



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0106911
(43) 공개일자 2012년09월27일

(51) 국제특허분류(Int. Cl.)

C12N 15/63 (2006.01) C12N 15/70 (2006.01)

C12N 15/31 (2006.01)

(21) 출원번호 10-2011-0023509

(22) 출원일자 2011년03월16일

심사청구일자 2011년03월16일

(71) 출원인

고려대학교 산학협력단

서울 성북구 안암동5가 1

(72) 발명자

최인걸

서울특별시 동작구 남부순환로263길 34 (사당동)

김경현

경기 성남시 분당구 삼평동 봇들마을 907동 170 4호

고혁진

서울특별시 성북구 길음로13길 39, 대우푸르지오 아파트 306동 1402호 (길음동)

(74) 대리인

특허법인다나

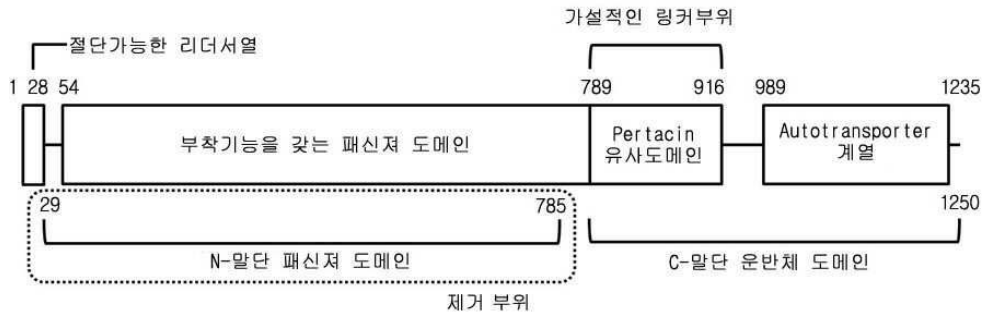
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 오토트랜스포터를 이용한 세포 표면 발현 시스템의 구축

(57) 요약

본 발명은 오토트랜스포터 단백질을 이용한 세포표면 발현 시스템의 구축에 관한 발명이다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 2010-0029734

부처명 한국연구재단

연구사업명 (원천)글로벌프론티어

연구과제명 [공동]화학 및 단백질촉매 융합시스템을 통한 생물친화형 목질계 바이오매스의 탈구조화
기술개발

주관기관 고려대학교 산학협력단

연구기간 2010.10.22 ~ 2011.08.21

특허청구의 범위

청구항 1

신호 서열, 패신저 도메인, 링커 부위 및 오토트랜스포터 슈퍼패밀리 부위로 구성된 박테리아의 오토트랜스포터 단백질로부터 패신저 도메인 부위가 제거된 클로닝 부위를 가지는 표면 발현 벡터.

청구항 2

제 1항에 있어서, 상기 벡터는 클로닝 부위로 TEV 프로테아제 절단 부위가 삽입된 벡터.

청구항 3

제 1항에 있어서, 상기 박테리아의 오토트랜스포터 단백질은 YfaL(NP_416736.1), 유전자은행 등록번호 NP_416736.1, CBJ01870.1, ZP_07247471.1, YP_001459034.1, ZP_07121532.1, ZP_02999058.1, YP_003045352.1, EGB57147.1, EGB37260.1, ZP_07146751.1, ZP_07498247.1, ZP_07787684.1, YP_001724408.1, YP_003035675.1, ZP_03068365.1, ZP_07523028.1, ZP_07102820.1, EGB90124.1, YP_002403506.1, EFZ69526.1, YP_002293769.1, ZP_07590844.1, ZP_06662998.1, ZP_07135193.1, YP_001463580.1, ZP_07687541.1, ZP_03028301.1, BAI31473.1, ZP_03043348.1, ZP_07098380.1, YP_002387709.1, NP_288807.1, EFW74417.1, BAI26434.1, ZP_07142137.1, BAI36814.1, YP_003078961.1, ZP_07614097.1, EFZ47577.1, ZP_06658169.1, ZP_07186279.1, EFX25099.1, ZP_06654172.1, ZP_02786344.1, EGC11783.1, EGB41107.1, ZP_02812144.1, ZP_02823402.1, ZP_07508510.1, ZP_02774656.1, YP_002413282.1, EFX30437.1, ZP_07619499.1, ZP_07518131.1, YP_001744429.1, YP_002408332.1, YP_002329881.1, NP_754661.1, YP_002398608.1, EGB75831.1, EGB51754.1, YP_541512.1, ZP_07175432.1, YP_670171.1, CAP76735.1, ZP_07624785.1, ZP_07503047.1, EFZ73335.1, ZP_07781748.1, ADN70533.1, ZP_03031601.1, ADR27681.1, ZP_07448816.1, ZP_02901808.1, YP_002382104.1, EGC06746.1, ZP_05431470.1 및 YP_004212039.1로 구성된 군으로부터 선택된 하나의 단백질을 특징으로 하는 표면 발현 벡터.

청구항 4

제 1항 또는 제2항에 있어서, 상기 벡터는 도 9의 개열 지도를 가지는 벡터.

청구항 5

제 1항 또는 제3항에 있어서, 상기 오토트랜스포터 단백질은 서열번호 1의 아미노산 서열을 가지는 것을 특징으로 하는 벡터.

청구항 6

제 1항에 있어서, 상기 패신저 도메인 부위는 서열번호 1의 아미노산 중 29번째 잔기부터 785번째 잔기의 서열을 가지는 것을 특징으로 하는 벡터.

청구항 7

제 1항 내지 제3항 중 어느 한 항의 벡터에 관심 있는 단백질을 코딩하는 유전자를 삽입하여 그 단백질을 세포 표면에 발현하는 방법.

명세서

기술분야

본 발명은 오토트랜스포터 단백질을 이용한 세포표면 발현 시스템의 구축에 관한 발명이다.

배경기술

박테리아 세포의 표면에서 관심있는 단백질 또는 이중 펩타이드의 발현은 많은 생명공학적인 적용에 이용되고 있다(Daugherty P.S. (2007) *Curr Opin Struct Biol*, **17**, 474-480, Lee S.Y., Choi J.H. and Xu Z. (2003) *Trends Biotechnol*, **21**, 45-52). 비록 세포 외부에 노출된 많은 단백질들이 표면 발현 시스템을 위한 스카폴드로 채택될 수 있었지만, 외막 단백질(Nakajima H., Shimbara N., Shimonishi Y., Mimori T., Niwa S. and

Saya H. (2000) *Gene*, **260**, 121-131), 리포프로틴(Dhillon J.K., Drew P.D. and Porter A.J. (1999) *Lett Appl Microbiol*, **28**, 350-354), 표면 부착기의 서브유닛(appendages) (Kjaergaard K., Sorensen J.K., Schembri M.A. and Klemm P. (2000) *Appl Environ Microbiol*, **66**, 10-14), 및 S-layer 단백질(Bingle W.H., Nomellini J.F. and Smit J. (1997) *Mol Microbiol*, **26**, 277-288)과 같은 일부 유전자들만이 세포의 표면에 관심 단백질을 발현하는 스카폴드 단백질로 모집되었다.

[0003] *E. coli* 표면 발현 시스템을 개발하기 위하여, 본 발명자들은 처음에는 가설 단백질로 해석되었다가, 추후에 *E. coli* 0126:H27의 AIDA-I의 운반체(translocator) 지역에 대한 서열 유사성에 근거하여 AIDA-I 계열로 예측된 yfaL 유전자(Laarmann S. and Schmidt M.A. (2003) *Microbiology*, **149**, 1871-1882;Yen M.R., Peabody C.R., Partovi S.M., Zhai Y., Tseng Y.H. and Saier M.H. (2002) *Biochim Biophys Acta*, **1562**, 6-31)에 주목하였다.

발명의 내용

해결하려는 과제

[0004] 본 발명은 상기의 필요성에 의하여 안출된 것으로서 본 발명의 목적은 미생물의 표면 발현 시스템을 제공하는 것이다.

과제의 해결 수단

[0005] 상기의 목적을 달성하기 위하여 본 발명은 신호 서열, 패신저 도메인, 링커 부위, 오토트랜스포터 슈퍼패밀리 부위로 구성된 박테리아의 오토트랜스포터 단백질에서 패신저 도메인 부위가 제거된 클로닝 부위를 가지는 표면 발현 벡터를 제공한다.

[0006] 본 발명에서 표면 발현 벡터란 목표 단백질이 박테리아의 세포막(내막) 및 외막을 통과하여 안정적으로 세포 표면에 고정되고 고유의 기능을 수행할 수 있게 하는 스카폴드 단백질을 제공하고, 그 스카폴드 단백질에 목표 단백질을 융합시킬 수 있는 클로닝 부위를 갖는 벡터를 의미한다.

[0007] 본 발명의 일 구현 예에 있어서, 상기 벡터는 클로닝 부위에 적절한 시점에서 목표 단백질의 방출을 가능하게 하기 위해서 신호서열(1-28번 아미노산) 뒤의 클로닝 부위인 SmaI(CCCGGG) 다음에 TEV 프로테아제 절단 부위(ENLYFQ(GAAAACCTTTACTTCCAG))가 삽입된 것이 바람직하나 이에 한정되지 아니한다.

[0008] 본 발명의 다른 구현 예에 있어서, 상기 박테리아의 오토트랜스포터 단백질은 YfaL(NP_416736.1), 유전자는 등록번호 NP_416736.1, CBJ01870.1, ZP_07247471.1, YP_001459034.1, ZP_07121532.1, ZP_02999058.1, YP_003045352.1, EGB57147.1, EGB37260.1, ZP_07146751.1, ZP_07498247.1, ZP_07787684.1, YP_001724408.1, YP_003035675.1, ZP_03068365.1, ZP_07523028.1, ZP_07102820.1, EGB90124.1, YP_002403506.1, EFZ69526.1, YP_002293769.1, ZP_07590844.1, ZP_06662998.1, ZP_07135193.1, YP_001463580.1, ZP_07687541.1, ZP_03028301.1, BAI31473.1, ZP_03043348.1, ZP_07098380.1, YP_002387709.1, NP_288807.1, EFW74417.1, BAI26434.1, ZP_07142137.1, BAI36814.1, YP_003078961.1, ZP_07614097.1, EFZ47577.1, ZP_06658169.1, ZP_07186279.1, EFX25099.1, ZP_06654172.1, ZP_02786344.1, EGC11783.1, EGB41107.1, ZP_02812144.1, ZP_02823402.1, ZP_07508510.1, ZP_02774656.1, YP_002413282.1, EFX30437.1, ZP_07619499.1, ZP_07518131.1, YP_001744429.1, YP_002408332.1, YP_002329881.1, NP_754661.1, YP_002398608.1, EGB75831.1, EGB51754.1, YP_541512.1, ZP_07175432.1, YP_670171.1, CAP76735.1, ZP_07624785.1, ZP_07503047.1, EFZ73335.1, ZP_07781748.1, ADN70533.1, ZP_03031601.1, ADR27681.1, ZP_07448816.1, ZP_02901808.1, YP_002382104.1, EGC06746.1, ZP_05431470.1 및 YP_004212039.1로 구성된 군으로부터 선택된 하나의 단백질인 것이 바람직하나 이에 한정되지 아니한다.

[0009] 본 발명의 또 다른 구현예에 있어서, 상기 벡터는 도 9의 개열 지도를 가지는 것이 바람직하나 이에 한정되지 아니한다.

[0010] 본 발명의 바람직한 구현예에 있어서, 상기 오토트랜스포터 단백질은 서열번호 1의 아미노산 서열을 가지는 것이 바람직하나 이에 한정되지 아니한다.

[0011] 본 발명의 바람직한 다른 구현 예에 있어서, 상기 패신저 도메인은 도 8에 목록화한 청구항 제3항에 열거한 단백질들의 패신저 도메인 부위이며, 서열번호 1의 아미노산 중 29번째 잔기부터 785번째 잔기의 서열을 가지

는 것이 바람직하나 이에 한정되지 아니한다.

[0012] 또한 본 발명은 상기 본 발명의 벡터에 관심 있는 단백질을 코딩하는 유전자를 삽입하여 세포 표면에 발현하는 방법을 제공한다.

[0013] 이하 본 발명을 설명한다.

[0014] 본 발명은 *yfaL* 유전자에 기초하여 대장균에서 신규한 표면 발현 시스템을 구축하고 고안하였다. 본 발명자들은 *YfaL*의 서열 분석에 의한 유전자내의 기능적 도메인의 구성을 예측하고, 세포 표면 발현에 필요하다고 추정되는 링커 부위(789-916)를 결정한 후 본래의 패신저 도메인(29-785)을 제거하고 N-말단과 링커부위만을 C-말단 운반체(translocator) 도메인에 융합하여서 자동수송(autotransporting) 시스템(pATLIC 벡터)를 구축하였다. 또한 이 시스템은 관심 있는 목표 단백질의 빠르고 쉬운 클로닝을 위하여 라이게이션 독립적인 클로닝(ligation independent cloning, LIC)에 기반하여 제작하였고 대장균 표면에 *Saccharophagus degradans* 2-40로부터 유래한 여섯 종류의 아가레이즈 효소의 활성발현시험을 통해 작동을 확인하였다.

[0015] 시험에 사용된 이들 아가레이즈들은 다양한 크기를 가지고 있고, 그 중 일부는 대장균의 원형질에서 성공적인 과발현이나 활성발현이 실패한 것으로 알려져 있으나, 자동수송시스템(pATLIC 벡터)에 융합하여 시험한 결과 모두 성공적으로 대장균의 표면에 발현되었고, 고유의 효소기능을 나타냄을 확인하였다. 결과적으로 본 발명에서 제작된 자동수송시스템(pATLIC 벡터)은 패신저 도메인으로 치환된 아가레이즈의 크기와는 크게 상관없이 (본 발명에서는 최대 143 kDa) 세포 표면에 성공적으로 노출되어 효소활성을 나타내었고, 특히 LIC 클로닝사이트에 TEV절단부위를 삽입함으로써 적절한 시점(또는 원하는 시점)에서 TEV 단백질분해효소에 의해 표면에 발현된 목표 단백질을 세포 밖으로 방출하는 것이 가능함을 확인시켜 주었다.

[0016] 이하 본 발명을 상세하게 설명한다.

[0017] *E. coli*의 *yfaL*에 기반한 표면발현 시스템의 고안

[0018] 본 발명자들은 목표 단백질이 패신저 도메인을 대체하여 표면 발현 시스템에서 융합될 때, 연결부위로 추정되는 링커구간(789-916)이 성공적인 표면발현에 필요하다는 것을 추정하였다. Pfam 도메인 데이터베이스에 의한 도메인 구조 분석 및 다른 자동수송체(예를 들어 PF03797)와 다중 서열 배열에 따라서, 본 발명자들은 *YfaL*에서 패신저 및 운반(translocation) 도메인을 연결하는 추정되는 링커 부위(789-916)를 예측하였다(도 1). 비록 링커 부위가 패신저 또는 운반 도메인에 속하지만, 이 부위가 목표 단백질의 외막을 통과하는 데 중요한 역할을 한다고 추정되기 때문에 본 발명자들은 발현 벡터 플라스미드가 구축을 위하여 링커부위를 남기기로 결정하였다.

[0019] 본 발명자들은 관심있는 단백질로 대체되는 패신저 도메인 지역으로 *YfaL*의 29-785 부위를 할당하였다.

[0020] 또한 링커 부위 결정에서, 본 발명자들은 표면 발현을 위한 세 종류의 벡터 후보를 고안하였다. 각각의 후보는 *YfaL*의 패신저 도메인(각각 29 아미노산부터 ¹⁾948, ²⁾785 및 ³⁾699 아미노산까지)을 제거한 후 형광단백질을 이들 벡터에 클로닝하였다. 결과적으로 *YfaL*의 29에서 948 아미노산이 제거된 후보의 형광은 다른 후보들의 것에 33% 정도이었고 다른 두 후보들의 형광은 비슷하였다. 따라서 본 발명자들은 *YfaL*의 패신저 도메인을 최대한으로 제거하면서 동시에 링커 부위로서 기능을 상실하지 않은 벡터 후보를 선택(*YfaL*의 29-785 아미노산이 제거된, pATLIC 벡터)하였다.

[0021] 표면 발현 시스템의 구축

[0022] 본 발명자들은 L-아라비노스-유도적인 프로모터 하에 조절되는 패신저 도메인이 제거된 *YfaL*을 포함하는 신규한 표면 발현 벡터를 구축하였다. 구축된 벡터 안의 *YfaL*은 예측된 시그널 펩타이드 (1-28 잔기), 링커 부분으로 추정되는 지역 (786-916 잔기) 및 운반체 지역 (917-1250 잔기)으로 구성되었다. 추가적으로, 여러 목표 유전자들을 클로닝할 때 가장 시간 소모적인 과정인 제한 효소의 사용 및 라이게이션 과정을 제거함으로써 보다 효율적인 클로닝을 제공하는 LIC 방법을 적용시켰으며, 이러한 LIC 서열은 신호서열과 운반체 지역 사이에 삽입하였다. 또한 본 발명자들은 또한 TEV 처리에 의하여 대장균의 세포 표면으로 발현된 단백질을 배지로 분비하기 위하여 LIC 클로닝 부위 뒤로 TEV 프로테이즈 절단 부위를 삽입하여 최종적으로 대장균에서 세포 표면 발현 및 분비를 위한 LIC 벡터를 고안하였다 (pATLIC 벡터) (도 2).

[0023] *E. coli*의 표면 상에 관심있는 단백질의 기능적인 발현

[0024] 본 발명자들은 SDS-PAGE에 의하여 pATLIC 벡터로 클로닝한 관심있는 유전자의 발현을 확인하였고 (도 3), 분비 신호서열의 확인을 위해서, 리포터 유전자로써 아가레이즈를 사용하여 세포 표면에서의 이중 단백질에 대한 기능적인 발현을 시도하였다. 아가는 대장균의 막을 통해서 확산이 불가능한 중합체이므로 세포 표면에 발현되거나 배지로 분비된 아가레이즈 만이 이러한 아가에 대한 분해 활성을 보일 수 있다. 대장균에서 고안된 표면 발현 시스템을 실험하기 위해서, 본 발명자들은 *S. degradans*로부터 유래한 다섯 종의 베타 아가레이즈 (Aga50A, Aga16B, Aga86C, Aga50D 및 Aga86E)를 pATLIC 벡터로 클로닝하였다. 대조군 실험으로 본 발명자들은 세포 파쇄로부터 효소의 누출에 의하여 야기되는 아가 분해 양상을 배제하기 위하여 원형질에서 Aga16B (pBAD-aga16B)를 과발현하는 균주를 사용하였다. 아가레이즈의 그 기능적인 발현은 아가 고체배지 상에서 스팟팅 분석에 의하여 증명하였다. 본 발명자들이 LB 아가(자가유도를 위하여 0.2% L(+)-아라비노스를 포함하는) 고체 배지 상에서 pATLIC-aga16B 또는 pBAD-aga16B를 가지는 하룻밤 배양시킨 50의 대장균 배양액을 위의 고체배지에 스팟팅 하였을 때, 오직 pATLIC-aga16B를 가지는 세포만이 아가 분해 양상을 나타내는 것을 요오드 염색을 통해 확인함으로써, 실제로 본 발명자들이 고안한 세포 표면 발현 시스템이 적절하게 작동하고 있는 것을 증명하였다 (도 4a). 본 발명자들은 또한 이전의 논문들에서 대장균에서 발현에 실패한 (Ekborg N.A., Taylor L.E., Longmire A.G., Henrissat B., Weiner R.M. and Hutcheson S.W. (2006) *Appl Environ Microbiol*, **72**, 3396-3405) Aga86C 및 Aga86E를 포함하는 *S. degradans* 유래한 다른 아가레이즈의 기능적인 표면 발현을 조사하였다. Aga50A를 제외한, 다른 아가레이즈(Aga16B, Aga86C, Aga50D 및 Aga86E)는 스팟팅 분석에 대하여 세포의 아가로스 분해활성을 나타내었다(도 4b). 본 발명자들은 또한 시험관에서 전체 세포 촉매작용을 이용하여 아가레이즈들의 기능적인 발현을 조사하였다.

[0025] 먼저, 아가레이즈들의 세포 표면 발현이 37℃에서 유도되었지만 몇몇 아가레이즈가 세포 표면 발현된 세포는 아가레이즈 활성을 가지지 않았다. 그러나, 본 발명자들이 유도 온도를 변화시킨 후(37℃에서 16℃), 본 발명자들은 기능적으로 아가레이즈들이 표면 발현된 세포들을 얻을 수 있었다. 24시간 동안 25℃에서 반응 후 Aga16B, Aga50D 및 Aga86E를 발현하는 세포들은 고체 아가로스를 액화시켰다. 본 발명자들은 도 5에 나타난 것과 같이 TLC에 의해서 발현된 아가레이즈들이 활성을 증명하고 반응산물을 분석하였다. 여러 아가레이즈들은 다양한 활성을 가지면서 특정 중합도(예를 들어, 아가 다당류, 올리고당 또는 이당류)를 가지는 기질에 대한 선호도를 가진다. Aga50A는 기질로써 아가로스에 대한 활성을 나타내지 않기 때문에, 본 발명자들은 Aga50A의 기질로 다른 아가로-올리고당을 테스트하였다. 그 중에서, 이러한 기질들 중에서 Aga50A는 네오아가로테트라오스 (neogaratetraose, DP4) 및 네오아가로헥소오스 (neogarohehexaose, DP6)에 대해 활성을 나타내었고 이러한 기질들을 네오아가로바이오스(NB)로 분해하였다.

[0026] 결과적으로 본 발명자들은 *S. degradans*의 다섯 종 아가레이즈들이 대장균 세포 성장에 치명적인 피해없이 높은 수준으로 기능적인 세포 표면 발현을 확인하였다.

[0027] 배지로 발현된 단백질의 분비

[0028] YfaL는 AIDA-I 패밀리와 서열 유사성을 보이기 때문에, 본 발명자들은 AIDA-I의 것에 기초하여 YfaL의 절단 부위를 처음 예측하였다. AIDA-I 계열에서 절단 부위는 846 및 847 잔기 사이인 것으로 알려졌다(Suhr, M., I. Benz, and M. A. Schmidt. 1996. *Mol Microbiol* 22:31-42). YfaL에서 해당 부위는 서열 정렬로부터 699 및 700 잔기 사이로 예측되었다(도시 안함). 예측된 절단 부위를 증명하기 위하여, 본 발명자들은 정제된 절단된 55 kDa 운반체 도메인의 N-말단 아미노산 시퀀싱을 수행하였다(도시 안함). N-말단의 서열은 예측된 부위에 근접한 696과 697 사이에 YfaL의 절단 부위를 나타낸 ILPYA (YfaL 아미노산 서열에서 697-701 아미노산에 해당)이었다. YfaL에서 절단 부위에 의한 배지로 분비를 테스트하기 위하여, 본 발명자들은 그 부위 (DVAD/ILPYA)를 가지는 분비 LIC 벡터(pATLICSec)를 구축하고 발현 시스템에서 기능적으로 발현된 아가레이즈를 클로닝하였다. 다양한 조건에서 *E. coli*에서 단백질의 오랜 발현에도 불구하고, 본 발명자들은 배지에 분비된 단백질의 활성을 발견할 수 없었다. 따라서 예측된 절단 부위는 작동을 하지 아니하고 미지의 다른 요인이 YfaL의 분비에 관여할 수 있다.

[0029] 이중 단백질의 분비 시스템에 적용이 필요하여, 본 발명자들은 pATLIC 벡터에 이중 단백질 및 YfaL의 운반체 도메인 사이에 LIC 부위 뿐만 아니라 링커로 TEV 절단 부위를 삽입하였다. pATLIC 벡터의 TEV 절단 부위가 실질적으로 인지되고 TEV 프로테아제에 의하여 절단되는지를 확인하기 위하여, 본 발명자들은 pATLIC 벡터에 NABH(neoagarobiose hydrolase)를 클로닝한 후 발현된 NABH에 대한 절단 테스트를 수행하였다. NABH는 기질로 NB(neoagarobiose)를 사용하고 그것을 D-galactose (D-Gal) 및 3,6-anhydro-L-galactose (L-AHG)로 분해한

다.

[0030] 다이머 NABH만이 기능적인 활성을 가지나 표면에 발현된 NABH YfaL은 모노머 AT이기 때문에 NB에 대한 활성을 가지지 않았다. 따라서 만약 삽입된 TEV 절단 부위가 절단되면, 발현된 NABH가 반응 혼합물로 분비되고 다이머화가 가능할 수 있다. 본 발명자들은 대장균의 표면에 NABH를 발현한 후 NABH를 발현하는 대장균 세포를 수집하여서 그 세포에 TEV 프로테아제를 처리하였다. 반응 후, TEV 프로테아제 반응에서 상등액을 분리하고 NB에 대한 NABH의 활성을 검출하였다. TEV 프로테아제에 의하여 처리된 상등액은 NABH의 활성을 나타내었고 NB를 L-AHG 및 D-Gal로 분해하였다(도 6).

[0031] 기능적인 표면 발현 시스템에 대한 발현 속도 효과 및 이중 단백질 크기

[0032] 대장균의 YfaL에서 유추된 패신저 도메인의 크기는 약 75 kDa인 반면에 성공적으로 발현된 *S. degradans* 유래 여섯 종 아가레이즈의 크기들은 25.3 kDa에서 143 kDa까지의 범위이었다. 이 결과는 고안된 발현 시스템의 크기 수용력은 적어도 본래의 패신저 도메인의 크기의 절반에서 두 배 정도의 크기를 가지는 이중 단백질을 세포 표면으로 발현하는 데 충분하다는 것을 나타내었다. 이러한 실험적 관찰을하면, YfaL 기반 세포 표면 발현 시스템은 다양한 크기의 이중 단백질들을 기능적인 형태를 가지고 세포 표면으로 발현이 가능하다는 것을 의미한다.

[0033] 대장균의 표면에 발현된 단백질의 수 계산

[0034] *E. coli*의 세포 표면에 발현된 단백질의 양을 조사하기 위하여 mRFP1(여기 584 nm/방출 607 nm)를 리포터 단백질로 사용하였다. 먼저 본 발명자들은 mRFP1을 정제하고 표면 발현된 mRFP1의 정량적 검출을 위한 mRFP1의 농도와 형광 강도의 표준 곡선을 결정하였다(도 11a). 유도된 세포와 대조군으로 유도되지 않은 세포의 600 nm에서 형광 강도 및 세포 밀도를 측정하였다(도 11b). 그 결과, 세포 당 약 200,000 카피의 mRFP1가 발현되었다. 베타-배럴의 AT 및 단일 *E. coli* 세포의 표면적을 고려할 때(Barnard, T. J., N. Dautin, P. Lukacik, H. D. Bernstein, and S. K. Buchanan. 2007. Nat Struct Mol Biol 14:1214-12202), 이 숫자는 *E. coli* 세포 표면에 발현된 mRFP1의 점유 면적이 약 1%이라는 것을 의미하고 이 값은 다른 표면 발현 시스템을 사용한 많은 스카폴드 단백질 뿐만 아니라 다른 AT들보다 더 높다는 것을 의미한다(Jose, J., R. Bernhardt, and F. Hannemann. 2001. Chembiochem 2:695-701; Jose, J., and S. von Schwichow. 2004. Chembiochem 5:491-499; Klauser, T., J. Kramer, K. Otzelberger, J. Pohlner, and T. F. Meyer. 1993. J Mol Biol 234:579-593). 따라서 그러한 높은 수준의 표면 발현은 여러 분야에 바람직하다.

발명의 효과

[0035] 본 발명에서, 본 발명자들은 *E. coli* 내에서 이중 단백질들을 발현하기 위한 스카폴드로서 YfaL에서 서열분석을 통해 기능적 도메인들을 추정하였고, 효율적인 클로닝 과정을 위해 라이게이션 독립적인 클로닝(ligation-independent cloning; LIC) 시스템을 적용시켰다. 본 발명자들은 기존에 활성형태로의 발현에 실패한 단백질들을 발현하거나 다양한 크기의 단백질들을 발현 가능하다는 점에서 이 신규한 시스템의 능력 및 기능성을 조사하였고, 또한 본 발명의 시스템에 TEV 절단 부위를 삽입함으로써 세포 표면에 발현된 목표 단백질들을 배지로 분비가 가능하다는 것을 확인하였다. 본 발명에서, *E. coli* YfaL을 사용한 대장균 표면 발현 시스템을 기술하였고 그것은 잠재적인 여러 생명공학적 응용에 사용될 수 있는 효과가 있다.

도면의 간단한 설명

[0036] 도 1은 YfaL의 서열 분석을 나타낸 그림이다. 부위(789 ~ 916)는 *Bordetella pertussis* 유래 pertactin과 상동성을 가지며 패신저 도메인과 운반체(translocator) 도메인 사이에 추정된 링커 부위를 가짐.

도 2는 pATLIC 벡터 구축 및 pATLIC 벡터에 대한 LIC 반응 도를 나타냄.

도 3은 발현된 아가레이즈의 SDS-PAGE를 나타낸 그림. 레인 1: 분자 크기 마커, lane 2: pATLIC empty 벡터를 가지는 *E. coli* BW25113 펠렛, lane 3: pATLIC-aga50A, lane 4: pATLIC-aga16B, lane 5: pATLIC-aga86C, lane 6: pATLIC-aga50D, lane 7: pATLIC-aga86E, lane 8: pATLIC-NABH

도 4는 다섯 종 베타-아가레이즈의 기능적 발현의 검출을 나타낸 그림. (a) (1) pBAD-aga16B; (2) pBAD-yfaL; (3) pATLIC empty vector; (4) pATLIC-aga16B를 가지는 *E. coli* BW25113를 나타내는 LB 아가 플레이트. (b) (1) pATLIC empty vector; (2) pATLIC-aga50A; (3) pATLIC-aga16B; (4) pATLIC-aga86C; (5) pATLIC-aga50D;

(6) pATLIC-aga86E를 가지는 *E. coli* BW25113를 나타내는 LB 아가 플레이트.

도 5는 발현된 다섯 종의 베타 아가레이즈에 의한 반응 산물의 TLC 크로마토그램을 나타냄. Lane 1: 0.2% NB (DP2), lane 2: 네오아가로올리고사카라이드(DP4 및 DP6), lane 3: pATLIC empty 벡터, lane 4: pATLIC-aga50A, lane 5: pATLIC-aga16B, lane 6: pATLIC-aga86C, lane 7: pATLIC-aga50D, lane 8: pATLIC-aga86E를 가지는 *E. coli* BW25113 세포 페이스트의 반응 산물.

도 6은 분비된 NABH에 의한 반응산물의 TLC 크로마토그램. Lane 1: 0.2% L-AHG, lane 2: 0.2% D-Gal, lane 3: 0.2% NB (DP2), lane 4: pATLIC-NABH를 가지는 *E. coli* BW25113의 반응산물(TEV 프로테아제 처리 안함), lane 5: TEV 프로테아제에 의하여 처리된 (pATLIC-NABH를 가지는 *E. coli* BW25113) 세포 페이스트의 반응 산물, lane 6: TEV 절단 반응 혼합물로부터 분리된 상등액의 반응 산물.

도 7은 본 발명의 시스템에 의하여 시도하였던 세포 표면발현이 가능한 단백질의 리스트이다.

도 8은 YfaL 과 서열상동성을 가지는 단백질을 스크리닝하여 정리하고 그 도메인에 대하여 추정한 표이다.

도 9는 본 발명의 pATLIC vector 지도를 나타낸 그림이다.

도 10은 다양한 세포 표면 발현 벡터에 의해 세포 표면으로 발현된 mRFP1 단백질의 상대적 형광을 나타낸 그래프이다. 도 10에서 pATLIC1 vector : YfaL에서 29-948번 아미노산까지 제거한 벡터에 클로닝한 mRFP1, pATLIC2 vector : YfaL에서 29-785번 아미노산까지 제거한 벡터에 클로닝한 mRFP1, pATLIC3 vector : YfaL에서 29-699번 아미노산까지 제거한 벡터에 클로닝한 mRFP1 및 pATLIC4 vector : YfaL에서 29-691번 아미노산까지 제거한 벡터에 클로닝한 mRFP1을 나타낸다.

도 11 a)는 mRFP1 단백질의 농도 및 형광 강도의 표준 곡선이고, b)는 *E. coli* 세포 상에 발현된 mRFP1 융합 단백질의 정량적 검출 그림이다. 유도된 세포는 수집하여 세척하고 200 μ l의 20 mM Tris-HCl (pH 8.0)에서 재부유하였다. 200 μ l의 샘플을 시험관에 첨가하고 600nm에서 세포 밀도 및 형광을 측정하였다. 모든 데이터는 3회 반복 실험으로부터 얻었다.

발명을 실시하기 위한 구체적인 내용

[0037] 이하 비한정적인 실시예를 통하여 본 발명을 더욱 상세하게 설명한다. 단 하기 실시예는 본 발명을 예시하기 위한 의도로 기재된 것으로서 본 발명의 목적은 하기 실시예에 의하여 제한되는 것으로 해석되지 아니한다.

[0038] 본 발명에서 사용된 모든 박테리아 균주 및 플라스미드는 표 1에 리스트하였다. *E. coli* DH5알파(인비트로젠)를 일반적인 클로닝 숙주로 사용하였고, *E. coli* BW25113(The Coli Genetic Stock Center 으로부터 구입)을 표면 발현 시스템을 위한 숙주 균주로 사용하였다.

[0039] 표 1은 본 발명에서 사용된 플라스미드를 리스트한 것이다.

표 1

플라스미드	특징
pBAD-yfaL	개량된 pBAD 벡터에 <i>E. coli</i> K12 yfaL 유전자의 클로닝
pBAD-aga16B	개량된 pBAD 벡터에 <i>Saccharophagus degradans</i> 2-40 ^T 유래 Aga16B (62.3 kDa)의 클로닝
pATLIC vector	새롭게 고안된 오토트랜스포팅 LIC 벡터
pATLIC-NABH	pATLIC 벡터에 <i>S. degradans</i> 2-40 ^T 유래 NABH (41.5 kDa)의 클로닝
pATLIC-aga50A	pATLIC 벡터에 <i>S. degradans</i> 2-40 ^T 유래 Aga50A (84.8 kDa)의 클로닝
pATLIC-aga16B	pATLIC 벡터에 <i>S. degradans</i> 2-40 ^T 유래 Aga16B (62.3 kDa)의 클로닝
pATLIC-aga86C	pATLIC 벡터에 <i>S. degradans</i> 2-40 ^T 유래 Aga86C (83.7 kDa)의 클로닝
pATLIC-aga50D	pATLIC 벡터에 <i>S. degradans</i> 2-40 ^T 유래 Aag50D (88.3 kDa)의 클로닝
pATLIC-aga86E	pATLIC 벡터에 <i>S. degradans</i> 2-40 ^T 유래 NABH (143 kDa)의 클로닝

[0040]

[0041]

실시예 1: 서열 분석 및 도메인 인식

[0042]

E. coli (uniprot id: P45508)의 yfaL과 유사성을 가지는 단백질 서열을 BLAST 프로그램(Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. (1990) *J Mol Biol*, **215**, 403-410)을 사용하여 NCBI 진뱅크로부터 검색하였다. 그 다중 서열 배열을 ClustalW 프로그램(Thompson J.D., Higgins D.G. and Gibson T.J. (1994) *Nucleic Acids Res*, **22**, 4673-4680)에 의하여 수행하였고 도메인 구조를 Pfam 데이터베이스(Finn R.D., 등. (2010) *Nucleic Acids Res*, **38**, D211-222)에 의하여 인식하였다. 그 신호 서열 분석을 SignalP 3.0 서버(Bendtsen J.D., Nielsen H., von Heijne G. and Brunak S. (2004) *J Mol Biol*, **340**, 783-795.)에 의해서 수행하였다.

[0043]

CSD 분석을 요약하면,

[0044]

1) Blastp에서 YfaL 단백질 서열로 E value (<0.00001) 기준으로 검색함 : 두 종류의 superfamily가 존재하며, 이 중 하나는 pertactin-like passenger domains이고, 나머지 하나는 autotransporter beta-domain이다.

- [0045] 2) 이 두 가지 superfamily 중 pertacin-like passenger domain의 경우 YfaL 단백질 서열 내에서 789 잔기부터 시작하여 915 잔기이었다.
- [0046] 3) Autotransporter beta-domain의 경우 YfaL 단백질 서열 내에서 989 잔기부터 시작하였다.
- [0047] 4) 선도서열을 예측하는 프로그램인 "SignalP 3.0"에서 YfaL 단백질 서열 중 N-말단 100개의 아미노산 서열 내의 cleavage site를 예측함 : 아미노산 28번과 29번 사이에서 잘리는 것으로 예측되었다.
- [0048] 위의 결과를 바탕으로 YfaL의 단백질 서열에서 1-28번 : 선도서열, 29-789번 : passenger domain, 790-916번 : pertacin-like domain (passenger domain의 세포 표면으로 발현될 수 있도록 대장균의 외막을 통과하는 데 가장 중요한 역할을 하는 부분으로 추정되고, 917-987번 : autotransporter와 pertacin-like domain 사이에서 대장균 외막을 가로지르는 역할을 하는 부분으로 추정되며, 실제로 이 부분의 서열은 [DTPE]-P 서열이 반복적으로 나타나며, 989-1235번 : autotransporter beta-domain이라는 것을 추정하였다.
- [0049] YfaL 이외에 다른 상동성 있는 단백질에 대해서는 일단 YfaL 아미노산 서열을 Blastp(E value < 0.0001)를 수행한 뒤, YfaL과 유사한 길이 및 도메인 구조(organization : signal sequence, passenger domain, linker region, autotransporter beta-barrel)를 가지는 단백질을 스크리닝하였다(도 8). 이와 같은 추론을 하게 된 근거로는 1) Blast 결과의 상동성, 2) YfaL과의 도메인 구조 유사성에 의한 정해진 부분의 passenger domain 을 제거하였을 때 다른 단백질에 의해서도 display이 되는 가능성이 충분하다.
- [0050] 실시예 2: 표면 발현을 위한 LIC 벡터의 구축
- [0051] 표면 발현 시스템을 위한 LIC 벡터를 가지는 관심있는 단백질의 용이하고 편리한 클로닝을 위해서, 본 발명자들은 서열 및 라이게이션 독립적인 클로닝(sequence and ligation independent cloning;SLIC) 방법(Li M.Z. and Elledge S.J. (2007) *Nat Methods*, **4**, 251-256)을 일부 변형하였다. 본 발명자들은 아라비노스 유도 프로 모터 하에서 변형된 pBAD 벡터(Lee J. and Kim S.H. (2009) *Protein Expr Purif*, **63**, 58-61)에 대장균의 *yfaL* 유전자를 클로닝하고 그것을 LIC 벡터 시스템의 구축을 위한 주형 플라스미드로 사용하였다. 그 골격 벡터를 YfaL의 패신저 도메인 부위를 우회하는 LIC 부위에서 SmaI 제한효소 부위를 포함하는 고안된 오버행을 가지는 프라이머 (CCCGGGCCAACGACCGAAACCTTTACTTCCAGCTTCCAACGTGACGGTTAATGGC, forward 프라이머, 서열번호 2 및 GTTGGCCCGGGCGGACACCGTTAGCAGAGAAAA, reverse 프라이머, 서열번호 3)(SmaI 제한 효소 부위는 밑줄)로 증폭하였다. 그 증폭된 산물을 주형을 제거하기 위하여 37℃에서 2 시간 동안 20 U DpnI (NEB, England)로 배양하였다. 그 PCR 산물을 25℃에서 15분간 0.5 U T4 DNA polymerase (New England Biolabs)로 배양하였다. T4 DNA polymerase를 70℃에서 20분간 가열하여 불활성화하였다. 그 산물을 단일 가닥 오버행의 어닐링을 위해서 25℃에서 30분간 더 배양하였고, *E. coli* DH5알파로 직접 형질전환하였다. 표면 발현을 위해서 구축된 LIC 벡터 (pATLIC)를 SmaI (New England Biolabs) 제한효소 및 DNA 시퀀싱으로 입증하였다. tobacco etch virus (TEV) 프로테아제로 절단하여 배지로 발현된 단백질을 배출하기 위하여, TEV 절단 부위 (ENLYFQ) 를 그 클로닝 부위에 삽입하였다.
- [0052] LIC ready 벡터(pATLIC)를 SmaI 제한효소 처리(37℃에서 2 시간), 37℃에서 30분간 dATP (Promega)를 가지는 T4 DNA polymerase 반응을 수반하여 제조하였다. 선형 LIC ready 벡터를 -20℃에서 저장하였다. 타겟 단백질의 클로닝을 위해서, 특정 쌍의 오버행을 PCR 프라이머에 부착하는 것이 필요하다. 본 발명에 사용된 모든 올리고머들을 표 2에 리스트하였다.

표 2

프라이머	서열
NABH-pATLIC_F	<u>CGGTGTCGCGCCC</u> AGCGATTCAAAAGTAAATAAAAAATTG
NABH-pATLIC_R	<u>CGGTCGTTGGCCCT</u> ACTGCTCCGGAATCGCCTG
aga50A-pATLIC_F	<u>CGGTGTCGCGCCC</u> GATAAAGACGAGCCGCAAGC
aga50A-pATLIC_R	<u>CGGTCGTTGGCCCT</u> CGCCAAAACGTCGACTATAT
aga16B-pATLIC_F	<u>CGGTGTCGCGCCC</u> GATTGGGACGGAATTCCTGTC
aga16B-pATLIC_R	<u>CGGTCGTTGGCCCT</u> GTTGCTAAGCGTGAACCTATCTAG
aga86C-pATLIC_F	<u>CGGTGTCGCGCCC</u> GGGGGCGGTAAATCCTCCAC
aga86C-pATLIC_R	<u>CGGTCGTTGGCCCT</u> GCATAGGACGAGTAATTGAG
aga50D-pATLIC_F	<u>CGGTGTCGCGCCC</u> GGTGCAATTGGAGGTCTCGT
aga50D-pATLIC_R	<u>CGGTCGTTGGCCCT</u> TTTGCTGCCTAGCCTTTCGG
aga86E-pATLIC_F	<u>CGGTGTCGCGCCC</u> GCCGATTATGTAATCGAAGCGG
aga86E-pATLIC_R	<u>CGGTCGTTGGCCCT</u> CTATTGGCTCAGAAGTAAATTCC

[0053]

[0054]

표 2는 본 발명의 클로닝에 사용된 올리고뉴클레오타이드 프라이머 서열이다.

[0055]

실시예 3: 표면 발현 LIC 벡터에서 여러 단백질의 클로닝

[0056]

발현된 단백질들의 크기 다양성 및 기능적 발현을 조사하기 위하여 본 발명자들은 mRFP1, *Saccharophagus degradans* 2-40으로부터 네오아가로바이오하이드로레이즈(neoagarobiohydrolase;NABH), 및 다섯 종의 베타 아가레이즈(Aga50A, Aga16B, Aga86C, Aga50D 및 Aga86E)을 조사하였다. 클로닝에 사용된 프라이머들은 표 2에 리스트하였다. 본 발명자들은 Aga50A (1-23), Aga16B (1-19), Aga86C (1-33), Aga50D (1-46), Aga86E (1-29)의 N-말단에서 예측된 시그널 서열을 제거하고 다음과 같이 pATLIC로 클로닝하였다. PCR 산물들을 정제하고 37℃에서 30분간 dTTP (Promega)를 가지는 T4 DNA polymerase로 처리하고 T4 DNA polymerase를 불활성화시키기 위해서 70℃에서 20분간 가열하였다. 그 반응 산물을 1:2 비율(벡터 대 삽입체)로 LIC ready 벡터와 혼합하고 *E. coli* DH5알파에 형질전환하였다(Lee J. and Kim S.H. (2009) *Protein Expr Purif*, **63**, 58-61).

[0057]

실시예 4: 단백질의 세포 표면 발현

[0058]

여섯 아가레이즈 및 mRFP1을 가지는 pATLIC를 가지는 재조합 세포들을 100 µg/ml의 앰피실린을 포함하는 LB broth에서 배양하였다. 0.5의 600nm에서 흡광도(OD₆₀₀)에서, 세포들을 재조합 단백질의 발현을 위해서 최종 농도 0.02 ~ 0.2% (w/v)가 되게 L(+)-arabinose (Sigma)로 유도하였다. 세포들을 유도 후 16℃에서 24시간 동안 배양하고 원심분리(3500 rpm, 10분, 4℃에서)에 의하여 수집하고, ice-cold 10 mM NaCl로 세척하고, 사용할 때까지 -20℃에서 저장하였다. 부드러운 세포 파쇄를 위해서 발현된 세포 페이스트를 BugBuster®

Master Mix (Novagen)에 재부유하였다. 4℃, 15000 rpm에서 20 분간 원심분리에 의하여 펠렛 분획을 분리하고 10% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)로 분석하였다.

[0059]

실시예 5: 발현된 효소 활성의 검출

[0060]

다른 타입의 다섯 종의 아가레이즈(Aga50A, Aga16B, Aga86C, Aga50D 및 Aga86E)의 기능적인 표면 발현을 조사하기 위해서, 각 단백질을 발현하는 50µl의 세포(~10⁸ 세포에 해당)를 대수기에 샘플링하여서 자가 유도를 위한 0.2% L(+)-arabinose를 포함하는 LB 아가 플레이트에 스팟팅하였다. 25℃에서 24시간 동안 배양 후, 아가

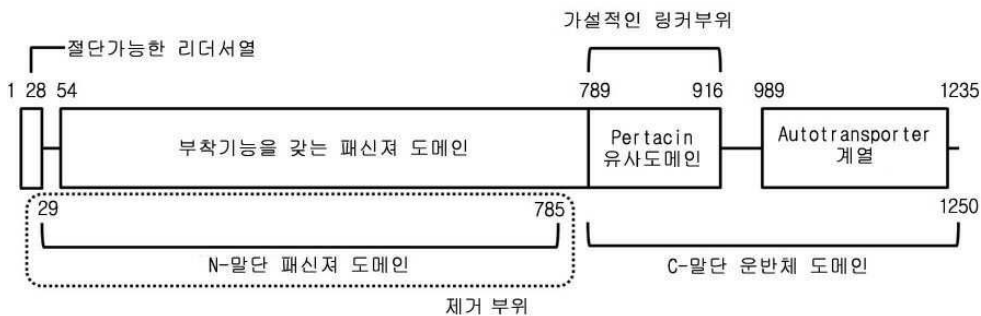
레이즈 활성을 요오드 용액으로 아가 염색을 한 후 스팟된 지역 주위에 hollow zone의 크기로 검출하였다. 발현된 세포에 의한 반응 산물들을 다음과 같이 thin-layer chromatography (TLC)로도 동정하였다. 100 ml의 발현된 세포를 16℃에서 24시간 유도한 후 수집하여 1.0% 아가로스를 가지는 20 mM Tris-HCl (pH 8.0)를 포함하는 2ml의 반응 혼합물에 농축하였다. 그 반응은 25℃에서 24시간 동안 수행하고 95℃에서 5분간 가열하여 중지하였다. 반응 혼합물(2 μ l)의 상등액을 실리카겔 60 TLC 플레이트(Cellulose F, Art. 5728; Merck, KG Darmstadt, Germany)에 로딩하고, *n*-부탄올: 아세트산: 물(2:1:1, v/v) 용매 시스템으로 2회 현상하고, 10% (v/v) H₂SO₄ and 0.2% naphthoresorcinol in ethanol에서 가열하여서 육안화하였다(Duckworth M. and Yaphe W. (1970) *J Chromatogr*, **49**, 482-487).

[0061] 실시예 6: tobacco etch virus (TEV) 프로테아제를 사용한 발현된 단백질의 절단

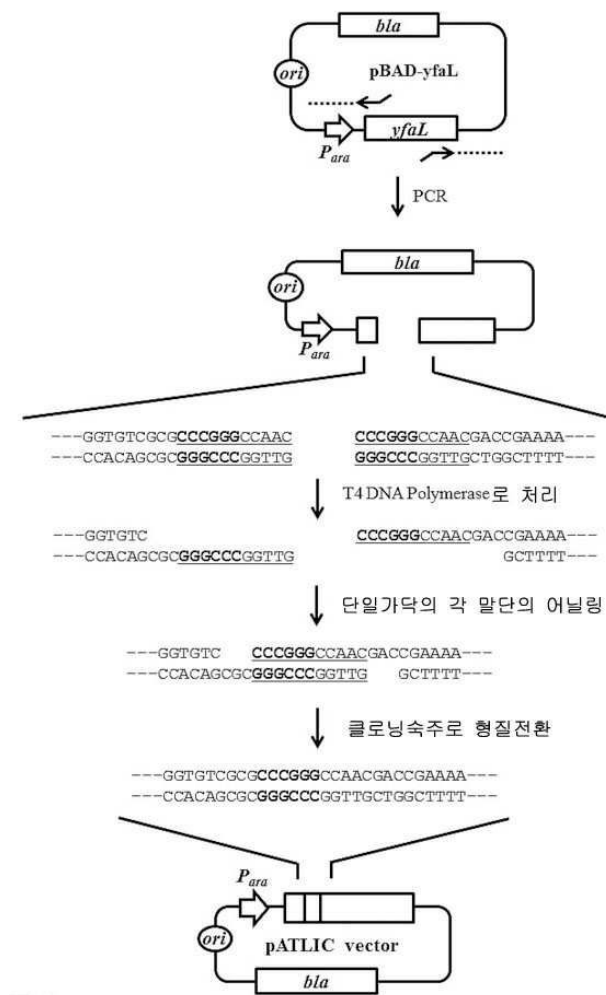
[0062] 배지로 발현된 단백질들을 분리하기 위하여 본 발명자들은 TEV 프로테아제(TurboTEV, Eton Bioscience)를 사용하여 절단 테스트를 수행하였다. NABH를 발현하는 100ml의 세포를 16℃에서 24 시간 유도한 후 수집하고 4℃에서 16 시간 동안 25 mM Tris-HCl (pH 8.0)를 포함하는 2ml 반응 혼합물에 존재하는 100U TEV 프로테아제로 처리하였다. 그 반응 혼합물을 원심분리(4℃에서, 10분간 3500 rpm)에 의해서 분리하였다. 각 분획(상등액 및 세포 페이스트)에서 NABH의 활성을 확인하기 위해서, 반응 혼합물의 상등액을 기질로 0.1% neoagarobiose (NB)를 첨가하고 세포 페이스트를 0.1% NB를 포함하는 2 ml의 25 mM Tris-HCl로 재부유하였다. 25℃에서 12 시간 동안 그들을 반응한 후, 반응 산물의 그 상등액(2 μ l)를 TLC 플레이트 상에 로딩하고, *n*-부탄올: 에탄올: 물(3:2:2, v/v) 용매 시스템으로 현상하고, 10% (v/v) H₂SO₄ and 0.2% naphthoresorcinol in ethanol에서 가열하여서 육안화하였다.

도면

도면1



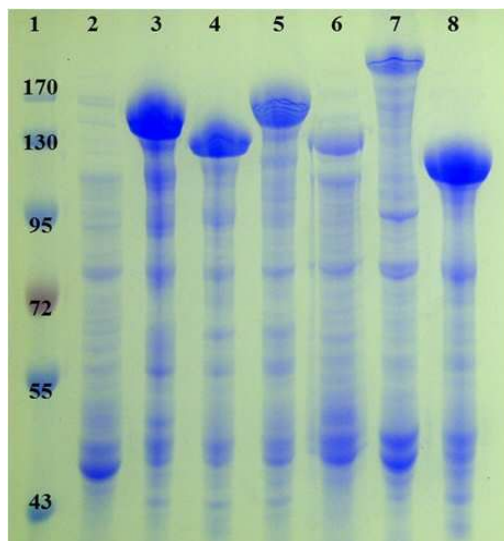
도면2a



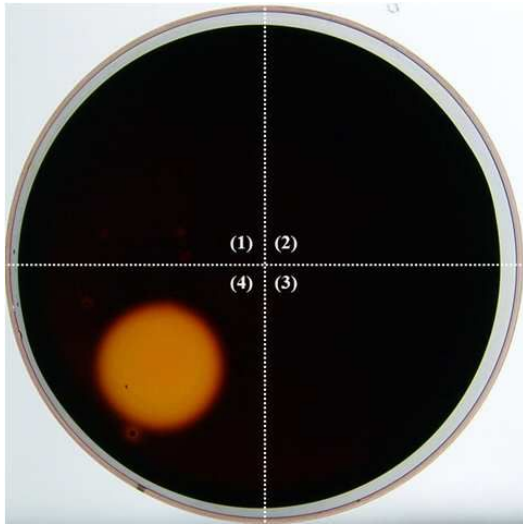
도면2b



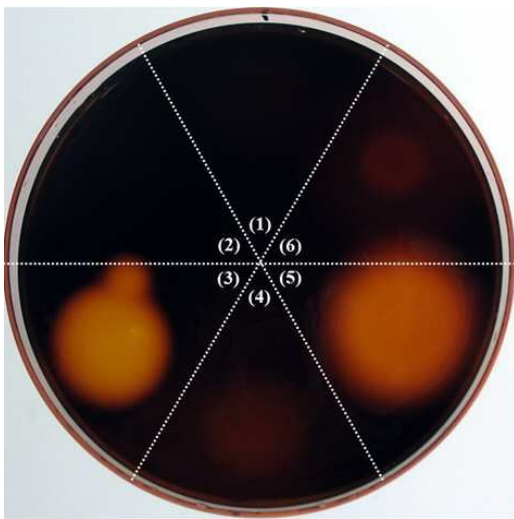
도면3



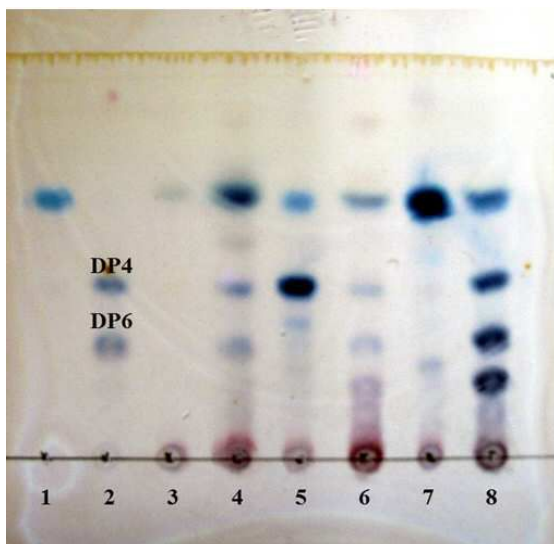
도면4a



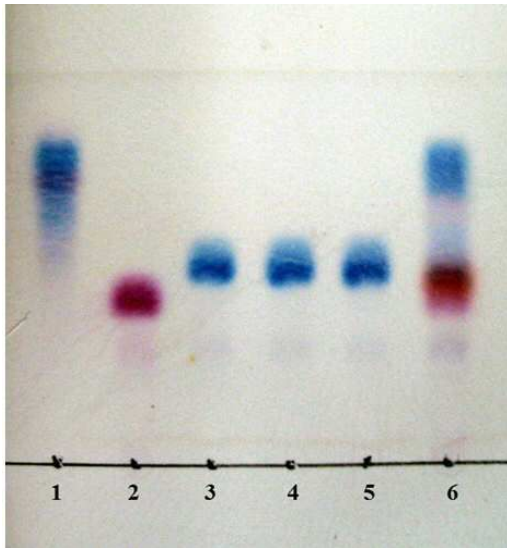
도면4b



도면5



도면6



#	미생물	Genbank ID	Gene Name/ID	설명
1	Saccharophagus degradans 2-40 ^T	YP_526649	Sde1175	agarase (GH16) : 1-19 잔기 제거
2		YP_526650.1	Sde1176	agarase (GH50) : 1-23 잔기 제거
3		YP_528116.1	Sde2644	agarase (GH50) : 1-46 잔기 제거
4		YP_528122.1	Sde2650	agarase (GH86) : 1-33 잔기 제거
5		YP_528127.1	Sde2655	agarase (GH86) : 1-29 잔기 제거
6		YP_528129.1	Sde2657	agarase (GH43)
7		YP_526868.1	Sde1394	cellulase (GH1)
8		YP_527744.1	Sde2272	cellulase (GH6) : 1-45 잔기 제거
9		YP_528706.1	Sde3237	cellulase (GH5) : 1-36 잔기 제거
10		YP_529070.1	Sde3603	cellulase (GH1)
11	Streptomyces coelicolor M600	NP_627674.1	Sco3471	agarase (GH16) : 1-30 잔기 제거
12	Cellvibrio japonicus	NP_627690.1	Sco3487	agarase (GH50) : 1-45 잔기 제거
13		YP_001983791.1	Cja3337	cellulase (GH5) : 1-25 잔기 제거
14	Clostridium cellulolyticum	YP_002504739.1	Ccel0374	cellulase (GH1)
15		YP_002500588.1	Ccel0729	cellulase (GH48) : 1-29 잔기 제거
16	Hahella chejuensis	YP_435061.1	Hch03914	cellulase (GH5) : 1-46 잔기 제거
17		YP_660406.1	Pat0824	agarase (GH16) : 1-28 잔기 제거
18		YP_661477.1	Pat1904	agarase (GH86) : 1-33 잔기 제거
19		YP_661543.1	Pat1971	agarase (GH50) : 1-22 잔기 제거
20	Pseudoalteromonas atlantica T6c	YP_661912.1	Pat12341	agarase (GH86) : 1-24 잔기 제거
21		238556185	GFPuv	green fluorescence protein
22		AF506027	mRFP1	red fluorescence protein
23	Soil bacterium	FJ755834	C3UWD1	aryl acylamidase

YfAL 서열상동성

Organism	Accession number	Identity	Signal seq.	Domain organization		
				1	2	3
<i>Escherichia coli</i>	NP_416736.1	100%	1~28	Passenger domain 29~785	Linker region 786~916	Autotransporter superfamily 989~1235
	CBJ01870.1	100%	1~28	29~785	786~916	989~1235
	ZP_0724747.1	100%	1~28	29~785	786~916	989~1235
	YP_001459034.1	100%	1~28	29~785	786~916	989~1235
	ZP_07121532.1	100%	1~28	29~785	786~916	989~1235
	ZP_02999058.1	100%	1~28	29~785	786~916	989~1235
	YP_003045352.1	100%	1~28	29~785	786~916	989~1235
	EGB57147.1	100%	1~28	29~785	786~916	989~1235
	EGB37260.1	98%	1~28	29~785	786~916	989~1235
	ZP_07146751.1	100%	1~12	13~769	770~900	973~1219
	ZP_07498247.1	98%	1~12	13~769	770~900	973~1219
	ZP_07787684.1	98%	1~28	29~785	786~916	989~1235
	YP_001724408.1	98%	1~12	13~769	770~900	973~1219
	YP_003035675.1	98%	1~12	13~769	770~900	973~1219
	ZP_03068365.1	98%	1~12	13~769	770~900	973~1219
	ZP_07523028.1	98%	1~12	13~769	770~900	973~1213
	ZP_07102820.1	100%	1~28	29~785	786~916	989~1235
	EGB90124.1	100%	1~28	29~785	786~916	989~1235
	YP_002403506.1	100%	1~28	29~785	786~916	989~1235
	EFZ69526.1	100%	1~28	29~785	786~916	989~1235
	YP_002293769.1	100%	1~28	29~785	786~916	989~1235
	ZP_07590844.1	100%	1~28	29~785	786~916	989~1235
	ZP_06662998.1	100%	1~28	29~785	786~916	989~1235
	ZP_07135193.1	100%	1~28	29~785	786~916	989~1235
	YP_001463580.1	100%	1~28	29~785	786~916	991~1237
	ZP_07687541.1	100%	1~28	29~785	786~916	989~1235
	ZP_03028301.1	100%	1~28	29~785	786~916	989~1235
	BAI31473.1	100%	1~28	29~785	786~916	989~1235
	ZP_03043348.1	100%	1~28	29~785	786~916	989~1235

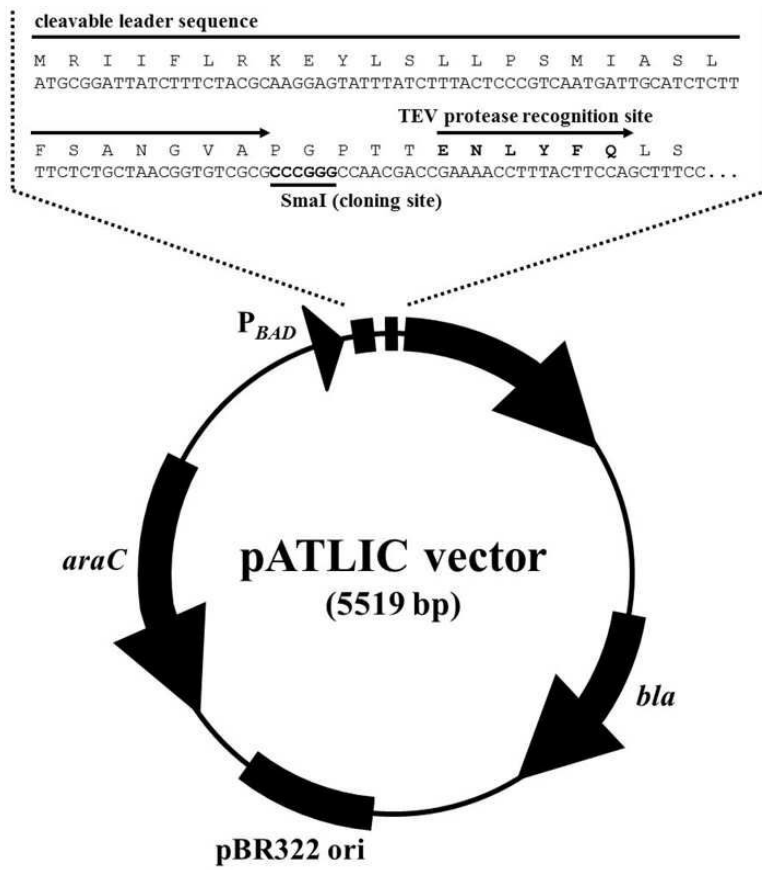
도면8b

ZP 07098380.1	100%	1~28	29~785	786~916	983~1229
YP 002387709.1	100%	1~28	29~785	786~916	983~1229
NP 288807.1	100%	1~28	29~785	786~916	989~1235
EFW74417.1	100%	1~28	29~785	786~916	979~1225
BAI26434.1	100%	1~28	29~785	786~916	983~1229
ZP 07142137.1	100%	1~28	29~785	786~916	985~1231
BAI36814.1	100%	1~28	29~785	786~916	989~1235
YP 003078961.1	100%	1~28	29~785	786~916	989~1235
ZP 07614097.1	98%	1~12	13~769	770~900	973~1213
EFZ47577.1	98%	1~12	13~769	770~900	973~1213
ZP 06658169.1	100%	1~28	29~785	786~916	983~1229
ZP 07186279.1	100%	1~28	29~785	786~916	989~1235
EFX25099.1	98%	1~12	13~769	770~900	973~1219
ZP 06654172.1	100%	1~28	29~785	786~916	981~1227
ZP 02786344.1	98%	1~12	13~769	770~900	973~1219
EGC11783.1	98%	1~12	13~769	770~900	973~1213
EB41107.1	98%	1~12	13~769	770~900	973~1213
ZP 02812144.1	98%	1~12	13~769	770~900	973~1219
ZP 02823402.1	98%	1~12	13~769	770~900	973~1219
ZP 07508510.1	98%	1~12	13~769	770~900	973~1219
ZP 02774656.1	98%	1~12	13~769	770~900	973~1219
YP 002413282.1	100%	1~28	29~772	773~903	976~1222
EFX30437.1	98%	1~12	13~775	776~906	979~1225
ZP 07619499.1	98%	1~12	13~769	770~900	975~1221
ZP 07518131.1	98%	1~12	13~769	770~900	965~1211
YP 001744429.1	100%	1~28	29~787	788~918	993~1239
YP 002408332.1	100%	1~28	29~787	788~918	993~1239
YP 002329881.1	100%	1~28	29~787	788~918	991~1237
NP 754661.1	100%	1~28	29~787	788~918	991~1237
YP 002398608.1	100%	1~28	29~787	788~918	991~1237
EG875831.1	100%	1~28	29~787	788~918	991~1237
EG861754.1	100%	1~28	29~787	788~918	991~1237
YP 541512.1	100%	1~28	29~787	788~918	991~1237
ZP 07175432.1	100%	1~28	29~787	788~918	991~1237

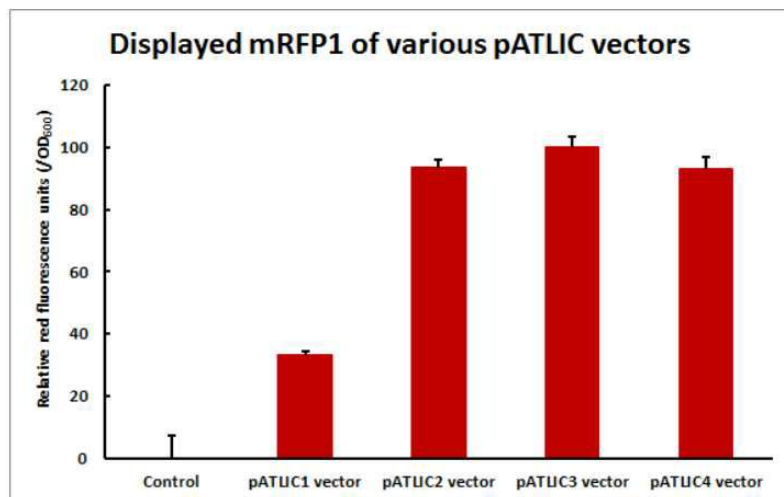
도면8c

	<u>YP 670171.1</u>	100%	1~28	29~787	788~918	991~1237
	<u>CAP76735.1</u>	100%	1~28	29~787	788~918	991~1237
	<u>ZP 07624785.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>ZP 07503047.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>EFZ73335.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>ZP 07781748.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>ADN70533.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>ZP 03031601.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>ADR27681.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>ZP 07448816.1</u>	98%	1~12	13~771	772~902	975~1221
	<u>Escherichia albertii</u> <u>ZP 02901808.1</u>	99%	1~26	27~785	786~916	989~1235
	<u>Escherichia fergusonii</u> <u>YP 002382104.1</u>	100%	1~28	29~787	788~918	991~1237
	<u>EGC06746.1</u>	100%	1~28	29~787	788~918	993~1239
<u>Shigella sp.</u>	<u>ZP 05431470.1</u>	100%	1~28	29~785	786~916	983~1229
<u>Rahnella sp.</u>	<u>YP 004212039.1</u>	99%	1~17	18~774	775~906	994~1252

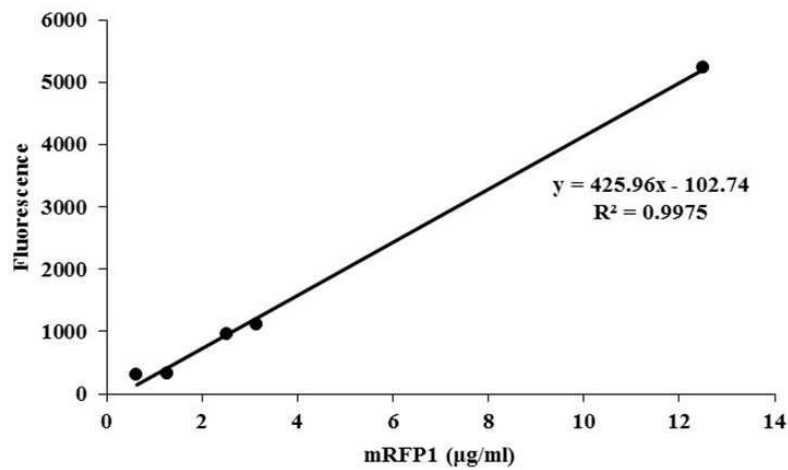
도면9



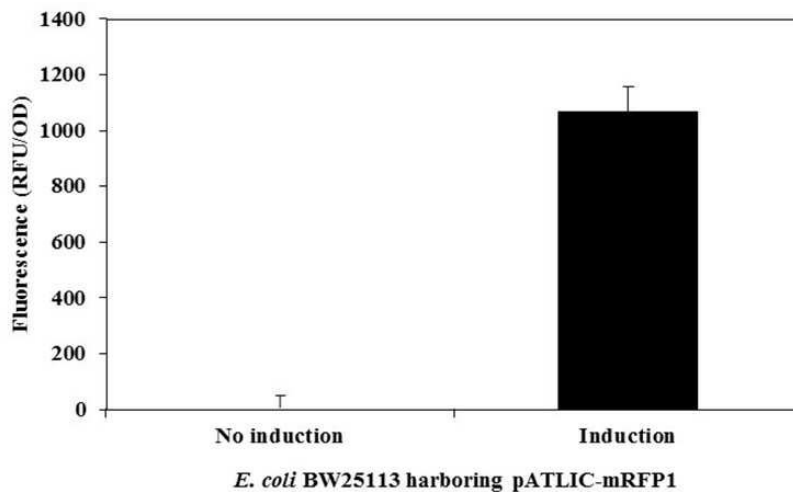
도면10



도면11a



도면11b



서열 목록

- <110> KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION
- <120> Construction of A Bacterial Surface Display System Based on A Putative Autotransporter
- <160> 3
- <170> Kopatent In 1.71
- <210> 1
- <211> 1250
- <212> PRT
- <213> *E. coli*
- <400> 1

Met Arg Ile Ile Phe Leu Arg Lys Glu Tyr Leu Ser Leu Leu Pro Ser

1 5 10 15

Met Ile Ala Ser Leu Phe Ser Ala Asn Gly Val Ala Ala Val Thr Asp

				20				25				30			
Ser	Cys	Gln	Gly	Tyr	Asp	Val	Lys	Ala	Ser	Cys	Gln	Ala	Ser	Arg	Gln
35				40				45							
Ser	Leu	Ser	Gly	Ile	Thr	Gln	Asp	Trp	Ser	Ile	Ala	Asp	Gly	Gln	Trp
50				55				60							
Leu	Val	Phe	Ser	Asp	Met	Thr	Asn	Asn	Ala	Ser	Gly	Gly	Ala	Val	Phe
65				70				75				80			
Leu	Gln	Gln	Gly	Ala	Glu	Phe	Ser	Leu	Leu	Pro	Glu	Asn	Glu	Thr	Gly
85				90				95							
Met	Thr	Leu	Phe	Ala	Asn	Asn	Thr	Val	Thr	Gly	Glu	Tyr	Asn	Asn	Gly
100				105				110							
Gly	Ala	Ile	Phe	Ala	Lys	Glu	Asn	Ser	Thr	Leu	Asn	Leu	Thr	Asp	Val
115				120				125							
Ile	Phe	Ser	Gly	Asn	Val	Ala	Gly	Gly	Tyr	Gly	Gly	Ala	Ile	Tyr	Ser
130				135				140							
Ser	Gly	Thr	Asn	Asp	Thr	Gly	Ala	Val	Asp	Leu	Arg	Val	Thr	Asn	Ala
145				150				155				160			
Met	Phe	Arg	Asn	Asn	Ile	Ala	Asn	Asp	Gly	Lys	Gly	Gly	Ala	Ile	Tyr
165				170				175							
Thr	Ile	Asn	Asn	Asp	Val	Tyr	Leu	Ser	Asp	Val	Ile	Phe	Asp	Asn	Asn
180				185				190							
Gln	Ala	Tyr	Thr	Ser	Thr	Ser	Tyr	Ser	Asp	Gly	Asp	Gly	Gly	Ala	Ile
195				200				205							
Asp	Val	Thr	Asp	Asn	Asn	Ser	Asp	Ser	Lys	His	Pro	Ser	Gly	Tyr	Thr
210				215				220							
Ile	Val	Asn	Asn	Thr	Ala	Phe	Thr	Asn	Asn	Thr	Ala	Glu	Gly	Tyr	Gly
225				230				235				240			
Gly	Ala	Ile	Tyr	Thr	Asn	Ser	Val	Thr	Ala	Pro	Tyr	Leu	Ile	Asp	Ile
245				250				255							
Ser	Val	Asp	Asp	Ser	Tyr	Ser	Gln	Asn	Gly	Gly	Val	Leu	Val	Asp	Glu
260				265				270							

Asn Asn Ser Ala Ala Gly Tyr Gly Asp Gly Pro Ser Ser Ala Ala Gly
 275 280 285
 Gly Phe Met Tyr Leu Gly Leu Ser Glu Val Thr Phe Asp Ile Ala Asp
 290 295 300
 Gly Lys Thr Leu Val Ile Gly Asn Thr Glu Asn Asp Gly Ala Val Asp
 305 310 315 320

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

 385 390 395 400
 Thr Gln Gln Thr Phe Val His Ala Leu Thr Gly Phe Gln Asn Gly Thr
 405 410 415
 Leu Asn Ile Asp Ala Gly Gly Asn Val Thr Val Asn Gln Gly Ser Phe
 420 425 430
 Ala Gly Ile Ile Glu Gly Ala Gly Gln Leu Thr Ile Ala Gln Asn Gly
 435 440 445
 Ser Tyr Val Leu Ala Gly Ala Gln Ser Met Ala Leu Thr Gly Asp Ile
 450 455 460

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

Gln Asp Val Val Asp Leu Ala Gly Gly Asp Asn Leu His Ile Gly Gly

530

535

540

Asp Gly Lys Asp Gly Val Tyr Val Val Val Asp Ala Ser Asp Gly Gln

545

550

555

560

Val Ser Leu Ala Asn Asn Asn Ser Tyr Leu Gly Thr Thr Gln Ile Ala

565

570

575

Ser Gly Thr Leu Met Val Ser Asp Asn Ser Gln Leu Gly Asp Thr His

580

585

590

Tyr Asn Arg Gln Val Ile Phe Thr Asp Lys Gln Gln Glu Ser Val Met

595

600

605

Glu Ile Thr Ser Asp Val Asp Thr Arg Ser Asp Ala Ala Gly His Gly

610

615

620

Arg Asp Ile Glu Met Arg Ala Asp Gly Glu Val Ala Val Asp Ala Gly

625

630

635

640

Val Asp Thr Gln Trp Gly Ala Leu Met Ala Asp Ser Ser Gly Gln His

645

650

655

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

660

665

670

Leu Thr Ala Ser Gly Thr Thr Gln Ser Ala Val Arg Val Glu Glu Gly

675

680

685

Thr Leu Lys Gly Asp Val Ala Asp Ile Leu Pro Tyr Ala Ser Ser Leu

690

695

700

Trp Val Gly Asp Gly Ala Thr Phe Val Thr Gly Ala Asp Gln Asp Ile

705

710

715

720

Gln Ser Ile Asp Ala Ile Ser Ser Gly Thr Ile Asp Ile Ser Asp Gly

725

730

735

Thr Val Leu Arg Leu Thr Gly Gln Asp Thr Ser Val Ala Leu Asn Ala

740

745

750

Ser Leu Phe Asn Gly Asp Gly Thr Leu Val Asn Ala Thr Asp Gly Val

755

760

765

Thr Leu Thr Gly Glu Leu Asn Thr Asn Leu Glu Thr Asp Ser Leu Thr

770 775 780
 Tyr Leu Ser Asn Val Thr Val Asn Gly Asn Leu Thr Asn Thr Ser Gly
 785 790 795 800
 Ala Val Ser Leu Gln Asn Gly Val Ala Gly Asp Thr Leu Thr Val Asn
 805 810 815
 Gly Asp Tyr Thr Gly Gly Gly Thr Leu Leu Leu Asp Ser Glu Leu Asn

 820 825 830
 Gly Asp Asp Ser Val Ser Asp Gln Leu Val Met Asn Gly Asn Thr Ala
 835 840 845
 Gly Asn Thr Thr Val Val Val Asn Ser Ile Thr Gly Ile Gly Glu Pro
 850 855 860
 Thr Ser Thr Gly Ile Lys Val Val Asp Phe Ala Ala Asp Pro Thr Gln
 865 870 875 880
 Phe Gln Asn Asn Ala Gln Phe Ser Leu Ala Gly Ser Gly Tyr Val Asn
 885 890 895

 Met Gly Ala Tyr Asp Tyr Thr Leu Val Glu Asp Asn Asn Asp Trp Tyr
 900 905 910
 Leu Arg Ser Gln Glu Val Thr Pro Pro Ser Pro Pro Asp Pro Asp Pro
 915 920 925
 Thr Pro Asp Pro Asp Pro Thr Pro Asp Pro Asp Pro Thr Pro Asp Pro
 930 935 940
 Glu Pro Thr Pro Ala Tyr Gln Pro Val Leu Asn Ala Lys Val Gly Gly
 945 950 955 960
 Tyr Leu Asn Asn Leu Arg Ala Ala Asn Gln Ala Phe Met Met Glu Arg

 965 970 975
 Arg Asp His Ala Gly Gly Asp Gly Gln Thr Leu Asn Leu Arg Val Ile
 980 985 990
 Gly Gly Asp Tyr His Tyr Thr Ala Ala Gly Gln Leu Ala Gln His Glu
 995 1000 1005
 Asp Thr Ser Thr Val Gln Leu Ser Gly Asp Leu Phe Ser Gly Arg Trp
 1010 1015 1020
 Gly Thr Asp Gly Glu Trp Met Leu Gly Ile Val Gly Gly Tyr Ser Asp

1025 1030 1035 1040

Asn Gln Gly Asp Ser Arg Ser Asn Met Thr Gly Thr Arg Ala Asp Asn

1045 1050 1055

Gln Asn His Gly Tyr Ala Val Gly Leu Thr Ser Ser Trp Phe Gln His

1060 1065 1070

Gly Asn Gln Lys Gln Gly Ala Trp Leu Asp Ser Trp Leu Gln Tyr Ala

1075 1080 1085

Trp Phe Ser Asn Asp Val Ser Glu Gln Glu Asp Gly Thr Asp His Tyr

1090 1095 1100

His Ser Ser Gly Ile Ile Ala Ser Leu Glu Ala Gly Tyr Gln Trp Leu

1105 1110 1115 1120

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

1125 1130 1135

Gly Val Gln Gln Asp Asp Phe Thr Ala Ala Asn Arg Ala Arg Val Ser

1140 1145 1150

Gln Ser Gln Gly Asp Asp Ile Gln Thr Arg Leu Gly Leu His Ser Glu

1155 1160 1165

Trp Arg Thr Ala Val His Val Ile Pro Thr Leu Asp Leu Asn Tyr Tyr

1170 1175 1180

His Asp Pro His Ser Thr Glu Ile Glu Glu Asp Gly Ser Thr Ile Ser

1185 1190 1195 1200

Asp Asp Ala Val Lys Gln Arg Gly Glu Ile Lys Val Gly Val Thr Gly

1205 1210 1215

Asn Ile Ser Gln Arg Val Ser Leu Arg Gly Ser Val Ala Trp Gln Lys

1220 1225 1230

Gly Ser Asp Asp Phe Ala Gln Thr Ala Gly Phe Leu Ser Met Thr Val

1235 1240 1245

Lys Trp

1250

<210> 2

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 2

cccgggccaa cgaccgaaaa ctttacttc cagctttcca acgtgacggt taatggc 57

<210> 3

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 3

gttgccccgg gcgcgacacc gtttagcagag aaaa 34