(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0092361
(43) 공개일자 2018년08월20일

(51) 국제특허분류(Int. Cl.)
C12N 9/48 (2006.01) *C12N 1/20* (2006.01)
C12N 15/72 (2006.01)
(52) CPC특허분류
C12N 9/48 (2013.01)
C12N 1/20 (2013.01)
(21) 출원번호 10-2017-0017899
(22) 출원일자 2017년02월09일
심사청구일자 2017년02월09일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
채영기
서울특별시 강남구 삼성로111길 8, 206동 1303호 (삼성동, 삼성동힐스테이트2단지아파트)
김경남
경기도 성남시 분당구 정자로76번길 10, 205동 1903호(정자동, 상록마을라이프2단지아파트)
(74) 대리인
두호특허법인

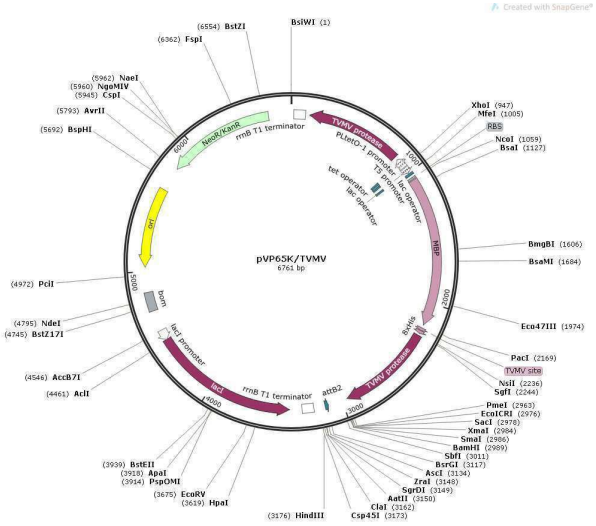
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 **TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 대량 생산 방법**

(57) 요 약

본 발명은 TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 대량 생산 방법에 관한 것으로, 담배 정맥 반점형성 바이러스 유래 (TVMV) 프로테아제 코딩 유전자를 T5 프로모터 하위(downstream)에 가지는 벡터 및 이를 이용함으로써, TVMV 프로테아제를 우수한 수율로 생산할 수 있으며, TVMV 프로테아제를 수용성 발현시킬 수 있다. 또한, TVMV 프로테아제를 우수한 활성을 가지도록 정제할 수 있으며, TVMV 프로테아제를 대량으로 생산할 수 있고, 배양 및 정제가 간편하여 배양 공정을 단순화시킬 수 있는, TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 대량 생산 방법에 관한 것이다.

대 표 도 - 도1



(52) CPC특허분류
C12N 15/72 (2013.01)
C07K 2319/21 (2013.01)
C07K 2319/24 (2013.01)
C12N 2770/34051 (2013.01)
C12N 2800/80 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 C0403064
부처명 중소기업청
연구관리전문기관 한국산학연협회
연구사업명 이공계전문가 기술개발서포터즈사업
연구과제명 유용 프로테아제 생산 및 정제 방법 개선을 통한 수율 제고
기 여 율 1/1
주관기관 세종대학교산학협력단
연구기간 2016.07.15 ~ 2016.11.14

명 세 서

청구범위

청구항 1

담배 정맥 반점형성 바이러스 (TVMV) 프로테아제 코딩 유전자를 T5 프로모터 하위(downstream)에 가지는, TVMV 프로테아제 대량 생산용 벡터.

청구항 2

청구항 1에 있어서, 상기 벡터는 TVMV 프로테아제 코딩 유전자를 PLtetO-1 프로모터 하위에 가지는, TVMV 프로테아제 대량 생산용 벡터.

청구항 3

청구항 1에 있어서, 상기 TVMV 프로테아제 코딩 유전자는 Sgf1 및 Pme1의 제한효소 부위 서열을 포함하는, TVMV 프로테아제 대량 생산용 벡터.

청구항 4

청구항 1에 있어서, 상기 TVMV 프로테아제 코딩 유전자는 서열번호 1의 염기서열을 포함하는, TVMV 프로테아제 대량 생산용 벡터.

청구항 5

청구항 1에 있어서, 상기 벡터는 lac 오퍼레이터를 상기 프로모터와 TVMV 프로테아제 사이에 가지는, TVMV 프로테아제 대량 생산용 벡터.

청구항 6

청구항 5에 있어서, 말토오스 결합 단백질 코딩 유전자를 상기 lac 오퍼레이터와 상기 TVMV 프로테아제 코딩 유전자 사이에 가지는, TVMV 프로테아제 대량 생산용 벡터.

청구항 7

청구항 6에 있어서, 8xHis tag를 코딩하는 유전자를 상기 말토오스 결합 단백질 코딩 유전자와 상기 TVMV 프로테아제 코딩 유전자 사이에 가지는, TVMV 프로테아제 대량 생산용 벡터.

청구항 8

청구항 1에 있어서, 상기 벡터는 도 1의 개열지도를 갖는 것인, TVMV 프로테아제 대량 생산용 벡터.

청구항 9

청구항 1에 있어서, 상기 벡터는 서열번호 8의 염기서열을 포함하는, TVMV 프로테아제 대량 생산용 벡터.

청구항 10

청구항 1의 벡터로 형질전환된 미생물을 배양하여 TVMV 프로테아제를 수득하는 단계를 포함하는, TVMV 프로테아제 대량 생산 방법.

청구항 11

청구항 10에서, 상기 미생물은 대장균인, TVMV 프로테아제 대량 생산 방법.

청구항 12

청구항 11에 있어서, 상기 대장균은 Rosetta2(DE3)pLysS인, TVMV 프로테아제 대량 생산 방법.

청구항 13

청구항 10에 있어서, 상기 배양은 Overnight Express LB 배지에서의 배양을 포함하는, TVMV 프로테아제 대량 생산 방법.

청구항 14

청구항 10에 있어서, 상기 TVMV 프로테아제를 수득하는 단계는

상기 배양으로 수득된 배양액을 원심분리하여 얻은 상층액에 완충용액을 처리하는 단계; 및

상기 완충용액이 처리된 상층액에 용출 완충액을 처리하는 단계;를 포함하는, TVMV 프로테아제 대량 생산 방법.

청구항 15

청구항 14에 있어서, 상기 완충용액은 50 내지 90 mM의 이미다졸을 포함하는, TVMV 프로테아제 대량 생산 방법.

청구항 16

청구항 14에 있어서, 상기 용출 완충액은 200 내지 300 mM의 이미다졸을 포함하는, TVMV 프로테아제 대량 생산 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 담배정맥반점바이러스(Tobacco vein mottling virus, TVMV) 유래의 프로테아제를 코딩하는 유전자를 포함하는 TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 대량 생산 방법에 관한 것이다.

배 경 기 술

[0002] 효소는 오래전부터 식품, 의약품 및 다양한 산업분야에서 유용하게 이용되고 있다. 산업적으로 활용되는 효소 중에서도 단백질 분해효소는 제빵, 치즈 등의 식품산업, 소화제 소염제 등의 의약품과 향장 및 세제를 만드는 등 여러 분야에서 상업적으로 중요한 역할을 하고 있다. 특히 기질의 특정부위를 인식해 절단하는 엔도펩티다제(endopeptidase)의 경우 고부가치의 단백질 의약품 및 산업용 효소를 고순도로 분리정제 하는데 있어서 산업적

으로 중요한 효소이다. 하지만 이러한 단백질분해효소는 세포 외부로 분비되어 단백질성 기질을 분해하는 특성으로 인하여 세포 내에서 매우 낮은 농도로 존재하고, 생산수율이 낮다는 단점이 있다. 이러한 문제점을 극복하기 위해 생산된 효소의 활성을 증가시키거나 안정성을 높이는 연구가 활발히 진행 중이나, 엔도펩티다제 중 하나인 TVMV 프로테아제를 높은 수율로 수득하고 수용성 발현이 가능하며, 수득한 TVMV 프로테아제의 활성을 최대로 저하시키면서 정제하는 방법에 관하여는 공지된 바가 없다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) 한국공개특허공보 제2015-0032440호

발명의 내용

해결하려는 과제

[0004] 본 발명은 목적단백질을 코딩하는 유전자와 보리유래 리보솜 억제 단백질을 코딩하는 유전자를 함유하고, 숙주 미생물에서 목적단백질과 보리유래 리보솜 억제 단백질을 융합단백질 형태로 발현시킬 수 있는 재조합 벡터 및 상기 재조합 벡터를 함유하는 미생물에 관한 것이다.

과제의 해결 수단

[0005] 1. 담배 정맥 반점형성 바이러스 (TVMV) 프로테아제 코딩 유전자를 T5 프로모터 하위(downstream)에 가지는, TVMV 프로테아제 대량 생산용 벡터.

[0006] 2. 위 1에 있어서, 상기 벡터는 TVMV 프로테아제 코딩 유전자를 PLtetO-1 프로모터 하위에 가지는, TVMV 프로테아제 대량 생산용 벡터.

[0007] 3. 위 1에 있어서, 상기 TVMV 프로테아제 코딩 유전자는 Sgf1 및 Pme1의 제한효소 부위 서열을 포함하는, TVMV 프로테아제 대량 생산용 벡터.

[0008] 4. 위 1에 있어서, 상기 TVMV 프로테아제 코딩 유전자는 서열번호 1의 염기서열을 포함하는, TVMV 프로테아제 대량 생산용 벡터.

[0009] 5. 위 1에 있어서, 상기 벡터는 lac 오퍼레이터를 상기 프로모터와 TVMV 프로테아제 사이에 가지는, TVMV 프로테아제 대량 생산용 벡터.

[0010] 6. 위 5에 있어서, 말토오스 결합 단백질 코딩 유전자를 상기 lac 오퍼레이터와 상기 TVMV 프로테아제 코딩 유전자 사이에 가지는, TVMV 프로테아제 대량 생산용 벡터.

[0011] 7. 위 6에 있어서, 8xHis tag를 코딩하는 유전자를 상기 말토오스 결합 단백질 코딩 유전자와 상기 TVMV 프로테아제 코딩 유전자 사이에 가지는, TVMV 프로테아제 대량 생산용 벡터.

[0012] 8. 위 1에 있어서, 상기 벡터는 도 1의 개열지도를 갖는 것인, TVMV 프로테아제 대량 생산용 벡터.

[0013] 9. 위 1에 있어서, 상기 벡터는 서열번호 8의 염기서열을 포함하는, TVMV 프로테아제 대량 생산용 벡터.

[0014] 10. 위 1의 벡터로 형질전환된 미생물을 배양하여 TVMV 프로테아제를 수득하는 단계를 포함하는, TVMV 프로테아제 대량 생산 방법.

[0015] 11. 위 10에서, 상기 미생물은 대장균인, TVMV 프로테아제 대량 생산 방법.

[0016] 12. 위 11에 있어서, 상기 대장균은 Rosetta2(DE3)pLysS인, TVMV 프로테아제 대량 생산 방법.

[0017] 13. 위 10에 있어서, 상기 배양은 Overnight Express LB 배지에서의 배양을 포함하는, TVMV 프로테아제 대량 생산 방법.

[0018] 14. 위 10에 있어서, 상기 TVMV 프로테아제를 수득하는 단계는

[0019] 상기 배양으로 수득된 배양액을 원심분리하여 얻은 상층액에 완충용액을 처리하는 단계; 및 상기 완충용액이 처

리된 상층액에 용출 완충액을 처리하는 단계;를 포함하는, TVMV 프로테아제 대량 생산 방법.

[0020] 15. 위 14에 있어서, 상기 완충용액은 50 내지 90 mM의 이미다졸을 포함하는, TVMV 프로테아제 대량 생산 방법.

[0021] 16. 위 14에 있어서, 상기 용출 완충액은 200 내지 300 mM의 이미다졸을 포함하는, TVMV 프로테아제 대량 생산 방법.

발명의 효과

[0022] 본 발명의 벡터 및 이를 사용하는 생산 방법은 TVMV 프로테아제를 우수한 수율로 생산할 수 있다.

[0023] 본 발명의 벡터 및 이를 사용하는 생산 방법은 TVMV 프로테아제를 수용성 발현시킬 수 있다.

[0024] 본 발명의 벡터 및 이를 사용하는 생산 방법은 TVMV 프로테아제의 활성 저하를 최소화하면서 높은 순도로 정제할 수 있다.

[0025] 본 발명의 벡터 및 이를 사용하는 생산 방법은 TVMV 프로테아제를 대량으로 생산할 수 있다.

[0026] 본 발명의 벡터 및 이를 사용하는 생산 방법은 저비용으로 TVMV 프로테아제를 생산할 수 있다.

[0027] 본 발명의 생산 방법은 배양 및 정제가 간편하여 TVMV 프로테아제 생산 공정을 단순화시킬 수 있다.

도면의 간단한 설명

[0028] 도 1은 본 발명의 벡터의 개열지도를 나타낸 것이다.

도 2는 pVP65K/TVMV로 형질전환된 Rosetta2(DE3) (Rosetta2(DE3)/pVP65K/TVMV) 및 Rostetta2(DE3)pLysS (Rostetta2(DE3)pLysS/pVP65K/TVMV)를 Overnight Express LB배지(Overnight Express) 또는 IPTG 유도 방법(IPTG)을 사용한 경우, 각각의 TVMV 프로테아제의 생산량을 PAGE로 나타낸 것이다. (C = cell, S = supernatant (상층액), p = cell pellet, OE = Overnight Express LB배지)

도 3은 TVMV 프로테아제의 정제과정을 보여주는 PAGE이다. IPTGx로 표시한 레인은 프로테아제 생산이 되지 않은 콘트롤이고, 상기 레인은 프로테아제가 생산된 경우와 비교하기 위해 사용하였다(C = cell, S = supernatant (상층액), FT = Filtration, p = cell pellet, elute = elution, old TVMV = 예비실험으로 분리한 TVMV 프로테아제).

도 4는 MBP-NusG를 발현하는 벡터인 pRK1036로 형질전환된 Rosetta2(DE3)pLysS를 30℃로 Overnight Express LB 배지에서 배양 (Rosetta2(DE3)/pLysS/pRK1036_OE_30℃)하여 MBP-NusG를 생산한 것을 PAGE로 나타낸 것이다.

도 5는 TVMV 프로테아제를 농도를 달리하여 처리하였을 경우 기질인 MPB-NusG 1mg이 1시간 내 완전히 분해되는 TVMV 프로테아제의 Unit을 결정하여 TVMV 프로테아제의 활성을 측정하기 위한 PAGE 실험을 나타낸 것이다(1, 2, 4, 8, 16, 32, 64, 128, 256, 512x TVMV는 0.75mg/mL의 TVMV 프로테아제를 각각 1/1, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128, 1/512배로 희석한 것을 의미함).

발명을 실시하기 위한 구체적인 내용

[0029] 본 발명은 TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 대량 생산 방법에 관한 것으로, 담배 정맥 반점형성 바이러스 유래 (TVMV) 프로테아제 코딩 유전자를 T5 프로모터 하위(downstream)에 가지는 벡터 및 이를 이용함으로써, TVMV 프로테아제를 우수한 수율로 생산할 수 있으며, TVMV 프로테아제를 수용성 발현시킬 수 있다. 또한, TVMV 프로테아제를 우수한 활성을 가지도록 정제할 수 있으며, TVMV 프로테아제를 대량으로 생산할 수 있고, 배양 및 정제가 간편하여 배양 공정을 단순화시킬 수 있는, TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 대량 생산 방법에 관한 것이다.

[0030] 이하, 본 발명을 상세히 설명한다.

[0031] 용어 "벡터"는 세포 내로 전달하는 DNA 단편(들), 핵산 분자를 지칭하는 것으로, DNA를 복제시키고, 숙주세포에서 독립적으로 재생산될 수 있는 것이다.

[0032] 본 발명의 TVMV 프로테아제 대량 생산용 벡터는 담배 정맥 반점형성 바이러스 유래 (TVMV) 프로테아제 코딩 유전자를 T5 프로모터 하위 (downstream)에 가진다.

[0033] TVMV (tobacco vein mottling virus)는 식물 바이러스인 포티비리디 (Potyviridae) 그룹에 속하는 포티 바이러

스(potyvirus)의 하나이다. TVMV (tobacco vein mottling virus) 프로테아제는 27kDa의 NIa(Nuclear Inclusion a) 단백질로서 TVMV에 의해 만들어지며, 이는 Xa인자, 트롬빈이나 엔테로키나제보다 훨씬 더 특이적으로 ENLYFQG(S)를 인식해 Q와 G(S)사이를 절단하는 단백질 분해 효소이다. 상기 TVMV 프로테아제 코딩 유전자는 서열번호 1의 염기서열을 포함한다.

[0034] 상기 TVMV 프로테아제 코딩 유전자는 해당 유전자의 염기 서열을 참조하여 핵산 합성기 등을 이용하여 인공적으로 합성하거나, 해당 유전자를 함유하는 벡터를 사용하여, 해당 유전자의 양 말단에 상보적인 서열을 가지는 올리고 뉴클레오타이드를 프라이머로 이용하여 PCR을 실시함으로써 제조할 수 있다.

[0035] 본 발명의 일 실시예에 따르면, 상기 제조된 TVMV 프로테아제 코딩 유전자는 Sgf1 및 Pme1의 제한효소 부위 서열을 포함한 것일 수 있다. 본 발명의 일 실시예에 따르면, TVMV 프로테아제 코딩 유전자는 전방향 프라이머 (예를 들면, 서열번호 2: 5'-GGT TGC GAT CGC Cat gtc taa ggc ttt gct gaa g-3')와 역방향 프라이머 (예를 들면, 서열번호 3: 5'-GTG TGT TTA AAC tta gtc cat gat ggc ggc aac agt-3')로 중합효소연쇄반응 (PCR)로 증폭되어 수득된 것일 수 있다.

[0036] 프라이머는 PCR에서 이용되는 타깃 핵산에 인접해 있는 5'말단 서열 및 3'말단 서열에 상보적인 올리고뉴클레오타이드를 의미한다. "전방향 프라이머" 및 "역방향 프라이머"는 중합효소연쇄반응에 의해 증폭되는 유전자의 일정 부위의 3'말단 및 5'말단에 각각 결합하는 프라이머를 의미한다. 프라이머로써 이용된 올리고뉴클레오티드는 뉴클레오티드 유사체(analogue), 예를 들면, 포스포로티오에이트(phosphorothioate), 알킬포스포로티오에이트 또는 펩티드 핵산(peptide nucleic acid)을 포함할 수 있으며, 또는 삽입 물질(intercalating agent)을 포함할 수도 있다.

[0037] 프라이머는 표적 서열에 혼성화되고, 표적 서열을 증폭할 수 있다. 혼성화는 DNA 주형과 비공유적으로 결합하는 것을 의미하며, 보다 상세하게는 프라이머가 DNA 증폭반응에서 본래의 기능을 할 수 있을 정도로 DNA 주형과 비공유적으로 결합되어 있는 상태를 의미하는 것이다. 즉, DNA 주형에 대하여 완전하게 왓슨/크릭 염기쌍을 이루는 것뿐만 아니라, 부분적으로 왓슨/크릭 염기쌍을 이루지 못한 부분이 있다 하더라도, 프라이머로 기능을 할 수 있을 정도로 DNA 주형에 대해 결합한 것을 의미한다. 프라이머를 이용해 표적 핵산을 증폭하는 방법에는 중합효소연쇄반응(PCR), 리가아제 연쇄반응(ligase chain reaction), 핵산 서열 기재 증폭(nucleic acid sequence-based amplification), 전사 기재 증폭 시스템 (transcription-based amplification system), 가닥치환 증폭(strand displacement amplification) 또는 Qβ복제효소(replicase)를 통한 증폭 또는 당업계에 알려진 핵산 분자를 증폭하기 위한 임의의 기타 적당한 방법이 가능하다.

[0038] 중합효소연쇄반응(PCR)은 표적 핵산을 선별적으로 증폭하는 방법으로서 이에 대한 내용은 Miller, H.I.(WO 89/06700) 및 Davey, C. et al. (EP 329,822))에 개시되어 있으며, 이 문헌은 본 명세서에 참조로써 삽입된다. 중합효소연쇄반응(PCR)에는 DNA 중합효소(예컨대, Thermus aquaticus(Taq), Thermus thermophilus(Tth), Thermus filiformis, Thermis flavus, Thermococcus literalis 또는 Pyrococcus furiosus(Pfu)로부터 수득한 열 안정성 DNA 중합효소, dNTP 혼합물(dATP, dCTP, dGTP, dTTP), PCR 완충액 (buffer) 및 DNA 중합효소 조인자를 포함할 수 있다.

[0039] 벡터는 적합한 숙주 내에서 목적 유전자를 발현시킬 수 있도록 적합한 프로모터에 작동 가능하게 연결된 유전자의 염기서열을 함유하는 유전자 작제물을 의미하는 것으로, 상기 조절 서열은 전사를 개시할 수 있는 프로모터, 그러한 전사를 조절하기 위한 임의의 오퍼레이터 서열, 및 전사 및 해독의 종결을 조절하는 서열을 포함할 수 있다.

[0040] 상기 벡터는 세포 내에서 복제 가능한 것이면 특별히 한정되지 않고 당업계에 알려진 임의의 벡터를 이용할 수 있으며, 예컨대 플라스미드, 코즈미드, 파지 입자, 바이러스 벡터 등일 수 있다. 예를 들어, 벡터는 pVP65K, pUC19, pSTV28, pBBR1MCS, pBluscriptII, pBAD, pTrc99A, pET, pACYC184, pBR322 계열 등의 당업계에 상용화된 벡터를 사용할 수 있으며, 구체적으로는 pVP65K 벡터(CESG, Univ. of Wisconsin - Madison, USA)를 사용할 수 있으나, 이로 제한되는 것은 아니다. 상기 pVP65K 벡터를 사용하여 클로닝을 수행하면, 발현되는 단백질의 말단에 His tag이 결합되어 발현되므로, 상기 단백질을 효과적으로 정제할 수 있다. 클로닝된 유전자로부터 발현된 융합단백질을 분리하기 위해서는 당업계에 공지된 일반적인 방법이 이용될 수 있으며, 구체적으로, 본 발명에서는 Ni-NTA 칼럼을 사용하는 크로마토그래피를 이용하여 분리할 수 있다.

[0041] 작동 가능하게 연결된 (operably linked)는 일반적 기능을 수행하도록 핵산 발현조절 서열과 목적하는 단백질을 코딩하는 핵산 서열이 기능적으로 연결되어 있는 것을 말한다. 재조합 벡터와의 작동적 연결은 당해 기술분야에

서 잘 알려진 유전자 재조합 기술을 이용하여 제조할 수 있으며, 부위-특이적 DNA 절단 및 연결은 당해 기술 분야에서 일반적으로 알려진 효소 등을 사용한다.

[0042] 목적 유전자의 발현은 상기 목적 유전자를 발현시켜 목적 유전자가 코딩하는 단백질을 생산하는 것을 의미할 수 있다. 본 발명에서 목적 유전자를 발현하는 방법은 상기 목적 유전자를 포함하는 벡터로 형질 전환된 미생물(숙주세포)을 배양하여 상기 목적 유전자가 코딩하는 단백질을 발현시키는 방법일 수 있으며, 이를 통해 상기 단백질이 관여되는 생합성 경로의 최종산물을 제조할 수 있다.

[0043] 프로모터는 구조 유전자로부터의 DNA 업스트림의 영역을 의미하며 전사를 개시하기 위하여 RNA 폴리머라아제가 결합하는 DNA 분자를 말하며, 본 발명에서 프로모터는 숙주 내에서 작동할 수 있으면 제한 없이 사용될 수 있다. 예컨대, 프로모터는 TATA 박스와 전사 개시위치를 특정하는 역할을 하는 짧은 DNA 서열인 최소 프로모터를 포함할 수 있으며, 발현 제어를 위한 조절성분이 더 첨가될 수 있으나, 이에 제한되는 것은 아니다. 예컨대, 프로모터는 전체가 천연 유전자로부터 유래될 수 있거나, 자연에서 발견된 상이한 프로모터들로부터 유래된 상이한 성분들로 이루어질 수 있으며, 또는 합성 DNA 절편으로 이루어질 수 있으나, 이에 제한되는 것은 아니다. 예컨대, 프로모터는 생리적 또는 성장 조건에 대한 반응으로 전사 개시의 유효성을 제어하는 단백질 인자의 결합에 관여하는 DNA 서열을 함유할 수 있으나, 이에 제한되는 것은 아니다. 예컨대, 프로모터는 모든 시간대에 상시적으로 목적 유전자의 발현을 유도하는 프로모터 또는 특정한 위치, 시기에 목적 유전자의 발현을 유도하는 프로모터를 사용할 수 있으며, 그 예로 T5 프로모터, U6 프로모터, CMV(cytomegalovirus) 프로모터, SV40 프로모터, CAG 프로모터, CaMV 35S 프로모터, Rsyn7 프로모터, 라이스 액틴(rice actin) 프로모터, 유비퀴틴 프로모터, ALS 프로모터, PLtetO-1 프로모터 등을 들 수 있으나, 이에 제한되는 것은 아니다.

[0044] 오퍼레이터는 전사억제물질(repressor)이 부착하는 부위로, 전사억제물질을 일정한 조건을 처리하여 오퍼레이터로부터 분리시킴으로써 전사를 시작하게 한다. 즉, 오퍼레이터는 프로모터에 의한 전사를 조절하는 역할을 한다.

[0045] 본 발명의 일 실시예에 따르면, 상기 오퍼레이터는 lac 오퍼레이터일 수 있으며, lac 오퍼레이터는 서열번호 2의 염기서열을 포함한다. 본 발명의 일 실시예에 따르면, 상기 벡터는 lac 오퍼레이터를 상기 프로모터와 TVMV 프로테아제 사이에 가질 수 있다.

[0046] 본 발명의 일 실시예에 따르면, 상기 벡터는 TVMV 프로테아제 코딩 유전자를 PLtetO-1 프로모터 하위에 가질 수 있다. 즉, 본 발명의 벡터에는 TVMV 프로테아제를 코딩하는 유전자 부위가 2개 또는 그 이상이며 서로 다른 위치 및 서로 다른 프로모터 하위에 위치할 수 있다. TVMV 프로테아제 코딩 유전자는 전술한 TVMV 프로테아제 코딩 유전자일 수 있으며, 상기 PLtetO-1 프로모터는 서열번호 4의 염기서열을 포함한다. PLtetO-1 프로모터에 의해 발현된 TVMV 프로테아제는 TVMV site를 특이적으로 식별하여 절단한다. 상기 TVMV site는 서열번호 5의 염기서열을 포함한다.

[0047] 상기 TVMV site는 말토오스 결합 단백질 코딩 유전자와 TVMV 프로테아제 코딩 유전자 사이에 위치할 수 있다. 말토오스 결합 단백질은 수용성 향상 태그로, 상기 수용성 향상 태그는 TVMV 프로테아제를 수용성 형태로 발현시킬 수 있는 것이면 제한 없이 사용 가능하다. 상기 말토오스 결합 단백질 외 예컨대, 티오레독신(Thioredoxin, Trx), 글루타치온 S-트랜스퍼라제(Glutatione S-transferase; GST), N-이용 기질 단백질 A(N-utilization substance Protein A; NusA) 또는 인간 단백질 이황화물 이성질화효소(human protein disulfide isomerase; hPDI)를 코딩하는 유전자를 상기 말토오스 결합 단백질 코딩 유전자 대신 사용할 수 있으나, 이에 제한되지 않는다. TVMV 프로테아제와 말토오스 결합 단백질이 결합된 형태로 발현되는 경우 수용성 발현이 가능해진다. 그리고, TVMV site가 TVMV 프로테아제와 말토오스 결합 단백질 사이에 위치하도록 함으로써 이들의 절단이 용이해져 TVMV 프로테아제의 정제가 용이해질 수 있다. 말토오스 결합 단백질 코딩 유전자는 서열번호 6의 염기서열을 포함한다.

[0048] 본 발명의 일 실시예에 따르면, 8xHis tag를 코딩하는 유전자를 상기 말토오스 결합 단백질 코딩 유전자와 상기 TVMV 프로테아제 코딩 유전자 사이에 가질 수 있다. 상기 8xHis tag를 코딩하는 유전자를 가짐으로써 TVMV 프로테아제를 보다 용이하게 정제할 수 있다. 즉, 8xHis tag에 의해 Ni-NTA 컬럼으로 TVMV 프로테아제를 정제할 수 있어 간편하게 정제가 가능하다. 8xHis tag를 코딩하는 유전자는 서열번호 7의 염기서열을 포함한다.

[0049] 본 발명의 일 실시예에 따르면, 상기 벡터는 도 1의 개열지도 또는 서열번호 8의 염기서열을 포함할 수 있다.

[0050] 본 발명의 일 실시예에 따르면. 본 발명의 벡터는 하나 이상의 선택성 마커를 포함할 수 있다. 상기 마커는 통상적으로 화학적인 방법으로 선택될 수 있는 특성을 갖는 핵산 서열로, 형질 전환된 세포를 비형질전환 세포로

부터 구별할 수 있는 모든 유전자가 이에 해당된다. 그 예로는 글리포세이트 (glyphosate) 또는 포스피노트리신과 같은 제초제 저항성 유전자, 카나마이신, G418, 블레오마이신 (Bleomycin), 하이그로마이신 (hygromycin), 클로람페니콜 (chloramphenicol)과 같은 항생제 내성 유전자가 있으나, 이에 한정되는 것은 아니다.

[0051] 본 발명의 벡터는 전술한 각 염기서열과 각각 70% 이상, 구체적으로는 80% 이상, 보다 더욱 구체적으로는 90% 이상, 가장 구체적으로는 95% 이상의 서열 상동성을 가지는 염기서열을 포함할 수 있다. 폴리뉴클레오티드에 대한 "서열 상동성의 %"는 두 개의 최적으로 배열된 서열과 비교 영역을 비교함으로써 확인되며, 비교 영역에서의 폴리뉴클레오티드 서열의 일부는 두 서열의 최적 배열에 대한 참고 서열(추가 또는 삭제를 포함하지 않음)에 비해 추가 또는 삭제(즉, 갭)를 포함할 수 있다.

[0052] 또한, 본 발명의 벡터로 형질전환된 미생물을 배양하여 TVMV 프로테아제를 수득하는 단계를 포함하는 TVMV 프로테아제를 대량 생산하는 방법을 제공한다.

[0053] TVMV 프로테아제는 전술한 TVMV 프로테아제일 수 있다.

[0054] 미생물은 원핵세포일 수 있으며, 구체적으로, 대장균일 수 있다. 상기 원핵세포로는, 이로 제한되지는 않으나, 예컨대, E. coli균주 BL21(DE3), E. coli균주 DH5a, E. coli균주 JM101, E. coli K12균주 294, E. coli균주 W3110, E. coli균주 X1776, E. coli XL1-Blue (Stratagene), Rosetta2(DE3), Rosetta2(DE3)pLysS 및 E. coli B 등을 들 수 있으며, 구체적으로는 Rosetta2(DE3)pLysS일 수 있다. 이 외 FMB101, NM522, NM538 및 NM539와 같은 E. coli균주 및 다른 원핵생물의 종(speices) 및 속 (genera)등 또한 사용될 수 있다. 전술한 E. coli에 덧붙여, 상기 원핵세포로 아그로박테리움 A4와 같은 아그로박테리움 속 균주. 바실루스 섭틸리스(Bacillus subtilis)와 같은 바실리(Bacilli), 살모넬라 타이피뮤리움(Salmonella typhimurium) 또는 세라티아 마르게센스(Serratia marcescens)와 같은 또 다른 장내세균 및 다양한 슈도모나스 (Pseudomonas) 속 균주 등을 추가로 더 들 수 있으나, 이에 제한되지 않는다.

[0055] 배양은 재조합 미생물을 적당히 인공적으로 조절한 환경조건에서 생육시키는 것을 의미한다. 본 발명의 일 실시예에 따르면, 상기 재조합 미생물을 배양하는 배지는 overnight express LB배지일 수 있다. Overnight express LB배지는 자가유도가 가능하여 LB배지에서처럼 IPTG 유도 시점을 잡기 위해 OD_{600}을 계속 측정할 필요가 없고, 글리세롤이 첨가되어 LB 배지에서보다 배양된 세포 수가 현저히 많아(최종 OD_{600} > 10), TVMV 프로테아제를 간편하고 우수한 수율로 수득할 수 있다.

[0056] 배양에 사용되는 배지는 적당한 탄소원, 질소원, 인원, 무기화합물, 아미노산, 비타민 등을 함유한 통상의 배지 내에서 온도, pH 등을 조절하면서 적절한 방식으로 특정 균주의 요건을 충족해야 한다.

[0057] 탄소원으로는 글루코즈 및 자일로즈의 혼합당을 주 탄소원으로 사용하며 이외에 수크로즈, 락토즈, 프락토즈, 말토즈, 전분, 셀룰로즈와 같은 당이 포함될 수 있다.

[0058] 또한, 탄수화물, 대두유, 해바라기유, 피마자유, 코코넛유 등과 같은 오일 및 지방, 팔미트산, 스테아린산, 리놀레산과 같은 지방산, 글리세롤, 에탄올과 같은 알코올, 아세트산과 같은 유기산이 포함될 수 있다. 이들 물질은 개별적 또는 혼합물로 포함될 수 있다.

[0059] 질소원으로 암모니아, 황산암모늄, 염화암모늄, 초산암모늄, 인산암모늄, 탄산안모늄, 및 질산암모늄과 같은 무기질소원; 글루탐산, 메티오닌, 글루타민과 같은 아미노산 및 펩톤, NZ-아민, 육류 추출물, 효모 추출물, 맥아 추출물, 옥수수 침지액, 카세인 가수분해물, 어류 또는 그의 분해생성물, 탈지 대두 케이크 또는 그의 분해생성물 등 유기질소원이 사용될 수 있다. 이들 질소원은 단독 또는 조합되어 사용될 수 있다.

[0060] 인원으로 인산이수소칼륨 또는 나트륨, 인산수소이칼륨 또는 나트륨 등이 사용될 수 있다.

[0061] 또한, 무기화합물로 염화나트륨, 염화칼슘, 염화철, 황산마그네슘, 황산철, 황산망간 및 탄산칼슘 등이 사용될 수 있다.

[0062] 아미노산으로 아르기닌, 시스테인, 글루타민, 히스티딘, 이소류신, 류신, 리신, 메티오닌, 페닐알라닌, 트레오닌, 트립토판, 티로신, 발린, 알라닌, 아스파라긴, 아스파라긴산, 글루탐산, 글리신, 프롤린, 세린, 이들의 이성질체 등이 사용될 수 있다.

[0063] 비타민으로 비타민 B 그룹 (엽산, 비오틴, 판토텐산 등), 비타민 C (아스코르브산), 비타민 E (토코페롤) 등이 사용될 수 있다.

[0064] 또한, 배양 배지에 첨가되기에 적절한 이들의 전구체들이 사용될 수 있다. 상기된 원료들은 배양과정에서 배양물에 적절한 방식에 의해 회분식, 유가식 또는 연속식으로 첨가될 수 있으나, 특별히 이에 제한되지는 않는다. 수산화나트륨, 수산화칼륨, 암모니아와 같은 기초 화합물 또는 인산 또는 황산과 같은 산 화합물을 적절한 방식으로 사용하여 배양물의 pH를 조절할 수 있다.

[0065] 본 발명의 일 실시예에 따르면, 본 발명은 상기 TVMV 프로테아제를 수득하는 단계는 상기 배양으로 수득된 배양액을 원심분리하여 얻은 상층액에 완충용액을 처리하는 단계; 및 상기 완충용액이 처리된 상층액에 용출 완충액을 처리하는 단계;를 포함할 수 있다.

[0066] 본 발명의 일 실시예에 따르면, 원심분리는 예를 들면 1000 rpm 이상 5000 rpm이하로 수행될 수 있다.

[0067] 완충용액으로는 DNA를 변성시키지 않으며 용액 내 pH를 유지시켜줄 수 있는 것이면 제한 없이 사용될 수 있다.

[0068] 완충 용액으로 TAPS, 바이신 (Bicine), Tris-HCl, PBS, TAPSO, HEPES, TES, MOPS, Tricine 등이 사용될 수 있으며, 상기 완충용액에 필요에 따라 이미다졸이 더 포함될 수 있다.

[0069] 본 발명의 일 실시예에 따르면, 상기 완충용액은 50 내지 90 mM의 이미다졸을 포함할 수 있으며, 구체적으로는 65 내지 85 nM의 이미다졸을 포함할 수 있다. 상기 범위내로 이미다졸을 포함하면, TVMV 프로테아제를 우수한 활성도를 가지도록 정제할 수 있다.

[0070] 용출 완충액은 형질 전환된 미생물에서 생산된 TVMV 프로테아제를 용출시키는 동안 단백질의 변성이나 분해를 막을 수 있게 pH를 유지할 수 있는 완충용액을 말하며, TVMV 프로테아제의 변성이나 분해를 막을 수 있는 것이면 제한 없이 사용될 수 있다. 본 발명의 일 실시예에 따르면, 상기 용출 완충액은 200 내지 300 mM의 이미다졸을 포함할 수 있으며, 구체적으로는 230 내지 270 mM의 이미다졸을 포함할 수 있으나, 이에 제한되지 않는다.

[0071] 본 발명의 일 실시예에 따르면, 상기 TVMV 프로테아제를 수득하는 단계를 수행 후, 버퍼를 교환 및 농축하고, 40 내지 60중량%의 글리세롤을 첨가하여 영하의 온도에서 보관하는 단계를 더 포함할 수 있다. 상기 범위 내의 글리세롤을 포함하면 TVMV 프로테아제를 영하의 온도에서도 동결되는 것을 방지해주는 효과가 극대화되며, 상기 범위보다 낮은 중량%의 글리세롤이 첨가되면 TVMV 프로테아제가 보관 중 결빙되는 문제가 발생할 수 있으며, 상기 범위보다 높은 중량%의 글리세롤은 경제적이지 않아 바람직하지 않은 문제점이 있다.

[0072] 영하의 온도는 -10 내지 -30℃일 수 있으며, 상기 범위 내인 경우 TVMV 프로테아제를 장기간 보관할 수 있다. 상기 범위보다 더 낮은 온도로 TVMV 프로테아제를 저장하는 경우 TVMV 프로테아제가 결빙될 수 있고, 상기 범위보다 더 높은 온도로 TVMV 프로테아제를 저장하는 경우 TVMV 프로테아제의 저장기간이 짧아질 수 있다.

[0073] 형질전환은 표적 단백질을 코딩하는 염기서열을 포함하는 벡터를 숙주세포 내에 도입하여 상기 염기서열이 코딩하는 단백질이 발현할 수 있도록 하는 것으로, 형질전환은 염화칼슘 용액 처리 후 열처리 또는 전기 충격하는 방법, 반응능 세포에 순간적인 열 충격을 주어 세포 내로 플라스미드 DNA를 들어가게 하는 방법인 열 충격 방법, 반응능 세포에 짧고 강한 전류를 흘려주어 전기 충격에 의해 박테리아 세포막을 불안정하게 하여 순간적인 구멍이 형성됨으로써 이곳을 통해 세포 내로 플라스미드 DNA를 들어가게 하는 방법 등으로 수행될 수 있으나, 이에 제한되지 않는다.

[0074] 이하, 본 발명을 하기 실시예로 더욱 구체적으로 설명한다. 하기 실시예는 본 발명을 예시하는 것일 뿐 본 발명의 권리범위를 제한하지 않는 것은 당연하다.

[0075] **실시예**

[0076] **1. TVMV 프로테아제 코딩 유전자 제조**

[0077] pVP65K 벡터 (CESG, Univ. of Wisconsin - Madison, USA, 서열번호 9)에 들어있는 TVMV protease 유전자를 전방향 프라이머 (서열번호 2: 5'-GGT TGC GAT CGC Cat gtc taa ggc ttt gct gaa g-3')과 역방향 프라이머 (서열번호 3: 5'-GTG TGT TTA AAC tta gtc cat gat ggc ggc aac agt-3')를 사용하여 PCR로 증폭하고 SgfI/PmeI으로 절단하고 정제하여 TVMV 프로테아제 코딩 유전자 (서열번호 1)를 수득하였다.

[0078] **2. TVMV 프로테아제 대량 생산용 벡터의 제조**

[0079] pVP65K 벡터를 SgfI/PmeI으로 절단하고 정제한 후, 상기 TVMV 프로테아제 코딩 유전자를 SgfI/PmeI로 이어붙였다. TVMV 프로테아제 유전자가 pVP65K에 삽입된 결과물을 DNA sequencing을 통해 확인하였으며, 상기 벡터를 pVP65K/TVMV로 명명하였다.

[0080] pVP65K/TVMV는 T5 프로모터와 lac 오퍼레이터로 발현 조절 가능하며, pVP65K/TVMV에서 TVMV 프로테아제는 말토오스 결합 단백질 (MBP)과 fusion protein의 형태로 생산된다. MBP 와 TVMV 프로테아제 사이에는 TVMV 프로테아제가 절단할 수 있는 아미노산 서열 (TVMV site)이 있기 때문에 fusion 형태로 생산된 후 세포 내에서 절단될 수 있다. MBP로부터 분리된 TVMV 프로테아제의 N- 말단에는 8xHis Tag가 달려 있어서 Ni-NTA 칼럼으로 정제 가능하다. 최종 pVP65K/TVMV벡터는 Rosetta2 (DE3)pLysS에 삽입되어 형질 전환되었다.

[0081] **3. TVMV 프로테아제 생산 방법 정립**

[0082] LB 배지와 IPTG 유도를 사용한 방법과 OvernightExpress LB 배지에서의 TVMV 생산량을 확인하였다 (도 2). OvernightExpress LB 배지는 자가유도가 가능하기 때문에 LB 배지에서처럼 IPTG 유도의 시점을 잡기 위해 OD_{600}을 계속 측정할 필요가 없고, 특히, OvernightExpress LB배지에는 글리세롤이 첨가되어 있어 LB 배지에서보다 세포 수가 훨씬 많았다 (최종 OD_{600} >10). LB 배지와 IPTG 유도를 사용한 방법과 OvernightExpress LB 배지에서의 TVMV 프로테아제 생산량을 확인하였다 (도 2). 배양을 37℃, 30℃ 및 25℃의 조건에서 수행하였으며, 수용성 TVMV 프로테아제의 생산량과 배양 시간을 고려하여 30℃로 배양 온도를 결정하였다.

[0083] 세포 배양이 끝난 후, 형질 전환된 세포주를 수확하고 10 mM TrisHCl pH 8.0 버퍼에 재현탁시킬 때, 프로테아제 억제자 칵테일 알약을 혼합하여 TVMV 프로테아제가 절단되는 것을 최소화하였다.

[0084] 효소 생산용 플라즈미드 (pVP65K/TVMV)를 집어넣은 대장균 (Rosetta2(DE3)pLysS, Novagen, Madison, WI, USA)을 100 mL OvernightExpress LB 배지가 들어있는 1L baffled flask에서 24시간 배양 후 원심분리를 통해 수확하였다. 배지에는 카나마이신 (kanamycin)과 클로람페니콜 (chloramphenicol)이 각각 50 ug/mL와 34 ug/mL의 농도로 들어있었다. 세포 pellet은 10 mL의 10 mM TrisHCl pH8.0 버퍼에 재현탁시켜 15mL 코니컬튜브에 옮긴 후 -80도 냉동고에 보관하였다.

[0085] **4. TVMV 프로테아제 정제 방법의 최적화**

[0086] 동결된 세포 샘플을 녹이고, Rosetta2(DE3)pLysS에서 미량 생산되는 T7 lysozyme을 활용하여 freeze-and-thaw의 방법으로 대장균 세포를 파괴한 후 DNaseI 1 mg을 실온에서 10분간 처리하여 DNA를 절단하고, TritonX-100을 1%로 넣어준 후, 섭씨 4도에서 10,000g로 원심분리하여 상층액 수집하였다. 수집된 상층액을 HisTrap HP 컬럼 (5mL, GE Healthcare, Marlborough, MA, USA)에 로딩하고, 10mM imidazole이 들어있는 PBS 버퍼 40mL와 75mM imidazole이 들어있는 PBS 버퍼 40mL로 순차적으로 씻어준 후, 250mM imidazole가 들어있는 PBS 버퍼 20mL로 TVMV 프로테아제를 용출시켰다 (도 3). 용출된 TVMV protease가 들어있는 분획은 VivaSpin-20 (Sartorius)을 사용하여 PBS 버퍼로 교환하고 2mL까지 농축한 후 글리세롤을 50%로 첨가하여 -20℃ 냉동고에 보관하였다.

[0087] **5. TVMV 프로테아제 수율**

[0088] 본 발명의 TVMV 프로테아제 생산 방법 및 정제 방법에 따라 수득한 TVMV 프로테아제를 BCA assay kit를 사용하여 정량한 결과 100mL OvernightExpress 배지에서 약 3 mg의 TVMV 프로테아제를 생산한 것을 확인하였다 (도 3). 이는 1L의 배지에서는 30mg의 TVMV 프로테아제를 생산할 수 있다는 것을 의미하므로, 이를 통해 대량의 TVMV 프로테아제 생산 가능성을 확인할 수 있었다.

[0089] **6. TVMV 프로테아제 활성 측정**

[0090] TVMV 프로테아제 활성은 TVMV 프로테아제의 활성을 1 Unit을 실온에서 1 시간 내에 MBP-NusG 1mg을 완전히 가수분해하는 TVMV 프로테아제 의 양으로 정의하여 TVMV 프로테아제 활성을 측정하였다.

[0091] MBP-NusG 단백질을 OvernightExpress 배지에서 약 50 mg을 생산하고 정제하여 활성을 측정을 위한 기질로 사용하였다. 활성은 하기의 표 1의 조성으로 측정하였다.

표 1

[0092]

조성	부피
TVMV 프로테아제 (0.75mg/mL)	25uL
MBP-NusG (50mg/mL)	2uL
bME (1.43M, 10x diluted stock)	1uL
1M Tris-HCl pH 8.0	2uL
Double-distilled water (ddH20)	70uL

총량	100uL

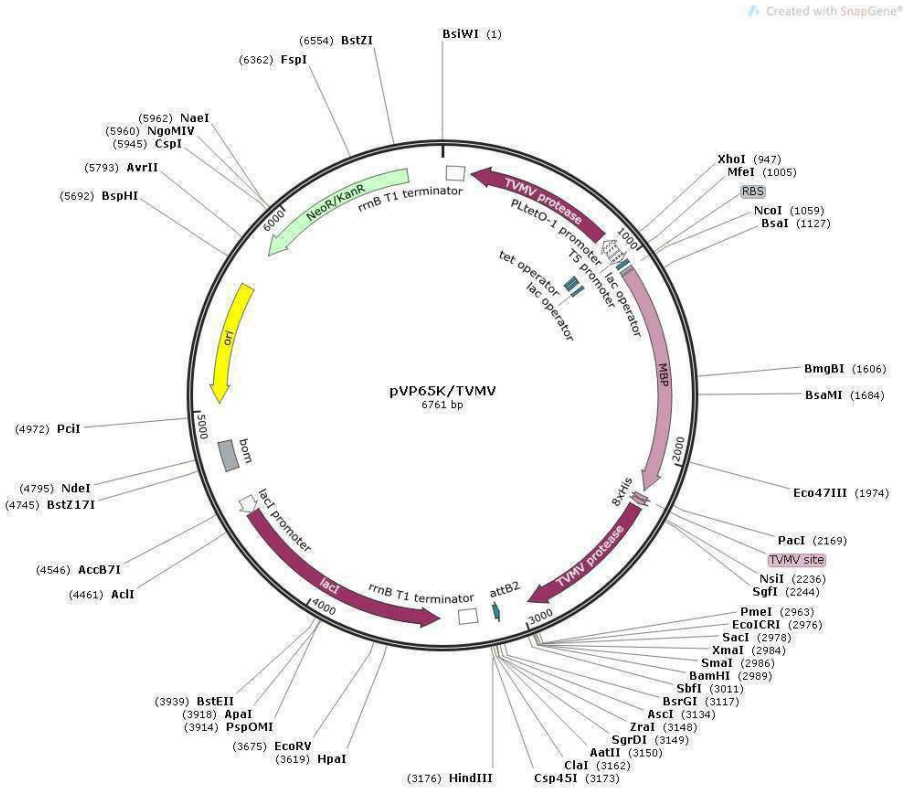
[0093] 활성을 측정한 결과, 정제된 TVMV 프로테아제의 활성은 13 Unit/㎍으로 확인되었다 (도 5).

[0094] **7. 결론**

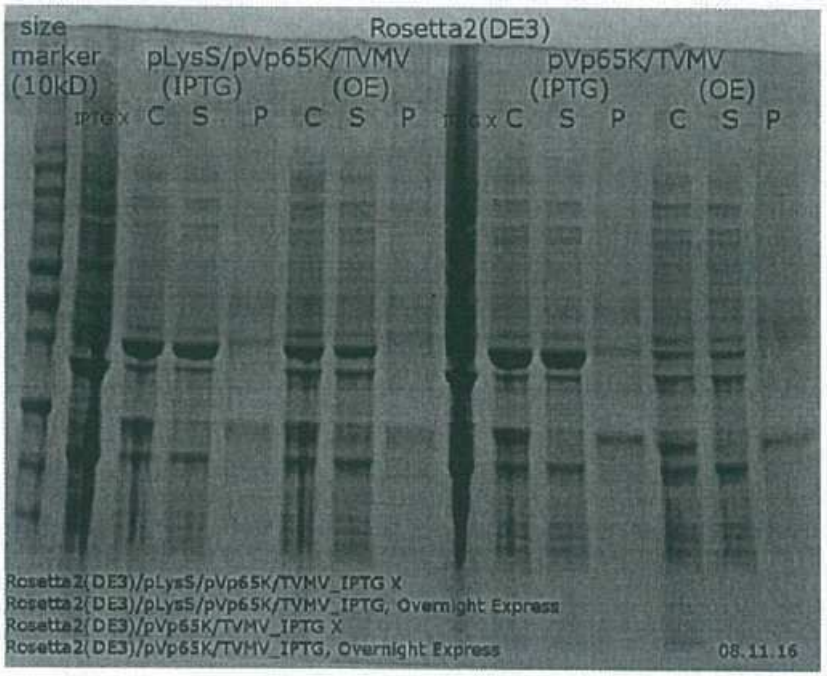
[0095] 본 발명의 TVMV 프로테아제 대량 생산용 벡터 및 이를 이용한 TVMV 프로테아제 생산 방법은 다량의 TVMV 프로테아제를 생산할 수 있으며, 말토오스 결합 단백질 등 수용성 향상 태그와 결합되어 발현됨으로써 수용성 발현이 가능하며, 최적화된 농도의 이미다졸을 TVMV 프로테아제 생산 과정 중 처리함에 따라 활성 저해를 최소화시킬 수 있다.

도면

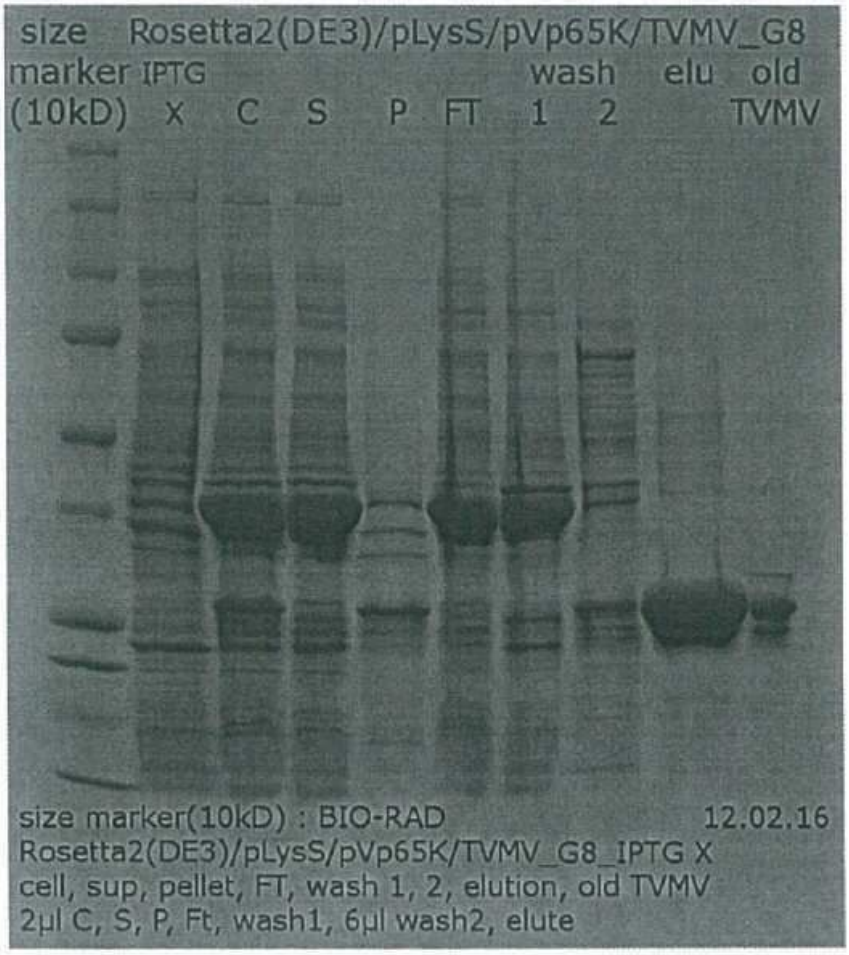
도면1



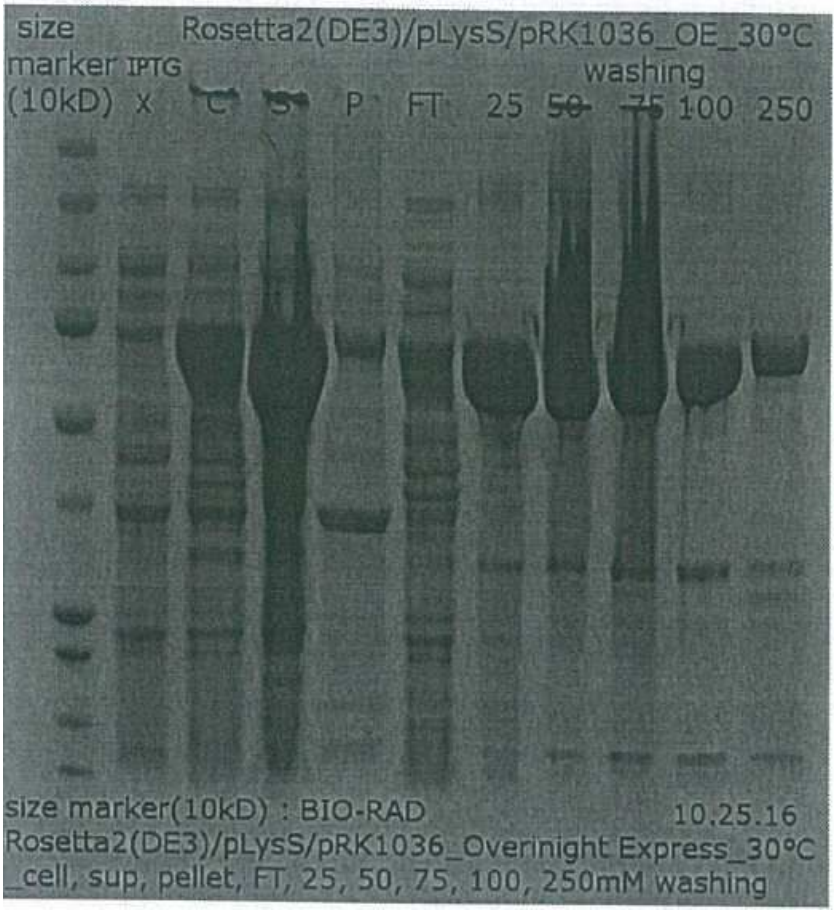
도면2


size
marker
(10kD)
Rosetta2(DE3)
pLysS/pVp65K/TVMV
(IPTG)
(OE)
pVp65K/TVMV
(IPTG)
(OE)
IPTG X C S P C S P IPTG X C S P C S P
Rosetta2(DE3)/pLysS/pVp65K/TVMV_IPTG X
Rosetta2(DE3)/pLysS/pVp65K/TVMV_IPTG, Overnight Express
Rosetta2(DE3)/pVp65K/TVMV_IPTG X
Rosetta2(DE3)/pVp65K/TVMV_IPTG, Overnight Express
08.11.16

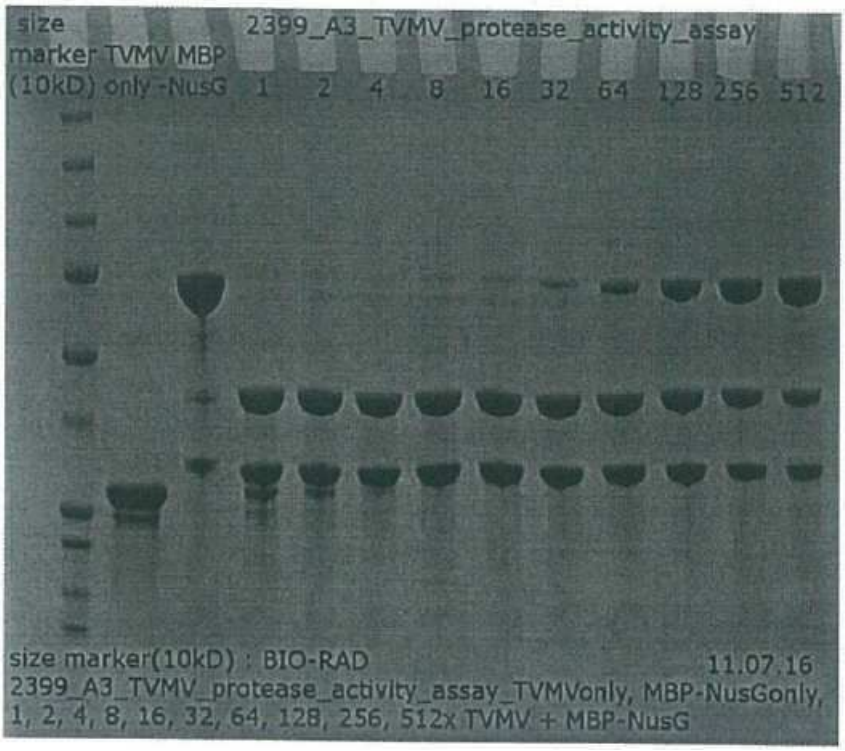
도면3


size
marker
(10kD)
Rosetta2(DE3)/pLysS/pVp65K/TVMV_G8
IPTG
X C S P FT
wash
1 2
elu
old
TVMV
size marker(10kD) : BIO-RAD
12.02.16
Rosetta2(DE3)/pLysS/pVp65K/TVMV_G8_IPTG X
cell, sup, pellet, FT, wash 1, 2, elution, old TVMV
2μl C, S, P, Ft, wash1, 6μl wash2, elute

도면4


size marker (10kD)
IPTG
X
C
S
P
FT
25
50
75
100
250
Rosetta2(DE3)/pLysS/pRK1036_OE_30°C
washing
size marker(10kD) : BIO-RAD
10.25.16
Rosetta2(DE3)/pLysS/pRK1036_Overinight Express_30°C
_cell, sup, pellet, FT, 25, 50, 75, 100, 250mM washing

도면5


size marker (10kD)
TVMV only
MBP -NusG
1
2
4
8
16
32
64
128
256
512
2399_A3_TVMV_protease_activity_assay
size marker(10kD) : BIO-RAD
11.07.16
2399_A3_TVMV_protease_activity_assay_TVMVonly, MBP-NusGonly,
1, 2, 4, 8, 16, 32, 64, 128, 256, 512x TVMV + MBP-NusG

서 열 목 록

<110> INDUSTRY-ACADEMIA COOPERATION GROUP OF SEJONG UNIVERSITY

<120> Vector for mass production of TVMV protease and method for mass production of TVMV protease using the same

<130> 16P12065

<160> 9

<170> KoPatentIn 3.0

<210> 1

<211> 712

<212> DNA

<213> Artificial Sequence

<220><223> TVMV protease coding sequence

<400> 1

aatgtctaag gctttgctga agggcgtgcg cgattttaat ccgatctctg cttgcgtatg 60

cctgctggaa aactcctcgg atggtcatag tgaacgtctg tttggcattg gttttggccc 120

gtatatcatt gccaaccagc atctgtttcg tcgtaacaat ggcgaactga ccatcaaaac 180

catgcacggt gaattcaaag tcaaaaactc tacccagctg cagatgaaac cggttgaagg 240

ccgtgacatt atcgttatca aaatggctaa agacttcccg ccgttcccgc agaaactgaa 300

attccgtcag ccgaccatca aagatcgtgt gtgcatggtg tccaccaact ttcagcagaa 360

aagcgtgtcc agcctggtgt ctgaatcctc tcacattgtg cataaagaag acacttcttt 420

ctggcagcac tggatcacca ctaaagatgg ccagtgtggc agcccactag tttccatcat 480

tgatggcaac attctgggca tccacagcct gactcatacc accaacggta gcaactactt 540

cgtggaattt ccggaaaaat tcgtggcgac ttatctcgat gccgcggatg gttggtgcaa 600

aaactggaaa ttcaacgcgg ataaaatcag ctggggttcc tttaccctgg ttgaagatgc 660

gccggaagat gacttcatgg ccaaaaaaac tgttgccgcc atcatggact aa 712

<210> 2

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> sense primer

<400> 2

ggttgcgatc gccatgtcta aggctttgct gaag 34

<210> 3

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223

> antisense primer

<400> 3

gtgtgtttaa acttagtcca tgatggcggc aacagt 36

<210> 4

<211> 74

<212> DNA

<213> Artificial Sequence

<220><223> PLtetO-1 promoter

<400> 4

ggtcagtgcg tcctgctgat gtgctcagta tctctatcac tgatagggat gtcaatctct 60

atcactgata ggga 74

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> TVMV site

<400> 5

gaaaccgtgc gtttccagtc t 21

<210> 6

<211> 1104

<212> DNA

<213> Artificial Sequence

<220><223> maltose binding protein coding DNA

<400> 6

atgggtaaaa tcgaagaagg taaactggta atctggatta acggcgataa aggctataac 60

ggtctcgctg aagtcggtaa gaaattcgag aaagataccg gaattaaagt caccgttgag 120

catccggata aactggaaga gaaattccca caggttgcgg caactggcga tggccctgac 180

attatcttct gggcacacga ccgctttggt ggctacgctc aatctggcct gttggctgaa 240

atcaccccgg acaaagcgtt ccaggacaag ctgtatccgt ttacctggga tgcggtacgt 300

tacaacggca agctgattgc ttacccgatc gctgttgaag cgttatcgct gatttataac 360

aaagatctgc tgccgaaccc gccaaaaacc tgggaagaga tcccggcgct ggataaagaa 420

ctgaaagcga aaggtaagag cgcgctgatg ttcaacctgc aagaaccgta cttcacctgg 480

ccgctgattg ctgctgacgg gggttatgcg ttcaagtatg aaaacggcaa gtacgacatt 540

aaagacgtgg gcgtggataa cgctggcgcg aaagcgggtc tgaccttcct ggttgacctg 600

attaaaaaca aacacatgaa tgcagacacc gattactcca tcgcagaagc tgcctttaat 660

aaaggcgaaa cagcgatgac catcaacggc ccgtgggcat ggtccaacat cgacaccagc 720

aaagtgaatt atggtgtaac ggtactgccg accttcaagg gtcaaccatc caaaccgttc 780

gttggcgtgc tgagcgcagg tattaacgcc gccagtccga acaaagagct ggcaaaagag 840

ttcctcgaaa actatctgct gactgatgaa ggtctggaag cggttaataa agacaaaccg 900

ctgggtgccg tagcgctgaa gtcttacgag gaagagttgg cgaaagatcc acgtattgcc 960

gccactatgg aaaacgccca gaaaggtgaa atcatgccga acatcccgca gatgtccgct 1020

ttctggtatg ccgtgcgtac tgcggtgatc aacgccgcca gcggtcgtca gactgtcgat 1080

gaagccctga aagacgcgca gact 1104

<210> 7

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> 8xHis tag coding DNA

<400> 7

catcaccatc atcaccatca ccat 24

<210> 8

<211> 6761

<212> DNA

<213> Artificial Sequence

<220><223> pVP65K/TVMV vector

<400> 8

cgtacgtcta gggcggcgga tttgtcctac tcaggagagc gttcaccgac aaacaacaga 60

taaaacgaaa ggcccagtct ttcgactgag cctttcgttt tatttgatgc ctctagcacg 120

cgtatcattg gatctttatt agtccatgat ggcggcaaca gttttttttgg ccatgaagtc 180

atcttccggc gcatcttcaa ccagggtaaa ggaaccccag ctgattttat ccgcgttgaa 240

tttccagttt ttgcaccaac catccgcggc atcgagataa gtcgccacga atttttccgg 300

aaattccacg aagtagttgc taccgttggt ggtatgagtc aggctgtgga tgcccagaat 360

gttgccatca atgatggaaa ctagtgggct gccacactgg ccatctttag tggtgatcca 420

gtgctgccag aaagaagtgt cttctttatg cacaatgtga gaggattcag acaccaggct 480

ggacacgctt ttctgctgaa agttggtgga caccatgcac acacgatctt tgatggtcgg 540

ctgacggaat ttcagtttct gcgggaacgg cgggaagtct ttagccattt tgataacgat 600

aatgtcacgg ccttcaaccg gtttcatctg cagctgggta gagtttttga ctttgaattc 660

accgtgcatg gttttgatgg tcagttcgcc attgttacga cgaaacagat gctggttggc 720

aatgatatac gggccaaaac caatgccaaa cagacgttca ctatgaccat ccgaggagtt 780

ttccagcagg catacgcaag cagagatcgg attaaaatcg cgcacgccct tcagcaaagc 840

cttagacata tcttacctcc ttaaatcaat tcggtcagtg cgtcctgctg atgtgctcag 900

tatctctatc actgataggg atgtcaatct ctatcactga tagggactcg agaaatcata 960

aaaaatttat ttgctttgtg agcggataac aattataata gattcaattg tgagcggata 1020

acaatttcac acagaattca ttaaagagga gaaattaacc atgggtaaaa tcgaagaagg 1080

taaactggta atctggatta acggcgataa aggctataac ggtctcgctg aagtcggtaa 1140

gaaattcgag aaagataccg gaattaaagt caccgttgag catccggata aactggaaga 1200

gaaattccca caggttgcgg caactggcga tggccctgac attatcttct gggcacacga 1260

ccgctttggt ggctacgctc aatctggcct gttggctgaa atcaccccgg acaaagcgtt 1320

ccaggacaag ctgtatccgt ttacctggga tgcggtacgt tacaacggca agctgattgc 1380

ttacccgatc gctgttgaag cgttatcgct gatttataac aaagatctgc tgccgaaccc 1440

gccaaaaacc tgggaagaga tcccggcgct ggataaagaa ctgaaagcga aaggtaagag 1500

cgcgctgatg ttcaacctgc aagaaccgta cttcacctgg ccgctgattg ctgctgacgg 1560

gggttatgcg ttcaagtatg aaaacggcaa gtacgacatt aaagacgtgg gcgtggataa 1620

cgctggcgcg aaagcgggtc tgaccttcct ggttgacctg attaaaaaca aacacatgaa 1680

tgcagacacc gattactcca tcgcagaagc tgcctttaat aaaggcgaaa cagcgatgac 1740

catcaacggc ccgtgggcat ggtccaacat cgacaccagc aaagtgaatt atggtgtaac 1800

ggtactgccg accttcaagg gtcaaccatc caaaccgttc gttggcgtgc tgagcgcagg 1860

tattaacgcc gccagtccga acaaagagct ggcaaaagag ttcctcgaaa actatctgct 1920

gactgatgaa ggtctggaag cggttaataa agacaaaccg ctgggtgccg tagcgctgaa 1980

gtcttacgag gaagagttgg cgaaagatcc acgtattgcc gccactatgg aaaacgccca 2040

gaaaggtgaa atcatgccga acatcccgca gatgtccgct ttctggtatg ccgtgcgtac 2100

tgcggtgatc aacgccgcca gcggtcgtca gactgtcgat gaagccctga aagacgcgca 2160

gactttaatt aacggcgacg gtgccgggga aaccgtgcgt ttccagtctc atcaccatca 2220

tcaccatcac catgcatccg cgatcgcaat gtctaaggct ttgctgaagg gcgtgcgcga 2280

ttttaatccg atctctgctt gcgtatgcct gctggaaaac tcctcggatg gtcatagtga 2340

acgtctgttt ggcattggtt ttggcccgta tatcattgcc aaccagcatc tgtttcgtcg 2400

taacaatggc gaactgacca tcaaaaccat gcacggtgaa ttcaaagtca aaaactctac 2460

ccagctgcag atgaaaccgg ttgaaggccg tgacattatc gttatcaaaa tggctaaaga 2520

cttcccgccg ttcccgcaga aactgaaatt ccgtcagccg accatcaaag atcgtgtgtg 2580

catggtgtcc accaactttc agcagaaaag cgtgtccagc ctggtgtctg aatcctctca 2640

cattgtgcat aaagaagaca cttctttctg gcagcactgg atcaccacta aagatggcca 2700

gtgtggcagc ccactagttt ccatcattga tggcaacatt ctgggcatcc acagcctgac 2760

tcataccacc aacggtagca actacttcgt ggaatttccg gaaaaattcg tggcgactta 2820

tctcgatgcc gcggatggtt ggtgcaaaaa ctggaaattc aacgcggata aaatcagctg 2880

gggttccttt accctggttg aagatgcgcc ggaagatgac ttcatggcca aaaaaactgt 2940

tgccgccatc atggactaag tttaaacgaa ttcgagctcg gtacccgggg atcctctaga 3000

gtcgacctgc aggcatgcaa gctgatccgg ctgctaacaa agcccgaaag gaagctgagt 3060

tggctgctgc caccgctgag caataactag caaactcgtt tctcgttcag ctttcttgta 3120

caaagtggtg atggcgcgcc tgtaggacgt cgacggtacc atcgatacgc gttcgaagct 3180

tcgcctgggg taatgactct ctagcttgag gcatcaaata aaacgaaagg ctcagtcgaa 3240

agactgggcc tttcgtttta tctgttgttt gtcggtgaac gctctcctga gtaggacaaa 3300

tccgccctct agattacgtg cagtcgatga taagctgtca aacatgagaa ttgtgcctaa 3360

tgagtgagct aacttacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac 3420

ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 3480

gggcgccagg gtggtttttc ttttcaccag tgagacgggc aacagctgat tgcccttcac 3540

cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca gcaggcgaaa 3600

atcctgtttg atggtggtta acggcgggat ataacatgag ctgtcttcgg tatcgtcgta 3660

tcccactacc gagatatccg caccaacgcg cagcccggac tcggtaatgg cgcgcattgc 3720

gcccagcgcc atctgatcgt tggcaaccag catcgcagtg ggaacgatgc cctcattcag 3780

catttgcatg gtttgttgaa aaccggacat ggcactccag tcgccttccc gttccgctat 3840

cggctgaatt tgattgcgag tgagatattt atgccagcca gccagacgca gacgcgccga 3900

gacagaactt aatgggcccg ctaacagcgc gatttgctgg tgacccaatg cgaccagatg 3960

ctccacgccc agtcgcgtac cgtcttcatg ggagaaaata tactgttga tgggtgtctg 4020

gtcagagaca tcaagaaata acgccggaac attagtgcag gcagcttcca cagcaatggc 4080

atcctggtca tccagcggat agttaatgat cagcccactg acgcgttgcg cgagaagatt 4140

gtgcaccgcc gctttacagg cttcgacgcc gcttcgttct accatcgaca ccaccacgct 4200

ggcacccagt tgatcggcgc gagatttaat cgccgcgaca atttgcgacg gcgcgtgcag 4260

ggccagactg gaggtggcaa cgccaatcag caacgactgt ttgcccgcca gttgttgtgc 4320

cacgcggttg ggaatgtaat tcagctccgc catcgccgct tccacttttt cccgcgtttt 4380

cgcagaaacg tggctggcct ggttcaccac gcgggaaacg gtctgataag agacaccggc 4440

atactctgcg acatcgtata acgttactgg tttcacattc accaccctga attgactctc 4500

ttccgggcgc tatcatgcca taccgcgaaa ggttttgcgc cattcgatgg tgtcggaacc 4560

tagagctgcc tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg 4620

gagacggtca cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg 4680

tcagcgggtg ttggcgggtg tcggggcgca gccatgaccc agtcacgtag cgatagcgga 4740

gtgtatactg gcttaactat gcggcatcag agcagattgt actgagagtg caccatatgc 4800

ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggcgc tcttccgctt 4860

cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact 4920

caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag 4980

caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 5040

ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 5100

cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg 5160

ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc 5220

tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg 5280

gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc 5340

ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga 5400

ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg 5460

gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 5520

aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg 5580

tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 5640

ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat 5700

tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 5760

aaagtatata tgagtaaact tggtctgaca gccctaggtc agaagaactc gtcaagaagg 5820

cgatagaagg cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg 5880

tcagcccatt cgccgccaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 5940

tagcggtccg ccacacccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc 6000

accatgatat tcggcaagca ggcatcgcca tgagtcacga cgagatcctc accatccggc 6060

atacgcgcct tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc 6120

agatcatcct gatcgacaag accggcttcc atccgagtgc gtgctcgctc gatgcgatgt 6180

ttcgcttggt ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca 6240

tcagccatga tggatacttt ctcggcagga gcaaggtgag atgacaggag atcctgcccc 6300

ggcacttcgc ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct 6360

gcgcaaggaa cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca 6420

ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc 6480

cggaacacgg cggcatcaga gcagccgatt gtctgttgtg cccagtcata gccgaatagc 6540

ctctccaccc aagcggccgg agaacctgcg tgcaatccat cttgttcaag catgcgaaac 6600

gaccgtcatc ctgtctcttg atcagatctt gatcccctgc gccatcagat ccttggcggc 6660

aagaaagcca tccagtttac tttgcagggc ttcccaacct taccagaggg cgccccagct 6720

ggcaattctt ttgaagctca cgctgccgca agcactcagg g 6761

<210> 9

<211> 7174

<212> DNA

<213> Artificial Sequence

<220><223> pVP65K vector

<400> 9

cgtacgtcta gggcggcgga tttgtcctac tcaggagagc gttcaccgac aaacaacaga 60

taaaacgaaa ggcccagtct ttcgactgag cctttcgttt tatttgatgc ctctagcacg 120

cgtatcattg gatctttatt agtccatgat ggcggcaaca gtttttttgg ccatgaagtc 180

atcttccggc gcatcttcaa ccagggtaaa ggaaccccag ctgattttat ccgcgttgaa 240

tttccagttt ttgcaccaac catccgcggc atcgagataa gtcgccacga atttttccgg 300

aaattccacg aagtagttgc taccgttggt ggtatgagtc aggctgtgga tgcccagaat 360

gttgccatca atgatggaaa ctagtgggct gccacactgg ccatctttag tggtgatcca 420

gtgctgccag aaagaagtgt cttctttatg cacaatgtga gaggattcag acaccaggct 480

ggacacgctt ttctgctgaa agttggtgga caccatgcac acacgatctt tgatggtcgg 540

ctgacggaat ttcagtttct gcgggaacgg cgggaagtct ttagccattt tgataacgat 600

aatgtcacgg ccttcaaccg gtttcatctg cagctgggta gagtttttga ctttgaattc 660

accgtgcatg gttttgatgg tcagttcgcc attgttacga cgaaacagat gctggttggc 720

aatgatatac gggccaaaac caatgccaaa cagacgttca ctatgaccat ccgaggagtt 780

ttccagcagg catacgcaag cagagatcgg attaaaatcg cgcacgccct tcagcaaagc 840

cttagacata tcttacctcc ttaaatcaat tcggtcagtg cgtcctgctg atgtgctcag 900

tatctctatc actgataggg atgtcaatct ctatcactga tagggactcg agaaatcata 960

aaaaatttat ttgctttgtg agcggataac aattataata gattcaattg tgagcggata 1020

acaatttcac acagaattca ttaaagagga gaaattaacc atgggtaaaa tcgaagaagg 1080

taaactggta atctggatta acggcgataa aggctataac ggtctcgctg aagtcggtaa 1140

gaaattcgag aaagataccg gaattaaagt caccgttgag catccggata aactggaaga 1200

gaaattccca caggttgcgg caactggcga tggccctgac attatcttct gggcacacga 1260

ccgctttggt ggctacgctc aatctggcct gttggctgaa atcaccccgg acaaagcgtt 1320

ccaggacaag ctgtatccgt ttacctggga tgcggtacgt tacaacggca agctgattgc 1380

ttacccgatc gctgttgaag cgttatcgct gatttataac aaagatctgc tgccgaaccc 1440

gccaaaaacc tgggaagaga tcccggcgct ggataaagaa ctgaaagcga aaggtaagag 1500

cgcgctgatg ttcaacctgc aagaaccgta cttcacctgg ccgctgattg ctgctgacgg 1560

gggttatgcg ttcaagtatg aaaacggcaa gtacgacatt aaagacgtgg gcgtggataa 1620

cgctggcgcg aaagcgggtc tgaccttcct ggttgacctg attaaaaaca aacacatgaa 1680

tgcagacacc gattactcca tcgcagaagc tgcctttaat aaaggcgaaa cagcgatgac 1740

catcaacggc ccgtgggcat ggtccaacat cgacaccagc aaagtgaatt atggtgtaac 1800

ggtactgccg accttcaagg gtcaaccatc caaaccgttc gttggcgtgc tgagcgcagg 1860

tattaacgcc gccagtccga acaaagagct ggcaaaagag ttcctcgaaa actatctgct 1920

gactgatgaa ggtctggaag cggttaataa agacaaaccg ctgggtgccg tagcgctgaa 1980

gtcttacgag gaagagttgg cgaaagatcc acgtattgcc gccactatgg aaaacgccca 2040

gaaaggtgaa atcatgccga acatcccgca gatgtccgct ttctggtatg ccgtgcgtac 2100

tgcggtgatc aacgccgcca gcggtcgtca gactgtcgat gaagccctga aagacgcgca 2160

gactttaatt aacggcgacg gtgccgggga aaccgtgcgt ttccagtctc atcaccatca 2220

tcaccatcac catgcatccg cgatcgcggc cgcgttggaa taagtaaagg aatcacatgg 2280

cacaggttat caacacgttt gacggggttg cggattatct tcagacatat cataagctac 2340

ctgataatta cattacaaaa tcagaagcac aagccctcgg ctgggtggca tcaaaaggga 2400

accttgcaga cgtcgctccg gggaaaagca tcggcggaga catcttctca aacaaggaag 2460

gcaaactccc gggcaaaagc ggacgaacat ggcgtgaagc ggatattaac tatacatcag 2520

gcttcagaaa ttcagaccgg attctttact caagcgactg gctgatttac aaaacaacgg 2580

accattatca gacctttaca aaaatcagat aattaggcac cccaggcttt acactttatg 2640

ctttcggctc gtataatgtg tggattttga gttaggatcc gtcgagattt tcaggagcta 2700

aggaagctaa aatggagaaa aaaatcactg gatataccac cgttgatata tcccaatggc 2760

atcgtaaaga acattttgag gcatttcagt cagttgctca atgtacctat aaccagaccg 2820

ttcagctgga tattacggcc tttttaaaga ccgtaaagaa aaataagcac aagttttatc 2880

cggcctttat tcacattctt gcccgcctga tgaatgctca tccggaattc cgtatggcaa 2940

tgaaagacgg tgagctggtg atatgggata gtgttcaccc ttgttacacc gttttccatg 3000

agcaaactga aacgttttca tcgctctgga gtgaatacca cgacgatttc cggcagtttc 3060

tacacatata ttcgcaagat gtggcgtgtt acggtgaaaa cctggcctat ttccctaaag 3120

ggtttattga gaatatgttt ttcgtctcag ccaatccctg ggtgagtttc accagttttg 3180

atttaaacgt ggccaatatg gacaacttct tcgcccccgt tttcacgatg ggcaaatatt 3240

atacgcaagg cgacaaggtg ctgatgccgc tggcgattca ggttcatcat gccgtttgtg 3300

atggcttcca tgtcggcaga atgcttaatg aattacaaca gtactgcgat gagtggcagg 3360

gcggggcgta atgtttaaac gaattcgagc tcggtacccg gggatcctct agagtcgacc 3420

tgcaggcatg caagctgatc cggctgctaa caaagcccga aaggaagctg agttggctgc 3480

tgccaccgct gagcaataac tagcaaactc gtttctcgtt cagctttctt gtacaaagtg 3540

gtgatggcgc gcctgtagga cgtcgacggt accatcgata cgcgttcgaa gcttcgcctg 3600

gggtaatgac tctctagctt gaggcatcaa ataaaacgaa aggctcagtc gaaagactgg 3660

gcctttcgtt ttatctgttg tttgtcggtg aacgctctcc tgagtaggac aaatccgccc 3720

tctagattac gtgcagtcga tgataagctg tcaaacatga gaattgtgcc taatgagtga 3780

gctaacttac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt 3840

gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgcc 3900

agggtggttt ttcttttcac cagtgagacg ggcaacagct gattgccctt caccgcctgg 3960

ccctgagaga gttgcagcaa gcggtccacg ctggtttgcc ccagcaggcg aaaatcctgt 4020

ttgatggtgg ttaacggcgg gatataacat gagctgtctt cggtatcgtc gtatcccact 4080

accgagatat ccgcaccaac gcgcagcccg gactcggtaa tggcgcgcat tgcgcccagc 4140

gccatctgat cgttggcaac cagcatcgca gtgggaacga tgccctcatt cagcatttgc 4200

atggtttgtt gaaaaccgga catggcactc cagtcgcctt cccgttccgc tatcggctga 4260

atttgattgc gagtgagata tttatgccag ccagccagac gcagacgcgc cgagacagaa 4320

cttaatgggc ccgctaacag cgcgatttgc tggtgaccca atgcgaccag atgctccacg 4380

cccagtcgcg taccgtcttc atgggagaaa ataatactgt tgatgggtgt ctggtcagag 4440

acatcaagaa ataacgccgg aacattagtg caggcagctt ccacagcaat ggcatcctgg 4500

tcatccagcg gatagttaat gatcagccca ctgacgcgtt gcgcgagaag attgtgcacc 4560

gccgctttac aggcttcgac gccgcttcgt tctaccatcg acaccaccac gctggcaccc 4620

agttgatcgg cgcgagattt aatcgccgcg acaatttgcg acggcgcgtg cagggccaga 4680

ctggaggtgg caacgccaat cagcaacgac tgtttgcccg ccagttgttg tgccacgcgg 4740

ttgggaatgt aattcagctc cgccatcgcc gcttccactt tttcccgcgt tttcgcagaa 4800

acgtggctgg cctggttcac cacgcgggaa acggtctgat aagagacacc ggcatactct 4860

gcgacatcgt ataacgttac tggtttcaca ttcaccaccc tgaattgact ctcttccggg 4920

cgctatcatg ccataccgcg aaaggttttg cgccattcga tggtgtcgga acctagagct 4980

gcctcgcgcg tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg 5040

tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg 5100

gtgttggcgg gtgtcggggc gcagccatga cccagtcacg tagcgatagc ggagtgtata 5160

ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga 5220

aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg cttcctcgct 5280

cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 5340

ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 5400

ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 5460

ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 5520

actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 5580

cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 5640

tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 5700

gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 5760

caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 5820

agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 5880

tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 5940

tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 6000

gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 6060

gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 6120

aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 6180

atatgagtaa acttggtctg acagccctag gtcagaagaa ctcgtcaaga aggcgataga 6240

aggcgatgcg ctgcgaatcg ggagcggcga taccgtaaag cacgaggaag cggtcagccc 6300

attcgccgcc aagctcttca gcaatatcac gggtagccaa cgctatgtcc tgatagcggt 6360

ccgccacacc cagccggcca cagtcgatga atccagaaaa gcggccattt tccaccatga 6420

tattcggcaa gcaggcatcg ccatgagtca cgacgagatc ctcaccatcc ggcatacgcg 6480

ccttgagcct ggcgaacagt tcggctggcg cgagcccctg atgctcttcg tccagatcat 6540

cctgatcgac aagaccggct tccatccgag tgcgtgctcg ctcgatgcga tgtttcgctt 6600

ggtggtcgaa tgggcaggta gccggatcaa gcgtatgcag ccgccgcatt gcatcagcca 6660

tgatggatac tttctcggca ggagcaaggt gagatgacag gagatcctgc cccggcactt 6720

cgcccaatag cagccagtcc cttcccgctt cagtgacaac gtcgagcaca gctgcgcaag 6780

gaacgcccgt cgtggccagc cacgatagcc gcgctgcctc gtcctgcagt tcattcaggg 6840

caccggacag gtcggtcttg acaaaaagaa ccgggcgccc ctgcgctgac agccggaaca 6900

cggcggcatc agagcagccg attgtctgtt gtgcccagtc atagccgaat agcctctcca 6960

cccaagcggc cggagaacct gcgtgcaatc catcttgttc aagcatgcga aacgaccgtc 7020

atcctgtctc ttgatcagat cttgatcccc tgcgccatca gatccttggc ggcaagaaag 7080

ccatccagtt tactttgcag ggcttcccaa ccttaccaga gggcgcccca gctggcaatt 7140

cttttgaagc tcacgctgcc gcaagcactc aggg 7174