



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2014-0107569  
(43) 공개일자 2014년09월04일

- (51) 국제특허분류(Int. Cl.)  
C07K 19/00 (2006.01) C07K 14/52 (2006.01)  
C07K 14/025 (2006.01) A61K 39/12 (2006.01)  
A61P 35/00 (2006.01)
- (21) 출원번호 10-2014-7020468
- (22) 출원일자(국제) 2012년12월20일  
심사청구일자 없음
- (85) 번역문제출일자 2014년07월21일
- (86) 국제출원번호 PCT/EP2012/076404
- (87) 국제공개번호 WO 2013/092875  
국제공개일자 2013년06월27일
- (30) 우선권주장  
61/578,542 2011년12월21일 미국(US)

- (71) 출원인  
백시바디 에이에스  
노르웨이 엔-0349 오슬로 가우스타달렌 21
- (72) 발명자  
브레케, 올레 헨릭  
노르웨이 오슬로 엔오-0476 토르쇼브그트. 10에이  
프레드릭센, 아그네데 브룬스빅  
노르웨이 랄린겐 엔오-2005 스트란드본. 5  
(뒷면에 계속)
- (74) 대리인  
박장원

전체 청구항 수 : 총 56 항

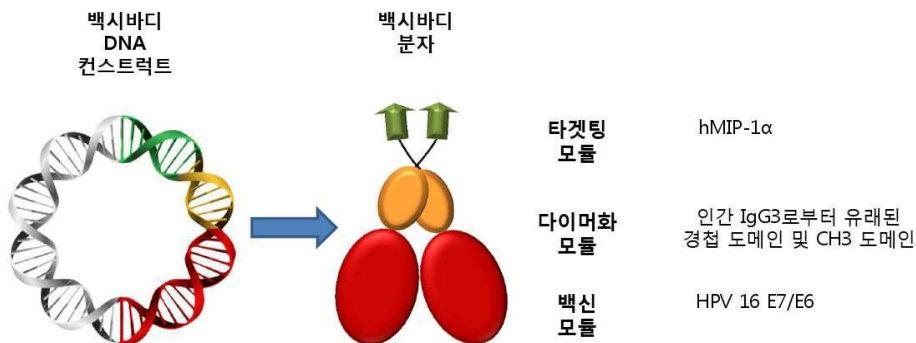
(54) 발명의 명칭 HPV에 대한 백신들

(57) 요약

본 발명은 인간 파필로마바이러스(HPV)에 대한 백신들 같은 치료적 화합물들, 특히 HPV16 또는 HPV18에 대한 DNA 백신들에 관한 것이다. 본 발명은 추가적으로 호모다имер성 펩타이드들을 인코딩하는 단백질 컨스트럭트에 관한 것으로, 상기 펩타이드들은 DNA 백신으로부터 배출되거나 또는 개별적으로 사용될 수 있다. 상기 백신들을 생산하기 위한 약제학적 제형물들, 숙주세포들 및 방법들, 뿐 아니라 적용에 의해 다양한 HPV 유도된 질환들, 예컨대 암들 및 감염성 질환들의 치료를 위한 방법들이 추가적으로 기재된다.

대표도 - 도1

백시바디 치료적 HPV DNA 백신



(72) 발명자

**아레파드, 알리**

노르웨이 오슬로 엔오-0575 셀제베이엔 11

**린더베르그, 모나 마리**

노르웨이 오슬로 엔오-0575 갈게베르그 3지

---

## 특허청구의 범위

### 청구항 1

두 개의 동일한 아미노산 체인들의 호모다имер성 단백질로, 각 아미노산 체인은 (1) 시그널 펩타이드, (2) 타겟팅 단위(unit), (3) 다이머화 모티프 및 (4) 항원성 단위를 포함하고, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열을 포함하고, 상기 항원성 단위는 인간 파필로마바이러스(HPV)의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 초기 단백질 E6 및/또는 E7로부터 유래된 항원성 단위를 포함하는 것인 호모다имер성 단백질.

### 청구항 2

제1항에 있어서, 상기 아미노산 체인 내 상기 타겟팅 단위, 다이머화 모티프 및 항원성 단위는 N-말단부터 C-말단 순서로 타겟팅 단위, 다이머화 모티프 및 항원성 단위로 존재하는 것인 호모다имер성 단백질.

### 청구항 3

제1항 및 제2항 중 어느 하나의 항에 있어서, 상기 시그널 펩타이드는 서열번호 1의 아미노산 서열 1-23에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다имер성 단백질.

### 청구항 4

제3항에 있어서, 상기 시그널 펩타이드는 서열번호 1의 아미노산 서열 1-23에 대해 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99%, 예컨대 최소 100% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다имер성 단백질.

### 청구항 5

제1항 내지 제4항 중 어느 하나의 항에 있어서, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다имер성 단백질.

### 청구항 6

제1항 내지 제5항 중 어느 하나의 항에 있어서, 상기 다이머화 모티프는 경첩부위(hinge region)를 포함하고 선택적으로 링커를 통해 연결되는 면역글로불린 도메인 같은 다이머화를 용이하게 하는 다른 도메인을 선택적으로 포함하는 것인 호모다имер성 단백질.

### 청구항 7

제6항에 있어서, 상기 경첩부위는 IgG3로부터 유래되는 경첩부위 같은 Ig 유래된 경첩부위인 호모다имер성 단백질.

### 청구항 8

제6항 및 제7항 중 어느 하나의 항에 있어서, 상기 경첩부위는 한 개, 두 개, 또는 여러 개의 공유결합들을 형성하는 능력을 가지는 것인 호모다имер성 단백질.

### 청구항 9

제8항에 있어서, 상기 공유결합은 이황화 브릿지(disulphide bridge)인 것인 호모다имер성 단백질.

### 청구항 10

제6항 내지 제9항 중 어느 하나의 항에 있어서, 상기 다이머화 모티프의 면역글로불린 도메인은 카르복시말단(carboxyterminal) C 도메인이거나, 또는 상기 C 도메인 또는 이의 변이체와 실질적으로 동일한 서열인 것인 호모다이머성 단백질.

**청구항 11**

제10항에 있어서, 상기 카르복시말단 C 도메인은 IgG로부터 유래되는 것인 호모다이머성 단백질.

**청구항 12**

제6항 내지 제11항 중 어느 하나의 항에 있어서, 상기 다이머화 모티프의 면역글로불린 도메인은 호모다이머화시키는 능력을 가지는 것인 호모다이머성 단백질.

**청구항 13**

제6항 내지 제12항 중 어느 하나의 항에 있어서, 상기 면역글로불린 도메인은 비공유결합성 상호작용들을 통해 호모다이머화시키는 능력을 가지는 것인 호모다이머성 단백질.

**청구항 14**

제13항에 있어서 상기 비공유결합성 상호작용들은 소수성 상호작용들인 것인 호모다이머성 단백질.

**청구항 15**

제1항 내지 제14항 중 어느 하나의 항에 있어서, 상기 다이머화 도메인은 CH2 도메인을 포함하지 않는 것인 호모다이머성 단백질.

**청구항 16**

제1항 내지 제15항 중 어느 하나의 항에 있어서, 상기 다이머화 모티프는 인간 IgG3의 C<sub>H</sub>3 도메인에 링커를 통해 연결되는 경첩 엑손들(hinge exons) h1 및 h4로 구성되는 것인 호모다이머성 단백질.

**청구항 17**

제1항 내지 제16항 중 어느 하나의 항에 있어서, 상기 다이머화 모티프는 서열번호 3의 아미노산 서열 94-237에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다이머성 단백질.

**청구항 18**

제2항 내지 제17항 중 어느 하나의 항에 있어서, 상기 링커는 G<sub>3</sub>S<sub>2</sub>G<sub>3</sub>SG 링커인 것인 호모다이머성 단백질.

**청구항 19**

제1항 내지 제18항 중 어느 하나의 항에 있어서, 상기 항원성 단위 및 다이머화 모티프는 링커, 예컨대 GLGGL 링커 또는 GLSGL 링커를 통해 연결되는 것인 호모다이머성 단백질.

**청구항 20**

제1항 내지 제19항 중 어느 하나의 항에 있어서, 상기 타겟팅 단위는 서열번호 1의 24-93의 아미노산들, 또는 이의 변이체로 구성되는 것인 호모다이머성 단백질.

**청구항 21**

제1항 내지 제20항 중 어느 하나의 항에 있어서, 호모다이머성 단백질은 서열번호 1의 24-93의 아미노산들, 또는 이의 변이체로 구성된 타겟팅 단위를 가지는 동일한 호모다이머성 단백질의 친화도와 비교 시 CCR1, CCR3 및 CCR5로부터 선택된 어느 하나의 케모카인 수용체에 대한 증가된 친화도를 가지는 것인 호모다이머성 단백질.

**청구항 22**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 3의 아미노산 서열 243-293에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대

최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함하는 것인 호모다이머성 단백질.

**청구항 23**

제1항 내지 제22항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 3의 아미노산 서열 243-293에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다이머성 단백질.

**청구항 24**

제22항 또는 제23항에 있어서, 상기 항원성 단위는 서열번호 22의 F47, L50, C63, C106 및 I128로 구성된 리스트로부터 선택된 위치에서 하나 이상의 아미노산 치환들, 또는 서열번호 22의 Y43-L50으로 구성된 리스트로부터 선택된 하나 이상의 아미노산을 포함하는 결실을 포함하는 것인 호모다이머성 단백질.

**청구항 25**

제1항 내지 제23항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 3, 서열번호 5, 서열번호 7, 또는 서열번호 9의 아미노산 서열 243-293, 또는 이의 변이체 또는 항원성 단편을 포함하는 것인 호모다이머성 단백질.

**청구항 26**

제1항 내지 제23항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 3, 서열번호 5, 서열번호 7, 또는 서열번호 9의 아미노산 서열 243-293, 또는 이의 변이체 또는 항원성 단편으로 구성되는 것인 호모다이머성 단백질.

**청구항 27**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 11의 아미노산 서열 243-340에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함하는 것인 호모다이머성 단백질.

**청구항 28**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 11의 아미노산 서열 243-340에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다이머성 단백질.

**청구항 29**

제27항 또는 제28항에 있어서, 상기 항원성 단위는 서열번호 23의 C24, E26, C58, C61, C91, 및 C94로 구성된 리스트로부터 선택된 위치에서 하나 이상의 아미노산 치환들, 또는 서열번호 23의 L22-E26 및/또는 C58-C61 및/또는 C91-S95으로 구성된 리스트로부터 선택된 하나 이상의 아미노산을 포함하는 결실을 포함하는 것인 호모다이머성 단백질.

**청구항 30**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 11, 서열번호 13, 서열번호 15, 또는 서열번호 17의 아미노산 서열 243-340, 또는 이의 변이체 또는 항원성 단편을 포함하는 것인 호모다이머성 단백질.

**청구항 31**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 11, 서열번호 13, 서열번호 15, 또는 서열번호 17의 아미노산 서열 243-340, 또는 이의 변이체 또는 항원성 단편으로 구성되는 것인 호모다имер성 단백질.

**청구항 32**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호 34의 아미노산 서열 243-501에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함하는 것인 호모다имер성 단백질.

**청구항 33**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호 34의 아미노산 서열 243-501에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다имер성 단백질.

**청구항 34**

제1항 내지 제22항, 제24항, 제25항, 제27항, 제29항, 제30항, 제32항 및 제33항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 초기 단백질 E6 및 E7 모두로부터 유래된 인간 파필로마바이러스 16(HPV16)의 아미노산 서열을 포함하는 것인 호모다имер성 단백질.

**청구항 35**

제32항 내지 제34항에 있어서, 상기 항원성 단위는 서열번호 22의 F47, L50, C63, C106 및 I128 그리고 서열번호 23의 C24, E26, C58, C61, C91, C94로 구성된 리스트로부터 선택된 위치에서 하나 이상의 아미노산 치환들을 포함하는 것인 호모다имер성 단백질.

**청구항 36**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호 34의 아미노산 서열 243-501, 또는 이의 변이체 또는 항원성 단편으로 구성되는 것인 호모다имер성 단백질.

**청구항 37**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 아미노산 체인은 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호 34로 구성되는 리스트로부터 선택된 아미노산 서열, 또는 이의 변이체 또는 항원성 단편으로 구성되는 것인 호모다имер성 단백질.

**청구항 38**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 22, 서열번호 23, 서열번호 24, 및 서열번호 25로부터 선택된 어느 하나의 아미노산 서열에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함하는 것인 호모다имер성 단백질.

**청구항 39**

제1항 내지 제21항 중 어느 하나의 항에 있어서, 상기 항원성 단위는 서열번호 22, 서열번호 23, 서열번호 24, 및 서열번호 25로부터 선택된 어느 하나의 아미노산 서열에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성되는 것인 호모다имер성 단백질.

**청구항 40**

제1항 내지 제39항 중 어느 하나의 항에 있어서, 상기 호모다имер성 단백질은 어느 하나의 시그널 펩타이드 서열이 없는 성숙한 형태인 것인 호모다имер성 단백질.

**청구항 41**

(1) 시그널 펩타이드, (2) 타겟팅 단위, (3) 다이머화 모티프 및 (4) 항원성 단위를 포함하는 아미노산 체인으로, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열을 포함하고, 항원성 단위는 인간 파필로마바이러스(HPV)의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 초기 단백질 E6 및/또는 E7로부터 유래된 항원성 단위를 포함하며, 상기 아미노산 체인은 제1항 내지 제40항 중 어느 하나의 항에 따른 호모다имер성 단백질을 형성할 수 있는 것인 아미노산 체인.

**청구항 42**

제41항에 따른 아미노산 체인을 인코딩하는 DNA 같은 핵산 분자.

**청구항 43**

제42항에 있어서, 핵산 분자는 인간 코돈 최적화된 것인 핵산 분자.

**청구항 44**

서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 31 및 서열번호 33으로 구성된 리스트로부터 선택된 뉴클레오타이드 서열들 중 어느 하나, 또는 이의 변이체를 포함하는 핵산 분자.

**청구항 45**

제42항 내지 제44항에 있어서, 상기 핵산은 벡터에 의해 포함되는 것인 핵산 분자.

**청구항 46**

제42항 내지 제45항 중 어느 하나의 항에 있어서, 상기 핵산 분자는 환자에서 호모다имер성 단백질의 생산을 유도하기 위해 상기 환자에게 투여하기 위해 제형화되는 것인 핵산 분자.

**청구항 47**

의약품(medicament)으로서 사용하기 위한 제1항 내지 제40항 중 어느 하나의 항에 따른 호모다имер성 단백질, 제41항에 따른 아미노산 체인, 또는 제42항 내지 제46항 중 어느 하나의 항에 따른 핵산 분자.

**청구항 48**

제1항 내지 제40항 중 어느 하나의 항에 따른 호모다имер성 단백질, 제41항에 따른 아미노산 체인, 또는 제42항 내지 제46항 중 어느 하나의 항에 따른 핵산 분자를 포함하는 약제학적 조성물.

**청구항 49**

제42항 내지 제46항 중 어느 하나의 항에 따른 핵산 분자를 포함하는 숙주세포.

**청구항 50**

제1항 내지 제40항 중 어느 하나의 항에 따른 호모다имер성 단백질, 또는 제41항의 아미노산 체인을 제조하는 방법으로, 상기 방법은

- a) 제41항 내지 제45항 중 어느 하나의 항에 따른 핵산 분자를 세포 개체군(cell population) 내로 트랜스펙션시키는 단계;
- b) 상기 세포 개체군을 배양하는 단계;
- c) 상기 세포 개체군으로부터 발현된 상기 호모다имер성 단백질, 또는 아미노산 체인을 수득하고 정제시키는 단계를 포함하는 것인 방법.

**청구항 51**

제42항 내지 제46항 중 어느 하나의 항에 따른 핵산 분자의 면역학적으로 효과적인 양을 포함하는 DNA 백신 같은 백신을 제조하는 방법으로, 상기 방법은

- a) 제41항 내지 제45항 중 어느 하나의 항에 따른 핵산 분자를 제조하는 단계;
- b) 약제학적으로 허용가능한 담체, 희석제 또는 완충액에 단계 a)에서 얻어진 상기 핵산 분자를 용해시키는 단계를 포함하는 것인 방법.

**청구항 52**

제1항 내지 제40항 중 어느 하나의 항에 따른 호모다имер성 단백질, 또는 제41항에 따른 아미노산 체인, 또는 제42항 내지 제46항 중 어느 하나의 항에 따른 DNA 같은 핵산 분자의 면역학적으로 효과적인 양을 포함하는 HPV에 대한 백신으로, 상기 백신은 T-세포- 및 B-세포 면역반응 모두를 유발시킬 수 있는 것인 백신.

**청구항 53**

제52항에 있어서, 상기 백신은 약제학적으로 허용가능한 담체 및/또는 보조제(adjuvant)를 추가적으로 포함하는 것인 백신.

**청구항 54**

환자에서 HPV에 의해 야기되는 암 또는 감염성 질병 같은 HPV 유도된 질병 또는 상태를 치료 또는 예방하는 방법으로, 상기 방법은 제1항 내지 제40항 중 어느 하나의 항에 따른 호모다имер성 단백질, 또는 제41항에 따른 아미노산 체인, 또는 제42항 내지 제46항 중 어느 하나의 항에 따른 DNA 같은 핵산 분자를 이를 필요로 하는 상기 환자에게 투여하는 단계를 포함하는 것인 방법.

**청구항 55**

제54항에 있어서, 상기 방법은 다음의 전기천공 단계와 함께 제42항 내지 제46항 중 어느 하나의 항에 따른 DNA 같은 핵산 분자를 이를 필요로 하는 상기 환자에게 투여한 후 추가적인 전기천공을 실시하는 단계를 포함하는 것인 방법.

**청구항 56**

제54항 또는 제55항에 있어서, 상기 투여는 피부 내 또는 근육 내로 실시되는 것인 방법.

**명세서**

**기술분야**

[0001] 본 발명은 인간 파필로마바이러스(HPV)에 대한 백신들 같은 치료적 화합물들, 특히 HPV16 및/또는 HPV18에 대한 DNA 백신들에 관한 것이다. 본 발명은 추가적으로 호모다имер성 펩타이드들을 인코딩하는 단백질 컨스트럭트에 관한 것으로, 상기 펩타이드들은 DNA 백신으로부터 배출되거나 또는 개별적으로 사용될 수 있다. 상기 백신들을 생산하기 위한 약제학적 제형물들, 숙주세포들 및 방법들, 뿐 아니라 적용에 의해 다양한 HPV 유도된 질환들, 예컨대 암들 및 감염성 질환들의 치료를 위한 방법들이 추가적으로 기재된다.

**배경 기술**

- [0002] 인간 파필로마바이러스(HPV)가 자궁경부암 및 항문생식기(항문, 외음부, 질 및 음경) 암들 및 두경부암들의 서브세트 같은 다른 HPV-연관된 악성종양들의 원인이라는 것이 현재 잘 알려져 있다. 특히, HPV16 및 HPV18은 전세계 모든 자궁경부암들의 약 70%를 책임진다.
- [0003] 현재까지, 두 개의 예방적 HPV 백신들이 시장에서 판매 중이다(가다실 및 서바릭스). 예방적 백신들의 목표는 HPV 바이러스 캡시드 단백질들인 L1 및 L2에 특이적인 중화 항체들의 생산을 자극함으로써 체액성 면역반응을 유도하는 것이다. 상기 예방적 백신들은 HPV 유도된 자궁경부암 및 가능하다면 다른 HPV-연관된 악성종양들의 조절에 대한 중요한 이정표일 지라도, 상술한 백신들의 효과는 20-40년 동안 유의하게 관찰되지 않을 것이다(Ma B et al., Current Cancer Therapy Reviews, 2010). 더욱이, 예방적 백신들을 위한 대량 백신화의 적용 범위(coverage)가 이미 HPV 감염된 전세계의 실질적인 개체군 외에도 지금까지는 제한되고 있기 때문에, HPV-연관된 악성종양들은 계속 진행될 것이다. 따라서, HPV-연관된 악성종양들의 사망률(mortality) 및 유병률(morbidity), 그리고 이의 전구체 병소들을 감소시키기 위한 HPV-특이적 치료 백신을 개발하는 것이 중요할 것이다(Ma B et al., Current Cancer Therapy Reviews, 2010).
- [0004] 다양한 암 백신들 및 암 면역치료법 전략들의 개발이 지난 20년에 걸쳐서 확대되어 왔다. 여전히, 프로벤지(Dendreon INC)로 불리는 오직 하나의 치료적 암 백신만이 이제까지 전립선암을 위한 표준 치료법으로 적용되는 것이 승인되었다. 특히, 윤리적 이유들로 인해 대부분의 치료적 암 백신들이 말기 단계의 종양을 가지는 환자 그룹에서 테스트된다. 상기 종양세포들이 면역시스템을 오랫동안 회피하여 발암(carcinogenesis)에 따른 종양에 대한 면역학적 내성을 유도하는 데 기여했다는 의미에서 이러한 환자 그룹은 실질적으로 면역억제되어 있다. 또한, 상기 환자들에서 심각한 부작용으로 이어질 수 있는 종양-특이적 면역반응들을 유도하고 건강한 세포들을 죽이는 것을 피하기 위해 백신들로서 적용되는 항원들(종양-특이적 대 종양-연관된)의 선택이 중요하다. 따라서, 암 면역치료법에서 주요 도전들은 면역학적 내성을 파괴하고 종양세포들을 인지하고 죽이는 종양-특이적 이펙터(effector) 기능들을 활성화시키는 것이다. 일부 케이스 보고서들이 말기 단계의 종양 환자들에서 치료적 암 백신들에 대한 임상적 반응을 보일 지라도, 대부분의 공통적 일차 종결점은 종래의 치료법(수술, 화학 치료법 및 방사선 치료법)과 비교하여 전체적인 생존 상에 영향(impact)을 관찰한다는 것이다. 하지만, 대부분의 연구들은 결론적이지 않거나 또는 이것을 보이는 것을 완전히 실패한다. 상기 음성 결과들에 대한 하나의 이유는 우선 치료하기 위해 도전하고 있는 말기-단계의(end-stage) 종양들을 가지는 환자 그룹에 있다. 가능한 전략은 치료적 백신 임상들에서 초기-단계의(early-stage) 종양들을 가진 환자들을 포함시키는 것일 수 있다.
- [0005] 하나의 전략은 전-암성(pre-cancerous) 병소들을 타겟팅하는 것이다. 이러한 전략에 대한 도전들은 주로 많은 조직들에 대한 전암성 병소들에 의해 특이적으로 발현되는 신뢰할만한 바이오마커들의 부재 및 좋지 않은 의학적 스크리닝(존재하지 않거나 또는 민감성의 부재로 고통받는 기존의 방법)이다. 예외적으로, 이것은 HPV-유도된 악성종양들에 대한 케이스에는 적용되지 않는다. 예를 들어, 서구 나라들의 대부분은 파파니콜라우 검사(papanicolaou test; 팝 도말 검사(Pap smear test))를 실시함으로써 자궁경부이형성증 및 자궁경부암에 대한 우수한 스크리닝 프로그램들을 가지고 있다. 팝 도말 검사로부터 불명확하거나 또는 비정상적인 결과들이 있다면, 질경 검사(colposcopy)가 실시될 것이다(National Cervical Cancer Coalition). 또한, 전암성 병소에서 "고-위험(high-risk)" HPV-타입의 존재를 검출하기 위해 HPV-테스팅이 일부 환자들에게 추천될 수 있다. 따라서, HPV는 HPV-연관된 전암성 병소들, 특히 자궁경부 상피내 이형성증(CIN)에 대한 잠재적인 바이오마커에 해당한다.
- [0006] DNA 백신들은 인간 질병들, 특히 암의 치료를 위해 증가하는 잠재성을 나타냈다. DNA 백신들은 강력한 항원-특이적 면역반응들을 유도하고 상기 타겟-특이적 면역반응들을 유지하기 위해 반복적으로 투여될 수 있다. 그러한 백신들은 다른 암 치료적 포맷들과 비교하여 대규모로 생산하기에 안전하고 단순하며 경제적이라고 고려된다. 수많은 면역치료적 중재들(interventions)이 면역학적 기억을 유도하는 데 실패한 상태다. 예외적으로, DNA 백신화가 인 *비보*에서 항원-특이적인 면역학적 기억을 증대시키는 백신 산물의 지속적 배출을 보장한다. 전문적 항원-제시 세포들(APCs)로의 항원들의 직접 운반은 CD4+ 및 CD8+ T 세포 면역반응들 모두를 인 *비보*에서 자극한다. 그런 강력한 세포내 면역반응들이 인 *비트로* 및 인 *비보* 모두에서 항원-양성 악성세포들을 특이적으로 인지하고 효과적으로 죽인다는 것이 증명되었다.
- [0007] 감염성 질병들 및 암 모두를 유발할 수 있는 HPV에 대한 강력하고 특이적인 면역반응들을 유도하기 위해 개선된 백신들에 대한 당업계의 요구가 여전히 존재한다.

**발명의 내용**

**해결하려는 과제**

- [0008] 발명의 목적
- [0009] HPV에 의해 야기되는 질병들 및 상태들에 대한 DNA 백신들 같은 특이적이고 매우 효과적인 치료적 화합물들을 제공하는 것이 본 발명의 구현예들의 목적이다.
- [0010] 발명의 요약
- [0011] HPV, 예컨대 HPV16 및/또는 HPV18로부터 유래되는 초기 유전자 산물인 E6 및 E7의 항원들과 hMIP-1 $\alpha$ 의 타겟팅 모듈의 조합을 통해 치료적 백신들이 제공된다는 것이 본 발명자들에 의해 발견되었으며, HPV 유전자 산물들의 상기 강력한 면역원성 에피토프들은 APCs에 높은 효율로 제시되어 특이적이고 강력한 면역반응을 유도한다. 본 발명에 따른 산물들은 일차적으로 DNA 백신들 같은 치료적 핵산 백신들로서 계획되고, 상기 백신바디 컨스트럭트(vaccibody construct)를 인코딩하는 핵산 컨스트럭트는 상기 백신을 받는 사람 안에서 상기 단백질 산물의 인 *비보* 생산을 유발하는 치료적 화합물로서 이용된다. 하지만, 택일적으로 상기 단백질 산물 자체가 백신으로 제형화되어 직접적으로 이용될 수 있다.
- [0012] 따라서, 첫 번째 양태에서 본 발명은 두 개의 동일한 아미노산 체인들의 호모다имер성 단백질에 관한 것으로, 각 아미노산 체인은 (1) 시그널 펩타이드, (2) 타겟팅 단위(unit), (3) 다이머화 모티프 및 (4) 항원성 단위를 포함하고, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열을 포함하고, 상기 항원성 단위는 인간 파필로마바이러스(HPV)의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 초기 단백질 E6 및/또는 E7로부터 유래된 항원성 단위를 포함한다.
- [0013] 두 번째 양태에서 본 발명은 (1) 시그널 펩타이드, (2) 타겟팅 단위, (3) 다이머화 모티프 및 (4) 항원성 단위를 포함하는 아미노산 체인에 관한 것으로, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열을 포함하고, 항원성 단위는 인간 파필로마바이러스(HPV)의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 초기 단백질 E6 및/또는 E7로부터 유래된 항원성 단위를 포함하며, 상기 아미노산 체인은 본 발명에 따른 호모다имер성 단백질을 형성할 수 있다.
- [0014] 세 번째 양태에서 본 발명은 (1) 시그널 펩타이드, (2) 타겟팅 단위, (3) 다이머화 모티프 및 (4) 항원성 단위를 포함하는 아미노산 체인을 인코딩하는 DNA 같은 핵산 분자에 관한 것으로, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열을 포함하고, 항원성 단위는 인간 파필로마바이러스(HPV)의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 초기 단백질 E6 및/또는 E7로부터 유래된 항원성 단위를 포함하며, 상기 아미노산 체인은 본 발명에 따른 호모다имер성 단백질을 형성할 수 있다.
- [0015] 추가적인 양태에서 본 발명은 의약품으로서 사용하기 위한 본 발명에 따른 호모다имер성 단백질, 또는 본 발명에 따른 아미노산 체인, 또는 본 발명에 따른 핵산 분자에 관한 것이다.
- [0016] 추가적인 양태에서 본 발명은 본 발명에 따른 호모다имер성 단백질, 또는 본 발명에 따른 아미노산 체인, 또는 본 발명에 따른 핵산 분자를 포함하는 약제학적 조성물에 관한 것이다.
- [0017] 추가적인 양태에서 본 발명은 본 발명에 따른 핵산 분자를 포함하는 숙주세포에 관한 것이다.
- [0018] 추가적인 양태에서 본 발명은 본 발명에 따른 호모다имер성 단백질, 또는 본 발명에 따른 아미노산 체인을 제조하는 방법에 관한 것으로, 상기 방법은 a) 본 발명에 따른 핵산 분자를 세포 개체군(cell population) 내로 트랜스펙션시키는 단계; b) 상기 세포 개체군을 배양하는 단계; c) 상기 세포 개체군으로부터 발현된 상기 호모다имер성 단백질, 또는 아미노산 체인을 수득하고 정제시키는 단계를 포함한다.
- [0019] 추가적인 양태에서 본 발명은 본 발명에 따른 핵산 분자의 면역학적으로 효과적인 양을 포함하는 DNA 백신 같은 백신을 제조하는 방법으로, 상기 방법은 a) 본 발명에 따른 핵산 분자를 제조하는 단계; b) 약제학적으로 허용 가능한 담체, 희석제 또는 완충액에 단계 a)에서 얻어진 상기 핵산 분자를 용해시키는 단계를 포함한다.
- [0020] 추가적인 양태에서 본 발명은 본 발명에 따른 호모다имер성 단백질, 또는 본 발명에 따른 아미노산 체인, 또는

본 발명에 따른 DNA 같은 핵산 분자의 면역학적으로 효과적인 양을 포함하는 HPV에 대한 백신에 관한 것으로, 상기 백신은 T-세포- 및 B-세포 면역반응 모두를 유발시킬 수 있다.

[0021] 추가적인 양태에서 본 발명은 환자에서 HPV에 의해 야기되는 암 또는 감염성 질병 같은 HPV 유도된 질병 또는 상태를 치료 또는 예방하는 방법에 관한 것으로, 상기 방법은 본 발명에 따른 호모다имер성 단백질, 또는 본 발명에 따른 아미노산 체인, 또는 본 발명에 따른 DNA 같은 핵산 분자를 이를 필요로 하는 상기 환자에게 투여하는 단계를 포함한다.

**과제의 해결 수단**

[0022] 본 발명의 발명자들에 의해 본 명세서에서 기재된 컨스트럭트들("백시바디(vaccibody)" 분자들/백신들/컨스트럭트들로도 언급됨) 및 DNA 백신 기술은 감염성 질병들 및 암 모두에 대한 강력하고 특이적인 면역반응들을 유도하는 신규한 백신 전략을 표현한다. HPV E6/E7, 예컨대 본 명세서에 기재된 HPV16 또는 HPV18 E6/E7 백신은 피부내 주입, 바람직하게는 피부내 주입 후 전기천공에 의해 DNA 백신으로서 투여될 수 있다. 이것은 수지상세포들(랑게르한스 세포들)을 포함하는 주입 위치(피부)의 세포에서 백시바디-HPV16 및/또는 HPV18 E6/E7 백신을 인코딩하는 DNA-컨스트럭트의 흡수(uptake)를 초래하여 상기 백시바디-E6/E7 분자의 인 비보 생산을 유도한다.

[0023] HPV16 및 18 같은 "고-위험" HPV 타입들로부터 유래되는 초기 유전자 산물들인 E6 및 E7은 기저-상피세포들의 형질전환 및 전암성 병소의 유도를 담당할 수 있다. 두 개의 단백질들은 높은 면역원성 에피토프들로 구성되고 강한 면역반응들을 유도하여 인 비트로 및 인 비보 모두에서 "고-위험" HPV 양성 종양세포들의 특이적 박멸을 유발하는 것으로 본 명세서에서 보여진다.

[0024] 본 명세서에 기재된 백시바디 분자는 3개의 모듈들로 구성된 호모다имер이다: 타겟팅 모듈, 다이머화 모듈 및 백신 모듈(도 1). 상기 3개의 모듈들을 인코딩하는 유전자들이 하나의 유전자로 발현되도록 유전학적으로 조작된다. 인 비보에서 발현되는 경우, 상기 백시바디 분자는 항원 제시 세포들(APCs)를 타겟팅하여 동일한 비-타겟팅된 항원들과 비교하여 증대된 백신 효능을 초래한다. 케모카인 인간 대식세포 염증성 단백질 1 알파(hMIP-1 $\alpha$  / LD78 $\beta$ )의 인 비보 발현은 상기 발현 위치로 CCR1 및 CCR5 수용체들을 운반하는 DCs, 호중구들 및 다른 면역 세포들의 유인(attraction)을 초래한다. 따라서, 타겟팅 모듈로서 hMIP-1 $\alpha$ 로 구성되는 백시바디 분자는 특이적 세포들에 대한 항원들을 타겟팅할 뿐 아니라 주입 위치로 특이적 면역세포들을 회귀시킴으로서 반응-증폭 효과(보조적 효과)를 추가적으로 제공한다. 이러한 독특한 기작은 백신 자체가 보조적 효과(adjutant effect)를 제공하기 때문에 환자들이 어떠한 추가적인 보조제들 없이 백신을 받을 수 있는 임상적 세팅에서 매우 중요할 것이다.

[0025] 본 발명의 발명자들은 HPV16 또는 HPV18 서브타입으로부터 기원하는 E6 전장 서열과 융합하는 E7 전장 유전적 서열로 구성되는 항원성 모듈의 백신 컨스트럭트들을 본 명세서에서 기재한다. 이러한 포맷의 장점은 E6 및 E7 모두가 하나의 컨스트럭트에 존재하고 인 비보에서 동일하게 발현될 것이라는 것이다. 결과적으로, 다중-항원성(multi-antigenic) 단위로 구성되는 하나의 백시바디 분자는 면역시스템에 대해 동일한 레벨의 E6 및 E7을 나타낼 수 있다. HPV16 E6 및 E7 유전자 산물들은 천연 형태에서 발암성이다. 이들의 발암성 특징들을 증화시키기 위해, 특이적 위치들에서의 돌연변이들이 E6 및 E7 유전적 서열에 도입될 수 있다.

[0026] 결실들을 포함하는 돌연변이들이 E6 및 E7의 발암성 특징들을 억제하는 것으로 알려진 특이적 위치들에 도입될 수 있는데, 예컨대 Dalal S 등, *J Virol*, 1996; Munger K 등, *EMBO*, 1989; Nakagawa S 등, *Virology*, 1995; Crook T 등, *Cell*, 1991; Munger K 등, *HPV Compendium Online*, 1997 ([http://www.stdgen.lanl.gov/COMPENDIUM\\_PDF/97PDF/3/E7.pdf](http://www.stdgen.lanl.gov/COMPENDIUM_PDF/97PDF/3/E7.pdf)); Nguyen, M 등, *J Virol*, 2002; Nomine Y 등, *Molecular Cell*, 2006; Moody C 등, *Nat Rev Cancer*, 2010, Polakova I 등, *Vaccine*, 2010; Xie Q, *Virologica Sinica*, 2011; Mesplede T 등, *J Virol*, 2012; US 2008/0102084 및 US6306397 중 어느 하나에 기재되어 있고 상기 레퍼런스들이 본 명세서에 참조로서 삽입된다. 따라서, 본 발명의 일부 양태들에서, 본 발명에 따른 컨스트럭트들은 Dalal S 등, *J Virol*, 1996; Munger K 등, *EMBO*, 1989; Nakagawa S 등, *Virology*, 1995; Crook T 등, *Cell*, 1991; Munger K 등, *HPV Compendium Online*, 1997 ([http://www.stdgen.lanl.gov/COMPENDIUM\\_PDF/97PDF/3/E7.pdf](http://www.stdgen.lanl.gov/COMPENDIUM_PDF/97PDF/3/E7.pdf)); Nguyen, M 등, *J Virol*, 2002; Nomine Y 등, *Molecular Cell*, 2006; Moody C 등, *Nat Rev Cancer*, 2010, Polakova I 등, *Vaccine*, 2010; Xie Q, *Virologica Sinica*, 2011; Mesplede T 등, *J Virol*, 2012; US 2008/0102084 또는 US6306397에 기재된 대로 발암성 특징들을 억제하는 것으로 알려진 위치에서 HPV16 E6, E7 또는 양자 모두 중 어느 것에서 하나 이상의 돌연변이들을 가지는 HPV16 E6, E7 또는 HPV16 E6/E7 키메라성 컨스트럭트들을 포함한다. 본 발명의 다른 양태들에서, 본 발명에 따른 컨스트럭트들은 Dalal S 등, *J Virol*, 1996; Munger K 등, *EMBO*, 1989; Nakagawa S 등,

Virology, 1995; Crook T 등, Cell, 1991; Munger K 등, HPV Compendium Online, 1997 ([http://www.stdgen.lanl.gov/COMPENDIUM\\_PDF/97PDF/3/E7.pdf](http://www.stdgen.lanl.gov/COMPENDIUM_PDF/97PDF/3/E7.pdf)); Moody C 등, Nat Rev Cancer, 2010, US 2008/0102084 및 US6306397에 기재된 대로 발암성 특징들을 억제하는 것으로 알려진 위치에서 HPV18 E6, E7 또는 양자 모두 중 어느 것에서 하나 이상의 돌연변이들을 가지는 HPV18 E6, E7 또는 HPV18 E6/E7 키메라성 컨스트럭트들을 포함한다.

[0027] 백신바디-모이어티(타겟팅 모듈 및 다이머화 모듈)는 최종 융합 단백질에서 E6 및 E7 야생형 단백질들의 발암성 특징들을 제거시킬 수 있는 가능성이 있다. 따라서, 백신바디 구축에서 야생형 전장 E6 및/또는 E7 서열들의 이용이 본 발명의 보다 다른 양태에 존재한다.

[0028] 본 발명은 도 1에 기재된 전체 포맷에 기반되는 여러 가지 변이체의 백신바디 HPV 치료적 DNA 백신들 모두를 기재하는데, 상기 치료적 백신바디-HPV DNA 백신들은 인간에서 천연적으로 발현되는 유전자들을 인코딩한다; 상기 타겟팅 모듈 유전자들은 APCs의 세포 표면에 발현되는 동족의(cognate) 수용체들인 CCR1 및 CCR5에 결합하는 케모카인 hMIP-1 $\alpha$ 를 인코딩한다. 상기 다이머화 모듈 유전자들은 예컨대 호모다имер 분자를 발생시키는 2개의 백신바디 노모머들을 연결하는 인간 IgG3로부터 유래된 경첩부위 및 중쇄의 불변부위(constant heavy chain) 3을 인코딩할 수 있다. 본 전략을 위한 상기 백신 모듈을 인코딩하는 유전자들은 HPV, 예컨대 HPV 16 및/또는 HPV18 E7 및 E6 항원들, 예컨대 전장 HPV16 E7 및 E6 항원들로 구성되고, 선택적으로 상기 발암성 특징들을 억제하는 하나 이상의 돌연변이를 포함한다. 일단 인 vivo에서 피부내(i.d.) 주입 후 전기천공에 따라 투여되는 경우, 상기 백신 컨스트럭트를 받아들인 진피세포(dermal cells)는 상기 백신바디-HPV 분자를 발현할 것이다. 인 vivo 생산된 백신바디 백신들은 피부 내 APCs, 특히 DCs의 표면 상에 발현되는 CCR1 및 CCR5로 타겟팅한다. 상기 백신바디 분자와 이의 동족 수용체들 간의 결합은 APC에서 복합체의 내재화, MHC 분자들 위로 로딩되어 HPV16 E6 및 E7 특이적 면역반응들을 유도하는 CD4<sup>+</sup> 및 CD8<sup>+</sup> T 세포들에 제시되는 작은 펩타이드들로의 상기 단백질들의 분해를 초래한다. 일단 자극되고 활성화된 CD4<sup>+</sup> T 세포들로부터 도움을 받는 경우, CD8<sup>+</sup> T 세포들은 HPV16 E6 및 E7을 발현하는 세포들을 타겟팅하여 죽일 것이다(도 2). "빌트-인(built-in)" 보조적 효과와 함께 백신에 대한 그러한 증대된 면역반응들은 면역학적 내성을 파괴하고 악성 세포들을 효과적으로 죽임으로써 종양-회피(tumor-escape)(종양 면역 감시(tumor immune surveillance))를 강력하게 극복시킬 것이다. hMIP-1 $\alpha$  타겟팅 단위는 다이머화 모티프, 예컨대 경첩부위를 통해 항원성 단위에 연결될 수 있고, 상기 연결은 상기 항원성 단위의 COOH-말단 또는 NH<sub>2</sub>-말단 중 어느 하나에 존재한다. 본 발명은 이러한 재조합 단백질을 코딩하는 DNA 서열, 뿐 아니라 상술한 DNA 서열들을 포함하는 발현 벡터들, 상기 발현 벡터들을 포함하는 세포주들, 우선적으로 백신바디 DNA, 백신바디 RNA, 또는 백신바디 단백질을 이용한 면역화에 의한 포유동물들의 치료, 및 마지막으로 상기 분자들을 포함하는 제약물 및 키트에 관한 것이다.

[0029] 본 발명에 따른 단백질들에서 다이머화 모티프는 경첩부위 및 면역글로불린 도메인(예컨대, C $\gamma$ 3 도메인), 예를 들어 카르복시말단 C 도메인(C $\mu$ 3 도메인), 또는 상기 C 도메인과 실질적으로 동일한 서열을 포함하도록 구축될 수 있다. 상기 경첩부위는 Ig-유래될 수 있고 체인 간(interchain) 공유결합(들), 예컨대 이황화 브릿지(들)(disulfide bridge(s))의 형성을 통한 다이머화에 기여한다. 또한, 상기 경첩부위는 도메인들 간의 탄력적 스페이서(flexible spacer)로서 기능하여 2개의 타겟팅 단위들이 다양한 거리를 가지고 발현되는 APC 상의 2개의 타겟 분자들에 동시에 결합하도록 허용한다. 상기 면역글로불린 도메인들은 비-공유결합성 상호작용들, 예컨대 소수성 상호작용들을 통한 호모다имер화에 기여한다. 바람직한 구현예에서, C $\mu$ 3 도메인은 IgG로부터 유래된다. 상술한 다이머화 모티프들은 다른 다중화(multimerization) 모이어티들(예컨대, 다른 Ig 이소형들/서브클래스들)로 교환될 수 있다. 바람직하게는, 상기 다이머화 모티프는 천연 인간 단백질들, 예컨대 인간 IgG로부터 유래된다.

[0030] 상기 다이머화 모티프는 항원성 단위 및 타겟팅 단위에 대해 어떠한 방향(orientation)을 가질 수 있다는 것이 자명할 것이다. 하나의 구현예에서 상기 다이머화 모티프의 N-말단에 존재하는 상기 타겟팅 단위와 함께 상기 항원성 단위는 상기 다이머화 모티프의 COOH-말단에 존재한다. 다른 구현예에서 상기 다이머화 모티프의 COOH-말단에 존재하는 상기 타겟팅 단위와 함께 상기 항원성 단위는 상기 다이머화 모티프의 N-말단에 존재한다.

[0031] 본 명세서에 의해 참조로서 삽입되는 국제출원 WO 2004/076489는, 본 발명에 따라 이용될 수 있는 핵산 서열들 및 벡터들을 개시하고 있다.

[0032] 본 발명에 따른 단백질들은 HPV, 예컨대 HPV16 E7 및 E6 항원들, 예컨대 전장 HPV16 E7 및 E6 항원들 뿐 아니라 이의 면역원성 단편들 또는 변이체들로부터 유래되는 항원성 단위를 포함한다. 상기 항원성 서열은 충분한 길이

로 구성될 것이다. 그러한 항원성 단위의 최소 길이는 약 9개의 아미노산 근처일 수 있다. 따라서, 일부 구현예들에서, HPV로부터 유래된 항원성 단위는 그러한 항원성 단위를 인코딩하는 핵산 서열에서 최소 약 27개의 뉴클레오타이드들에 상응하는 최소 9개의 아미노산들의 아미노산 서열을 포함한다. 바람직하게는 HPV로부터 유래되는 항원성 단위는 상당히 더 길 수 있는데, 예컨대 전장 HPV16 E7 및 E6 항원들이다. 다양성(diversity)은 코딩 부위에서 제한된 뉴클레오타이드 변화들(<2%의 빈도에서) 및 비-코딩 부위에서 제한된 뉴클레오타이드 변화들(<5%의 빈도에서)를 통해 주어진 HPV 유전형 내에서 발생한다(Bernard, HU et al., Int J Cancer, 2006). 그러한 변이체들은 지리학적인 기원에 기반하여 계통발생학적으로 분리되어 유럽인, 아프리카인, 아시아인, 아시아계 미국인 및 북미 사람으로 표지된다. 백신바디 포맷에서 그러한 서열들의 삽입은 면역반응의 양팔들의 활성화를 유발시킬 수 있다.

[0033] 백신바디 단백질, 백신바디 DNA, 또는 백신바디 RNA에 의한 면역화, 예를 들어 마지막 두 개는 근육내 또는 피부내 주입 후 전기천공을 실시하거나 또는 실시함이 없이 실행되는 면역화는 모두 본 발명에 따른 실현 가능한 방법들이다.

[0034] 상기 논의된 대로, 본 발명은 HPV에 의해 야기되는 암 또는 감염성 질병들에 대한 백신 조성물에 관한 것으로, 상기 백신 조성물은 본 발명의 분자를 인코딩하는 핵산 또는 이의 축퇴성(degenerate) 변이체들의 면역학적으로 효과적인 양을 포함한다. 상기 백신은 T-세포- 및 B-세포 면역반응 모두를 유발시킬 수 있다. 또한, 본 발명은 진단적, 의학적 또는 과학적 목적을 위한 백신바디 DNA, RNA, 또는 단백질을 포함하는 키트에 관한 것이다.

[0035] 본 발명은 추가적으로 본 발명의 분자를 포함하는 벡터를 세포 개체군 내로 트랜스펙션시키는 단계; 상기 세포 개체군을 배양하는 단계; 상기 세포 개체군으로부터 발현된 재조합 단백질을 수득하는 단계; 및 상기 발현된 단백질을 정제하는 단계를 포함하는 본 발명의 재조합 분자를 제조하는 방법에 관한 것이다.

[0036] 상기 기재된 뉴클레오타이드 서열들은 유전자 치료법에 맞춰진 벡터, 예컨대 특이적 프로모터의 조절 하의 벡터 내로 삽입되어 세포 내로 도입될 수 있다. 일부 구현예들에서 상기 DNA 서열을 포함하는 벡터는 바이러스, 예컨대 아데노바이러스, 백신니아 바이러스 또는 아데노-연관된 바이러스이다. 일부 구현예들에서 레트로바이러스는 벡터로서 이용된다. 적합한 레트로바이러스의 예들은, 예컨대 MoMuLV 또는 HaMuSV이다. 유전자 치료법의 목적을 위해, 본 발명에 따른 DNA/RNA 서열들도 콜로이드 분산액들의 형태에서 상기 타겟 세포들로 전달될 수 있다. 상기 콜로이드 분산액들은, 예를 들어 리포솜들 또는 리포플렉스들(lipoplexes)을 포함할 수 있다.

[0037] 본 발명은 본 명세서에서 특정된 아미노산 서열(들) 또는 본 명세서에서 특정된 특이적 특징들을 가지는 폴리펩타이드와 최소 정도의 서열 동일성 또는 서열 상동성을 가지는 항원성 단위 뿐 아니라 타겟팅 단위의 용도를 포괄한다. 특히, 본 발명은 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 32, 또는 서열번호 34 중 어느 하나와 어느 정도의 서열 동일성을 가지는 본 발명에 따른 컨스트럭트들에 이용되는 펩타이드 변이체들 또는 펩타이드 단위들의 용도를 포괄한다. 여기에서, 상기 용어 "변이체(variant)"는 대상 아미노산 서열들 또는 대상 뉴클레오타이드 서열들과 어느 정도의 서열 동일성을 가지는 실체(entity)를 의미하고, 상기 대상 아미노산 서열은 바람직하게는 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 13, 서열번호 14, 서열번호 15, 서열번호 16, 서열번호 17, 서열번호 18, 서열번호 19, 서열번호 20, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 31, 서열번호 32, 서열번호 33, 또는 서열번호 34이다.

[0038] 하나의 양태에서, 상기 변이체 또는 단편 아미노산 서열 및/또는 뉴클레오타이드 서열은 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 32, 또는 서열번호 34의 폴리펩타이드의 기능적 활성을 보유하고/보유하거나 상기 폴리펩타이드의 활성을 증대시키는 폴리펩타이드를 제공하고/제공하거나 인코딩한다.

[0039] 본 발명의 맥락에서, 변이체 서열은 상기 대상 서열과 최소 80%, 최소 85%, 최소 90%, 최소 95%, 최소 96%, 최소 97%, 최소 98% 또는 최소 99% 동일한 아미노산 서열을 포함하도록 형성된다. 전형적으로, 본 발명에 따라 이용되는 변이체들은 대상 아미노산 서열처럼 동일한 활성 위치들 등을 포함할 것이다. 본 발명의 맥락에서 상동성도 유사성(즉, 유사한 화학적 특성들/기능들을 가지는 아미노산 잔기들)의 측면에서 고려될 지라도, 서열 동일성의 측면에서 상동성을 표현하는 것이 바람직하다.

[0040] 서열 동일성 비교들은 눈으로, 또는 보다 일반적으로 이미 이용가능한 서열 비교 컴퓨터 프로그램들의 도움으로 실시될 수 있다. 상술한 상업적으로 이용가능한 컴퓨터 프로그램들은 복잡한 비교 알고리즘들을 이용하여 진화적 이벤트들을 가장 잘 반영하는 상기 2개 이상의 서열들을 정렬시켜 상기 2개 이상의 서열들 간의 차이(들)로 이어졌다. 따라서, 상술한 알고리즘들은 동일하거나 또는 유사한 아미노산들의 정렬을 보상하고 비-유사성 아미노산들의 갭, 갭 연장 및 정렬의 삽입을 페널티로 주는 스코어링 시스템으로 작동한다. 상기 비교 알고리즘들의 스코어링 시스템은:

- [0041] i) 갭이 삽입될 때마다 페널티 스코어의 할당(갭 페널티 스코어),
- [0042] ii) 존재하는 갭이 부가 위치(extra position)와 함께 연장되는 때마다 페널티 스코어의 할당(연장 페널티 스코어),
- [0043] iii) 동일한 아미노산들의 정렬 상의 높은 스코어들의 할당, 및
- [0044] iv) 비-동일성 아미노산들의 정렬 상의 다양한 스코어들의 할당을 포함한다.

[0045] 대부분의 정렬 프로그램들은 변형되는 갭 페널티들을 허용한다. 하지만, 서열 비교를 위한 그러한 소프트웨어를 이용하는 경우 내정값(default values)을 이용하는 것이 바람직하다.

[0046] 비-동일성 아미노산들의 정렬을 위해 주어지는 스코어들은 치환 매트릭스으로도 불리는 스코어링 매트릭스에 따라 할당된다. 그러한 치환 매트릭스들에서 제공되는 스코어들은 진화과정 동안 다른 아미노산으로 치환되는 하나의 아미노산의 가능성이 다양하고 치환되는 아미노산의 물리적/화학적 특성에 의존적이라는 사실을 반영하고 있다. 예를 들어, 다른 극성 아미노산으로 치환되는 극성 아미노산의 가능성이 소수성 아미노산으로 치환되는 것과 비교하여 더 높다. 따라서, 상기 스코어링 매트릭스는 동일한 아미노산들에 대해 가장 높은 스코어, 비-동일성이지만 유사한 아미노산에 대해 더 낮은 스코어 및 비-동일성 비-유사한 아미노산들에 대해 매우 낮은 스코어를 할당할 것이다. 가장 빈번하게 이용되는 스코어링 매트릭스는 PAM 매트릭스(Dayhoff et al. (1978), Jones et al. (1992)), BLOSUM 매트릭스(Henikoff and Henikoff (1992)) 및 Gonnet 매트릭스(Gonnet et al. (1992))이다.

[0047] 그러한 정렬을 실시하기 위해 적합한 컴퓨터 프로그램들은 Vector NTI(Invitrogen Corp.) 및 ClustalV, ClustalW 및 ClustalW2 프로그램들(Higgins DG & Sharp PM (1988), Higgins et al. (1992), Thompson et al. (1994), Larkin et al. (2007)을 포함하지만, 이에 한정되는 것은 아니다. 다른 정렬 수단들의 선택은 ExPASy 프로테오믹스 서버인 [www.expasy.org](http://www.expasy.org)로부터 이용가능하다. 서열 정렬을 실시할 수 있는 소프트웨어의 다른 예는 Altschul 등 (1990) J. Mol. Biol. 215: 403-410에 처음으로 기재되었고 현재 <http://www.ncbi.nlm.nih.gov/>에서 발견될 수 있는 미국 국립생물공학정보센터(National Center for Biotechnology Information)의 웹페이지로부터 이용가능한 BLAST(Basic Local Alignment Search Tool)이다.

[0048] 상기 소프트웨어가 정렬을 만들면, % 유사성 및 % 서열 동일성을 계산하는 것이 가능하다. 전형적으로 상기 소프트웨어는 서열 비교의 일부로서 이것을 행해 수치적 결과를 발생시킨다.

[0049] 하나의 구현예에서, 서열 정렬을 실시하기 위해 ClustalW 소프트웨어를 이용하는 것이 바람직하다. 바람직하게는, ClustalW를 이용한 정렬은 쌍 정렬을 위해 다음의 매개변수들로 실시된다:

|           |            |
|-----------|------------|
| 치환 매트릭스:  | Gonnet 250 |
| 갭 오픈 페널티: | 20         |
| 갭 연장 페널티: | 0.2        |
| 갭 끝 페널티:  | 없음         |

[0050]

[0051] 예를 들어, ClustalW2는 유럽 생물정보학 연구소(European Bioinformatics Institute) 내 수단-서열분석-ClustalW2 하의 EMBL-EBI 웹사이트인 [www.ebi.ac.uk](http://www.ebi.ac.uk)에 의한 인터넷 상에서 이용가능하다. 현재, 상기 ClustalW2 수단의 정확한 주소는 [www.ebi.ac.uk/Tools/clustalw2](http://www.ebi.ac.uk/Tools/clustalw2)이다.

[0052] 다른 구현예에서, 서열 정렬을 실시하기 위해 Vector NTI(Invitrogen) 내 프로그램 Align X를 이용하는 것이 바람직하다. 하나의 구현예에서, 사용되었던 Exp10은 내정값 세팅으로 이용될 수 있다:

[0053] 갭 오픈 페널티: 10

- [0054] 갭 연장 페널티: 0.05
- [0055] 갭 분리 페널티 범위: 8
- [0056] 스코어 매트릭스: blosum62mt2
- [0057] 따라서, 본 발명은 또한 본 명세서에서 특정된 대로 단백질, 폴리펩타이드, 모티프 또는 도메인의 어느 하나의 아미노산 서열, 특히 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 32, 또는 서열번호 34의 어느 하나의 아미노산 서열의 변이체들, 단편들, 및 유도체들의 용도를 포함한다.
- [0058] 상기 서열들, 특히 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 32, 또는 서열번호 34의 변이체들, 단편들, 및 유도체들의 서열들도 침묵 변화(silent change)를 생산하고 기능적으로 동등한 물질을 초래하는 아미노산 잔기들의 결실들, 삽입들 또는 치환들을 가질 수 있다. 의도적 아미노산 치환들이 물질의 이차 결합 활성이 보유되는 한 잔기들의 극성, 전하, 용해성, 소수성, 친수성, 및/또는 양친매성 특징에서의 유사성에 기반하여 실시될 수 있다. 예를 들어, 음적 전하된 아미노산들은 아스파르트산 및 글루탐산을 포함하고; 양적 전하된 아미노산들은 라이신 및 아르기닌을 포함하고; 유사한 친수성 값을 가지는 전하를 나타내지 않는 극성 머리기들을 가지는 아미노산들은 루이신, 이소루이신, 발린, 글라이신, 알라닌, 아스파라긴, 글루타민, 세린, 트레오닌, 페닐알라닌, 및 타이로신을 포함한다.
- [0059] 또한, 본 발명은 보존적 치환(치환 및 대체는 택일적 잔기로 존재하는 아미노산 잔기의 교환을 나타내기 위해 본 명세서에서 모두 이용됨), 다시 말하면 염기성에 대해 염기성, 산성에 대해 산성, 극성에 대해 극성 등과 같은 유사한 것에 대해 유사한 것(like-for-like)으로의 치환을 일으키는 것을 포함한다. 또한, 비-보존적 치환은, 다시 말하면 하나의 클래스의 잔기로부터 다른 클래스의 잔기 또는 택일적으로 오르니틴(이후 본 명세서에서 Z로 언급됨), 디아미노부틸산(diaminobutyric acid) 오르니틴(이후 본 명세서에서 B로 언급됨), 노르루이신 오르니틴(이후 본 명세서에서 O로 언급됨), 피리일알라닌(pyriylalanine), 티에닐알라닌, 나프틸알라닌 및 페닐글라이신 같은 비정상적인(unnatural) 아미노산들의 함유를 포함할 수 있다.
- [0060] 실행될 수 있는 보존적 치환들은, 예를 들어 염기성 아미노산들(아르기닌, 라이신 및 히스티딘), 산성 아미노산들(글루탐산 및 아스파르트산), 양친매성 아미노산(알라닌, 발린, 루이신, 이소루이신), 극성 아미노산들(글루타민, 아스파라긴, 세린, 트레오닌), 방향족성 아미노산들(페닐알라닌, 트립토판 및 타이로신), 하이드록실 아미노산들(세린, 트레오닌), 큰 아미노산들(페닐알라닌 및 트립토판) 및 작은 아미노산들(글라이신, 알라닌)의 그룹들 내에 존재한다.
- [0061] 또한, 대체들은 알파\* 및 알파-이중치환된(disubstituted)\* 아미노산들, N-알킬 아미노산들\*, 젯산\*, 트리플루오로타이로신\*, p-Cl-페닐알라닌\*, p-Br-페닐알라닌\*, p-I-페닐알라닌\*, L-알릴-글라이신\*, β-알라닌\*, L-α-아미노 부틸산\*, L-γ-아미노 부틸산\*, L-α-아미노 이소부틸산\*, L-ε-아미노 카프로산#, 7-아미노 헵탄산\*, L-메티오닌 술폰<sup>#</sup>, L-노르루이신\*, L-노르발린\*, p-니트로-L-페닐알라닌\*, L-하이드록시프로린<sup>#</sup>, L-티오프롤린\*, 같은 천연 아미노산들의 할라이드 유도체들, 4-메틸-Phe\*, 펜타메틸-Phe\*, L-Phe (4-아미노)<sup>#</sup>, L-Tyr (메틸)\*, L-Phe (4-이소프로필)\*, L-Tic (1,2,3,4-테트라하이드로이소퀴놀린-3-카르복실산)\*, L-디아미노프로피온산<sup>#</sup> 및 L-Phe (4-벤질)\* 같은 페닐알라닌(Phe)의 메틸 유도체들을 포함하는 비정상적인 아미노산들에 의해 실시될 수 있다. 표기 \*는 상기 유도체의 소수성 특성을 지시하기 위해 상기 논의사항(상동성 또는 비-보존적 치환에 관련하여)의 목적을 위해 이용되는 반면에, 표기 #은 상기 유도체의 친수성 특성을 나타내기 위해 이용되었고, 표기 \*\*는 상기 유도체의 양친매성 특징을 나타내기 위해 이용되었다.
- [0062] 변이체 아미노산 서열들은 글라이신 또는 β-알라닌 잔기들 같은 아미노산 스페이서들 외에도 메틸, 에틸 또는 프로필 같은 알킬기들을 포함하는 서열의 어떠한 두 개의 아미노산 잔기들 사이에 삽입될 수 있는 적합한 스페이서 기들을 포함할 수 있다. 변이(variation)의 추가적인 형태는 당업자에 의해 잘 이해될 것인 펩타이드 형태에서 하나 이상의 아미노산 잔기들의 존재를 포함한다. 의문을 피하기 위해, "펩타이드 형태(peptoid form)"는 변이체 아미노산 잔기를 언급하는 데 이용되고, α-탄소 치환기는 α-탄소보다는 오히려 상기 잔기의 질소 원자 상에 있다. 펩타이드 형태로 펩타이드들을 제조하는 과정들은 당업계, 예를 들어 Simon RJ *et al.* (1992),

HorwellDC. (1995)에 알려져 있다.

- [0063] 하나의 구현예에서, 본 발명에 따른 호모다имер성 단백질에서 이용되는 상기 변이체 타겟팅 단위는 타겟팅 단위와 최소 80%, 최소 85%, 최소 90%, 최소 95%, 최소 96%, 최소 97%, 최소 98% 또는 최소 99% 아미노산 서열 상동성의 아미노산들의 서열을 가지는 변이체이다.
- [0064] 하나의 양태에서, 본 발명에서 이용되는 단백질 또는 서열은 바람직하게는 정제된 형태로 존재한다. 상기 용어 "정제된(purified)"은 주어진 구성성분(component)이 높은 레벨로 존재하는 것을 의미한다. 바람직하게는 상기 구성성분은 조성물 내에 존재하는 우세한 활성 구성성분이다.
- [0065] "변이체(variant)" 또는 "변이체들(variants)"는 단백질들, 폴리펩타이드들, 단위들, 모티프들, 도메인들 또는 핵산들을 의미한다. 상기 용어 "변이체"는 상기 용어 "돌연변이(mutant)"와 상호호환가능하게 이용될 수 있다. 변이체들은 각각 아미노산 또는 뉴클레오타이드 서열에서 하나 이상의 위치들에서의 삽입들, 치환들, 염기전환들(transversions), 절단들(truncations), 및/또는 도치들(inversions)을 포함한다. 상기 구 "변이체 폴리펩타이드(variant polypeptide)", "폴리펩타이드", "변이체" 및 "변이체 효소(variant enzyme)"는 서열번호 1의 아미노산 서열로부터 변형되었던 아미노산 서열을 가지는 폴리펩타이드/단백질을 의미한다. 상기 변이체 폴리펩타이드들은 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 32, 또는 서열번호 34의 아미노산 서열과 특정 퍼센트, 예컨대 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 가지는 폴리펩타이드를 포함한다.
- [0066] "변이체 핵산들(Variant nucleic acids)"은 본 명세서에 제시된 뉴클레오타이드 서열들과 혼성화할 수 있는 서열들에 상보적인 서열들을 포함할 수 있다. 예를 들어, 변이체 서열은 본 명세서에 제시된 뉴클레오타이드 서열들에 대해 엄격 조건들, 예컨대 50°C 및 0.2X SSC (1X SSC = 0.15 M NaCl, 0.015 M 나트륨 시트레이트, pH 7.0) 하에서 혼성화할 수 있는 서열들에 대해 상보적이다. 보다 특이적으로는, 상기 용어 변이체는 본 명세서에 제시된 뉴클레오타이드 서열들에 대해 엄격 조건들, 예컨대 65°C 및 0.1X SSC 하에서 혼성화할 수 있는 서열들에 대해 상보적인 서열들을 포함한다. 변이체 핵산의 녹는 점(Tm)은 야생형 핵산의 Tm보다 약 1, 2, 또는 3°C 더 낮을 수 있다. 상기 변이체 핵산들은 본 발명에 따른 호모다имер성 단백질을 형성할 수 있는 모노머성 단백질을 인코딩하는 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 서열번호 26, 서열번호 27, 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 32, 또는 서열번호 34을 인코딩하는 핵산과 특정 퍼센트, 예컨대 80%, 85%, 90%, 95%, 또는 99% 서열 동일성을 가지는 폴리펩타이드를 포함한다.
- [0067] 돌연변이들의 특이적 범주는 E6 및 E7에서의 돌연변이들이다:
- [0068] E6 단백질은 p53 결합을 불가능하게 함으로써 해독될 수 있다. 전장 HPV E6 단백질 내 다섯 위치들은 E6 기능성(functionality)의 불활성화에 대한 돌연변이를 위한 위치들로, F47, L50, C63, C106 및 I128이다. 상술한 위치들에서 어떠한 아미노산 치환이 E6의 불활성화를 유도할 수 있고 중앙 저해를 유도한다. 상술한 위치들 중 어느 하나에서 다른 하나의 어떠한 아미노산으로의 치환들이 잠재적으로 이용될 수 있다. 잠재적인 돌연변이들을 위한 위치들이 서열번호 22에 보여지고 있다.
- [0069] E7 단백질에서 N-말단 Rb (retinoblastoma binding protein) 결합-위치 모티프(LXCXE) 및 Zn-결합 모티프(CXXC)와 함께 2개의 보존된 부위들 3(업스트림 및 다운스트림)을 포함하는 발암성 특징들과 연관되는 보존된 부위들이 있다(참고, Phelps et al J. Virol. April 1992, vol. 66, no. 42418-242; Gulliver et al J Virol. 1997, August; 71(8)). LXCXE-모티프에서 바람직한 돌연변이 위치들은 C24 및 E26이다. 두 개의 CXXC 모티프들에서 선호되는 위치들은 C58, C61, C91 및 C94이다. 하지만, 상술한 부위들에서 어떠한 돌연변이들도 결합 기능들의 감소를 위해 치환되고, 이에 따라 E7의 발암성 효과들을 제거하기 위해 고려될 수 있다. 잠재적인 돌연변이들을 위한 위치들은 서열번호 23에 보여지고 있다.
- [0070] 시그널 펩타이드:
- [0071] 형성 초기의(nascent) 폴리펩타이드의 N-말단에 위치하는 시그널 펩타이드는 골지 복합체 내로 운반하기 전에 상기 분자를 ER 내로 지시한다. 상기 시그널 펩타이드는 상기 단백질을 ER로 타겟팅하여 운반하는 목적에 도움을 준 경우 시그널 펩티다제에 의해 절단된다. 상술한 시그널 펩타이드들은 일반적으로 15 내지 30개의 아미노산들 사이이지만, 50개 이상의 잔기들일 수도 있다(Martoglio, B. et al., Trends in Cell Biology, 1998,

*Knappskog, S. et al., J Biotechnol, 2007*). 원래의 시그널 펩타이드는 어떠한 포유동물, 원핵 또는 해양 기원으로부터 유래된 시그널 펩타이드들에 의해 대체될 수 있다. 일반적으로 시그널 펩타이드들은 많은 양의 단백질을 분비하는 천연 능력으로 인해, 예를 들어 인간 IL-2 및 인간 알부민이 이용된다. 시그널 펩타이드의 선택은 합성된 및 분비된 단백질의 양에 상당한 영향을 줄 수 있다.

- [0072] 일부 구현예들에서, 본 발명에 따른 단백질 컨스트럭트에서 이용되는 시그널 펩타이드는 케모카인 단백질, 예컨대 LD78베타의 시그널 서열로부터 유래된다.
- [0073] 일부 구현예들에서 상기 시그널 펩타이드는 국제출원번호: PCT/EP2011/060628에 개시된 pLNOH2(B1-8 변이가능한 면역글로불린 리더)로부터 유래되지 않는다.
- [0074] 일부 구현예들에서 상기 시그널 펩타이드는 면역글로불린 유전자로부터 유래되지 않는다.
- [0075] 본 명세서에서 사용되는 용어 "호모다이머성 단백질(homodimeric protein)"은 두 개의 개별 동일한 가닥들의 아미노산들, 또는 수소 결합, 이온성(전하된) 상호작용들, 실제 공유결합성 이황화 결합(bonding), 또는 상술한 상호작용들의 어떤 조합에 의해 단일 다이머성 단백질로서 함께 결합되는 서브유니트들을 포함하는 단백질을 의미한다.
- [0076] 본 명세서에서 사용되는 용어 "다이머화 모티프(dimerization motif)"는 항원성 단위와 타겟팅 단위 사이의 아미노산들의 서열로, 다이머화에 기여할 수 있는 경첩부위 및 선택적인 두 번째 도메인을 포함한다. 이러한 두 번째 도메인은 면역글로불린 도메인일 수 있고, 선택적으로 경첩부위 및 두 번째 도메인이 링커를 통해 연결된다. 따라서 상기 다이머화 모티프는 항원성 단위와 타겟팅 단위를 연결시키는 데 기능할 뿐 아니라 본 발명에 따른 호모다이머성 단백질 내로 두 개의 모노머성 단백질들의 다이머화를 용이하게 하는 경첩부위를 포함한다.
- [0077] 본 명세서에서 사용되는 용어 "타겟팅 단위(targeting unit)"는 CD4+ T 세포들로 MHC 클래스 II-제한된 제시 또는 MHC 클래스 I 제한에 의해 CD8+ T 세포들로의 교차 제시(cross presentation)를 제공하기 위해 항원을 가진 단백질을 마우스 또는 인간 APC로 운반하는 단위를 의미한다. 본 발명에 따른 컨스트럭트들에서 이용되는 타겟팅 단위는 성숙한 LD78-베타로부터 유래되거나 또는 이와 동일하다.
- [0078] 본 명세서에서 사용되는 용어 "항원성 단위(antigenic unit)"는 T-세포들 위의 표면 수용체 같은 면역시스템의 항체 또는 다른 구성성분에 의해 특이적으로 인지될 수 있는 펩타이드 같은 어떠한 분자를 의미한다. 또한, 면역시스템을 유도할 수 있는 면역원들(immunogens)이 이러한 정의 안에 포함된다. 상기 용어들 "에피토프(epitope)" 또는 "항원성 에피토프"는 항원성 단위 내 짧은 펩타이드 서열에 의해 제공되는 분자적 표면 같은 독특한 분자적 표면을 언급하는 데 사용된다. 일부 구현예들에서 상기 항원성 단위는 2개 이상의 항원성 에피토프들을 포함한다. 본 발명에 따른 컨스트럭트들에서 이용되는 항원성 단위는 HPV, 예컨대 HPV16 또는 HPV18로부터 초기 유전자 산물들인 E6 및 E7로부터 유래되거나 또는 이와 동일하다.
- [0079] 상기 용어 "경첩부위(hinge region)"는, 예를 들어 체인 간 공유결합(들), 예컨대 이황화 브릿지(들)(disulfide bridge(s))의 형성을 통해 다이머화를 용이하게 하는 호모다이머성 단백질의 펩타이드 서열을 의미한다. 상기 경첩부위는 Ig로부터 유래될 수 있으며, 예를 들어 IgG3 같은 Ig의 경첩 엑손들(hinge exons) h1+h4로부터 유래될 수 있다.
- [0080] 본 발명의 특이적 구현예들:
- [0081] 상기 기재된 대로, 본 발명은 두 개의 동일한 아미노산 체인들의 호모다이머성 단백질에 관한 것으로, 각 아미노산 체인은 (1) 시그널 펩타이드, (2) 타겟팅 단위(unit), (3) 다이머화 모티프 및 (4) 항원성 단위를 포함하고, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열을 포함하고, 상기 항원성 단위는 인간 파필로마바이러스(HPV)의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 아미노산 서열을 포함하는 항원성 단위, 예컨대 HPV16 및/또는 HPV18의 초기 단백질 E6 및/또는 E7로부터 유래된 항원성 단위를 포함한다. 본 발명에 따른 일부 구현예들에서, 상기 아미노산 체인 내 상기 타겟팅 단위, 다이머화 모티프 및 항원성 단위는 N-말단부터 C-말단 순서로 타겟팅 단위, 다이머화 모티프 및 항원성 단위로 존재한다.
- [0082] 일부 구현예들에서, 본 발명에 따른 컨스트럭트에 이용되는 상기 항원성 단위는 HPV16, 예컨대 초기 단백질인 E6 및/또는 E7으로부터 유래된다.
- [0083] 일부 구현예들에서, 본 발명에 따른 컨스트럭트들에 이용되는 상기 항원성 단위는 HPV16의 E6로부터 유래된다.

- [0084] 일부 구현예들에서, 본 발명에 따른 컨스트럭트들에 이용되는 상기 항원성 단위는 HPV16의 E7로부터 유래된다.
- [0085] 일부 구현예들에서, 본 발명에 따른 컨스트럭트에 이용되는 상기 항원성 단위는 HPV18, 예컨대 초기 단백질인 E6 및/또는 E7으로부터 유래된다.
- [0086] 일부 구현예들에서, 본 발명에 따른 컨스트럭트들에 이용되는 상기 항원성 단위는 HPV18의 E6로부터 유래된다.
- [0087] 일부 구현예들에서, 본 발명에 따른 컨스트럭트들에 이용되는 상기 항원성 단위는 HPV18의 E7로부터 유래된다.
- [0088] 본 발명에 따른 일부 구현예들에서, 상기 시그널 펩타이드는 서열번호 1의 아미노산 서열 1-23에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0089] 본 발명에 따른 일부 구현예들에서, 상기 시그널 펩타이드는 서열번호 1의 아미노산 서열 1-23에 대해 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99%, 예컨대 최소 100% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0090] 본 발명에 따른 일부 구현예들에서, 상기 타겟팅 단위는 서열번호 1의 아미노산 서열 24-93에 대해 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99%, 예컨대 최소 100% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0091] 본 발명에 따른 일부 구현예들에서, 상기 다이머화 모티프는 경첩부위를 포함하고 선택적으로 링커를 통해 연결되는 면역글로불린 도메인 같은 다이머화를 용이하게 하는 다른 도메인을 선택적으로 포함한다.
- [0092] 본 발명에 따른 일부 구현예들에서, 상기 경첩부위는 IgG3로부터 유래되는 경첩부위 같은 Ig 유래된 경첩부위이다.
- [0093] 본 발명에 따른 일부 구현예들에서, 상기 경첩부위는 한 개, 두 개, 또는 여러 개의 공유결합들을 형성하는 능력을 가진다. 본 발명에 따른 일부 구현예들에서, 상기 공유결합은 이황화 브릿지이다.
- [0094] 본 발명에 따른 일부 구현예들에서, 상기 다이머화 모티프의 면역글로불린 도메인은 카르복시말단(carboxyterminal) C 도메인이거나, 또는 상기 C 도메인 또는 이의 변이체와 실질적으로 동일한 서열이다.
- [0095] 본 발명에 따른 일부 구현예들에서, 상기 카르복시말단 C 도메인은 IgG로부터 유래된다.
- [0096] 본 발명에 따른 일부 구현예들에서, 상기 다이머화 모티프의 면역글로불린 도메인은 호모다이머화시키는 능력을 가진다.
- [0097] 본 발명에 따른 일부 구현예들에서, 상기 면역글로불린 도메인은 비공유결합성 상호작용들을 통해 호모다이머화시키는 능력을 가진다. 본 발명에 따른 일부 구현예들에서, 상기 비공유결합성 상호작용들은 소수성 상호작용들이다.
- [0098] 본 발명에 따른 일부 구현예들에서, 상기 다이머화 도메인은 CH2 도메인을 포함하지 않는다.
- [0099] 본 발명에 따른 일부 구현예들에서, 상기 다이머화 모티프는 인간 IgG3의 C<sub>H</sub>3 도메인에 링커를 통해 연결되는 경첩 엑손들 h1 및 h4로 구성된다.
- [0100] 본 발명에 따른 일부 구현예들에서, 상기 다이머화 모티프는 서열번호 3의 아미노산 서열 94-237에 대해 최소 80% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0101] 본 발명에 따른 일부 구현예들에서, 상기 링커는 G<sub>3</sub>S<sub>2</sub>G<sub>3</sub>SG 링커이다.
- [0102] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위 및 다이머화 모티프는 링커, 예컨대 GLGGL 링커 또는 GLSGL 링커를 통해 연결된다.
- [0103] 본 발명에 따른 일부 구현예들에서, 상기 타겟팅 단위는 서열번호 1의 24-93의 아미노산들, 또는 이의 변이체로 구성된다.
- [0104] 본 발명에 따른 일부 구현예들에서, 상기 호모다이머성 단백질은 서열번호 1의 24-93의 아미노산들, 또는 이의 변이체로 구성된 타겟팅 단위를 가지는 동일한 호모다이머성 단백질의 친화도와 비교 시 CCR1, CCR3 및 CCR5로

부터 선택된 어느 하나의 케모카인 수용체에 대한 증가된 친화도를 가진다.

- [0105] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 3의 아미노산 서열 243-293에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함한다.
- [0106] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 3의 아미노산 서열 243-293에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0107] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 22의 F47, L50, C63, C106 및 I128로 구성된 리스트로부터 선택된 위치에서 하나 이상의 아미노산 치환들, 또는 서열번호 22의 Y43-L50으로 구성된 리스트로부터 선택된 하나 이상의 아미노산을 포함하는 결실을 포함한다.
- [0108] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 22에 대해 상대적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18 또는 20개 미만의 아미노산 치환들 및/또는 결실들을 포함한다.
- [0109] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 3, 서열번호 5, 서열번호 7, 또는 서열번호 9의 아미노산 서열 243-293, 또는 이의 변이체 또는 항원성 단편을 포함한다.
- [0110] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 3, 서열번호 5, 서열번호 7, 또는 서열번호 9의 아미노산 서열 243-293, 또는 이의 변이체 또는 항원성 단편으로 구성된다.
- [0111] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 11의 아미노산 서열 243-340에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함한다.
- [0112] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 11의 아미노산 서열 243-340에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0113] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 23의 C24, E26, C58, C61, C91, 및 C94로 구성된 리스트로부터 선택된 위치에서 하나 이상의 아미노산 치환들, 또는 서열번호 23의 L22-E26 및/또는 C58-C61 및/또는 C91-S95으로 구성된 리스트로부터 선택된 하나 이상의 아미노산을 포함하는 결실을 포함한다.
- [0114] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 23에 대해 상대적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18 또는 20개 미만의 아미노산 치환들 및/또는 결실들을 포함한다.
- [0115] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 11, 서열번호 13, 서열번호 15, 또는 서열번호 17의 아미노산 서열 243-340, 또는 이의 변이체 또는 항원성 단편을 포함한다.
- [0116] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 11, 서열번호 13, 서열번호 15, 또는 서열번호 17의 아미노산 서열 243-340, 또는 이의 변이체 또는 항원성 단편으로 구성된다.
- [0117] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호 34의 아미노산 서열 243-501에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함한다.
- [0118] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호

호 34의 아미노산 서열 243-501에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성된다.

- [0119] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 초기 단백질 E6 및 E7 모두로부터 유래된 인간 파필로마바이러스 16(HPV16)의 아미노산 서열을 포함한다.
- [0120] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 초기 단백질 E6 및 E7 모두로부터 유래된 인간 파필로마바이러스 18(HPV18)의 아미노산 서열을 포함한다.
- [0121] 본 발명에 따른 일부 구현예들에서, 상기 항원성 단위는 서열번호 22의 F47, L50G, C63, C106 및 I128T 그리고 서열번호 23의 C24, E26, C58, C61, C91, C94로 구성된 리스트로부터 선택된 위치에서 하나 이상의 아미노산 치환들을 포함한다.
- [0122] 본 발명에 따른 일부 구현예들에서 상기 항원성 단위는 서열번호 22 및 서열번호 23에 대해 상대적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18 또는 20개 미만의 아미노산 치환들 및/또는 결실들을 포함한다.
- [0123] 본 발명에 따른 일부 구현예들에서 상기 항원성 단위는 서열번호 19, 서열번호 21, 서열번호 32, 또는 서열번호 34의 아미노산 서열 243-501, 또는 이의 변이체 또는 항원성 단편으로 구성된다.
- [0124] 본 발명에 따른 일부 구현예들에서 상기 아미노산 체인은 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 서열번호 32, 및 서열번호 34로 구성되는 리스트로부터 선택된 아미노산 서열, 또는 이의 변이체 또는 항원성 단편으로 구성된다.
- [0125] 본 발명에 따른 일부 구현예들에서 상기 항원성 단위는 서열번호 22, 서열번호 23, 서열번호 24, 및 서열번호 25로부터 선택된 어느 하나의 아미노산 서열에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열을 포함한다.
- [0126] 본 발명에 따른 일부 구현예들에서 상기 항원성 단위는 서열번호 22, 서열번호 23, 서열번호 24, 및 서열번호 25로부터 선택된 어느 하나의 아미노산 서열에 대해 최소 80%, 예컨대 최소 81%, 예컨대 최소 82%, 예컨대 최소 83%, 예컨대 최소 84%, 예컨대 최소 85%, 예컨대 최소 86%, 예컨대 최소 87%, 예컨대 최소 88%, 예컨대 최소 89%, 예컨대 최소 90%, 예컨대 최소 91%, 예컨대 최소 92%, 예컨대 최소 93%, 예컨대 최소 94%, 예컨대 최소 95%, 예컨대 최소 96%, 예컨대 최소 97%, 예컨대 최소 98%, 예컨대 최소 99% 서열 동일성을 가지는 아미노산 서열로 구성된다.
- [0127] 일부 구현예들에서 본 발명에 따른 상기 호모다имер성 단백질은 어느 하나의 시그널 펩타이드 서열이 없는 성숙한 형태로 존재한다.
- [0128] 일부 구현예들에서 본 발명에 따른 상기 핵산 분자는 인간 코돈 최적화되어 있다.
- [0129] 본 발명에 따른 인간 코돈 최적화된 핵산 분자는 야생형 서열과 비교 시 하나 이상의 핵산 치환을 포함하는 것으로 이해될 것이고, 상기 치환은 인간 코딩 부위들에서 더 높은 빈도의 이용을 보이는 코돈을 위해 제공한다. 호모 사피엔스 내 코돈 사용의 빈도는 [http://biowiki.edu-wiki.org/en/codon\\_table](http://biowiki.edu-wiki.org/en/codon_table)에서 발견될 수 있다.
- [0130] 일부 구현예들에서 본 발명에 따른 상기 핵산 분자는 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 31 및 서열번호 33으로 구성된 리스트로부터 선택된 뉴클레오타이드 서열들 중 어느 하나, 또는 이의 변이체를 포함하고 있다.
- [0131] 일부 구현예들에서 본 발명에 따른 상기 핵산 분자는 박테리에 의해 포함된다.
- [0132] 일부 구현예들에서 본 발명에 따른 상기 핵산 분자는 환자에서 호모다имер성 단백질의 생산을 유도하기 위해 상기 환자에게 투여하기 위해 제형화된다.
- [0133] 일부 구현예들에서 본 발명에 따른 상기 백신은 약제학적으로 허용가능한 담체 및/또는 보조제(adjuvant)를 추가적으로 포함한다.

[0134] 일부 구현예들에서, 본 발명에 따라 환자에서 HPV에 의해 야기되는 암 또는 감염성 질병 같은 HPV 유도된 질병 또는 상태를 치료 또는 예방하는 방법은 본 발명에 따른 DNA 같은 핵산 분자를 이를 필요로 하는 상기 환자에게 투여투여한 후 전기천공의 이후 단계를 실시하는 단계를 포함한다. 일부 구현예들에서 상기 투여는 피부 내 또는 근육 내로 실시된다.

**도면의 간단한 설명**

[0135] 도 1: E7/E6 융합 항원과 함께 백신바디 백신들의 전체 구조. DNA 및 단백질 포맷들 모두가 제시되고 있다. 상기 백신바디는 세 개의 기능성 모듈들로 구성된다: 타겟팅 모듈 내 케모카인 인간 MIP-1 $\alpha$  (LD78 $\beta$ ), 다이머화 모듈 내 인간 IgG3로부터 유래된 경첩서열 및 CH3 서열 그리고 백신 모듈 내 전장 E7 및/또는 E6 융합.

도 2: HPV-유도된 악성종양들에 대한 백신바디 DNA 백신의 제시된 활성 모드. 백신바디를 인코딩하는 네이키드 DNA 플라스미드가 피부내 주입된 후 전기천공된다. 상기 플라스미드는 국소 세포들로 유입되고 백신바디 단백질들이 생산되어 분비된다. 화학주성 타겟팅 모듈들이 CCR1 및 CCR5 발현 항원 제시 세포들(APC)을 유인하고 수지상세포들(DC)과 결합하고 흡수되는 것을 보장한다. 상기 DC는 CD4+ 및 CD8+ T 세포들에 항원성 펩타이드들을 제시할 것이고 CD8+ T 세포들이 자궁 경관 내 HPV-감염된 및 HPV-형질전환된 세포들을 죽일 것이다.

도 3: 투여된 다른 양의 백신의 작용(function)으로서 E7 및 E6 특이적 T 세포 반응들의 수를 보여주는 ELISPOT 결과들. C57BL/6 마우스들에 VB1009 및 VB1016을 인코딩하는 네이키드 DNA 플라스미드들 및 이들에 상응하는 대조군들을 피부내 주입한 후 0일 및 7일째에 전기천공(Collectis, France)을 실시하였다. 비장세포들이 21일째에 수득되어 MHC 클래스 I-제한된 E7 또는 E6 펩타이드로 24시간 동안 자극되었다. IFN $\gamma$  분비성 비장세포들의 수가 ELISPOT에 의해 계산되었다. (A) 25  $\mu$ g의 VB1009, 대조군 1(항원 단독) 및 pUMVC4a(공백터)로 피부내 백신화시킨 후의 E7-특이적 반응들. (B) 12.5 및 1.4  $\mu$ g의 VB1016, 대조군 2(항원 단독) 및 pUMVC4a(공백터)로 피부내 백신화시킨 후의 E7-특이적 반응들. (C) 12.5 및 1.4  $\mu$ g의 VB1016, 대조군 2(항원 단독) 및 pUMVC4a(공백터)로 피부내 백신화시킨 후의 E6-특이적 반응들.

도 4. 측정된 종양 부피에 의해 보여지는 VB1016의 치료 효과.  $5 \times 10^5$  TC-1 세포들이 0일째에 C57BL/6 마우스들에 피하 주입되었다. 3일 및 10일째에, 상기 마우스들에 12.5  $\mu$ g의 VB1016을 인코딩하는 네이키드 DNA 플라스미드, 대조군 2 또는 공백터를 피부내 주입하고 전기천공(Collectis, France)을 실시하였다. 상기 종양 크기들이 캘리퍼스에 의해 한 주에 2 내지 3번씩 측정되고 종양 부피가 계산되었다.

도 5. 측정된 종양 부피에 의해 보여지는 VB1016의 치료 효과.  $5 \times 10^4$  TC-1 세포들이 0일째에 C57BL/6 마우스들의 목 부위에 피하 주입되었다. 3일, 7일 및 10일째에, 상기 마우스들에 20  $\mu$ g 또는 2  $\mu$ g의 VB1016을 인코딩하는 네이키드 DNA 플라스미드, 대조군 2 또는 공백터를 피부내 주입하고 전기천공(Collectis, France)을 실시하였다. 상기 종양 크기들이 캘리퍼스에 의해 한 주에 2 내지 3번씩 측정되고 종양 부피가 계산되었다.

도 6. 측정된 종양 부피에 의해 보여지는 VB1020 및 VB1021의 치료 효과.  $5 \times 10^4$  TC-1 세포들이 0일째에 C57BL/6 마우스들의 넓적다리 부위에 피하 주입되었다. 3일 및 10일째에, 상기 마우스들에 10  $\mu$ g의 VB1016, VB1020 또는 VB1021을 인코딩하는 네이키드 DNA 플라스미드, 또는 공백터를 피부내 주입하고 전기천공(Collectis, France)을 실시하였다. 상기 종양 크기들이 캘리퍼스에 의해 한 주에 2 내지 3번씩 측정되고 종양 부피가 계산되었다.

**발명을 실시하기 위한 구체적인 내용**

[0136] 실시예 1

[0137] 백신들의 구축 및 발현.

[0138] 유전자 서열들은 다음의 구조에 따라 디자인되었다: 1: 인간 LD78 b의 고유 리더 서열, 2: 전장 LD78b 서열. 3: IgG3로부터 유래되는 인간 경첩-부위 1. 4: IgG3로부터 유래되는 인간 경첩-부위 4. 5: 글라이신-세린 링커. 6: IgG3로부터 유래되는 인간 CH3 도메인. 7: 글라이신-루이신 링커. 8: 야생형 및 돌연변이 인간 파필로마 바이러스 발암유전자들 E6, E7 및 글라이신-세린 링커에 의해 분리되는 E6와 E7 모두로 구성된 융합 단백질들. 상기 컨스트럭트들은 E6 및/또는 E7 조성물에 따라 다음과 같이 지시된다:

[0139] VB1001: 백신바디-E6 야생형;

- [0140] VB1005: 백시바디-E7 야생형;
- [0141] 돌연변이들은 상응하는 고유 E6 또는 E7 서열 내 아미노산 위치에 따라 지시된다.
- [0142] VB1002: 백시바디-E6 C63R;
- [0143] VB1003: 백시바디-E6 C106R;
- [0144] VB1004: 백시바디-E6 F47R,C63R,C106R;
- [0145] VB1006: 백시바디-E7 C24G, E26G;
- [0146] VB1007: 백시바디-E7 C24G, E26G, C58G, C61G;
- [0147] VB1008: 백시바디-E7 C24G, E26G, C91G, C94G;
- [0148] VB1009: 백시바디-E7 C24G, E26G/ E6 F47R,C63R,C106R;
- [0149] VB1016: 백시바디-E7 C24G, E26G/ E6 C63R,C106R;
- [0150] VB1020: 백시바디-E7 C24G, E26G/ E6 F47R, C63R, C106R 인간 코돈 최적화됨
- [0151] VB1021: 백시바디-E7 C24G, E26G/ E6 F47R, L50G, C106R, I128T 인간 코돈 최적화됨
- [0152] 오직 항원들로만 구성된 대조군 백신들이 포함되었다:
- [0153] 대조군 1: E7 C24G, E26G/ E6 F47R,C63R,C106R;
- [0154] 대조군 2: E7 C24G, E26G/ E6 C63R,C106R
- [0155] 모든 유전자 서열들이 Aldevron(Fargo ND, USA) 또는 Eurofins MWG GmbH로부터 주문되어 발현벡터 pUMVC4a 내로 클로닝되었다.
- [0156] 모든 컨스트럭트들이 293E 세포들로 트랜스펙션되었고 온전한 백시바디 단백질들의 발현은 도트 블랏(dot blot) 및 ELISA에 의해 실시되어 확인되었다(데이터를 보이지 않음). 대조군 1 및 대조군 2를 제외한 모든 아미노산 서열들이 서열번호들로 제시되어 있다.
- [0157] 실시예 2.
- [0158] **면역반응 연구들**
- [0159] VB1009, VB1016, VB1020 및 VB1021이 각각 이에 상응하는 대조군 1 및 대조군 2와 함께 백신 후보들로 선택되었다. 음성 대조군으로서 공벡터 pUMVC4a가 이용되었다.
- [0160] 각 후보의 플라스미드 DNA 25, 12.5 및 1.4  $\mu$ g가 C57B1/6 마우스들의 등 아래 부분에 피부내 주입된 후 전기천공(Dermavax, Collectis, Paris, France)되었다. 7일 후에 상기 마우스들이 유사한 양의 백신들 및 대조군 플라스미드들로 촉진되었다. 21일째에 상기 마우스들이 희생되어 비장이 수득되었다.
- [0161] T 세포반응들이 ELISPOT에 의해 계산되었다(도 3a, 3b 및 3c).
- [0162] 실시예 3.
- [0163] **치료 효과**
- [0164] 상응하는 대조군 1 및 대조군 2와 함께 VB1016, VB1020 및 VB1021이 치료 백신 연구들을 위한 백신 후보로서 선택되었다.
- [0165]  $5 \times 10^4$  또는  $5 \times 10^5$  TC-1 세포들(Johns Hopkins University, Baltimore, USA, *Lin KY et al., Cancer Res, 1996*)이 C57B1/6 마우스들의 목 또는 넓적다리 부위에 주입되었다.
- [0166] 3일 및 10일 또는 3일, 7일 및 10일 후에, 상기 마우스들이 2  $\mu$ g, 10  $\mu$ g, 12.5  $\mu$ g 또는 20  $\mu$ g의 플라스미

드 DNA 플라스미드로 백신화되어 전기천공(Dermavax, Collectis France)을 실시하였다. TC-1 세포 주입 후 49 일까지, 중앙 크기들이 한 주에 2 내지 3번씩 측정되었다(도 4, 5 및 6).

- [0167] 실시예 4.
- [0168] 이용될 치료 DNA 백신은 GMP 세포 은행(cell banking), 약물 물질 및 약물 산물의 GMP 제조(manufacturing), ICH 안정성 연구들 및 DNA 백신의 Fill & Finish를 포함하는 규제기관의 가이드라인들에 따른 플라스미드 백신의 GMP 제조에 의해 준비될 수 있다. DNA 백신은 10 mM Tris, 1 mM EDTA 같은 염 용액에 2-5 mg/ml의 농도로 용해시켜 제형화될 수 있다. 상기 백신은 피부-내 또는 근육-내 중 어느 하나로 투여되고 이후에 전기천공을 실시하거나 또는 실시하지 않을 수 있다.
- [0169] 서열들:
- [0170] 시그널 펩타이드(굵은 글씨체의 아미노산 1-23) 및 성숙한 펩타이드(LD78-베타), 아미노산 24-93(서열번호 1)을 포함하는 C-C 모티프 케모카인 3-유사 1 전구체:
- [0171] **MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNF IADYFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA**
- [0172] 백신바디 HPV 컨스트럭트들의 특이적 DNA 및 상응하는 아미노산 서열들:
- [0173] **E6 또는 E7 단일 컨스트럭트들:**
- [0174] 오직 예시의 목적을 위해, 다른 도메인들의 컨스트럭트들은 다음의 순서에서의 도메인들과 함께 "|"로 분리된다: 시그널 펩타이드|인간 MIP-1 $\alpha$ |경첩h1|경첩h4|Gly-Ser 링커 또는 Gly-Leu 링커|hCH3 IgG3|Gly-Ser 링커 또는 Gly-Leu 링커|야생형 또는 돌연변이 전장 E6 또는 E7. 굵은 글씨의 아미노산들 또는 뉴클레오타이드들은 돌연변이들의 위치들을 예시한다.
- [0175] VB1001의 DNA 서열(서열번호 2):
- [0176] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTCTGCAACCATGGCTCTCTGCAACCAGGTCCTCTCT|GCACCACTTGCTGCTGACACGCCGACCCCTGCTGCTTCAGTACACCTCCCGACAGATTCCACAGAATTTATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCTTAACCAA GAGAGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTCAGTGACCTGGAGCTGAGTGCC|GAGCTCAAACCCCACTGGTGACA CAACTCACACA|GAGCCAAATCTGTGACACACCTCCCGGTGCCAAGGTGCCA|GGCGGTGAAGCAGCGGAGGTGGAAGTGA|GGACAGCCCGAG AACACAGGTGTACACCTGCCCCATCCCGGAGGAGATGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG TGGAGTGGGAGAGCAGCGGGCAGCCGAGACAACAACACACAGCCTCCATGTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGG ACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGTAA |GGCTCGGTGGCTG|ATGTTTCAGGACCCACAGGAGCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAAAACAATAACATGATATAATATT AGAATGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAATCCATATGCTGT ATGTGATAAATGTTAAAGTTTTATTCTAAAATTAGTGAGTATAGACATTATGTTATAGTTTGTATGGAACAACATTAGAACAGCAATAACAACAAACCGTT GTGTGATTTGTTAATTAGGTGTATTAAGTGTCAAAAAGCCACTGTGCTGAAAGAAAAGCAAAGACATCTGGACAAAAGCAAAGATTCCATAATATAAGGGG TCGGTGGACCGGTGATGTATGCTTGTGAGATCATCAAGAACACGTAGAGAAAACCCAGCTGTAA
- [0177] VB1001의 단백질 서열(E6와 함께 본 발명에 따른 호모다이머성 컨스트럭트, 서열번호 3): 아미노산 서열, 393개의 아미노산들.
- [0178] MQVSTAALAVLLCTMALCNQVLS|APLAADTPTACCFSYTSRQIPQNF IAD
- [0179] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA|ELKTPLG
- [0180] DTHT|EPKSCDTPPPCRCP|GGSSGGGSG|GQPREPQVYTLPPSREEMTK
- [0181] NQVSLTCLVKGFYPSDIAVEWESSQPENNYNTTPMLDSGSAFLYSKL
- [0182] TVDKSRWQGNIFSCSVMEALHNRFTQKLSLSPGK|GLGGL|MFQDPQER

- [0183] PRKLPQLCTELQTTIHD I ILECVYCKQQLLRREVYDFAFRDL C I VYRDGN
- [0184] PYAVCDKCLKFYSKI SEYRHYCYSLYGTTLEQQYNKPLCDLL IRC INCQK
- [0185] PLCPEEKQRHLDDKKQRFHN IRGRWTGRCMSCCRSSRTRRETQL\*
  
- [0186] VB1002의 DNA 서열(서열번호 4):
- [0187] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTCTGCACCATGGCTCTCTGCAACCAGGTCTCTCT | GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGCTACACCTCCCGACAGATTCCACAGAATTTTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCCTAACCAA  
 GAGAGGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTGACGTGACCTGGAGCTGAGTGCC | GAGCTCAAACCCCACTTGGTGACA  
 CAACTCACACA | GAGCCCAAATCTGTGACACACCTCCCGCTGCCAAGGTGCCA | GCGGTGGAAGCAGCGGAGGTGGAAGTGGA | GGACAGCCCGGAG  
 AACCACAGGTGTACACCTGCCCCATCCCGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG  
 TGGAGTGGGAGAGCAGCGGGCAGCCGAGAACAAC TACAACACCACGCTCCCATGCTGGACTCCGACGGCTCCTTCTCCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA  
 AA | GGCTCGGTGGCCTG | ATGTTTCAGGACCCACAGGAGCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAAACTATAACATGATATAATATT  
 AGAATGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAATCCATATGCTGT  
 ACGAGATAAATGTTTAAAGTTTTATTCTAAAATTAGTGAGTATAGACATTATGTTATAGTTGTATGGAACAACATTAGAACAGCAATAACAACAAACCGTT  
 GTGTGATTTGTTAATTAGGTGTATTAACGTCAAAAGCCACTGTGCTGAAGAAAAGCAAAGACATCTGGACAAAAGCAAAGATTCCATAATATAAGGGG  
 TCGGTGACCGGTCGATGTATGTCTTGTGTCAGATCATCAAGAACACGTAGAGAAAACCCAGCTGTAA
- [0188] VB1002의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 5): 아미노산 서열, 393개의 아미노산들.
- [0189] MQVSTAALAVLLCTMALCNQVLS | APLAADTPTACCF SYTSRQIPQNF IAD
- [0190] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA | ELKTPLG
- [0191] DTTHT | EPKSCDTPPPCPRCP | GGGSSGGGSG | GQPREPQVYTLPPSREEMTK
- [0192] NQVSL TCLVKGFYPSDIAVEWESSQPENNYNTTPMLSDSGSFFLYSKL
- [0193] TVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK | GLGGL | MFQDPQER
- [0194] PRKLPQLCTELQTTIHD I ILECVYCKQQLLRREVYDFAFRDL C I VYRDGN
- [0195] PYAVRDKCLKFYSKI SEYRHYCYSLYGTTLEQQYNKPLCDLL IRC INCQK
- [0196] PLCPEEKQRHLDDKKQRFHN IRGRWTGRCMSCCRSSRTRRETQL\*
  
- [0197] VB 1003의 DNA 서열(서열번호 6):
- [0198] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTCTGCACCATGGCTCTCTGCAACCAGGTCTCTCTCT | GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGCTACACCTCCCGACAGATTCCACAGAATTTTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCCTAACCAA  
 GAGAGGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTGACGTGACCTGGAGCTGAGTGCC | GAGCTCAAACCCCACTTGGTGACA  
 CAACTCACACA | GAGCCCAAATCTGTGACACACCTCCCGCTGCCAAGGTGCCA | GCGGTGGAAGCAGCGGAGGTGGAAGTGGA | GGACAGCCCGGAG  
 AACCACAGGTGTACACCTGCCCCATCCCGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG  
 TGGAGTGGGAGAGCAGCGGGCAGCCGAGAACAAC TACAACACCACGCTCCCATGCTGGACTCCGACGGCTCCTTCTCCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA  
 AA | GGCTCGGTGGCCTG | ATGTTTCAGGACCCACAGGAGCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAAACTATAACATGATATAATATT  
 AGAATGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAATCCATATGCTGT  
 ATGTGATAAATGTTTAAAGTTTTATTCTAAAATTAGTGAGTATAGACATTATGTTATAGTTGTATGGAACAACATTAGAACAGCAATAACAACAAACCGTT  
 GTGTGATTTGTTAATTAGGTGTATTAACCGA CAAAAGCCACTGTGCTGAAGAAAAGCAAAGACATCTGGACAAAAGCAAAGATTCCATAATATAAGGGG  
 TCGGTGACCGGTCGATGTATGTCTTGTGTCAGATCATCAAGAACACGTAGAGAAAACCCAGCTGTAA
- [0199] VB1003의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 7): 아미노산 서열, 393개의 아미노산들.

- [0200] MQVSTAALAVLLCTMALCNQVLS|APLAADTPTACCFSYTSRQIPQNF IAD
- [0201] YFETSSQCSKPSVIFLTRGRQVCADPSEEWVQKYVSDLELSA|ELKTPLG
- [0202] DTTHT|EPKSCDTPPPCPRCP|GGGSSGGGSG|GQPREPQVYTLPPSREEMTK
- [0203] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLSDGSAFFLYSKL
- [0204] TVDKSRWQQGNIFSCVMHEALHNRFTQKLSLSPGK|GLGGL|MFQDPQER
- [0205] PRKLPQLCTELQTTIHDII ILECVYCKQQLLRREVYDFAFRDLCI VYRDGN
- [0206] PYAVCDKCLKFYSKI SEYRHYCYSLYGTTLQYQYKPLCDLL IRCINRQK
- [0207] PLCPEEKQRHLDDKQRFHNI RGRWTGRCMCCRSSRTRRETQL\*

[0208] VB1004의 DNA 서열(서열번호 8):

[0209] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTGCAACCATGGCTCTCTGCAACCATGGCTCTCT|GCACCACTGCTGCTGACACGCCGACCGCTGCTGCTTACAGCTACACCTCCCGACAGATTCCACAGAATTTATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCTTAACCAAGAGAGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTGACGTGACCTGGAGCTGAGTGCC|GAGCTAAAACCCCACTTGGTGACA CAACTCACACA|GAGCCCAAATCTTGTGACACACCTCCCCGTGCCAAGGTGCCA|GGCGGTGGAAGCAGCGGAGGTGGAAGTGA|GGACAGCCCGGAG AACACAGGTGTACACCTGCCCCATCCCGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGGCTGCAAGGCTTCTACCCAGCGACATCGCCG TGGAGTGGGAGAGCAGCGGGCAGCCGAGAAACAACACTACAACACCACGCTCCCATGTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGG ACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA AA|GGCCTCGGTGGCCTG|ATGTTTCAGGACCCACAGGAGCAGCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAAAACAATAACATGATATAATATT AGAATGTGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTCGACGGGATTTATGCATAGTATATAGAGATGGGAATCCATATGCTGT ACGAGATAAATGTTAAAGTTTTATTCTAAAATTAGTGAGTATAGACATTATGTTATAGTTGTATGGAACAACATTAGAACAGCAATAACAACAACCGTT GTGTGATTTGTAATTAGGTGATTAACCGA AAAAGCCACTGTGCTGAAGAAAAGCAAAGACATCTGGACAAAAGCAAAGATTCCATAATATAAGGGG TCGGTGGACCGGTGATGTATGTCTTGTGAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAA

[0210] VB1004의 단백질 서열(본 발명에 따른 호모다이어머성 컨스트럭트, 서열번호 9): 아미노산 서열, 393개의 아미노 산들.

- [0211] MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNF IAD
- [0212] YFETSSQCSKPSVIFLTRGRQVCADPSEEWVQKYVSDLELSAELKTPLG
- [0213] DTTHTPEKSCDTPPPCPRCPGGSSGGGSGGQPREPQVYTLPPSREEMTK
- [0214] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLSDGSAFFLYSKL
- [0215] TVDKSRWQQGNIFSCVMHEALHNRFTQKLSLSPGKGLGGLMFQDPQER
- [0216] PRKLPQLCTELQTTIHDII ILECVYCKQQLLRREVYDFAFRRDLCI VYRDGN
- [0217] PYAVRDKCLKFYSKI SEYRHYCYSLYGTTLQYQYKPLCDLL IRCINRQK
- [0218] PLCPEEKQRHLDDKQRFHNI RGRWTGRCMCCRSSRTRRETQL\*

[0219] VB1005의 DNA 서열(서열번호 10):

[0220] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTGCAACCATGGCTCTCTGCAACCATGGCTCTCT|GCACCACTGCTGCTGACACGCCGACCGCTGCTGCTTACAGCTACACCTCCCGACAGATTCCACAGAATTTATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCTTAACCAAGAGAGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTGACGTGACCTGGAGCTGAGTGCC|GAGCTAAAACCCCACTTGGTGACA CAACTCACACA|GAGCCCAAATCTTGTGACACACCTCCCCGTGCCAAGGTGCCA|GGCGGTGGAAGCAGCGGAGGTGGAAGTGA|GGACAGCCCGGAG AACACAGGTGTACACCTGCCCCATCCCGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGGCTGCAAGGCTTCTACCCAGCGACATCGCCG TGGAGTGGGAGAGCAGCGGGCAGCCGAGAAACAACACTACAACACCACGCTCCCATGTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGG ACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA

AA|GGCCTCGGTGGCCTG|ATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTGCAACCAGAGACAACCTGATCTCTACTGTTATGAGCAATT  
 AAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCAGGACAGAGCCATTACAATATTGTAACCTTTTGTGCAAGTGTGA  
 CTCTACGCTTCGGTTGTGCGTACAAAGCACACACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCATCTGTTCTCA  
 GAAACCATAA

[0221] VB1005의 단백질 서열(E7 함께 본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 11): 아미노산 서열, 340개의 아미노산들.

[0222] MQVSTAALAVLLCTMALCNQVLS|APLAADTPTACCFSYTSRQIPQNF IAD

[0223] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA|ELKTPLG

[0224] DTTHT|EPKSCDTPPPCPRCP|GGGSSGGGSG|GQPREPQVYTLPPSREEMTK

[0225] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLSDGSFFLYSKL

[0226] TVDKSRWQGNIFSCVMHEALHNRFTQKLSLSPGK|GLGGL|MHGDTPTL

[0227] HEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFC

[0228] CKCDSTLRLCVQSTHVDIRTLEDLLMGTGLIVCPICSQKP\*

[0229] VB1006의 DNA 서열(서열번호 12):

[0230] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTCTGCACCATGGCTCTCTGCAACCAGGTCCTCTCT|GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGTACACCTCCCGACAGATTCCACAGAATTTATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCTTAACCAA  
 GAGAGCCCGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTACGTGACCTGGAGCTGAGTGCC|GAGCTCAAACCCCACTTGGTGACA  
 CAACTCACACA|GAGCCCAAATCTTGTGACACACCTCCCCCGTGCCCAAGGTGCCCA|GGCGGTGGAAGCAGCGGAGGTGGAAGTGA|GGACAGCCCGGAG  
 AACCACAGGTGTACACCTGCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGCCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG  
 TGGAGTGGGAGAGCAGCGGGCAGCCGAGACAACAACACACCGCTCCATGTGGACTCCGACGGCTCCTTCTCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA  
 AA|GGCCTCGGTGGCCTG|ATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTGCAACCAGAGACAACCTGATCTCTACGGATATGGACAATT  
 AAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCAGGACAGAGCCATTACAATATTGTAACCTTTTGTGCAAGTGTGA  
 CTCTACGCTTCGGTTGTGCGTACAAAGCACACACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCATCTGTTCTCA  
 GAAACCATAA

[0231] VB1006의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 13): 아미노산 서열, 340개의 아미노산들.

[0232] MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNF IAD

[0233] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSAELKTPLG

[0234] DTTHTPEKSCDTPPPCPRCPGGSSGGGSGGQPREPQVYTLPPSREEMTK

[0235] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLSDGSFFLYSKL

[0236] TVDKSRWQGNIFSCVMHEALHNRFTQKLSLSPGKGLGGLMHGDTPTL

[0237] HEYMLDLQPETTDLYGQQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFC

[0238] CKCDSTLRLCVQSTHVDIRTLEDLLMGTGLIVCPICSQKP\*

[0239] VB1007의 DNA 서열(서열번호 14):

[0240] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTCTGCACCATGGCTCTCTGCAACCAGGTCCTCTCT|GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGTACACCTCCCGACAGATTCCACAGAATTTATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCTTAACCAA  
 GAGAGCCCGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTACGTGACCTGGAGCTGAGTGCC|GAGCTCAAACCCCACTTGGTGACA  
 CAACTCACACA|GAGCCCAAATCTTGTGACACACCTCCCCCGTGCCCAAGGTGCCCA|GGCGGTGGAAGCAGCGGAGGTGGAAGTGA|GGACAGCCCGGAG  
 AACCACAGGTGTACACCTGCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGCCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG

TGGAGTGGGAGAGCAGCGGGCAGCCGAGAGAACTACAACACCAGCCTCCCATGCTGGACTCCGACGGCTCCTTCTCCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA  
 AA|GGCCTCGGTGGCCTG|ATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACCTGATCTCTACGGATATGGACAATT  
 AAATGACAGCTCAGAGGAGGAGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTACAATATTGTAACCTTTGGATGCAAGGGAGA  
 CTCTACGCTTCGGTGTGCGTACAAAGCACACAGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCATCTGTTCTCA  
 GAAACCATAA

[0241] VB1007의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 15): 아미노산 서열, 340개의 아미노산들.

[0242] MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNF IAD

[0243] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSAELKTPLG

[0244] DTTHTPEKSCDTPPPCPRCPGGSSGGGSGGQPREPVYTLPPSREEMTK

[0245] NQVSLTCLVKGFYPSDIAVEWESSQPENNYNTTPMLSDGSAFFLYSKL

[0246] TVDKSRWQQGNIFSCVMHEALHNRFTQKLSLSPGKGLGGLMHGDTPTL

[0247] HEYMLDLQPETTDLYGYGQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFG

[0248] CKGDSTLRLCVQSTHVDIRTLEDLLMGTGIVCPICSQKP\*

[0249] VB1008의 DNA 서열(서열번호 16):

[0250] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTGACCATGGCTCTGCAACCAGGTCCTCTCT|GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGTACACCTCCCGACAGATTCACAGAAATTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGCATCTTCTTAACCAA  
 GAGAGCCCGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTCAGTGACCTGGAGCTGAGTGCC|GAGCTCAAAACCCACTTGGTGACA  
 CAACTCACACA|GAGCCAAATCTGTGACACACCTCCCCCGTGCCCAAGGTGCCA|GGCGGTGGAAGCAGCGGAGGTGGAAGTGA|GGACAGCCCGAG  
 AACCACAGGTGTACACCTGCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTACCCAGCGACATCGCCG  
 TGGAGTGGGAGAGCAGCGGGCAGCCGAGAGAACTACAACACCAGCCTCCCATGCTGGACTCCGACGGCTCCTTCTCCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA  
 AA|GGCCTCGGTGGCCTG|ATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACCTGATCTCTACGGATATGGACAATT  
 AAATGACAGCTCAGAGGAGGAGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTACAATATTGTAACCTTTGTGCAAGTGTGA  
 CTCTACGCTTCGGTGTGCGTACAAAGCACACAGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGGACCCATCGGATCTCA  
 GAAACCATAA

[0251] VB1008의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 17): 아미노산 서열, 340개의 아미노산들.

[0252] MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNF IAD

[0253] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSAELKTPLG

[0254] DTTHTPEKSCDTPPPCPRCPGGSSGGGSGGQPREPVYTLPPSREEMTK

[0255] NQVSLTCLVKGFYPSDIAVEWESSQPENNYNTTPMLSDGSAFFLYSKL

[0256] TVDKSRWQQGNIFSCVMHEALHNRFTQKLSLSPGKGLGGLMHGDTPTL

[0257] HEYMLDLQPETTDLYGYGQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFC

[0258] CKCDSTLRLCVQSTHVDIRTLEDLLMGTGIVGPIGSQKP\*

[0259] E6 및 E7을 포함하는 컨스트럭트들:

[0260] 오직 예시의 목적을 위해, 다른 도메인들의 컨스트럭트들은 다음의 순서의 도메인들과 함께 "|"로 분리된다: 시그널 펩타이드|인간 MIP-1 α|경첩h1|경첩h4|Gly-Ser 링커 또는 Gly-Leu 링커|hCH3 IgG3|Gly-Ser 링커 또는 Gly-Leu 링커|E7 돌연변이|Gly-Ser 링커 또는 Gly-Leu 링커|E6 돌연변이. 굵은 글씨의 아미노산들 또는 뉴클레

오타이드들은 돌연변이들의 위치들을 예시한다.

[0261] VB1009의 DNA 서열(서열번호 18):

[0262] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTGCACCATGGCTCTCTGCAACCAGGTCCTCTCT | GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGCTACACCTCCCGACAGATTCCACAGAATTTTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGCATCTTCTTAACCAA  
 GAGAGGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTACGTGACCTGGAGCTGAGTGCC | GAGCTCAAAACCCCACTGGTGACA  
 CAACTCACACA | GAGCCCAAATCTTGTGACACACCTCCCCCGTGCCCAAGGTGCCCA | GCGGTGGAAGCAGCGGAGGTGGAAGTGA | GGACAGCCCGAG  
 AACCACAGGTGTACACCTGCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG  
 TGGAGTGGGAGAGCAGCGGGCAGCCGAGAACAAC TACAACACCACGCTCCCATGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGTA  
 AA | GGCTCGGTGGCCTG | ATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTGCAACCAGAGACAACCTGATCTCTACGGATATGGACAATT  
 AAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCAGGACAGAGCCATTACAATATTGTAACCTTTTGTGCAAGTGGA  
 CTCTACGCTTCGGTTGTGCGTACAAAGCACACGTAGACATTCGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTTGTGTGCCCATCTGTTCTCA  
 GAAACA | GCGGTGGAAGCAGCGGAGGTGGAAGTGA | ATGTTTCAGGACCCAGGAGCGACCCAGAAAGTACCACAGTTATGCACAGAGCTGCAAAACA  
 ACTATACATGATATAATATTAGAATGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTCGACGGGATTTATGCATAGTATATAGA  
 GATGGGAATCCATATGCTGTACGAGATAAATGTTTAAAGTTTTATTCTAAAATTAGTGAGTATAGACATTATTGTTATAGTTTGTATGGAACAACATTAGAA  
 CAGCAATACAACAAACCGTTGTGTGATTGTTAATTAGGTGATTAACCGACAAAAGCCACTGTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAA  
 AGATTCCATAATATAAGGGTCGGTGGACCGGTGATGTATGCTTGTGTCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAA

[0263] VB1009의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 서열번호 19): 아미노산 서열, 501개의 아미노산들.

[0264] MQVSTAALAVLLCTMALCNQVLS | APLAADTPTACCFSYTSRQIPQNF IAD  
 [0265] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA | ELKTPLG  
 [0266] DTTHT | EPKSCDTPPPCPRCP | GGGSSGGSG | GQPREPQVYTLPPSREEMTK  
 [0267] NQVSLTCLVKGFYPSDI AVEWESSGQPENNYNTTPMLDSGDSFFLYSKL  
 [0268] TVDKSRWQQGNIFSCVMHEALHNRFTQKSLSLSPGK | GLGGL | MHGDTPTL  
 [0269] HEYMLDLQPETDLYGYGQLNDSSEEEDEIDGPAGQAEPRAHYNI VTFC  
 [0270] CKCDSTLRLCVQSTHVDIRTELDLLMGLGIVCPICSQKP | GGGSSGGSG |  
 [0271] MFQDPQERPRKLPQLCTELQTTI HDI ILECVYCKQLLRREVDFAARDL  
 [0272] CIVYRDGNPYAVRDKCLKFYSKI SEYRHYCYSLYGTITLEQQYNKPLCDLL  
 [0273] IRCINRQKPLCPPEEKQRHLDDKQRFHNIRGRWTGRCMCCRSSRTRRETQ  
 [0274] L\*

[0275] VB1016의 DNA 서열(서열번호 20):

[0276] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCTCTGCACCATGGCTCTCTGCAACCAGGTCCTCTCT | GCACCACTTGCTGCTGACACGCCGACCGCCTG  
 CTGCTTCAGCTACACCTCCCGACAGATTCCACAGAATTTTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGCATCTTCTTAACCAA  
 GAGAGGCCGGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTACGTGACCTGGAGCTGAGTGCC | GAGCTCAAAACCCCACTGGTGACA  
 CAACTCACACA | GAGCCCAAATCTTGTGACACACCTCCCCCGTGCCCAAGGTGCCCA | GCGGTGGAAGCAGCGGAGGTGGAAGTGA | GGACAGCCCGAG  
 AACCACAGGTGTACACCTGCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTACCCAGCGACATCGCCG  
 TGGAGTGGGAGAGCAGCGGGCAGCCGAGAACAAC TACAACACCACGCTCCCATGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGG  
 ACAAGAGCAGGTGGCAGCAGGGAACATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGTA  
 AA | GGCTCGGTGGCCTG | ATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTGCAACCAGAGACAACCTGATCTCTACGGATATGGACAATT  
 AAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCAGGACAGAGCCATTACAATATTGTAACCTTTTGTGCAAGTGGA

CTCTACGCTTCGGTTGTGCGTACAAAGCACACACGTAGACATTCGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCA  
 GAAACCA|GGCGGTGGAAGCAGCGGAGGTGGAAGTGA|ATGTTTCAGGACCCACAGGAGCGACCCAGAAAAGTTACCACAGTTATGCACAGAGCTGCAAACA  
 ACTATACATGATATAATATTAGAATGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTTTTCGGGATTATGCATAGTATATAGA  
 GATGGGAATCCATATGCTGTACGAGATAAAATGTTTAAAGTTTTATTCTAAAATTAGTGAGTATAGACATTATTGTTATAGTTTGTATGGAACAACATTAGAA  
 CAGCAATACAACAAACCGTTGTGTGATTGTAAATAGGTGATTAAACCGACAAAAGCCACTGTGTCTGAAGAAAAGCAAAGACATCTGGACAAAAGCAA  
 AGATTCCATAATATAAGGGGTCGGTGGACCGGTGATGTCTTGTGTCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAA

[0277] VB1016의 단백질 서열(본 발명에 따른 호모다이머성 퀴스트럭트, 서열번호 21): 아미노산 서열, 501개의 아미노산들

[0278] MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFYSYTSRQIPQNF IAD

[0279] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSAELKTPLG

[0280] DTHTEPKSCDTPPPCPRCPGGSSGGGSGGQPREPQVYTLPPSREEMTK

[0281] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLSDGSFFLYSKL

[0282] TVDKSRWQGNIFSCSVMHEALHNRFTQKSLSLSPGKGLGGLMHGDTPTL

[0283] HEYMLDLQPETTDLYGYQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFC

[0284] CKCDSTLRLCVQSTHVDIRTLEDLLMGTGIVCPICSQKPGGGSSGGGSG

[0285] MFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVDFAFRDL

[0286] CIVYRDGNPYAVRDKCLKFYSKI SEYRHYCYSLYGTTLEQQYNKPLCDLL

[0287] IRCINRQKPLCPEEKQRHLDKKQRFHNIIRGRWTGRCMSCCRSSRTRRETQ

[0288] L\*

[0289] 서열번호 22:

[0290] >tr|Q778I6|Q778I6\_HP16 E6 단백질 OS=인간 파필로마바이러스 타입 16 GN=E6 PE=4 SV=1; (밑줄친 아미노산들은 결실될 아미노산들을 지시한다; 돌연변이될 수 있는 잠재적인 아미노산들은 강조된다)

[0291] MFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVDFAFRDLICIVYRDGNPYAVCDKCLKFYSKI SEYRHYCYSLYGTTLEQQYNKPLCDLLIR  
 CINQKPLCPEEKQRHLDKKQRFHNIIRGRWTGRCMSCCRSSRTRRETQL

[0292] 서열번호 23:

[0293] >sp|P03129|VE7\_HP16 단백질 E7 OS=인간 파필로마바이러스 타입 16 GN=E7 PE=1 SV=1; (밑줄친 아미노산들은 결실될 아미노산들을 지시한다; 돌연변이될 수 있는 잠재적인 아미노산들은 강조된다)

[0294] MHGDTPTLHEYMLDLQPETTDLYCYEQNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTGIVCPICSQKPL

[0295] 서열번호 24:

[0296] >sp|P06463|VE6\_HP18 단백질 E6 OS=인간 파필로마바이러스 타입 18 GN=E6 PE=1 SV=1

[0297] MARFEDPTRRYPKLPDLCTELNTSLQDIEITCVYCKTVLELTVFVFAFKDLFVVYRDSI

[0298] PHAACHKCIDFYSRIRELRHYSVYGDITLEKLTNTGLYNLLIRCLRCQKPLNPAEKL RH

[0299] LNEKRRFHNIAHYRGQCHSCCNRRARQERLQRRRETQV

- [0300] 서열번호 25:
- [0301] >sp|P06788|VE7\_HP18 단백질 E7 OS=인간 파필로마바이러스 타입 18 GN=E7 PE=3 SV=2
- [0302] MHGPKATLQDIVLHLEPQNEIPVDLLCHEQLSDSEEENDEIDGVNHQHLPARRAEPQRHT
- [0303] MLCMCCKCEARIKLVVSSADDLRAFQQLFLNLSFVCPWCASQQ
- [0304] 서열번호 26:
- [0305] 경첩부위들 (IgG3 UH 경첩), 12개의 아미노산들: ELKTPLGDTTHT
- [0306] 서열번호 27:
- [0307] 경첩부위 (IgG3, MH 경첩, 15개의 아미노산들): EPKSCDTPPPCRCP
- [0308] 서열번호 28:
- [0309] Gly-Ser 링커: GGSSSGGGSG
- [0310] 서열번호 29: hCH3 IgG3:
- [0311] GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGPENNYNTPPMLDSGDSFFLYSKLTVDKSRWQQGNIFSCSVMEALHNRFQKLSLS  
LSPGK
- [0312] 서열번호 30: 링커: GLGGL
- [0313] 서열번호 31: VB1020의 DNA 서열:
- [0314] ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTCTGCACCATGGCTCTCTGCAACCAGGTCTCTCT | GCACCACTTGCTGCTGACACGCCGACCCGCTG  
CTGCTTCAGTACACCTCCCACAGATTCCACAGAAATTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCCTAACCAA  
GAGAGCCCGCAGGTCTGTGCTGACCCAGTGAGGAGTGGGTCCAGAAATACGTGACGTGACCTGGAGCTGAGTGCC | GAGCTCAAACCCCACTTGGTGACA  
CAACTCACACA | GAGCCCAAATCTGTGACACACCTCCCCCGTGCCCAAGGTGCCCA | GCGGTGGAAGCAGCGGAGGTGGAAGTGA | GGACAGCCCGAG  
AACCACAGGTGTACACCTGCCCCATCCCGGAGGAGATGACCAAGAACCAGGTGACGCTGACCTGCCTGGTCAAAGGCTTACCCACGACATCGCCG  
TGGAGTGGGAGAGCAGCGGCAGCCGAGAACAAC TACAACACCACGCTCCCATGCTGGACTCCGACGGCTCCTTCTCCTCTACAGCAAGCTCACCGTGG  
ACAAGAGCAGGTGGCAGCAGGGAAATCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCGCTTACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA  
AA | GGCTCGGTGGCTG / ATGCATGGCGATACCCCAACTCCATGAGTACATGCTGGACCTTACGCCGAGACTACGGATCTGTATGGCTATGGCAGTT  
GAATGACTCATCTGAGGAGGAGACGAAATAGACGGCCAGCTGGTCAAGCCGAACCGATAGAGCCACTACAACATTGTGACCTTTGTGTAAAGTGTGA  
CAGCACTCTGAGACTGTGTGTTCACTCCACTCATGTCGACATACGCACATTGGAGGATCTCTGATGGGAACACTGGGAATTGTGTGCCATCTGTCCCA  
AAAGCCT/GGAGGTGGAAGCAGTGGAGGCGTTTCAGGC / ATGTTCAAGATCCTCAAGAACGTCCTCGTAAGCTGCCACAGCTGTGTACCGAGCTTCAGACC  
ACCATTCACGACATCATCTGGAGTGCCTTATTGCAAACAGCAGCTCCTTAGAAGGGAAGTGTACGATTTTGCACGGAGGGACCTCTGCATCGTGTATCGG  
GACGGCAATCCCTATGCGGTACGGGATAAATGCCTGAAGTTCTACAGCAAAATCTCCGAGTACCGGCACTACTGCTACTCTCTATATGGGACGACTCTGGAA  
CAGCAGTACAACAAGCCCTGTGCGATCTGTGATTTCGCTGCATTAATCGCCAGAAACCTCTGTGCCAGAAGAGAAGCAAAGACACCTGGACAAGAAACAG  
CGATTCCACAACATCCGAGGAGATGGACAGGGAGGTGTATGAGCTGCTGTCGGAGTTCTAGGACAAGGCGGAAACCCAGCTTTGA
- [0315] 서열번호 32: VB1020의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트, 아미노산 서열, 501개의 아미노  
산들:

[0316] MQVSTAALAVLLCTMALCNQVLS|APLAADTPTACCFSYTSRQIPQNF IAD  
 [0317] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA|ELKTPLG  
 [0318] DTTHT|EPKSCDTPPPCPRCP|GGGSSGGGSG|GQPREPQVYTLPPSREEMTK  
 [0319] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLDSDGSFFLYSKL  
 [0320] TVDKSRWQQGNIFSCVMHEALHNRFTQKLSLSPGK|GLGGL|MHGDTPTL  
 [0321] HEYMLDLQPETTDLYGYQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFC  
 [0322] CKCDSTLRLCVQSTHVDIRTLEDLLMGTGIVCPICQKP|GGGSSGGGSG|  
 [0323] MFQDPQERPRKLPQLCTELQTTIHDIIIECVYCKQQLLRREVDFA**R**RDL  
 [0324] CIVYRDGNPYAV**R**DKCLKFYSKI SEYRHYCYSLYGTTLQYQNKPLCDLL  
 [0325] IRCIN**R**QKPLCPEEKQRHLDDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQ  
 [0326] L\*

[0327] 서열번호 33: VB1021의 DNA 서열:

[0328] CA/GGCGGAGGATCTTCAGGAGCGGGAGTGGC/ATGTTCCAAGACCCTCAGGAACGCCCTCGGAAACTGCCCAATTGTGTACTGAGCTCCAGACAACGAT  
 ACACGACATAATCCTGGAGTGCCTGTATTGCAAGCAGCAGCTTCTGAGGAGGAAGTGTACGATTTTCCAGGAGAGATGGCTGCATTGTCTACCGAGATGG  
 CAATCCCTATGCGGTGTGTGATAAGTGTCTGAAGTCTATTCCAAAATCAGCGAATATCGGCATTATTGCTACTACTGTACGAACTACCTCGAACAGCA  
 GTACAACAAACCGCTCTGTGATCTGCTGATCAGATGCATCAATCGGCAGAAACCCCTTTGTCCCGAAGAGAAGCAAAGACACCTGGACAAGAAGCAGAGGTT  
 CCACAATACCGAGGTCGTTGGACTGGGCGCTGCATGTCCTGTGTGCTCCTCTCGCACAAAGGAGAGAGACACAACCTGTGA

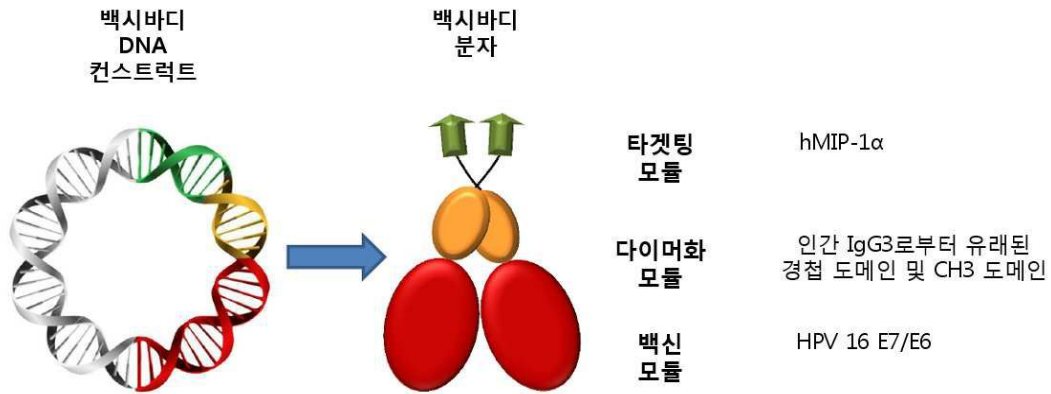
[0329] 서열번호 34: VB1021의 단백질 서열(본 발명에 따른 호모다이어성 컨스트럭트. 아미노산 서열, 501개의 아미노산들):

[0330] MQVSTAALAVLLCTMALCNQVLS|APLAADTPTACCFSYTSRQIPQNF IAD  
 [0331] YFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA|ELKTPLG  
 [0332] DTTHT|EPKSCDTPPPCPRCP|GGGSSGGGSG|GQPREPQVYTLPPSREEMTK  
 [0333] NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPMLDSDGSFFLYSKL  
 [0334] TVDKSRWQQGNIFSCVMHEALHNRFTQKLSLSPGK|GLGGL|MHGDTPTL  
 [0335] HEYMLDLQPETTDLYGYQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFC  
 [0336] CKCDSTLRLCVQSTHVDIRTLEDLLMGTGIVCPICQKP|GGGSSGGGSG|  
 [0337] MFQDPQERPRKLPQLCTELQTTIHDIIIECVYCKQQLLRREVDFA**R**RD**G**  
 [0338] CIVYRDGNPYAVCDKCLKFYSKI SEYRHYCYSLYGTTLQYQNKPLCDLL  
 [0339] IRCIN**R**QKPLCPEEKQRHLDDKKQRFHN**T**RGRWTGRCMSCCRSSRTRRETQ  
 [0340] L\*

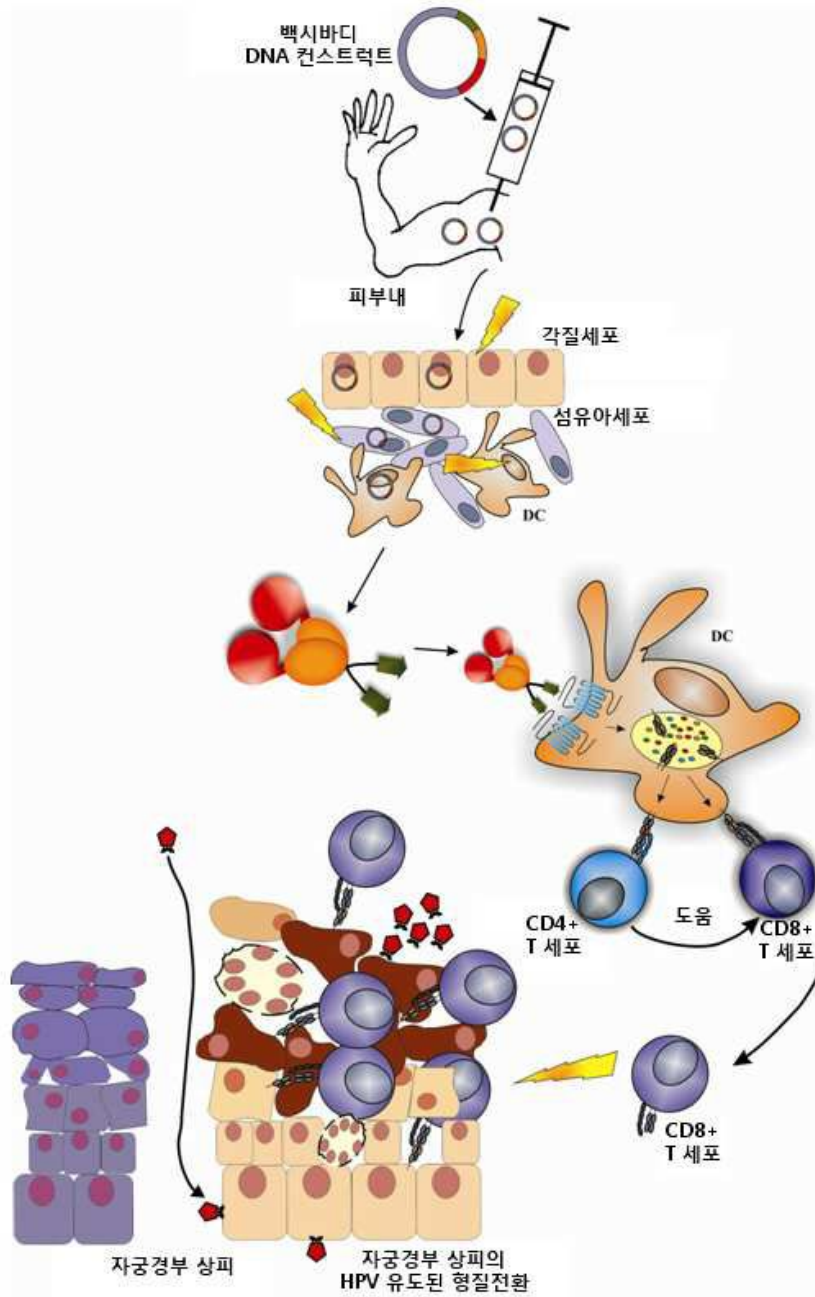
도면

도면1

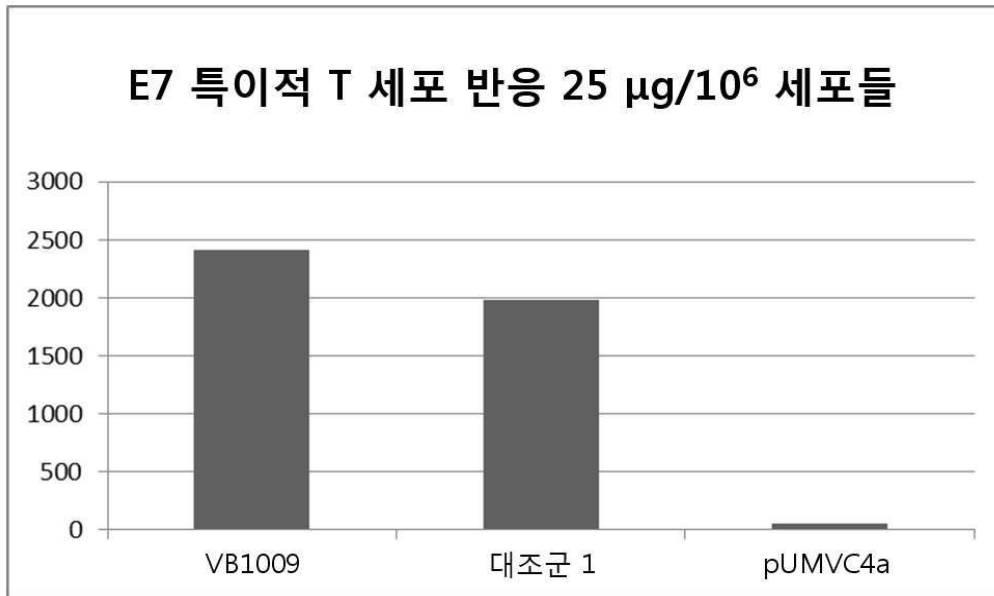
### 백시바디 치료적 HPV DNA 백신



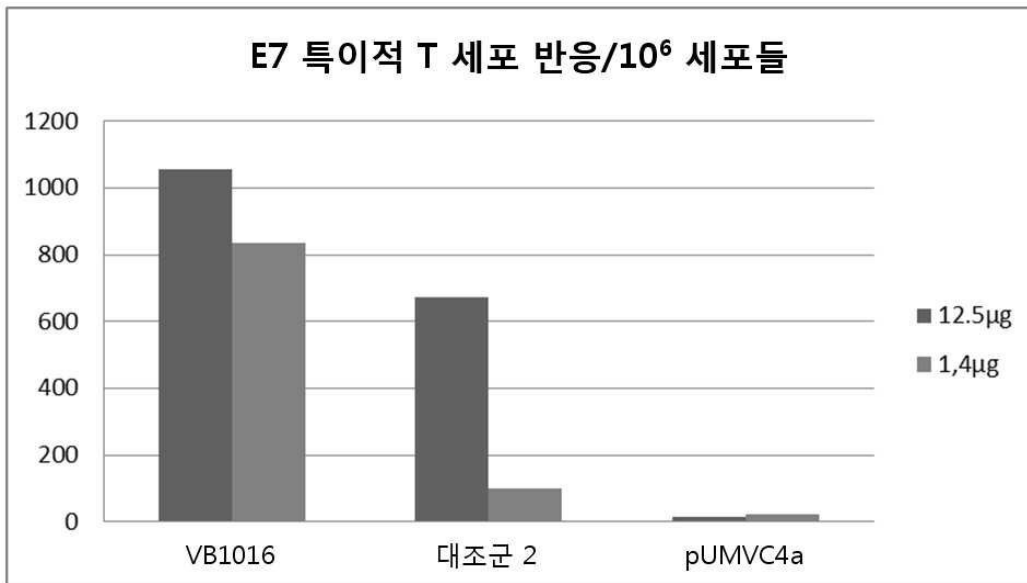
도면2



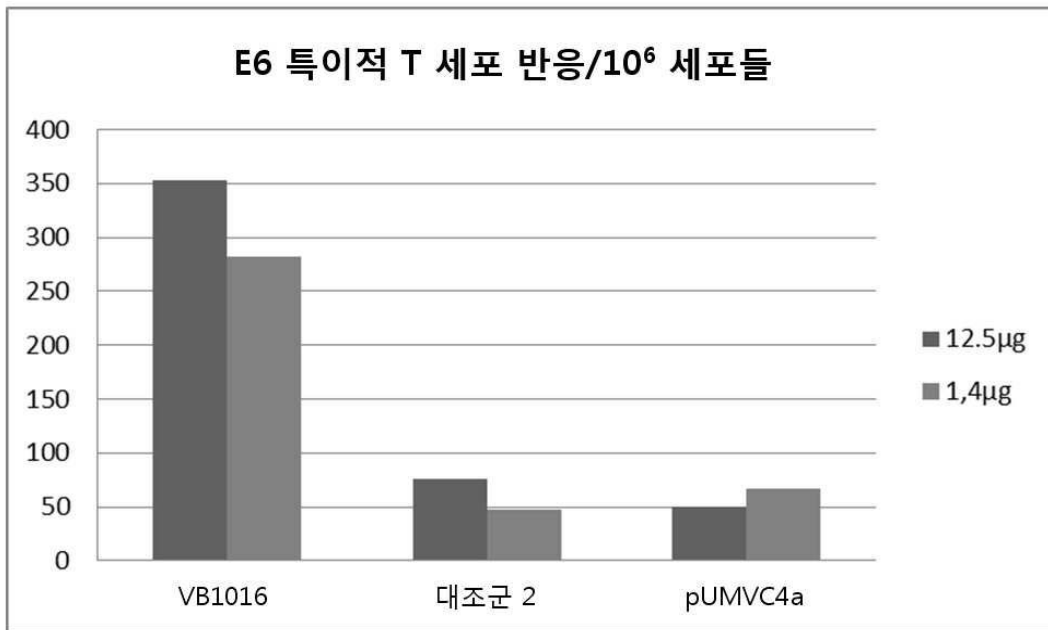
도면3a



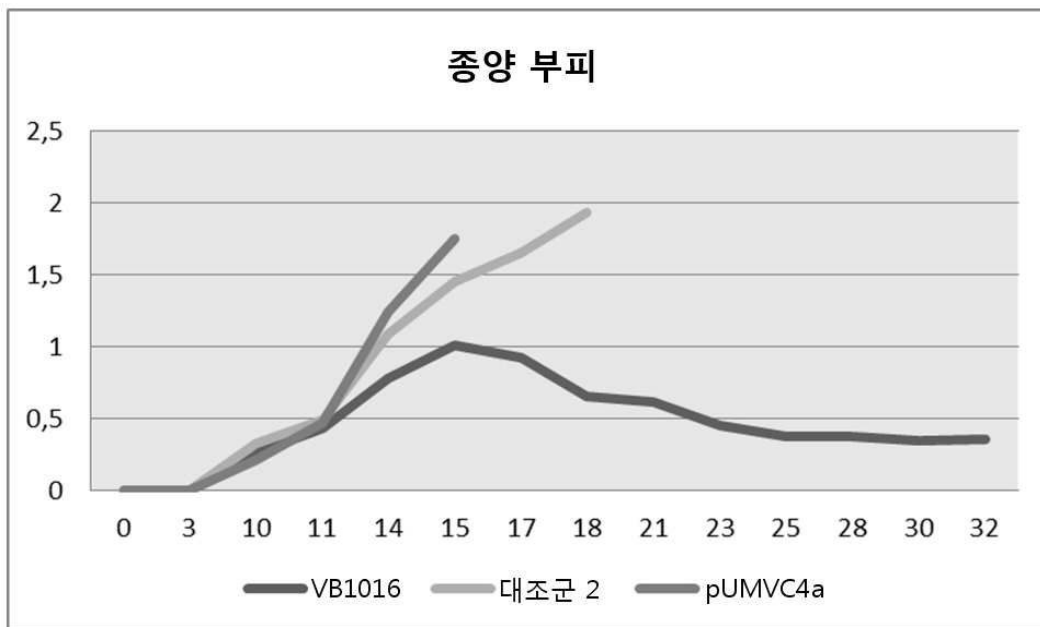
도면3b



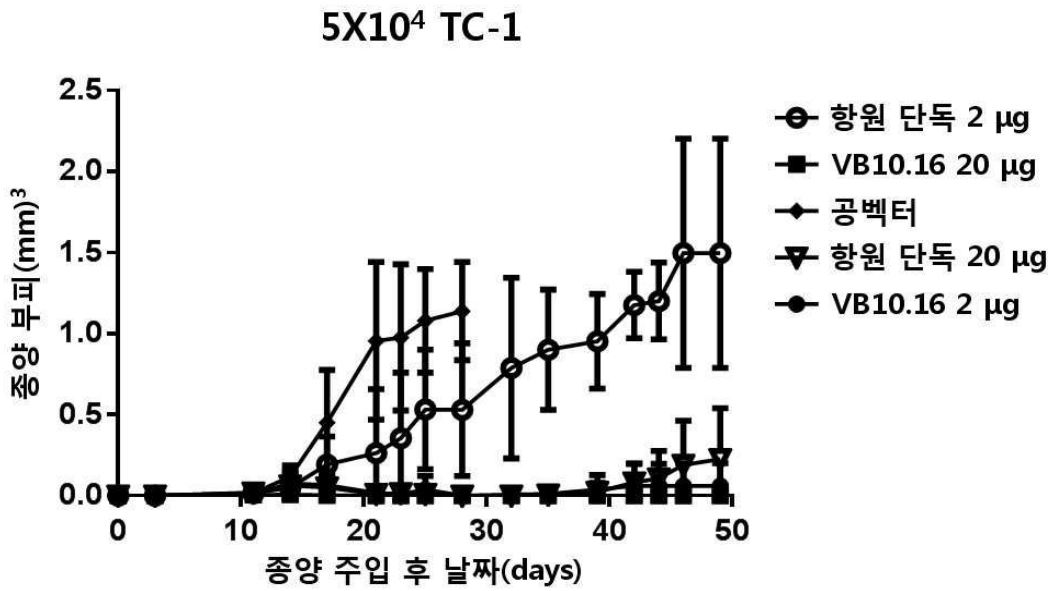
도면3c



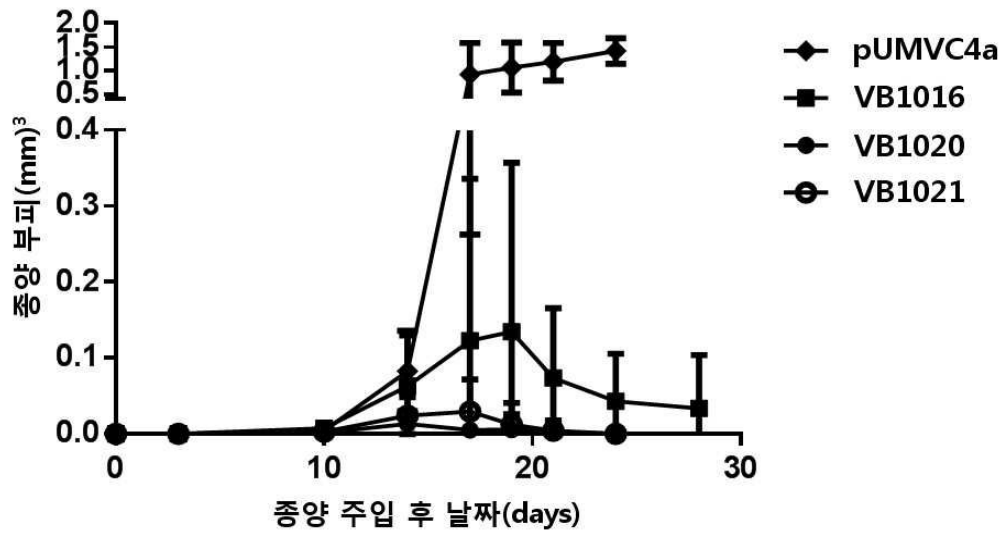
도면4



도면5



도면6



**서열목록**

<110> VACCIBODY AS

<120> VACCINES AGAINST HPV

<130> KP-CH-142169

<150> US 61/578,542

<151> 2011-12-21

<160> 34

<170> Kopatent In 2.0

<210> 1

<211> 93

<212> PRT

<213> Artificial Sequence

<220><223> C-C motif chemokine 3-like 1 precursor

<400> 1

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr

20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile

35 40 45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile

50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu

65 70 75 80

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

85 90

<210> 2

<211> 1182

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of VB1001

<400> 2

atgcaggtct cactgctgc cttgccgtc ctctctgca ccatggctct ctgcaaccag 60

gtctctctg caccacttg tgctgacacg cggaccgct gctgcttcag ctacacctcc 120

cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180

cccagtgtca tcttctaac caagagaggc cggcaggtct gtgctgacct cagtgaggag 240

tgggtccaga aatacgtcag tgacctggag ctgagtgccg agctcaaac cccacttggc 300

gacacaactc acacagagcc caaatcttgt gacacacctc cccctgccc aaggtgccc 360

ggcggtgga gacagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc 420

ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggcctcaa 480

ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540

tacaacacca cgctcccat gctggactcc gacggctcct tcttctcta cagcaagctc 600

accgtggaca agagcagggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660  
 gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcgg 720  
 ggcctgatgt ttcaggaccc acaggagcga cccagaaaagt taccacagtt atgcacagag 780

ctgcaaaca ctatacatga tataatatta gaatgtgtgt actgcaagca acagttactg 840  
 cgacgtgagg tatatgactt tgcttttcgg gatttatgca tagtatatag agatgggaat 900  
 ccatatgctg taigtgataa atgtttaaag ttttattcta aaattagtga gtatagacat 960  
 tattgttata gtttgtatgg aacaacatta gaacagcaat acaacaaacc gttgtgtgat 1020  
 ttgttaatta ggtgtattaa ctgtcaaaag ccaactgtgtc ctgaagaaaa gcaaagacat 1080  
 ctggacaaaa agcaaagatt ccataatata aggggtcggg ggaccggtcg atgtatgtct 1140  
 tgttgcatg catcaagaac acgtagagaa acccagctgt aa 1182

<210> 3  
 <211> 393  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Protein sequence of VB1001  
 <400> 3

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala  
 1 5 10 15  
 Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr  
 20 25 30  
 Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile  
 35 40 45  
 Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
 50 55 60  
 Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
 65 70 75 80  
 Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys  
 85 90 95  
 Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr  
 100 105 110  
 Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Ser Ser Gly Gly Gly



Asn Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser  
 370 375 380

Ser Arg Thr Arg Arg Glu Thr Gln Leu  
 385 390

<210> 4  
 <211> 1182  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> DNA sequence of VB1002  
 <400> 4

atgcaggtct ccaactgctgc cttgcccgtc ctctctgca ccatggctct ctgcaaccag 60

gtcctctctg caccacttgc tgctgacacg ccgaccgcct gctgcttcag ctacacctcc 120

cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180

cccagtgta tcttctaac caagagaggc cggcaggtct gtgctgacct cagtgaggag 240

tgggtccaga aatcagtcag tgacctggag ctgagtccg agctcaaac cccacttgg 300

gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtgccc aaggtgccca 360

ggcggtgga gcagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc 420

ctgccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggtcaaa 480

ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540

tacaacacca cgctcccat gctggactcc gacggctcct tcttctcta cagcaagctc 600

accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660

gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggctcgg 720

ggcctgatgt ttcaggacc acagagagcga cccagaaagt taccacagtt atgcacagag 780

ctgcaaaaa ctatacatga tataatatta gaatgtgtgt actgcaagca acagttactg 840

cgacgtgagg tatatgactt tgcttttcgg gatttatgca tagtatatag agatgggaat 900

ccatatgctg tacgagataa atgtttaaag ttttattcta aaattagtga gtatagacat 960

tattgttata gttgtatgg aacaacatta gaacagcaat acaacaaacc gttgtgtgat 1020

ttgttaatta ggtgtattaa ctgtcaaaaag ccaactgtgtc ctgaagaaaa gcaaagacat 1080

ctggacaaaa agcaaagatt ccataatata aggggtcggg ggaccggtcg atgtatgtct 1140

tgttgcagat catcaagaac acgtagagaa acccagctgt aa 1182

<210> 5  
 <211> 393

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1002

<

400> 5

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr

20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile

35 40 45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile

50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu

65 70 75 80

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

85 90 95

Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr

100 105 110

Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly

115 120 125

Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser

130 135 140

Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys

145 150 155 160

Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln

165 170 175

Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly

180 185 190

Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln

195 200 205

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

210                                 215                                 220  
 Arg Phe Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Leu Gly  
 225                                 230                                 235                                 240  
 Gly Leu Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln  
                                   245                                 250                                 255  
 Leu Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys  
                                   260                                 265                                 270  
 Val Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala  
                                   275                                 280                                 285  
  
 Phe Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val  
                                   290                                 295                                 300  
 Arg Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His  
 305                                 310                                 315                                 320  
 Tyr Cys Tyr Ser Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys  
                                   325                                 330                                 335  
 Pro Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro Leu  
                                   340                                 345                                 350  
 Cys Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His  
  
                                   355                                 360                                 365  
 Asn Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser  
                                   370                                 375                                 380  
 Ser Arg Thr Arg Arg Glu Thr Gln Leu  
 385                                 390  
 <210>     6  
 <211>     1182  
 <212>     DNA  
 <213>     Artificial Sequence  
 <220><223>     DNA sequence of VB1003  
 <400>     6  
 atgcaggtct ccaactgtgc cttgcccgtc ctctctgtca ccatggctct ctgcaaccag                 60  
 gtctctctg caccacttgc tgctgacacg cggaccgcct gctgcttcag ctacacctec                 120  
  
 cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag                 180

cccagtgta ttttctaac caagagaggc cggcaggtct gtgctgacct cagtgaggag 240  
 tgggtccaga aatacgtcag tgacctggag ctgagtcccg agctcaaac cccacttggt 300  
 gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtgccc aaggtgccca 360  
 ggcggtggaa gcagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc 420  
 ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggtcaaa 480  
 ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540

tacaacacca cgctcccat gctggactcc gacggctcct ttttctcta cagcaagctc 600  
 accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660  
 gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcggt 720  
 ggctgatgt ttcaggacct acaggagcga ccagaaaagt taccacagtt atgcacagag 780  
 ctgcaaaaaa ctatacatga tataatatta gaatgtgtgt actgcaagca acagttactg 840  
 cgacgtgagg tatatgactt tgcttttcgg gatttatgca tagtatatag agatgggaat 900  
 ccatatgctg tatgtgataa atgtttaaag ttttattcta aaattagtga gtatagacat 960

tattgttata gtttgtatgg aacaacatta gaacagcaat acaacaaacc gttgtgtgat 1020  
 ttgttaatta ggtgtattaa cgcacaaaag cactgtgtc ctgaagaaaa gcaaagacat 1080  
 ctggacaaaa agcaaagatt ccataatata aggggtcggt ggaccggtcg atgtatgtct 1140  
 tgttgcatg catcaagaac acgtagagaa acccagctgt aa 1182

<210> 7

<211> 393

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1003

<400> 7

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr

20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile

35 40 45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile

50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu



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

370 375 380  
 Ser Arg Thr Arg Arg Glu Thr Gln Leu  
 385 390

<210> 8  
 <211> 1182  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> DNA sequence of VB1004  
 <400> 8

atgcaggctc ccaactgetgc ccttgccgtc ctctctgca ccatggctct ctgcaaccag 60  
 gtctctctg caccacttgc tgctgacacg cggaccgct gctgcttcag ctacacctcc 120  
 cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180  
 cccagtgtca tcttctaac caagagaggc cggcaggctc gtgctgacct cagtgaggag 240  
  
 tgggtccaga aatacgtcag tgacctggag ctgagtccg agctcaaac cccacttggc 300  
 gacacaactc acacagagcc caaatcttgt gacacacctc ccccgctgcc aaggtgcca 360  
 ggcggtggaa gcagcggagg tggaaagtga ggacagcccc gagaaccaca ggtgtacacc 420  
 ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctgggtcaa 480  
 ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540  
 tacaacacca cgctcccat gctggactcc gacggctcct tcttctcta cagcaagctc 600  
 accgtggaca agagcagggt gcagcagggg aacatcttct catgctccgt gatgcatgag 660  
  
 gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcggt 720  
 ggcttgatgt ttcaggacct acaggagcga ccagaaagt taccacagt atgcacagag 780  
 ctgcaaacaa ctatacatga tataatatta gaatgtgtgt actgcaagca acagttactg 840  
 cgacgtgagg tatatgactt tgctcgacgg gatttatgca tagtatatag agatgggaat 900  
 ccatatgctg tacgagataa atgtttaaag ttttattcta aaattagtga gtatagacat 960

tattgttata gtttgtatgg aacaacatta gaacagcaat acaacaaacc gttgtgtgat 1020  
 ttgttaatta ggtgtattaa cgcacaaaag ccaactgtgc ctgaagaaaa gcaaagacat 1080

ctggacaaaa agcaaagatt ccataatata aggggtcggt ggaccggtcg atgtatgtct 1140  
 tgttcagat catcaagaac acgtagagaa acccagctgt aa 1182

<210> 9

<211> 393

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1004

<400> 9

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr

20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile

35 40 45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile

50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu

65 70 75 80

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

85 90 95

Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr

100 105 110

Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly

115 120 125

Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser

130 135 140

Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys

145 150 155 160

Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln



<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of VB1005

<400> 10

```

atgcaggtct ccaactgctgc cttgcccgtc ctctctgca ccatggctct ctgcaaccag      60
gtcctctctg caccacttgc tgctgacacg cggaccgctt gctgcttcag ctacacctcc      120
cgacagattc cacagaatth catagctgac tactttgaga cgagcagcca gtgctccaag      180
cccagtgtca tcttctaac caagagaggg cggcaggtct gtgctgacct cagtgaggag      240
tgggtccaga aatacgtcag tgacctggag ctgagtgccg agctcaaac cccacttgg      300

gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtgccc aaggtgccca      360
ggcgggtggaa gcagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc      420
ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctgggcaaa      480
ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac      540
tacaacacca cgcctcccat gctggactcc gacggctcct tcttctctca cagcaagctc      600
accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag      660
gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcgg      720

ggcctgatgc atggagatac acctacattg catgaatata tgttagattt gcaaccagag      780
acaactgata tctactgta tgagcaatta aatgacagct cagaggagga ggatgaata      840
gatggtccag ctggacaagc agaaccggac agagcccatt acaatattgt aacctttgt      900
tgcaagtgtg actctacgct tcggttgtgc gtacaaagca cacacgtaga cattcgtact      960
ttggaagacc tgtaatggg cacactagga attgtgtgcc ccatctgttc tcagaaacca      1020
taa                                                                                   1023
    
```

<210> 11

<211> 340

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1005

<400> 11

```

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala
  1           5           10           15
Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr
          20           25           30
    
```

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile  
 35 40 45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
 50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
 65 70 75 80

Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys  
 85 90 95

Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr  
 100 105 110

Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly  
 115 120 125

Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser  
 130 135 140

Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys  
 145 150 155 160

Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln  
 165 170 175

Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly  
 180 185 190

Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln  
 195 200 205

Gln Gly Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn  
 210 215 220

Arg Phe Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Leu Gly  
 225 230 235 240

Gly Leu Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp  
 245 250 255

Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp  
 260 265 270

Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu

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

340  
 <210> 12  
 <211> 1023  
 <212> DNA  
 <213> Artificial Sequence

<220><223> DNA sequence of VB1006

<400> 12  
 atgcaggtct ccaactgctgc cttgcccgtc ctctcttgca ccatggctct ctgcaaccag 60  
 gtctctctg caccacttgc tgctgacacg ccgaccgcct gctgcttcag ctacacctec 120  
 cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180  
 cccagtgtca tcttctaac caagagaggc cggcaggtct gtgctgacct cagtgaggag 240  
 tgggtccaga aatacgtcag tgacctggag ctgagtgccg agctcaaac cccacttgg 300  
  
 gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtgccc aaggtgcca 360  
 ggccgtggaa gcagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc 420  
 ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctgggcaaa 480  
 ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540  
 tacaacacca cgcctccat gctggactcc gacggctcct tcttcteta cagcaagctc 600  
 accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660  
 gctctgcaca accgcttac gcagaagagc ctctccctgt ctccgggtaa aggcctcgg 720  
  
 ggctgatgc atggagatac acctacattg catgaatata tgttagattt gcaaccagag 780  
 acaactgac tctacggata tggacaatta aatgacagct cagaggagga ggatgaata 840  
 gatggtccag ctggacaagc agaaccggac agagccatt acaatattgt aacctttgt 900  
 tgcaagtgtg actctacgt tcggttgtgc gtacaaagca cacacgtaga cattcgtact 960  
 ttggaagacc tgtaatggg cacactagga attgtgtgcc ccatctgttc tcagaaacca 1020

taa

1023

<210> 13

<211> 340

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1006

<400> 13

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr

20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile

35 40 45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile

50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu

65 70 75 80

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

85 90 95

Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr

100 105 110

Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly

115 120 125

Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser

130 135 140

Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys

145 150 155 160

Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln

165 170 175

Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly

180 185 190

Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln

195 200 205

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

210 215 220

Arg Phe Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Leu Gly

225 230 235 240

Gly Leu Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp

245 250 255

Leu Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gly Gln Leu Asn Asp

260 265 270

Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu

275 280 285

Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp

290 295 300

Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr

305 310 315 320

Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys

325 330 335

Ser Gln Lys Pro

340

<210> 14

<211> 1023

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of VB1007

<400> 14

atgcaggtct ccaactgctgc cttgcccgtc ctctctgca ccatggctct ctgcaaccag 60

gtcctctctg caccacttgc tgctgacacg ccgaccgcct gctgcttcag ctacactec 120

cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180

cccagtgtca tcttctaacc caagagagge cggcaggtct gtgctgaccc cagtgaggag 240

tgggtccaga aatacgtcag tgacctggag ctgagtgccg agctcaaac cccacttggt 300

gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtagcc aaggtgccca 360  
 ggcggtggaa gcagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc 420  
 ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggtcaaa 480  
 ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540  
 tacaacacca cgctcccat gctggactcc gacggctcct tcttcctcta cagcaagctc 600  
 accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660  
 gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcggt 720

ggcctgatgc atggagatac acctacattg catgaatata tgtagattt gcaaccagag 780  
 acaactgata tctacggata tggacaatta aatgacagct cagaggagga ggatgaata 840  
 gatggtccag ctggacaagc agaaccggac agagcccatt acaatattgt aacctttgga 900  
 tgcaaggag actctacgct tcggttgtgc gtacaaagca cacacgtaga cattcgtact 960  
 ttggaagacc tgtaaatggg cacactagga attgtgtgcc ccatctgttc tcagaaacca 1020  
 taa 1023

<210> 15  
 <211> 340  
 <212> PRT

<213> Artificial Sequence  
 <220><223> Protein sequence of VB1007

<400> 15  
 Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala  
 1 5 10 15  
 Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr  
 20 25 30  
 Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile  
 35 40 45  
 Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
 50 55 60  
 Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
 65 70 75 80  
 Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys  
 85 90 95  
 Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr

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

340

<210> 16  
 <211> 1023  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> DNA sequence of VB1008  
 <400> 16  
 atgcaggtct ccaactgctgc ccttgccgic ctctctgca ccatggctct ctgcaaccag 60  
 gtctctctg caccacttgc tgctgacacg cggaccgct gctgttcag ctacacctcc 120  
 cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180  
 cccagtgtca tcttcctaac caagagaggc cggcaggtct gtgctgacct cagtgaggag 240  
 tgggtccaga aatacgtcag tgacctggag ctgagtgccg agctcaaac cccacttgg 300  
  
 gacacaactc acacagagcc caaatcttgt gacacacctc ccccgcgccc aaggtgceca 360  
 ggcggtggaa gcagcggagg tggaaagtga ggacagcccc gagaaccaca ggtgtacacc 420  
 ctgcccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctgggcaaa 480  
 ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540  
 tacaacacca cgctcccat gctggactcc gacggctcct tcttctcta cagcaagctc 600  
 accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660  
 gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcgg 720  
  
 ggctgatgc atggagatac acctacattg catgaatata tgtagattt gcaaccagag 780  
 acaactgac tctacggata tggacaatta aatgacagct cagaggagga ggatgaaata 840  
 gatggtccag ctggacaagc agaaccggac agagcccatt acaatattgt aacctttgt 900  
 tgcaagtgtg actctacgt tcggttgtgc gtacaaagca cacacgtaga cattcgtact 960  
 ttggaagacc tgtaatggg cacactagga attgtgggac ccatcggatc tcagaaacca 1020  
 taa 1023

<210> 17  
 <211> 340  
 <212> PRT

<213> Artificial Sequence  
 <220><223> Protein sequence of VB1008

<400> 17  
 Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala  
 1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr  
 20 25 30  
 Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile  
 35 40 45  
 Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
 50 55 60  
 Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
 65 70 75 80  
 Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys  
 85 90 95  
 Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr  
 100 105 110  
 Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly  
 115 120 125  
 Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser  
 130 135 140  
 Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys  
 145 150 155 160  
 Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln  
 165 170 175  
 Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly  
 180 185 190  
 Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln  
 195 200 205  
 Gln Gly Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn  
 210 215 220  
 Arg Phe Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Leu Gly  
 225 230 235 240  
 Gly Leu Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp  
 245 250 255  
 Leu Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gly Gln Leu Asn Asp



gatgggccag ctggacaagc agaaccggac agagcccatt acaatattgt aaccttttgt 900  
 tgcaagtgtg actctacgct tcggttgtgc gtacaaagca cacacgtaga cattcgtact 960  
 ttggaagacc tgftaatggg cacactagga attgtgtgcc ccatctgttc tcagaaacca 1020  
 ggcggtggaa gcagcggagg tggaagtgga atgtttcagg acccacagga gcgaccaga 1080  
 aagttaccac agttatgcac agagctgcaa acaactatac atgatataat attagaatgt 1140

gtgtactgca agcaacagtt actgcgacgt gaggtatatg actttgctcg acgggattta 1200  
 tgcatagtat atagagatgg gaatccatat gctgtacgag ataaatgttt aaagttttat 1260  
 tctaaaatta gtgagtatag acattattgt tatagtttgt atggaacaac attagaacag 1320  
 caatacaaca aaccgttgtg tgatttgtaa attaggtgta ttaaccgaca aaagccactg 1380  
 tgtcctgaag aaaagcaaag acatctggac aaaaagcaaa gattccataa tataagggtt 1440  
 cgggtggaccg gtcgatgtat gtcttgttgc agatcatcaa gaacacgtag agaaaccag 1500  
 ctgtaa 1506

<210> 19  
 <211> 501  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Protein sequence of VB1009  
 <400> 19

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala  
 1 5 10 15  
 Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr  
 20 25 30  
 Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile  
 35 40 45  
 Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
 50 55 60  
 Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
 65 70 75 80  
 Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys  
 85 90 95  
 Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr  
 100 105 110

Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly  
 115 120 125

Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser  
 130 135 140

Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys  
 145 150 155 160

Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln  
 165 170 175

Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly  
 180 185 190

Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln  
 195 200 205

Gln Gly Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn  
 210 215 220

Arg Phe Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Leu Gly  
 225 230 235 240

Gly Leu Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp  
 245 250 255

Leu Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gly Gln Leu Asn Asp  
 260 265 270

Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu  
 275 280 285

Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp  
 290 295 300

Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr  
 305 310 315 320

Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys  
 325 330 335

Ser Gln Lys Pro Gly Gly Gly Ser Ser Gly Gly Gly Ser Gly Met Phe  
 340 345 350

Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu Cys Thr Glu

355 360 365  
 Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val Tyr Cys Lys

370 375 380  
 Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Arg Arg Asp Leu  
 385 390 395 400

Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Arg Asp Lys Cys  
 405 410 415

Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr Cys Tyr Ser  
 420 425 430

Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro Leu Cys Asp  
 435 440 445

Leu Leu Ile Arg Cys Ile Asn Arg Gln Lys Pro Leu Cys Pro Glu Glu  
 450 455 460

Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn Ile Arg Gly  
 465 470 475 480

Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser Arg Thr Arg  
 485 490 495

Arg Glu Thr Gln Leu  
 500

<210> 20

<211> 1506

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of VB1016

<400> 20

atgcaggctc ccaactgctgc ccttgccgtc ctctctgca ccatggctct ctgcaaccag 60  
 gtctctctg caccacttgc tgctgacacg cggaccgect gctgcttcag ctacacctcc 120  
 cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180  
 cccagtgtca tcttctaac caagagaggc cggcaggctc gtgctgacct cagtgaggag 240

tgggtccaga aatacgtcag tgacctggag ctgagtgccg agctcaaaac ccaacttggc 300  
 gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtgccc aaggtgceca 360  
 ggcggtggaa gcagcggagg tggaagtgga ggacagcccc gagaaccaca ggtgtacacc 420

ctgccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctgggtcaaa 480  
 ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540  
 tacaacacca cgcctcccat gctggactcc gacggctcct tcttctctca cagcaagctc 600  
 accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660

gctctgcaca accgcttcac gcagaagagc ctctccctgt ctccgggtaa aggcctcggg 720  
 ggccctgatgc atggagatac acctacattg catgaatata tgttagattt gcaaccagag 780  
 acaactgata tctacggata tggacaatta aatgacagct cagaggagga ggatgaaata 840  
 gatgggccag ctggacaagc agaaccggac agagcccatt acaatattgt aaccttttgt 900  
 tgcaagtgtg actctacgct tcggttgtgc gtacaaagca cacacgtaga cattcgtact 960  
 ttggaagacc tgtaaatggg cacactagga attgtgtgcc ccatctgttc tcagaaacca 1020  
 ggccggtggaa gcagcggagg tggaaagtga atgtttcagg acccacagga gcgaccaga 1080

aagttaccac agttatgcac agagctgcaa acaactatac atgatataat attagaatgt 1140  
 gtgtactgca agcaacagtt actgcgacgt gaggtatatg actttgcttt tcgggattta 1200  
 tgcatagtat atagagatgg gaatccatat gctgtacgag ataaatgttt aaagttttat 1260  
 tctaaaatta gtgagtatag acattattgt tatagtttgt atggaacaac attagaacag 1320  
 caatacaaca aaccgttgtg tgatttgta attaggtgta ttaaccgaca aaagccactg 1380  
 tgtcctgaag aaaagcaaag acatctggac aaaaagcaaa gattccataa tataaggggt 1440  
 cgggtggaccg gtcgatgtat gtcttgttgc agatcatcaa gaacacgtag agaaaccag 1500

ctgtaa 1506

<210> 21

<211> 501

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1016

<400> 21

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr

20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile

35

40

45

Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
 50 55 60

Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
 65 70 75 80

Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys  
 85 90 95

Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr  
 100 105 110

Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly  
 115 120 125

Ser Gly Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser  
 130 135 140

Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys  
 145 150 155 160

Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln  
 165 170 175

Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly  
 180 185 190

Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln  
 195 200 205

Gln Gly Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn  
 210 215 220

Arg Phe Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Leu Gly  
 225 230 235 240

Gly Leu Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp  
 245 250 255

Leu Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gly Gln Leu Asn Asp  
 260 265 270

Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu  
 275 280 285

Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp  
 290 295 300

Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr  
 305 310 315 320  
 Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys  
 325 330 335  
 Ser Gln Lys Pro Gly Gly Gly Ser Ser Gly Gly Gly Ser Gly Met Phe  
 340 345 350  
 Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu Cys Thr Glu  
 355 360 365  
 Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val Tyr Cys Lys  
 370 375 380  
 Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe Arg Asp Leu  
 385 390 395 400  
 Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Arg Asp Lys Cys  
 405 410 415  
 Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr Cys Tyr Ser  
 420 425 430  
 Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro Leu Cys Asp  
 435 440 445  
 Leu Leu Ile Arg Cys Ile Asn Arg Gln Lys Pro Leu Cys Pro Glu Glu  
 450 455 460  
 Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn Ile Arg Gly  
 465 470 475 480  
 Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser Arg Thr Arg  
 485 490 495  
 Arg Glu Thr Gln Leu  
 500  
 <210> 22  
 <211> 151  
 <212> PRT  
 <213> Human papillomavirus type 16  
 <400> 22  
 Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu Cys

1                    5                    10                    15  
 Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val Tyr  
  
                   20                    25                    30  
 Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe Arg  
                   35                    40                    45  
 Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys Asp  
                   50                    55                    60  
 Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr Cys  
                   65                    70                    75                    80  
 Tyr Ser Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro Leu  
                   85                    90                    95

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

- <210> 23
- <211> 98
- <212> PRT
- <213> Human papillomavirus type 16
- <400> 23

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

Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu

65 70 75 80

Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln

85 90 95

Lys Pro

<210> 24

<211> 158

<212> PRT

<213> Human papillomavirus type 18

<400> 24

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

1 5 10 15

Leu Cys Thr Glu Leu Asn Thr Ser Leu Gln Asp Ile Glu Ile Thr Cys

20 25 30

Val Tyr Cys Lys Thr Val Leu Glu Leu Thr Glu Val Phe Glu Phe Ala

35 40 45

Phe Lys Asp Leu Phe Val Val Tyr Arg Asp Ser Ile Pro His Ala Ala

50 55 60

Cys His Lys Cys Ile Asp Phe Tyr Ser Arg Ile Arg Glu Leu Arg His

65 70 75 80

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

85 90 95

Gly Leu Tyr Asn Leu Leu Ile Arg Cys Leu Arg Cys Gln Lys Pro Leu

100 105 110

Asn Pro Ala Glu Lys Leu Arg His Leu Asn Glu Lys Arg Arg Phe His

115 120 125

Asn Ile Ala Gly His Tyr Arg Gly Gln Cys His Ser Cys Cys Asn Arg

130 135 140

Ala Arg Gln Glu Arg Leu Gln Arg Arg Arg Glu Thr Gln Val

145 150 155

<210> 25

<211> 105

<212> PRT

<213> Human papillomavirus type 18

<400> 25

Met His Gly Pro Lys Ala Thr Leu Gln Asp Ile Val Leu His Leu Glu

1 5 10 15  
 Pro Gln Asn Glu Ile Pro Val Asp Leu Leu Cys His Glu Gln Leu Ser  
 20 25 30  
 Asp Ser Glu Glu Glu Asn Asp Glu Ile Asp Gly Val Asn His Gln His  
 35 40 45  
 Leu Pro Ala Arg Arg Ala Glu Pro Gln Arg His Thr Met Leu Cys Met  
 50 55 60  
 Cys Cys Lys Cys Glu Ala Arg Ile Lys Leu Val Val Glu Ser Ser Ala  
 65 70 75 80

Asp Asp Leu Arg Ala Phe Gln Gln Leu Phe Leu Asn Thr Leu Ser Phe  
 85 90 95  
 Val Cys Pro Trp Cys Ala Ser Gln Gln  
 100 105

<210> 26

<211> 12

<212> PRT

<213> Homo sapiens

<400> 26

Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr

1 5 10

<210> 27

<211> 15

<212> PRT

<213> Homo sapiens

<400> 27

Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys Pro

1 5 10 15

<210> 28  
 <211> 10  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Gly-Ser linker  
 <400> 28

Gly Gly Gly Ser Ser Gly Gly Gly Ser Gly  
 1 5 10

<210> 29  
 <211> 107  
 <212> PRT  
 <213> Homo sapiens  
 <400> 29

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu  
 1 5 10 15

Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe

20 25 30  
 Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln Pro Glu

35 40 45  
 Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe

50 55 60  
 Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly

65 70 75 80  
 Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Phe

85 90 95

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
 100 105

<210> 30  
 <211> 5  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Linker  
 <400> 30

Gly Leu Gly Gly Leu

1 5

<210> 31

<211> 1506

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of VB1020

<400> 31

atgcaggtct ccaactgetgc ccttgccgtc ctctctgca ccatggctct ctgcaaccag 60

gtcctctctg caccacttgc tgctgacacg ccgaccgctt gctgcttcag ctacacctcc 120

cgacagattc cacagaattt catagctgac tactttgaga cgagcagcca gtgctccaag 180

cccagtgta ttttctaac caagagaggc cggcaggtct gtgctgacct cagtgaggag 240

tgggtccaga aatacgtcag tgacctggag ctgagtccg agctcaaac cccacttgg 300

gacacaactc acacagagcc caaatcttgt gacacacctc ccccgtgccc aaggtgcccc 360

ggcggtgga gacgaggagg tgaagtga ggacagcccc gagaaccaca ggtgtacacc 420

ctgccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggtcaaa 480

ggcttctacc ccagcgacat cgccgtggag tgggagagca gcgggcagcc ggagaacaac 540

tacaacacca cgctcccat gctggactcc gacggctcct tcttctcta cagcaagctc 600

accgtggaca agagcaggtg gcagcagggg aacatcttct catgctccgt gatgcatgag 660

gctctgcaca accgcttac gcagaagagc ctctcctgt ctccgggtaa aggcctcgg 720

ggcctgatgc atggcgata cccaacactc catgagtaca tgctggacct tcagcccag 780

actacggatc tgtatggcta tgggcagttg aatgactcat ctgaggagga ggacgaata 840

gacggcccag ctggtcaagc cgaaccgat agagcccact acaacattgt gaccttttgc 900

tgtaagtgtg acagcactct gagactgtgt gttcagcca ctcatgtcga catacgaca 960

ttggaggatc tctgatggg aacctggga attgtgtgtc ccatctgttc ccaaaagcct 1020

ggaggtgga gacgtggagg cggttcaggc atgtccaag atcctcaaga acgtcctcgt 1080

aagctccac agctgtgtac cgagcttcag accaccattc acgacatcat cctggagtgc 1140

gtctattgca aacagcagct cttagaagg gaagtgtacg attttgcag gagggacctc 1200

tgcactgtgt atcgggacgg caatcctat gcggtacggg ataaatgcct gaagtctac 1260

agcaaaatct ccgagtaccg gcaactactgc tactctctct atgggacgac tctggaacag 1320

cagtacaaca agcccttgtg cgatctgctg attcgctgca ttaatcgcca gaaacctctg 1380

tgcccagaag agaagcaaag acacctggac aagaaacagc gattccacaa catccgaggg 1440  
 agatggacag ggaggtgtat gagctgctgt cggagtctta ggacaaggcg cgaaccag 1500  
 ctttga 1506

<210> 32  
 <211> 501  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Protein sequence of VB1020  
 <400> 32

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





ctggaggatc tgctcatggg gactctgggg attgtgtgtc ctatttgag ccagaaacca 1020  
 ggccgaggat cttcaggagg cgggagtggc atgttccaag accctcagga acgccctcgg 1080  
 aaactgcccc aattgtgtac tgagctccag acaacgatac acgacataat cctggagtgc 1140  
 gtgtattgca agcagcagct tctgaggagg gaagtgtacg attttgccag gagagatggc 1200  
 tgcattgtct accgagatgg caatccctat gcggtgtgtg ataagtgtct gaagtctat 1260  
 tccaaaatca gcaaatatcg gcattattgc tactcactgt acggaactac cctcgaacag 1320

cagtacaaca aaccgctctg tgatctgctg atcagatgca tcaatcggca gaaccctt 1380  
 tgtcccgaag agaagcaaag acacctggac aagaagcaga ggttcacaa tacccgaggt 1440  
 cgttggactg ggcgctgcat gtcctgttgt cgctcctctc gcacaaggag agagacacaa 1500  
 ctgtga 1506

<210> 34

<211> 501

<212> PRT

<213> Artificial Sequence

<220><223> Protein sequence of VB1021

<400> 34

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

|   |     |     |     |
|---|-----|-----|-----|
| 1   | 5   | 10  | 15  |
| Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr |     |     |     |
|   | 20  | 25  | 30  |
| Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile |     |     |     |
|   | 35  | 40  | 45  |
| Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile |     |     |     |
|   | 50  | 55  | 60  |
| Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu |     |     |     |
| 65  | 70  | 75  | 80  |
| Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala Glu Leu Lys |     |     |     |
|   | 85  | 90  | 95  |
| Thr Pro Leu Gly Asp Thr Thr His Thr Glu Pro Lys Ser Cys Asp Thr |     |     |     |
|   | 100 | 105 | 110 |
| Pro Pro Pro Cys Pro Arg Cys Pro Gly Gly Gly Ser Ser Gly Gly Gly |     |     |     |
|   | 115 | 120 | 125 |

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

