



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0129425
(43) 공개일자 2009년12월16일

(51) Int. Cl.

C07K 14/19 (2006.01) C12N 9/54 (2006.01)
C12N 15/63 (2006.01)

(21) 출원번호 10-2009-7018911

(22) 출원일자 2008년03월12일

심사청구일자 없음

(85) 번역문제출일자 2009년09월10일

(86) 국제출원번호 PCT/US2008/003291

(87) 국제공개번호 WO 2008/112258

국제공개일자 2008년09월18일

(30) 우선권주장

60/906,734 2007년03월12일 미국(US)

(71) 출원인

다니스코 유에스 인크.

미합중국 캘리포니아 (우편번호 94304) 팔로 알토 페이지 밀 로드 925

(72) 발명자

에스텔 데이비드 에이

미국 94304 캘리포니아주 팔로 알토 페이지 밀 로드 925 다니스코 유에스 인크. 제넨코 디비전

페라리 유지니오

미국 94304 캘리포니아주 팔로 알토 페이지 밀 로드 925 다니스코 유에스 인크. 제넨코 디비전

(74) 대리인

특허법인코리아나

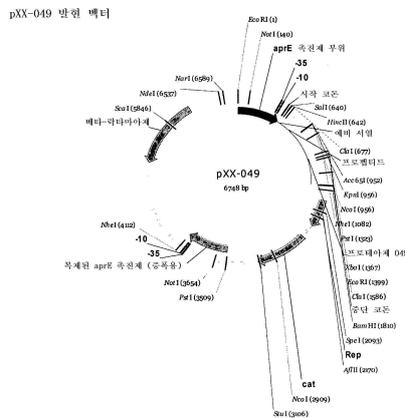
전체 청구항 수 : 총 26 항

(54) 개질된 프로테아제

(57) 요약

본 발명은 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드, 및 미생물에서의 프로테아제의 생성을 변경시키기 위한 방법에 관한 것이다. 특히, 본 발명은 바실러스 종과 같은 미생물에서의 프로테아제의 발현을 변경시키는 방법에 관한 것이다. 본 발명은 개질된 폴리뉴클레오티드, 벡터, 개질된 폴리펩티드, 및 프로테아제의 생성 증강 방법을 개시한다.

대표도 - 도6



특허청구의 범위

청구항 1

개질된 전체 길이 프로테아제를 인코딩하는 단리된 개질 폴리뉴클레오티드로서, 상기 전체 길이 프로테아제의 pro 부위를 인코딩하는 상기 개질된 폴리뉴클레오티드의 부분이 SEQ ID NO:5, 13 또는 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함하는 폴리뉴클레오티드.

청구항 2

제 1 항에 있어서, 상기 개질된 전체 길이 프로테아제가 세린 프로테아제인, 단리된 개질 폴리뉴클레오티드.

청구항 3

제 1 항에 있어서, 상기 개질된 전체 길이 프로테아제가 야생형 또는 변종 전구체 알칼리 세린 프로테아제에서 유래한 알칼리 세린 프로테아제인, 단리된 개질 폴리뉴클레오티드.

청구항 4

제 3 항에 있어서, 상기 전구체 알칼리 세린 프로테아제가 *B. 클라우시* (*B. clausii*) 또는 *B. 렌투스* (*B. lentus*) 알칼리 세린 프로테아제인, 단리된 개질 폴리뉴클레오티드.

청구항 5

제 1 항에 있어서, SEQ ID NOS: 1, 9 또는 240에서 나타난 폴리뉴클레오티드 서열을 포함하는, 단리된 개질 폴리뉴클레오티드.

청구항 6

제 1 항에 있어서, pro 부위를 인코딩하는 상기 개질된 폴리뉴클레오티드의 상기 부분이 SEQ ID NO:5, 13 또는 244의 E33, E43, A44, E47, V49, E57, A59, E63, E70, E74, E84 및 E88에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함하는, 단리된 개질 폴리뉴클레오티드.

청구항 7

제 6 항에 있어서, E33의 상기 치환이 E33D, E33I, E33S, E33N, E33K, E33H, E33Q 및 E33R에서 선택되는, 단리된 폴리뉴클레오티드.

청구항 8

제 6 항에 있어서, E57의 상기 치환이 E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57S, E57H 및 E57N에서 선택되는, 단리된 폴리뉴클레오티드.

청구항 9

SEQ ID NO:5, 13 또는 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하도록 돌연변이된 전구체 pro 서열을 포함하는 단리된 개질 폴리뉴클레오티드.

청구항 10

제 9 항에 있어서, 성숙 부위를 포함하는 전체 길이 프로테아제를 인코딩하는 폴리뉴클레오티드를 추가로 포함하는 단리된 개질 폴리뉴클레오티드로서, 상기 성숙 부위를 인코딩하는 상기 폴리뉴클레오티드 서열이 SEQ ID NOS:8, 16 또는 247과 약 70% 이상 동일한 폴리뉴클레오티드.

청구항 11

제 1 항 또는 제 9 항의 개질된 폴리뉴클레오티드를 포함하는 벡터.

청구항 12

제 11 항의 벡터로 형질전환된 숙주 세포.

청구항 13

제 12 항에 있어서, 상기 숙주 세포가 미생물인 숙주 세포.

청구항 14

제 12 항에 있어서, 상기 숙주 세포가 바실러스 종 (*Bacillus* sp.), 스트렙토마이세스 종 (*Streptomyces* sp.), 대장균 종 (*Escherichia* sp.) 및 아스퍼질러스 종 (*Aspergillus* sp.)으로 이루어지는 군에서 선택된 미생물인 숙주 세포.

청구항 15

제 12 항에 있어서, 상기 숙주 세포가 *B. 서브틸리스* (*B. subtilis*) 세포인 숙주 세포.

청구항 16

제 12 항의 숙주 세포에 의해 생성된 프로테아제.

청구항 17

하기 단계를 포함하는, 미생물에서 이종 프로테아제를 생성시키는 방법:

(a) 적절한 조건 하에 바실러스 숙주 세포를 배양하는 단계, 여기서 상기 바실러스 숙주 세포는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함; 및

(b) 미생물에 의해 상기 프로테아제를 생성시키는 단계.

청구항 18

제 17 항에 있어서, 상기 바실러스 숙주에 의해 생성된 상기 이종 프로테아제가 회수되는 방법.

청구항 19

제 17 항에 있어서, 상기 프로테아제가 알칼리 세린 프로테아제인 방법.

청구항 20

제 17 항에 있어서, 상기 개질된 폴리뉴클레오티드가 SEQ ID NOS:8,16 또는 247과 약 70% 이상 동일한 성숙 부위를 포함하는 프로테아제를 인코딩하는 방법.

청구항 21

제 17 항에 있어서, 개질된 프로테아제를 인코딩하는 상기 폴리뉴클레오티드가 SEQ ID NO:5, 13 또는 240의 아미노산 위치 28~108 및 109와 동등한 위치로 이루어지는 군에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함하는 pro 부위를 포함하는 방법.

청구항 22

제 21 항에 있어서, 상기 하나 이상의 아미노산 치환부가 E33D, E33I, E33S, E33N, E33K, E33H, E33Q, E33R, E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57S, E57H 및 E57N에서 선택되는 방법.

청구항 23

제 17 항에 있어서, 상기 숙주 세포가 *B. 리체니포르미스* (*B. licheniformis*), *B. 렌투스* (*B. lentus*), *B. 서브틸리스* (*B. subtilis*), *B. 아밀로리쿠에파시엔스* (*B. amyloliquefaciens*), *B. 브레비스* (*B. brevis*), *B. 스테아로써모필루스* (*B. stearothermophilus*), *B. 클라우시* (*B. clausii*), *B. 알칼로필루스* (*B. alkalophilus*), *B. 할로두란스* (*B. halodurans*), *B. 코아굴란스* (*B. coagulans*), *B. 써큘란스* (*B. circulans*), *B. 푸밀루스* (*B. pumilus*) 및 *B. 투링기엔시스* (*B. thuringiensis*)로 이루어지는 군에서 선택되는 방법.

청구항 24

제 17 항에 있어서, 상기 이중 프로테아제가 B. 클라우시 또는 B. 렌투스 프로테아제인 방법.

청구항 25

제 17 항에 있어서, 상기 숙주 세포가 B. 서브틸리스 숙주 세포인 방법.

청구항 26

제 17 항에 있어서, 상기 이중 프로테아제가 1 이상의 생성 비율을 나타내는 방법.

명세서

기술분야

<1> 본 발명은 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드, 및 미생물에서의 프로테아제의 생성을 변경시키는 방법을 제공한다. 특히, 본 발명은 바실러스 종 (Bacillus species)과 같은 미생물에서의 프로테아제의 발현을 변경시키는 방법에 관한 것이다. 본 발명은 개질된 폴리뉴클레오티드, 벡터, 개질된 폴리펩티드, 및 프로테아제 생성 증강 방법을 개시한다.

배경기술

<2> 미생물, 예컨대 바실러스 속 (genus Bacillus)의 멤버인 그람-양성 미생물은 일부 이들의 배양 배지로의 발효 생성물 분비 능력으로 인해, 대규모 산업 발효에 사용되어왔다. 분비된 단백질은 세포막 및 세포벽을 가로질러 배출된 후, 외부 배지로 방출된다. 폴리펩티드의 주변세포질 공간 또는 이의 배양 배지로의 분비는 산업 발효에서 주의깊게 고려될 필요가 있는 매개변수의 다양성의 영향을 받기 쉽다.

<3> 실제로, 이중 폴리펩티드의 분비는 산업에서 널리 사용되는 기술이다. 전형적으로 세포는, 발현 및 분비되어 대량의 원하는 폴리펩티드를 생성시킬 관심 있는 이중 폴리펩티드를 인코딩하는 핵산으로 형질전환된다. 이러한 기술은 대량의 폴리펩티드를 생성하는데 사용되어왔다. 원하는 폴리펩티드의 발현 및 분비는 원하는 단백질을 인코딩하는 폴리펩티드의 유전적 조작을 통해 제어되어왔다. 단백질 생성 방법의 다양한 진전에도 불구하고, 세포의 단백질 분비를 위한 효율적인 방법을 제공할 필요성이 당업계에 남아있다.

<4> 발명의 요약

<5> 본 발명은 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드, 및 미생물에서의 프로테아제의 생성을 변경시키기 위한 방법에 관한 것이다. 특히, 본 발명은 바실러스 종과 같은 미생물에서의 프로테아제의 생성을 변경시키기 위한 방법에 관한 것이다. 본 발명은 개질된 폴리뉴클레오티드, 벡터, 개질된 폴리펩티드, 및 프로테아제 생성 증강 방법을 개시한다.

<6> 본 발명은 폴리뉴클레오티드, 폴리펩티드 및 세포 (개질된 단백질의 생성이 증강되도록 유전적으로 조작된)에 관한 것이다. 특히, 본 발명은 이에 도입된 외생적 핵산 서열을 갖는 그람-양성 미생물, 및 이러한 숙주 세포 예컨대 바실러스 속의 멤버에서의 단백질 생성 방법에 관한 것이다. 보다 특별히, 본 발명은 프로테아제의 생성, 및 발현된 단백질을 생성하기 위한 변경된 용량을 갖도록 유전적으로 조작된 세포에 관한 것이다. 특히, 본 발명은 미생물에 의한 프로테아제의 증강된 생성을 제공한다.

<7> 일부 구현예에서, 본 발명은 개질된 전체 길이 프로테아제를 인코딩하는 단리된 개질 폴리뉴클레오티드를 제공하는데, 여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 치환부를 인코딩 치환부를 인코딩하는 돌연변이를 포함한다. 일부 구현예에서 개질된 전체 길이 프로테아제는 세린 프로테아제이며, 여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함한다. 다른 구현예에서, 전체 길이 프로테아제는 야생형 또는 변종 전구체 알칼리 세린 프로테아제 유래 알칼리 세린 프로테아제이며, 여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함한다. 다른 구현

예에서, 전체 길이 프로테아제는 야생형 또는 변종 전구체 알칼리 세린 프로테아제 유래 알칼리 세린 프로테아제, 즉 B. 클라우시 (B. Clausii) 또는 B. 렌투스 (B. lentus) 알칼리 세린 프로테아제이며, 여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함한다. 일부 구현예에서, 단리된 개질 폴리뉴클레오티드는 SEQ ID NOS:1, 9 또는 240에 나타난 폴리뉴클레오티드 서열을 포함한다.

- <8> 일부 구현예에서, 본 발명은 개질된 전체 길이 프로테아제를 인코딩하는 단리된 개질 폴리뉴클레오티드를 제공하는데, 여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함하며, 여기서 하나 이상의 치환부는 SEQ ID NO:5, 13 또는 244의 E33, E43, A44, E47, V49, E57, A59, E63, E70, E74, E84, 및 E88에서 선택된 하나의 아미노산 위치에서 생성된다. 일부 다른 구현예에서, E33의 아미노산 치환부는 E33D, E33I, E33S, E33N, E33K, E33H, E33Q 및 E33R에서 선택된다. 다른 구현예에서, E57의 아미노산 치환부는 E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57S, E57H 및 E57N에서 선택된다.
- <9> 일부 구현예에서, 본 발명은 SEQ ID NO:5, 13 또는 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하도록 돌연변이된 전구체 pro 서열을 포함하는 단리된 개질 폴리뉴클레오티드를 제공한다. 일부 추가적인 구현예에서, 단리된 개질 폴리뉴클레오티드는 성숙 부위를 포함하는 전체 길이 프로테아제를 인코딩하는 폴리뉴클레오티드를 추가로 포함하며, 여기서 프로테아제의 성숙 부위를 인코딩하는 폴리뉴클레오티드는 SEQ ID NOS:8, 16 또는 247과 약 70% 이상 동일하다.
- <10> 일부 구현예에서, 본 발명은 개질된 전체 길이 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함하는 하나 이상의 벡터를 제공하며, 여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이, 또는 SEQ ID NO:5, 13 또는 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하도록 돌연변이된 전구체 pro 서열을 포함하는 단리된 개질 폴리뉴클레오티드를 포함한다.
- <11> 다른 구현예에서, 본 발명은 본 발명의 개질된 폴리뉴클레오티드를 포함하는 벡터로 형질전환된 숙주 세포를 제공한다. 일부 바람직한 구현예에서, 형질전환 숙주 세포는 미생물이다. 예를 들어, 일부 구현예에서 본 발명은 개질된 전체 길이 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함하는 하나 이상의 벡터로 형질전환되거나 (여기서 전체 길이 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 및 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함함), 또는 SEQ ID NO:5, 13 또는 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하도록 돌연변이된 전구체 pro 서열을 포함하는 단리된 개질 폴리뉴클레오티드로 형질전환된 숙주 세포를 제공한다. 일부 구현예에서, 본 발명의 폴리뉴클레오티드로 형질전환된 숙주 세포는 바실러스 종 (Bacillus sp.), 스트렙토마이세스 종 (Streptomyces sp.), 대장균 종 (Escherichia sp.) 및 아스퍼질러스 종 (Aspergillus sp.)으로 이루어지는 군에서 선택된 미생물이다. 일부 다른 구현예에서, 숙주 세포는 B. 서브틸리스 (B. subtilis)이다.
- <12> 일부 구현예에서, 본 발명은 본 발명의 형질전환된 숙주 세포에 의해 생성된 프로테아제를 제공한다.
- <13> 본 발명은 또한 하기 단계를 포함하는, 미생물에서의 이중 프로테아제 생성 방법을 제공한다: (a) 적절한 조건 하에 바실러스 숙주 세포를 배양하는 단계 (여기서 상기 바실러스 숙주 세포는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계. 일부 구현예에서, 바실러스 숙주에 의해 생성된 프로테아제가 회수된다. 본원에 제공된 개질된 폴리뉴클레오티드 중 임의 하나를 본 발명의 방법에서 사용한다.
- <14> 일부 구현예에서, 본 발명은 하기 단계를 포함하는, 미생물에서의 이중 알칼리 세린 프로테아제의 생성 방법을 제공한다: (a) 적절한 조건 하에 바실러스 숙주 세포를 배양하는 단계 (여기서 상기 바실러스 숙주 세포는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계. 본 발명은 또한 하기 단계를 포함하는, 미생물에서의 이중 프로테아제의 생성 방법을 제공한다: (a) 적절한 조건 하에 바실러스 숙주 세포를 배양하는 단계 (여기서 상기 바실러스 숙주 세포는 SEQ ID NOS:8, 16 또는 247과 약 70% 이상 동일한 성숙 부위를 포함하는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오

티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계. 일부 구현예에서, 바실러스 속주에 의해 생성된 프로테아제가 회수된다.

<15> 일부 구현예에서, 본 발명은 하기 단계를 포함하는, 미생물에서의 이중 프로테아제의 생성 방법을 제공한다: (a) 적절한 조건 하에 바실러스 속주 세포를 배양하는 단계 (여기서 상기 바실러스 속주 세포는 SEQ ID NO:5, 13 또는 244의 아미노산 위치 28~108 및 109와 동등한 위치에서 선택된 하나 이상의 아미노산 위치에서 치환부를 인코딩하는 돌연변이를 포함하는 pro 부위를 포함하는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계. 다른 구현예에서, 하나 이상의 아미노산 치환부는 E33D, E33I, E33S, E33N, E33K, E33H, E33Q, E33R, E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57S, E57H 및 E57N에서 선택된다.

<16> 본 발명은 또한 하기 단계를 포함하는, 이중 프로테아제, 즉 B. 클라우시 또는 B. 렌투스 프로테아제의 미생물에서의 생성 방법을 제공한다: (a) 적절한 조건 하에 바실러스 속주 세포를 배양하는 단계 (여기서 상기 바실러스 속주 세포는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계.

<17> 본 발명은 또한 하기의 단계를 포함하는, 미생물에서의 이중 프로테아제의 생성 방법을 제공한다: (a) 적절한 조건 하에 B. 리체니포르미스 (B. licheniformis), B. 렌투스, B. 서브틸리스, B. 아밀로리쿠에파시엔스 (B. amyloliquefaciens), B. 브레비스 (B. brevis), B. 스테아로써모필루스 (B. stearothermophilus), B. 클라우시, B. 알칼로필루스 (B. alkalophilus), B. 할로두란스 (B. halodurans), B. 코아굴란스 (B. coagulans), B. 써쿨란스 (B. circulans), B. 푸밀루스 (B. pumilus), 및 B. 투링기엔시스 (B. thuringiensis)에서 선택된 바실러스 속주 세포를 배양하는 단계 (여기서 상기 바실러스 속주 세포는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계. 일부 구현예에서, 속주 세포는 B. 서브틸리스 속주 세포이다.

<18> 본 발명은 또한 하기 단계를 포함하며, 이중 프로테아제가 1 이상의 생성 비율을 나타내는, 미생물에서의 이중 프로테아제의 생성 방법을 제공한다: (a) 적절한 조건 하에 바실러스 속주 세포를 배양하는 단계 (여기서 상기 바실러스 속주 세포는 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 포함함); 및 (b) 미생물에 의해 프로테아제를 생성시키는 단계.

발명의 상세한 설명

<65> 본 발명은 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드, 및 미생물에서의 프로테아제의 생성을 변경시키는 방법에 관한 것이다. 특히, 본 발명은 바실러스 종과 같은 미생물에서의 프로테아제의 생성을 변경시키는 방법에 관한 것이다. 본 발명은 개질된 폴리뉴클레오티드, 벡터, 개질된 플라스미드, 및 프로테아제의 생성 증강 방법을 개시한다.

<66> 일부 구현예에서, 본 발명은 동물, 식물성 또는 미생물 기원의 야생형 또는 변종 전구체 프로테아제 유래의 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드를 제공한다. 미생물 유래의 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드가 바람직하며, 야생형 형태에서 유래된 변종 전구체 프로테아제 및 야생형 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 본 발명의 개질된 폴리뉴클레오티드는 미생물 기원의 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드에서 유래된다. 본 발명은 또한 단백질 공학 제조된 전구체를 인코딩하는 폴리뉴클레오티드 유래의 개질된 프로테아제를 인코딩하는 폴리뉴클레오티드를 포함한다. 세린 프로테아제를 인코딩하는 폴리뉴클레오티드가 바람직한 전구체 프로테아제 폴리뉴클레오티드이며, 이 중 알칼리 미생물 프로테아제 폴리뉴클레오티드가 특히 바람직하다. 세린 프로테아제는 활성 위치에 본질적인 세린 잔기가 있는 펩티드 결합의 가수분해를 촉매하는 효소이다. 세린 프로테아제는 약 25,000 내지 30,000 범위의 분자량을 갖는다 (Priest, Bacteriol. Rev., 41 :711-753 [1977] 참조). 일부 바람직한 구현예에서, 서브틸리신 (subtilisin) 및 서브틸리신 변종을 인코딩하는 폴리뉴클레오티드가 바람직한 전구체 세린 프로테아제 폴리뉴클레오티드이다. 일부 구현예에서, 본 발명은 미생물 예컨대 B. 리체니포르미스, B. 서브틸리스, B. 아밀로리쿠에파시엔스, B. 클라우시, B. 렌투스 및 B. 할로두란스의 폴리뉴클레오티드 유래의 개질된 프로테아제를 인코딩하는 폴리뉴클레오티드를 포함한다. 서브틸리신을 인코딩하는 다양한 바실러스 폴리뉴클레오티드, 예를 들어 서브틸리신 168, 서브틸리신 BPN', 서브틸리신 칼스버그, 서브틸리신 DY, 서브틸리신 147 및 서브틸리신 309가 확인되고 서열화되었다 (예를 들어 각각 전문이 본원에 참고문헌으로 포함된 EP 414279 B; WO 89/06279; 및 Stahl 등, J. Bacteriol., 159:811-818 [1984] 참조). 본 발명의 일부 구현예에서, 돌연변이 (예를 들어 변종) 프로테아제 폴리뉴클레오티드는 본 발명의 개질된 프로테아제가 유

래되는 단백질을 인코딩하는 전구체 폴리뉴클레오티드로서 역할한다. 많은 참고문헌이 변종 프로테아제의 예를 제공한다 (예를 들어 각각 전문이 본원에 참고문헌으로 포함된 WO 99/20770; WO 99/20726; WO 99/20769; WO 89/06279; RE 34,606; 미국 특허 제 4,914,031호; 미국 특허 제 4,980,288호; 미국 특허 제 5,208,158호; 미국 특허 제 5,310,675호; 미국 특허 제 5,336,611호; 미국 특허 제 5,399,283호; 미국 특허 제 5,441,882호; 미국 특허 제 5,482,849; 미국 특허 제 5,631,217호; 미국 특허 제 5,665,587호; 미국 특허 제 5,700,676호; 미국 특허 제 5,741,694호; 미국 특허 제 5,858,757호; 미국 특허 제 5,880,080호; 미국 특허 제 6,197,567호; 및 미국 특허 제 6,218,165호 참조). 다른 구현예에서, 본 발명에 의해 포함되는 폴리뉴클레오티드에는 시판되는 프로테아제를 인코딩하는 폴리뉴클레오티드에서 유래되는 개질된 폴리뉴클레오티드가 포함된다. 예를 들어, 시판되는 프로테아제에는 ALCALASE™, SAVINASE™, PRIMASE™, DURALASE™, ESPERASE™ 및 KANNASE™ (Novo Nordisk A/S사제), MAXATASE™, MAXACAL™, MAXAPEM™, PROPERASE™, PURAFECT™ 및 PURAFECT OXP™ (Genencor International Inc.사제)가 포함되나 이에 제한되지는 않는다. 일부 구현예에서, 본 발명은 B. 클라우시로부터의 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드 유래의 개질된 폴리뉴클레오티드를 포함한다. 다른 구현예에서, 본 발명은 B. 렌투스로부터의 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드 유래의 개질된 폴리뉴클레오티드를 포함한다.

<67> 일부 구현예에서, 본 발명의 개질된 전체 길이 폴리뉴클레오티드는 돌연변이된 전구체 프로테아제의 pro 부위를 인코딩하는 서열을 포함한다. 임의 적절한 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열이 본 발명의 하나 이상의 개질된 폴리뉴클레오티드의 생성에 사용된다. 일부 바람직한 구현예에서, pro 부위를 인코딩하는 전구체 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:5, 13 또는 244의 프로테아제의 위치 1~109와 동등한 위치에서 하나 이상의 아미노산 치환부를 인코딩하도록 돌연변이된다. 일부 구현예에서, 개질된 전구체 폴리뉴클레오티드는 SEQ ID NO:5, 13 또는 244의 E33, E43, A44, E47, V49, E57, A59, E63, E70, E74, E84 및 E88과 동등한 위치에서 pro 부위 중 하나 이상의 아미노산 치환부를 인코딩한다. 일부 구현예에서, E33과 동등한 위치에서 아미노산을 인코딩하는 폴리뉴클레오티드 서열은 치환부 E33D, E33I, E33S, E33N, E33K, E33H, E33Q 또는 E33R 중 하나 이상을 인코딩하도록 돌연변이된다. 일부 다른 구현예에서, E57과 동등한 위치에서 아미노산을 인코딩하는 폴리뉴클레오티드 서열은 치환부 E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57S, E57H 및/또는 E57N 중 하나를 인코딩하도록 돌연변이된다.

<68> 일부 구현예에서, 전구체 B. 클라우시 프로테아제 폴리뉴클레오티드 (예를 들어 SEQ ID NO:1)는 야생형 프로테아제 (MAXACAL™; SEQ ID NO:5)를 인코딩한다.

<69> (SEQ ID NO:1)

ATGAAGAAACCGTTGGGGAAAATTGTCGCAAGCACCGCACTACTCATTCTGTTGCTTTTA
 GTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAGAAAATATTTAATTGGCTTTAATGA
 GCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCT
 CTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTCCTGTTTTAT
 CCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACCGATCCAGCGATTTCTTATAT
 TGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCAGTGCCATGGGGAATTAGCCGTGT
 GCAAGCCCAGCTGCCATAACCGTGGATTGACAGGTTCTGGTGTAAAAGTTGCTGTCCCT
 CGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTGGCGCTAGCTTTGTACCA
 GGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCTGGGACGATTGC
 TGCTTTAAACAATTGATTGGCGTTCTTGGCGTAGCACCGAACGCGGAACACTATACGCTGTT
 AAAGTATTAGGGGCGAGCGGTTCCAGGTTCCGGTCAGCTCGATTGCCAAGGATTGGAATGG
 GCAGGGAACAATGGCATGCACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCC
 ACACTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTG
 GGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCCGTTATGCGAACGCAATGGCAGTCG
 GAGCTACTGACCAAACAACAACCGCGCCAGCTTTTACAGTATGGCGCAGGGCTTGACA
 TTGTGCGACCAGGTGTAAACGTGCAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAA
 ACGGTACATCGATGGCTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAA
 CCCATCTTGGTCCAATGTACAAATCCGCAATCATCTAAAGAATACGGCAACGAGCTTAGGA
 AGCACGAACTTGTATGGAAGCGGACTTGTCAATGCAGAAGCGGCAACACGCTAA

<70>

<71> (SEQ ID NO:5)

MKKPLGKIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEE
 VEIELLHEFETIPVLSVELSPEDVDALELDPAISYIEEDA EVTTMRVQAPAAHNRGLTGSGVKVA
 VLDTGISTHPDLNIRGGASFVPGEPTQDGNHGHGTHVAGTIAALNNSIGVLGVAPNAELYAVKV
 LGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPPSATLEQAVNSATSRGVLVVAASGNSG
 AGSISYPARYANAMAVGATDQNNNRASFSQYAGLDIVAPGVNVQSTYPGSTYASLNGTSM
 A
 TPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

<72>

<73>

<74> 다른 구현예에서, 전구체 B. 클라우시 프로테아제 폴리뉴클레오티드 (예를 들어 SEQ ID NO:9)는 변종 프로테아제 (예를 들어 SEQ ID NO:13)를 인코딩한다.

<75> (SEQ ID NO:13)

VRSKKLWIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEE
 VEIELLHEFETIPVLSVELSPEDVDALELDPAISYIEEDA EVTTMAQSVPWGISRVQAPAAHNRG
 LTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPTQDGNHGHGTHVAGTIAALNNSIGVLGVAP
 NAELYAVKVLGASGSGSVSSIAQGLEWAGNVMHVNLSLGLQAPSATLEQAVNSATSRGVL
 VVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYAGLDIVAPGVNVQSTYPGSTY
 ASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

<76>

<77> 다른 구현예에서, 전구체 프로테아제 폴리뉴클레오티드는 SEQ ID NO:244의 프로테아제 GG36을 인코딩하는 B. 렌투스 폴리뉴클레오티드 (예를 들어 SEQ ID NO:240)이다.

<78> (SEQ ID NO:240)

GTGAGAAGCAAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTTAGTTCAT
 CGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGT
 CAGTGAGTTTGTAGAACAAGTAGAGGCAAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAGGAAGT
 CGAAATTGAATTGCTTCATGAATTTGAAACGATTCCTGTTTTATCCGTTGAGTTAAGCCCAGAAGATG
 TGGACGCGCTTGAACCTCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGAAGTAACGACAATGGC
 GCAATCAGTGCCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTGGATTGACAGG
 TTCTGGTGAAAAGTTGCTGTCCTCGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTG
 GCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCC
 GGGACGATTGCTGCTCTAAACAATTCGATTGGCGTTCTTGGCGTAGCGCCGAGCGCGGAACCTATAC
 GCTGTTAAAGTATTAGGGGCGAGCGGTTTCCAGGCTCGGTCAGCTCGATTGCCCAAGGATTGGAATGG
 GCAGGGAACAATGGCATGCACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCCACACTT
 GAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGAAATTCAGGT
 GCAGGCTCAATCAGCTATCCGGCCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAAC
 AACACC GCGCCAGCTTTTACAGTATGGCGCAGGGCTTGACATTGTGCGACCAGGTGTAACGTTG
 CAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTTG
 CAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCATCTTGGTCCAATGTACAAATCCGCAATCATCT
 AAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGTATGGAAGCGGACTTGTCAATGCAGAAGC
 TGCAACTCGT

<79>

<80> (SEQ ID NO:244)

MRSKKLWIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEVEIELLH
 EFETIPVLSVELSPEDVDALELDPAISYIEEDA EVTTMAQSVPWGISRVQAPAAHNRGLTGSVVKVAVLDT
 GISTHPDLNIRGGASFVPGEPSTQDGNHGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVS
 SIAQGLEWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAV
 GATDQNNNRASFQYAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNV
 QIRNHLKNTATSLGSTNLYGSGLVNAEAATR

<81>

<82> 일부 구현예에서, 본 발명은 SEQ ID NO:9의 전체 길이 전구체 폴리뉴클레오티드 유래의 개질된 전체 길이 폴리 뉴클레오티드를 제공한다.

<83> (SEQ ID NO:9)

GTGAGAAGCAAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTTA
 GTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGA
 GCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCT
 CTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTAT
 CCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACCGATCCAGCGATTTCTTATAT
 TGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCGGTACCATGGGGAATTAGCCGTGT
 GCAAGCCCCAGCTGCCATAACCGTGGATTGACAGGTTCTGGTGTAAAAGTTGCTGTCTT
 CGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTGGCGTAGCTTTGTACCA
 GGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCTGGGACGATTGC
 TGCTTTAAACAATTGATTGGCGTTCTTTGGCGTAGCACCGAACCGGAACTATACGCTGTT
 AAAGTATTAGGGGCGAGCGGTTTCAGGTTTCGGTCAGCTCGATTGCCAAGGATTGGAATGG
 GCAGGGAACAATGTTATGCACGTTGCTAATTTGAGTTTAGGACTGCAGGCACCAAGTGCC
 ACACCTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTG
 GGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCCCGTTATGCGAACGCAATGGCAGTCG
 GAGCTACTGACCAAAACAACAACCGCGCCAGCTTTTCACAGTATGGCGCAGGGCTTGACA
 TTGTGCGACCAGGTGTAACCGTGCAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAA
 ACGGTACATCGATGGCTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAA
 CCCATCTTGGTCCAATGTACAAATCCGCAATCATCTAAAGAATACGGCAACGAGCTTAGGA
 AGCACGAACTTGTATGGAAGCGGACTTGTCAATGCAGAAGCGGCAACACGTTAA

<84>

<85> SEQ ID NO:9의 폴리뉴클레오티드 서열은 서열 (SEQ ID NO:10)을 포함하는데, 이는 발현되는 경우 SEQ ID NO:13의 아미노산 1~27에 미치는 신호 서열 펩티드 (SEQ ID NO:14); SEQ ID NO:13의 아미노산 잔기 28~111에 미치는 pro 부위 서열 (SEQ ID NO:15)을 인코딩하는 N-말단 pro 서열 (SEQ ID NO:11); 및 SEQ ID NO:13의 아미노산 잔기 112~380을 인코딩하는 성숙 세린 프로테아제 서열 (SEQ ID NO:12) (즉 SEQ ID NO:16)을 인코딩하도록 의도된다. SEQ ID NO:13의 프로테아제의 신호 펩티드의 첫 8개 아미노산을 인코딩하는 폴리뉴클레오티드 서열은 B. 서브틸리스 AprE 프로테아제의 첫 8개 아미노산을 인코딩하는 서열이다.

<86> (SEQ ID NO:10)

GTGAGAAGCAAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTTA
 GTTCATCGATCGCATCGGCT

<87>

<88> (SEQ ID NO:15)

AEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEEEVEIELLHEFETIPVLSVELSPEDVDAL
 ELDPAISYIEEDA EVTTM

<89>

<90> (SEQ ID NO:14)

VRSKKLWIVASTALLISVAFSSSIASA

<91>

<92> (SEQ IS NO:11)

GCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTCAGTGAGT
 TTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAGGAAGTCG
 AAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTATCCGTTGAGTTAAGCCCAGAA
 GATGTGGACGCGCTTGAACCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGAAGTAA
 CGACAATG

<93>

<94> (SEQ ID NO:12)

GCGCAATCGGTACCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTGG
 ATTGACAGGTTCTGGTGTAAAAGTTGCTGTCTCGATACAGGTATTTCCACTCATCCAGAC
 TAAATATTCGTGGTGGCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGGAATG
 GGCATGGCACGCATGTGGCTGGGACGATTGCTGCTTTAAACAATTCGATTGGCGTTCTTG
 GCGTAGCACCGAACGCGGAACCTATACGCTGTTAAAGTATTAGGGGCGAGCGGTTGAGTT
 CGGTCAGCTCGATTGCCAAGGATTGGAATGGGCAGGGAACAATGTTATGCACGTTGCTA
 ATTTGAGTTTAGGACTGCAGGCACCAAGTGCCACACTTGAGCAAGCTGTTAATAGCGCGA
 CTTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCAGGCTCAATCAGCT
 ATCCGGCCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAACAACAACCGCG
 CCAGCTTTTACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTGTAAACGTGCAGA
 GCACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGT
 TGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCAATGTACAAATCCGC
 AATCATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACCTTGTATGGAAGCGGACTT
 GTCAATGCAGAAGCGGCAACACGTTAA

<95>

<96>

<97> (SEQ ID NO:16)

AQSVPWGISRVQAPAAHNRGLTGSVGVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNH
 GTHVAGTIAALNNSIGVLGVAPNAELYAVKVLGASGSGSVSSIAQGLEWAGNNVMHVNLSLG
 LQAPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYG
 AGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATS
 LGSTNLYGSLVNAEAATR

<98>

<99> 다른 구현예에서, 본 발명은 SEQ ID NO:1의 야생형 전체 길이 전구체 폴리뉴클레오티드 유래의 개질된 전체 길
 이 폴리뉴클레오티드를 제공한다. SEQ ID NO:1의 폴리뉴클레오티드는 서열 SEQ ID NO:2를 포함하는데, 이는
 발현되는 경우 SEQ ID NO:5의 아미노산 1~27을 인코딩하는 신호 서열 펩티드 (SEQ ID NO:6); SEQ ID NO:5의 아
 미노산 잔기 28~111을 인코딩하는 N-말단 pro 서열 (SEQ ID NO:3) (즉 SEQ ID NO:7); 및 SEQ ID NO:5의 아미노
 산 잔기 112~380을 인코딩하는 야생형 성숙 세린 프로테아제 서열 (SEQ ID NO:4) (즉 SEQ ID NO:8)을 인코딩하
 도록 의도된다.

<100> (SEQ ID NO:2)

ATGAAGAAACCGTTGGGGAAAATTGTGCAAGCACCGCACTACTCATTCTGTTGCTTTTA
 GTTCATCGATCGCATCGGCT

<101>

<102> (SEQ ID NO:6)

MKKPLGKIVASTALLISVAFSSSIASA

<103>

<104> (SEQ ID NO:3)

GCTGAAGAAGCAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTCAGTGAGT
 TTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAGGAAGTCG
 AAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTATCCGTTGAGTTAAGCCCAGAA
 GATGTGGACGCGCTTGAACCTCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGAAGTAA
 CGACAATG

<105>

<106> (SEQ ID NO:7)

AEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEEEVEIELLHEFETIPVLSVELSPEDVDAL
 ELDPAISYIEEDA EVTTM

<107>

<108> (SEQ ID NO:4)

GCGCAATCAGTGCCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTGG
 ATTGACAGGTTCTGGTGTAAAAGTTGCTGTCTCGATACAGGTATTTCCACTCATCCAGAC
 TTAATATTCGTGGTGGCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGGAATG
 GGCATGGCACGCATGTGGCTGGGACGATTGCTGCTTTAAACAATTGATTGGCGTTCTTG
 GCGTAGCACCGAACGCGGAACATACGCTGTAAAGTATTAGGGGCGAGCGGTTGAGTT
 CGGTGAGCTCGATTGCCAAGGATTGGAATGGGCAGGGAACAATGGCATGCACGTTGCTA
 ATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCCACACTTGAGCAAGCTGTTAATAGCGCGAC
 TTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCAGGCTCAATCAGCTAT
 CCGGCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAACAACAACCGCGC
 CAGCTTTTACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTGTAACGTGCAGAG
 CACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTT
 GCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCA
 ATCATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACCTTGATGGAAGCGGACTTGT
 CAATGCAGAAGCGGCAACACGCTAA

<109>

<110> (SEQ ID NO:8)

RVQAPAAHNRGLTGSVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNHGHVAGTIAA
 LNNSIGVLGVAPNAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPPSATLEQ
 AVNSATSRGLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFQYAGLDIVAPGV
 NVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTAT
 SLGSTNLYGSLVNAEAATR

<111>

<112> 다른 구현예에서, 본 발명은 SEQ ID NO:240의 변종 전체 길이 전구체 폴리뉴클레오티드 유래의 개질된 전체 길
 이 폴리뉴클레오티드를 제공한다. SEQ ID NO:240의 폴리뉴클레오티드는 서열 SEQ ID NO:241을 포함하는데,
 이는 발현되는 경우 SEQ ID NO:244의 아미노산 1~27을 인코딩하는 신호 서열 펩티드 (SEQ ID NO:245); SEQ ID
 NO:244의 아미노산 잔기 28~111을 인코딩하는 N-말단 pro 서열 (SEQ ID NO:242) (즉 SEQ ID NO:246); 및 SEQ
 ID NO:244의 아미노산 잔기 112~380을 인코딩하는 야생형 성숙 세린 프로테아제 서열 (SEQ ID NO:243) (즉 SEQ
 ID NO:247)을 인코딩하도록 의도된다. SEQ ID NO:244의 전구체 프로테아제의 신호 펩티드의 첫 8개 아미노
 산을 인코딩하는 폴리뉴클레오티드 서열은 B. 서브틸리스 AprE 프로테아제의 첫 8개 아미노산을 인코딩하는 서
 열이다. 전구체 GG36의 pro 부분 및 성숙 부분은 GG36 야생형 서열에 의해 인코딩된다.

<113> (SEQ ID NO:241)

GTGAGAAGCAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTTAGTTCAT
 CGATCGCATCGGCT

<114>

<115> (SEQ ID NO:245)

MRSKKLWIVASTALLISVAFSSSIASA

<116>

<117> (SEQ ID NO:242)

GCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTCAGTGAGTTTGTAG
 AACAAAGTAGAGGCAATGACGAGGTCGCCATTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGC
 TTCATGAATTTGAAACGATTCCCTGTTTTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAA
 CTCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGAAGTAACGACAATG

<118>

<119> (SEQ ID NO:246)

AEEAKEKYLIGFNEQEA VSEFVEQVEANDEVAILSEEEVEI ELLHEFETIPVLSVELSPEDVDALELDPAIS
 YIEEDA EVTTM

<120>

<121> (SEQ ID NO:243)

GCGCAATCAGTGCCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTGGATTGACA
 GGTCTGGTGTAAAAGTTGCTGTCCTCGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGG
 TGGCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGG
 CCGGGACGATTGCTGCTCTAAACAATTCGATTGGCGTTCTTGGCGTAGCGCCGAGCGCGGAAGTAT
 ACGCTGTTAAAGTATTAGGGGCGAGCGGTTCAAGGCTCGGTCAGCTCGATTGCCCAAGGATTGGAAT
 GGGCAGGGAAACAATGGCATGCACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCCACAC
 TTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGAAATTCAGG
 TGCAGGCTCAATCAGCTATCCGGCCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAAA
 CAACAACCGCGCCAGCTTTTCACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTGTAACGT
 GCAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTT
 GCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCTATTGGTCCAATGTACAAATCCGCAATCATC
 TAAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGATGGAAGCGGACTTGTCATGCAGAAG
 CTGCAACTCGTTA

<122>

<123> (SEQ ID NO:247)

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNHGHGTHVAG
 TIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPPSATLEQAVN
 SATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFQYAGLDIVAPGVNVQSTYPGS
 TYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSLVNAEAATR

<124>

<125> 바람직한 구현예에서, 본 발명의 개질된 전체 길이 폴리뉴클레오티드는 돌연변이된 전구체 프로테아제의 pro 부위를 인코딩하는 서열을 포함한다. 임의 적절한 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열이 본 발명의 하나 이상의 개질된 폴리뉴클레오티드의 생성에 사용된다. 일부 바람직한 구현예에서, pro 부위를 인코딩하는 전구체 폴리뉴클레오티드 서열의 부분은 SEQ ID NO:13의 프로테아제의 위치 1~109와 동등한 위치에서 하나 이상의 아미노산 치환부를 인코딩하도록 돌연변이된다. 다른 구현예에서, 치환부는 SEQ ID NO:5의 프로테아제의 1~109와 동등한 위치에서 만들어진다. 다른 구현예에서, 치환부는 SEQ ID NO:244의 프로테아제의 위치 1~109와 동등한 위치에서 만들어진다. 일부 구현예에서, 개질된 전구체 폴리뉴클레오티드는 SEQ ID NO:5, 13 또는 244의 E33, E43, A44, E47, V49, E57, A59, E63, E70, E74, E84 및 E88과 동등한 위치에서 하나 이상의 아미노산 치환부를 인코딩한다. 일부 구현예에서, E33과 동등한 위치에서 아미노산을 인코딩하는 폴리뉴클레오티드 서열은 치환부 E33D, E33I, E33S, E33N, E33K, E33H, E33Q 또는 E33R 중 하나 이상을 인코딩하도록 돌연변이된다. 다른 일부 구현예에서, E57과 동등한 위치에서 아미노산을 인코딩하는 폴리뉴클레오티드 서열은 치환부 E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57S, E57H 및/또는 E57N 중 하나를 인코딩하도록 돌연변이된다.

<126> 상기 토의한 바와 같이, 일부 구현예에서, 임의의 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드는 돌연변이된 pro 부위를 갖는 하나 이상의 개질된 프로테아제를 인코딩하는 하나 이상의 개질된 폴리뉴클레오티드를 생성하도록 돌연변이된다. 일부 특히 바람직한 구현예에서, 세린 프로테아제를 인코딩하는 전체 길이 폴리뉴클레오티드는 본 발명의 개질된 폴리뉴클레오티드의 생성에 사용된다. 일부 대안적인 특히 바람직한 구현예에서, 알칼리 세린 프로테아제를 인코딩하는 전체 길이 폴리뉴클레오티드가 사용된다. 상기 토의한 바와 같이, 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드는 B. 리체니포르미스, B. 서브틸리스, B. 아밀로리쿠에파시엔스, B. 클라우시, B. 렌투스 및 B. 할로두란스를 포함하나 이에 제한되지 않는 미생물로부터의 것이다.

<127> 일부 구현예에서, 전체 길이 개질된 프로테아제를 인코딩하는 본 발명의 개질된 폴리뉴클레오티드는 전구체 프로테아제의 성숙 형태의 아미노산 서열과의 약 65% 이상의 아미노산 서열 동일성, 바람직하게는 약 70% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 75% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 80% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 85% 이상의 아미노산 서열 동일성, 더욱 보다 바람직하게는 약 90% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 92% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 95% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 97% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 98% 이상의 아미노산 서열 동일성, 및 가장 바람직하게는 약 99% 이상의 아미노산 서열 동일성을 공유하는 프로테아제의 성숙 형태를 인코딩하는 서열을 포함하고, 전구체 폴리펩티드와 비교하여 비교

할만한 또는 증강된 생성 활성을 갖는다. 일부 구현예에서, 전체 길이 개질된 프로테아제를 인코딩하는 본 발명의 개질된 폴리뉴클레오티드는 SEQ ID NO:8, SEQ ID NO:16 또는 SEQ ID NO:247의 아미노산 서열과의 약 65% 이상의 아미노산 서열 동일성, 바람직하게는 약 70% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 75% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 80% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 85% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 90% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 92% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 95% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 97% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 98% 이상의 아미노산 서열 동일성, 및 가장 바람직하게는 약 99% 이상의 아미노산 서열 동일성을 공유하는 프로테아제의 성숙 형태를 인코딩하는 서열을 포함한다.

<128> 일부 구현예에서, 개질된 폴리뉴클레오티드는 전구체 프로테아제의 아미노산 서열과의 65% 이하의 아미노산 서열 동일성, 바람직하게는 약 70% 이하의 아미노산 서열 동일성, 보다 바람직하게는 약 75% 이하의 아미노산 서열 동일성, 보다 더 바람직하게는 약 80% 이하의 아미노산 서열 동일성, 보다 바람직하게는 약 85% 이하의 아미노산 서열 동일성, 보다 더 바람직하게는 약 90% 이하의 아미노산 서열 동일성, 보다 바람직하게는 약 92% 이하의 아미노산 서열 동일성, 보다 더 바람직하게는 약 95% 이하의 아미노산 서열 동일성, 보다 바람직하게는 약 97% 이하의 아미노산 서열 동일성, 보다 더 바람직하게는 약 98% 이하의 아미노산 서열 동일성, 및 가장 바람직하게는 약 99% 이하의 아미노산 서열 동일성을 공유하는 전체 길이 아미노산 서열을 인코딩하며, 전구체 폴리펩티드와 비교하여 비교할만한 또는 증강된 생성 활성을 갖는다.

<129> 숙련자에 의해 이해될 것으로서, 유전적 코드의 퇴화로 인해, 다양한 개질된 폴리뉴클레오티드는 개질된 프로테아제를 인코딩한다. 본 발명의 일부 다른 구현예에서, SEQ ID NOS:4, 12 또는 243의 폴리뉴클레오티드 서열에 대해 약 70% 이상의 서열 동일성, 약 75% 이상의 서열 동일성, 약 80% 이상의 서열 동일성, 약 85% 이상의 서열 동일성, 약 90% 이상의 서열 동일성, 약 92% 이상의 서열 동일성, 약 95% 이상의 서열 동일성, 약 97% 이상의 서열 동일성, 약 98% 이상의 서열 동일성 및 약 99% 이상의 서열 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드가 제공된다.

<130> 일부 구현예에서, 폴리뉴클레오티드 서열에 의해 공유된 퍼센트 동일성은, 서열을 정렬하고 당업계에 공지된 방법에 의해 동일성을 측정함으로써 분자 사이의 서열 정보의 직접적인 비교에 의해 측정된다. 일부 구현예에서, 퍼센트 동일성 (예를 들어 아미노산 서열, 핵산 서열 및/또는 유전자 서열)은 서열을 정렬하고, 두 정렬된 서열 사이 일치 정렬의 정확한 수를 세고, 더 짧은 서열의 길이로 나누고, 결과를 100으로 곱함으로써, 두 분자 사이의 서열 정보의 직접적인 비교에 의해 측정된다. 쉽게 활용가능한 컴퓨터 프로그램이 상기 기재된 것들을 포함하는 이들 분석에 사용된다. 뉴클레오티드 서열 동일성을 측정하기 위한 프로그램이 Wisconsin Sequence Analysis Package, Version 8 (Genetics Computer Group, Madison, WI) 예를 들어 BESTFIT, FASTA 및 GAP 프로그램에서 활용가능한데, 이는 또한 Smith and Waterman 알고리즘에 의존한다. 이들 프로그램은 생산자에 의해 추천되고 상기 나타낸 Wisconsin Sequence Analysis Package에 기재된 디폴트 매개변수로 쉽게 사용된다.

<131> 서열 유사성을 측정하기에 적절한 알고리즘의 예는 Altschul, 등, J. Mol. Biol., 215:403-410 (1990)에 기재된 BLAST 알고리즘이다. BLAST 분석을 수행하기 위한 소프트웨어는 국립생명공학센터를 통해 공식적으로 활용가능하다. 이러한 알고리즘은 데이터베이스 서열에서 동일한 길이의 단어로 정렬되는 경우 일부 양성-값 역치 스코어 T와 일치 또는 이를 만족하는 문의 서열에서 길이 W의 짧은 단어를 확인함으로써 높은 스코어 서열 쌍 (HSP)을 1차 확인하는 것을 포함한다. 이들 초기 이웃 단어 적중은 이를 함유하는 더 긴 HSP를 발견하기 위한 출발점으로서 역할한다. 단어 적중은 누적 정렬 스코어가 증가될 수 있는 한, 비교되는 각각의 두 서열을 따르는 모든 방향으로 증폭된다. 단어 적중의 증폭은 다음의 경우 중단된다: 누적 정렬 스코어가 최대 획득 값으로부터 양 X에 의해 감소되는 경우; 누적 스코어가 0 이하가 되는 경우; 또는 서열의 말단에 도달하는 경우. BLAST 알고리즘 매개변수 W, T 및 X는 정렬의 속도 및 감수성을 결정한다. BLAST 프로그램은 단어길이 (W) 11, BLOSUM62 스코어 매트릭스 (Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989) 참조) 정렬 (B) 50, 기대값 (E) 10, M'5, N'-4, 및 두 가닥의 비교를 디폴트로서 사용한다.

<132> BLAST 알고리즘은 이후 두 서열 사이 유사성의 통계학적 분석을 수행한다 (예를 들어 Karlin 및 Altschul, Proc. Natl Acad. Sci. USA 90:5873-5787 [1993] 참조). BLAST 알고리즘에 의해 제공된 유사성의 한 측정치는, 두 뉴클레오티드 또는 아미노산 서열 사이의 일치 발생 확률의 표시를 제공하는 최소 합 확률 (P(N))이다. 예를 들어, 핵산은 세린 프로테아제 핵산에 대한 시험 핵산의 비교에서 최소 합 확률이 약 0.1 미만, 보다 바람직하게는 약 0.01 미만, 및 가장 바람직하게는 약 0.001 미만이라면 본 발명의 세린 프로테

아제 핵산과 유사한 것으로 여겨진다. 시험 핵산이 세린 프로테아제 폴리펩티드를 인코딩하는 경우, 비교 결과가 약 0.5 미만, 및 보다 바람직하게는 약 0.2 미만의 최소 합 확률 중이라면 구체화된 세린 프로테아제 핵산과 유사한 것으로 여겨진다.

- <133> 본 발명의 일부 구현예에서, BLAST 및 단백질 번역 서열 톨로 서열을 분석하였다. 일부 실험에서, 바람직한 버전은 BLAST (Basic BLAST version 2.0)였다. 선택된 프로그램은 "BlastX"였고, 선택된 데이터베이스는 "nr"이었다. 표준/디폴트 매개변수 값을 사용하였다.
- <134> 본 발명의 개질된 폴리뉴클레오티드 서열을 생성시키기에 적절한 당업계에 공지된 몇몇 방법에는 위치-포화 돌연변이 유발, 스캐닝 돌연변이 유발, 삽입 돌연변이 유발, 결실 돌연변이 유발, 무작위 돌연변이 유발, 위치-지정 돌연변이 유발 및 지정된 진화 뿐 아니라 다양한 다른 재조합적 접근이 포함되나 이에 제한되지는 않는다.
- <135> 일부 구현예에서, 본 발명의 개질된 폴리뉴클레오티드 서열은 하나 이상의 코돈에서 위치 지정 돌연변이 유발로 생성된 프로테아제의 pro 부위를 인코딩하는 서열의 개질을 포함한다. 다른 바람직한 구현예에서, 위치 지정 돌연변이 유발은 둘 이상의 코돈에 대해 수행된다. 일부 추가적인 구현예에서, 개질된 프로테아제의 pro 부위를 인코딩하는 개질된 폴리뉴클레오티드 서열은 전구체 폴리펩티드의 pro 부위를 인코딩하는 폴리뉴클레오티드 예를 들어 SEQ ID NOS:3, 11 또는 242와 약 40% 이하, 약 45% 이하, 약 50% 이하, 약 55% 이하, 약 60% 이하, 약 65% 이하, 약 70% 이하, 약 75% 이하, 약 80% 이하, 약 85% 이하, 약 90% 이하, 약 95% 이하, 약 98% 이하 또는 약 99% 이하의 상동성을 갖는다. 일부 대안적인 구현예에서, 개질된 폴리뉴클레오티드는 임의 공지된 돌연변이성 방법 예컨대 방사선, 니트로소구아니딘 등을 사용하여 체내에서 생성된다. 원하는 개질 폴리뉴클레오티드 서열을 이후 단리하고 본원에 제공되는 방법에 사용한다.
- <136> 일부 바람직한 구현예에서, 전구체 프로테아제 폴리뉴클레오티드의 pro 부위의 위치 포화 돌연변이 유발은, 돌연변이성 프라이머를 포함하는 조성물을 사용하여 달성된다. 돌연변이성 프라이머는 전구체 폴리뉴클레오티드와 정확히 일치되지는 않으며, 프라이머 중 불일치(들)는 주형 폴리뉴클레오티드에 원하는 돌연변이를 도입하는데 사용된다. 전구체 폴리뉴클레오티드와 정확히 일치되는 비돌연변이성 프라이머가 또한 프라이머 조성물에 포함된다. 돌연변이성 프라이머 중 하나 이상에 상응하는 돌연변이성 프라이머 및 비돌연변이성 프라이머의 혼합물을 첨가함으로써, 다양한 돌연변이 패턴이 존재하는 핵산 라이브러리가 생성된다. 예를 들어, 돌연변이 핵산 라이브러리의 일부 멤버가 특정 위치에서 이의 전구체 서열을 유지하면서 다른 멤버가 이러한 위치에서 돌연변이인 것이 요구된다면, 비돌연변이성 프라이머는 주어진 잔기에 대한 핵산 라이브러리 내 비돌연변이 멤버의 특정 수준을 수득하기 위한 능력을 제공한다. 본 발명의 방법은 일반적으로 약 10 내지 약 50 염기 길이, 보다 바람직하게는 약 15 내지 약 45 염기 길이인 돌연변이성 및 비돌연변이성 올리고뉴클레오티드를 사용한다. 상응하는 돌연변이성 및 비돌연변이성 프라이머에 관해, 상응하는 올리고뉴클레오티드가 동일한 길이의 것일 필요는 없으나, 첨가될 돌연변이에 상응하는 부위에서 중복이 있을 뿐이다.
- <137> 일부 구현예에서, 프라이머는 예비 정의된 비율로 첨가된다. 예를 들어, 생성된 라이브러리가 큰 수준의 특정한 특이적 돌연변이 및 더 적은 양의 상이한 돌연변이를 동일 또는 상이한 위치에서 가지는 것이 요구된다면, 첨가될 프라이머의 양을 조정함으로써 원하는 편향된 라이브러리가 생성된다. 대안적으로, 더 적은 또는 많은 양의 비돌연변이성 프라이머를 첨가함으로써, 상응하는 돌연변이(들)가 돌연변이 핵산 라이브러리에서 생성되는 주기가 조정된다. 위치 지정 돌연변이 유발에 대한 프라이머의 조성물을 포함하는 키트가 시판되며 QuikChange[®] 위치 지정 돌연변이 유발 키트 (Stratagene, San Diego, CA)가 포함된다. 일부 구현예에서, 전구체 폴리뉴클레오티드는 pro 부위에서 둘 이상의 아미노산 치환부를 포함하는 개질된 프로테아제를 인코딩하도록 추가로 개질된다. 다중-위치 지정 돌연변이 유발을 수행하기 위한 키트도 또한 시판된다 (예를 들어 QuikChange[®] Multisite, Stratagene, San Diego, CA).
- <138> 일부 구현예에서, 본 발명의 개질된 폴리뉴클레오티드는 위치-포화 방법을 사용하여 생성된다. 일부 다른 구현예에서, 각각의 개질된 폴리뉴클레오티드는 pro 부위에서 하나 이상의 아미노산 치환부를 포함한다.
- <139> 본 발명은 본 발명의 개질된 폴리뉴클레오티드 중 임의 하나에 의해 인코딩되는 개질된 전체 길이 전구체 프로테아제를 제공한다. 본 발명은 프로테아제의 성숙 형태를 제공하는데 이는 개질된 전구체 프로테아제로부터 가공되는 경우 비개질된 전구체 프로테아제의 가공에 의해 이루어지는 이들 수준보다 더 큰 수준에서 생성될 수 있다.
- <140> 본 발명은 전구체 프로테아제의 pro 부위 중 하나 이상의 아미노산을 돌연변이시킴으로써 개질된 전체 길이 프로테아제를 포함한다. 일부 구현예에서, 전구체 프로테아제는 SEQ ID NO:13의 전체 길이 전구체 프로테아제

의 pro 부위 (SEQ ID NO:15)의 아미노산 28-109, SEQ ID NO:5의 전체 길이 전구체 프로테아제의 pro 부위 (SEQ ID NO:7) 또는 SEQ ID NO:244의 전체 길이 전구체 프로테아제의 pro 부위 (SEQ ID NO:246)와 동등한 pro 부위 중 하나 이상의 아미노산에서 돌연변이된다. 일부 구현예에서, 아미노산 치환부는 B. 클라우시 V049 전구체 프로테아제 (SEQ ID NO:13), B. 클라우시 야생형 프로테아제 (SEQ ID NO:5) 또는 야생형 B. 렌투스 프로테아제 (SEQ ID NO:244)의 E33, E43, A44, E47, V49, E57, A59, E63, E70, E74, E84 및/또는 E88과 동등한 위치에서 만들어진다.

<141> 일부 구현예에서, E33과 동등한 위치에서의 아미노산의 치환에는 E33D, E33I, E33S, E33N, E33K, E33H, E33Q 또는 E33R이 포함되나 이에 제한되지는 않는다. 다른 구현예에서, E57의 치환에는 E57F, E57W, E57K, E57R, E57D, E57M, E57C, E57Q, E57G, E57S, E57H 또는 E57N이 포함되나 이에 제한되지는 않는다. 본 발명은 개질된 프로테아제를 포함하는데 여기서 SEQ ID NO:5, 13 또는 244의 아미노산 28-109와 동등한 각각의 아미노산 위치에서의 임의 하나의 아미노산 치환부는 전구체 프로테아제의 것과 비교하여 개질된 프로테아제의 생성/분비가 증강되도록, 자연적으로 발생하는 L-아미노산 19개 중 임의 하나로 만들어진다 (표 2 및 실시예 4 참조). 흔히 사용되고 이해되는 한 글자 코드로써 아미노산을 본원에 나타낸다 (예를 들어 Dale, M.W. (1989), *Molecular Genetics of Bacteria*, John Wiley & Sons, Ltd.).

<142> 일부 구현예에서, 개질된 전구체 프로테아제 폴리펩티드는 SEQ ID NOS:8, 16 또는 247에서 나타낸 아미노산 서열에 대해 약 65% 이상의 아미노산 서열 동일성, 바람직하게는 약 70% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 75% 이상의 아미노산 서열 동일성, 보다 더 약 80% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 85% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 90% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 92% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 95% 이상의 아미노산 서열 동일성, 보다 바람직하게는 약 97% 이상의 아미노산 서열 동일성, 보다 더 바람직하게는 약 98% 이상의 아미노산 서열 동일성, 및 가장 바람직하게는 약 99% 이상의 아미노산 서열 동일성을 공유하는 성숙 부위를 포함하며, 전구체 폴리펩티드에 대해 비교할만한 또는 증강된 생성 활성을 갖는다.

<143> 상기 나타낸 바와 같이, 일부 구현예에서, 본 발명은 전술한 폴리뉴클레오티드를 포함하는 벡터를 제공한다. 일부 바람직한 구현예에서, 벡터는 본 발명의 프로테아제를 인코딩하는 DNA 서열이 DNA의 전사에 필요한 추가적인 단편에 작동적으로 연결되는 발현 벡터이다. 일부 바람직한 구현예에서, 발현 벡터는 플라스미드 또는 바이러스 DNA에서 유래하거나, 또는 대안적인 구현예에서, 둘 모두의 요소를 함유한다. 예시적인 벡터에는 pXX, pC194, pJH101, pE194, pHP13이 포함되나 이에 제한되지는 않는다 (Harwood and Cutting (eds), *Molecular Biological Methods for Bacillus*, John Wiley & Sons, [1990], 특히 3장; 92페이지에 실린 것을 포함하는 B. 서브틸리스용 적절한 복제 플라스미드; Perego, M. (1993) *Integrational Vector for Genetic Manipulations in Bacillus subtilis*, p. 615-624; A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis and other Gram-positive bacteria: biochemistry, physiology and molecular genetics*, American Society for Microbiology, Washington, D.C.).

<144> 일부 바람직한 구현예에서, 벡터 pXX가 본원에 기재된 폴리뉴클레오티드를 포함하는 벡터의 구성에 사용된다 (예를 들어 pXX-049; 도 6 참조). 본원에 기재된 각각의 벡터가 본 발명에 사용되는 것으로 의도된다. 일부 구현예에서 구성물은 복제 플라스미드 (예를 들어 pHP13)에 존재하는 반면, 다른 구현예에서는 하나 이상의 복제물에서 염색체에 통합된다. 통합을 위한 위치의 예에는 aprE, amyE, veg 또는 pps 부위가 포함되나 이에 제한되지는 않는다. 실제로, 당업자에게 공지된 다른 위치가 본 발명에 사용될 것으로 예견된다. 일부 구현예에서, 촉진제는 선택된 전구체 프로테아제에 대한 야생형 촉진제이다. 일부 다른 구현예에서, 촉진제는 전구체 프로테아제에 대해 이중이나, 숙주 세포에서 기능적이다. 특별히, 세균성 숙주 세포에서 사용하기 위한 적절한 촉진제의 예에는 pSPAC, pAprE, pAmyE, pVeg, pHpaII 촉진제, B. 스테아로써모필루스 말토제닉 아밀라아제 유전자, B. 아밀로리쿠에파시엔스 (BAN) 아밀라아제 유전자, B. 서브틸리스 알칼리 프로테아제 유전자, B. 클라우시 알칼리 프로테아제 유전자, B. 푸밀루스 자일로스이다아제 유전자, B. 투링기엔시스 cryIIIA 및 B. 리체니포르미스 알파-아밀라아제 유전자의 촉진제가 포함되나 이에 제한되지는 않는다. 추가적인 촉진제에는 A4 촉진제 뿐 아니라 파지 랩다 P_R 또는 P_L 촉진제, 및 대장균 lac, trp 또는 tac 촉진제를 포함하나 이에 제한되지는 않는다.

<145> 일부 바람직한 구현예에서, 발현 벡터는 바람직하게는 벡터에 대해 고유한 하나 이상의 제한 엔도뉴클레아제 위치를 포함하는 다중 클로닝 위치 카세트르를 함유하여, 핵산 조각의 용이함을 촉진시킨다. 일부 추가적인 바람직한 구현예에서, 벡터는 또한 하나 이상의 선택가능한 표식 (예를 들어 항미생물 표식 예컨대 에리스로마이

신, 악티노마이신, 클로람페니콜 및/또는 테트라시클린)을 포함한다. 다른 구현예에서, 당업계에서 공지된 방법을 사용하여 다중복제물 복제 플라스미드가 바실러스 게놈 DNA로의 플라스미드 통합에 사용된다.

<146> 관심 있는 단백질(들) 예를 들어 프로테아제의 세포에서의 발현 및 생성을 위해, 개질된 프로테아제를 인코딩하는 폴리뉴클레오티드의 하나 이상의 복제물, 바람직하게는 다중 복제물을 포함하는 하나 이상의 발현 벡터가 단백질(들)의 발현에 적절한 조건 하에 세포에 형질전환된다. 일부 특히 바람직한 구현예에서 관심 있는 단백질 예를 들어 프로테아제를 인코딩하는 서열 (뿐만 아니라 벡터에 포함되는 다른 서열)이 숙주 세포의 게놈에 통합되는 반면, 다른 구현예에서는 플라스미드가 자율 염색체의 요소로서 세포 내에 잔존한다. 따라서, 본 발명은 염색체의 요소 뿐만 아니라 숙주 세포 게놈에 통합되는 도입 서열을 모두 제공한다.

<147> 전구체 및 개질된 프로테아제는 세균 및 진균을 포함하는 임의 적절한 그람-양성 미생물의 숙주 세포에서 생성된다. 예를 들어 일부 구현예에서, 개질된 프로테아제는 진균 및/또는 세균 기원의 숙주 세포에서 생성된다. 일부 구현예에서, 숙주 세포는 바실러스 종, 스트렙토마이세스 종, 대장균 종 또는 아스피릴러스 종이다. 일부 바람직한 구현예에서, 개질된 프로테아제는 바실러스 속의 숙주 세포에 의해 생성된다. 본 발명의 개질된 단백질의 생성에 사용되는 바실러스 숙주 세포의 예에는 B. 리체니포르미스, B. 렌투스, B. 서브틸리스, B. 아밀로리쿠에파시엔스, B. 렌투스, B. 브레비스, B. 스테아로써모필루스, B. 알칼로필루스, B. 코아굴란스, B. 썬콜란스, B. 푸밀루스, B. 투링기엔시스, B. 클라우시 및 B. 메가테리움을 포함하나 이에 제한되는 것을 뿐만 아니라 바실러스 속의 다른 생물도 포함된다. 일부 특히 바람직한 구현예에서, B. 서브틸리스 숙주 세포가 사용된다. 미국 특허 제 5,264,366호 및 제 4,760,025호 (RE 34,606)는 본 발명에서 사용되는 다양한 바실러스 숙주 균주를 기재하지만, 다른 적절한 균주가 본 발명에서 사용된다.

<148> 본 발명에서 사용되는 산업 균주에는 비재조합 (즉 야생형) 바실러스 균주 뿐만 아니라 자연적으로 발생하는 균주의 변종 및/또는 재조합 균주가 포함된다. 일부 바람직한 구현예에서, 숙주 균주는 재조합 균주이며, 여기서 관심 있는 폴리펩티드를 인코딩하는 폴리뉴클레오티드가 숙주로 도입된다. 일부 바람직한 구현예에서, 숙주 균주는 B. 서브틸리스 숙주 균주, 및 특히 재조합 바실러스 서브틸리스 숙주 균주이다. 많은 B. 서브틸리스 균주가 공지되어 있으며, 1A6 (ATCC 39085), 168 (1A01), SB19, W23, Ts85, B637, PB1753부터 PB1758, PB3360, JH642, 1A243 (ATCC 39,087), ATCC 21332, ATCC 6051, MI113, DE100 (ATCC 39,094), GX4931, PBT 110 및 PEP 211 균주 (예를 들어 Hoch 등, Genetics, 73:215-228 [1973] 참조) (또한 각각의 전문가 본원에 참고문헌으로 포함되는 미국 특허 제 4,450,235호; 미국 특허 제 4,302,544호; 및 EP 0134048 참조)까지 포함되나 이에 제한되는 것은 아니다. 발현 숙주로서의 B. 서브틸리스의 사용이 당업계에 잘 공지되어 있다 (예를 들어 Palva 등, Gene 19:81-87 [1982]; Fahnestock and Fischer, J. Bacteriol., 165:796-804 [1986]; 및 Wang 등, Gene 69:39-47 [1988] 참조).

<149> 일부 구현예에서, 바람직한 바실러스 숙주는 degU, degS, degR 및 degQ의 유전자 중 하나 이상에서의 돌연변이 또는 결실을 포함하는 바실러스 종이다. 바람직하게는 돌연변이는 degU 유전자이고, 보다 바람직하게는 돌연변이는 degU(Hy)32이다 (예를 들어 Msadek 등, J. Bacteriol., 172:824-834 [1990]; 및 Olmos 등, Mol. Gen. Genet., 253:562-567 [1997] 참조). 보다 특히 바람직한 숙주 균주는 degU32(Hy) 돌연변이를 수반하는 바실러스 서브틸리스이다. 일부 추가적인 구현예에서, 바실러스 균주는 scoC4 (예를 들어 Caldwell 등, J. Bacteriol., 183:7329-7340 [2001] 참조); spoIIE (Arigoni 등, Mol. Microbiol., 31:1407-1415 [1999] 참조); 및/또는 oppA 또는 opp 오페론의 다른 유전자 (예를 들어 Perego 등, Mol. Microbiol., 5:173-185 [1991] 참조)에서의 돌연변이 또는 결실을 포함한다. 실제로, oppA 유전자에서의 돌연변이로서 동일한 표현형을 일으키는 opp 오페론에서의 임의 돌연변이가 본 발명의 변경된 바실러스 균주의 일부 구현예에서 사용될 것임이 예견된다. 일부 구현예에서 이들 돌연변이는 단독으로 발생하는 반면, 다른 구현예에서는 돌연변이의 조합이 존재한다. 일부 구현예에서, 본 발명의 개질된 프로테아제를 생성시키는데 사용될 수 있는 변경된 바실러스는, 상술한 유전자 중 하나 이상에서의 돌연변이를 이미 포함하는 바실러스 숙주 균주이다 (일부 구현예에서, 다른 유전자에서의 돌연변이가 또한 존재한다). 일부 대안적인 구현예에서, 상술한 유전자 중 하나 이상의 돌연변이가 포함되도록 추가적으로 공학처리된 변경된 바실러스가 사용된다.

<150> 숙주 세포는 당업계에 공지된 임의 적절한 방법을 사용하여, 본 발명의 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드로 형질전환된다. 개질된 폴리뉴클레오티드는 벡터에 혼입되거나, 플라스미드 DNA 없이 사용되어, 이는 미생물, 일부 구현예에서 바람직하게는 대장균 세포 또는 적격 바실러스 세포에 도입된다. 플라스미드 구성물 및 대장균으로의 플라스미드의 형질전환을 포함하는, 바실러스 세포에 DNA를 도입하는 방법이 잘 공지되어 있다. 일부 구현예에서, 플라스미드는 이후 대장균으로부터 단리되고 바실러스에 형질전환된다. 그러나, 대장균과 같은 중개 미생물을 사용하는 것이 필요한 것은 아니며, 일부 구현예에서

DNA 구성물 또는 벡터는 바실러스 숙주에 직접적으로 도입된다.

- <151> 당업자는 바실러스 세포에 폴리뉴클레오티드 서열을 도입하는 적절한 방법을 알고 있다 (예를 들어 Ferrari 등, "Genetics," in Harwood 등, (ed.), *Bacillus*, Plenum Publishing Corp. [1989], 페이지 57-72; Saunders 등, *J. Bacteriol.*, 157:718-726 [1984]; Hoch 등, *J. Bacteriol.*, 93:1925-1937 [1967]; Mann 등, *Current Microbiol.*, 13:131-135 [1986]; 및 Holubova, *Folia Microbiol.*, 30:97 [1985]; Chang 등, *Mol. Gen. Genet.*, 168:11-115 [1979]; Vorobjeva 등, *FEMS Microbiol. Lett.*, 7:261-263 [1980]; Smith 등, *Appl. Env. Microbiol.*, 51:634 [1986]; Fisher 등, *Arch. Microbiol.*, 139:213-217 [1981]; 및 McDonald, *J. Gen. Microbiol.*, 130:203 [1984] 참조). 실제로, 원형질체 형질전환 및 회합, 형질도입 및 원형질체 융합을 포함하는 형질전환과 같은 방법이 잘 공지되며, 본 발명에서의 사용에 알맞다. 형질전환의 방법은 본 발명에 의해 제공되는 DNA 구성물을 숙주 세포에 도입시키기에 특히 바람직하다.
- <152> 흔히 사용되는 방법에 추가적으로, 일부 구현예에서는 숙주 세포가 직접적으로 형질전환된다 (즉 증폭시키는데 중간 세포가 사용되지 않거나, 또는 다른 방법에서, 숙주 세포에의 도입 전 DNA 구성물이 사용됨). DNA 구성물의 숙주 세포로의 도입에는 플라스미드 또는 벡터에의 삽입 없이 숙주 세포에 DNA를 도입시키기 위한 당업계에 공지된 이들 물리적 및 화학적 방법이 포함된다. 이러한 방법에는 염화칼슘 침전, 전기 천공, 네이키드 DNA (naked DNA), 리포솜 등이 포함되나 이에 제한되지는 않는다. 추가적인 구현예에서, DNA 구성물은 플라스미드에 삽입되지 않고 플라스미드로 공동 형질전환된다. 추가적인 구현예에서, 선택적 표식은 당업계에 공지된 방법에 의해 변경된 바실러스 균주로부터 결실된다 (Stahl 등, *J. Bacteriol.*, 158:411-418 [1984]; 및 Palmeros 등, *Gene* 247:255-264 [2000] 참조).
- <153> 상기 나타낸 바와 같이, 본 발명의 일부 구현예에서, 관심 있는 하나 이상의 개질된 폴리펩티드를 인코딩하는 핵산은 숙주 세포 내에서 복제할 수 있는 발현 벡터를 통해 숙주 세포에 도입된다. 바실러스용 적절한 복제 및 통합 플라스미드가 공지되어 있다 (예를 들어 Harwood and Cutting (eds), *Molecular Biological Methods for Bacillus*, John Wiley & Sons, [1990], 특히 3장; 92페이지에 기재된 것을 포함하는 B. 서브틸리스용 적절한 복제 플라스미드). 기술적인 장애물이 있지만, 당업자는 바실러스에서의 DNA의 직접적인 클로닝을 위한 여러 전략이 있음을 알고 있다.
- <154> 바실러스를 형질전환시키기 위한 당업계에 공지된 방법에는, 부분 상동인 상주 플라스미드를 수반하는 적격 세포에 의해 공여 플라스미드를 취하는 것을 포함하는, 플라스미드 표식 구출 형질전환과 같은 방법이 포함된다 (Contente 등, *Plasmid* 2:555-571 [1979]; Haima 등, *Mol. Gen. Genet.*, 223:185-191 [1990]; Weinrauch 등, *J. Bacteriol.*, 154:1077-1087 [1983]; 및 Weinrauch 등, *J. Bacteriol.*, 169:1205-1211 [1987]). 상기 방법에서, 도입 공여 플라스미드는 염색체 형질전환을 모방하는 방법 중 상주 "헬퍼" 플라스미드의 상동 부위와 재조합된다.
- <155> 원형질체 형질전환에 의한 형질전환을 포함하는 다른 방법이 당업계에 공지되어 있다 (예를 들어 Chang and Cohen, *Mol. Gen. Genet.*, 168:111-115 [1979]; Vorobjeva 등, *FEMS Microbiol. Lett.*, 7:261-263 [1980]; Smith 등, *Appl. Env. Microbiol.*, 51:634 [1986]; Fisher 등, *Arch. Microbiol.*, 139:213-217 [1981]; McDonald [1984] *J. Gen. Microbiol.*, 130:203 [1984]; 및 Bakhiet 등, 49:577 [1985] 참조). 추가적으로, Mann 등, (Mann 등, *Curr. Microbiol.*, 13:131-135 [1986])은 바실러스 원형질체의 형질전환을 기재하며, Holubova (Holubova, *Microbiol.*, 30:97 [1985])는 DNA 함유 리포솜을 사용하여 DNA를 원형질체에 도입시키는 방법을 기재한다. 일부 바람직한 구현예에서, 관심 있는 유전자가 숙주 세포에 존재하는지 아닌지를 나타내기 위해 표식 유전자가 사용된다.
- <156> 이들 방법에 추가적으로, 다른 구현예에서 숙주 세포는 직접적으로 형질전환된다. "직접 형질전환"에서, 증폭시키는데 중간 세포가 사용되지 않거나, 다른 방법에서, 숙주 (즉 바실러스) 세포에의 도입 전 개질된 폴리뉴클레오티드가 사용된다. 개질된 폴리뉴클레오티드의 숙주 세포에의 도입에는 플라스미드 또는 벡터에의 삽입 없이 숙주 세포에 개질된 폴리뉴클레오티드를 도입시키기 위한 당업계에 공지된 이들 물리적 및 화학적 방법이 포함된다. 이러한 방법에는 적격 세포 뿐 아니라 DNA를 세포에 도입시키기 위한 염화칼슘 침전, 전기 천공 등과 같은 "인공 수단"의 사용이 포함되나 이에 제한되지는 않는다. 따라서, 본 발명에서는 네이키드 DNA, 리포솜 등이 사용된다. 다른 구현예에서, 개질된 폴리뉴클레오티드는 플라스미드에 삽입되지 않고 플라스미드로 공동 형질전환된다.
- <157> 보다 특히, 본 발명은 구성물, 본원에 기재된 폴리뉴클레오티드를 포함하는 벡터, 이러한 벡터로 형질전환된 숙주 세포, 이러한 숙주 세포에 의해 발현된 프로테아제, 발현 방법 및 미생물 (특히 바실러스 속의 멤버) 유래의

상동 또는 이종 세린 프로테아제 효소를 생성시키기 위한 시스템을 제공한다. 일부 구현예에서, 개질된 세린 프로테아제(들)를 인코딩하는 개질된 폴리뉴클레오티드(들)가 개질된 세린 프로테아제(들)의 발현에 적절한 재조합 숙주 세포를 생성시키는데 사용된다. 일부 바람직한 구현예에서, 발현 숙주는 개질된 프로테아제(들)의 성숙 형태의 분비를 증가시킬 수 있어, 이에 따라 프로테아제의 상업적 생산을 증가시킨다.

<158> 일부 구현예에서, 본 발명의 숙주 세포 및 형질전환된 세포는 통상적인 영양 배지에서 배양된다. 온도, pH 등과 같은 적절한 특정 배양 조건이 당업자에게 공지되어 있다. 추가적으로, 일부 바람직한 배양 조건이 [Hopwood (2000) Practical Streptomyces Genetics, John Innes Foundation, Norwich UK; Hardwood 등, (1990) Molecular Biological Methods for Bacillus, John Wiley and from the American Type Culture Collection (ATCC)]과 같은 과학 문헌에서 발견될 수 있다.

<159> 일부 구현예에서, 개질된 프로테아제를 인코딩하는 폴리뉴클레오티드 서열로 형질전환된 숙주 세포는 세포 배양 물로부터의 인코딩된 단백질의 발현 및 회수에 적절한 조건 하에 배양된다. 본 발명의 개질된 프로테아제를 포함하는 재조합 숙주 세포에 의해 생성된 단백질은 배양 배지에 분비된다. 일부 구현예에서, 다른 재조합 구성물은 가용성 단백질의 정제를 촉진하는 프로테아제 폴리펩티드 도메인을 인코딩하는 뉴클레오티드 서열에 대해 이종 또는 상동인 폴리뉴클레오티드 서열과 결합한다 (Kroll DJ 등 (1993) DNA Cell Biol 12:441-53).

<160> 이러한 정제는 비고정화된 금속 상에서 정제되게 하는 히스티딘-트립토판 모듈과 같은 금속 킬레이트 펩티드 (Porath J (1992) Protein Expr Purif 3:263-281), 비고정화된 면역글로불린 상에서 정제되게 하는 단백질 A 도메인, 및 FLAGS 연장/친화성 정제 시스템 (Immunex Corp, Seattle WA)에서 사용된 도메인을 포함하나 이에 제한되지는 않는다. 정제 도메인 및 이종 단백질 사이의 요소 XA 또는 엔테로키나아제 (Invitrogen, San Diego CA)와 같은 절단가능한 연결체 서열의 세포 함유물이 또한 정제를 촉진시키는데 사용된다.

<161> 일부 바람직한 구현예에서, 이종 또는 상동 단백질을 인코딩하거나 또는 상기 단백질을 내생적으로 갖는 폴리뉴클레오티드 서열로 형질전환된 세포는 세포 배양 배지로부터의 인코딩된 단백질의 발현 및 회수에 적절한 조건 하에 배양된다. 일부 구현예에서, 다른 재조합 구성물은 가용성 단백질 (예를 들어 다양한 종류의 태그)의 정제를 촉진시키는 폴리펩티드 도메인을 인코딩하는 뉴클레오티드 서열에 대한 이종 또는 상동 폴리뉴클레오티드 서열을 포함한다 (Kroll 등, DNA Cell. Biol., 12:441-53 [1993]).

<162> 이러한 정제 촉진 도메인에는 비고정화된 금속 상에서 정제되게 하는 히스티딘-트립토판 모듈과 같은 금속 킬레이트 펩티드 (Porath, Prot. Expr. Purif., 3:263-281 [1992]), 비고정화된 면역글로불린 상에서 정제되게 하는 단백질 A 도메인, 및 FLAGS 연장/친화성 정제 시스템 (Immunex Corp, Seattle WA)에서 사용된 도메인이 포함되나 이에 제한되지는 않는다. 정제 도메인 및 이종 단백질 사이의 요소 XA 또는 엔테로키나아제 (Invitrogen, San Diego CA)와 같은 절단가능한 연결체 서열의 세포 함유물이 정제를 촉진시키는데 사용된다.

<163> 일부 바람직한 구현예에서, 본 발명의 형질전환된 숙주 세포는 생성된 프로테아제가 배양물로부터 회수된 후 본 프로테아제의 발현을 허용하는 조건 하에 적절한 영양 배지에서 배양된다. 세포를 배양하는데 사용된 배지는 숙주 세포를 성장시키는데 적절한 임의 통상적인 배지 예컨대 적합한 보충물을 함유하는 최소 또는 복합 배지를 포함한다. 적절한 배지는 상업적 공급자로부터 활용가능하거나, 또는 공개된 방법 {예를 들어 미국 미생물 보존 센터 (American Type Culture Collection)의 카탈로그 중}에 따라 제조될 수 있다. 일부 구현예에서, 세포에 의해 생성된 프로테아제는 통상적인 방법에 의해 배양 배지로부터 회수되는데, 이는 원심분리 또는 여과에 의한 배지로부터의 숙주 세포의 분리, 염을 사용함에 의한 (예를 들어 암모늄 설페이트) 여과액 또는 상청액의 단백질성 성분의 침전, 크로마토그래피적 정제 (예를 들어 이온 교환, 겔 여과, 친화성 등)가 포함되나 이에 제한되지는 않는다. 따라서, 본 발명의 프로테아제(들)를 회수하기에 적절한 임의 방법이 본 발명에 사용된다. 실제로, 본 발명이 임의 특정 정제 방법에 제한되도록 의도되지는 않는다.

<164> 상기 나타낸 바와 같이, 본 발명의 폴리펩티드는 이의 상응하는 비개질된 전구체 폴리펩티드로부터 가공된 성숙 효소보다 더 큰 수준에서 성숙 효소로서 생성된다. 본 발명의 바람직한 구현예에서, 전구체 폴리펩티드의 pro 부위 내의 돌연변이는, 동일한 조건 하에 바실러스 균주에 의해 생성되는 경우의 전구체 프로테아제로부터 가공된 상응하는 성숙 프로테아제와 비교하는 경우 성숙 폴리펩티드의 분비/발현을 증가시킨다.

<165> 증강의 측정은 활성 비율로서 측정될 수 있는데, 이는 개질된 프로테아제로부터 가공된 성숙 형태의 효소적 활성 대 전구체 프로테아제로부터 가공된 성숙 형태의 효소적 활성의 비율로서 표현될 수 있다. 1 이상의 비율은, 개질된 프로테아제의 성숙 형태가 전구체 프로테아제의 성숙 형태가 생성된 수준 이상의 수준에서 생성된다는 것을 나타낸다. 예를 들어, 1.5의 활성 비율은 개질된 프로테아제로부터 가공된 성숙 프로테아제가 전

구체 프로테아제로부터 가공된 성숙 프로테아제에서의 수준의 1.5배에서 생성된다는 것을 나타내며, 즉 개질된 프로테아제는 비개질된 전구체 프로테아제보다 50% 초과된 성숙 프로테아제를 산출시킨다. 일부 구현예에서, 활성 비율은 1 이상, 약 1.05 이상, 약 1.1 이상, 약 1.2 이상, 약 1.3 이상, 약 1.4 이상, 약 1.5 이상, 약 1.6 이상, 약 1.7 이상, 약 1.8 이상, 약 1.9 이상 및 약 2 이상이다. 다른 구현예에서, 활성 비율은 약 2.1 이상, 약 2.2 이상, 약 2.3 이상, 약 2.4 이상, 약 2.5 이상, 약 2.6 이상, 약 2.7 이상, 약 2.8 이상, 약 2.9 이상 및 약 3 이상이다. 다른 구현예에서, 활성 비율은 약 3.5 이상, 약 4.0 이상 및 약 5 이상이다. 따라서, 일부 구현예에서 개질된 프로테아제로부터 가공된 성숙 프로테아제의 생성은 비개질된 전구체 프로테아제로부터 가공된 상응하는 성숙 프로테아제와 비교하여 약 0.5% 이상, 약 1.0%, 약 1.5%, 약 2.0%, 약 2.5%, 약 3.0%, 약 4.0%, 약 5.0%, 약 8.0%, 약 10%, 약 15%, 약 20%, 약 25%, 약 30%, 약 40%, 약 50%, 약 60% 이상, 약 70% 이상, 약 80% 이상, 약 90% 이상, 약 100% 이상으로 증강된다. 다른 구현예에서, 개질된 프로테아제로부터 가공된 성숙 형태의 프로테아제의 생성은 비개질된 전구체 프로테아제로부터 가공된 성숙 형태의 프로테아제의 상응하는 생성과 비교하여 약 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190% 이상, 및 최대 약 200% 이상으로 증강된다. 일부 구현예에서, 개질된 프로테아제의 증강된 생성은 상응하는 비개질된 전구체 프로테아제의 성숙 형태의 단백질분해적 활성과 비교하여 개질된 프로테아제로부터 가공된 성숙 형태의 단백질분해적 활성의 비율을 기준으로 측정된다.

<166> 숙주 세포에서의 이중 또는 상동 단백질의 분비 수준을 측정하고 분비된 단백질을 검출하기 위한 다른 수단에는 단백질에 대해 특이적인 다클론성 또는 단클론성 항체의 사용이 포함되나 이에 제한되지는 않는다. 예에는 효소-연결 면역흡착제 분석 (ELISA), 방사선면역분석 (RIA), 형광 면역분석 (FIA) 및 형광 활성 세포 분류 (FACS)가 포함되나 이에 제한되지는 않는다. 이들 및 다른 분석이 당업계에서 잘 공지되어 있다 (예를 들어 Maddox 등, *J. Exp. Med.*, 158:1211 [1983] 참조). 본 발명의 일부 바람직한 구현예에서, 동일한 방법 또는 조성물을 사용하는 경우보다 본원에 제공된 방법 및 조성물을 사용하는 것이 분비가 더 높으나, 펩티드 수송 단백질 또는 펩티드 수송 오페론의 유전자 생성물은 도입되지 않는다.

<167> 본 발명의 폴리펩티드의 활성을 검출 및 측정하기 위한, 당업계의 숙련자에게 공지된 다양한 분석이 있다. 특히, 카세인 또는 헤모글로빈으로부터의 산-가용성 펩티드의 방출을 기준으로 하는 프로테아제 활성 측정에 활용가능한 분석은 폴린 (Folin) 방법을 사용하여 280 nm에서의 흡광도로서 또는 측정적으로 측정된다 (예를 들어 Bergmeyer 등, "Methods of Enzymatic Analysis" vol. 5, Peptidases, Proteinases and their Inhibitors, Verlag Chemie, Weinheim [1984] 참조). 일부 다른 분석에는 발색성 기질의 가용화가 포함된다 (예를 들어 Ward, "Proteinases," in Fogarty (ed.), *Microbial Enzymes and Biotechnology*, Applied Science, London, [1983], pp 251-317 참조). 다른 예시적인 분석에는 숙시닐-Ala-Ala-Pro-Phe-파라 니트로아닐라이드 분석 (SAAPFpNA) 및 2,4,6-트리니트로벤젠 설포네이트 나트륨 염 분석 (TNBS 분석)이 포함되나 이에 제한되지는 않는다. 당업자에게 공지된 많은 추가적인 참고 문헌이 적절한 방법을 제공한다 (예를 들어 Wells 등, *Nucleic Acids Res.* 11:7911-7925 [1983]; Christianson 등, *Anal. Biochem.*, 223:119-129 [1994]; 및 Hsia 등, *Anal Biochem.*, 242:221-227 [1999] 참조). 본 발명이 임의 특정 분석 방법(들)에 제한되도록 의도되지는 않는다.

<168> 일부 구현예에서, 미생물에 의한 개질된 프로테아제의 생성은 비개질된 전구체 프로테아제로부터 가공된 성숙 프로테아제의 활성과 비교된 개질된 전구체 프로테아제로부터 가공된 성숙 프로테아제의 활성의 비율을 사용함으로써 측정된다. 일부 특히 바람직한 구현예에서, 1 이상의 비율이 바람직하다.

<169> 관심 있는 단백질, 예를 들어 프로테아제의 숙주 세포에서의 생성 수준을 측정하고 발현된 단백질을 검출하기 위한 다른 수단에는 단백질에 대해 특이적인 다클론성 또는 단클론성 항체로의 면역분석의 사용이 포함된다. 예에는 효소-연결 면역흡착제 분석 (ELISA), 방사선면역분석 (RIA), 형광 면역분석 (FIA) 및 형광 활성 세포 분류 (FACS)가 포함된다. 그러나, 다른 방법이 당업자에게 공지되며 관심 있는 단백질을 평가하는데 사용된다 (예를 들어 Hampton 등, *Serological Methods, A Laboratory Manual*, APS Press, St. Paul, MN [1990]; 및 Maddox 등, *J. Exp. Med.*, 158:1211 [1983] 참조). 일부 바람직한 구현예에서, 관심 있는 단백질의 분비는, 상응하는 비변경된 숙주에서보다 본 발명을 사용하여 수득된 변경된 균주에서 더 높다. 당업계에 공지된 바와 같이, 본 발명을 사용하여 생성된 변경된 바실러스 세포는 세포 배양물로부터 관심 있는 폴리펩티드를 발현 및 회수하기에 적절한 조건 하에 유지 및 성장된다 (예를 들어 Hardwood and Cutting (eds.) *Molecular Biological Methods for Bacillus*, John Wiley & Sons [1990] 참조). 본 발명이 임의 특정 분석 방법(들)에 제한되도록 의도되지는 않는다.

<170> 본원에 언급된 모든 발행물 및 특허는 참고 문헌으로 본원에 포함된다. 본 발명의 기재된 방법 및 시스템의

다양한 수정 및 변형은 본 발명의 범주 및 취지에서 벗어남이 없이 당업자에게 명백할 것이다. 본 발명이 특정 바람직한 구현예와 관련되어 기재되지만, 발명은 이러한 특정 구현예에 과도하게 제한되지 않는 것으로서 이해되어야 한다. 실제로, 당업자 및/또는 관련 분야에 명백한, 본 발명을 실행하기 위해 기재된 방식의 다양한 수정은 본 발명의 범주 내에 있는 것으로 의도된다.

실시예

<171> 하기의 실시예는 본 발명의 측면 및 특정 바람직한 구현예를 추가로 설명하고 논증하기 위해 제공되며, 이의 범주를 제한하는 것으로서 파악되지 않는다.

<172> 하기의 실험적 개시에서, 하기의 약어가 적용된다: ppm (100만분의 1); M (몰라); mM (밀리몰라); μ M (마이크로몰라); nM (나노몰라); mol (몰); mmol (밀리몰); μ mol (마이크로몰); nmol (나노몰); gm (그램); mg (밀리그램); μ g (마이크로그램); pg (피코그램); L (리터); ml 및 mL (밀리리터); μ l 및 μ L (마이크로리터); cm (센티미터); mm (밀리미터); μ m (마이크로미터); nm (나노미터); U (단위); V (볼트); MW (분자량); sec (초); min (분); h 및 hr (시간); $^{\circ}$ C (섭씨 온도); QS (충분량); ND (비수행); NA (비해당); rpm (분당 회전수); H₂O (물); dH₂O (탈이온수); HCl (염산); aa (아미노산); bp (염기쌍); kb (킬로염기쌍); kD (킬로달톤); cDNA (복제물 또는 상보성 DNA); DNA (데옥시리보핵산); ssDNA (단일 가닥 DNA); dsDNA (이중 가닥 DNA); dNTP (데옥시리보뉴클레오티드 트리포스페이트); RNA (리보핵산); MgCl₂ (염화마그네슘); NaCl (염화나트륨); w/v (중량 대 부피); v/v (부피 대 부피); g (중력); OD (광학 밀도); 둘베코의 포스페이트 완충 용액 (DPBS); OD₂₈₀ (280 nm에서의 광학 밀도); OD₆₀₀ (600 nm에서의 광학 밀도); A₄₀₅ (405 nm에서의 흡광도); PAGE (폴리아크릴아미드 겔 전기영동); PBS (포스페이트 완충 식염수 [150 mM NaCl, 10 mM 나트륨 포스페이트 완충액, pH 7.2]); PBST (PBS+0.25% TWEEN[®]-20); PEG (폴리에틸렌 글리콜); PCR (중합효소 연쇄 반응); SDS (나트륨 도데실 설페이트); Tris (트리스(하이드록시메틸)아미노메탄); HEPES (N-[2-하이드록시에틸]피페라진-N-[2-에탄설포산]); HBS (HEPES 완충 식염수); SDS (나트륨 도데실설페이트); bME, BME 및 β ME (베타-머캅토에탄올 또는 2-머캅토에탄올); 트리스-HCl (트리스[하이드록시메틸]아미노메탄-하이드로클로라이드); 트리신 (N-[트리스-(하이드록시메틸)-메틸]-글리신); DMSO (디메틸 설펝시드); Taq (테르무스 아쿠아티쿠스 DNA 중합효소); 클레나우 (DNA 중합효소 I 대(클레나우) 절편); rpm (분당 회전수); EGTA (에틸렌 글리콜-비스(β -아미노에틸 에테르) N,N,N',N'-테트라아세트산); EDTA (에틸렌디아민테트라아세트산); bla (β -락타마아제 또는 암피실린-저항성 유전자); DNA2.0 (DNA2.0, Menlo Park, CA); OXOID (Oxoid, Basingstoke, Hampshire, UK); Corning (Corning Life Sciences, Corning, NY); ATCC (미국 미생물 보존 센터, Rockville, MD); Gibco/BRL (Gibco/BRL, Grand Island, NY); Sigma (Sigma Chemical Co., St. Louis, MO); Pharmacia (Pharmacia Biotech, Pisacataway, NJ); NCBI (국립생물정보센터); Applied Biosystems (Applied Biosystems, Foster City, CA); Clontech (CLONTECH Laboratories, Palo Alto, CA); Operon Technologies (Operon Technologies, Inc., Alameda, CA); Bachem (Bachem Bioscience, Inc., King of Prussia, PA); Difco (Difco Laboratories, Detroit, MI); GIBCO BRL 또는 Gibco BRL (Life Technologies, Inc., Gaithersburg, MD); Millipore (Millipore, Billerica, MA); Bio-Rad (Bio-Rad, Hercules, CA); Invitrogen (Invitrogen Corp., San Diego, CA); NEB (New England Biolabs, Beverly, MA); Sigma (Sigma Chemical Co., St. Louis, MO); Pierce (Pierce Biotechnology, Rockford, IL); Takara (Takara Bio Inc. Otsu, Japan); Roche (Hoffmann-La Roche, Basel, Switzerland); EM Science (EM Science, Gibbstown, NJ); Qiagen (Qiagen, Inc., Valencia, CA); Molecular Devices (Molecular Devices, Corp., Sunnyvale, CA); R&D Systems (R&D Systems, Minneapolis, MN); Stratagene (Stratagene Cloning Systems, La Jolla, CA); 및 Microsoft (Microsoft, Inc., Redmond, WA).

<173> **실시예 1. 바실러스 클라우시 알칼리 프로테아제 V049의 Pro 서열 및 SigP의 위치-스캐닝 돌연변이 유발**

<174> QuikChange[®] 위치-지정 돌연변이 유발 키트 (QC; Stratagene)을 사용하여 생산자에 의해 제공된 지침에 따라 바실러스 클라우시 알칼리 프로테아제 V049 전구체 프로테아제의 pro 서열의 위치-포화 돌연변이 유발을 수행하였다. B. 클라우시 프로테아제 변종 V049를 인코딩하는 전구체 폴리뉴클레오티드의 개질된 폴리뉴클레오티드 서열이 포함되도록 위치-포화 라이브러리를 생성시켰다. pXX-V049 플라스미드에 포함된 V049 폴리뉴클레오티드의 pro 부위를 인코딩하는 서열 (도 5 및 6; SEQ ID NO:17)을 돌연변이시켜 각각 전구체 프로테아제의 pro 부위의 한 코돈의 돌연변이를 포함하는 폴리뉴클레오티드의 라이브러리를 생성시켰다. NNG/C에 의해 예시된 V049의 pro 부위의 각각의 코돈을, 20개의 자연적으로 발생하는 아미노산을 인코딩하는 32개의 가능한 뉴클레오티드 삼중선에 의해 치환되도록 돌연변이시켰다. NNS에 측면 위치하는 18개의 염기로 관심 있는 각각

의 코돈에 대한 상보성 중복 프라이머를 설계하였고, 프라이머의 서열 (SEQ ID NO:18~239)을 표 1에 나타내었다.

표 1

프라이머*	프라이머 서열	염기 좌측**	염기 우측***	SEQ ID NO
V049-1F	TTAAAAGGAGAGGGTAAAGANNSAGAAGCAA AAAATTGTG	20	17	18
V049-2F	GGAGAGGGTAAAGAGTGNNSAGCAAAAATT GTGGATC	17	18	19
V049-3F	GAGGGTAAAGAGTGAGANNSAAAAAATTGTG GATCGTC	17	18	20
V049-4F	GGTAAAGAGTGAGAAGCNNSAAATTGTGGAT CGTCGC	17	17	21
V049-5F	GTAAGAGTGAGAAGCAAANNSTTGTGGATC GTCGCGTC	19	17	22
V049-6F	GAGTGAGAAGCAAAAANNSTGGATCGTCGC GTCGAC	17	17	23
V049-7F	GTGAGAAGCAAAAATTGNNSATCGTCGCGT CGACCGC	18	17	24
V049-8F	GAAGCAAAAATTGTGGNNSGTCGCGTCGAC CGCACTAC	17	19	25
V049-9F	GCAAAAATTGTGGATCNNSGCGTCGACCGC ACTACTC	17	18	26
V049-10F	CAAAAATTGTGGATCGTCNNSTCGACCGCA CTACTATTC	19	20	27
V049-11F	AAAAATTGTGGATCGTCGCGNNSACCGCACT ACTATTC	20	17	28
V049-12F	AATTGTGGATCGTCGCGTCGNNSGCACTACT CATTTCTGTTG	20	19	29
V049-13F	GGATCGTCGCGTCGACCNNSCTACTCATTC TGTTGC	17	17	30
V049-14F	GATCGTCGCGTCGACCGCANNSTCATTTCT GTTGCTTTAG	19	20	31
V049-15F	GTCGCGTCGACCGCACTANNSATTTCTGTTG CTTTAG	18	17	32
V049-16F	CGTCGACCGCACTACTCNNSCTGTTGCTTT TAGTTC	17	17	33

<176>

V049-17F	CGACCGCACTACTCATTNNSGTTGCTTTTAGT TCATC	17	17	34
V049-18F	CCGCACTACTCATTCTNNSGCTTTTAGTTCA TCGATC	17	18	35
V049-19F	CACTACTCATTCTGTTNNSSTTTAGTTCATCG ATCGC	17	17	36
V049-20F	CTACTCATTCTGTTGCTNNSAGTTCATCGAT CGCATC	18	17	37
V049-21F	CTCATTCTGTTGCTTTTNNSTCATCGATCGC ATCGGC	18	17	38
V049-22F	CATTTCTGTTGCTTTTAGTNNSTCGATCGCAT CGGCTGC	19	17	39
V049-23F	CTGTTGCTTTTAGTTCANNSATCGCATCGGCT GCTGAAG	17	19	40
V049-24F	GTTGCTTTTAGTTCATCGNNSGCATCGGCTG CTGAAGAAG	18	19	41
V049-25F	CTTTTAGTTCATCGATCNNSTCGGCTGCTGAA GAAGC	17	17	42
V049-26F	CTTTTAGTTCATCGATCGCANNSGCTGCTGA AGAAGCAAAAG	20	19	43
V049-27F	GTTCATCGATCGCATCGNNSGCTGAAGAAGC AAAAGAAA	17	20	44
V049-28F	CATCGATCGCATCGGCTNNSGAAGAAGCAAA AGAAAAATA	17	20	45
V049-29F	CGATCGCATCGGCTGCTNNSGAAGCAAAAGA AAAATATTT	17	20	46
V049-30F	GATCGCATCGGCTGCTGAANNSGCAAAAGAA AAATATTTAAT	19	20	47
V049-31F	CATCGGCTGCTGAAGAANNSAAAGAAAAATA TTTAATTG	17	19	48
V049-32F	CGGCTGCTGAAGAAGCANNSGAAAAATATTT AATTGG	17	17	49
V049-33F	CTGCTGAAGAAGCAAAANNSAAATATTTAATT GGCTTTAA	17	20	50
V049-34F	CTGAAGAAGCAAAAGAANNSTATTTAATTGGC TTTAATG	17	19	51

<177>

V049-35F	GAAGAAGCAAAAGAAAAANNSTTAATTGGCTT TAATGAG	18	18	52
V049-36F	GAAGCAAAAGAAAAATATNNSATTGGCTTTAA TGAGCAG	18	18	53
V049-37F	CAAAAGAAAAATATTTANNSGGCTTTAATGAG CAGGAAG	17	19	54
V049-38F	CAAAAGAAAAATATTTAATTNNSTTTAATGAG CAGGAAGC	20	17	55
V049-39F	GAAAAATATTTAATTGGCNNSAATGAGCAGG AAGCTGTC	18	18	56
V049-40F	AAAAATATTTAATTGGCTTTNNSGAGCAGGAA GCTGTCAG	20	17	57
V049-41F	AATATTTAATTGGCTTTAATNNSCAGGAAGCT GTCAGTGAG	20	18	58
V049-42F	ATTTAATTGGCTTTAATGAGNNSGAAGCTGTC AGTGAGTTTG	20	19	59
V049-43F	TAATTGGCTTTAATGAGCAGNNSGCTGTCAG TGAGTTTGAG	20	19	60
V049-44F	GCTTTAATGAGCAGGAANNSGTCAGTGAGTT TGAGAAC	17	19	61
V049-45F	CTTTAATGAGCAGGAAGCTNNSAGTGAGTTT GTAGAACAAG	19	19	62
V049-46F	TTAATGAGCAGGAAGCTGTCNNSGAGTTTGT AGAACAAGTAG	20	19	63
V049-47F	GAGCAGGAAGCTGTCAGTNNSTTTGTAGAAC AAGTAGAG	18	18	64
V049-48F	CAGGAAGCTGTCAGTGAGNNSGTAGAACAAG TAGAGGC	18	17	65
V049-49F	GAAGCTGTCAGTGAGTTNNSGAACAAGTAG AGGCAAATG	18	19	66
V049-50F	CTGTCAGTGAGTTTGTANNSCAAGTAGAGGC AAATGAC	17	18	67
V049-51F	GTCAGTGAGTTTGTAGAANNSGTAGAGGCAA ATGACGAG	18	18	68
V049-52F	GTGAGTTTGTAGAACAANNSGAGGCAAATGA CGAGGTC	17	18	69

<178>

V049-53F	GAGTTTGTAGAACAAGTANNNGCAAATGACG AGGTCGC	18	17	70
V049-54F	GTTTGTAGAACAAGTAGAGNNSAATGACGAG GTCGCCATTC	19	19	71
V049-55F	GTAGAACAAGTAGAGGCANNSGACGAGGTC GCCATTCTC	18	18	72
V049-56F	GAACAAGTAGAGGCAAATNNSGAGGTCGCCA TTCTCTC	18	17	73
V049-57F	CAAGTAGAGGCAAATGACNNSGTCGCCATTC TCTCTGAG	18	18	74
V049-58F	GTAGAGGCAAATGACGAGNNSGCCATTCTCT CTGAGGAAG	18	19	75
V049-59F	GAGGCAAATGACGAGGTCNNSATTCTCTCTG AGGAAGAG	18	18	76
V049-60F	CAAATGACGAGGTCGCCNNSCTCTCTGAGGA AGAGGAAG	17	19	77
V049-61F	CAAATGACGAGGTCGCCATTNNSTCTGAGGA AGAGGAAGTC	20	18	78
V049-62F	GACGAGGTCGCCATTCTCNNSGAGGAAGAG GAAGTCGAAAT	18	20	79
V049-63F	GAGGTCGCCATTCTCTNNSGAGGAAGGAAG TCGAAATTG	18	19	80
V049-64F	GTCGCCATTCTCTCTGAGNNSGAGGAAGTCG AAATTGAATT	18	20	81
V049-65F	CCATTCTCTCTGAGGAANNSGAGTCGAAAT TGAATTG	17	18	82
V049-66F	CATTCTCTCTGAGGAAGAGNNSGTCGAAATT GAATTGCTTC	19	19	83
V049-67F	CTCTCTGAGGAAGAGGAANNSGAAATTGAAT TGCTTCATG	18	19	84
V049-68F	CTGAGGAAGAGGAAGTCNNSATTGAATTGCT TCATGAATT	17	20	85
V049-69F	GAGGAAGAGGAAGTCGAANNSGAAATTGCTTC ATGAATTTG	18	19	86
V049-70F	GAAGAGGAAGTCGAAATTNNSTTGCTTCATG AATTTGAAAC	18	20	87

<179>

V049-71F	GAGGAAGTCGAAATTGAANNSCTTCATGAAT TTGAAAC	18	17	88
V049-72F	GAAGTCGAAATTGAATTGNNSCATGAATTTGA AACGATTC	18	19	89
V049-73F	GTCGAAATTGAATTGCTTNNNGAATTTGAAAC GATTCC	18	17	90
V049-74F	GAAATTGAATTGCTTCATNNSTTTGAAACGAT TCCTGTTTT	18	20	91
V049-75F	AAATTGAATTGCTTCATGAANNSGAAACGATT CCTGTTTTATC	20	20	92
V049-76F	GAATTGCTTCATGAATTTNNSACGATTCCTGT TTTTATC	18	17	93
V049-77F	AATTGCTTCATGAATTTGAANNSATTCCTGTT TTATCCGTTG	20	19	94
V049-78F	CTTCATGAATTTGAAACGNNSCCTGTTTTATC CGTTGAG	18	18	95
V049-79F	CATGAATTTGAAACGATTNNSGTTTTATCCGT TGAGTTAAG	18	20	96
V049-80F	GAATTTGAAACGATTCCTNNSTTATCCGTTGA GTTAAG	18	17	97
V049-81F	AATTTGAAACGATTCCTGTTNNSTCCGTTGAG TTAAGCCC	20	17	98
V049-82F	GAAACGATTCCTGTTTTANNSGTTGAGTTAAG CCCAGAAG	18	19	99
V049-83F	CGATTCCTGTTTTATCCNNSGAGTTAAGCCCA GAAGATG	17	19	100
V049-84F	GATTCCTGTTTTATCCGTTNNSTTAAGCCCAG AAGATGTG	19	18	101
V049-85F	CTGTTTTATCCGTTGAGNNSAGCCCAGAAGA TGTGGAC	17	18	102
V049-86F	GTTTTATCCGTTGAGTTANNSCCAGAAGATGT GGACGC	18	17	103
V049-87F	TTTTATCCGTTGAGTTAAGCNNSGAAGATGTG GACGCGCTTG	20	19	104
V049-88F	CCGTTGAGTTAAGCCCANNSGATGTGGACGC GCTTGAAC	17	19	105

V049-89F	GTTGAGTTAAGCCCAGAANNSTGGACGCG CTTGAACTC	18	18	106
V049-90F	GAGTTAAGCCCAGAAGATNNSGACGCGCTTG AACTCGATC	18	19	107
V049-91F	GTTAAGCCCAGAAGATGTGNNSGCGCTTGAA CTCGATCC	19	17	108
V049-92F	GCCCAGAAGATGTGGACNNSCTTGAACCTGA TCCAGC	17	17	109
V049-93F	CAGAAGATGTGGACGCGNNSGAACCTCGATC CAGCGATTTC	17	20	110
V049-94F	GAAGATGTGGACGCGCTTNSCTCGATCCAG CGATTTC	18	17	111
V049-95F	GATGTGGACGCGCTTGAANNSTCCAGCG ATTTCTTATAT	18	20	112
V049-96F	GTGGACGCGCTTGAACCTNNSCCAGCGATT CTTATATTG	18	19	113
V049-97F	GACGCGCTTGAACCTCGATNNSGCGATTTC ATATTGAAG	18	19	114
V049-98F	CGCTTGAACCTCGATCCANNSTTTCTTATAT GAAGAG	17	18	115
V049-99F	CTTGAACCTCGATCCAGCGNNSCTTATATTGA AGAGGATG	18	19	116
V049-100F	GAACCTCGATCCAGCGATTNNSATATTGAAG AGGATGC	18	17	117
V049-101F	CTCGATCCAGCGATTCTNNSATTGAAGAGG ATGCAGAAG	18	19	118
V049-102F	GATCCAGCGATTCTTATNNSGAAGAGGATG CAGAAGTAAC	18	20	119
V049-103F	CAGCGATTCTTATATTNNSGAGGATGCAGA AGTAAC	17	17	120
V049-104F	CGATTCTTATATTGAANNSTGCAGAAGTA ACGAC	17	17	121
V049-105F	GATTCTTATATTGAAGAGNNSGCAGAAGTAA CGACAATG	19	18	122
V049-106F	CTTATATTGAAGAGGATNNSGAAGTAACGAC AATGGC	17	17	123

V049-107F	CTTATATTGAAGAGGATGCANNSGTAACGAC AATGGCGCAATC	20	20	124
V049-108F	ATATTGAAGAGGATGCAGAANNSACGACAAT GGCGCAATC	20	17	125
V049-109F	GAAGAGGATGCAGAAGTANNSACAATGGCG CAATCGGTAC	18	19	126
V049-110F	GAGGATGCAGAAGTAACGNN SATGGCGCAAT CGGTACC	18	17	127
V049-111F	GATGCAGAAGTAACGACANNSGCGCAATCG GTACCATG	18	17	128
V049-1R	CACAATTTTTTGCTTCTSNNTCTTTACCCTCT CCTTTTAA	17	20	129
V049-2R	GATCCACAATTTTTTGCTSN NCACTCTTTACC CTCTCC	18	17	130
V049-3R	GACGATCCACAATTTTTSNNTCTCACTCTTT ACCCTC	18	17	131
V049-4R	GCGACGATCCACAATTTSNNGCTTCTCACTC TTTACC	17	17	132
V049-5R	GACGCGACGATCCACAASNNTTTGCTTCTCA CTCTTTAC	17	19	133
V049-6R	GTCGACGCGACGATCCASNNTTTTTGCTTC TCACTC	17	17	134
V049-7R	GCGGTGACGCGACGATSNNCAATTTTTGC TTCTCAC	17	18	135
V049-8R	GTAGTGCGGTGACGCGACSNCCACAATTT TTTGCTTC	19	17	136
V049-9R	GAGTAGTGCGGTGACGCSNNGATCCACAAT TTTTTG	18	17	137
V049-10R	GAAATGAGTAGTGCGGTGASNNGACGATCC ACAATTTTTG	20	19	138
V049-11R	GAAATGAGTAGTGCGGTSNNCGCGACGATC CACAATTTTT	17	20	139
V049-12R	CAACAGAAATGAGTAGTGCSNNGACGCGAC GATCCACAATT	19	20	140
V049-13R	GCAACAGAAATGAGTAGSNNGGTCGACGCG ACGATCC	17	17	141

V049-14R	CTAAAAGCAACAGAAATGAGSNNTGCGGTCG ACGCGACGATC	20	19	142
V049-15R	CTAAAAGCAACAGAAATSNNTAGTGC GGTCG ACGCGAC	17	18	143
V049-16R	GAACTAAAAGCAACAGASNNAGTAGTGCGG TCGACG	17	17	144
V049-17R	GATGAACTAAAAGCAACSNNAAATGAGTAGTG CGGTCG	17	17	145
V049-18R	GATCGATGAACTAAAAGCSNNAGAAATGAGT AGTGCGG	18	17	146
V049-19R	GCGATCGATGAACTAAASNNAACAGAAATGA GTAGTG	17	17	147
V049-20R	GATGCGATCGATGAACTSNNAGCAACAGAAA TGAGTAG	17	18	148
V049-21R	GCCGATGCGATCGATGASNNAAAAGCAACAG AAATGAG	17	18	149
V049-22R	GCAGCCGATGCGATCGASNNACTAAAAGCAA CAGAAATG	17	19	150
V049-23R	CTTCAGCAGCCGATGCGATSNNNTGAACTAAA AGCAACAG	19	17	151
V049-24R	CTTCTTCAGCAGCCGATGCSNNCGATGAACT AAAAGCAAC	19	18	152
V049-25R	GCTTCTTCAGCAGCCGASNNGATCGATGAAC TAAAAG	17	17	153
V049-26R	CTTTTGCTTCTTCAGCAGCSNNTGCGATCGA TGACTAAAAG	19	20	154
V049-27R	TTTTCTTTTGCTTCTTCAGCSNNCGATGCGAT CGATGAAC	20	17	155
V049-28R	TATTTTCTTTTGCTTCTCSNNAGCCGATGC GATCGATG	20	17	156
V049-29R	AAATATTTTCTTTTGCTTCSNNAGCAGCCGA TGCGATCG	20	17	157
V049-30R	ATTAATATTTTCTTTTGCSNNTTCAGCAGC CGATGCGATC	20	19	158
V049-31R	CAATTAATATTTTCTTTSNNTTCTTCAGCAG CCGATG	19	17	159

V049-32R	CCAATTAAATATTTTTCSNNTGCTTCTTCAGC AGCCG	17	17	160
V049-33R	TTAAAGCCAATTAATATTTSNNTTTTGCTTCT TCAGCAG	20	17	161
V049-34R	CATTAAGCCAATTAATASNNTTCTTTTGCTT CTTCAG	19	17	162
V049-35R	CTCATTAAAGCCAATTAASNNTTTTCTTTTGC TTCTTC	18	18	163
V049-36R	CTGCTCATTAAAGCCAATSNNATATTTTCTTT TGCTTC	18	18	164
V049-37R	CTTCCTGCTCATTAAAGCCSNNTAAATATTTT TCTTTTG	19	17	165
V049-38R	GCTTCCTGCTCATTAAASNNAATTAATATTTT TCTTTTG	17	20	166
V049-39R	GACAGCTTCCTGCTCATTSNNGCCAATTAAT ATTTTTC	18	18	167
V049-40R	CTGACAGCTTCCTGCTCSNNAAGCCAATTA AATATTTT	17	20	168
V049-41R	CTCACTGACAGCTTCCTGSNNATTAAGCCA ATTAATATT	18	20	169
V049-42R	CAAACACTGACAGCTCSNNTCATTAAAG CCAATTAAT	19	20	170
V049-43R	CTACAAACTCACTGACAGCSNNTGCTCATT AAAGCCAATTA	19	20	171
V049-44R	GTTCTACAAACTCACTGACSNNTTCTGCTCA TTAAAGC	19	17	172
V049-45R	CTTGTTCTACAAACTCACTSNNAGCTTCTGC TCATTAAG	19	19	173
V049-46R	CTACTTGTTCTACAAACTCSNNGACAGCTTCC TGCTCATTAA	19	20	174
V049-47R	CTCTACTTGTTCTACAAASNNACTGACAGCTT CCTGCTC	18	18	175
V049-48R	GCCTCTACTTGTTCTACSNNTCACTGACAG CTTCCTG	17	18	176
V049-49R	CATTTGCCTCTACTTGTTCSNNAACTCACTG ACAGCTTC	19	18	177

V049-50R	GTCATTTGCCTCTACTTGSNNTACAAACTCAC TGACAG	18	17	178
V049-51R	CTCGTCATTTGCCTCTACSNNTTCTACAAACT CACTGAC	18	18	179
V049-52R	GACCTCGTCATTTGCCTCSNNTTGTCTACAA ACTCAC	18	17	180
V049-53R	GCGACCTCGTCATTTGCSNNTACTTGTCTAC AAACTC	17	18	181
V049-54R	GAATGGCGACCTCGTCATTSNNCTCTACTTG TTCTACAAAC	19	19	182
V049-55R	GAGAATGGCGACCTCGTCSNNTGCCTCTACT TGTTCTAC	18	18	183
V049-56R	GAGAGAATGGCGACCTCSNNTTGCCTCTA CTTGTTTC	17	18	184
V049-57R	CTCAGAGAGAATGGCGACSNNGTCATTTGCC TCTACTTG	18	18	185
V049-58R	CTTCTCAGAGAGAATGGCSNNTCGTCATT TGCCTCTAC	19	18	186
V049-59R	CTTTCTCAGAGAGAATSNNGACCTCGTCA TTTGCCTC	18	18	187
V049-60R	CTTCTCTTCTCAGAGAGSNNGGCGACCTC GTCATTTG	19	17	188
V049-61R	GACTTCTCTTCTCAGASNNAATGGCGACC TCGTCATTTG	18	20	189
V049-62R	ATTTGACTTCTCTTCTCSNNGAGAATGG CGACCTCGTC	20	18	190
V049-63R	CAATTTGACTTCTCTTCSNAGAGAGAATG GCGACCTC	19	18	191
V049-64R	AATTCAATTTGACTTCTCSNNTCAGAGAG AATGGCGAC	20	18	192
V049-65R	CAATTCAATTTGACTTCSNNTTCTCAGAGA GAATGG	18	17	193
V049-66R	GAAGCAATTCAATTTGACSNNTCTTCTCA GAGAGAATG	19	19	194
V049-67R	CATGAAGCAATTCAATTTCSNNTTCTTCC TCAGAGAG	19	18	195

V049-68R	AATTCATGAAGCAATTCATSNNGACTTCCTC TTCCTCAG	20	17	196
V049-69R	CAAATTCATGAAGCAATTCNNNTTCGACTTCC TCTTCCTC	19	18	197
V049-70R	GTTTCAAATTCATGAAGCAASNNAATTTTCGAC TTCCTCTTC	20	18	198
V049-71R	GTTTCAAATTCATGAAGSNNTTCAATTTTCGAC TTCCTC	17	18	199
V049-72R	GAATCGTTTCAAATTCATGSNNCAATTCAATT TCGACTTC	19	18	200
V049-73R	GGAATCGTTTCAAATTCNNAAGCAATTCAAT TTCGAC	17	18	201
V049-74R	AAAACAGGAATCGTTTCAAASNNATGAAGCA ATTCAATTC	20	18	202
V049-75R	GATAAACAGGAATCGTTTCSNNTTCATGAAG CAATTCATTT	20	20	203
V049-76R	GATAAACAGGAATCGTSNNAATTCATGAA GCAATTC	17	18	204
V049-77R	CAACGGATAAACAGGAATSNNTTCAAATTC TGAAGCAATT	19	20	205
V049-78R	CTCAACGGATAAACAGGSNNCGTTTCAAAT TCATGAAG	18	18	206
V049-79R	CTTAACTCAACGGATAAACSNNAATCGTTTC AAATTCATG	20	18	207
V049-80R	CTTAACTCAACGGATAASNAGGAATCGTTTC AAATTC	17	18	208
V049-81R	GGGCTTAACTCAACGGASNNAACAGGAATCG TTTCAAATT	17	20	209
V049-82R	CTTCTGGGCTTAACTCAACSNNTAAACAGG AATCGTTTC	19	18	210
V049-83R	CATCTTCTGGGCTTAACTCSNNGGATAAAC AGGAATCG	19	17	211
V049-84R	CACATCTTCTGGGCTTAAASNNAACGGATAAA ACAGGAATC	18	19	212
V049-85R	GTCCACATCTTCTGGGCTSNNTCAACGGAT AAAACAG	18	17	213

V049-86R	GCGTCCACATCTTCTGGSNNTAACTCAACGG ATAAAAC	17	18	214
V049-87R	CAAGCGCGTCCACATCTTCSNNGCTTAACTC AACGGATAAAA	19	20	215
V049-88R	GTTCAAGCGCGTCCACATCSNNTGGGCTTAA CTCAACGG	19	17	216
V049-89R	GAGTTCAAGCGCGTCCACSNNTTCTGGGCTT AACTCAAC	18	18	217
V049-90R	GATCGAGTTCAAGCGCGTCSNNATCTTCTGG GCTTAACTC	19	18	218
V049-91R	GGATCGAGTTCAAGCGCSNNCACATCTTCTG GGCTTAAAC	17	19	219
V049-92R	GCTGGATCGAGTTCAAGSNNGTCCACATCTT CTGGGC	17	17	220
V049-93R	GAAATCGCTGGATCGAGTTCSNNCGGCTCCA CATCTTCTG	20	17	221
V049-94R	GAAATCGCTGGATCGAGSNNAAGCGCGTCC ACATCTTC	17	18	222
V049-95R	ATATAAGAAATCGCTGGATCSNNTTCAAGCG CGTCCACATC	20	18	223
V049-96R	CAATATAAGAAATCGCTGGSNNGAGTTCAAG CGCGTCCAC	19	18	224
V049-97R	CTTCAATATAAGAAATCGCSNNATCGAGTTCA AGCGCGTC	19	18	225
V049-98R	CTCTTCAATATAAGAAATSNNTGGATCGAGTT CAAGCG	18	17	226
V049-99R	CATCCTCTTCAATATAAGASNNGCTGGATC GAGTTCAAG	19	18	227
V049-100R	GCATCCTCTTCAATATASNNAATCGCTGGATC GAGTTC	17	18	228
V049-101R	CTTCTGCATCCTTCAATSNNAGAAATCGCT GGATCGAG	19	18	229
V049-102R	GTTACTTCTGCATCCTCTTCSNNATAAGAAAT CGCTGGATC	20	18	230
V049-103R	GTTACTTCTGCATCCTCSNNAATATAAGAAAT CGCTG	17	17	231

V049-104R	GTCGTTACTTCTGCATCSNNTTCAATATAAGA AATCG	17	17	232
V049-105R	CATTGTCGTTACTTCTGCSNNCTCTTCAATAT AAGAAATC	18	19	233
V049-106R	GCCATTGTCGTTACTTCSNNATCCTCTTCAAT ATAAG	17	17	234
V049-107R	GATTGCGCCATTGTCGTTACSNNTGCATCCT CTTCAATATAAG	20	20	235
V049-108R	GATTGCGCCATTGTCGTSNNTTCTGCATCCT CTTCAATAT	17	20	236
V049-109R	GTACCGATTGCGCCATTGTSNNTACTTCTGC ATCCTCTTC	19	18	237
V049-110R	GGTACCGATTGCGCCATSNNCGTTACTTCTG CATCCTC	17	18	238
V049-111R	CATGGTACCGATTGCGCSNNTGTCGTTACTT CTGCATC	17	18	239

<188>

<189>

* 제공된 프라이머 이름은 치환의 아미노산 위치를 반영하고; "R"은 역방향 프라이머인 프라이머를 나타내고 "F"는 정방향 프라이머인 프라이머를 나타낸다. 예를 들어, V049~108F는 V049 전구체 프로테아제의 위치 108에서의 아미노산의 치환부에서 사용된 정방향 프라이머이다.

<190>

"염기 좌측" 및 *"염기 우측"은 프라이머에 존재하는 돌연변이 코돈의 좌측 및 우측에 대한 염기 수를 나타낸다. 이들 염기는 주형 전구체 폴리뉴클레오티드 염기의 염기 (즉 V049)에 대해 상보성이다.

<191>

SEQ ID NO:17 - 벡터 pXX-049의 폴리뉴클레오티드 서열

AATTCCTCCATTTTCTTCTGCTATCAAATAACAGACTCGTGATTTTCAAACGAGCTTTCAA
 AAAAGCCTCTGCCCTTGCAAATCGGATGCCTGTCTATAAAATCCCGATATTGGCTTAAA
 CAGCGGCGCAATGGCGGCCGCATCTGATGTCTTTGCTTGGCGAATGTTTCATCTTATTTCTT
 CCTCCCTCTCAATAATTTTTTCATTCTATCCCTTTTCTGTAAGTTTATTTTTTCAGAATACTTT
 TATCATCATGCTTTGAAAAAATATCACGATAATATCCATTGTTCTCACGGAAGCACACGCAG
 GTCATTTGAACGAATTTTTTCGACAGGAATTTGCCGGGACTCAGGAGCATTTAACCTAAAA
 AAGCATGACATTTTCAGCATAATGAACATTTACTCATGTCTATTTTCGTTCTTTTCTGTATGAA
 AATAGTTATTTGAGTCTCTACGGAAATAGCGAGAGATGATATACCTAAATAGAGATAAAAT
 CATCTCAAAAAATGGGTCTACTAAAATATTATTCATCTATTACAATAAAATTCACAGAATAG

<192>

TCTTTAAGTAAGTCTACTCTGAATTTTTTAAAAGGAGAGGGTAAAGAGTGAGAAGCAAAA
 AATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTTAGTTCATCGATCGC
 ATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTAAATGAGCAGGAAGCTGTC
 AGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAG
 GAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTATCCGTTGAGTTAAG
 CCCAGAAGATGTGGACGCGCTTGAACCTCGATCCAGCGATTTCTTATATTGAAGAGGATGCA
 GAAGTAACGACAATGGCGCAATCGGTACCATGGGGAATTAGCCGTGTGCAAGCCCCAGCT
 GCCCATAACCGTGGATTGACAGGTTCTGGTGTAAAAGTTGCTGTCCTCGATACAGGATTTT
 CCACTCATCCAGACTTAAATATTCGTGGTGGCGCTAGCTTTGTACCAGGGGAACCATCCAC
 TCAAGATGGGAATGGGCATGGCACGCATGTGGCTGGGACGATTGCTGCTTTAAACAATTC
 GATTGGCGTTCTTGGCGTAGCACCGAACGCGGAACTATACGCTGTTAAAGTATTAGGGGC
 GAGCGGTTGAGGTTGCGTCAGCTCGATTGCCCAAGGATTGGAATGGGCAGGGAACAATGT
 TATGCACGTTGCTAATTTGAGTTTAGGACTGCAGGCACCAAGTGCCACACTTGAGCAAGCT
 GTTAATAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCA
 GGCTCAATCAGCTATCCGGCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAA
 AACAAACAACCGCGCCAGCTTTTCACAGTATGGCGCAGGGCTTGACATTGTCGACCAGGT
 GTAAACGTGCAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATG
 GCTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCA
 ATGTACAAATCCGCAATCATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGTA
 TGGAAGCGGACTTGCAATGCAGAAGCGGCAACACGTTAATCAATAAAAAAACGCTGTGC
 GGTTAAAGGGCACAGCGTTTTTTTTGTGTATGAATCGGGATCCTCGATCGAGACTAGAGTCG
 ATTTTTACAAGAATTAGCTTTATATAATTTCTGTTTTTCTAAAGTTTTATCAGCTACAAAAGAC
 AGAAATGTATTGCAATCTTCAACTAAATCCATTTGATTCTCTCCAATATGACGTTTAATAAT
 TTCTGAAATACTTGATTTCTTTGTTTTTCTCAGTATACTTTTCCATGTTATAACACATAAAAA
 CAACTTAGTTTTACAAACTATGACAATAAAAAAGTTGCTTTTTCCCTTTCTATGTATGTT
 TTTTACTAGTCATTTAAAACGATACATTAATAGGTACGAAAAAGCAACTTTTTTTGCGCTTAA
 AACCAGTCATACCAATAACTTAAGGGTAACTAGCCTCGCCGGCAATAGTTACCCTTATTAT
 CAAGATAAGAAAGAAAAGGATTTTTCGCTACGCTCAAATCCTTTAAAAAACACAAAAGACC
 ACATTTTTAATGTGGTCTTTATTCTTCAACTAAAGCACCCATTAGTTCAACAAACGAAAATT
 GGATAAAGTGAGATTTTTTAAAATATATTTTATGTTACAGTAATATTGACTTTTAAAAAAG
 GATTGATTCTAATGAAGAAAGCAGACAAGTAAGCCTCCTAAATTCACTTTAGATAAAAAATTT
 AGGAGGCATATCAAATGAACTTTAATAAAATTGATTTAGACAATTGGAAGAGAAAAGAGATA
 TTTAATCATTATTTGAACCAACAAACGACTTTTAGTATAACCACAGAAATTGATATTAGTGT
 TTATACCGAAACATAAAACAAGAAGGATATAAATTTTACCCTGCATTTATTTTCTTAGTGACA
 AGGGTGATAAACTCAAATACAGCTTTTAGAAGTGGTTACAATAGCGACGGAGAGTTAGGTT
 ATTGGGATAAGTTAGAGCCACTTTTATACAATTTTTGATGGTGTATCTAAAACATTCTCTGGT

ATTTGGACTCCTGTAAAGAATGACTTCAAAGAGTTTTATGATTTATACCTTTCTGATGTAGA
 GAAATATAATGGTTCGGGGAAATTGTTTCCAAAACACCTATACCTGAAAATGCTTTTTCTC
 TTTCTATTATTCCATGGACTTCATTTACTGGGTTAACTTAAATATCAATAATAATAGTAATTA
 CCTTCTACCCATTATTACAGCAGGAAAATTCATTAATAAAGGTAATTCAATATATTTACCGCT
 ATCTTTACAGGTACATCATTCTGTTTGTGATGGTTATCATGCAGGATTGTTTATGAACTCTAT
 TCAGGAATTGTCAGATAGGCCTAATGACTGGCTTTTATAATATGAGATAATGCCGACTGTA
 CTTTTTACAGTCGGTTTTCTAATGTCACCTAACCTGCCCGTTAGTTGAAGAAGTTTTTATA
 TTACAGCTCCAGATCCATATCCTTCTTTTTCTGAACCGACTTCTCCTTTTTCGCTTCTTTATT
 CCAATTGCTTTATTGACGTTGAGCCTCGGAACCTTAAACAATCCAAAACCTTGTGCAATGG
 TCGGCTTAATAGCTCACGCTATGCCGACATTCGTCTGCAAGTTTAGTTAAGGGTCTTCTC
 AACGCACAATAAATTTTTCTCGGCATAAATGCGTGGTCTAATTTTTATTTTTAATAACCTTGAT
 AGCAAAAAATGCCATTCCAATACAAAACACATACCTATAATCGACCTGCAGGAATTAATTC
 CTCCATTTTCTTCTGCTATCAAATAACAGACTCGTGATTTTCAAACGAGCTTTCAAAAA
 GCCTCTGCCCTTGCAAATCGGATGCCTGTCTATAAAATCCCGATATTGGCTTAAACAGC
 GCGCAATGGCGGCCGCATCTGATGTCTTTGCTTGGCGAATGTTTATTTCTTCTCCTC
 CCTCTCAATAATTTTTTCTTCTATCCCTTTTCTGTAAGTTTATTTTTCAGAATACTTTTATC
 ATCATGCTTTGAAAAAATATCACGATAATATCCATTGTTCTCACGGAAGCACACGCAGGTCA
 TTTGAACGAATTTTTTCGACAGGAATTTGCCGGGACTCAGGAGCATTAACTAAAAAGC
 ATGACATTTACAGCATAATGAACATTTACTCATGTCTATTTTCGTTCTTTTCTGTATGAAAAA
 GTTATTTGAGTCTCTACGAAATAGCGAGAGATGATATACCTAAATAGAGATAAAATCATC
 TCAAAAAATGGGTCTACTAAAAATTATTCCATCTATTACAATAAATTCACAGAATAGTCTT
 TTAAGTAAGTCTACTCTGAATTTTTTATCAAGCTAGCTTGGCGTAATCATGGTCATAGCTG
 TTTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAA
 AGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT
 GCCCGCTTTCAGTCGGGAAACCTGTCTGTCAGCTGCATTAATGAATCGGCCAACGCGC
 GGGGAGAGGCGGTTTTCGCTATTGGGCGCTTTCGCTTCTCGCTCACTGACTCGCTGC
 GCTCGGTCGTTGCGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
 CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCC
 AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAG
 CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC
 CAGGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC
 GGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGT
 AGGTATCTCAGTTCCGGTGTAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
 GTTCAGCCCCGACCGCTGCGCCTTATCCGTAACACTATCGTCTTGAGTCCAACCCGGTAAGA
 CACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA
 GCGCGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTA

TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT
 CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGC
 GCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG
 GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
 ATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT
 GACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATC
 CATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
 CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAAT
 AAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCA
 TCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGC
 CAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTTGGTATGGCTTCA
 TTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAG
 CGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCACT
 CATGGTTATGGCAGCACTGCATAATTCTTACTGTATGCCATCCGTAAGATGCTTTTCTG
 TGAAGTGGTACTCAACCAAGTCACTTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT
 CTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAGTTTAAAAGTGTCTCAT
 CATTGAAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCGCTGTTGAGATCCAGT
 TCGATGTAACCCACTCGTGACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTC
 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGAATAAGGGCGACACGGA
 AATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTC
 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCAC
 ATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
 AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCTTTTCGGTGATGACGGTGAAAACC
 TCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGC
 AGACAAGCCCGTCAGGGCGCGTCAAGCGGTGTTGGCGGGTGTGGGGCTGGCTTAACTA
 TGCGGCATCAGAGCAGATTGTAAGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAG
 ATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTGCCATTGAGGCTGCGCAACTGTTG
 GGAAGGGCGATCGGTGCGGGCCTTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTG
 CTGCAAGGCGATTAAGTTGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAACCGA
 CGGCCAGTG

<195>

<196>

QC 반응은, 총 50 μ L에 대해, 40.25 μ L의 멸균 증류수, 키트로부터의 5 μ L의 PfuTurbo 10x 완충액, 키트로부터의 1 μ L의 dNTP, 1.25 μ L의 정방향 프라이머 (100ng/ μ L), 1.25 μ L 역방향 프라이머 (100ng/ μ L), 주형으로서 0.25 μ L의 pMSAT-NcoI 미니프랩 DNA (~50ng), 및 키트로부터의 1 μ L의 PfuTurbo로 이루어졌다. 싸이클 조건은 1 min 동안 95 $^{\circ}$ C 1회, 이후 1 min 동안 95 $^{\circ}$ C, 1 min 동안 55 $^{\circ}$ C, 및 12 min 동안 68 $^{\circ}$ C의 19~20 싸이클이었다. 반응을 분석하기 위해, 종결 시 5 μ L의 반응물을 1.2% E-gel (Invitrogen사제) 상에서 실험하였다. 다음으로, 37 $^{\circ}$ C에서 2 내지 8시간 동안 1 μ L 및 1 μ L의 효소로 DpnI 소화를 순차적으로 2회 실행하였다. 유사한 조건 하에, 프라이머 없이 음성 대조군을 수행하였다. 이후, 1 μ L의 DpnI-소화된 반응 생성물을 BioRad 전기천공기를 사용하여 50 μ L의 원샷 TOP10 전기적격 세포 (Invitrogen사제)에 형질전환시켰다. 이후, TOP10 세포 (Invitrogen사제)와 함께 제공된 1 ml의 SOC를 전기천공된 세포에 첨가하고, 5 ppm 클로람페니콜을 함유하는 LA 플레이트에 플레이팅하기 전 1시간 동안 진탕 인큐베이션하였다. 플레이트를 37 $^{\circ}$ C에서 하룻밤 동안 인큐베이션하였다. 이러한 인큐베이션 후, 각각의 라이브러리 (즉 각각의 위치)로부터의 96개 콜로니를, 96 웰 마이크로타이터 플레이트 중 10~50 ppm의 클로람페니콜을 함유하는 200 μ L의 LB에 접종하고 37 $^{\circ}$ C에서 하룻밤 동안 성장시켰다. 다음 날 20% 최종 농도로 클리세롤을 첨가한 후 플레이트를 -80 $^{\circ}$ C에서 냉각시키고, 이를 V049SEQ-R2 프라이머로의 고작업량 서열화에 사용하였다.

<197>

유사하게, QuikChange[®] 다중 위치-지정 돌연변이 유발 (QCMS;Stratagene사제)을 사용하여 전구체 프로테아제 V049의 pro 부위 중 둘 이상의 아미노산을 인코딩하는 코돈의 돌연변이를 수행하였다. 총 25 μ L에 대해, 19.25 μ L의 멸균 증류수, 키트로부터의 2.5 μ L의 10x 완충액, 키트로부터의 1 μ L dNTP, 1 μ L의 5'인산화 정방향 프라이머 (100ng/ μ L), 주형으로서 0.25 μ L의 pMSAT-NcoI 미니프랩 DNA (~50ng), 및 키트로부터의 1 μ L

의 효소 블렌드를 사용하여 QCMS 반응을 수행하였다. 싸이클 조건은 1 min 동안 95°C 1회, 이후 1 min 동안 95°C, 1 min 동안 55°C, 및 12 min 동안 65°C의 30 싸이클이었다. 반응 생성물을 분석하기 위해, 종결시 2.5 µL의 반응물을 1.2% E-gel (Invitrogen사제) 상에서 실험하였다. 다음으로, 37°C에서 2 내지 8시간 동안 1 µL, 이후 0.5 µL의 효소로 DpnI 소화를 순차적으로 2회 실행하였다. 대조군, 형질전환 및 서열화를 상기 기재된 QC 방법에서와 같이 수행하였다.

<198> **실시예 2. 숙주 세포 형질전환 및 개질된 프로테아제의 발현**

<199> 관심 있는 개질된 단백질을 인코딩하는 폴리뉴클레오티드를 함유하는 플라스미드 pXX-049를 37°C에서 3~5 시간 동안 DpnI으로 2회 소화시켰다.

<200> 대장균의 형질전환 및 스크리닝 - 1 µl의 DpnI 소화된 플라스미드 DNA를 사용하여 전기천공으로 대장균 Top10 세포를 형질전환시켰다. 형질전환된 세포를 5 ppm CMP (클로람페니콜)을 함유하는 LB 한천 배지에 플레이팅 하고, 콜로니를 하룻밤 동안 성장시켰다. 96개의 개별적인 콜로니를 고르고, LB + 5 ppm CMP + 50 ppm 카르 베니실린을 함유하는 96웰 마이크로타이터 플레이트의 상응하는 웰에 옮겼다. 배양물을 250 rpm에서 진탕하면서 37°C에서 하룻밤 동안 성장시켰다. 글리세롤 저장액을 10%의 최종 농도로 배양물에 첨가하였다. 플라스미드 DNA를 96개 대장균 배양물로부터 생성시키고, 플라스미드 DNA 생성물의 부분을 서열화시켰다 (Cogenics, Morrisville, NC). Phrep, Phrap, Consed, Custal W 소프트웨어를 사용하여 자동화된 서열 분석을 수행하였다.

<201> 바실러스 서브틸리스로의 형질전환 - 플라스미드 DNA의 2차 부분을 사용하여 B. 서브틸리스 숙주 세포를 형질전환시켰다. 적합한 돌연변이를 수반하는 각각의 96개 대장균 배양물로부터의 2 마이크로타이터의 플라스미드 DNA를 사용하여 100 µl의 B. 서브틸리스 ComK 적격 세포를 형질전환시켰다. 세포를 250 rpm에서 진탕하면서 37°C에서 45분 동안 인큐베이션하였다. 96개 형질전환 혼합물로부터의 세포를 1.6% 탈지유 및 5 ppm CMP를 함유하는 LA에 플레이팅하고, 37°C 인큐베이터에서 하룻밤 동안 인큐베이션하였다.

<202> 각각의 96개 형질전환물로부터 4개 콜로니를 고르고, 웰 당 150 µl의 LB 및 5 ppm CMP를 함유하는 마이크로타이터 플레이트에 개별적으로 옮겼다. 상당수의 마이크로타이터 플레이트의 웰은 적합한 대조군을 함유하였다. 이후 마이크로타이터 플레이트를 250 rpm에서 회전시키면서 37°C에서 4시간 동안 인큐베이션하였다. 각각의 배양물 10 µl를 140 µl의 Grant II 배지 + 5 ppm CMP (pH 7.3)를 함유하는 새로운 마이크로타이터 플레이트에 옮겼다. GrantII 배지를 하기와 같이 제조하였다: 용액 I: 10g의 소이톤을 500 ml 물에 용해시키고 20~25 분 동안 오토클레이브하였다; 용액 II: 3 ml의 1 M K₂HPO₄, 75 g 글루코오스, 3.6 g 우레아, 100 ml Grant 10X MOPS를 400 ml 물에 희석시켰다. 용액 I 및 II를 혼합하고, HCl/NaOH로 pH를 pH 7.3으로 조정하였다. 최종 부피를 1 L로 조정하고, 최종 용액을 0.22 µm PES 필터를 통해 멸균시켰다.

<203> 마이크로타이터 플레이트 배양물을 37°C, 250 rpm에서 진탕기에서 인큐베이션하였다. 분석법 분석에 대해 규칙적인 간격 (40시간 이하)으로 시료를 취하였다.

<204> **실시예 3. 개질된 프로테아제 생성의 측정: 프로테아제 활성의 AAPF 분석**

<205> 실시예 2에서 기재된 바와 같이 수득된 각각의 B. 서브틸리스 배양물을, 개질된 프로테아제의 생성에 대해 분석하였다. 생성된 효소를 기질, 숙시닐-L-Ala-L-Ala-L-Pro-L-Phe-p-니트로아날라이드 (AAPF)에 대한 활성에 대해 분석하였다. 상기 분석은, p-니트로아날린의 방출 및 가수분해로 야기되는 405 nm/min에서의 흡광도에서의 증가로서, 개질된 프로테아제의 생성을 측정하였다 (Estell 등, J Biol Chem., 260:6518-6521 (1985)).

Softmax Pro 소프트웨어를 사용하여 측정하고, 구체화된 조건을 하기와 같이 설정하였다: 유형: 키네틱; 감소: Vmax 점 (15/28 점에서 최적 리딩); Lml : 405 nm; 시간: 5분; 및 간격: 11초. 각각의 B. 서브틸리스 배양물의 10 마이크로리터를, 10 mM 트리스 + 0.005% TWEEN[®]-80, pH 8.6; 및 25 µl의 100 mg/ml AAPF를 함유하는 100 µl의 트리스 완충액에 희석시켰다. 각각의 개질된 프로테아제의 상대 활성을 계산하고, 각각의 개질된 프로테아제로부터 가공된 성숙 프로테아제의 활성 대 비개질된 V049 전구체 프로테아제로부터 가공된 성숙 프로테아제의 활성의 비율로서 상응하는 개질된 프로테아제의 생성에 대한 각각의 아미노산 치환부의 효과를 측정하였다. 결과를 표 2에 나타내었다.

<206> DNA 구성물이 적격 바실러스 서브틸리스 균주에 안정적으로 통합되고 나면, 개질된 프로테아제의 활성을 마이크로타이터 분석에서 측정하고, 상응하는 전구체 프로테아제의 활성과 상기 활성을 비교하였다.

<207> 하룻밤 배양한 Grant II 배지 세포 배양물 10 마이크로리터를, 10 mM 트리스 + 0.005% TWEEN[®]-80 (pH 8.6)을

함유하는 100 μ l의 트리스 완충액으로 희석시키고; 25 μ l의 100 mg/ml AAPF 기질을 프로테아제 활성화에 대한 분석에 사용하였다. 분석을 마이크로타이터 플레이트에서 수행하고 Softmax Pro 소프트웨어를 사용하였다.

<208> 결과는 전구체 V049 프로테아제의 대부분의 아미노산의 아미노산 치환부가 프로테아제의 성숙 형태의 증강된 생성을 일으킨다는 것을 나타내었다. 추가적으로, 각각의 치환된 아미노산의 위치 포화는, 비개질된 pro 부위를 갖는 전구체 프로테아제로부터 수득된 것에 관해 성숙 형태의 생성을 증가시키도록 각각의 아미노산이 동일한 위치에서 둘 이상의 아미노산에 의해 치환될 수 있다는 것을 나타내었다.

<209> 표 2

위치*	변종 코드	변종 아미노산	단백질 농도	활성 비율#
1	V001D	D	102.37	1.02
1	V001Q	Q	98.52	0.99
1	V001F	F	96.85	0.97
1	V001L	L	58.70	0.59
1	V001A	A	58.67	0.59
1	V001H	H	36.54	0.37
1	V001I	I	35.67	0.36
1	V001G	G	29.93	0.30
1	V001Y	Y	14.97	0.15
1	V001E	E	9.30	0.09
1	V001T	T	6.99	0.07
1	V001C	C	6.75	0.07
1	V001W	W	6.61	0.07
1	V001S	S	5.54	0.06

<210>

1	V001R	R	5.19	0.05
2	R002M	M	122.01	1.22
2	R002W	W	104.55	1.05
2	R002K	K	103.46	1.03
2	R002C	C	101.20	1.01
2	R002S	S	101.19	1.01
2	R002L	L	98.36	0.98
2	R002F	F	93.71	0.94
2	R002H	H	93.65	0.94
2	R002N	N	89.23	0.89
2	R002A	A	82.16	0.82
2	R002G	G	73.95	0.74
2	R002D	D	60.97	0.61
2	R002V	V	53.99	0.54
2	R002T	T	40.33	0.40
2	R002Y	Y	13.66	0.14
3	S003F	F	133.20	1.33
3	S003M	M	127.94	1.28
3	S003R	R	121.00	1.21
3	S003T	T	116.45	1.16
3	S003Q	Q	114.41	1.14
3	S003N	N	105.57	1.06
3	S003I	I	104.04	1.04
3	S003V	V	101.09	1.01
3	S003W	W	99.63	1.00
3	S003G	G	99.50	0.99
3	S003D	D	98.94	0.99
3	S003H	H	98.52	0.99
3	S003A	A	97.60	0.98
3	S003P	P	93.74	0.94
3	S003Y	Y	65.38	0.65
3	S003L	L	48.58	0.49
3	S003C	C	10.43	0.10
4	K004T	T	121.17	1.21
4	K004V	V	116.61	1.17

<211>

4	K004Y	Y	111.33	1.11
4	K004I	I	109.36	1.09
4	K004C	C	109.34	1.09
4	K004R	R	107.32	1.07
4	K004F	F	103.89	1.04
4	K004H	H	96.72	0.97
4	K004A	A	95.03	0.95
4	K004Q	Q	92.30	0.92
4	K004P	P	91.49	0.91
4	K004N	N	86.73	0.87
4	K004S	S	78.31	0.78
4	K004G	G	53.68	0.54
4	K004E	E	29.18	0.29
4	K004L	L	7.86	0.08
5	K005W	W	123.75	1.24
5	K005N	N	118.59	1.19
5	K005Q	Q	114.87	1.15
5	K005Y	Y	112.41	1.12
5	K005V	V	112.29	1.12
5	K005H	H	111.16	1.11
5	K005G	G	110.84	1.11
5	K005S	S	109.37	1.09
5	K005D	D	108.39	1.08
5	K005C	C	103.96	1.04
5	K005R	R	100.76	1.01
5	K005T	T	99.54	1.00
5	K005A	A	91.66	0.92
5	K005L	L	90.96	0.91
5	K005P	P	49.40	0.49
5	K005F	F	6.26	0.06
5	K005M	M	5.65	0.06
6	L006M	M	122.55	1.23
6	L006S	S	118.01	1.18
6	L006G	G	116.94	1.17
6	L006N	N	115.98	1.16

<212>

6	L006V	V	115.39	1.15
6	L006P	P	115.11	1.15
6	L006H	H	114.99	1.15
6	L006D	D	114.30	1.14
6	L006K	K	111.33	1.11
6	L006E	E	109.38	1.09
6	L006A	A	109.11	1.09
6	L006T	T	105.11	1.05
6	L006I	I	105.10	1.05
6	L006R	R	103.58	1.04
6	L006F	F	101.84	1.02
6	L006C	C	101.69	1.02
6	L006W	W	77.03	0.77
6	L006Y	Y	7.99	0.08
7	W007V	V	125.70	1.26
7	W007M	M	121.99	1.22
7	W007S	S	121.49	1.21
7	W007R	R	121.42	1.21
7	W007P	P	113.12	1.13
7	W007T	T	112.42	1.12
7	W007N	N	111.62	1.12
7	W007Q	Q	110.40	1.10
7	W007F	F	109.54	1.10
7	W007K	K	99.28	0.99
7	W007C	C	99.28	0.99
7	W007G	G	98.93	0.99
7	W007H	H	95.65	0.96
7	W007A	A	92.24	0.92
7	W007L	L	80.60	0.81
7	W007Y	Y	61.72	0.62
7	W007I	I	6.21	0.06
8	I008F	F	114.74	1.15
8	I008S	S	111.06	1.11
8	I008P	P	107.32	1.07
8	I008L	L	106.57	1.07

<213>

8	I008V	V	105.39	1.05
8	I008T	T	105.22	1.05
8	I008Y	Y	103.53	1.04
8	I008M	M	100.20	1.00
8	I008E	E	99.53	1.00
8	I008A	A	95.64	0.96
8	I008D	D	75.18	0.75
8	I008R	R	4.60	0.05
9	V009R	R	130.14	1.30
9	V009P	P	122.65	1.23
9	V009M	M	114.91	1.15
9	V009I	I	114.34	1.14
9	V009Y	Y	111.23	1.11
9	V009S	S	107.52	1.08
9	V009C	C	105.80	1.06
9	V009T	T	105.04	1.05
9	V009E	E	100.29	1.00
9	V009W	W	98.90	0.99
9	V009L	L	97.67	0.98
9	V009A	A	93.84	0.94
9	V009H	H	86.35	0.86
9	V009N	N	85.33	0.85
9	V009K	K	6.32	0.06
10	A010M	M	127.85	1.28
10	A010R	R	106.71	1.07
10	A010I	I	106.36	1.06
10	A010S	S	106.34	1.06
10	A010Q	Q	106.33	1.06
10	A010P	P	105.05	1.05
10	A010C	C	104.29	1.04
10	A010H	H	103.11	1.03
10	A010N	N	101.39	1.01
10	A010G	G	100.93	1.01
10	A010W	W	97.00	0.97
10	A010T	T	96.44	0.96

10	A010K	K	59.66	0.60
10	A010D	D	45.09	0.45
10	A010L	L	30.80	0.31
10	A010F	F	8.41	0.08
11	S011G	G	124.00	1.24
11	S011M	M	114.00	1.14
11	S011P	P	113.00	1.13
11	S011C	C	112.00	1.12
11	S011F	F	112.00	1.12
11	S011V	V	104.16	1.04
11	S011N	N	104.00	1.04
11	S011A	A	100.00	1.00
11	S011Y	Y	96.00	0.96
11	S011D	D	95.00	0.95
11	S011T	T	95.00	0.95
11	S011L	L	92.00	0.92
11	S011Q	Q	92.00	0.92
11	S011I	I	90.00	0.90
11	S011W	W	88.00	0.88
11	S011K	K	66.00	0.66
11	S011R	R	43.00	0.43
12	T012A	A	142.37	1.42
12	T012G	G	134.84	1.35
12	T012H	H	131.21	1.31
12	T012C	C	131.00	1.31
12	T012W	W	127.05	1.27
12	T012S	S	126.74	1.27
12	T012V	V	125.13	1.25
12	T012M	M	123.79	1.24
12	T012P	P	120.33	1.20
12	T012I	I	117.21	1.17
12	T012Q	Q	111.02	1.11
12	T012F	F	105.34	1.05
12	T012N	N	97.46	0.97
12	T012E	E	82.85	0.83

<215>

12	T012K	K	76.37	0.76
12	T012R	R	51.07	0.51
12	T012D	D	50.05	0.50
12	T012L	L	7.07	0.07
13	A013G	G	127.43	1.27
13	A013V	V	119.50	1.20
13	A013S	S	106.50	1.07
13	A013Q	Q	105.87	1.06
13	A013F	F	101.75	1.02
13	A013C	C	100.86	1.01
13	A013T	T	95.70	0.96
13	A013M	M	91.53	0.92
13	A013N	N	86.10	0.86
13	A013W	W	85.59	0.86
13	A013E	E	83.94	0.84
13	A013P	P	72.98	0.73
13	A013D	D	50.95	0.51
13	A013R	R	27.55	0.28
13	A013H	H	7.61	0.08
13	A013L	L	7.22	0.07
14	L014S	S	150.94	1.51
14	L014V	V	144.29	1.44
14	L014A	A	141.26	1.41
14	L014F	F	129.64	1.30
14	L014W	W	125.51	1.26
14	L014M	M	117.13	1.17
14	L014G	G	107.37	1.07
14	L014I	I	95.68	0.96
14	L014H	H	89.97	0.90
14	L014N	N	78.82	0.79
14	L014Q	Q	67.18	0.67
14	L014Y	Y	62.88	0.63
14	L014K	K	41.53	0.42
14	L014E	E	40.85	0.41
14	L014R	R	32.05	0.32

14	L014P	P	8.63	0.09
15	L015G	G	144.04	1.44
15	L015T	T	134.29	1.34
15	L015M	M	128.01	1.28
15	L015C	C	125.90	1.26
15	L015V	V	119.25	1.19
15	L015Y	Y	118.30	1.18
15	L015W	W	116.98	1.17
15	L015A	A	109.90	1.10
15	L015F	F	102.58	1.03
15	L015S	S	98.62	0.99
15	L015P	P	84.79	0.85
15	L015Q	Q	53.19	0.53
15	L015K	K	49.84	0.50
15	L015N	N	43.69	0.44
15	L015H	H	41.81	0.42
15	L015E	E	41.21	0.41
15	L015R	R	10.80	0.11
16	I016W	W	153.84	1.54
16	I016S	S	129.21	1.29
16	I016G	G	122.70	1.23
16	I016A	A	116.84	1.17
16	I016C	C	109.27	1.09
16	I016V	V	103.40	1.03
16	I016Y	Y	101.35	1.01
16	I016T	T	88.85	0.89
16	I016H	H	75.02	0.75
16	I016N	N	71.58	0.72
16	I016F	F	57.93	0.58
16	I016E	E	38.94	0.39
16	I016P	P	23.34	0.23
16	I016R	R	18.70	0.19
16	I016L	L	11.62	0.12
17	S017M	M	129.94	1.30
17	S017A	A	129.04	1.29

<217>

17	S017R	R	126.92	1.27
17	S017L	L	118.09	1.18
17	S017H	H	116.37	1.16
17	S017K	K	114.85	1.15
17	S017V	V	112.73	1.13
17	S017D	D	112.49	1.12
17	S017C	C	110.18	1.10
17	S017G	G	108.87	1.09
17	S017E	E	108.53	1.09
17	S017Y	Y	104.68	1.05
17	S017I	I	91.72	0.92
17	S017T	T	88.65	0.89
17	S017F	F	86.66	0.87
17	S017P	P	83.37	0.83
18	V018M	M	141.00	1.41
18	V018T	T	118.00	1.18
18	V018A	A	117.48	1.17
18	V018W	W	112.69	1.13
18	V018E	E	110.30	1.10
18	V018S	S	104.38	1.04
18	V018R	R	102.00	1.02
18	V018C	C	99.00	0.99
18	V018D	D	92.82	0.93
18	V018F	F	90.00	0.90
18	V018	8	86.00	0.86
18	V018P	P	80.92	0.81
18	V018H	H	76.00	0.76
18	V018L	L	53.62	0.54
18	V018I	I	50.32	0.50
18	V018G	G	49.92	0.50
18	V018Y	Y	42.71	0.43
19	A019E	E	142.51	1.43
19	A019M	M	121.02	1.21
19	A019C	C	114.96	1.15
19	A019S	S	113.37	1.13

19	A019L	L	112.60	1.13
19	A019W	W	109.73	1.10
19	A019V	V	109.69	1.10
19	A019G	G	108.40	1.08
19	A019T	T	103.33	1.03
19	A019I	I	91.94	0.92
19	A019K	K	91.82	0.92
19	A019F	F	84.12	0.84
19	A019R	R	81.36	0.81
19	A019P	P	80.92	0.81
20	F020N	N	123.06	1.23
20	F020Q	Q	118.20	1.18
20	F020M	M	115.31	1.15
20	F020V	V	107.59	1.08
20	F020S	S	107.24	1.07
20	F020T	T	104.65	1.05
20	F020L	L	100.66	1.01
20	F020G	G	99.06	0.99
20	F020D	D	94.64	0.95
20	F020I	I	93.21	0.93
20	F020E	E	89.45	0.89
20	F020K	K	89.03	0.89
20	F020R	R	71.76	0.72
20	F020A	A	6.86	0.07
21	S021T	T	98.31	0.98
21	S021E	E	90.76	0.91
21	S021Q	Q	86.03	0.86
21	S021M	M	82.73	0.83
21	S021G	G	81.75	0.82
21	S021A	A	81.13	0.81
21	S021W	W	80.36	0.80
21	S021V	V	79.71	0.80
21	S021R	R	79.66	0.80
21	S021K	K	77.77	0.78
21	S021L	L	72.62	0.73

21	S021N	N	71.73	0.72
21	S021I	I	66.14	0.66
21	S021C	C	39.26	0.39
21	S021Y	Y	39.16	0.39
21	S021P	P	30.33	0.30
21	S021F	F	8.13	0.08
21	S021H	H	7.67	0.08
22	S022M	M	154.51	1.55
22	S022V	V	121.71	1.22
22	S022D	D	119.67	1.20
22	S022W	W	116.41	1.16
22	S022G	G	115.06	1.15
22	S022T	T	109.16	1.09
22	S022R	R	108.79	1.09
22	S022P	P	108.47	1.08
22	S022Y	Y	106.89	1.07
22	S022A	A	102.30	1.02
22	S022L	L	97.24	0.97
22	S022F	F	95.17	0.95
22	S022E	E	93.31	0.93
22	S022H	H	92.95	0.93
22	S022C	C	89.96	0.90
23	S023K	K	128.31	1.28
23	S023G	G	119.02	1.19
23	S023E	E	114.55	1.15
23	S023H	H	113.30	1.13
23	S023R	R	113.12	1.13
23	S023T	T	112.61	1.13
23	S023A	A	91.93	0.92
23	S023M	M	91.58	0.92
23	S023P	P	91.42	0.91
23	S023V	V	88.72	0.89
23	S023C	C	86.61	0.87
23	S023Q	Q	83.15	0.83
23	S023L	L	77.16	0.77

<220>

23	S023I	I	74.48	0.74
23	S023W	W	36.43	0.36
23	S023D	D	7.46	0.07
24	I024T	T	140.91	1.41
24	I024V	V	128.01	1.28
24	I024N	N	110.91	1.11
24	I024R	R	110.89	1.11
24	I024L	L	105.32	1.05
24	I024H	H	102.41	1.02
24	I024P	P	101.05	1.01
24	I024S	S	100.69	1.01
24	I024C	C	98.49	0.98
24	I024A	A	95.72	0.96
24	I024W	W	94.63	0.95
24	I024K	K	94.07	0.94
24	I024M	M	93.64	0.94
24	I024Y	Y	8.75	0.09
24	I024E	E	6.70	0.07
25	A025V	V	88.01	0.88
25	A025T	T	63.92	0.64
25	A025G	G	63.07	0.63
25	A025S	S	57.55	0.58
25	A025I	I	51.29	0.51
25	A025L	L	43.81	0.44
25	A025E	E	32.21	0.32
25	A025M	M	31.37	0.31
25	A025P	P	29.29	0.29
25	A025K	K	27.25	0.27
25	A025N	N	25.98	0.26
25	A025H	H	23.74	0.24
25	A025D	D	19.49	0.19
25	A025F	F	16.89	0.17
25	A025W	W	11.14	0.11
26	S026Y	Y	129.81	1.30
26	S026L	L	120.14	1.20

26	S026A	A	112.01	1.12
26	S026R	R	111.71	1.12
26	S026T	T	107.06	1.07
26	S026N	N	101.40	1.01
26	S026K	K	99.57	1.00
26	S026M	M	94.81	0.95
26	S026D	D	93.37	0.93
26	S026G	G	90.33	0.90
26	S026W	W	87.38	0.87
26	S026V	V	82.41	0.82
26	S026H	H	81.00	0.81
26	S026I	I	76.78	0.77
26	S026P	P	47.45	0.47
27	A027G	G	67.18	0.67
27	A027S	S	44.42	0.44
27	A027K	K	40.64	0.41
27	A027C	C	39.47	0.39
27	A027P	P	39.46	0.39
27	A027R	R	37.92	0.38
27	A027H	H	36.42	0.36
27	A027E	E	31.42	0.31
27	A027F	F	30.21	0.30
27	A027T	T	21.37	0.21
27	A027V	V	20.42	0.20
27	A027W	W	11.82	0.12
27	A027Q	Q	11.60	0.12
27	A027L	L	4.86	0.05
28	A028Q	Q	109.07	1.09
28	A028T	T	107.38	1.07
28	A028S	S	83.96	0.84
28	A028R	R	80.35	0.80
28	A028M	M	79.43	0.79
28	A028E	E	71.65	0.72
28	A028G	G	68.05	0.68
28	A028D	D	61.77	0.62

28	A028N	N	57.89	0.58
28	A028L	L	55.70	0.56
28	A028V	V	51.81	0.52
28	A028F	F	31.30	0.31
28	A028P	P	27.66	0.28
28	A028W	W	24.87	0.25
29	E029T	T	133.40	1.33
29	E029Y	Y	124.28	1.24
29	E029V	V	117.64	1.18
29	E029Q	Q	115.48	1.15
29	E029K	K	114.06	1.14
29	E029M	M	113.42	1.13
29	E029P	P	111.96	1.12
29	E029N	N	111.04	1.11
29	E029S	S	107.45	1.07
29	E029G	G	107.33	1.07
29	E029A	A	107.09	1.07
29	E029C	C	87.77	0.88
29	E029L	L	85.82	0.86
29	E029I	I	80.97	0.81
29	E029H	H	72.39	0.72
29	E029W	W	65.30	0.65
29	E029F	F	58.62	0.59
30	E030Q	Q	125.96	1.26
30	E030K	K	120.69	1.21
30	E030M	M	117.64	1.18
30	E030A	A	112.55	1.13
30	E030G	G	104.25	1.04
30	E030R	R	103.70	1.04
30	E030S	S	103.37	1.03
30	E030H	H	103.05	1.03
30	E030N	N	100.98	1.01
30	E030T	T	100.94	1.01
30	E030V	V	99.31	0.99
30	E030L	L	97.86	0.98

30	E030F	F	87.10	0.87
30	E030C	C	71.99	0.72
30	E030W	W	69.05	0.69
30	E030P	P	6.54	0.07
31	A031R	R	123.14	1.23
31	A031G	G	118.11	1.18
31	A031S	S	110.90	1.11
31	A031K	K	108.32	1.08
31	A031P	P	105.73	1.06
31	A031T	T	99.73	1.00
31	A031M	M	96.10	0.96
31	A031V	V	95.95	0.96
31	A031L	L	75.13	0.75
31	A031N	N	67.25	0.67
31	A031W	W	64.44	0.64
31	A031H	H	10.14	0.10
32	K032C	C	67.52	0.68
32	K032H	H	55.09	0.55
32	K032T	T	54.10	0.54
32	K032W	W	50.65	0.51
32	K032N	N	47.14	0.47
32	K032L	L	44.25	0.44
32	K032R	R	37.80	0.38
32	K032F	F	37.58	0.38
32	K032V	V	33.13	0.33
32	K032S	S	30.77	0.31
32	K032P	P	27.44	0.27
32	K032I	I	15.39	0.15
32	K032G	G	12.63	0.13
32	K032Y	Y	10.28	0.10
33	E033R	R	197.92	1.98
33	E033Q	Q	194.17	1.94
33	E033G	G	146.63	1.47
33	E033H	H	132.79	1.33
33	E033N	N	129.84	1.30

<224>

33	E033S	S	128.25	1.28
33	E033D	D	125.01	1.25
33	E033I	I	124.33	1.24
33	E033K	K	116.99	1.17
33	E033M	M	105.94	1.06
33	E033L	L	94.42	0.94
33	E033T	T	89.95	0.90
33	E033Y	Y	72.82	0.73
33	E033F	F	71.77	0.72
33	E033P	P	38.46	0.38
34	K034R	R	111.00	1.11
34	K034H	H	86.37	0.86
34	K034Q	Q	64.00	0.64
34	K034N	N	56.38	0.56
34	K034P	P	55.00	0.55
34	K034T	T	55.00	0.55
34	K034M	M	52.00	0.52
34	K034C	C	48.00	0.48
34	K034V	V	46.00	0.46
34	K034L	L	35.00	0.35
34	K034G	G	34.00	0.34
34	K034D	D	27.00	0.27
34	K034F	F	25.00	0.25
34	K034Y	Y	24.00	0.24
34	K034S	S	5.00	0.05
35	Y035F	F	92.05	0.92
35	Y035W	W	29.00	0.29
35	Y035V	V	23.00	0.23
35	Y035L	L	11.96	0.12
35	Y035C	C	11.74	0.12
35	Y035A	A	9.31	0.09
35	Y035R	R	8.51	0.09
35	Y035K	K	8.00	0.08
35	Y035S	S	7.60	0.08
35	Y035Q	Q	7.00	0.07

35	Y035P	P	6.25	0.06
35	Y035D	D	6.15	0.06
35	Y035E	E	6.10	0.06
35	Y035N	N	6.00	0.06
35	Y035T	T	6.00	0.06
35	Y035G	G	5.85	0.06
36	L036M	M	97.00	0.97
36	L036G	G	33.00	0.33
36	L036T	T	33.00	0.33
36	L036Y	Y	25.04	0.25
36	L036C	C	24.00	0.24
36	L036N	N	20.00	0.20
36	L036F	F	19.00	0.19
36	L036V	V	17.00	0.17
36	L036W	W	9.00	0.09
36	L036R	R	8.00	0.08
36	L036A	A	7.00	0.07
36	L036S	S	7.00	0.07
36	L036Q	Q	6.21	0.06
36	L036H	H	6.08	0.06
36	L036P	P	5.71	0.06
36	L036D	D	3.67	0.04
37	I037V	V	59.70	0.60
37	I037H	H	22.95	0.23
37	I037C	C	9.56	0.10
37	I037S	S	9.48	0.09
37	I037N	N	8.54	0.09
37	I037A	A	8.15	0.08
37	I037F	F	7.80	0.08
37	I037P	P	7.57	0.08
37	I037L	L	7.27	0.07
37	I037Y	Y	7.00	0.07
37	I037T	T	6.86	0.07
37	I037G	G	6.86	0.07
37	I037W	W	6.53	0.07

37	I037D	D	6.46	0.06
37	I037Q	Q	6.45	0.06
37	I037R	R	5.83	0.06
38	G038A	A	56.00	0.56
38	G038H	H	34.76	0.35
38	G038I	I	32.00	0.32
38	G038S	S	31.00	0.31
38	G038V	V	31.00	0.31
38	G038T	T	24.00	0.24
38	G038M	M	21.00	0.21
38	G038L	L	15.10	0.15
38	G038N	N	15.00	0.15
38	G038R	R	13.00	0.13
38	G038K	K	12.55	0.13
38	G038C	C	12.00	0.12
38	G038P	P	10.00	0.10
38	G038W	W	9.87	0.10
38	G038D	D	8.00	0.08
38	G038E	E	8.00	0.08
38	G038Y	Y	7.09	0.07
39	F039L	L	14.25	0.14
39	F039M	M	14.10	0.14
39	F039W	W	10.45	0.10
39	F039Y	Y	9.72	0.10
39	F039S	S	8.78	0.09
39	F039R	R	8.71	0.09
39	F039P	P	8.06	0.08
39	F039D	D	7.82	0.08
39	F039E	E	7.79	0.08
39	F039V	V	7.59	0.08
39	F039C	C	7.51	0.08
39	F039A	A	6.77	0.07
39	F039K	K	6.00	0.06
39	F039H	H	5.90	0.06
39	F039G	G	5.87	0.06

<227>

39	F039N	N	5.76	0.06
39	F039Q	Q	5.54	0.06
39	F039T	T	5.39	0.05
40	N040V	V	126.00	1.26
40	N040K	K	118.87	1.19
40	N040E	E	110.47	1.10
40	N040Q	Q	109.00	1.09
40	N040H	H	107.23	1.07
40	N040Y	Y	106.81	1.07
40	N040T	T	106.00	1.06
40	N040A	A	104.00	1.04
40	N040L	L	104.00	1.04
40	N040W	W	102.81	1.03
40	N040F	F	102.53	1.03
40	N040R	R	102.00	1.02
40	N040C	C	98.00	0.98
40	N040P	P	98.00	0.98
40	N040I	I	88.00	0.88
40	N040G	G	82.00	0.82
40	N040D	D	72.46	0.72
40	N040M	M	70.39	0.70
40	N040S	S	19.00	0.19
41	E041A	A	142.97	1.43
41	E041S	S	125.74	1.26
41	E041N	N	113.99	1.14
41	E041T	T	113.64	1.14
41	E041H	H	108.25	1.08
41	E041R	R	104.86	1.05
41	E041G	G	97.74	0.98
41	E041Y	Y	95.99	0.96
41	E041Q	Q	92.70	0.93
41	E041F	F	63.24	0.63
41	E041L	L	60.52	0.61
41	E041I	I	52.81	0.53
41	E041V	V	45.47	0.45

41	E041W	W	32.22	0.32
41	E041P	P	9.27	0.09
42	Q042M	M	148.09	1.48
42	Q042R	R	130.84	1.31
42	Q042P	P	121.40	1.21
42	Q042W	W	113.30	1.13
42	Q042S	S	95.75	0.96
42	Q042Y	Y	86.74	0.87
42	Q042K	K	83.66	0.84
42	Q042F	F	75.21	0.75
42	Q042H	H	74.84	0.75
42	Q042L	L	73.83	0.74
42	Q042A	A	73.23	0.73
42	Q042I	I	67.65	0.68
42	Q042E	E	58.12	0.58
42	Q042D	D	57.23	0.57
42	Q042T	T	57.19	0.57
42	Q042V	V	42.92	0.43
42	Q042C	C	42.91	0.43
42	Q042N	N	27.71	0.28
43	E043S	S	163.96	1.64
43	E043R	R	124.40	1.24
43	E043D	D	120.22	1.20
43	E043H	H	105.43	1.05
43	E043A	A	104.90	1.05
43	E043L	L	95.08	0.95
43	E043N	N	91.28	0.91
43	E043G	G	90.19	0.90
43	E043F	F	69.79	0.70
43	E043Y	Y	66.91	0.67
43	E043C	C	66.68	0.67
43	E043W	W	64.35	0.64
43	E043V	V	53.37	0.53
43	E043P	P	42.96	0.43
43	E043M	M	8.82	0.09

43	E043T	T	5.27	0.05
44	A044D	D	136.49	1.36
44	A044Q	Q	113.72	1.14
44	A044E	E	108.45	1.08
44	A044S	S	102.67	1.03
44	A044N	N	90.51	0.91
44	A044R	R	82.63	0.83
44	A044G	G	82.29	0.82
44	A044C	C	80.18	0.80
44	A044L	L	79.49	0.79
44	A044H	H	69.28	0.69
44	A044V	V	50.75	0.51
44	A044Y	Y	44.57	0.45
44	A044T	T	42.35	0.42
44	A044I	I	29.22	0.29
44	A044P	P	16.75	0.17
44	A044F	F	8.08	0.08
45	V045L	L	128.46	1.28
45	V045A	A	122.17	1.22
45	V045C	C	117.75	1.18
45	V045R	R	89.35	0.89
45	V045T	T	83.00	0.83
45	V045I	I	81.29	0.81
45	V045Q	Q	69.47	0.69
45	V045F	F	42.91	0.43
45	V045Y	Y	38.60	0.39
45	V045S	S	31.46	0.31
45	V045D	D	30.61	0.31
45	V045H	H	15.52	0.16
45	V045P	P	13.16	0.13
45	V045N	N	8.45	0.08
46	S046H	H	180.17	1.80
46	S046R	R	163.56	1.64
46	S046T	T	145.04	1.45
46	S046A	A	134.77	1.35

<230>

46	S046Q	Q	133.68	1.34
46	S046Y	Y	126.03	1.26
46	S046W	W	108.60	1.09
46	S046V	V	99.68	1.00
46	S046L	L	95.29	0.95
46	S046G	G	94.57	0.95
46	S046D	D	94.34	0.94
46	S046E	E	80.44	0.80
46	S046F	F	68.99	0.69
46	S046P	P	46.63	0.47
47	E047D	D	161.68	1.62
47	E047R	R	127.01	1.27
47	E047K	K	105.12	1.05
47	E047Q	Q	103.73	1.04
47	E047H	H	102.20	1.02
47	E047A	A	99.30	0.99
47	E047Y	Y	93.27	0.93
47	E047N	N	88.69	0.89
47	E047S	S	82.06	0.82
47	E047T	T	81.19	0.81
47	E047V	V	78.37	0.78
47	E047L	L	68.89	0.69
47	E047G	G	55.86	0.56
47	E047F	F	50.48	0.50
47	E047C	C	49.85	0.50
47	E047W	W	41.91	0.42
47	E047P	P	22.22	0.22
48	F048W	W	100.77	1.01
48	F048M	M	20.43	0.20
48	F048L	L	16.60	0.17
48	F048I	I	11.37	0.11
48	F048V	V	10.65	0.11
48	F048R	R	9.36	0.09
48	F048C	C	9.30	0.09
48	F048P	P	8.69	0.09

48	F048S	S	8.60	0.09
48	F048K	K	8.47	0.08
48	F048H	H	8.40	0.08
48	F048N	N	8.39	0.08
48	F048E	E	8.29	0.08
48	F048D	D	7.97	0.08
48	F048G	G	7.50	0.07
49	V049G	G	135.00	1.35
49	V049C	C	95.00	0.95
49	V049R	R	87.00	0.87
49	V049K	K	58.91	0.59
49	V049M	M	58.41	0.58
49	V049A	A	54.56	0.55
49	V049W	W	46.00	0.46
49	V049T	T	38.00	0.38
49	V049S	S	25.00	0.25
49	V049Y	Y	20.00	0.20
49	V049Q	Q	17.26	0.17
49	V049P	P	15.00	0.15
49	V049L	L	14.00	0.14
49	V049F	F	13.93	0.14
49	V049E	E	10.00	0.10
49	V049N	N	9.00	0.09
49	V049D	D	8.00	0.08
49	V049H	H	8.00	0.08
49	V049I	I	8.00	0.08
50	E050H	H	112.00	1.12
50	E050A	A	96.00	0.96
50	E050S	S	83.00	0.83
50	E050D	D	82.58	0.83
50	E050Q	Q	77.00	0.77
50	E050R	R	77.00	0.77
50	E050K	K	73.00	0.73
50	E050M	M	71.98	0.72
50	E050T	T	64.00	0.64

50	E050G	G	61.95	0.62
50	E050Y	Y	52.00	0.52
50	E050L	L	49.00	0.49
50	E050F	F	47.00	0.47
50	E050I	I	46.83	0.47
50	E050W	W	42.01	0.42
50	E050N	N	32.42	0.32
50	E050V	V	29.83	0.30
50	E050P	P	6.00	0.06
51	Q051M	M	119.81	1.20
51	Q051R	R	108.69	1.09
51	Q051K	K	97.54	0.98
51	Q051A	A	84.10	0.84
51	Q051T	T	83.15	0.83
51	Q051E	E	75.80	0.76
51	Q051N	N	75.33	0.75
51	Q051H	H	69.22	0.69
51	Q051S	S	68.17	0.68
51	Q051G	G	66.07	0.66
51	Q051L	L	65.45	0.65
51	Q051D	D	52.77	0.53
51	Q051I	I	45.84	0.46
51	Q051F	F	44.57	0.45
51	Q051C	C	43.93	0.44
51	Q051V	V	43.56	0.44
51	Q051P	P	5.94	0.06
52	V052L	L	108.19	1.08
52	V052W	W	104.92	1.05
52	V052M	M	88.46	0.88
52	V052I	I	79.27	0.79
52	V052A	A	72.78	0.73
52	V052H	H	52.94	0.53
52	V052F	F	52.32	0.52
52	V052C	C	46.72	0.47
52	V052T	T	39.52	0.40

52	V052G	G	39.28	0.39
52	V052Q	Q	34.36	0.34
52	V052R	R	18.76	0.19
52	V052N	N	15.55	0.16
52	V052S	S	15.02	0.15
52	V052E	E	7.54	0.08
52	V052P	P	6.39	0.06
52	V052D	D	5.86	0.06
53	E053A	A	140.37	1.40
53	E053K	K	127.20	1.27
53	E053R	R	119.65	1.20
53	E053H	H	109.57	1.10
53	E053N	N	98.87	0.99
53	E053L	L	98.50	0.98
53	E053P	P	97.72	0.98
53	E053V	V	94.43	0.94
53	E053F	F	94.05	0.94
53	E053S	S	89.81	0.90
53	E053Y	Y	86.04	0.86
53	E053T	T	83.88	0.84
53	E053C	C	81.12	0.81
53	E053W	W	77.76	0.78
53	E053D	D	68.39	0.68
53	E053G	G	67.56	0.68
53	E053Q	Q	13.97	0.14
54	A054L	L	137.74	1.38
54	A054S	S	136.79	1.37
54	A054R	R	97.90	0.98
54	A054D	D	97.17	0.97
54	A054G	G	90.30	0.90
54	A054T	T	88.68	0.89
54	A054C	C	83.34	0.83
54	A054N	N	83.10	0.83
54	A054I	I	81.67	0.82
54	A054K	K	81.00	0.81

54	A054H	H	77.17	0.77
54	A054Y	Y	74.93	0.75
54	A054F	F	60.70	0.61
54	A054V	V	60.08	0.60
54	A054P	P	23.50	0.24
55	N055H	H	85.10	0.85
55	N055A	A	72.93	0.73
55	N055M	M	69.99	0.70
55	N055Q	Q	63.43	0.63
55	N055E	E	59.81	0.60
55	N055R	R	53.48	0.53
55	N055G	G	46.03	0.46
55	N055K	K	45.82	0.46
55	N055S	S	41.16	0.41
55	N055Y	Y	38.86	0.39
55	N055I	I	35.72	0.36
55	N055V	V	29.77	0.30
55	N055F	F	28.77	0.29
55	N055L	L	26.96	0.27
55	N055P	P	22.56	0.23
56	D056G	G	124.69	1.25
56	D056S	S	116.41	1.16
56	D056K	K	112.18	1.12
56	D056R	R	101.95	1.02
56	D056N	N	100.33	1.00
56	D056H	H	98.12	0.98
56	D056A	A	86.03	0.86
56	D056T	T	68.54	0.69
56	D056Y	Y	56.65	0.57
56	D056F	F	49.98	0.50
56	D056P	P	47.08	0.47
56	D056V	V	46.58	0.47
56	D056L	L	38.46	0.38
56	D056I	I	31.73	0.32
56	D056W	W	30.14	0.30

<235>

57	E057N	N	252.24	2.52
57	E057H	H	157.45	1.57
57	E057S	S	145.56	1.46
57	E057G	G	145.37	1.45
57	E057Q	Q	143.17	1.43
57	E057C	C	139.04	1.39
57	E057M	M	128.97	1.29
57	E057D	D	128.47	1.28
57	E057R	R	120.64	1.21
57	E057K	K	118.25	1.18
57	E057W	W	116.28	1.16
57	E057F	F	107.79	1.08
57	E057V	V	100.08	1.00
57	E057L	L	99.38	0.99
57	E057P	P	65.08	0.65
57	E057I	I	61.84	0.62
57	E057T	T	26.62	0.27
58	V058R	R	129.15	1.29
58	V058Y	Y	123.00	1.23
58	V058H	H	123.00	1.23
58	V058M	M	112.92	1.13
58	V058K	K	110.16	1.10
58	V058C	C	109.43	1.09
58	V058L	L	106.58	1.07
58	V058A	A	104.97	1.05
58	V058F	F	102.77	1.03
58	V058N	N	95.69	0.96
58	V058S	S	93.83	0.94
58	V058I	I	93.50	0.94
58	V058T	T	91.46	0.91
58	V058G	G	87.22	0.87
58	V058D	D	74.09	0.74
58	V058E	E	67.17	0.67
58	V058P	P	61.56	0.62
58	V058Q	Q	6.39	0.06

59	A059R	R	187.27	1.87
59	A059Y	Y	172.73	1.73
59	A059V	V	159.93	1.60
59	A059I	I	136.28	1.36
59	A059F	F	135.64	1.36
59	A059W	W	125.18	1.25
59	A059M	M	123.19	1.23
59	A059L	L	121.10	1.21
59	A059H	H	110.74	1.11
59	A059S	S	91.95	0.92
59	A059T	T	87.06	0.87
59	A059Q	Q	73.05	0.73
59	A059G	G	71.74	0.72
59	A059D	D	69.29	0.69
59	A059N	N	66.70	0.67
59	A059E	E	64.01	0.64
59	A059P	P	59.36	0.59
59	A059K	K	33.09	0.33
60	I060L	L	103.80	1.04
60	I060V	V	97.35	0.97
60	I060R	R	96.55	0.97
60	I060M	M	90.48	0.90
60	I060K	K	75.87	0.76
60	I060Y	Y	73.84	0.74
60	I060Q	Q	72.24	0.72
60	I060N	N	67.92	0.68
60	I060E	E	65.70	0.66
60	I060D	D	65.27	0.65
60	I060A	A	61.85	0.62
60	I060G	G	61.08	0.61
60	I060S	S	51.51	0.52
60	I060T	T	46.40	0.46
60	I060H	H	36.92	0.37
60	I060P	P	19.01	0.19
60	I060W	W	5.42	0.05

<237>

61	L061M	M	129.95	1.30
61	L061T	T	124.71	1.25
61	L061A	A	123.36	1.23
61	L061I	I	110.37	1.10
61	L061V	V	108.74	1.09
61	L061G	G	108.57	1.09
61	L061P	P	102.67	1.03
61	L061D	D	102.28	1.02
61	L061S	S	97.83	0.98
61	L061K	K	95.81	0.96
61	L061Q	Q	95.57	0.96
61	L061E	E	91.81	0.92
61	L061R	R	85.77	0.86
61	L061H	H	66.88	0.67
61	L061N	N	35.28	0.35
62	S062V	V	161.32	1.61
62	S062W	W	157.12	1.57
62	S062M	M	150.12	1.50
62	S062P	P	139.57	1.40
62	S062N	N	134.16	1.34
62	S062H	H	133.29	1.33
62	S062Q	Q	124.88	1.25
62	S062D	D	120.10	1.20
62	S062G	G	118.35	1.18
62	S062I	I	116.17	1.16
62	S062A	A	115.62	1.16
62	S062L	L	107.95	1.08
62	S062C	C	107.34	1.07
62	S062R	R	107.31	1.07
62	S062T	T	105.80	1.06
62	S062K	K	96.88	0.97
62	S062F	F	77.46	0.77
63	E063A	A	151.93	1.52
63	E063G	G	151.54	1.52
63	E063D	D	143.23	1.43

63	E063R	R	138.35	1.38
63	E063T	T	133.00	1.33
63	E063Q	Q	126.78	1.27
63	E063C	C	122.66	1.23
63	E063K	K	114.56	1.15
63	E063H	H	107.55	1.08
63	E063I	I	103.97	1.04
63	E063S	S	101.87	1.02
63	E063Y	Y	96.83	0.97
63	E063L	L	94.72	0.95
63	E063W	W	72.62	0.73
63	E063V	V	68.49	0.68
63	E063F	F	54.52	0.55
64	E064T	T	118.91	1.19
64	E064R	R	110.53	1.11
64	E064S	S	109.23	1.09
64	E064A	A	102.92	1.03
64	E064H	H	93.59	0.94
64	E064Q	Q	91.03	0.91
64	E064C	C	90.99	0.91
64	E064Y	Y	88.33	0.88
64	E064G	G	79.36	0.79
64	E064F	F	67.32	0.67
64	E064V	V	66.57	0.67
64	E064L	L	62.47	0.62
64	E064W	W	60.29	0.60
64	E064P	P	44.81	0.45
64	E064K	K	6.18	0.06
64	E064I	I	5.10	0.05
65	E065M	M	123.46	1.23
65	E065R	R	116.00	1.16
65	E065Q	Q	109.80	1.10
65	E065A	A	107.63	1.08
65	E065V	V	104.79	1.05
65	E065Y	Y	104.29	1.04

65	E065K	K	100.28	1.00
65	E065C	C	94.55	0.95
65	E065S	S	93.73	0.94
65	E065T	T	89.80	0.90
65	E065L	L	89.07	0.89
65	E065I	I	87.89	0.88
65	E065G	G	80.18	0.80
65	E065D	D	71.30	0.71
65	E065W	W	70.97	0.71
65	E065P	P	56.54	0.57
66	E066S	S	135.22	1.35
66	E066A	A	128.91	1.29
66	E066T	T	127.73	1.28
66	E066P	P	123.84	1.24
66	E066R	R	121.88	1.22
66	E066Y	Y	115.52	1.16
66	E066K	K	108.25	1.08
66	E066G	G	105.02	1.05
66	E066V	V	104.05	1.04
66	E066M	M	103.51	1.04
66	E066Q	Q	100.84	1.01
66	E066W	W	99.26	0.99
66	E066L	L	93.75	0.94
66	E066F	F	91.41	0.91
66	E066N	N	87.34	0.87
66	E066D	D	71.95	0.72
66	E066H	H	36.85	0.37
67	V067I	I	113.21	1.13
67	V067M	M	73.82	0.74
67	V067L	L	72.32	0.72
67	V067Y	Y	70.71	0.71
67	V067F	F	58.59	0.59
67	V067C	C	56.98	0.57
67	V067R	R	55.35	0.55
67	V067A	A	54.90	0.55

67	V067T	T	41.46	0.41
67	V067P	P	34.73	0.35
67	V067S	S	28.45	0.28
67	V067E	E	20.66	0.21
67	V067K	K	20.54	0.21
67	V067N	N	17.25	0.17
67	V067G	G	15.25	0.15
67	V067D	D	13.59	0.14
67	V067H	H	6.18	0.06
68	E068K	K	102.16	1.02
68	E068Q	Q	101.87	1.02
68	E068S	S	85.61	0.86
68	E068D	D	83.72	0.84
68	E068R	R	80.31	0.80
68	E068N	N	80.22	0.80
68	E068P	P	74.49	0.74
68	E068T	T	71.78	0.72
68	E068G	G	68.18	0.68
68	E068A	A	68.12	0.68
68	E068C	C	67.44	0.67
68	E068Y	Y	58.18	0.58
68	E068M	M	48.30	0.48
68	E068V	V	46.31	0.46
68	E068I	I	25.39	0.25
68	E068L	L	20.98	0.21
69	I069V	V	123.93	1.24
69	I069L	L	79.94	0.80
69	I069N	N	42.27	0.42
69	I069S	S	27.06	0.27
69	I069F	F	22.49	0.22
69	I069A	A	19.56	0.20
69	I069T	T	10.74	0.11
69	I069Y	Y	9.65	0.10
69	I069W	W	9.01	0.09
69	I069P	P	7.84	0.08

<241>

69	I069R	R	7.15	0.07
69	I069Q	Q	7.13	0.07
69	I069G	G	6.86	0.07
69	I069D	D	6.38	0.06
70	E070T	T	135.22	1.35
70	E070R	R	127.41	1.27
70	E070N	N	119.02	1.19
70	E070S	S	118.79	1.19
70	E070V	V	102.83	1.03
70	E070G	G	92.33	0.92
70	E070D	D	89.84	0.90
70	E070C	C	87.77	0.88
70	E070A	A	80.65	0.81
70	E070I	I	77.03	0.77
70	E070Y	Y	73.97	0.74
70	E070L	L	73.91	0.74
70	E070P	P	64.34	0.64
70	E070F	F	60.29	0.60
70	E070W	W	43.57	0.44
70	E070K	K	6.39	0.06
71	L071V	V	127.43	1.27
71	L071I	I	122.69	1.23
71	L071D	D	77.31	0.77
71	L071M	M	31.70	0.32
71	L071T	T	25.45	0.25
71	L071P	P	24.11	0.24
71	L071R	R	15.22	0.15
71	L071K	K	14.69	0.15
71	L071S	S	13.41	0.13
71	L071W	W	11.88	0.12
71	L071Q	Q	10.19	0.10
71	L071E	E	9.80	0.10
71	L071C	C	9.52	0.10
71	L071G	G	9.32	0.09
71	L071H	H	7.50	0.08

72	L072V	V	98.37	0.98
72	L072C	C	73.50	0.73
72	L072K	K	72.20	0.72
72	L072R	R	61.79	0.62
72	L072H	H	47.06	0.47
72	L072M	M	44.46	0.44
72	L072N	N	43.42	0.43
72	L072T	T	38.32	0.38
72	L072G	G	34.67	0.35
72	L072Q	Q	30.17	0.30
72	L072Y	Y	29.73	0.30
72	L072W	W	28.66	0.29
72	L072D	D	27.32	0.27
72	L072E	E	24.56	0.25
72	L072S	S	23.76	0.24
72	L072A	A	8.70	0.09
72	L072P	P	6.61	0.07
73	H073W	W	75.35	0.75
73	H073S	S	63.17	0.63
73	H073F	F	60.20	0.60
73	H073K	K	32.91	0.33
73	H073C	C	31.74	0.32
73	H073R	R	27.97	0.28
73	H073T	T	25.37	0.25
73	H073G	G	24.90	0.25
73	H073Q	Q	24.68	0.25
73	H073V	V	22.30	0.22
73	H073N	N	20.69	0.21
73	H073M	M	20.67	0.21
73	H073I	I	19.80	0.20
73	H073L	L	19.63	0.20
73	H073D	D	16.98	0.17
73	H073P	P	6.69	0.07
73	H073Y	Y	6.36	0.06
73	H073A	A	5.49	0.05

74	E074Q	Q	161.32	1.61
74	E074R	R	142.60	1.43
74	E074V	V	136.24	1.36
74	E074K	K	127.86	1.28
74	E074A	A	116.09	1.16
74	E074I	I	105.81	1.06
74	E074S	S	97.24	0.97
74	E074T	T	86.73	0.87
74	E074C	C	86.04	0.86
74	E074H	H	63.24	0.63
74	E074W	W	59.27	0.59
74	E074L	L	51.36	0.51
74	E074Y	Y	41.17	0.41
74	E074F	F	37.16	0.37
74	E074G	G	33.60	0.34
74	E074D	D	30.16	0.30
74	E074P	P	29.71	0.30
75	F075Y	Y	103.89	1.04
75	F075W	W	16.17	0.16
75	F075M	M	16.15	0.16
75	F075P	P	9.44	0.09
75	F075L	L	9.32	0.09
75	F075H	H	7.78	0.08
75	F075C	C	7.76	0.08
75	F075T	T	7.40	0.07
75	F075K	K	7.11	0.07
75	F075S	S	6.90	0.07
75	F075R	R	6.89	0.07
75	F075N	N	6.79	0.07
75	F075G	G	6.63	0.07
75	F075E	E	6.57	0.07
76	E076K	K	125.00	1.25
76	E076S	S	111.17	1.11
76	E076T	T	110.72	1.11
76	E076H	H	110.48	1.10

76	E076N	N	110.33	1.10
76	E076D	D	100.51	1.01
76	E076R	R	95.92	0.96
76	E076Y	Y	78.30	0.78
76	E076L	L	68.68	0.69
76	E076C	C	62.01	0.62
76	E076M	M	60.90	0.61
76	E076G	G	60.13	0.60
76	E076A	A	58.13	0.58
76	E076F	F	56.00	0.56
76	E076V	V	53.16	0.53
76	E076I	I	41.46	0.41
76	E076W	W	36.12	0.36
76	E076P	P	12.90	0.13
77	T077N	N	137.52	1.38
77	T077H	H	121.65	1.22
77	T077Y	Y	89.82	0.90
77	T077S	S	86.50	0.87
77	T077R	R	86.26	0.86
77	T077W	W	76.37	0.76
77	T077E	E	68.68	0.69
77	T077C	C	56.68	0.57
77	T077M	M	53.12	0.53
77	T077F	F	52.13	0.52
77	T077A	A	48.97	0.49
77	T077G	G	43.04	0.43
77	T077K	K	38.82	0.39
77	T077I	I	28.21	0.28
77	T077V	V	25.58	0.26
77	T077L	L	21.11	0.21
77	T077P	P	5.57	0.06
78	I078T	T	105.05	1.05
78	I078F	F	86.64	0.87
78	I078L	L	48.68	0.49
78	I078V	V	42.29	0.42

78	I078H	H	32.83	0.33
78	I078C	C	14.99	0.15
78	I078Y	Y	13.55	0.14
78	I078S	S	10.19	0.10
78	I078W	W	9.59	0.10
78	I078A	A	8.51	0.09
78	I078G	G	8.10	0.08
78	I078K	K	7.77	0.08
78	I078P	P	7.76	0.08
78	I078R	R	6.63	0.07
79	P079S	S	18.58	0.19
79	P079Q	Q	10.04	0.10
79	P079M	M	9.48	0.09
79	P079K	K	9.48	0.09
79	P079R	R	8.79	0.09
79	P079G	G	8.40	0.08
79	P079E	E	8.01	0.08
79	P079T	T	7.60	0.08
79	P079V	V	7.38	0.07
79	P079L	L	6.94	0.07
79	P079C	C	6.59	0.07
79	P079A	A	6.54	0.07
79	P079I	I	5.84	0.06
80	V080C	C	57.46	0.57
80	V080I	I	21.08	0.21
80	V080T	T	20.20	0.20
80	V080S	S	10.00	0.10
80	V080L	L	9.00	0.09
80	V080E	E	7.00	0.07
80	V080M	M	7.00	0.07
80	V080Y	Y	7.00	0.07
80	V080N	N	6.44	0.06
80	V080W	W	6.32	0.06
80	V080P	P	6.04	0.06
80	V080D	D	6.00	0.06

<246>

80	V080G	G	6.00	0.06
80	V080H	H	6.00	0.06
80	V080K	K	6.00	0.06
80	V080R	R	6.00	0.06
81	L081V	V	49.01	0.49
81	L081A	A	41.20	0.41
81	L081M	M	26.29	0.26
81	L081T	T	15.28	0.15
81	L081F	F	11.66	0.12
81	L081W	W	9.72	0.10
81	L081Y	Y	8.23	0.08
81	L081S	S	7.53	0.08
81	L081P	P	7.23	0.07
81	L081D	D	6.93	0.07
81	L081G	G	6.62	0.07
81	L081E	E	6.60	0.07
81	L081H	H	5.90	0.06
81	L081R	R	5.68	0.06
82	S082L	L	129.13	1.29
82	S082Q	Q	128.15	1.28
82	S082V	V	103.94	1.04
82	S082H	H	97.35	0.97
82	S082Y	Y	91.12	0.91
82	S082W	W	88.77	0.89
82	S082A	A	87.22	0.87
82	S082E	E	85.81	0.86
82	S082F	F	81.39	0.81
82	S082C	C	79.80	0.80
82	S082N	N	61.37	0.61
82	S082R	R	58.97	0.59
82	S082G	G	54.28	0.54
82	S082T	T	54.07	0.54
82	S082D	D	18.02	0.18
82	S082P	P	6.19	0.06
83	V083N	N	109.58	1.10

<247>

83	V083Y	Y	102.00	1.02
83	V083L	L	97.61	0.98
83	V083G	G	42.23	0.42
83	V083H	H	33.13	0.33
83	V083K	K	11.63	0.12
83	V083C	C	10.69	0.11
83	V083T	T	7.51	0.08
83	V083Q	Q	7.00	0.07
83	V083D	D	6.61	0.07
83	V083I	I	6.32	0.06
83	V083M	M	6.20	0.06
83	V083R	R	6.04	0.06
83	V083E	E	5.77	0.06
83	V083P	P	5.67	0.06
83	V083S	S	4.97	0.05
83	V083W	W	4.69	0.05
83	V083F	F	4.62	0.05
83	V083A	A	4.61	0.05
84	E084L	L	153.83	1.54
84	E084K	K	150.95	1.51
84	E084V	V	148.36	1.48
84	E084M	M	139.82	1.40
84	E084S	S	135.46	1.35
84	E084A	A	127.39	1.27
84	E084G	G	124.12	1.24
84	E084T	T	123.30	1.23
84	E084R	R	116.66	1.17
84	E084F	F	111.03	1.11
84	E084Q	Q	101.46	1.01
84	E084N	N	100.11	1.00
84	E084W	W	98.81	0.99
84	E084I	I	98.06	0.98
84	E084C	C	95.71	0.96
84	E084Y	Y	89.76	0.90
84	E084H	H	75.48	0.75

<248>

84	E084D	D	50.64	0.51
84	E084P	P	7.96	0.08
85	L085F	F	100.71	1.01
85	L085V	V	87.45	0.87
85	L085A	A	66.38	0.66
85	L085G	G	41.13	0.41
85	L085T	T	27.10	0.27
85	L085Q	Q	12.11	0.12
85	L085N	N	11.63	0.12
85	L085R	R	10.96	0.11
85	L085W	W	8.95	0.09
85	L085H	H	7.59	0.08
85	L085E	E	6.71	0.07
85	L085K	K	6.71	0.07
86	S086R	R	133.45	1.33
86	S086D	D	124.54	1.25
86	S086K	K	120.11	1.20
86	S086N	N	118.58	1.19
86	S086A	A	114.90	1.15
86	S086C	C	94.83	0.95
86	S086L	L	80.83	0.81
86	S086M	M	77.93	0.78
86	S086Q	Q	77.62	0.78
86	S086E	E	61.08	0.61
86	S086I	I	45.96	0.46
86	S086G	G	43.12	0.43
86	S086P	P	40.52	0.41
87	P087K	K	89.04	0.89
87	P087R	R	62.86	0.63
87	P087T	T	58.88	0.59
87	P087H	H	58.22	0.58
87	P087S	S	56.23	0.56
87	P087N	N	51.90	0.52
87	P087V	V	47.19	0.47
87	P087Q	Q	46.90	0.47

87	P087I	I	38.44	0.38
87	P087G	G	35.49	0.35
87	P087M	M	32.83	0.33
87	P087D	D	27.06	0.27
87	P087L	L	25.98	0.26
87	P087W	W	15.17	0.15
88	E088P	P	128.01	1.28
88	E088D	D	103.64	1.04
88	E088G	G	98.73	0.99
88	E088I	I	76.72	0.77
88	E088Y	Y	73.31	0.73
88	E088R	R	71.94	0.72
88	E088T	T	71.19	0.71
88	E088V	V	69.78	0.70
88	E088W	W	59.58	0.60
88	E088S	S	57.12	0.57
88	E088Q	Q	45.20	0.45
88	E088N	N	32.31	0.32
88	E088H	H	31.97	0.32
88	E088L	L	28.70	0.29
88	E088K	K	24.60	0.25
89	D089K	K	110.11	1.10
89	D089H	H	101.55	1.02
89	D089A	A	97.39	0.97
89	D089N	N	88.63	0.89
89	D089C	C	86.69	0.87
89	D089S	S	68.40	0.68
89	D089M	M	60.16	0.60
89	D089T	T	57.25	0.57
89	D089R	R	56.85	0.57
89	D089F	F	56.61	0.57
89	D089Q	Q	56.61	0.57
89	D089G	G	50.16	0.50
89	D089V	V	49.23	0.49
89	D089E	E	38.22	0.38

<250>

89	D089L	L	10.85	0.11
89	D089P	P	7.52	0.08
90	V090K	K	93.59	0.94
90	V090A	A	77.43	0.77
90	V090L	L	74.46	0.74
90	V090R	R	70.69	0.71
90	V090I	I	68.15	0.68
90	V090C	C	56.98	0.57
90	V090T	T	51.24	0.51
90	V090S	S	23.79	0.24
90	V090M	M	23.60	0.24
90	V090H	H	21.17	0.21
90	V090Y	Y	14.44	0.14
90	V090F	F	13.27	0.13
90	V090W	W	11.98	0.12
90	V090G	G	10.52	0.11
90	V090P	P	7.80	0.08
90	V090N	N	7.52	0.08
90	V090D	D	7.09	0.07
91	D091E	E	148.69	1.49
91	D091P	P	144.01	1.44
91	D091A	A	142.94	1.43
91	D091T	T	136.68	1.37
91	D091N	N	124.34	1.24
91	D091K	K	118.47	1.18
91	D091W	W	108.37	1.08
91	D091R	R	107.58	1.08
91	D091F	F	104.13	1.04
91	D091M	M	98.03	0.98
91	D091C	C	97.38	0.97
91	D091V	V	87.03	0.87
91	D091L	L	81.60	0.82
91	D091G	G	18.66	0.19
91	D091S	S	5.81	0.06
92	A092R	R	139.58	1.40

<251>

92	A092K	K	135.47	1.35
92	A092T	T	134.11	1.34
92	A092M	M	132.58	1.33
92	A092V	V	130.35	1.30
92	A092Q	Q	121.93	1.22
92	A092I	I	119.20	1.19
92	A092L	L	117.04	1.17
92	A092S	S	105.71	1.06
92	A092E	E	97.94	0.98
92	A092C	C	89.25	0.89
92	A092H	H	88.18	0.88
92	A092Y	Y	84.08	0.84
92	A092D	D	83.51	0.84
92	A092F	F	76.69	0.77
92	A092W	W	70.16	0.70
92	A092P	P	8.53	0.09
92	A092G	G	7.42	0.07
93	L093F	F	60.69	0.61
93	L093I	I	53.14	0.53
93	L093M	M	51.15	0.51
93	L093V	V	44.68	0.45
93	L093Y	Y	12.17	0.12
93	L093W	W	11.12	0.11
93	L093A	A	11.07	0.11
93	L093T	T	9.55	0.10
93	L093N	N	7.93	0.08
93	L093R	R	7.73	0.08
93	L093Q	Q	7.40	0.07
93	L093K	K	7.36	0.07
93	L093G	G	7.31	0.07
93	L093E	E	7.18	0.07
93	L093P	P	7.17	0.07
93	L093H	H	6.92	0.07
93	L093S	S	6.62	0.07
93	L093C	C	6.43	0.06

94	E094A	A	130.00	1.30
94	E094M	M	127.00	1.27
94	E094T	T	125.00	1.25
94	E094R	R	115.00	1.15
94	E094S	S	111.00	1.11
94	E094L	L	103.00	1.03
94	E094F	F	96.41	0.96
94	E094V	V	94.00	0.94
94	E094C	C	88.52	0.89
94	E094N	N	86.00	0.86
94	E094D	D	76.35	0.76
94	E094G	G	71.00	0.71
94	E094P	P	15.00	0.15
94	E094I	I	6.00	0.06
94	E094K	K	6.00	0.06
95	L095E	E	144.50	1.44
95	L095R	R	143.47	1.43
95	L095K	K	132.53	1.33
95	L095S	S	127.82	1.28
95	L095C	C	126.67	1.27
95	L095A	A	123.21	1.23
95	L095T	T	119.38	1.19
95	L095D	D	112.59	1.13
95	L095G	G	107.90	1.08
95	L095V	V	101.75	1.02
95	L095H	H	101.25	1.01
95	L095Y	Y	72.15	0.72
95	L095I	I	66.43	0.66
95	L095F	F	55.60	0.56
95	L095W	W	54.88	0.55
95	L095M	M	49.85	0.50
95	L095P	P	43.73	0.44
95	L095N	N	7.87	0.08
96	D096Y	Y	154.21	1.54
96	D096H	H	154.17	1.54

96	D096C	C	131.97	1.32
96	D096S	S	125.64	1.26
96	D096L	L	122.29	1.22
96	D096E	E	115.11	1.15
96	D096I	I	111.95	1.12
96	D096W	W	111.27	1.11
96	D096V	V	103.05	1.03
96	D096T	T	91.78	0.92
96	D096F	F	81.46	0.81
96	D096G	G	66.73	0.67
96	D096R	R	54.25	0.54
96	D096P	P	30.89	0.31
96	D096K	K	24.76	0.25
96	D096M	M	7.93	0.08
96	D096N	N	7.71	0.08
97	P097E	E	73.24	0.73
97	P097S	S	66.40	0.66
97	P097D	D	63.49	0.63
97	P097A	A	58.83	0.59
97	P097K	K	57.21	0.57
97	P097Q	Q	54.30	0.54
97	P097N	N	47.24	0.47
97	P097G	G	46.89	0.47
97	P097T	T	45.56	0.46
97	P097R	R	45.09	0.45
97	P097M	M	33.57	0.34
97	P097C	C	29.31	0.29
97	P097V	V	22.34	0.22
97	P097Y	Y	21.56	0.22
97	P097F	F	20.98	0.21
97	P097L	L	19.17	0.19
97	P097I	I	18.76	0.19
97	P097W	W	15.10	0.15
98	A098R	R	132.47	1.32
98	A098K	K	131.64	1.32

98	A098D	D	102.97	1.03
98	A098L	L	102.06	1.02
98	A098T	T	91.56	0.92
98	A098G	G	89.92	0.90
98	A098V	V	70.86	0.71
98	A098M	M	48.09	0.48
98	A098F	F	40.13	0.40
98	A098E	E	36.13	0.36
99	I099V	V	77.11	0.77
99	I099F	F	13.83	0.14
99	I099L	L	13.75	0.14
99	I099T	T	10.36	0.10
99	I099P	P	7.86	0.08
99	I099Q	Q	7.27	0.07
99	I099E	E	7.21	0.07
99	I099K	K	6.92	0.07
99	I099A	A	6.85	0.07
99	I099H	H	6.53	0.07
99	I099D	D	6.44	0.06
99	I099S	S	6.42	0.06
99	I099G	G	6.26	0.06
99	I099W	W	6.22	0.06
99	I099R	R	5.85	0.06
99	I099Y	Y	5.59	0.06
100	S100A	A	103.00	1.03
100	S100V	V	97.27	0.97
100	S100R	R	91.59	0.92
100	S100E	E	87.00	0.87
100	S100T	T	81.00	0.81
100	S100C	C	75.00	0.75
100	S100L	L	61.00	0.61
100	S100H	H	59.00	0.59
100	S100I	I	53.00	0.53
100	S100W	W	43.04	0.43
100	S100G	G	35.00	0.35

100	S100Q	Q	21.00	0.21
100	S100F	F	13.00	0.13
100	S100M	M	7.00	0.07
101	Y101F	F	79.08	0.79
101	Y101H	H	34.40	0.34
101	Y101R	R	8.50	0.09
101	Y101T	T	7.52	0.08
101	Y101C	C	7.50	0.08
101	Y101Q	Q	7.46	0.07
101	Y101S	S	7.43	0.07
101	Y101L	L	7.42	0.07
101	Y101N	N	7.33	0.07
101	Y101G	G	7.03	0.07
101	Y101E	E	6.76	0.07
101	Y101P	P	6.39	0.06
102	I102L	L	32.85	0.33
102	I102T	T	12.31	0.12
102	I102A	A	11.19	0.11
102	I102F	F	7.57	0.08
102	I102P	P	7.53	0.08
102	I102G	G	6.21	0.06
102	I102K	K	5.78	0.06
102	I102S	S	5.77	0.06
102	I102R	R	5.60	0.06
102	I102Y	Y	5.40	0.05
102	I102H	H	5.15	0.05
102	I102E	E	5.10	0.05
102	I102Q	Q	4.91	0.05
103	E103T	T	113.38	1.13
103	E103P	P	111.36	1.11
103	E103S	S	86.19	0.86
103	E103G	G	81.52	0.82
103	E103D	D	56.43	0.56
103	E103H	H	33.35	0.33
103	E103Q	Q	10.36	0.10

103	E103I	I	8.00	0.08
103	E103A	A	7.62	0.08
103	E103K	K	7.55	0.08
103	E103L	L	7.25	0.07
103	E103W	W	7.22	0.07
103	E103C	C	7.10	0.07
103	E103F	F	6.98	0.07
103	E103V	V	6.88	0.07
103	E103R	R	6.04	0.06
104	E104P	P	124.98	1.25
104	E104Q	Q	98.83	0.99
104	E104L	L	95.33	0.95
104	E104Y	Y	91.66	0.92
104	E104H	H	87.15	0.87
104	E104S	S	81.74	0.82
104	E104A	A	76.42	0.76
104	E104F	F	71.58	0.72
104	E104T	T	69.53	0.70
104	E104C	C	64.06	0.64
104	E104V	V	57.67	0.58
104	E104N	N	48.45	0.48
104	E104W	W	43.19	0.43
104	E104G	G	33.84	0.34
105	D105M	M	25.47	0.25
105	D105T	T	16.88	0.17
105	D105E	E	14.77	0.15
105	D105P	P	7.10	0.07
105	D105C	C	6.97	0.07
105	D105G	G	6.51	0.07
105	D105S	S	6.19	0.06
105	D105F	F	5.83	0.06
105	D105L	L	5.71	0.06
105	D105W	W	5.66	0.06
105	D105Y	Y	5.36	0.05
106	A106F	F	108.54	1.09

<257>

106	A106K	K	106.67	1.07
106	A106V	V	106.21	1.06
106	A106L	L	105.12	1.05
106	A106M	M	102.59	1.03
106	A106C	C	95.21	0.95
106	A106S	S	94.82	0.95
106	A106I	I	88.65	0.89
106	A106E	E	84.86	0.85
106	A106N	N	80.65	0.81
106	A106G	G	66.53	0.67
106	A106D	D	44.78	0.45
107	E107Q	Q	112.74	1.13
107	E107P	P	102.05	1.02
107	E107V	V	99.68	1.00
107	E107S	S	99.06	0.99
107	E107T	T	93.40	0.93
107	E107H	H	89.38	0.89
107	E107R	R	87.40	0.87
107	E107L	L	82.19	0.82
107	E107G	G	65.90	0.66
107	E107C	C	34.35	0.34
107	E107D	D	6.32	0.06
107	E107Y	Y	5.97	0.06
107	E107I	I	5.76	0.06
108	V108I	I	96.87	0.97
108	V108A	A	81.56	0.82
108	V108L	L	73.95	0.74
108	V108C	C	73.59	0.74
108	V108T	T	54.76	0.55
108	V108G	G	34.72	0.35
108	V108E	E	13.15	0.13
108	V108S	S	6.94	0.07
108	V108H	H	6.42	0.06
108	V108P	P	6.03	0.06
108	V108Q	Q	6.03	0.06

<258>

108	V108R	R	5.94	0.06
109	T109Y	Y	151.13	1.51
109	T109E	E	140.77	1.41
109	T109G	G	135.65	1.36
109	T109N	N	135.23	1.35
109	T109S	S	125.47	1.25
109	T109M	M	123.65	1.24
109	T109R	R	121.74	1.22
109	T109Q	Q	105.15	1.05
109	T109W	W	102.35	1.02
109	T109K	K	98.87	0.99
109	T109L	L	95.89	0.96
109	T109A	A	31.57	0.32
110	T110V	V	117.34	1.17
110	T110L	L	113.17	1.13
110	T110P	P	112.26	1.12
110	T110M	M	110.23	1.10
110	T110C	C	98.41	0.98
110	T110K	K	97.01	0.97
110	T110E	E	80.71	0.81
110	T110G	G	80.05	0.80
110	T110R	R	79.50	0.80
110	T110N	N	74.55	0.75
110	T110Q	Q	57.23	0.57
110	T110S	S	5.27	0.05
110	T110W	W	4.82	0.05
111	M111N	N	145.51	1.46
111	M111T	T	138.47	1.38
111	M111Q	Q	134.85	1.35
111	M111L	L	134.45	1.34
111	M111W	W	132.12	1.32
111	M111E	E	130.67	1.31
111	M111C	C	113.25	1.13
111	M111G	G	101.14	1.01
111	M111I	I	100.83	1.01

111	M111V	V	93.69	0.94
111	M111R	R	74.46	0.74
111	M111K	K	70.17	0.70
111	M111P	P	31.70	0.32

<259>

<260>

<261> * "위치"는 SEQ ID NO:13의 전체 길이 V049 전구체 프로테아제에서의 숫자 매겨진 바와 같은 아미노산 위치를 나타낸다. 위치 1~27은 V049 전구체 효소의 신호 펩티드 부분 중 위치에서의 아미노산을 나타내고, 위치 28~111은 V049 전구체 효소의 pro 부위 중 위치에서의 아미노산을 나타낸다.

<262> #활성 비율은, 비개질된 전구체 프로테아제로부터 가공된 성숙 프로테아제의 활성으로 나눈 개질된 프로테아제로부터 가공된 성숙 프로테아제의 활성이다.

<263> 실시예 4

<264> 이 실시예에서, 상기 실시예에서 기재된 동일한 실험적 프로토콜을 사용하여, 바실러스 렌투스로부터의 서브틸

리신의 pro 부위 (GG36으로 불림) (SEQ ID NO:244)를 위치 33 및 위치 59에서 돌연변이화시켰다. 전구체 폴리뉴클레오티드 서열 (SEQ ID NO:240)을 돌연변이시키고, 구성물을 생성시키고, 상기 기재된 바와 같이 프로테아제 발현을 증가시키기 위한 개질을 시험하였다.

- <265> 결과는 아미노산 치환부 E33G, E33Q 및 E33V가 모두 야생형 성숙 GG36 프로테아제의 생성을 비개질된 전구체의 적어도 50% 초과로 증가시켰음을 나타내었다. 돌연변이 H32K는 야생형 성숙 프로테아제의 생성을 20%로 증가시켰다.
- <266> 이들 결과는 전체 길이 프로테아제의 pro 부위의 돌연변이가 변종 전구체 단백질 뿐 아니라 자연적으로 발생하는 효소의 발현을 개선시킬 수 있음을 나타낸다.

도면의 간단한 설명

- <19> 도 1A는 B. 클라우시로부터의 야생형 프로테아제의 전체 길이 폴리뉴클레오티드 서열을 나타낸다 (SEQ ID NO:1). 프로테아제의 pro 부위를 인코딩하는 서열 부분은 굵은 글씨로 나타낸다.
- <20> 도 1 B, C 및 D는 각각 SEQ ID NO:5의 전체 길이 B. 클라우시 프로테아제의 신호 펩티드 (SEQ ID NO:6), pro 부위 (SEQ ID NO:7) 및 성숙 형태 (SEQ ID NO:8)를 인코딩하는 폴리뉴클레오티드 서열 (SEQ ID NO:2, 3 및 4)을 나타낸다.
- <21> 도 2A는 B. 클라우시로부터의 야생형 프로테아제의 전체 길이 폴리펩티드 서열을 나타낸다 (SEQ ID NO:5). 프로테아제의 pro 부위를 인코딩하는 서열 부분은 굵은 글씨로 나타낸다.
- <22> 도 2 B, C 및 D는 각각 SEQ ID NO:5의 전체 길이 B. 클라우시 프로테아제의 신호 펩티드 (SEQ ID NO:6), pro 부위 (SEQ ID NO:7) 및 성숙 형태 (SEQ ID NO:8)를 인코딩하는 폴리펩티드 서열을 나타낸다.
- <23> 도 3A는 B. 클라우시로부터의 변종 프로테아제 V049의 전체 길이 폴리뉴클레오티드 서열을 나타낸다 (SEQ ID NO:9). 프로테아제의 pro 부위를 인코딩하는 서열 부분은 굵은 글씨로 나타낸다.
- <24> 도 3 B, C 및 D는 각각 SEQ ID NO:13의 전체 길이 변종 B. 클라우시 프로테아제의 신호 펩티드 (SEQ ID NO:14), pro 부위 (SEQ ID NO:15) 및 성숙 형태 (SEQ ID NO:16)를 인코딩하는 폴리뉴클레오티드 서열 (SEQ ID NO:10, 11 및 12)을 나타낸다.
- <25> 도 4A는 B. 클라우시로부터의 변종 프로테아제 V049의 전체 길이 폴리펩티드 서열을 나타낸다 (SEQ ID NO:13). 프로테아제의 pro 부위를 인코딩하는 서열 부분은 굵은 글씨로 나타낸다.
- <26> 도 4 B, C 및 D는 각각 SEQ ID NO:13의 전체 길이 변종 B. 클라우시 프로테아제의 신호 펩티드 (SEQ ID NO:14), pro 부위 (SEQ ID NO:15) 및 성숙 형태 (SEQ ID NO:16)를 인코딩하는 폴리펩티드 서열을 나타낸다.
- <27> 도 5는 SEQ ID NO:5의 B. 클라우시 야생형 세린 프로테아제, SEQ ID NO:13의 B. 클라우시 변종 세린 프로테아제 및 SEQ ID NO:244의 B. 렌투스 세린 프로테아제의 아미노산 서열 정렬을 나타낸다. 아미노산에서의 차이를 밑줄그었다.
- <28> 도 6은 SEQ ID NO:9의 변종 세린 프로테아제를 포함하는 pXX-049 플라스미드 벡터의 맵을 제공한다.
- <29> 도 7A-D는 본 발명에 사용된 pXX-V049 플라스미드 벡터의 폴리뉴클레오티드 서열을 나타낸다 (SEQ ID NO:17).
- <30> 도 8A는 B. 렌투스로부터의 프로테아제의 전체 길이 폴리뉴클레오티드 서열을 나타낸다 (SEQ ID NO:240). 프로테아제의 pro 부위를 인코딩하는 서열 부분은 굵은 글씨로 나타낸다.
- <31> 도 8B, C 및 D는 각각 SEQ ID NO:244의 B. 렌투스 전구체 프로테아제의 신호 펩티드 (SEQ ID NO:245), pro 부위 (SEQ ID NO:246) 및 성숙 형태 (SEQ ID NO:247)를 인코딩하는 폴리뉴클레오티드 서열 (SEQ ID NO:241, 242 및 243)을 나타낸다.
- <32> 도 9A는 B. 클라우시로부터의 변종 프로테아제 GG36의 전체 길이 폴리펩티드 서열을 나타낸다 (SEQ ID NO:244). 프로테아제의 pro 부위를 인코딩하는 서열 부분은 굵은 글씨로 나타낸다.
- <33> 도 9 B, C 및 D는 각각 SEQ ID NO:244의 전체 길이 B. 렌투스 프로테아제의 신호 펩티드 (SEQ ID NO:245), pro 부위 (SEQ ID N:246) 및 성숙 형태 (SEQ ID NO:247)를 인코딩하는 폴리펩티드 서열을 나타낸다.
- <34> **발명의 설명**

- <35> 본 발명은 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드, 및 미생물에서의 프로테아제의 생성을 변경시키는 방법에 관한 것이다. 특히, 본 발명은 바실러스 종과 같은 미생물에서의 프로테아제의 생성을 변경시키는 방법에 관한 것이다. 본 발명은 개질된 폴리뉴클레오티드, 벡터, 개질된 폴리펩티드, 및 프로테아제의 생성 증강 방법을 개시한다.
- <36> 본 발명은 돌연변이된 pro 부위를 갖는 프로테아제 뿐 아니라 개질된 폴리뉴클레오티드에 의해 인코딩된 개질된 프로테아제를 인코딩하는 개질된 폴리뉴클레오티드, 및 이의 생성 방법을 제공한다. 개질된 프로테아제 폴리뉴클레오티드는 미생물에서 개질된 프로테아제를 발현시키고, 상응하는 전구체 프로테아제 초과 수준에서 성숙 형태를 가공하기에 적절하다. 제조된 프로테아제는 세척, 동물 사육 및 식물 가공 산업을 포함하나 이에 제한되지 않는, 다양한 산업에서 사용하기에 적절한, 효소의 산업 생성에 사용된다. 본 발명은 또한 이러한 효소를 생성시키기 위한 수단을 제공한다. 일부 바람직한 구현예에서, 본 발명의 프로테아제는 순수 또는 상대적으로 순수한 형태의 것이다.
- <37> 다르게 표현하지 않는 한, 본 발명의 실행에는 분자 생물학, 미생물학, 단백질 정제, 단백질 공학, 단백질 및 DNA 서열화, 및 당업계의 지식 내에 있는 재조합 DNA 분야에 흔히 사용되는 통상적인 기술이 포함된다. 이러한 기술은 당업자에게 공지되어 있으며, 많은 문헌 및 참조물에 기재된다 (예를 들어 Sambrook 등, "Molecular Cloning: A Laboratory Manual", Second Edition (Cold Spring Harbor), [1989]); 및 Ausubel 등, "Current Protocols in Molecular Biology" [1987] 참조). 상기 및 하기 모두에서 본원에 언급된 모든 특허, 특허 출원, 논문 및 발행물은 이로써 명백하게 본원에 참고문헌으로 포함된다.
- <38> 본원에서 다르게 정의하지 않는 한, 본원에 사용된 모든 기술적 및 과학적 용어는 이러한 발명에 관계되는 당업계의 통상의 지식 중 하나에 의해 흔히 이해되는 바와 같은 동일한 의미를 갖는다. 예를 들어, [Singleton and Sainsbury, Dictionary of Microbiology and Molecular Biology, 2d Ed., John Wiley 및 Sons, NY (1994); 및 Hale and Markham, The Harper Collins Dictionary of Biology, Harper Perennial, NY (1991)]은 본 발명에 사용된 많은 용어의 일반적인 사전을 당업자에게 제공한다. 본원에 기재된 것과 유사 또는 동등한 임의 방법 및 물질이 본 발명의 실행에서 사용되지만, 바람직한 방법 및 물질이 본원에 기재된다. 따라서, 하기에 바로 정의된 용어는 그 전체가 본 명세서에 대한 참조로써 보다 충분히 기재된다. 또한 본원에서 사용된 바와 같이, 문맥에서 명백히 다르게 나타내지 않는 한 단수형 표현은 복수형 참조를 포함한다. 수 범위는 범위를 정의하는 수를 포함한다. 다르게 나타내지 않는 한, 각각 핵산은 5'에서 3' 방향에서, 왼쪽에서 오른쪽으로 적고; 아미노산 서열은 아미노에서 카복시 방향에서, 왼쪽에서 오른쪽으로 적는다. 당업자에 의해 사용되는 문맥에 따라 다양할 수 있기 때문에, 본 발명은 특정 방법론, 프로토콜 및 기재된 시약에 제한되지 않는 것으로 이해될 것이다.
- <39> 더욱이, 본원에 제공된 표제는 그 전체가 본 명세서에 대한 참조가 될 수 있는 본 발명의 다양한 측면 또는 구현예의 제한이 아니다. 따라서, 하기에 바로 정의된 용어는 그 전체가 본 명세서에 대한 참조로써 보다 충분히 정의된다. 그럼에도 불구하고, 본 발명의 이해를 촉진시키기 위해, 많은 용어가 하기에 정의된다.
- <40> **정의**
- <41> "개질된 프로테아제"는 전체 길이 "전구체 프로테아제"의 아미노산 서열에서 유래한 아미노산 서열을 갖는 전체 길이 프로테아제이다. 전구체 프로테아제는 또한 "비개질된 프로테아제"로서 나타낼 수 있다. 개질된 프로테아제는 pro 부위에서의 이의 전구체 프로테아제와 상이하다. 전구체 프로테아제는 자연적으로 발생하는 즉 야생형 프로테아제이거나, 또는 변종 프로테아제일 수 있다. 이는 개질된 프로테아제를 생성시키도록 개질된 변종 프로테아제 또는 야생형 프로테아제의 pro 부위이다. 개질된 프로테아제의 아미노산 서열은 전구체 아미노산 서열의 pro 부위 중 하나 이상의 아미노산의 치환, 결실 또는 삽입에 의해 전구체 프로테아제 아미노산 서열로부터 "유래"되는 것으로 여겨진다. 바람직한 구현예에서, 전구체 프로테아제의 pro 부위 중 하나 이상의 아미노산은 개질된 프로테아제를 생성시키도록 치환된다. 이러한 개질은 본질적으로 전구체 프로테아제의 조작이라기보다는 전구체 프로테아제의 아미노산 서열을 인코딩하는 "전구체 DNA 서열"의 개질이다. 본원의 개질된 프로테아제는 전구체 프로테아제의 pro 부위의 아미노산 잔기 중 임의 하나에서 자연적으로 발생하는 임의의 19개 아미노산의 치환부를 포함한다. 본 문맥에서, "개질된" 및 "전구체" 프로테아제는 모두 신호 펩티드, pro 부위 및 성숙 부위를 포함하는 전체 길이 프로테아제이다. 개질된 서열을 인코딩하는 폴리뉴클레오티드는 "개질된 폴리뉴클레오티드"로서 나타내고, 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드는 "전구체 폴리뉴클레오티드"로서 나타낸다. "전구체 폴리펩티드" 및 "전구체 폴리뉴클레오티드"는 각각 교환가능하게 "비개질된 전구체 폴리펩티드" 또는 "비개질된 전구체 폴리뉴클레오티드"로서 나타낼 수 있다.

- <42> "자연적으로 발생하는" 또는 "야생형"은 자연에서 발견되는 것과 동일한 비개질된 아미노산 서열을 갖는 프로테아제를 인코딩하는 프로테아제 또는 폴리뉴클레오티드를 나타낸다. 자연적으로 발생하는 효소에는 선천적 효소 (native enzyme)가 포함되며, 이들 효소는 특정 미생물에서 발견되거나 자연적으로 발현된다. 야생형 또는 자연적으로 발생하는 서열은 변종이 유래되는 서열을 나타낸다. 야생형 서열은 상동 또는 이종 단백질을 인코딩할 수 있다.
- <43> 본원에서 사용된 바와 같이 "변종"은, C- 및 N-말단 끝으로의 하나 이상의 아미노산 첨가, 아미노산 서열에서의 하나 이상의 상이한 위치에서 하나 이상의 아미노산의 치환, 단백질의 양 말단 또는 아미노산 서열에서의 하나 이상의 위치에서 하나 이상의 아미노산의 결실, 및/또는 아미노산 서열에서의 하나 이상의 위치에서 하나 이상의 아미노산의 삽입에 의한, 이의 상응하는 야생형 단백질과 상이한 전구체 단백질을 나타낸다. 본 발명의 문맥에서의 변종 단백질은, 자연적으로 발생하는 단백질 Maxacal (SEQ ID NO:5)의 변종인 B. 클라우시 프로테아제 V049 (SEQ ID NO:13)로써 예시된다. 변종의 전구체 단백질은 야생형 또는 변종 단백질일 수 있다.
- <44> 본원에서 사용된 바와 같이, "동등한"은 단백질 또는 펩티드에서의 열거된 위치에서 잔기, 또는 단백질 또는 펩티드에서의 열거된 잔기와 유사, 상동 또는 상응하는 잔기를 나타낸다.
- <45> 프로테아제과 관련된 용어 "생성"은 하기를 포함하는 전체 길이 프로테아제의 두 가공 단계를 포함한다: 1. 단백질 분비 동안 발생하는 것으로 공지된 신호 펩티드의 제거; 및 2. 효소의 활성 성숙 형태를 만들어내며 성숙 과정 동안 발생하는 것으로 공지된 pro 부위의 제거 (Wang 등, Biochemistry 37:3165-3171 (1998); Power 등, Proc Natl Acad Sci USA 83:3096-3100 (1986)).
- <46> 성숙 프로테아제와 관련된 용어 "가공된"은 전체 길이 단백질 예를 들어 프로테아제가 활성 성숙 효소가 되도록 거치는 성숙 과정을 나타낸다.
- <47> 용어 "활성 비율" 및 "생성 비율"은, 개질된 프로테아제로부터 가공된 성숙 프로테아제의 효소적 활성 대 비개질된 프로테아제로부터 가공된 성숙 프로테아제의 효소적 활성의 비율을 나타내도록 교환가능하게 사용된다.
- <48> 본원의 용어 "전체 길이 단백질"은 유전자의 1차 유전자 생성물을 나타내며, 신호 펩티드, pro 서열 및 성숙 서열을 포함한다.
- <49> 용어 "신호 서열" 또는 "신호 펩티드"는 단백질의 성숙 또는 전구체 형태의 분비에 참여할 수 있는 뉴클레오티드 및/또는 아미노산의 임의 서열을 나타낸다. 신호 서열의 이러한 정의는 기능적인 것이며, 단백질의 분비 실시예에 참여하는 단백질 유전자의 N-말단 부분에 의해 인코딩된 모든 아미노산 서열을 포함하는 것으로 의도된다.
- <50> 용어 "pro 서열" 또는 "pro 부위"는 프로테아제의 분비/생성에 필요한 신호 서열 및 성숙 프로테아제 사이의 아미노산 서열이다. pro 서열의 절단은 성숙 활성 프로테아제를 야기할 것이다. 예시하기 위해, 본 발명의 프로테아제의 pro 부위는 적어도 SEQ ID NO:5, 13 또는 244의 잔기 28~111과 동일한 아미노산 서열을 포함한다.
- <51> 용어 "성숙 형태" 또는 "성숙 부위"는 단백질의 최종 기능적 부분을 나타낸다. 예시하기 위해, 본 발명의 프로테아제의 성숙 형태는 적어도 SEQ ID NO:5, 13 또는 244의 잔기 112~380과 동일한 아미노산 서열을 포함한다. 본 문맥에서, "성숙 형태"는 전체 길이 프로테아제로부터 "가공"되며, 여기서 전체 길이 프로테아제의 가공은 신호 펩티드의 제거 및 pro 부위의 제거를 포함한다.
- <52> 본원에서 사용된 바와 같이, 용어 "이종 단백질"은 숙주 세포에서 자연적으로 발생하지 않는 단백질 또는 폴리펩티드를 나타낸다. 유사하게, "이종 폴리뉴클레오티드"는 숙주 세포에서 자연적으로 발생하지 않는 폴리뉴클레오티드를 나타낸다.
- <53> 본원에서 사용된 바와 같이, "상동 단백질"은 세포에서 선천적 또는 자연적으로 발생하는 단백질 또는 폴리펩티드를 나타낸다. 유사하게, "상동 폴리뉴클레오티드"는 세포에서 선천적 또는 자연적으로 발생하는 폴리뉴클레오티드를 나타낸다.
- <54> 본원에서 사용된 바와 같이, 용어 "촉진제"는 하류 유전자의 직접 전사에 대해 기능하는 핵산 서열을 나타낸다. 바람직한 구현예에서, 촉진제는 목표 유전자가 발현되는 숙주 세포에 적합하다. 촉진제와 함께, 다른 전사적 및 번역적 조절 핵산 서열 (또는 "제어 서열")이 주어진 유전자를 발현시키기 위해 필요하다. 일반적으로, 전사적 및 번역적 조절 서열에는 촉진제 서열, 리보솜 결합 위치, 전사 시작 및 중단 서열, 번역 시작 및

중단 서열, 및 증강제 또는 활성화제 서열이 포함되나 이제 제한되지는 않는다.

- <55> 핵산은, 또 다른 핵산 서열과 기능적 관계에 놓여지는 경우 "작동적으로 연결된다". 예를 들어, 서열의 전사에 영향을 미친다면 촉진제 또는 증강제는 코딩 서열에 작동적으로 연결되고; 또는 번역을 촉진하도록 위치된다면 리보솜 결합 위치는 코딩 서열에 작동적으로 연결된다. 일반적으로, "작동적으로 연결된"은 연결된 DNA 서열이 인접하다는 것을 의미한다.
- <56> 본원에서 사용된 바와 같이, 용어 "프로테아제," 및 "단백질분해적 활성화"은 단백질 또는 펩티드가 펩티드 연결을 갖는 펩티드 또는 기질을 가수분해하기 위한 능력을 나타냄을 나타낸다. 단백질분해적 활성을 측정하기 위한 많은 잘 공지된 방법이 존재한다 (Kalisz, "Microbial Proteinases," In: Fiechter (ed.), *Advances in Biochemical Engineering/Biotechnology*, [1988]). 예를 들어, 단백질분해적 활성화는 시판되는 기질을 가수분해시키기 위한 생성된 프로테아제의 능력을 분석하는 비교 분석에 의해 확인될 수 있다. 프로테아제 또는 단백질분해적 활성화의 이러한 분석에서 유용한 예시적 기질에는 디-메틸 카세인 (Sigma C-9801), 소과 콜라겐 (Sigma C-9879), 소과 엘라스틴 (Sigma E-1625) 및 소과 케라틴 (ICN Biomedical 902111)이 포함되나 이에 제한되지는 않는다. 이러한 기질을 사용하는 측색 분석은 당업계에 잘 공지되어 있다 (예를 들어, 모두 본원에 참고문헌으로 포함되는 WO 99/34011; 및 미국 특허 제 6,376,450호 참조). AAPF 분석 (예를 들어 Del Mar 등, *Anal. Biochem.*, 99:316-320 [1979] 참조)이 또한 성숙 프로테아제의 생성을 측정하는데 사용된다. 상기 분석은 가용성 합성 기질, 숙시닐-알라닌-알라닌-프롤린-페닐알라닌-p-니트로아닐라이드 (sAAPF-pNA)를 가수분해하는 효소로서 p-니트로아닐린이 방출되는 속도를 측정한다. 가수분해 반응으로부터 황색이 생성되는 속도는 분광 광도계에서 410 nm에서 측정되며, 활성화 효소 농도에 대해 비례적이다.
- <57> 본원에서 사용된 바와 같이, "바실러스 속"에는 당업자에게 공지된 바와 같은 "바실러스" 속 내의 모든 종이 포함되는데, B. 서브틸리스, B. 리체니포르미스, B. 렌투스, B. 브레비스, B. 스테아로써모필루스, B. 알칼로필루스, B. 아밀로리쿠에파시엔스, B. 클라우시, B. 할로두란스, B. 메가테리움 (B. megaterium), B. 코아굴란스, B. 썬쿨란스, B. 라우투스 (B. lautus) 및 B. 투링기엔시스가 포함되나 이에 제한되지는 않는다. 바실러스 속이 분류학적 재현성을 거치도록 지속된다는 것이 인지된다. 따라서, 속이 재분류된 종을 포함하는 것으로 의도되며, 현재 "지오바실러스 스테아로써모필루스 (Geobacillus stearothermophilus)"로 불리는 B. 스테아로써모필루스와 같은 생물이 포함되나 이에 제한되지는 않는다. 산소의 존재 하의 저항성 내생포자의 생성은 바실러스 속을 정의하는 특성으로 여겨지지만, 상기 특징은 또한 최근 알리시클로바실러스 (Alicyclobacillus), 암피바실러스 (Amphibacillus), 아뉴리니바실러스 (Aneurinibacillus), 아녹시바실러스 (Anoxybacillus), 브레비바실러스 (Brevibacillus), 필로바실러스 (Filobacillus), 그라실리바실러스 (Gracilibacillus), 할로바실러스 (Halobacillus), 파에니바실러스 (Paenibacillus), 살리바실러스 (Salibacillus), 썬모바실러스 (Thermobacillus), 우레이바실러스 (Ureibacillus) 및 버기바실러스 (Virgibacillus)로 불리는 것에도 적용된다.
- <58> 본원에서 교환가능하게 사용된 용어 "폴리뉴클레오티드" 및 "핵산"은, 임의 길이의 뉴클레오티드의 중합체성 형태를 나타낸다. 이들 용어에는 단일-, 이중-나선 DNA, 게놈 DNA, cDNA, 또는 푸린 및 피리미딘 염기를 포함하는 중합체, 또는 다른 자연적, 화학적, 생화학적으로 개질된, 비자연적 또는 유도화된 뉴클레오티드 염기가 포함되나 이제 제한되지는 않는다. 폴리뉴클레오티드의 비제한적인 예에는 유전자, 유전자 절편, 염색체 절편, EST, 엑손, 인트론, mRNA, tRNA, rRNA, 리보자임, cDNA, 재조합 폴리뉴클레오티드, 분지된 폴리뉴클레오티드, 플라스미드, 벡터, 임의 서열의 단리된 DNA, 임의 서열의 단리된 RNA, 핵산 프로브 및 프라이머가 포함된다.
- <59> 본원에서 사용된 바와 같이, 용어 "DNA 구성물" 및 "DNA 형질전환"은 숙주 세포 또는 생물에 서열을 도입하는데 사용되는 DNA를 나타내도록 교환가능하게 사용된다. DNA는 PCR 또는 당업자에게 공지된 임의 다른 적절한 기술(들)에 의해 체외 생성될 수 있다. 특히 바람직한 구현예에서, DNA 구성물은 관심있는 서열 (예를 들어 개질된 서열)을 포함한다. 일부 구현예에서, 서열은 부가적인 요소 예컨대 제어 요소 (예를 들어 촉진제 등)에 작동적으로 연결된다. 일부 구현예에서, DNA 구성물은 숙주 세포 염색체와 상동인 서열을 포함한다. 다른 구현예에서, DNA 구성물은 비상동 서열을 포함한다. DNA 구성물이 체외에서 어셈블링되고나면, 이는 숙주 세포 염색체의 부위를 돌연변이화시키는데 (즉, 내생 서열을 이중 서열로 대체) 사용될 수 있다.
- <60> 본원에서 사용된 바와 같이, 용어 "벡터"는 하나 이상의 세포 유형에 핵산을 도입시키도록 설계된 폴리뉴클레오티드 구성물을 나타낸다. 벡터에는 클로닝 벡터, 발현 벡터, 서플 벡터 및 플라스미드가 포함된다. 일부 구현예에서, 폴리뉴클레오티드 구성물은 전체 길이 프로테아제를 인코딩하는 DNA 서열을 포함한다 (예를 들

어, 개질된 프로테아제 또는 비개질된 전구체 프로테아제).

- <61> 본원에서 사용된 바와 같이, 용어 "플라스미드"는 클로닝 벡터로서 사용된 환형 이중-나선 (ds) DNA 구성물을 나타내며, 이는 일부 진핵생물 또는 원핵생물에서 염색체의 자가복제 유전적 요소를 형성하거나, 숙주 염색체로 통합된다.
- <62> 세포에 핵산 서열을 도입하는 문맥 중 본원에서 사용된 바와 같이, 용어 "도입된"은 세포에 핵산 서열을 이동시키는 데 적절한 임의 방법을 나타낸다. 이러한 도입 방법에는 원형질체 융합, 트랜스펙션, 형질전환, 건주계 이식 및 형질도입을 포함하나 이에 제한되지는 않는다 (예를 들어 Ferrari 등, "Genetics," in Hardwood 등, (eds.), Bacillus, Plenum Publishing Corp., pages 57-72, [1989] 참조).
- <63> 본원에서 사용된 바와 같이, 용어 "형질전환된" 및 "안정적으로 형질전환된"은 이의 계놈에 통합된 비-선천적 (이중) 폴리뉴클레오티드 서열을 갖거나 또는 두 세대 이상 동안 유지되는 에피솜 플라스미드를 갖는 세포를 나타낸다.
- <64> 본 발명은 아미노산 서열을 인코딩하고 개질된 프로테아제를 인코딩하는 단리된 개질 폴리뉴클레오티드를 제공한다. 개질된 프로테아제는 전구체 프로테아제의 폴리뉴클레오티드 서열을 돌연변이시킴으로써 수득된다. 특별하게는, 전구체 프로테아제의 pro 부위를 인코딩하는 폴리뉴클레오티드 서열 중 하나 이상의 돌연변이가 만들어져 본 발명의 개질된 폴리뉴클레오티드가 제공된다.

도면

도면1A

B. 클라우시 Maxacal 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드

SEQ ID NO: 1
 ATGAAGAAACCGTTGGGGAAAATTGTCGCAAGCACCCGCACTACTCATTCTGTTGCTTTT
 AGTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAA
 TGAGCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGC
 CATTCTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTC
 CTGTTTTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACCTCGATCCAGC
 GATTTCTTATATTGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCAGTGCCATGG
 GGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTGGATTGACAGGTTCTGGTGT
 AAAAGTTGCTGTCCTCGATACAGGTATTTCCAATCATCCAGACTTAAATATTCGTGGTGG
 CGCTAGCTTTGTACCAGGGGAACCATCCAATCAAGATGGGAATGGGCATGGCACGCAT
 GTGGCTGGGACGATTGCTGCTTTAAACAATTCGATTGGCGTTCCTGGCGTAGCACCGAA
 CGCGGAACATACGCTGTAAAGTATTAGGGGCGAGCGGTTGAGTTCCGGTCAGCTCGA
 TTGCCAAGGATTGGAATGGGCAGGGAACAATGGCATGCACGTTGCTAATTTGAGTTTA
 GGAAGCCCTTCGCCAAGTGCCACACTTGAGCAAGCTGTTAATAGCGGACTTCTAGAGG
 CGTTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCC
 GTTATGCGAACGCAATGGCAGTCCGAGCTACTGACCAAAACAACAACCGCGCCAGCTTT
 TCACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTGTAACGTGCAGAGCACATA
 CCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTTGCAG
 GTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCAATC
 ATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGTATGGAAGCGGACTTGT
 AATGCAGAAGCGGCAACACGCTAA

도면1B

SEQ ID NO: 2
ATGAAGAAACCGTTGGGGAAAATTGTCGCAAGCACCGCACTACTCATTCTGTTGCTTTT
AGTTCATCGATCGCATCGGCT

도면1C

SEQ ID NO: 3
GCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTCAGTGAG
TTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAGGAAGT
CGAAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTATCCGTTGAGTTAAGCCCA
GAAGATGTGGACGCGCTTGAACCTCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGA
AGTAACGACAATG

도면1D

SEQ ID NO: 4
GCGCAATCAGTGCCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTG
GATTGACAGGTTCTGGTGAAAAGTTGCTGTCCTCGATACAGGTATTTCCACTCATCCAG
ACTTAAATATTCGTGGTGGCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGG
AATGGGCATGGCACGCATGTGGCTGGGACGATTGCTGCTTTAAACAATTCGATTGGCGT
TCTTGGCGTAGCACCGAACGCGGAECTATACGCTGTTAAAGTATTAGGGGCGAGCGGTT
CAGGTTCCGTCAGCTCGATTGCCCAAGGATTGGAATGGGCAGGGAACAATGGCATGCA
CGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCCACACTTGAGCAAGCTGTTAA
TAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCAGGCT
CAATCAGCTATCCGGCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAAC
AACAAACCGCGCCAGCTTTTCACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTGT
AAACGTGCAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGG
CTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCA
ATGTACAAATCCGCAATCATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGT
ATGGAAGCGGACTTGTCAATGCAGAAGCGGCAACACGCTAA

도면2A

B. 클라우시 Maxaca1 전구체 프로테아제의 아미노산 서열

SEQ ID NO: 5
MKKPLGKIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEE
EEVEIELLHEFETIPVLSVELSPEDVDALELDP AISYIEEDA EVTTMRVQAPAAHNRGLTGSG
VKVAVLDTGISTHPDLNIRGGASFVPGEPTQDGNHGHGTHVAGTIAALNNSIGVLGVAPNAEL
YAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVA
ASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYGAGLDIVPGVNVQSTYPGSTYAS
LNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSLVNAEAATR

도면2B

SEQ ID NO: 6

MKKPLGKIVASTALLISVAFSSSIASA

도면2C

SEQ ID NO: 7

AEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEVEIEILLHEFETIPVLSVELSPEDVD
ALELDPAISYIEEDA EVTTM

도면2D

SEQ ID NO: 8

RVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNHGHGTHVAGTI
AALNNSIGVLGVAPNAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPSPSA
TLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFQYAGLDI
VAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTAT
SLGSTNLYGSGLVNAEAATR

도면3A

B. 클라우시 049 진구체 프로테아제를 인코딩하는 폴리뉴클레오티드

SEQ ID NO: 9

GTGAGAAGCAAAAAATTGTGGATCGTCGCGTCCGACCGCACTACTCATTCTGTTGCTTTT
AGTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAA
TGAGCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCCG
CATTCTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTC
CTGTTTTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACGATCCAGC
GATTTCTTATATTGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCGGTACCATGG
GGAATTAGCCGTGTGCAAGCCCAGCTGCCATAACCGTGGATTGACAGGTTCTGGTGT
AAAAGTTGCTGTCCCTCGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTGG
CGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCAT
GTGGCTGGGACGATTGCTGCTTTAAACAATTCGATTGGCGTTCTTGGCGTAGCACCGAA
CGCGGAACTATACGCTGTAAAGTATTAGGGGCGAGCGTTTCAAGTTCTGGTCAAGTCCGA
TTGCCCAAGGATTGGAATGGGCAGGGAACAATGTTATGCACGTTGCTAATTTGAGTTTAG
GACTGCAGGCACCAAGTGCCACACTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGC
GTTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCCCG
TTATGCCAACGCAATGGCAGTCCGAGCTACTGACCAAAACAACAACCGCGCCAGCTTTT
CACAGTATGGCGCAGGGCTTGACATTGTCCGACCAGGTGTAACGTGCAGAGCACATAC
CCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTTGCAGG
TGCAGCAGCCCTTGTAAACAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCAATCA
TCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGTATGGAAGCGGACTTGTCA
ATGCAGAAGCGGCAACACGTTAA

도면3B

SEQ ID NO: 10

GTGAGAAGCAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTT
AGTTCATCGATCGCATCGGCT

도면3C

SEQ ID NO: 11

GCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTCAGTGAG
TTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAGGAAGT
CGAAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTATCCGTTGAGTTAAGCCCA
GAAGATGTGGACGCGCTTGAACCTCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGA
AGTAACGACAATG

도면3D

SEQ ID NO: 12

GCGCAATCGGTACCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTG
GATTGACAGGTTCTGGTGAAAAGTTGCTGTCCCTCGATACAGGTATTTCCACTCATCCAG
ACTTAAATATTCGTGGTGGCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGG
AATGGGCATGGCACGCATGTGGCTGGGACGATTGCTGCTTAAACAATTGATTGGCGT
TCTTGGCGTAGCACCGAACGCGGAECTATACGCTGTTAAAGTATTAGGGGCGAGCGGTT
CAGGTTCCGTCAGCTCGATTGCCAAGGATTGGAATGGGCAGGGAACAATGTTATGCAC
GTTGCTAATTTGAGTTTAGGACTGCAGGCACCAAGTGCCACACTTGAGCAAGCTGTTAAT
AGCGCGACTTCTAGAGGCGTCTTGTGTAGCGGCATCTGGGAATTCAGGTGCAGGCTC
AATCAGCTATCCGGCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAAACA
ACAACCGCGCCAGCTTTTACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTGTA
AACGTGCAGAGCACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGC
TACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCAA
TGTAACAAATCCGCAATCATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACCTGTA
TGGAAGCGGACTTGTCAATGCAGAAGCGGCAACACGTTAA

도면4A

B. 클라우시 049 전구체 프로테아제의 아미노산 서열

SEQ ID NO:13

VRSKLWIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEE
EEVEIELLHEFETIPVLSVELSPEDVDALELDPAISYIEEDA EVTTMAQSVPWGISRVQAPAAH
NRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNHGHGTHVAGTIAALNNSIGV
LGVAPNAELYAVKVLGASGSGSVSSIAQGLEWAGNNVMHVNLSLGLQAPSATLEQAVNSA
TSRGLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYGAGLDIVPGVNVQS
TYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLV
NAEAATR

도면4B

SEQ ID NO:14
VRSKKLWIVASTALLISVAFSSSIASA

도면4C

SEQ ID NO:15
AEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEVEIELLHEFETIPVLSVELSPEDVD
ALELDPAISYIEEDA EVTTM

도면4D

SEQ ID NO:16
AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNG
HGTHVAGTIAALNNSIGVLGVAPNAELYAVKVLGASGSGSVSSIAQGLEWAGNNVMHVNLS
LGLQAPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFS
QYGAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLK
NTATSLGSTNLYGSLVNAEAATR

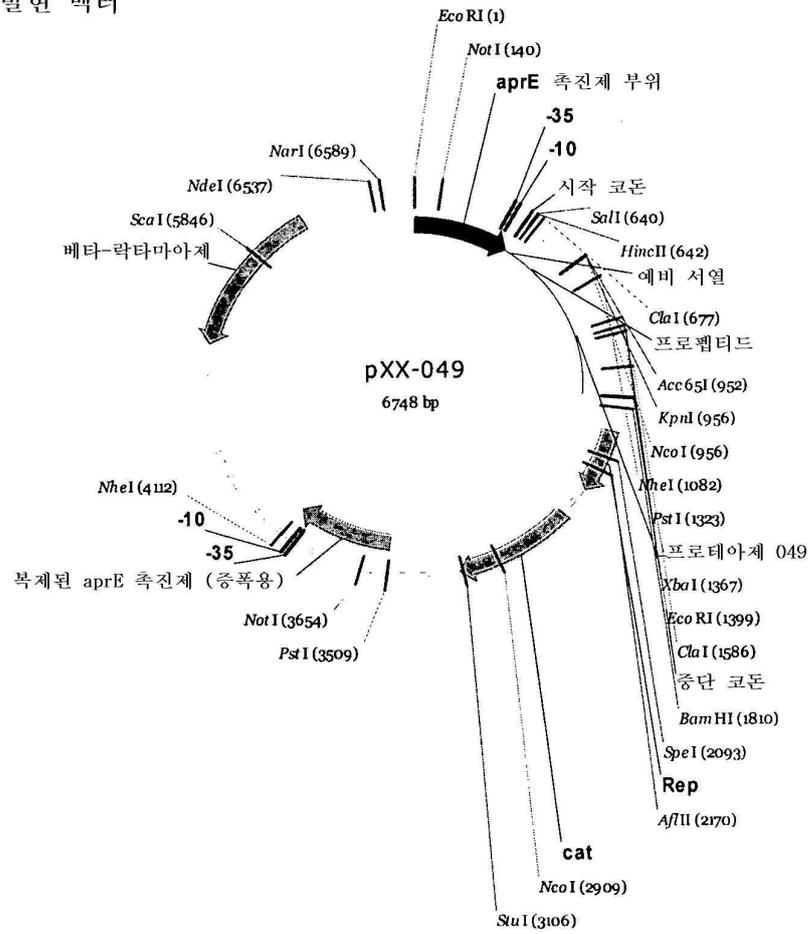
도면5

SEQ ID NO:5 (Maxacal), SEQ ID NO:13 (V049) 및 SEQ ID NO:244 (GG36)의
아미노산 서열의 정렬

		1		50
V049	(1)	MRSK	KLWIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVE	
GG36	(1)	MRSK	KLWIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVE	
Maxacal	(1)	<u>MKKPLGKIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVE</u>		
		51		100
V049	(51)	QVEANDEVAILSEEEEVEIELLHEFETIPVLSVELSPEDVDALELDPAIS		
GG36	(51)	QVEANDEVAILSEEEEVEIELLHEFETIPVLSVELSPEDVDALELDPAIS		
Maxacal	(51)	QVEANDEVAILSEEEEVEIELLHEFETIPVLSVELSPEDVDALELDPAIS		
		101		150
V049	(101)	YIEEDA	EVTTTMAQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHP	
GG36	(101)	YIEEDA	EVTTTMAQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHP	
Maxacal	(101)	YIEEDA	EVTTTMAQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHP	
		151		200
V049	(151)	DLNIRGGAS	FVPGEPSTQDGNHGHGTHVAGTIAALNNSIGVLGVAPNAELY	
GG36	(151)	DLNIRGGAS	FVPGEPSTQDGNHGHGTHVAGTIAALNNSIGVLGVAPSAELY	
Maxacal	(151)	DLNIRGGAS	FVPGEPSTQDGNHGHGTHVAGTIAALNNSIGVLGVAPNAELY	
		201		250
V049	(201)	AVKVLGAS	SGSGSVSSIAQGLEWAGNNMHVANLSLGLQAPSATLEQAVNS	
GG36	(201)	AVKVLGAS	SGSGSVSSIAQGLEWAGNNMHVANLSLGSPPSATLEQAVNS	
Maxacal	(201)	AVKVLGAS	SGSGSVSSIAQGLEWAGNNMHVANLSLGSPPSATLEQAVNS	
		251		300
V049	(251)	ATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYGAG		
GG36	(251)	ATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYGAG		
Maxacal	(251)	ATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYGAG		
		301		350
V049	(301)	LDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQ		
GG36	(301)	LDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQ		
Maxacal	(301)	LDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQ		
		351		381
V049	(351)	IRNHLKNTATSLGSTNLYGSLVNAAEATR-	(SEQ ID NO:13)	
GG36	(351)	IRNHLKNTATSLGSTNLYGSLVNAAEATR-	(SEQ ID NO:244)	
Maxacal	(351)	IRNHLKNTATSLGSTNLYGSLVNAAEATR-	(SEQ ID NO:5)	

도면6

pXX-049 발현 벡터



도면7A

pXX-049 플라스미드의 폴리뉴클레오티드 서열

SEQ ID NO:17

AATTCCTCCATTTTCTTCTGCTATCAAATAACAGACTCGTGATTTTCCAAACGAGCTTTC
 AAAAAAGCCTCTGCCCTTGCAAATCGGATGCCTGTCTATAAAATCCCGATATTGGCTT
 AACAGCGGCGCAATGGCGGCCGCATCTGATGTCTTTGCTTGGCGAATGTTTCATCTTAT
 TTCTTCTCCCTCTCAATAATTTTTTCATTCTATCCCTTTTCTGTAAAGTTATTTTTTCAGAA
 TACTTTTATCATCATGCTTTGAAAAATATCACGATAATATCCATTGTTCTCACGGAAGCA
 CACGCAGGTCATTTGAACGAATTTTTTCGACAGGAATTTGCCGGGACTCAGGAGCATTTA
 ACCTAAAAAGCATGACATTTTCAGCATAATGAACATTTACTCATGTCTATTTTCGTTCTTTT
 CTGTATGAAAATAGTTATTTTCGAGTCTCTACGGAAATAGCGAGAGATGATATACCTAAATA
 GAGATAAAATCATCTCAAAAAATGGGTCTACTAAAATATTATTCCATCTATTACAATAAAT
 TCACAGAATAGTCTTTTAAGTAAGTCTACTCTGAATTTTTTTAAAAGGAGAGGGTAAAGAG
 TGAGAAGCAAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTTA
 GTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATTTAATTGGCTTTAATG
 AGCAGGAAGCTGTCAAGTGGTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATT
 CTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTCTCTGTT
 TTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACCTCGATCCAGCGATTTCT
 TATATTGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCGGTACCATGGGGAATTAG
 CCGTGTGCAAGCCCCAGCTGCCATAACCGTGGATTGACAGGTTCTGGTGAAAAGTTG
 CTGTCCTCGATACAGGATTTCCACTCATCCAGACTTAAATATTCGTGGTGGCGCTAGCT
 TTGTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCTGG
 GACGATTGCTGCTTTAAACAATTTCGATTGGCGTTCTTGGCGTAGCACCGAACCGGGAAC
 TATACGCTGTTAAAGTATTAGGGGCGAGCGTTTCAGGTTCCGGTCAGCTCGATTGCCCAA
 GGATTGGAATGGGCAGGGAACAATGTTATGCACGTTGCTAATTTGAGTTTAGGACTGCA
 GGCACCAAGTGCCCACTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTG
 TTGTAGCGGCATCTGGGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCCCGTTATGCG
 AACGCAATGGCAGTCGGAGCTACTGACCAAAACAACAACCGCGCCAGCTTTTTCACAGTA
 TGGCGCAGGGCTTGACATTGTCGCACCAGGTGTAACGTGCAGAGCACATAACCCAGGTT
 CAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTTGCAGGTGCAGCA
 GCCCTTGTTAAACAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCAATCATCTAAAG
 AATACGGCAACGAGCTTAGGAAGCACGAACCTTGATGGAAGCGGACTTGTCAATGCAGA
 AGCGGCAACACGTTAATCAATAAAAAACGCTGTGCGGTTAAAGGGCACAGCGTTTTTTT
 GTGTATGAATCGGGATCCTCGATCGAGACTAGAGTCGATTTTTACAAGAATTAGCTTTAT
 ATAATTTCTGTTTTTCTAAAGTTTTATCAGCTACAAAAGACAGAAATGTATTGCAATCTTCA
 ACTAAATCCATTTGATTCTCTCCAATATGACGTTTAAATTTCTGAAATACTTGATTCT
 TTGTTTTTCTCAGTATACTTTTCCATGTTATAACACATAAAAACAACTTAGTTTTACAAA
 CTATGACAATAAAAAAAGTTGCTTTTTCCCTTTTCTATGTATGTTTTTACTAGTCATTTAA

도면7B

AACGATACATTAATAGGTACGAAAAAGCAACTTTTTTTGCGCTTAAAACCAGTCATACCAA
 TAACCTAAGGGTAACTAGCCTCGCCGGCAATAGTTACCCTTATTATCAAGATAAGAAAGA
 AAAGGATTTTTCGCTACGCTCAAATCCTTTAAAAAACACAAAAGACCACATTTTTTAATG
 TGGTCTTTATTCTTCAACTAAAGCACCCATTAGTTCAACAAACGAAAATTGGATAAAGTGG
 GATATTTTTAAAAATATATTTTATGTTACAGTAATATTGACTTTTTAAAAAAGGATTGATTCTA
 ATGAAGAAAGCAGACAAGTAAGCCTCCTAAATTCACCTTAGATAAAAAATTTAGGAGGCAT
 ATCAAATGAACTTTAATAAAAATTGATTTAGACAATTGGAAGAGAAAAGAGATATTTAATCAT
 TATTTGAACCAACAAACGACTTTTAGTATAACCACAGAAAATTGATATTAGTGTTTTATACC
 GAAACATAAAAAAAGAAGGATATAAATTTTACCCTGCATTTATTTTCTTAGTGACAAGGGT
 GATAAACTCAAATACAGCTTTTGAAGCTGTTACAATAGCGACGGAGAGTTAGGTTATTG
 GGATAAGTTAGAGCCACTTTATACAATTTTTGATGGTGTATCTAAAACATTCTCTGGTATT
 TGGACTCCTGTAAGAATGACTTCAAAGAGTTTTATGATTTATACCTTTCTGATGTAGAGA
 AATAAATGGTTCGGGGAAAATTGTTCCCAAAACACCTATACCTGAAAATGCTTTTTCTCT
 TTCTATTATCCATGGACTTCATTTACTGGGTTAACTTAAATATCAATAATAATAGTAATT
 ACCTTCTACCCATTATTACAGCAGGAAAATTCATTAATAAAGGTAATTCAATATATTTACC
 GCTATCTTTACAGGTACATCATTCTGTTTGTGATGGTTATCATGCAGGATTGTTTATGAAC
 TCTATTACAGGAATTGTCAGATAGGCCTAATGACTGGCTTTTATAATATGAGATAATGCCGA
 CTGTACTTTTTACAGTCGGTTTTCTAATGTCCTAACCTGCCCGTTAGTTGAAGAAGGTT
 TTTATATTACAGCTCCAGATCCATATCCTTCTTTTTCTGAACCGACTTCTCCTTTTTCGCTT
 CTTTATTCCAATTGCTTTATTGACGTTGAGCCTCGGAACCCTTAAACAATCCCAAAACTGT
 CGAATGGTCGGCTTAATAGCTCACGCTATGCCGACATTCGTCTGCAAGTTTAGTTAAGG
 GTTCTTCTCAACGCACAATAAATTTTCTCGGCATAAATGCGTGGTCTAATTTTTATTTTTAA
 TAACCTTGATAGCAAAAAATGCCATTCCAATACAAAACCACATACCTATAATCGACCTGCA
 GGAATTAATTCCTCCATTTTCTTCTGCTATCAAAATAACAGACTCGTGATTTTCCAAACGA
 GCTTTCAAAAAGCCTCTGCCCTTGCAAATCGGATGCCTGTCTATAAAATTCCCGATAT
 TGGCTTAAACAGCGGCGCAATGGCGGCCGCATCTGATGTCTTTGCTTGGCGAATGTTCA
 TCTTATTTCTTCCCTCCTCAATAATTTTTTCACTTATCCCTTTCTGTAAAGTTTATTTT
 TCAGAATACTTTTATCATCATGCTTTGAAAAAATATCACGATAATATCCATTGTTCTCACG
 GAAGCACACGCAGGTCATTTGAACGAATTTTTTCGACAGGAATTTGCCGGGACTCAGGA
 GCATTTAACCTAAAAAAGCATGACATTTACAGCATAATGAACATTTACTCATGTCTATTTTC
 GTTCTTTTCTGTATGAAAATAGTTATTTTCGAGTCTCTACGGAAATAGCGAGAGATGATATA
 CCTAAATAGAGATAAAATCATCTCAAAAAAATGGGTCTACTAAAATATTATCCATCTATTA
 CAATAAATTCACAGAATAGTCTTTAAGTAAGTCTACTCTGAATTTTTTTATCAAGCTAGCT
 TGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCAC
 ACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCTAATGAGTGAGCTAA
 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCTGCCA
 GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCATTTGGGCGCTCT
 TCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCCGGCTGCGGGCAGCGGTAT

도면7C

CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG
AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGG
CGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC
GTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCT
TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA
GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATC
TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACG
TTAAGGGATTTTGGTCATGAGATTACAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTGCCT
GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCC
AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCTCCATCCAGTCT
ATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTT
GTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTAG
CTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGG
TTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAAGTTGGCCGAGTGTATCACTCA
TGGTTATGGCAGCACTGCATAATTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGT
GACTGGTGTAGTACTCAACCAAGTCACTTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT
CTTGCCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAGTTTAAAGTGCTC
ATCATTGAAAAAGTCTTCCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCC
AGTTGATGTAACCCACTCGTGACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGC
GTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGA
CACGGAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTATCAGGG
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACA

도면7D

TTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTTCGTCTCGCGCGTTTCGGTGATGAC
GGTGAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGA
TGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGGGG
CTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAGAGTGACCATATGCGGTGTG
AAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTGCGCATTGAGG
CTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGG
CGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTC
ACGACGTTGTAAAACGACGGCCAGTG

도면8A

B. 렌투스 GG36 전구체 프로테아제를 인코딩하는 폴리뉴클레오티드

SEQ ID NO:240

GTGAGAAGCAAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTT
 AGTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAAGAAAAATATTTAATTGGCTTTAA
TGAGCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGC
CATTCTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTC
CTGTTTTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACCTCGATCCAGC
GATTTCTTATATTGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCAGTGCCATGG
 GGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTGGATTGACAGGTTCTGGTGT
 AAAAGTTGCTGTCCCTCGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTGG
 CGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCAGCAT
 GTGGCCGGGACGATTGCTGCTCTAAACAATTGATTGGCGTTCTTGGCGTAGCGCCGAG
 CGCGGAACTATACGCTGTTAAAGTATTAGGGGCGAGCGGTTCAAGGCTCGGTCAGCTCG
 ATTGCCAAGGATTGGAATGGGCAGGGAACAATGGCATGCACGTTGCTAATTTGAGTTT
 AGGAAGCCCTTCGCCAAGTGCCACACTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAG
 GCGTTCTTGTGTAGCGGCATCTGGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCC
 CGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAAACAACAACCGCGCCAGCTT
 TTCACAGTATGGCGCAGGGCTTGACATTGTCCGACCAGGTGTAACGTGCAGAGCACAT
 ACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTACTCCTCATGTTGCA
 GGTGCAGCAGCCCTTGTTAAACAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCAAT
 CATCTAAAGAATACGGCAACGAGCTTAGGAAGCACGAACTTGTATGGAAGCGGACTTGT
 CAATGCAGAAGCTGCAACTCGT

도면8B

SEQ ID NO: 241

GTGAGAAGCAAAAAATTGTGGATCGTCGCGTCGACCGCACTACTCATTCTGTTGCTTTT
 AGTTCATCGATCGCATCGGCT

도면8C

SEQ ID NO: 242

GCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTCAGTGAG
 TTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTCTGAGGAAGAGGAAGT
 CGAAATTGAATTGCTTCATGAATTTGAAACGATTCTGTTTTATCCGTTGAGTTAAGCCCA
 GAAGATGTGGACGCGCTTGAACCTCGATCCAGCGATTTCTTATATTGAAGAGGATGCAGA
 AGTAACGACAATG

도면8D

SEQ ID NO: 243
 GCGCAATCAGTGCCATGGGGAATTAGCCGTGTGCAAGCCCCAGCTGCCATAACCGTG
 GATTGACAGGTTCTGGTGTAAGTTGCTGTCCTCGATACAGGTATTTCCACTCATCCAG
 ACTTAAATATTCGTGGTGGCGCTAGCTTTGTACCAGGGGAACCATCCACTCAAGATGGG
 AATGGGCATGGCACGCATGTGGCCGGGACGATTGCTGCTCTAAACAATTGATTGGCGT
 TCTTGGCGTAGCGCCGAGCGCGGAAGTATACGCTGTTAAAGTATTAGGGGCGAGCGGT
 TCAGGCTCGGTCAGCTCGATTGCCCAAGGATTGGAATGGGCAGGGAACAATGGCATGC
 ACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCCACACTTGAGCAAGCTGTTA
 ATAGCGCGACTTCTAGAGGCGTTCTTGTGTAGCGGCATCTGGAAATTCAGGTGCAGGC
 TCAATCAGCTATCCGGCCCGTTATGCGAACGCAATGGCAGTCGGAGCTACTGACCAAAA
 CAACAACCGCGCCAGCTTTTACAGTATGGCGCAGGGCTTGACATTGTCGCACCAGGTG
 TAAACGTGCAGAGCACATAACCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATG
 GCTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCATCTTGGTCC
 AATGTACAAATCCGCAATCATCTAAGAATACGGCAACGAGCTTAGGAAGCACGAACCTTG
 TATGGAAGCGGACTTGCAATGCAGAAGCTGCAACTCGTTA

도면9A

B. 렌투스 GG36 전구체 프로테아제의 아미노산 서열

SEQ ID NO:244
 MRSKKLWIVASTALLISVAFSSSIASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEE
 EEVEIELLHEFETIPVLSVELSPEDVDALELDPAISYIEEDA EVTTMAQSVPWGISRVQAPAAH
 NRGLTGSVGVKAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNHGHGTHVAGTIAALNNSIGV
 LGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLSLGSPSPSATLEQAVNSA
 TSRGLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFSQYGAGLDIVAPGVNVQS
 TYPGSTYASLNGTSMATPHVAGAAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSLV
 NAEAATR

도면9B

SEQ ID NO:245
 MRSKKLWIVASTALLISVAFSSSIASA

도면9C

SEQ ID NO:246
 AEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAILSEEEVEIELLHEFETIPVLSVELSPEDVD
 ALELDPAISYIEEDA EVTTM

도면9D

SEQ ID NO:247

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNG
HGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVANLS
LGSPSPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNRASFS
QYGAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQIRNHLK
NTATSLGSTNLYGSLVNAEAATR