



등록특허 10-2748840



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2025년01월02일

(11) 등록번호 10-2748840

(24) 등록일자 2024년12월26일

- (51) 국제특허분류(Int. Cl.)
C07K 14/005 (2006.01) *A61K 39/145* (2006.01)
A61P 31/16 (2006.01) *C12N 15/70* (2006.01)
- (52) CPC특허분류
C07K 14/005 (2013.01)
A61K 39/145 (2013.01)
- (21) 출원번호 10-2018-7025483
- (22) 출원일자(국제) 2017년02월02일
 심사청구일자 2022년02월03일
- (85) 번역문제출일자 2018년09월03일
- (65) 공개번호 10-2018-0115717
- (43) 공개일자 2018년10월23일
- (86) 국제출원번호 PCT/US2017/016251
- (87) 국제공개번호 WO 2017/136575
 국제공개일자 2017년08월10일
- (30) 우선권주장
 62/290,894 2016년02월03일 미국(US)
- (56) 선행기술조사문헌
 KR1019930013112 A*
 Ekiert, D. C., et al., Science, 2011, vol. 333, no. 6044, pp. 843-850(2011.08.12.) 1부.*
 *는 심사관에 의하여 인용된 문헌

- (73) 특허권자
 씨지 디스커버리, 인코포레이티드
 미국 캘리포니아 92130 샌디에고 페어포트 웨이 4838
- (72) 발명자
 루오, 춘
 미국 캘리포니아 92130 샌디에고 페어포트 웨이 4838 씨지 디스커버리 인코포레이티드 내
- (74) 대리인
 박장원

전체 청구항 수 : 총 18 항

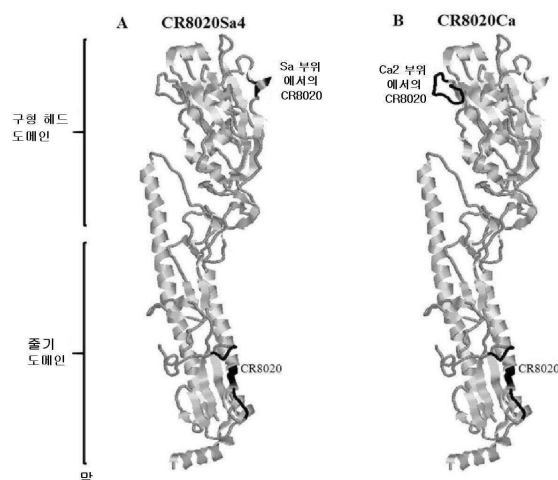
심사관 : 강덕희

(54) 발명의 명칭 이종 에피토프 및/또는 성숙 절단 부위를 갖는 인플루엔자 헤마글루티닌 조성물

(57) 요약

변형된 면역우성 영역들 및 변형된 성숙 절단 부위들을 갖는 것들을 비롯한, 헤마글루티닌 (HA) 단백질의 변형된 형태, 및 이들을 함유하는 바이러스 및 바이러스-유사 입자들.

대표도 - 도6



(52) CPC특허분류

A61P 31/16 (2018.01)

C12N 15/70 (2013.01)

C12N 2760/16134 (2013.01)

C12N 2760/16234 (2013.01)

명세서

청구범위

청구항 1

다음을 포함하는 인플루엔자 바이러스의 변형된 헤마글루티닌 (HA) 단백질로서:

(i) 구형 헤드 도메인 및 줄기 도메인으로서, 상기 HA 단백질의 구형 헤드 도메인의 적어도 하나의 면역우성 영역은 인플루엔자 줄기 에피토프를 함유하고, 상기 인플루엔자 줄기 에피토프는 동일한 HA의 하나 이상의 줄기 펩타이드를 포함하고, 인플루엔자 바이러스에 대한 광범위 중화 항체 (broadly neutralizing antibody, bnAb)에 의하여 인식되는 것인, 구형 헤드 도메인 및 줄기 도메인, 및

(ii) 변형된 성숙 절단 부위 (maturation cleavage site, MCS)로서, 상기 MCS는 임의의 동물 프로테아제에 의해서는 절단가능하지 않지만 비-동물 프로테아제에 의해서는 절단가능하도록 변형되고, 절단시, HA2의 N-말단의 글리신이 유지되는 것인, 변형된 MCS,

여기서, 인플루엔자 HA 줄기 에피토프를 포함하는 구형 헤드 도메인의 상기 면역우성 영역은, 각각 SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:58, SEQ ID NO:60, 또는 SEQ ID NO:62 중에서 선택되는 핵산에 의하여 인코딩되는, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:59, 또는 SEQ ID NO:61의 아미노산 서열을 갖는 것인, 변형된 HA 단백질.

청구항 2

제1항에 있어서, 구형 헤드 도메인의 적어도 하나의 면역우성 영역은:

상기 도메인으로의 삽입에 의해 상기 인플루엔자 줄기 에피토프를 함유하거나, 또는 상기 도메인을 상기 인플루엔자 줄기 에피토프로 대체함으로써 상기 줄기 에피토프를 함유하는 것인, 변형된 HA 단백질.

청구항 3

제1항에 있어서, 구형 헤드 도메인의 적어도 2개, 적어도 3개, 적어도 4개, 또는 적어도 5개의 면역우성 영역에 함유된 하나 이상의 줄기 에피토프는 구별되는 (distinct) 것인, 변형된 HA 단백질.

청구항 4

제1항에 있어서, 구형 헤드 도메인의 적어도 2개, 적어도 3개, 적어도 4개, 또는 적어도 5개의 면역우성 영역에 함유된 하나 이상의 줄기 에피토프는 동일한 (identical) 것인, 변형된 HA 단백질.

청구항 5

제1항에 있어서, 상기 변형된 HA 단백질은 살아있거나 약화된 바이러스 내에 또는 VLP 내에 함유되는 것인, 변형된 HA 단백질.

청구항 6

제1항에 기재된 변형된 HA 단백질을 인코딩하는 폴리뉴클레오타이드.

청구항 7

재조합 인플루엔자 바이러스로서,

상기 재조합 인플루엔자 바이러스는 제6항에 기재된 폴리뉴클레오타이드를 포함함으로써, 변형된 HA 단백질이 바이러스의 표면에 발현되는, 재조합 인플루엔자 바이러스.

청구항 8

제6항에 기재된 폴리뉴클레오타이드를 포함하는 단리된 재조합 숙주 세포.

청구항 9

인플루엔자 바이러스 감염을 예방하기 위한 용도의, 제1항에 기재된 변형된 HA 단백질을 포함하는 면역원성 조성물로서,

상기 면역원성 조성물은 인플루엔자 바이러스에 대한 항체를 생성하는 것인, 면역원성 조성물.

청구항 10

제9항에 있어서, 조성물은 살아있거나 약화된 바이러스 또는VLP를 포함하는 것인, 면역원성 조성물.

청구항 11

인플루엔자 바이러스 감염을 예방하기 위한 용도의, 제6항에 기재된 폴리뉴클레오타이드를 포함하는 면역원성 조성물로서,

상기 면역원성 조성물은 인플루엔자 바이러스에 대한 항체를 생성하는 것인, 면역원성 조성물.

청구항 12

변형된 인플루엔자 바이러스로서, 상기 바이러스는 변형된 인플루엔자 헤마글루티닌 (HA) 단백질을 포함하고, 상기 변형된 HA 단백질은 다음을 포함하고:

(i) 구형 헤드 도메인 및 줄기 도메인으로서, 상기 HA 단백질의 구형 헤드 도메인의 적어도 하나의 면역우성 영역은 인플루엔자 줄기 에피토프를 함유하고, 상기 인플루엔자 줄기 에피토프는 동일한 HA의 하나 이상의 줄기 펩타이드를 포함하고, 인플루엔자 바이러스에 대한 광범위 중화 항체 (broadly neutralizing antibody, bnAb) 에 의하여 인식되는 것인, 구형 헤드 도메인 및 줄기 도메인, 및

(ii) 변형된 성숙 절단 부위 (maturation cleavage site, MCS)로서, 상기 MCS는 임의의 동물 프로테아제에 의해서는 절단가능하지 않지만 비-동물 프로테아제에 의해서는 절단가능하도록 변형되고, 절단시, HA2의 N-말단의 글리신이 유지되는 것인, 변형된 MCS,

여기서, 인플루엔자 HA 줄기 에피토프를 포함하는 구형 헤드 도메인의 상기 면역우성 영역은, 각각 SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:58, SEQ ID NO:60, 또는 SEQ ID NO:62 중에서 선택되는 핵산에 의하여 인코딩되는, SEQ ID NO:24, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:59, 또는 SEQ ID NO:61의 아미노산 서열을 갖는 것인, 변형된 인플루엔자 바이러스.

청구항 13

제12항에 있어서,

MCS는 담배 식각 바이러스 (TEV) 프로테아제에 의해 절단가능하도록 변형된 것이고,

전체 MCS가 TEV 절단 부위에 의해 대체되거나, MCS의 일부 잔기가 TEV 절단 부위에 의해 대체되거나, 또는 분할 부위에 대한 아르기닌 N-말단만이 TEV 절단 부위에 의해 대체되는 것인, 변형된 인플루엔자 바이러스.

청구항 14

제12항 또는 제13항에 기재된 변형된 인플루엔자 바이러스를 포함하는 백신.

청구항 15

변형된 인플루엔자 HA 단백질을 생산하는 방법으로서, 상기 방법은:

(i) 변형된 HA 단백질의 발현에 적합한 조건 하에, 상기 변형된 HA 단백질 아미노산 서열을 인코딩하는 핵산 발

현 박터를 포함하는 배양 배지의 존재 하에서 단리된 재조합 세포를 배양하는 단계로서,
 상기 변형된 HA 단백질은 구형 헤드 도메인 및 변형된 성숙 절단 부위 (MCS)을 포함하고,
 여기서 상기 HA 단백질의 구형 헤드 도메인의 적어도 하나의 면역우성 영역은 인플루엔자 줄기 에피토프를 함유
 하고, 여기서 상기 인플루엔자 줄기 에피토프는 동일한 HA의 하나 이상의 줄기 펩타이드를 포함하고 인플루엔자
 바이러스에 대한 광범위 중화 항체 (broadly neutralizing antibody, bnAb)에 의하여 인식되고,
 여기서 상기 MCS는 임의의 동물 프로테아제에 의해서는 절단가능하지 않지만 비-동물 프로테아제에 의해서는 절
 단가능하도록 변형되고, 절단시, HA2의 N-말단의 글리신이 유지되는 것인, 단계; 및
 (ii) 단리된 세포 및/또는 배양 배지로부터 변형된 HA 단백질을 단리하는 단계
 를 포함하는, 방법.

청구항 16

대상자에서 하나 이상의 인플루엔자 균주 또는 분리주에 대한 면역 반응을 유도하기 위한, 제15항에 기재된 방
 법에 의하여 생산된 변형된 HA 단백질의 치료적 유효량을 포함하는 의약 조성물.

청구항 17

하나 이상의 인플루엔자 균주 또는 분리주에 의하여 대상자의 감염을 억제하기 위한, 제15항에 기재된 방법
 에 의하여 생산된 변형된 HA 단백질의 치료적 유효량을 포함하는 의약 조성물.

청구항 18

제1항에 있어서,

MCS는 담배 식각 바이러스 (TEV) 프로테아제에 의하여 절단가능하도록 변형되고,

전체 MCS가 TEV 절단 부위로 대체되거나, MCS의 몇몇 잔기가 TEV 절단 부위로 대체되거나, 또는 분리 부위
 (scission site)에 대한 아르기닌 N-말단만이 TEV 절단 부위에 의해 대체되는 것인, 변형된 HA 단백질.

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호-참조

[0002] [1] 이 출원은 2016년 2월 3일자로 "이종 에피토프 및/또는 변형된 성숙 절단 부위를 갖는 인플루엔자 헤마글루티닌의 조성물 및 이의 사용방법"이라는 명칭으로 출원된 미국 가특허출원 제62/290,894호에 기초한 우선권 주장 출원으로서, 상기 출원의 내용은 그 전체가 본 발명에 참조 병합된다.

[0003] 서열목록의 참조 병합

[0004] [2] 본원은 전자 포맷 형태의 서열목록과 함께 출원되었다. 서열목록은 2017년 1월 30일에 758252000140SeqList.TXT라는 파일명으로 생성되었으며 크기는 173,061 바이트이다. 서열목록의 전자 포맷에 담긴 정보는 그 전체가 본 발명에 참조 병합된다.

[0005] 기술분야

[0006] [3] 본 발명은 면역학 및 바이러스학 분야에 관한 것이다. 더욱 구체적으로 본 발명은 변형된 이종의 (heterologous) 에피토프 및/또는 변형된 성숙한 절단 부위를 갖는 제조함 인플루엔자 헤마글루티닌 (HA) 단백질의 조성물에 관한 것이다.

[0007] 다음의 약어들이 이용될 수 있다. HA, 헤마글루티닌; NA, 뉴라미니다제; NAbs, 중화 항체; bNAbs, 광범한 중화 항체; TEV, 담배식각 바이러스; DNA, 데옥시리보핵산; cDNA, 상보적 DNA; RNA, 리보핵산; kb, 킬로베이스; kDa, 킬로 달톤; CHO 세포, 차이니즈 햄스터 난소 세포; HEK 293 세포, 인간 배아 신장 293 세포; VLP, 바이러

스-유사 입자; BEVS, 배콜로바이러스 발현 벡터 시스템; AcNPV, 오토그래프 캘리포니카뉴클리어 폴리헤드로시스 바이러스; BmNPV, 봄빅스 모리 핵 폴리헤드로시스 바이러스; TIPS, titerless infected-cells preservation and scale-up; BIIC, 배콜로바이러스 감염된 곤충 세포; PCR, 폴리머라제 연쇄 반응; FBS, 소태아 혈청; PDB, 월드와이드웹 상의 (rcsb.org/pdb/home/home.do)에서 입수가 가능한 단백질 데이터 은행; MOI, 감염 다중도; MCS, 성숙 절단 부위; HPAI, 고병원성 조류 인플루엔자; HI, HA 억제; WT, 야생형; TIV, 불활성화된 3가 인플루엔자 백신; PBS, 포스포이트 완충 염수.

배경 기술

- [0008] [4] 흔히 "독감(flu)"라고 알려진 인플루엔자는 인간 또는 동물의 인플루엔자 바이러스의 감염에 기인하는 감염성 호흡기 질환이다. 인플루엔자의 창궐은 매년 추운 계절에 발생하며, 이 시기는 흔히 "독감의 계절"이라 칭해진다. 비록 대부분의 인플루엔자 감염 증상은 가볍지만, 연례적으로 일어나는 이 전염병에서 전세계적으로 350만명은 심각한 증상을 겪게 되고 최대 500,000명은 사망에 이르고 있다 (온라인 who.int/mediacentre/factsheets/fs211/en/에 따른 World Health Organization: Influenza (Seasonal) 참조). 매 세기마다, 수 차례씩 상당한 비율의 인류가 인플루엔자 바이러스에 감염되어 전세계적으로 인플루엔자가 만연하게 되어 심각한 위병률과 치사율을 나타낸다. 전세계 인구의 3 내지 5 퍼센트를 죽인 1918년의 스페인 독감은 역사상 가장 치명적인 전염병으로 기록되었다.
- [0009] 줄기 도메인 반응성 항체를 유도하기 위한 현행 전략
- [0010] [5] 모든 인플루엔자 바이러스에 대해 광범위 중화 항체(broadly neutralizing antibodies: bNAbs)를 유도해내는 것은 독감 백신 설계의 궁극적인 목표가 되어왔다. 단리된 bNAbs의 대다수는 HA 줄기 도메인에서 보존된 에피토프를 인식한다. 그러나, 천연 인플루엔자 바이러스 감염 또는 현재의 독감 백신에 의해 유도되는 항체들은 대부분 HA 구형 헤드 도메인(globular head domain) 내의 수용체 결합 부위를 둘러싸는 항원 부위를 인식하는데 이는 이들 항원 부위의 면역우성(immunodominance)에 기인한다. 이들 항체들은 일반적으로, 인플루엔자 바이러스 스트레인들 중에서도 이들 항원 부위의 고도의 다양성의 결과로 인해 스트레인-특이적이다. 그러므로 현재의 독감 백신은 독감 바이러스들에 대해 보편적인 면역성을 유도하지 못한다.
- [0011] [6] 헤마글루티닌(HA)은 일체형 막 당단백질이다. 이것은 인플루엔자 바이러스 지질 이중층 엔벨롭에서 가장 풍부한 단백질이다. HA의 약 500개 분자는 비리온의 표면에 존재하는 것으로 여겨진다. HA는 바이러스가 숙주 세포 내로 흡착 및 침투하는데 필요한 유일한 단백질이다. 이것은 숙주 세포 수용체에 결합하여 비리온의 지질 엔벨롭과 숙주 세포막 사이의 융합을 가능케 한다. 숙주 세포 수용체에 대한 그의 결합 능력을 통해, HA는 인플루엔자 스트레인의 숙주 범위를 결정한다.
- [0012] [7] HA는 세포질 세막에서 HAO라 명명된 단일 사슬 전구체로서 합성된다(Stevens, J., *et al.*, *Science*(2006) 312:404-410). HAO 전구체는 그의 N-말단에서 시그널 펩타이드를, 그리고 그의 C-말단에서 막 앵커 서열을 갖는다. N-말단 시그널 펩타이드는 HA가 숙주 세포막으로 전달되어 숙주 세포막 내에 앵커되는 과정에서 제거된다. HA는 세포질 세막에서 동일한 서브유닛들의 삼량체로 조립된 다음 골지 네트워크를 통해 세포 표면으로 배출된다. HAO는 특이적인 숙주 트립신-유사 프로테아제에 의해 성숙 절단 부위 (maturation cleavage site: MCS)에서 절단되어 2개의 디설파이드-결합 연결된 폴리펩타이드인 HA1 및 HA2로 이루어지는 성숙한 형태로 전환된다. HA1은 HAO의 크기가 더 큰 N-말단 부위이고 HA2는 크기가 더 작은 C-말단 부위이다. HA2는 짧은 C-말단 세포내 꼬리(intracellular tail)와 함께 막관통 영역을 가지며 HA를 막에 결합(anchors)시킨다. 각각의 성숙한 HA는 총 분자량이 약 70 kDa인데 HA1이 약 45 kDa이고 HA2가 약 25 kDa이다. 많은 인플루엔자 스트레인들로부터의 HA의 세포외 부분의 원자 구조가 탐구되어 왔다. 성숙한 HA의 구조 모델이 도 1에 도시되어 있다. 모든 HA는 HAO로서든 또는 그의 성숙한 형태로서든, 줄기에 구형 헤드(globular head)를 갖는 전체적으로 동일한 삼량체의 구조를 갖는다. 각각의 HA 단량체는 짧은 친수성 세포질 꼬리를 갖는 단일의 C-말단 막관통 영역에 의해 막에 결합된다.
- [0013] [8] 구형 헤드 도메인은 전적으로 HA1으로 만들어지며 면역우성(immunodominant) 에피토프들을 함유한다. 줄기 도메인은 대부분 HA2로 만들어지고 면역학적으로 준우성(subdominant)인 보존된 영역을 함유한다. 구형 헤드 도메인은 그 중심에 표면 루프 및 헬릭스와 함께 8-가닥의 β -시트 구조를 갖는다. 막 근위 줄기 도메인은 HA1과 HA2 양자 모두로부터의 잔기들의 삼중-코일형 코일 구조의 왼손잡이 수퍼헬릭스로 구성된다. 각각의 HA 단량체는 분자량이 약 13 kDa이거나 또는 총 HA 분자량의 19%인 총 탄수화물을 갖는 다수의 글리코실화 부위를 갖는다. 대부분의 글리코실화는 막 표면 근처의 줄기 도메인에서 일어난다. HA는 또한 세포질 꼬리상의 시스템

인 잔기의 팔미토일화에 의해 변형된다 (Veit, M., *et al.*, *J. Virol*(1991) 65:2491-2500).

- [0014] [9] 구형 헤드 도메인의 면역우성을 회피하기 위해, 면역원들은 구형 헤드 도메인 없이 HA 줄기 도메인 상에 기초하도록 설계되어 왔다 (헤드레스 HA). 이러한 시도는 적절한 3차 구조를 갖는 그러한 분자의 생산이 어렵기 때문에 크게 성공적이지 못했다 (Krammer, F. 및 Palese, P., *Curr Opin Virol*(2013) 3:521-530; Eckert, D. M., 및 Kay, M. S., *PNAS*(2010) 107:13563-13564). HA2 그 자체는 E. coli에서 수용성 형태로 발현되어 그 의 가장 안정한 융합후 저-pH 유도된 배열로 폴딩된다 (Chen, J., *et al.*, *PNAS*(1995) 92:12205-12209). HA2의 저-pH 배열을 탈안정화시키기 위하여, 설계된 돌연변이들을 통합시킴으로써, H3 HA에 기반한 또 다른 HA2 구조물(HA6)이 E. coli에서 발현되어 목적하는 중성 pH의 예비-융합 배열로 리폴딩되었다 (Bommakanti, G., *et al.*, *PNAS*(2010) 107:13701-13706). HA6는 마우스에서 고도로 면역원성이었으며 동종 A형 인플루엔자 바이러스에 의한 감염에 대해 마우스들을 보호하였다. 그러나, HA6 면역화된 마우스들로부터의 혈청은 시험관내에서 바이러스를 중화시키는데 실패했는데, 이는 구형 헤드 도메인을 인식하는 항체들의 바이러스-중화 활성만을 감지하는 분석법의 한계에 기인하는 것일 수 있다. 줄기 도메인 내에 HA1 영역과 HA2를 갖는 또 다른 "헤드레스 (headless)" HA 줄기 도메인 구조물이 설명된 바 있다 (Steel, J., *et al.*, *mBio*(2010)1:e00018-10). 이 면역원은 동종 인플루엔자 켈린지로부터 마우스를 보호하고 동일 그룹 내 이중 HA에 대해 항혈청 교차-반응을 이끌어냈다. HA6의 경우에서처럼, 이 항혈청은 시험관내에서 중화 활성을 나타내지 않았다. 이러한 설계물들은 모두 면역우성 영역을 이끌어냄으로써 단백질 최소화에 기반하는 것들이다.
- [0015] [10] 최근, 최근 구조 기반의 합리적인 설계와 구조 라이브러리의 반복적인 bNAb 선택을 통해 안정한 삼량체 H1 HA 줄기만의(stem-only) 면역원 (헤드레스 미니-HA)가 만들어졌다 (Impagliazzo, A., *et al.*, *Science*(2015) 349:1301-1306; Yassing, H., M., *et al.*, *Nat Med*(2015) 21:1065-1070; 특허출원 W02014/191435_A1). 이들 줄기만의 미니-HA 면역원들은 HA 줄기 도메인에 대해 예상된 항체들을 이끌어냈다. 이들 미니-HA 면역원들에 의해 면역화된 마우스 및 페렛들은 고도로 병원성인 H5 바이러스의 치명적인 켈린지로부터 보호되었다. 이에 더해, 미니-HA 면역원은 시노물구스 원숭이에서 H5 중화 항체를 유도하였다. 이들 미니-HA 면역원들은 특이적인 bNAbs에 대한 이들의 결합에 대해 선택되었다. 길고 반복적인 bNAb 선택 공정은 상이한 유형의 bNAbs에 대해 반복될 필요가 있다. 비록 bNAb 에피토프의 배열이 이들 미니-HA 면역원들에 있어서 보존된다 해도, 전체적인 구조는 천연 HA의 줄기 도메인과 달랐다. 이러한 구조적 변화로 인해, 이들 헤드레스 미니-HA 면역원들이 인플루엔자 바이러스 또는 인플루엔자 바이러스-유사 입자(VLPs)로 조립될 확률은 적다. 또한, 이들 헤드레스 미니-HA 면역원들은 숙주 세포 결합을 위한 수용체 결합 부위를 갖지 않는다.
- [0016] [11] 키메라 HAs는 숙주 면역 반응을 줄기 도메인에 지향시키도록 설계되었다. 세계적인 유행성 H1N1 감염에 의해 유도된 대부분의 중화 항체들은 복수종의 인플루엔자 스트레인들의 HA 줄기 도메인 및 구형 헤드 도메인에서 에피토프에 대해 광범하게 교차-반응성이었다 (Wrammert, J., *et al.*, *J. Exp Med*(2011) 208:181-193). 사람에서 순환되지 않는 H5N1 인플루엔자 스트레인으로부터 유래된 HA (그룹 1 HA)에 의한 면역화는 그룹 1 순환 계절성 스트레인에 대한 HA 줄기-특이적 반응을 실질적으로 증가시킨다 (Ellebedy, A. H., *et al.*, *PNAS*(2014) 111:13133-13138; Nachbagauer, R., *et al.*, *J Virol*(2014) 88:13260-13268; Whittle, J. R. R., *et al.*, *J. Virol*(2014) 88:4047-4057). 1957년, 1968년, 및 2009년의 세계적인 유행성 인플루엔자 바이러스 스트레인들의 각 출현 후, 기존의 계절성 바이러스 스트레인들은 이들 신규한 유행성 스트레인에 의해 인간 집단에서 대체되었다.
- [0017] [12] 다른(divergent) 구형 헤드 도메인을 갖는 새로운 유행성 인플루엔자 바이러스 스트레인들에 대한 노출이 줄기 도메인에서의 보존된 에피토프에 대한 친화도 성숙한 메모리 반응을 이끌어내는 것으로 가정되었다 (Palese, P. 및 Wang, T. T., *mBio*(2011) 2:e00150-11). 이러한 가설을 뒷받침하는 것으로, H6, H9, 또는 H5 구형 헤드 도메인 및 H1 줄기 도메인의 조작된 재조합 키메라 HA는 고역가 줄기-특이적 중화 항체를 생성하였다 (Pica, N., *et al.*, *PNAS*(2012) 109:253-2578; Krammer, F., *et al.*, *J. Virol*(2013) 87:6542-6550). 이들 경우에서, H1 HA의 구형 헤드 도메인은 대부분의 인가 집단이 그에 대해 나이브한 동일한 그룹 1 HA로부터의 H6, H9, 또는 H5의 구형 헤드 도메인에 의해 치환되었다. 이들 치환은 HA1의 시스테인 52과 시스테인 277 사이의 H1 HA 서열을, H6, H9, 또는 H5의 대응하는 서열로 치환함으로써 이루어졌다. 시스테인 52 및 시스테인 277은 구형 헤드 도메인과 줄기 도메인 사이의 힌지 영역에서 디설파이드 결합을 형성한다. 그룹 2 인플루엔자 스트레인에 대한 백신 개발을 위해 유사한 키메라 HA들이 H3 줄기 도메인을 이용하여 만들어졌다 (Krammer, F., *et al.*, *J. Virol*(2014) 88:2340-2343; Margine, I., *et al.*, *J. Virol*(2013) 87:10435-10446). 이들 키메라 HA들로 동물을 면역화시키자 각 바이러스 그룹에 대해 광범한 중화 활성을 갖는 줄기-특이적 항체들이 유도되었다. 인간 집단은 순환 H1 그룹 1), H3 그룹 2), 및 B형 인플루엔자 바이러스 스트레인에 대해 기존의 면

역성을 갖는다. 키메라 HA를 이용한 예방접종은 순환 스트레인의 HA 및 키메라 HA에 공통적인 줄기 도메인에 대한 항체 수준을 부스팅시켰다. 오직 일차 반응만이, 인간이 그에 대해 나이브한 키메라 HA 상의 신규한 구형 헤드 도메인에 대해 유도되었다. 동일한 줄기 도메인을 갖지만 헤드 도메인은 상이한 제2의 키메라 HA에 의한 후속 부스트는 줄기-특이적 항체 수준을 더 증가시켰다. 이러한 결과는 구형 헤드 도메인 내의 면역우성 에피토프에 대한 숙주 노출에 변화를 줌으로써 줄기 도메인내 면역준우성(immunosubdominant) 에피토프에 대한 광범하게 보호적인 면역 반응이 증가될 수 있음을 시사하는 것이다.

[0018] [13] A/WSN/33(H1N1) 헤마글루티닌의 항원 부위 B의 6-아미노산 루프는 A/Japan/57(H2N2) 및 A/Hong Kong/8/68(H3N2)로부터의 HA들의 동종 항원 부위 B 잔기들에 의해 치환될 수 있다 (Li, S., *et al.*, *J. Virol*(1992) 66: 399-404). 이러한 치환은 HA의 수용체 결합 기능을 방해하지 않는다. 이들 키메라 HA를 갖는 재조합 인플루엔자 바이러스가 MDCK (Madin-Darby Canine Kidney Epithelial Cells) 세포 배양체에서 복제되었다. A/WSN/33(H1N1) 및 A/Hong Kong/8/68(H3N2)의 키메라 HA를 갖는 바이러스들은 A/WSN/33(H1N1) 및 A/Hong Kong/8/68(H3N2) 양자 모두에 대한 항체를 유도하였다. 이러한 결과는 HA의 면역우성 항원 부위가 상이한 스트레인의 다른 HA들의 대응하는 동종 면역우성 항원 부위에 의해 치환될 수 있음을 시사하는 것이다. 이러한 치환은 결과적인 키메라 HA들의 항원 특이성을 변화시킨다.

[0019] [14] 전술한 문단들에서, 이에 따라, 하나의 인플루엔자 스트레인의 면역우성 에피토프가 다른 스트레인의 동종 면역우성 에피토프에 의해 치환되었다.

[0020] [15] 또 다른 전략은 면역우성 에피토프를 약화시키고 줄기 도메인에 대한 숙주 면역 반응에 재집중하는 것이다. 구형 헤드 도메인 내 면역우성 항원 부위는 하이퍼글리코실화를 위해 부가적인 글리코실화 부위를 도입함으로써 보호될 수 있다 (Eggink, D., *et al.*, *J. Virol*(2014) 88:699-704, Patent application US2014/0004149_A1). 구형 헤드 도메인 내 하이퍼글리코실화는 줄기-반응성 항체의 결합 친화성을 변화시키지 않았다. 하이퍼글리코실화된 HA를 이용한 마우스의 면역화는 줄기-반응성 항체의 높은 역가 및 상이한 계절 바이러스들에 의한 켈린지시 이환율과 사망률에 대한 보호를 유도하였다. 미국특허출원 2013/0315929_A1에는 면역우성 에피토프들의 일부 잔기들을 에피토프의 일부일 가능성이 덜한 다른 아미노산으로 치환시킴으로써 구형 헤드 도메인 내 면역우성 에피토프를 약화시키는 또 다른 방법이 개시되어 있다. 이러한 결과는 구형 헤드 도메인 내에서 면역우성 에피토프들을 보호함으로써 줄기 도메인 내 면역준우성 에피토프들에 대한 광범한 보호적 면역 반응을 증가시킬 수 있음을 시사한다.

[0021] 외래 에피토프를 제시하기 위한 캐리어로서의 HA

[0022] [16] 인플루엔자 HA는 HIV-1 엔벨롭 단백질의 V3-루프의 에피토프를 위한 캐리어로서 사용되어 왔다. HIV-1 엔벨롭 단백질 gp120으로부터의 12 내지 22 잔기 길이의 면역우성 에피토프 펩타이드를 HA 면역우성 항원 부위에 (부위 A 또는 부위 B 중 아무거나에) 삽입하자, 구형 헤드 도메인 내에 개별적인 HIV-1 에피토프를 갖는 키메라 HA들이 생성되었다. 부위 A 및 부위 B의 어떠한 잔기도 제거되지 않았다. 면역우성 HIV-1 에피토프들을 이들 부위에 삽입하였다. 키메라 HA들은 동물에서 HIV-1V3-루프 에피토프에 대한 면역 반응을 유도하였다 (Kalyan, N. K., *et al.*, *Vaccine*(1994) 12:753-760; 미국특허 U.S. 5,591,823_A; Li, S., *et al.*, *J. Virol*(1993) 67:6659-6666). 삽입된 에피토프들의 면역원성은 HA에 의해 증강된 것으로 보였는데, 이는 키메라 HA 단백질의 매우 낮은 투여량으로도 삽입된 에피토프에 특이적인 항체를 유도하는데 충분했기 때문이다. 이러한 결과는 HA 이외의 단백질로부터의 면역우성 외래 에피토프들이 HA의 면역우성 항원 부위에 삽입될 수 있음을 시사하는 것이다. 외래 면역우성 에피토프가 삽입된 이들 키메라 HA 분자들은 상기 삽입된 외래 면역우성 에피토프에 대한 면역 반응을 유도할 수 있다.

발명의 내용

해결하려는 과제

[0023] 발명의 개시

[0024] [17] 본 발명은 인플루엔자 헤마글루티닌 단백질의 변형된 형태, 이를 함유하는 백신, 바이러스-유사 입자, 및 바이러스, 그리고 이의 제조를 위한 재조합 방법 및 물질에 관한 것이다. 일반적으로, 이들 변형은 HA 단백질의 줄기 도메인에서 항체 생성 및/또는 성숙 절단 부위 (MCS)의 변형을 위한 대안적 에피토프로 HA 단백질의 구형 헤드 도메인의 면역우성 영역을 대체하는 것을 포함한다. MCS와 융합 펩타이드 사이의 HAO를 절단하여 HA2의 유리 N-말단을 생성하는 것은 세포 진입에 있어 필수적이다. 동물성 프로테아제에 의한 MCS의 절단을 방지하도록

MCS를 변경함으로써, MCS가 이미 절단된 바이러스는 그 바이러스가 증식할 수 있는 숙주 세포의 감염은 허용하지만 그 자손은 감염성이 아니도록 할 수 있다. 따라서, 면역우성 부위를 접하는 것들을 포함한 임의의 에피토프의 존재가 자손 바이러스에 의한 숙주의 추가 감염 없이 증폭된다.

과제의 해결 수단

- [0025] [18] 일 측면에서, 본 발명은 변형된 HA 또는 이를 함유하는 바이러스 또는 바이러스-유사 입자를 포함하는 독감 백신에 관한 것으로서, 여기서 상기 HA 단백질의 면역우성 영역이, 이 영역에 삽입된 동일한 인플루엔자 스트레인 또는 다른 인플루엔자 스트레인의 보존된 대안적인 에피토프를 함유하는 것인 독감 백신에 관한 것이다. 이에 따라 그의 본래 위치에서 면역준우성인 삽입된 대안적인 에피토프의 보다 성공적인 면역원성 형태가 얻어진다. 대안적인 에피토프들은 다른 HA의 구형 헤드 도메인으로부터 유래하지 않는 것들이다. 일반적으로 이들은 HA 줄기 도메인이나 또는 M2와 같은 비-HA 인플루엔자 단백질로부터 유래한다.
- [0026] [19] 또 다른 측면에서, 본 발명은 동물 프로테아제에 의해 절단되지 않는 MCS를 함유하도록 변형된, 변형된 인플루엔자 바이러스(백신으로도 사용가능함)에 관한 것이다. 전술한 바와 같이, 이는 그의 감염성 형태를 생성시키지 않으면서 바이러스의 증폭을 허용한다. 그러한 바이러스는 하나 이상의 면역우성 영역을, 외래 에피토프 예컨대 다른 바이러스, 박테리아 또는 종양 관련 항원에 특징적인 에피토프를 포함하는 외래 에피토프 또는 이중 에피토프로 치환함으로써 추가 변형될 수 있다.
- [0027] [20] 또 다른 측면에서, 본 발명은 재조합 물질 및 변형된 HA 단백질의 바이러스 또는 바이러스-유사 입자, 단백질의 제조방법, 및 이러한 단백질, 바이러스-유사 입자 또는 변형된 바이러스를 이용하여 항체를 생성하는 방법에 관한 것이다.

도면의 간단한 설명

- [0028] [21] 도 1은 돼지 기원의 A형 인플루엔자 바이러스 H1N1 스트레인 (PDB:3LZG)인, A/California/07/2009의 성숙한 HA의 세포외 부분의 X-선 결정 구조를 나타낸 도면이다. 이 구조는 본 발명에서 구축물 설계를 가이드하는데 이용된다. 이 HA의 펩타이드 서열은 모계(parental) 야생형(WT) H1 HA 구축물을 만드는데 사용된다. 도 1A에는 리본 드로잉 내 HA 단량체 하나와 백본 드로잉 내의 다른 두 개의 단량체로 된 HA 삼량체가 도시되어 있다. 막과 관련한 방향성이, 구형 헤드 도메인, 및 줄기 도메인의 위치가 나타나 있다. HA1은 옅은 색으로 색칠된 리본으로 도시되어 있고 HA2는 어두운 색으로 칠해진 리본으로서 도시되어 있다. 도 1B는 HA1 및 HA2의 위치, 그리고 HA1 및 HA2의 N-말단과 C-말단이 표시된 하나의 HA 단량체를 도시한다. 융합 펩타이드에는 동그라미로 표시하였다.
- [22] 도 2는 A/California/07/2009의 H1 HA 단량체에 매핑된 H3 HA의 면역우성 항원 부위의 대략적인 위치를 나타낸 도면이다. 도 2A는 도 1에서와 같은 리본 드로잉을 보여준다. H3 HA 면역우성 항원 부위 A, B, C, D, 및 E는 어둡게 색칠되어 있다. 도 2B는 막에 대해 원위부인 구형 헤드 도메인을 나타내는 상기 도 2A의 평면도이다.
- [23] 도 3은 A/California/07/2009의 H1 HA 단량체에 매핑된 H1 HA의 면역우성 항원 부위의 대략적인 위치를 나타낸 도면이다. 도 3A는 도 1에서와 같은 리본 드로잉을 보여준다. H1 HA 면역우성 항원 부위 Ca1, Ca2, Cb, Sa, 및 Sb는 어둡게 색칠되어 있다. 도 3B는 막에 대해 원위부인 구형 헤드 도메인을 나타내는 상기 도 3A의 평면도이다. 이들 면역우성 항원 부위의 많은 부위들이 화살표로 표시된 바와 같이 숙주 세포 수용체 결합 부위를 둘러싸고 있다.
- [24] 도 4는 H1 HA 구축물의 개략적인 설계도이다. 도 4A는 구축물의 핵산 특성을 나타내는 개략도이다. 서브클로닝을 위한 제한효소 부위의 위치들을 표시하였다. 도 4B는 도 4A에 도시된 핵산에 상대적 비율로 구축물의 단백질 특징을 나타낸 개략도이다. GP67ss는 N-말단에서 GP67 분비 시그널 (시그널 펩타이드)를 나타낸다. HA1은 HA HA1을 나타낸다. MCS는 HA1의 일부인 HA 성숙 절단 부위를 나타낸다. HA2는 HA HA2를 나타낸다. TEV는 TEV 절단 부위를 나타낸다. Foldon은 foldon 서열을 나타낸다. His는 C-말단에서 10-히스티딘-태그 (10xHis-태그)를 나타낸다.
- [25] 도 5는 이중 에피토프가 HA 구형 헤드 도메인에 위치하고, 변형된 성숙 절단 부위가 있는 위치를 도시한다. 모계의 구축물 펩타이드 서열 (WT)을 H3 (A/Aichi/2/1968 H3N2) 및 H1 (A/Puerto Rico/8/1934 H1N1)의 HA 서열과 정렬시킨다. 시그널 서열 제거 후 첫 번째 잔기를 위치 1로 하는 HAO 넘버링에 기초하여 우측에 각 서열의 아미노산 위치를 나타내었다. 성숙 절단 부위 (MCS)은 화살표 및 수직선을 이용하여 표시하였다. H3

HA의 면역우성 항원 부위 A, B, C, D 및 E를 H3 서열 위에 마킹하였다. H1 HA의 면역우성 항원 부위 Ca1, Ca2, Cb, Sa 및 Sb는 H1 서열 밑에 마킹하였다. WT 중 밀줄친 서열들을 서열 밑에 표시된 이중 에피토프로 대체한다. 변경된 성숙 절단 부위는 성숙 절단 부위 (MCS) 밑에 표시하였다. 광범위 중화 항체(bNAbs), F16, 1C9, 및 CR8020에 대한 천연의 면역준우성 항원 부위를 각각 서열 위에 줄을 그어 표시하였다.

[26] 도 6은 구형 헤드 도메인의 면역우성 항원 부위에 위치한 복합 CR8020 에피토프를 갖는 HA 단량체 모델을 도시한 도면이다. 이 HA 단량체는 도 1에 도시된 바와 같이 리본 드로잉으로 도시되어 있다. 복합 CR8020 에피토프에 의해 치환된 면역우성 항원 부위는 어둡게 색칠되어 있다. 줄기 도메인 내의 본래의 천연 CR8020 에피토프들 역시 어둡게 색칠되어 있다. 도 6A는 CR8020Sa4의 모델을 나타낸다. 도 6B는 CR8020Ca의 모델을 나타낸다.

[27] 도 7은 본 발명에서 설명된 HA 구축물의 발현 데이터를 나타낸다. 표시된 구축물의 재조합 배컬로바이러스로 감염된 SF9 세포의 세포 배양 배지를 수확하여 Ni-NTA 수지(Ni 풀-다운)에 의해 포획하였다. Ni-NTA 수지를 1xPBS로 세척하여 미결합 단백질을 제거하였다. Ni-NTA 수지에 결합된 단백질들을 Coomassie® 염색을 이용한 SDS-PAGE (도 7A, B, C, 및 D) 또는 항-His 웨스턴 블롯(도 7E)에 의해 분석하였다. 화살표는 전장 HA0를 나타낸다.

[28] 도 8은 본 발명에 사용된 정제된 HA 구축물을 나타낸다. 표시된 구축물의 재조합 배컬로바이러스로 감염된 SF9 세포의 세포 배양 배지를 수확하여 Ni-NTA 수지(Ni 풀-다운)에 의해 포획하였다. Ni-NTA 수지를 1xPBS로 세척하였다. 결합된 단백질 이미다졸을 이용하여 용리시키고 Coomassie® 염색을 이용한 SDS-PAGE에 의해 분석하였다. 화살표는 전장 HA0를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0029] 발명을 수행하기 위한 모드
- [0030] [29] 본 발명은 HA 단백질의 변형에 관한 것이므로, 이 단백질 및 상기 설명된 것에 추가적인 그의 기능에 대한 설명이 도움이 될 수 있다.
- [0031] [30] HA0의 프로테아제 절단은 A형 인플루엔자 바이러스의 감염성의 선결조건이다. 숙주에서의 이들 프로테아제의 분포는 조직 방향성(tissue tropism)의 결정인자 중 하나이며 따라서 병원성의 결정인자이기도 하다. 일염기성 성숙 절단 부위를 갖는 HA의 대부분을 절단하는 몇몇 트립신-유사 프로테아제가 호흡기도 및 폐에서 동정되었다 (Kido, H., *et al.*, *J. Biol Chem*(1992) 267:13573-13579; Peitsch, C., *et al.*, *J. Virol*(2014)88:282-291; Zhirnov, O., P., *et al.*, *J. Virol*(2002) 76: 8682-8689). 이들 프로테아제들은 세린 프로테아제이다. 이들 중 한 가지는 쥐의 기관지 상피 클라라(Clara) 세포로부터 최초로 단리된, 트립신-유사 프로테아제 클라라이다. HA0를 절단하는 이들 프로테아제의 좁은 조직 분포는 인플루엔자 바이러스 감염을 포유동물의 호흡기도 및 폐로 국한시킨다.
- [0032] [31] HA 성숙에 책임이 있는 프로테아제는 아직까지 잘 특징화되어 있지 않다. 트립신과 마찬가지로, 이들 프로테아제들은 펩타이드 결합 C-말단을 절단하여 아르기닌(R) 또는 리신(K)과 같은 염기성 잔기로 만든다. 성숙 절단 부위 (MCS)는 HA1의 C-말단에 위치하며 MCS의 C-말단에서의 절단은 감염성에 필수적이다.
- [0033] [32] H5 및 H7 아형(subtypes)의 성숙 절단 부위 내의 다염기성 서열은 퓨린 및 서브티리신형 프로테아제와 같은 보다 광범위한 세포 프로테아제에 의한 절단 감수성을 유도하는데, 이는 포유동물에 있어 이들 바이러스들의 보다 광범한 조직 방향성 및 보다 높은 병원성과 상관성이 있다 (Stieneke-Grøfber, A., *et al.*, *EMBO J*(1992) 11:2407-2414; Maines, T. R., *et al.*, *J. Virol*(2005) 79:11788-11800). 다수의 고병원성 조류 인플루엔자 (HPAI) 아형들은 H5 및 H7이다. 인간의 HPAI 발병 중 보고된 인명 피해 사망률은 유행성 인플루엔자 바이러스 및 계절성 인플루엔자 바이러스보다 높다 (Morens, D. M., *et al.*, *Nature* (2012) 486:335-340). MCS의 이들 다염기성 서열은 또한 조류 종의 여러 장기에서 바이러스 복제를 유도하여 이들 조류 종에서 높은 사망률을 초래한다.
- [0034] [33] 계절성 독감 바이러스 및 비병원성 조류 인플루엔자 바이러스의 HA는 호흡기관과 폐에서 특이적인 프로테아제에 의해 세포의 절단되는데 이로 인해 이들의 조직 방향성이 제한된다. 다른 한편, 고병원성 바이러스의 HA는 어디에서나 발생하는(ubiquitously occurring) 프로테아제에 의해 세포내 절단된다. 이들 고도로 병원성인 바이러스들은 다양한 조직에서 복수-사이클의 복제를 수행하여 전신 감염을 일으킨다 (Steinhauer, D. A., *Virus. Virol*(1999) 258:1-20; Taubenberger, J. K., *PNAS*(1998) 95:9713-9715). 1918년의 대유행성 인간

인플루엔자 바이러스 스트레인 역시 광범한 세포 프로테아제를 이용하고 그의 뉴라미니다제(NA)를 사용하여 HA 절단을 위한 플라스미노젠을 동원(recruit)한다 (Goto, H. 및 Kawaoka, Y., *PNAS*(1998) 95:10224-10228; Chaipan, C., *et al.*, *J. Virol*(2009) 83:3200-3211).

- [0035] [34] HA0의 절단은 숙주 세포와의 비리온 융합에 필수적인 새로운 유리 N-말단을 갖는 HA2를 생성시킨다. 절단된 형태는 수용체 결합 및 막 융합의 기능을 완전히 갖는 HA의 성숙한 형태이다. HA0는 수용체에 결합하는 능력은 있지만 막 융합을 매개하지는 않는다. HA의 전구체 형태와 성숙한 형태 양자 모두 비리온 표면에 존재한다. HA0만을 갖는 바이러스들은 융합 활성이 없어 감염을 일으키지 않는다.
- [0036] [35] HA2의 N-말단 12개 잔기는 "융합 펩타이드"라 칭해진다 (Skehel, *et al.*, *Biochem Soc Trans*(2001) 29:623-626). 상기 융합 펩타이드는 소수성 서열을 갖는다. HA0에서, MCS 및 융합 펩타이드는 표면 루프를 형성한다. 절단 후, 새로이 생성된 N-말단 융합 펩타이드는 HA 삼량체 인터페이스 내로 삽입된다. 숙주 세포 수용체에 결합한 후, 부착된 인플루엔자 비리온은 숙주 세포에 의해 세포내섭취되어 엔도솜으로 유입된다. 성숙한 HA의 융합 가능성은 특정 인플루엔자 바이러스 스트레인에 따라 엔도솜 pH, 즉 pH 5 내지 6에서 활성화된다. 낮은 pH에서의 HA 구조의 집중적인 변화는 세포 막으로의 융합 펩타이드의 탈출을 일으킨다. 숙주 엔도솜 막으로의 융합 펩타이드의 삽입은 숙주의 엔도솜 막과 HA의 C-말단 막 앵커 영역을 함유하는 바이러스 막의 융합을 야기한다. 융합은 바이러스 RNA 세그먼트들을 숙주 세포의 세포질 내로 방출시켜 바이러스 복제가 일어나는 숙주 세포의 핵으로 유입시킨다.
- [0037] [36] HA는 인플루엔자 바이러스 스트레인의 숙주 범위를 결정한다. HA는 숙주 세포 표면 당단백질과 당지질의 말단 시알산에 결합한다. 인간 인플루엔자 스트레인으로부터의 HA는 SA α 2,6Gal에 의해 종결된 시알릴올리고당에만 배타적으로 결합하는 반면, 조류 및 말의 인플루엔자 스트레인으로부터의 HA는 SA α 2,3Gal에 결합한다. 인간은 대부분 SA α 2,6Gal 시알릴올리고당을 갖고 조류 종은 대체로 SA α 2,3Gal 시알릴올리고당을 갖는다. 도 3에 개략적으로 도시된 바와 같이, HA 상의 숙주 수용체 결합 부위는 HA1으로만 만들어진 구형 헤드 도메인 내에 위치한다. 인플루엔자 바이러스 HA의 수용체 특이성은 잘 특징화되어 있다. SA α 2,6Gal 또는 SA α 2,3Gal을 인식하는데 책임이 있는 아미노산 잔기들이 매핑되어 있다. 수용체 결합 포켓 내에서 하나의 아미노산 잔기만 치환되어도 수용체 결합 특이성이 변하는 것으로 나타났다 (Rogers, G. N., *et al.*, *Nature*(1983) 304:76-78). 1918년의 인플루엔자 대유행 바이러스의 다른 분리 스트레인들은 HA의 수용체 결합 부위에서의 하나의 아미노산 치환 결과, 상이한 수용체 결합 특이성을 가졌었다 (Glaser, L., *et al.*, *J. Virol*(2005) 79:11533-11536). 스트레인 A/South Carolina/1/18 HA는 우선적으로 인간 세포 수용체 SA α 2,6Gal에 결합하는 반면, 스트레인 A/New York/1/18 HA는 인간 세포 수용체 SA α 2,6Gal 및 조류 세포 수용체 SA α 2,3Gal 양자 모두에 결합하는데, 이는 인플루엔자 바이러스들의 숙주 적응의 역동적 특성을 나타내는 것이다.
- [0038] [37] 비록 HA의 수용체 결합 부위에서 하나의 돌연변이만으로도 조류로부터 인간으로 수용체 결합 특이성이 변화한다고 해도, 인간에서의 H5 조류 인플루엔자 바이러스 스트레인의 효과적인 전달을 위해서는 수용체 결합 부위 외부에서의 HA의 몇몇 다른 변화 및 다른 인플루엔자 단백질에서의 변화가 요구된다 (Imai, M., *et al.*, *Nature*(2012)486:420-428; Herfst, S., *et al.*, *Science*(2012) 336:1534-1541; Chen, L-M., *et al.*, *Virol*(2012) 422:105-113; Russell, C. A., *et al.*, *Science*(2012) 336:1541-1547). 다른 변화와 함께, H5 HA에서의 단지 4개의 아미노산 치환만으로도 호흡기 비말 감염을 통해 페렛들에게 돌연변이 조류 H5 바이러스를 전달하는데 충분하다. 종들 간의 인플루엔자 바이러스 전달에 대한 이해에 더해, 이들 연구는 포유동물에서의 비-포유동물 인플루엔자 바이러스 전달에 대한 실험실 공정을 정립시켰는데, 이를 이용하여 범 인플루엔자 백신 후보를 평가할 수 있다.
- [0039] 항-독감 항체 및 HA의 항원 부위에 대한 설명
- [0040] [38] 인플루엔자 바이러스 지질층 엔벨롭의 가장 풍부한 단백질로서, HA는 인플루엔자 바이러스의 주요 항원이며 중화 항체에 대한 일차 에피토프를 산생한다. 대부분의 항-독감 항체는 HA 및 NA에 대한 것이다. 예방접종에 의해 유발된 인플루엔자-반응성 항체의 육십 퍼센트(60%)는 HA에 반응한다 (Wrammert, J., *et al.*, *Nature*(2008) 453:667-671). HA-반응 항체의 대부분은 구형 헤드 도메인 내의 수용체 결합 부위를 둘러싸고 있다. 이들 항체의 일부는 바이러스-중화 항체이다. 이들 중화 항체 (NAbs)는 일반적으로 숙주 세포에 대한 HA 결합을 간섭하고 HA 억제(HI) 활성을 나타낸다. 이들은 이들 항원 부위의 높은 다양성으로 인해 일반적으로 스트레인-특이적이고, 따라서 열망되는 광범한 중화 활성을 결여한다 (Wang, T. T., 및 Palese, P., *Nat Struct Mol Biol*(2009) 16:233-234).
- [0041] [39] 인플루엔자 바이러스의 광범한 감염 확산은 항원 속성을 변경할 수 있는 바이러스의 능력의 결과이다. 인

플루엔자 바이러스의 항원 속성의 변화는 오류가 발생하기 쉬운 바이러스성 RNA 의존성 RNA 폴리머라제 복합체에 의한 바이러스 게놈의 저충실도(low fidelity)의 결과이다. 높은 돌연변이율은 HA의 항원 속성의 점진적인 변화인 항원 소변이(antigenic drift)로 이어진다. 항원 소변이는 인플루엔자 바이러스의 모든 유형에서 일어난다. HA 유전자의 서열 분석 결과, 침묵 핵산 서열 변화가 전체 HA 유전자에 퍼져 있음에도 불구하고 아미노산 서열의 변화의 대부분이 HA1에 위치한다는 것을 보여준다 (Palese, P., 및 Young, J. F., *Science*(1982) 215:1468-1474). HA2는 HA1보다 더 보존되어 있다.

[0042] [40] 바이러스 게놈의 분절화(segmented) 속성은 바이러스 게놈의 분열된 성질은 또한 두 가지 다른 스트레인의 바이러스가 숙주를 동시에 감염시킬 때 바이러스 게놈 분절들의 재-조립(re-assortment)을 유도한다. HA의 항원 속성의 갑작스런 변화인 항원 대변이(antigenic shift)로 이끄는 다양한 바이러스 표면 단백질을 갖는 새로운 바이러스 스트레인이 이 재-조립으로부터 발생한다. 항원 대변이(antigenic shift)는 A형 인플루엔자 바이러스에서만 일어난다. 항원 소변이(antigenic drift) 및 항원 대변이는 인플루엔자 바이러스를 기존 항체에 의한 중화를 방지할 수 있다.

[0043] H3형 인플루엔자

[0044] [41] 지난 35년 동안 탐구된 HA형 인플루엔자 의 최초의 상세한 구조는 HA의 항체-결합 부위를 규명하고 항원 소변이 및 항원 대변이에 대한 분자 수준의 설명을 제시해준다 (Wilson, I. A., et al., *Nature*(1981)289:366-373; Wiley, D. C., et al., *Nature*(1981) 289:373-378; Wiley, D. C., 및 Skehel, J. J., *Ann Rev Biochem*(1987) 6:365-394).

[0045] [42] 항원적으로 구별되는 바이러스들의 HA 서열 비교 결과 H3 HA, 그룹 2 아형 HA의 구형 헤드 도메인 내의 면역우성 항원 부위가 동정되었다. H3 HA의 면역우성 항원 부위에 대한 설명은 다른 사람들에 의해 입증된다 (Both, G. W., et al., *J. Virol*(1983) 48:52-60). HA 구조에서 이들 항원 부위의 개략적인 위치가 도 2에 도시되어 있다. 이들 항원 부위의 서열 위치들이 도 5에 나타나 있다. 부위 A는 A/Aichi/2/1968 H3N2 스트레인의 HA의 잔기 133 내지 148의 표면 루프 내에 위치한다 (도 5에서 H3으로서 넘버링). 140-루프로 알려진 이 루프는, 구형 헤드 도메인으로부터 돌출되며 수용체 결합 포켓의 하부 림 상에 위치한다. 부위 B는 구형 헤드 도메인의 상부(top)에 위치하며 잔기 187 내지 196의 α 헬릭스의 표면과 수용체 결합 포켓의 상부 림을 따라 잔기 155 내지 160의 인접 표면 루프를 포함한다. 부위 C는 Cys52와 Cys277 사이의 디설파이드 결합을 감싸고 있다. Cys52를 중심으로 하는 루프 (잔기 46 내지 55)와 Cys277를 중심으로 하는 루프 (잔기 271 내지 280)의 교차(crossing)에 의해 구형 헤드 도메인과 줄기 도메인 사이의 힌지부에 벌지(bulge)가 형성된다. 부위 D는 HA 삼량체의 HA 단량체 서브유닛들 간의 인터페이스 영역에 위치한다. 이 부위는 구형 헤드 도메인의 핵심부의 8-가닥 β -시트 구조 내의 잔기 200 내지 214의 2개의 β -가닥을 중심으로 한다. 다른 β -가닥의 턴(turns)에 존재하는 잔기들 역시 부위 D의 일부일 수 있다. 부위 D는 대개 HA 삼량체 인터페이스에 매립되어 있다. 부위 D가 항원 부위로서 어떤 방식으로 기능하는지는 분명치 않다. 부위 E는 구형 헤드 도메인 측면의 부위 A와 부위 C 사이에 위치하며 8-가닥의 β -시트의 가장자리 상의 잔기 119 내지 122의 β -가닥, 및 잔기 78 내지 83, 잔기 62 내지 63의 표면 루프로 만들어진다. 부위 E의 이들 잔기들과 함께 구형 헤드 도메인의 측면에 연속 표면이 형성된다. 1968년에서 2003년 사이의 기간 동안 출현한 인플루엔자 바이러스의 아미노산 치환 및 항원 속성을 비교한 결과 주요 항원 변화에 책임이 있는 치환들이 부위 A 및 부위 B에만 존재하는 것으로 나타났다 (Smith, D. J., et al., *Science* (2004) 305:371-376; Koel, B. F., et al., *Science*(2013) 342:976-979). 부위 C, D 및 E에서의 치환은 마이너한 항원 변화를 야기하는 것으로 보인다. 이러한 결과는 대부분의 스트레인-특이적 중화 항체가 구형 헤드 도메인의 수용체 결합 부위 주변의 부위 A 및 부위 B에 결합함을 시사하는 것이다.

[0046] H1 인플루엔자

[0047] [43] 바이러스 스트레인 A/PR/8/34의 항원성으로 구별되는 그룹 1 H1 HA의 유전 분석에 의해 구형 헤드 도메인에서 Ca1, Ca2, Cb, Sa 및 Sb 부위라고 불리는 상이한 항원 부위들이 동정되었다 (Caton, A. J., et al., *Cell*(1982) 31:417-427; Gerhard, W., et al., *Nature*(1981) 290:713-717). HA 구조에서 이들 항원 부위의 대략적인 위치가 도 3에 개략적으로 도시되어 있다. 이들 항원 부위의 서열 위치는 도 5에 도시되어 있다. Ca1 부위는 8-가닥의 β -시트 구조의 β -가닥들의 턴(turn) 부위에 위치한다. 잔기 165 내지 169의 턴들 중 하나 (도 5에서 H1으로 넘버링됨)가 표면에 노출된다. Ca1 부위의 잔기 207은 H3 HA의 부위 D에 해당하는 두 개의 β -가닥을 연결하는 턴 부분에 위치한다. H1 HA의 Ca1 부위는 일반적으로 H3 HA의 부위 D에 상응한다. Ca2 부위 역시 또한 1차 구조에서는 분리되지만 3차 구조에서는 함께인 2 개의 세그먼트로 구성된다. Ca2의 하나의 세그먼트는 H3 HA의 부위 A의 상응하는 위치에서 잔기 136 내지 141의 표면 루프에 존재한다. Ca2의 또 다른 세그먼트는 긴

표면 루프 상에 잔기 220 내지 221로 이루어진다. Ca2 부위는 HA 단량체의 구형 헤드 도메인에서 Ca1의 반대쪽에 있지만 HA 삼량체 내 또 다른 HA 단량체의 Ca1 부위에 인접해 있다. 하나의 HA 단량체의 Ca1 부위는 삼량체 구조의 또 다른 HA 단량체의 Ca2 부위와 연속적인 표면을 형성한다. Cb 부위는 8-가닥의 β -시트 구조 다음에 표면 루프를 형성하는 잔기 70 내지 75의 선형 에피토프이다. 이것은 H3 HA의 부위 E에 해당한다. 부위 Sa 및 Sb는 H3 HA의 부위 B에 해당하는 하위 부위(subsites)로 간주될 수 있다. 잔기 154 내지 163의 부위 Sa는 H3 HA의 부위 B 루프에 대응하는 루프와 중첩된다. Sa 부위의 또 다른 세그먼트는 잔기 124에서 125 부근의 턱 부분이다. 부위 Sb는 H3 HA의 부위 B α -헬릭스에 상응하고 대부분 α -헬릭스 잔기 188 내지 194를 갖는다.

[0048] 광범위 중화 항체 및 보존된 에피토프

[44] 자연적인 인플루엔자 바이러스 감염 또는 불활성화된 3가 인플루엔자 백신(TIV: trivalent inactivated influenza)에 의한 예방접종 역시, 대체로 막 근위 줄기 도메인에 위치하는 보존된 HA 에피토프에 대한 낮은 수준의 항체를 유도한다 (Ellebedy, AH 등, PNAS (2014) 111: 1313313138). 이들 항체들은 많은 인플루엔자 바이러스 스트레인들 간에 보존된 에피토프를 인식한다. 여러 스트레인들에 대해 보호를 제공하는 것들을 광범위 중화 항체 (bNAbs)라 칭한다. 이들 bNAb는 일반적으로 HA 저해 활성을 가지지 않으며, 숙주 세포 표면 수용체에 대한 HA 결합을 막지 못한다. 첫 교차-스트레인 항체 C179는 20년 전에 특징지어졌다 (Okuno, Y., *et al.*, *J. Virol*(1993) 67:2552-2558). 이것은 H1과 H2 스트레인 간에 보존되어있는 HA1의 잔기 318 내지 322 (TGLRN) 및 HA2의 잔기 47 내지 58 (GITNKNVSVIEK)의 구조적 에피토프를 인식합니다. C179는 HA의 융합 활성을 억제하여 바이러스 중화를 일으킨다.

[0050] [45] 많은 bNAb가 인플루엔자 바이러스에 감염된 환자로부터 검출되거나 분리되었다 (Ekiert, D. C., 및 Wilson, I. A., *Curr Opin Virol*(2012) 2:134-141). bNAb는 그룹 1, 그룹 2 또는 그룹 1 및 그룹 2 A형 인플루엔자 바이러스를 중화시키는 것으로 입증되었다. bNAb에 의해 인식되는 다수의 항원 에피토프가 동정되고 특징화되었다. 이들 에피토프들은 HA 서열의 선형 세그먼트이거나 또는 HA 서열의 다중 선형 세그먼트들을 갖는 구조적 에피토프이다. bNAb에 대한 다수의 에피토프는 HA 단백질의 덜 가변적인 줄기 도메인에 위치한다. 이들 에피토프는 MCS 주위의 융합 펩타이드 및 펩타이드 서열을 포함하지만 이에 한정되지 않는다.

[0051] [46] 많은 bNAb 및 이들의 에피토프가 구조적으로 특징화되어 왔다. 예를 들어, bNAb FI6은 HA0에서 MCS 및 융합 펩타이드의 잔기를 인식한다 (Corti, D., *et al.*, *Science*(2011) 333:850-856). HA의 펩타이드 매핑에 의해 FI6 에피토프로서 두 개의 펩타이드가 동정되었는데, RCSK는 MCS와 대부분의 융합 펩타이드로 구성되었고, KESTQKAIDGVTNKVNS는 MCS 및 대부분의 융합 펩타이드로 구성된 RKKRGLFGAIAFGIE 및 HA2의 헬릭스 코일된-코일된 펩타이드의 KESTQKAIDGVTNKVNS가 그것이다. 제단된 FI6의 중화 메카니즘은 HA0의 MCS에 대한 프로테아제의 접근을 차단함으로써 HA 성숙을 방지하는 것 뿐만 아니라 막 융합을 억제하는 것이다. 또 다른 bNAb F10은 성숙한 HA 형태의 융합 펩타이드 주변의 구조적 에피토프를 인식한다 (Su, J., *et al.*, *Nat Struct Mol Biol*(2009) 16:265-273). F10은 아마도 막 융합을 방지함으로써 모든 그룹 1 A형 인플루엔자 바이러스를 억제하는 것으로 여겨진다.

[0052] [47] 모노클로날 bNAb 1C9는 시험관내에서 세포 융합을 억제한다 (미국특허 8,540,995; Prabhu, N., *et al.*, *J. Virol*(2009) 83:2553-2562). 1C9는 HPAI H5N1의 H5 HA2의 융합 펩타이드의 N-말단, GLFGAIAFG의 선형 에피토프를 인식한다 (Immune Epitope Database web address: iedb.org/assay_details.php?assayId=1599077). 1C9는 고병원성 조류 인플루엔자 (HPAI) H5N1 바이러스 감염에 대해 마우스를 보호하는 것으로 나타났다. 모노클로날 bNAb는 또한 기억 B 세포로부터도 분리되었다 (Hu, W., *et al.*, *Virol*(2013) 435:320-328). 이 모노클로날 항체들 중 몇몇은 2009 년의 유행성 H1N1 인플루엔자 바이러스로부터 HA2의 FIEGGWTGMVDGWYGYHH의 선형 에피토프를 인식한다. 이 에피토프는 융합 펩타이드의 1C9 에피토프에 대한 C- 말단이다. HA 융합 펩타이드의 14 개 잔기의 서열은 A형 및 B형 인플루엔자 바이러스를 통해 고도로 보존된다.

[0053] [48] HA 융합 펩타이드의 보존된 성질이 범용 인플루엔자 백신을 개발하기 위해 연구되었다. 절단 쇠(scission bond) 양측에서 HA2의 융합 펩타이드 및 MCS 를 함유하는 HA1의 마지막 9개 아미노산 잔기를 포함하는 B형 인플루엔자 바이러스의 HA0의 고도로 보존된 서열에 기초한 펩타이드 컨쥬게이트 백신은, B형 인플루엔자 바이러스의 항원적으로 상이한 바이러스 스트레인들에 의한 치명적인 켈린지에 대해 보호적인 면역 반응을 유도하였다 (Bianchi, E., *et al.*, *J. Virol*(2005) 79:7380-7388).

[0054] [49] 모노클로날 bNAb CR6261은 줄기 도메인의 HA2 헬릭스 A 및 HA1 잔기에서 고도로 보존된 영역을 인식한다 (Ekiert, D. C., *et al.*, *Science*(2009) 324:246-251). CR6261은 HA가 융합 후 (post-fusion) 형태로 전환되는 것을 방지함으로써 그룹 1 인플루엔자 바이러스를 중화시킨다. CR6261은 동일한 VH1-69 생식계열 항체 중쇄

를 사용하는 bNAb에 속한다. 다른 VH1-69 모노클로날 항체인 CR8020은 그룹 2 인플루엔자 바이러스를 중화시킨다. CR8020은 HIV gp41 서브유닛으로부터의 막-근위 외부 영역(membrane-proximal external region: MPER)을 인식하는 HIV에 대한 항체와 유사하게 막에 아주 근접한 (~ 15-20Å) 줄기 도메인의 기부에서 HA에 결합한다 (Ekiert, D. C., *et al.*, *Science*(2011) 333:843-850). CR8020 에피토프의 2개의 주요 성분들은 융합 펩타이드의 C-말단 부위(HA2 잔기 15 내지 19, H3의 EGMID) 및 줄기 도메인의 기부에 근접한 5-가닥 β 시트의 최외곽 가닥(HA2 잔기 30 내지 36, H3의 EGTGQAA)으로 구성된다. 이들 두 성분들은 H3의 일차 구조에서 10 잔기만큼 떨어져 있다. HA의 줄기 도메인에 결합하는 대부분의 bNAb는 A형 인플루엔자 바이러스의 HA의 그룹 1 (H1, H2, H5, H6, H8, H9, H11, H12, H13, 및 H16) 또는 그룹 (H3, H4, H7, H10, H14, 및 H15)를 중화시킨다. 이들 항체들은 숙주 세포들에 결합하는 HA를 억제하지 않지만 바이러스 막 및 숙주 세포막의 융합을 방지할 수도 있다.

[0055] [50] HA 수용체 결합 부위는 구형 헤드 도메인의 최정상(top)에 있는 포켓이다(Wilson, I. A., *et al.*, *Nature*(1981) 289:366-373; Wiley, D. C., 및 Skehel, J. J., *Ann Rev Biochem*(1987) 56:365-394). 이 포켓은 많은 인플루엔자 스트레인들에서 고도로 보존되는 아미노산 잔기들에 의해 형성된다. 포켓 테두리는 전술한 바와 같은, H3 내 부위 A 및 부위 B와 같은 면역우성 항원 부위에 의해 형성된다. 건강한 인간 대상자로부터의 인간 모노클로날 항체(mAb)의 클로닝에 의해 H1, H2 및 H3 스트레인으로부터의 HA의 구형 헤드 도메인 내 수용체 결합 부위에 매우 근접한 보존된 잔기를 인식하는 bNAb가 확인되었다 (Krause, J. C., *et al.*, *J. Virol*(2011) 85:10905-10908; Krause, J. C., *et al.*, *J. Virol*(2012) 86:6334-6340). 구조 연구에 의하면, 이들 bNAb 중 적어도 일부는 수용체 결합 포켓에 대한 시알산의 상호 작용을 모방하는 것으로 밝혀졌다 (Whittle, J. R. R., *et al.*, *PNAS*(2011) 108:14216-14221; Ekiert, D. C., *et al.*, *Nature*(2012) 489:526-532).

[0056] [51] 줄기-반응 항체는 자연 감염에서 드문 경우이며, 현재 계절 독감 백신으로 예방 접종을 하는 경우에는 더 더욱 드물다. 오직 이들 항체의 서브세트만이 중화 항체이다. 줄기 도메인의 보존된 특성으로 인해, 대부분의 중화 항체는 bNAb이다. 인간 대상자로부터 만들어진 항체 라이브러리로부터의 클로닝에 의해 이제 이들 희귀한 광범위 중화 줄기-반응성 항체를 일상적으로 동정할 수 있게 되었다 (Kashyap, A. K., *et al.*, *PNAS*(2008) 105:5986-5991; Wrammert, J., *et al.*, *Nature*(2008) 453:667-671). 이러한 줄기-반응성 항체의 희귀한 발생은 줄기 도메인 내의 이들 에피토프는 면역준우성인 반면, 구형 헤드 영역 내의 에피토프는 대부분의 항체가 지향되는 면역우성이라는 가설에 이르게 한다 (Krammer, F., 및 Palese, P., *Nature Rev Drug Disc*(2015) 14:167-182). 유행성 계절 독감 바이러스에 대한 반복 노출 또는 계절 독감 백신 접종은 구형 헤드 도메인 내 면역우성 항원 부위에 대한 항체 생성을 유도한다. 구형 헤드 도메인 내 면역우성 에피토프들의 존재는 줄기 도메인으로부터의 이차 반응을 왜곡할 수 있다고 제안되어 왔다 (Russell, C. J., *N Engl J Med*(2011) 365:1541-1542). 현행 스트레인들에 감염되었거나 이들에 대해 예방접종된 사람들은 면역학적으로 나이브한 사람들보다 보편적인 반응을 하기가 더 힘들 수 있다.

[0057] M2 단백질 및 그의 에피토프

[0058] [52] M2는 바이러스 엔벨롭에서 동종사량체(homotetramer) (homotetramer) 양성자 채널을 형성하는 단일-통과(single-pass) 막 단백질이다 (Lamb, R. A., *et al.*, *Cell*(1985)40:627-633; Pielak, R. M., 및 Chou, J. J., *Biochim Biophys Acta*.(2011) 1808:522-529). 이것은 바이러스 엔벨롭에서 HA 에 비해 훨씬 낮은 농도로 존재한다 (1:10 내지 1:100의 M2:HA 비율). M2 양성자 채널 기능은 바이러스 성 단백질을 숙주 세포의 세포질 내로 방출하고 세포 표면으로의 HA 수송을 위한 골지 내강의 pH를 조절하기 위한, 바이러스 내부의 pH를 조절하는데 있어 중요하다. A형 인플루엔자 바이러스 M2(AM2) 단백질은 M2e로 알려진 잔기 1 내지 23의 세포외 N-말단 도메인, 잔기 24 내지 46의 막관통(TM: transmembrane) 도메인 및 잔기 47 내지 97의 세포 내 C-말단 도메인으로 된 97개의 잔기를 갖는다. 4개의 M2 분자들로부터의 TM 도메인은 세포 유입시 바이러스 막에서의 pH를 조절하고 바이러스 조립 및 배출 시 감염된 세포의 트랜스-골지 막에서의 pH를 조절하는 pH-민감성 양자 채널로서 기능하는 4개의 나선 번들을 형성한다. 큰 세포질 도메인은 안정한 사량체 형성에 결정적이며 바이러스 내부 껍질의 M1 단백질과의 결합을 통해 바이러스 조립에 일익을 담당한다. 양성자 채널 기능에 필수적인 TM 도메인의 HXXXW 서열 모티프를 제외하면 A, B 및 C형 인플루엔자의 M2 단백질은 서열 상동성을 거의 공유하지 않는다. 그러나 AM2의 N-말단의 세포외 잔기 10개는 모든 A형 인플루엔자 바이러스에서 보존된다.

[0059] [53] A형 인플루엔자 바이러스 M2(AM2)는 AM2 양성자 채널을 차단하는 항바이러스제인 아만타딘과 리만타딘의 표적이다. 양성자 채널의 폐쇄 상태를 안정화시키는 약물인 리만타딘은 채널 도메인의 C-말단 근처의 지질-포켓에 결합한다. 아다만틴은 A형 인플루엔자 바이러스에 대해서는 항바이러스제이지만 B형 인플루엔자에는 효과가 없다. 이러한 채널 차단제의 사용은 AM2의 채널 도메인에서 돌연변이에 기인한 광범위한 약물 내성으로 인해 중

단되었다. 이러한 돌연변이의 대부분은 야생형(WT) 바이러스보다 전파성이 덜한 다소 약화된 바이러스를 발생시킨다. 이러한 약제 내성 돌연변이주들은 약물 선택 압력이 없을 때 WT로 되돌아 갈 수 있다.

[0060] [54] M2 단백질은 인플루엔자 바이러스에 감염된 숙주 세포의 표면 상에 풍부하게 발현되는 필수 막 단백질이다 (Lamb, R.A., 등, Cell (1985) 40: 627-633). M2는 인플루엔자 바이러스에 대한 세포 독성 T 림프구(CTL) 반응을 위한 세포 표면 항원인 것으로 제안된다. A형 인플루엔자 바이러스 감염은 M2에 대한 낮은 역가의 항체만을 유도할 뿐이다(Feng, J., et al., Virology (2006) 3:102-115). M2e의 고도의 구조 보존은 부분적으로는, 빈약한 M2e-특이 항체 반응 및 그에 따른 변화에 대한 압력의 부재로 인한 것일 수 있다. 항-M2 항체 반응은 M2 단백질에 대한 기존의 항체가 있는 개체에서 더욱 강하게 나타났다 (Zhong, W., et al., J. Infect Dis (2014) 209:986-994). 2009년의 유행성 H1N1 A형 인플루엔자 바이러스 감염으로 유발된 항-M2 항체는 계절성 A형 A형 인플루엔자 바이러스의 M2 단백질과 교차 반응을 일으켰다. 마우스를 항-M2e 항체로 처리하면 이 병의 진행이 현저하게 지연되고 M2e 탈출 돌연변이가 분리되는데, 이는 M2e를 A형 인플루엔자 바이러스 감염에 대한 백신으로 사용할 가능성이 있음을 시사하는 것이다 (Zharikova, D., et al., J. Virology (2005) 79:6644-6654).

[0061] [55] M2e와 HA의 융합에 기초한 DNA 백신이 설명되나 바 있다 (Park, K. S., et al., Vaccine (2011) 29:5481-5487). 이 융합 단백질은 HA 단백질의 N 말단에 하나의 인간 M2e 펩타이드와 하나의 조류 M2e 펩타이드를 갖는데 이들 사이에는 20-잔기의 링커가 위치한다. 코딩된 M2e-HA 융합 단백질의 발현이 확인되었다. M2e-HA 융합 DNA 백신으로 면역화된 마우스는 M2e에 대해 향상된 T 세포 반응 및 이중 조류 인플루엔자 바이러스의 치명적인 감염으로부터의 완전한 보호를 나타내었다.

[0062] [56] 미코박테리움 투베르쿨로시스 HSP70 (mHSP70) 단백질의 C-말단에 유전적으로 융합된 M2e의 4개의 탠덤 반복체를 포함하는 재조합 융합 단백질은 마우스에서 인플루엔자 바이러스의 복수종 스트레인들에 대해 방어를 나타냈다 (Ebrahimi, S. M., et al., Virology (2012) 430:63-72). M2e 펩타이드인 SLLTEVETPIRNEWGCRNDSSD는 다른 인플루엔자 단백질의 펩타이드와 백신으로서 B형 인플루엔자 M2e의 상동체인 BM2와 함께 양이온성 리포솜 전달 비히클에 컨쥬게이트되었다 (특허출원 US2010/086584A1). 백신은 마우스에서 M2e에 대한 면역 반응을 유도했다. M2e 펩타이드를 함유한 백신으로 예방접종된 마우스들은 치명적인 인플루엔자 바이러스 감염으로부터 보호되었다. M2e VLP가 보강된 비활성화된 인플루엔자 백신은 마우스에서 항원적으로 상이한 A형 인플루엔자 바이러스에 대해 개선되고도 오래 지속되는 교차 방어능을 부여하였다 (Song, J-M., et al., PNAS (2011) 108:757-761).

[0063] 면역우성 영역의 변형

[0064] [57] 본 발명은 헤마글루티닌(HA)의 구형 헤드 도메인의 면역우성 영역을 대체 에피토프로 치환하거나 대체 에피토프를 그 안에 삽입하여 이들 에피토프에 대한 숙주 면역 반응을 유도함으로써 결과된, 유전적으로 변형된 재조합 인플루엔자 헤마글루티닌(HA) 유전자 및 단백질을 포함한다. 이러한 에피토프 중 일부는 HA에 대한 bNAbs에 의해 인식된다. 본 발명은 또한 변형된 HA 단백질 및 변형된 MCSfms 갖는 유전자 및 이들 변형체의 조합에 관한 것이다.

[0065] [58] 일 구체예에서, 변형된 재조합 HA는 구형 헤드 도메인의 면역우성 항원 부위가 M2 단백질의 세포외 도메인(M2e)으로 대체되거나 M2e가 면역우성 항원 부위에 삽입된다.

[0066] [59] HA 단백질의 3차원 구조에 의해 유도된, 특정 표면 펩타이드 또는 면역력우성 부위의 그러한 표면 펩타이드는 동일한 HA의 다른 영역의 이중 펩타이드 또는 여러 개의 이중 펩타이드들 또는 인플루엔자 바이러스의 다른 서브타입 또는 다른 스트레인의 HA로부터의 이중 펩타이드 또는 여러 개의 이중 펩타이드들에 의해 대체된다. 이들 이중 펩타이드들은 면역우성 영역 내로 삽입될 수 있다. 또한, HA의 특정 표면 펩타이드 또는 펩타이드들은 HA와 관련이 없는 다른 단백질로부터의 이중 펩타이드로 대체되거나, 또는 그러한 펩타이드가 그 안에 삽입된다. 이러한 이중 펩타이드들은 천연 단백질이거나 인위적으로 설계된 것이다. 이들은 현재 알려진 항체에 의해 인식되지 않을 수도 있다.

[0067] [60] HA의 면역우성 영역은 HA의 구형 헤드 도메인의 표면에 존재한다. 변형될 이러한 표면 영역들은 표면-노출된 나선형, β -가닥 또는 루프이다. 이러한 표면 영역은 바람직하게는 면역우성 항원 부위 및 에피토프이거나 또는 그러한 항원 부위와 에피토프의 일부이거나 또는 그러한 항원 부위와 에피토프에 인접하는 것이 좋다.

[0068] [61] 일부 구체예에서, HA 구형 헤드 도메인의 면역우성 영역은 HA에 대한 bNAbs에 의해 인식되는 HA의 면역준우성 에피토프의 펩타이드에 의해 대체된다. 또 다른 구체예에서, HA 구형 헤드 도메인의 면역우성 영역은 일군의 인플루엔자 바이러스들을 중화시키는 항체에 의해 인식되는 면역준우성 에피토프의 펩타이드로 대체된다. 이러

한 면역준우성 에피토프들은 HA 줄기 도메인으로부터 선택되거나 또는 HA 융합 펩타이드 또는 HA 성숙 절단 부위를 포함한다.

- [0069] [62] 본 발명의 재조합 변형된 HA 단백질은 세포 배양체에서 발현되어 재조합 야생형 HA와 유사한 수준으로 세포 배양 배지에 분비된다. 이러한 설계는 인체에 감염되는 인플루엔자 바이러스 및 다른 포유동물 및 조류 종을 감염시키는 인플루엔자 바이러스를 비롯하여 (그러나 이에 한정되지 않음) A, B 및 C형 인플루엔자 바이러스의 모든 스트레인들의 인플루엔자 HA에 적용 가능하다. 일 구체예에서, 이들 변형된 HA 단백질은 인간 또는 동물에서의 인플루엔자 바이러스 감염에 대한 백신을 제조하기 위한 면역원으로서 사용될 수 있다.
- [0070] [63] 이들 면역우성 영역은 부위-지향적 돌연변이에 의해 변화될 수 있다. 항원부위(antigenic site)의 용매-노출된 잔기는 HA 단백질의 3차원 구조에 의해 확인된다. 항원부위의 용매-노출 잔기 또는 수개의 용매-노출 잔기들은 부위-지향적 돌연변이에 의해 변형되거나 또는 이들 잔기들의 특이적 변화를 포함하는 펩타이드로 대체된다. 새로운 부위는 원래 부위와 동일한 2차 구조를 갖는다.
- [0071] 이중 에피토프로 대체된 항원 부위가 있는 HA의 예
- [0072] [64] 도 3에 도시된 HA 구조체 기초하여, 예를 들어 HA의 줄기 도메인으로부터의 이중 펩타이드에 의한 구형 헤드 도메인의 면역우성 항원 부위를 대체 할 가능성을 설명하기 위해 HA 내의 위치들을 몇 개 선택한다.
- [0073] [65] 잔기 119 내지 122의 펩타이드 KTSS (도 5에서 WT로서의 넘버링)는 Sa 부위 잔기 124 내지 125 부근의 표면-노출된 나선형 구조이다. Sa 부위는 HA 삼량 체의 구형 헤드 도메인 쪽에 있고, 수용체 결합 부위로부터 멀리 떨어져 있다. 이 부위에서의 변형은 수용체 결합 부위를 변형시키지 않을 듯하고 따라서 변형된 HA에 의한 숙주 수용체 결합도 변화시키지 않을 듯하다. 상기 4-잔기 짜리 펩타이드 KTSS는 줄기 도메인의 bNAb 에피토프에 의해 대체된다. H3 HA의 부위 A에 상응하는 Ca2 부위의 잔기 137-142의 HAGAKS는 주요 면역우성 부위이다. 잔기 153 내지 164에서의 KKGNSYPKLSKS는 H3 HA의 부위 B에 상응하는 Sa 부위의 표면 루프이다. 잔기 153 내지 157의 이 펩타이드 KKGNS의 N-말단 부분은 일부 구조체에서 대체된다. Sb 부위의 잔기 184 내지 195의 TSADQSLYQNA는 HA 구형 헤드 도메인의 최정상부에서 나선(헬릭스)을 형성한다.
- [0074] [66] The Sa 부위의 표면 루프와 Sb 부위의 헬릭스는 HA의 구형 헤드 도메인 내에 서로 인접해 있으며 H3의 B 부위에 해당한다. Sa 부위와 Sb 부위는 함께, 루프 및 나선 구조를 갖는 구조적 에피토프를 수용할 수 있다. 그러나, 이들 부위들은 수용체 결합 부위 근방에 있기 때문에, 이들 펩타이드를 치환하면 수용체 결합 부위가 파괴될 수 있다. Ca2 부위의 잔기 213 내지 224에서의 EIAIRPKVRDQE는 2 개의 HA 단량체들 간의 인터페이스 부근의 루프이다. 이 루프는 H3 HA의 부위 D에 상응하는 인접한 HA 단량체의 β -가닥들과 접촉되어 있다. 루프의 일부는 표면에 노출되어 있다. Ca2 부위는 이 펩타이드 내에 위치한다. 일부 구조물에서, 이 펩타이드가 부분적으로 대체된다.
- [0075] [67] 이에 더해, 이중 펩타이드는 도 5에 도시된 바와 같은 임의의 다른 항원 내 또는 그 근처 및 도 1 내지 도 3에 도시된 바와 같은 표면 루프 또는 헬릭스 내 또는 그 근처에 배치될 수 있다. 뿐만 아니라, 이중 펩타이드는 항원 부위의 잔기들의 어떠한 제거 없이도 이들 중 어떤 위치로도 삽입될 수 있다.
- [0076] [68] bNAb CR8020 에피토프 bNAb CR8020 에피토프는 줄기 도메인에 위치한다. 도 5 및 도 6에 도시된 바와 같이, CR8020 에피토프는 줄기 도메인의 기부 근처의 5-가닥 β -시트의 최외곽 β -가닥 (HA2 잔기 30 내지 36, H3 HA의 EGTGQAA, SEQ ID NO: 26) 및 융합 펩타이드의 C-말단 부위 (HA2 잔기 15 내지 19, H3 HA의 EGMID, SEQ ID NO: 25)로 이루어진 2 가지 주요 성분을 갖는다. 이들 2 성분들은 1차 구조에서 10 개의 아미노산 잔기만큼 분리되어 있다. 2 개의 에피토프 성분이 티로신 (Y)에 의해 연결된 복합 CR8020 에피토프 EGMIDYEGTGQAA (SEQ ID NO: 27)가 설계된다. 이 복합 CR8020 에피토프는 이중 펩타이드로 사용되어 면역우성 부위를 대체한다.
- [0077] [69] 또 다른 bNAb1C9 에피토프는 줄기 도메인에 존재하는 HA2의 N-말단에 있는 융합 펩타이드이다. GIFGAIAGFIEG (SEQ ID NO: 36)의 변형된 1C9 에피토프 펩타이드는 면역우성 부위에 디스플레이될 펩타이드로서 설계되었다. I1C9로 표시된 변형된 1C9 펩타이드는 bNAb 1C9에 의해 인식되는 H5 융합 펩타이드의 위치 2에서 류신(L)이 이소류신(I)으로 치환된 것이다. 이 I1C9 융합 펩타이드는 돼지 H1 HA 및 일부 H6 및 H9 바이러스의 HA에 존재한다. 이 I1C9 펩타이드는 다른 면역우성 부위에 위치한다. I1C9를 사용하여 시프트된 위치를 갖는 항원 부위에 대한 다중 구조물이 만들어졌다.
- [0078] [70] bNAb FI6 에피토프는 HA 전구체 HA0의 3차 구조 상에서 연속적인 표면을 형성하지만 1차 구조에서는 분리된 2개의 펩타이드를 함유하는 구조적 에피토프이다(도 5). 하나의 FI6 에피토프 펩타이드 RKKRGLFGAIAGFIE는 성숙 절단 부위 및 HA0의 융합 펩타이드이다. 또 다른 FI6 펩타이드 KESTQKAIDGVTKVNS는 HA2에서 코일형 코일

(coiled-coil) 구조를 갖는다. 이들 2개의 FI6 에피토프 펩타이드가 구형 헤드 도메인 내에 위치하는 구조물이 만들어졌다. FI6 에피토프 펩타이드 RKKRGLFGAIGFIE는 잔기 153 내지 164의 Sa 부위에 위치하고 코일형 코일 FI6 에피토프 펩타이드 KESTQKAIDGVTNKVNS는 잔기 184 내지 195의 Sb 부위에 위치한다. 또 다른 구조물에서 FI6 에피토프 펩타이드 RKKRGLFGAIGFIE는 두 번째 FI6 에피토프 펩타이드 없이 Sa 부위에 위치한다.

- [0079] [71] 많은 bNAb 에피토프가 특징화되었다. 이들 중 어느 것이든 전술한 바와 같이 구형 헤드 도메인의 면역우성 항원 부위 또는 표면 루프에 또는 그 근처에 위치시킬 수 있다.
- [0080] [72] 다른 구체예에서, 이중 펩타이드는 M2 단백질의 세포외 도메인의 펩타이드(M2e 펩타이드)이다. M2e 펩타이드는 A형 인플루엔자 바이러스들 사이에서 보존되고 덜 가변적이다. 비록 A형 인플루엔자 바이러스와 B형 인플루엔자 바이러스로부터의 M2e 펩타이드가 다르긴 하지만, M2e 펩타이드는 B형 인플루엔자 바이러스들 사이에서도 보존되고 덜 가변적이다.
- [0081] [73] 일부 구체예에서, 이중 펩타이드는 부위-지향된 돌연변이에 의해 면역우성 항원 부위를 특이적으로 변화시키는 인공적으로 설계된 펩타이드이다. 일부 구체예에서, 인공 펩타이드는 면역우성 항원 부위와 이중 펩타이드의 바람직한 특징을 조합한 것이다. 일부 구체예에서, 인공 펩타이드는 HA 구형 헤드 도메인의 3 차원 구조를 유지하기 위해 원래의 면역우성 항원 부위의 잔기 및 bNAb와의 상호반응을 위한 잔기를 갖는다. 이러한 인공 펩타이드는 HA 3차원 구조를 기반으로 합리적으로 설계되거나 무작위로 생성된 펩타이드의 라이브러리를 스크리닝 함으로써 만들어진다.
- [0082] [74] 배콜로바이러스 발현 시스템 또는 포유동물 발현 시스템을 이용하는 발현을 위해 Ca2 부위나 또는 Sb 부위에 M2e 펩타이드 SLLTEVETPTRNGWECKCSDS가 위치된 구조물이 제작되었다. M2e 펩타이드는 상이한 인플루엔자 스트레인들로부터의 컨센서스 서열에 기초하여 더욱 최적화될 수 있다.
- [0083] 변형된 성숙 절단 부위를 갖는 HA
- [0084] [75] 본 발명은 결과적인 HA가 상이한 여러 부류의 프로테아제에 대해 민감하고 인플루엔자 바이러스의 모든 스트레인들의 천연(native) HA 성숙 절단 부위들을 절단하는 트립신-유사 프로테아제에 대해 내성을 갖도록, HA 성숙 절단 부위(MCS)의 프로테아제 감수성을 변화시킨 구체예들을 포함한다. 변형된 MCS는 인플루엔자 바이러스의 공지의 천연 숙주에서는 존재하지 않는 특이적인 프로테아제에 의해 인식되도록 설계되었다. 결과적인 HA는 이로 인해, 인플루엔자 바이러스의 모든 자연 숙주에서 성숙에 대해 내성을 갖게 된다. 이러한 변경된 성숙 절단 부위를 인식하는 이 특이적인 프로테아제의 존재하에, 이들 결과적인 HA는 HA1 및 HA2를 함유하는 성숙한 형태로 절단된다. 이 특이적 프로테아제의 존재 하에서 결과적인 HA로부터 만들어진 재조합 인플루엔자 바이러스는 성숙한 HA를 형성하고 천연 인플루엔자 숙주에게 감염성이 된다. 그러나, 감염된 천연 숙주에서 복제된 바이러스 자손은, 천연 숙주에서 적절한 프로테아제가 없기 때문에 감염되지 않는다.
- [0085] [76] H1 HA 분자는 종종 MCS에서 단일 염기성 잔기를 갖는다. 이 단일 염기성 잔기는 통상 절단 결합(scission bond)의 N-말단 아르기닌 잔기이다. H1 HA의 MCS에 아르기닌(R) 또는 라이신(K)과 같은 복수개의 염기성 잔기들을 첨가하면 변형된 HA를 함유하는 재조합 인플루엔자 바이러스의 감염성이 증가한다. 천연 H1 MCS를 다염기(polybasic) 잔기를 함유하는 H5 MCS로 대체하면 변형된 HA를 함유하는 재조합 인플루엔자 바이러스의 감염성이 증가한다 (Kong, W-p., *et al.*, *PNAS*(2006) 103:15987-15991). 이들 다염기 잔기는 많은 세포내 트립신-유사 프로테아제에 의한 절단에 대해 보다 민감하다. 일부 구체예에서, H1 HA MCS는 H5 HA MCS 또는 다염기 잔기로 대체됨으로써 변형된다. H1 HA MCS는 또한 HA1의 마지막 잔기(아르기닌)와 HA2의 첫 번째 잔기(글리신) 사이에 다염기 잔기를 삽입함으로써 변형된다.
- [0086] [77] 유전적으로 변형된 재조합 인플루엔자 HA 유전자 및 변형된 MCS 서열을 갖는 단백질 역시도 개시된다. 천연 MCS는 또 다른 HA의 MCS로부터의 서열들로 대체한다. 상기 변화는 아마도 결과적인 HA를 미확인 숙주 프로테아제에 의한 성숙 절단에 보다 민감하게 만듦으로써, 재조합 인플루엔자 바이러스를 감염성이 더 높게 만드는 것으로 알려진 다른 HA의 또 다른 MCS로 HA의 천연 MCS를 대체하는 것을 포함한다. 바람직한 구체예에서, H1 HA의 천연 MCS는 H5 HA의 MCS로 대체되거나 아르기닌(R) 및 라이신(K)의 다염기 잔기들로 대체된다.
- [0087] [78] 인자 Xa 및 엔테로키나제와 같은 몇몇 포유동물 프로테아제는 절단 결합의 N-말단 부분에 전적으로 위치하는 그들의 대응하는 절단 인식 부위를 갖는다. 이들의 절단 인식 부위는 또한 천연 HA 성숙 절단 부위를 대체하는데 사용될 수도 있다. 천연 HA의 경우에서와 같이, 이들 프로테아제 절단 부위의 절단 결합의 C-말단 쪽은 글리신(G)이다. 트롬빈 절단 부위 역시도 절단 결합의 C-말단 쪽에 글리신 (G)을 갖는다. 트롬빈 절단 후, 글리신 (G)의 새로운 N-말단이 생성된다. 이들 프로테아제는 모두 수많은 천연 인플루엔자 숙주 내에 존재한다. 그러나

이들 프로테아제는 대개 재조합 HA 생산에 사용되는 세포 배양 배지 또는 세포주에는 아예 존재하지 않거나 미량 존재할 뿐이다.

[0088] [79] 담배 식각 바이러스(TEV)의 TEV 프로테아제는 ENLYFQG의 절단 부위를 인식하고 글루타메이트(Q)와 글리신(G) 사이를 절단한다. TEV 프로테아제 절단에 의해 생성된 유리 N-말단은, HA 성숙 후 HA2의 N-말단 잔기와 동일한 글리신(G)이다. 천연 HA의 MCS와 달리, TEV 절단 부위에는 염기성 잔기가 없다. 일부 구체예에서, 전체 HA MCS는 TEV 절단 부위로 대체되거나, MCS의 몇몇 잔기가 TEV 절단 부위로 대체되거나, 또는 예컨대 절단 결합에 대한 아르기닌(R) N-말단 만이 TEV 절단 부위에 의해 대체된다. TEV 절단 부위에 의해 대체된 MCS를 갖는 HA 구조물들은 배컬로 바이러스 발현 시스템 또는 포유동물 발현 시스템을 사용하여 발현되며, 야생형 HA와 동일한 수준으로 발현되는 구조물이 동정되었다.

[0089] [80] 변형된 MCS를 절단하는 TEV 프로테아제의 존재하에, 상기 변형된 HA는 성숙해지고 융합 펩타이드의 천연 N-말단을 갖는 HA1 및 HA2로 전환된다. 성숙한 변형된 재조합 HA는 상응하는 숙주 세포 수용체에 결합하여 막 융합을 수행하는 능력을 갖는다. TEV 프로테아제가 없는 경우, 새로 합성된 상기 변형된 HA는 절단되지 않은 HA0 전구체로 남아있다. 융합 펩타이드의 유리 N-말단의 부재는 상기 HA가 막 융합하는 것을 방지한다. 감염된 숙주 세포에서 생성된 자손 HA 단백질은 HA0로 남아 있으며 숙주 내 TEV 프로테아제의 결여로 인해, 성숙한 기능성 형태로 가공되지 않는다. HA0를 갖는 새롭게 제조된 바이러스는 막 융합 능력이 없으므로 비-감염성이다.

[0090] 항원 부위와 성숙 절단 부위의 조합

[0091] [81] HA의 구형 도메인 내 항원 부위의 변화와 HA의 줄기 도메인에서의 MCS의 변화는 여하한 종류의 조합일 수 있다. 특정한 이중 펩타이드가 상이한 변경된 성숙 절단 부위를 갖는 HA 단백질의 구형 헤드 도메인의 항원 부위에 도입될 수 있다. 또한, 2중 이상의 이중 펩타이드가 상이한 여러 항원 부위에서 단일 HA 단백질에 도입될 수 있다. 일 구체예에서, 유전자 변형된 재조합 인플루엔자 HA는 MCS로서 TEV 프로테아제 부위 및 HA의 구형 헤드 도메인의 항원 부위를 대체하는 M2e 펩타이드를 갖는다. 또 다른 구체예에서, 유전자 변형된 재조합 인플루엔자 H1 HA는 H1 천연 MCS 대신 H5 MCS 및 HA의 구형 헤드 도메인의 항원 부위를 대체하는 M2e 펩타이드를 갖는다.

[0092] 생산 방법

[0093] [82] 변형된 HA의 생산을 위해, 단백질은 세포에 의해 가용성 형태로 세포 배양 배지에 분비되고, 통합 막 단백질도 아니고, 세포막이나 세포 표면에 부착되지도 않는다. 일부 구체예에서, 천연 HA 시그널 서열은 곤충 세포 및 포유동물 세포에서의 분비에 사용된다. 다른 구체예에서, 천연 HA 시그널 서열은 곤충 세포에서의 분비를 위해 곤충 세포 시그널 서열로 대체되거나 또는 포유동물 세포에서의 분비를 위해 포유동물 시그널 서열로 대체된다. HA 막관통 도메인 및 세포내 도메인 대신에, HA 세포외 도메인의 C-말단에 프로테아제 절단 부위가 배치되고, 이어서 에 HA 삼량체를 안정화시키는 박테리오파지 T4 피브리틴 유래의 "폴드온(foldon)"서열, 그리고 정제를 용이하게 하는 C-말단 His-태그가 배치된다. 재조합 HA 단백질은 시그널 서열이 제거된, 전장의 절단되지 않은 전구체 HA0로서 제조되거나 또는 HA2의 N-말단에 야생형 융합 펩타이드를 갖는 HA1 및 HA2 서브 유닛을 함유하는 성숙한 형태로서 제조된다. 바람직한 구체예에서, 최종 정제된 유전자 변형된 HA는 야생형 HA로서 삼량체를 형성한다.

[0094] [83] 설계된 재조합 인플루엔자 HA의 유전자는 드 노보 유전자 합성에 의해 만들어진다. 유전자 합성 기술은 잘 확립되어 있으며 광범위하게 검토되었다 (Kosuri, S., 및 Church, G. M., *Nat Methods*(2014) 11:499-507). 길이가 10kb가 넘는 유전자는 시중 공급자에 의해 일상적으로 제조되어 왔다. 유전자 합성은 발현 숙주 또는 특정 유전 공학적 요구에 따라 코돈-최적화 기회의 특정 HA 서열을 변형할 수 있는 능력을 부여한다. 제한효소 부위(restriction sites)는 합성된 HA 유전자의 특정 위치에서 서로 다른 구조물 사이에서 HA 단편의 교환을 용이하게 하도록 설계된다. 시그널 서열 및 프로테아제 부위, 폴드온 서열 및 C-말단 His-태그를 갖는 합성된 HA 유전자는 잘 확립된 프로토콜로 배컬로바이러스 게놈에 혼입되어, 변형된 HA 단백질의 발현이 배컬로바이러스 폴리헤드린 프로모터에 의해 지시되도록 한다. 다른 구체예에서, 포유동물 세포에서의 발현을 위해 동일한 세트의 HA 유전자들이 포유동물에 클로닝된다.

[0095] [84] 일부 구현예에서, HA 구조물은 배컬로바이러스 벡터를 사용하여 곤충 세포에서 발현하기 위해 코돈-최적화된다. 또 다른 구체예에서, HA 구조물은 비제한적인 예로서 CHO 세포 (차이니스 햄스터 난소 세포) 및 HEK 293 세포 (인간 배아 신장 293 세포)와 같은 포유동물 세포에서의 발현을 위해 코돈-최적화된다. 모든 유기체에 대한 코돈 최적화는 OPTIMIZER (월드 와이드 웹 상의 genomes.urv.es/OPTIMIZER/에 위치)와 같은 온라인 웹 서버

인 DNASTAR, Inc. (3801 Regent Street, Madison, WI 53705 USA)의 Lasergene 소프트웨어 패키지와 같은 상용 소프트웨어를 사용하거나, 또는 독점적인 알고리즘을 사용하는 유전자 합성 서비스 제공자에 의해 수행될 수 있다. 코돈 최적화에는 G/C 함량, 암호 스프라이스 부위 및 RNA 불안정화 서열 요소의 제거, 안정한 RNA 2차 구조의 회피가 고려된다. 코돈은 또한 Codon Usage Database(월드 와이드 웹 상의 kazusa.or.jp/codon/에 위치)에 기반하여 수동으로 조정되기도 한다. 코돈 축퇴로 인해, 유전자 서열은 코딩된 아미노산 서열을 변화시키지 않고 변경될 수 있다. 제한효소 부위는 아미노산 서열을 변화시키지 않고 코돈을 변화시킴으로써 특정 위치에 도입된다. 이러한 방법 중 어느 하나를 사용함으로써, 특정 유기체에 대한 코돈 최적화가 되거나 되지 않은 유전자 서열이, 컴퓨터 알고리즘을 사용하거나 또는 수동적으로 단백질 서열의 역번역함으로써 일상적으로 생성된다.

- [0096] [85] 다른 구체예는 확립된 세포 배양법을 이용하여, MCS가 TEV 프로테아제 인식 부위로 변경된 유전자 변형된 HA 유전자로부터 재조합 비-감염성 인플루엔자 바이러스를 제조하는 방법을 포함한다. 클로닝된 cDNA를 갖는 플라스미드 또는 RNA 세그먼트로부터 인플루엔자 바이러스의 생산을 허용하는 역-유전학 시스템(reverse-genetics systems)이, 헬퍼 바이러스를 사용하거나 사용하지 않고 개발되었다 (Luytjes, W., *et al.*, *Cell*(1989) 59:1107-1113; Neumann, G., *et al.*, *PNAS*(1999) 96:9345-9350; Fodor, E., *et al.*, *J. Virol*(1999) 73:9679-9682; de Wit, E., *et al.*, *J. Gen Virol*(2007) 88:1281-1287). 감염성 인플루엔자 바이러스는 인플루엔자 RNA 세그먼트 또는 인플루엔자 RNA의 cDNA를 갖는 플라스미드로 포유동물 세포를 일시적으로 형질감염시킴으로써 만들어진다. 세포 배양 배지에서 분리된 이들 바이러스는 백신 제조를 위한 살아있는 전염성 인플루엔자 바이러스를 생산하기 위해 배아가 있는(embryonated) 달걀을 감염시키는데 이용된다. 일부 인플루엔자 RNA 세그먼트는 외래 유전자로 대체된다. 재조합 A형 인플루엔자 바이러스는 HA 단백질을 대체하는 C형 인플루엔자 바이러스 HEF 단백질을 사용하여 만들어졌다 (Gao, Q., *et al.*, *J. Virol*(2008) 82:6419-6426). 이들 구체예에서, 변형된 HA의 cDNA 또는 RNA는 표준 분자 기술 및 유전자 합성에 의해 제조된 다른 7개의 인플루엔자 RNA 또는 다른 7개의 RNA 세그먼트의 cDNA로 포유동물 세포에 동시-형질감염된다. 확립된 방법 또는 유사한 방법을 이용하여, 다수의 동일하거나 상이한 HA 단백질을 단일 비리온 내로 패키징할 수 있다 (Uraki, R., *et al.*, *J. Virol*(2013) 87:7874-7881).
- [0097] [86] 유전적으로 변형된 HA 유전자로부터 재조합 감염성 인플루엔자 바이러스를 만드는 방법은 전술한 바와 같은 확립된 세포 배양 방법을 사용한다. 이러한 구체예에서, MCS에서의 H5 HA MCS 또는 다염기 서열을 갖는 유전적으로 변형된 HA 유전자는 상기 기재된 표준 역 유전 시스템에 의해 제조된 다른 7개의 인플루엔자 RNA 또는 다른 7개의 RNA 세그먼트의 cDNA로 포유동물 세포에 동시-형질감염된다. 이들 변형된 성숙 절단 부위는 인플루엔자 천연 숙주 또는 세포 배지에서 HA 성숙의 효율을 증가시킨다 (Kong, W-p., *et al.*, *PNAS*(2006) 103:15987-15991). 성숙한 기능성 HA 단백질은 숙주 세포에 결합하여 숙주 세포막과 융합하는 능력을 갖는다.
- [0098] [87] 다른 구현예는 포유동물 세포, 곤충 세포 및 식물 세포에서 확립된 세포 배양 방법을 사용하여 유전자 변형된 HA 유전자로부터 재조합 인플루엔자 바이러스-유사 입자 (VLP: virus-like particles)를 제조하는 방법을 포함한다 (Chen, B. J., *et al.*, *J. Virol*(2007) 81:7111-7123; Smith, G. E., *et al.*, *Vaccine*(2013) 31:4305-4313; D'Aoust, M. A., *et al.*, *Plant Biotech*(2010) 8:607-619).
- [0099] [88] 다른 구체예는 확립된 방법을 사용하여 유전자 변형된 HA 유전자로부터 DNA 백신을 제조하는 방법을 포함한다 (Jiang, Y., *et al.*, *Antiviral Res*(2007) 75:234-241; Alexander, J., *et al.*, *Vaccine*(2010) 28:664-672; Rao, S. S., *et al.*, *PLoS ONE*(2010) 5:e9812).
- [0100] [89] 본원에 개시된 발현 결과는 H1 HA의 MCS가, 결과적인 HA의 발현에 영향을 미치지 않고 변형될 수 있음을 보여준다. H5 HAMCS 또는 다염기 MCS를 갖는 구조물들은 야생형 H1 HA와 동일한 수준으로 발현되는 것으로 입증되었다. 또한, HA의 MCS는 염기성 잔기를 갖지 않는 TEV 프로테아제 절단 부위로 대체된다. TEV 프로테아제 절단 부위의 위치를 변화시킴으로써, 야생형 H1 HA와 동일한 수준으로 발현되는, MCS로서 TEV 절단 부위를 갖는 구조물이 만들어졌다.
- [0101] [90] 발현 결과는 HA1의 구형 헤드 도메인의 많은 항원 부위가 동일한 HA 또는 상이한 HA의 줄기 도메인으로부터의 이중 펩타이드로 대체될 수 있다는 것을 보여준다. 일부 구조물은 야생형 HA와 동일한 수준으로 발현되는 반면, 다른 구조물은 훨씬 덜 발현된다. 구형 헤드 도메인 내 이중 펩타이드의 성질 및 이중 펩타이드의 위치는 각각의 결과적인 HA 구조물의 발현 수준에 유의한 영향을 미친다. 가장 최적한 발현 구조물을 동정하기 위해 동일한 방법으로 더 많은 구조물을 만들고 테스트할 수 있다.
- [0102] [91] 또한, HA 구형 헤드 도메인의 특정 면역우성 항원 부위가 인플루엔자 M2 단백질의 M2e 에피토프에 의해 대

체된 재조합 HA가 제조된다. 가능하게는, 다른 인플루엔자 단백질의 임의의 에피토프가 HA 구형 헤드 도메인의 면역우성 항원 부위를 대체하는 대체 펩타이드로 작용할 수 있다. 또한, 인플루엔자 바이러스와 관련이 없는 다른 단백질로부터 이중 대체 펩타이드가 유래될 수 있다. 대체 펩타이드의 고유한 구조적 특징을 가능한 한 많이 유지하여 대체 펩타이드를 숙주 면역계에 제공할 수 있도록 HA 구형 헤드 도메인의 특정 항원 부위를 대체 펩타이드로 선택할 수 있다.

[0103] 배큘로바이러스 발현 시스템

[92] 약 30년 전 도입된 이래 (미국특허 4,745,051; Summers, M. D., and Smith, G. E., (1987) A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experiment Station Bulletin No. 1555.), 배큘로바이러스 발현 벡터 시스템 (BEVS)은 세포내 단백질, 막 단백질 및 분비된 단백질을 비롯한 많은 상이한 유형의 인간 및 바이러스 단백질을 발현하는데 사용되어 왔다. BEVS는 바이러스-유사 입자(VLP)를 생산하는 데 사용되어 왔다. BEVS는 곤충 세포를 숙주로 사용하여 포유동물 세포와 유사한 단백질의 번역 후 변형을 제공하는 진핵세포 발현 시스템이다 (Jarvis, D. L., "Baculovirus Expression Vectors" in The Baculoviruses, ed. Miller, L. K. (1997) pp. 389-420 Plenum Press, New York). 오토그라파 칼리포니카(*Autographa californica*) 핵 다면체병 바이러스(AcNPV: *Autographa californica* nuclear polyhedrosis virus)에 근거한 BEVS는 잘 확립되어 있다. 많은 상업용 키트가 재조합 배큘로바이러스를 생산하는데 이용 가능하다. BEVS는 미국에서 판매된 재조합 백신 생산에 성공적으로 사용되어 왔다. 2종의 FDA 승인된 백신, 즉 자궁 경부암에 대한 비-감염성 바이러스-유사 입자(VLP) 형태의 재조합 인간 유두종 2가 바이러스(타입 16 및 18)인 Cervarix™ (특허출원 W02010/012780 A1) 및 인플루엔자에 대한 막 결합 헤마글루티닌 전구체(HA0)으로 만들어진 3가 인플루엔자 백신인 Flublok®(미국특허 5,762,939 A 및 5,858,368 A)이 BEVS를 사용하여 제조된다.

[0105] [93] BEVS를 이용하여 표적 단백질을 발현시키기 위해, 먼저 표적 단백질을 코딩하는 관심 유전자를 배큘로바이러스 폴리헤드린 유전자에 인접한 배큘로바이러스 서열을 포함하는 대장균 플라스미드인 전달 벡터로 서브클로닝한다. 폴리헤드린은 야생형 배큘로바이러스에서 풍부한 바이러스성 단백질로서 세포 배양시 배큘로바이러스 증식에 필요하지 않다. 전달 벡터는 표준 분자 생물학 기술을 사용하여 대장균에서 증식되고 분리된다 (Sambrook, J., *et al.*, (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor). 관심 유전자를 갖는 전달 벡터는 대장균 또는 곤충 세포에서 재조합 배큘로바이러스 게놈을 생성하기 위해 배큘로바이러스 게놈 DNA와 재조합된다. 재조합 배큘로 바이러스 게놈은 재조합 배큘로바이러스의 생성 및 표적 단백질의 발현을 지시한다. 본래의 방법은 곤충 세포의 바이러스 게놈 DNA와 전달 벡터의 폴리헤드린 유전자에 인접한 동일한 서열들 간의 동종 재조합에 의존한다. 분리된 전달 벡터 DNA는 배큘로바이러스 게놈 DNA와 함께 곤충 세포 내로 형질감염된다. 재조합 배큘로바이러스는 폴리헤드린 결핍을 기준으로 선택된다. Clontech의 BacPAK™ Baculovirus Expression System 및 Millipore의 BacMagic™ System과 같은 현재의 상업용 키트는 재조합 없이는 생존 가능한 바이러스를 생성하지 않는 선형화된 배큘로바이러스 게놈 DNA를 이용한다. 이들 키트는 재조합 배큘로바이러스를 고효율로 생산하고 비 재조합 바이러스에 의한 오염은 거의 없다. 바이러스 게놈에 대한 재조합은 Gateway 재조합 반응 (BaculoDirect™ Baculovirus Expression System, Life Technologies)을 이용하여 시험관내에서 만들 수 있다. 재조합 바이러스를 만드는 또 다른 방법은 대장균에서 부위-특이적 전이를 통한 방법이다(Bac-to-Bac System, Life Technologies). pFastBac-기반 전달 벡터는 대장균 (미국 특허 제 5,348,886 호)에서 바크미드(bacmids: 배큘로바이러스의 게놈을 함유하는 커다란 플라스미드)를 만드는 트랜스포존을 함유한다. 바크미드에 대한 관심 유전자의 재조합은 곤충 세포에 대한 바크미드의 형질감염에 앞서 폴리메라제 연쇄 반응(PCR)에 의해 쉽게 확인된다.

[0106] [94] 재조합 배큘로바이러스는 곤충 세포 배양체에서 증식한다. 곤충 세포를 감염시키기 위해 소량의 바이러스가 사용된다. 수일 후, 증폭된 바이러스를 함유하는 조건화된 배지를 바이러스 스톱으로서 수확한다. 이 증폭 과정은 대용량 바이러스 스톱을 생성하기 위해 종종 여러 번 반복된다. 바이러스 스톱은 대개 수개월 및 심지어 수년간 냉동되어 암실에서 저장된다. 바이러스 스톱의 보존을 위해 5-10% 소태아 혈청 (fetal bovine serum, FBS)이 바이러스 스톱에 보강되는데 이에 의해 이들의 저장 수명이 연장되는 것으로 믿어진다. 바이러스 스톱은 때때로 -70℃에서 장기간 보관을 위해 동결된다. 바이러스는 또한 증폭되어 배큘로바이러스에 감염된 곤충 세포(BIIC)로서 저장된다 (Wasilko, D. J., *et al.*, *Prot Exp Purif* (2009) 65:122-132). 감염된 곤충 세포는 그의 용해 전 BIIC 스톱으로서 수확되어 표준 세포 동결 과정에 따라 동결된다. BIIC 스톱은 액체 질소에서 또는 -65℃ 내지 -85℃의 초저온에서 장기간 저장되며 단백질 발현을 위해 곤충 세포를 감염시키는 바이러스 스톱으로서 사용된다. 동결된 BIIC 스톱은 액체 형태의 바이러스 스톱에 비해 더 긴 저장 시간을 제공한다.

- [0107] [95] BEVS를 위해 일반적으로 사용되는 곤충 세포주는 양배추 루퍼 에서 유래된 T.ni 세포(*Trichoplusia ni*) 또는 Hi5 및 가을 아충 (*Spodoptera frugiperda*)으로부터 유래된 SF9 및 SF21 세포이다. 누에 (*Bombyx mori*), 벌집 나방 (*Galleria mellonella*), 및 집시 나방 (*Lymantria dispar*) 중으로부터 유래된 다른 곤충 세포들도 사용되어 왔다. SF9, SF21, 및 Hi5 (또는 T. ni) 세포는 무혈청 배지에서 현탁 배양에 적응시켰다. 이들 세포주 및 배지는 많은 상업적 출처로부터 입수 가능하다. 이들 세포주의 세포 배양은 22-28℃ 범위의 온도에서 가스를 보충하지 않고 주변 대기에서 셰이커 배양기에서 최대 1 또는 2 리터의 소용량으로 셰이커 플라스크에서 일상적으로 유지된다. 세포 배양은 교반 탱크 바이오 리액터 또는 일회용 생물반응기에서 증폭된다. 생물반응기에서 대규모 곤충 세포 배양을 위한 조건은 잘 확립되어 있다 (WAVE Bioreactor Systems - 세포 배양 절차, GE Healthcare). 산소 보충은 또한 대규모 곤충 세포 배양에서 세포 밀도를 증가시키기 위해 일상적으로 사용된다.
- [0108] [96] 특정 표적 단백질의 발현을 위한 일련의 조건을 시험하여 감염 후 바이러스 대 세포 비율 및 수확 시간을 변화시킴으로써 상이한 세포주를 사용하여 표적 단백질의 발현을 최적화한다. SF9, SF21 및 Hi5 (또는 T. ni) 세포는 BEVS 단백질 발현에 사용되는 가장 일반적인 세포주이다. 표적 단백질은 이들 세포주들 중 하나에서 다른 것보다 더 잘 발현될 수 있다. 일부 보고서에 의하면 Hi5 (또는 T. ni) 세포가 특정 분비 단백질의 발현을 더 많이 한다는 것을 보여준다. 일반적으로 감염 다중도 (MOI)로 알려진 바이러스 대 세포 비율을 테스트하여 표적 단백질 발현을 위한 최상의 감염 조건을 구한다. 감염 후 다양한 시점에서 배양 샘플을 채취한다. 세포와 조건화 배지를 원심분리 또는 여과에 의해 분리한다. 단백질 발현 수준은 표준 방법에 의해 결정된다. 세포 성장과 배양 조건에 대한 정보를 제공하는 세포 밀도, 세포 생존능 및 세포 크기에 대해 배양물을 모니터링한다. 포도당 수준, 용존 산소 및 배양물의 pH를 때때로 배지에서 영양소의 소비 관점에서 모니터링한다. 최상의 표적 단백질 발현을 유도하는 조건을 단백질의 대량 생산을 위해 선택한다.
- [0109] [97] 바이러스 스톡 및 MOI의 역가를 결정하는 것이 BEVS를 이용한 바이러스 증폭 및 단백질 발현의 감염 조건을 결정하는데 널리 사용되기는 하지만, 전술한 TIPS 방법은 바이러스 증폭 및 단백질 발현에 대한 최상의 감염 조건을 결정하는 보다 빠른 방법을 제공한다. 곤충 세포는 재조합 배컬로바이러스 감염 후 크기가 커진다. 세포 분열 역시 감염 후에 중단된다. 재조합 배컬로바이러스는 감염 후 약 48 시간 동안 세포 용해를 일으킨다. 세포 분열은 세포를 배양체의 세포 밀도로서 계수함으로써 모니터링된다. 세포 용해는 배양체 내의 살아있는 세포와 죽은 세포를 계수함으로써 세포 생존능으로써 모니터링된다. 생존 가능한 세포 크기는 세포 지름으로서 측정된다. 세포 밀도, 세포 생존능 및 생존 가능한 세포 크기는 세포 계수 장치의 많은 모델에 의해 일상적으로 측정된다. 세포 밀도와 세포 생존능은 혈구 미터를 사용하여 수동으로 측정할 수도 있다. 총체적으로, 세포 밀도, 세포 생존능 및 생존 가능한 세포 크기는 감염 동력학에 대한 정보를 제공한다. TIPS 방법은 특정 바이러스 스톡을 사용하여 특정 표적 단백질의 발현을 위한 최적 조건을 결정하기 위해 감염 동력학을 사용한다. 이에 의해, MOI 산출을 위한 바이러스 역가의 시간 소모적인 측정 필요성이 제거된다. TIPS 방법의 사용에 의해, 일관된 바이러스 증폭 및 단백질 발현이 일상적으로 달성된다.
- [0110] [98] 곤충 세포 외에도, 재조합 배컬로바이러스에 감염된 살아있는 곤충이 많은 분비 단백질 및 막 단백질을 발현시키는데 사용되어왔다. 성숙한 HA는 담배 벌레(*Heliothis virescens*)의 유충에서 발현되었다 (Kuroda, K., et al., J. Virol (1989) 63:1677-1685).
- [0111] [99] 봄빅스 모리 핵 다면체병 바이러스(BmNPV: *Bombyx mori* nuclear polyhedrosis virus)와 같은 다른 유형의 배컬로바이러스를 이용한 발현 시스템도 개발되었다. BmNPV는 더 좁은 숙주 범위를 가지며 야생에서 유해 곤충으로 자라지 않는다는 측면에서 AcNPV보다 더 우수한 생물학적 안전성 프로파일을 갖는 것일 수 있다. BmNPV-기반 배컬로바이러스 발현 시스템은 세포 배양 및 누에(*Bombyx mori*) 유충에서 기능성 단백질을 발현하는데 사용되어 왔다 (Maeda, S., et al., Nature(1985) 315:592-594).
- [0112] [100] BEVS는 단백질 복합체 및 VLP를 생산하는데 일상적으로 사용되어 왔다. 두 개 이상의 유전자가 다중 단백질의 발현을 위해 단일 벡터에 클로닝된다. 각각 단일 단백질의 발현을 지시하는 두 개 이상의 바이러스 스톡이 단백질 복합체 또는 VLP를 생성하기 위한 곤충 세포의 동시 감염(co-infection)에 사용된다. 인플루엔자 VLP는 배컬로바이러스 발현 시스템을 사용하여 만들어졌다 (Bright, R. A., et al., PLoS ONE(2008) 3:e1501).
- [0113] 배컬로바이러스 발현을 위한 HA 전달 벡터
- [0114] [101] BEVS를 사용하여 발현되는 유전자는 일반적으로, 배컬로바이러스-유도된 곤충 세포 사멸 전에 단백질 발현을 유도하는 배컬로바이러스의 강력한 후기 프로모터인 폴리헤드린 프로모터의 제어하에 있다. 다른 초기 또는 후기 프로모터 역시도 단백질 발현에 사용된다. 곤충 세포에 의해 재조합 표적 단백질을 세포 배양 배지에 분비하기 위해, 시그널 펩타이드를 코딩하는 DNA 세그먼트가 관심 유전자에 대한 5'-말단에 인-프레임

조작된다. 곤충 세포에 일반적으로 사용되는 두 가지 시그널 펩타이드는 꿀벌 멜리틴 시그널 서열 또는 AcNPV 엔벨롭 표면 당 단백질 GP67 시그널 서열이다. 시그널 서열 데이터베이스(Signal Sequence Database: World Wide Web상의 signalpeptide.de/index.php)는 고려해야 할 다른 많은 시그널 펩타이드를 수록하고 있다. 분비 후 시그널 펩타이드는 시그널 펩타이드를 처리하는 세포 프로테아제에 의해 제거된다. 인플루엔자 바이러스는 곤충 바이러스가 아니지만, HA 시그널 펩타이드는 BEVS에서 재조합 단백질의 분비 또는 막 삽입을 위한 분비 시그널 펩타이드로서 사용되어 왔다. MKTIIALSYIFCLVFA의 HA 시그널 펩타이드는 곤충 세포에서 재조합 막관통 G 단백질 커플링된 수용체 (GPCR: G protein coupled receptors)의 발현에 흔히 사용된다 (Rosenbaum, DM, *et al.*, *Science* (2007) 318:1266-1273; Zou, Y., *et al.*, *PLoS One* (2012) 7:e46039). HA는 곤충 세포에서 그의 천연 시그널 펩타이드를 이용하여 발현되어 왔다. 그 특별한 경우, 총 발현량은 곤충 세포 시그널 펩타이드를 사용하는 것보다 낮지만, 발현된 HA는 HA1 및 HA2를 함유하는 성숙한 형태였다 (미국특허 5,858,368 A). HA 시그널 펩타이드는 HA들 사이에서 보존되지 않는다. 예를 들어, 전술한 HA 시그널 펩타이드는 SEQ ID NO: 7의 H1 HA 시그널 펩타이드와 다르다. 경험적으로 HA 시그널 펩타이드를 선택하면 곤충 세포에서 HA 발현을 향상시킬 수 있다.

[0115] [102] Bac-to-Bac 시스템 또는 Gateway 시스템 (Life Technologies)을 사용하여 상동성 재조합을 통해 재조합 배칼로바이러스를 제조하는데, 상업용 벡터를 이용할 수 있다. 프로모터 및 시그널 서열과 같은 이들 전달 벡터의 속성을 표준 분자생물학 기술에 의해 커스터마이즈시킬 수 있다. 이제는 전체 전달 벡터를 유전자 합성에 의해 완전히 합성할 수 있다. 유전자 합성에 의해 특정 서열 및 특징을 갖는 전달 벡터를 설계할 수 있다.

[0116] [103] BEVS를 사용하여 인플루엔자 HA를 발현하는 방법은 실험실 규모에서 잘 확립되어 있다 (Stevens, J., *et al.*, *Science*(2004) 303:1866-1870). 예를 들어, 전달 벡터는 폴리헤드린 프로모터의 제어하에 있는 1918 인플루엔자 바이러스로부터의 HA를 함유한다. 1918 인플루엔자 바이러스 HA 발현 구조물은 분비를 위해 N-말단에 HA 시그널 펩타이드 대신 GP67 시그널 펩타이드를 갖는다. 막관통 도메인 대신, 트롬빈 절단 부위가 HA 세포의 도메인의 C-말단에 도입되고, 그 다음에 HA 삼량체를 안정화시키기 위해 박테리오파지 T4 피프리틴으로부터 유래된 "폴드온(foldon)"서열, 그리고 정제를 용이하게 하기 위한 C-말단 His-태그가 삽입된다. 폴드온 서열과 His-tag는 트롬빈 절단에 의해 제거될 수 있다. 많은 인플루엔자 스트레인으로부터 상이한 HA를 발현시키기 위해 HA 발현 전달 벡터의 유사한 디자인이 사용되어 왔다 (Stevens, J., *et al.*, *Science*(2006) 312:404-410; Xu, R., *et al.*, *Science*(2010)328:357-360; Whittle, J. R., *et al.*, *PNAS*(2011) 108:14216-14221).

[0117] [104] 폴드온에 더해, 다른 삼량체화 도메인이 재조합 HA의 삼량체를 안정화 시키는데 사용될 수 있다. 예컨대, 절단된(truncated) HA 구조물의 삼량체화를 돕기 위해 열안정성 HIV 1 당단백질 41(gp41) 삼량체화 도메인 또는 GCN4 류신 지퍼 서열의 31개 잔기의 초기 16개가 사용되었다 (Impagliazzo, A., *et al.*, *Science*(2015) 349:1301-1306; Yassing, H. M., *et al.*, *Nat Med*(2015) 21:1065-1070).

[0118] [105] 막관통 도메인 및 세포내 도메인을 갖는, A형 인플루엔자 및 B형 인플루엔자 양자 모두로부터의 재조합 HA 전구체(HA0)가, 미국특허 5,762,939A, 5,858,368A 및 6,245,532 B1에 개시된 바와 같이 상업적 규모로 BEVS를 사용하여 제조되었다. 이들 재조합 HA 단백질은 인플루엔자 백신으로서 FDA 승인을 얻은 Flublok®의 성분이다. Flublok®의 HA0는 HA 시그널 펩타이드를 배칼로바이러스 키티나제 시그널 펩타이드 (61K 시그널 펩타이드라고 함)로 대체하여 제조되었다. 재조합 HA는 곤충 세포의 말초 막과 연관되어있다. 이것은 세제를 사용하여 막으로부터 추출되고 더 정제된다.

[0119] 포유동물 발현

[0120] [106] 차이나이즈 햄스터 난소(CHO) 세포 및 인간 배아 신장 293 (HEK293) 세포는 재조합 단백질의 일시적 형질감염 유전자 발현을 위해 일반적으로 사용되는 포유동물 세포 숙주이다. CHO 세포 및 HEK293 세포 양자 모두 현탁 세포주 및 부착 세포주를 갖는다. 이들 세포들은 FBS가 존재 또는 부재하는 배지에서 일반적으로 배양된다. FBS 부재 하에 세포 배양시, 화학적으로 제한된 배지가 종종 사용된다. 모든 현탁 및 부착 세포주는 HA 또는 다른 분비 단백질을 발현하는데 사용될 수 있다.

[0121] [107] 일시적 형질감염(transient transfection)은 단백질 발현을 위해 DNA를 세포에 도입하는데 있어 잘 확립된 방법이다. 많은 시중의 형질감염 시약을 이용할 수 있다. 배칼로바이러스 발현을 위한 발현 최적화 공정과 유사하게, 표준 방법을 이용하여, 단백질 발현 분석을 위해 형질감염 후 배지 샘플을 주기적으로 채취한다. HA 단백질은 HA 구조 내의 His-tag를 통한 친화력 포획법 또는 상업적으로 이용 가능한 많은 항-HA 항체를 이용한 웨스턴 블롯법에 의해 쉽게 검출된다. 단백질 발현은 일반적으로 형질감염으로부터 약 48시간 후에 개시되며 형질감염 후 며칠 동안 증가할 수 있다. 때때로 표적 단백질 발현을 증가시키기 위해 형질감염 후 보충제가 첨가

된다. 아프리카 녹색 원숭이 신장 조직 유래의 섬유 아세포-유사 세포주인 COS-1과 같은 다른 포유동물 세포주도 또한 HA 발현에 사용되어 왔다. 인플루엔자 VLP는 모든 10 개의 인플루엔자 바이러스-코딩된 단백질의 cDNA가 포함된 여러 플라스미드의 동시-형질감염에 의해 생성되어 왔다 (Mena, I., *et al.*, *J. Virol*(1996) 70: 5016-5024; Chen, B. J., *et al.*, *J. Virol*(2007) 81:7111-7123).

[0122] [108] 포유동물 세포에 유전자를 전달하는 또 다른 방법은 포유동물 세포 내로의 배클로바이러스 유전자 전달을 위한 BacMam으로 알려진 재조합 배클로바이러스를 이용한다 (Boyce, F. M., 및 Bucher, N. L., *PNAS*(1996) 93:2348-2352; Dukkkipatia. A., *et al.*, *Protein Expr Purif* (2008) 62:160-170). 포유동물 발현 카세트를 배클로바이러스 발현 벡터에 도입하여 포유동물 세포에서 유전자변형 발현을 위해 배클로바이러스를 변형시켰다. 재조합 BacMam 배클로바이러스가 배클로바이러스 생산을 위한 표준 방법에 의해 생성된다. 단백질을 발현을 위해 CHO 또는 HEK293 세포에 포유동물 발현 카세트를 전달하기 위해 CHO 또는 HEK293 배양체에 증폭된 재조합 BacMam 배클로바이러스 바이러스 스톱을 첨가한다. BacMam 플랫폼은 대량의 포유동물 세포의 형질도입을 용이하게 할 수 있다.

[0123] [109] 포유동물 세포에 형질감염된 유전자 발현 벡터는 표적 단백질의 발현을 위한 안정한 세포주를 확립을 위해 세포 염색체에 통합될 수 있다. 안정한 CHO 세포주는 모노클로날 항체와 같은 치료용 생물 의약품 생산에 가장 많이 사용되는 숙주이다. 이 기술은 치료용 단백질과 백신의 대량 생산에 잘 적합화되었다. 안정한 세포주를 확립하기 위해, 발현 벡터로 형질감염된 세포를 발현 벡터 상의 선택 마커에 기초하여 선택한다. 표적 유전자의 다중 복제물이 종종 안정한 세포주의 계통에 통합된다.

[0124] 포유동물 발현을 위한 HA 벡터

[0125] [110] 포유동물 발현을 위한 HA 발현 구조물은 배클로바이러스 발현의 경우와 동일한 디자인 및 아미노산 서열을 갖는다. 코돈을 CHO 세포 발현 또는 HEK293 세포 발현에 최적화시킨다. 배클로바이러스 코돈은 또한 포유동물 세포에서 잘 작동하며 그 반대도 마찬가지이다. HA 시그널 펩타이드는 포유동물 발현 시스템에서 재조합 단백질의 분비 또는 막 삽입을 위한 시그널 펩타이드로 사용되어 왔다. 일반적으로 사용되는 다른 포유동물 시그널 펩타이드로는 인간 IL2 시그널 펩타이드, 조직 플라스미노겐 활성화제(tPA) 시그널 펩타이드, 및 시그널 서열 데이터베이스 (signal sequence database) (월드 와이드 웹의 signalpeptide.de/index.php에 기재)에서 발견되는 많은 다른 시그널 펩타이드를 들 수 있다. 많은 포유동물 발현 벡터가 상업적으로 이용 가능하다. 각 벡터는 대개 고수준 발현을 위한 인핸서-프로모터, mRNA 안정성을 위한 폴리아데닐화 시그널 및 전사 종결 서열, 에피솜 복제를 위한 SV40 원점, 항생제 내성 유전자 및 대장균에서의 선택 및 유지를 위한 pUC 원점을 갖는다. 포유동물 발현에 일반적으로 사용되는 프로모터로는 CMV (시토메갈로바이러스) 프로모터, hEF1-HTLV 프로모터, 신장 인자 1 α (EF1 α) 코어 프로모터 및 R 세그먼트를 포함하는 복합 프로모터 및 인간 T 세포 백혈병 바이러스 (HTLV) 1형 롱 터미널 반복체의 U5 서열의 일부 (R-U5') 또는 MPromDb (월드 와이드 웹의 mpromdb.wistar.upenn.edu/에 있는 포유동물 프로모터 데이터베이스) 또는 진핵 생물 프로모터 데이터베이스 (월드 와이드 웹의 epd.vital-it.ch/)에서 발견되는 기타 프로모터를 들 수 있다. 전달 벡터는 종종 안정한 세포주를 생성하기 위한 또 다른 선택 마커를 갖는다.

[0126] [111] 다음 실시예를 들어 본 발명을 더욱 상세히 설명하나, 본 발명이 이들 실시예로 한정되는 것은 아니다.

[0127] 실시예 1

[0128] 재조합 배클로바이러스 생산을 위한 H1 HA의 전달 벡터의 구축

[0129] [112] 이 실시예에서는 본 발명의 변형된 HA 단백질을 제조하기 위해 변형 될 수 있는 모(parental) 플라스미드를 설명한다. 이 플라스미드의 기본 개요는 도 4에 도시되어 있으며 이 도면에는 배클로바이러스에서의 발현을 위한 표준 플라스미드로의 삽입체가 그러한 삽입체에 대한 제한효소 부위에 의해 브래킷 표시되어 있고 HA의 구형 헤드 도메인의 면역우성 영역의 치환 및 HA 단백질의 MCS의 치환을 허용하는 제한효소 부위가 존재한다. 코딩된 단백질은 정제를 위해 10개의 히스티딘 (10xHis- 태그)의 친화성 태그에 융합되는데 이들은 정제 후 삽입되는 프로테아제 절단 부위에 의해 제거될 수 있다. 이 벡터는 H1WT로 표시된다.

[0130] [113] 2009년 인플루엔자 백신 생산을 위해 WHO가 권장한 A형 돼지 인플루엔자 바이러스 H1N1인 A/California/07/2009의 HA 서열 (SEQ ID NO: 1) (월드 와이드 웹의 who.int/csr/resources/publications/swineflu/vaccine_recommendations/en/에 기재됨)이 이중 펩타이드를 통합하기 위한 모(parental) 서열로 사용되었다. 이 HA 서열은 UniProt:C3W5X2의 수탁 번호를 갖는다. 모든 구조물들이 도 4에 개략적으로 도시된 것과 동일한 디자인을 갖는다. N-말단 HA 시그널 펩타이드 ((SEQ ID NO:

7)은 GP67 시그널 펩타이드 MVSAIVLYVLLAAAAHSAFA (SEQ ID NO:2)에 의해 대체되었다. Stevens, J., *et al.*, *Science* (2006) 312:404-410에 의해 최초로 설명된 HA 구조물에 기초하여, HA의 서열 ILAIYSTVASSLVLVVSLGAISFWMC를 갖는 C-말단 막관통(TM) 도메인 및 서열 ILAIYSTVASSLVLVVSLGAISFWMC를 갖는 세포내 도메인을, TEV 절단 부위 (SEQ ID NO: 4) 및 이어서 폴드온 (SEQ ID NO: 5) 그리고 10 xHis-tag (SEQ ID NO: 6), 즉 GAENLYFQGGSGYIPEAPRDGQAYVRKDGWVLLSTFLGHHHHHHHHH (SEQ ID NO:3)으로 대체하였다. 폴드온은 재조합 HA를 안정화시키는 반면, 10xHis-tag는 세포 배양 배지로부터 HA의 정제를 용이하게 한다. 폴드온과 10xHis-tag는 HA 서열과 폴드온 서열 사이에 위치하는 TEV 절단 부위를 인식하는 TEV 프로테아제에 의해 제거될 수 있다. 이 유전자를 배콜로바이러스 발현을 위해 코돈 최적화하고 표준 방법에 의해 합성하였다 (GENEWIZ, South Plainfield, NJ 07080, USA).

[0131] [114] 도 4에 개략적으로 도시된 바와 같이, 이격된 독특한 제한효소 부위 인 ClaI 부위, NsiI 부위 및 BamHI 부위를 HA 서열 내로 도입하여 서열 변화 및 교환을 용이화하였다. "H1WT"(SEQ ID NO: 9)로 표시되는 결과적인 구조물 GP67-H1WT-TEV-폴드온-10His를 NcoI 부위와 HindIII 부위 사이의 pFastBacplasmid에 서브클로닝하였다. ClaI 부위와 NsiI 부위 사이의 서열은 구형 헤드 도메인의 면역우성 항원 부위의 대부분을 코딩한다. NsiI와 BamHI 사이의 서열은 MCS를 코딩한다. H1WT 구조물의 발현은 배콜로바이러스 폴리헤드린 프로모터의 제어하에 있다. 결과적인 플라스미드는 A/California/07/2009의 재조합 야생형 HA의 발현을 위한 전달 벡터이다. H1WT 구조물은 그로부터 유전적으로 변형된 HA 구조물이 만들어지는 모(parental) 구조물이다.

[0132] [115] 이중 펩타이드로 유전적으로 변형된 HA를 제조하기 위해, 전술한 제한효소를 이용하여 H1WT의 야생형 단편을 HA에 특이적인 변화를 코딩하는 합성된 DNA 단편으로 대체하였다. MCS를 변형시키기 위해, H1WT의 NsiI-BamHI 단편을, 변형된 성숙 절단 부위를 코딩하는 DNA 단편으로 대체하였다. HA1 구형 헤드 도메인의 항원 부위를 변경하기 위해, H1WT의 ClaI-NsiI 단편을, 이중 펩타이드로 변형된 항원 부위를 코딩하는 DNA 단편으로 대체하였다. 성숙 절단 부위와 항원 부위의 상이한 여러 조합을 갖는 구조물을, 상이한 플라스미드들의 제한 단편의 간단한 교환에 의해 만들었다. 결과적인 구조물을 DNA 시퀀싱에 의해 확인하였다.

[0133] 실시예 2

[0134] 변형된 성숙 절단 부위를 갖는 H1 HA 구조물의 설계 및 구축

[0135] [116] 이 실시예에서는, MCS를 브래킷하는 NsiI 및 BamHI 제한효소 부위들 사이의 H1WT 세그먼트를 이 MCS를 변형시키는 대안적인 서열로 대체한다. 몇몇 경우에, TEV 절단 부위는 변형된 MCS에 포함된다.

[0136] [117] 변형된 성숙 절단 부위가 실시예 1에 기재된 H1 HA의 전달 벡터 내로 도입되었다. 변형된 성숙 절단 부위를 갖는 구조물을 표 1에 나타내었다.

표 1

표 1 List of 변형된 성숙 절단 부위를 갖는 H1 HA 구조물 리스트

리스트 No.	구조물 명칭	구조물 설명	SEQ ID Number
	H1MCS ^a	GP67ss-H1MCS-TEV-폴드온-10His ^b	
1	H1WT ^c	GP67ss-H1WT-TEV-폴드온-10His	SEQ ID NO:9
2	H1H5cs	GP67ss-H1H5cs-TEV-폴드온-10His	SEQ ID NO:11
3	H1R5cs	GP67ss-H1R5cs-TEV-폴드온-10His	SEQ ID NO:14
4	H1IR5cs	GP67ss-H1IR5cs-TEV-폴드온-10His	SEQ ID NO:16
5	H1TEV1	GP67ss-H1TEV1-TEV-폴드온-10His	SEQ ID NO:18
6	H1TEV2	GP67ss-H1TEV2-TEV-폴드온-10His	SEQ ID NO:20
7	H1TEV3	GP67ss-H1TEV3-TEV-폴드온-10His	SEQ ID NO:22

^a 구조물 명칭은 성숙 절단 부위에 기초한다. H1 은 H1 HA 를 의미한다. MCS 는 성숙 절단 부위를 의미한다.

^b 구조물은 그의 서열 속성에 의해 설명된다. GP67ss 은 GP67 시그널 서열을 나타낸다. TEV-폴드온-10H 는 구조물의 C-말단의 TEV 절단 부위, 폴드온 서열, 및 10xHis-tag 서열을 나타낸다. 구조물의 개략적 도면을 도 4 에 나타내었다.

^c H1WT 는 천연 MCS 를 갖는 야생형 H1 HA 구조물이다.

[0137]

[0138]

[119] 구조물 H1H5cs (SEQ ID NO:11)은 H5 MCS (SEQ ID NO:12)에 의해 대체된 야생형 H1 HA의 MCS (SEQ ID NO:9)를 갖는다. 구조물 H1R5cs (SEQ ID NO:14)는 5개의 아르기닌에 의해 대체된 야생형 H1의 MCS를 갖는다. 구조물 H1IR5cs (SEQ ID NO:16)는 천연 HA1 잔기들 모두를 유지하면서 HA1 서열과 HA2 서열 사이에 삽입된 4 개의 아르기닌을 갖는다. 이러한 변화를 갖는 NsiI-BamHI 단편을 합성하여 H1WT의 NsiI 부위와 BamHI 부위 사이에 서브클로닝하여 야생형 단편을 대체하였다.

[0139]

[120] H1 HA MCS가 어떠한 염기성 잔기도 함유하지 않는 TEV 절단 부위로 변형될 수 있는지 시험하기 위해, H1 HA MCS의 몇몇 잔기들을 변형시킴으로써 수 개의 구조물을 제작하였다. 구조물 H1TEV1 (SEQ ID NO:18)은 H1 HA MCS의 7개 잔기인 PSIQSRG가 7-잔기 TEV 절단 부위인 ENLYFQG (SEQ ID NO:4)에 의해 대체되어 있다. 구조물 H1TEV2 (SEQ ID NO:20)은 H1 HA MCS의 8개 잔기인 IPSIQSRG가 TEV 절단 부위를 함유하는 SPENLYFQG에 의해 대체되어 있다. 구조물 H1TEV3 (SEQ ID NO:22)은 H1 HA MCS (SEQ ID NO:8)의 마지막 3개 잔기(QSR)가 TEV 절단 부위에 의해 대체되어 있다.

[0140]

실시예 3

[0141]

구형 헤드 도메인 내 이중 에피토프를 갖는 H1 H2 구조물의 설계 및 구축

[0142]

[121] 이 실시예에서는, 기본 플라스미드와 변형된 MCS를 갖는 플라스미드를, 면역우성 부위들이 이중 에피토프에 의해 대체되어 있는 대안적인 뉴클레오타이드 서열에 의해, ClaI 부위와 NsiI 부위 사이에서 변형시켰다.

[0143]

[122] 이중 에피토프들을 실시예 1과 실시예 2에 설명된 구조물 내로 도입하였다. 이들 이중 에피토프들을 HA 구형 헤드 도메인의 면역우성 항원 부 주변에 삽입하거나 또는 HA 구형 헤드 도메인의 면역우성 항원 부위를 대체하였다. 변형된 항원 부위(들)을 갖는 구조물들을 하기 표 2에 나타내었다.

표 2

표 2 구형 헤드 도메인 내 변형된 면역우성 항원 부위를 갖는 H1 HA 구조물의 리스트

리스트 No.	구조물 명칭	구조물 설명	SEQ IDNo.
	H1MCS-AS ^a	GP67ss-H1MCS - AS - TEV-폴드온-10His ^b	
8	H1H5cs-CR8020Ca	GP67ss-H1H5cs-CR8020Ca-TEV-폴드온-10His	SEQ ID NO:24
9	H1TEV2-CR8020Ca	GP67ss-H1TEV2-CR8020Ca-TEV-폴드온-10His	SEQ ID NO:29
10	H1TEV2-CR8020Sa3	GP67ss-H1TEV2-CR8020Sa3-TEV-폴드온-10His	SEQ ID NO:31
11	H1TEV2-CR8020Sa4	GP67ss-H1TEV2-CR8020Sa4-TEV-폴드온-10His	SEQ ID NO:33
12	H1TEV2-I1C9Ca1	GP67ss-H1TEV2-I1C9Ca1-TEV-폴드온-10His	SEQ ID NO:35
13	H1TEV2-I1C9Ca2	GP67ss-H1TEV2-I1C9Ca2-TEV-폴드온-10His	SEQ ID NO:38
14	H1TEV2-I1C9Sa3	GP67ss-H1TEV2-I1C9Sa3-TEV-폴드온-10His	SEQ ID NO:40
15	H1TEV2-I1C9Sa4	GP67ss-H1TEV2-I1C9Sa4-TEV-폴드온-10His	SEQ ID NO:42
16	H1H5cs-I1C9Sb	GP67ss-H1H5cs-I1C9Sb-TEV-폴드온-10His	SEQ ID NO:44
17	H1H5cs-I1C9Ca	GP67ss-H1H5cs-I1C9Ca-TEV-폴드온-10His	SEQ ID NO:46
18	H1H5cs-FI6Sab	GP67ss-H1H5cs-FI6Sab-TEV-폴드온-10His	SEQ ID NO:48
19	H1WT-FI6Sab	GP67ss-H1WT-FI6Sab-TEV-폴드온-10His	SEQ ID NO:52
20	H1TEV1-FI6Sa	GP67ss-H1TEV1-FI6Sa-TEV-폴드온-10His	SEQ ID NO:54
21	H1H5cs-M2eCa	GP67ss-H1H5cs-M2eCa2-TEV-폴드온-10His	SEQ ID NO:56
22	H1H5cs-M2eSb	GP67ss-H1H5cs-M2eSb-TEV-폴드온-10His	SEQ ID NO:59
23	H1TEV2-M2eCa	GP67ss-H1TEV2-M2eCa2-TEV-폴드온-10His	SEQ ID NO:61

^a 구조물 명칭은 성숙 절단 부위 및 항원 부위 변형에 기초한다. H1 은 H1 HA 를 의미한다. MCS 는 성숙 절단 부위를 의미한다. AS 는 이중 에피토프를 갖는 H1 HA 의 구형 헤드 도메인 내 변형된 면역우성 항원 부위를 의미한다. 각각의 AS 는 이중 에피토프가 위치한 H1 HA 면역우성 항원 부위와 이중 에피토프에 의해 지시된다.

^b 구조물은 그의 서열 속성에 의해 설명된다. GP67ss 은 GP67 시그널 서열을 나타낸다. TEV-폴드온-10H 는 구조물의 C-말단의 TEV 절단 부위, 폴드온 서열, 및 10xHis-tag 서열을 나타낸다. 구조물의 개략적 도면을 도 4 에 나타내었다.

[0144]

[0145]

[123] 이들 변형 위치를 도 5에 나타내었다.

[0146]

[124] 13개 아미노산 잔기의 복합 CR8020 에피토프 펩타이드는 항원 부위 Ca2 주위의 12 잔기 EIAIRPKVRDQE의 표면 루프를 대체 하였다. 복합 CR8020 에피토프 펩타이드 치환체를 갖는 ClaI-NsiI 단편을 합성하고 H1H5cs의 ClaI와 NsiI 부위 사이에 서브클로닝하여 H1H5cs-CR8020Ca (SEQ ID NO: 24)로 명명된 HA 구조물을 생성하였다. H1TEV2-CR8020Ca (SEQ ID NO: 29)를 제조하기 위해, H1H5cs-CR8020Ca의 ClaI-NsiI 단편을 분리하고 H1TEV2의 ClaI와 NsiI 부위 사이에 서브클로닝 하였다.

[0147]

[125] H1TEV2-CR8020Sa3 구조물 (SEQ ID NO: 31)는 구형 헤드 도메인 상부의 H3 HA의 부위 B 루프에 상응하는 Sa 부위 중 하나의 잔기 KKGNS를 대체하는 복합 CR8020 에피토프 펩타이드를 갖는다. CR8020 변형을 갖는 ClaI-NsiI 단편을 합성하고 H1TEV2의 ClaI와 NsiI 부위 사이에서 서브클로닝하여 H1TEV2-CR8020Sa3로 명명된 HA 구조물을 생성하였다.

[0148]

[126] H1TEV2-CR8020Sa4 구조물 (SEQ ID NO: 33)은 Sa 부위 중 어느 하나 부근의 잔기 KTSS의 나선형 구조를 대체하는 복합 CR8020 에피토프 펩타이드를 갖는다. 이 위치는 HA 삼량체의 구형 헤드 도메인 쪽에 있고 수용체 결합 부위로부터 떨어져있다. 이러한 변형은 수용체 결합을 변화시키지 않을 듯하다. CR8020 변형을 갖는 ClaI-NsiI 단편을 합성하고 H1TEV2의 ClaI와 NsiI 부위 사이에 서브클로닝하여 H1TEV2-CR8020Sa4로 명명된 HA 구조

물을 생성하였다.

- [0149] [127] 구조체 CR8020Sa4 및 CR8020Ca의 HA 단량체 모델을 도 6에 나타내었다. 각각의 HA 단량체는 구형 헤드 도메인에 복합 CR8020 에피토프 펩타이드 및 줄기 도메인에 고유 CR8020 에피토프를 갖는다.
- [0150] [128] GIFGAIAGFIEG (SEQ ID NO: 36)의 I1C9 에피토프 펩타이드를 갖는 몇몇 HA 구조물을 제조하였다. 하나의 구조물 H1TEV2-I1C9Ca1 (SEQ ID NO: 35)은 Ca2 부위 주변의 217-224 잔기에서 RPKVRDQE를 대체하는 I1C9 펩타이드를 갖는다. 또 다른 구조물 H1H5cs-I1C9Ca (SEQ ID NO: 46)은 Ca2의 잔기 213 내지 224에서 보다 긴 펩타이드 EIAIRPKVRDQE를 대체하는 I1C9 펩타이드를 갖는다. 다른 구조물은 H3 H1의 부위 A에 상응하는 Ca2 부위 HAGAKS, Sa 부위 KKGNS 및 Sa 부위 KTSS를 개별적으로 대체하는 I1C9 펩타이드를 갖는다. 이러한 변화를 갖는 각각의 ClaI-NsiI 단편을 합성하고 H1TEV2의 ClaI와 NsiI 부위 사이로 서브클로닝하여 각각 H1TEV2-I1C9Ca1 (SEQ ID NO: 35), H1TEV2-I1C9Ca2 (SEQ ID NO: 38), H1TEV2 1C9Sa3 (SEQ ID NO: 40), 및 H1TEV2-I1C9Sa4 (SEQ ID NO: 42)로 명명된 HA 구조물을 생성한다. I1C9Ca의 변화를 갖는 ClaI-NsiI 단편을 합성하고 H1H5cs의 ClaI와 NsiI 부위 사이에 서브클로닝하여 HA 구조물 H1H5cs-I1C9Ca를 생성시켰다. 또한, I1C9 펩타이드는 동일한 방법에 의해 H1H5cs의 Sb 부위의 잔기 TSADQQSLYQNA를 H1H5cs-I1C9Sb (SEQ ID NO:44)로서 대체하였다. Sb 부위는 H3 HA의 부위 B 헬릭스에 해당한다.
- [0151] [129] ClaI-NsiI 단편은 Sa 부위에 위치한 FI6 에피토프 펩타이드 RKKRGLFGAIAGFIE (SEQ ID NO: 49) 및 Sb 부위에 위치한 코일형 코일 FI6 에피토프 펩타이드 KESTQKAIDGVTNKVNS (SEQ ID NO: 50)로 유전자 합성되었다. FI6 치환체를 갖는 이 ClaI-NsiI 단편을 H1H5cs 및 H1WT의 ClaI 및 NsiI 부위 사이에 서브 클로닝시켜 각각 H1H5cs-FI6Sab (SEQ ID NO:48) 및 H1WT-FI6Sab (SEQ ID NO:52) 구조체를 생성시켰다. 또 다른 ClaI-NsiI 단편은 Sa 부위에 위치한 FI6 에피토프 펩타이드 RKKRGLFGAIAGFIE (SEQ ID NO: 49)로 유전자 합성하였다. 이 단편으로 H1TEV1의 ClaI-NsiI 단편을 대체하여 H1TEV1-FI6Sa 구조물 (SEQ ID NO: 54)을 생성시켰다.
- [0152] 실시예 4
- [0153] 구형 헤드 도메인 내 항원 부위에 M2e 펩타이드를 갖는 H1 Ha 구조물의 설계 및 구축
- [0154] [130] 이 실시예에서는 다시, Ca2 부위의 일부 또는 Sb 부위의 일부가 M2e 펩타이드를 코딩하는 뉴클레오타이드 서열에 의해 대체된 뉴클레오타이드 서열로 ClaI-NsiI 단편의 뉴클레오타이드 서열을 대체시켰다. M2e 펩타이드를 갖는 구조물들을 표 2에 나타내었다. 이들 변형의 위치는 도 5에 도시되어 있다.
- [0155] [131] 유전자 합성에 의해 Ca2 부위에서 EIAIRPKVRDQE를 대체하거나 또는 Sb 부위의 잔기 TSADQQSLYQNA의 헬릭스를 대체하는 M2e 펩타이드(SEQ ID NO:57)를 갖는 ClaI-NsiI 단편을 생성하였다. H1H5cs의 ClaI-NsiI 단편이 M2e 펩타이드를 함유하는 개별적인 ClaI-NsiI 단편에 의해 대체되어 각각 H1H5cs-M2eCa (SEQ ID NO:56) 및 H1H5cs-M2eSb (SEQ ID NO:59)이 결과된 H1H5cs에서 M2e 펩타이드를 갖는 HA 구조물을 먼저 만들었다. H1H5cs-M2eCa의 ClaI-NsiI 단편을 H1TEV2 구조물에 전달하기 위해, H1H5cs-M2eCa의 ClaI-NsiI 단편을 분리하고 이를 H1TEV2의 ClaI와 NsiI 부위 사이에 서브클로닝시켜, H1TEV2-M2eCa (SEQ ID NO:61)라 명명된 구조물을 얻었다.
- [0156] 실시예 5
- [0157] Bac-to-Bac 배콜로바이러스 발현 시스템을 이용한 재조합 배콜로바이러스의 생성
- [0158] [132] 이 실시예는 실시예 1-4에서 제조된 구조물에 함유된 발현 시스템들을 포함하는 곤충 세포들을 제조하기 위한 기술을 설명한다.
- [0159] [133] 재조합 배콜로바이러스는 제조자의 지시 (Bac-to-Bac Baculovirus Expression Systems, Life Technology, Carlsbad, CA, USA)에 따라 구조물의 pFastBac 기반 전달 벡터로부터 생성하였다. 간략히 설명하면, 전달 벡터를 대장균 DH10Bac 화학적으로 능력있는 세포로 형질전환시켰다. 재조합 바크미드(bacmid)를 갖는 백색 콜로니를 Blue-gal을 이용하여 플레이트에서 선택 하였다. 바크미드 DNA는 제조자의 지시에 따라 표준 알칼리 용해 미니 프랩 방법으로 분리하였다. M13 정방향(-40) 프라이머 (GTTTCCAGTCACGAC) 및 M13 역방향 프라이머 (CAGGAACAGCTATGAC)를 사용하여 PCR에 의해 재조합 바크미드를 동정하였다. 재조합 바크미드는 약 4kb의 PCR 생성물을 제공하였다.
- [0160] [134] 12-웰 플레이트에 부착된 SF9 곤충 세포를 Cellfectin[®] Reagent, 다른 상업적으로 이용가능한 형질감염 시약 또는 폴리에틸렌 이민 (PEI)을 사용하여 재조합 바크미드로 형질감염시켰다. 형질감염 4-7일 후 또는 바이

러스 감염으로 인해 SF9 세포가 커진 후 세포 배양 배지를 P0 바이러스 스톡으로서 수확했다. 각 P0 바이러스 스톡을 이용하여 진탕 플라스크에서 2×10^6 세포/ml의 밀도로 SF9 세포 50 ML을 감염시켜 바이러스를 증폭시켰다. Cedex Cell Counter (Roche Diagnostics Corporation, Indianapolis, IN, USA)를 사용하여 배양체를 모니터링 하였다. 수확 시간은 평균 생존 가능한 세포 크기 및 세포 생존능에 의해 결정되었다. 평균 생존가능 세포 크기가 감염되지 않은 세포보다 4-7 μm 더 크고 감염 후 4-5일 내에 생존율이 50-70 % 범위에 있는 경우 세포를 원심 분리로 제거하고 세포 배양 상등액을 수집하여 4°C 암실에서 P1 바이러스 스톡으로서 보관하였다. 바이러스 스톡의 멸균성을 유지하기 위해 때로 바이러스 스톡을 0.2 μm 멸균 필터를 통해 여과시켰다.

[0161] [135] 바이러스 스톡을 추가로 증폭하기 위해, 2×10^6 세포/ml의 밀도로 500 ml의 SF9 세포를 250 μl P1 바이러스 스톡으로 감염시켰다. 배양체를 Cedex Cell Counter를 사용하여 모니터링 하였다. 평균 생존가능 세포 크기가 미감염 세포보다 4 ~ 7 μm 더 크고 감염 4 ~ 5일 후에 생존율이 50 ~ 70% 범위인 경우 세포를 원심분리로 제거하고 세포 배양 상등액을 수집하여 4°C 암실에서 P2 바이러스 스톡으로서 보관하였다. 바이러스 스톡의 멸균성을 유지하기 위해 때로 바이러스 스톡을 0.2 μm 멸균 필터를 통해 여과시켰다.

[0162] 실시예 6

[0163] 재조합 HA 단백질의 발현

[0164] [136] 이 실시예는 실시예 1-4에 설명된 다양한 구조물들의 곤충 세포내에서의 발현을 입증한다. 특정 구조물에 따라, 다양한 표적 단백질 발현 수준이 결과되었다.

[0165] [137] 재조합 HA의 발현을 검출하기 위해, 재조합 배칼로바이러스 스톡, 예를 들어 실시예 5에 기재된 P1 또는 P2 바이러스 스톡으로 감염된 SF9 세포의 세포 배양 배지를 수집하고, 재조합 HA의 10xHis-tag를 통한 친화성 포획을 위해 Ni-NTA 수지 (Qiagen, Germantown, MD USA)와 함께 인큐베이션시켰다. 인산염 완충 염수 (1xPBS)로 Ni-NTA 수지를 세척한 후, Ni-NTA 수지를 항-His 항체를 사용하여 SDS-PAGE 또는 웨스턴 블롯에 의한 분석을 위해 환원제의 존재 또는 부재하에 겔 충전 완충액과 함께 비등시켰다. Coomassie® 염색된 겔 또는 항-His 웨스턴 블롯 상에서 약 63 kDa의 단백질 밴드가 나타났는데 이는 분비된 재조합 HA 전장 단백질의 발현을 가리키는 것이다.

[0166] [138] 도 7 및 도 8에 나타낸 바와 같이, 모 구조물인 H1WT의 H1 HA 단백질 발현이 예상대로 검출되었다. 다염기 잔기의 성숙 절단 부위가 변경된 HA 단백질 H1H5cs, H1R5cs 및 H1IR5cs는 H1WT와 동일한 수준으로 발현되었다. 전체 H1 HA 성숙 절단 부위가 TEV 절단 부위로 대체된 H1TEV1은 H1WT보다 훨씬 적게 발현되었다. 성숙 절단 부위에서 보다 많은 천연 잔기를 유지함으로써, H1TEV2 및 H1TEV3은 H1WT와 비교하여 매우 잘 발현되었다 (도 7 및 도 8). 모든 대체물이 같은 수준으로 발현되지는 않았다. 구조물 내의 TEV 절단 부위의 위치를 조절함으로써, 야생형 HA의 발현 수준 정도의 발현 수준이 달성되었다.

[0167] [139] 도 7 및 도 8에도 도시된 바와 같이, 상이한 면역우성 항원 부위에 동일한 면역준우성 bNAb 에피토프의 배치하자 상이한 발현 수준이 결과되었다. Sa 부위 (H1TEV2-CR8020Sa4) 또는 Ca2 부위 (H1H5cs-CR8020Ca)에서의 복합 CR8020 에피토프 펩타이드의 구조물들은 야생형 HA와 동일한 수준으로 발현된 반면, H1TEV2-CR8020Sa3 구조물은 잘 발현되지 않았다. 따라서, 이종 에피토프 배치의 정확한 위치는 결과적인 HA의 발현에 중요할 수 있다. 면역준우성 bNAb 에피토프의 특성이 생성된 HA 단백질의 발현 수준에 영향을 주는 것일 수 있다.

[0168] [140] 몇몇 항원 부위에서의 I1C9 에피토프의 개별 배치 결과 역시도, 결과적인 HA 단백질의 상이한 수준의 발현을 나타냈다 (도 7 및 도 8). H1H5cs-I1C9Ca 구조물은 H1TEV2-I1C9Ca1 구조물에 비해 더 높은 발현을 나타냈다. 이들 두 가지 구조물에서 I1C9 에피토프 펩타이드에 의해 대체된 항원 부위의 펩타이드 서열은 약간 다르다. MCS와 같은 TEV 절단 부위를 갖는 구조물과 마찬가지로, 치환 위치는 발현 수준에 영향을 미친다.

[0169] [141] 성숙 절단 부위에 관계없이 Ca2 부위에 M2e 펩타이드를 갖는 구조물들, 즉 H1H5cs-M2eCa 및 H1TEV2-M2eCa는 H1WT만큼 양호한 발현을 나타냈다(도 7 및 도 8). 그러나 Sb 부위에 M2e 펩티드를 갖는 H1H5cs-M2eSb 구조물은 잘 발현되지 않았다. Sb 부위는 HA의 적절한 접힘(folding)을 위해 요구될 수 있는 헬릭스 구조로 되어 있다.

[0170] [142] 많은 H1 HA 구조물의 성공적인 발현은 HA 구형 헤드 도메인의 면역우성 항원 부위가 줄기 도메인으로부터의 면역준우성 bNAb 에피토프 또는 HA와 관련이 없는 EH 다른 단백질로부터의 M2e 펩티드로 대체될 수 있음을 입증하였다. 이러한 결과는 구형 헤드 도메인의 가소성을 드러내는 것으로서, 구형 헤드 도메인에서 이종 에피토프를 제시하기 위한 기능성 HA의 제조 가능성을 입증하는 것이다. 이와 함께, 실시예 1-6은 이종 에피토프 및

/또는 변형된 MCS를 갖는 HA를 제조하고 우수한 발현을 나타내는 변형된 HA를 선택하는 방법을 입증하였다.

[0171] 실시예 7

[0172] 제조합 HA 단백질의 정제

[0173] [143] This 이 실시예는 실시예 1-4에 설명된 다양한 구조물들의 정제를 입증한다.

[0174] [144] 진탕 플라스크에서 2×10^6 세포/ml의 밀도로 50 ml 내지 1 L의 SF9 세포를, 세포 1 리터 당 1-4 ml P1 또는 P2 바이러스 스톱의 체적비로 배컬로바이러스 스톱으로 감염시켰다. 배양체를 CedexCell Counter를 사용하여 모니터링하였다. 수확 시간은 평균 생존 가능한 세포 크기 및 세포 생존능에 의해 결정되었다. 평균 생존 세포 크기가 감염되지 않은 세포보다 4 ~ 7 μm 더 크고 감염 2 ~ 3 일 후에 생존율이 약 80% 있을 때 세포를 원심 분리로 제거하고 컨디셔닝된 배양 배지를 수집하여 4°C에서 보관했다. 컨디셔닝된 세포 배양 배지를 4°C에서 2-4 시간 동안 로터서리 플랫폼에서 흔들어 주면서 Ni-NTA 수지와 함께 인큐베이션하였다. Ni-NTA 수지를 원심분리에 의해 수집하고 1xPBS로 세척하였다. 세척된 Ni-NTA 수지를 중력 흐름 컬럼 내로 포장하고, 50 mM 이미다졸이 보강된 1 x PBS로 추가 세척하였다. 그 후, Ni-NTA 수지를 중력 흐름에 의해 1xPBS 중의 400 mM 이미다졸로 용리시켰다. 용리된 HA 단백질을 1xPBS로 완충액 교환하고 한외여과에 의해 농축시켰다. 정제된 HA 단백질을 도 8에 나타난 바와 같이 SDS-PAGE로 분석하였다.

[0175] 부록

[0176] 인플루엔자 바이러스, 유형 및 숙주에 대한 설명

[0177] [145] 인플루엔자 바이러스는 A형, B형 및 C형 인플루엔자 바이러스들인, 오르토믹소비리대(Orthomyxoviridae) 바이러스 패밀리의 3가지 속으로 구성된다. 각 속은 오직 단일 종의 바이러스, 즉 A형, B형 및 C형 인플루엔자만을 각각 갖는다. A형, B형 및 C형 인플루엔자 바이러스는 각각 A형, B형 및 C형 인플루엔자 바이러스라고도 합니다. 인플루엔자 바이러스는 엔벨로프된 음성-센스 단일 가닥 RNA 바이러스이다. 바이러스 게놈은 별도의 RNA 세그먼트에 코딩된다. A형 및 B형 인플루엔자 바이러스는 각각 10개의 단백질을 코딩하는 8개의 RNA 세그먼트를 갖는다. C형 인플루엔자 바이러스들은 각각 9개의 단백질을 코딩하는 7개의 RNA 세그먼트를 갖는다. A형 인플루엔자 바이러스의 10가지 특성화된 단백질은 PB2, PB1 및 PA 폴리머라제, 헤마글루티닌(HA), 핵 단백질(NP), 뉴라미니다제(NA), 매트릭스 단백질 M1 및 M2, 비구조 단백질 NS1 및 NS2이다 (Webster, RG, *et al.*, Microbiol Rev (1992) 56:152-179). C형 인플루엔자 바이러스는 HA 및 NA의 기능이 조합된 헤마글루티닌-에스테라제-융합(HEF) 단백질을 갖는다 (Herrler, G., *et al.*, J. Gen Virol (1988) 69:839-846). A형, B형 및 C형 인플루엔자 바이러스의 분류는 NP 및 매트릭스 단백질의 항원 차이에 기반한다.

[0178] [146] A형 인플루엔자 바이러스는 바이러스-코딩된 막 단백질 HA, NA 및 M2리 박혀있는 숙주-유래 지질 이중층 엔벨롭, M1 매트릭스 단백질로 제조된 내부 쉘 및 중심부에 개별적인 RNA 세그먼트의 바이러스 게놈의 뉴클레오캡시드로 구성된 직경 80-120 nm의 소립자이다 (Webster, RG, *et al.*, Microbiol Rev (1992) 56:152-179). RNA 세그먼트는 다중 NP 분자에 의해 느슨하게 캡슐화된다. 3종의 바이러스 폴리머라제 단백질(PB1, PB2 및 PA)의 복합체들이 뉴클레오캡시드의 말단에 위치한다. 모든 RNA 세그먼트들이 감염성 바이러스 입자를 생성하는 데 필요하다.

[0179] [147] 겔 전기영동에 의해 가장 느리게 이동하는 RNA 종인 RNA 세그먼트 1은 PB2 RNA 폴리머라제를 코딩한다. RNA 세그먼트 2는 PB1 RNA 폴리머라제와, PB1-N40 및 PB1-F2 단백질에 대한 두 개의 다른 전사체를 동일한 RNA 서열의 다른 리딩 프레임을 사용하여 코딩한다. PB1-N40 및 PB1-F2 단백질은 숙주 세포의 아포토시스를 유도한다. RNA 세그먼트 3은 PB1과 PB2와 RNA-의존성 RNA 폴리머라제 복합체를 형성하는 PA RNA 폴리머라제를 코딩한다. RNA 세그먼트 4는 인플루엔자 비리온의 주요 표면 항원인 HA를 코딩한다. 각 비리온에는 비리온 표면에 균일하게 분포된 스파이크로 나타나는 약 500개의 HA 분자가 있다 (Ruigrok, RWH, *et al.*, J. Gen Virol (1984) 65:799-902, Murti, KG, Webster, RG, Virol 1986) 149:36-43). HA는숙주 수용체에 결합하여 숙주 범위를 결정한다. RNA 세그먼트 5는 NP 핵 단백질을 코딩한다. NP는 바이러스 RNA를 캡슐화하고 숙주 세포 핵으로 전달된다. NP는 감염된 세포에서 풍부하게 합성되며 인플루엔자 바이러스 비리온에서 두 번째로 풍부한 단백질이다. 이것은 숙주 세포독성 T-세포 면역 반응의 주요 표적이다. RNA 세그먼트 6은 인플루엔자 바이러스 비리온의 두 번째 주요 표면 항원인 뉴라미니다제 (NA)를 코딩한다. NA는 당단백질 또는 당지질로부터 말단 시알산을 절단하여 숙주 세포 수용체로부터 바이러스 입자를 방출시키는 효소이다. 그의 기능은 바이러스 확산에 요구된다. 각 비리온은 그 표면에 약 100개의 NA 분자를 갖는다. NA는 사랑 체를 형성하고 비리온 엔벨롭 상의 불연속적 패치에 위치한다 (Murti, K.G., 및 Webster, R.G., Virol (1986) 149:36-43). RNA 세그먼트 7은 바이시스트

로닉 (bicistronic)이며, 매트릭스 단백질 M1 및 M2 양자 모두를 코딩한다. M1은 비리 온 엔벨롭 밑에서 비리온 뉴클레오펙시드를 감싸는 셸을 형성하며 인플루엔자 바이러스 비리온에서 가장 풍부한 단백질이다. M2는 M1의 동일한 전사체의 다른 스플라이싱 형태로 만들어진다. M2는 필수 막 단백질이며 숙주 세포에서 인플루엔자 바이러스가 생성되는 동안 골지 네트워크의 pH를 조절하는 양성자 채널로서 기능한다. M2는 대체 스플라이싱 변이체 M42로 부분적으로 대체될 수 있다. 하나의 비리온을 만들기 위해서는 약 3000 개의 매트릭스 단백질 분자가 필요하다. RNA 세그먼트 8은 바이러스 복제를 위해 두 개의 비구조적 단백질인 NS1과 NS2를 코딩한다. NS2는 NS1과는 다른 리딩 프레임이다. 두 단백질 모두 감염된 세포에는 풍부하지만 자손 비리온으로는 통합되지 않는다.

[0180] [148] 인플루엔자 바이러스의 3가지 속 중, A형 인플루엔자 바이러스는 가장 독성이 강한 인간 병원체이며 인간에서 가장 심각한 질병을 일으킨다. A형 인플루엔자 바이러스는 그의 표면 당단백질 HA 및 NA의 항원 성질에 기초하여 아형으로 세분된다. H1에서 H18로 명명되는 총 18개의 HA 아형과 N1에서 N10으로 명명되는 10개의 NA 아형이 확인되었다 (Tong, S., *et al.*, *PLoS Pathogens* (2013) 9:e1003657). A형 인플루엔자 아형에 대한 공통적인 명명법은 H1N1, H3N2 및 H7N9와 같은 HA 아형 및 NA 아형의 상이한 여러 조합으로부터 유도된다 (*Bull World Health Organ* (1980) 58:585-591). 물새(waterfowl)는 아형 H1 내지 H16 및 N1 내지 N9의 천연 저장소이다. 아형 H17N10 및 H18N10은 최근 미국산 과일 박쥐에서 발견되었다. 본질적으로 A형 인플루엔자 바이러스의 세그먼트화된 게놈은 A형 인플루엔자 바이러스의 두 가지 아형에 숙주가 동시에 감염되었을 때 RNA 세그먼트의 재조합을 허용한다. 2가지 상이한 아형으로부터의 RNA 세그먼트를, 새로운 아형을 생성하기 위해 단일 비리온에 패키징할 수 있다. A형 인플루엔자 바이러스의 이러한 많은 아형 중에서 H1N1, H2N2 및 H3N2만이 인간에서의 효율적인 전염을 위한 능력을 발달시켰다. 이들 세 가지 아형은 계절 독감 바이러스의 유행하는 바이러스 아형이다. 따라서 인류 집단은 많은 A형 인플루엔자 아형에 대해 면역학적으로 나이브하다. 때로는 다른 아형의 A형 인플루엔자 바이러스가 종을 초월하여 사람을 감염시켜 높은 사망률로 유행성 전염병을 일으킬 수 있다.

[0181] [149] 그들의 계통 발생 관계에 기초하여, A형 인플루엔자 HA 아형은 2 개의 뚜렷한 군으로 분류된다. 그룹 1은 16가지 아형 중 10가지인 H1, H2, H5, H6, H8, H9, H11, H12, H13 및 H16을 포함합니다. 그룹 2는 나머지 6가지 아형, 즉 H3, H4, H7, H10, H14 및 H15에 해당한다.

[0182] [150] B형 인플루엔자 바이러스는 A형 인플루엔자 바이러스와 항원 상으로 상이하다. B형 인플루엔자 바이러스는 A형 바이러스와 함께 순환하며 사람에게 전염병을 유발한다. B형 인플루엔자 바이러스는 알려진 동물 저장소가 없어도 사람에게 안정적으로 적응한다. A형 인플루엔자 바이러스의 광범위한 유전적 변이와는 달리 HA 유전자의 계통학적 관계에 의해 세 계통이 정의되지만 B형 인플루엔자 바이러스에는 한 가지 혈청형만 보고되었다. B형 인플루엔자 바이러스는 A형 인플루엔자 바이러스와 재조합하는 것으로 보이지 않는데, 이는 A형 인플루엔자 바이러스와 B형 인플루엔자 바이러스 사이의 RNA 세그먼트의 재분류에 대한 보고가 없기 때문이다.

[0183] [151] A형 및 B형 인플루엔자 바이러스는 숙주 세포 표면의 말단 시알산을 세포 수용체로 사용한다. 바이러스의 두 가지 유형의 HA는 동일한 구조적 특징을 공유하고 숙주 세포에의 유입을 위해 시알산 수용체에 결합한다.

[0184] [152] A형 인플루엔자 및 B형 인플루엔자 바이러스의 아형들은 스트레인으로 추가 분류된다. A형 인플루엔자 및 B형 인플루엔자에는 많은 여러가지 상이한 스트레인들이 있다. 각각의 독감 계절은 이전의 독감 계절의 계통과 유전적으로 다른 A형 및 B형 독감 바이러스에 의해 지배된다. 서로 다른 지리적 위치 또는 독감 계절 시기에 분리된 한 스트레인의 바이러스는 분리주(isolate)로 알려져 있으며 종종 유전적 변화를 갖는다.

[0185] [153] C형 인플루엔자 바이러스의 감염은 일반적으로 무증상이거나 주로 어린이 또는 젊은 성인과 관련된 경미한 질병을 일으킨다. C형 인플루엔자 바이러스는 산발적인 경우 그리고 국소적인 마이너한 발병과 연관되어 있다. C형 인플루엔자 바이러스는 A형 및 B형 인플루엔자 바이러스보다 질병 부담이 훨씬 적다. 인간 개체군의 상당 부분은 혈청 전환을 보여 주며 이는 인간 집단에서 C형 인플루엔자 바이러스의 광범위한 순환을 시사하는 것이다. C형 인플루엔자 바이러스도 동물로부터 분리되었다. C형 인플루엔자 바이러스의 헤마글루티닌-에스테라제-융합 (HEF) 단백질은 A형 및 B형 인플루엔자 바이러스의 HA 및 NA의 기능을 조합한다. A형 및 B형 인플루엔자 바이러스의 HA와는 달리, C형 인플루엔자 바이러스의 HEF는 세포 수용체로서 말단 9-O-아세틸-N-아세틸 뉴라민산 (9-O-Ac-NeuAc)을 사용한다 (Herrler, G., *et al.*, *EMBO J.* (1985) 4:1503-1506).

[0186] 바이러스 감염 프로세스 및 수명 주기의 설명

[0187] [154] 인플루엔자 바이러스는 호흡기 방울(droplets) 또는 포뮿(fomites) (감염성 유기체를 전달할 수 있는 모든 대상체 또는 물질)을 통해 사람에게서 사람으로 확산된다. 바이러스는 호흡기의 상피 세포를 감염시킨다. 세포 표면 수용체에 결합한 후 부착된 비리온은 숙주 세포에 의해 엔도솜 내로 세포내 섭취되는데 이 엔도솜에서

낮은 pH는 HA의 공간배열에 변화를 유발하고 이에 의해 그의 소수성 융합 펩타이드의 숙주 세포의 소낭막으로의 삽입이 유도되어 바이러스와 소낭막과의 융합이 개시된다. 융합은 비리온의 내용물을 감염된 세포의 세포질로 방출시킨다. 바이러스의 뉴클레오캡시드는 숙주 세포 핵으로 이동하고, 관련 폴리머라제 복합체는 바이러스 단백질의 번역을 위해 mRNA의 1차 전사를 개시한다. 이와 동시에, 숙주 mRNA의 번역이 차단된다. 새로 합성된 바이러스 RNA는 캡시드화되고 바이러스 구조 단백질이 합성되어 숙주 세포 표면으로 운반되어 여기서 이들이 숙주 세포막으로 통합된다. 인플루엔자 바이러스는 기관지 상피세포와 같은 극성 상피세포의 최첨단 표면으로부터 폐 내로 분출되며, 이에 따라 일반적으로 폐친화성(pneumotropic)이다 (Roth, M. G., *et al.*, *PNAS*(1979) 76: 6430-6434; Nayak, D. P., *et al.*, *Virus Res*(2009) 143:147-161). 돼지와 인간 간의 인플루엔자 바이러스 전염이 입증되었다. 한 개체가 감염되면 면역 시스템이 인플루엔자 바이러스에 대한 항체를 생성한다. 이것은 신체의 주요 보호 기전이다.

[0188] 서열 목록

[0189] SEQ ID NO:1: A/California/07/2009의 HA의 펩타이드 서열, 돼지-기원 인플루엔자 A 바이러스 H1N1 스트레인 (UniProt:C3W5X2)

[0190] MKAILVLLYTFATANADTLICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGKLCCLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSS
DNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIIHPSTSAD
QQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINTS
LPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAAGIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFT
AVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEA
KLNREEIDGVKLESTRIYQILAIYSTVASSLVLVSLGAISFWMCSNGSLQCRICI

[0191] SEQ ID NO:2: GP67 분비 시그널의 펩타이드 서열 (GP67ss)

[0192] MVSAILVYVLLAAAAHSAFA

[0193] SEQ ID NO:3: TEV 절단 부위, 폴드온, 및 10xHis-tag의 펩타이드 서열

[0194] GAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0195] SEQ ID NO:4: TEV 절단 부위의 펩타이드 서열

[0196] ENLYFQG

[0197] SEQ ID NO:5: Peptide서열 of 폴드온의 펩타이드 서열

[0198] GSGYIPEAPRDGQAYVRKDGEWVLLSTFL

[0199] SEQ ID NO:6: Peptide서열 of 10xHis-tag 의 펩타이드 서열

[0200] HHHHHHHHHH

[0201] SEQ ID NO:7: Peptide서열 of 시그널 펩타이드 of HA of A/California/07/2009의 HA의 시그널 펩타이드의 펩타이드 서열

[0202] (UniProt:C3W5X2)

[0203] MKAILVLLYTFATANA

[0204] SEQ ID NO:8: Peptide서열 of H1 HA 성숙 절단 부위 (MCS) 의 펩타이드 서열

[0205] IPSIQSR

[0206] SEQ ID NO:9: GP67ss-H1WT-TEV-폴드온-10His (H1WT)의 펩타이드 서열

[0207] MVSAILVYVLLAAAAHSAFADTLICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGKLCCLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAI
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAAGIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNT
QFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSE
EAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0208] SEQ ID NO:10: GP67ss-H1WT-TEV-폴드온-10His (H1WT)의 뉴클레오타이드 서열

[0209] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTCCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGTAAGCTGAGAGGAGTTG
CCCCCTCTGCACcCTGGGCAAAATGTAATATTGCCGGCTGGATtCTGGGAAATCCTGAATGTGAagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAaA
CCcctAgcAGCGAcAATGGCACCTGTTACCCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCAGCTTCGAGAGaTTCGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGAttCtAatAAGGGAGTGACAGCCGCTGTcCTcATGCCGGAGCCAagTCCTTTTACAagAACCTGaTcT
GGCTGgtgAAGAAGGGCAaCAGCTACCCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAaTcCACCACCCTagcA
CAAGcgCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATgTgTtTGTGGCAGCAGCAgATACAgCAaAAagTTCaagCCTGAAATTGCCattAGAC
CCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCCTGgTGGAACCTGGCGATAagATCACATTTGAGGCCACCGAAATCTGgTgGtgCCTa
GaTATGCattTGCtTGgAgAGAAatgCTGGCTCTGGCATCATtAtcTCTGATaCCCTGTGCACGACTGTAATACCAccTGTCAGACACCTaagGGGCCA
TTAATAcCAGCCTGCCCTTCCAGAATATTCACCCTaTcACCATCGGCAAGTGTCTTAAGTATGTGAaGAGCACCAAgcTgAGAcTGGCtACCGGtCTGAGAA
ATATCCCTAGCATCCAGAGCAGAGGcCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGAATGgTgGATGGCTGGTACGGCTATCACCACC
AGAATGAGCAGGGATCcGGATATGCCGCCGATCtGAAGTCTACACAGAACGCCATCGACGagATCACAAACAAGTGAAACAGCGTGATCGAGAAGATGAACA
CCCAGTTTACAGCTGTGGCAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGACGACGGCTTCTCGGATATTTGGACCTACA
ATGCCGAGCTGTCTGTCTCCTGGAGAATGAGAGAACCTGGACTaCCACGACAGCAATGTGAagAACCTGTACGAGAAGGTGAgAGCCAGCTGAAGAACA
ATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAcaCCTGTATGGAGTCTGTGAagAACGGCACCTACGACTACCCTAAGTATA
GCGAGGAGgcCAAGCTGAATAGAGAGAGATCGACGGCGTGAACTGGAAAGCAcaaGaaTcTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTG
GTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTACCATCATCACCACCATC
ACCATCATCACTGATAAagctt

[0210] SEQ ID NO:11: GP67ss-H1H5cs-TEV-폴드온-10His (H1H5cs)의 펩타이드 서열

[0211] MVSAILVLYVLLAAAAHSFAFDTLICGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKKNSYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAI
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAFGFIEGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITKNVNSVIEK
MNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYP
KYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDQAYVRKDGWVLLSTFLGHHHHHHHHHHH

[0212] SEQ ID NO:12: H5 성숙 절단 부위의 펩타이드 서열

[0213] QRERRKKR

[0214] SEQ ID NO:13: GP67ss-H1H5cs-TEV-폴드온-10His (H1H5cs)의 뉴클레오타이드 서열

[0215] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTCCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAAGCACAAATGGCAAGCTGTGTAAGCTGAGAGGAGTTG
CCCCCTCTGCACCTGGGCAAAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAaA
CCCCTAGCAGCGACAATGGCACCTGTTACCCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCAGCTTCGAGAGATTGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACCACCCTAGCA
CAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGCAGCAGCAGATACAGCAAAAAGTTCaagCCTGAAATTGCCATTAGAC
CCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTGAGGCCACCGAAATCTGGTGGTGCCTA
GATATGCATTGTCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGGCCA
TTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCTTAAGTATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAA
ATAGCCCTCAGAGGGAGAGACGAAGAAGAGAGGCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGAATGGTGGATGGCTGGTACGGCT
ATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCGAGA
AGATGAACACCCAGTTTACAGCTGTGGGAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGACGACGGCTTCTCGGATATTT
GGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGC
TGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACC
CTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCACAGAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGG
GCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTACCATCATC
ACCACCATCACCATCATCACTGATAAagctt

[0216] SEQ ID NO:14: GP67ss-H1R5cs-TEV-폴드온-10His (H1R5cs)의 펩타이드 서열

[0217] MVSAILVLYVLLAAAAHSADFATLCIGYHANNSTDVTDVLEKNVTVTHSVNLEDDKHNGKLCCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKGNSYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPRRRRRLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNT
QFTAVGKEFNHLEKRIENLNKKVDDGFLDIWYTNAELLVLENERLTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDPKYS
EEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0218] SEQ ID NO:15: GP67ss-H1R5cs-TEV-폴드온-10His (H1R5cs)의 뉴클레오타이드 서열

[0219] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGTACATCAACGACAAGGGCAAAGAAGTGTGGTGTGTGGGAATCCACCACCCTAGCA
CAAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCTATGTGTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGAC
CCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTA
GATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGCGCCA
TTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCTAAGTATGTGAAGAGCACCAGCTGAGACTGGCTACCGGTCTGAGAA
ATATCCCTAGGAGACGAGAAGAGGCTGTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGAATGGTGGATGGCTGGTACGGCTATCACCACC
AGAATGAGCAGGGATCCGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCGAGAAGATGAACA
CCCAGTTTACAGCTGTGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTGGATATTTGGACCTACA
ATGCCGAGCTGCTCGTCTCCTGGAGAATGAGAGAACCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACA
ATGCCAAGGAGATCGGAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTAAGTATA
GCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCTGAAACTGGAAGCACAAGAATCTATCAGGGCGCTGAAAACCTGTATTTTACGGGCGGTTCTG
GTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTCACCATCATCACCACCATC
ACCATCATCACTGATAAaagctt

[0220] SEQ ID NO:16: GP67ss-H1R5cs-TEV-폴드온-10His (H1R5cs)의 펩타이드 서열

[0221] MVSAILVLYVLLAAAAHSADFATLCIGYHANNSTDVTDVLEKNVTVTHSVNLEDDKHNGKLCCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKGNSYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPS IQSRRRRRLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIE
KMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWYTNAELLVLENERLTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDY
PKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0222] SEQ ID NO:17: GP67ss-H1R5cs-TEV-폴드온-10His (H1R5cs)의 뉴클레오타이드 서열

[0223] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGTACATCAACGACAAGGGCAAAGAAGTGTGGTGTGTGGGAATCCACCACCCTAGCA
CAAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCTATGTGTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGAC
CCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTA
GATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGCGCCA
TTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCTAAGTATGTGAAGAGCACCAGCTGAGACTGGCTACCGGTCTGAGAA
ATATCCCTAGCATCCAGAGCAGGAGACGAGAAGAGGCTGTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGAATGGTGGATGGCTGGTACG
GCTATCACCACCAGAATGAGCAGGATCCGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCG
AGAAGATGAACCCAGTTTACAGCTGTGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTGGATA
TTTGGACCTACAATGCCGAGTGCTCGTGCTCTGGAGAATGAGAGAACCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCC

AGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACT
ACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGTGAAAACCTGTATTTTC
AGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTCACCATC
ATCACCACCATCACCATCATCTGATAAaagctt

[0224] SEQ ID NO:18: GP67ss-H1TEV1-TEV-폴드온-10His (H1TEV1)의 펩타이드 서열

[0225] MVSAILVLYVLLAAAAHSAFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKKGNISYPKLSKSYINDKGKEVLVLWGIHHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNT
QFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYS
EEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0226] SEQ ID NO:19: GP67ss-H1TEV1-TEV-폴드온-10His (H1TEV1)의 뉴클레오타이드 서열

[0227] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGAAGCTGAGAGGAgTTG
CCCCTCTGCAcCTGGCAAAATGAATATTGCCGGCTGGaTtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCAGAttCtAatAAGGGAGTGACAGCCGCTGTcTcATGCCGGAGCCAagTCCTTTTACAagAACCTGaTcT
GGCTGgtgAAGAAGGGCAaCAGTACCCTAAGCTGTCTAAGAGTACaTcAACgACAAGGGCAAAGAAGTGTGGTGTGTGGGGAaTcCACCACCCTagcA
CAaGcgCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCTATgTgTgTGGGCAGCAGCAGaTACAgCAaAAagTTCaagCCTGAAATTGCCatAGAC
CCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCCTGgTGGAACCTGGCGATAagATCACATTTGAGGCCACCGAAATCTGgTgTgCCTa
GaTATGCaTTTGctATgAgAGAAatgCTGGCTCTGGCATCATtAtcTCTGATaCCCCTGTGCACGACTGTAATACCAccTGTCAGACACCTaagGGGCCA
TTAATAcCAGCCTGCCCTTCCAGAATATTACCCCTaTcACCATCGGCaAGTGTCTAAGTATGTGAaGAGCACCAAgcTgAGAcTGGctACCGgtCTGAGAA
ATATCGAAAACCTGTATTTTCAaGGcCTGTTTGGAGCCATCGCCGGCTTTATTTGAGGGAGGATGGACCGAATGgTgGATGGCTGGTACGGCTATCACCACC
AGAATGAGCAGGGATCgGATATGCCGCCGATCtGAAGTCTACACAGAACGCCATCGACGagATCACAACAAGGTGAACAGCGTGATCGAGAAGATGAACA
CCCAGTTTACAGCTGTGGCAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTGGATATTTGGACCTACA
ATGCCGAGCTGCTCGTGCTCCTGGAGAATGAGAGAACCTGGACTaCCACGACAGCAATGTGAaAACCTGTACGAGAAGGTGAgAGCCAGCTGAAGAACA
ATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAcaCCTGTATGGAGTCTGTGAaAACGGCACCTACGACTACCCTAAGTATA
GCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAGACaaGaaTcTATCAGGGCGCTGAAAACCTGTATTTTACAGGGCGGTTCTG
GTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTCACCATCATCACCACCATC
ACCATCATCACTGATAAaagctt

[0228] SEQ ID NO:20: GP67ss-H1TEV2-TEV-폴드온-10His (H1TEV2)의 펩타이드 서열

[0229] MVSAILVLYVLLAAAAHSAFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKKGNISYPKLSKSYINDKGKEVLVLWGIHHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNT
QFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKY
SEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0230] SEQ ID NO:21: GP67ss-H1TEV2-TEV-폴드온-10His (H1TEV2)의 뉴클레오타이드 서열

[0231] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGAAGCTGAGAGGAgTTG
CCCCTCTGCAcCTGGCAAAATGAATATTGCCGGCTGGaTtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCAGAttCtAatAAGGGAGTGACAGCCGCTGTcTcATGCCGGAGCCAagTCCTTTTACAagAACCTGaTcT
GGCTGgtgAAGAAGGGCAaCAGTACCCTAAGCTGTCTAAGAGTACaTcAACgACAAGGGCAAAGAAGTGTGGTGTGTGGGGAaTcCACCACCCTagcA
CAaGcgCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCTATgTgTgTGGGCAGCAGCAGaTACAgCAaAAagTTCaagCCTGAAATTGCCatAGAC
CCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCCTGgTGGAACCTGGCGATAagATCACATTTGAGGCCACCGAAATCTGgTgTgCCTa
GaTATGCaTTTGctATgAgAGAAatgCTGGCTCTGGCATCATtAtcTCTGATaCCCCTGTGCACGACTGTAATACCAccTGTCAGACACCTaagGGGCCA

TTAATAcCAGCCTGCCCTTCCAGAATATTCACCCTaTcACCATCGGCaAGTGTCTTAAGTATGTGAaGAGCACCAAgcTgAGAcTGGCt ACCGGt CTGAGAA
ATAgCCCTGAAAACTGTATTTTCAaGGcCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGgTgGATGGCTGGTACGGCTATCACC
ACCAGAATGAGCAGGGATcGGATATGCCGCCGATCtGAAGTCTACACAGAACGCCATCGACGagATCACAACAAGGTGAACAGCGTGATCGAGAAGATGA
ACACCCAGTTTACAGCTGTGGGAAGGAGTTCAACcACCTGGAGaGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTTGGACCT
ACAATGCCGAGTGTCTGTGCTCCTGGAGAATGAGAGAACCCTGGACTaCCACGACAGCAATGTGAagAACCTGTACGAGAAGGTGAgaAGCCAGCTGAAGA
ACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAaCCTGTATGGAGTCTGTGAagAACGGCACCTACGACTACCTAAGT
ATAGCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCAaaGaaTcTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCGGTT
CTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTACCATCATCACCACC
ATCACCATCATCACTGATAAaagctt

[0232] SEQ ID NO:22: GP67ss-H1TEV3-TEV-폴드온-10His (H1TEV3)의 펩타이드 서열

[0233] MVSAIVLYVLLAAAAHSAFADTLICGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKKGNsYPKLskSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYFVGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
NTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIENLYFQGLFGAIAAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEK
MNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYP
KYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0234] SEQ ID NO:23: GP67ss-H1TEV3-TEV-폴드온-10His (H1TEV3)의 뉴클레오타이드 서열

[0235] ccATGGTAAGCGCTATTGTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGAAATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGTGGTGTGTGGGGAATCCACCACCTAGCA
CAAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGAC
CCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCTGGTGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGCCTA
GATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGCGCCA
TTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCTAAGTATGTGAAGAGCACCAGCTGAGACTGGCTACCGGTCTGAGAA
ATATCCCTAGCATCGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCT
ATCACCACCAAGATGAGCAGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAAGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCGAGA
AGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTT
GGACCTACAATGCCGAGCTGCTCGTGTCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGC
TGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACC
CTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCACAGAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGG
GCGGTCTGTTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGGGTACCATCATC
ACCACCATCACCATCATCACTGATAAaagctt

[0236] SEQ ID NO:24: GP67ss-H1H5cs-CR8020Ca-TEV-폴드온-10His (H1H5cs-CR8020Ca)의 펩타이드 서열

[0237] MVSAIVLYVLLAAAAHSAFADTLICGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKKGNsYPKLskSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYFVGSSRYSKFKPEGMIDYEGTGQAAGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
INTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIE
KMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDY
PKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0238] SEQ ID NO:25: CR8020 에피토프 of HA2 residues 15-19의 펩타이드 서열

[0239] EGMID

[0240] SEQ ID NO:26: CR8020 에피토프 of HA2 residues 30-36의 펩타이드 서열

[0241] EGTGQAA

[0242] SEQ ID NO:27: composite C8020 에피토프의 펩타이드 서열

[0243] EGMIDYEGTGQAA

[0244] SEQ ID NO:28: GP67ss-H1H5cs-CR8020Ca-TEV-폴드온-10His (H1H5cs-CR8020Ca)의 뉴클레오타이드 서열

[0245] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTTCGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACCACCCTAGCA
CAAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAGGCATGATTGATT
ACGAAGGCACAGGCCAGGCGAGCCGGCAGAATGAATTACTACTGGACCCTGGTGGAACTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGC
CTAGATATGCATTTGCTATGGAGAGAAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTGAGACACCTAAGGGCG
CCATTAATACCAGCCTGCCCTTCCAGAATATTCACCTATCACCATCGGCAAGTGCTCTAAGTATGTGAAGAGCACCAGCTGAGACTGGCTACCGGTCTGA
GAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACG
GCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCCGCATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCG
AGAAGATGAACACCCAGTTTACAGCTGTGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTGGATA
TTTGGACCTACAATGCCGAGCTGCTCGTGCTCTGGAGAATGAGAGAACCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCC
AGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACT
ACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGTGAAAACCTGTATTTTC
AGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTCTGTGTCTACCTTCTGGGTCACCATC
ATCACCACCATCACCATCATCTGATAAAaagctt

[0246] SEQ ID NO:29: GP67ss-H1TEV2-CR8020Ca-TEV-폴드온-10His (H1TEV2-CR8020Ca)의 펩타이드 서열

[0247] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGKCLKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLVKKNSYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQLYQNADAYVFGSSRYSKFKFPEGMIDYEGTGQAAGRNNYWTLPVPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGA
INTSLPFQNIHPIITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKM
NTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKKNNAKEIGNGCFEFYHKDCNTCMESVKNGTIDYYPK
YSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDQAYVRKDGWVLLSTFLGHHHHHHHHHHH

[0248] SEQ ID NO:30: GP67ss-H1TEV2-CR8020Ca-TEV-폴드온-10His (H1TEV2-CR8020Ca)의 뉴클레오타이드 서열

[0249] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTTCGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACCACCCTAGCa
CaAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAGGCATGATTGATT
ACGAAGGCACAGGCCAGGCGAGCCGGCAGAATGAATTACTACTGGACCCTGGTGGAACTGGCGATAAGATCACATTTGAGGCCACCGGAAATCTGGTGGTGC
CTAGATATGCATTTGCTATGGAGAGAAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTGAGACACCTAAGGGCG
CCATTAATACCAGCCTGCCCTTCCAGAATATTCACCTATCACCATCGGCAAGTGCTCTAAGTATGTGAAGAGCACCAGCTGAGACTGGCTACCGGTCTGA
GAAATAGCCCTGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATC
ACCACCAGAATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGtGAACAGCGTGATCGAGAAGA
TGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTGGATATTTGGA
CCTACAATGCCGAGCTGCTCGTGCTCTGGAGAATGAGAGAACCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGA
AGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACCCTA
AGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAAACCTGTATTTTCAGGGCG
GTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTCTGTGTCTACCTTCTGGGTCACCATCATCACC
ACCATCACCATCATCACTGATAAAaagctt

- [0250] SEQ ID NO:31: GP67ss-H1TEV2-CR8020Sa3-TEV-폴드온-10His (H1TEV2-CR8020Sa3)의 펩타이드 서열
- [0251] MVSAILVLYVLLAAAAHSADFATLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAACPHAGAKSFYKNLIWLVEGMIDYEGTGQAAYPKLSKSYINDKGKEVLVL
WGIHHPSTSADQQLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTT
CQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEITNKV
NSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERTLDDYHDSNVKNLYEKVRSQKLNNAKEIGNGCFEFYHKCDNTCMESVK
GTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH
- [0252] SEQ ID NO:32: GP67ss-H1TEV2-CR8020Sa3-TEV-폴드온-10His (H1TEV2-CR8020Sa3)의 뉴클레오타이드 서열
- [0253] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGAGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGAGGATAAgCACAAATGGCAAGCTGTGAAGCTGAGAGGAgTTG
CCCCTCTGCAcCTGGCAAAATGAATATTGCCGGCTGGAAtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAaA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCAGAttCtAatAAGGGAGTGACAGCCGCTGTcCTcATGCGGAGCCAagTCCTTTTACAagAACCTGaTcT
GGCTGgtgGAAGGATGATTGATTACGAAGGCACAGGCCAGGCAGCCTACCCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAAGAAGTGTGGTGC
TGTGGGGAaTcCACCACCCTagcACAaGcgCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATgTgTtTGTGGGCAGCAGCAGaTACAGCAaAAagT
TCAagCCTGAAATTGCCatAGACCCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCCTGgTGAACCTGGCGATAagATCACATTTGAGG
CCACCGGAAATCTGgTgTgCCTaGaTATGCaTTTGctATgGgAGAAaagCTGGCTCTGGCATCATtAtcTCTGATaCCCCTGTGCACGACTGTAATACCA
ccTGTGAGACACCTaagGGCGCCATTAATAcCAGCCTGCCCTTCAGAAATATTCACCTaTcACCATCGGCaAGTGTCTTAAGTATGTGAaGAGACCAAGc
TgAGAcTGGCtACCGGtCTGAGAAATAgCCCTGAAAACCTGTATTTTCAaGgcCTGTTTGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGg
TgGATGGCTGGTACGGCTATCACCACAGAATGAGCAGGATCcgGATATGCCGCCGATCtGAAGTCTACACAGAAGCCATCGACGagATCACAACAAGG
TGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGGACG
ACGGCTTCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGTCTCTGGAGAATGAGAGAACCCTGGACTaCCACGACAGCAATGTGAagAACCTGTACG
AGAAGGTGAgaAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAaCCTGTATGGAGTCTGTGAagA
ACGGCACCTACGACTACCTAAGTATAGCGAGGAgcCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAGCAaGaaGaaTcTATCAGGGCGCTG
AAAACCTGTATTTTACGGGCGGTCTGgTTACATCCCGAAGCTCCGCGTGACGGTCAAGGCTTACGTTCGTAAAGACGGTGAATGGGTCTGCTGTCTACCT
TCCTGGGTCAACATCATCACCACCATCACCATCATCACTGATAAaaagctt
- [0254] SEQ ID NO:33: GP67ss-H1TEV2-CR8020Sa4-TEV-폴드온-10His (H1TEV2-CR8020Sa4)의 펩타이드 서열
- [0255] MVSAILVLYVLLAAAAHSADFATLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPEGMIDYEGTGQAAPNHDSNKGVTAACPHAGAKSFYKNLIWLVKKGSYPKLSKSYINDKGKEVLV
LWGIHHPSTSADQQLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNT
TCQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEITNK
VNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERTLDDYHDSNVKNLYEKVRSQKLNNAKEIGNGCFEFYHKCDNTCMESVK
NGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH
- [0256] SEQ ID NO:34: GP67ss-H1TEV2-CR8020Sa4-TEV-폴드온-10His (H1TEV2-CR8020Sa4)의 뉴클레오타이드 서열
- [0257] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGAGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGAGGATAAgCACAAATGGCAAGCTGTGAAGCTGAGAGGAgTTG
CCCCTCTGCAcCTGGCAAAATGAATATTGCCGGCTGGAAtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAaA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCTGAAGGATGATTGATTACGAAGGCACAGGCCAGGCAGCCTGGCCTAATCAGAttCtAatAAGGGAGTGACAGCCGCTGTcCTcATGCCGGAG
CCAagTCCTTTTACAagAACCTGaTcTGGCTGgtgAAGAAGGGCAaCAGCTACCCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAAGAAGTGTGG
TGCTGTGGGGAaTcCACCACCCTagcACAaGcgCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATgTgTtTGTGGGCAGCAGCAGaTACAGCAaAA
agTTCaagCCTGAAATTGCCatAGACCCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCCTGgTGAACCTGGCGATAagATCACATTTG
AGGCCACCGAAATCTGgTgTgCCTaGaTATGCaTTTGctATgGgAGAAaagCTGGCTCTGGCATCATtAtcTCTGATaCCCCTGTGCACGACTGTAATA
CCAacTGTACAGACCTaagGGCGCCATTAATAcCAGCCTGCCCTTCCAGAATATTCACCTaTcACCATCGGCaAGTGTCTAAGTATGTGAaGAGACCA
AgcTgAGAcTGGCtACCGGtCTGAGAAATAgCCCTGAAAACCTGTATTTTCAaGgcCTGTTTGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAA
TGgTgGATGGCTGGTACGGCTATCACCACAGAATGAGCAGGATCcgGATATGCCGCCGATCtGAAGTCTACACAGAAGCCATCGACGagATCACAACA
AGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGCAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGG
ACGACGGCTTCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGTCTCTGGAGAATGAGAGAACCCTGGACTaCCACGACAGCAATGTGAagAACCTGT

ACGAGAAGGTGAgAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAcaCCTGTATGGAGTCTGTGA
agAACGGCACCTACGACTACCTAAGTATAGCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCaaGaaTcTATCAGGGCG
CTGAAAACCTGTATTTTCAGGGCGGTCTGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTCTGCTGTCTA
CCTTCCTGGGTCACCATCATCACCACCATCACCATCATCTACTGATAAaagctt

[0258] SEQ ID NO:35: GP67ss-H1TEV2-I1C9Ca1-TEV-폴드온-10His (H1TEV2-I1C9Ca1)의 펩타이드 서열

[0259] MVSAILVLYVLLAAAAHSAFADTLCIGYHANNSTDVTDVLEKNVTVTHSVNLEDDKHNGKLCRLGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIPKTSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIHHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIGIFGAIAGFIEGGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTP
KGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVI
EKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYD
YPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0260] SEQ ID NO:36: I1C9의 펩타이드 서열

[0261] GIFGAIAGFIEG

[0262] SEQ ID NO:37: GP67ss-H1TEV2-I1C9Ca1-TEV-폴드온-10His (H1TEV2-I1C9Ca1)의 뉴클레오타이드 서열

[0263] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGCGGCGCATTCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGTAAGCTGAGAGGAgTTG
CCCCTCTGCAcCTGGGCAATGTAATATTGCCGGCTGGaTtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCACGAttCtAatAAGGGAGTGACAGCCGCTGTcTcATGCCGGAGCCAagTCCTTTTACAagAACCTGaTcT
GGCTGgtgAAGAAGGGCAaCAGCTACCCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAAGAAGTGTGGTGTGTGGGGAaTcCACCACCCTagcA
CAAGcgCgGATCAGCAGAGCCTGTACCAGAAATGCCGATGCTATgTgTtTGTGGGCAGCAGCagaTACAgCAaAAagTTCaagCCTgaaATTGCCatTgGca
TtTtGcGcGcTATGCCCGGCTTcATTGAGGGAGGCAAGATGAATTACTAGTGACCTGgTGGAACCTGGCGATAagATCACATTTAGGGCCACCGGAAATC
TGgTgTgCCTaGaTATGCaTTTGctTgAgAGaAatgCTGGCTCTGGCATCATtAtcTCTGATaCCCCGTGTGCACGACTGTAATACCAcTGTGAGACAC
CTaagGGCGCCATTAATAcCAGCTGCCCTTCCAGAATATTACCTTaTcACCATCGGCAAGTGTCTAAGTATGTGAaGAGACCAAGcTgAGAcTGGCTA
CCGGtCTGAGAAATAgCCCTGAAAACCTGTATTTTCAaGgcCTGTTTGAGCCATCGCCGCTTTATTGAGGGAGGATGGACCGGAATGgTgGATGGCTGGT
ACGGCTATCACCACCAGAATGAGCAGGGATCcGGATATGCCCGCGATCtGAAGTCTACACAGAACGCCATCGACGagATCACAAACAAGGTGAACACGGTGA
TCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCTGG
ATATTTGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCTGGACTaCCACGACAGCAATGTGAaagAACCTGTACGAGAAGGTGAgaA
GCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAcaCCTGTATGGAGTCTGTGAaagAACGGCACCTACG
ACTACCCTAAGTATAGCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCAaaGaaTcTATCAGGGCGCTGAAAACCTGTATT
TTCAGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCGTAAAGACGGTGAATGGGTCTGCTGTCTACCTCTCTGGGTACCC
ATCATCACCACCATCACCATCATCTACTGATAAaagctt

[0264] SEQ ID NO:38: GP67ss-H1TEV2-I1C9Ca2-TEV-폴드온-10His (H1TEV2-I1C9Ca2)의 펩타이드 서열

[0265] MVSAILVLYVLLAAAAHSAFADTLCIGYHANNSTDVTDVLEKNVTVTHSVNLEDDKHNGKLCRLGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIPKTSWPNHDSNKGVTAAACPGIFGAIAGFIEGFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIHHPST
SADQQSLYQNADAYVFGSSRYSKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQ
TPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNS
VIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGT
YDPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0266] SEQ ID NO:39: GP67ss-H1TEV2-I1C9Ca2-TEV-폴드온-10His (H1TEV2-I1C9Ca2)의 뉴클레오타이드 서열

[0267] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGCGGCGCATTCTGCCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGTAAGCTGAGAGGAgTTG
CCCCTCTGCAcCTGGGCAATGTAATATTGCCGGCTGGaTtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCACGAttCtAatAAGGGAGTGACAGCCGCTGTcTcTGGcaTtTtGcGcGcTATGCCCGGCTTcATTGAGGGAT
TTTACAagAACCTGaTcTGGCTGgtgAAGAAGGGCAaCAGCTACCCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAAGAAGTGTGGTGTGTGGG
GAaTcCACCACCCTagcACAaGcgCgGATCAGCAGAGCCTGTACCAGAAATGCCGATGCTATgTgTtTGTGGGCAGCAGCagaTACAgCAaAAagTTCaagC

CTGAAATTGCCat tAGACCCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCTGgTGGAACTGGCGATAagATCACATTGAGGCCACCG
GAAATCTGgTgGt gCCTaGaTATGCaTTTGCTaTgGAgAGAAat gCTGGCTCTGGCATCATt At cTCTGATaCCCCTGTGCACGACTGTAATACCAcTGTC
AGACACCTaagGGCGCCATTAATAcCAGCCTGCCCTTCCAGAATATTACCCCTaTcACCATCGGCaAGTGTCTAAGTATGTGAaGAGACCAAgcTgAGAc
TGGCt ACCGt CTGAGAAATAgCCCTGAAAACCTGTATTTTCAaGGcCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGgTgGATG
GCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCcgGATATGCCGCGGATCtGAAGTCTACACAGAACGCCATCGACGagATCACAACAAGGTGAACA
GCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCT
TCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCCTGGACTaCCACGACAGCAATGTGAagAACCTGTACGAGAAGG
TGagaAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAcaCCTGTATGGAGTCTGTGAagAACGGCA
CCTACGACTACCCTAAGTATAGCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCTGAAACTGGAAAGCAcaaGaaTcTATCAGGGCGCTGAAAACC
TGTATTTTCAGGGCGGTTCTGTTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCCTGG
GTCACCATCATCACCACCATCACCATCATCACTGATAAagct t

[0268] SEQ ID NO:40: GP67ss-H1TEV2-I1C9Sa3-TEV-폴드온-10His (H1TEV2-I1C9Sa3)의 펩타이드 서열

[0269] MVSaIVLYVLLAAAAHSAFADTLcIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGKLCKLRGVAPLHLGKcNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAcPHAGAKSFYKNLlWLVGIFGAIAGFIEGYPKLSKSYINDKGKEVLVLW
GIHHPSTSADQQLYQNADAYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTC
QTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEITNKVN
SVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRSQKLNNAKEIGNGCFEFYHKCDNTCMESVKNG
TYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0270] SEQ ID NO:41: GP67ss-H1TEV2-I1C9Sa3-TEV-폴드온-10His (H1TEV2-I1C9Sa3)의 뉴클레오타이드 서열

[0271] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGCGCGCATTCCTGCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGTAAGCTGAGAGGAGTTG
CCCCTCTGCAcCTGGGCAAAATGTAATATTGCCGGCTGGAAtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAaA
CCcctAgcAGCGAcAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGAttCtAatAAGGGAGTGACAGCCGCTGTcCTcATGCCGGAGCCAagTCCTTTTACAagAACCTGaTcT
GGCTGgtgGgcaTtTtCGGcGcTATCGCCGGCTTcATTGAGGGATACCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAAGAAGTGTGTTGCTGT
GGGGAaTcCACACCCTAgcACAaGcgCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATgTgTtTGTGGGCAGCAGCAGaTACAgCAaAAagTTCA
agCCTGAAATTGCCat tAGACCCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCTGgTGGAACTGGCGATAagATCACATTGAGGCCA
CCGGAATCTGgTgGtgCCTaGaTATGCaTTTGCTaTgGAgAGAAat gCTGGCTCTGGCATCATtAtcTCTGATaCCCCTGTGCACGACTGTAATACCAcT
GTCAGACACCTaagGGCGCCATTAATAcCAGCCTGCCCTTCCAGAATATTACCCCTaTcACCATCGGCaAGTGTCTAAGTATGTGAaGAGACCAAgcTgA
GAcTGGCtACCGGtCTGAGAAATAgCCCTGAAAACCTGTATTTTCAaGGcCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGgTgG
ATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCcgGATATGCCGCGGATCtGAAGTCTACACAGAACGCCATCGACGagATCACAACAAGGTGA
ACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGGACGACG
GCTTCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGTCTCTGGAGAATGAGAGAACCCTGGACTaCCACGACAGCAATGTGAagAACCTGTACGAGA
AGGTGAgaAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAcaCCTGTATGGAGTCTGTGAaAGACG
GCACCTACGACTACCCTAAGTATAGCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCTGAAACTGGAAAGCAcaaGaaTcTATCAGGGCGCTGAAA
ACCTGTATTTTCAGGGCGGTTCTGTTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCC
TGGGTCACCATCATCACCACCATCACCATCATCACTGATAAagct t

[0272] SEQ ID NO:42: GP67ss-H1TEV2-I1C9Sa4-TEV-폴드온-10His (H1TEV2-I1C9Sa4)의 펩타이드 서열

[0273] MVSaIVLYVLLAAAAHSAFADTLcIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGKLCKLRGVAPLHLGKcNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPGIFGAIAGFIEGWPNHDSNKGVTAAcPHAGAKSFYKNLlWL VKKNSYPKLSKSYINDKGKEVLVL
WGIHHPSTSADQQLYQNADAYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTT
CQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEITNKV
NSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRSQKLNNAKEIGNGCFEFYHKCDNTCMESVKN
GTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0274] SEQ ID NO:43: GP67ss-H1TEV2-I1C9Sa4-TEV-폴드온-10His (H1TEV2-I1C9Sa4)의 뉴클레오타이드 서열

[0275] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGCGCGCATTCCTGCTTTGCGGATACAcTgTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAagCACAAATGGCAAGCTGTGTAAGCTGAGAGGAGTTG

CCCCTCTGCAcCTGGGCAATGTAATATTGCCGGCTGGaTtCTGGGAAATCCTGAATGTGAaagcCTGtctACAGCCAGCAGCTGGTCTTATATCGTGGAAaCCcctAgcAGCGAcAATGGCACCTGTTACCTTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGaTTCGAGATCTTCCCTGGcaTtTTcGGcGcTATCGCCGGCTTcATTGAGGGATGGCCTAATCACGAttCtAatAAGGGAGTGACAGCCGCTGTcCTcATGCCGGAGCCAagTCCTTTTACAagAACCTGaTcTGGCTGgtgAAGAAGGGCAaCAGCTACCCTAAGCTGTCTAAGAGCTACaTcAACgACAAGGGCAAGAAGTGTGGTGTGTGGGGAAaTcCACCACCTAgcACAAgGcCcGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATgTgTtTGTGGGCAGCAGCAgATACAGCAaAAagTTCaagCCTGAAATTGCCattAGACCCAAAGTGAGAGATCAGGaaGGCAGAATGAATTACTACTGGACCCTGgTGGAACTGGCGTAaagATCACATTTGAGGCCACCGGAAATCTGgTgTgCCTaGaTATGCaTTTGCTaTGgAgAGAaatgCTGGCTCTGGCATCATtAtcTCTGATaCCCTGTGCACGACTGTAATACCAccTGTCAGACACCTaagGGCGCCATTAATAcCAGCCTGCCCTTCCAGAATATTCACCCTaTcACCATCGGCaAGTGTCTTAAGTATGTGAaGAGACCAAGcTgAGAcTGGCTACCGGtCTGAGAAATAgCCCTGAAAACCTGTATTTTCAaGGcCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGgTgGATGGCTGGTACGGCTATCACCACCAAGATGAGCAGGGATcCGATATGCCGCCGATCtGAAGTCTACACAGAAGCCATCGACGagATCACAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACcACCTGGAGaaGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCCTGGACTaCCACGACAGCAATGTGAagAACCTGTACGAGAAGGTGAaAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACaAaCCTGTATGGAGTCTGTGAaGACGGCACCTACGACTACCCTAAGTATAGCGAGGAGgcCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCAaaGaaTcTATCAGGGCGCTGAAAACCTGTATTTTACGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0276] SEQ ID NO:44: GP67ss-H1H5cs-I1C9Sb-TEV-폴드온-10His (H1H5cs-I1C9Sb)의 펩타이드 서열

[0277] MVSAILVLYVLLAAAAHSAFADTLICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIPKTSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIIHPSGIFGAIAGFIEGDAYVFGSSRSYKFKFPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRRRKKRGLFGAIAGFIEGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITKNVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGWVLLSTFLGHHHHHHHHHH

[0278] SEQ ID NO:45: GP67ss-H1H5cs-I1C9Sb-TEV-폴드온-10His (H1H5cs-I1C9Sb)의 뉴클레오타이드 서열

[0279] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCGATACCGTGGATACAGTGCTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTGCCCTCTGCACCTGGGCAAAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAACCCCTAGCAGCGACAATGGCACCTGTTACCTTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTTCGAGATCTTCCCTAAGACAAGCAGCTGGCCTAATCAGATTCTAATAAGGGAGTGACAGCCGCTGCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAACAGTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAGAAGTGTGGTGTGTGGGGAATCCACCACCTAGCGGCATTTTCGCGCTATCGCCGGCTTCATtGAGGGAGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTGAGGCCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGTGGCTCTGGCATCATTATCTGTATACCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGGCCATTAATACCAGCCTGCCCTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCTAAGTATGTGAAGAGCACCAAGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGAGAGACGCAAGAAGAGAGGCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGGCTATCACCACCAAGATGAGCAGGGATCCGATATGCCGCCGATCTGAAGTCTACACAGAAGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGACGACGGCTTCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAAGGCACCTACGACTACCTAAGTATAGCGAGGAGGCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAGAATCTATCAGGGCGTGAAAACCTGTATTTTACGGCGGTTCTGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGCTGTCTACCTTCTGGGTCACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0280] SEQ ID NO:46: GP67ss-H1H5cs-I1C9Ca-TEV-폴드온-10His (H1H5cs-I1C9Ca)의 펩타이드 서열

[0281] MVSAILVLYVLLAAAAHSAFADTLICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIPKTSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIIHPST SADQQLYQNADAYVFGSSRSYKFKFPGIFGAIAGFIEGGRMNYYWTLVEPGDKITFEATGNLVVPYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRRRKKRGLFGAIAGFIEGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITKNVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEYHKCDNTCMESVKNGTYDYP

KYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0282] SEQ ID NO:47: GP67ss-H1H5cs-I1C9Ca-TEV-폴드온-10His (H1H5cs- I1C9Ca)의 뉴클레오타이드 서열

[0283] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAAATAGCACCG ATACCGTGGATACAGTGTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTG CCCCTCTGCACCTGGGCAAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAA CCCCTAGCAGCGACAATGGCACCTGTTACCCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCT GGCTGGTGAAGAAGGGCAACAGCTACCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACCACCCTAGCA CAAGCGCCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCAAGCCTGGCATTTTCGGCGCTA TCGCCGGCTTCATTGAGGGAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTGAGGCCACCGGAAATCTGGTGGTGCCTA GATATGCATTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCTGTGCACGACTGTAATACCACCTGTCAGACACCTAAGGGGCCA TTAATACCAGCCTGCCCTCCAGAATATTCACCCTATCACCATCGGCAAGTGCTCTAAGTATGTGAAGAGCACCAGCTGAGACTGGCTACCGGTCTGAGAA ATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACGgcT ATCACCACCAGAATGAGCAGGATCCGATATGCCGCCGATCtGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCGAGA AGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCCTGGATATTT GGACCTACAATGCCGAGCTGCTCGTGTCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCCAGC TGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACTACC CTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCACAAGAATCTATCAGGGCGTGAAAACCTGTATTTTCAGG GCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAGACGGTGAATGGGTTCTGTGTCTACCTTCTGGGTCACCATCATC ACCACCATCACCATCATCACTGATAAaagctt

[0284] SEQ ID NO:48: GP67ss-H1H5cs-FI6Sab-TEV-폴드온-10His (H1H5cs-FI6Sab)

[0285] MVSAIVLYVLLAAAAHSAFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGKLCCKLRGVAPLHLGKCN IAGWILGNPECESLSTASSWSYIVET PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNL IWLVRKKRGLFGAIAGFIE IYINDKKEVLVLWGIIH PSKESQKAIDGVTNKVNSDAYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGI IISDTPVHDCNTT CQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRRRRKKRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITN KVNSEIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKDNTCMESV KNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0286] SEQ ID NO:49: 성숙 절단 부위 of FI6 에피토프의 성숙 절단 부위의 펩타이드 서열

[0287] RKKRGLFGAIAGFIE

[0288] SEQ ID NO:50: helix coiled-coil 펩타이드 of FI6 에피토프의 헬릭스 코일형-코일 펩타이드의 펩타이드 서열

[0289] KESTQKAIDGVTNKVNS

[0290] SEQ ID NO:51: GP67ss-H1H5cs-FI6Sab-TEV-폴드온-10His (H1H5cs-FI6Sab)의 뉴클레오타이드 서열

[0291] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAAATAGCACCG ATACCGTGGATACAGTGTGGAGAAGAATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGTAAGCTGAGAGGAGTTG CCCCTCTGCACCTGGGCAAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAA CCCCTAGCAGCGACAATGGCACCTGTTACCCCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA TCTTCCCTAAGACAAGCAGCTGGCCTAATCACGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGGAGCCAAGTCCTTTTACAAGAACCTGATCT GGCTGGTGAGAAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACC ACCCTAGCAAGGAGTCTACACAGAAGGCCATTGATGGCGTTACAAAATAAGGTCAATtCTGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGT TCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGAACCTGGCGATAAGATCACATTTGAGG CCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCTGTGCACGACTGTAATACCA CCTGTGACACACCTAAGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGCAAGTGTCTAAGTATGTGAAGAGCACAAGC TGAGACTGGCTACCGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCG GAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGATCCGATATGCCCGGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACA ACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAG TGGACGACGGCTTCCTGGATATTTGGACCTACAATGCCGAGCTGCTCGTGTCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACC TGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTG

TGAAGAACGGCACCTACGACTACCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGG
GCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAGACGGTGAATGGGTTCTGCTGT
CTACCTTCTGGGTCACCATCATCACCACCATCACCATCATCTAGTATAaagctt

[0292] SEQ ID NO:52: GP67ss-H1WT-FI6Sab-TEV-폴드온-10His의 펩타이드 서열 (H1WT-FI6Sab)

[0293] MVSAILVLYVLLAAAAHSAFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLVRKKRGLFGAIAGFIEYINDKGKEVLVLWGIHH
PSKESTQKAIDGVTNKVNSDAYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTT
CQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGSYAADLKSTQNAIDEITNKVN
SVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNG
TYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0294] SEQ ID NO:53: GP67ss-H1WT-FI6Sab-TEV-폴드온-10His의 뉴클레오타이드 서열 (H1WT-FI6Sab)

[0295] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGAGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCAGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCGGAGCCAAGTCTTTTACAAGAACCTGATCT
GGCTGGTGAGAAAGAAGAGAGGCTGTTTGAGGCCATCGCCGGCTTTATTGAGTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACC
ACCCTAGCAAGGAGTCTACACAGAAGGCCATTGATGGCGTTACAAATAAGGTCAATTCTGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGT
TCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCTGGTGAACCTGGCGATAAGATCACATTTGAGG
CCACCGGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCA
CCTGTGACACCTAAGGGCGCCATTAATACCAGCTGCCCTTCCAGAATATTCACCTATCACCATCGCAAGTGTCTAAGTATGTGAAGAGACCAAGC
TGAGACTGGCTACCGGTCTGAGAAATATCCCTAGCATCCAGAGCAGAGGCTGTTTGAGGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGG
ATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAAACAAGGTGA
ACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACG
GCTTCTGGATATTTGGACCTACAATGCCGAGCTGCTGCTGCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGA
AGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAAG
GCACCTACGACTACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCACAAGAATCTATCAGGGCGCTGAAA
ACCTGTATTTTACAGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTTCGTAAAGACGGTGATGGGTTCTGCTGTCTACCTTCC
TGGGTCACCATCATCACCACCATCACCATCATCTAGTATAaagctt

[0296] SEQ ID NO:54: GP67ss-H1TEV1-FI6Sa-TEV-폴드온-10His의 펩타이드 서열 (H1TEV1-FI6Sa)

[0297] MVSAILVLYVLLAAAAHSAFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLVRKKRGLFGAIAGFIEGYINDKGKEVLVLWGIH
HPSTSADQQLYQNDAYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTP
KGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIENTLYFQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGSYAADLKSTQNAIDEITNKVNSVIE
KMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDY
PKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHHH

[0298] SEQ ID NO:55: GP67ss-H1TEV1-FI6Sa-TEV-폴드온-10His의 뉴클레오타이드 서열 (H1TEV1-FI6Sa)

[0299] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGTGAGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA
TCTTCCTTAAGACAAGCAGCTGGCCTAATCAGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCGGAGCCAAGTCTTTTACAAGAACCTGATCT
GGCTGGTGAGAAAGAAGAGAGGCTGTTTCGGCGCTATCGCCGGCTTCATTGAGGGATACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCC
ACCACCCTAGCACAAAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAAAGTTCAAGCCTGAAA
TTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCTGGTGAACCTGGCGATAAGATCACATTTGAGGCCACCGGAAATC
TGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATCATTATCTCTGATACCCCTGTGCACGACTGTAATACCACCTGTGACACAC
CTAAGGGCGCCATTAATACCAGCTGCCCTTCCAGAATATTCACCTATCACCATCGCAAGTGTCTAAGTATGTGAAGAGACCAAGCTGAGACTGGCTA

CCGGTCTGAGAAATATCGAAAACCTGTATTTTCAAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAATGGTGGATGGCTGGTACG
GCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCCGCATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACAAGGTGAACAGCGTGATCG
AGAAGATGAACACCCAGTTTACAGCTGTGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGGACGACGGCTTCTCGGATA
TTTGGACCTACAATGCCGAGCTGCTCGTGCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGTACGAGAAGGTGAGAAGCC
AGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGAAGAACGGCACCTACGACT
ACCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCACAAGAATCTATCAGGGCGTGAAAACCTGTATTTTC
AGGGCGGTCTGTTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTCTGCTGTCTACCTTCTGGGTACCATC
ATCACCACCATCACCATCATCTAGTATAAagctt

[0300] SEQ ID NO:56: GP67ss-H1H5cs-M2eCa2-TEV-폴드온-10His의 펩타이드 서열 (H1H5cs-M2eCa)

[0301] MVSAILVYVLLAAAAHSAFADTLICGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYVFVGSSRYSKFKPSLLTEVETPTRNGWECKSDSGRMNYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNT
TCQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEIT
NKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKDNTCMES
VKNGTYDYPKYSEEAKLNREEIDGVKLESTRIFYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGWVLLSTFLGHHHHHHHHHH

[0302] SEQ ID NO:57: M2e 펩타이드의 펩타이드 서열

[0303] SLLTEVETPTRNGWECKSDS

[0304] SEQ ID NO:58: GP67ss-H1H5cs-M2eCa2-TEV-폴드온-10His의 뉴클레오타이드 서열 (H1H5cs-M2eCa)

[0305] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAAGCAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTTCGAGA
TCTTCCTCAAGACAAGCAGCTGGCCTAATCAGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGCAACAGTACCCTAAGCTGTCTAAGAGTACATCAACGACAAGGGCAAAGAAGTGTGGTGTGTGGGAATCCACCACCTAGCA
CAAGCGCGATCAGCAGAGCCTGTACCAGAATGCCGATGCTATGTGTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTTCCCTGTGACCGAGG
TGGAGACCCACCAGGAACGGCTGGGAGTGCAAGTGCTCCGACTCCGGCAGAATGAATTACTACTGGACCCTGGTGGAACTGGCGATAAGATCACATTTG
AGGCCACCGAAATCTGGTGGTGCTAGATATGCATTTGCTATGGAGAGAAATGTGGCTCTGGCATATTATCTGTATACCCCTGTGCACGACTGTAATA
CCACCTGTACAGACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTACCCTATCACCATCGCAAGTGTCTTAAGTATGTGAAGAGCACCA
AGCTGAGACTGGCTACCGTCTGAGAAATAGCCCTCAGAGGAGAGACGCAAGAAGAGAGGCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGA
CCGGAATGGTGGATGGCTGTACGGCTATCACCACCAGAATGAGCAGGATCCGGATATGCCCGGATCTGAAGTCTACACAGAACGCCATCGACGAGATCA
CAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGA
AAGTGGACGACGGCTTCTGGATATTGGACCTACAATGCCGAGCTGCTGCTGCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGA
ACCTGTACGAGAAGGTGAGAAGCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGT
CTGTGAAGAACGGCACCTACGACTACCTAAGTATAGCGAGGAGCCAAAGCTGAATAGAGAGGAGATCGACGGCGTGAACTGGAAAGCACAAGAATCTATC
AGGGCGGTGAAAACCTGTATTTTACGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTCTGC
TGTCTACCTTCTGGGTACCATCATCACCACCATCACCATCATCTAGTATAAagctt

[0306] SEQ ID NO:59: GP67ss-H1H5cs-M2eSb-TEV-폴드온-10His의 펩타이드 서열 (H1H5cs-M2eSb)

[0307] MVSAILVYVLLAAAAHSAFADTLICGYHANNSTDVDTVLEKNVTVTHSVNLEDKHNGLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIIHPSS
LLTEVETPTRNGWECKSDSDAYVFVGSSRYSKFKPEIAIRPKVRDQEGRMNYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNT
TCQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPQRERRKKRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEIT
NKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKDNTCMES
VKNGTYDYPKYSEEAKLNREEIDGVKLESTRIFYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGWVLLSTFLGHHHHHHHHHH

[0308] SEQ ID NO:60: GP67ss-H1H5cs-M2eSb-TEV-폴드온-10His의 뉴클레오타이드 서열 (H1H5cs-M2eSb)

[0309] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGCGCATTTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGCTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGGAAA

CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCAGCTTCGAGAGATTGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCAGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACCACCTAGCT
CCCTGCTGACCGAGGTGGAGACCCCCACAGGAACGGCTGGGAGTGCAAGTGCTCCGACTCCGATGCCTATGTGTTTGTGGGCAGCAGCAGATACAGCAAAA
AGTTCAAGCCTGAAATTGCCATTAGACCCAAAGTGAGAGATCAGGAAGGCAGAATGAATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTG
AGGCCACCGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATATTATCTCTGATACCCCTGTGCACGACTGTAATA
CCACCTGTGACAGACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACC
AGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTCAGAGGGAGAGACGCAAGAAGAGAGGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGA
CCGAATGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCA
CAAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGA
AAGTGGACGACGGCTTCCTGGATATTGGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGA
ACCTGTACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAAACCTGTATGGAGT
CTGTGAAGAACGGCACCTACGACTACCCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATC
AGGGCGCTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGC
TGTCTACCTTCTGGGTACCATCATCACCACCATCACCATCATCACTGATAAaagctt

[0310] SEQ ID NO:61: GP67ss-H1TEV2-M2eCa2-TEV-폴드온-10His의 펩타이드 서열 (H1TEV2-M2eCa)

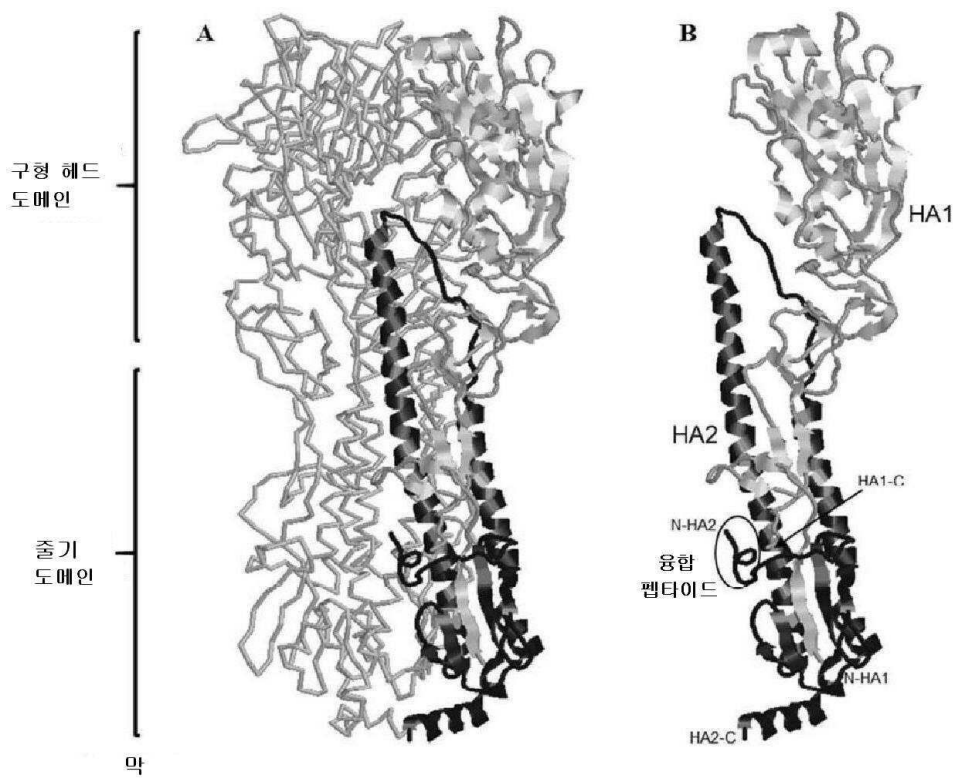
[0311] MVSAILVLYVLLAAAAHSADFATLCIGYHANNSTDVDTVLEKNVTVTHSVNLEDDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVET
PSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWL VKKGNISYPKLSKSYINDKGKEVLVLWGIIHPST
SADQQSLYQNADAYVFGSSRYSKKFKPSLLTEVETPTRNGWECKSDSGRMNYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNT
TCQTPKGAINTSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNSPENLYFQGLFGAIAAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNK
VNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKLKNAKEIGNGCFEFYHKCDNTCMESVK
NGTYDYPKYSEEAKLNREEIDGVKLESTRIYQGAENLYFQGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHHHHH

[0312] SEQ ID NO:62: GP67ss-H1TEV2-M2eCa2-TEV-폴드온-10His의 뉴클레오타이드 서열 (H1TEV2-M2eCa)

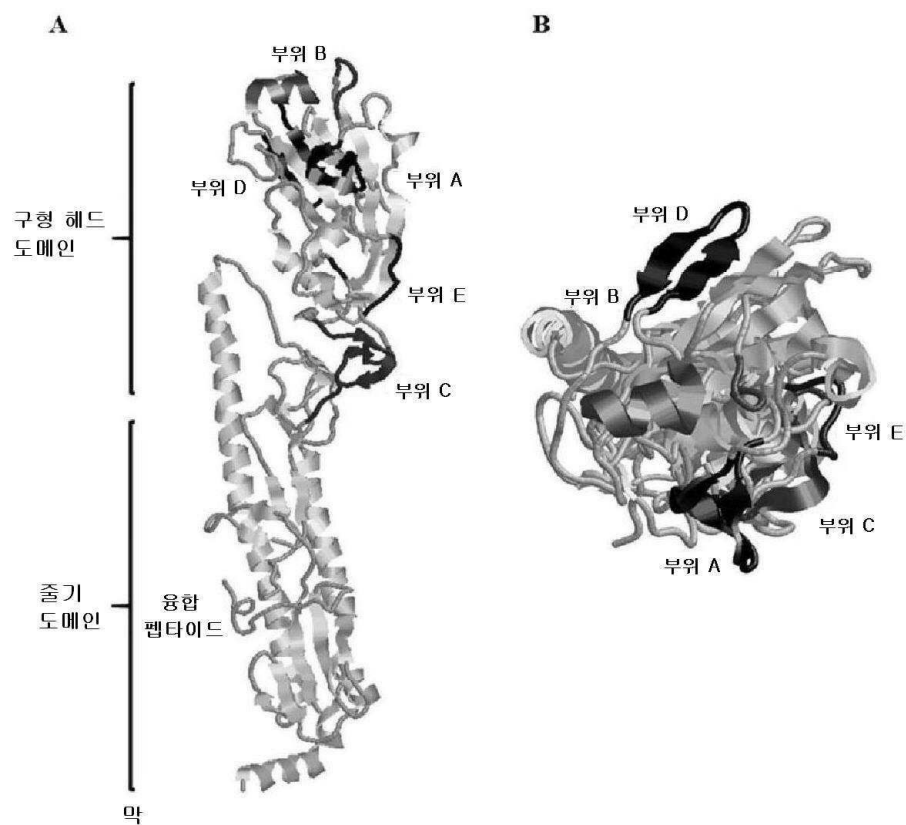
[0313] ccATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGCGCGCATTCTGCCTTTGCGGATACACTGTGTATTGGCTACCACGCCAACAAATAGCACCG
ATACCGTGGATACAGTGCTGGAGAAGATGTGACCGTGACCCACTCTGTGAATCTGTGGAGGATAAGCACAATGGCAAGCTGTGAAGCTGAGAGGAGTTG
CCCCTCTGCACCTGGGCAATGTAATATTGCCGGCTGGATTCTGGGAAATCCTGAATGTGAAAGCCTGTCTACAGCCAGCAGCTGGTCTTATATCGTGAAA
CCCCTAGCAGCGACAATGGCACCTGTTACCTGGCGACTTCATCGATTACGAGGAGCTGAGAGAACAGCTGTCTAGCGTGTCCAGCTTCGAGAGATTGAGA
TCTTCCCTAAGACAAGCAGCTGGCCTAATCAGATTCTAATAAGGGAGTGACAGCCGCTGTCTCATGCCGAGCCAAGTCCTTTTACAAGAACCTGATCT
GGCTGGTGAAGAAGGGCAACAGCTACCCCTAAGCTGTCTAAGAGCTACATCAACGACAAGGGCAAAGAAGTGCTGGTGTGTGGGGAATCCACCACCTAGCA
CAAGCGCGATCAGCAGACCTGTACCAGAATGCCGATGCCTATGTGTTTGTGGGAGCAGCAGATACAGCAAAAAGTTCAAGCCTTCCCTGTGACCGAGG
TGGAGACCCACCAGGAACGGCTGGGAGTGCAAGTGCTCCGACTCCGGCAGAATGAATTACTACTGGACCCTGGTGGAACCTGGCGATAAGATCACATTTG
AGGCCACCGAAATCTGGTGGTGCCTAGATATGCATTTGCTATGGAGAGAAATGCTGGCTCTGGCATATTATCTCTGATACCCCTGTGCACGACTGTAATA
CCACCTGTGACAGACCTAAGGGCGCCATTAATACCAGCCTGCCCTTCCAGAATATTCACCCTATCACCATCGGCAAGTGTCCTAAGTATGTGAAGAGCACC
AGCTGAGACTGGCTACCGGTCTGAGAAATAGCCCTGAAAACCTGTATTTTCAAGCCTGTTTGGAGCCATCGCCGGCTTTATTGAGGGAGGATGGACCGGAA
TGGTGGATGGCTGGTACGGCTATCACCACCAGAATGAGCAGGGATCCGATATGCCGCCGATCTGAAGTCTACACAGAACGCCATCGACGAGATCACAACA
AGGtGAACAGCGTGATCGAGAAGATGAACACCCAGTTTACAGCTGTGGGCAAGGAGTTCAACCACCTGGAGAAGAGAATCGAGAACCTGAACAAGAAAGTGG
ACGACGGCTTCTGGATATTGGACCTACAATGCCGAGCTGCTCGTCTCTGGAGAATGAGAGAACCCTGGACTACCACGACAGCAATGTGAAGAACCTGT
ACGAGAAGGTGAGAAGCCAGCTGAAGAACAATGCCAAGGAGATCGGCAACGGCTGCTTTGAGTTCTACCACAAGTGTGACAACACCTGTATGGAGTCTGTGA
AGAACGGCACCTACGACTACCCCTAAGTATAGCGAGGAGGCCAAGCTGAATAGAGAGGAGATCGACGGCGTGAAACTGGAAAGCACAAGAATCTATCAGGGCG
CTGAAAACCTGTATTTTCAGGGCGGTTCTGGTTACATCCCGAAGCTCCGCGTGACGGTCAGGCTTACGTTCTGTAAGACGGTGAATGGGTTCTGTGTCTA
CCTTCTGGGTACCATCATCACCACCATCACCATCATCACTGATAAaagctt

도면

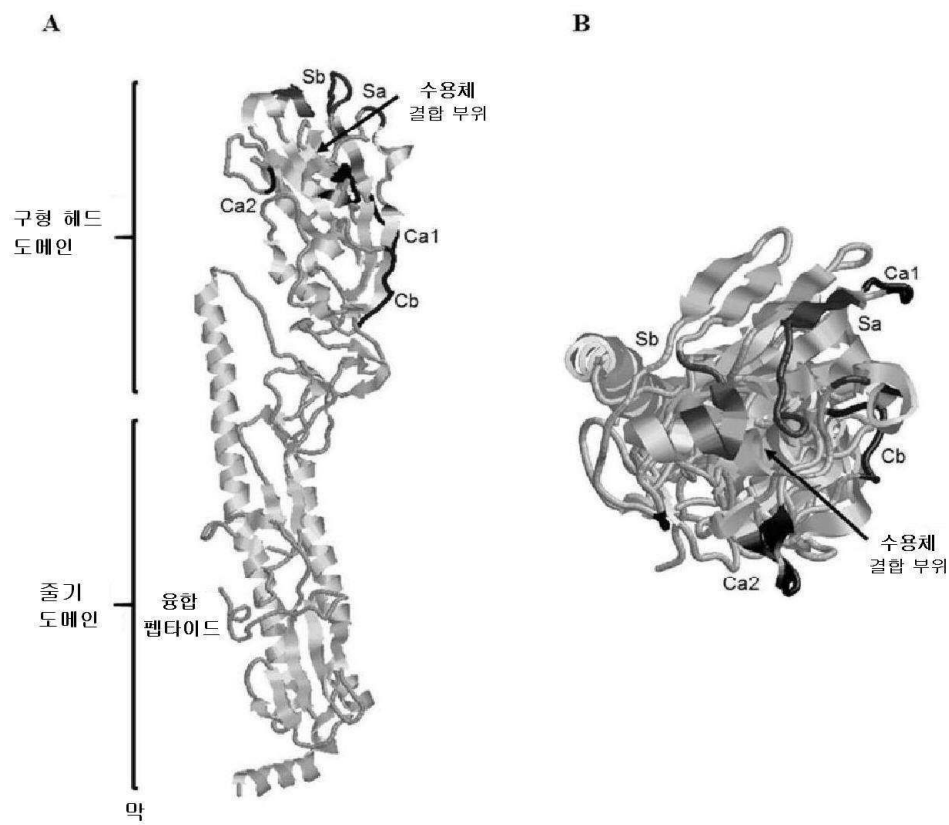
도면1



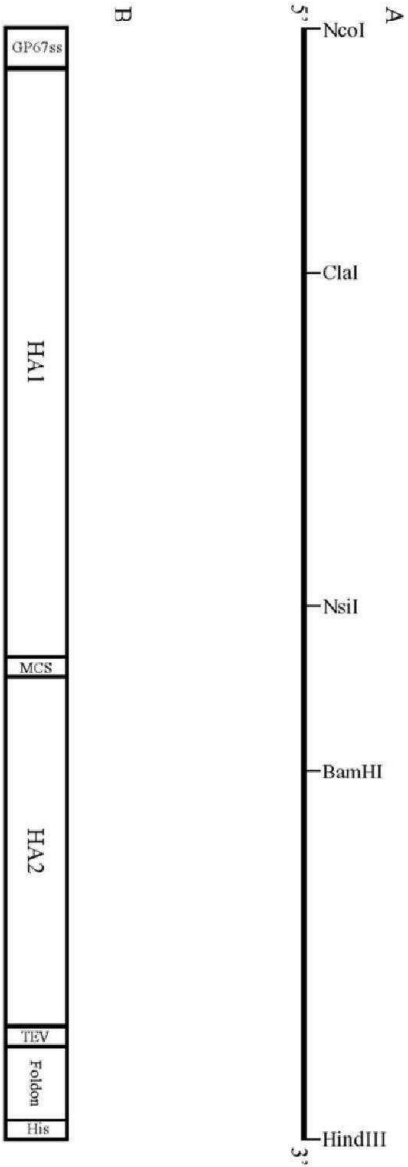
도면2



도면3



도면4



H3 QDLPGNDNSTATTCIGHHAVPNTLVKTIITDDQIEVNTATELVQSSSTGKICNNP-HRIIDGIDCTTIDALIGDPHCD-VFONEIWDLPVER-SKAFSNC 97
WT -----DTLCIGYHANNSTVDVTELENNVTYTHSVNLLIEDNENKLCIKRGVAPLHLGKCNLAGWILGNPECBLSITASWSYIVETPSSDNGTC 90
H1 -----DTICIGYHANNSTVDVTELENNVTYTHSVNLLIEDSHNGKILRLKGIAPLQLGKCNLAGWILGNPECDPLIPVRSWSYIVETPNSENGIC 90

H3 IRYVDVDAISIRSI VASG---TLEFITEGTWTGVTQN-GGSNACKRGPGSGFFSKLNNITKSGSTYPVNVNTWPNNDNFDKLYIWGIIHHPSTINQEQTS 193
WT YPGDFIDYEELRQJSSVSFEFEIPEFISFNSNDNKNLTAACFAGAGSFYKMLIWLIVKNGSYKPLSKSYINDRGKLVYLMGIIHHPSTADQOS 190
H1 YPGDFIDYEELRQJSSVSFEFEIPEFISFNSNPNNNTN-GTTLACSHHGGSFYKMLILITLTKSGSYKPLKNSYNNKKGKGVLYLMGIIHHPNSKEQN 189

CR8020s^{a4} Ca2
HIC9Sa^d HIC9Ca² CR8020s^{a3} HIC9Sa³ F16Sa F16Sab
HIC9Sa^d F16Sa^b

H3 LVYQASGRVYSTRRSQOTIINIGSRFWVBGLSRSISITWITVPGDVLVNSGNMLIAERGFPMRTGR-SSIMRSDAPIDTGISCEITPNSIPNDR 292
WT LYQNDADAVYVSSRSKSKFKFPLAIRKVDQGRNNTYWTIVPEGDKITFATGNLVVPRYAFAMENNASGIIISDTPYHDCNTTCQTPKGAINTSL 290
H1 LYQNDNAVSVTSNRRFTPELIERPRVDDQGRNNTYWTLLKPGDTIIFEANGMLIAEMYAFALSRGFSGLITSNASMHCNTKCQTPLAGAINSSL 289

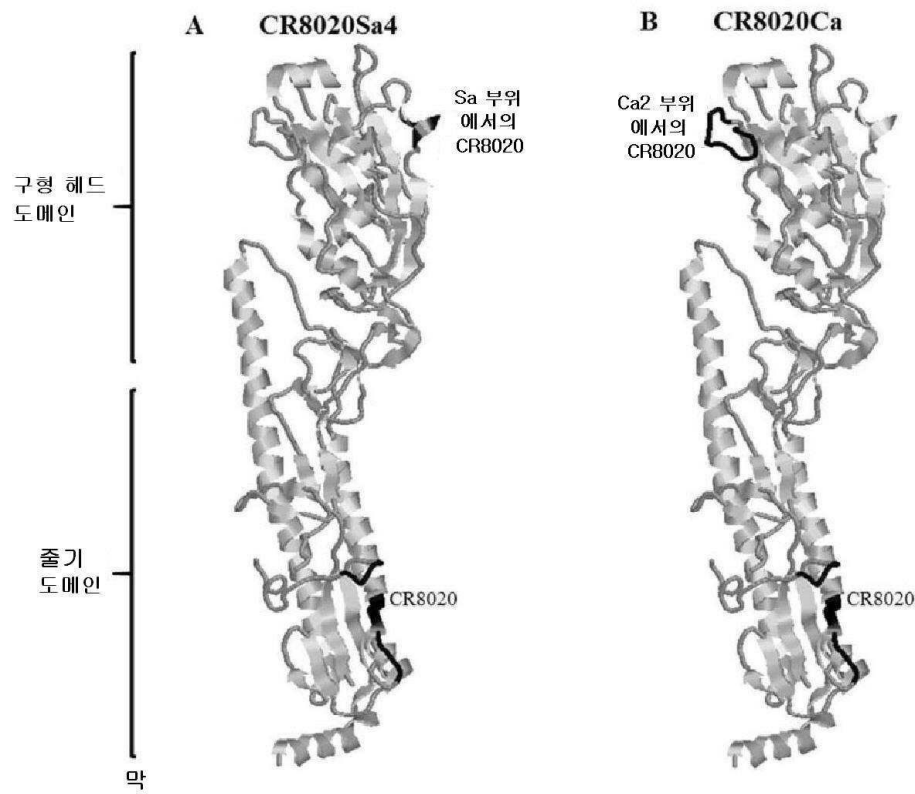
CR8020Ca^a Ca²
HIC9Ca¹ HIC9Ca¹ MeCa^a HIC9Ca^a

H3 PRQNVNKITYACPKYKQNTIKLATGKRNVEKQTHLEFALAGFIENGNGMIDWYGRHQNSEGTQAAADLKSTQAAIDQINGKLNRYIEKTNKEF 392
WT PRQNIHPTTIGCKPKYKSTIKLALGRLNVEISIOSGLFAGIAGFIENGWAGWYGHQNSQSGYAADLKSTQAAIDETNKNVSVTEKMTQF 390
H1 PRQNIHPTTIGCKPKYKSAKLAWTGLRNLISIOSGLFAGIAGFIENGWGMIDWYGHQNSQSGYAADQKSTQANAINGITNKNVTVIEKMTQF 389

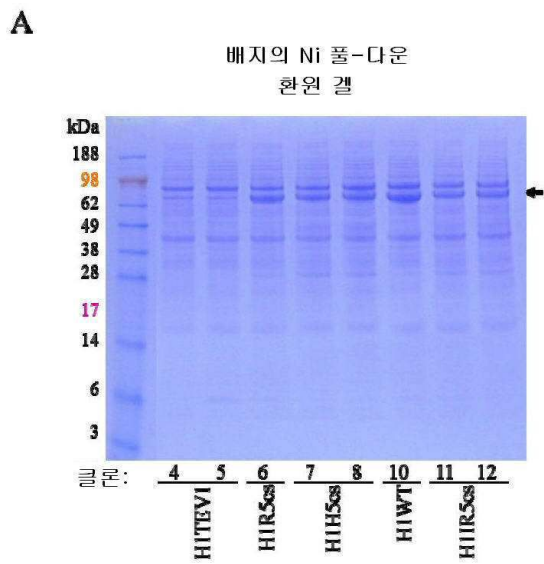
H5G, K5G, IRCS⁸
TEV1, TEV2, TEV3
성숙 결단 부위 (MCS)

H3 HQIEREFSVEGRIOJLERVEEDTKIDMSYNELIVALENQHTIIDNDEENKLFEKTRQJRENAEEMNGCFKIYHKQDNACIESIRNGTYDHDVYR 492
WT TAVGKEFVHLEKRIEMLNKNKVDGFLDWTYTNELIYVLENEREDYHDSNRKL YERYASQJKNNAKEIUNGCEFEYHKDNTCMESYINGTYD YPKYS 490
H1 TAVGKEFVHLEKRMENINRKNVDGFLDWTYNELIYVLENEREDYHDSNRKL YERYASQJKNNAKEIUNGCEFEYHKDNECMESYVNGTYD YPKYS 489
H3 DEALNNRPQIKVEIKSG--YKDW 514
WT EAKLNREIDGVLESIRIYQ 512
H1 ESKLNREKVDGVLESIGIYQ 511

도면6



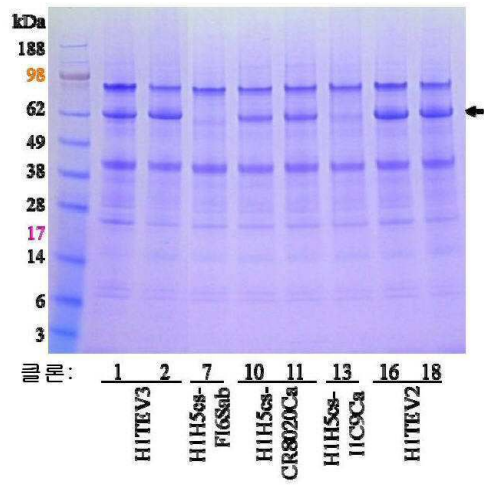
도면7a



도면7b

B

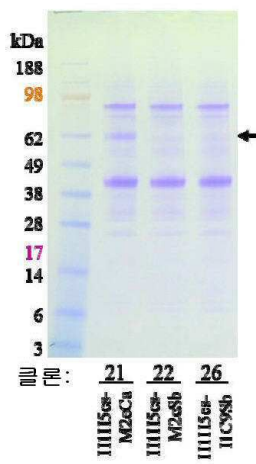
배지의 Ni 풀-다운
비-환원 겔



도면7c

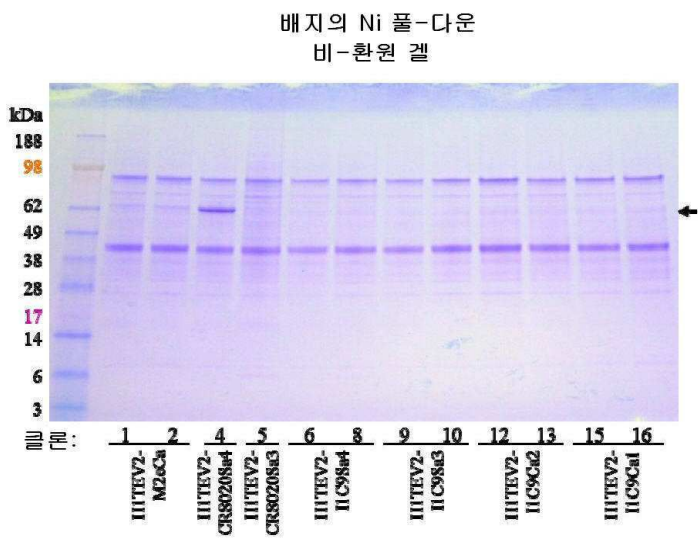
C

배지의 Ni 풀-다운
비-환원 겔



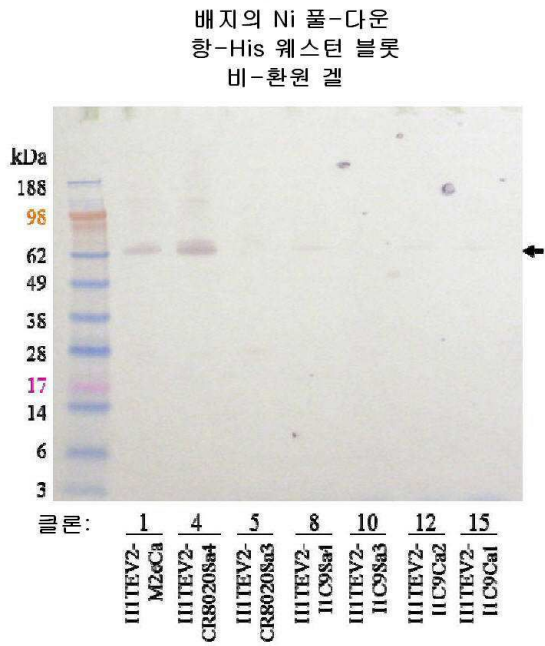
도면7d

D

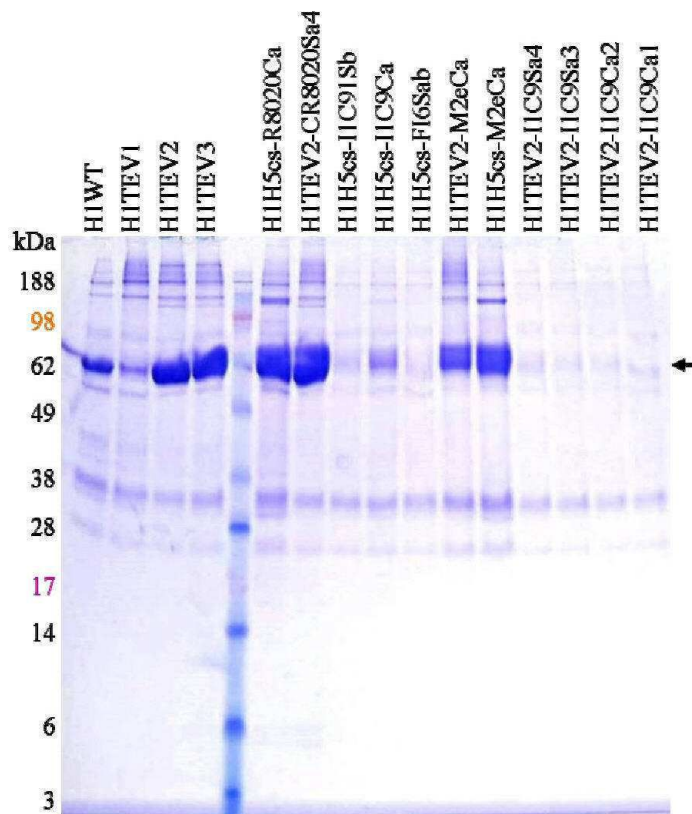


도면7e

E



도면8



서열 목록

SEQUENCE LISTING

<110> CG Discovery, Inc.

LUO, Chun

<120> COMPOSITIONS OF INFLUENZA HEMAGGLUTININ WITH HETEROLOGOUS EPITOPES AND/OR ALTERED MATURATION CLEAVAGE SITES AND METHODS OF USE THEREOF

<130> 758252000140

<140> Not Yet Assigned

<141> Concurrently Herewith

<150> US 62/290,894

<151> 2016-02-03

<160> 62

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 566

<212> PRT

<213> Influenza A virus

<400> 1

Met Lys Ala Ile Leu Val Val Leu Leu Tyr Thr Phe Ala Thr Ala Asn

1 5 10 15
Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr

20 25 30
Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
35 40 45

Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu Arg Gly Val
50 55 60
Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp Ile Leu Gly

65 70 75 80
Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp Ser Tyr Ile
85 90 95

Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro Gly Asp Phe
100 105 110
Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe

115 120 125
Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro Asn His Asp

130 135 140
Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly Ala Lys Ser
145 150 155 160

Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn Ser Tyr Pro
165 170 175
Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu Val Leu Val

180 185 190
Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu

195 200 205
Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser
210 215 220

Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln
225 230 235 240

Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys
 245 250 255
 Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe
 260 265 270
 Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro
 275 280 285
 Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn
 290 295 300
 Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys
 305 310 315 320
 Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg
 325 330 335
 Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
 340 345 350
 Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
 355 360 365
 His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser
 370 375 380
 Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile
 385 390 395 400
 Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His
 405 410 415
 Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe
 420 425 430
 Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn
 435 440 445
 Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu
 450 455 460
 Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly
 465 470 475 480
 Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val

485 490 495
Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu
500 505 510
Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr

515 520 525
Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Val
530 535 540
Val Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu
545 550 555 560
Gln Cys Arg Ile Cys Ile
565

<210> 2

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 2

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15
Ser Ala Phe Ala
20

<210> 3

<211> 49

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 3

Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu
1 5 10 15
Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val
20 25 30

Leu Leu Ser Thr Phe Leu Gly His His His His His His His His

35

40

45

His

<210> 4

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 4

Glu Asn Leu Tyr Phe Gln Gly

1

5

<210> 5

<211> 29

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 5

Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val

1

5

10

15

Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu

20

25

<210> 6

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 6

His His His His His His His His His His

1

5

10

<210> 7

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 7

Met Lys Ala Ile Leu Val Val Leu Leu Tyr Thr Phe Ala Thr Ala Asn

1 5 10 15

Ala

<210> 8

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 8

Ile Pro Ser Ile Gln Ser Arg

1 5

<210> 9

<211> 581

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 9

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
 50 55 60
 Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
 65 70 75 80
 Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
 85 90 95
 Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
 100 105 110
 Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
 115 120 125
 Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
 130 135 140
 Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
 145 150 155 160
 Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
 165 170 175
 Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
 180 185 190
 Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
 195 200 205
 Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
 210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
 225 230 235 240
 Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
 245 250 255
 Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
 260 265 270
 Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser
 275 280 285
 Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly

290 295 300
 Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
 305 310 315 320
 Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
 325 330 335
 Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala
 340 345 350
 Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp

 355 360 365
 Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp
 370 375 380
 Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn
 385 390 395 400
 Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu
 405 410 415
 Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp

 420 425 430
 Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu
 435 440 445
 Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn
 450 455 460
 Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile
 465 470 475 480
 Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met

 485 490 495
 Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu
 500 505 510
 Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr
 515 520 525
 Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly
 530 535 540

Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp

545 550 555 560

Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His His

565 570 575

His His His His His

580

<210> 10

<211> 1757

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 10

ccatggttaag cgctattgtt ttatatgtgc ttttggcggc ggccggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420
caagcagctg gcctaatac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accacctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660

ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgcct atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atatccctag catccagagc agaggcctgt ttggagccat cgccggcttt attgaggag 1080

gatggaccgg aatggtggat ggctggtacg gctatcacca ccagaatgag cagggatccg 1140

gatatgccgc cgatctgaag tctacacaga acgccatcga cgagatcaca aacaaggatga 1200

acagcgtgat cgagaagatg aacacccagt ttacagctgt gggcaaggag ttcaaccacc 1260

tggagaagag aatcgagaac ctgaacaaga aagtggacga cggcttcctg gatatttga 1320

cctacaatgc cgagctgctc gtgctcctgg agaatgagag aacctggac taccacgaca 1380

gcaatgtgaa gaacctgtac gagaaggatga gaagccagct gaagaacaat gccaaggaga 1440

tcggcaacgg ctgctttgag ttctaccaca agtgtgacaa cacctgtatg gagtctgtga 1500

agaacggcac ctacgactac cctaagtata gcgaggaggc caagctgaat agagaggaga 1560

tcgacggcgt gaaactggaa agcacaagaa tctatcaggg cgctgaaaac ctgtattttc 1620

agggcggttc tggttacatc ccggaagctc cgcgtgacgg tcaggcttac gttcgtaaag 1680

acggtgaatg ggttctgtctg tctaccttcc tgggtcacca tcacaccac catcaccatc 1740

atcactgata aaagctt 1757

<210> 11

<211> 584

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 11

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

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

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

 210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
 225 230 235 240
 Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
 245 250 255
 Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
 260 265 270
 Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser

 275 280 285
 Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
 290 295 300
 Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
 305 310 315 320
 Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
 325 330 335
 Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly Leu

340 345 350
Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
355 360 365
Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
370 375 380
Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn
385 390 395 400
Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val

405 410 415
Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys
420 425 430
Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
435 440 445
Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn
450 455 460
Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala

465 470 475 480
Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn
485 490 495
Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
500 505 510
Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu
515 520 525
Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly

530 535 540
Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
545 550 555 560
Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His
565 570 575
His His His His His His His His
580

<210> 12

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 12

Gln Arg Glu Arg Arg Lys Lys Arg

1

5

<210> 13

<211> 1766

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 13

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggaggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcacg gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420
caagcagctg gcctaatac gattctaata agggagtgac agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accacccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatag agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcttagata tgcatittgt atggagagaa 840

atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgcccttcca gaatattcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020

atagccctca gagggagaga cgcaagaaga gaggcctgtt tggagccatc gccggcttta 1080
 ttgaggagg atggaccgga atggtgatg gctggtacgg ctatcaccac cagaatgagc 1140
 agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcacia 1200
 acaagggtgaa cagcgtgac gagaagatga acaccagtt tacagctgtg ggcaaggagt 1260
 tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320

atatttgac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
 accacgacag caatgtgaag aacctgtacg agaaggtag aagccagctg aagaacaatg 1440
 ccaaggagat cggcaacggc tgctttgagt tctaccacia gtgtgacaac acctgtatgg 1500
 agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
 gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
 tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680
 ttcgtaaaga cggatgaatgg gttctgctgt ctaccttcct gggtcaccat catcaccacc 1740
 atcaccatca tcaatgataa aagctt 1766

<210> 14

<211> 581

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 14

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140
Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160
Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175
Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190
Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205
Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220
Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
225 230 235 240
Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
245 250 255
Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
260 265 270
Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser
275 280 285
Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
290 295 300
Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
305 310 315 320
Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
325 330 335
Gly Leu Arg Asn Ile Pro Arg Arg Arg Arg Gly Leu Phe Gly Ala

340 345 350
Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp

355 360 365
Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp

370 375 380
Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn

385 390 395 400
Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu

405 410 415
Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp

420 425 430
Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu

435 440 445
Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn

450 455 460
Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile

465 470 475 480
Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met

485 490 495
Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu

500 505 510
Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr

515 520 525
Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly

530 535 540
Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp

545 550 555 560
Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His His

565 570 575
His His His His His

580

<210> 15

<211> 1757

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 15

ccatggaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacac aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accacctag cacaagcgc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatag agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780

tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgc atggagagaa 840
atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgccttcca gaatattcac cctatcacca 960
tcggcaagt tccaaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atatccctag gagacgcaga agaggcctgt ttggagccat cgccggcttt attgaggag 1080
gatggaccgg aatggtggat ggctggtacg gctatcacca ccagaatgag cagggatccg 1140
gatatgccgc cgatctgaag tctacacaga acgcatcga cgagatcaca aacaagtgta 1200
acagcgtgat cgagaagatg aacacccagt ttacagctgt gggcaaggag ttcaaccacc 1260

tggagaagag aatcgagaac ctgaacaaga aagtggacga cggcttcctg gatatttga 1320
cctacaatgc cgagctgtct gtgctcctgg agaagtagag aacctggac taccacgaca 1380
gcaatgtgaa gaacctgtac gagaagtgta gaagccagct gaagaacaat gccaaggaga 1440
tcggcaacgg ctgctttgag ttctaccaca agtgtgacaa cacctgtatg gagtctgtga 1500

agaacggcac ctacgactac cctaagtata gcgaggaggc caagctgaat agagaggaga 1560
 tcgacggcgt gaaactggaa agcacaagaa tctatcaggg cgctgaaaac ctgtattttc 1620
 agggcggttc tggttacatc ccggaagctc cgcgtgacgg tcaggcttac gttcgtaaag 1680
 acggtgaatg gggtctgctg tctaccttcc tgggtcacca tcatcaccac catcaccatc 1740

atcactgata aaagctt 1757

<210> 16

<211> 585

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 16

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
225 230 235 240

Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
245 250 255

Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
260 265 270

Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser
275 280 285

Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
290 295 300

Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
305 310 315 320

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
325 330 335

Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Arg Arg Arg Gly
340 345 350

Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met
355 360 365

Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly
370 375 380

Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr
385 390 395 400

Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala

405 410 415
Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
420 425 430

Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
435 440 445

Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
450 455 460

Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
465 470 475 480

Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
485 490 495

Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
500 505 510

Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
515 520 525

Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln
530 535 540

Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
545 550 555 560

Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His
565 570 575

His His His His His His His His His
580 585

<210> 17

<211> 1769

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 17

ccatggttaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatac gattetaata agggagtac agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540

agctgtctaa gagctacac aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600

accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660

ttgtgggcag cagcagatag agcaaaaagt tcaagcctga aattgccatt agacccaaag 720

tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780

tcacatttga ggccaccgga aatctgggtg tgcttagata tgcatctgct atggagagaa 840

atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900

agacaccta gggcgcatt aataccagcc tgccttcca gaatttcac cctatcacca 960

tcggcaagtg tcctaaglat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020

atatccctag catccagagc aggagacgca gaagaggcct gtttgagacc atcgccggct 1080

ttattgaggg aggatggacc ggaatgggtg atggctggta cggctatcac caccagaatg 1140

agcagggatc cgatatgcc gccgatctga agtctacaca gaacccatc gacgagatca 1200

caaacaaggt gaacagcgtg atcgagaaga tgaacacca gtttacagct gtgggcaagg 1260

agttcaacca cctggagaag agaatcgaga acctgaaca gaaagtggac gacggcttcc 1320

tggatatttg gacctacaat gccgagctgc tcgtgctcct ggagaatgag agaaccctgg 1380

actaccagca cagcaatgtg aagaacctgt acgagaaggt gagaagccag ctgaagaaca 1440

atgccaagga gatcggcaac ggctgctttg agttctacca caagtgtgac aacacctgta 1500

tggagtctgt gaagaacggc acctacgact accctaagta tagcgaggag gccaagctga 1560

atagagagga gatcgacggc gtgaaactgg aaagcacaag aatctatcag ggcgctgaaa 1620

acctgtattt tcaggcggt tctggttaca tcccgaagc tccgctgac ggtcaggctt 1680

acgttcgtaa agacggtgaa tgggttctgc tgtctacct cctgggtcac catcatcacc 1740

accatcacca tcatcactga taaaagctt 1769

<210> 18

<211> 581

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 18

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
 225 230 235 240
 Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
 245 250 255
 Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
 260 265 270
 Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser

 275 280 285
 Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
 290 295 300
 Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
 305 310 315 320
 Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
 325 330 335
 Gly Leu Arg Asn Ile Glu Asn Leu Tyr Phe Gln Gly Leu Phe Gly Ala

 340 345 350
 Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp
 355 360 365
 Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp
 370 375 380
 Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn
 385 390 395 400
 Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu

 405 410 415
 Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp
 420 425 430
 Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu
 435 440 445
 Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn
 450 455 460

Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile

465 470 475 480

Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met

485 490 495

Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu

500 505 510

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

515 520 525

Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly

530 535 540

Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp

545 550 555 560

Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His His

565 570 575

His His His His His

580

<210> 19

<211> 1757

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 19

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctgggtgaa gaagggcaac agctacccta 540

agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctgggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatag agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctgggtgg tgcctagata tgcatttctt atggagagaa 840
atgtggctc tggcatcatt atctctgata ccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgccttcca gaatttcac cctatcacca 960
tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020

atatcgaata cctgtatttt caaggcctgt ttggagccat cgccggcttt attgaggag 1080
gatggaccgg aatggtggat ggctgggtacg gctatcacca ccagaatgag cagggatccg 1140
gatatgccgc cgatctgaag tctacacaga acgcatcga cgagatcaca aacaaggatga 1200
acagcgtgat cgagaagatg aacacccagt ttacagctgt gggcaaggag ttcaaccacc 1260
tggagaagag aatcgagaac ctgaacaaga aagtggacga cggttctctg gatatttga 1320
cctacaatgc cgagctgctc gtgctcctgg agaattgagag aacctggac taccacgaca 1380
gcaatgtgaa gaacctgtac gagaaggatga gaagccagct gaagaacaat gccaaggaga 1440
tcggcaacgg ctgctttgag ttctaccaca agtgtgacaa cacctgtatg gagtctgtga 1500

agaacggcac ctacgactac cctaagtata gcgaggaggc caagctgaat agagaggaga 1560
tcgacggcgt gaaactggaa agcacaagaa tctatcaggg cgctgaaaac ctgtattttc 1620
agggcgggtc tggttacatc ccggaagctc cgctgacgg tcaggcttac gttcgtaaag 1680
acggtgaatg ggttctgctg tctaccttcc tgggtcacca tcatcaccac catcaccatc 1740
atcactgata aaagctt 1757

<210> 20

<211> 582

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 20

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

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

 65 70 75 80
 Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
 85 90 95
 Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
 100 105 110
 Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
 115 120 125
 Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

 130 135 140
 Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
 145 150 155 160
 Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
 165 170 175
 Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
 180 185 190
 Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

 195 200 205
 Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
 210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
 225 230 235 240
 Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
 245 250 255
 Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg

260	265	270	
Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser			
275	280	285	
Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly			
290	295	300	
Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile			
305	310	315	320
Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr			
325	330	335	
Gly Leu Arg Asn Ser Pro Glu Asn Leu Tyr Phe Gln Gly Leu Phe Gly			
340	345	350	
Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly			
355	360	365	
Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala			
370	375	380	
Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val			
385	390	395	400
Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys			
405	410	415	
Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val			
420	425	430	
Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val			
435	440	445	
Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys			
450	455	460	
Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu			
465	470	475	480
Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys			
485	490	495	
Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu			
500	505	510	

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

515 520 525
 Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser
 530 535 540
 Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys
 545 550 555 560
 Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His His
 565 570 575
 His His His His His His
 580

<210> 21

<211> 1760

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 21

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60
 cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
 tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
 agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
 ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
 tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcacg gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
 caagcagctg gcctaatac gattetaata agggagtgac agccgcctgt cctcatgceg 480
 gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
 agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
 accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
 ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
 tgagagatca ggaagcgaga atgaattact actggaccct ggtggaacct ggcgataaga 780
 tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatittgt atggagagaa 840

atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900
 agacacctaa gggcgccatt aataccagcc tgccttcca gaatattcac cctatcacca 960
 tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
 atagccctga aaacctgtat tttcaaggcc tgtttggagc catcgccggc tttattgagg 1080
 gaggatggac cggaatggtg gatggctggt acggctatca ccaccagaat gaggaggat 1140
 ccggatatgc cgccgatctg aagtctacac agaacgcat cgacgagatc acaacaagg 1200
 tgaacagcgt gatcgagaag atgaacaccc agtttacagc tgtgggcaag gatttcaacc 1260
 acctggagaa gagaatcgag aacctgaaca agaaagtga cgacggcttc ctggatattt 1320

ggacctacaa tgccgagctg ctctgtctcc tggagaatga gagaacctg gactaccagc 1380
 acagcaatgt gaagaacctg tacgagaagg tgagaagcca gctgaagaac aatgccaagg 1440
 agatcgcaa cgctgcttt gagttctacc acaagtgtga caacacctgt atggagtctg 1500
 tgaagaacgg cacctacgac taccctaagt atagcgagga ggccaagctg aatagagagg 1560
 agatcgacgg cgtgaaactg gaaagcaca gaatctatca gggcgctgaa aacctgtatt 1620
 ttcaggggcg tcttggttac atcccgaag ctccgctga cggtcaggct tacgttcgta 1680
 aagacggtga atgggttctg ctgtctacct tctgggtca ccatcatcac caccatcacc 1740
 atcatcactg ataaaagctt 1760

<210> 22

<211> 584

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 22

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65						70						75						80	
Ile	Leu	Gly	Asn	Pro	Glu	Cys	Glu	Ser	Leu	Ser	Thr	Ala	Ser	Ser	Trp				
					85					90					95				
Ser	Tyr	Ile	Val	Glu	Thr	Pro	Ser	Ser	Asp	Asn	Gly	Thr	Cys	Tyr	Pro				
					100					105					110				
Gly	Asp	Phe	Ile	Asp	Tyr	Glu	Glu	Leu	Arg	Glu	Gln	Leu	Ser	Ser	Val				
					115					120					125				
Ser	Ser	Phe	Glu	Arg	Phe	Glu	Ile	Phe	Pro	Lys	Thr	Ser	Ser	Trp	Pro				
					130					135					140				
Asn	His	Asp	Ser	Asn	Lys	Gly	Val	Thr	Ala	Ala	Cys	Pro	His	Ala	Gly				
145					150					155					160				
Ala	Lys	Ser	Phe	Tyr	Lys	Asn	Leu	Ile	Trp	Leu	Val	Lys	Lys	Gly	Asn				
					165					170					175				
Ser	Tyr	Pro	Lys	Leu	Ser	Lys	Ser	Tyr	Ile	Asn	Asp	Lys	Gly	Lys	Glu				
					180					185					190				
Val	Leu	Val	Leu	Trp	Gly	Ile	His	His	Pro	Ser	Thr	Ser	Ala	Asp	Gln				
					195					200					205				
Gln	Ser	Leu	Tyr	Gln	Asn	Ala	Asp	Ala	Tyr	Val	Phe	Val	Gly	Ser	Ser				
210					215					220									
Arg	Tyr	Ser	Lys	Lys	Phe	Lys	Pro	Glu	Ile	Ala	Ile	Arg	Pro	Lys	Val				
225					230					235					240				
Arg	Asp	Gln	Glu	Gly	Arg	Met	Asn	Tyr	Tyr	Trp	Thr	Leu	Val	Glu	Pro				
					245					250					255				
Gly	Asp	Lys	Ile	Thr	Phe	Glu	Ala	Thr	Gly	Asn	Leu	Val	Val	Pro	Arg				
					260					265					270				
Tyr	Ala	Phe	Ala	Met	Glu	Arg	Asn	Ala	Gly	Ser	Gly	Ile	Ile	Ile	Ser				
					275					280					285				
Asp	Thr	Pro	Val	His	Asp	Cys	Asn	Thr	Thr	Cys	Gln	Thr	Pro	Lys	Gly				
290					295					300									
Ala	Ile	Asn	Thr	Ser	Leu	Pro	Phe	Gln	Asn	Ile	His	Pro	Ile	Thr	Ile				
305					310					315					320				

Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr
 325 330 335
 Gly Leu Arg Asn Ile Pro Ser Ile Glu Asn Leu Tyr Phe Gln Gly Leu
 340 345 350
 Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
 355 360 365
 Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
 370 375 380
 Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn
 385 390 395 400
 Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
 405 410 415
 Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys
 420 425 430
 Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
 435 440 445
 Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn
 450 455 460
 Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala
 465 470 475 480
 Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn
 485 490 495
 Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
 500 505 510
 Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu
 515 520 525
 Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly
 530 535 540
 Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
 545 550 555 560
 Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His

565

570

575

His His His His His His His His

580

<210> 23

<211> 1766

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 23

ccatggttaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420

caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540

agctgtctaa gagctacac aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600

accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660

ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt agacccaaag 720

tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780

tcacatttga ggccaccgga aatctggtgg tgcttagata tgcatcttgc atggagagaa 840

atgctggctc tggcatcatt atctctgata cccctgtgca cgactgtaat accacctgtc 900

agacaccta gggcgccatt aataccagcc tgcccttcca gaatttcac cctatcacca 960

tcggcaagtg tcctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020

atatccctag catcgaaaac ctgtattttc aaggcctgtt tggagccatc gccggcttta 1080

ttgaggagg atggaccgga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140

agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcacia 1200

acaagtgtaa cagcgtgac gagaagatga acaccagtt tacagctgtg ggcaaggagt 1260

tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320

atatttgac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
accacgacag caatgtgaag aacctgtacg agaagtgtag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680

ttcgtaaaga cggatgaatgg gttctgctgt ctaccttctt gggtcaccat catcaccacc 1740
atcaccatca tcaactgataa aagctt 1766

<210> 24

<211> 585

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 24

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140
 Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
 145 150 155 160

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

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

 Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys
 290 295 300
 Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr
 305 310 315 320
 Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala
 325 330 335
 Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly
 340 345 350

 Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met
 355 360 365
 Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly
 370 375 380

Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr
 385 390 395 400
 Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala
 405 410 415

Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
 420 425 430
 Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
 435 440 445
 Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
 450 455 460
 Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
 465 470 475 480

Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
 485 490 495
 Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
 500 505 510
 Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
 515 520 525
 Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln
 530 535 540

Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr
 545 550 555 560
 Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His
 565 570 575
 His His His His His His His His His
 580 585

<210> 25

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 25

Glu Gly Met Ile Asp

1 5

<210> 26

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 26

Glu Gly Thr Gly Gln Ala Ala

1 5

<210> 27

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 27

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

1 5 10

<210> 28

<211> 1770

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 28

ccatggtaag cgtattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
 agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
 caagcagctg gcctaatac gattctaata agggagtac agccgcctgt cctcatgccg 480
 gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
 agctgtctaa gagctacac aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600

 accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
 ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aggcattgatt gattacgaag 720
 gcacaggcca ggcagccggc agaataatt actactggac cctggtggaa cctggcgata 780
 agatcacatt tgaggccacc ggaaatctgg tggcgcctag atatgcattt gctatggaga 840
 gaaatgctgg ctctggcatc attatctctg ataccctgt gcacgactgt aataccacct 900
 gtcagacacc taagggcgcc attaatacca gcctgccctt ccagaatatt caccctatca 960
 ccatcgga gttgcctaag tatgtgaaga gcaccaagct gagactggct accggtctga 1020
 gaaatagccc tcagaggag agacgaaga agagaggcct gtttgagacc atcgccggt 1080

 ttattgaggg aggatggacc ggaatggtgg atggctggta cgctatcac caccagaatg 1140
 agcagggatc cgcatatgcc gccgatctga agtctacaca gaacccatc gacgagatca 1200
 caaacaaggt gaacagcgtg atcgagaaga tgaacacca gtttacagct gtgggcaagg 1260
 agttcaacca cctggagaag agaatcgaga acctgaacaa gaaagtggac gacggcttc 1320
 tggatatattg gacctacaat gccgagctgc tcgtgctcct ggagaatgag agaaccctgg 1380
 actaccacga cagcaatgtg aagaacctgt acgagaaggt gagaagccag ctgaagaaca 1440
 atgccaagga gatcggaac ggctgctttg agttctacca caagtgtgac aacacctgta 1500
 tggagtctgt gaagaacggc acctacgact accctaagta tagcgaggag gccaagctga 1560

 atagagagga gatcgacggc gtgaaactgg aaagcacaag aatctatcag ggcgctgaaa 1620
 acctgtattt tcagggcggt tctggttaca tcccggaagc tccgcgtgac ggtcaggctt 1680
 acgttcgtaa agacggtgaa tgggttctgc tgtctacctt cctgggtcac catcatcacc 1740
 accatcacca tcatcactga taaaaagctt 1770

 <210> 29
 <211> 583
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic
 <400> 29

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His
1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser
20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His
35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Glu Gly Met Ile Asp Tyr Glu Gly
225 230 235 240

Thr Gly Gln Ala Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu

245 250 255
Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro
260 265 270

Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile
275 280 285
Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys
290 295 300
Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr
305 310 315 320
Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala
325 330 335

Thr Gly Leu Arg Asn Ser Pro Glu Asn Leu Tyr Phe Gln Gly Leu Phe
340 345 350
Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val Asp
355 360 365
Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala
370 375 380
Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys
385 390 395 400

Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly
405 410 415
Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys
420 425 430
Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu
435 440 445
Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser Asn Val
450 455 460

Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn Ala Lys
465 470 475 480
Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Thr
485 490 495

Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser
500 505 510

Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu
515 520 525

Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly
530 535 540

Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg
545 550 555 560

Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His His
565 570 575

His His His His His His His
580

<210> 30

<211> 1763

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 30

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggccggcgcac tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacac aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cacaagcgc gatacagaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatag agcaaaaagt tcaagcctga aggcattgatt gattacgaag 720
gcacaggcca ggcagccggc agaataaatt actactggac cctgggtgaa cctggcgata 780

agatcacatt tgaggccacc ggaaatctgg tggtagcctag atatgcattt gctatggaga 840
gaaatgctgg ctctggcatc attatctctg ataccctgt gcacgactgt aataccacct 900

gtcagacacc taagggcgcc attaatacca gctgcccctt ccagaatatt caccctatca 960
ccatcggcaa gtgtcctaag tatgtgaaga gcaccaagct gagactggct accggtctga 1020
gaaatagccc tgaaacctg tattttcaag gcctgtttgg agccatcgcc ggctttattg 1080
agggaggatg gaccggaatg gtggatggct ggtacggcta tcaccaccag aatgagcagg 1140
gatccggata tgccgccgat ctgaagtcta cacagaacgc catcgacgag atcacaaca 1200
aggtgaacag cgtgatcgag aagatgaaca cccagtttac agctgtgggc aaggagtcca 1260
accacctgga gaagagaatc gagaacctga acaagaaagt ggacgacggc ttcctggata 1320
tttgaccta caatgccgag ctgctcgtgc tcctggagaa tgagagaacc ctggactacc 1380

acgacagcaa tgtgaagaac ctgtacgaga aggtgagaag ccagctgaag aacaatgcca 1440
aggagatcgg caacggctgc tttgagttct accacaagtg tgacaacacc tgtatggagt 1500
ctgtgaagaa cggcacctac gactacccta agtatagcga ggaggccaag ctgaatagag 1560
aggagatcga cggcgtgaaa ctggaaagca caagaatcta tcagggcgct gaaaacctgt 1620
atcttcaggg cggttctggt tacatcccgg aagctccgcg tgacggtcag gcttacgttc 1680
gtaaagacgg tgaatgggtt ctgctgtcta ctttctggg tcaccatcat caccaccatc 1740
accatcatca ctgataaaag ctt 1763

<210

> 31

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 31

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60
 Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
 65 70 75 80
 Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
 85 90 95
 Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
 100 105 110
 Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

 115 120 125
 Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
 130 135 140
 Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
 145 150 155 160
 Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Glu Gly Met Ile
 165 170 175
 Asp Tyr Glu Gly Thr Gly Gln Ala Ala Tyr Pro Lys Leu Ser Lys Ser

 180 185 190
 Tyr Ile Asn Asp Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His
 195 200 205
 His Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp
 210 215 220
 Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro
 225 230 235 240
 Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn

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

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

305 310 315 320

Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys

325 330 335

Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu Asn

340 345 350

Leu Tyr Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly

355 360 365

Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn

370 375 380

Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala

385 390 395 400

Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn

405 410 415

Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg

420 425 430

Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp

435 440 445

Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu

450 455 460

Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser

465 470 475 480

Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe

485 490 495

Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr

500 505 510

Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu

515 520 525

Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu

530 535 540

Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg

Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser

Thr Phe Leu Gly His His His His His His His His His His

580 585 590

<210> 32

<211> 1785

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Synthetic

<400> 32

ccatggtaag cgctatigtt ttatatgtgc ttttggcggc ggcggcgcgt tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccttc tgcacctggg caaatgtaat attgccgget 240

ggattctctggg aaatccigaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcacg gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctgggtgga aggcatgatt gattacgaag 540
gcacaggcca ggcagcctac cctaagctgt ctaagagcta catcaacgac aaggggcaag 600
aagtgtctggt gctgtgggga atccaccacc ctagcacaag cgccgatcag cagagcctgt 660
accagaatgc cgatgcctat gtgtttgtgg gcagcagcag atacagcaaa aagttcaagc 720

ctgaaattgc cattagaccg aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaaatctg gtggtgccta 840
gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gatacccctg 900
tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgcctt 960
tccagaatat tcacctatc accatcggca agtgtcctaa gtatgtgaag agcaccaagc 1020
tgagactggc taccggtctg agaaatagcc ctgaaaacct gtattttcaa ggccctgtttg 1080
gagccatcgc cggttttatt gagggaggat ggaccggaat ggtggatggc tggtagcgct 1140

atcaccacca gaatgagcag ggatccggat atgccgccga tctgaagtct acacagaacg 1200

ccatcgacga gatcacaac aaggtgaaca gcgtgatcga gaagatgaac acccagtta 1260

cagctgtggg caaggagttc aaccacctgg agaagagaat cgagaacctg aacaagaaag 1320

tggacgacgg cticctggat atttggacct acaatgccga gctgctcgtg ctcttgaga 1380

atgagagaac cctggactac cagcagcga atgtgaagaa cctgtacgag aaggtgagaa 1440

gccagctgaa gaacaatgcc aaggagatcg gcaacggctg ctttgagttc taccacaagt 1500

gtgacaacac ctgtatggag tctgtgaaga acggcaccta cgactaccct aagtatagcg 1560

aggaggccaa gctgaataga gaggagatcg acggcgtgaa actggaaaagc acaagaatct 1620

atcagggcgc tgaaaacctg tattttcagg gcggttctgg ttacatcccg gaagctccgc 1680

gtgacgggtca ggcttacgtt cgtaaagacg gtgaatgggt tctgctgtct accttcctgg 1740

gtcaccatca tcaccacat caccatcatc actgataaaa agctt 1785

<210> 33

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 33

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100	105	110	
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val			
115	120	125	
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Glu Gly Met Ile Asp Tyr			
130	135	140	
Glu Gly Thr Gly Gln Ala Ala Trp Pro Asn His Asp Ser Asn Lys Gly			
145	150	155	160
Val Thr Ala Ala Cys Pro His Ala Gly Ala Lys Ser Phe Tyr Lys Asn			
165	170	175	
Leu Ile Trp Leu Val Lys Lys Gly Asn Ser Tyr Pro Lys Leu Ser Lys			
180	185	190	
Ser Tyr Ile Asn Asp Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile			
195	200	205	
His His Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala			
210	215	220	
Asp Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys			
225	230	235	240
Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met			
245	250	255	
Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu			
260	265	270	
Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg			
275	280	285	
Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys			
290	295	300	
Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro			
305	310	315	320
Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val			
325	330	335	
Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu			
340	345	350	

Asn Leu Tyr Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
 355 360 365
 Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln
 370 375 380
 Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn
 385 390 395 400
 Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met
 405 410 415

 Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys
 420 425 430
 Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile
 435 440 445
 Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr
 450 455 460
 Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg
 465 470 475 480

 Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu
 485 490 495
 Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly
 500 505 510
 Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu
 515 520 525
 Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala
 530 535 540

 Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro
 545 550 555 560
 Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
 565 570 575
 Ser Thr Phe Leu Gly His His His His His His His His His
 580 585 590

<210> 34

<211> 1787

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 34

ccatggttaag cgctattgtt ttataatgtgc ttttggcggc ggcggcgcat tctgcctttg 60

cggatacact gtgtatttggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctgaag 420

gcatgattga ttacgaaggc acaggccagg cagcctggcc taatcacgat tctaataagg 480

gagtgcagc cgctgtcct catgccggag ccaagtctt ttacaagaac ctgatctggc 540

tggtgaagaa gggcaacagc taccctaagc tgtctaagag ctacatcaac gacaaggga 600

aagaagtgtt ggtgctgtgg ggaatccacc accctagcac aagcgccgat cagcagagcc 660

tgtaccagaa tgccgatgcc tatgtgtttg tgggcagcag cagatacagc aaaaagttca 720

agcctgaaat tgccattaga cccaaagtga gagatcagga aggcagaatg aattactact 780

ggaccctggt ggaacctggc gataagatca catcttaggc caccggaaat ctggtggtgc 840

ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900

ctgtgcagca ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960

ccttcagaa tttcacctc atcacatcg gcaagtgtcc taagtatgtg aagagcacca 1020

agctgagact ggctaccggt ctgagaaata gcctgaaaa cctgtatctt caaggcctgt 1080

ttggagccat cgccggtctt attgaggag gatggaccgg aatggtggat ggctggtacg 1140

gctatcacca ccagaatgag cagggatecg gatatgccgc cgatctgaag tctacacaga 1200

acgccatcga cgagatcaca aacaaggtga acagcgtgat cgagaagatg aacacccagt 1260

ttacagctgt gggcaaggag ttcaaccacc tggagaagag aatcgagaac ctgaacaaga 1320

aagtggacga cggcttctct gatatttga cctacaatgc cgagctgctc gtgctcctgg 1380

agaatgagag aaccctggac taccacgaca gcaatgtgaa gaacctgtac gagaaggatga 1440

gaagccagct gaagaacaat gccaaaggaga tcggcaacgg ctgctttgag ttctaccaca 1500

agtgtgacaa cacctgtatg gagtctgtga agaacggcac ctacgactac cctaagtata 1560

gcgaggaggc caagctgaat agagaggaga tcgacggcgt gaaactggaa agcacaagaa 1620
tctatcaggg cgctgaaaac ctgtattttc agggcggttc tggttacatc ccggaagctc 1680
cgcgtagcgg tcaggcttac gttcgtaaag acggtgaatg ggttctgctg tctaccttcc 1740
tgggtcacca tcataccacc catcaccatc atcactgata aaagctt 1787

<210> 35

<211> 586

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 35

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1	5	10	15
Ser	Ala	Phe	Ala
Asp	Thr	Leu	Cys
Ile	Gly	Tyr	His
Ala	Asn	Asn	Ser
20	25	30	
Thr	Asp	Thr	Val
Asp	Thr	Val	Leu
Glu	Lys	Asn	Val
Thr	Val	Thr	His
35	40	45	
Ser	Val	Asn	Leu
Leu	Glu	Asp	Lys
His	Asn	Gly	Lys
Leu	Cys	Lys	Leu
50	55	60	
Arg	Gly	Val	Ala
Pro	Leu	His	Leu
Gly	Lys	Cys	Asn
Ile	Ala	Gly	Trp

65	70	75	80
Ile	Leu	Gly	Asn
Pro	Glu	Cys	Glu
Ser	Leu	Ser	Thr
Ala	Ser	Ser	Trp
85	90	95	
Ser	Tyr	Ile	Val
Glu	Thr	Pro	Ser
Ser	Asp	Asn	Gly
Thr	Cys	Tyr	Pro
100	105	110	
Gly	Asp	Phe	Ile
Asp	Tyr	Glu	Glu
Leu	Arg	Glu	Gln
Leu	Ser	Ser	Val
115	120	125	
Ser	Ser	Phe	Glu
Arg	Phe	Glu	Ile
Phe	Pro	Lys	Thr
Ser	Ser	Trp	Pro

130	135	140
Asn	His	Asp
Ser	Asn	Lys
Gly	Val	Thr
Ala	Ala	Cys
Pro	His	Ala
Gly		
145	150	155
Ala	Lys	Ser
Phe	Tyr	Lys
Asn	Leu	Ile
Trp	Leu	Val
Lys	Lys	Gly
Asn		

165 170 175
 Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
 180 185 190
 Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

 195 200 205
 Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
 210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Gly Ile Phe Gly
 225 230 235 240
 Ala Ile Ala Gly Phe Ile Glu Gly Gly Arg Met Asn Tyr Tyr Trp Thr
 245 250 255
 Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu

 260 265 270
 Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly
 275 280 285
 Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln
 290 295 300
 Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His
 305 310 315 320
 Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu

 325 330 335
 Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu Asn Leu Tyr Phe Gln
 340 345 350
 Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly
 355 360 365
 Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser
 370 375 380
 Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile

 385 390 395 400
 Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr
 405 410 415

Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu
420 425 430
Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala
435 440 445
Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp
450 455 460
Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn
465 470 475 480
Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys
485 490 495
Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro
500 505 510
Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val
515 520 525
Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe
530 535 540
Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala
545 550 555 560
Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly
565 570 575
His His His His His His His His His His
580 585

<210> 36

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 36

Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly

1 5 10

<210> 37

<211> 1772

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 37

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgccccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaccccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacac aacgacaagg gcaaagaagt gctgggtgctg tggggaatcc 600
accacccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660

ttgtgggcag cagcagatac agcaaaaagt tcaagcctga aattgccatt ggcattttcg 720
gcgctatcgc cggttcatt gagggaggca gaatgaatta ctactggacc ctggtggaac 780
ctggcgataa gatcacattt gagggcaccg gaaatctggt ggtgcctaga tatgcatttg 840
ctatggagag aaatgctggc tctggcatca ttatctctga taccctgtg cagcactgta 900
ataccacctg tcagacacct aaggcgcca ttaataccag cctgcccttc cagaatattc 960
accctatcac catcggaag tgtcctaagt atgtgaagag caccaagctg agactggcta 1020
ccggtctgag aaatagccct gaaaacctgt attttcaagg cctgtttgga gccatcgccg 1080
gctttattga gggaggatgg accggaatgg tggatggctg gtacggctat caccaccaga 1140

atgagcaggg atccggatat gccgccgatc tgaagtctac acagaacgcc atcgacgaga 1200
tcacaacaa ggtgaacagc gtgatcgaga agatgaacac ccagtttaca gctgtgggca 1260
aggagttaa ccacctggag aagagaatcg agaacctgaa caagaaagtg gacgacggct 1320
tcctggatat ttggacctac aatgccgagc tgctcgtgct cctggagaat gagagaaccc 1380
tggactacca cgacagcaat gtgaagaacc tgtacagaa ggtgagaagc cagctgaaga 1440
acaatgcaa ggagatcggc aacggctgct ttgagttcta ccacaagtg gacaacacct 1500
glatggagtc tgtgaagaac ggcacctacg actaccctaa gtatagcgag gaggccaagc 1560

tgaatagaga ggagatcgac ggcgtgaaac tggaaagcac aagaatctat cagggcgctg 1620

aaaacctgta ttttcagggc ggttctgggtt acatcccgga agctccgcgt gacggtcagg 1680

cttacgttcg taaagacggt gaatgggttc tgcgtgtctac cttcctgggt caccatcatc 1740

accacatca ccatcatcac tgataaaagc tt 1772

<210> 38

<211> 588

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 38

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

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

145 150 155 160

Gly Ala Ile Ala Gly Phe Ile Glu Gly Phe Tyr Lys Asn Leu Ile Trp
 165 170 175
 Leu Val Lys Lys Gly Asn Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile
 180 185 190
 Asn Asp Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro
 195 200 205
 Ser Thr Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr
 210 215 220
 Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile
 225 230 235 240
 Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr
 245 250 255
 Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly
 260 265 270
 Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly
 275 280 285
 Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr
 290 295 300
 Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn
 305 310 315 320
 Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr
 325 330 335
 Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu Asn Leu Tyr
 340 345 350
 Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp
 355 360 365
 Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln
 370 375 380
 Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp
 385 390 395 400
 Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln

405 410 415
Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu
420 425 430
Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr
435 440 445
Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr
450 455 460
His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu

465 470 475 480
Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His
485 490 495
Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp
500 505 510
Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp
515 520 525
Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu

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

580 585

<210> 39

<211> 1778

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 39

ccatggttaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcgat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
 agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
 ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
 tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
 agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

 caagcagctg gcctaatac gattetaata agggagtgc agccgcctgt cctggcattt 480
 tcggcgctat cgccggcttc attgagggat ttacaagaa cctgatctgg ctgggaaga 540
 agggcaacag ctaccctaag ctgtctaaga gctacatcaa cgacaagggc aaagaagtgc 600
 tgggtgctgtg gggaatccac caccctagca caagcgccga tcagcagagc ctgtaccaga 660
 atgccgatgc ctatgtgttt gtgggcagca gcagatacag caaaaagttc aagcctgaaa 720
 ttgccattag acccaaagtg agagatcagg aaggcagaat gaattactac tggaccctgg 780
 tggaaacctgg cgataagatc acatttgagg ccaccggaaa tctggtggtg cctagatatg 840
 catttgctat ggagagaaat gctggtctg gcatcattat ctctgatacc cctgtgcacg 900

 actgtaatac cacctgtcag acacctaagg gcgccattaa taccagcctg cccttcaga 960
 atattcacc tatcaccatc ggcaagtgtc ctaagtatgt gaagagcacc aagctgagac 1020
 tggctaccgg tctgagaaat agccctgaaa acctgtattt tcaaggcctg tttggagcca 1080
 tcgccggctt tattgagga ggatggaccg gaatggtgga tggctggtac ggctatcacc 1140
 accagaatga gcaggatcc ggatatgccg ccgatctgaa gtctacacag aacgccatcg 1200
 acgagatcac aaacaagtg aacagcgtga tcgagaagat gaacaccag ttacagctg 1260
 tgggcaagga gttaaccac ctggagaaga gaatcgagaa cctgaacaag aaagtggacg 1320
 acggcttctt ggatatttgg acctacaatg ccgagctgct cgtgctcctg gagaatgaga 1380

 gaacctgga ctaccagac agcaatgtga agaacctgta cgagaaggtg agaagccagc 1440
 tgaagaacaa tgccaaggag atcggaacg gctgctttga gttctaccac aagtgtgaca 1500
 acacctgtat ggagtctgtg aagaacggca cctacgacta ccctaagtat agcaggagg 1560
 ccaagctgaa tagagaggag atcgacggcg tgaaactgga aagcacaaga atctatcagg 1620
 gcgctgaaaa cctgtatttt caggcggtt ctggttcat cccggaagct ccgctgacg 1680
 gtcaggctta cgttcgtaaa gacgtgta ggttctgct gtctacctc ctgggtcacc 1740
 atcatcacca ccatcaccat catcactgat aaaagctt 1778

<210

> 40

<211> 589

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 40

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Gly Ile Phe Gly

165 170 175

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

180 185 190

Ile Asn Asp Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His

195 200 205

Pro Ser Thr Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Ala

210 215 220

Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu
 225 230 235 240
 Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr

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

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

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

 435 440 445
 Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp
 450 455 460
 Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln

465 470 475 480
 Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr
 485 490 495
 His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr

 500 505 510
 Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile
 515 520 525
 Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn
 530 535 540
 Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp
 545 550 555 560
 Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr

 565 570 575
 Phe Leu Gly His His His His His His His His His His His

 580 585

<210> 41

<211> 1781

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 41

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
 cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
 tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
 agctgtgtaa gctgagagga gttgcccttc tgcacctggg caaatgtaat attgccggct 240

 ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
 tggaaccacc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
 agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420
 caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
 gagccaagtc cttttacaag aacctgatct ggctgggtgg cattttcggc gctatcgccg 540
 gcttcattga gggataccct aagctgtcta agagctacat caacgacaag ggcaaagaag 600

tgctggtgct gtggggaatc caccacccta gcacaagcgc cgatcagcag agcctgtacc 660
 agaatgccga tgcctatgtg ttgtgggca gcagcagata cagcaaaaag ttcaagcctg 720

aaattgccat tagaccecaa gtgagagatc aggaaggcag aatgaattac tactggacce 780
 tgggtggaacc tggcgataag atcacatttg aggccaccgg aaatctggtg gtgcctagat 840
 atgcatttgc tatggagaga aatgctggct ctggcatcat tatctctgat acccctgtgc 900
 acgactgtaa taccacctgt cagacaccta agggcgccat taataccagc ctgcccttcc 960
 agaatattca ccctatcacc atcggaagc gtccctaagta tgtgaagagc accaagctga 1020
 gactggctac cggctctgaga aatagccctg aaaacctgta ttttcaaggc ctgtttggag 1080
 ccatcgccgg ctttattgag ggaggatgga ccggaatggt ggatggctgg tacggctatc 1140
 accaccagaa tgagcaggga tccggatatg ccgccgatct gaagtctaca cagaacgcca 1200

tcgacgagat cacaacaag gtgaacagcg tgatcgagaa gatgaacacc cagtttacag 1260
 ctgtgggcaa ggagttcaac cacctggaga agagaatcga gaacctgaac aagaaagtgg 1320
 acgacggctt cctggatatt tggacctaca atgccgagct gctcgtgctc ctggagaatg 1380
 agagaacctt ggactaccac gacagcaatg tgaagaacct gtacgagaag gtgagaagcc 1440
 agctgaagaa caatgccaag gagatcggca acggctgctt tgagttctac cacaagtgtg 1500
 acaacacctg tatggagtct gtgaagaacg gcacctacga ctaccctaag tatagcgagg 1560
 aggccaagct gaatagagag gagatcgacg gctgaaaact ggaaagcaca agaattctatc 1620
 agggcgctga aaacctglat tttcagggcg gttctggtta catcccgaa gctccgcgtg 1680

acggtcaggc ttacgttcgt aaagacggtg aatgggttct gctgtctacc ttcctgggtc 1740
 accatcatca ccaccatcac catcatcact gataaaagct t 1781

<210> 42

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 42

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His
 35 40 45
 Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu
 50 55 60
 Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp
 65 70 75 80
 Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
 85 90 95

 Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
 100 105 110
 Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
 115 120 125
 Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Gly Ile Phe Gly Ala Ile
 130 135 140
 Ala Gly Phe Ile Glu Gly Trp Pro Asn His Asp Ser Asn Lys Gly Val
 145 150 155 160

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

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

275 280 285
 Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn
 290 295 300
 Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe
 305 310 315 320
 Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys
 325 330 335
 Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu Asn
 340 345 350

 Leu Tyr Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly
 355 360 365
 Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn
 370 375 380
 Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala
 385 390 395 400
 Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn
 405 410 415

 Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg
 420 425 430
 Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp
 435 440 445
 Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu
 450 455 460
 Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser
 465 470 475 480

 Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe
 485 490 495
 Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr
 500 505 510
 Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu
 515 520 525

Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu
530 535 540

Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg
545 550 555 560
Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser
565 570 575
Thr Phe Leu Gly His His His His His His His His His His
580 585 590

<210> 43

<211> 1784

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 43

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctggca 420
ttttcggcgc tatcgccggc ttcatgtagg gatggcctaa tcacgattct aataaggag 480
tgacagccgc ctgtcctcat gccggagcca agtcctttta caagaacctg atctggcttg 540

tgaagaaggg caacagctac cctaagctgt ctaagagcta catcaacgac aagggcaaag 600
aagtgttgtt gctgtgggga atccaccacc ctagcacaag cgccgatcag cagagcctgt 660
accagaatgc cgatgcctat gtgtttgtgg gcagcagcag atacagcaaa aagttcaagc 720
ctgaaattgc cattagacc aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaatctg gtggcgctta 840
gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gataccctg 900
tgcacgactg taataccacc tgtcagacac ctaaggcgc cattaatacc agcctgccct 960

tccagaatat tcaccciatc accatcggca agtgtcctaa gtatgtgaag agcaccaagc 1020

tgagactggc taccggtctg agaaatagcc ctgaaaacct gtatittcaa ggctgtttg 1080

gagccatcgc cggttttatt gagggaggat ggaccggaat ggtggatggc tggtagcgct 1140

atcaccacca gaatgagcag ggatccggat atgccgccga tctgaagtct acacagaacg 1200

ccatcgacga gatcacaac aaggtgaaca gctgatcga gaagatgaac acccagttta 1260

cagctgtggg caaggagtgc aaccacctgg agaagagaat cgagaacctg aacaagaaag 1320

tggacgacgg ctctctggat atttggacct acaatgccga gctgctctg ctctctggaga 1380

atgagagaac cctggactac cagcagcga atgtgaagaa cctgtacgag aaggtgagaa 1440

gccagctgaa gaacaatgcc aaggagatcg gcaacggctg ctttgagtgc taccacaagt 1500

gtgacaacac ctgtatggag tctgtgaaga acggcaccta cgactaccct aagtatagcg 1560

aggaggccaa gctgaataga gaggagatcg acggcgtgaa actggaaagc acaagaatct 1620

atcagggcgc tgaaaacctg tatittcagg gcggttctgg ttacatcccg gaagctccgc 1680

gtgacggcca ggcttacgtt cgtaaagacg gtgaatgggt tctgctgtct accttcctgg 1740

gtcaccatca tcaccacat caccatcatc actgataaaa gctt 1784

<210> 44

<211> 584

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 44

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15
Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30
Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45
Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60
Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp
 85 90 95
 Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro
 100 105 110
 Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
 115 120 125
 Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro
 130 135 140
 Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
 145 150 155 160
 Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
 165 170 175
 Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
 180 185 190
 Val Leu Val Leu Trp Gly Ile His His Pro Ser Gly Ile Phe Gly Ala
 195 200 205
 Ile Ala Gly Phe Ile Glu Gly Asp Ala Tyr Val Phe Val Gly Ser Ser
 210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile Arg Pro Lys Val
 225 230 235 240
 Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro
 245 250 255
 Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg
 260 265 270
 Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly Ile Ile Ile Ser
 275 280 285
 Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln Thr Pro Lys Gly
 290 295 300
 Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His Pro Ile Thr Ile
 305 310 315 320
 Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Leu Ala Thr

325 330 335
 Gly Leu Arg Asn Ser Pro Gln Arg Glu Arg Arg Lys Lys Arg Gly Leu
 340 345 350
 Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
 355 360 365
 Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
 370 375 380
 Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn

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

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

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

580

<210> 45

<211> 1766

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 45

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttataatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacac aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
accaccctag cggcatttcc ggcgctatcg ccggcttcat tgaggagat gcctatgtgt 660
ttgtgggcag cagcagatag agcaaaaagt tcaagcctga aattgccatt agacccaaag 720
tgagagatca ggaaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccgga aatctggtgg tgcctagata tgcatttgc atggagagaa 840

atgctggctc tggcatcatt atctctgata ccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgccttcca gaatttcac cctatcacca 960
tcggcaagtg tctaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atagccctca gagggagaga cgcaagaaga gaggcctgtt tggagccatc gccggcttta 1080
ttgaggagg atggaccgga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140
agggatccgg atatgccgcc gatctgaagt ctacacagaa cgccatcgac gagatcaca 1200
acaagtgaa cagcgtgac gagaagatga acaccagtt tacagctgtg ggcaaggagt 1260
tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320

atatttggac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380

accacgacag caatgtgaag aacctgtacg agaaggtgag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcgggtct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680
ttcgtaaaga cggatgaatg gttctgctgt ctaccttctt gggtcaccat catcaccacc 1740
atcaccatca tctactgataa aagctt 1766

<210> 46

<211> 584

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 46

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly
 145 150 155 160
 Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn
 165 170 175
 Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

 180 185 190
 Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln
 195 200 205
 Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
 210 215 220
 Arg Tyr Ser Lys Lys Phe Lys Pro Gly Ile Phe Gly Ala Ile Ala Gly
 225 230 235 240
 Phe Ile Glu Gly Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro

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

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

 370 375 380
 Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr Asn

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

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

 500 505 510
 Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys Leu
 515 520 525
 Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln Gly
 530 535 540
 Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val
 545 550 555 560
 Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His His

 565 570 575
 His His His His His His His His
 580

<210> 47

<211> 1766

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 47

ccatggttaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtatttggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgccccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tgaaacccc tagcagcgac aatggcacct gttacctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacac aacgacaagg gcaaagaagt gctgggtgctg tggggaatcc 600
accacctag cacaagcgc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatc agcaaaaagt tcaagcctgg cattttcggc gctatcgccg 720

gcttcattga gggaggcaga atgaattact actggaccct ggtggaacct ggcgataaga 780
tcacatttga ggccaccga aatctggtgg tgcctagata tgcatittgt atggagagaa 840
atgctggctc tggcatcatt atctctgata ccctgtgca cgactgtaat accacctgtc 900
agacacctaa gggcgccatt aataccagcc tgccttcca gaatattcac cctatcacca 960
tcggcaagtg tectaagtat gtgaagagca ccaagctgag actggctacc ggtctgagaa 1020
atagccctca gaggagaga cgcaagaaga gaggcctgtt tggagccatc gccggcttta 1080
ttgaggagg atggaccga atggtggatg gctggtacgg ctatcaccac cagaatgagc 1140
agggatccgg atatgccgc gatctgaagt ctacacagaa cgccatcgac gagatcaca 1200

acaagtgaa cagcgtgac gagaagatga acaccagtt tacagctgtg ggcaaggagt 1260
tcaaccacct ggagaagaga atcgagaacc tgaacaagaa agtggacgac ggcttcctgg 1320
atatttggac ctacaatgcc gagctgctcg tgctcctgga gaatgagaga accctggact 1380
accacgacag caatgtgaag aacctgtacg agaaggtgag aagccagctg aagaacaatg 1440
ccaaggagat cggcaacggc tgctttgagt tctaccacaa gtgtgacaac acctgtatgg 1500
agtctgtgaa gaacggcacc tacgactacc ctaagtatag cgaggaggcc aagctgaata 1560
gagaggagat cgacggcgtg aaactggaaa gcacaagaat ctatcagggc gctgaaaacc 1620
tgtattttca gggcggttct ggttacatcc cggaagctcc gcgtgacggt caggcttacg 1680

ttcgtaaaga cggtgaatgg gttctgctgt ctaccttct gggtcaccat catcaccacc 1740
atcaccatca tcaactgataa aagctt 1766

<210> 48

<211> 592

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 48

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Arg Lys Lys Arg

165 170 175

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

180 185 190

Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Lys Glu

195 200 205

Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Asp
210 215 220

Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro
225 230 235 240

Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn
245 250 255

Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala
260 265 270

Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn
275 280 285

Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn
290 295 300

Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe
305 310 315 320

Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys
325 330 335

Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg
340 345 350

Glu Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
355 360 365

Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His
370 375 380

Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln
385 390 395 400

Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys
405 410 415

Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu
420 425 430

Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp
435 440 445

Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg

450 455 460
 Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
 465 470 475 480

Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
 485 490 495
 Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn
 500 505 510
 Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg
 515 520 525
 Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly
 530 535 540

Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala
 545 550 555 560
 Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu
 565 570 575
 Leu Ser Thr Phe Leu Gly His His His His His His His His His
 580 585 590

<210> 49

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 49

Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu

1 5 10 15

<210> 50

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 50

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

1 5 10 15

Ser

<210> 51

<211> 1790

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 51

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120

tggagaagaa tggaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcacg gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctgggtgag aaagaagaga ggcctgtttg 540

gagccatcgc cgcttttatt gagtacatca acgacaaggg caaagaagtg ctggtgctgt 600

ggggaatcca ccaccctagc aaggagtcta cacagaaggc cattgatggc gttacaaata 660

aggtcaattc tgatgcctat gtgtttgtgg gcagcagcag atacagcaa aagttcaagc 720

ctgaaattgc cattagacc aaagtgagag atcaggaagg cagaatgaat tactactgga 780

ccctgggtga acctggcgat aagatcacat ttgaggccac cggaaatctg gtgggtgccta 840

gatatgcatt tgctatggag agaaatgtg gctctggcat cattatctct gataccctg 900

tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgccct 960

tccagaatat tcacctatc accatcgga agtgcctaa gtatgtgaag agcaccaagc 1020

tgagactggc taccggtctg agaaatagcc ctgagaggga gagacgcaag aagagaggcc 1080

tgtttggagc catcgccggc tttattgagg gaggatggac cggaatggtg gatggctggt 1140

acggctatca ccaccagaat gagcagggat ccggatatgc cgccgatctg aagtctacac 1200

agaacgcat cgacgagatc acaaacaagg tgaacagcgt gatcgagaag atgaacaccc 1260
 agtttacagc tgtgggcaag gagttcaacc acctggagaa gagaatcgag aacctgaaca 1320
 agaaagtgga cgacggcttc ctggatattt ggacctacaa tgccgagctg ctcgtgctcc 1380
 tggagaatga gagaacctg gactaccacg acagcaatgt gaagaacctg tacgagaagg 1440
 tgagaagcca gctgaagaac aatgccaagg agatcggcaa cggtgcttt gagttctacc 1500
 acaagtgtga caacacctgt atggagtctg tgaagaacgg cacctacgac taccctaagt 1560

atagcgagga ggccaagctg aatagagagg agatcgacgg cgtgaaactg gaaagcaca 1620
 gaatctatca gggcgctgaa aacctgtatt ttcaggcggtt ttctggttac atcccgaag 1680
 ctccgctga cggtcaggct tacgttcgta aagacggtga atgggttctg ctgtctacct 1740
 tcctgggtca ccatcatcac caccatcacc atcatcactg ataaaagctt 1790

<210> 52

<211> 589

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 52

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115	120	125	
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro			
130	135	140	
Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly			
145	150	155	160
Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Arg Lys Lys Arg			
	165	170	175
Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Tyr Ile Asn Asp Lys			
	180	185	190
Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Lys Glu			
195	200	205	
Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Asp			
210	215	220	
Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro			
225	230	235	240
Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn			
	245	250	255
Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala			
260	265	270	
Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn			
275	280	285	
Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn			
290	295	300	
Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe			
305	310	315	320
Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys			
	325	330	335
Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ile Pro Ser Ile			
340	345	350	
Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly			
355	360	365	

Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu
370 375 380

Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile
385 390 395 400

Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr
405 410 415

Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile
420 425 430

Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr
435 440 445

Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp
450 455 460

Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln
465 470 475 480

Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr
485 490 495

His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr
500 505 510

Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile
515 520 525

Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn
530 535 540

Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp
545 550 555 560

Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr
565 570 575

Phe Leu Gly His His His His His His His His His
580 585

<210> 53

<211> 1781

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 53

```

ccatggaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgag aaagaagaga ggcctgtttg 540
gagccatcgc cggttttatt gagtacatca acgacaaggg caaagaatg ctggtgctgt 600
ggggaatcca ccacctagc aaggagtcta cacagaaggc cattgatggc gttacaaata 660
aggtcaattc tgatgcctat gtgtttgtgg gcagcagcag atacagcaa aagttcaagc 720
ctgaaattgc cattagacc aaagtgagag atcaggaagg cagaatgaat tactactgga 780
ccctggtgga acctggcgat aagatcacat ttgaggccac cggaatctg gtggtgccta 840

gatatgcatt tgctatggag agaaatgctg gctctggcat cattatctct gataccctg 900
tgcacgactg taataccacc tgtcagacac ctaagggcgc cattaatacc agcctgccct 960
tccagaatat tcacctatc accatcgga agtgcctaa gtatgtgaag agcaccaagc 1020
tgagactggc taccggtctg agaaatatcc ctagcatcca gagcagaggc ctgtttggag 1080
ccatcgccgg ctttattgag ggaggatgga ccggaatggt ggatggctgg tacggctatc 1140
accaccagaa tgagcagga tccgatatg cgcctgatct gaagtctaca cagaacgcca 1200
tcgacgagat cacaacaag gtgaacagcg tgatcgagaa gatgaacacc cagtttacag 1260
ctgtgggcaa ggagttcaac cacctggaga agagaatcga gaacctgaac aagaaagtgg 1320

acgacggctt cctggatatt tggacctaca atgccgagct gctcgtgctc ctggagaatg 1380
agagaacctt ggactaccac gacagcaatg tgaagaacct gtacgagaag gtgagaagcc 1440
agctgaagaa caatgccaag gagatcgga acggctgctt tgagttctac cacaagtgtg 1500
acaacacctg tatggagtct gtgaagaacg gcacctacga ctaccctaag tatagcgagg 1560
aggccaagct gaatagagag gagatcgacg gcgtgaaact ggaaagcaca agaattctatc 1620
agggcgctga aaacctglat tttcaggcg gttctggtta catcccgaa gctccgcgtg 1680

```

acggtcaggc ttacgttcgt aaagacgggtg aatgggttct gctgtctacc ttctgggtc 1740
accatcatca ccaccatcac catcatcact gataaaagct t 1781

<210> 54

<211> 585

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 54

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Arg Lys Lys Arg

165 170 175

Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Tyr Ile Asn Asp

180 185 190
Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr

195 200 205
Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe

210 215 220
Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile

225 230 235 240
Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr

245 250 255
Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu

260 265 270
Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly

275 280 285
Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln

290 295 300
Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His

305 310 315 320
Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu

325 330 335
Arg Leu Ala Thr Gly Leu Arg Asn Ile Glu Asn Leu Tyr Phe Gln Gly

340 345 350
Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met

355 360 365
Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly

370 375 380
Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr

385 390 395 400
Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala

405 410 415
Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn

420 425 430
Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu

435 440 445
Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser

450 455 460
Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn

465 470 475 480
Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp

485 490 495
Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys

500 505 510
Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys

515 520 525
Leu Glu Ser Thr Arg Ile Tyr Gln Gly Ala Glu Asn Leu Tyr Phe Gln

530 535 540
Gly Gly Ser Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr

545 550 555 560
Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu Gly His

565 570 575
His His His His His His His His His

580 585

<210> 55

<211> 1769

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 55

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcac tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttacctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatac gattctaata agggagtac agccgctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctggtag aaagaagaga ggcctgttcg 540

gcgctatcg cggttcatt gagggataca tcaacgaca gggcaaagaa gtgctggtag 600

tgtggggaat ccaccacct agcacaagcg ccgacgagc gagcctgtac cagaatgccg 660

atgcctatgt gtttgtgggc agcagcagat acagcaaaaa gttcaagcct gaaattgcc 720

ttagacccaa agtgagagat caggaaggca gaatgaatta ctactggacc ctggtggaac 780

ctggcgataa gatcacattt gagggcacccg gaaatctggt ggtgcctaga tatgcatttg 840

ctatggagag aaatgctggc tctggcatca ttatctctga taccctgtg cagcactgta 900

ataccacctg tcagacacct aaggcgcca ttaataccag cctgcccttc cagaatattc 960

acctatcac catcggcaag tgtcctaagt atgtgaagag caccaagctg agactggcta 1020

ccggtctgag aaatatcgaa aacctgtatt ttcaaggcct gtttgagacc atcgccggct 1080

ttattgaggg aggatggacc ggaatggtgg atggctgga cggctatcac caccagaatg 1140

agcagggatc cggatatgcc gccgatctga agtctacaca gaacgccatc gacgagatca 1200

caaacaaggt gaacagcgtg atcgagaaga tgaacacca gtttacagct gtgggcaagg 1260

agttcaacca cctggagaag agaatcgaga acctgaacaa gaaagtggac gacggcttc 1320

tggatatttg gacctacaat gccgagctgc tctgtctcct ggagaatgag agaaccctgg 1380

actaccagc cagcaatgtg aagaacctgt acgagaaggt gagaagccag ctgaagaaca 1440

atgccaagga gatcggcaac ggctgctttg agttctacca caagtgtgac aacacctgta 1500

tggagtctgt gaagaacggc acctacgact accctaagta tagcgaggag gccaaagtga 1560

atagagagga gatcgacggc gtgaaactgg aaagcacaag aatctatcag ggcgctgaaa 1620

acctgtatct tcagggcggt tctggttaca tcccgaagc tccgctgac ggtcaggctt 1680

acgttcgtaa agacggtgaa tgggttctgc tgtctacctt cctgggtcac catcatcacc 1740

accatcacca tcatcactga taaaagctt 1769

<210> 56

<211> 593

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 56

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser

210 215 220

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

 Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys
 290 295 300
 Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro
 305 310 315 320
 Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val
 325 330 335
 Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln
 340 345 350

 Arg Glu Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
 355 360 365
 Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 370 375 380
 His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr
 385 390 395 400
 Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu
 405 410 415

 Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu
 420 425 430
 Glu Lys Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
 435 440 445
 Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
 450 455 460
 Arg Thr Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys

465 470 475 480

Val Arg Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys

485 490 495

Phe Glu Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys

500 505 510

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ala Lys Leu Asn

515 520 525

Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Thr Arg Ile Tyr Gln

530 535 540

Gly Ala Glu Asn Leu Tyr Phe Gln Gly Gly Ser Gly Tyr Ile Pro Glu

545 550 555 560

Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val

565 570 575

Leu Leu Ser Thr Phe Leu Gly His His His His His His His His His

580 585 590

His

<210> 57

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 57

Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Cys

1 5 10 15

Lys Cys Ser Asp Ser

20

<210> 58

<211> 1793

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 58

```

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggct 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaacccc tagcagcgac aatggcacct gttaccctgg cgacttcac gattacgagg 360
agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttccctaaga 420
caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480
gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
agctgtctaa gagctacac aacgacaagg gcaaagaagt gctgggtgctg tggggaatcc 600
accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
ttgtgggcag cagcagatac agcaaaaagt tcaagccttc cctgctgacc gaggtggaga 720
ccccaccag gaacggctgg gagtgcgaagt gctccgactc cggcagaatg aattactact 780

ggaccctggt ggaacctggc gataagatca catttgaggc caccggaaat ctggtggtgc 840
ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900
ctgtgcacga ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960
ccttcagaa tattcacctc atcaccatcg gcaagtgtcc taagtatgtg aagagcacca 1020
agctgagact ggctaccggt ctgagaaata gcctcagag ggagagacgc aagaagagag 1080
gcctgtttgg agccatcgcc ggctttattg agggaggatg gaccggaatg gtggatggct 1140
ggtacggcta tcaccaccag aatgagcagg gatccggata tgccgccgat ctgaagtcta 1200
cacagaacgc catcgacgag atcacaaaca aggtgaacag cgtgatcgag aagatgaaca 1260

cccagtttac agctgtgggc aaggagttca accacctgga gaagagaatc gagaacctga 1320
acaagaaagt ggacgacggc ttcttgata ttggaccta caatgccgag ctgctcgtgc 1380
tcctggagaa tgagagaacc ctggactacc acgacagcaa tgtgaagaac ctgtacgaga 1440
aggtgagaag ccagctgaag aacaatgcca aggagatcgg caacggctgc ttgagttct 1500
accacaagtg tgacaacacc tgtatggagt ctgtgaagaa cggcacctac gactacccta 1560
agtatagcga ggaggccaag ctgaatagag aggagatcga cggcgtgaaa ctggaaagca 1620
caagaatcta tcagggcgct gaaaacctgt attttcaggg cggttctggt tacatcccgg 1680

```

aagctccgcg tgacggctcag gcttacgttc gtaaagacgg tgaatgggtt ctgctgtcta 1740

ccttcctggg tcaccatcat caccaccatc accatcatca ctgataaaag ctt 1793

<210> 59

<211> 593

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 59

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu
180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Ser Leu Leu Thr Glu
195 200 205

Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Cys Lys Cys Ser Asp Ser
210 215 220

Asp Ala Tyr Val Phe Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys
225 230 235 240

Pro Glu Ile Ala Ile Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met
245 250 255

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

Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg
275 280 285

Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys
290 295 300

Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro
305 310 315 320

Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val
325 330 335

Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln
340 345 350

Arg Glu Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
355 360 365

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
370 375 380

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr
385 390 395 400

Gln Asn Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu
405 410 415

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu

				420				425				430			
Glu	Lys	Arg	Ile	Glu	Asn	Leu	Asn	Lys	Lys	Val	Asp	Asp	Gly	Phe	Leu
435				440				445							
Asp	Ile	Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Leu	Glu	Asn	Glu
450				455				460							
Arg	Thr	Leu	Asp	Tyr	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Glu	Lys
465				470				475				480			
Val	Arg	Ser	Gln	Leu	Lys	Asn	Asn	Ala	Lys	Glu	Ile	Gly	Asn	Gly	Cys
485				490				495							
Phe	Glu	Phe	Tyr	His	Lys	Cys	Asp	Asn	Thr	Cys	Met	Glu	Ser	Val	Lys
500				505				510							
Asn	Gly	Thr	Tyr	Asp	Tyr	Pro	Lys	Tyr	Ser	Glu	Glu	Ala	Lys	Leu	Asn
515				520				525							
Arg	Glu	Glu	Ile	Asp	Gly	Val	Lys	Leu	Glu	Ser	Thr	Arg	Ile	Tyr	Gln
530				535				540							
Gly	Ala	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gly	Ser	Gly	Tyr	Ile	Pro	Glu
545				550				555				560			
Ala	Pro	Arg	Asp	Gly	Gln	Ala	Tyr	Val	Arg	Lys	Asp	Gly	Glu	Trp	Val
565				570				575							
Leu	Leu	Ser	Thr	Phe	Leu	Gly	His	His	His	His	His	His	His	His	His
580				585				590							
His															

<210> 60

<211> 1793

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Synthetic

 $\langle 400 \rangle$ 60

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcgggcgc tctgcctttg 60

cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtc 120

tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180

agctgtgtaa gctgagagga gttgcccctc tgcacctggg caaatgtaat attgccggt 240

ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300

tggaaacccc tagcagcgac aatggcacct gttacctgg cgacttcac gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatac gattctaata agggagtgc agccgcctgt cctcatgccg 480

gagccaagtc cttttacaag aacctgatct ggctgggtgaa gaagggaac agctacccta 540

agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctgggtgctg tggggaatcc 600

accaccctag ctccctgctg accgaggtgg agacccccac caggaacggc tgggagtgc 660

agtgtccga ctccgatgcc tatgtgtttg tgggcagcag cagatacagc aaaaagtcca 720

agcctgaaat tgccattaga cccaaagtga gagatcagga aggcagaatg aattactact 780

ggacctggt ggaacctggc gataagatca cttttaggc caccggaaat ctggtggtgc 840

ctagatatgc atttctatg gagagaaatg ctggctctgg catcattatc tetgataccc 900

ctgtgcagca ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960

ccttcagaa tattaccct atcaccatcg gcaagtgtcc taagtatgtg aagagacca 1020

agctgagact ggctaccgt ctgagaaata gcctcagag ggagagacgc aagaagagag 1080

gcctgtttgg agccatgcc ggctttattg agggaggatg gaccggaatg gtggatggct 1140

ggtacggcta tcaccaccag aatgagcagg gatccggata tgccgccgat ctgaagtcta 1200

cacagaacgc catcgacgag atcacaaca aggtgaacag cgtgatcgag aagatgaaca 1260

cccagtttac agctgtgggc aaggagtcca accacctgga gaagagaatc gagaacctga 1320

acaagaaagt ggacgacggc ttcttgata tttggaccta caatgccgag ctgctcgtgc 1380

ttctggagaa tgagagaacc ctggactacc acgacagcaa tgtgaagaac ctgtacgaga 1440

aggtgagaag ccagctgaag aacaatgcca aggagatcgg caacggctgc tttgagttct 1500

accacaagtg tgacaacacc tgtatggagt ctgtgaagaa cggcacctac gactacccta 1560

agtatagcga ggaggccaag ctgaatagag aggagatcga cggcgtgaaa ctggaaagca 1620

caagaatcta tcagggcgt gaaaacctgt attttcaggc cggttctggt tacatcccg 1680

aagctccgcg tgacggtcag gcttacgttc gtaaagacgg tgaatgggtt ctgctgtcta 1740

ccttcctggg tcaccatcat caccaccatc accatcatca ctgataaaag ctt 1793

<210> 61

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic

<400> 61

Met Val Ser Ala Ile Val Leu Tyr Val Leu Leu Ala Ala Ala Ala His

1 5 10 15

Ser Ala Phe Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser

20 25 30

Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His

35 40 45

Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys Leu Cys Lys Leu

50 55 60

Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp

65 70 75 80

Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr Ala Ser Ser Trp

85 90 95

Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly Thr Cys Tyr Pro

100 105 110

Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val

115 120 125

Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro

130 135 140

Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys Pro His Ala Gly

145 150 155 160

Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn

165 170 175

Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp Lys Gly Lys Glu

180 185 190

Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr Ser Ala Asp Gln

195 200 205

Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe Val Gly Ser Ser
210 215 220

Arg Tyr Ser Lys Lys Phe Lys Pro Ser Leu Leu Thr Glu Val Glu Thr
225 230 235 240

Pro Thr Arg Asn Gly Trp Glu Cys Lys Cys Ser Asp Ser Gly Arg Met
245 250 255

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

Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg
275 280 285

Asn Ala Gly Ser Gly Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys
290 295 300

Asn Thr Thr Cys Gln Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro
305 310 315 320

Phe Gln Asn Ile His Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val
325 330 335

Lys Ser Thr Lys Leu Arg Leu Ala Thr Gly Leu Arg Asn Ser Pro Glu
340 345 350

Asn Leu Tyr Phe Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
355 360 365

Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Gln
370 375 380

Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn
385 390 395 400

Ala Ile Asp Glu Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met
405 410 415

Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn His Leu Glu Lys
420 425 430

Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp Ile
435 440 445

Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg Thr

450 455 460

Leu Asp Tyr His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val Arg

465 470 475 480

Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu

485 490 495

Phe Tyr His Lys Cys Asp Asn Thr Cys Met Glu Ser Val Lys Asn Gly

500 505 510

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

515 520 525

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

<210> 62

<211> 1787

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Synthetic

<400> 62

ccatggtaag cgctattgtt ttatatgtgc ttttggcggc ggcggcgcat tctgcctttg 60
cggatacact gtgtattggc taccacgcca acaatagcac cgataccgtg gatacagtgc 120
tggagaagaa tgtgaccgtg acccactctg tgaatctgct ggaggataag cacaatggca 180
agctgtgtaa gctgagagga gttgcccttc tgcacctggg caaatgtaat attgccgget 240
ggattctggg aaatcctgaa tgtgaaagcc tgtctacagc cagcagctgg tcttatatcg 300
tggaaacccc tagcagcgac aatggcacct gttaccttgg cgacttcacg gattacgagg 360

agctgagaga acagctgtct agcgtgtcca gcttcgagag attcgagatc ttcctaaga 420

caagcagctg gcctaatacac gatttctaata agggagtgac agccgcctgt cctcatgccg 480
 gagccaagtc cttttacaag aacctgatct ggctggtgaa gaagggaac agctacccta 540
 agctgtctaa gagctacatc aacgacaagg gcaaagaagt gctggtgctg tggggaatcc 600
 accaccctag cacaagcgcc gatcagcaga gcctgtacca gaatgccgat gcctatgtgt 660
 ttgtgggcag cagcagatag agcaaaaagt tcaagccttc cctgtgacc gaggtggaga 720
 cccccaccag gaacggctgg gagtgcgaagt gctccgactc cggcagaatg aattactact 780
 ggaccctggt ggaacctggc gataagatca catttgaggc caccggaaat ctggtggtgc 840

 ctagatatgc atttgctatg gagagaaatg ctggctctgg catcattatc tctgataccc 900
 ctgtgcacga ctgtaatacc acctgtcaga cacctaaggg cgccattaat accagcctgc 960
 ccttcagaa taitcacctc atcaccatcg gcaagtgtcc taagtatgtg aagagcacca 1020
 agctgagact ggctaccgtg ctgagaaata gccctgaaaa cctgtatctt caaggcctgt 1080
 ttggagccat cgccggcttt attgaggag gatggaccgg aatggtgat ggctggtacg 1140
 gctatcacca ccagaatgag cagggatccg gatatgccgc cgatctgaag tctacacaga 1200
 acgccatcga cgagatcaca aacaaggtga acagcgtgat cgagaagatg aacacccagt 1260
 ttacagctgt gggcaaggag ttcaaccacc tggagaagag aatcgagaac ctgaacaaga 1320

 aagtggacga cggttcctg gatatttga cctacaatgc cgagctgctc gtgctcctgg 1380
 agaatgagag aacctggac taccacgaca gcaatgtgaa gaacctgtac gagaaggtga 1440
 gaagccagct gaagaacaat gccaaaggaga tcggcaacgg ctgctttgag ttctaccaca 1500
 agtgtgacaa cacctgtatg gagtctgtga agaacggcac ctacgactac cctaagtata 1560
 gcgaggaggc caagctgaat agagaggaga tcgacggcgt gaaactggaa agcacaagaa 1620
 tctatcaggg cgctgaaaac ctgtatcttc agggcggttc tggttacatc ccggaagctc 1680
 cgctgacgg tcaggcttac gtctgtaaag acggtgaatg ggttctgctg tctaccttcc 1740
 tgggtcacca tcatcaccac catcaccatc atcactgata aaagctt 1787