyr Val Ser Trp Ser Asp Gly Asp Gly
420 425 430
Ile Thr Gly Val Pro Gln Pro Val Glu Gly Val Ser Ser Phe Met Asn
435 440 445
Val Thr Leu Asp Lys Cys Thr Lys Tyr Asn Ile Tyr Asp Val Ser Gly
450 455 460
Val Gly Val Ile Arg Val Ser Asn Asp Thr Phe Leu Asn Gly Ile Thr

465 470 475 480
Tyr Thr Ser Thr Ser Gly Asn Leu Leu Gly Phe Lys Asp Val Thr Lys
485 490 495
Gly Thr Ile Tyr Ser Ile Thr Pro Cys Asn Pro Pro Asp Gln Leu Val
500 505 510
Val Tyr Gln Gln Ala Val Val Gly Ala Met Leu Ser Glu Asn Phe Thr
515 520 525
Ser Tyr Gly Phe Ser Asn Val Val Glu Leu Pro Lys Phe Phe Tyr Ala

530 535 540
Ser Asn Gly Thr Tyr Asn Cys Thr Asp Ala Val Leu Thr Tyr Ser Ser
545 550 555 560
Phe Gly Val Cys Ala Asp Gly Ser Ile Ile Ala Val Gln Pro Ala Asn

565 570 575

Val Ser Tyr Asp Ser Val Ser Ala Ile Val Thr Ala Asn Leu Ser Ile

580 585 590

Pro Ser Asn Trp Thr Thr Ser Val Gln Val Glu Tyr Leu Gln Ile Thr

595 600 605

Ser Thr Pro Ile Val Val Asp Cys Ser Thr Tyr Val Cys Asn Gly Asn

610 615 620

Val Arg Cys Val Glu Leu Leu Lys Gln Tyr Thr Ser Ala Cys Lys Thr

625 630 635 640

Ile Glu Asp Ala Leu Arg Asn Ser Ala Arg Leu Glu Ser Ala Asp Val

645 650 655

Ser Glu Met Leu Thr Phe Asp Lys Lys Ala Phe Thr Leu Ala Asn Val

660 665 670

Ser Ser Phe Gly Asp Tyr Asn Leu Ser Ser Val Ile Pro Ser Leu Pro

675 680 685

Thr Ser Gly Ser Arg Val Ala Gly Arg Ser Ala Ile Glu Asp Ile Leu

690 695 700

Phe Ser Lys Leu Val Thr Ser Gly Leu Gly Thr Val Asp Ala Asp Tyr

705 710 715 720

Lys Lys Cys Thr Lys Gly Leu Ser Ile Ala Asp Leu Ala Cys Ala Gln

725 730 735

Tyr Tyr Asn Gly Ile Met Val Leu Pro Gly Val Ala Asp Ala Glu Arg

740 745 750

Met Ala Met Tyr Thr Gly Ser Leu Ile Gly Gly Ile Ala Leu Gly Gly

755 760 765

Leu Thr Ser Ala Val Ser Ile Pro Phe Ser Leu Ala Ile Gln Ala Arg

770 775 780

Leu Asn Tyr Val Ala Leu Gln Thr Asp Val Leu Gln Glu Asn Gln Lys

785 790 795 800

Ile Leu Ala Ala Ser Phe Asn Lys Ala Met Thr Asn Ile Val Asp Ala

805 810 815

```
Phe Thr Gly Val Asn Asp Ala Ile Thr Gln Thr Ser Gln Ala Leu Gln
            820                 825                 830
Thr Val Ala Thr Ala Leu Asn Lys Ile Gln Asp Val Val Asn Gln Gln
        835                 840                 845
Gly Asn Ser Leu Asn His Leu Thr Ser Gln Leu Arg Gln Asn Phe Gln


    850                 855                 860
Ala Ile Ser Ser Ser Ile Gln Ala Ile Tyr Asp Arg Leu Asp Pro Pro
865                 870                 875                 880
Gln Ala Asp Gln Gln Val Asp Arg Leu Ile Thr Gly Arg Leu Ala Ala
                885                 890                 895
Leu Asn Val Phe Val Ser His Thr Leu Thr Lys Tyr Thr Glu Val Arg
            900                 905                 910
Ala Ser Arg Gln Leu Ala Gln Gln Lys Val Asn Glu Cys Val Lys Ser


        915                 920                 925
Gln Ser Lys Arg Tyr Gly Phe Cys Gly Asn Gly Thr His Ile Phe Ser
    930                 935                 940
Ile Val Asn Ala Ala Pro Glu Gly Leu Val Phe Leu His Thr Val Leu
945                 950                 955                 960
Leu Pro Thr Gln Tyr Lys Asp Val Glu Ala Trp Ser Gly Leu Cys Val
                965                 970                 975
Asp Gly Thr Asn Gly Tyr Val Leu Arg Gln Pro Asn Leu Ala Leu Tyr


            980                 985                 990
Lys Glu Gly Asn Tyr Tyr Arg Ile  Thr Ser Arg Ile Met  Phe Glu Pro
        995                 1000                1005
Arg Ile  Pro Thr Met Ala Asp  Phe Val Gln Ile Glu  Asn Cys Asn
    1010                1015                1020
Val Thr  Phe Val Asn Ile Ser  Arg Ser Glu Leu Gln  Thr Ile Val
    1025                1030                1035
Pro Glu  Tyr Ile Asp Val Asn  Lys Thr Leu Gln Glu  Leu Ser Tyr


    1040                1045                1050
Lys Leu  Pro Asn Tyr Thr Val  Pro Asp Leu Val Val  Glu Gln Tyr
```

1055 1060 1065

Asn Gln Thr Ile Leu Asn Leu Thr Ser Glu Ile Ser Thr Leu Glu

1070 1075 1080

Asn Lys Ser Ala Glu Leu Asn Tyr Thr Val Gln Lys Leu Gln Thr

1085 1090 1095

Leu Ile Asp Asn Ile Asn Ser Thr Leu Val Asp Leu Lys Trp Leu

1100 1105 1110

Asn Arg Val Glu Thr Tyr Ile Lys Trp Pro Trp Trp Val Trp Leu

1115 1120 1125

Cys Ile Ser Val Val Leu Ile Phe Val Val Ser Met Leu Leu Leu

1130 1135 1140

Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys

1145 1150 1155

Ile

<210> 158

<211> 1360

<212> PRT

<

213> OC43-Coronavirus

<400> 158

Met Phe Leu Ile Leu Leu Ile Ser Leu Pro Thr Ala Phe Ala Val Ile

1 5 10 15

Gly Asp Leu Asn Cys Thr Leu Asp Pro Arg Leu Lys Gly Ser Phe Asn

20 25 30

Asn Arg Asp Thr Gly Pro Pro Ser Ile Ser Ile Asp Thr Val Asp Val

35 40 45

Thr Asn Gly Leu Gly Thr Tyr Tyr Val Leu Asp Arg Val Tyr Leu Asn

50 55 60

Thr Thr Leu Phe Leu Asn Gly Tyr Tyr Pro Thr Ser Gly Ser Thr Tyr

65 70 75 80

Arg Asn Met Ala Leu Lys Gly Thr Asp Leu Leu Ser Thr Leu Trp Phe

85 90 95

Lys Pro Pro Phe Leu Ser Asp Phe Ile Asn Gly Ile Phe Ala Lys Val
100 105 110
Lys Asn Thr Lys Val Phe Lys Asp Gly Val Met Tyr Ser Glu Phe Pro
115 120 125

Ala Ile Thr Ile Gly Ser Thr Phe Val Asn Thr Ser Tyr Ser Val Val
130 135 140
Val Gln Pro Arg Thr Ile Asn Ser Thr Gln Asp Gly Val Asn Lys Leu
145 150 155 160
Gln Gly Leu Leu Glu Val Ser Val Cys Gln Tyr Asn Met Cys Glu Tyr
165 170 175
Pro His Thr Ile Cys His Pro Asn Leu Gly Asn His Phe Lys Glu Leu
180 185 190

Trp His Tyr Asp Thr Gly Val Val Ser Cys Leu Tyr Lys Arg Asn Phe
195 200 205
Thr Tyr Asp Val Asn Ala Thr Tyr Leu Tyr Phe His Phe Tyr Gln Glu
210 215 220
Gly Gly Thr Phe Tyr Ala Tyr Phe Thr Asp Thr Gly Phe Val Thr Lys
225 230 235 240
Phe Leu Phe Asn Val Tyr Leu Gly Met Ala Leu Ser His Tyr Tyr Val
245 250 255

Met Pro Leu Thr Cys Ile Arg Arg Pro Lys Asp Gly Phe Ser Leu Glu
260 265 270
Tyr Trp Val Thr Pro Leu Thr Pro Arg Gln Tyr Leu Leu Ala Phe Asn
275 280 285
Gln Asp Gly Ile Ile Phe Asn Ala Val Asp Cys Met Ser Asp Phe Met
290 295 300
Ser Glu Ile Lys Cys Lys Thr Gln Ser Ile Ala Pro Pro Thr Gly Val
305 310 315 320

Tyr Glu Leu Asn Gly Tyr Thr Val Gln Pro Val Ala Asp Val Tyr Arg
325 330 335
Arg Lys Pro Asp Leu Pro Asn Cys Asn Ile Glu Ala Trp Leu Asn Asp

340 345 350

Lys Ser Val Pro Ser Pro Leu Asn Trp Glu Arg Lys Thr Phe Ser Asn

355 360 365

Cys Asn Phe Asn Met Ser Ser Leu Met Ser Phe Ile Gln Ala Asp Ser

370 375 380

Phe Thr Cys Asn Asn Ile Asp Ala Ala Lys Ile Tyr Gly Met Cys Phe

385 390 395 400

Ser Ser Ile Thr Ile Asp Lys Phe Ala Ile Pro Asn Arg Arg Lys Val

405 410 415

Asp Leu Gln Leu Gly Asn Leu Gly Tyr Leu Gln Ser Ser Asn Tyr Arg

420 425 430

Ile Asp Thr Thr Ala Thr Ser Cys Gln Leu Tyr Tyr Asn Leu Pro Ala

435 440 445

Ala Asn Val Ser Val Ser Arg Phe Asn Pro Ser Thr Trp Asn Lys Arg

450 455 460

Phe Gly Phe Ile Glu Asp Ser Val Phe Val Pro Gln Pro Thr Gly Val

465 470 475 480

Phe Thr Asn His Ser Val Val Tyr Ala Gln His Cys Phe Lys Ala Pro

485 490 495

Lys Asn Phe Cys Pro Cys Ser Ser Cys Ser Cys Pro Gly Lys Asn Asn

500 505 510

Gly Ile Gly Thr Cys Pro Ala Gly Thr Asn Ser Leu Thr Cys Asp Asn

515 520 525

Leu Cys Thr Leu Asp Pro Ile Thr Leu Lys Ala Pro Asp Thr Tyr Lys

530 535 540

Cys Pro Gln Ser Lys Ser Leu Val Gly Ile Gly Glu His Cys Ser Gly

545 550 555 560

Leu Ala Val Lys Ser Asp Tyr Cys Gly Asn Asn Ser Cys Thr Cys Gln

565 570 575

Pro Gln Ala Phe Leu Gly Trp Ser Ala Asp Ser Cys Leu Gln Gly Asp

580 585 590

Lys Cys Asn Ile Phe Ala Asn Phe Ile Leu His Asp Val Asn Asn Gly
595 600 605
Leu Thr Cys Ser Thr Asp Leu Gln Lys Ala Asn Thr Glu Ile Glu Leu
610 615 620
Gly Val Cys Val Asn Tyr Asp Leu Tyr Gly Ile Ser Gly Gln Gly Ile
625 630 635 640

Phe Val Glu Val Asn Ala Thr Tyr Tyr Asn Ser Trp Gln Asn Leu Leu
645 650 655
Tyr Asp Ser Asn Gly Asn Leu Tyr Gly Phe Arg Asp Tyr Ile Thr Asn
660 665 670
Arg Thr Phe Met Ile His Ser Cys Tyr Ser Gly Arg Val Ser Ala Ala
675 680 685
Tyr His Ala Asn Ser Ser Glu Pro Ala Leu Leu Phe Arg Asn Ile Lys
690 695 700

Cys Asn Tyr Val Phe Asn Asn Ser Leu Thr Arg Gln Leu Gln Pro Ile
705 710 715 720
Asn Tyr Ser Phe Asp Ser Tyr Leu Gly Cys Val Val Asn Ala Tyr Asn
725 730 735
Ser Thr Ala Ile Ser Val Gln Thr Cys Asp Leu Thr Val Gly Ser Gly
740 745 750
Tyr Cys Val Asp Tyr Ser Lys Asn Arg Arg Ser Arg Arg Ala Ile Thr
755 760 765

Thr Gly Tyr Arg Phe Thr Asn Phe Glu Pro Phe Thr Val Asn Ser Val
770 775 780
Asn Asp Ser Leu Glu Pro Val Gly Gly Leu Tyr Glu Ile Gln Ile Pro
785 790 795 800
Ser Glu Phe Thr Ile Gly Asn Met Glu Glu Phe Ile Gln Thr Ser Ser
805 810 815
Pro Lys Val Thr Ile Asp Cys Ala Ala Phe Val Cys Gly Asp Tyr Ala
820 825 830

Ala Cys Lys Leu Gln Leu Val Glu Tyr Gly Ser Phe Cys Asp Asn Ile

835 840 845

Asn Ala Ile Leu Thr Glu Val Asn Glu Leu Leu Asp Thr Thr Gln Leu

850 855 860

Gln Val Ala Asn Ser Leu Met Asn Gly Val Thr Leu Ser Thr Lys Leu

865 870 875 880

Lys Asp Gly Val Asn Phe Asn Val Asp Asp Ile Asn Phe Ser Pro Val

885 890 895

Leu Gly Cys Leu Gly Ser Glu Cys Ser Lys Ala Ser Ser Arg Ser Ala

900 905 910

Ile Glu Asp Leu Leu Phe Asp Lys Val Lys Leu Ser Asp Val Gly Phe

915 920 925

Val Glu Ala Tyr Asn Asn Cys Thr Gly Gly Ala Glu Ile Arg Asp Leu

930 935 940

Ile Cys Val Gln Ser Tyr Lys Gly Ile Lys Val Leu Pro Pro Leu Leu

945 950 955 960

Ser Glu Asn Gln Ile Ser Gly Tyr Thr Leu Ala Ala Thr Ser Ala Ser

965 970 975

Leu Phe Pro Pro Trp Thr Ala Ala Ala Gly Val Pro Phe Tyr Leu Asn

980 985 990

Val Gln Tyr Arg Ile Asn Gly Leu Gly Val Thr Met Asp Val Leu Ser

995 1000 1005

Gln Asn Gln Lys Leu Ile Ala Asn Ala Phe Asn Asn Ala Leu His

1010 1015 1020

Ala Ile Gln Gln Gly Phe Asp Ala Thr Asn Ser Ala Leu Val Lys

1025 1030 1035

Ile Gln Ala Val Val Asn Ala Asn Ala Glu Ala Leu Asn Asn Leu

1040 1045 1050

Leu Gln Gln Leu Ser Asn Arg Phe Gly Ala Ile Ser Ala Ser Leu

1055 1060 1065

Gln Glu Ile Leu Ser Arg Leu Asp Ala Leu Glu Ala Glu Ala Gln

1070 1075 1080

Ile Asp Arg Leu Ile Asn Gly Arg Leu Thr Ala Leu Asn Ala Tyr
1085 1090 1095
Val Ser Gln Gln Leu Ser Asp Ser Thr Leu Val Lys Phe Ser Ala
1100 1105 1110
Ala Gln Ala Met Glu Lys Val Asn Glu Cys Val Lys Ser Gln Ser
1115 1120 1125
Ser Arg Ile Asn Phe Cys Gly Asn Gly Asn His Ile Ile Ser Leu
1130 1135 1140

Val Gln Asn Ala Pro Tyr Gly Leu Tyr Phe Ile His Phe Asn Tyr
1145 1150 1155
Val Pro Thr Lys Tyr Val Thr Ala Lys Val Ser Pro Gly Leu Cys
1160 1165 1170
Ile Ala Gly Asn Arg Gly Ile Ala Pro Lys Ser Gly Tyr Phe Val
1175 1180 1185
Asn Val Asn Asn Thr Trp Met Tyr Thr Gly Ser Gly Tyr Tyr Tyr
1190 1195 1200

Pro Glu Pro Ile Thr Glu Asn Asn Val Val Val Met Ser Thr Cys
1205 1210 1215
Ala Val Asn Tyr Thr Lys Ala Pro Tyr Val Met Leu Asn Thr Ser
1220 1225 1230
Ile Pro Asn Leu Pro Asp Phe Lys Glu Glu Leu Asp Gln Trp Phe
1235 1240 1245
Lys Asn Gln Thr Ser Val Ala Pro Asp Leu Ser Leu Asp Tyr Ile
1250 1255 1260

Asn Val Thr Phe Leu Asp Leu Gln Val Glu Met Asn Arg Leu Gln
1265 1270 1275
Glu Ala Ile Lys Val Leu Asn His Ser Tyr Ile Asn Leu Lys Asp
1280 1285 1290
Ile Gly Thr Tyr Glu Tyr Tyr Val Lys Trp Pro Trp Tyr Val Trp
1295 1300 1305
Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val Leu Leu Phe

1310 1315 1320

Phe Ile Cys Cys Cys Thr Gly Cys Gly Thr Ser Cys Phe Lys Lys
1325 1330 1335
Cys Gly Gly Cys Cys Asp Asp Tyr Thr Gly Tyr Gln Glu Leu Val
1340 1345 1350
Ile Lys Thr Ser His Asp Asp
1355 1360

<210> 159
<211> 1173
<212> PRT
<213> 229E Coronavirus
<400> 159

Met Phe Val Leu Leu Val Ala Tyr Ala Leu Leu His Ile Ala Gly Cys
1 5 10 15

Gln Thr Thr Asn Gly Leu Asn Thr Ser Tyr Ser Val Cys Asn Gly Cys
20 25 30
Val Gly Tyr Ser Glu Asn Val Phe Ala Val Glu Ser Gly Gly Tyr Ile
35 40 45
Pro Ser Asp Phe Ala Phe Asn Asn Trp Phe Leu Leu Thr Asn Thr Ser
50 55 60
Ser Val Val Asp Gly Val Val Arg Ser Phe Gln Pro Leu Leu Leu Asn
65 70 75 80

Cys Leu Trp Ser Val Ser Gly Leu Arg Phe Thr Thr Gly Phe Val Tyr
85 90 95
Phe Asn Gly Thr Gly Arg Gly Asp Cys Lys Gly Phe Ser Ser Asp Val
100 105 110
Leu Ser Asp Val Ile Arg Tyr Asn Leu Asn Phe Glu Glu Asn Leu Arg
115 120 125
Arg Gly Thr Ile Leu Phe Lys Thr Ser Tyr Gly Val Val Val Phe Tyr
130 135 140

Cys Thr Asn Asn Thr Leu Val Ser Gly Asp Ala His Ile Pro Phe Gly

145 150 155 160

Thr Val Leu Gly Asn Phe Tyr Cys Phe Val Asn Thr Thr Ile Gly Asn

165 170 175

Glu Thr Thr Ser Ala Phe Val Gly Ala Leu Pro Lys Thr Val Arg Glu

180 185 190

Phe Val Ile Ser Arg Thr Gly His Phe Tyr Ile Asn Gly Tyr Arg Tyr

195 200 205

Phe Thr Leu Gly Asn Val Glu Ala Val Asn Phe Asn Val Thr Thr Ala

210 215 220

Glu Thr Thr Asp Phe Cys Thr Val Ala Leu Ala Ser Tyr Ala Asp Val

225 230 235 240

Leu Val Asn Val Ser Gln Thr Ser Ile Ala Asn Ile Ile Tyr Cys Asn

245 250 255

Ser Val Ile Asn Arg Leu Arg Cys Asp Gln Leu Ser Phe Asp Val Pro

260 265 270

Asp Gly Phe Tyr Ser Thr Ser Pro Ile Gln Ser Val Glu Leu Pro Val

275 280 285

Ser Ile Val Ser Leu Pro Val Tyr His Lys His Thr Phe Ile Val Leu

290 295 300

Tyr Val Asp Phe Lys Pro Gln Ser Gly Gly Gly Lys Cys Phe Asn Cys

305 310 315 320

Tyr Pro Ala Gly Val Asn Ile Thr Leu Ala Asn Phe Asn Glu Thr Lys

325 330 335

Gly Pro Leu Cys Val Asp Thr Ser His Phe Thr Thr Lys Tyr Val Ala

340 345 350

Val Tyr Ala Asn Val Gly Arg Trp Ser Ala Ser Ile Asn Thr Gly Asn

355 360 365

Cys Pro Phe Ser Phe Gly Lys Val Asn Asn Phe Val Lys Phe Gly Ser

370 375 380

Val Cys Phe Ser Leu Lys Asp Ile Pro Gly Gly Cys Ala Met Pro Ile

385 390 395 400

Val Ala Asn Trp Ala Tyr Ser Lys Tyr Tyr Thr Ile Gly Ser Leu Tyr
405 410 415
Val Ser Trp Ser Asp Gly Asp Gly Ile Thr Gly Val Pro Gln Pro Val
420 425 430
Glu Gly Val Ser Ser Phe Met Asn Val Thr Leu Asp Lys Cys Thr Lys
435 440 445
Tyr Asn Ile Tyr Asp Val Ser Gly Val Gly Val Ile Arg Val Ser Asn
450 455 460

Asp Thr Phe Leu Asn Gly Ile Thr Tyr Thr Ser Thr Ser Gly Asn Leu
465 470 475 480
Leu Gly Phe Lys Asp Val Thr Lys Gly Thr Ile Tyr Ser Ile Thr Pro
485 490 495
Cys Asn Pro Pro Asp Gln Leu Val Val Tyr Gln Gln Ala Val Val Gly
500 505 510
Ala Met Leu Ser Glu Asn Phe Thr Ser Tyr Gly Phe Ser Asn Val Val
515 520 525

Glu Leu Pro Lys Phe Phe Tyr Ala Ser Asn Gly Thr Tyr Asn Cys Thr
530 535 540
Asp Ala Val Leu Thr Tyr Ser Ser Phe Gly Val Cys Ala Asp Gly Ser
545 550 555 560
Ile Ile Ala Val Gln Pro Arg Asn Val Ser Tyr Asp Ser Val Ser Ala
565 570 575
Ile Val Thr Ala Asn Leu Ser Ile Pro Ser Asn Trp Thr Thr Ser Val
580 585 590

Gln Val Glu Tyr Leu Gln Ile Thr Ser Thr Pro Ile Val Val Asp Cys
595 600 605
Ser Thr Tyr Val Cys Asn Gly Asn Val Arg Cys Val Glu Leu Leu Lys
610 615 620
Gln Tyr Thr Ser Ala Cys Lys Thr Ile Glu Asp Ala Leu Arg Asn Ser
625 630 635 640
Ala Arg Leu Glu Ser Ala Asp Val Ser Glu Met Leu Thr Phe Asp Lys

645 650 655

Lys Ala Phe Thr Leu Ala Asn Val Ser Ser Phe Gly Asp Tyr Asn Leu
660 665 670

Ser Ser Val Ile Pro Ser Leu Pro Thr Ser Gly Ser Arg Val Ala Gly
675 680 685

Arg Ser Ala Ile Glu Asp Ile Leu Phe Ser Lys Leu Val Thr Ser Gly
690 695 700

Leu Gly Thr Val Asp Ala Asp Tyr Lys Lys Cys Thr Lys Gly Leu Ser
705 710 715 720

Ile Ala Asp Leu Ala Cys Ala Gln Tyr Tyr Asn Gly Ile Met Val Leu
725 730 735

Pro Gly Val Ala Asp Ala Glu Arg Met Ala Met Tyr Thr Gly Ser Leu
740 745 750

Ile Gly Gly Ile Ala Leu Gly Gly Leu Thr Ser Ala Val Ser Ile Pro
755 760 765

Phe Ser Leu Ala Ile Gln Ala Arg Leu Asn Tyr Val Ala Leu Gln Thr
770 775 780

Asp Val Leu Gln Glu Asn Gln Lys Ile Leu Ala Ala Ser Phe Asn Lys
785 790 795 800

Ala Met Thr Asn Ile Val Asp Ala Phe Thr Gly Val Asn Asp Ala Ile
805 810 815

Thr Gln Thr Ser Gln Ala Leu Gln Thr Val Ala Thr Ala Leu Asn Lys
820 825 830

Ile Gln Asp Val Val Asn Gln Gln Gly Asn Ser Leu Asn His Leu Thr
835 840 845

Ser Gln Leu Arg Gln Asn Phe Gln Ala Ile Ser Ser Ser Ile Gln Ala
850 855 860

Ile Tyr Asp Arg Leu Asp Thr Ile Gln Ala Asp Gln Gln Val Asp Arg
865 870 875 880

Leu Ile Thr Gly Arg Leu Ala Ala Leu Asn Val Phe Val Ser His Thr
885 890 895

Leu Thr Lys Tyr Thr Glu Val Arg Ala Ser Arg Gln Leu Ala Gln Gln
900 905 910

Lys Val Asn Glu Cys Val Lys Ser Gln Ser Lys Arg Tyr Gly Phe Cys
915 920 925
Gly Asn Gly Thr His Ile Phe Ser Ile Val Asn Ala Ala Pro Glu Gly
930 935 940
Leu Val Phe Leu His Thr Val Leu Leu Pro Thr Gln Tyr Lys Asp Val
945 950 955 960
Glu Ala Trp Ser Gly Leu Cys Val Asp Gly Thr Asn Gly Tyr Val Leu
965 970 975

Arg Gln Pro Asn Leu Ala Leu Tyr Lys Glu Gly Asn Tyr Tyr Arg Ile
980 985 990
Thr Ser Arg Ile Met Phe Glu Pro Arg Ile Pro Thr Met Ala Asp Phe
995 1000 1005
Val Gln Ile Glu Asn Cys Asn Val Thr Phe Val Asn Ile Ser Arg
1010 1015 1020
Ser Glu Leu Gln Thr Ile Val Pro Glu Tyr Ile Asp Val Asn Lys
1025 1030 1035

Thr Leu Gln Glu Leu Ser Tyr Lys Leu Pro Asn Tyr Thr Val Pro
1040 1045 1050
Asp Leu Val Val Glu Gln Tyr Asn Gln Thr Ile Leu Asn Leu Thr
1055 1060 1065
Ser Glu Ile Ser Thr Leu Glu Asn Lys Ser Ala Glu Leu Asn Tyr
1070 1075 1080
Thr Val Gln Lys Leu Gln Thr Leu Ile Asp Asn Ile Asn Ser Thr
1085 1090 1095

Leu Val Asp Leu Lys Trp Leu Asn Arg Val Glu Thr Tyr Ile Lys
1100 1105 1110
Trp Pro Trp Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe
1115 1120 1125
Val Val Ser Met Leu Leu Leu Cys Cys Cys Ser Thr Gly Cys Cys

1130 1135 1140

Gly Phe Phe Ser Cys Phe Ala Ser Ser Ile Arg Gly Cys Cys Glu

1145 1150 1155

Ser Thr Lys Leu Pro Tyr Tyr Asp Val Glu Lys Ile His Ile Gln

1160 1165 1170

<210> 160

<211> 1346

<212> PRT

<213> OC43 Coronavirus

<400> 160

Val Ile Gly Asp Leu Asn Cys Thr Leu Asp Pro Arg Leu Lys Gly Ser

1 5 10 15

Phe Asn Asn Arg Asp Thr Gly Pro Pro Ser Ile Ser Ile Asp Thr Val

20 25 30

Asp Val Thr Asn Gly Leu Gly Thr Tyr Tyr Val Leu Asp Arg Val Tyr

35 40 45

Leu Asn Thr Thr Leu Phe Leu Asn Gly Tyr Tyr Pro Thr Ser Gly Ser

50 55 60

Thr Tyr Arg Asn Met Ala Leu Lys Gly Thr Asp Leu Leu Ser Thr Leu

65 70 75 80

Trp Phe Lys Pro Pro Phe Leu Ser Asp Phe Ile Asn Gly Ile Phe Ala

85 90 95

Lys Val Lys Asn Thr Lys Val Phe Lys Asp Gly Val Met Tyr Ser Glu

100 105 110

Phe Pro Ala Ile Thr Ile Gly Ser Thr Phe Val Asn Thr Ser Tyr Ser

115 120 125

Val Val Val Gln Pro Arg Thr Ile Asn Ser Thr Gln Asp Gly Val Asn

130 135 140

Lys Leu Gln Gly Leu Leu Glu Val Ser Val Cys Gln Tyr Asn Met Cys

145 150 155 160

Glu Tyr Pro His Thr Ile Cys His Pro Asn Leu Gly Asn His Phe Lys

165 170 175

Glu Leu Trp His Tyr Asp Thr Gly Val Val Ser Cys Leu Tyr Lys Arg

180 185 190

Asn Phe Thr Tyr Asp Val Asn Ala Thr Tyr Leu Tyr Phe His Phe Tyr

195 200 205

Gln Glu Gly Gly Thr Phe Tyr Ala Tyr Phe Thr Asp Thr Gly Phe Val

210 215 220

Thr Lys Phe Leu Phe Asn Val Tyr Leu Gly Met Ala Leu Ser His Tyr

225 230 235 240

Tyr Val Met Pro Leu Thr Cys Ile Arg Arg Pro Lys Asp Gly Phe Ser

245 250 255

Leu Glu Tyr Trp Val Thr Pro Leu Thr Pro Arg Gln Tyr Leu Leu Ala

260 265 270

Phe Asn Gln Asp Gly Ile Ile Phe Asn Ala Val Asp Cys Met Ser Asp

275 280 285

Phe Met Ser Glu Ile Lys Cys Lys Thr Gln Ser Ile Ala Pro Pro Thr

290 295 300

Gly Val Tyr Glu Leu Asn Gly Tyr Thr Val Gln Pro Val Ala Asp Val

305 310 315 320

Tyr Arg Arg Lys Pro Asp Leu Pro Asn Cys Asn Ile Glu Ala Trp Leu

325 330 335

Asn Asp Lys Ser Val Pro Ser Pro Leu Asn Trp Glu Arg Lys Thr Phe

340 345 350

Ser Asn Cys Asn Phe Asn Met Ser Ser Leu Met Ser Phe Ile Gln Ala

355 360 365

Asp Ser Phe Thr Cys Asn Asn Ile Asp Ala Ala Lys Ile Tyr Gly Met

370 375 380

Cys Phe Ser Ser Ile Thr Ile Asp Lys Phe Ala Ile Pro Asn Arg Arg

385 390 395 400

Lys Val Asp Leu Gln Leu Gly Asn Leu Gly Tyr Leu Gln Ser Ser Asn

405 410 415

Tyr Arg Ile Asp Thr Thr Ala Thr Ser Cys Gln Leu Tyr Tyr Asn Leu

420 425 430

Pro Ala Ala Asn Val Ser Val Ser Arg Phe Asn Pro Ser Thr Trp Asn

435 440 445

Lys Arg Phe Gly Phe Ile Glu Asp Ser Val Phe Val Pro Gln Pro Thr

450 455 460

Gly Val Phe Thr Asn His Ser Val Val Tyr Ala Gln His Cys Phe Lys

465 470 475 480

Ala Pro Lys Asn Phe Cys Pro Cys Ser Ser Cys Ser Cys Pro Gly Lys

485 490 495

Asn Asn Gly Ile Gly Thr Cys Pro Ala Gly Thr Asn Ser Leu Thr Cys

500 505 510

Asp Asn Leu Cys Thr Leu Asp Pro Ile Thr Leu Lys Ala Pro Asp Thr

515 520 525

Tyr Lys Cys Pro Gln Ser Lys Ser Leu Val Gly Ile Gly Glu His Cys

530 535 540

Ser Gly Leu Ala Val Lys Ser Asp Tyr Cys Gly Asn Asn Ser Cys Thr

545 550 555 560

Cys Gln Pro Gln Ala Phe Leu Gly Trp Ser Ala Asp Ser Cys Leu Gln

565 570 575

Gly Asp Lys Cys Asn Ile Phe Ala Asn Phe Ile Leu His Asp Val Asn

580 585 590

Asn Gly Leu Thr Cys Ser Thr Asp Leu Gln Lys Ala Asn Thr Glu Ile

595 600 605

Glu Leu Gly Val Cys Val Asn Tyr Asp Leu Tyr Gly Ile Ser Gly Gln

610 615 620

Gly Ile Phe Val Glu Val Asn Ala Thr Tyr Tyr Asn Ser Trp Gln Asn

625 630 635 640

Leu Leu Tyr Asp Ser Asn Gly Asn Leu Tyr Gly Phe Arg Asp Tyr Ile

645 650 655

Thr Asn Arg Thr Phe Met Ile His Ser Cys Tyr Ser Gly Arg Val Ser

660 665 670

Ala Ala Tyr His Ala Asn Ser Ser Glu Pro Ala Leu Leu Phe Arg Asn

675 680 685

Ile Lys Cys Asn Tyr Val Phe Asn Asn Ser Leu Thr Arg Gln Leu Gln

690 695 700

Pro Ile Asn Tyr Ser Phe Asp Ser Tyr Leu Gly Cys Val Val Asn Ala

705 710 715 720

Tyr Asn Ser Thr Ala Ile Ser Val Gln Thr Cys Asp Leu Thr Val Gly

725 730 735

Ser Gly Tyr Cys Val Asp Tyr Ser Lys Asn Arg Arg Ser Arg Arg Ala

740 745 750

Ile Thr Thr Gly Tyr Arg Phe Thr Asn Phe Glu Pro Phe Thr Val Asn

755 760 765

Ser Val Asn Asp Ser Leu Glu Pro Val Gly Gly Leu Tyr Glu Ile Gln

770 775 780

Ile Pro Ser Glu Phe Thr Ile Gly Asn Met Glu Glu Phe Ile Gln Thr

785 790 795 800

Ser Ser Pro Lys Val Thr Ile Asp Cys Ala Ala Phe Val Cys Gly Asp

805 810 815

Tyr Ala Ala Cys Lys Leu Gln Leu Val Glu Tyr Gly Ser Phe Cys Asp

820 825 830

Asn Ile Asn Ala Ile Leu Thr Glu Val Asn Glu Leu Leu Asp Thr Thr

835 840 845

Gln Leu Gln Val Ala Asn Ser Leu Met Asn Gly Val Thr Leu Ser Thr

850 855 860

Lys Leu Lys Asp Gly Val Asn Phe Asn Val Asp Asp Ile Asn Phe Ser

865 870 875 880

Pro Val Leu Gly Cys Leu Gly Ser Glu Cys Ser Lys Ala Ser Ser Arg

885 890 895

Ser Ala Ile Glu Asp Leu Leu Phe Asp Lys Val Lys Leu Ser Asp Val

900 905 910

Gly Phe Val Glu Ala Tyr Asn Asn Cys Thr Gly Gly Ala Glu Ile Arg
915 920 925
Asp Leu Ile Cys Val Gln Ser Tyr Lys Gly Ile Lys Val Leu Pro Pro

930 935 940
Leu Leu Ser Glu Asn Gln Ile Ser Gly Tyr Thr Leu Ala Ala Thr Ser
945 950 955 960
Ala Ser Leu Phe Pro Pro Trp Thr Ala Ala Ala Gly Val Pro Phe Tyr
965 970 975
Leu Asn Val Gln Tyr Arg Ile Asn Gly Leu Gly Val Thr Met Asp Val
980 985 990
Leu Ser Gln Asn Gln Lys Leu Ile Ala Asn Ala Phe Asn Asn Ala Leu

995 1000 1005
His Ala Ile Gln Gln Gly Phe Asp Ala Thr Asn Ser Ala Leu Val
1010 1015 1020
Lys Ile Gln Ala Val Val Asn Ala Asn Ala Glu Ala Leu Asn Asn
1025 1030 1035
Leu Leu Gln Gln Leu Ser Asn Arg Phe Gly Ala Ile Ser Ala Ser
1040 1045 1050
Leu Gln Glu Ile Leu Ser Arg Leu Asp Ala Leu Glu Ala Glu Ala

1055 1060 1065
Gln Ile Asp Arg Leu Ile Asn Gly Arg Leu Thr Ala Leu Asn Ala
1070 1075 1080
Tyr Val Ser Gln Gln Leu Ser Asp Ser Thr Leu Val Lys Phe Ser
1085 1090 1095
Ala Ala Gln Ala Met Glu Lys Val Asn Glu Cys Val Lys Ser Gln
1100 1105 1110
Ser Ser Arg Ile Asn Phe Cys Gly Asn Gly Asn His Ile Ile Ser

1115 1120 1125
Leu Val Gln Asn Ala Pro Tyr Gly Leu Tyr Phe Ile His Phe Asn
1130 1135 1140
Tyr Val Pro Thr Lys Tyr Val Thr Ala Lys Val Ser Pro Gly Leu

1145 1150 1155

Cys Ile Ala Gly Asn Arg Gly Ile Ala Pro Lys Ser Gly Tyr Phe

1160 1165 1170

Val Asn Val Asn Asn Thr Trp Met Tyr Thr Gly Ser Gly Tyr Tyr

1175 1180 1185

Tyr Pro Glu Pro Ile Thr Glu Asn Asn Val Val Val Met Ser Thr

1190 1195 1200

Cys Ala Val Asn Tyr Thr Lys Ala Pro Tyr Val Met Leu Asn Thr

1205 1210 1215

Ser Ile Pro Asn Leu Pro Asp Phe Lys Glu Glu Leu Asp Gln Trp

1220 1225 1230

Phe Lys Asn Gln Thr Ser Val Ala Pro Asp Leu Ser Leu Asp Tyr

1235 1240 1245

Ile Asn Val Thr Phe Leu Asp Leu Gln Val Glu Met Asn Arg Leu

1250 1255 1260

Gln Glu Ala Ile Lys Val Leu Asn His Ser Tyr Ile Asn Leu Lys

1265 1270 1275

Asp Ile Gly Thr Tyr Glu Tyr Tyr Val Lys Trp Pro Trp Tyr Val

1280 1285 1290

Trp Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val Leu Leu

1295 1300 1305

Phe Phe Ile Cys Cys Cys Thr Gly Cys Gly Thr Ser Cys Phe Lys

1310 1315 1320

Lys Cys Gly Gly Cys Cys Asp Asp Tyr Thr Gly Tyr Gln Glu Leu

1325 1330 1335

Val Ile Lys Thr Ser His Asp Asp

1340 1345

<210> 161

<211> 1157

<212> PRT

<213> 229E Coronavirus

<400> 161

Gln Thr Thr Asn Gly Leu Asn Thr Ser Tyr Ser Val Cys Asn Gly Cys

1 5 10 15

Val Gly Tyr Ser Glu Asn Val Phe Ala Val Glu Ser Gly Gly Tyr Ile

20 25 30

Pro Ser Asp Phe Ala Phe Asn Asn Trp Phe Leu Leu Thr Asn Thr Ser

35 40 45

Ser Val Val Asp Gly Val Val Arg Ser Phe Gln Pro Leu Leu Leu Asn

50 55 60

Cys Leu Trp Ser Val Ser Gly Leu Arg Phe Thr Thr Gly Phe Val Tyr

65 70 75 80

Phe Asn Gly Thr Gly Arg Gly Asp Cys Lys Gly Phe Ser Ser Asp Val

85 90 95

Leu Ser Asp Val Ile Arg Tyr Asn Leu Asn Phe Glu Glu Asn Leu Arg

100 105 110

Arg Gly Thr Ile Leu Phe Lys Thr Ser Tyr Gly Val Val Val Phe Tyr

115 120 125

Cys Thr Asn Asn Thr Leu Val Ser Gly Asp Ala His Ile Pro Phe Gly

130 135 140

Thr Val Leu Gly Asn Phe Tyr Cys Phe Val Asn Thr Thr Ile Gly Asn

145 150 155 160

Glu Thr Thr Ser Ala Phe Val Gly Ala Leu Pro Lys Thr Val Arg Glu

165 170 175

Phe Val Ile Ser Arg Thr Gly His Phe Tyr Ile Asn Gly Tyr Arg Tyr

180 185 190

Phe Thr Leu Gly Asn Val Glu Ala Val Asn Phe Asn Val Thr Thr Ala

195 200 205

Glu Thr Thr Asp Phe Cys Thr Val Ala Leu Ala Ser Tyr Ala Asp Val

210 215 220

Leu Val Asn Val Ser Gln Thr Ser Ile Ala Asn Ile Ile Tyr Cys Asn

225 230 235 240

Ser Val Ile Asn Arg Leu Arg Cys Asp Gln Leu Ser Phe Asp Val Pro

245 250 255

Asp Gly Phe Tyr Ser Thr Ser Pro Ile Gln Ser Val Glu Leu Pro Val

260 265 270

Ser Ile Val Ser Leu Pro Val Tyr His Lys His Thr Phe Ile Val Leu

275 280 285

Tyr Val Asp Phe Lys Pro Gln Ser Gly Gly Gly Lys Cys Phe Asn Cys

290 295 300

Tyr Pro Ala Gly Val Asn Ile Thr Leu Ala Asn Phe Asn Glu Thr Lys

305 310 315 320

Gly Pro Leu Cys Val Asp Thr Ser His Phe Thr Thr Lys Tyr Val Ala

325 330 335

Val Tyr Ala Asn Val Gly Arg Trp Ser Ala Ser Ile Asn Thr Gly Asn

340 345 350

Cys Pro Phe Ser Phe Gly Lys Val Asn Asn Phe Val Lys Phe Gly Ser

355 360 365

Val Cys Phe Ser Leu Lys Asp Ile Pro Gly Gly Cys Ala Met Pro Ile

370 375 380

Val Ala Asn Trp Ala Tyr Ser Lys Tyr Tyr Thr Ile Gly Ser Leu Tyr

385 390 395 400

Val Ser Trp Ser Asp Gly Asp Gly Ile Thr Gly Val Pro Gln Pro Val

405 410 415

Glu Gly Val Ser Ser Phe Met Asn Val Thr Leu Asp Lys Cys Thr Lys

420 425 430

Tyr Asn Ile Tyr Asp Val Ser Gly Val Gly Val Ile Arg Val Ser Asn

435 440 445

Asp Thr Phe Leu Asn Gly Ile Thr Tyr Thr Ser Thr Ser Gly Asn Leu

450 455 460

Leu Gly Phe Lys Asp Val Thr Lys Gly Thr Ile Tyr Ser Ile Thr Pro

465 470 475 480

Cys Asn Pro Pro Asp Gln Leu Val Val Tyr Gln Gln Ala Val Val Gly

485 490 495

Ala Met Leu Ser Glu Asn Phe Thr Ser Tyr Gly Phe Ser Asn Val Val
500 505 510
Glu Leu Pro Lys Phe Phe Tyr Ala Ser Asn Gly Thr Tyr Asn Cys Thr

515 520 525
Asp Ala Val Leu Thr Tyr Ser Ser Phe Gly Val Cys Ala Asp Gly Ser
530 535 540
Ile Ile Ala Val Gln Pro Arg Asn Val Ser Tyr Asp Ser Val Ser Ala
545 550 555 560
Ile Val Thr Ala Asn Leu Ser Ile Pro Ser Asn Trp Thr Thr Ser Val
565 570 575
Gln Val Glu Tyr Leu Gln Ile Thr Ser Thr Pro Ile Val Val Asp Cys

580 585 590
Ser Thr Tyr Val Cys Asn Gly Asn Val Arg Cys Val Glu Leu Leu Lys
595 600 605
Gln Tyr Thr Ser Ala Cys Lys Thr Ile Glu Asp Ala Leu Arg Asn Ser
610 615 620
Ala Arg Leu Glu Ser Ala Asp Val Ser Glu Met Leu Thr Phe Asp Lys
625 630 635 640
Lys Ala Phe Thr Leu Ala Asn Val Ser Ser Phe Gly Asp Tyr Asn Leu

645 650 655
Ser Ser Val Ile Pro Ser Leu Pro Thr Ser Gly Ser Arg Val Ala Gly
660 665 670
Arg Ser Ala Ile Glu Asp Ile Leu Phe Ser Lys Leu Val Thr Ser Gly
675 680 685
Leu Gly Thr Val Asp Ala Asp Tyr Lys Lys Cys Thr Lys Gly Leu Ser
690 695 700
Ile Ala Asp Leu Ala Cys Ala Gln Tyr Tyr Asn Gly Ile Met Val Leu

705 710 715 720
Pro Gly Val Ala Asp Ala Glu Arg Met Ala Met Tyr Thr Gly Ser Leu
725 730 735
Ile Gly Gly Ile Ala Leu Gly Gly Leu Thr Ser Ala Val Ser Ile Pro

740 745 750

Phe Ser Leu Ala Ile Gln Ala Arg Leu Asn Tyr Val Ala Leu Gln Thr

755 760 765

Asp Val Leu Gln Glu Asn Gln Lys Ile Leu Ala Ala Ser Phe Asn Lys

770 775 780

Ala Met Thr Asn Ile Val Asp Ala Phe Thr Gly Val Asn Asp Ala Ile

785 790 795 800

Thr Gln Thr Ser Gln Ala Leu Gln Thr Val Ala Thr Ala Leu Asn Lys

805 810 815

Ile Gln Asp Val Val Asn Gln Gln Gly Asn Ser Leu Asn His Leu Thr

820 825 830

Ser Gln Leu Arg Gln Asn Phe Gln Ala Ile Ser Ser Ser Ile Gln Ala

835 840 845

Ile Tyr Asp Arg Leu Asp Thr Ile Gln Ala Asp Gln Gln Val Asp Arg

850 855 860

Leu Ile Thr Gly Arg Leu Ala Ala Leu Asn Val Phe Val Ser His Thr

865 870 875 880

Leu Thr Lys Tyr Thr Glu Val Arg Ala Ser Arg Gln Leu Ala Gln Gln

885 890 895

Lys Val Asn Glu Cys Val Lys Ser Gln Ser Lys Arg Tyr Gly Phe Cys

900 905 910

Gly Asn Gly Thr His Ile Phe Ser Ile Val Asn Ala Ala Pro Glu Gly

915 920 925

Leu Val Phe Leu His Thr Val Leu Leu Pro Thr Gln Tyr Lys Asp Val

930 935 940

Glu Ala Trp Ser Gly Leu Cys Val Asp Gly Thr Asn Gly Tyr Val Leu

945 950 955 960

Arg Gln Pro Asn Leu Ala Leu Tyr Lys Glu Gly Asn Tyr Tyr Arg Ile

965 970 975

Thr Ser Arg Ile Met Phe Glu Pro Arg Ile Pro Thr Met Ala Asp Phe

980 985 990

Val Gln Ile Glu Asn Cys Asn Val Thr Phe Val Asn Ile Ser Arg Ser

995 1000 1005

Glu Leu Gln Thr Ile Val Pro Glu Tyr Ile Asp Val Asn Lys Thr

1010 1015 1020

Leu Gln Glu Leu Ser Tyr Lys Leu Pro Asn Tyr Thr Val Pro Asp

1025 1030 1035

Leu Val Val Glu Gln Tyr Asn Gln Thr Ile Leu Asn Leu Thr Ser

1040 1045 1050

Glu Ile Ser Thr Leu Glu Asn Lys Ser Ala Glu Leu Asn Tyr Thr

1055 1060 1065

Val Gln Lys Leu Gln Thr Leu Ile Asp Asn Ile Asn Ser Thr Leu

1070 1075 1080

Val Asp Leu Lys Trp Leu Asn Arg Val Glu Thr Tyr Ile Lys Trp

1085 1090 1095

Pro Trp Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe Val

1100 1105 1110

Val Ser Met Leu Leu Leu Cys Cys Cys Ser Thr Gly Cys Cys Gly

1115 1120 1125

Phe Phe Ser Cys Phe Ala Ser Ser Ile Arg Gly Cys Cys Glu Ser

1130 1135 1140

Thr Lys Leu Pro Tyr Tyr Asp Val Glu Lys Ile His Ile Gln

1145 1150 1155

<210> 162

<211> 56

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-OC43-COV wtTMCT-AA

<400> 162

Trp Tyr Val Trp Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val

1 5 10 15

Leu Leu Phe Phe Ile Cys Cys Cys Thr Gly Cys Gly Thr Ser Cys Phe

20 25 30

Lys Lys Cys Gly Gly Cys Cys Asp Asp Tyr Thr Gly Tyr Gln Glu Leu

35 40 45

Val Ile Lys Thr Ser His Asp Asp

50 55

<210> 163

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-OC43-COV H5iTMCT-AA

<400> 163

Trp Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala

1 5 10 15

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly

20 25 30

Ser Leu Gln Cys Arg Ile Cys Ile

35 40

<210> 164

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-OC43-COV H5iCT-AA

<400> 164

Trp Tyr Val Trp Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val

1 5 10 15

Leu Leu Phe Phe Ile Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln

20 25 30

Cys Arg Ile Cys Ile

35

<210> 165

<211> 34

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-OC43-COV H5iCT(V4)-AA

<400> 165

Trp Tyr Val Trp Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val

1 5 10 15

Leu Leu Phe Phe Ile Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile

20 25 30

Cys Ile

<210> 166

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-OC43-COV H1cCT-AA

<400> 166

Trp Tyr Val Trp Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val

1 5 10 15

Leu Leu Phe Phe Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln

20 25 30

Cys Arg Ile Cys Ile

35

<210> 167

<211> 58

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-229E-wtTMCT-AA

<400> 167

Trp Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe Val Val Ser

1 5 10 15

Met Leu Leu Leu Cys Cys Cys Ser Thr Gly Cys Cys Gly Phe Phe Ser

20 25 30

Cys Phe Ala Ser Ser Ile Arg Gly Cys Cys Glu Ser Thr Lys Leu Pro

35 40 45

Tyr Tyr Asp Val Glu Lys Ile His Ile Gln

50 55

<210> 168

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-229E-H5iTMCT-AA

<400> 168

Trp Trp Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
1 5 10 15

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
20 25 30

Ser Leu Gln Cys Arg Ile Cys Ile
35 40

<210> 169

<211> 36

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-229E-H5iCT-AA

<400> 169

Trp Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe Val Val Ser

1 5 10 15

Met Leu Leu Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
20 25 30

Arg Ile Cys Ile
35

<210> 170

<211> 33

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-229E-H5iCT(V4)-AA

<400> 170

Trp Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe Val Val Ser
1 5 10 15

Met Leu Leu Leu Cys Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys

20 25 30

Ile

<210> 171

<211> 35

<212> PRT

<213> Artificial Sequence

<220><223> TMCT region of modified PDI-229E-H1cCT-AA

<400> 171

Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe Val Val Ser Met

1 5 10 15

Leu Leu Leu Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg

20 25 30

Ile Cys Ile

35

<210

> 172

<211> 34

<212> PRT

<213> Artificial Sequence

<220><223> TM/CT Region of Modified OC43-CoV S protein with intervening

peptide sequence Xn

<220><221> MISC_FEATURE

<222> (22)..(22)

<223> Xaa may be any combination of 0 to 10 amino acids

<220><221> misc_feature

<222> (28)..(29)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (31)..(31)

<223> Xaa can be any naturally occurring amino acid

<400> 172

Trp Tyr Val Trp Leu Leu Ile Cys Leu Ala Gly Val Ala Met Leu Val

1 5 10 15

Leu Leu Phe Phe Ile Xaa Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa Ile

20 25 30

Cys Ile

<210> 173

<211> 33

<212> PRT

<213> Artificial Sequence

<220><223> TM/CT Region of Modified OC43-CoV S protein with intervening peptide sequence Xn

<220><221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa may be any combination of 0 to 10 amino acids

<220><221> misc_feature

<222> (27)..(28)

<223> Xaa can be any naturally occurring amino acid

<220><221> misc_feature

<222> (30)..(30)

<223> Xaa can be any naturally occurring amino acid

<400> 173

Trp Trp Val Trp Leu Cys Ile Ser Val Val Leu Ile Phe Val Val Ser

1 5 10 15

Met Leu Leu Leu Xaa Cys Ser Asn Gly Ser Xaa Xaa Cys Xaa Ile Cys

20 25 30

Ile