



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0072201
(43) 공개일자 2013년07월01일

- (51) 국제특허분류(Int. Cl.)
C12N 15/78 (2006.01) C12N 15/31 (2006.01)
C07K 14/00 (2006.01) C07K 19/00 (2006.01)
(21) 출원번호 10-2012-7027780
(22) 출원일자(국제) 2011년03월28일
심사청구일자 없음
(85) 번역문제출일자 2012년10월24일
(86) 국제출원번호 PCT/US2011/030227
(87) 국제공개번호 WO 2011/126811
국제공개일자 2011년10월13일
(30) 우선권주장
61/319,152 2010년03월30일 미국(US)
(뒷면에 계속)
(71) 출원인
피페넥스 인크.
미국 캘리포니아주 92121 샌디에고 로젤레 스트리트 10790
(72) 발명자
레탈락 다이안 엠
미국 92064 캘리포니아주 포웨이 캐리지 로드 13420
츄 로렌스
미국 92130 캘리포니아주 샌디에고 윌로미어 레인 5590
(뒷면에 계속)
(74) 대리인
김진희, 김성기

전체 청구항 수 : 총 36 항

(54) 발명의 명칭 **제조합 독소 단백질의 고수준 발현**

(57) 요약

본 발명은 박테리아 숙주에서의 제조합 독소 단백질 제조 분야에 관한 것이다. 특히, 본 발명은 박테리아 숙주로부터 제조합 CRM197, 디프테리아 독소, 백일해 독소, 파상풍 독소이드 단편 C, 콜레라 독소 B, 콜레라 홀로톡신, 및 슈도모나스 외독소 A를 고수준으로 수득하기 위한 제조 방법에 관한 것이다.

대표도 - 도1



(72) 발명자

진 홍판

미국 92128 캘리포니아주 샌디에고 캐너리지 로드
11978

탈보트 헨리 더블유

미국 92130 캘리포니아주 샌디에고 스타전 코트
4111

(30) 우선권주장

61/325,235 2010년04월16일 미국(US)

PCT/US2010/030573 2010년04월09일 세계지적재
산권기구(WIPO)(WO)

특허청구의 범위

청구항 1

슈도모나드(*Pseudomonad*) 숙주 세포에서 재조합 독소 단백질을 제조하는 방법으로서,

독소 단백질을 코딩하는 뉴클레오티드 서열을 발현 벡터에 결합시키는 단계;

슈도모나드 숙주 세포를 상기 발현 벡터로 형질전환시키는 단계; 및

재조합 독소 단백질의 발현에 적합한 배양 배지 중에서 상기 형질전환된 슈도모나드 숙주 세포를 배양하는 단계를 포함하며;

상기 재조합 독소 단백질이 CRM197, 디프테리아 독소, 콜레라 홀로톡신, 콜레라 독소 B, 백일해 독소, 파상풍 독소 단편 C, C. 디피실레(*C. difficile*) 독소 B, 및 P. 아에루기노사(*P. aeruginosa*) 외독소 A로부터 선택되거나,

상기 재조합 독소 단백질이 콜레라 독소 B, 콜레라 홀로톡신, 백일해 독소, 파상풍 독소 단편 C, C. 디피실레 독소 B, 및 P. 아에루기노사 외독소 A로부터 선택되거나, 또는

상기 재조합 독소 단백질이 콜레라 독소 B, 콜레라 홀로톡신, 백일해 독소, 파상풍 독소 단편 C, 및 C. 디피실레 독소 B로부터 선택되는 것인 제조 방법.

청구항 2

제1항에 있어서, 재조합 단백질이 0.2 g/L~약 12 g/L의 가용성 및/또는 활성 독소 단백질의 수율로 제조되는 것인 제조 방법.

청구항 3

제2항에 있어서, 가용성 및/또는 활성 독소 단백질의 수율이 약 0.2 g/L~약 12 g/L이고, 약 0.2 g/L, 약 0.3 g/L, 약 0.4 g/L, 약 0.5 g/L, 약 0.6 g/L, 약 0.7 g/L, 약 0.8 g/L, 약 0.9 g/L, 약 1 g/L, 약 1.5 g/L, 약 2 g/L, 약 2.5 g/L, 약 3 g/L, 약 3.5 g/L, 약 4 g/L, 약 4.5 g/L, 약 5 g/L, 약 5.5 g/L, 약 6 g/L, 약 6.5 g/L, 약 7 g/L, 약 7.5 g/L, 약 8 g/L, 약 8.5 g/L, 약 9 g/L, 약 9.5 g/L, 약 10 g/L, 약 10.5 g/L, 약 11 g/L, 약 12 g/L, 약 0.2 g/L~약 0.5 g/L, 약 0.2 g/L~약 1 g/L, 약 0.2~약 2 g/L, 약 0.3 g/L~약 0.6 g/L, 약 0.3 g/L~약 1 g/L, 약 0.3~약 2 g/L, 약 0.4~약 0.7 g/L, 약 0.4~약 1 g/L, 약 0.4~약 2 g/L, 약 0.4~약 3 g/L, 약 0.5 g/L~약 1 g/L, 약 0.5 g/L~약 2 g/L, 약 0.5 g/L~약 3 g/L, 약 0.5 g/L~약 4 g/L, 약 0.5 g/L~약 5 g/L, 약 0.5 g/L~약 6 g/L, 약 0.5 g/L~약 7 g/L, 약 0.5 g/L~약 8 g/L, 약 0.5 g/L~약 9 g/L, 약 0.5 g/L~약 10 g/L, 약 0.5 g/L~약 11 g/L, 약 0.5 g/L~약 12 g/L, 약 1 g/L~약 2 g/L, 약 1 g/L~약 3 g/L, 약 1 g/L~약 4 g/L, 약 1 g/L~약 5 g/L, 약 1 g/L~약 6 g/L, 약 1 g/L~약 7 g/L, 약 1 g/L~약 8 g/L, 약 1 g/L~약 9 g/L, 약 1 g/L~약 10 g/L, 약 1 g/L~약 11 g/L, 약 1 g/L~약 12 g/L, 약 2 g/L~약 3 g/L, 약 2 g/L~약 4 g/L, 약 2 g/L~약 5 g/L, 약 2 g/L~약 6 g/L, 약 2 g/L~약 7 g/L, 약 2 g/L~약 8 g/L, 약 2 g/L~약 9 g/L, 약 2 g/L~약 10 g/L, 약 2 g/L~약 11 g/L, 약 2 g/L~약 12 g/L, 약 3 g/L~약 4 g/L, 약 3 g/L~약 5 g/L, 약 3 g/L~약 6 g/L, 약 3 g/L~약 7 g/L, 약 3 g/L~약 8 g/L, 약 3 g/L~약 9 g/L, 약 3 g/L~약 10 g/L, 약 3 g/L~약 11 g/L, 약 3 g/L~약 12 g/L, 약 4 g/L~약 5 g/L, 약 4 g/L~약 6 g/L, 약 4 g/L~약 7 g/L, 약 4 g/L~약 8 g/L, 약 4 g/L~약 9 g/L, 약 4 g/L~약 10 g/L, 약 4 g/L~약 11 g/L, 약 4 g/L~약 12 g/L, 약 5 g/L~약 6 g/L, 약 5 g/L~약 7 g/L, 약 5 g/L~약 8 g/L, 약 5 g/L~약 9 g/L, 약 5 g/L~약 10 g/L, 약 5 g/L~약 11 g/L, 약 5 g/L~약 12 g/L, 약 6 g/L~약 7 g/L, 약 6 g/L~약 8 g/L, 약 6 g/L~약 9 g/L, 약 6 g/L~약 10 g/L, 약 6 g/L~약 11 g/L, 약 6 g/L~약 12 g/L, 약 7 g/L~약 8 g/L, 약 7 g/L~약 9 g/L, 약 7 g/L~약 10 g/L, 약 7 g/L~약 11 g/L, 약 7 g/L~약 12 g/L, 약 8 g/L~약 9 g/L, 약 8 g/L~약 10 g/L, 약 8 g/L~약 11 g/L, 약 8 g/L~약 12 g/L, 약 9 g/L~약 10 g/L, 약 9 g/L~약 11 g/L, 약 9 g/L~약 12 g/L, 약 10 g/L~약 11 g/L, 약 10 g/L~약 12 g/L, 또는 약 11 g/L~약 12 g/L인 것인 제조 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 독소 단백질을 코딩하는 뉴클레오티드 서열이, 발현 시 독소 단백질의 주변세포질로의 이동을 지시하는 분비 신호 코딩 서열에 융합되어 있는 것인 제조 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 숙주 세포가 하나 이상의 프로테아제의 발현에 결함이 있는 것이거나, 또는 숙주 세포가 하나 이상의 폴딩 조절 인자를 과발현하는 것이거나, 또는 이들의 조합인 것인 제조 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 CRM197이고, 숙주 세포가 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결함이 있는 것인 제조 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 재조합 독소 단백질이 Azu, IbpS31A, CupA2, PbpA20V, 또는 Pbp인 분비 리더에 융합되어 있는 것인 제조 방법.

청구항 8

제1항 내지 제6항 중 어느 한 항에 있어서, 재조합 독소 단백질이 CRM197이고, 숙주 세포가 세라리신(Serralsin), HslU, HslV, Prc1, DegP1, DegP2, AprA, 또는 이들의 임의의 조합의 발현에 결함이 있는 것이거나, 또는 숙주 세포가 DsbA, DsbB, DsbC, 및 DsbD를 과발현하는 것이고, 추가로, 재조합 독소 단백질이 Azu, Pbp, 또는 네이티브(native) 분비 리더에 융합되어 있는 것인 제조 방법.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 숙주 세포가 DsbA, DsbB, DsbC, 및 DsbD를 과발현하는 것이고, 재조합 독소 단백질이 분비 리더 Azu에 융합되어 있는 것인 제조 방법.

청구항 10

제1항 내지 제8항 중 어느 한 항에 있어서, 숙주 세포가 세라리신의 발현에 결함이 있는 것이고, 재조합 독소 단백질이 분비 리더 Pbp 또는 Azu에 융합되어 있는 것인 제조 방법.

청구항 11

제1항 내지 제8항 중 어느 한 항에 있어서, 숙주 세포가 HslU 및 HslV의 발현에 결함이 있는 것이고, 재조합 독소 단백질이 분비 리더 Pbp 또는 Azu에 융합되어 있는 것인 제조 방법.

청구항 12

제1항 내지 제4항 중 어느 한 항에 있어서, 재조합 독소 단백질이 CRM197이고, 숙주 세포가 야생형이고, 재조합 독소 단백질이 분비 리더 Pbp 또는 Azu에 융합되어 있는 것인 제조 방법.

청구항 13

제1항 내지 제8항 중 어느 한 항에 있어서, 재조합 독소 단백질이 CRM197이고, 여기서 재조합 독소 단백질이 분비 리더 Azu, Pbp, IbpS31A, CupA2, 또는 PbpA20V에 융합되어 있는 것인 제조 방법.

청구항 14

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 콜레라 독소 B이고, 숙주 세포가 Lon, La, 및 AprA의 발현에 결함이 있는 것이거나; 또는 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결함이 있는 것이고, DegP2 S219A를 과발현하는 것인 제조 방법.

청구항 15

제1항 내지 제5항 및 제14항 중 어느 한 항에 있어서, 숙주 세포가 Lon, La, 및 AprA의 발현에 결함이 있는 것

이고, 재조합 독소 단백질이 분비 리더 Pbp A20V에 융합되어 있는 것이거나; 또는 숙주 세포가 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결함이 있는 것이고, DegP2 S219A를 과발현하는 것이며, 재조합 독소 단백질이 분비 리더 DsbA에 융합되어 있는 것인 제조 방법.

청구항 16

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 백일해 독소 S1 E129A R9K이고, 숙주 세포가 Lon, La, 및 AprA의 발현에 결함이 있는 것이거나; GrpE, DnaK, 및 DnaJ를 과발현하는 것이거나; HtpX의 발현에 결함이 있는 것이거나; RXF01590의 발현에 결함이 있는 것이거나; 또는 ppiB(RXF05345)의 발현에 결함이 있는 것인 제조 방법.

청구항 17

제1항 내지 제5항 및 제16항 중 어느 한 항에 있어서, 재조합 독소 단백질이 그의 네이티브 분비 리더에 융합되어 있는 것인 제조 방법.

청구항 18

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 파상풍 독소 C이고, 숙주 세포가 Lon, La 및 AprA의 발현에 결함이 있는 것이거나, 또는 숙주 세포가 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결함이 있는 것이거나, 또는 숙주 세포가 dsbABCD를 과발현하는 것이거나, 또는 숙주 세포가 GrpE, DnaK, 및 DnaJ를 과발현하는 것인 제조 방법.

청구항 19

제1항 내지 제5항 및 제18항 중 어느 한 항에 있어서, 재조합 독소 단백질이 분비 리더 TolB, DsbA, DsbC, Pbp A20V, NikA, 또는 CupA2에 융합되어 있는 것인 제조 방법.

청구항 20

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 파상풍 독소 C이고, 숙주 세포가 Lon, La, 및 AprA의 발현에 결함이 있는 것인 제조 방법.

청구항 21

제1항 내지 제5항, 제18항 및 제20항 중 어느 한 항에 있어서, 재조합 독소 단백질이 분비 리더 DsbA에 융합되어 있는 것인 제조 방법.

청구항 22

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 파상풍 독소 C이고, 숙주 세포가 GrpE, DnaK, 및 DnaJ의 발현에 결함이 있는 것인 제조 방법.

청구항 23

제1항 내지 제5항, 제18항, 제20항 및 제22항 중 어느 한 항에 있어서, 재조합 독소 단백질이 분비 리더 NikA에 융합되어 있는 것인 제조 방법.

청구항 24

제1항 내지 제5항 중 어느 한 항에 있어서, 재조합 독소 단백질이 C. 디피실레 독소 B이고, 숙주 세포가 HtpX의 발현에 결함이 있는 것이거나; DegP1의 발현에 결함이 있는 것이거나; HslU, HslV, Prc1 및 Prc2의 발현에 결함이 있는 것이거나; Lon, La, 및 DegP2의 발현에 결함이 있는 것이거나; 또는 숙주 세포가 Lon, Prc1, DegP2, AprA의 발현에 결함이 있는 것이고, DegP2 S219A를 과발현하는 것인 제조 방법.

청구항 25

제1항 내지 제24항 중 어느 한 항에 있어서, 재조합 독소 단백질의 활성을 활성 검정법으로 측정하는 단계를 추가로 포함하며, 여기서, 제조된 가용성 독소 단백질의 약 40%~약 100%가 활성을 띠는 것으로 측정되는 것인 제

조 방법.

청구항 26

제25항에 있어서, 활성 검정법이 면역학적 검정법, 수용체 결합 검정법, 또는 효소 검정법인 것인 제조 방법.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 발현 벡터가 단백질 코딩 서열에 작동 가능하게 연결된 lac 유도체 프로모터를 포함하고, 여기서, 배양 단계는 약 0.02 mM~약 1.0 mM 농도의 IPTG를 사용하여 프로모터를 유도하는 것을 포함하며, 유도 시 세포 밀도는 약 40~약 200 흡광 단위(AU)의 광학 밀도이고, 배양물의 pH는 약 6~약 7.5이며, 성장 온도는 약 20℃~약 35℃인 것인 제조 방법.

청구항 28

제1항 내지 제27항 중 어느 한 항에 있어서, 숙주 세포가 슈도모나스(*Pseudomonas*) 세포인 것인 제조 방법.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서, 숙주 세포가 슈도모나스 플루오레센스(*Pseudomonas fluorescens*)인 것인 제조 방법.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 뉴클레오티드 서열이 슈도모나스 숙주 세포에서의 발현을 위해 최적화된 것인 제조 방법.

청구항 31

제28항에 있어서, 뉴클레오티드 서열이 슈도모나스 숙주 세포에서의 발현을 위해 최적화된 것인 제조 방법.

청구항 32

제29항에 있어서, 뉴클레오티드 서열이 슈도모나스 플루오레센스 숙주 세포에서의 발현을 위해 최적화된 것인 제조 방법.

청구항 33

제1항 내지 제4항 및 제25항 내지 제32항 중 어느 한 항에 있어서, 재조합 독소 단백질이 백일해 독소이고, 여기서 백일해 독소가 야생형 또는 S1 E129A R9K인 것인 제조 방법.

청구항 34

제1항 내지 제4항 및 제25항 내지 제32항 중 어느 한 항에 있어서, 재조합 독소 단백질이 P. 아에루기노사 외독소 A이고, 여기서 P. 아에루기노사 외독소 A가 야생형, CRM66, 또는 rEPA인 것인 제조 방법.

청구항 35

제4항 내지 제34항 중 어느 한 항에 있어서, 발현 벡터가 분비 신호에 대한 코딩 서열에 인접한 태그 서열을 추가로 포함하는 것인 제조 방법.

청구항 36

제1항 내지 제35항 중 어느 한 항에 있어서, 발현 벡터가 독소 단백질에 대한 코딩 서열에 인접한 태그 서열을 추가로 포함하는 것인 제조 방법.

명 세 서

기술 분야

우선권 주장

[0001]

[0002] 본 출원은 35 U.S.C. § 119(e)하에 2010년 4월 16일에 출원된 미국 가출원 일련 번호 제61/325,235호, 2010년 4월 9일에 출원된 PCT/US10/30573, 및 2010년 3월 30일 출원된 미국 가출원 일련 번호 제61/319,152호를 우선권 주장한다. 이들 가출원은 그 전문이 본원에 참고로 포함된다.

[0003] 서열 목록

[0004] 본 출원은 EFS-Web을 통해 ASCII 포맷으로 제출되고, 본 발명에 그 전문이 본원에 참고로 포함된 서열 목록을 포함한다. 2011년 3월 16일에 작성된 상기 ASCII 복사본의 파일명은 38194201.txt이고, 그 크기는 156,975 바이트이다.

배경 기술

[0005] 미생물 독소 단백질은 약제에서는 독소 생산 미생물에 대한 백신접종을 위한 면역원으로서 및 다른 백신을 위한 캐리어 단백질 및 애주번트로서, 및 과학 연구에서는 분자 경로를 연구하는 도구로서 사용된다.

[0006] 디프테리아 독소(DT: Diphtheria toxin)는 독소생산 균주인 코리네박테리움 디프테리아에(*Corynebacterium diphtheriae*)가 합성 및 분비하는 단백질성 독소이다. 독소생산 균주는 독소 유전자를 보유하는 박테리오파지 용원을 포함한다. DT는 535 아미노산 폴리펩티드로서 합성되고, 단백질분해를 통해 성숙한 독소로 형성된다. 성숙한 독소는 이황화 브릿지에 의해 연결된 2개의 서브유닛 A 및 B를 포함한다. 온전한 DT의 C 말단부로부터 형성된 B 서브유닛은 DT가 결합한 후, 세포막을 통해 세포질로 진입할 수 있게 한다. 세포 진입시, 온전한 DT의 N 말단부로부터 형성된 효소 A 서브유닛은 신장 인자 2(EF-2: Elongation Factor 2)의 ADP 리보실화를 촉진한다. 그 결과, EF-2는 불활성화되고, 단백질 합성이 중단되어, 세포는 사멸하게 된다. 디프테리아 독소는 세포독성이 매우 높으며; 단일 분자조차 세포에는 치명적일 수 있고, 10 ng/kg의 용량은 동물 및 인간을 치사시킬 수 있다.

[0007] CRM197 단백질은 DT의 비독성의 면역학적 교차반응 형태이다. 이 단백질은 DT 추가접종 또는 백신 항원으로서의 잠재적 용도에 대해 연구되고 있다. CRM197은 독소생산 코리네박테리움 β의 니트로소구아니딘 돌연변이유발에 의해 형성된 비독소생산 파지 β 197^{tox-}에 의해 감염된 C. 디프테리아에에 의해 생산된다. CRM197 단백질은 DT와 분자량은 같지만, A 서브유닛의 단일 염기 변화(구아닌에서 아데닌)에 차이가 있다. 단일 염기 변화는 아미노산 치환(글리신의 경우 글루탐산)을 유발하여 DT의 독성 특성을 제거한다.

[0008] 캐리어 단백질로서 CRM197을 사용하는 컨쥬게이트된 다당류 백신이 인간용으로 승인을 받았다. 백신으로는 네이세리아 메닝기티디스(*Neisseria meningitidis*) 서브군 A, C, Y 및 W-135에 의해 유발된 침입성 수막구균 질환 예방용으로 지시되는 백신인 멘비오(Menveo)[®](Novartis Vaccines and Diagnostics); 수막구균 그룹 C 컨쥬게이트 백신인 멘쥬게이트(Menjugate)(Novartis Vaccines); 및 스트렙토코커스 뉴모니아에(*Streptococcus pneumoniae*)의 7가지 혈청형을 표적으로 하는 소아기 뉴모니아 백신인 프리베나(Prevnar)[®](Wyeth Pharmaceuticals, Inc.) 및 헤모필루스 인플루엔자에(*Haemophilus influenzae*) b형 백신인 HibTITER[®](Wyeth)를 포함한다. 또한, CRM197은 C. 디프테리아 백신접종을 위한 추가 항원으로서의 잠재적 용도를 가지며, 다른 백신에 사용하기 위한 캐리어 단백질로서 연구되고 있다.

[0009] 승인받은 치료제 및 조사 용도를 위한 CRM197을 고수준으로 발현시키는 방법에 대해서는 보고된 바 없다. CRM197은 예를 들어, C. 디프테리아에(*C. diphtheriae*), B. 섭틸리스(*B. subtilis*), 및 E. 콜라이(*E. coli*)에서 수십 mg/L 범위의 수준으로 발현되었었다. 단일 용량의 프리베나 컨쥬게이트 백신은 약 20 μg의 CRM197을 함유한다. 그러므로, CRM197을 약 1 g/L 이상의 수준으로 경제적으로 제조하는 방법이 백신 연구 및 제조를 매우 용이하게 만들 것이다.

[0010] 설사와 구토를 특징으로 하는 감염을 유발하는 박테리아 병원균인 비브리오 콜레라(*Vibrio cholera*)에 의해 생산되는 콜레라 독소(CTX: Cholera Toxin) 또한 ADP 리보실화 독소이다. CTX는, 단일 카피의 콜레라 독소 A 서브유닛(CTA: Cholera toxin A subunit)과 5 카피의 콜레라 독소 B 서브유닛(CTB: Cholera toxin B subunit)로 이루어진 6개의 단백질 서브유닛으로 구성된 올리고머 복합체이다. 각각의 분자량이 12 kDa인 5개의 B 서브유닛이 5원 환을 형성한다. A 서브유닛은 G 단백질을 ADP 리보실화하는 구상 효소인 A1부인 CTA1, 및 B 서브유닛 환의 중심 포어에 편안하게 안치된 연장형 알파 나선을 형성하는 A2 쇠인 CTA2를 가진다. 상기 환은 숙주 세포 표면에 존재하는 GM1 ganglioside 수용체에 결합하여 전체 복합체를 내재화시킨다. 일단 내재화되고 나면, 이황화 브릿지 환원에 의해 CTA1 쇠가 유리된다. 이어서, CTA1은 활성화되고, 아데닐레이트 사이클라제에 의한

ADP 리보실화를 촉매화한다. 그 결과, 아데닐레이트 사이클라제 활성이 증가하면 사이클릭 AMP 합성도 증가하게 되고, 이로써 대량의 체액 및 전해질 유출이 유발됨으로써 설사가 일어나게 된다.

[0011] 상대적으로 무해하기는 하나, CTX의 B 서브유닛은 GM1 강글리오사이드 수용체에 결합할 수 있는 그의 능력을 보유한다. 그러므로, CTB는 화학적으로 또는 유전적으로 컨쥬게이트된 외래 항원의 점막 흡수를 촉진시키는 데 사용되는 것으로 알려져 있다. 점막 및 전신 면역, 둘 모두를 유도하는 것으로 입증되어 있으며, 식용 백신 생산에 사용하기 위한 후보물질이 된다. CTB는 그의 결합상의 선호도에 기인한 뉴런의 추적자로서의 용도도 알려져 있다.

[0012] 백일해 독소(PTX: Pertussis toxin)는 백일해라는 질환을 유발하는 인간 기도의 박테리아 병원균인 보르데텔라 페르투스시스(*Bordetella pertussis*)에 의해 생산되는 외독소이자 발병 인자이다. 백일해 홀로톡신은 AB 5 구조를 가진 다중 서브유닛 복합체이다. 효소적으로 활성을 띠는 A 서브유닛(S1)이 포유동물 세포에서 수개의 이중삼량체 G 단백질을 변형시키는 ADP 리보실트랜스퍼라제이고, B 올리고머(S2, S3, 2 개의 S4, 및 S5)는 세포 상의 글리코컨쥬게이트 수용체에 결합한다. 독소의 5개의 서브유닛은 백일해 독소이드 오페론으로부터 발현된다.

[0013] 백일해 독소의 비독성 변이체가 방어용 백신에서 및 백신 애주번트로서의 사용을 위해 조사되어 왔다. 또한, 예를 들어, G 단백질 신호전달 경로 연구와 같은 연구에서 사용하기 위한 백일해 독소 단백질이 요구되고 있다.

[0014] 클로스트리디움 테타니(*Clostridium tetani*)에 의해 생산된 파상풍 독소는 분자량이 150 kDa인, 신경독소이다. 이는 2개의 부분: 100 kDa의 중쇄 또는 B쇄 및 50 kDa의 경쇄 또는 A쇄로 구성된다. 상기 쇠는 이황화 결합에 의해 연결되어 있다. B쇄는 뉴런 막 상의 디시알로강글리오사이드(GD2 및 GD1b)에 결합한다. 아연 엔도펩티다제인 A쇄는 소포체 결합 막 단백질(VAMP: vesicle-associated membrane protein)을 공격한다.

[0015] A쇄의 작용은 단백질 시냅토프레빈을 분해함으로써 영향을 받은 뉴런이 억제성 신경전달물질인 GABA(감마-아미노부티르산) 및 글리신을 유리시키는 것을 중단시킨다. 이러한 결과는 감각 자극에 의한 운동 반사 억제 실패라는 가작 작은 자극으로부터 발생하는 근육에서의 위험한 과다활성이다. 이는 파상풍 연축이라 명명되는, 주동근 및 길항근 조직의 전신 수축을 유발한다.

[0016] 파상풍 독소 단편 C(Tet C 또는 TTC)는 파상풍 독소의 (예를 들어, 파파인에 의한) 프로테아제 절단에 의해, 또는 상기 단편의 재조합 발현을 통해 생성된 50 kD 폴리펩티드이다. 이는 C 말단의 451 아미노산(아미노산 위치 865-1,315)에 상응한다.

[0017] 단편 C는 비독성인 것으로 밝혀져 있다. 이는 뉴런에 고 특이도 및 고 친화도로 결합하기 때문에, TTC는 뉴런의 약물 전달을 위한, 또는 연구 목적용의 표적한 분자로서의 용도가 있는 것으로 알려져 있다. TTC 단백질은 또한 잠재적으로는 백신 캐리어 단백질로서, 및 C. 테타니(*C. tetani*) 감염으로부터 방어하기 위한 백신에서 사용하기 위한 것으로 조사되어 왔다

[0018] 클로스트리디움 디피실레 독소 B(TcdB: *Clostridium difficile* Toxin B)는 병원내 설사 및 위막성 대장염을 유발하는 클로스트리디움 디피실레(*Clostridium difficile*)에 의해 생산되는 발병 인자이다. TcdB, 및 두번째로 큰 클로스트리디움계 독소, TcdA는 위막성 대장염 발병에 관여한다.

[0019] TcdB는 약 270 kD의 글리코실화 독소이며, 이는 효소 도메인, 전위 도메인, 및 수용체 결합 도메인으로 구분될 수 있다. TcdB의 처음 546 아미노산은 효소 영역을 포함하고, 그 다음으로 추정 전위 도메인 및 수용체 결합 도메인 순으로 포함한다. TcdB는 C 디피실레 감염에 대한 방어용 백신으로서 뿐만 아니라, 진단 테스트 및 그의 개발에서 잠재적인 용도를 가진다.

[0020] 슈도모나스 아에루기노사(*Pseudomonas aeruginosa*)의 외독소 A(ETA: Exotoxin A 또는 PE)는 II형 ADPRT이다. 그의 패밀리 구성원 디프테리아 독소 및 콜레라 독소와 같이, 세포 신장 인자 2의 ADP 리보실화에 의해 단백질 합성을 억제시킨다. P. 아에루기노사(*P. aeruginosa*) 외독소 A는 613 아미노산으로 된 단일 폴리펩티드 쇠 (66 Kd)로 구성된 단량체로서 존재한다.

[0021] ETA는 잠재적으로 백신 컨쥬게이트로서 유용하다. ETA의 비독성 돌연변이체는 스타필로코커스 아우레우스(*Staphylococcus aureus*), 말라리아, 및 살모넬라 타이피(*Salmonella Typhi*)에 대해 방어하는 백신접종의 백신 컨쥬게이트로서 연구되어 왔다.

[0022] 확대된 요구를 충족시키는 데 충분한 양으로 상기와 같은 독소를 제조하는 것이 중요한 도전 과제를 제시하였다. 종래의 단백질 과발현 시스템에서 제조될 경우, 독소 단백질은 독소의 특이적인 특징들, 예를 들어, 크기 및 2차 구조에 따른 분해, 부적절한 폴딩, 또는 그 둘 모두에 기인하여 단지 극히 낮은 농도로만 활

성 형태로 회수된다. 그러므로, 저비용으로 상기 독소를 가용성 및/또는 활성 형태로 대량 제조하는 방법이 요구되고 있다.

발명의 내용

- [0023] 본 발명은 독소 단백질을 코딩하는 뉴클레오티드 서열을 발현 벡터에 결찰시키는 단계; 슈도모나스 속주 세포를 상기 발현 벡터로 형질전환시키는 단계; 및 재조합 독소 단백질의 발현에 적합한 배양 배지 중에서 상기 형질전환된 슈도모나스 속주 세포를 배양하는 단계를 포함하며; 여기서, 재조합 독소 단백질이 CRM197, 디프테리아 독소, 콜레라 홀로톡신, 콜레라 독소 B, 백일해 독소, 파상풍 독소 단편 C, C. 디피실레(*C. difficile*) 독소 B, 또는 P. 아에루기노사 외독소 A인 것인, 슈도모나드(*Pseudomonad*) 속주 세포에서 재조합 독소 단백질을 제조하는 방법에 관한 것이다.
- [0024] 실시양태에서, 재조합 독소 단백질은 콜레라 독소 B, 콜레라 홀로톡신, 백일해 독소, 파상풍 독소 단편 C, C. 디피실레 독소 B, 또는 P. 아에루기노사 외독소 A이다.
- [0025] 다른 실시양태에서, 재조합 독소 단백질은 콜레라 독소 B, 콜레라 홀로톡신, 백일해 독소, 파상풍 독소 단편 C, 또는 C. 디피실레 독소 B이다.
- [0026] 다른 실시양태에서, 재조합 독소 단백질은 CRM197, 디프테리아 독소, 콜레라 홀로톡신, 콜레라 독소 B, 백일해 독소, 파상풍 독소 단편 C, 또는 C. 디피실레 독소 B이다.
- [0027] 특정 실시양태에서, 재조합 단백질은 약 0.2 g/L~약 12 g/L의 가용성 및/또는 활성 독소 단백질의 수율로 제조된다. 구체적인 실시양태에서, 가용성 및/또는 활성 독소 단백질의 수율은 약 0.2 g/L, 약 0.3 g/L, 약 0.4 g/L, 약 0.5 g/L, 약 0.6 g/L, 약 0.7 g/L, 약 0.8 g/L, 약 0.9 g/L, 약 1 g/L, 약 1.5 g/L, 약 2 g/L, 약 2.5 g/L, 약 3 g/L, 약 3.5 g/L, 약 4 g/L, 약 4.5 g/L, 약 5 g/L, 약 5.5 g/L, 약 6 g/L, 약 6.5 g/L, 약 7 g/L, 약 7.5 g/L, 약 8 g/L, 약 8.5 g/L, 약 9 g/L, 약 9.5 g/L, 약 10 g/L, 약 10.5 g/L, 약 11 g/L, 약 12 g/L, 약 0.2 g/L~약 0.5 g/L, 약 0.2 g/L~약 1 g/L, 약 0.2~약 2 g/L, 약 0.3 g/L~약 0.6 g/L, 약 0.3 g/L~약 1 g/L, 약 0.3~약 2 g/L, 약 0.4~약 0.7 g/L, 약 0.4~약 1 g/L, 약 0.4~약 2 g/L, 약 0.4~약 3 g/L, 약 0.5 g/L~약 1 g/L, 약 0.5 g/L~약 2 g/L, 약 0.5 g/L~약 3 g/L, 약 0.5 g/L~약 4 g/L, 약 0.5 g/L~약 5 g/L, 약 0.5 g/L~약 6 g/L, 약 0.5 g/L~약 7 g/L, 약 0.5 g/L~약 8 g/L, 약 0.5 g/L~약 9 g/L, 약 0.5 g/L~약 10 g/L, 약 0.5 g/L~약 11 g/L, 약 0.5 g/L~약 12 g/L, 약 1 g/L~약 2 g/L, 약 1 g/L~약 3 g/L, 약 1 g/L~약 4 g/L, 약 1 g/L~약 5 g/L, 약 1 g/L~약 6 g/L, 약 1 g/L~약 7 g/L, 약 1 g/L~약 8 g/L, 약 1 g/L~약 9 g/L, 약 1 g/L~약 10 g/L, 약 1 g/L~약 11 g/L, 약 1 g/L~약 12 g/L, 약 2 g/L~약 3 g/L, 약 2 g/L~약 4 g/L, 약 2 g/L~약 5 g/L, 약 2 g/L~약 6 g/L, 약 2 g/L~약 7 g/L, 약 2 g/L~약 8 g/L, 약 2 g/L~약 9 g/L, 약 2 g/L~약 10 g/L, 약 2 g/L~약 11 g/L, 약 2 g/L~약 12 g/L, 약 3 g/L~약 4 g/L, 약 3 g/L~약 5 g/L, 약 3 g/L~약 6 g/L, 약 3 g/L~약 7 g/L, 약 3 g/L~약 8 g/L, 약 3 g/L~약 9 g/L, 약 3 g/L~약 10 g/L, 약 3 g/L~약 11 g/L, 약 3 g/L~약 12 g/L, 약 4 g/L~약 5 g/L, 약 4 g/L~약 6 g/L, 약 4 g/L~약 7 g/L, 약 4 g/L~약 8 g/L, 약 4 g/L~약 9 g/L, 약 4 g/L~약 10 g/L, 약 4 g/L~약 11 g/L, 약 4 g/L~약 12 g/L, 약 5 g/L~약 6 g/L, 약 5 g/L~약 7 g/L, 약 5 g/L~약 8 g/L, 약 5 g/L~약 9 g/L, 약 5 g/L~약 10 g/L, 약 5 g/L~약 11 g/L, 약 5 g/L~약 12 g/L, 약 6 g/L~약 7 g/L, 약 6 g/L~약 8 g/L, 약 6 g/L~약 9 g/L, 약 6 g/L~약 10 g/L, 약 6 g/L~약 11 g/L, 약 6 g/L~약 12 g/L, 약 7 g/L~약 8 g/L, 약 7 g/L~약 9 g/L, 약 7 g/L~약 10 g/L, 약 7 g/L~약 11 g/L, 약 7 g/L~약 12 g/L, 약 8 g/L~약 9 g/L, 약 8 g/L~약 10 g/L, 약 8 g/L~약 11 g/L, 약 8 g/L~약 12 g/L, 약 9 g/L~약 10 g/L, 약 9 g/L~약 11 g/L, 약 9 g/L~약 12 g/L, 약 10 g/L~약 11 g/L, 약 10 g/L~약 12 g/L, 또는 약 11 g/L~약 12 g/L이다.
- [0028] 실시양태에서, 독소 단백질을 코딩하는 뉴클레오티드 서열은, 발현 시 독소 단백질의 주변세포질로의 이동을 지시하는 분비 신호 코딩 서열에 융합되어 있다. 실시양태에서, 속주 세포는 하나 이상의 프로테아제의 발현에 결합이 있는 것이거나, 또는 속주 세포는 하나 이상의 폴딩 조절 인자를 과발현하는 것이거나, 또는 이들의 조합이다.
- [0029] 실시양태에서, 재조합 독소 단백질은 CRM197이고, 속주 세포는 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결합이 있는 것이다. 관련 실시양태에서, 재조합 독소 단백질은 Azu, IbpS31A, CupA2, PbpA20V, 또는 Pbp인 분비 리더에 융합되어 있다. 실시양태에서, 재조합 독소 단백질은 CRM197이고, 속주 세포는 HslU 및 HslV, 또는 Prc1, 또는 DegP1, 또는 DegP2, 또는 AprA의 발현에 결합이 있는 것이다. 구체적인 실시양태에서, 재조합 독소 단백질은 CRM197이고, 속주 세포는 세라리신(Serralysin), HslU, HslV, Prc1, DegP1, DegP2, 또는 AprA의

발현에 결함이 있는 것이거나, 또는 DsbA, DsbB, DsbC, 및 DsbD를 과발현하는 것이다. 실시양태에서, 숙주 세포는 DsbA, DsbB, DsbC, 및 DsbD를 과발현하는 것이고, 재조합 독소 단백질은 분비 리더 Azu에 융합되어 있다. 실시양태에서, 숙주 세포는 세라리신의 발현에 결함이 있는 것이고, 재조합 독소 단백질은 분비 리더 Pbp 또는 Azu에 융합되어 있다. 실시양태에서, 숙주 세포는 HslU 및 HslV의 발현에 결함이 있는 것이고, 재조합 독소 단백질은 분비 리더 Pbp 또는 Azu에 융합되어 있다. 실시양태에서, 재조합 독소 단백질은 CRM197이고, 숙주 세포는 야생형이고, 여기서, 재조합 독소 단백질은 분비 리더 Pbp 또는 Azu에 융합되어 있다. 실시양태에서, 재조합 독소 단백질은 CRM197이고, 재조합 독소 단백질은 분비 리더 Azu, Pbp, IbpS31A, CupA2, 또는 PbpA20V에 융합되어 있다.

[0030] 다른 실시양태에서, 재조합 독소 단백질은 콜레라 독소 B이고, 숙주 세포는 Lon, La, 및 AprA의 발현에 결함이 있는 것이거나, 또는 숙주 세포는 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결함이 있는 것이다. 관련 실시양태에서, 숙주 세포는 Lon, La, 및 AprA의 발현에 결함이 있는 것이고, 재조합 독소 단백질은 분비 리더 Pbp A20V에 융합되어 있다.

[0031] 다른 실시양태에서, 재조합 독소 단백질은 백일해 독소 S1 E129A R9K이고, 숙주 세포는 Lon, La, 및 AprA; GrpE, DnaK, 및 DnaJ; HtpX; RXF01590; 또는 ppiB(RXF05345)의 발현에 결함이 있는 것이다. 관련 실시양태에서, 재조합 독소 단백질은 그의 네이티브(native) 분비 리더에 융합되어 있다.

[0032] 다른 실시양태에서, 재조합 독소 단백질은 파상풍 독소 C이고, 숙주 세포는 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA의 발현에 결함이 있는 것이다. 관련 실시양태에서, 재조합 독소 단백질은 분비 리더 DsbC, Pbp A20V, 또는 CupA2에 융합되어 있다.

[0033] 다른 실시양태에서, 재조합 독소 단백질은 파상풍 독소 C이고, 숙주 세포는 Lon, La, 및 AprA의 발현에 결함이 있는 것이다. 관련 실시양태에서, 재조합 독소 단백질은 분비 리더 DsbA에 융합되어 있다.

[0034] 다른 실시양태에서, 재조합 독소 단백질은 파상풍 독소 C이고, 숙주 세포는 GrpE, DnaK, 및 DnaJ의 발현에 결함이 있는 것이다. 관련 실시양태에서, 재조합 독소 단백질은 분비 리더 NikA에 융합되어 있다.

[0035] 다른 실시양태에서, 재조합 독소 단백질은 C. 디피실레 독소 B이고, 숙주 세포는 HtpX; DegP1; HslU, HslV, Prc1 및 Prc2; 또는 Lon 및 DegP2에 결함이 있는 것이거나, 숙주 세포는 Lon, Prc1, DegP2, AprA에 결함이 있는 것이면서, 동시에 DegP2 S219A를 과발현하는 것이다.

[0036] 실시양태에서, 재조합 독소 단백질의 활성은 활성 검정법으로 측정되는데, 여기서, 제조된 가용성 독소 단백질의 약 40%~약 100%가 활성을 띠는 것으로 측정된다. 관련 실시양태에서, 활성 검정법은 면역학적 검정법, 수용체 결합 검정법, 또는 효소 검정법이다.

[0037] 본 발명의 실시양태에서, 발현 벡터는 단백질 코딩 서열에 작동 가능하게 연결된 lac 유도체 프로모터를 포함하고, 여기서, 배양 단계는 약 0.02 mM~약 1.0 mM 농도의 IPTG를 사용하여 프로모터를 유도하는 것을 포함하며, 유도 시 세포 밀도는 약 40~약 200 흡광 단위(AU: absorbance unit)의 광학 밀도이고, 배양물의 pH는 약 6~약 7.5이며, 성장 온도는 약 20℃~약 35℃이다.

[0038] 실시양태에서, 숙주 세포는 슈도모나스 세포이다. 관련 실시양태에서, 숙주 세포는 슈도모나스 플루오레센스(*Pseudomonas fluorescens*)이다.

[0039] 본 발명의 실시양태에서, 뉴클레오타이드 서열은 슈도모나스 숙주 세포에서 발현될 수 있도록 최적화되었다. 관련 실시양태에서, 뉴클레오타이드 서열은 슈도모나스 숙주 세포에서 발현될 수 있도록 최적화되었다. 다른 관련 실시양태에서, 뉴클레오타이드 서열은 슈도모나스 플루오레센스 숙주 세포에서 발현될 수 있도록 최적화되었다.

[0040] 실시양태에서, 백일해 독소는 야생형 또는 S1 E129A R9K이다. 실시양태에서, P. 아에루기노사 외독소 A는 야생형, CRM66, 또는 rEPA이다.

[0041] 본 발명의 실시양태에서, 발현 벡터는 분비 신호에 대한 코딩 서열에 인접한 태그 서열을 추가로 포함한다. 실시양태에서, 발현 벡터는 독소 단백질에 대한 코딩 서열에 인접한 태그 서열을 추가로 포함한다.

[0042] 본 발명은 또한 본원에 기술된 방법에 따라 제조된 재조합 독소 단백질을 제공한다. 실시양태에서, 재조합 독소 단백질은 CRM197, 디프테리아 독소, 콜레라 홀로톡신, 콜레라 독소 B, 백일해 독소, 파상풍 독소 단편 C, C. 디피실레 독소 B, 또는 P. 아에루기노사 외독소 A이다. 실시양태에서, 외독소 A는 야생형, CRM66, 또는 rEPA이다. 특정 실시양태에서, 재조합 독소 단백질은 본원에서 고수율로 독소를 생산하거나, 또는 고 정질의 독소를 생산

하는 것으로 확인된 P. 플루오레센스 균주에서 제조된다. 특정 실시양태에서, 제조합 독소 단백질은 본원에서 고수율로 독소를 생산하는 것으로 기술된 P. 플루오레센스 균주에서 제조된다. 다른 실시양태에서, 제조합 독소 단백질은 본원에서 독소의 발효 생산을 위해 사용되는 것으로 기술된 균주에서 제조된다.

[0043] **참고 문헌 인용**

[0044] 본 명세서 중에 언급된 모든 공개문헌, 특허 및 특허 출원은, 각 개별 공개문헌, 특허 또는 특허 출원이 구체적이고 개별적으로 참고로 포함되었다고 제시된 것과 동일한 정도로 본원에 참고로 포함된다.

도면의 간단한 설명

[0045] 본 발명의 신규한 특징은 하기 특허청구범위에서 구체적으로 설명된다. 본 발명의 원리를 이용한 예시적 실시양태를 설명하는 하기 상세한 설명과 첨부된 도면을 참고하면 본 발명의 특징 및 이점을 더욱 명확하게 이해할 수 있을 것이다.

도 1. CRM197의 고처리량 발현 분석. 도 1B에 제시된 DNA 서열을 사용하여 발현된 CRM197 단백질을 모세관 겔 전기영동(SDS-CGE)을 사용하여 분석하였다. 테스트된 40종의 CRM197 발현 균주의 가용성 분획은 SDS-CGE 데이터로부터 작성된 겔 유사 이미지로 제시되어 있다. 하기 표 10에 기술되어 있는 균주명이 각 레인 상에 열거되어 있다. P. 플루오레센스 발현 CRM197은 SDS-CGE에서 ~58 kDa의 단일 밴드로서 이동하였다(좌측 화살표). 분자량 마커는 첫번째 레인부터 마지막 레인까지 순서대로 16, 20, 29, 48, 69 및 119 kDa이다.

도 2. 콜레라 독소 B의 고처리량 발현 분석. 서열 번호: 23에 제시된 DNA 서열을 사용하여 발현된 콜레라 독소 B 단백질을 모세관 겔 전기영동(SDS-CGE)을 사용하여 분석하였다. 테스트된 40종의 콜레라 독소 발현 균주의 가용성 분획은 SDS-CGE 데이터로부터 작성된 겔 유사 이미지로 제시되어 있다. 하기 표 11에 기술되어 있는 균주명이 각 레인 상에 열거되어 있다. 유도된 CTB는 SDS-CGE에서 ~11.5 kDa의 단일 밴드로서 이동하였다(좌측 화살표). 분자량 마커는 첫번째 레인부터 마지막 레인까지 순서대로 16, 20, 29, 48, 69 및 119 kDa이다.

도 3. 백일해 특소이드 오페론. 4,210 염기쌍을 가지는 BPETOX_S1-R9K & E129A가 제시되어 있다.

도 4. 백일해 특소이드의 DNA 서열. 번역과 함께 백일해 독소 S1 R9K E129A DNA 서열이 제시되어 있다(서열 번호: 24). 서열은 진뱅크 엔트리 M13223으로부터 유래된 것이다. 서브유닛 S1-S5 및 신호 서열은 서열 위에 표시되어 있다. S1 중 R9K 및 E129A 돌연변이는 밑줄로 표시되어 있다. 코딩된 단백질은 출현 순서대로 서열 번호 25, 26, 28, 29, 및 27로 개시되어 있다.

도 5. 백일해 특소이드 서브유닛의 아미노산 서열. 분비 신호는 밑줄로 표시되어 있다. **A.** S1 서브유닛(R9K E129A)(서열 번호: 25). **B.** S2 서브유닛(서열 번호: 26). **C.** S3 서브유닛(서열 번호: 27). **D.** S4 서브유닛(서열 번호: 28). **E.** S5 서브유닛(서열 번호: 29).

도 6. 백일해 특소이드 발현 샘플의 웨스턴 블롯 분석. 균주명이 각 레인 상에 열거되어 있다. 유도된 Ptx는 11 ~26 kDa 범위의 다중 밴드로서 이동하였다(S1: 26.1 Kda, S2: 20.9 Kda, S3: 21.8 KDa, S4(2x): 12 KDa, S5: 11 KDa). **A.** 환원성 샘플. **B.** 비환원성 샘플. 두 패널 모두: 레인 1 - 분자량 마커(10, 15, 20, 25, 37, 50, 75, 100, 150, 250 kDa); 레인 2 - 무효값; 레인 3 - 균주 321; 레인 4 - 균주 322; 레인 5 - 균주 323; 레인 6 - 균주 324; 레인 7 - 균주 325; 레인 8 - 균주 326; 레인 9 - 균주 327; 레인 10 - 균주 328.

도 7. 파상풍 독소 C 단편 발현. P. 플루오레센스에서 발현된 파상풍 독소 C 단편을 모세관 겔 전기영동(SDS-CGE)을 사용하여 분석하였다. 테스트된 40종의 파상풍 독소 발현 균주로부터의 가용성 분획은 SDS-CGE 데이터로부터 작성된 겔 유사 이미지로 제시되어 있다. 하기 표 15에 기술되어 있는 균주명이 각 레인 상에 열거되어 있다. 유도된 파상풍 독소 C 단편은 SDS-CGE에서 ~51.6 kDa의 단일 밴드로서 이동하였다(좌측 화살표). 분자량 마커는 첫번째 레인부터 마지막 레인까지 순서대로 16, 20, 29, 48, 69 및 119 kDa이다.

도 8. TcdB 발현. P. 플루오레센스에서 발현된 TcdB를 모세관 겔 전기영동(SDS-CGE)을 사용하여 분석하였다. 테스트된 24종의 TcdB 발현 균주로부터의 가용성 분획은 SDS-CGE 데이터로부터 작성된 겔 유사 이미지로 제시되어 있다. 하기 표 18에 기술되어 있는 균주명 뿐만 아니라, 무효 추출값 및 기준 표준(List Biologicals)도 각 레인 상에 열거되어 있다. 유도된 TcdB는 SDS-CGE에서 ~300 kDa의 단일 밴드로서 이동하였다(좌측 화살표). 분자량 마커는 첫번째 레인부터 마지막 레인까지 순서대로 16, 20, 29, 48, 69 및 119 kDa이다.

도 9. 외독소 A 아미노산 서열. P. 아에루기노사 외독소 A의 아미노산 서열이 제시되어 있다(서열 번호: 34). 3 가지 외독소 A 단백질: 야생형, CRM66, 및 rEPA가 도면에 제시되어 있다. 변이체 CRM66에서, His 426(굵은 밑줄

체로 표시)은 서열 상에 표시되어 있는 바와 같이, Tyr로 치환되어 있다. rEPA에서, Glu 553(붉은 밑줄체로 표시)은 서열 상에 표시되어 있는 바와 같이, 결실되어 있다.

도 10. P. 플루오레센스 발효 배양물 중에서의 가용성 파상풍 독소 C 및 콜레라 독소 B 제조. SDS-CGE 분석. 레인 1 - 16, 20, 29, 48, 69 및 119 kDa 분자량 마커. 우측 화살표로 표시된 바와 같이, 각각 레인 2 및 4 - 콜레라 독소 B를 발현하는 PS538-088 U5 및 U6 발효물의 유도 전 샘플 및 레인 3 및 5 - 콜레라 독소 B를 발현하는 PS538-088 U5 및 U6 발효물의 유도 후 샘플.

도 11. P. 플루오레센스 발효 배양물 중에서의 가용성 파상풍 독소 단편 C 제조. A. SDS-CGE 분석. 레인 1 - 16, 20, 29, 48, 69 및 119 kDa 분자량 마커. 레인 2, 3 및 4는 우측 화살표로 표시된 바와 같이, 각각, 파상풍 독소 단편 C를 발현하는 PS538-529 U1, PS538-546 U5 및 PS538-547 U7 발효물의 유도 후 샘플이다. B. 웨스턴 블롯 분석. 균주 PS538-538(U1 및 U2), PS538-548(U3 및 U4), PS538-558(U5 및 U6) 및 PS538-568(U7 및 U8)로부터의 발효 샘플을 웨스턴 블롯으로 평가하였다. 발효물 단위 및 유도 후 경과시간(hr)(I0, I8, I24)은 각 레인 위에 열거하였다. 분자량(MW) 표준은 블롯 좌측에 제시되어 있고, 파상풍 독소 C 기준 표준(Std; List Biological, 카탈로그 번호 193)은 우측에 제시되어 있다. 블롯을 래빗에서 유래된 폴리클로날 항파상풍 독소 C 단편(Abeam, 카탈로그 번호: ab34890)으로 프로빙한 후, 염소에서 유래된 항래빗 IgG 퍼옥시다제(Pierce, 카탈로그 번호: 31460)로 프로빙하였다. 검출하는 데 이뮤노퓨어 메탈 인핸스드 DAB(Immunopure Metal Enhanced DAB)(Pierce 34065)를 사용하였다.

도 12. 플루오레센스 발효 배양물에서의 가용성 C. 디피실레 B 독소 단백질 제조. 레인 1 - 16, 20, 29, 48, 69 및 119 kDa 분자량 마커. 마커 크기 또한 레인 1에서의 이동에 기초하여 우측의 각 위치에 표시되어 있다. 레인 2, 3 및 4는 우측 화살표로 표시된 바와 같이, 각각, C. 디피실레 B 독소 단백질을 발현하는 PS538-671 U5 및 U6, 및 PS538-674 U7 발효물의 유도 후 샘플이다.

도 13. 야생형 백일해 독소이드의 DNA 서열. 번역과 함께 야생형 백일해 독소 DNA 서열이 제시되어 있다(서열 번호: 35). 서열은 진뱅크 엔트리 M13223으로부터 유래된 것이다. 서브유닛 S1-S5 및 신호 서열은 서열 위에 표시되어 있다. 코딩된 단백질은 출현 순서대로 서열 번호 41-45로 개시되어 있다.

도 14. 콜레라 홀로톡신의 아미노산 및 DNA 서열. A. CTA 아미노산 서열(서열 번호: 38)(분비 리더는 밑줄로 표시)(AE003852; 단백질 ID AAF94614.1). B. CTB 아미노산 서열 (서열 번호: 39)(분비 리더는 밑줄로 표시)(진뱅크 AE003852; 단백질 ID AAF94613.1). C. 번역과 함께 제시된, A 및 B 서브유닛을 나타내는 CTX DNA 서열(서열 번호: 40)(진뱅크 AE003852). 코딩된 단백질은 출현 순서대로 서열 번호 38 및 39로 개시되어 있다.

도 15. P. 플루오레센스 발효 배양물 중 가용성 rEPA 제조에 관한 SDS-CGE 겔 유사 이미지. P. 플루오레센스의 발효 배양물 중에서 발현된 가용성 rEPA를 모세관 겔 전기영동(SDS-CGE)을 사용하여 분석하였다. 유도 후 0 및 24시간째에 테스트된 발현 균주 PS538-1633(u1 및 u2), PS538-1640(u3 및 u5) 및 PS538-1670(u6, u7 및 u8)의 발효물로부터의 가용성 분획은 SDS-CGE 데이터로부터 작성된 겔 유사 이미지로 제시되어 있다. Mw = 분자량 표준(16, 20, 29, 48, 및 69 kDa).

도 16. P. 플루오레센스 발효 배양물 중 가용성 rEPA 제조 추세. 그의 각 발효물(u1, u2, u3, u6, u7 및 u8)에서 균주(PS538-1633, PS538-1640 및 PS538-1670)의 SDS-CGE 분석에 의해 측정된 바, 가용성 rEPA 발현 수준을 유도 후 경과시간에 대해 플롯팅하였다.

도 17. P. 플루오레센스 발효 배양물 중 가용성 rEPA 제조의 웨스턴 블롯. 웨스턴 블롯 분석을 사용하여 P. 플루오레센스의 발효 배양물 중에서 발현된 가용성 rEPA를 분석하였다. 유도 후 0 및 24시간째에 발현 균주 PS538-1633(u1), PS538-1640(u3 및 u5) 및 PS538-1670(u6 및 u8)의 발효물로부터의 가용성 분획을, P. 아에루기노사 외독소 A에 특이적인 항체를 사용하여 웨스턴 블롯 분석으로 나타내었다. Mw = 분자량 표준, std = rEPA 표준.

도 18. P. 플루오레센스 발효 배양물 중 가용성 CRM197 제조에 관한 SDS-CGE 겔 유사 이미지. P. 플루오레센스의 발효 배양물 중에서 발현된 CRM197을 모세관 겔 전기영동(SDS-CGE)을 이용하여 분석하였다. 유도 후 다양한 시점(0, 16, 21 및 23시간째)에 테스트된 발현 균주 PS538-772(u1 및 u2), PS538-776(u3 및 u5) 및 PS538-782(u6 및 u7)의 다양한 발효물로부터의 가용성 분획은 SDS-CGE 데이터로부터 작성된 겔 유사 이미지로 제시되어 있다. Mw = 분자량 표준(16, 20, 29, 48, 68, 및 119 kDa).

도 19. P. 플루오레센스 발효 배양물 중 가용성 CRM197 제조 추세. 그의 각 발효물(u1, u2, u3, u6 및 u7)에서 상이한 균주(PS538-772, PS538-776 및 PS538-782)로부터의 SDS-CGE에 의해 측정된 바, 가용성 CRM197 발현 수

준을 유도 후 경과시간에 대해 플롯팅하였다.

도 20. P. 플루오레센스 발효 배양물 중 가용성 CRM197 제조의 웨스턴 블롯. 웨스턴 블롯 분석을 사용하여 P. 플루오레센스의 발효 배양물 중에서 발현된 CRM197을 분석하였다. 유도 후 다양한 시점(0, 16, 21 및 23시간째)에 테스트된 발현 균주 PS538-772(u1 및 u2), PS538-776(u3 및 u5) 및 PS538-782(u6 및 u7)의 다양한 발효 물로부터의 가용성 분획을, 디프테리아 독소 특이 항체를 사용하여 웨스턴 블롯 분석으로 나타내었다. Mw = 분자량 표준(37, 50, 75, 100, 150, 및 250 kDa). STD = CRM197 표준.

발명을 실시하기 위한 구체적인 내용

[0046] 독소

[0047] ADP 리보실화 독소

[0048] ADP 리보실화 독소(ADPRT: ADP-ribosylating toxin)는 NAD의 니코틴아미드와 N-리보스 사이의 N-글리코실 결합 절단이 원활하게 일어나도록 하며, ADP 리보스 모이어티를 표적 단백질로 전달한다. ADPRT는 그의 각 표적에 기초하여 4가지 패밀리로 분류된다. I형 ADPRT는 헤테로머 GTP 결합 단백질을 표적화한다. 이는 콜레라 독소(CTX), 백일해 독소(PTX), 및 에스케리키아 콜라이(*Escherichia coli*) 이열성 내독소(LT: heat-labile enterotoxin)를 포함한다. II형 ADPRT(디프테리아 독소 및 슈도모나스 외독소 A)는 신장 인자 2(EF2)를 변형시킨다. III형 ADPRT(클로스트리디움 보툴리눔(*Clostridium botulinum*) C3 세포외효소)는 소형 GTP 결합 단백질을 ADP 리보실화한다. IV형 ADPRT는 액틴을 ADP 리보실화한다. 이러한 액틴 특이 ADPRT로는 C. 보툴리눔(*C. botulinum*) C2 독소, C. 퍼프링겐스(*C. perfringens*) v-독소, C. 디피실레 독소(TcdA 및 TcdB와는 완전히 다른 독소, 문헌 [Popoff, et al., 1988, "Actin-specific ADP-ribosyltransferase produced by a *Clostridium difficile* strain," Infection and Immunity 56(9):2299-2306](상기 문헌은 본원에 참고로 포함된다)에 기재)), C. 스피로포르메(*C. spiroforme*) 독소, 및 바실러스 세레우스(*Bacillus cereus*) 영양 살충 단백질(VIP: vegetative insecticidal protein)을 포함하는 이원 독소 패밀리를 포함한다.

[0049] 각 유형의 ADPRT로부터의 여러 효소 성분들의 구조는 NAD의 잔부에 따라 결정되었고, 이는 예를 들어, 문헌 [Tsuge, et al., 2008, "Structural basis of actin recognition and arginine ADP-ribosylation by *Clostridium perfringens*-toxin," PNAS 105(21):7399-7404](상기 문헌은 본원에 참고로 포함된다)에서 논의되고 있다. 전형적인 액틴 특이 ADPRT는 2개의 유사 도메인: 촉매 활성화에 필수적인 C 도메인; 및 결합 및 전위 서브유닛과의 상호작용에 중요한 N 도메인을 소지한다. 대조적으로, 살모넬라로부터의 SpvB 및 III형 ADPRT C3은 단 하나의 ADP 리보실트랜스퍼라제 도메인을 가지고 있으며, N 말단 어댑터 도메인은 가지고 있지 않다. 모든 IV형 ADPRT에서, 중요 글루타메이트 잔기 2개를 포함하는 EXE 모티프는 촉매 중심에 존재한다. EXE 모티프 중 전자의 글루타메이트는 ADP 리보실트랜스퍼라제에 중요한 잔기인 것으로 여겨지는 것으로서, 이는 액틴 중 Arg-177로부터 탈양성자화된다. 후자의 글루타메이트는 N-리보스 상의 O'2와 수소 결합을 형성하는 데, 이것이 옥소 카르베늄 양이온을 안정화시키는 것으로 여겨진다.

[0050] ADPRT는 추가로 문헌 ([Barth, et al., 2004, "Binary Bacterial Toxins: Biochemistry, Biology, and Application of Common *Clostridium* and *Bacillus* Proteins," Microbiology and Molecular Biology Reviews 68(3):373-402]; [Mueller-Dieckmann, et al., "Structure of mouse ADP-ribosylhydrolase 3 (mARH3)," Acta Cryst F64: 156-162]; [Kulich, et al., 1995, "Expression of Recombinant Exoenzyme S of *Pseudomonas aeruginosa* Infection and Immunity 63(1): 1-8]; [Sakurai, et al., 2009, "*Clostridium perfringens* Iota-Toxin: Structure and Function," Toxins 1 :208-228]; 및 [Schirmer, et al., 2002, "The ADP-ribosylating Mosquitocidal Toxin from *Bacillus sphaericus*," The Journal of Biological Chemistry 277(14): 11941-11948])(상기 문헌들은 모두 본원에 참고로 포함된다)에 기술되어 있다.

[0051] 본 발명의 실시양태에서, ADPRT를 포함하는 균으로부터 선택되는 재조합 독소 단백질이 제조된다. 실시양태에서, ADPRT의 균은 CTX(CTA 및/또는 CTB), PTX, DT(CRM197 및/또는 WT), 및 슈도모나스 외독소 A로 구성된다. 다른 실시양태에서, I형 ADPRT를 포함하는 균으로부터 선택되는 재조합 독소 단백질이 제조된다. 실시양태에서, I형 ADPRT의 균은 CTX(CTA 및/또는 CTB), 및 PTX로 구성된다. 다른 실시양태에서, II형 ADPRT를 포함하는 균으로부터 선택되는 재조합 독소 단백질이 제조된다. 실시양태에서, II형 ADPRT의 균은 DT(CRM197 및/또는 WT), 및 슈도모나스 외독소 A로 구성된다. 다른 실시양태에서, IV형 ADPRT를 포함하는 균으로부터 선택되는 재조합 독소 단백질이 제조된다. 실시양태에서, IV형 ADPRT는 TcdB이다.

- [0052] CRM197 및 DT
- [0053] 교차 반응 물질 197(CRM197: cross-reacting material 197)은 미스센스 돌연변이를 가진 DT 유전자로부터 생산된 디프테리아 독소(DT) 변이체이다. DT는 ADP 리보실화 독소이며; CRM197에는 DT의 ADP 리보실트랜스퍼라제(ADPRT) 활성이 없는 바, 비독성이다. CRM197에 대한 유전자는 단일 염기 치환을 가지며, 그 결과 52번 잔기의 글리신 잔기 대신 글루탐산이 도입된다(예를 들어, 문헌 [Bishai, et al., 1987, "High-Level Expression of a Proteolytically Sensitive Diphtheria toxin Fragment in *Escherichia coli*," J. Bact. 169(11):5140-51], [Giannini, et al., 1984, "The Amino-Acid Sequence of Two Non-Toxic Mutants of Diphtheria toxin: CRM45 and CRM 197," Nucleic Acids Research 12(10): 4063-9], 및 진뱅크 수탁 번호 1007216A)(상기 문헌들은 모두 본원에 참고로 포함된다) 참조).
- [0054] CRM197 단백질은 당업계에 공지된 방법에 의해, 또는 C. 디프테리아에 또는 다른 미생물에서의 발현에 의해 낮은 수준으로 제조될 수 있다. 천연적으로 발생된 디프테리아 독소, 또는 야생형 디프테리아 독소는 아메리칸 타입 컬처 컬렉션(American Type Culture Collection)을 비롯한 각종 공개 공급원으로부터 이용가능한 독소 생산 균주로부터 취득할 수 있다. C. 디프테리아에서 CRM197 단백질을 생산하는 플라스미드 시스템은 예를 들어, 미국 특허 번호 제5,614,382호(발명의 명칭: "Plasmid for Production of CRM Protein and Diphtheria toxin")(상기 문헌은 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.
- [0055] 뉴클레오티드 서열은, 코리네박테리오파지 β 가 보유한 디프테리아 독소에 대한 야생형 구조 유전자의 공지된 DT 뉴클레오티드 서열에 기초하여, (예를 들어, 문헌 [Sambrook et al, Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989]에 기술된 바와 같은) 재조합 DNA 기술 기법을 사용하고, 또한 부위 지정 돌연변이유발법에 제조될 수 있다(예를 들어, 문헌 [Greenfield, et al., 1993, "Nucleotide Sequence of the Structural Gene for Diphtheria toxin Carried by Corynebacteriophage 18," Proc Nat Acad Sci 80:6953-7)(상기 문헌은 본원에 참고로 포함된다) 참조). 본원 다른 부분 어디에나 기술되어 있는 바와 같이, 뉴클레오티드 서열은 최적화될 수 있다.
- [0056] 본 발명의 실시양태에서, CRM197 또는 DT는 실시예 1에 기술된 발현 벡터(플라스미드) 중 어느 것과 조합하여, 본원 실시예 1에 기술된 숙주 균주 중 어느 것을 사용함으로써 제조된다. 실시양태에서, 핵산 서열은 슈도모나스 숙주 세포에서의 발현을 위해 최적화된다. 실시양태에서, 사용되는 발현 벡터는 표 8 및 표 3에 기술된 것과 같은, 재조합 CRM197 또는 DT 단백질에 융합된 분비 리더 중 어느 것을 발현하는 구성체를 포함한다. 실시양태에서, 네이티브 분비 리더가 사용된다. 특정 실시양태에서, CRM197 또는 DT 단백질은 태그, 예를 들어, 정제 태그와 함께 발현된다. 실시양태에서, 본 발명의 방법을 사용하여 약 0.5 g/L~약 12 g/L 이상의 수율로 CRM197 또는 DT를 제조할 수 있다.
- [0057] 콜레라 독소
- [0058] 비브리오 콜레라에 의해 생산되는 콜레라 독소(CTX) 또한 ADP 리보실화 독소이다. 콜레라 독소(CTX)는 단일 카피의 콜레라 독소 A 서브유닛(CTA)과 5 카피의 콜레라 독소 B 서브유닛(CTB)로 이루어진 6개의 단백질 서브유닛으로 구성된 올리고머 복합체이다. 각각의 분자량이 12 kDa인 5개의 B 서브유닛이 5원 환을 형성한다. A 서브유닛은 G 단백질을 ADP 리보실화하는 구상 효소인 A1부인 CTA1, 및 B 서브유닛 환의 중심 포어에 편안하게 안치된 연장형 알파 나선을 형성하는 A2 쇠인 CTA2를 가진다. 상기 환은 숙주 세포 표면 상에 존재하는 GM1 강글리오사이드 수용체에 결합하여 전체 복합체를 내재화시킨다. 일단 내재화되고 나면, 이황화 브릿지 환원에 의해 CTA1 쇠가 유리된다. 이어서, CTA1은 활성화되고, 아데닐레이트 사이클라제에 의한 ADP 리보실화를 촉매화한다. 그 결과, 아데닐레이트 사이클라제 활성이 증가하면 사이클릭 AMP 합성도 증가하게 되고, 이로써 대량의 체액 및 전해질 유출이 유발됨으로써 설사가 일어나게 된다.
- [0059] 상대적으로 무해하기는 하나, CTX의 B 서브유닛은 GM1 강글리오사이드 수용체에 결합할 수 있는 그의 능력을 보유한다. 그러므로, CTB는 화학적으로 또는 유전적으로 컨주게이트된 외래 항원의 점막 흡수를 촉진시키는 데 사용되는 것으로 알려져 있다. 점막 및 전신 면역, 둘 모두를 유도하는 것으로 입증되어 있으며, 식용 백신 생산에 사용하기 위한 후보물질이 된다. CTB는 그의 결합상의 선호도에 기인한 뉴런의 추적자로서의 용도도 알려져 있다.
- [0060] CTB의 용도 뿐만 아니라, 그의 구조상의 특징은 예를 들어, 문헌 ([Nozoye, et al., 2009, "Production of *Ascaris suum* As 14 Protein and Its Fusion Protein with Cholera Toxin B Subunit in Rice Seeds," Parasitology 995-1000]; [Harakuni, et al., 2005, "Heteropentameric Cholera Toxin B Subunit Chimeric

Molecules Genetically Fused to a Vaccine Antigen Induce Systemic and Mucosal Immune Responses: a Potential New Strategy to Target Recombinant Vaccine Antigens to Mucosal Immune Systems," *Infection and Immunity* 73(9):5654-5665]; [Price, et al., 2005, "Intranasal Administration of Recombinant *Neisseria gonorrhoeae* Transferrin Binding Proteins A and B Conjugated to the Cholera Toxin B Subunit Induces Systemic and Vaginal Antibodies in Mice," *Infection and Immunity* 73(7):3945-3953]; 및 [Sun, et al., 1999, "Intranasal Administration of a *Schistosoma mansoni* Glutathione S-Transferase-Cholera Toxoid Conjugate Vaccine Evokes Antiparasitic and Antipathological Immunity in Mice," *J. Immunol.* 163: 1045-1052])(상기 문헌들은 모두 본원에 참고로 포함된다에 기술되어 있다.

[0061] 본 발명의 실시양태에서, CTB 또는 CTX는 실시예 3에 기술된 발현 벡터 중 어느 것과 조합하여, 본원 실시예 1에 기술된 숙주 균주 중 어느 것을 사용함으로써 제조된다. 실시양태에서, 핵산 서열은 슈도모나스 숙주 세포에서의 발현을 위해 최적화된다. 실시양태에서, 사용되는 발현 벡터는 표 8 및 표 3에 기술된 것과 같은, 재조합 CTB 또는 CTX 단백질에 융합된 분비 리더 중 어느 것을 발현하는 구성체를 포함한다. 실시양태에서, 네이티브 분비 리더가 사용된다. 특정 실시양태에서, CTB 또는 CTX 단백질은 태그, 예를 들어, 정제 태그와 함께 발현된다. 실시양태에서, 본 발명의 방법을 사용하여 약 0.2 g/L~약 5 g/L 이상의 수율로 CCTB 또는 CTX를 제조할 수 있다.

[0062] 백일해 독소

[0063] 백일해 독소는 백일해라는 질환을 유발하는 인간 기도의 박테리아 병원균인 보르데텔라 페르투스시에 의해 생산되는 외독소이자 발병 인자이다. 백일해 홀로톡신은 AB 5 구조를 가진 다중 서브유닛 복합체이다. 효소적으로 활성을 띠는 A 서브유닛(S1)이 포유동물 세포에서 수개의 이중삼량체 G 단백질을 변형시키는 ADP 리보실트랜스퍼라제이고, B 올리고머(S2, S3, 2 개의 S4, 및 S5)는 세포 상의 글리코컨쥬게이트 수용체에 결합한다. S1은 세포 진입 후, 단백질 분해에 의해 프로세싱된다. 문헌 [Carbonetti, et al., 2005, "Proteolytic Cleavage of Pertussis Toxin S 1 Subunit is Not Essential for Its Activity in Mammalian Cells," *BMC Microbiology* 5:7](상기 문헌은 본원에 참고로 포함된다)에는 S1의 프로세싱이 그의 포유동물 세포에서의 세포독성 활성화에는 필수적인 것은 아니라고 기록되어 있다.

[0064] 백일해 독소의 비독성 변이체가 백신에서의 사용을 위해 조사되어 왔다. 본 발명의 방법을 사용하여 제조된 백일해 독소 단백질은 백일해로부터의 방어를 위한 백신에서 사용하기 위한 것으로 고려된다. 백일해 독소는 또한 예를 들어, 문헌 [Roberts, et al., 1995, "A Mutant Pertussis Toxin Molecule That Lacks ADP-Ribosyltransferase Activity, PT-9K/129G, Is an Effective Mucosal Adjuvant for Intranasally Delivered Proteins," *Infection and Immunity* 63(6):2100-2108](상기 문헌은 본원에 참고로 포함된다)에 기술된 바와 같이, 백신 애주번트로서 테스트되어 왔다. 추가로, 백일해 독소는 또한 연구 목적으로, 예를 들어, G 단백질 신호전달 경로에 관한 연구에(예를 들어, 문헌 [McCoy, et al., 2010, "PAR1 and PAR2 couple to overlapping and distinct sets of G proteins and linked signaling pathways to differentially regulate cell physiology," *Molecular Pharmacology Fast Forward* MOL 62018](상기 문헌은 본원에 참고로 포함된다)), 및 자가면역 질환, 예를 들어, 실험적 자가면역 뇌수막염(EAE: experimental autoimmune encephalomyelitis), 실험적 자가면역 고환염, 실험적 자가면역 포도막염 등의 유도를 증진시키는 애주번트로서 유용하다(문헌 [Su, et al., 2001, "Pertussis Toxin Inhibits Induction of Tissue-Specific Autoimmune Disease by Disrupting G Protein-Coupled Signals," *J Immunol* 167:250-256](상기 문헌은 본원에 참고로 포함된다)).

[0065] 도 3에 제시된 바와 같이, 독소의 5개의 서브유닛은 백일해 독소이드 오페론으로부터 발현된다. 특정 변이체를 비롯한, 백일해 독소 단백질의 발현 및 구조는 상기 인용된 레포트 뿐만 아니라, 문헌 ([Burnette, et al., 1992, "Properties of Pertussis Toxin B Oligomer Assembled In Vitro from Recombinant Polypeptides Produced by *Escherichia coli*," *Infection and Immunity* 60(6):2252-2256]; 미국 특허 번호 제5,085,862호(발명의 명칭: "Genetic detoxification of pertussis toxin"); 및 [Kaslow, et al., 1987, "Structure-Activity Analysis of the Activation of Pertussis Toxin," *Biochemistry* 26(1): 123-7])(상기 문헌은 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.

[0066] 본원에서 사용되는 바, 백일해 독소 또는 PTX는 백일해 독소 돌연변이체 S1 R9K E129A 또는 야생형 단백질을 지칭한다. 야생형 백일해 독소 및 백일해 독소 돌연변이체 S1 R9K E129A는 예를 들어, 문헌 ([Roberts, et al., 1995](상기 인용됨); 미국 특허 번호 제7,427,404호 및 미국 특허 번호 제7,666,436호(둘 모두의 발명의 명칭: "Pertussis Toxin Mutants, *Bordetella* Strains Capable of Producing Such Mutants and Their Use in the

Development of Antipertussis Vaccines"; 미국 특허 번호 제5,935,580호(발명의 명칭: "Recombinant Mutants for Inducing Specific Immune Responses"); 미국 특허 번호 제7,169,399호(발명의 명칭: "Non-Toxic Double Mutant Forms of Pertussis Toxin as Adjuvants"); 미국 특허 번호 제5,785,971호 및 미국 특허 번호 제5,427,788호(둘 모두의 발명의 명칭: "Pertussis Toxin and Use in Vaccines") 및 미국 특허 번호 제5,773,600호(발명의 명칭: "DNA Encoding Pertussis Toxin Muteins"))(상기 문헌들은 모두 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.

[0067] 본 발명의 실시양태에서, 백일해 독소 돌연변이체 S1 E129A 또는 야생형 백일해 독소는 본원 실시예 1, 5 및 7에 기술된 숙주 균주 중 어느 것을 사용함으로써 제조된다. 실시양태에서, 사용되는 발현 벡터는 표 8 및 표 3에 기술된 것과 같은, 재조합 PTX 단백질에 융합된 분비 리더 중 어느 것을 발현하는 구성체를 포함한다. 실시양태에서, 네이티브 분비 리더가 사용된다. 실시양태에서, 본원 다른 부분 어디에나 기술되어 있는 바와 같이, 서브유닛 코딩 서열 중 어느 것 또는 그들 모두는 선택된 숙주모나드 숙주에서의 발현을 위해 최적화된다. 특정 실시양태에서, 서브유닛은 예를 들어, 당업계에 주지된 방법에 따라 개별 서열을 서브클로닝시킴으로써 2개 이상의 구성체로부터 발현된다. 특정 실시양태에서, PTX 단백질은 태그, 예를 들어, 정제 태그와 함께 발현된다. 실시양태에서, 본 발명의 방법을 사용하여 약 0.2 g/L~약 5 g/L 이상의 수율로 PTX의 각 개별 서브유닛을 제조할 수 있다.

[0068] 파상풍 독소 단편 C

[0069] 클로스트리디움 테타니에 의해 생산된 파상풍 독소는 분자량이 150 kDa인, 신경독소이다. 이는 2개의 부분: 100 kDa의 중쇄 또는 B쇄 및 50 kDa의 경쇄 또는 A쇄로 구성된다. 상기 쇄는 이황화 결합에 의해 연결되어 있다. B쇄는 뉴런 막 상의 디시알로강글리오사이드(GD2 및 GD1b)에 결합한다. 아연 엔도펩티다제인 A쇄는 소포체 결합 막 단백질(VAMP)을 공격한다.

[0070] A쇄의 작용은 단백질 시냅토프레빈을 분해함으로써 영향을 받은 뉴런이 억제성 신경전달물질인 GABA(감마-아미노부티르산) 및 글리신을 유리시키는 것을 중단시킨다. 이러한 결과는 감각 자극에 의한 운동 반사 억제 실패라는 가작 작은 자극으로부터 발생하는 근육에서의 위험한 과다활성이다. 이는 파상풍 연축이라 명명되는, 주동근 및 길항근 조직의 전신 수축을 유발한다.

[0071] 파상풍 독소 단편 C(Tet C 또는 TTC)는 파상풍 독소의 (예를 들어, 파파인에 의한) 프로테아제 절단에 의해, 또는 상기 단편의 재조합 발현을 통해 생성된 50 kD 폴리펩티드이다. 이는 C 말단의 451 아미노산(아미노산 위치 865-1,315)에 상응한다. 단편 C의 재조합 발현은 예를 들어, 미국 특허 번호 제5,443,966호(발명의 명칭: "Expression of Tetanus Toxin Fragment C"), WO/2005/000346(발명의 명칭: "Carrier Proteins for Vaccines") 및 미국 특허 번호 제6,010,871호(발명의 명칭: "Modification of Peptide and Protein")(상기 문헌들은 모두 그 전문이 본원에 참고로 포함된다)에 개시되어 있다.

[0072] 단편 C는 마우스 및 기니아 피그에서 비독성이며, 방어 면역 반응을 자극시킬 수 있는 것으로 밝혀져 있다. 미국 특허 번호 제5,443,966호에는 파상풍 독소의 서열 및 E. 콜라이에서의 단편 C의 제조가 기술되어 있다. 효모에서의 재조합 TTC의 발현은 예를 들어, 미국 특허 번호 제5,571,694호(발명의 명칭: "Expression of Tetanus Toxin Fragment C in Yeast")(상기 문헌은 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.

[0073] 이는 뉴런에 고 특이도 및 고 친화도로 결합하기 때문에, TTC는 뉴런의 약물 전달을 위한, 또는 연구 목적용의 표적화 분자로서의 용도가 있는 것으로 알려져 있다. 상기과 같은 용도에 대해서는 예를 들어, 문헌 [Townsend, et al., 2007, "Tetanus toxin C fragment conjugated nanoparticles for targeted drug delivery to neurons," Biomaterials 28(34):5176-5184](상기 문헌은 본원에 참고로 포함된다)에 기술되어 있다.

[0074] TTC 단백질은 또한 잠재적으로는 예를 들어, WO/2005/000346에 기술되어 있는 바와 같이, 백신 캐리어 단백질로서 유용하고, C. 테타니 감염으로부터 방어하기 위한 백신에서 사용하기 위한 것으로 조사되어 왔다.

[0075] 본 발명의 실시양태에서, TTC는 실시예 8에 기술된 발현 벡터 중 어느 것과 조합하여, 본원 실시예 1에 기술된 숙주 균주 중 어느 것을 사용함으로써 제조된다. 실시양태에서, 핵산 서열은 숙주모나드 숙주 세포에서의 발현을 위해 최적화된다. 실시양태에서, 사용되는 발현 벡터는 표 8 및 표 3에 기술된 것과 같은, 재조합 TTC 단백질에 융합된 분비 리더 중 어느 것을 발현하는 구성체를 포함한다. 특정 실시양태에서, TTC 단백질은 태그, 예를 들어, 정제 태그와 함께 발현된다. 실시양태에서, 네이티브 분비 리더가 사용된다. 실시양태에서, 본 발명의 방법을 사용하여 약 0.5 g/L~약 12 g/L 이상의 수율로 TTC를 제조할 수 있다.

- [0076] C. 디피실레 독소 B
- [0077] 클로스트리디움 디피실레 독소 B(TcdB)는 병원내 설사 및 위막성 대장염을 유발하는 클로스트리디움 디피실레에 의해 생산되는 발병 인자이다. TcdB, 및 두번째로 큰 클로스트리디움계 독소, TcdA는 위막성 대장염 발병에 관여한다.
- [0078] TcdB는 약 270 kD의 글리코실화 독소이며, 이는 효소 도메인, 전위 도메인, 및 수용체 결합 도메인으로 구분될 수 있다. TcdB의 처음 546 아미노산은 효소 영역을 포함하고, 그 다음으로 추정 전위 도메인 및 수용체 결합 도메인 순으로 포함한다. 상기 단편의 아미노 또는 카르복시 말단의 결실이 활성을 감소시키는 바, 효소적 활성을 위해서는 아미노 말단의 546 잔기가 필요하다고 보고되어 있다. 효소 영역내의 트립토판 102는 UDP 글루코스 결합을 위해 필수적인 것으로 밝혀져 있다. LCT 내의 보존되는 DXD 모티프는 LCT 글루코실트랜스퍼라제 활성을 위해 필수적이다. TcdB 및 TcsL 효소 도메인으로 이루어진 키메라에 관한 분석을 포함한 연구를 통해 잔기 364~516이 기질 특이성을 부여한다는 것이 제안되었다.
- [0079] TcdB의 구조 및 그의 발현, 및 C. 디피실레 감염에 대한 방어용 백신으로서의 잠재적인 용도는 예를 들어, 문헌 (미국 특허 번호 제7,226,597호(발명의 명칭: "Mutants of *Clostridium Difficile* Toxin B and Methods of Use"); [Jank, et al., 2008, "Structure and mode of action of clostridial glucosylating toxins: the ABCD model," Trends in Microbiology 16(5):222-229]; [Sullivan, et al., 1982, "Purification and Characterization of Toxins A and B of *Clostridium difficile*," Infection and Immunity 35(3): 1032-1040]; 및 [Yang, et al., 2008, "Expression of recombinant *Clostridium difficile* toxin A and B in *Bacillus megaterium*," BMC Microbiology 8: 192])(상기 문헌들은 모두 그 전문이 본원에 참고로 포함된다)에서 논의되고 있다.
- [0080] 본 발명의 실시양태에서, TcdB는 본원 실시예 1, 5 및 7에 기술된 숙주 균주 중 어느 것을 사용함으로써 제조된다. 실시양태에서, 핵산 서열은 슈도모나스 숙주 세포에서의 발현을 위해 최적화된다. 실시양태에서, 사용되는 발현 벡터는 표 8 및 표 3에 기술된 것과 같은, 재조합 TcdB 단백질에 융합된 분비 리더 중 어느 것을 발현하는 구성체를 포함한다. 실시양태에서, 네이티브 분비 리더가 사용된다. 특정 실시양태에서, TcdB 단백질은 태그, 예를 들어, 정제 태그와 함께 발현된다. 실시양태에서, 본 발명의 방법을 사용하여 약 0.5 g/L~약 10 g/L 이상의 수율로 TcdB를 제조할 수 있다.
- [0081] 슈도모나스 아에루기노사 외독소 A
- [0082] 슈도모나스 아에루기노사의 외독소 A(ETA 또는 PE)는 II형 ADPRT이다. 이는 촉매 도메인을 포유동물 세포 내로 전위시키고, 세포 신장 인자 2의 ADP 리보실화에 의해 단백질 합성을 억제시킬 수 있는, 분비된 박테리아 독소 패밀리의 일원이다. 단백질은 613 아미노산으로 된 단일 폴리펩티드 쇠 (66 Kd)로 구성된 단량체로서 존재한다. 3.0-Å 해상도로 측정한 바, 외독소 A의 x선 결정학적 구조는 주로 역평행 베타 구조로 구성되고, 상기 분자의 대략 절반 가량을 포함하는 아미노 말단 도메인; 알파 나선으로 구성된 중간 도메인; 및 상기 분자의 대략 1/3 가량을 포함하는 카르복실 말단 도메인을 나타낸다. 카르복실 말단 도메인이 독소의 ADP 리보실트랜스퍼라제이다. 나머지 두 도메인은 가정컨대, 세포 수용체 결합 및 막 전위에 관여할 것으로 추정된다.
- [0083] 독소는 세포 표면 상의 특이 수용체를 통해 세포에 결합한 후, 독소-수용체 복합체가 세포 내로 내재화된다. 마지막으로, ETA는 사이토졸로 전달되고, 그곳에서 단백질 합성을 효소적으로 억제시킨다. 예를 들어, NH₄⁺와 같은 약 염기를 통해 세포 증독으로 막을 수 있고, 이로써 산성 소포의 pH는 상승하는 바, 상기 전달 과정은 산성 구획으로부터 발생하는 것으로 사료된다. 산성 조건에 노출되면, PE의 소수성 도메인은 막으로 진입하게 되고, 그 결과 채널이 형성되고, 상기 채널을 통해 신장형의 효소 도메인이 사이토졸로 전달된다. PE 및 독성이 감소된 돌연변이체는 예를 들어, (미국 특허 번호 제4,892,827호(발명의 명칭: "Recombinant *Pseudomonas* Exotoxins: Construction of an Active Immunotoxin with Low Side Effects") 및 [Lukac, et al., 1988, "Toxoid of *Pseudomonas aeruginosa* Exotoxin A Generated by Deletion of an Active-Site Residue," Infection and Immunity 56(12): 3095-3098])(상기 두 문헌 모두 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.
- [0084] 외독소 A 돌연변이체 rEPA의 백신 컨주게이트로서의 용도는 예를 들어, 문헌 ([Fattom, et al., 1993, "Laboratory and Clinical Evaluation of Conjugate Vaccines Composed of *Staphylococcus aureus* Type 5 and Type 8 Capsular Polysaccharides Bound to *Pseudomonas aeruginosa* Recombinant Exoprotein A," Infection and Immunity 61(3): 1023-1032]; [Qian, et al., 2007, "Conjugating recombinant proteins to *Pseudomonas*

aeruginosa ExoProtein A: a strategy for enhancing immunogenicity of malaria vaccine candidates," Vaccine 25(20):3923-3933]; 및 [Lin, et al., 2001. "The Efficacy of a *Salmonella Typhi* Vi Conjugate Vaccine in Two-To-Five-Year-Old Children," N Engl J Med 344(17): 1263-1269]](상기 두 문헌은 모두 본원에 참고로 포함된다)에 기술되어 있다.

[0085] 본원에서 사용되는 바, 슈도모나스 아에루기노사 외독소 A는 슈도모나스 아에루기노사 외독소 A 돌연변이체 CRM66, 결실 rEPA, 또는 야생형 단백질을 지칭한다. 본 발명의 실시양태에서, 외독소 A는 본원 실시예 1, 5 및 7에 기술된 숙주 균주 중 어느 것을 사용하고, 표 8 및 표 3에 기술된 것과 같은, 재조합 외독소 A 단백질에 융합된 분비 리더 중 어느 것을 발현하는 구성체를 포함하는 발현 벡터를 사용하여 제조된다. 실시양태에서, 핵산 서열은 슈도모나스 숙주 세포에서의 발현을 위해 최적화된다. 실시양태에서, 네이티브 분비 리더가 사용된다. 특정 실시양태에서, ETA 단백질은 태그, 예를 들어, 정제 태그와 함께 발현된다. 실시양태에서, 본 발명의 방법을 사용하여 약 0.5 g/L~약 12 g/L 이상의 수율로 외독소 A를 제조할 수 있다.

[0086] 본 발명의 방법을 사용하여 제조되는 예시적인 독소 단백질을 하기 표 1에 열거되어 있다. 이러한 목록이 제한적인 것은 아니라는 것을 이해할 것이다. 본 발명의 실시양태에서, 본 발명의 방법을 사용하는 제조를 위해, 본원에 기술된 독소의 핵산 서열 중 임의의 것은 선택된 슈도모나스 숙주 세포에서의 발현을 위해 최적화될 수 있다. 본원 다른 부분 어디에나 기술되어 있는 바와 같이, 주어진 임의의 서열을 최적화시키는 것에 관해서는 여러 옵션이 존재한다. 기술된 바와 같이 옵션들 중 임의의 것이 본 발명의 방법을 사용하여 제조되는 독소의 서열을 최적화시키는 데 사용하기 위한 것으로 고려되어 진다. 본원에서 제공하는 최적화된 서열은 본 발명의 방법에서 유용한 최적화된 서열의 비제한적인 일례이다.

표 1

표 1. 예시적인 독소 단백질

표적	예시적인 서열 공급원/참고	기원
CRM197	진뱅크 수탁 번호 1007216A	코리네박테리움 디프테리아에 NCTC 13129
디프테리아 독소 (WT)	진뱅크 NC_002935.2 진뱅크 CAA00374.1	코리네박테리움 디프테리아에
콜레라 хол로독신	진뱅크 NC_002505.1; NP231099.1 및 NP23110.1	비브리오 콜레라
콜레라 독소 B	진뱅크 ACH70471 (E1 Tor 균주)	비브리오 콜레라 O1 바이오바 El tor
백일해 독소	진뱅크 M13223.1 (S1 에 돌연변이 포함)	보르데텔라 페르투스스
파상풍 독소 C 단편	진뱅크 1A8D_A	클로스트리디움 테타니
C. 디피실레 Tox B VPI (TcdB)	진뱅크 CAA63562	클로스트리디움 디피실레
P. 아에루기노사 외독소 A	진뱅크 NP_249839	슈도모나스 아에루기노사 PAO1

[0087]

[0088] 코돈 최적화

[0089] 이중성 발현 시스템에서, 최적화 단계는 외래 단백질을 제조할 수 있는 숙주의 능력을 개선시킬 수 있다. 단백질 발현은, 전사, mRNA 프로세싱 및 번역의 안정성 및 개시에 영향을 주는 인자들을 포함하는 숙주의 인자들에 의해 지배를 받는다. 폴리뉴클레오티드 최적화 단계는 외래 단백질을 생성하는 숙주의 능력을 향상시키는 단계 뿐만 아니라, 발현 구성체를 효과적으로 디자인하는 데 있어서 연구진들을 보조하는 단계를 포함할 수 있다. 최적화 전략법은, 예를 들어, 번역 개시 영역의 변형, mRNA 구조 성분의 변경 및 상이한 코돈 편향성의 이용을 포함할 수 있다. 박테리아 숙주에서 이중성 단백질의 발현을 개선시키기 위한 핵산 서열의 최적화 방법은 당업계에 공지되어 있으며, 문헌에 기술되어 있다. 예를 들어, 슈도모나스 숙주 균주의 발현을 위한 코돈 최적화는 예를 들어, 미국 특허 출원 공개 번호 2007/0292918(발명의 명칭: "Codon Optimization Method")(상기 문헌은 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.

[0090] 따라서, 최적화는 이중성 유전자의 다수의 서열 특징 중 임의의 것을 처리할 수 있다. 한 구체적인 예로서, 희귀 코돈을 통해 번역 중단을 유도하면 이중성 단백질 발현을 감소시킬 수 있다. 희귀 코돈을 통한 번역 중단 유도는 이용가능한 tRNA 풀에서의 이의 결핍으로 인하여 단백질 번역에 부정적인 영향을 줄 수 있는 것으로서, 숙주 유기체에서 드물게 사용되는 코돈을 관심의 대상이 되는 폴리뉴클레오티드 중에 존재하도록 하는 것을 포함한다. 숙주 유기체에서 최적 번역을 개선시키는 한가지 방법은 합성 폴리뉴클레오티드 서열로부터 희귀 숙주 코돈을 제거할 수 있는 코돈 최적화를 수행하는 것을 포함한다.

- [0091] 선택적 번역 개시 또한 이중성 단백질의 발현을 감소시킬 수 있다. 선택적인 번역 개시는 리보솜 결합 부위(RBS: ribosome binding site)로서의 기능을 할 수 있는 모티프를 우연히 함유하는 합성 폴리뉴클레오타이드 서열을 포함할 수 있다. 이들 부위는 유전자 내부 부위로부터 절두형 단백질의 번역 개시를 유발할 수 있다. 정제 과정에서는 제거하기 곤란할 수 있는 절두형 단백질의 생성 가능성을 줄이는 한 가지 방법은 최적화된 폴리뉴클레오타이드 서열로부터 추정 내부 RBS 서열을 제거하는 것을 포함한다.
- [0092] 반복부 유도성 폴리머라제 저하는 이중성 단백질의 발현을 감소시킬 수 있다. 반복부 유도성 폴리머라제 저하는 프레임쉬프트 돌연변이를 초래할 수 있는 DNA 폴리머라제의 저하 또는 장애를 유발하는 것으로 밝혀진 뉴클레오타이드 서열 반복부를 포함한다. 그러한 반복부는 또한 RNA 폴리머라제 저하를 유발할 수 있다. G + C 함량이 높은 것으로 편중된 유기체에는 G 또는 C 뉴클레오타이드 반복부로 구성된 고도의 반복부가 존재할 수 있다. 따라서, RNA 폴리머라제 저하를 유도할 수 있는 가능성을 감소시키는 한 방법은 G 또는 C 뉴클레오타이드의 연장 반복부를 변경시키는 것을 포함한다.
- [0093] 또한, 간접 2차 구조가 이중성 단백질의 발현을 감소시킬 수 있다. 2차 구조는 RBS 서열 또는 개시 코돈을 격리시킬 수 있으며, 단백질 발현 감소와 관련이 있다. 또한, 스템 루프 구조는 전사 중단 및 약화에 관여할 수 있다. 최적화된 폴리뉴클레오타이드 서열은 전사 및 번역을 개선시킬 수 있는 뉴클레오타이드 서열의 유전자 코딩 영역과 RBS 내의 최소 2차 구조물을 함유할 수 있다.
- [0094] 이중성 단백질 발현에 영향을 미칠 수 있는 또 다른 특징은 제한 부위의 존재이다. 전사 단위의 숙주 발현 벡터의 후속 서브클로닝을 저해하는 제한 부위를 제거함으로써, 폴리뉴클레오타이드 서열을 최적화할 수 있다.
- [0095] 예를 들어, 숙주에 의해 이중성으로 발현되는 원하는 아미노산 서열을 확인함으로써 최적화 과정을 개시할 수 있다. 아미노산 서열로부터 후보 폴리뉴클레오타이드 또는 DNA 서열을 디자인할 수 있다. 합성 DNA 서열을 디자인하는 동안, 코돈 사용의 빈도를 숙주 발현 유기체의 코돈 사용과 비교할 수 있으며, 희귀 숙주 코돈을 합성 서열로부터 제거할 수 있다. 추가로, 합성 후보 DNA 서열을 변형시켜 비바람직한 효소 제한 부위를 제거하고, 임의의 원하는 신호 서열, 링커 또는 비번역된 영역을 부가 또는 제거할 수 있다. 번역 과정을 저해할 수 있는 2차 구조, 예를 들어, G/C 반복부 및 스템-루프 구조의 존재에 대해 합성 DNA 서열을 분석할 수 있다. 후보 DNA 서열을 합성하기 전에, 최적화 서열 디자인을 체크하여 서열이 원하는 아미노산 서열을 바르게 코딩하는지를 확인할 수 있다. 마지막으로, DNA 합성 기법, 예를 들어, 당업계에 공지된 기법을 이용하여 후보 DNA 서열을 합성할 수 있다.
- [0096] 본 발명의 또 다른 실시양태에서, 숙주 유기체, 예를 들어, P. 플루오레센스에서의 일반적 코돈 사용을 이용하여 이중성 폴리뉴클레오타이드 서열의 발현을 최적화할 수 있다. 좀처럼 숙주 발현 시스템에서 특정 아미노산을 선호하지 않는 것으로 간주되는 코돈의 비율(%) 및 분포를 평가할 수 있다. 5% 및 10% 사용 값을 희귀 코돈 결정을 위한 컷오프 값으로서 사용할 수 있다. 예를 들어, 하기 표 2에 열거된 코돈의 P. 플루오레센스 MB214 계놈에서의 존재율에 대한 계산치는 5% 미만이며, 상기 코돈은 P. 플루오레센스 숙주에서 발현되는 최적화된 유전자에서는 일반적으로 회피된다.

표 2

표 2. P. 플루오레센스 MB214 중 5% 미만으로 존재하는 코돈

아미노산(들)	사용되는 코돈(들)	존재율(%)
G Gly	GGA	3.26
I Ile	ATA	3.05
L Leu	CTA	1.78
	CTT	4.57
	TTA	1.89
R Arg	AGA	1.39
	AGG	2.72
	CGA	4.99
S Ser	TCT	4.28

[0097]

- [0098] 본 발명은 사용되는 슈도모나스 숙주 세포에서의 발현을 위해 최적화된 임의의 서열을 비롯한, 제조되는 독소에 대한 임의의 코딩 서열의 사용을 고려한다. 사용이 고려되어지는 서열은 비제한적으로, 슈도모나스 숙주 세포에서 5% 미만으로 존재하는 코돈, 슈도모나스 숙주 세포에서 10% 미만으로 존재하는 코돈, 희귀 코돈을 통한 번역

중단 유도, 추정 내부 RBS 서열, G 또는 C 뉴클레오티드의 확장 반복부, 간섭 2차 구조, 제한 부위, 또는 이들의 조합을 제거하는 최적화를 비롯하여, 원하는 바와 같은 임의의 정도로 최적화할 수 있다.

[0099] 추가로, 본 발명의 방법을 실시하는 데유용한 임의의 분비 리더의 아미노산 서열은 임의의 적절한 핵산 서열에 의해 코딩될 수 있다.

[0100] 발현 시스템

[0101] 슈도모나스 속주 세포 뿐만 아니라, 본 발명의 방법에 유용한 속주 세포에서 유용한 조절 서열(예를 들어, 프로 모터, 분비 리더, 및 리보솜 결합 부위)을 비롯한 이종성 단백질을 발현하는 방법은 예를 들어, (미국 특허 출원 공개 번호 2008/0269070 및 미국 특허 출원 일련 번호 제12/610,207호(둘 모두의 발명의 명칭: "Method for Rapidly Screening Microbial Hosts to Identify Certain Strains with Improved Yield and/or Quality in the Expression of Heterologous Proteins"), 미국 특허 출원 공개 번호 2006/0040352(발명의 명칭: "Expression of Mammalian Proteins in *Pseudomonas Fluorescens*") 및 미국 특허 출원 공개 번호 2006/0110747(발명의 명칭: "Process for Improved Protein Expression by Strain Engineering"))(상기 문헌들은 모두 그 전문이 본원에 참고로 포함된다)에 기술되어 있다. 이들 공개문헌에는 또한 본 발명의 방법을 실시하는 데유용하며, 폴딩 조절 인자를 과발현하도록 조작하거나, 또는 결실을 포함한 프로테아제 돌연변이를 도입하여 이종성 단백질 발현을 증가시킬 수 있는 박테리아 속주 균주가 기술되어 있다.

[0102] 리더

[0103] 서열 리더는 미국 특허 출원 공개 번호 2008/0193974 및 2010/0048864(둘 모두의 발명의 명칭: "Bacterial Leader Sequences for Increased Expression") 및 미국 특허 출원 공개 번호 2006/0008877(발명의 명칭: "Expression systems with Sec-secretion")(상기 문헌들은 모두 그 전문이 본원에 참고로 포함된다) 뿐만 아니라, 미국 특허 출원 공개 번호 2008/0269070 및 미국 특허 출원 일련 번호 제12/610,207호에도 상세하게 기술되어 있다.

[0104] 실시양태에서, 분비 리더를 코딩하는 서열이 독소 단백질을 코딩하는 서열에 융합된다. 실시양태에서, 분비 리더는 주변세포질 분비 리더이다. 실시양태에서, 분비 리더는 네이티브 분비 리더이다.

표 3

표 3. 예시적인 분비 리더 서열

분비 리더	아미노산 서열	서열 번호
DsbA	MRNLILSAALVTASLFGMTAQA	3
Azu	MFAKLVAVSLTLASGQLLA	4
Ibp-S31A	MIRDNRLKTSLLRGLTLTLTLSPAHA	5
Tpr	MNRSSALLAFVFLSGCQAMA	6
CupB2	MLFRTLASLTFAVIAGLPSTAHA	7
CupA2	MSCTRAFKPLLIIGLATLMCSHAFA	8
NikA	MRLAALPLLLAPLFIAPMAVA	9
Pbp A20V	MKLKRLMAAMTFVAAGVATNAVA	10
DsbC	MRLTQIIAAAAIALVSTFALA	11
TolB	MRNLLRGMLVVICCMAGIAAA	12
Pbp	MKLKRLMAAMTFVAAGVATANAVA	13
Lao	MQNYKKFLAAAVSMAFSATAMA	14
CupC2	MPPRSIAACGLLGLMATQAAA	15
PorE	MKKSTLAVAVTLGAIAQQAGA	16
Pbp	MKLKRLMAAMTFVAAGVATANAVA	17
FlgI	MKFKQLMAMALLALSVAQA	18
ttg2C	MQNRTVEIGVGLFLLAGILALLLALRVGLSA	19
CRM197 네이티브 리더	MSRKLFASXLIGALLGIGAPPSAHA	20

[0105]

[0106] 본 발명의 방법에 유용한 분비 리더는 표 3에 개시된 것에 한정되지 않음을 이해해야 한다.

[0107] 실시양태에서, 분비 리더는 Azu, IbpS31A, CupA2, 또는 PbpA20V이다. 다른 실시양태에서, 분비 리더는 Azu, IbpS31A, CupA2, PbpA20V, 또는 Pbp이다.

- [0108] 네이티브 CRM197은 절단되는 분비 리더를 통해 *C. 디프테리아에(C. diphtheriae)*로부터 세포의 공간으로 수송되고, GADD(서열 번호 21)의 아미노 말단 서열을 남긴다. *P. 플루오레센스*에서의 발현 후에 CRM197의 천연 아미노 말단을 보존하고 이황화 결합을 형성하기 위해서, 단백질은 주변세포질 공간으로 표적화된다.
- [0109] 프로모터
- [0110] 본 발명에 따라 사용되는 프로모터는 구성적 프로모터 또는 조절 프로모터일 수 있다. 유용한 조절 프로모터의 일반적인 예는 lac 프로모터(즉, lacZ 프로모터)로부터 유도된 패밀리의 것, 특히, 미국 특허 번호 제4,551,433호(DeBoer)에 기술된 tac 및 trc 프로모터 뿐만 아니라, Ptac16, Ptac17, PtacII, PlacUV5, 및 T7lac 프로모터를 포함한다. 한 실시양태에서, 프로모터는 숙주 세포 유기체로부터 유도되지 않은 것이다. 특정 실시양태에서, 프로모터는 *E. 콜라이* 유기체로부터 유도된 것이다.
- [0111] 유도성 프로모터 서열을 이용하여 본 발명의 방법에 따른 독소의 발현을 조절할 수 있다. 실시양태에서, 본 발명의 방법에 유용한 유도성 프로모터는 lac 프로모터(즉, lacZ 프로모터)로부터 유도된 패밀리의 것, 특히, 미국 특허 번호 제4,551,433호(DeBoer)에 기술된 tac 및 trc 프로모터 뿐만 아니라, Ptac16, Ptac17, PtacII, PlacUV5, 및 T7lac 프로모터를 포함한다. 한 실시양태에서, 프로모터는 숙주 세포 유기체로부터 유도되지 않은 것이다. 특정 실시양태에서, 프로모터는 *E. 콜라이* 유기체로부터 유도된 것이다.
- [0112] 본 발명에 따른 발현 시스템에 유용한 lac 유형이 아닌 프로모터의 일반적인 예로는 예를 들어, 하기 표 4에 열거된 것들을 포함한다.

표 4

표 4. lac 유형이 아닌 프로모터의 예

프로모터	인두서
P _R	고온
P _L	고온
P _m	알킬- 또는 할로-벤조에이트
P _u	알킬- 또는 할로-톨루엔
P _{sal}	살리실레이트

- [0113]
- [0114] 예를 들어, 문헌 ([J. Sanchez-Romero & V. De Lorenzo (1999) Manual of Industrial Microbiology and Biotechnology (A. Demain & J. Davies, eds.) pp. 460-74 (ASM Press, Washington, D.C.); [H. Schweizer (2001) Current Opinion in Biotechnology, 12:439-445]; 및 [R. Slater & R. Williams (2000 Molecular Biology and Biotechnology (J. Walker & R. Rapley, eds.) pp. 125-54 (The Royal Society of Chemistry, Cambridge, UK))])를 참조한다. 선택된 박테리아 숙주 세포에 대해 천연인 프로모터의 뉴클레오티드 서열을 가지는 프로모터를 이용하여 표적 폴리펩티드, 예를 들어, 슈도모나스 안트라닐레이트 또는 벤조에이트 오피론 프로모터(Pant, Pben)를 코딩하는 트랜스진의 발현을 제어할 수 있다. 또한, 서열이 동일하든 또는 상이하든, 예를 들어, Pant-Pben 직렬 프로모터(프로모터간 하이브리드) 또는 Plac-Plac 직렬 프로모터, 또는 동일한 유기체 또는 상이한 유기체로부터 유도된 것이든 상관없이 1 초과인 프로모터가 상호 공유 결합된 직렬 프로모터가 이용될 수 있다
- [0115] 조절된 프로모터는 프로모터가 일부인 유전자의 전사를 제어하기 위해서 프로모터 조절 단백질을 이용한다. 본원에서 조절된 프로모터가 이용되는 경우, 상응하는 프로모터 조절 단백질 또한 본 발명에 따른 발현 시스템의 일부가 될 것이다. 프로모터 조절 단백질의 예는 활성화 인자 단백질, 예를 들어, *E. 콜라이* 이화대사산물 활성화 인자 단백질, MalT 단백질; AraC 패밀리 전사 활성화 인자; 억제 단백질, 예를 들어, *E. 콜라이* LacI 단백질; 및 이중 기능 조절 단백질, 예를 들어, *E. 콜라이* NagC 단백질을 포함한다. 다수의 조절된 프로모터/프로모터 조절 단백질 쌍이 당업계에 공지되어 있다. 한 실시양태에서, 관심의 대상이 되는 이중성 단백질 및 표적 단백질(들)에 대한 발현 구성체는 동일한 조절 성분의 제어하에 있다.
- [0116] 프로모터 조절 단백질은 이펙터 화합물, 즉, 단백질을 프로모터의 제어하에 있는 유전자의 하나 이상의 DNA 전사 조절 영역으로 방출시키거나 또는 상기 영역에 결합시키는 조절 단백질과 가역적으로 또는 비가역적으로 결합하는 화합물과 상호작용하여, 유전자의 전사를 개시하는 전사 효소를 작용시키거나 또는 상기 전사 효소의 작용을 차단한다. 이펙터 화합물은 인두서 또는 보조 리프레서로서 분류되며, 이들 화합물은 네이티브 이펙터 화합물 및 무상 인두서 화합물을 포함한다. 조절된 프로모터/프로모터 조절 단백질/이펙터 화합물의 트리오 다수

가 당업계에 공지되어 있다. 세포 배양 또는 발효 동안에 계속해서 이펙터 화합물이 사용될 수 있지만, 조절된 프로모터가 사용되는 바람직한 실시양태에서는 숙주 세포 바이오매스를 원하는 정량 또는 밀도로 성장시킨 후에, 적절한 이펙터 화합물을 배양물에 첨가하여 직접 또는 간접적으로 관심의 대상이 되는 단백질 또는 폴리펩티드를 코딩하는 원하는 유전자(들)를 발현시킨다.

[0117] lac 패밀리의 프로모터가 이용되는 실시양태에서, *lacI* 유전자가 시스템 내에 존재할 수 있다. *lacI* 유전자는 보통 항상적으로 발현되는 유전자로서, lac 패밀리의 프로모터의 lac 오퍼레이터에 결합하는 Lac 리프레서 단백질 LacI 단백질을 코딩한다. 따라서, lac 패밀리의 프로모터가 이용되는 경우, *lacI* 유전자가 또한 발현 시스템에 포함되어 발현될 수 있다.

[0118] 슈도모나스에 유용한 프로모터 시스템은 문헌, 예를 들어, 상기 언급된 미국 특허 출원 공개 번호 2008/0269070에 기술되어 있다.

[0119] 다른 조절 요소

[0120] 실시양태에서, 생성 과정에서 세포의 주변세포질 또는 세포질에 가용성 단백질이 존재한다. 단백질 표적화에 유용한 분비 리더는 본원 다른 부분 어디에나 기술되어 있고, 미국 특허 출원 공개 번호 2008/0193974, 미국 특허 출원 공개 번호 2006/0008877, 및 미국 특허 출원 일련 번호 제12/610,207호에 기술되어 있다.

[0121] 다른 요소로는 전사 인핸서 서열, 번역 인핸서 서열, 다른 프로모터, 활성화 인자, 번역 출발 및 정지 신호, 전사 종결 인자, 시스트론 조절 인자, 폴리시스트론 조절 인자, 태그 서열, 예를 들어, 뉴클레오티드 서열 "태그" 및 "태그" 폴리펩티드 코딩 서열(이는 발현된 폴리펩티드의 확인, 분리, 정제, 및/또는 단리가 원활하게 일어날 수 있도록 한다)을 포함하나, 이에 한정되지 않는다.

[0122] 실시양태에서, 발현 벡터는 추가로 분비 신호에 대한 코딩 서열에, 또는 관심의 대상이 되는 단백질 또는 폴리펩티드에 대한 코딩 서열에 인접한 태그 서열을 포함한다. 한 실시양태에서, 상기 태그 서열을 통해 단백질은 정제될 수 있다. 태그 서열은 친화 태그, 예를 들어, 헥사-히스티딘 친화 태그(서열 번호: 46)일 수 있다. 또 다른 실시양태에서, 친화 태그는 글루타티온-S-트랜스퍼라제 분자일 수 있다. 태그는 또한 형광성 분자, 예를 들어, YFP 또는 GFP, 또는 상기 형광성 단백질의 유사체일 수 있다. 태그는 또한 항체 분자의 일부분, 또는 정제에 유용한 공지의 결합 파트너에 대한 공지의 항원 또는 리간드일 수 있다.

[0123] 본 발명의 발명을 실시하는 데 유용한 발현 구성체는 단백질 코딩 서열 이외에도 여기에 작동 가능하게 연결된 하기의 조절 요소들을 포함할 수 있다: 프로모터, 리보솜 결합 부위(RBS), 전사 종결 인자, 및 번역 출발 및 정지 신호. 유용한 RBS는 예를 들어, 미국 특허 출원 공개 번호 2008/0269070 및 미국 특허 출원 일련 번호 제12/610,207호에 따른 발현 시스템에서 숙주 세포로서 유용한 임의의 종으로부터 수득될 수 있다. 다수의 특이적인 각종 컨센서스 RBS가 알려져 있으며, 예를 들어, 문헌 ([D. Frishman et al., Gene 234(2):257-65 (8 Jul. 1999)]; 및 [B. E. Suzek et al., Bioinformatics 17(12):1123-30 (December 2001)])에 기술되어 있고, 참조되어 있는 것들이 있다. 또한, 네이티브 또는 합성 RBS, 예를 들어, 문헌 (EP 0207459(합성 RBS); [O. Ikehata et al., Eur. J. Biochem. 181(3):563-70 (1989)](네이티브 RBS 서열인 AAGGAAG))에 기술된 것들이 사용될 수 있다. 본 발명에 유용한 방법, 벡터와, 번역 및 전사 요소 및 요소들의 추가 예는 당업계에 주지되어 있으며, 예를 들어, 미국 특허 번호 제5,055,294호(Gilroy) 및 미국 특허 번호 제5,128,130호(Gilroy 등); 미국 특허 번호 제5,281,532호(Rammler 등); 미국 특허번호 제4,695,455호 및 제4,861,595호(Barnes 등); 미국 특허번호 제4,755,465호(Gray 등); 및 미국 특허번호 제5,169,760호(Wilcox)(상기 문헌들은 모두 본원에 참고로 포함된다) 뿐만 아니라, 다른 다수의 공개 문헌들(상기 문헌은 본원에 참고로 포함된다)에 기술되어 있다.

[0124] 숙주 균주

[0125] 슈도모나드를 비롯한 박테리아 숙주 및 밀접하게 관련된 박테리아 유기체가 본 발명의 방법의 실시를 위해 고려된다. 특정 실시양태에서, 슈도모나드 숙주 세포는 슈도모나스 플루오레센스이다. 숙주 세포는 또한 E. 콜라이 세포일 수 있다.

[0126] 본 발명의 방법을 실시하는 데 유용한 숙주 세포 및 구성체는 당업계에 공지되어 있고, 문헌, 예를 들어, 미국 특허 출원 공개 번호 2009/0325230(발명의 명칭: "Protein Expression Systems")(상기 문헌은 그 전문이 본원에 참고로 포함된다)에 기술되어 있는 것과 같은 시약 및 방법을 사용하여 확인되거나, 제조될 수 있다. 상기 공개 문헌에는 염색체 *lacI* 유전자 삽입체를 포함하는 영양요구성 슈도모나스 플루오레센스 숙주 세포 내로 핵산 구성체를 도입함으로써 재조합 폴리펩티드를 제조하는 것이 기술되어 있다. 핵산 구성체는 숙주 세포에서 핵산의 발현을 지시하는 프로모터에 작동 가능하게 연결된, 재조합 폴리펩티드를 코딩하는 뉴클레오티드 서열을

포함하고, 또한, 영양요구성 선별 마커를 코딩하는 뉴클레오티드 서열도 포함한다. 영양요구성 선별 마커는 영양요구성 숙주 세포에 대한 원영양성을 복구하는 폴리펩티드이다. 실시양태에서, 세포는 프롤린, 우라실, 또는 이들의 조합에 대해 영양요구성이다. 실시양태에서, 숙주 세포는 MB101(ATCC 기탁 PTA-7841)로부터 유래된 것이다. (미국 특허 출원 공개 번호 2009/0325230(발명의 명칭: "Protein Expression Systems") 및 [Schneider, et al., 2005, "Auxotrophic markers pyrF and proC can replace antibiotic markers on protein production plasmids in high-cell-density *Pseudomonas fluorescens* fermentation," Biotechnol. Progress 21(2): 343-8])(상기 두 문헌 모두 그 전문이 본원에 참고로 포함된다)에는 균주 MB101에서 *pyrF* 유전자를 결실시킴으로써 구성된, 우라실에 대해 영양요구성인 숙주를 제조하는 것이 기술되어 있다. 균주 MB214(ATCC 기탁 PTA-7840)로부터 *pyrF* 유전자를 클로닝하여 *pyrF* 결실을 보완하여 원영양성을 복구할 수 있는 플라스미드를 생성하였다. 특정 실시양태에서, P. 플루오레센스 숙주 세포에서는 이중 *pyrF*-*proC* 이중 영양요구성 선별 마커 시스템이 사용된다. 기술된 바와 같은 *PyrF* 생산 숙주 균주는, 본 발명의 방법을 실시하는 데 유용한 것으로서 본원에 기술된 것을 비롯한, 다른 원하는 게놈 변화를 도입하기 위한 배경으로서 사용될 수 있다.

[0127] 실시양태에서, 숙주 세포는 슈도모나달레스(*Pseudomonadales*) 목의 것이다. 숙주 세포가 슈도모나달레스 목의 것인 경우, 이는 슈도모나스 속을 비롯한 슈도모나다세아에(*Pseudomonadaceae*) 과의 구성원일 수 있다. 감마 프로테오박테리아 숙주로는 에스케리키아 콜라이 종의 구성원, 및 슈도모나스 플루오레센스 종의 구성원을 포함한다.

[0128] 다른 슈도모나스 유기체 또한 유용할 수 있다. 슈도모나드 및 밀접하게 관련된 종은 그람 음성 프로테오박테리아 서브군 1을 포함하는데, 이는 문헌 ["Gram-Negative Aerobic Rods and Cocci" by R. E. Buchanan and N.E. Gibbons (eds.), Bergey's Manual of Determinative Bacteriology, pp. 217-289 (8th ed., 1974) (The Williams & Wilkins Co., Baltimore, Md., USA)] (이하 ["Bergey (1974)"])에 기술되어 있는 것과 같은 과 및 /또는 속에 속하는 박테리아 군을 포함한다. 하기 표 5에는 이러한 과 및 속의 유기체를 제시한다.

표 5

표 5. "그람 음성 호기성 간균 및 구균" 파트 (문헌 [Bergey, 1974])에 열거되어 있는 과 및 속

I. 슈도모나다세아에과	글루코노박터(<i>Gluconobacter</i>) 슈도모나스 크산토모나스(<i>Xanthomonas</i>) 주글로에아(<i>Zoogloea</i>)
II. 아조토박테라세아에(<i>Azotobacteraceae</i>)과	아조모나스(<i>Azomonas</i>) 아조토박터(<i>Azotobacter</i>) 베이제린키아(<i>Beijerinckia</i>) 데르시아(<i>Derxia</i>)
III. 리조비아세아에(<i>Rhizobiaceae</i>)과	아그로박테리움(<i>Agrobacterium</i>) 리조비움(<i>Rhizobium</i>)
IV. 메틸로모나다세아에(<i>Methylomonadaceae</i>)과	메틸로코커스(<i>Methylococcus</i>) 메틸로모나스(<i>Methylomonas</i>)
V. 할로박테리아세아에(<i>Halobacteriaceae</i>)과	할로박테리움(<i>Halobacterium</i>) 할로코커스(<i>Halococcus</i>)
다른 속	아세토박터(<i>Acetobacter</i>) 알칼리게네스(<i>Alcaligenes</i>) 보르데텔라(<i>Bordetella</i>) 브루셀라(<i>Brucella</i>) 프란시셀라(<i>Francisella</i>) 써머스(<i>Thermus</i>)

[0129]

[0130] 슈도모나스 및 밀접하게 관련된 박테리아는 일반적으로 "그람 음성 프로테오박테리아 서브군 1" 또는 "그람 음성 호기성 간균 및 구균"(문헌 [Buchanan and Gibbons (eds.) (1974) Bergey's Manual of Determinative Bacteriology, pp. 217-289])으로 정의되는 그룹의 일부이다. 슈도모나스 숙주 균주는 문헌, 예를 들어, 상기 인용된 미국 특허 출원 공개 번호 2006/0040352에 기술되어 있다.

[0131] "그람 음성 프로테오박테리아 서브군 1"은 분류에 사용된 기준에 따라 이러한 표제로 분류되는 프로테오박테리아를 또한 포함한다. 상기 표제는 또한 종전에는 이러한 섹션으로 분류되었지만 더 이상은 아닌 군, 예를 들어,

애시도보락스(*Acidovorax*) 속, 브레분디모나스(*Brevundimonas*) 속, 부르크홀데리아(*Burkholderia*) 속, 히드로제노파가(*Hydrogenophaga*) 속, 오세아니모나스(*Oceanimonas*) 속, 랄스토니아(*Ralstonia*) 속 및 스테노트로포모나스(*Stenotrophomonas*) 속; 크산토모나스(*Xanthomonas*) 속에 속하는 (그리고 종전에는 크산토모나스 속의 종으로 칭해진) 유기체를 재분류하여 생성된 스프링고모나스(*Sphingomonas*) 속 (및 이로부터 유래된 블라스토모나스(*Blastomonas*) 속); 문헌 [Bergey (1974)]에서 정의된 아세토박터 속에 속하는 생물을 재분류하여 생성된 애시도모나스(*Acidomonas*) 속을 포함한다. 추가로, 속주는 슈도모나스 속, 슈도모나스 에날리아(*Pseudomonas enalia*)(ATCC 14393), 슈도모나스 니그리파시엔스(*Pseudomonas nigrifaciens*)(ATCC 19375) 및 슈도모나스 푸트레파시엔스(*Pseudomonas putrefaciens*)(ATCC 8071)로부터의 세포를 또한 포함할 수 있고, 이들은 각각 알테로모나스 할로플란크티스(*Alteromonas haloplanktis*), 알테로모나스 니그리파시엔스(*Alteromonas nigrifaciens*) 및 알테로모나스 푸트레파시엔스(*Alteromonas putrefaciens*)로 재분류되었다. 유사하게, 예를 들어, 슈도모나스 애시도보란스(*Pseudomonas acidovorans*)(ATCC 15668) 및 슈도모나스 테스트스테로니(*Pseudomonas testosteroni*)(ATCC 11996)는 각각 코마모나스 애시도보란스(*Comamonas acidovorans*) 및 코마모나스 테스트스테로니(*Comamonas testosteroni*)로 각각 재분류되었고; 슈도모나스 니그리파시엔스(ATCC 19375) 및 슈도모나스 피스시다(*Pseudomonas piscicida*)(ATCC 15057)는 각각 슈도알테로모나스 니그리파시엔스(*Pseudoalteromonas nigrifaciens*) 및 슈도알테로모나스 피스시다(*Pseudoalteromonas piscicida*)로 재분류되었다. "그람 음성 프로테오박테리아 서브군 1"은 임의의 하기 과에 속하는 것으로 분류된 프로테오박테리아를 또한 포함한다: 슈도모나다세아에, 아조토박터라세아에(현재는 종종 슈도모나다세아에의 "아조토박터 군"이라는 동의어로 명명된다), 리조비아세아에(*Rhizobiaceae*), 및 메틸로모나다세아에(*Methylomonadaceae*)(현재는 종종 "메틸로코카세아에(*Methylococcaceae*)"라는 동의어로 명명된다). 따라서, 본원에 달리 기술된 속들 이외에도, "그람 음성 프로테오박테리아 서브군 1"에 속하는 추가적인 프로테오박테리아 속은: 1) 아조리조필루스(*Azorhizophilus*) 속의 아조토박터 군 박테리아; 2) 셀비브리오(*Cellvibrio*) 속, 올리겔라(*Oligella*) 속 및 테레디니박터(*Teredinibacter*) 속의 슈도모나다세아에 과 박테리아; 3) 켈라토박터(*Chelatobacter*) 속, 엔시러(*Ensifer*) 속, 리베리박터(*Liberibacter*) 속(또한 "칸디다투스 리베리박터(*Candidatus Liberibacter*)"로도 명명된다), 및 시노리조비움(*Sinorhizobium*) 속의 리조비아세아에 과 박테리아; 및 4) 메틸로박터(*Methylobacter*) 속, 메틸로칼둠(*Methylocaldum*) 속, 메틸로마이크로비움(*Methylomicrobium*) 속, 메틸로사르시나(*Methylosarcina*) 속 및 메틸로스파에라(*Methylosphaera*) 속의 메틸로코카세아에 과 박테리아를 포함한다.

[0132] 속주 세포는 "그람 음성 프로테오박테리아 서브군 16"으로부터 선택될 수 있다. "그람 음성 프로테오박테리아 서브군 16"은 하기의 슈도모나스 종의 프로테오박테리아의 군으로 정의된다(예시적인 군주(들)의 ATCC 또는 다른 기탁 번호가 괄호 안에 기재): 슈도모나스 아비에타니필라(*Pseudomonas abietaniphila*)(ATCC 700689); 슈도모나스 아에루지노사(*Pseudomonas aeruginosa*)(ATCC 10145); 슈도모나스 알칼리게네스(*Pseudomonas alcaligenes*)(ATCC 14909); 슈도모나스 안귄리셉티카(*Pseudomonas anguilliseptica*)(ATCC 33660); 슈도모나스 시트로넬로리스(*Pseudomonas citronellolis*)(ATCC 13674); 슈도모나스 플라베센스(*Pseudomonas flavescens*)(ATCC 51555); 슈도모나스 멘도시나(*Pseudomonas mendocina*)(ATCC 25411); 슈도모나스 니트로레두센스(*Pseudomonas nitroreducens*)(ATCC 33634); 슈도모나스 올레오보란스(*Pseudomonas oleovorans*)(ATCC 8062); 슈도모나스 슈도알칼리게네스(*Pseudomonas pseudoalcaligenes*)(ATCC 17440); 슈도모나스 레시노보란스(*Pseudomonas resinovorans*)(ATCC 14235); 슈도모나스 스트라미네아(*Pseudomonas straminea*)(ATCC 33636); 슈도모나스 아가리시(*Pseudomonas agarici*)(ATCC 25941); 슈도모나스 알칼리필라(*Pseudomonas alcaliphila*); 슈도모나스 알기노보라(*Pseudomonas alginovora*); 슈도모나스 안데르소니이(*Pseudomonas andersonii*); 슈도모나스 아스플레니이(*Pseudomonas asplenii*)(ATCC 23835); 슈도모나스 아젤라이카(*Pseudomonas azelaica*)(ATCC 27162); 슈도모나스 베이제린키이(*Pseudomonas bejerinckii*)(ATCC 19372); 슈도모나스 보레알리스(*Pseudomonas borealis*); 슈도모나스 보레오폴리스(*Pseudomonas boreopolis*)(ATCC 33662); 슈도모나스 브라시카세아룸(*Pseudomonas brassicacearum*); 슈도모나스 부타노보라(*Pseudomonas butanovora*)(ATCC 43655); 슈도모나스 셀룰로사(*Pseudomonas cellulosa*)(ATCC 55703); 슈도모나스 아우란티아카(*Pseudomonas aurantiaca*)(ATCC 33663); 슈도모나스 클로로라피스(*Pseudomonas chlororaphis*)(ATCC 9446, ATCC 13985, ATCC 17418, ATCC 17461); 슈도모나스 프라기(*Pseudomonas fragi*)(ATCC 4973); 슈도모나스 룬덴시스(*Pseudomonas lundensis*)(ATCC 49968); 슈도모나스 타에트롤렌스(*Pseudomonas taetrolens*)(ATCC 4683); 슈도모나스 시스시콜라(*Pseudomonas cissicola*)(ATCC 33616); 슈도모나스 코로나파시엔스(*Pseudomonas coronafaciens*); 슈도모나스 디테르페니필라(*Pseudomonas diterpeniphila*); 슈도모나스 엘롱가타(*Pseudomonas elongata*)(ATCC 10144); 슈도모나스 플렉텐스(*Pseudomonas flectens*)(ATCC 12775); 슈도모나스 아조토포르만스(*Pseudomonas azotoformans*); 슈도모나스 브렌네리(*Pseudomonas brenneri*); 슈도모나스 세드렐라(*Pseudomonas cedrella*); 슈도모나스 코루가

타(*Pseudomonas corrugata*)(ATCC 29736); 슈도모나스 엑스트레모리엔탈리스(*Pseudomonas extremorientalis*); 슈도모나스 플루오레센스(ATCC 35858); 슈도모나스 게사르디이(*Pseudomonas gessardii*); 슈도모나스 리바넨시스(*Pseudomonas libanensis*); 슈도모나스 만델리이(*Pseudomonas mandelii*)(ATCC 700871); 슈도모나스 마르지날리스(*Pseudomonas marginalis*)(ATCC 10844); 슈도모나스 미굴라에(*Pseudomonas migulae*); 슈도모나스 무시돌렌스(*Pseudomonas mucidolens*)(ATCC 4685); 슈도모나스 오리엔탈리스(*Pseudomonas orientalis*); 슈도모나스 로테시아에(*Pseudomonas rhodesiae*); 슈도모나스 신크잔타(*Pseudomonas synxantha*)(ATCC 9890); 슈도모나스 톨라아시이(*Pseudomonas tolaasii*)(ATCC 33618); 슈도모나스 베로니이(*Pseudomonas veronii*)(ATCC 700474); 슈도모나스 프레데릭스베르겐시스(*Pseudomonas frederiksbergensis*); 슈도모나스 게니쿨라타(*Pseudomonas geniculata*)(ATCC 19374); 슈도모나스 긴게리(*Pseudomonas gingeri*); 슈도모나스 그라미니스(*Pseudomonas graminis*); 슈도모나스 그리몬티(*Pseudomonas grimontii*); 슈도모나스 할로테니트리피칸스(*Pseudomonas halodenitrificans*); 슈도모나스 할로필라(*Pseudomonas halophila*); 슈도모나스 히비스시콜라(*Pseudomonas hibiscicola*)(ATCC 19867); 슈도모나스 후티엔시스(*Pseudomonas huttiensis*)(ATCC 14670); 슈도모나스 히드로제노보라(*Pseudomonas hydrogenovora*); 슈도모나스 제세니이(*Pseudomonas jessenii*)(ATCC 700870); 슈도모나스 킬로넨시스(*Pseudomonas kilonensis*); 슈도모나스 란세올라타(*Pseudomonas lanceolata*)(ATCC 14669); 슈도모나스 리니(*Pseudomonas lini*); 슈도모나스 마르기나타(*Pseudomonas marginata*)(ATCC 25417); 슈도모나스 메피티카(*Pseudomonas mephitica*)(ATCC 33665); 슈도모나스 테니트리피칸스(*Pseudomonas denitrificans*)(ATCC 19244); 슈도모나스 페르투치노게나(*Pseudomonas pertucinogena*)(ATCC 190); 슈도모나스 픽토룸(*Pseudomonas pictorum*)(ATCC 23328); 슈도모나스 사이크로필라(*Pseudomonas psychrophila*); 슈도모나스 필바(*Pseudomonas filva*)(ATCC 31418); 슈도모나스 몬테일리이(*Pseudomonas monteilii*)(ATCC 700476); 슈도모나스 모셀리이(*Pseudomonas mosselii*); 슈도모나스 오리지하비탄스(*Pseudomonas oryzihabitans*)(ATCC 43272); 슈도모나스 플레코글로시시다(*Pseudomonas plecoglossicida*)(ATCC 700383); 슈도모나스 푸티다(*Pseudomonas putida*)(ATCC 12633); 슈도모나스 레악탄스(*Pseudomonas reactans*); 슈도모나스 스피노사(*Pseudomonas spinosa*)(ATCC 14606); 슈도모나스 발레아리카(*Pseudomonas balearica*); 슈도모나스 루테올라(*Pseudomonas luteola*)(ATCC 43273); 슈도모나스 스투트제리(*Pseudomonas stutzeri*)(ATCC 17588); 슈도모나스 아미그달리(*Pseudomonas amygdali*)(ATCC 33614); 슈도모나스 아벨라나에(*Pseudomonas avellanae*)(ATCC 700331); 슈도모나스 카리카파파야에(*Pseudomonas caricapapayae*)(ATCC 33615); 슈도모나스 시초리이(*Pseudomonas cichorii*)(ATCC 10857); 슈도모나스 피쿠세렉타에(*Pseudomonas ficuserectae*)(ATCC 35104); 슈도모나스 푸스코바기나에(*Pseudomonas fuscovaginae*); 슈도모나스 멜리아에(*Pseudomonas meliae*)(ATCC 33050); 슈도모나스 시린가에(*Pseudomonas syringae*)(ATCC 19310); 슈도모나스 비리디플라바(*Pseudomonas viridiflava*)(ATCC 13223); 슈도모나스 테르모카록시도보란스(*Pseudomonas thermocarboxydovorans*)(ATCC 35961); 슈도모나스 테르모톨레란스(*Pseudomonas thermotolerans*); 슈도모나스 티베르발렌시스(*Pseudomonas thivervalensis*); 슈도모나스 반코우베넨시스(*Pseudomonas vancouverensis*)(ATCC 700688); 슈도모나스 위스콘시넨시스(*Pseudomonas wisconsinensis*); 및 슈도모나스 시아메넨시스(*Pseudomonas xiamenensis*). 한 실시양태에서, 숙주 세포는 슈도모나스 플루오레센스이다.

[0133] 숙주 세포는 "그람 음성 프로테오박테리아 서브군 17"로부터 선택될 수 있다. "그람 음성 프로테오박테리아 서브군 17"은 예를 들어, 하기의 슈도모나스 종에 속하는 것들을 비롯한, "형광성 슈도모나드"로 당업계에서 공지된 프로테오박테리아의 군으로 정의된다: 슈도모나스 아조토포르만스; 슈도모나스 브렌네리; 슈도모나스 세드렐라; 슈도모나스 코루가타; 슈도모나스 엑스트레모리엔탈리스; 슈도모나스 플루오레센스; 슈도모나스 게사르디이; 슈도모나스 리바넨시스; 슈도모나스 만델리이; 슈도모나스 마르지날리스; 슈도모나스 미굴라에; 슈도모나스 무시돌렌스; 슈도모나스 오리엔탈리스; 슈도모나스 로테시아에; 슈도모나스 신크잔타; 슈도모나스 톨라아시이; 및 슈도모나스 베로니이.

[0134] 실시양태에서, 슈도모나스 숙주 세포는 HslU, HslV, Prc1, DegP1, DegP2, AprA, 또는 이들의 조합의 발현에 결합이 있는 것이다. 실시양태에서, 숙주 세포는 프로테아제 HslU, HslV, Prc1, DegP1, DegP2, 및 AprA에 결합이 있고, DegP2 S219A를 과발현하는 것이다. 그러한 균주의 예는 본원에서 숙주 균주 2로서 개시되어 있다. 이들 프로테아제는 당업계에서 공지되어 있으며, 예를 들어, 미국 특허 출원 공개 번호 2006/0110747에 기술되어 있다. 세포의 세라리신형 메탈로프로테아제 메탈로프로테이나제인 AprA는 예를 들어, 문헌 ([Maunsell, et al., 2006, "Complex regulation of AprA metalloprotease in *Pseudomonas fluorescens* M114: evidence for the involvement of iron, the ECF sigma factor, PbrA and pseudobactin M114 siderophore, Microbiology 152(Pt 1):29-42](상기 문헌은 본원에 참고로 포함된다) 및 미국 특허 출원 공개 번호 2008/0193974 및 2010/0048864에 기술되어 있다.

- [0135] 다른 실시양태에서, 슈도모나스 속주 세포는 DsbA, DsbB, DsbC, 및 DsbD를 과발현하는 것이다. DsbA, B, C, 및 D는 이황화 결합 이소머라제로서, 이는 예를 들어, 미국 특허 출원 공개 번호 2008/0269070 및 미국 특허 출원 일련 번호 제12/610,207호에 기술되어 있다.
- [0136] 다른 실시양태에서, 슈도모나스 속주 세포는 야생형, 즉, 프로테아제 발현에 결함이 없고, 임의의 폴딩 조절 인자를 과발현하지 않는 것이다.
- [0137] 프로테아제 발현에 결함이 있는 속주 세포는 야생형 속주에 비해 프로테아제의 정상적인 활성 또는 발현 수준의 감소를 초래하는 임의의 변형을 포함할 수 있다. 예를 들어, 미스센스 또는 넌센스 돌연변이는 활성을 띠지 않는 단백질의 발현을 유도할 수 있으며, 유전자 결실은 전혀 단백질을 발현시키지 않을 수 있다. 유전자의 상류 조절 영역의 변화는 단백질 발현을 감소시키거나 또는 단백질을 발현시키지 않을 수 있다. 다른 유전자 결함은 단백질 번역에 영향을 줄 수 있다. 프로테아제 프로세싱에 필요한 단백질의 활성에 결함이 있는 경우에는 또한 프로테아제 발현에 결함이 있을 수 있다.
- [0138] 본 발명의 방법에 유용한 프로테아제 및 폴딩 조절 인자의 예는 각각 하기 표 6 및 7에 제시되어 있다. RXF 번호는 오픈 리딩 프레임의 의미한다(예를 들어, 미국 특허 출원 공개 번호 2008/0269070 및 미국 특허 출원 일련 번호 제12/610,207호를 참조한다).

표 6

표 6. P. 플루오레센스 균주 MB214 프로테아제

부류	패밀리	RXF	유전자	보완된(curated) 기능	위치
아스파르트산 펩티다제					
A8 (신호 펩티다제 II 패밀리)					
		RXF05383.2		지단백질 신호 펩티다제 (ec 3.4.23.36)	세포질 막
A24 (IV 형 프리필린 펩티다제 패밀리)					
		RXF05379.1		4 형 프리필린 펩티다제 pild (ec 3.4.99.-)	세포질 막
시스테인 펩티다제					
C15 (피로글루타밀 펩티다제 I 패밀리)					
		RXF02161.1		피롤리돈-카르복실레이트 펩티다제 (ec 3.4.19.3)	세포질
C40					
		RXF01968.1		침입 관련 단백질, P60	신호 펩티드
		RXF04920.1		침입 관련 단백질, P60	세포질
		RXF04923.1		포스파타제 관련 단백질 papq	신호 펩티드
C56 (PfpI 엔도펩티다제 패밀리)					
		RXF01816.1		프로테아제 I (ec 3.4.-.-)	비분비성
메탈로펩티다제					
M1					
		RXF08773.1		막 알라닌 아마노펩티다제 (ec 3.4.11.2)	비분비성
M3					
		RXF00561.2	prlC	올리고펩티다제 A (ec 3.4.24.70)	세포질
		RXF04631.2		Zn 의존성 올리고펩티다제	세포질
M4 (싸모리신 패밀리)					
		RXF05113.2		세포외 메탈로프로테아제 전구체 (ec 3.4.24.-)	세포외
M41 (FtsH 엔도펩티다제 패밀리)					
		RXF05400.2		세포 분열 단백질 ftsH (ec 3.4.24.-)	세포질 막
M10					
		RXF04304.1		세라리신 (ec 3.4.24.40)	세포외
		RXF04500.1		세라리신 (ec 3.4.24.40)	세포외
		RXF01590.2		세라리신 (ec 3.4.24.40)	세포외
		RXF04497.2		세라리신 (ec 3.4.24.40)	세포외
		RXF04495.2		세라리신 (ec 3.4.24.40)	세포외
		RXF02796.1		세라리신 (ec 3.4.24.40)	세포외
M14 (카르복시펩티다제 A 패밀리)					
		RXF09091.1		아연 카르복시펩티다제 전구체 (ec 3.4.17.-)	세포질
M16 (피트릴리신 패밀리)					
		RXF03441.1		조효소 pqq 합성 단백질 F (ec 3.4.99.-)	비분비성

[0139]

	RXF01918.1	아연 프로테아제(ec 3.4.99.-)	신호 펩티드
	RXF01919.1	아연 프로테아제(ec 3.4.99.-)	주변세포질
	RXF03699.2	프로세싱 펩티다제 (ec 3.4.24.64)	신호 펩티드
M17 (류실 아미노펩티다제 패밀리)			
	RXF00285.2	사이토솔 아미노펩티다제 (ec 3.4.11.1)	비분비성
M18			
	RXF07879.1	아스파르틸 아미노펩티다제 (ec 3.4.11.21)	세포질
M20			
	RXF00811.1	dapE 속시닐-디아미노피멜레이트 데속시닐라제(ec 3.5.1.18)	세포질
	RXF04052.2	Xaa-His 디펩티다제 (ec 3.4.13.3)	신호 펩티드
	RXF01822.2	카르복시펩티다제 G2 전구체 (ec 3.4.17.11)	신호 펩티드
	RXF09831.2::RXF 04892.1	N-아실-L-아미노산 아미도히드롤라제 (ec 3.5.1.14)	신호 펩티드
M28 (아미노펩티다제 Y 패밀리)			
	RXF03488.2	알칼리성 포스파타제 이소자임 전환 단백질 전구체 (ec 3.4.11.-)	외막
M42 (글루타밀 아미노펩티다제 패밀리)			
	RXF05615.1	탈차단 아미노펩티다제 (ec 3.4.11.-)	비분비성
M22			
	RXF05817.1	O-시알로당단백질 엔도펩티다제 (ec 3.4.24.57)	세포외
	RXF03065.2	글리코프로테아제 단백질 패밀리	비분비성
M23			
	RXF01291.2	세포벽 엔도펩티다제, 패밀리 M23/M37	신호 펩티드
	RXF03916.1	메탈로엔도펩티다제와 관련된 막 단백질	신호 펩티드
	RXF09147.2	세포벽 엔도펩티다제, 패밀리 M23/M37	신호 펩티드
M24			
	RXF04693.1	메티오닌 아미노펩티다제 (ec 3.4.11.18)	세포질
	RXF03364.1	메티오닌 아미노펩티다제 (ec 3.4.11.18)	비분비성
	RXF02980.1	Xaa-Pro 아미노펩티다제 (ec 3.4.11.9)	세포질
	RXF06564.1	Xaa-Pro 아미노펩티다제 (ec 3.4.11.9)	세포질
M48 (Ste24 엔도펩티다제 패밀리)			
	RXF05137.1	열 쇼크 단백질 HtpX	세포질 막
	RXF05081.1	아연 메탈로프로테아제 (ec 3.4.24.-)	신호 펩티드

[0140]

M50 (S2P 프로테아제 패밀리)			
	RXF04692.1	막 메탈로프로테아제	세포질 막
세린 펩티다제			
S1 (키모트립신 패밀리)			
	RXF01250.2	프로테아제 do (ec 3.4.21.-)	주변세포질
	RXF07210.1	프로테아제 do (ec 3.4.21.-)	주변세포질
S8 (서브틸리신 패밀리)			
	RXF06755.2	세린 프로테아제 (ec 3.4.21.-)	비분비성
	RXF08517.1	세린 프로테아제 (ec 3.4.21.-)	세포외
	RXF08627.2	세포외 세린 프로테아제 (ec 3.4.21.-)	신호 펩티드
	RXF06281.1	세포외 세린 프로테아제 전구체 (ec 3.4.21.-)	비분비성
	RXF08978.1	세포외 세린 프로테아제 (ec 3.4.21.-)	외막
	RXF06451.1	세린 프로테아제 (ec 3.4.21.-)	신호 펩티드
S9 (프롤릴 올리고펩티다제 패밀리)			
	RXF02003.2	프로테아제 ii (ec 3.4.21.83)	주변세포질
	RXF00458.2	히드롤라제	비분비성
S11 (D-Ala-D-Ala 카르복시펩티다제 A 패밀리)			
	RXF04657.2	D-알라닐-D-알라닌-엔도펩티다제 (ec 3.4.99.-)	주변세포질
	RXF00670.1	D-알라닐-D-알라닌 카르복시펩티다제 (ec 3.4.16.4)	세포질 막
S13 (D-Ala-D-Ala 펩티다제 C 패밀리)			
	RXF00133.1	D-알라닐-메소-디아미노피멜레이트 엔도펩티다제 (ec 3.4.-.-)	외막
	RXF04960.2	D-알라닐-메소-디아미노피멜레이트 엔도펩티다제 (ec 3.4.-.-)	신호 펩티드
S14 (ClpP 엔도펩티다제 패밀리)			
	RXF04567.1	clpP atp 의존성 Clp 프로테아제 단백질 분해 서브유닛 (ec 3.4.21.92)	비분비성
	RXF04663.1	clpP atp 의존성 Clp 프로테아제 단백질 분해 서브유닛 (ec 3.4.21.92)	세포질
S16 (lon 프로테아제 패밀리)			
	RXF04653.2	atp 의존성 프로테아제 La (ec 3.4.21.53)	세포질
	RXF08653.1	atp 의존성 프로테아제 La (ec 3.4.21.53)	세포질
	RXF05943.1	atp 의존성 프로테아제 La (ec 3.4.21.53)	세포질
S24 (LexA 패밀리)			
	RXF00449.1	LexA 리프레서 (ec 3.4.21.88)	비분비성
	RXF03397.1	LexA 리프레서 (ec 3.4.21.88)	세포질

[0141]

S26 (신호 펩티다제 I 패밀리)			
	RXF01181.1	신호 펩티다제 I (ec 3.4.21.89)	세포질 막
S33			
	RXF05236.1	pip3	프롤린 이미노펩티다제 (ec 3.4.11.5)
	RXF04802.1	pip1	프롤린 이미노펩티다제 (ec 3.4.11.5)
	RXF04808.2	pip2	프롤린 이미노펩티다제 (ec 3.4.11.5)
S41 (C 말단 프로세싱 펩티다제 패밀리)			
	RXF06586.1	테일 특이 프로테아제 (ec 3.4.21.-)	신호 펩티드
	RXF01037.1	테일 특이 프로테아제 (ec 3.4.21.-)	신호 펩티드
S45			
	RXF07170.1	pacB2	페니실린 아실라제 (ec 3.5.1.11)
	RXF06399.2	pacB1	페니실린 아실라제 ii (ec 3.5.1.11)
S49 (프로테아제 IV 패밀리)			
	RXF06993.2	가능한 프로테아제 sohb (ec 3.4.-.)	비분비성
	RXF01418.1	프로테아제 iv (ec 3.4.-.)	비분비성
S58 (DmpA 아미노펩티다제 패밀리)			
	RXF06308.2	D-아미노펩티다제 (ec 3.4.11.19)	세포질 막
트레오닌 펩티다제			
T1 (프로테아좀 패밀리)			
	RXF01961.2	hslV	atp 의존성 프로테아제 hslV (ec 3.4.25.-)
T3 (감마 글루타밀 트랜스퍼라제 패밀리)			
	RXF02342.1	ggt1	감마 글루타밀트랜스펩티다제 (ec 2.3.2.2)
	RXF04424.2	ggt2	감마 글루타밀트랜스펩티다제 (ec 2.3.2.2)
미분류펩티다제			
U32			
	RXF00428.1		프로테아제 (ec 3.4.-.)
	RXF02151.2		프로테아제 (ec 3.4.-.)
U61			
	RXF04715.1		유라모일테트라펩티드 카르복시펩티다제 (ec 3.4.17.13)
U62			
	RXF04971.2	pmbA	PmbA 단백질
	RXF04968.2		TldD 단백질
비 MEROPS 프로테아제			
	RXF00325.1		리프레서 단백질 C2
	RXF02689.2		미세소체 디펩티다제 (ec 3.4.13.19)
	RXF02739.1		막 디펩티다제 (3.4.13.19)
	RXF03329.2		가상 사이토솔 단백질

[0142]

RXF02492.1	Xaa-Pro 디펩티다제 (ec 3.4.13.9)	세포질
RXF04047.2	caax 아미노 말단 프로테아제 패밀리	세포질 막
RXF08136.2	프로테아제 (트랜스글루타미나제 유사 단백질)	세포질
RXF09487.1	아연 메탈로프로테아제 (ec 3.4.24.-)	비분비성

[0143]

[0144]

특정 프로테아제는 프로테아제 및 샤페론 유사 활성, 둘 모두를 가질 수 있다. 이들 프로테아제가 단백질 수율 및/또는 품질에 부정적인 영향을 미치는 경우, 이들을 제거하는 것이 유용할 수 있으며, 이들의 샤페론 활성이 단백질 수율 및/또는 품질에 긍정적인 영향을 줄 수 있는 경우에는 이들을 과발현시킬 수 있다. 이들 프로테아제로는 Hsp100(Clp/Hsl) 패밀리 구성원 RXF04587.1(clpA), RXF08347.1, RXF04654.2(clpX), RXF04663.1, RXF01957.2(hslU), RXF01961.2(hslV); 펩티달-프롤린 시스-트랜스 이소머라제 패밀리 구성원 RXF05345.2(ppiB); 메탈로펩티다제 M20 패밀리 구성원 RXF04892.1(아미노히드롤라제); 메탈로펩티다제 M24 패밀리 구성원 RXF04693.1(메티오닌 아미노펩티다제) 및 RXF03364.1(메티오닌 아미노펩티다제); 및 세린 펩티다제 S26 신호 펩티다제 I 패밀리 구성원 RXF01181.1(신호 펩티다제)을 포함하나, 이에 한정되지 않는다.

표 7

표 7. P. 플루오레센스 균주 MB214 단백질 폴딩 조절 인자

ORF ID	유전자	기능	패밀리	위치
GroES/EL				
RXF02095.1	groES	샤페론	Hsp10	세포질
RXF06767.1:: Rxf02090	groEL	샤페론	Hsp60	세포질
RXF01748.1	ibpA	소형 열 쇼크 단백질 (sHSP) Ibpa PA3126; GroESL 폴딩을 위한 홀더로서 작용	Hsp20	세포질
RXF03385.1	hscB	샤페론 단백질 hscB	Hsp20	세포질
Hsp70 (DnaK/J)				
RXF05399.1	dnaK	샤페론	Hsp70	주변세포질
RXF06954.1	dnaK	샤페론	Hsp70	세포질
RXF03376.1	hscA	샤페론	Hsp70	세포질
RXF03987.2	cbpA	만곡형 dna 결합 단백질, dnaJ 유사 활성	Hsp40	세포질
RXF05406.2	dnaJ	샤페론 단백질 dnaJ	Hsp40	세포질
RXF03346.2	dnaJ	분자 샤페론 (DnaJ 패밀리)	Hsp40	비분비성
RXF05413.1	grpE	열 쇼크 단백질 GrpE PA4762	GrpE	세포질
Hsp100 (Clp/Hsl)				
RXF04587.1	clpA	atp 의존성 clp 프로테아제 atp 결합 서브유닛 clpA	Hsp100	세포질
RXF08347.1	clpB	ClpB 단백질	Hsp100	세포질
RXF04654.2	clpX	atp 의존성 clp 프로테아제 atp 결합 서브유닛 clpX	Hsp100	세포질
RXF04663.1	clpP	atp 의존성 Clp 프로테아제 단백질 분해 서브유닛 (ec 3.4.21.92)	MEROPS 펩티다제 패밀리 S14	세포질
RXF01957.2	hslU	atp 의존성 hsl 프로테아제 atp 결합 서브유닛 hslU	Hsp100	세포질
RXF01961.2	hslV	atp 의존성 hsl 프로테아제 단백질 분해 서브유닛	MEROPS 펩티다제 서브패밀리 T18	세포질
Hsp33				
RXF04254.2	yrfI	33 kDa 샤페론인 (열 쇼크 단백질 33 호모로그) (HSP33).	Hsp33	세포질
Hsp90				
RXF05455.2	htpG	샤페론 단백질 htpG	Hsp90	세포질
SecB				
RXF02231.1	secB	분비 특이 샤페론 SecB	SecB	비분비성
이황화 결합 이소머라제				
RXF07017.2	dsbA	이황화 이소머라제	DSBA 옥시도 리덕타제	세포질
RXF08657.2	dsbA/ dsbC/ dsbG/ fernA	이황화 이소머라제	DSBA 옥시도 리덕타제	세포질

[0145]

RXF01002.1	dsbA/ dsbC	이황화 이소머라제	DSBA 옥시도 리덕타제/ 티오레독신	주변세포질
RXF03307.1	dsbC	이황화 이소머라제	글루타레독신/ 티오레독신	주변세포질
RXF04890.2	dsbG	이황화 이소머라제	글루타레독신/ 티오레독신	주변세포질
RXF03204.1	dsbB	이황화 결합 형성 단백질 B (이황화 옥시도리덕타제).	DSBA 옥시도 리덕타제	주변세포질
RXF04886.2	dsbD	티올:이황화 상호교환 단백질 dsbD	DSBA 옥시도 리덕타제	주변세포질
펩티달-프롤릴 시스-트랜스 이소머라제				
RXF03768.1	ppiA	펩티달-프롤릴 시스-트랜스 이소머라제 A (ec 5.2.1.8)	PPI 아제: 시클로필린형	주변세포질
RXF05345.2	ppiB	펩티달-프롤릴 시스-트랜스 이소머라제 B.	PPI 아제: 시클로필린형	세포질
RXF06034.2	fkIB	펩티달-프롤릴 시스-트랜스 이소머라제 FkIB.	PPI 아제: FKBP 형	외막
RXF06591.1	fkIB/ fkBP	fk506 결합 단백질 펩티달-프롤릴 시스-트랜스 이소머라제 (EC 5.2.1.8)	PPI 아제: FKBP 형	주변세포질
RXF05753.2	fkIB; fkBP	펩티달-프롤릴 시스-트랜스 이소머라제 (ec 5.2.1.8)	PPI 아제: FKBP 형	외막
RXF01833.2	slyD	펩티달-프롤릴 시스-트랜스 이소머라제 SlyD.	PPI 아제: FKBP 형	비분비성
RXF04655.2	tig	촉발 인자, PPI 아제 (ec 5.2.1.8)	PPI 아제: FKBP 형	세포질
RXF05385.1	yaad	가능 FKBP 형 16 kDa 펩티달-프롤릴 시스-트랜스 이소머라제 (EC 5.2.1.8) (PPI 아제) (로타마제).	PPI 아제: FKBP 형	비분비성
RXF00271.1		펩티달-프롤릴 시스-트랜스 이소머라제 (ec 5.2.1.8)	PPI 아제: FKBP 형	비분비성
pili 어셈블리샤페론 (papD 유사)				
RXF06068.1	cup	샤페론 단백질 cup	pili 어셈블리 papD	주변세포질
RXF05719.1	ecpD	샤페론 단백질 ecpD	pili 어셈블리 papD	신호 펩티드
RXF05319.1	ecpD	Hnr 단백질	pili 어셈블리 샤페론	주변세포질
RXF03406.2	ecpD; csuC	샤페론 단백질 ecpD	pili 어셈블리 papD	신호 펩티드
RXF04296.1	ecpD; cup	샤페론 단백질 ecpD	pili 어셈블리 papD	주변세포질
RXF04553.1	ecpD; cup	샤페론 단백질 ecpD	pili 어셈블리 papD	주변세포질
RXF04554.2	ecpD; cup	샤페론 단백질 ecpD	pili 어셈블리 papD	주변세포질
RXF05310.2	ecpD;	샤페론 단백질 ecpD	pili 어셈블리	주변세포질

[0146]

RXF05304.1	cup ecpD;	샤페론 단백질 ecpD	papD pili 어셈블리	주변세포질
RXF05073.1	cup gltF	그람 음성 pili 어셈블리샤페론 주변세포질 기능	papD pili 어셈블리 papD	신호 펩티드
II 형 분비 복합체				
RXF05445.1	YacJ	히스티딘-포스페이트 아미노트랜스퍼라제 (ec 2.6.1.9)	부류 II 피리독살 포스페이트 의존성 아미노 트랜스퍼라제 패밀리. 히스티딘 포스페이트 아미노 트랜스퍼라제 서브패밀리.	막
RXF05426.1	SecD	단백질 트랜스로카제 서브유닛 secD	II 형 분비 복합체	막
RXF05432.1	SecF	단백질 트랜스로카제 서브유닛 secF	II 형 분비 복합체	막
이항화 결합 리덕타제				
RXF08122.2	trxC	티오레독신 2	이항화 결합 리덕타제	세포질
RXF06751.1	Gor	글루타민 리덕타제 (EC 1.8.1.7) (GR) (GR 아제) PA2025	이항화 결합 리덕타제	세포질
RXF00922.1	gshA	글루타메이트 시스테인 리가제 (ec 6.3.2.2) PA5203	이항화 결합 리덕타제	세포질

[0147]

[0148]

고처리량 스크린

[0149]

일부 실시양태에서, 가용성 재조합 독소 단백질을 발현시키는 데 최적의 조건을 측정하기 위해 고처리량 스크린을 수행할 수 있다. 스크린에서 달리할 수 있는 조건으로는 예를 들어, 숙주 세포, 숙주 세포의 유전적 배경(예를 들어, 상이한 프로테아제의 결실), 발현 구성체 중 프로모터의 유형, 재조합 단백질을 코딩하는 서열에 융합되는 분비 리더의 유형, 성장 온도, 유도성 프로모터를 사용하였을 때, 유도 시 OD, lacZ 프로모터가 사용될 때, 유도를 위해 사용되는 IPTG의 농도, 단백질 유도 지속 기간, 배양물에 유도체를 첨가한 후의 성장 온도, 배양물 교반 속도, 플라스미드 유지를 위한 선별 방법, 배설 중 배양물 부피, 및 세포 용해 방법을 포함한다.

[0150]

일부 실시양태에서, 숙주 세포 중의 하나 이상의 표적 유전자의 발현을 조정하기 위해 라이브러리 중의 각 균주(또는 "숙주 세포 집단")가 유전적으로 변형되어 있는 것인, 숙주 균주의 라이브러리(또는 "어레이")를 제공한다. "최적의 숙주 균주" 또는 "최적의 발현 시스템"은 어레이 중 표현형이 상이한 다른 숙주 세포 집단과 비교하여 관심의 대상이 되는 발현 단백질의 정량, 정질, 및/또는 위치에 기초하여 확인 또는 선택될 수 있다. 따라서, 최적의 숙주 균주는 원하는 요구 조건에 따라 관심의 대상이 되는 폴리펩티드를 생산하는 균주이다. 원하는 요구 조건은 생산하고자 하는 폴리펩티드에 따라 달라지겠지만, 요구 조건으로는 단백질의 정질 및/또는 정량, 예를 들어, 단백질이 격리 또는 분비된 것인지 여부, 및 정량이 얼마나 되는지, 단백질이 적절히 또는 바람직하게 프로세싱 및/또는 폴딩되었는지 여부 등을 포함한다. 실시양태에서, 개선된 또는 바람직한 정질이란 분비 리더가 높은 신뢰도로 절단되고, 낮은 수준으로 분해되는 독소 단백질이 생산되는 것일 수 있다. 실시양태에서, 최적의 숙주 균주 또는 최적의 발현 시스템은 가용성 이중성 단백질의 양 또는 정량, 재생가능한 이중성 단백질의 양 또는 정량, 적절히 프로세싱된 이중성 단백질의 양 또는 정량, 적절히 폴딩된 이중성 단백질의 양 또는 정량, 활성 이중성 단백질의 양 또는 정량, 및/또는 이중성 단백질의 총량 또는 총 정량을 특징으로 하는 수율을 인디케이터 균주, 즉, 비교용으로 사용되는 균주에 의해 생산되는 것에 비교되는 특성의 절대 수준 또는 특정 수준으로 얻는다.

[0151]

이중성 단백질 발현에 있어 수율 및/또는 정질이 개선된 균주를 확인하기 위해 미생물 숙주를 스크리닝하는 방법은 예를 들어, 미국 특허 출원 공개공보 20080269070에 기술되어 있다.

[0152]

발효 방식

[0153]

본 발명에 따른 발현 시스템을 임의의 발효 방식으로 배양할 수 있다. 예를 들어, 본원에서는 회분식, 공급형

회분식, 반연속식 및 연속식 발효 모드를 사용할 수 있다.

- [0154] 실시양태에서, 발효 배지는 부영양 배지(rich media), 최소영양 배지 및 미네랄 염 배지로부터 선택될 수 있다. 다른 실시양태에서, 최소영양 배지 또는 미네랄 염 배지가 선택된다. 특정 실시양태에서, 미네랄 염 배지가 선택된다.
- [0155] 미네랄 염 배지는 미네랄 염과 탄소원, 예를 들어, 글루코스, 수크로스 또는 글리세롤로 구성된다. 미네랄 염 배지의 예는, 예를 들어, M9 배지, 슈도모나스 배지(ATCC 179)와, 아비스 및 민지오리(Davis and Mingioli) 배지를 포함한다(문헌 [B D Davis & E S Mingioli (1950) J. Bact. 60: 17-28] 참조). 미네랄 염 배지를 제조하는 데 사용되는 미네랄 염은 예를 들어, 인산칼륨, 황산암모늄 또는 염화암모늄, 황산마그네슘 또는 염화마그네슘, 및 미량 미네랄, 예를 들어, 염화칼슘, 붕산염 및 철, 구리, 망간 및 아연의 황산염으로부터 선택된 것들을 포함한다. 전형적으로, 예를 들어, 펩톤, 트립톤, 아미노산 또는 효모 추출물과 같은 유기 질소원은 미네랄 염 배지 중에 포함되지 않는다. 대신, 무기 질소원이 사용되며, 이는 예를 들어, 암모늄염, 수성 암모니아 및 기상 암모니아로부터 선택될 수 있다. 미네랄 염 배지는 전형적으로 글루코스 또는 글리세롤을 탄소원으로 포함한다. 미네랄 염 배지와 비교하여, 최소영양 배지는 또한 미네랄 염과 탄소원을 포함할 수 있지만, 비록 매우 매우 최소의 농도로 첨가될지라도, 예를 들어, 낮은 농도의 아미노산, 비타민, 펩톤 또는 다른 성분들로 보충될 수는 있다. 당업계에 기술된 방법, 예를 들어, 미국 특허 출원 공개공보 2006/0040352호(이는 상기에 참고 문헌으로서 참조되고 포함되어 있다)에 개시된 방법을 사용하여 배지를 제조할 수 있다. 본 발명의 방법에 유용한 배양 절차 및 미네랄 염 배지의 상세한 설명은 문헌 [Riesenberg, D et al., 1991, "High cell density cultivation of *Escherichia coli* at controlled specific growth rate," J. Biotechnol. 20 (1): 17-27]에 개시되어 있다.
- [0156] 실시양태에서, 생물반응기 배양으로 제조될 수 있다. 예를 들어, 미네랄 염 배지를 함유하는 최대 2 ℓ 생물반응기에서 배양물을 배양하고, 암모니아를 첨가하여 32℃ 및 pH 6.5에서 유지할 수 있다. 용존 산소는 발효기 내의 살포 공기와 산소의 흐름 및 교반 증가를 통해 과량으로 유지할 수 있다. 과량의 수준을 유지하기 위해 발효 과정 동안 글리세롤을 배양물에 전달할 수 있다. 실시양태에서, 유도를 위한 표적 배양물 세포 밀도, 예를 들어, 575 nm에서의 광학 밀도(A_{575})에 도달할 때까지 이들 조건을 유지하고, 이때 IPTG를 첨가하여 표적 단백질 생성을 개시한다. 유도 시 세포 밀도, IPTG 농도, pH 및 온도, 이들 각각을 다양하게 하여 최적의 발현 조건을 결정할 수 있다는 것을 이해할 수 있을 것이다. 실시양태에서, 유도 시 세포 밀도는 A_{575} 40~200 흡광 단위(AU)로 다양할 수 있다. IPTG 농도는 0.02~1.0 mM 범위로, pH는 6~7.5 범위로, 온도는 20~35℃ 범위로 다양할 수 있다. 16~24시간 후에, 각 생물반응기로부터의 배양물을 원심분리로 수거하고, -80℃에서 세포 펠릿을 냉동시킬 수 있다. 이어서, 예를 들어, SDS-CGE에 의해 생성물 제조에 대해 샘플을 분석할 수 있다.
- [0157] 발효는 어느 규모로든 수행될 수 있다. 본 발명에 따른 발현 시스템은 임의의 규모의 재조합 단백질 발현에 유용하다. 따라서, 예를 들어, 마이크로리터 규모, 밀리리터 규모, 센티리터 규모 및 데시리터 규모의 발효 부피가 사용될 수 있고, 1 리터 이상의 발효 부피도 사용가능하다.
- [0158] 실시양태에서, 발효 부피는 약 1 ℓ 이상이다. 실시양태에서, 발효 부피는 약 1 ℓ ~ 약 100 ℓ 이다. 실시양태에서, 발효 부피는 약 1 ℓ, 약 2 ℓ, 약 3 ℓ, 약 4 ℓ, 약 5 ℓ, 약 6 ℓ, 약 7 ℓ, 약 8 ℓ, 약 9 ℓ, 또는 약 10 ℓ 이다. 실시양태에서, 발효 부피는 약 1 ℓ ~ 약 5 ℓ, 약 1 ℓ ~ 약 10 ℓ, 약 1 ℓ ~ 약 25 ℓ, 약 1 ℓ ~ 약 50 ℓ, 약 1 ℓ ~ 약 75 ℓ, 약 10 ℓ ~ 약 25 ℓ, 약 25 ℓ ~ 약 50 ℓ, 또는 약 50 ℓ ~ 약 100 ℓ 이다. 다른 실시양태에서, 발효 부피는 5 ℓ, 10 ℓ, 15 ℓ, 20 ℓ, 25 ℓ, 50 ℓ, 75 ℓ, 100 ℓ, 200 ℓ, 500 ℓ, 1,000 ℓ, 2,000 ℓ, 5,000 ℓ, 10,000 ℓ, 또는 50,000 ℓ 이상이다.
- [0159] **박테리아 성장 조건**
- [0160] 제공하는 본 발명의 방법에서 유용한 성장 조건은 약 4℃ ~ 약 42℃의 온도 및 약 5.7 ~ 약 8.8의 pH를 포함할 수 있다. lacZ 프로모터를 포함하는 발현 구성체가 사용되는 경우, 발현은 IPTG를 약 0.01 mM ~ 약 1.0 mM의 최종 농도로 첨가함으로써 유도될 수 있다.
- [0161] 배양물의 pH는 pH 완충액 및 당업자에게 공지된 방법을 사용함으로써 유지될 수 있다. 또한, 수성 암모니아를 사용하여 배양하는 동안 pH를 조정할 수도 있다. 실시양태에서, 배양물의 pH는 약 5.7 ~ 약 8.8이다. 특정 실시양태에서, pH는 약 5.7, 5.8, 5.9, 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8.0, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 또는 8.8이다. 다른 실시양태에서, pH는 약 5.7~5.9, 5.8~6.0, 5.9~6.1, 6.0~6.2, 6.1~6.3, 6.2~6.5, 6.4~6.7, 6.5~6.8, 6.6~6.9, 6.7~

7.0, 6.8~7.1, 6.9~7.2, 7.0~7.3, 7.1~7.4, 7.2~7.5, 7.3~7.6, 7.4~7.7, 7.5~7.8, 7.6~7.9, 7.7~8.0, 7.8~8.1, 7.9~8.2, 8.0~8.3, 8.1~8.4, 8.2~8.5, 8.3~8.6, 8.4~8.7, 또는 8.5~8.8이다. 추가의 다른 실시양태에서, pH는 약 5.7~6.0, 5.8~6.1, 5.9~6.2, 6.0~6.3, 6.1~6.4, 또는 6.2~6.5이다. 특정 실시양태에서, pH는 약 5.7~약 6.25이다.

[0162] 실시양태에서, 성장 온도는 약 4℃~약 42℃로 유지된다. 특정 실시양태에서, 성장 온도는 약 4℃, 약 5℃, 약 6℃, 약 7℃, 약 8℃, 약 9℃, 약 10℃, 약 11℃, 약 12℃, 약 13℃, 약 14℃, 약 15℃, 약 16℃, 약 17℃, 약 18℃, 약 19℃, 약 20℃, 약 21℃, 약 22℃, 약 23℃, 약 24℃, 약 25℃, 약 26℃, 약 27℃, 약 28℃, 약 29℃, 약 30℃, 약 31℃, 약 32℃, 약 33℃, 약 34℃, 약 35℃, 약 36℃, 약 37℃, 약 38℃, 약 39℃, 약 40℃, 약 41℃, 또는 약 42℃이다. 다른 실시양태에서, 성장 온도는 약 25℃~약 27℃, 약 25℃~약 28℃, 약 25℃~약 29℃, 약 25℃~약 30℃, 약 25℃~약 31℃, 약 25℃~약 32℃, 약 25℃~약 33℃, 약 26℃~약 28℃, 약 26℃~약 29℃, 약 26℃~약 30℃, 약 26℃~약 31℃, 약 26℃~약 32℃, 약 27℃~약 29℃, 약 27℃~약 30℃, 약 27℃~약 31℃, 약 27℃~약 32℃, 약 26℃~약 33℃, 약 28℃~약 30℃, 약 28℃~약 31℃, 약 28℃~약 32℃, 약 29℃~약 31℃, 약 29℃~약 32℃, 약 29℃~약 33℃, 약 30℃~약 32℃, 약 30℃~약 33℃, 약 31℃~약 33℃, 약 31℃~약 32℃, 약 30℃~약 33℃, 또는 약 32℃~약 33℃로 유지된다. 다른 실시양태에서, 온도는 배양하는 동안 변동된다. 한 실시양태에서, 온도는 구성체로부터 발현을 유도하는 제제, 예를 들어, IPTG를 배양물에 첨가하기 전, 약 30℃로 유지된다. 유도제 첨가 후, 온도는 약 25℃로 강해진다.

[0163] 유도

[0164] 본원에 전역에 걸쳐 기술된 바와 같이, 재조합 독소 단백질의 발현을 제어하기 위해 발현 구성체 중 유도성 프로모터, 예를 들어, lac 프로모터가 사용될 수 있다. lac 프로모터 유도체 또는 패밀리 구성원, 예를 들어, tac 프로모터인 경우, 이펙터 화합물은 인듀서, 예를 들어, 무상 인듀서, 예를 들어, IPTG(이소프로필-β-D-1-티오갈락토피라노시드, 이는 또한 "이소프로필티오갈락토시드"로도 명명된다)이다. 실시양태에서, lac 프로모터 유도체가 사용되고, 재조합 단백질 발현은, 세포 밀도가 OD575가 약 80~약 160인 것으로 확인되는 수준에 도달하였을 때 발현은 IPTG를 약 0.01 mM~약 1.0 mM의 최종 농도로 첨가함으로써 유도된다. 실시양태에서, 재조합 단백질을 위한 배양 유도 시점에서의 OD575는 약 80, 약 90, 약 100, 약 110, 약 120, 약 130, 약 140, 약 150, 약 160, 약 170 약 180일 수 있다. 다른 실시양태에서, OD575는 약 80~약 100, 약 100~약 120, 약 120~약 140, 약 140~약 160이다. 다른 실시양태에서, OD575는 약 80~약 120, 약 100~약 140, 또는 약 120~약 160이다. 다른 실시양태에서, OD575는 약 80~약 140, 또는 약 100~160이다. 세포 밀도는 다른 방법에 의해 측정될 수 있고, 다른 단위로, 예를 들어, 단위 부피당 세포수로 표시될 수 있다. 예를 들어, 슈도모나스 플루오레센스 배양물의 OD575가 약 80~약 160인 것은 대략 1 mL당 8×10^{10} ~약 1.6×10^{11} 콜로니 형성 단위, 또는 건식 세포 중량으로 1 L당 35~70 g인 것과 등가이다. 실시양태에서, 배양 유도 시점에서의 세포 밀도는 세포 밀도를 측정하는 데 사용되는 방법, 또는 그의 측정 단위와는 상관없이, 본원에서 OD575에서의 흡광도에 의해 명시된 것과 같은 세포 밀도와 등가이다. 당업자는 임의의 세포 배양을 위해 적절히 전환시키는 방법을 알고 있을 것이다.

[0165] 실시양태에서, 배양물 중 IPTG의 최종 농도는 약 0.01 mM, 약 0.02 mM, 약 0.03 mM, 약 0.04 mM, 약 0.05 mM, 약 0.06 mM, 약 0.07 mM, 약 0.08 mM, 약 0.09 mM, 약 0.1 mM, 약 0.2 mM, 약 0.3 mM, 약 0.4 mM, 약 0.5 mM, 약 0.6 mM, 약 0.7 mM, 약 0.8 mM, 약 0.9 mM, 또는 약 1 mM이다. 다른 실시양태에서, 배양물 중 IPTG의 최종 농도는 약 0.08 mM~약 0.1 mM, 약 0.1 mM~약 0.2 mM, 약 0.2 mM~약 0.3 mM, 약 0.3 mM~약 0.4 mM, 약 0.2 mM~약 0.4 mM, 약 0.08~약 0.2 mM, 또는 약 0.1~1 mM이다.

[0166] 실시양태에서, lac 프로모터가 아닌 다른 유형의 프로모터가 사용되는 경우에는, 본원 및 문헌에 기술된 바와 같이, 다른 인듀서 또는 이펙터가 사용될 수 있다. 한 실시양태에서, 프로모터는 구성적 프로모터이다.

[0167] 유도제를 첨가한 후, 배양물을 일정 기간 동안, 예를 들어, 약 24시간 동안 성장시킬 수 있는데, 이 기간 동안 재조합 단백질은 발현된다. 유도제를 첨가한 후, 배양물을 약 1 hr, 약 2 hr, 약 3 hr, 약 4 hr, 약 5 hr, 약 6 hr, 약 7 hr, 약 8 hr, 약 9 hr, 약 10 hr, 약 11 hr, 약 12 hr, 약 13 hr, 약 14 hr, 약 15 hr, 약 16 hr, 약 17 hr, 약 18 hr, 약 19 hr, 약 20 hr, 약 21 hr, 약 22 hr, 약 23 hr, 약 24 hr, 약 36 hr, 또는 약 48 hr 동안 성장시킬 수 있다. 유도제를 배양물에 첨가한 후, 배양물을 약 1~48 hrs, 약 1~24 hrs, 약 10~24 hrs, 약 15~24 hrs, 또는 약 20~24 hrs 동안 성장시킬 수 있다. 세포 배양물을 원심분리에 의해 농축시킬 수 있고, 추후 용해 절차에 적합한 완충액 또는 용액 중에 배양물 펠릿을 재현탁시킬 수 있다.

[0168] 실시양태에서, 고압 기계적 세포 파괴용 장치(이는 상업적으로 이용가능하며, 그 예로는 마이크로플루이딕스 마

이크로플로다이어저(Microfluidics Microfluidizer), 콘스턴트 셀 디스rupter(Constant Cell Disruptor), 니로-소아비(Niro-Soavi) 균질화기 또는 APV-가울린(APV-Gaulin) 균질화기)를 사용하여 세포를 파괴시킨다. 초음파 처리를 사용함으로써 제조합 단백질 발현 세포를 파괴시킬 수 있다. 당업계에 공지된, 세포를 용해시키는 데 적절한 임의의 방법을 사용하여 가용성 분획을 유리시킬 수 있다. 예를 들어, 실시양태에서, 화학적 및/또는 효소적 세포 용해 시약, 예를 들어, 세포벽 용해 효소 및 EDTA가 사용될 수 있다. 본 발명의 방법에서는 또한 냉동 배양물 또는 미리 보관된 배양물을 사용하는 것 또한 고려된다. 용해하기 전에 배양물을 OD로 정규화할 수 있다. 예를 들어, 세포를 OD600이 약 10, 약 11, 약 12, 약 13, 약 14, 약 15, 약 16, 약 17, 약 18, 약 19, 또는 약 20인 것으로 정규화될 수 있다.

[0169] 임의의 적절한 장치 및 방법을 사용하여 원심분리를 수행할 수 있다. 불용성 분획으로부터 가용성 분리시키고자 하는 목적을 위해 세포 배양물 또는 용해물을 원심분리시키는 것은 당업계에 주지되어 있다. 예를 들어, 용해된 세포를 20분(4℃) 동안 20,800 x g로 원심분리하고, 수동 또는 자동 액상 취급기를 사용하여 상청액을 제거할 수 있다. 펠릿(불용성) 분획을 완충처리된 용액, 예를 들어, 인산염 완충처리된 염수(PBS: phosphate buffered saline, pH 7.4) 중에 재현탁시켰다. 오버헤드 혼합기, 자기 교반 막대, 진동식 진탕기 등에 연결된 임펠러와 같은 장치를 사용하여 재현탁을 수행하였다.

[0170] "가용성 분획," 즉, 용해물의 원심분리 후 수득된 가용성 상청액, 및 "불용성 분획," 즉, 용해물의 원심분리 후 수득된 펠릿은 배양물의 용해 및 원심분리로부터 생성된 것이다. 이들 두 분획은 모두 "제1 가용성 분획" 및 "제1 불용성 분획"이라고 지칭될 수 있다.

[0171] 생성물 평가

[0172] 단백질의 특징을 규명하는 분석법들 다수가 당업계에 공지되어 있다. 제조합 독소 단백질의 수율 또는 정질의 특징을 규명하는 임의의 적절한 방법을 사용하는 것도 본원에서 고려된다.

[0173] 단백질 수율

[0174] 본원에 기술된 바와 같이, 임의의 정제 분획의 단백질 수율은 당업자에게 공지된 방법, 예를 들어, 모세관 겔 전기영동(CGE: capillary gel electrophoresis) 및 웨스턴 블롯 분석으로 측정될 수 있다. 본원에 기술되어 있고, 당업계에 공지된 바와 같이, 활성 검정법은 또한 단백질 수율에 관한 정보를 제공할 수 있다. 실시양태에서, 상기 방법 또는 당업계에 공지된 임의의 다른 방법을 사용하여 단백질의 적절한 프로세싱, 예를 들어, 적절한 분비 리더 절단을 평가할 수 있다.

[0175] 단백질 수율의 유용한 척도로는 예를 들어, 배양물 부피당 제조합 단백질의 양(예를 들어, 배양물 1 ℓ당 단백질의 g 또는 mg), 세포 용해 후 수득된 불용성 펠릿 중에서 측정되는 제조합 단백질의 백분율 또는 분율(예를 들어, 추출 상청액 중 제조합 단백질의 양/불용성 분획 중 단백질의 양), 활성 단백질의 백분율 또는 분율(예를 들어, 활성 단백질의 양/검정에 사용된 단백질의 양), 총 세포 단백질(tcp: total cell protein)의 백분율 또는 분율, 세포 1개당 단백질의 양, 및 건조 바이오매스의 백분율 또는 비율을 포함한다. 실시양태에서, 본원에 기술된 단백질 수율의 척도는 수득된, 가용성 단백질의 양, 또는 활성 단백질의 양, 또는, 그 둘 모두를 기초로 한다.

[0176] 수율이 배양물 부피로 표시되는 실시양태에서, 특히 상이한 배양물 간의 수율을 비교하는 경우에는 배양물의 세포 밀도를 고려할 수 있다.

[0177] 실시양태에서, 본 발명의 방법을 사용하여 가용성 및/또는 활성 및/또는 적절히 프로세싱된(예를 들어, 분비 리더가 적절히 절단된) 제조합 독소 단백질 또는 서브유닛 단백질을 약 0.2 g/L~약 12 g/L의 수율로 수득할 수 있다. 실시양태에서, 수율은 약 0.5 g/L~약 12 g/L이다. 특정 실시양태에서, 제조합 단백질 또는 서브유닛 단백질 수율은 약 0.2 g/L, 약 0.3 g/L, 약 0.4 g/L, 약 0.5 g/L, 약 0.6 g/L, 약 0.7 g/L, 약 0.8 g/L, 약 0.9 g/L, 약 1 g/L, 약 1.5 g/L, 약 2 g/L, 약 2.5 g/L, 약 3 g/L, 약 3.5 g/L, 약 4 g/L, 약 4.5 g/L, 약 5 g/L, 약 5.5 g/L, 약 6 g/L, 약 6.5 g/L, 약 7 g/L, 약 7.5 g/L, 약 8 g/L, 약 8.5 g/L, 약 9 g/L, 약 9.5 g/L, 약 10 g/L, 약 10.5 g/L, 약 11 g/L, 약 12 g/L, 약 0.2 g/L~약 0.5 g/L, 약 0.2 g/L~약 1 g/L, 약 0.2~약 2 g/L, 약 0.3 g/L~약 0.6 g/L, 약 0.3 g/L~약 1 g/L, 약 0.3~약 2 g/L, 약 0.4~약 0.7 g/L, 약 0.4~약 1 g/L, 약 0.4~약 2 g/L, 약 0.4~약 3 g/L, 약 0.5 g/L~약 1 g/L, 약 0.5 g/L~약 1 g/L, 약 0.5 g/L~약 2 g/L, 약 0.5 g/L~약 3 g/L, 약 0.5 g/L~약 4 g/L, 약 0.5 g/L~약 5 g/L, 약 0.5 g/L~약 6 g/L, 약 0.5 g/L~약 7 g/L, 약 0.5 g/L~약 8 g/L, 약 0.5 g/L~약 9 g/L, 약 0.5 g/L~약 10 g/L, 약 0.5 g/L~약 11 g/L, 약 0.5 g/L~약 12 g/L, 약 1 g/L~약 2 g/L, 약 1 g/L~약 3 g/L, 약 1 g/L~약 4 g/L, 약 1 g/L~약 5 g/L,

약 1 g/L~약 6 g/L, 약 1 g/L~약 7 g/L, 약 1 g/L~약 8 g/L, 약 1 g/L~약 9 g/L, 약 1 g/L~약 10 g/L, 약 1 g/L~약 11 g/L, 약 1 g/L~약 12 g/L, 약 2 g/L~약 3 g/L, 약 2 g/L~약 4 g/L, 약 2 g/L~약 5 g/L, 약 2 g/L~약 6 g/L, 약 2 g/L~약 7 g/L, 약 2 g/L~약 8 g/L, 약 2 g/L~약 9 g/L, 약 2 g/L~약 10 g/L, 약 2 g/L~약 11 g/L, 약 2 g/L~약 12 g/L, 약 3 g/L~약 4 g/L, 약 3 g/L~약 5 g/L, 약 3 g/L~약 6 g/L, 약 3 g/L~약 7 g/L, 약 3 g/L~약 8 g/L, 약 3 g/L~약 9 g/L, 약 3 g/L~약 10 g/L, 약 3 g/L~약 11 g/L, 약 3 g/L~약 12 g/L, 약 4 g/L~약 5 g/L, 약 4 g/L~약 6 g/L, 약 4 g/L~약 7 g/L, 약 4 g/L~약 8 g/L, 약 4 g/L~약 9 g/L, 약 4 g/L~약 10 g/L, 약 4 g/L~약 11 g/L, 약 4 g/L~약 12 g/L, 약 5 g/L~약 6 g/L, 약 5 g/L~약 7 g/L, 약 5 g/L~약 8 g/L, 약 5 g/L~약 9 g/L, 약 5 g/L~약 10 g/L, 약 5 g/L~약 11 g/L, 약 5 g/L~약 12 g/L, 약 6 g/L~약 7 g/L, 약 6 g/L~약 8 g/L, 약 6 g/L~약 9 g/L, 약 6 g/L~약 10 g/L, 약 6 g/L~약 11 g/L, 약 6 g/L~약 12 g/L, 약 7 g/L~약 8 g/L, 약 7 g/L~약 9 g/L, 약 7 g/L~약 10 g/L, 약 7 g/L~약 11 g/L, 약 7 g/L~약 12 g/L, 약 8 g/L~약 9 g/L, 약 8 g/L~약 10 g/L, 약 8 g/L~약 11 g/L, 약 8 g/L~약 12 g/L, 약 9 g/L~약 10 g/L, 약 9 g/L~약 11 g/L, 약 9 g/L~약 12 g/L, 약 10 g/L~약 11 g/L, 약 10 g/L~약 12 g/L, 또는 약 11 g/L~약 12 g/L이다.

[0178] 실시양태에서, 제조함 독소 단백질 또는 서브유닛 단백질의 생산량은 총 세포 단백질의 약 1%~75%이다. 특정 실시양태에서, 독소 단백질 또는 서브유닛 단백질의 생산량은 총 세포 단백질의 약 1%, 약 2%, 약 3%, 약 4%, 약 5 %, 약 10%, 약 15 %, 약 20%, 약 25%, 약 30%, 약 35%, 약 40%, 약 45%, 약 50%, 약 55%, 약 60%, 약 65%, 약 70%, 약 75%, 약 1%~약 5%, 약 1%~약 10%, 약 1%~약 20%, 약 1%~약 30%, 약 1%~약 40%, 약 1%~약 50%, 약 1%~약 60%, 약 1%~약 75%, 약 2%~약 5%, 약 2%~약 10%, 약 2%~약 20%, 약 2%~약 30%, 약 2%~약 40%, 약 2%~약 50%, 약 2%~약 60%, 약 2%~약 75%, 약 3%~약 5%, 약 3%~약 10%, 약 3%~약 20%, 약 3%~약 30%, 약 3%~약 40%, 약 3%~약 50%, 약 3%~약 60%, 약 3%~약 75%, 약 4%~약 10%, 약 4%~약 20%, 약 4%~약 30%, 약 4%~약 40%, 약 4%~약 50%, 약 4%~약 60%, 약 4%~약 75%, 약 5%~약 10%, 약 5%~약 20%, 약 5%~약 30%, 약 5%~약 40%, 약 5%~약 50%, 약 5%~약 60%, 약 5%~약 75%, 약 10%~약 20%, 약 10%~약 30%, 약 10%~약 40%, 약 10%~약 50%, 약 10%~약 60%, 약 10%~약 75%, 약 20%~약 30%, 약 20%~약 40%, 약 20%~약 50%, 약 20%~약 60%, 약 20%~약 75%, 약 30%~약 40%, 약 30%~약 50%, 약 30%~약 60%, 약 30%~약 75%, 약 40%~약 50%, 약 40%~약 60%, 약 40%~약 75%, 약 50%~약 60%, 약 50%~약 75%, 약 60%~약 75%, 또는 약 70%~약 75%이다.

[0179] 특정 실시양태에서, 다중 단백질이 같은 숙주 세포로부터 생산된다. 예를 들어, 실시양태에서, 백일해 독소의 5 개의 서브유닛은 모두 단일의 배양물 중에서 성장된 같은 숙주 세포로부터 제조된 것이다. 상기 실시양태에서, 농도, 총 세포 단백질에 대한 백분율(%), 또는 관찰되는 활성은 각각의 개별 독소 서브유닛에 대한, 또는 서브 유닛들 모두에 대한 것이다. 즉, 실시양태에서, 본 발명의 방법을 사용하여 백일해 독소 단백질의 S1, S2, S3, S4, 또는 S5 서브유닛을 약 1 g/L~약 12 g/L의 수율로 수득할 수 있다. 실시양태에서, S1, S2, S3, S4, 또는 S5 서브유닛 단백질의 생산량은 총 세포 단백질의 1%~75%이다. 별법으로, 본 발명의 방법을 사용하여 S1, S2, S3, S4, 및 S5 서브유닛 단백질을 약 1 g/L~약 12 g/L의 수율로 수득한다. 실시양태에서, S1, S2, S3, S4, 및 S5 서브유닛 단백질의 생산량은 총 세포 단백질의 총 세포 단백질의 1%~75%이다. 특정 실시양태에서, 각 서브 유닛의 수득량(g/L, 또는 총 세포 단백질에 대한 백분율(%))은 거의 동일하다.

[0180] 단백질의 "가용성" 및 "활성"은 관련된 성질이기는 하지만, 일반적으로는 상이한 방법으로 측정된다. 단백질, 특히, 소수성 단백질의 가용성은 전형적으로 단백질의 폴딩과 관련이 있으며; 불용성은 소수성 아미노산 잔기가 폴딩된 단백질 외부에 부적절하게 위치함을 나타내는 것이다. 단백질 활성은, 예를 들어, 하기 기술하는 것과 같은 방법을 이용하여 평가될 수 있는데, 이는 적당한 단백질 구조의 또 다른 인디케이터이다. 본원에서 사용되는 바, "가용성, 활성 또는 둘 모두"는 당업자에게 공지되고, 본원에 기술된 방법으로 가용성, 활성 또는 가용 성이면서 활성인 것으로 확인되는 단백질을 의미한다. 주어진 단백질의 "활성"은 예를 들어, 수용체, 특이적 항 체 또는 또 다른 공지의 기질의 결합에 의해 나타나거나, 또는 관련된 경우, 효소 활성에 의해 나타나는 결합 활성을 포함할 수 있다. 활성 수준은 절대값으로, 또는 예를 들어, 표준 또는 대조군 샘플이나, 또는 기준으로 서 사용되는 임의의 샘플의 활성과 비교하였을 때 상대값으로 나타낼 수 있다.

[0181] 독소를 평가하는 활성 검정법은 당업계에 공지되어 있고, 문헌에 기술되어 있다. 활성 검정법으로는 면역학적 또는 항체 결합 검정법, 예를 들어, 웨스턴 블롯 분석 및 ELISA 뿐만 아니라, 수용체 결합 검정법을 포함하는데, 예를 들어, CRM197은 디프테리아 독소 수용체(proHB-EGF)에 의해 평가될 수 있다. 활성 검정법은 또한 효소 활성 검정법을 포함한다. 야생형 DT는 당업계에 공지되어 있고, 본원 전역에 걸쳐 기술되어 있는, P. 아에루기노사 외독소 A.에 대한 방법을 사용하여 면역학적으로 및 또한 ADP 리보실화 활성에 의해 검정될 수 있

다.

- [0182] 예를 들어, CTB의 웨스턴 블롯 분석은 예를 들어, 미국 특허 번호 제6,140,082호(발명의 명칭: "Expression of Gene Products from Genetically Manipulated Strains of *Bordetella*") (이 출원은 본원에 참고로 포함된다)에 기술된 바와 같이 수행될 수 있다. 상기 특허에는 보르데텔라에서의 CTB 발현이 기술되어 있다. 배양물 상청액 으로부터의 단백질을 SDS-PAGE에 의해 분석하거나, CTB가 분해되어 오량체에서 단량체 형태로 전환되기 전에 비 등시켰다. 상기 단백질을 나일론 막 상에 놓고, 염소 항콜레라게노이드 IgG 항체(항CTB, List Biologicals #GAC-01C)로 프로빙하였다. dig 화학 발광(Boehringer Mannheim)을 사용하여 알칼리성 포스파타제 컨주게이트된 당나귀 항염소 IgG로 검출을 수행하였다. 비교를 위해 CTA 및 CTB, 둘 모두를 함유하는 콜레라 독소 표준 (Sigma)을 사용하였다.
- [0183] 상업적으로 이용가능한 항체를 사용하여, 예를 들어, 본원 실시예에 기술되어 있는 바와 같이, PTX의 웨스턴 블롯 분석을 수행할 수 있다. 모노클로날 항체는 어빔(Abeam: 미국 매사추세츠주 캄브릿지 소재)으로부터 이용가능하다.
- [0184] 파상풍 독소 C 단편은 웨스턴 블롯 분석에 의해, 또는 미국 특허 번호 제5,443,966호(발명의 명칭: "Expression of tetanus toxin fragment C") (이 출원은 본원에 참고로 포함된다)에 기술된 바와 같이, ELISA에 의해 평가될 수 있다. 항체는 여러 상업적 공급업체, 예를 들어, 어빔(미국 매사추세츠주 캄브릿지 소재)으로부터 이용가능하다.
- [0185] TcdB 활성은 당업계에 기술된 바와 같이, 웨스턴 블롯 또는 다른 검출 분석에 의해 평가될 수 있다. 효소적 활성은 당업계에 기술된, 예를 들어, 미국 특허 번호 제7,226,597호(이 출원은 그 전문이 본원에 참고로 포함된다)에 기술된 글루코실히드롤라제/글루코실화 검정 방법을 사용하여 검정될 수 있다. 구체적으로, 글루코실화 반응은 50 mM n-2히드록시에틸피페라진-n'-2-에탄 술폰산, 100 mM KCl, 1 mM MnCl₂, 1 mM MgCl₂, 100 μg/ml BSA, 0.2 mM GDP, 40 μM [¹⁴C]UDP-글루코스 (303 Ci/mol; ICN Pharmaceuticals), 100 μM UDP-글루코스 및 3 pmol의 TcdB 또는 10 pmol의 각각의 융합 단백질을 함유하는 반응 믹스 중에서 수행될 수 있다. 37℃에서 밤새도록 인큐베이션함으로써 검정을 수행하고, AG1-X2 음이온 교환 수지를 사용하여 절단된 글루코스를 분리하고, 액체 섬광 계수기를 사용하여 계수한다.
- [0186] P. 아에루기노사 외독소 A 활성은 면역학적 방법, 예를 들어, 웨스턴 블롯 분석을 사용하여 평가될 수 있다. ETA가 ADP 리보실화 독소이기 때문에, 미국 특허 번호 제4,892,827호(이 출원은 본원에 참고로 포함된다)에 기술된 바와 같이, ADP 리보실화 활성에 대해 검정할 수 있다. 구체적으로, 신장 인자 2(EF-2)가 풍부한 토끼 망상적혈구 시료 또는 밀 배아 추출물을 EF-2의 공급원으로 사용한다. 검정(총 부피 500 μl)은 약 10 pmole의 EF-2, 37 pmole의 ¹⁴C-NAD(0.06 μCi), 0.25~1.25 μg의 ETA 및 완충액(40 mM DTT, 1 mM EDTA, 및 50 mM 트리스, pH 8.1)을 함유한다. 30분 동안 EF-2로 전달된 NAD(pmole)로서 활성을 측정한다. 공지된 농도의 PE 표준 곡선을 확립하고, 이를 사용하여 E. 콜라이로부터의 추출물 중 PE의 활성을 측정한다. 37℃에서 30분 동안 인큐베이션한 후, 0.5 ml 12% TCA를 각 검정 혼합물에 첨가한다. 이어서, 검정 혼합물을 15분 동안 얼음 배스에 놓은 후, 10분 동안 4℃, 3,000 x g로 원심분리한다. 펠릿을 1 ml 6% TCA로 세척하고, 상기와 같이 원심분리한다. 이어서, ADP 리보실화 활성의 지표로서 액체 섬광 계수기에서 ¹⁴C 방사능에 대해 펠릿을 측정한다.
- [0187] 그러므로, 활성 척도는 항체 또는 수용체 결합능, (칼럼 물질과 관련하여) 기질 결합능, 또는 효소 활성을 나타낼 수 있다.
- [0188] 실시양태에서, 활성은 검정된 총량과 비교하여 추출물 상청액 중 활성 재조합 독소 단백질(%)로 나타낸다. 이는 검정에 사용된 재조합 독소 단백질의 총량과 비교하여 검정에 의해 활성인 것으로 측정된 재조합 독소 단백질의 양에 기초한다. 다른 실시양태에서, 활성은 표준, 예를 들어, 네이티브 단백질과 비교하여 단백질의 활성 수준(%)으로 표시된다. 이는 표준 샘플 중 활성 단백질의 양에 대한 상청액 추출물 샘플 중 활성 재조합 독소 단백질의 양에 기초한 것이다(여기서, 각 샘플로부터 동량의 단백질이 검정에 사용된다).
- [0189] 실시양태에서, 독소 단백질 또는 서브유닛의 약 40%~약 100%가 활성인 것으로 측정된다. 실시양태에서, 재조합 독소 단백질 또는 단백질 서브유닛의 약 40%, 약 50%, 약 60%, 약 70%, 약 80%, 약 90%), 또는 약 100%가 활성인 것으로 측정된다. 실시양태에서, 재조합 독소 단백질 또는 서브유닛의 약 40%~약 50%, 약 50%~약 60%, 약 60%~약 70%, 약 70%~약 80%, 약 80%~약 90%, 약 90%~약 100%, 약 50%~약 100%, 약 60%~약 100%, 약 70%~약 100%, 약 80%~약 100%, 약 40%~약 90%, 약 40%~약 95%, 약 50%~약 90%, 약 50%~약 95%, 약 50%~약

100%, 약 60%~약 90%, 약 60%~약 95%, 약 60%~약 100%, 약 70%~약 90%, 약 70%~약 95%, 약 70%~약 100%, 또는 약 70%~약 100%가 활성인 것으로 측정된다.

[0190] 다른 실시양태에서, 재조합 독소 단백질 또는 단백질 서브유닛의 약 75%~약 100%가 활성인 것으로 측정된다. 실시양태에서, 재조합 독소 단백질 또는 서브유닛의 약 75%~약 80%, 약 75%~약 85%, 약 75%~약 90%, 약 75%~약 95%, 약 80%~약 85%, 약 80%~약 90%, 약 80%~약 95%, 약 80%~약 100%, 약 85%~약 90%, 약 85%~약 95%, 약 85%~약 100%, 약 90%~약 95%, 약 90%~약 100%, 또는 약 95%~약 100%가 활성인 것으로 측정된다.

[0191] 유도된 단백질의 정체를 확인하는 수단 또한 당업계에 공지되어 있다. 예를 들어, MALDI-TOF 질량 분광분석법, N 말단 서열 분석, 또는 펩티드 지도화를 이용한 펩티드 질량 핑거프린트로 단백질을 분석할 수 있다.

[0192] 본원에서는 본 발명의 바람직한 실시양태를 제시하고 기술하였지만, 그러한 실시양태들은 단지 예시를 위해 제공된 것이라는 것이 당업자에게는 명백할 것이다. 본 발명으로부터 벗어나지 않고 당업자들은 다수의 변형, 변화 및 치환을 실시할 수 있다. 본원에 기술된 본 발명의 실시양태들의 각종 대안을 본 발명을 실시하는 데 사용할 수 있음을 이해하여야 한다. 하기 특허청구범위가 본 발명의 범주를 규정하며, 이들 특허청구범위 및 그의 등가물 내의 방법 및 구성을 본원에 의해 포괄하고자 한다.

[0193] 실시예

[0194] 실시예 1: 재조합 CRM197 단백질의 고처리량 발현

[0195] CRM197 발현 균주를 구성하고, 균주에서 생산되는 가용성 CRM197 단백질의 생산량을 모세관 겔 전기영동(SDS-CGE)을 사용하여 분석하였다. 결과 데이터에 기초하여, 대규모 발현에 사용할 특정 균주를 선별하였다.

[0196] CRM197 발현 균주의 구성 및 성장

[0197] CRM197 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 CRM197 코딩 서열을 구성하였다. 서열 번호: 1은 최적화된 발현된 합성 CRM197 유전자에 의해 코딩된 아미노산 서열을 나타내고, 서열 번호: 2는 최적화된 발현된 합성 CRM197 유전자의 DNA 서열을 나타내는 것이다.

[0198] 하기 표 8에 기재된 바와 같은 10개의 P. 플루오레센스 분비 리더에 융합된, 최적화된 CRM197 서열을 보유하는 플라스미드를 구성하였다. 적절히 폴딩된 형태 및 활성 형태로 회수될 수 있도록 주변세포질로 단백질을 표적화시키는 P. 플루오레센스 분비 리더의 것과 함께 CRM197 코딩 서열을 프레임내에서 융합하였다.

표 8

표 8. CRM197 발현 스크린에 사용된 분비 리더

	분비 리더
1	DsbA
2	Azu
3	Ibp-S31A
4	Tpr
5	CupB2
6	CupA2
7	NikA
8	Pbp A20V
9	DsbC
10	TolB

[0199]

[0200] 재조합 CRM197 코딩 서열에 융합된 10개의 분비 리더를 함유하는 구성체를 P. 플루오레센스 숙주에서 테스트하였다. 각 발현 플라스미드를 사용하여 하기 표 9에 열거된 4가지의 숙주를 테스트하였다. 숙주 세포를 제시된 플라스미드로 전기천공시키고, 미량의 미네랄 및 5% 글리세롤을 포함하는 HTP 배양 배지 중에 재현탁시킨 후, 400 μ l M9 염 1% 글루코스 배지 및 미량 성분을 함유하는 96 웰 딥 웰 플레이트로 옮겼다. 96 웰 플레이트를 48 시간 동안 진탕하면서 30°C에서 인큐베이션하였다. 각 40개의 시드 배양물 10 μ l씩을 삼중 96 웰 딥 웰 플레이트로 옮겼으며(각 웰은 미량 원소 및 5% 글리세롤로 보충된 HTP 배지 500 μ l를 함유), 24시간 동안 이전과 같이 인큐베이션하였다.

표 9

표 9. CRM197 발현 스크린에 사용된 숙주 균주

숙주 균주	유전자형	유형
1	lon, la, aprA	PD
2	hsUV prc1 degP1 degP2 aprA 결실; DegP2 S219A 를 과발현	PD+FMO
3	dsbABCD	FMO
4	grpE, dnaKJ	FMO

PD = 프로테아제 결실 (열거된 프로테아제는 결실되어 있다);

FMO = 폴딩 조절 인자 과발현인자 (열거된 폴딩 조절 인자는 과발현되어 있다)

[0201]

[0202]

이소프로필-β-D-1-티오갈락토피라노시드(IPTG)를 최종 농도가 0.3 mM이 되도록 각 웰에 첨가하여 표적 단백질의 발현을 유도하였다. 만넛톨(시그마, M1902)을 최종 농도가 1%가 되도록 각 웰에 첨가하여 폴딩 조절 인자 과발현 균주에서 폴딩 조절 인자의 발현을 유도하였으며, 온도를 25℃로 강하시켰다. 유도 후 24시간 경과하였을 때, 400 μl 부피의 PBS를 사용하여 세포를 OD600 = 15로 정규화하였다. 초음파 처리 및 원심분리함으로써 가용성 및 불용성 분획을 생성하는 추후 프로세싱을 위해 샘플을 동결시켰다.

[0203]

샘플 제조 및 SDS-CGE 분석

[0204]

정규화된 배양물을 초음파처리한 후, 원심분리하여 가용성 및 불용성 세포 분획을 제조하였다. 냉동된 정규화된 배양물 브로쓰(400 μl)를 해동시키고 3.5분 동안 초음파처리하였다. 용해물을 20,800 x g에서 20분(4℃)동안 원심분리하고, 수동 또는 자동화된 액상 취급기(가용성 분획)를 이용하여 상청액을 제거하였다. 펠릿(불용성 분획)을 냉동시킨 후, 20,080 x g에서 20분 동안 4℃에서 재원심분리를 위해 해동시켜 잔류 상청액을 제거하였다. 이어서, 펠릿을 400 μl의 1X 인산염 완충처리된 염수(PBS)(pH 7.4)에 재현탁시켰다. SDS-CGE 분석을 위해 가용성 및 불용성 샘플을 1X 인산염 완충 염수(PBS)(pH 7.4)에서 추가로 희석시켰다. 디티오프레이톨(DTT)의 존재하에 SDS 모세관 겔 전기영동(CGE)(Caliper Life Sciences, Protein Express LabChip Kit, Part 760301)을 위한 가용성 및 불용성 샘플을 제조하였다.

[0205]

각 균주로부터의 가용성 분획의 환원성 SDS-CGE 분석 결과를 보여주는 대표적인 겔 유사 이미지는 도 1에 도시되어 있다. 하기 표 10은 구성된 CRM197 발현 균주 각각에 대한 3 복제물의 가용성 CRM197의 평균 수율 및 표준편차를 제시한다. 각 균주에 대해 스크리닝된 숙주 균주 및 분비 리더 또한 제시되어 있다.

[0206]

분비 리더 및 숙주 균주, 둘 모두 CRM197 발현에 유의적인 영향을 미치는 것으로 나타났다. 발현 범위는 0.5 mL 규모에서 검출불가능한 수율~1.2 g/L 초과이며, 최고 발현 수준은 숙주 균주 2 배경에서 관찰되었다. PS538-776에서 관찰되는 수율은 1,263 mg/L이며, PS538-772에서 관찰되는 수율은 1,241 mg/L이며, 둘모두 평균 수율이 340 mg/L를 초과하였다. 사용된 리더에 따라서 동일한 숙주 균주에서도 높은 수율과 낮은 수율이 모두 관찰되었으며, 상이한 숙주 균주에서 동일한 리더 서열을 사용해도 높은 수율 및 낮은 수율이 모두 관찰되었다.

[0207]

대규모 발효에서의 평가를 위해 PS538-772, PS538-773, PS538-776, PS538-778, PS538-782를 선택하였다.

표 10

표 10. CRM197 발현 균주에 대한 CRM197 의 평균 수율

균주 번호	미국 특허 출원 일련 번호 제 61/325,235 호에 서의 상응 균주 번호	숙주 균주	리더	평균 수율 (mg/L)	Std Dev (3 복제물)
PS538-761	PS538-731	1	DsbA	205	162
PS538-762	PS538-732	1	Azu	427	186
PS538-763	PS538-733	1	Ibp-S31A	361	183
PS538-764	PS538-734	1	Tpr	298	106
PS538-765	PS538-735	1	CupB2	105	109
PS538-766	PS538-736	1	CupA2	175	99
PS538-767	PS538-737	1	NikA	314	85
PS538-768	PS538-738	1	Pbp A20V	291	204
PS538-769	PS538-739	1	DsbC	148	91
PS538-770	PS538-740	1	TolB	213	36
PS538-771	PS538-741	2	DsbA	407	218
PS538-772	PS538-742	2	Azu	1241	372
PS538-773	PS538-743	2	Ibp-S31A	1107	219
PS538-774	PS538-744	2	Tpr	280	285
PS538-775	PS538-745	2	CupB2	192	219
PS538-776	PS538-746	2	CupA2	1263	474
PS538-777	PS538-747	2	NikA	699	259
PS538-778	PS538-748	2	Pbp A20V	914	416
PS538-779	PS538-749	2	DsbC	567	141
PS538-780	PS538-750	2	TolB	382	217
PS538-781	PS538-751	3	DsbA	591	230
PS538-782	PS538-752	3	Azu	1094	543
PS538-783	PS538-753	3	Ibp-S31A	323	143
PS538-784	PS538-754	3	Tpr	419	70
PS538-785	PS538-755	3	CupB2	75	74
PS538-786	PS538-756	3	CupA2	309	214
PS538-787	PS538-757	3	NikA	52	73
PS538-788	PS538-758	3	Pbp A20V	356	295
PS538-789	PS538-759	3	DsbC	319	117
PS538-790	PS538-760	3	TolB	69	88
PS538-791	PS538-761	4	DsbA	270	106
PS538-792	PS538-762	4	Azu	0	14
PS538-793	PS538-763	4	Ibp-S31A	0	6
PS538-794	PS538-764	4	Tpr	0	0
PS538-795	PS538-765	4	CupB2	18	39
PS538-796	PS538-766	4	CupA2	118	134
PS538-797	PS538-767	4	NikA	0	9
PS538-798	PS538-768	4	Pbp A20V	0	0
PS538-799	PS538-769	4	DsbC	0	0
PS538-800	PS538-770	4	TolB	0	0

[0208]

[0209] 실시예 2: 제조합 CRM197 단백질의 대규모 발현

[0210] 제조합 CRM197 단백질은 2 ℓ 발효기에서 슈도모나스 플루오레센스 균주 PS538-772, PS538-776, 및 PS538-782에서 제조하였다. 본원에 기술된 바와 같고, 또한, 예를 들어, 문헌 [Riesenber, D., et al., 1991]에 기술된 바와 같은 미네랄 염 배지를 함유하는 2 ℓ 발효기에서 배양물을 성장시키고, 암모니아를 첨가하여 pH 6.5 및 32℃에서 유지시켰다. 발효기에서 살포 공기 및 산소의 흐름 및 교반을 증가시켜 용존 산소를 과량으로 유지시켰다. 발효 과정 중에 글리세롤을 배양물에 전달하여 과량 수준으로 유지시켰다. 유도를 위한 표적 배양물 세포 밀도(575 nm에서의 광학 밀도(A_{575}))에 도달할 때까지 이들 조건을 유지시켰고, 이때 IPTG를 첨가하여 CRM197 이 제조될 수 있도록 개시시켰다. 유도 시 세포 밀도는 A_{575} 40~200 흡광 단위(AU)로 다양할 수 있다. IPTG 농도는 0.02~0.4 mM 범위로, pH는 6~7.5 범위로, 온도는 20~35℃ 범위로 다양할 수 있다. 16-24시간 후에, 각 생물반응기로부터의 배양물을 원심분리로 수거하고, -80℃에서 세포 펠릿을 냉동시켰다. SDS-CGE에 의해 생성물

제조에 대해 샘플을 분석하였다.

[0211] 다중 발효 조건을 평가함으로써 SDS-CGE에 의해 측정된 바, 1~2 g/L인 상위 CRM197 발현을 수득하였다(도 18 및 19 참조). 유도된 단백질의 정체는 디프테리아 독소 특이 항체를 사용하는 웨스턴 블롯 분석에 의해 확인하였다(도 20).

[0212] 실시예 3: 재조합 콜레라 독소 B 단백질의 고처리량 발현

[0213] 콜레라 독소 B 발현 균주의 구성 및 성장

[0214] 콜레라 독소 B 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 콜레라 독소 B 코딩 서열을 구성하였다. 서열 번호: 22는 발현된 합성 콜레라 독소 B 유전자에 의해 코딩된 아미노산 서열을 나타내고, 서열 번호: 23은 최적화된 발현된 합성 콜레라 독소 B 유전자의 DNA 서열을 나타내는 것이다.

[0215] (표 8에 기재된) CRM197과 함께 사용된 것과 동일한 10개의 P. 플루오레센스 분비 리더 코딩 서열에 융합된, 최적화된 콜레라 독소 B 서열을 보유하는 플라스미드를 구성하였다. 적절히 폴딩된 형태 및 활성 형태로 회수될 수 있도록 주변세포질로 단백질을 표적화시키기 위해 분비 리더를 포함하였다.

[0216] 재조합 콜레라 독소 B 단백질에 융합된 10개의 분비 리더를 발현하는 구성체를 P. 플루오레센스 숙주에서 테스트하였다. 각 발현 플라스미드를 사용하여 표 9에 열거된 4가지의 숙주를 테스트하였다. 숙주 세포를 제시된 플라스미드로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

[0217] 각 균주로부터의 가용성 분획의 환원성 SDS-CGE 분석 결과를 보여주는 대표적인 겔 유사 이미지는 도 2에 도시되어 있다. 하기 표 11은 구성된 콜레라 독소 B-발현 균주 각각에 대한 3 복제물의 가용성 콜레라 독소 B의 평균 수율 및 표준 편차를 제시한다.

[0218] 분비 리더 및 숙주 균주, 둘 모두 콜레라 독소 B 발현에 유의적인 영향을 미치는 것으로 나타났다. 발현 범위는 0.5 mL 규모에서 검출불가능한 수율~0.2 g/L 초과이며, 최고 발현 수준은 *hslUV prc1 degP1 degP2 aprA* 결실/DegP2 S219A 과발현 숙주 배경에서 관찰되었다. 리더 6(CupA2) 및 8(PbpA20V)에 융합된 콜레라 독소 B의 발현은 4가지의 균주 모두에서 일관되게 높게 나타났다.

표 11

표 11. 콜레라 독소 B 발현 요약

균주 번호	숙주 균주	플라스미드	리더	평균 수율 (mg/L)	Std Dev (3 복제물)
PS538-081	1	p538-021	DsbA	25	8
PS538-082	1	p538-022	Azu	1	8
PS538-083	1	p538-023	Ibp-S31A	0	5
PS538-084	1	p538-024	Tpr	35	14
PS538-085	1	p538-025	CupB2	10	9
PS538-086	1	p538-026	CupA2	138	18
PS538-087	1	p538-027	NikA	0	5
PS538-088	1	p538-028	Pbp A20V	213	23
PS538-089	1	p538-029	DsbC	0	6
PS538-090	1	p538-030	TolB	0	4
PS538-091	2	p538-021	DsbA	133	62
PS538-092	2	p538-022	Azu	83	56
PS538-093	2	p538-023	Ibp-S31A	50	44
PS538-094	2	p538-024	Tpr	61	55
PS538-095	2	p538-025	CupB2	62	19
PS538-096	2	p538-026	CupA2	147	57
PS538-097	2	p538-027	NikA	31	28
PS538-098	2	p538-028	Pbp A20V	223	78
PS538-099	2	p538-029	DsbC	41	24
PS538-100	2	p538-030	TolB	6	5
PS538-101	3	p538-021	DsbA	1	7
PS538-102	3	p538-022	Azu	1	2
PS538-103	3	p538-023	Ibp-S31A	19	17
PS538-104	3	p538-024	Tpr	28	36
PS538-105	3	p538-025	CupB2	5	9
PS538-106	3	p538-026	CupA2	40	12
PS538-107	3	p538-027	NikA	5	10
PS538-108	3	p538-028	Pbp A20V	45	19
PS538-109	3	p538-029	DsbC	0	6
PS538-110	3	p538-030	TolB	0	3
PS538-111	4	p538-021	DsbA	0	5
PS538-112	4	p538-022	Azu	0	3
PS538-113	4	p538-023	Ibp-S31A	0	2
PS538-114	4	p538-024	Tpr	13	3
PS538-115	4	p538-025	CupB2	2	4
PS538-116	4	p538-026	CupA2	15	16
PS538-117	4	p538-027	NikA	0	2

[0219]

PS538-118	4	p538-028	Pbp A20V	35	15
PS538-119	4	p538-029	DsbC	0	2
PS538-120	4	p538-030	TolB	0	2

[0220]

[0221]

실시예 4: 재조합 콜레라 독소 B 단백질의 대규모 발현

[0222]

재조합 콜레라 독소 B 단백질은 슈도모나스 플루오레센스 페넥스 익스프레션 테크놀로지(Pfenex Expression Technology)™ 균주 PS538-088 및 PS538-091에서 제조하였다. 선택된 균주를 본원에 기술된 바와 같고, 또한, 예를 들어, 문헌 [Riesenberg, D., et al., 1991]에 기술된 바와 같은 미네랄 염 배지를 함유하는 2 ℓ 발효기에서 배양물을 성장시키고, 암모니아를 첨가하여 pH 6.5 및 32℃에서 유지시켰다. 발효기에서 살포 공기 및 산소의 흐름 및 교반을 증가시켜 용존 산소를 과량으로 유지시켰다. 발효 과정 중에 글리세롤을 배양물에 전달하

여 과량 수준으로 유지시켰다. 유도를 위한 표적 배양물 세포 밀도(575 nm에서의 광학 밀도(A_{575}))에 도달할 때까지 이들 조건을 유지시켰고, 이때 IPTG를 첨가하여 표적 단백질이 제조될 수 있도록 개시시켰다. IPTG를 첨가하여 CTB가 제조될 수 있도록 개시시켰다. 16-24시간 후에, 각 생물반응기로부터의 배양물을 원심분리로 수거하고, -80°C 에서 세포 펠릿을 냉동시켰다.

[0223] 다중 발효 조건을 평가함으로써 SDS-CGE에 의해 측정된 바, 0.6~1.0 g/L인 상위 CTB 발현을 수득하였다. pH 6.5-7.2 및 32°C 에서 0.2 mM IPTG를 첨가하였을 때 대략 80-160 OD로 고성능의 상위 발효 배양물이 유도되었다(도 14 및 하기 표 12 참조). 유도된 단백질의 정체는 MALDI-TOF 질량 분광분석법을 사용하는 펩티드 질량 핑거프린트에 의해 확인하였다.

표 12

표 12. 가용성 콜레라 독소 B 역가

균주	발효물	생성물	생성물 농도 (g/L)
PS538-088	U5	CTB	0.94 ± 0.03
PS538-088	U6	CTB	0.59 ± 0.01
PS538-091	U3	CTB	0.81 ± 0.09

[0224]

[0225] 실시예 5: 재조합 백일해 독소 단백질의 고처리량 발현

[0226] 백일해 독소이드 S1 E129A R9K 발현 균주의 구성 및 성장

[0227] 재조합 백일해 독소 발현을 위해, S1 돌연변이 E129A 및 R9K를 포함하고, 서브유닛 S1, S2, S3, S4 및 S5를 코딩하는 백일해 독소이드 오페론의 서열을 사용하였다. 도 3은 상기 오페론의 지도이다. 도 4는 번역과 함께 상기 오페론의 DNA 서열을 제시한다(서열 번호: 24). 도 5는 S1, S2, S3, S4 및 S5 개개의 아미노산 서열을 제시한다.

[0228] 하기 표 13에 제시된 바와 같이, 8개의 P. 플루오레센스 숙주에서 구성체를 발현시켰다. p538-081을 사용하여 숙주 세포를 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

표 13

표 13. 백일해 독소 S1 E129A R9K 발현 균주

균주 번호	숙주	유전자형	플라스미드	유형
PS538-321	1	lon, la, aprA	p538-081	PD
PS538-322	2	hslUV, prc1, degP1, degP2 and aprA	p538-081	PD
PS538-323	3	dsbABCD	p538-081	FMO
PS538-324	4	grpE, dnaKJ	p538-081	FMO
PS538-325	5	htpX	p538-081	PD
PS538-326	6	RXF01590	p538-081	PD
PS538-327	7	lon, la, aprA 결실; grpE 및 dnaKJ를 과발현	p538-081	PD + FMO
PS538-328	8	ppiB (RFX05345)	p538-081	FMO

PD = 프로테아제 결실(열거된 프로테아제는 결실되어 있다); FMO = 폴딩 조절 인자 과발현인자(열거된 폴딩 조절 인자는 과발현되어 있다)

[0229]

[0230] 발현된 백일해 독소의 웨스턴 블롯 분석

[0231] 상기 기술된 8개의 배양물로부터의 가용성 분획을 웨스턴 블롯에 의해 분석하여 백일해 독소이드 발현에 대해 평가하였다. 20 μl 의 가용성 분획(2X 희석, 환원성 및 비환원성)을 1X 바이오 래드(Bio Rad) MES 전개용 완충액 중 바이오-래드 12% 비스-트리스 겔(Bio-Rad 12% Bis-Tris Gel) 상에서 전개시켰다. 환원성 웨스턴 분석을 위해, 1X XT 환원제를 첨가하였다. 20% 메탄올을 포함하는 1X NuPAGE 전달용 완충액을 사용하여 단백질을 60분 동안 100 V에서 SDS-PAGE로부터 0.2 μm 니트로셀룰로스 막(Bio Rad, 162 0232)으로 옮겨 놓았다. 1시간 동안

실온에서 막을 PBS 중 블로커™ 1% 카제인(Blocker™ 1% Casein)(Pierce, 37528)에서 차단시켰다. 검출을 위해 희석제를 따라 내고, 각각이 보르테텔라 백일해 독소 S4 및 S1(Abeam, 카탈로그 번호 ab37686 및 #37547)에 대한 것인 모노클로날 항체의 1:1,000으로 희석된 희석액 조합물을 함유하는 추가의 희석제를 첨가하였다. 블롯을 4℃에서 밤새도록 진동시키면서 인큐베이션하였다. 블롯을 각각 5분 동안 PBS-트윈(PBS-Tween)을 사용하여 3회에 걸쳐 세척한 후, 염소에서 유래된 항마우스 IgG-퍼옥시다제(Sigma, 카탈로그 번호 A4416)의 1:5,000으로 희석된, 추가의 희석제를 함유하는 희석액 중에서 1시간 동안 실온에서 인큐베이션하였다. 블롯을 각각 5분 동안 PBS-트윈(Sigma, P3563)을 사용하여 3회에 걸쳐 세척한 후, 이뮤노뷰어 메탈 인헨스드 DAB 기질(Pierce, 34065)을 사용하여 발색시켰다. 환원성 조건 및 비환원성 조건, 둘 모두하에서 항S1 및 항S4 항체에 의해 다중 서브유닛을 검출하였다(도 6). 발현된 독소이드의 환원된 및 비환원된 샘플의 밴딩 패턴에 관한 관찰 결과는 문헌 [Sekura, et al. (J. Biological Chemistry 258: 14647, 1983)]에 의해 보고된 바와 같이, 균주 165로부터 정제된 백일해 독소에 대해 관찰된 결과와 일치하였다.

[0232] 실시예 6: 재조합 백일해 독소 단백질의 대규모 발현

[0233] 재조합 백일해 독소 단백질은 슈도모나스 플루오레센스 페넥스 익스프레스 테크놀로지™ 균주 PS538-321, PS538-324, PS538-325, PS538-326, 및 PS538-328에서 제조하였다. 선택된 균주를 CTB 대규모 발현에 대해 상기 기술된 바와 같이, 2 ℓ 발효기에서 성장시키고, IPTG로 유도시키고, 분석을 위해 샘플을 제조하였다. 샘플을 SDS-CGE에 의해 분석하고, 생성물 형성 및 그의 활성에 대해서는 웨스턴 블롯에 의해 분석하였다.

[0234] 실시예 7: 재조합 야생형 백일해 독소이드의 고처리량 발현

[0235] 백일해 독소이드 발현 균주의 구성 및 성장

[0236] 재조합 백일해 독소이드 발현을 위해, S1을 포함하고, 서브유닛 S1, S2, S3, S4 및 S5를 코딩하는 야생형 백일해 독소 오페론의 서열을 사용하였다. 도 13은 번역과 함께 상기 야생형 오페론의 DNA 서열을 제시한다(서열 번호: 35).

[0237] 하기 표 14에 제시된 바와 같이 P. 플루오레센스 숙주에서 구성체를 발현시켰다. 과발현 플라스미드가 없는, 열거된 각각의 균주를 a) 기술된 바와 같이(과발현 플라스미드 없는 상태로); b) GrpE DnaKJ 과발현 플라스미드를 포함한 상태에서, 및 c) DsbABCD 과발현 플라스미드를 포함한 상태에서 테스트하였다. PTX WT 발현 플라스미드를 사용하여 숙주 세포를 전기천공시키고, PTX S1 R9K E129A 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. 또한 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

표 14

표 14. 백일해 특소이드 야생형 발현 균주

숙주*	유전자형	유형
9	hslUV prc2	PD
10	hslUV degP1	PD
11	la	PD
12	hslUV prc1 prc2	PD
13	lon la prc1 prc2	PD
14	RXF01590	PD
1	lon la aprA	PD
7	lon la prc1 degP2 aprA; GrpE DnaKJ 를 과발현	PD + FMO
15	RXF02151 RXF00428	PD
16	lon la degP2	PD
17	DsbAB 를 과발현	FMO
18	DsbCD 를 과발현	FMO
19	prc1 degP2; degP2 S219A 를 과발현	PD + FMO
20	prc1 degP2 clp1 aprA; degP2 S219A 를 과발현	PD + FMO
21	prc1 degP2 lon aprA; degP2 S219A 를 과발현	PD + FMO
22	prc1 degP2 degP1 aprA; degP2 S219A 를 과발현	PD + FMO
23	lon prc1 degP2 degP1 aprA; degP2 S219A 를 과발현	PD + FMO
2	hslUV prc1 degP2 degP1 aprA; degP2 S219A 를 과발현	PD + FMO
25	lon la degP2 prc1 aprA; degP2 S219A 를 과발현	PD + FMO
26	degP2; SecB 를 과발현	PD + FMO
27	degP2; FkpP 를 과발현	PD + FMO
28	degP2; GroELES 를 과발현	PD + FMO
29	lon la aprA; SecB 를 과발현	PD + FMO
30	lon la aprA; FkpP 를 과발현	PD + FMO
31	lon la aprA; GroELES 를 과발현	PD + FMO
32	dsbC	PD
33	dsbC; DsbAB 를 과발현	PD + FMO
3	DsbABCD 를 과발현	FMO
35	lexA aprA	PD
36	SlyD 를 과발현	FMO
37	lon hslUV	PD
38	Wt	---
39	aprA	PD
4	GrpE DnaKJ 를 과발현	FMO
5	htpX	PD
40	lon	PD
41	prc1	PD
42	hslUV	PD
43	degP2	PD
44	degP1	PD

[0238]

45	prc2	PD
46	RXF6451	PD
6	RXF1590	PD
48	RXF4692	PD
49	hslUV mic	PD
50	RXF2161	PD
51	RXF00133	PD
52	RXF2796	PD
53	RXF4968	PD
54	DsbC 를 과발현	FMO
55	DsbAC 를 과발현	FMO
56	LepB 를 과발현	FMO
57	SecB 를 과발현	FMO
58	ClpA 를 과발현	FMO
59	FliB2 를 과발현	FMO
60	DnaK 유사를 과발현	FMO
61	FkbP 를 과발현	FMO
62	PpiA 를 과발현	FMO
8	PpiB 를 과발현	FMO
63	HscA 를 과발현	FMO
64	GshA 를 과발현	FMO
65	Gor 을 과발현	FMO
66	TrxC 를 과발현	FMO
67	DsbG 를 과발현	FMO
68	Ppi 를 과발현	FMO
69	GroELES 를 과발현	FMO
70	prc1 aprA; GrpE DnaKJ 를 과발현	PD + FMO
71	과다분비	
72	DsbD 를 과발현	FMO
73	과다분비	
74	과다분비	
75	prc1 prc2	PD
76	hslUV clpA	PD

* 과발현 플라스미드가 없는, 열거된 각각의 균주를 a) 기술된 바와 같이(과발현 플라스미드 없는 상태로); b) GrpE DnaKJ 과발현 플라스미드를 포함한 상태에서, 및 c) DsbABCD 과발현 플라스미드를 포함한 상태에서 테스트하였다. PD = 프로테아제 결실(열거된 프로테아제는 결실되어 있다); FMO = 폴딩 조절 인자 과발현인자(열거된 폴딩 조절 인자는 과발현되어 있다).

[0239]

[0240]

과다 소포 형성 균주로도 알려져 있는 과다분비 균주는 예를 들어, W02010/008764(발명의 명칭: "W02010/008764, "Pseudomonas Fluorescens Strains for Production of Extracellular Recombinant Protein")(이 출원은 그 전문이 본원에 참고로 포함된다)에 기술되어 있다.

[0241]

실시예 8: 재조합 파상풍 독소 단편 C 단백질의 고처리량 발현

[0242]

CRM197 발현 균주의 구성 및 성장

[0243]

파상풍 독소 C 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 파상풍 독소 C 코딩 서열을 구성하였다. 서열 번호: 30은 발현된 합성 파상풍 독소 C 유전자에 의해 코딩된 아미노산 서열을 나타내고, 서열 번호: 31은 최적화된 발현된 합성 파상풍 독소 C 유전자의 DNA 서열을 나타내는 것이다.

[0244]

(표 8에 기재된) CRM197과 함께 사용된 것과 동일한 10개의 P. 플루오레센스 분비 리더 코딩 서열에 융합된, 최적화된 파상풍 독소 C 서열을 보유하는 플라스미드를 구성하였다. 적절히 폴딩된 형태 및 활성 형태로 회수될 수 있도록 주변세포질로 단백질을 표적화시키기 위해 분비 리더를 포함하였다.

[0245]

재조합 파상풍 독소 C 단백질에 융합된 10개의 분비 리더를 발현하는 구성체를 P. 플루오레센스 숙주에서 테스트하였다. 각 리더를 사용하여 표 9에 열거된 4가지의 숙주를 테스트하였다. 숙주 세포를 제시된 플라스미드로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

[0246] 각 균주로부터의 가용성 분획의 환원성 SDS-CGE 분석 결과를 보여주는 대표적인 겔 유사 이미지는 도 7에 도시되어 있다. 하기 표 15는 구성된 과상풍 독소 C-발현 균주 각각에 대한 3 복제물의 가용성 과상풍 독소 C의 평균 수율 및 표준 편차를 제시한다. 과상풍 독소 C 단편은 테스트된 모든 균주에서 잘 발현된 것으로 나타났으며, 최고수율은 *hslUV prc1 degP1 degP2 aprA* 결실/DegP2 S219A 과발현 발현 숙주에서 최대 600 mg/L 이하 범위였다. 추가의 평가를 위해 균주 PS538-529, PS538-538, PS538-544, PS538-546, PS538-547, PS538-548, PS538-558, PS538-565 및 PS538-568을 선택하였다.

표 15

표 15. 과상풍 독소 C 발현 요약.

균주 번호	숙주	플라스미드	리더	평균 수율 (mg/L)	Std Dev (3 복제물)
PS538-529	1	p538-132	DsbA	261	75
PS538-530	1	p538-133	Azu	200	82
PS538-531	1	p538-134	Ibp-S31A	165	64
PS538-532	1	p538-135	Tpr	207	107
PS538-533	1	p538-136	CupB2	205	128
PS538-534	1	p538-137	CupA2	200	117
PS538-535	1	p538-138	NikA	174	96
PS538-536	1	p538-139	Pbp A20V	311	156
PS538-537	1	p538-140	DsbC	188	97
PS538-538	1	p538-141	TolB	129	63
PS538-539	2	p538-132	DsbA	486	89
PS538-540	2	p538-133	Azu	495	93
PS538-541	2	p538-134	Ibp-S31A	568	68
PS538-542	2	p538-135	Tpr	589	364
PS538-543	2	p538-136	CupB2	534	318
PS538-544	2	p538-137	CupA2	504	134
PS538-545	2	p538-138	NikA	444	145
PS538-546	2	p538-139	Pbp A20V	637	280
PS538-547	2	p538-140	DsbC	574	68
PS538-548	2	p538-141	TolB	438	61
PS538-549	3	p538-132	DsbA	174	37
PS538-550	3	p538-133	Azu	180	58
PS538-551	3	p538-134	Ibp-S31A	88	58
PS538-552	3	p538-135	Tpr	247	134
PS538-553	3	p538-136	CupB2	199	39
PS538-554	3	p538-137	CupA2	165	69
PS538-555	3	p538-138	NikA	97	90
PS538-556	3	p538-139	Pbp A20V	297	112
PS538-557	3	p538-140	DsbC	151	52
PS538-558	3	p538-141	TolB	35	13
PS538-559	4	p538-132	DsbA	39	39
PS538-560	4	p538-133	Azu	40	43
PS538-561	4	p538-134	Ibp-S31A	36	40
PS538-562	4	p538-135	Tpr	35	39
PS538-563	4	p538-136	CupB2	54	26
PS538-564	4	p538-137	CupA2	42	36
PS538-565	4	p538-138	NikA	44	37

[0247]

PS538-566	4	p538-139	Pbp A20V	37	40
PS538-567	4	p538-140	DsbC	39	43
PS538-568	4	p538-141	TolB	45	38

[0248]

[0249] 실시예 9: 재조합 파상풍 독소 단편 C 단백질의 대규모 발현

[0250] 재조합 파상풍 독소 C 단백질은 슈도모나스 플루오레센스 페넥스 익스프레스 테크놀로지™ 균주 PS538-529, PS538-538, PS538-544, PS538-546, PS538-547, PS538-548, PS538-558, PS538-565 및 PS538-568에서 제조하였다. 선택된 균주를 CRM197에 대해 상기 기술된 바와 같이, 미네랄 염 배지를 함유하는 2 ℓ 발효기에서 배양물을 성장시켰다.

[0251] 다중 발효 조건을 평가함으로써 SDS-CGE에 의해 측정된 바, 6~10 g/L인, 균주 PS538-529, PS538-546, 및 PS538-547로부터 상위 가용성 TTC 발현을 수득하였다(도 11A 및 하기 표 16 참조). pH 7.2 및 32℃에서 0.2 mM IPTG를 첨가하였을 때 대략 160 OD로 고성능의 상위 발효 배양물이 유도되었다. 유도된 단백질의 정체는 MALDI-TOF 질량 분광분석법을 사용하는 펩티드 질량 핑거프린트 및 웨스턴 블롯에 의해 확인하였다. PS538-529, PS538-546 및 PS538-547의 분비 리더(각각 (DsbA, Pbp A20V 및 DsbC)가 이러한 발현 조건하에서는 100% 발현된 단백질로부터 프로세싱되지 않았다는 것이 질량 분광분석법 및 웨스턴 블롯 분석을 통해 밝혀졌다. 그러나, TolB 리더는 선택된 단백질로부터 정확하게 절단된 것으로 확인되었다(데이터는 나타내지 않음). 분비 리더가 높은 신뢰도로 절단되고, 낮은 수준으로 분해되도록 하면서 TTC를 생산할 수 있는 균주를 확인하기 위해, 상기 개략적으로 설명된 조건을 사용하여 2 ℓ 발효 규모로 TolB-TTC 발현 균주 PS538-538, PS538-548, PS538-558 및 PS538-568을 스크리닝하였다. 웨스턴 블롯 분석에 의해 균주 PS538-538, PS538-548 및 PS538-558은 유사한 정질 및 수율로 물질을 생산한다는 것이 관찰되었다(도 11B).

표 16

표 16. 가용성 파상풍 독소 C (TTC) 역가

균주	발효물	생성물	생성물 농도 (g/L)
PS538-529	U1 - 도 11A	TTC	5.7 ± 1.3
PS538-546	U7 - 도 11A	TTC	9.5 ± 1.1
PS538-547	U5 - 도 11A	TTC	6.2 ± 1.9
PS538-538	U1 - 도 11B	TTC	2.5 ± 0.09
PS538-538	U2 - 도 11B	TTC	1.8 ± 0.2
PS538-548	U3 - 도 11B	TTC	5.3 ± 0.6
PS538-548	U4 - 도 11B	TTC	4.5 ± 0.2
PS538-558	U5 - 도 11B	TTC	1.1 ± 0.8
PS538-558	U6 - 도 11B	TTC	1.9 ± 0.1
PS538-568	U7 - 도 11B	TTC	0.2 ± 0.01
PS538-568	U8 - 도 11B	TTC	0.2 ± 0.01

[0252]

[0253] 실시예 10: 재조합 C. 디피실레 B 단백질의 고처리량 발현

[0254] TcdB 발현 균주의 구성 및 성장

[0255] TcdB 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 TcdB 코딩 서열을 구성하였다. 서열 번호: 32는 발현된 합성 TcdB 유전자에 의해 코딩된 아미노산 서열을 나타내고, 서열 번호: 33은 최적화된 발현된 합성 TcdB 유전자의 DNA 서열을 나타내는 것이다.

[0256] 하기 표 17에 열거된 유전자형을 가진 P. 플루오레센스 숙주에서 최적화된 TcdB 서열을 보유하는 플라스미드를 테스트하였다. 숙주 세포를 세포질 발현 플라스미드 p538-211로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

표 17

표 17. TcdB 속주 균주

속주 균주	유전자형	유형
37	hslUV lon	PD
38		WT
4	dnaKJ grpE	FMO
5	htpX	PD
40	lon	PD
41	prc1	PD
42	hslUV	PD
43	degP2	PD
44	degP1	PD
45	prc2	PD
47	RXF01590	PD
49	hslUV mic	PD
53	RXF04968	PD
55	dsbAC	FMO
61	fkbp	FMO
66	trxC	FMO
72	dsbD	FMO
76	hslUV clpA	PD
12	hslUV prc1 prc2	PD
1	lon la aprA	PD
16	lon la degP2	PD
2	hslUV prc1 degP1 degP2 aprA 결실; degP2 S219A 를 과발현	PD + FMO
3	dsbABCD	FMO
21	lon prc1 degP2 aprA 결실; degP2 S219A 를 과발현	PD + FMO

[0257]

[0258] 테스트된 24개의 각 균주로부터의 가용성 분획의 환원성 SDS-CGE 분석 결과를 보여주는 대표적인 겔 유사 이미지는 도 8에 도시되어 있다. 하기 표 18은 구성된 TcdB 발현 균주 각각에 대한 3 복제물의 가용성 TcdB의 평균 수율 및 표준 편차를 제시한다. 추가 평가를 위해 균주 PS538-654, PS538-659, PS538-669, PS538-671, 및 PS538-674를 선택하였다.

표 18

표 18. TcdB 발현 요약

균주 번호	숙주 균주	플라스미드	평균 수율 (mg/L)	Std Dev (3 복제물)
PS538-651	37	p538-211	103	7
PS538-652	38	p538-211	55	4
PS538-653	4	p538-211	57	1
PS538-654	5	p538-211	166	13
PS538-655	40	p538-211	88	3
PS538-656	41	p538-211	68	5
PS538-657	42	p538-211	90	14
PS538-658	43	p538-211	68	2
PS538-659	44	p538-211	109	8
PS538-660	45	p538-211	78	4
PS538-661	6	p538-211	98	15
PS538-662	49	p538-211	106	10
PS538-663	53	p538-211	91	6
PS538-664	55	p538-211	45	4
PS538-665	61	p538-211	63	6
PS538-666	66	p538-211	56	8
PS538-667	72	p538-211	70	8
PS538-668	76	p538-211	80	6
PS538-669	12	p538-211	117	39
PS538-670	1	p538-211	108	18
PS538-671	16	p538-211	247	65
PS538-672	2	p538-211	32	6
PS538-673	3	p538-211	52	2
PS538-674	21	p538-211	145	12

[0259]

[0260] 실시예 11: 재조합 C. 디피실레 독소 B 단백질의 대규모 발현

[0261] 재조합 C. 디피실레 독소 B 단백질은 슈도모나스 플루오레센스 페넥스 익스프레스 테크놀로지™ 균주 PS538-654, PS538-659, PS538-669, PS538-671, 및 PS538-674에서 제조하였다. 선택된 균주를 CTB 대규모 발현에 대해 상기 기술된 바와 같이, 2 ℓ 발효기에서 성장시키고, IPTG로 유도시키고, 분석을 위해 샘플을 제조하였다.

[0262] 다중 발효 조건을 평가함으로써 SDS-CGE에 의해 측정된 바, 대략 2 g/L인 상위 C. 디피실레 B 독소 발현을 수득하였다. pH 6.5 및 32℃에서 0.08 mM IPTG를 첨가하였을 때 대략 160 OD로 고성능의 상위 발효 배양물이 유도되었다. 가용성 C. 디피실레 B 독소 농도를 SDS-CGE로 측정하였다(도 12 및 하기 표 19 참조). 유도된 단백질의 정제는 웨스턴 블롯에 의해 확인하였다.

표 19

표 19. 가용성 C. 디피실레 B 독소 (TcdB) 역가

균주	발효물	생성물	생성물 농도 (g/L)
PS538-671	U5	TcdB	1.6 ± 0.4
PS538-671	U6	TcdB	2.1 ± 0.2
PS538-674	U7	TcdB	1.8 ± 0.2

[0263]

[0264] 실시예 12: 재조합 외독소 A 단백질의 고처리량 발현

[0265] P. 아에루기노사 외독소 A 발현 균주의 구성 및 성장

[0266] rEPA 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 P. 아에루기노사 외독소 A 돌연변이체 rEPA 코딩 서열을 구성하였다. 도 13은 발현된 합성 rEPA 유전자의 아미노산 및 DNA 서열을 나타내는 것이다.

[0267] (표 8에 기재된) CRM197과 함께 사용된 것과 동일한 10개의 P. 플루오레센스 분비 리더 코딩 서열에 융합된, 도

13에 명시된 것과 같은 결실 돌연변이체 rEPA 중 어느 것을 코딩하는 최적화된 서열을 보유하는 플라스미드를 구성하였다. 적절히 폴딩된 형태 및 활성 형태로 회수될 수 있도록 주변세포질로 단백질을 표적화시키기 위해 분비 리더 코딩 서열을 포함하였다.

[0268] rEPA 단백질에 융합된 10개의 분비 리더를 발현하는 구성체를 하기 표 20에 열거된, 8개의 P. 플루오레센스 숙주에서 테스트하였다. 숙주 세포를 제시된 플라스미드로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다. 최고수율은 4.7-6.7 g/L의 가용성 rEPA 범위였다.

표 20

표 20. 외독소 A 숙주 균주

숙주 균주	유전자형	유형
5	htpX	PD
6	세라리신	PD
7	Lon La1 aprA ; grpE dnaKJ 를 과발현	PD/FMO
8	ppiB	PD
1	Lon-, La-, aprA-	PD
2	hslUV-, prc1-, degP1-, degP2-, aprA-; degP2 S219A 를 과발현	PD + FMO
3	dsbABCD 를 과발현	FMO
4	grpE, dnaKJ 를 과발현	FMO

[0269]

[0270] 숙주 세포를 제시된 플라스미드로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다. 최고수율은 1.6~2.2 g/L의 가용성 외독소 A 단백질 범위였다. 하기 표 21에는 추가 테스트를 위해 선택된 각 발현 균주의 가용성 rEPA 수율이 제시되어 있다.

표 21

표 21. rEPA HTP 발현 요약

균주 번호	숙주	플라스미드	분비 리더	부피 수율 (g/L)
PS538-1670	3	p538-250	DsbC	6.7
PS538-1663	3	p538-243	Ibp-s31a	5.7
PS538-1633	1	p538-243	Ibp-s31a	5.7
PS538-1640	1	p538-249	Pbp-A20V	4.7
PS538-1662	3	p538-242	Azu	4.2
PS538-1632	1	p538-242	Azu	3.2
PS538-1671	4	p538-241	DsbA	2.9
PS538-1665	3	p538-245	Tpr	2.7
PS538-1667	3	p538-247	CupA2	2.6
PS538-1674	4	p538-244	TolB	2.3
PS538-1672	4	p538-242	Azu	2.2
PS538-1676	4	p538-246	CupB2	2.2
PS538-1677	4	p538-247	CupA2	2.1
PS538-1635	1	p538-245	Tpr	2.0
PS538-1675	4	p538-245	Tpr	2.0
PS538-1673	4	p538-243	Ibp-s31a	2.0
PS538-1680	4	p538-250	DsbC	1.9
PS538-1679	4	p538-249	Pbp-A20V	1.7
PS538-1669	3	p538-249	Pbp-A20V	1.6
PS538-1678	4	p538-248	NikA	1.5
PS538-1652	2	p538-242	Azu	1.5
PS538-1653	2	p538-243	Ibp-s31a	1.4
PS538-1660	2	p538-250	DsbC	1.4
PS538-1637	1	p538-247	CupA2	1.3
PS538-1666	3	p538-246	CupB2	1.1
PS538-1636	1	p538-246	CupB2	1.0
PS538-1634	1	p538-244	TolB	1.0
PS538-1627	8	p538-247	CupA2	0.8
PS538-1631	1	p538-241	DsbA	0.8
PS538-1622	8	p538-242	Azu	0.8
PS538-1661	3	p538-241	DsbA	0.7
PS538-1603	5	p538-243	Ibp-s31a	0.6
PS538-1630	8	p538-250	DsbC	0.6
PS538-1602	5	p538-242	Azu	0.6
PS538-1605	5	p538-245	Tpr	0.6
PS538-1623	8	p538-243	Ibp-s31a	0.6
PS538-1664	3	p538-244	TolB	0.5
PS538-1668	3	p538-248	NikA	0.5
PS538-1610	5	p538-250	DsbC	0.5
PS538-1606	5	p538-246	CupB2	0.4
PS538-1659	2	p538-249	Pbp-A20V	0.4
PS538-1607	5	p538-247	CupA2	0.4
PS538-1626	8	p538-246	CupB2	0.4

[0271]

PS538-1625	8	p538-245	Tpr	0.4
PS538-1638	1	p538-248	NikA	0.3
PS538-1609	5	p538-249	Pbp-A20V	0.3
PS538-1604	5	p538-244	TolB	0.3
PS538-1629	8	p538-249	Pbp-A20V	0.3
PS538-1657	2	p538-247	CupA2	0.2
PS538-1651	2	p538-241	DsbA	0.2
PS538-1601	5	p538-241	DsbA	0.2
PS538-1624	8	p538-244	TolB	0.2
PS538-1621	5	p538-241	DsbA	0.2
PS538-1608	5	p538-248	NikA	0.2
PS538-1654	2	p538-244	TolB	0.2
PS538-1628	8	p538-248	NikA	0.1
PS538-1658	2	p538-248	NikA	0.1
PS538-1655	2	p538-245	Tpr	0.1
PS538-1641	7	p538-241	DsbA	0.1
PS538-1611	6	p538-241	DsbA	NQ
PS538-1612	6	p538-242	Azu	NQ
PS538-1613	6	p538-243	Ibp-s31a	NQ
PS538-1614	6	p538-244	TolB	NQ
PS538-1615	6	p538-245	Tpr	NQ
PS538-1616	6	p538-246	CupB2	NQ
PS538-1617	6	p538-247	CupA2	NQ
PS538-1618	6	p538-248	NikA	NQ
PS538-1619	6	p538-249	Pbp-A20V	NQ
PS538-1620	6	p538-250	DsbC	NQ
PS538-1642	7	p538-242	Azu	NQ
PS538-1643	7	p538-243	Ibp-s31a	NQ
PS538-1644	7	p538-244	TolB	NQ
PS538-1645	7	p538-245	Tpr	NQ
PS538-1646	7	p538-246	CupB2	NQ
PS538-1647	7	p538-247	CupA2	NQ
PS538-1648	7	p538-248	NikA	NQ
PS538-1649	7	p538-249	Pbp-A20V	NQ
PS538-1650	7	p538-250	DsbC	NQ
PS538-1656	2	p538-246	CupB2	NQ

NQ= 정량화 불가

[0272]

[0273] 실시예 13: 재조합 슈도모나스 아에루기노사 외독소 A 단백질의 대규모 발현

[0274]

재조합 P. 아에루기노사 외독소 A 단백질(rEPA)을 2 ℓ 발효기에서 슈도모나스 플루오레센스 균주 PS538-1633, PS538-1640 및 PS538-1670에서 제조하였다. 본원에 기술된 바와 같고, 또한, 예를 들어, 문헌 [Riesenberg, D., et al., 1991]에 기술된 바와 같은 미네랄 염 배지를 함유하는 2 ℓ 발효기에서 배양물을 성장시키고, 암모니아를 첨가하여 pH 6.5 및 32℃에서 유지시켰다. 발효기에서 살포 공기 및 산소의 흐름 및 교반을 증가시켜 용존 산소를 과량으로 유지시켰다. 발효 과정 중에 글리세롤을 배양물에 전달하여 과량 수준으로 유지시켰다. 유도를 위한 표적 배양물 세포 밀도(575 nm에서의 광학 밀도(A575))에 도달할 때까지 이들 조건을 유지시켰고, 이때 IPTG를 첨가하여 rEPA가 제조될 수 있도록 개시시켰다. 유도 시 세포 밀도는 A575 40~200 흡광 단위(AU)로 다양할 수 있다. IPTG 농도는 0.02~0.4 mM 범위로, pH는 6~7.5 범위로, 온도는 20~35℃ 범위로 다양할 수 있다. 16-24시간 후에, 각 생물반응기로부터의 배양물을 원심분리로 수거하고, -80℃에서 세포 펠릿을 냉동시켰다. SDS-CGE에 의해 생성물 제조에 대해 샘플을 분석하였다.

[0275]

다중 발효 조건을 평가함으로써 SDS-CGE에 의해 측정된 바, 최대 32 g/L 이하인 상위 rEPA 발현을 수득하였다(도 15 및 16 참조). 유도된 단백질의 정제는 P. 아에루기노사 외독소 A에 특이적인 항체를 사용하는 웨스턴 블롯 분석에 의해 확인하였다(도 17). 수득된 수율은 하기 표 22에 제시되어 있다.

표 22

표 22. rEPA 발효물 분석

균주 번호	발효물	수율(g/L)
PS538-1633	U1	15.5+/- 0.7
PS538-1633	U2	11.1 +/- 0.6
PS538-1640	U3	20.1 +/- 1.7
PS538-1640	U5	31.9 +/- 1.6
PS538-1670	U6	20.0 +/- 0.7
PS538-1670	U7	14.6 +/- 1.1
PS538-1670	U8	31.0 +/- 1.7

[0276]

[0277] 실시예 14: 재조합 야생형 디프테리아 독소 단백질의 고처리량 발현

[0278] 야생형 디프테리아 독소 발현 균주의 구성 및 성장

[0279] 야생형 디프테리아 독소 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 디프테리아 독소 코딩 서열을 구성하였다. 서열 번호: 36은 발현된 합성 디프테리아 독소 유전자에 의해 코딩된 아미노산 서열을 나타내고, 서열 번호: 37은 최적화된 발현된 합성 디프테리아 독소 유전자의 DNA 서열을 나타내는 것이다.

[0280] (표 8에 기재된) CRM197과 함께 사용된 것과 동일한 10개의 P. 플루오레센스 분비 리더 코딩 서열에 융합된, 디프테리아 독소를 코딩하는 최적화된 서열을 보유하는 플라스미드를 구성하였다. 적절히 폴딩된 형태 및 활성 형태로 회수될 수 있도록 주변세포질로 단백질을 표적화시키기 위해 분비 리더 코딩 서열을 포함하였다.

[0281] 재조합 디프테리아 독소 단백질에 융합된 10개의 분비 리더를 발현하는 구성체를 P. 플루오레센스 숙주에서 테스트하였다. 각 리더를 사용하여 표 9에 열거된 4가지의 숙주를 테스트하였다. 숙주 세포를 제시된 플라스미드로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

[0282] 실시예 15: 재조합 야생형 디프테리아 독소 단백질의 대규모 발현

[0283] 재조합 야생형 디프테리아 독소 단백질은 선택된 슈도모나스 플루오레센스 페넥스 익스프레션 테크놀로지™ 균주에서 제조하였다. 선택된 균주를 CRM197 대규모 발현에 대해 상기 기술된 바와 같이, 2 ℓ 발효기에서 성장시키고, IPTG로 유도시키고, 분석을 위해 샘플을 제조하였다. 샘플을 SDS-CGE에 의해 분석하였다.

[0284] 실시예 16: 재조합 콜레라 홀로톡신 단백질의 고처리량 발현

[0285] CTX 발현 균주의 구성 및 성장

[0286] CTX 아미노산 서열을 코딩하는 P. 플루오레센스의 선호 코돈을 사용하여 CTX 코딩 서열을 구성하였다. 코딩 서열은 도 14에 제시된 CTX 유전자의 아미노산 및 DNA 서열에 기초한 것이다.

[0287] (표 8에 기재된) CRM197과 함께 사용된 것과 동일한 10개의 P. 플루오레센스 분비 리더 코딩 서열에 융합된, 최적화된 CTX 서열을 보유하는 플라스미드를 구성하였다. 적절히 폴딩된 형태 및 활성 형태로 회수될 수 있도록 주변세포질로 단백질을 표적화시키기 위해 분비 리더를 포함하였다.

[0288] 재조합 CTX 단백질에 융합된 10개의 분비 리더를 발현하는 구성체를 P. 플루오레센스 숙주에서 테스트하였다. 각 발현 플라스미드를 사용하여 표 9에 열거된 4가지의 숙주를 테스트하였다. 숙주 세포를 제시된 플라스미드로 전기천공시키고, CRM197 고처리량 발현에 대해 상기 기술된 바와 같이 96 웰 방식으로 성장시키고, 유도하였다. CRM197 고처리량 발현 샘플에 대해 상기 기술된 바와 같이 샘플을 제조하고, SDS-CGE에 의해 분석하였다.

[0289] 실시예 17: 재조합 콜레라 홀로톡신 단백질의 대규모 발현

[0290] 재조합 콜레라 홀로톡신 단백질은 선택된 슈도모나스 플루오레센스 페넥스 익스프레션 테크놀로지™ 균주에서 제조하였다. 선택된 균주를 CRM197에 대해 상기 기술된 바와 같이, 2 ℓ 발효기에서 성장시키고, IPTG로 유도시키고, 분석을 위해 샘플을 제조하였다. 샘플을 SDS-CGE에 의해 분석하였다.

표 23

표 23. 서열 목록

서열 번호	설명
1	CRM197 아미노산 서열
2	최적화된 CRM197 DNA 서열
3	DsbA 분배 리더
4	Azu
5	Ibp-S31A
6	Tpr
7	CupB2
8	CupA2
9	NikA
10	Pbp A20V
11	DsbC
12	TolB
13	Pbp
14	Lao
15	CupC2
16	PorE
17	Pbp
18	FlgI
19	ttg2C
20	CRM197 네이티브 리더
21	절단 생성물 GADD
22	콜레라 독소 B 아미노산 서열
23	최적화된 콜레라 독소 B DNA 서열(번역과 함께); 서열 번호: 22의 아미노산 서열을 코딩
24	백일해 독소 S1 R9K E129A DNA 서열
25	백일해 독소 S1 R9K E129A 아미노산 서열
26	백일해 독소 S2 아미노산 서열
27	백일해 독소 S3 아미노산 서열
28	백일해 독소 S4 아미노산 서열
29	백일해 독소 S5 아미노산 서열
30	파상풍 독소 C 아미노산 서열
31	최적화된 파상풍 독소 C DNA 서열
32	TcdB 아미노산 서열
33	최적화된 TcdB DNA 서열
34	외독소 A 아미노산 서열
35	야생형 백일해 독소이드의 DNA 서열
36	야생형 디프테리아 독소 아미노산 서열
37	최적화된 야생형 디프테리아 독소 DNA 서열
38	콜레라 독소 A 아미노산 서열
39	콜레라 독소 B 아미노산 서열
40	콜레라 할로톡신 (CTX) DNA 서열
41	야생형 백일해 독소 S1 아미노산 서열
42	백일해 독소 S2 아미노산 서열
43	백일해 독소 S4 아미노산 서열

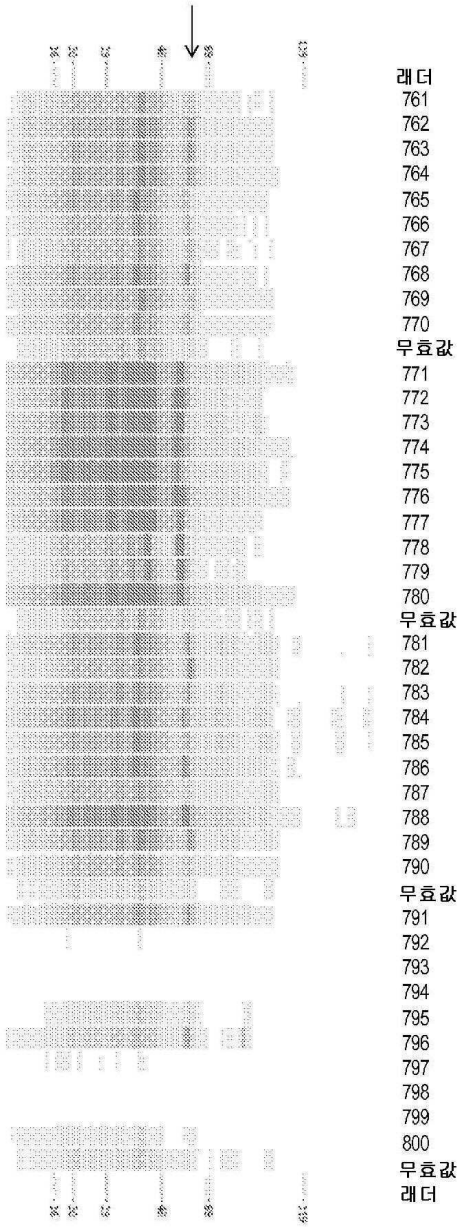
[0291]

44	백일해 독소 S5 아미노산 서열
45	백일해 독소 S3 아미노산 서열
46	헥사-히스티딘 친화 태그

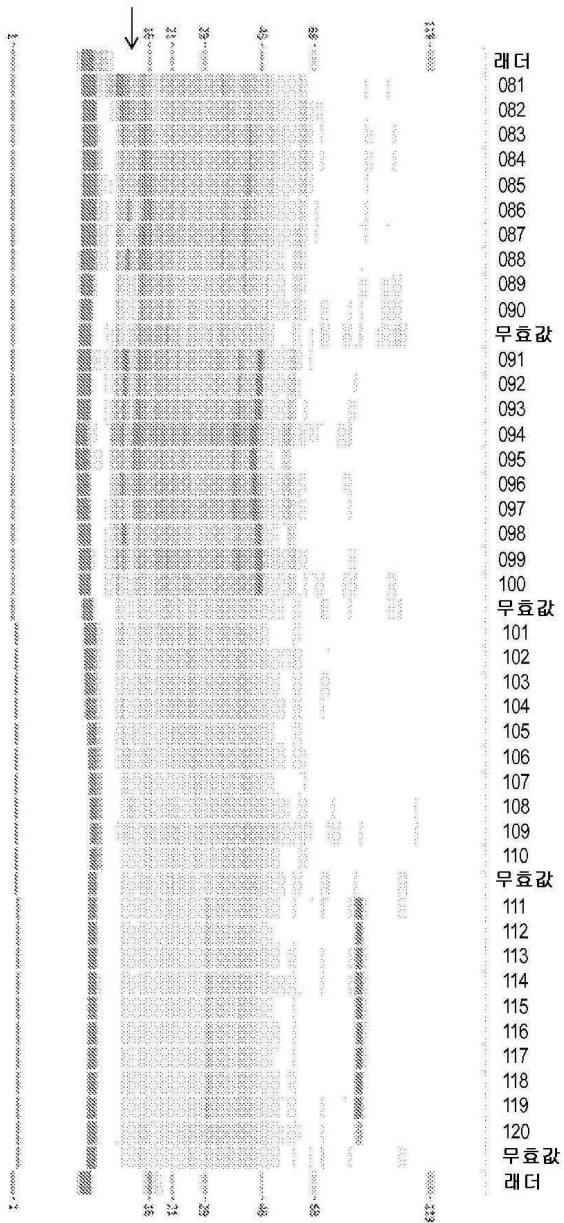
[0292]

도면

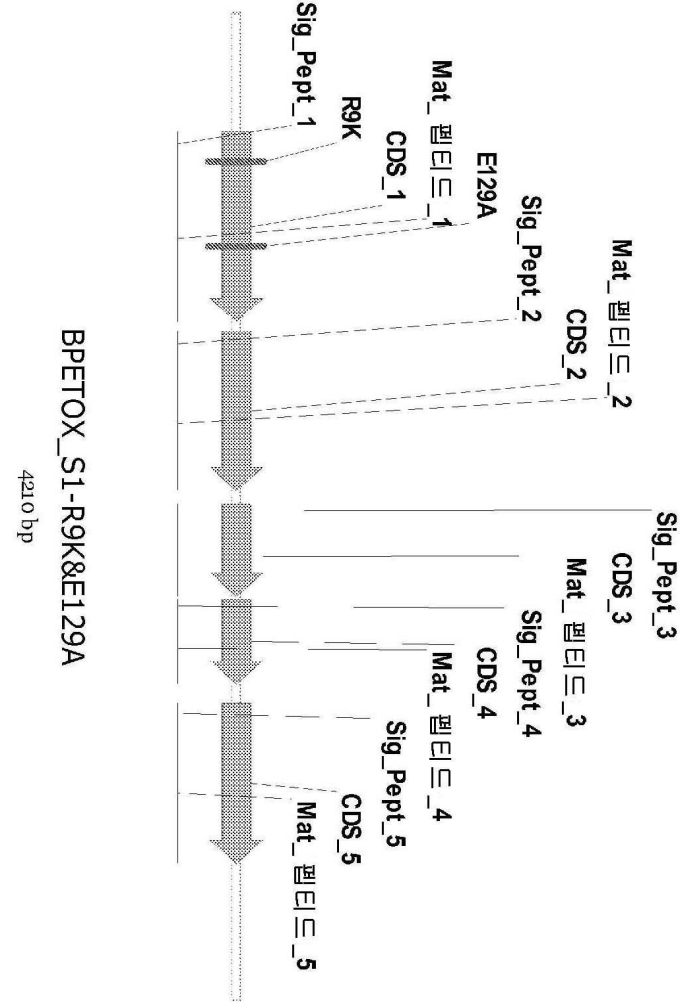
도면1



도면2



도면3



도면4a

신호 서열

~~~~~  
S1  
~~~~~  
Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp Leu Thr Trp Leu Ala
1 ATG CGT TGC ACT CGG GCA ATT CGC CAA ACC GCA AGA ACA GGC TGG CTG ACG TGG CTG CGC
signal sequence
~~~~~  
S1  
~~~~~  
Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala Trp Ala Asp Asp Pro Pro Ala Thr
61 ATT CTT GCC GTC ACG GCG CCC GTG ACT TCG CCG GCA TGG GCC GAC GAT CCT CCC GCC ACC
S1
~~~~~  
Val Tyr Lys Tyr Asp Ser Arg Pro Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp  
121 GTA TAC AAA TAT GAC TCC CGC CCG CCG GAG GAC GTT TTC CAG AAC GGA TTC ACG GCG TGG  
S1  
~~~~~  
Gly Asn Asn Asp Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser
181 GGA AAC AAC GAC AAT GTG CTC GAC CAT CTG ACC GGA CGT TCC TGC CAG GTC GCG AGC AGC
S1
~~~~~  
Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu Val Tyr Leu Glu His  
241 AAC AGC GCT TTC GTC TCC ACC AGC AGC AGC CGG CGC TAT ACC GAG GTC TAT CTC GAA CAT  
S1  
~~~~~  
Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly Arg Gly Thr Gly His Phe Ile Gly
301 CGC ATG CAG GAA GCG GTC GAG GCC GAA CGC GCC GGC AGG GGC ACC GGC CAC TTC ATC GGC
S1
~~~~~  
Tyr Ile Tyr Glu Val Arg Ala Asp Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu  
361 TAC ATC TAC GAA GTC CGC GCC GAC AAC AAT TTC TAC GGC GCC GCC AGC TCG TAC TTC GAA  
S1  
~~~~~  
Tyr Val Asp Thr Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr
421 TAC GTC GAC ACT TAT GGC GAC AAT GCC GGC CGT ATC CTC GCC GGC GCG CTG GCC ACC TAC
S1
~~~~~  
Gln Ser Ala Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg Arg Val Thr Arg  
481 CAG AGC GCA TAT CTG GCA CAC CCG CGC ATT CCG CCC GAA AAC ATC CGC AGG GTA ACG CGG  
S1  
~~~~~  
Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr Thr Glu Tyr Ser Asn Ala Arg Tyr
541 GTC TAT CAC AAC GGC ATC ACC GGC GAG ACC ACG ACC ACG GAG TAT TCC AAC GCT CGC TAC

도면4b

신호 서열

~~~~~  
S1  
~~~~~  
Val Ser Gln Gln Thr Arg Ala Asn Pro Asn Pro Tyr Thr Ser Arg Arg Ser Val Ala Ser
601 GTC AGC CAG CAG ACT CGC GCC AAT CCC AAC CCC TAC ACA TCG CGA AGG TCC GTA GCG TCG
S1
~~~~~  
Ile Val Gly Thr Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys Met Ala Arg Gln Ala  
661 ATC GTC GGC ACA TTG GTG CGC ATG GCG CCG GTG ATA GGC GCT TGC ATG GCG CGG CAG GCC  
S1  
~~~~~  
Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala Met Val Leu Val
721 GAA AGC TCC GAG GCC ATG GTC GCA GCC TGG TCC GAA CGC GCC GGC GAG GCG ATG GTT CTC GTG
S1
~~~~~  
Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe  
781 TAC TAC GAA AGC ATC GCG TAT TCG TTC TAG ACC TGG CCC AGC CCC GCC CAA CTC CGG TAA  
~~~~~  
신호 서열
~~~~~  
S2  
~~~~~  
Met Pro Ile Asp Arg Lys Thr Leu Cys His Leu Leu Ser Val Leu Pro Leu
841 TTG AAC AGC ATG CCG ATC GAC CGC AAG ACG CTC TGC CAT CTC CTG TCC GTT CTG CCG TTG
~~~~~  
신호 서열  
~~~~~  
S2
~~~~~  
Ala Leu Leu Gly Ser His Val Ala Arg Ala Ser Thr Pro Gly Ile Val Ile Pro Pro Gln  
901 GCC CTC CTC GGA TCT CAC GTG GCG CGG GCC TCC ACG CCA GGC ATC GTC ATT CCG CCG CAG  
S2  
~~~~~  
Glu Gln Ile Thr Gln His Gly Ser Pro Tyr Gly Arg Cys Ala Asn Lys Thr Arg Ala Leu
961 GAA CAG ATT ACC CAG CAT GGC AGC CCC TAT GGA CGC TGC GCG AAC AAG ACC CGT GCC CTG
S2
~~~~~  
Thr Val Ala Glu Leu Arg Gly Ser Gly Asp Leu Gln Glu Tyr Leu Arg His Val Thr Arg  
1021 ACC GTG GCG GAA TTG CGC GGC AGC GGC GAT CTG CAG GAG TAC CTG CGT CAT GTG ACG CGC  
S2  
~~~~~  
Gly Trp Ser Ile Phe Ala Leu Tyr Asp Gly Thr Tyr Leu Gly Gly Glu Tyr Gly Gly Val
1081 GGC TGG TCA ATA TTT GCG CTC TAC GAT GGC ACC TAT CTC GGC GGC GAA TAT GGC GGC GTG
S2
~~~~~  
Ile Lys Asp Gly Thr Pro Gly Gly Ala Phe Asp Leu Lys Thr Thr Phe Cys Ile Met Thr  
1141 ATC AAG GAC GGA ACA CCC GGC GGC GCA TTC GAC CTG AAA ACG ACG TTC TGC ATC ATG ACC  
S2  
~~~~~  
Thr Arg Asn Thr Gly Gln Pro Ala Thr Asp His Tyr Tyr Ser Asn Val Thr Ala Thr Arg
1201 ACG CGC AAT ACG GGT CAA CCC GCA ACG GAT CAC TAC TAC AGC AAC GTC ACC GCC ACT CGC

도면4c

```

S2
~~~~~
1261 Leu Leu Ser Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val Arg Ser Gly Gln Pro Val
CTG CTC TCC AGC ACC AAC AGC AGG CTA TGC GCG GTC TTC GTC AGA AGC GGG CAA CCG GTC
S2
~~~~~
1321 Ile Gly Ala Cys Thr Ser Pro Tyr Asp Gly Lys Tyr Trp Ser Met Tyr Ser Arg Leu Arg
ATT GGC GCC TGC ACC AGC CCG TAT GAC GGC AAG TAC TGG AGC ATG TAC AGC CGG CTG CGG
S2
~~~~~
1381 Lys Met Leu Tyr Leu Ile Tyr Val Ala Gly Ile Ser Val Arg Val His Val Ser Lys Glu
AAA ATG CTT TAC CTG ATC TAC GTG GCC GGC ATC TCC GTA GCG GTC CAT GTC AGC AAG GAA
S2
~~~~~
1441 Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe Glu Thr Tyr Ala Leu Thr Gly Ile Ser
GAA CAG TAT TAC GAC TAT GAG GAC GCA ACG TTC GAG ACT TAC GCC CTT ACC GGC ATC TCC
S2
~~~~~
1501 Ile Cys Asn Pro Gly Ser Ser Leu Cys
ATC TGC AAT CCT GGA TCA TCC TTA TGC TGA GAC GCT TCC CCA CTC GAA CCA CCG CCC CGG
신호 서열
~~~~~
S4
~~~~~
1561 Val Arg Ala Leu Ala Trp Leu Leu Ala Ser Gly
GAC AGG GCG GCG CCC GGC GGT CCG GC GTG CCG GCC CTG GCG TGG TTG CTG GCA TCC GGC
신호 서열
~~~~~
S4
~~~~~
1620 Ala Met Thr His Leu Ser Pro Ala Leu Ala Asp Val Pro Tyr Val Leu Val Lys Thr Asn
GCG ATG ACG CAT CTT TCC CCC GGC CTG GCC GAC GTT CCT TAT GTG CTG GTG AAG ACC AAT
S4
~~~~~
1680 Met-Val Val Thr Ser Val Ala Met Lys Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val
ATG GTG GTC ACC AGC GTA GCC ATG AAG CCG TAT GAA GTC ACC CCG ACG CGC ATG CTG GTC
S4
~~~~~
1740 Cys Gly Ile Ala Ala Lys Leu Gly Ala Ala Ala Ser Ser Pro Asp Ala His Val Pro Phe
TGC GGC ATC GCC GCC AAA CTG GGC GCC GCG GCC AGC AGC CCG GAC GCG CAC GTG CCG TTC
S4
~~~~~
1800 Cys Phe Gly Lys Asp Leu Lys Arg Pro Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala
TGC TTC GGC AAG GAT CTC AAG CGT CCC GGC AGC AGT CCC ATG GAA GTC ATG TTG CGC GCC
S4
~~~~~
1860 Val Phe Met Gln Gln Arg Pro Leu Arg Met Phe Leu Gly Pro Lys Gln Leu Thr Phe Glu
GTC TTC ATG CAA CAA CCG CCG CTG CGC ATG TTT CTG GGT CCC AAG CAA CTC ACT TTC GAA

```

도면4d

```

S4
~~~~~
1920 Gly Lys Pro Ala Leu Glu Leu Ile Arg Met Val Glu Cys Ser Gly Lys Gln Asp Cys Pro
GGC AAG CCC GCG CTC GAA CTG ATC CCG ATG GTC GAA TGC AGC GGC AAG CAG GAT TGC CCC T
신호 서열
~~~~~
S5
~~~~~
1981 Met His Thr Ile Ala Ser Ile Leu Leu Ser Val Leu Gly Ile Tyr Ser
GAA GGC GAA CCC C ATG CAT ACC ATC GCA TCC ATC CTG TTG TCC GTG CTC GGC ATA TAC AG
신호 서열
~~~~~
S5
~~~~~
2041 Pro Ala Asp Val Ala Gly Leu Pro Thr His Leu Tyr Lys Asn Phe Thr Val Gln Glu Leu
C CCG GCT GAC GTC GCC GGC TTG CCG ACC CAT CTG TAC AAG AAC TTC ACT GTC CAG GAG CT
S5
~~~~~
2101 Ala Leu Lys Leu Lys Gly Lys Asn Gln Glu Phe Cys Leu Thr Ala Phe Met Ser Gly Arg
G GCC TTG AAA CTG AAG GGC AAG AAT CAG GAG TTC TGC CTG ACC GCC TTC ATG TCG GGC AG
S5
~~~~~
2161 Ser Leu Val Arg Ala Cys Leu Ser Asp Ala Gly His Glu His Asp Thr Trp Phe Asp Thr
A AGC CTG GTC CCG GCG TGC CTG TCC GAC GCG GGA CAC GAG CAC GAC ACG TGG TTC GAC AC
S5
~~~~~
2221 Met Leu Gly Phe Ala Ile Ser Ala Tyr Ala Leu Lys Ser Arg Ile Ala Leu Thr Val Glu
C ATG CTT GGC TTT GCC ATA TCC GCG TAT GCG CTC AAG AGC CCG ATC GCG CTG ACG GTG GA
S5
~~~~~
2281 Asp Ser Pro Tyr Pro Gly Thr Pro Gly Asp Leu Leu Glu Leu Gln Ile Cys Pro Leu Asn
A GAC TCG CCG TAT CCG GGC ACT CCC GGC GAT CTG CTC GAA CTG CAG ATC TGC CCG CTC AA
S5
~~~~~
2341 Gly Tyr Cys Glu
C GGA TAT TGC GAA TG AAC CCT TCC GGA GGT TTC GAC GTT TCC GCG CAA TCC GCT TGA GAC
신호 서열
~~~~~
S3
~~~~~
2401 Met Leu Ile Asn Asn Lys Lys Leu
CAT CTT CCG CCC TGG TTC CAT TCC GGG AAC ACC GCA AC ATG CTG ATC AAC AAC AAG AAG C
신호 서열
~~~~~
S3
~~~~~
2461 Leu His His Ile Leu Pro Ile Leu Val Leu Ala Leu Leu Gly Met Arg Thr Ala Gln Ala
TG CTT CAT CAC ATT CTG CCC ATC CTG GTG CTC GCC CTG CTG GGC ATG CGC ACG GCC CAG G

```

도면4e

신호 서열

```

~~~~~
S3
~~~~~
2521 Val Ala Pro Gly Ile Val Ile Pro Pro Lys Ala Leu Phe Thr Gln Gln Gly Gly Ala Tyr
CC GTT GCG CCA GGC ATC GTC ATC CCG CCG AAG GCA CTG TTC ACC CAA CAG GGC GGC GCC T
S3
~~~~~
2581 Gly Arg Cys Pro Asn Gly Thr Arg Ala Leu Thr Val Ala Glu Leu Arg Gly Asn Ala Glu
AT GGA CGC TGC CCG AAC GGA ACC CGC GGC TTG ACC GTG GCC GAA CTG CGC GGC AAC GCC G
S3
~~~~~
2641 Leu Gln Thr Tyr Leu Arg Gln Ile Thr Pro Gly Trp Ser Ile Tyr Gly Leu Tyr Asp Gly
AA TTG CAG ACG TAT TTG CGC CAG ATA ACG CCC GGC TGG TCC ATA TAC GGT CTC TAT GAC G
S3
~~~~~
2701 Thr Tyr Leu Gly Gln Ala Tyr Gly Gly Ile Ile Lys Asp Ala Pro Pro Gly Ala Gly Phe
GT ACG TAC CTG GGC CAG GCG TAC GGC GGC ATC ATC AAG GAC GCG CCG CCA GGC GCG GGC T
S3
~~~~~
2761 Ile Tyr Arg Glu Thr Phe Cys Ile Thr Thr Ile Tyr Lys Thr Gly Gln Pro Ala Ala Asp
TC ATT TAT CGC GAA ACT TTC TGC ATC ACG ACC ATA TAC AAG ACC GGG CAA CCG GCT GCG G
S3
~~~~~
2821 His Tyr Tyr Ser Lys Val Thr Ala Thr Arg Leu Leu Ala Ser Thr Asn Ser Arg Leu Cys
AT CAC TAC TAC AGC AAG GTC ACG GCC ACG CGC CTG CTC GCC AGC ACC AAC AGC AGG CTG T
S3
~~~~~
2881 Ala Val Phe Val Arg Asp Gly Gln Ser Val Ile Gly Ala Cys Ala Ser Pro Tyr Glu Gly
GC GCG GTA TTC GTC AGG GAC GGG CAA TCG GTC ATC GGA GCC TGC GCC AGC CCG TAT GAA G
S3
~~~~~
2941 Arg Tyr Arg Asp Met Tyr Asp Ala Leu Arg Arg Leu Leu Tyr Met Ile Tyr Met Ser Gly
GC AGG TAC AGA GAC ATG TAC GAC GCG CTG CGG CGC CTG CTG TAC ATG ATC TAT ATG TCC G
S3
~~~~~
3001 Leu Ala Val Arg Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr
GC CTT GCC GTA CGC GTC CAC GTC AGC AAG GAG GAA CAG TAT TAC GAC TAC GAG GAC GCC A
S3
~~~~~
3061 Phe Gln Thr Tyr Ala Leu Thr Gly Ile Ser Leu Cys Asn Pro Ala Ala Ser Ile Cys
CA TTC CAG ACC TAT GCC CTC ACC GGC ATT TCC CTC TGC AAC CCG GCA GCG TCG ATA TGC

```

도면5a

1 Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp
Leu Thr Trp Leu Ala

21 Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala Trp Ala Asp
 Asp Pro Pro Ala Thr

41 Val Tyr Lys Tyr Asp Ser Arg Pro Pro Glu Asp Val Phe Gln Asn
 Gly Phe Thr Ala Trp

61 Gly Asn Asn Asp Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys
 Gln Val Gly Ser Ser

81 Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu
 Val Tyr Leu Glu His

101 Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly Arg Gly Thr
 Gly His Phe Ile Gly

121 Tyr Ile Tyr Glu Val Arg Ala Asp Asn Asn Phe Tyr Gly Ala Ala
 Ser Ser Tyr Phe Glu

141 Tyr Val Asp Thr Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly
 Ala Leu Ala Thr Tyr

161 Gln Ser Ala Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile
 Arg Arg Val Thr Arg

181 Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr Thr Glu Tyr
 Ser Asn Ala Arg Tyr

201 Val Ser Gln Gln Thr Arg Ala Asn Pro Asn Pro Tyr Thr Ser Arg
 Arg Ser Val Ala Ser

221 Ile Val Gly Thr Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys
 Met Ala Arg Gln Ala

241 Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu
 Ala Met Val Leu Val

261 Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe

도면5b

1 Met Pro Ile Asp Arg Lys Thr Leu Cys His Leu Leu Ser Val Leu
Pro Leu Ala Leu Leu

21 Gly Ser His Val Ala Arg Ala Ser Thr Pro Gly Ile Val Ile Pro
Pro Gln Glu Gln Ile

41 Thr Gln His Gly Ser Pro Tyr Gly Arg Cys Ala Asn Lys Thr Arg
Ala Leu Thr Val Ala

61 Glu Leu Arg Gly Ser Gly Asp Leu Gln Glu Tyr Leu Arg His Val
Thr Arg Gly Trp Ser

81 Ile Phe Ala Leu Tyr Asp Gly Thr Tyr Leu Gly Gly Glu Tyr Gly
Gly Val Ile Lys Asp

101 Gly Thr Pro Gly Gly Ala Phe Asp Leu Lys Thr Thr Phe Cys Ile
Met Thr Thr Arg Asn

121 Thr Gly Gln Pro Ala Thr Asp His Tyr Tyr Ser Asn Val Thr Ala
Thr Arg Leu Leu Ser

141 Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val Arg Ser Gly Gln
Pro Val Ile Gly Ala

161 Cys Thr Ser Pro Tyr Asp Gly Lys Tyr Trp Ser Met Tyr Ser Arg
Leu Arg Lys Met Leu

181 Tyr Leu Ile Tyr Val Ala Gly Ile Ser Val Arg Val His Val Ser
Lys Glu Glu Gln Tyr

201 Tyr Asp Tyr Glu Asp Ala Thr Phe Glu Thr Tyr Ala Leu Thr Gly
Ile Ser Ile Cys Asn

221 Pro Gly Ser Ser Leu Cys

도면5c

1 Met Leu Ile Asn Asn Lys Lys Leu Leu His His Ile Leu Pro
Ile Leu Val Leu Ala Leu

21 Leu Gly Met Arg Thr Ala Gln Ala Val Ala Pro Gly Ile Val
 Ile Pro Pro Lys Ala Leu

41 Phe Thr Gln Gln Gly Gly Ala Tyr Gly Arg Cys Pro Asn Gly
 Thr Arg Ala Leu Thr Val

61 Ala Glu Leu Arg Gly Asn Ala Glu Leu Gln Thr Tyr Leu Arg
 Gln Ile Thr Pro Gly Trp

81 Ser Ile Tyr Gly Leu Tyr Asp Gly Thr Tyr Leu Gly Gln Ala
 Tyr Gly Gly Ile Ile Lys

101 Asp Ala Pro Pro Gly Ala Gly Phe Ile Tyr Arg Glu Thr Phe
 Cys Ile Thr Thr Ile Tyr

121 Lys Thr Gly Gln Pro Ala Ala Asp His Tyr Tyr Ser Lys Val
 Thr Ala Thr Arg Leu Leu

141 Ala Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val Arg Asp
 Gly Gln Ser Val Ile Gly

161 Ala Cys Ala Ser Pro Tyr Glu Gly Arg Tyr Arg Asp Met Tyr
 Asp Ala Leu Arg Arg Leu

181 Leu Tyr Met Ile Tyr Met Ser Gly Leu Ala Val Arg Val His
 Val Ser Lys Glu Glu Gln

201 Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe Gln Thr Tyr Ala Leu
 Thr Gly Ile Ser Leu Cys

221 Asn Pro Ala Ala Ser Ile Cys

도면5d

1 Val Arg Ala Leu Ala Trp Leu Leu Ala Ser Gly Ala Met Thr
His Leu Ser Pro Ala Leu

21 Ala Asp Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val
 Thr Ser Val Ala Met Lys

41 Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val Cys Gly Ile
 Ala Ala Lys Leu Gly Ala

61 Ala Ala Ser Ser Pro Asp Ala His Val Pro Phe Cys Phe Gly
 Lys Asp Leu Lys Arg Pro

81 Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala Val Phe Met
 Gln Gln Arg Pro Leu Arg

101 Met Phe Leu Gly Pro Lys Gln Leu Thr Phe Glu Gly Lys Pro
 Ala Leu Glu Leu Ile Arg

121 Met Val Glu Cys Ser Gly Lys Gln Asp Cys Pro

도면5e

1 Met His Thr Ile Ala Ser Ile Leu Leu Ser Val Leu Gly Ile
Tyr Ser Pro Ala Asp Val

21 Ala Gly Leu Pro Thr His Leu Tyr Lys Asn Phe Thr Val Gln
Glu Leu Ala Leu Lys Leu

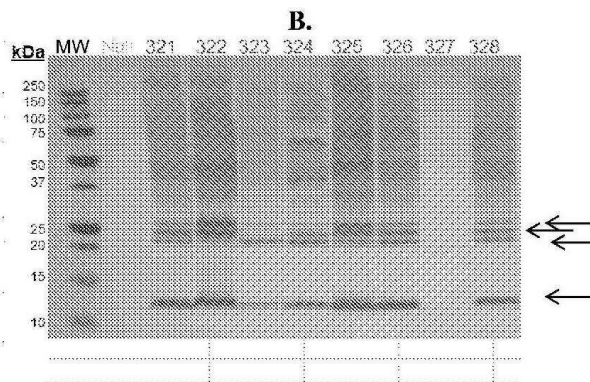
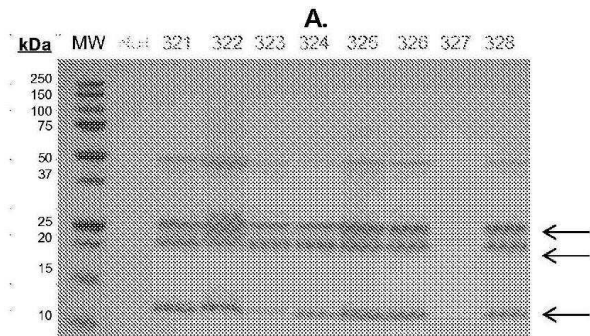
41 Lys Gly Lys Asn Gln Glu Phe Cys Leu Thr Ala Phe Met Ser
Gly Arg Ser Leu Val Arg

61 Ala Cys Leu Ser Asp Ala Gly His Glu His Asp Thr Trp Phe
Asp Thr Met Leu Gly Phe

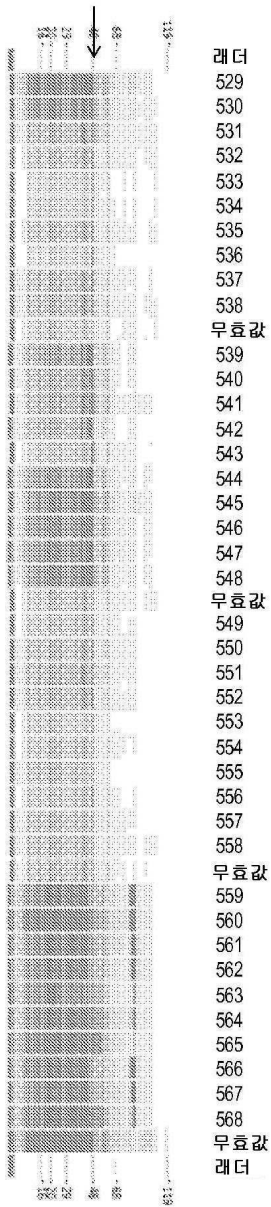
81 Ala Ile Ser Ala Tyr Ala Leu Lys Ser Arg Ile Ala Leu Thr
Val Glu Asp Ser Pro Tyr

101 Pro Gly Thr Pro Gly Asp Leu Leu Glu Leu Gln Ile Cys Pro
Leu Asn Gly Tyr Cys Glu

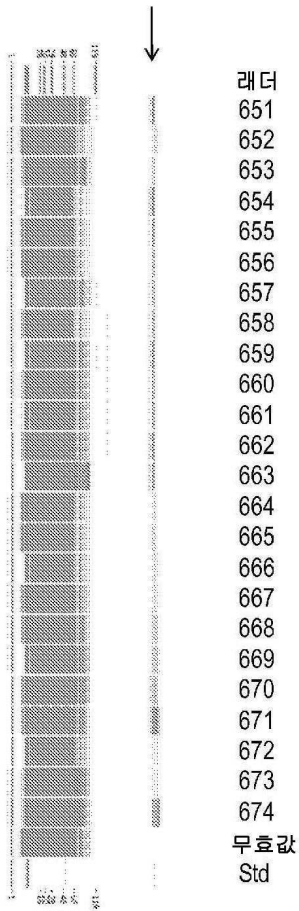
도면6



도면7



도면8



도면9a

슈도모나스 외독소 A 아미노산 서열

1 Ala Glu Glu Ala Phe Asp Leu Trp Asn Glu Cys Ala Lys Ala Cys Val Leu
Asp Leu Lys

21 Asp Gly Val Arg Ser Ser Arg Met Ser Val Asp Pro Ala Ile Ala Asp Thr
Asn Gly Gln

41 Gly Val Leu His Tyr Ser Met Val Leu Glu Gly Gly Asn Asp Ala Leu Lys
Leu Ala Ile

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

81 Pro Asn Lys Pro Val Arg Tyr Ser Tyr Thr Arg Gln Ala Arg Gly Ser Trp
Ser Leu Asn

101 Trp Leu Val Pro Ile Gly His Glu Lys Pro Ser Asn Ile Lys Val Phe Ile
His Glu Leu

121 Asn Ala Gly Asn Gln Leu Ser His Met Ser Pro Ile Tyr Thr Ile Glu Met
Gly Asp Glu

141 Leu Leu Ala Lys Leu Ala Arg Asp Ala Thr Phe Phe Val Arg Ala His Glu
Ser Asn Glu

161 Met Gln Pro Thr Leu Ala Ile Ser His Ala Gly Val Ser Val Val Met Ala
Gln Ala Gln

181 Pro Arg Arg Glu Lys Arg Trp Ser Glu Trp Ala Ser Gly Lys Val Leu Cys
Leu Leu Asp

201 Pro Leu Asp Gly Val Tyr Asn Tyr Leu Ala Gln Gln Arg Cys Asn Leu Asp
Asp Thr Trp

221 Glu Gly Lys Ile Tyr Arg Val Leu Ala Gly Asn Pro Ala Lys His Asp Leu
Asp Ile Lys

241 Pro Thr Val Ile Ser His Arg Leu His Phe Pro Glu Gly Gly Ser Leu Ala
Ala Leu Thr

261 Ala His Gln Ala Cys His Leu Pro Leu Glu Thr Phe Thr Arg His Arg Gln
Pro Arg Gly

281 Trp Glu Gln Leu Glu Gln Cys Gly Tyr Pro Val Gln Arg Leu Val Ala Leu
Tyr Leu Ala

301 Ala Arg Leu Ser Trp Asn Gln Val Asp Gln Val Ile Arg Asn Ala Leu Ala
Ser Pro Gly

321 Ser Gly Gly Asp Leu Gly Glu Ala Ile Arg Glu Gln Pro Glu Gln Ala Arg
Leu Ala Leu

341 Thr Leu Ala Ala Ala Glu Ser Glu Arg Phe Val Arg Gln Gly Thr Gly Asn
Asp Glu Ala

361 Gly Ala Ala Ser Ala Asp Val Val Ser Leu Thr Cys Pro Val Ala Ala Gly
Glu Cys Ala

도면9b

381 Gly Pro Ala Asp Ser Gly Asp Ala Leu Leu Glu Arg Asn Tyr Pro Thr Gly Ala
Glu Phe

401 Leu Gly Asp Gly Gly Asp Ile Ser Phe Ser Thr Arg Gly Thr Gln Asn Trp Thr
Val Glu

Tyr - CRM66

421 Arg Leu Leu Gln Ala His Arg Gln Leu Glu Glu Arg Gly Tyr Val Phe Val Gly
Tyr His

441 Gly Thr Phe Leu Glu Ala Ala Gln Ser Ile Val Phe Gly Gly Val Arg Ala Arg
Ser Gln

461 Asp Leu Asp Ala Ile Trp Arg Gly Phe Tyr Ile Ala Gly Asp Pro Ala Leu Ala
Tyr Gly

481 Tyr Ala Gln Asp Gln Glu Pro Asp Ala Arg Gly Arg Ile Arg Asn Gly Ala Leu
Leu Arg

501 Val Tyr Val Pro Arg Ser Ser Leu Pro Gly Phe Tyr Arg Thr Gly Leu Thr Leu
Ala Ala

521 Pro Glu Ala Ala Gly Glu Val Glu Arg Leu Ile Gly His Pro Leu Pro Leu Arg
Leu Asp

Δ - rEPA

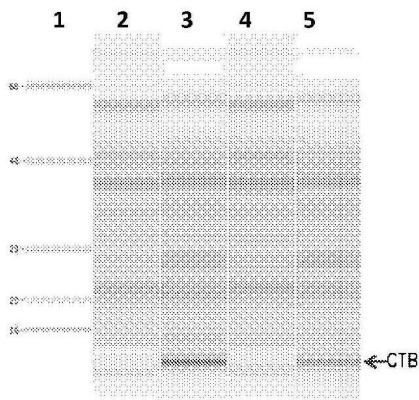
541 Ala Ile Thr Gly Pro Glu Glu Glu Gly Gly Arg Leu Glu Thr Ile Leu Gly Trp
Pro Leu

561 Ala Glu Arg Thr Val Val Ile Pro Ser Ala Ile Pro Thr Asp Pro Arg Asn Val
Gly Gly

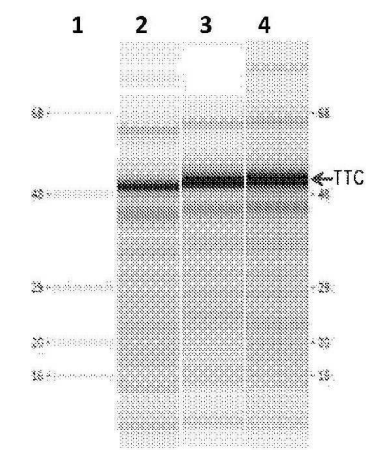
581 Asp Leu Asp Pro Ser Ser Ile Pro Asp Lys Glu Gln Ala Ile Ser Ala Leu Pro
Asp Tyr

601 Ala Ser Gln Pro Gly Lys Pro Pro Arg Glu Asp Leu Lys

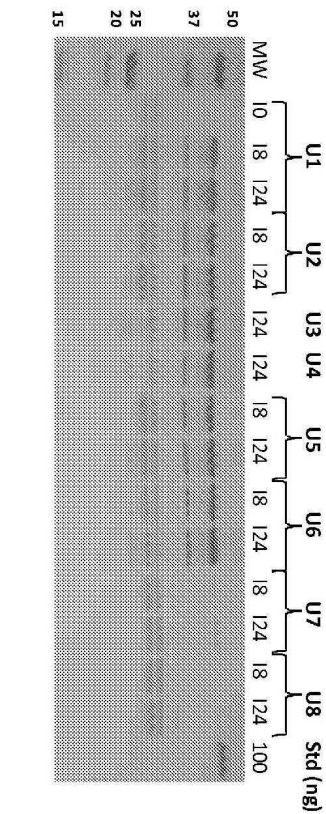
도면10



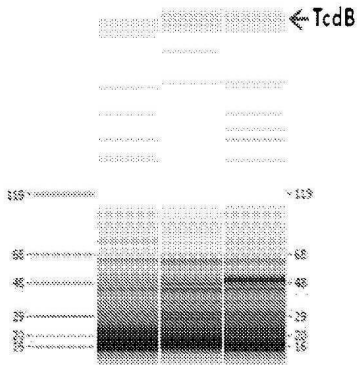
도면11a



도면11b



도면12



도면13a

신호 서열

S1

Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp Leu Thr Trp Leu Ala
 1 ATG CGT TGC ACT CCG GCA ATT CCG CAA ACC GCA AGA ACA GGC TGG CTG ACG TGG CTG GCG

신호 서열

S1

Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala Trp Ala Asp Asp Pro Pro Ala Thr
 61 ATT CTT GCC GTC ACG GCG CCC GTG ACT TCG CCG GCA TGG GCC GAC GAT CCT CCC GCC ACC

S1

Val Tyr Arg Tyr Asp Ser Arg Pro Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp
 121 GTA TAC CGC TAT GAC TCC CGC CCG CCG GAG GAC GTT TTC CAG AAC GGA TTC ACG GCG TGG

S1

Gly Asn Asn Asp Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser
 181 GGA AAC AAC GAC AAT GTG CTC GAC CAT CTG ACC GGA CGT TCC TGC CAG GTC GGC AGC AGC

S1

Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu Val Tyr Leu Glu His
 241 AAC AGC GCT TTC GTC TCC ACC AGC AGC AGC CCG CGC TAT ACC GAG GTC TAT CTC GAA CAT

S1

Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly Arg Gly Thr Gly His Phe Ile Gly
 301 CGC ATG CAG GAA GCG GTC GAG GCC GAA CCG GCC GGC AGG GGC ACC GGC CAC TTC ATC GGC

S1

Tyr Ile Tyr Glu Val Arg Ala Asp Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu
 361 TAC ATC TAC GAA GTC CGC GCC GAC AAC AAT TTC TAC GGC GCC GCC AGC TCG TAC TTC GAA

S1

Tyr Val Asp Thr Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr
 421 TAC GTC GAC ACT TAT GGC GAC AAT GCC GGC CGT ATC CTC GCC GGC GCG CTG GCC ACC TAC

S1

Gln Ser Glu Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg Arg Val Thr Arg
 481 CAG AGC GAA TAT CTG GCA CAC CCG CCG ATT CCG CCC GAA AAC ATC CGC AGG GTA ACG CCG

S1

Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr Thr Glu Tyr Ser Asn Ala Arg Tyr
 541 GTC TAT CAC AAC GGC ATC ACC GGC GAG ACC ACG ACC ACG GAG TAT TCC AAC GCT CGC TAC

도면13b

```

~~~~~
S1
~~~~~
601 Val Ser Gln Gln Thr Arg Ala Asn Pro Asn Pro Tyr Thr Ser Arg Arg Ser Val Ala Ser
    GTC AGC CAG CAG ACT CGC GCC AAT CCC AAC CCC TAC ACA TCG CGA AGG TCC GTA GCG TCG
    S1
~~~~~
661 Ile Val Gly Thr Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys Met Ala Arg Gln Ala
    ATC GTC GGC ACA TTG GTG CGC ATG GCG CGG GTG ATA GGC GCT TGC ATG GCG CGG CAG GCC
    S1
~~~~~
721 Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala Met Val Leu Val
    GAA AGC TCC GAG GCC ATG GCA GCC TGG TCC GAA CGC GCC GGC GAG GCG ATG GTT CTC GTG
    S1
~~~~~
781 Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe
    TAC TAC GAA AGC ATC GCG TAT TCG TTC TAG ACC TGG CCC AGC CCC GCC CAA CTC CGG TAA
    신호 서열
~~~~~
S2
~~~~~
841 Met Pro Ile Asp Arg Lys Thr Leu Cys His Leu Leu Ser Val Leu Pro Leu
    TTG AAC AGC ATG CCG ATC GAC CGC AAG ACG CTC TGC CAT CTC CTG TCC GTT CTG CCG TTG
    신호 서열
~~~~~
S2
~~~~~
901 Ala Leu Leu Gly Ser His Val Ala Arg Ala Ser Thr Pro Gly Ile Val Ile Pro Pro Gln
    GCC CTC CTC GGA TCT CAC GTG GCG CGG GCC TCC ACG CCA GGC ATC GTC ATT CCG CCG CAG
    S2
~~~~~
961 Glu Gln Ile Thr Gln His Gly Ser Pro Tyr Gly Arg Cys Ala Asn Lys Thr Arg Ala Leu
    GAA CAG ATT ACC CAG CAT GGC AGC CCC TAT GGA CGC TGC GCG AAC AAG ACC CGT GCC CTG
    S2
~~~~~
1021 Thr Val Ala Glu Leu Arg Gly Ser Gly Asp Leu Gln Glu Tyr Leu Arg His Val Thr Arg
    ACC CTG GCG GAA TTG GCG GGC AGC GGC GAT CTG CAG GAG TAC CTG CGT CAT GTG ACG CGC
    S2
~~~~~
1081 Gly Trp Ser Ile Phe Ala Leu Tyr Asp Gly Thr Tyr Leu Gly Gly Glu Tyr Gly Gly Val
    GGC TGG TCA ATA TTT GCG CTC TAC GAT GGC ACC TAT CTC GGC GGC GAA TAT GGC GGC GTG
    S2
~~~~~
1141 Ile Lys Asp Gly Thr Pro Gly Gly Ala Phe Asp Leu Lys Thr Thr Phe Cys Ile Met Thr
    ATC AAG GAC GGA ACA CCC GGC GGC GCA TTC GAC CTG AAA ACG ACG TTC TGC ATC ATG ACC
    S2
~~~~~
1201 Thr Arg Asn Thr Gly Gln Pro Ala Thr Asp His Tyr Tyr Ser Asn Val Thr Ala Thr Arg
    ACG CGC AAT ACG GGT CAA CCC GCA ACG GAT CAC TAC TAC AGC AAC GTC ACC GCC ACT CGC

```

도면13c

```

~~~~~
S2
~~~~~
1261 Leu Leu Ser Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val Arg Ser Gly Gln Pro Val
    CTG CTC TCC AGC ACC AAC AGC AGG CTA TGC GCG GTC TTC GTC AGA AGC GGG CAA CCG GTC
    S2
~~~~~
1321 Ile Gly Ala Cys Thr Ser Pro Tyr Asp Gly Lys Tyr Trp Ser Met Tyr Ser Arg Leu Arg
    ATT GGC GCC TGC ACC AGC CCG TAT GAC GGC AAG TAC TGG AGC ATG TAC AGC CGG CTG CGG
    S2
~~~~~
1381 Lys Met Leu Tyr Leu Ile Tyr Val Ala Gly Ile Ser Val Arg Val His Val Ser Lys Glu
    AAA ATG CTT TAC CTG ATC TAC GTG GCC GGC ATC TCC GTA CGC GTC CAT GTC AGC AAG GAA
    S2
~~~~~
1441 Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe Glu Thr Tyr Ala Leu Thr Gly Ile Ser
    GAA CAG TAT TAC GAC TAT GAG GAC GCA ACG TTC GAG ACT TAC GCC CTT ACC GGC ATC TCC
    S2
~~~~~
1501 Ile Cys Asn Pro Gly Ser Ser Leu Cys
    ATC TGC AAT CCT GGA TCA TCC TTA TGC TGA GAC GCT TCC CCA CTC GAA CCA CCG CCC CGG
    신호 서열
~~~~~
S4
~~~~~
1561 Val Arg Ala Leu Ala Trp Leu Leu Ala Ser Gly
    GAC AGG GCG GCG CCC GGC GGT CGC GC GTG CGC GCC CTG GCG TGG TTG CTG GCA TCC GGC
    신호 서열
~~~~~
S4
~~~~~
1620 Ala Met Thr His Leu Ser Pro Ala Leu Ala Asp Val Pro Tyr Val Leu Val Lys Thr Asn
    GCG ATG ACG CAT CTT TCC CCC GCC CTG GCC GAC GTT CCT TAT GTG CTG GTG AAG ACC AAT
    S4
~~~~~
1680 Met-Val Val Thr Ser Val Ala Met Lys Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val
    ATG GTG GTC ACC AGC GTA GCC ATG AAG CCG TAT GAA GTC ACC CCG ACG CGC ATG CTG GTC
    S4
~~~~~
1740 Cys Gly Ile Ala Ala Lys Leu Gly Ala Ala Ala Ser Ser Pro Asp Ala His Val Pro Phe
    TGC GGC ATC GCC GCC AAA CTG GGC GCG GCG GCC AGC AGC CCG GAC GCG CAC GTG CCG TTC
    S4
~~~~~
1800 Cys Phe Gly Lys Asp Leu Lys Arg Pro Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala
    TGC TTC GGC AAG GAT CTC AAG CGT CCC GCG AGC AGT CCC ATG GAA GTC ATG TTG CGC GCC
    S4
~~~~~
1860 Val Phe Met Gln Gln Arg Pro Leu Arg Met Phe Leu Gly Pro Lys Gln Leu Thr Phe Glu
    GTC TTC ATG CAA CAA CCG CCG CTG CGC ATG TTT CTG GGT CCC AAG CAA CTC ACT TTC GAA

```

도면13d

```

      S4
      ~~~~~
      Gly Lys Pro Ala Leu Glu Leu Ile Arg Met Val Glu Cys Ser Gly Lys Gln Asp Cys Pro
1920 GGC AAG CCC GCG CTC GAA CTG ATC CGG ATG GTC GAA TGC AGC GGC AAG CAG GAT TGC CCC T
      신호 서열
      ~~~~~
      S5
      ~~~~~
      Met His Thr Ile Ala Ser Ile Leu Leu Ser Val Leu Gly Ile Tyr Ser
1981 GAA GGC GAA CCC C ATG CAT ACC ATC GCA TCC ATC CTG TTG TCC GTG CTC GGC ATA TAC AG
      신호 서열
      ~~~~~
      S5
      ~~~~~
      Pro Ala Asp Val Ala Gly Leu Pro Thr His Leu Tyr Lys Asn Phe Thr Val Gln Glu Leu
2041 C CCG GCT GAC GTC GCC GGC TTG CCG ACC CAT CTG TAC AAG AAC TTC ACT GTC CAG GAG CT
      S5
      ~~~~~
      Ala Leu Lys Leu Lys Gly Lys Asn Gln Glu Phe Cys Leu Thr Ala Phe Met Ser Gly Arg
2101 G GCC TTG AAA CTG AAG GGC AAG AAT CAG GAG TTC TGC CTG ACC GCC TTC ATG TCG GGC AG
      S5
      ~~~~~
      Ser Leu Val Arg Ala Cys Leu Ser Asp Ala Gly His Glu His Asp Thr Trp Phe Asp Thr
2161 A AGC CTG GTC CCG GCG TGC CTG TCC GAC GCG GGA CAC GAG CAC GAC ACG TGG TTC GAC AC
      S5
      ~~~~~
      Met Leu Gly Phe Ala Ile Ser Ala Tyr Ala Leu Lys Ser Arg Ile Ala Leu Thr Val Glu
2221 C ATG CTT GGC TTT GCC ATA TCC GCG TAT GCG CTC AAG AGC CCG ATC GCG CTG ACG GTG GA
      S5
      ~~~~~
      Asp Ser Pro Tyr Pro Gly Thr Pro Gly Asp Leu Leu Glu Leu Gln Ile Cys Pro Leu Asn
2281 A GAC TCG CCG CAG TAT CCG GGC ACT CCC GGC GAT CTG CTC GAA CTG CAG ATC TGC CCG CTC AA
      S5
      ~~~~~
      Gly Tyr Cys Glu
2341 C GGA TAT TGC GAA TG AAC CCT TCC GGA GGT TTC GAC GTT TCC GCG CAA TCC GCT TGA GAC
      신호 서열
      ~~~~~
      S3
      ~~~~~
      Met Leu Ile Asn Asn Lys Lys Leu
2401 GAT CTT CCG CCC TGG TTC CAT TCC GGG AAC ACC GCA AC ATG CTG ATC AAC AAC AAG AAG C
      신호 서열
      ~~~~~
      S3
      ~~~~~
      Leu His His Ile Leu Pro Ile Leu Val Leu Ala Leu Leu Gly Met Arg Thr Ala Gln Ala
2461 TG CTT CAT CAC ATT CTG CCC ATC CTG GTG CTC GCC CTG CTG GGC ATG CCG ACG GCC CAG G
  
```

도면13e

신호 서열

```

      S3
      ~~~~~
      Val Ala Pro Gly Ile Val Ile Pro Pro Lys Ala Leu Phe Thr Gln Gln Gly Gly Ala Tyr
2521 CC GPT GCG CCA GGC ATC GTC ATC CCG CCG AAG GCA CTG TTC ACC CAA CAG GGC GGC GCC T
      S3
      ~~~~~
      Gly Arg Cys Pro Asn Gly Thr Arg Ala Leu Thr Val Ala Glu Leu Arg Gly Asn Ala Glu
2581 AT GGA GCG TGC CCG AAC GGA ACC GCG GGC ACC TTG ACC GTG GCC GAA CTG CCG GGC AAC GCC G
      S3
      ~~~~~
      Leu Gln Thr Tyr Leu Arg Gln Ile Thr Pro Gly Trp Ser Ile Tyr Gly Leu Tyr Asp Gly
2641 AA TTG CAG ACG TAT TTG CCG CAG ATA ACG CCC GGC TGG TCC ATA TAC GGT CTC TAT GAC G
      S3
      ~~~~~
      Thr Tyr Leu Gly Gln Ala Tyr Gly Gly Ile Ile Lys Asp Ala Pro Pro Gly Ala Gly Phe
2701 GT ACG TAC CTG GGC CAG GCG TAC GGC GGC ATC ATC AAG GAC GCG CCG CCA GGC GCG GGG T
      S3
      ~~~~~
      Ile Tyr Arg Glu Thr Phe Cys Ile Thr Thr Ile Tyr Lys Thr Gly Gln Pro Ala Ala Asp
2761 TC ATT TAT CCG GAA ACT TTC TGC ATC ACG ACC ATA TAC AAG ACC GGG CAA CCG GCT GCG G
      S3
      ~~~~~
      His Tyr Tyr Ser Lys Val Thr Ala Thr Arg Leu Leu Ala Ser Thr Asn Ser Arg Leu Cys
2821 AT CAC TAC TAC AGC AAG GTC ACG GCC ACG CCG CTG CTC GCC AGC ACC AAC AGC AGG CTG T
      S3
      ~~~~~
      Ala Val Phe Val Arg Asp Gly Gln Ser Val Ile Gly Ala Cys Ala Ser Pro Tyr Glu Gly
2881 GC GCG GTA TTC GTC AGG GAC GGG CAA TCG GTC ATC GGA GCC TGC GCC AGC CCG TAT GAA G
      S3
      ~~~~~
      Arg Tyr Arg Asp Met Tyr Asp Ala Leu Arg Arg Leu Leu Tyr Met Ile Tyr Met Ser Gly
2941 GC AGG TAC AGA GAC ATG TAC GAC GCG CTG CCG GCG CTG CTG TAC ATG ATC TAT ARG TCC G
      S3
      ~~~~~
      Leu Ala Val Arg Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr
3001 GC CTT GCC GTA CCG GTC CAC CTC AGC AAG GAG GAA CAG TAT TAC GAC TAC GAG GAC GCC A
      S3
      ~~~~~
      Phe Gln Thr Tyr Ala Leu Thr Gly Ile Ser Leu Cys Asn Pro Ala Ala Ser Ile Cys
3061 CA TTC CAG ACC TAT GCC CTC ACC GGC ATT TCC CTC TGC AAC CCG GCA GCG TCG ATA TGC
  
```

도면14a

A. 콜레라 독소 A 서브유닛 (AE003852; 단백질 ID AAF94614.1)
(분비 리더는 밑줄로 표시)

1 Met Val Lys Ile Ile Phe Val Phe Phe Ile Phe Leu Ser Ser
Phe Ser Tyr Ala Asn Asp

21 Asp Lys Leu Tyr Arg Ala Asp Ser Arg Pro Pro Asp Glu Ile
Lys Gln Ser Gly Gly Leu

41 Met Pro Arg Gly Gln Ser Glu Tyr Phe Asp Arg Gly Thr Gln
Met Asn Ile Asn Leu Tyr

61 Asp His Ala Arg Gly Thr Gln Thr Gly Phe Val Arg His Asp
Asp Gly Tyr Val Ser Thr

81 Ser Ile Ser Leu Arg Ser Ala His Leu Val Gly Gln Thr Ile
Leu Ser Gly His Ser Thr

101 Tyr Tyr Ile Tyr Val Ile Ala Thr Ala Pro Asn Met Phe Asn
Val Asn Asp Val Leu Gly

121 Ala Tyr Ser Pro His Pro Asp Glu Gln Glu Val Ser Ala Leu
Gly Gly Ile Pro Tyr Ser

141 Gln Ile Tyr Gly Trp Tyr Arg Val His Phe Gly Val Leu Asp
Glu Gln Leu His Arg Asn

161 Arg Gly Tyr Arg Asp Arg Tyr Tyr Ser Asn Leu Asp Ile Ala
Pro Ala Ala Asp Gly Tyr

181 Gly Leu Ala Gly Phe Pro Pro Glu His Arg Ala Trp Arg Glu
Glu Pro Trp Ile His His

201 Ala Pro Pro Gly Cys Gly Asn Ala Pro Arg Ser Ser Met Ser
Asn Thr Cys Asp Glu Lys

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

241 Phe Ser Gly Tyr Gln Ser Asp Ile Asp Thr His Asn Arg Ile
Lys Asp Glu Leu

도면14b

B. 콜레라 독소 B 아미노산 서열 (AE003852; 단백질 ID AAF94613.1)
(분비 리더는 밑줄로 표시)

1 Met Ile Lys Leu Lys Phe Gly Val Phe Phe Thr Val Leu Leu Ser
Ser Ala Tyr Ala His

21 Gly Thr Pro Gln Asn Ile Thr Asp Leu Cys Ala Glu Tyr His Asn
Thr Gln Ile Tyr Thr

41 Leu Asn Asp Lys Ile Phe Ser Tyr Thr Glu Ser Leu Ala Gly Lys
Arg Glu Met Ala Ile

61 Ile Thr Phe Lys Asn Gly Ala Ile Phe Gln Val Glu Val Pro Gly
Ser Gln His Ile Asp

81 Ser Gln Lys Lys Ala Ile Glu Arg Met Lys Asp Thr Leu Arg Ile
Ala Tyr Leu Thr Glu

101 Ala Lys Val Glu Lys Leu Cys Val Trp Asn Asn Lys Thr Pro His
Ala Ile Ala Ala Ile

121 Ser Met Ala Asn

도면14c

C. 콜레라 독소 DNA 서열 (A 및 B 서브유닛 코딩 영역 진뱅크 AE003852)

분비 리더

```

~~~~~
CTA
~~~~~
Met Val Lys Ile Ile Phe Val Phe Phe Ile Phe Leu Ser Ser Phe Ser Tyr Ala Asn Asp
1 ATG GTA AAG ATA ATA TTT GTG TTT TTT ATT TTC TTA TCA TCA TTT TCA TAT GCA AAT GAT
CTA
~~~~~
Asp Lys Leu Tyr Arg Ala Asp Ser Arg Pro Pro Asp Glu Ile Lys Gln Ser Gly Gly Leu
61 GAT AAG TTA TAT CGG GCA GAT TCT AGA CCT OCT GAT GAA ATA AAG CAG TCA GGT GGT CTT
CTA
~~~~~
Met Pro Arg Gly Gln Ser Glu Tyr Phe Asp Arg Gly Thr Gln Met Asn Ile Asn Leu Tyr
121 ATG CCA AGA GGA CAG ACT GAG TAC TTT GAC CGA GGT ACT CAA ATG AAT ATC AAC CTT TAT
CTA
~~~~~
Asp His Ala Arg Gly Thr Gln Thr Gly Phe Val Arg His Asp Asp Gly Tyr Val Ser Thr
181 GAT CAT GCA AGA GGA ACT CAG ACG GGA TTT GTT AGG CAC GAT GAT GGA TAT GTT TCC ACC
CTA
~~~~~
Ser Ile Ser Leu Arg Ser Ala His Leu Val Gly Gln Thr Ile Leu Ser Gly His Ser Thr
241 TCA ATT AGT TTG AGA AGT GCC CAC TTA GTG GGT CAA ACT ATA TTG TCT GGT CAT TCT ACT
CTA
~~~~~
Tyr Tyr Ile Tyr Val Ile Ala Thr Ala Pro Asn Met Phe Asn Val Asn Asp Val Leu Gly
301 TAT TAT ATA TAT GTT ATA GCC ACT GCA CCC AAC ATG TTT AAC GTT AAT GAT GTA TTA GGG
CTA
~~~~~
Ala Tyr Ser Pro His Pro Asp Glu Gln Glu Val Ser Ala Leu Gly Gly Ile Pro Tyr Ser
361 GCA TAC AGT CCT CAT CCA GAT GAA CAA GAA GTT TCT GCT TTA GGT GGT ATT CCA TAC TCC
CTA
~~~~~
Gln Ile Tyr Gly Trp Tyr Arg Val His Phe Gly Val Leu Asp Glu Gln Leu His Arg Asn
421 CAA ATA TAT GGA TGG TAT CGA GTT CAT TTT GGG GTG CTT GAT GAA CAA TTA CAT CGT AAT
CTA
~~~~~
Arg Gly Tyr Arg Asp Arg Tyr Tyr Ser Asn Leu Asp Ile Ala Pro Ala Ala Asp Gly Tyr
481 AGG GGC TAC AGA GAT AGA TAT TAC AGT AAC TTA GAT ATT GCT CCA GCA GCA GAT GGT TAT
CTA
~~~~~
Gly Leu Ala Gly Phe Pro Pro Glu His Arg Ala Trp Arg Glu Glu Pro Trp Ile His His
541 GGA TTG GCA GGT TTC CCT CCG GAG CAT AGA GCT TGG AGG GAA GAG CCG TGG ATT CAT CAT
CTA
~~~~~
Ala Pro Pro Gly Cys Gly Asn Ala Pro Arg Ser Ser Met Ser Asn Thr Cys Asp Glu Lys
601 GCA CCG CCG GGT TGT GGG AAT GCT CCA AGA TCA TCG ATG AGT AAT ACT TGC GAT GAA AAA

```

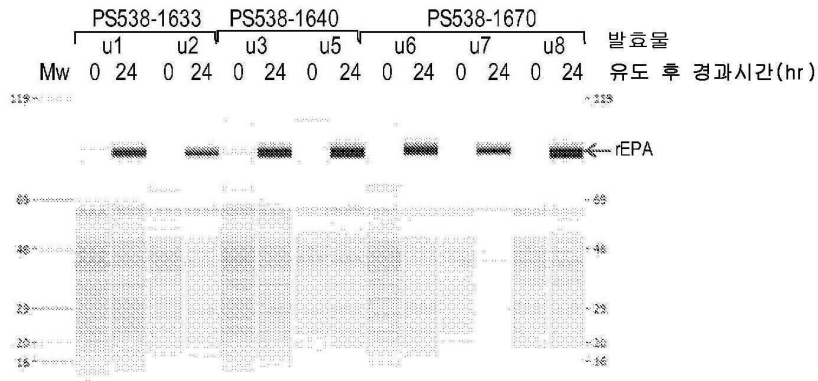
도면14d

```

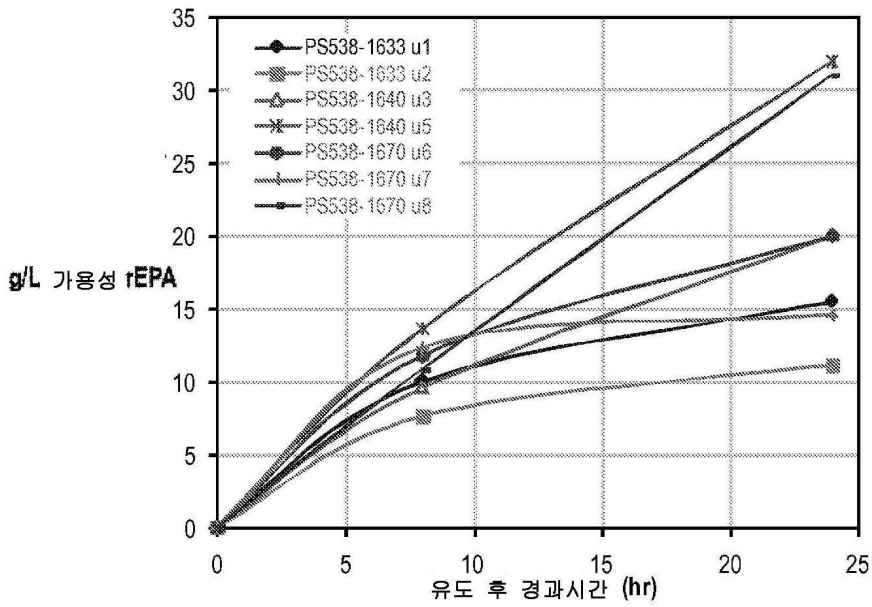
~~~~~
CTA
~~~~~
Thr Gln Ser Leu Gly Val Lys Phe Leu Asp Glu Tyr Gln Ser Lys Val Lys Arg Gln Ile
661 ACC CAA AGT CTA GGT GTA AAA TTC CTT GAC GAA TAC CAA TCT AAA GTT AAA AGA CAA ATA
CDS_1
~~~~~
CTA
~~~~~
~~~~~
분비 리더
~~~~~
CTB
~~~~~
Phe Ser Gly Tyr Gln Ser Asp Ile Asp Thr His Asn Arg Ile Lys Asp Glu Leu ***
721 TTT TCA GGC TAT CAA TCT GAT ATT GAT ACA CAT AAT AGA ATT AAG GAT GAA TTA TGA TTA
분비 리더
~~~~~
CTB
~~~~~
LLeu Lys Phe Gly Val Phe Phe Thr Val Leu Leu Ser Ser Ala Tyr Ala His Gly Thr Pro
781 AAT TAA AAT TTG GTG TTT TTT TTA CAG TTT TAC TAT CTT CAG CAT ATG CAC ATG GAA CAC
CTB
~~~~~
PGLn Asn Ile Thr Asp Leu Cys Ala Glu Tyr His Asn Thr Gln Ile Tyr Thr Leu Asn Asp
841 CTC AAA ATA TTA CTG ATT TGT GTG CAG AAT ACC ACA ACA CAC AAA TAT ATA CGC TAA ATG
CTB
~~~~~
ALys Ile Phe Ser Tyr Thr Glu Ser Leu Ala Gly Lys Arg Glu Met Ala Ile Ile Thr Phe
901 ATA AGA TAT TTT CGT ATA CAG AAT CTC TAG CTG GAA AAA GAG AGA TGG CTA TCA TTA CTT
CTB
~~~~~
PLys Asn Gly Ala Ile Phe Gln Val Glu Val Pro Gly Ser Gln His Ile Asp Ser Gln Lys
961 TTA AGA ATG GTG CAA TTT TTC AAG TAG AAG TAC CAG GTA GTC AAC ATA TAG ATT CAC AAA
CTB
~~~~~
LLys Ala Ile Glu Arg Met Lys Asp Thr Leu Arg Ile Ala Tyr Leu Thr Glu Ala Lys Val
1021 AAA AAG CGA TTG AAA GGA TGA AGG ATA CCC TGA GGA TTG CAT ATC TTA CTG AAG CTA AAG
CTB
~~~~~
VGlu Lys Leu Cys Val Trp Asn Asn Lys Thr Pro His Ala Ile Ala Ala Ile Ser Met Ala
1081 TCG AAA AGT TAT GTG TAT GGA ATA ATA AAA CGC CTC ATG CGA TTG CCG CAA TTA GTA TGG
CTB
~~~~~
AAsn
1141 CAA ATT AA

```

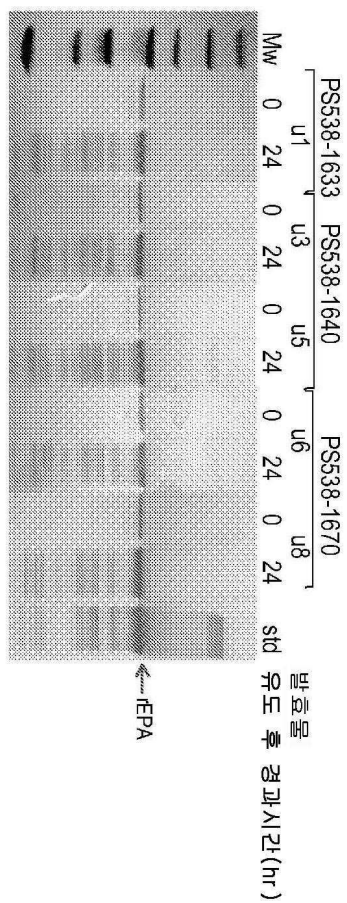
도면15



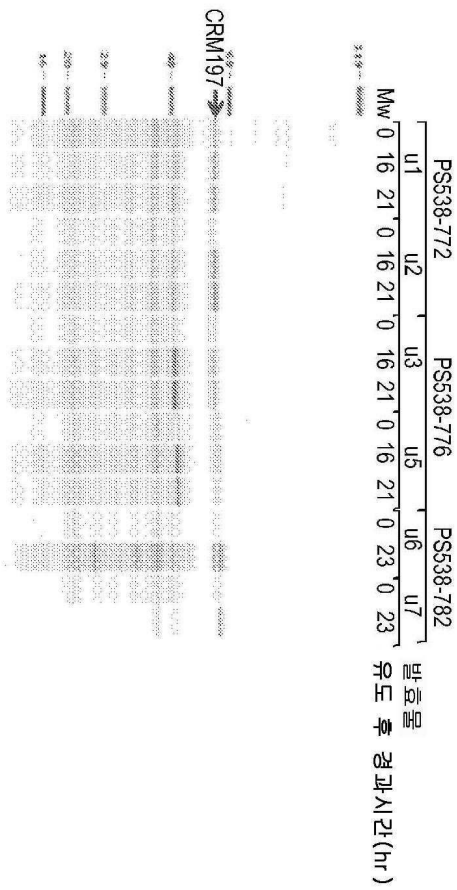
도면16



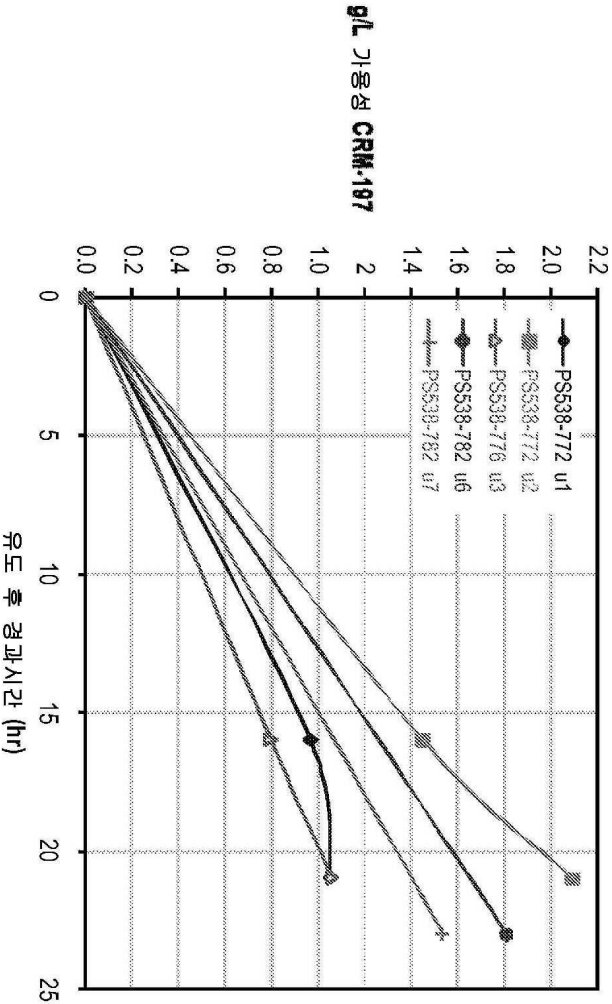
도면17



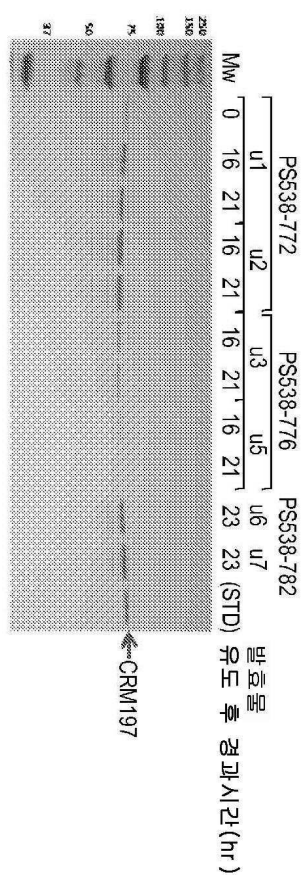
도면18



도면19



도면20



서열목록

SEQUENCE LISTING

<110> RETALLACK, DIANE M.

CHEW, LAWRENCE

JIN, HONGFAN

TALBOT, HENRY W.

<120> HIGH LEVEL EXPRESSION OF RECOMBINANT TOXIN PROTEINS

<130> 38194-735.201

<140><141><150> 61/325,235

<151> 2010-04-16

<150> PCT/US10/30573

<151> 2010-04-09

<150> 61/319,152

<151> 2010-03-30

<160> 46

<170> PatentIn version 3.5

<210> 1

<211> 535

<212> PRT

<213> *Corynebacterium diphtheriae*

<400> 1

Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn

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

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

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

195 200 205

Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly
210 215 220

Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu
225 230 235 240

Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu
245 250 255

His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val

260 265 270

Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val
275 280 285

Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu
290 295 300

Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala
305 310 315 320

Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser

325 330 335

Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp
340 345 350

Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe
355 360 365

Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His
370 375 380

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

385 390 395 400

Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser Gly His
405 410 415

Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala Gly Val
420 425 430

Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser Lys Thr
435 440 445

His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg Ala Ile

450 Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr Val Gly
 465 Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser Ser Ser
 485 Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly Val Leu
 500 Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys Leu Ser

515 Leu Phe Phe Glu Ile Lys Ser
 530
 <210> 2
 <211> 1605
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<220><221> CDS
 <222> (1)..(1605)
 <400> 2

ggg gcg gac gat gtg gtg gat tcc tcc aag tcg ttt gtc atg gaa aat 48
 Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn
 1 5 10 15

ttc tcg tcg tac cat ggc act aag cca ggc tac gtg gat agc att caa 96
 Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser Ile Gln
 20 25 30

aag ggc atc cag aag ccc aag agc ggt act cag ggg aac tat gac gac 144
 Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn Tyr Asp Asp
 35 40 45

gac tgg aag gaa ttt tac agc acc gac aat aag tac gat gct gcg ggc 192

Asp Trp Lys Glu Phe Tyr Ser Thr Asp Asn Lys Tyr Asp Ala Ala Gly
 50 55 60

tat	agc	gtg	gac	aac	gaa	aac	cca	ttg	tcg	ggc	aag	gcc	ggt	ggc	gtg	240
Tyr	Ser	Val	Asp	Asn	Glu	Asn	Pro	Leu	Ser	Gly	Lys	Ala	Gly	Gly	Val	
65			70			75			80							
gtg	aag	gtg	acc	tat	cct	ggg	ctg	acg	aaa	gtt	ctg	gcg	ttg	aaa	gtg	288
Val	Lys	Val	Thr	Tyr	Pro	Gly	Leu	Thr	Lys	Val	Leu	Ala	Leu	Lys	Val	
85			90			95										
gac	aac	gcc	gag	act	atc	aag	aaa	gaa	ttg	ggc	ttg	agt	ttg	acc	gag	336
Asp	Asn	Ala	Glu	Thr	Ile	Lys	Lys	Glu	Leu	Gly	Leu	Ser	Leu	Thr	Glu	
100			105			110										
ccg	ctg	atg	gaa	cag	gtg	ggg	acc	gaa	gaa	ttc	att	aaa	cgt	ttt	ggg	384
Pro	Leu	Met	Glu	Gln	Val	Gly	Thr	Glu	Glu	Phe	Ile	Lys	Arg	Phe	Gly	
115			120			125										
gac	ggc	gcg	tcg	cgc	gtg	gtc	ctg	tcg	ttg	ccg	ttc	gcc	gaa	ggg	tcg	432
Asp	Gly	Ala	Ser	Arg	Val	Val	Leu	Ser	Leu	Pro	Phe	Ala	Glu	Gly	Ser	
130			135			140										
tcg	tcg	gtg	gaa	tat	atc	aac	aac	tgg	gaa	cag	gcc	aag	gcg	ctg	tcg	480
Ser	Ser	Val	Glu	Tyr	Ile	Asn	Asn	Trp	Glu	Gln	Ala	Lys	Ala	Leu	Ser	
145			150			155			160							
gtg	gaa	ctg	gaa	att	aac	ttc	gaa	acg	cgg	ggc	aaa	cgg	ggc	cag	gac	528
Val	Glu	Leu	Glu	Ile	Asn	Phe	Glu	Thr	Arg	Gly	Lys	Arg	Gly	Gln	Asp	
165			170			175										
gcc	atg	tac	gaa	tac	atg	gcg	cag	gcg	tgc	gcc	ggg	aac	cgg	gtg	cgg	576
Ala	Met	Tyr	Glu	Tyr	Met	Ala	Gln	Ala	Cys	Ala	Gly	Asn	Arg	Val	Arg	
180			185			190										
cgc	agc	gtg	ggc	agt	tcc	ttg	agt	tgc	atc	aat	ctg	gac	tgg	gac	gtc	624
Arg	Ser	Val	Gly	Ser	Ser	Leu	Ser	Cys	Ile	Asn	Leu	Asp	Trp	Asp	Val	
195			200			205										
atc	cgc	gat	aag	acg	aag	acg	aaa	atc	gag	tcg	ctc	aaa	gag	cac	ggc	672
Ile	Arg	Asp	Lys	Thr	Lys	Thr	Lys	Ile	Glu	Ser	Leu	Lys	Glu	His	Gly	
210			215			220										
ccg	atc	aaa	aac	aaa	atg	agc	gag	tcg	ccg	aat	aaa	acg	gtg	tcc	gag	720

Pro	Ile	Lys	Asn	Lys	Met	Ser	Glu	Ser	Pro	Asn	Lys	Thr	Val	Ser	Glu	
225			230			235			240							
gag	aag	gcg	aag	caa	tac	ctg	gag	gaa	ttc	cac	cag	acg	gct	ctg	gag	768
Glu	Lys	Ala	Lys	Gln	Tyr	Leu	Glu	Glu	Phe	His	Gln	Thr	Ala	Leu	Glu	
245			250			255										
cac	ccg	gag	ctg	agc	gaa	ctc	aaa	acc	gtt	acc	ggc	acg	aac	ccg	gtg	816
His	Pro	Glu	Leu	Ser	Glu	Leu	Lys	Thr	Val	Thr	Gly	Thr	Asn	Pro	Val	
260			265			270										
ttt	gcc	ggg	gca	aac	tat	gca	gct	tgg	gcc	gtc	aac	gtg	gcc	caa	gtg	864
Phe	Ala	Gly	Ala	Asn	Tyr	Ala	Ala	Trp	Ala	Val	Asn	Val	Ala	Gln	Val	
275			280			285										
atc	gac	tcc	gaa	acg	gcc	gac	aac	ctg	gaa	aag	act	acc	gcc	gcg	ttg	912
Ile	Asp	Ser	Glu	Thr	Ala	Asp	Asn	Leu	Glu	Lys	Thr	Thr	Ala	Ala	Leu	
290			295			300										
tcg	atc	ctc	ccg	ggc	atc	ggg	agc	gtc	atg	ggc	att	gcc	gat	ggc	gcg	960
Ser	Ile	Leu	Pro	Gly	Ile	Gly	Ser	Val	Met	Gly	Ile	Ala	Asp	Gly	Ala	
305			310			315			320							
gtg	cat	cac	aac	acc	gaa	gag	atc	gtc	gcg	cag	tcg	atc	gca	ttg	tcc	1008
Val	His	His	Asn	Thr	Glu	Glu	Ile	Val	Ala	Gln	Ser	Ile	Ala	Leu	Ser	
325			330			335										
tcc	ctg	atg	gtc	gcc	caa	gct	atc	ccg	ctg	gtc	ggc	gag	ctg	gtc	gat	1056
Ser	Leu	Met	Val	Ala	Gln	Ala	Ile	Pro	Leu	Val	Gly	Glu	Leu	Val	Asp	
340			345			350										
atc	ggc	ttt	gcc	gct	tat	aac	ttt	gtt	gaa	tcg	atc	att	aac	ctg	ttc	1104
Ile	Gly	Phe	Ala	Ala	Tyr	Asn	Phe	Val	Glu	Ser	Ile	Ile	Asn	Leu	Phe	
355			360			365										
cag	gtg	gtg	cat	aac	agc	tac	aac	cgg	cca	gcg	tac	tcg	ccc	ggc	cac	1152
Gln	Val	Val	His	Asn	Ser	Tyr	Asn	Arg	Pro	Ala	Tyr	Ser	Pro	Gly	His	
370			375			380										

aag acc cag ccc ttt ctc cac gac ggc tat gcc gtg tcg tgg aac acc	1200
Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr	
385 390 395 400	
gtg gag gac agc atc atc cgc acc ggt ttc cag ggc gag agc ggc cat	1248
Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser Gly His	
405 410 415	
gac atc aaa att acc gcg gaa aac acg ccc ttg ccg atc gct ggc gtg	1296
Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala Gly Val	
420 425 430	
ttg ctc ccg acg atc ccg ggt aag ctc gac gtc aac aag tcc aag acc	1344
Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser Lys Thr	
435 440 445	
cat atc agc gtc aat ggc cgc aag atc cgc atg cgc tgt cgg gcc att	1392
His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg Ala Ile	
450 455 460	
gat ggc gac gtc acg ttt tgc cgg ccg aag agt ccc gtc tat gtc ggg	1440
Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr Val Gly	
465 470 475 480	
aac ggt gtc cat gcc aac ctg cac gtc gca ttc cac cgg agc agc tcg	1488
Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser Ser Ser	
485 490 495	
gaa aag atc cac agc aat gag atc agc agc gac agc atc ggc gtg ctg	1536
Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly Val Leu	
500 505 510	
ggg tat caa aag acg gtc gat cac acc aag gtg aac agt aaa ctg agc	1584
Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys Leu Ser	
515 520 525	
ttg ttc ttt gaa atc aag tcg	1605
Leu Phe Phe Glu Ile Lys Ser	

530

535

<210> 3

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 3

Met Arg Asn Leu Ile Leu Ser Ala Ala Leu Val Thr Ala Ser Leu Phe

1 5 10 15

Gly Met Thr Ala Gln Ala

20

<210> 4

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 4

Met Phe Ala Lys Leu Val Ala Val Ser Leu Leu Thr Leu Ala Ser Gly

1 5 10 15

Gln Leu Leu Ala

20

<210> 5

<211> 31

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 5

Met Ile Arg Asp Asn Arg Leu Lys Thr Ser Leu Leu Arg Gly Leu Thr

1 5 10 15

Leu Thr Leu Leu Ser Leu Thr Leu Leu Ser Pro Ala Ala His Ala

20 25 30

<210> 6

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 6

Met Asn Arg Ser Ser Ala Leu Leu Leu Ala Phe Val Phe Leu Ser Gly

1 5 10 15

Cys Gln Ala Met Ala

20

<210> 7

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 7

Met Leu Phe Arg Thr Leu Leu Ala Ser Leu Thr Phe Ala Val Ile Ala

1 5 10 15

Gly Leu Pro Ser Thr Ala His Ala

20

<210> 8

<211> 25

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 8

Met Ser Cys Thr Arg Ala Phe Lys Pro Leu Leu Leu Ile Gly Leu Ala

1 5 10 15

Thr Leu Met Cys Ser His Ala Phe Ala

20

25

<210> 9

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 9

Met Arg Leu Ala Ala Leu Pro Leu Leu Leu Ala Pro Leu Phe Ile Ala

1 5 10 15

Pro Met Ala Val Ala

20

<210> 10

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 10

Met Lys Leu Lys Arg Leu Met Ala Ala Met Thr Phe Val Ala Ala Gly

1 5 10 15

Val Ala Thr Val Asn Ala Val Ala

20

<210> 11

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 11

Met Arg Leu Thr Gln Ile Ile Ala Ala Ala Ala Ile Ala Leu Val Ser

1 5 10 15

Thr Phe Ala Leu Ala

20

<210> 12

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 12

Met Arg Asn Leu Leu Arg Gly Met Leu Val Val Ile Cys Cys Met Ala

1 5 10 15

Gly Ile Ala Ala Ala

20

<210> 13

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 13

Met Lys Leu Lys Arg Leu Met Ala Ala Met Thr Phe Val Ala Ala Gly

1 5 10 15

Val Ala Thr Ala Asn Ala Val Ala

20

<210> 14

<211> 23

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 14

Met Gln Asn Tyr Lys Lys Phe Leu Leu Ala Ala Ala Val Ser Met Ala

1 5 10 15

Phe Ser Ala Thr Ala Met Ala

20

<210> 15

<211> 23

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 15

Met Pro Pro Arg Ser Ile Ala Ala Cys Leu Gly Leu Leu Gly Leu Leu

1 5 10 15

Met Ala Thr Gln Ala Ala Ala

20

<210> 16

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 16

Met Lys Lys Ser Thr Leu Ala Val Ala Val Thr Leu Gly Ala Ile Ala

1 5 10 15

Gln Gln Ala Gly Ala

20

<210> 17

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 17

Met Lys Leu Lys Arg Leu Met Ala Ala Met Thr Phe Val Ala Ala Gly

1 5 10 15

Val Ala Thr Ala Asn Ala Val Ala

20

<210> 18

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400>

> 18

Met Lys Phe Lys Gln Leu Met Ala Met Ala Leu Leu Leu Ala Leu Ser

1 5 10 15

Ala Val Ala Gln Ala

20

<210> 19

<211> 33

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 19

Met Gln Asn Arg Thr Val Glu Ile Gly Val Gly Leu Phe Leu Leu Ala

1 5 10 15

Gly Ile Leu Ala Leu Leu Leu Leu Ala Leu Arg Val Ser Gly Leu Ser

20

25

30

Ala

<210> 20

<211> 25

<212> PRT

<213> Corynebacterium diphtheriae

<220><221> MOD_RES

<222> (9)..(9)

<223> Any amino acid

<400> 20

Met Ser Arg Lys Leu Phe Ala Ser Xaa Leu Ile Gly Ala Leu Leu Gly

1 5 10 15

Ile Gly Ala Pro Pro Ser Ala His Ala

20 25

<210> 21

<211> 4

<212> PRT

<213> Corynebacterium diphtheriae

<400> 21

Gly Ala Asp Asp

1

<210> 22

<211> 102

<212> PRT

<213> Vibrio cholerae

<400> 22

Thr Pro Gln Asn Ile Thr Asp Leu Cys Ala Glu Tyr His Asn Thr Gln

1 5 10 15

Ile His Thr Leu Asn Asp Lys Ile Phe Ser Tyr Thr Glu Ser Leu Ala

20 25 30

Gly Lys Arg Glu Met Ala Ile Ile Thr Phe Lys Asn Gly Ala Thr Phe

35 40 45

Gln Val Glu Val Pro Gly Ser Gln His Ile Asp Ser Gln Lys Lys Ala

50 55 60

Ile Glu Arg Met Lys Asp Thr Leu Arg Ile Ala Tyr Leu Thr Glu Ala

65 70 75 80

Lys Val Glu Lys Leu Cys Val Trp Asn Asn Lys Thr Pro His Ala Ile

85 90 95

Ala Ala Ile Ser Met Ala

100

<210> 23

<211> 306

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<220><221> CDS

<222> (1)..(306)

<400> 23

acg ccg caa aat atc acc gac ctg tgc gca gaa tat cac aat acc caa 48

Thr Pro Gln Asn Ile Thr Asp Leu Cys Ala Glu Tyr His Asn Thr Gln

1 5 10 15

atc cat act ctg aac gac aaa atc ttc agc tac acc gag agc ctg gct 96

Ile His Thr Leu Asn Asp Lys Ile Phe Ser Tyr Thr Glu Ser Leu Ala

20 25 30

ggc aag cgc gag atg gcg atc att acg ttc aaa aac ggt gcg acc ttt 144

Gly Lys Arg Glu Met Ala Ile Ile Thr Phe Lys Asn Gly Ala Thr Phe

35 40 45

cag gtg gaa gtc ccc ggc agt cag cac atc gat tcc cag aaa aag gcc 192

Gln Val Glu Val Pro Gly Ser Gln His Ile Asp Ser Gln Lys Lys Ala

50 55 60

att gaa cgg atg aag gac acc ctc cgt atc gcc tac ttg acc gaa gcc 240

Ile Glu Arg Met Lys Asp Thr Leu Arg Ile Ala Tyr Leu Thr Glu Ala

65 70 75 80

aag gtg gag aag ctg tgc gtt tgg aac aac aaa acc ccg cac gcc atc 288

Lys Val Glu Lys Leu Cys Val Trp Asn Asn Lys Thr Pro His Ala Ile

85 90 95

gcg gcc atc tcg atg gcc 306

Ala Ala Ile Ser Met Ala

100

<210> 24

<211> 3119

<212> DNA

<213> Bordetella pertussis

<220><221> CDS

<222> (1)..(807)

<220><221> CDS

<222> (850)..(1527)

<220><221> CDS

<222> (1587)..(1979)

<220><221> CDS

<222> (1994)..(2353)

<220><221> CDS

<222> (2439)..(3119)

<400> 24

atg cgt tgc act cgg gca att cgc caa acc gca aga aca ggc tgg ctg 48

Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp Leu

1 5 10 15

acg tgg ctg gcg att ctt gcc gtc acg gcg ccc gtg act tcg ccg gca 96

Thr Trp Leu Ala Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala

20 25 30

tgg gcc gac gat cct ccc gcc acc gta tac aaa tat gac tcc cgc ccg 144

Trp Ala Asp Asp Pro Pro Ala Thr Val Tyr Lys Tyr Asp Ser Arg Pro

35 40 45

ccg gag gac gtt ttc cag aac gga ttc acg gcg tgg gga aac aac gac 192

Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp Gly Asn Asn Asp

50 55 60

aat gtg ctc gac cat ctg acc gga cgt tcc tgc cag gtc ggc agc agc 240

Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser

65 70 75 80

aac agc gct ttc gtc tcc acc agc agc agc cgg cgc tat acc gag gtc 288

Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu Val

85 90 95

tat ctc gaa cat cgc atg cag gaa gcg gtc gag gcc gaa cgc gcc ggc 336

Tyr Leu Glu His Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly

100 105 110

agg ggc acc ggc cac ttc atc ggc tac atc tac gaa gtc cgc gcc gac 384

Arg Gly Thr Gly His Phe Ile Gly Tyr Ile Tyr Glu Val Arg Ala Asp	
115 120 125	
aac aat ttc tac ggc gcc gcc agc tcg tac ttc gaa tac gtc gac act	432
Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu Tyr Val Asp Thr	
130 135 140	
tat ggc gac aat gcc ggc cgt atc ctc gcc ggc gcg ctg gcc acc tac	480
Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr	
145 150 155 160	
cag agc gca tat ctg gca cac cgg cgc att ccg ccc gaa aac atc cgc	528
Gln Ser Ala Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg	
165 170 175	
agg gta acg cgg gtc tat cac aac ggc atc acc ggc gag acc acg acc	576
Arg Val Thr Arg Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr	
180 185 190	
acg gag tat tcc aac gct cgc tac gtc agc cag cag act cgc gcc aat	624
Thr Glu Tyr Ser Asn Ala Arg Tyr Val Ser Gln Gln Thr Arg Ala Asn	
195 200 205	
ccc aac ccc tac aca tcg cga agg tcc gta gcg tcg atc gtc ggc aca	672
Pro Asn Pro Tyr Thr Ser Arg Arg Ser Val Ala Ser Ile Val Gly Thr	
210 215 220	
ttg gtg cgc atg gcg ccg gtg ata ggc gct tgc atg gcg cgg cag gcc	720
Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys Met Ala Arg Gln Ala	
225 230 235 240	
gaa agc tcc gag gcc atg gca gcc tgg tcc gaa cgc gcc ggc gag gcg	768
Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala	
245 250 255	
atg gtt ctc gtg tac tac gaa agc atc gcg tat tcg ttc tagacctggc	817
Met Val Leu Val Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe	
260 265	
ccagccccgc ccaactccgg taattgaaca gc atg ccg atc gac cgc aag acg	870

Met Pro Ile Asp Arg Lys Thr

	270	275	
ctc tgc cat ctc ctg tcc gtt ctg ccg ttg gcc ctc ctc gga tct cac			918
Leu Cys His Leu Leu Ser Val Leu Pro Leu Ala Leu Leu Gly Ser His			
280	285	290	
gtg gcg cgg gcc tcc acg cca ggc atc gtc att ccg ccg cag gaa cag			966
Val Ala Arg Ala Ser Thr Pro Gly Ile Val Ile Pro Pro Gln Glu Gln			
295	300	305	
att acc cag cat ggc agc ccc tat gga cgc tgc gcg aac aag acc cgt			1014
Ile Thr Gln His Gly Ser Pro Tyr Gly Arg Cys Ala Asn Lys Thr Arg			
310	315	320	
gcc ctg acc gtg gcg gaa ttg cgc ggc agc ggc gat ctg cag gag tac			1062
Ala Leu Thr Val Ala Glu Leu Arg Gly Ser Gly Asp Leu Gln Glu Tyr			
325	330	335	340
ctg cgt cat gtg acg cgc ggc tgg tca ata ttt gcg ctc tac gat ggc			1110
Leu Arg His Val Thr Arg Gly Trp Ser Ile Phe Ala Leu Tyr Asp Gly			
345	350	355	
acc tat ctc ggc ggc gaa tat ggc ggc gtg atc aag gac gga aca ccc			1158
Thr Tyr Leu Gly Gly Glu Tyr Gly Gly Val Ile Lys Asp Gly Thr Pro			
360	365	370	
ggc ggc gca ttc gac ctg aaa acg acg ttc tgc atc atg acc acg cgc			1206
Gly Gly Ala Phe Asp Leu Lys Thr Thr Phe Cys Ile Met Thr Thr Arg			
375	380	385	
aat acg ggt caa ccc gca acg gat cac tac tac agc aac gtc acc gcc			1254
Asn Thr Gly Gln Pro Ala Thr Asp His Tyr Tyr Ser Asn Val Thr Ala			
390	395	400	
act cgc ctg ctc tcc agc acc aac agc agg cta tgc gcg gtc ttc gtc			1302
Thr Arg Leu Leu Ser Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val			
405	410	415	420

aga agc ggg caa ccg gtc att ggc gcc tgc acc agc ccg tat gac ggc 1350
 Arg Ser Gly Gln Pro Val Ile Gly Ala Cys Thr Ser Pro Tyr Asp Gly
 425 430 435

aag tac tgg agc atg tac agc cgg ctg cgg aaa atg ctt tac ctg atc 1398
 Lys Tyr Trp Ser Met Tyr Ser Arg Leu Arg Lys Met Leu Tyr Leu Ile
 440 445 450

tac gtg gcc ggc atc tcc gta cgc gtc cat gtc agc aag gaa gaa cag 1446

Tyr Val Ala Gly Ile Ser Val Arg Val His Val Ser Lys Glu Glu Gln
 455 460 465

tat tac gac tat gag gac gca acg ttc gag act tac gcc ctt acc ggc 1494
 Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe Glu Thr Tyr Ala Leu Thr Gly
 470 475 480

atc tcc atc tgc aat cct gga tca tcc tta tgc tgagacgctt cccactcga 1547
 Ile Ser Ile Cys Asn Pro Gly Ser Ser Leu Cys

485 490 495

accaccgccc cgggacaggg cggcgcccg cggtcgcgc gtg cgc gcc ctg gcg 1601
 Val Arg Ala Leu Ala
 500

tgg ttg ctg gca tcc ggc gcg atg acg cat ctt tcc ccc gcc ctg gcc 1649
 Trp Leu Leu Ala Ser Gly Ala Met Thr His Leu Ser Pro Ala Leu Ala
 505 510 515

gac gtt cct tat gtg ctg gtg aag acc aat atg gtg gtc acc agc gta 1697
 Asp Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val Thr Ser Val
 520 525 530

gcc atg aag ccg tat gaa gtc acc ccg acg cgc atg ctg gtc tgc ggc 1745
 Ala Met Lys Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val Cys Gly
 535 540 545

atc gcc gcc aaa ctg ggc gcc gcg gcc agc agc ccg gac gcg cac gtg 1793

Ile Ala Ala Lys Leu Gly Ala Ala Ala Ser Ser Pro Asp Ala His Val
 550 555 560

ccg ttc tgc ttc ggc aag gat ctc aag cgt ccc ggc agc agt ccc atg 1841

Pro	Phe	Cys	Phe	Gly	Lys	Asp	Leu	Lys	Arg	Pro	Gly	Ser	Ser	Pro	Met	
565				570				575				580				
gaa gtc atg ttg cgc gcc gtc ttc atg caa caa cgg ccg ctg cgc atg																1889
Glu	Val	Met	Leu	Arg	Ala	Val	Phe	Met	Gln	Gln	Arg	Pro	Leu	Arg	Met	
585				590				595								
ttt ctg ggt ccc aag caa ctc act ttc gaa ggc aag ccc gcg ctc gaa																1937
Phe	Leu	Gly	Pro	Lys	Gln	Leu	Thr	Phe	Glu	Gly	Lys	Pro	Ala	Leu	Glu	
600				605				610								
ctg atc cgg atg gtc gaa tgc agc ggc aag cag gat tgc ccc																1979
Leu	Ile	Arg	Met	Val	Glu	Cys	Ser	Gly	Lys	Gln	Asp	Cys	Pro			
615				620				625								
tgaaggcgaa cccc atg cat acc atc gca tcc atc ctg ttg tcc gtg ctc																2029
				Met	His	Thr	Ile	Ala	Ser	Ile	Leu	Leu	Ser	Val	Leu	
				630				635								
ggc ata tac agc ccg gct gac gtc gcc ggc ttg ccg acc cat ctg tac																2077
Gly	Ile	Tyr	Ser	Pro	Ala	Asp	Val	Ala	Gly	Leu	Pro	Thr	His	Leu	Tyr	
640				645				650								
aag aac ttc act gtc cag gag ctg gcc ttg aaa ctg aag ggc aag aat																2125
Lys	Asn	Phe	Thr	Val	Gln	Glu	Leu	Ala	Leu	Lys	Leu	Lys	Gly	Lys	Asn	
655				660				665				670				
cag gag ttc tgc ctg acc gcc ttc atg tcg ggc aga agc ctg gtc cgg																2173
Gln	Glu	Phe	Cys	Leu	Thr	Ala	Phe	Met	Ser	Gly	Arg	Ser	Leu	Val	Arg	
675				680				685								
gcg tgc ctg tcc gac gcg gga cac gag cac gac acg tgg ttc gac acc																2221
Ala	Cys	Leu	Ser	Asp	Ala	Gly	His	Glu	His	Asp	Thr	Trp	Phe	Asp	Thr	
690				695				700								
atg ctt ggc ttt gcc ata tcc gcg tat gcg ctc aag agc cgg atc gcg																2269
Met	Leu	Gly	Phe	Ala	Ile	Ser	Ala	Tyr	Ala	Leu	Lys	Ser	Arg	Ile	Ala	
705				710				715								
ctg acg gtg gaa gac tcg ccg tat ccg ggc act ccc ggc gat ctg ctc																2317
Leu	Thr	Val	Glu	Asp	Ser	Pro	Tyr	Pro	Gly	Thr	Pro	Gly	Asp	Leu	Leu	

720	725	730	
gaa ctg cag atc tgc ccg ctc aac gga tat tgc gaa tgaacccttc			2363
Glu Leu Gln Ile Cys Pro Leu Asn Gly Tyr Cys Glu			
735	740	745	
cggagggtttc gacgtttccg cgcaatccgc ttgagacgat cttccgcctt ggttcattc			2423
cgggaacacc gcaac atg ctg atc aac aac aag aag ctg ctt cat cac att			2474
Met Leu Ile Asn Asn Lys Lys Leu Leu His His Ile			
	750	755	
ctg ccc atc ctg gtg ctc gcc ctg ctg ggc atg cgc acg gcc cag gcc			2522
Leu Pro Ile Leu Val Leu Ala Leu Leu Gly Met Arg Thr Ala Gln Ala			
760	765	770	
gtt gcg cca ggc atc gtc atc ccg ccg aag gca ctg ttc acc caa cag			2570
Val Ala Pro Gly Ile Val Ile Pro Pro Lys Ala Leu Phe Thr Gln Gln			
775	780	785	790
ggc ggc gcc tat gga cgc tgc ccg aac gga acc cgc gcc ttg acc gtg			2618
Gly Gly Ala Tyr Gly Arg Cys Pro Asn Gly Thr Arg Ala Leu Thr Val			
	795	800	805
gcc gaa ctg cgc ggc aac gcc gaa ttg cag acg tat ttg cgc cag ata			2666
Ala Glu Leu Arg Gly Asn Ala Glu Leu Gln Thr Tyr Leu Arg Gln Ile			
	810	815	820
acg ccc ggc tgg tcc ata tac ggt ctc tat gac ggt acg tac ctg ggc			2714
Thr Pro Gly Trp Ser Ile Tyr Gly Leu Tyr Asp Gly Thr Tyr Leu Gly			
	825	830	835
cag gcg tac ggc ggc atc atc aag gac gcg ccg cca ggc gcg ggg ttc			2762
Gln Ala Tyr Gly Gly Ile Ile Lys Asp Ala Pro Pro Gly Ala Gly Phe			
	840	845	850
att tat cgc gaa act ttc tgc atc acg acc ata tac aag acc ggg caa			2810
Ile Tyr Arg Glu Thr Phe Cys Ile Thr Thr Ile Tyr Lys Thr Gly Gln			
855	860	865	870

ccg gct gcg gat cac tac tac agc aag gtc acg gcc acg cgc ctg ctc 2858
Pro Ala Ala Asp His Tyr Tyr Ser Lys Val Thr Ala Thr Arg Leu Leu
875 880 885
gcc agc acc aac agc agg ctg tgc gcg gta ttc gtc agg gac ggg caa 2906
Ala Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val Arg Asp Gly Gln
890 895 900
tcg gtc atc gga gcc tgc gcc agc ccg tat gaa ggc agg tac aga gac 2954

Ser Val Ile Gly Ala Cys Ala Ser Pro Tyr Glu Gly Arg Tyr Arg Asp
905 910 915
atg tac gac gcg ctg cgg cgc ctg ctg tac atg atc tat atg tcc ggc 3002
Met Tyr Asp Ala Leu Arg Arg Leu Leu Tyr Met Ile Tyr Met Ser Gly
920 925 930
ctt gcc gta cgc gtc cac gtc agc aag gag gaa cag tat tac gac tac 3050
Leu Ala Val Arg Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr

935 940 945 950
gag gac gcc aca ttc cag acc tat gcc ctc acc ggc att tcc ctc tgc 3098
Glu Asp Ala Thr Phe Gln Thr Tyr Ala Leu Thr Gly Ile Ser Leu Cys
955 960 965
aac ccg gca gcg tcg ata tgc 3119
Asn Pro Ala Ala Ser Ile Cys
970

<210> 25

<211> 269

<212> PRT

<213> Bordetella pertussis

<400> 25

Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp Leu
1 5 10 15
Thr Trp Leu Ala Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala
20 25 30
Trp Ala Asp Asp Pro Pro Ala Thr Val Tyr Lys Tyr Asp Ser Arg Pro
35 40 45

Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp Gly Asn Asn Asp

50 55 60

Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser

65 70 75 80

Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu Val

85 90 95

Tyr Leu Glu His Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly

100 105 110

Arg Gly Thr Gly His Phe Ile Gly Tyr Ile Tyr Glu Val Arg Ala Asp

115 120 125

Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu Tyr Val Asp Thr

130 135 140

Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr

145 150 155 160

Gln Ser Ala Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg

165 170 175

Arg Val Thr Arg Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr

180 185 190

Thr Glu Tyr Ser Asn Ala Arg Tyr Val Ser Gln Gln Thr Arg Ala Asn

195 200 205

Pro Asn Pro Tyr Thr Ser Arg Arg Ser Val Ala Ser Ile Val Gly Thr

210 215 220

Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys Met Ala Arg Gln Ala

225 230 235 240

Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala

245 250 255

Met Val Leu Val Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe

260 265

<210> 26

<211> 226

<212> PRT

<213> Bordetella pertussis

<400> 26

Met Pro Ile Asp Arg Lys Thr Leu Cys His Leu Leu Ser Val Leu Pro
1 5 10 15

Leu Ala Leu Leu Gly Ser His Val Ala Arg Ala Ser Thr Pro Gly Ile
20 25 30

Val Ile Pro Pro Gln Glu Gln Ile Thr Gln His Gly Ser Pro Tyr Gly
35 40 45

Arg Cys Ala Asn Lys Thr Arg Ala Leu Thr Val Ala Glu Leu Arg Gly
50 55 60

Ser Gly Asp Leu Gln Glu Tyr Leu Arg His Val Thr Arg Gly Trp Ser
65 70 75 80

Ile Phe Ala Leu Tyr Asp Gly Thr Tyr Leu Gly Gly Glu Tyr Gly Gly
85 90 95

Val Ile Lys Asp Gly Thr Pro Gly Gly Ala Phe Asp Leu Lys Thr Thr
100 105 110

Phe Cys Ile Met Thr Thr Arg Asn Thr Gly Gln Pro Ala Thr Asp His
115 120 125

Tyr Tyr Ser Asn Val Thr Ala Thr Arg Leu Leu Ser Ser Thr Asn Ser
130 135 140

Arg Leu Cys Ala Val Phe Val Arg Ser Gly Gln Pro Val Ile Gly Ala
145 150 155 160

Cys Thr Ser Pro Tyr Asp Gly Lys Tyr Trp Ser Met Tyr Ser Arg Leu
165 170 175

Arg Lys Met Leu Tyr Leu Ile Tyr Val Ala Gly Ile Ser Val Arg Val
180 185 190

His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe
195 200 205

Glu Thr Tyr Ala Leu Thr Gly Ile Ser Ile Cys Asn Pro Gly Ser Ser
210 215 220

Leu Cys

225

<210> 27

<211> 227

<212> PRT

<213> Bordetella pertussis

<400> 27

Met Leu Ile Asn Asn Lys Lys Leu Leu His His Ile Leu Pro Ile Leu

1 5 10 15

Val Leu Ala Leu Leu Gly Met Arg Thr Ala Gln Ala Val Ala Pro Gly

20 25 30

Ile Val Ile Pro Pro Lys Ala Leu Phe Thr Gln Gln Gly Gly Ala Tyr

35 40 45

Gly Arg Cys Pro Asn Gly Thr Arg Ala Leu Thr Val Ala Glu Leu Arg

50 55 60

Gly Asn Ala Glu Leu Gln Thr Tyr Leu Arg Gln Ile Thr Pro Gly Trp

65 70 75 80

Ser Ile Tyr Gly Leu Tyr Asp Gly Thr Tyr Leu Gly Gln Ala Tyr Gly

85 90 95

Gly Ile Ile Lys Asp Ala Pro Pro Gly Ala Gly Phe Ile Tyr Arg Glu

100 105 110

Thr Phe Cys Ile Thr Thr Ile Tyr Lys Thr Gly Gln Pro Ala Ala Asp

115 120 125

His Tyr Tyr Ser Lys Val Thr Ala Thr Arg Leu Leu Ala Ser Thr Asn

130 135 140

Ser Arg Leu Cys Ala Val Phe Val Arg Asp Gly Gln Ser Val Ile Gly

145 150 155 160

Ala Cys Ala Ser Pro Tyr Glu Gly Arg Tyr Arg Asp Met Tyr Asp Ala

165 170 175

Leu Arg Arg Leu Leu Tyr Met Ile Tyr Met Ser Gly Leu Ala Val Arg

180 185 190

Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr

195 200 205

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

210 215 220

Ser Ile Cys

225

<210> 28

<211> 131

<212> PRT

<213> Bordetella pertussis

<400> 28

Val Arg Ala Leu Ala Trp Leu Leu Ala Ser Gly Ala Met Thr His Leu

1 5 10 15

Ser Pro Ala Leu Ala Asp Val Pro Tyr Val Leu Val Lys Thr Asn Met

20 25 30

Val Val Thr Ser Val Ala Met Lys Pro Tyr Glu Val Thr Pro Thr Arg

35 40 45

Met Leu Val Cys Gly Ile Ala Ala Lys Leu Gly Ala Ala Ala Ser Ser

50 55 60

Pro Asp Ala His Val Pro Phe Cys Phe Gly Lys Asp Leu Lys Arg Pro

65 70 75 80

Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala Val Phe Met Gln Gln

85 90 95

Arg Pro Leu Arg Met Phe Leu Gly Pro Lys Gln Leu Thr Phe Glu Gly

100 105 110

Lys Pro Ala Leu Glu Leu Ile Arg Met Val Glu Cys Ser Gly Lys Gln

115 120 125

Asp Cys Pro

130

<210> 29

<211> 120

<212> PRT

<213> Bordetella pertussis

<400> 29

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

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

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

<210> 30

<211> 451

<212> PRT

<213> Clostridium tetani

<400> 30

Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp Ile Asp Val Ile

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

65 70 75 80

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

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

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

450

<210> 31

<211> 1353

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<220><221> CDS

<222> (1)..(1353)

<400> 31

aaa aac ctg gac tgt tgg gtt gac aac gaa gaa gat atc gat gtc atc 48

Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp Ile Asp Val Ile

1 5 10 15

ctg aag aaa tcc acc att ttg aac ctc gac atc aac aat gac atc att 96

Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile Asn Asn Asp Ile Ile

20

25

30

tcc gac att agc ggt ttc aac tcg tcc gtg att acg tac cca gat gct	144
Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile Thr Tyr Pro Asp Ala	
35 40 45	
cag ctg gtg ccc ggg att aac ggc aag gct atc cac ctc gtc aac aac	192
Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile His Leu Val Asn Asn	
50 55 60	
gag tcg tcg gaa gtc atc gtc cat aaa gcg atg gac atc gag tat aac	240
Glu Ser Ser Glu Val Ile Val His Lys Ala Met Asp Ile Glu Tyr Asn	
65 70 75 80	
gac atg ttt aat aat ttc acc gtg tcc ttt tgg ctg cgc gtg ccc aag	288
Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys	
85 90 95	
gtg tcc gcc tcc cac ctg gaa cag tac ggg acc aac gag tac agc atc	336
Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr Asn Glu Tyr Ser Ile	
100 105 110	
atc agc tcg atg aag aag cac tcg ttg agc atc ggc agc ggc tgg tcg	384
Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile Gly Ser Gly Trp Ser	
115 120 125	
gtt agt ctc aaa ggg aac aac ctg att tgg acc ctg aaa gat agc gcc	432
Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr Leu Lys Asp Ser Ala	
130 135 140	
ggc gag gtg cgt cag atc act ttc cgg gac ctg ccg gat aag ttc aac	480
Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu Pro Asp Lys Phe Asn	
145 150 155 160	
gcc tac ctg gca aac aaa tgg gtg ttc att acc atc acg aac gac cgc	528
Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr Ile Thr Asn Asp Arg	
165 170 175	
ctg agt agc gcg aat ctc tac atc aat ggc gtg ctg atg ggc agc gcg	576
Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val Leu Met Gly Ser Ala	
180 185 190	
gaa atc acg ggc ttg ggt gcc atc cgc gaa gat aac aat atc acc ttg	624

Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp Asn Asn Ile Thr Leu	
195 200 205	
aag ctg gac cgc tgc aac aac aac aac caa tac gtg tcc att gat aag	672
Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr Val Ser Ile Asp Lys	
210 215 220	
ttc cgc atc ttt tgc aag gcc ctg aac ccg aaa gag atc gaa aag ctc	720
Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys Glu Ile Glu Lys Leu	
225 230 235 240	
tac acc agc tac ttg agt atc acc ttc ctg cgc gac ttt tgg ggt aat	768
Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg Asp Phe Trp Gly Asn	
245 250 255	
ccg ttg cgt tat gac acc gag tat tat ctg atc ccc gtg gcc agc agc	816
Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile Pro Val Ala Ser Ser	
260 265 270	
agc aag gac gtc cag ctg aag aac atc acc gac tac atg tac ttg act	864
Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp Tyr Met Tyr Leu Thr	
275 280 285	
aac gcg ccc tcg tat acc aat ggc aaa ctg aac att tac tac cgc cgg	912
Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn Ile Tyr Tyr Arg Arg	
290 295 300	
ctc tac aac ggg ctc aag ttc atc atc aaa cgc tat acg ccg aat aat	960
Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn	
305 310 315 320	
gaa atc gac agt ttt gtc aag agc ggc gac ttc atc aag ttg tac gtg	1008
Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe Ile Lys Leu Tyr Val	
325 330 335	
agc tac aat aac aac gag cac atc gtt ggt tac cct aag gat ggc aac	1056
Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr Pro Lys Asp Gly Asn	
340 345 350	
gct ttc aac aac ctc gat cgt atc ctg cgg gtt ggc tac aac gca cca	1104

Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn Ala Pro

355 360 365
ggc att ccg ctg tat aag aag atg gaa gcg gtc aaa ctg cgt gac ctg 1152
Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu Arg Asp Leu

370 375 380
aaa acc tac tcc gtg caa ctg aag ctg tac gac gac aag aat gcc tgc 1200
Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn Ala Ser

385 390 395 400

ttg ggt ctg gtc ggc acg cat aat ggt cag att ggc aac gac ccg aac 1248
Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn Asp Pro Asn

405 410 415
cgg gac atc ctg atc gcc agc aac tgg tat ttc aat cac ctg aag gat 1296
Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His Leu Lys Asp

420 425 430
aag atc ttg ggc tgc gat tgg tat ttc gtc cct acc gat gag ggc tgg 1344

Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr Asp Glu Gly Trp
435 440 445

act aat gac 1353

Thr Asn Asp

450

<210> 32

<211> 2366

<212> PRT

<213> Clostridium difficile

<400> 32

Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg

1 5 10 15
Phe Arg Thr Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu

20 25 30
Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys

35 40 45

Leu Lys Asp Ile Asn Ser Leu Thr Asp Ile Tyr Ile Asp Thr Tyr Lys
 50 55 60
 Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val

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

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

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

 260 265 270
 Arg Ile Ser Ala Leu Lys Glu Ile Gly Gly Met Tyr Leu Asp Val Asp
 275 280 285
 Met Leu Pro Gly Ile Gln Pro Asp Leu Phe Glu Ser Ile Glu Lys Pro

290 295 300
 Ser Ser Val Thr Val Asp Phe Trp Glu Met Thr Lys Leu Glu Ala Ile
 305 310 315 320
 Met Lys Tyr Lys Glu Tyr Ile Pro Glu Tyr Thr Ser Glu His Phe Asp

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

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

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

 515 520 525
 Ala Gln Phe Glu Glu Tyr Lys Arg Asn Tyr Phe Glu Gly Ser Leu Gly
 530 535 540

Glu Asp Asp Asn Leu Asp Phe Ser Gln Asn Ile Val Val Asp Lys Glu
 545 550 555 560
 Tyr Leu Leu Glu Lys Ile Ser Ser Leu Ala Arg Ser Ser Glu Arg Gly
 565 570 575
 Tyr Ile His Tyr Ile Val Gln Leu Gln Gly Asp Lys Ile Ser Tyr Glu

 580 585 590
 Ala Ala Cys Asn Leu Phe Ala Lys Thr Pro Tyr Asp Ser Val Leu Phe
 595 600 605
 Gln Lys Asn Ile Glu Asp Ser Glu Ile Ala Tyr Tyr Tyr Asn Pro Gly
 610 615 620
 Asp Gly Glu Ile Gln Glu Ile Asp Lys Tyr Lys Ile Pro Ser Ile Ile
 625 630 635 640
 Ser Asp Arg Pro Lys Ile Lys Leu Thr Phe Ile Gly His Gly Lys Asp

 645 650 655
 Glu Phe Asn Thr Asp Ile Phe Ala Gly Phe Asp Val Asp Ser Leu Ser
 660 665 670
 Thr Glu Ile Glu Ala Ala Ile Asp Leu Ala Lys Glu Asp Ile Ser Pro
 675 680 685
 Lys Ser Ile Glu Ile Asn Leu Leu Gly Cys Asn Met Phe Ser Tyr Ser
 690 695 700
 Ile Asn Val Glu Glu Thr Tyr Pro Gly Lys Leu Leu Leu Lys Val Lys

 705 710 715 720
 Asp Lys Ile Ser Glu Leu Met Pro Ser Ile Ser Gln Asp Ser Ile Ile
 725 730 735
 Val Ser Ala Asn Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly Arg Arg
 740 745 750
 Glu Leu Leu Asp His Ser Gly Glu Trp Ile Asn Lys Glu Glu Ser Ile
 755 760 765
 Ile Lys Asp Ile Ser Ser Lys Glu Tyr Ile Ser Phe Asn Pro Lys Glu

 770 775 780
 Asn Lys Ile Thr Val Lys Ser Lys Asn Leu Pro Glu Leu Ser Thr Leu

785 790 795 800
 Leu Gln Glu Ile Arg Asn Asn Ser Asn Ser Ser Asp Ile Glu Leu Glu
 805 810 815
 Glu Lys Val Met Leu Thr Glu Cys Glu Ile Asn Val Ile Ser Asn Ile
 820 825 830
 Asp Thr Gln Ile Val Glu Glu Arg Ile Glu Glu Ala Lys Asn Leu Thr

 835 840 845
 Ser Asp Ser Ile Asn Tyr Ile Lys Asp Glu Phe Lys Leu Ile Glu Ser
 850 855 860
 Ile Ser Asp Ala Leu Cys Asp Leu Lys Gln Gln Asn Glu Leu Glu Asp
 865 870 875 880
 Ser His Phe Ile Ser Phe Glu Asp Ile Ser Glu Thr Asp Glu Gly Phe
 885 890 895
 Ser Ile Arg Phe Ile Asn Lys Glu Thr Gly Glu Ser Ile Phe Val Glu

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

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

 1025 1030 1035

Val Ser Leu Gly Ala Ala Ile	Lys Glu Leu Ser Glu Thr Ser Asp	
1040	1045	1050
Pro Leu Leu Arg Gln Glu Ile	Glu Ala Lys Ile Gly Ile Met Ala	
1055	1060	1065
Val Asn Leu Thr Thr Ala Thr	Thr Ala Ile Ile Thr Ser Ser Leu	
1070	1075	1080
Gly Ile Ala Ser Gly Phe Ser	Ile Leu Leu Val Pro Leu Ala Gly	
1085	1090	1095
Ile Ser Ala Gly Ile Pro Ser	Leu Val Asn Asn Glu Leu Val Leu	
1100	1105	1110
Arg Asp Lys Ala Thr Lys Val	Val Asp Tyr Phe Lys His Val Ser	
1115	1120	1125
Leu Val Glu Thr Glu Gly Val	Phe Thr Leu Leu Asp Asp Lys Ile	
1130	1135	1140
Met Met Pro Gln Asp Asp Leu	Val Ile Ser Glu Ile Asp Phe Asn	
1145	1150	1155
Asn Asn Ser Ile Val Leu Gly	Lys Cys Glu Ile Trp Arg Met Glu	
1160	1165	1170
Gly Gly Ser Gly His Thr Val	Thr Asp Asp Ile Asp His Phe Phe	
1175	1180	1185
Ser Ala Pro Ser Ile Thr Tyr	Arg Glu Pro His Leu Ser Ile Tyr	
1190	1195	1200
Asp Val Leu Glu Val Gln Lys	Glu Glu Leu Asp Leu Ser Lys Asp	
1205	1210	1215
Leu Met Val Leu Pro Asn Ala	Pro Asn Arg Val Phe Ala Trp Glu	
1220	1225	1230
Thr Gly Trp Thr Pro Gly Leu	Arg Ser Leu Glu Asn Asp Gly Thr	
1235	1240	1245
Lys Leu Leu Asp Arg Ile Arg	Asp Asn Tyr Glu Gly Glu Phe Tyr	
1250	1255	1260
Trp Arg Tyr Phe Ala Phe Ile	Ala Asp Ala Leu Ile Thr Thr Leu	

1265	1270	1275
Lys Pro Arg Tyr Glu Asp Thr Asn Ile Arg Ile Asn Leu Asp Ser		
1280	1285	1290
Asn Thr Arg Ser Phe Ile Val Pro Ile Ile Thr Thr Glu Tyr Ile		
1295	1300	1305
Arg Glu Lys Leu Ser Tyr Ser Phe Tyr Gly Ser Gly Gly Thr Tyr		
1310	1315	1320
Ala Leu Ser Leu Ser Gln Tyr Asn Met Gly Ile Asn Ile Glu Leu		
1325	1330	1335
Ser Glu Ser Asp Val Trp Ile Ile Asp Val Asp Asn Val Val Arg		
1340	1345	1350
Asp Val Thr Ile Glu Ser Asp Lys Ile Lys Lys Gly Asp Leu Ile		
1355	1360	1365
Glu Gly Ile Leu Ser Thr Leu Ser Ile Glu Glu Asn Lys Ile Ile		
1370	1375	1380
Leu Asn Ser His Glu Ile Asn Phe Ser Gly Glu Val Asn Gly Ser		
1385	1390	1395
Asn Gly Phe Val Ser Leu Thr Phe Ser Ile Leu Glu Gly Ile Asn		
1400	1405	1410
Ala Ile Ile Glu Val Asp Leu Leu Ser Lys Ser Tyr Lys Leu Leu		
1415	1420	1425
Ile Ser Gly Glu Leu Lys Ile Leu Met Leu Asn Ser Asn His Ile		
1430	1435	1440
Gln Gln Lys Ile Asp Tyr Ile Gly Phe Asn Ser Glu Leu Gln Lys		
1445	1450	1455
Asn Ile Pro Tyr Ser Phe Val Asp Ser Glu Gly Lys Glu Asn Gly		
1460	1465	1470
Phe Ile Asn Gly Ser Thr Lys Glu Gly Leu Phe Val Ser Glu Leu		
1475	1480	1485
Pro Asp Val Val Leu Ile Ser Lys Val Tyr Met Asp Asp Ser Lys		
1490	1495	1500

Pro Ser Phe Gly Tyr Tyr Ser Asn Asn Leu Lys Asp Val Lys Val

1505 1510 1515

Ile Thr Lys Asp Asn Val Asn Ile Leu Thr Gly Tyr Tyr Leu Lys

1520 1525 1530

Asp Asp Ile Lys Ile Ser Leu Ser Leu Thr Leu Gln Asp Glu Lys

1535 1540 1545

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

1550 1555 1560

Glu Ile Leu Lys Phe Met Asn Arg Lys Gly Asn Thr Asn Thr Ser

1565 1570 1575

Asp Ser Leu Met Ser Phe Leu Glu Ser Met Asn Ile Lys Ser Ile

1580 1585 1590

Phe Val Asn Phe Leu Gln Ser Asn Ile Lys Phe Ile Leu Asp Ala

1595 1600 1605

Asn Phe Ile Ile Ser Gly Thr Thr Ser Ile Gly Gln Phe Glu Phe

1610 1615 1620

Ile Cys Asp Glu Asn Asp Asn Ile Gln Pro Tyr Phe Ile Lys Phe

1625 1630 1635

Asn Thr Leu Glu Thr Asn Tyr Thr Leu Tyr Val Gly Asn Arg Gln

1640 1645 1650

Asn Met Ile Val Glu Pro Asn Tyr Asp Leu Asp Asp Ser Gly Asp

1655 1660 1665

Ile Ser Ser Thr Val Ile Asn Phe Ser Gln Lys Tyr Leu Tyr Gly

1670 1675 1680

Ile Asp Ser Cys Val Asn Lys Val Val Ile Ser Pro Asn Ile Tyr

1685 1690 1695

Thr Asp Glu Ile Asn Ile Thr Pro Val Tyr Glu Thr Asn Asn Thr

1700 1705 1710

Tyr Pro Glu Val Ile Val Leu Asp Ala Asn Tyr Ile Asn Glu Lys

1715 1720 1725

Ile Asn Val Asn Ile Asn Asp Leu Ser Ile Arg Tyr Val Trp Ser

1730	1735	1740
Asn Asp Gly Asn Asp Phe Ile	Leu Met Ser Thr Ser	Glu Glu Asn
1745	1750	1755
Lys Val Ser Gln Val Lys Ile	Arg Phe Val Asn Val	Phe Lys Asp
1760	1765	1770
Lys Thr Leu Ala Asn Lys Leu	Ser Phe Asn Phe Ser	Asp Lys Gln
1775	1780	1785
Asp Val Pro Val Ser Glu Ile	Ile Leu Ser Phe Thr	Pro Ser Tyr
1790	1795	1800
Tyr Glu Asp Gly Leu Ile Gly	Tyr Asp Leu Gly Leu	Val Ser Leu
1805	1810	1815
Tyr Asn Glu Lys Phe Tyr Ile	Asn Asn Phe Gly Met	Met Val Ser
1820	1825	1830
Gly Leu Ile Tyr Ile Asn Asp	Ser Leu Tyr Tyr Phe	Lys Pro Pro
1835	1840	1845
Val Asn Asn Leu Ile Thr Gly	Phe Val Thr Val Gly	Asp Asp Lys
1850	1855	1860
Tyr Tyr Phe Asn Pro Ile Asn	Gly Gly Ala Ala Ser	Ile Gly Glu
1865	1870	1875
Thr Ile Ile Asp Asp Lys Asn	Tyr Tyr Phe Asn Gln	Ser Gly Val
1880	1885	1890
Leu Gln Thr Gly Val Phe Ser	Thr Glu Asp Gly Phe	Lys Tyr Phe
1895	1900	1905
Ala Pro Ala Asn Thr Leu Asp	Glu Asn Leu Glu Gly	Glu Ala Ile
1910	1915	1920
Asp Phe Thr Gly Lys Leu Ile	Ile Asp Glu Asn Ile	Tyr Tyr Phe
1925	1930	1935
Asp Asp Asn Tyr Arg Gly Ala	Val Glu Trp Lys Glu	Leu Asp Gly
1940	1945	1950
Glu Met His Tyr Phe Ser Pro	Glu Thr Gly Lys Ala	Phe Lys Gly
1955	1960	1965

Leu Asn Gln Ile Gly Asp Tyr	Lys Tyr Tyr Phe Asn Ser Asp Gly
1970	1975 1980
Val Met Gln Lys Gly Phe Val	Ser Ile Asn Asp Asn Lys His Tyr
1985	1990 1995
Phe Asp Asp Ser Gly Val Met	Lys Val Gly Tyr Thr Glu Ile Asp
2000	2005 2010
Gly Lys His Phe Tyr Phe Ala	Glu Asn Gly Glu Met Gln Ile Gly
2015	2020 2025
Val Phe Asn Thr Glu Asp Gly	Phe Lys Tyr Phe Ala His His Asn
2030	2035 2040
Glu Asp Leu Gly Asn Glu Glu	Gly Glu Glu Ile Ser Tyr Ser Gly
2045	2050 2055
Ile Leu Asn Phe Asn Asn Lys	Ile Tyr Tyr Phe Asp Asp Ser Phe
2060	2065 2070
Thr Ala Val Val Gly Trp Lys	Asp Leu Glu Asp Gly Ser Lys Tyr
2075	2080 2085
Tyr Phe Asp Glu Asp Thr Ala	Glu Ala Tyr Ile Gly Leu Ser Leu
2090	2095 2100
Ile Asn Asp Gly Gln Tyr Tyr	Phe Asn Asp Asp Gly Ile Met Gln
2105	2110 2115
Val Gly Phe Val Thr Ile Asn	Asp Lys Val Phe Tyr Phe Ser Asp
2120	2125 2130
Ser Gly Ile Ile Glu Ser Gly	Val Gln Asn Ile Asp Asp Asn Tyr
2135	2140 2145
Phe Tyr Ile Asp Asp Asn Gly	Ile Val Gln Ile Gly Val Phe Asp
2150	2155 2160
Thr Ser Asp Gly Tyr Lys Tyr	Phe Ala Pro Ala Asn Thr Val Asn
2165	2170 2175
Asp Asn Ile Tyr Gly Gln Ala	Val Glu Tyr Ser Gly Leu Val Arg
2180	2185 2190
Val Gly Glu Asp Val Tyr Tyr	Phe Gly Glu Thr Tyr Thr Ile Glu

2195 2200 2205
 Thr Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr
 2210 2215 2220
 Phe Asn Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile

2225 2230 2235
 Asp Asp Ile Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr
 2240 2245 2250
 Gly Leu Ile Ser Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn
 2255 2260 2265
 Gly Glu Met Gln Phe Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe
 2270 2275 2280
 Tyr Phe Gly Glu Asp Gly Val Met Gln Ile Gly Val Phe Asn Thr

2285 2290 2295
 Pro Asp Gly Phe Lys Tyr Phe Ala His Gln Asn Thr Leu Asp Glu
 2300 2305 2310
 Asn Phe Glu Gly Glu Ser Ile Asn Tyr Thr Gly Trp Leu Asp Leu
 2315 2320 2325
 Asp Glu Lys Arg Tyr Tyr Phe Thr Asp Glu Tyr Ile Ala Ala Thr
 2330 2335 2340
 Gly Ser Val Ile Ile Asp Gly Glu Glu Tyr Tyr Phe Asp Pro Asp

2345 2350 2355
 Thr Ala Gln Leu Val Ile Ser Glu
 2360 2365

<210> 33

<211> 7098

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<220><221> CDS

<222> (1)..(7098)

<400> 33

atg tcc ctc gtc aat cgc aag cag ctg gag aag atg gcc aac gtt cgt	48
Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg	
1 5 10 15	
ttc cgc acc caa gag gac gaa tac gtc gcc atc ctc gac gcc ctg gaa	96
Phe Arg Thr Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu	
20 25 30	
gaa tac cat aac atg agc gaa aac acc gtt gtc gag aag tac ctc aag	144
Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys	
35 40 45	
ctg aag gac atc aac agc ctg acg gac atc tat atc gac acg tac aag	192
Leu Lys Asp Ile Asn Ser Leu Thr Asp Ile Tyr Ile Asp Thr Tyr Lys	
50 55 60	
aag tcc ggt cgc aac aag gca ctc aag aag ttc aaa gag tac ctg gtc	240
Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val	
65 70 75 80	
acc gaa gtg ttg gaa ctc aag aac aac aac ctc acg ccg gtg gaa aag	288
Thr Glu Val Leu Glu Leu Lys Asn Asn Asn Leu Thr Pro Val Glu Lys	
85 90 95	
aac ctg cat ttc gtg tgg att ggc ggc cag atc aac gac acc gcc atc	336
Asn Leu His Phe Val Trp Ile Gly Gly Gln Ile Asn Asp Thr Ala Ile	
100 105 110	
aac tac att aac cag tgg aaa gac gtc aac tcg gac tac aat gtg aat	384
Asn Tyr Ile Asn Gln Trp Lys Asp Val Asn Ser Asp Tyr Asn Val Asn	
115 120 125	
gtg ttc tat gac tcg aac gcc ttt ttg atc aac acg ctg aag aaa acc	432
Val Phe Tyr Asp Ser Asn Ala Phe Leu Ile Asn Thr Leu Lys Lys Thr	
130 135 140	
gtc gtg gaa tcg gcc atc aat gac acc ctg gaa agc ttc cgt gaa aac	480
Val Val Glu Ser Ala Ile Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn	
145 150 155 160	

ctc aac gat cct cgg ttt gac tac aac aag ttt ttc cgc aag cgc atg 528

Leu Asn Asp Pro Arg Phe Asp Tyr Asn Lys Phe Phe Arg Lys Arg Met
165 170 175

gag atc atc tat gac aag cag aaa aac ttt atc aac tac tac aaa gcg 576

Glu Ile Ile Tyr Asp Lys Gln Lys Asn Phe Ile Asn Tyr Tyr Lys Ala
180 185 190

cag cgc gaa gag aac ccg gag ctg atc atc gac gat atc gtg aaa acc 624

Gln Arg Glu Glu Asn Pro Glu Leu Ile Ile Asp Asp Ile Val Lys Thr
195 200 205

tat ctg tcc aac gag tat agt aaa gaa atc gat gag ctg aac acg tat 672

Tyr Leu Ser Asn Glu Tyr Ser Lys Glu Ile Asp Glu Leu Asn Thr Tyr
210 215 220

atc gaa gag agt ctg aac aag atc act cag aac agc ggt aac gac gtc 720

Ile Glu Glu Ser Leu Asn Lys Ile Thr Gln Asn Ser Gly Asn Asp Val
225 230 235 240

cgg aac ttt gaa gag ttc aaa aac ggc gag tcg ttc aac ctc tac gaa 768

Arg Asn Phe Glu Glu Phe Lys Asn Gly Glu Ser Phe Asn Leu Tyr Glu
245 250 255

cag gag ctg gtc gag cgc tgg aac ctg gca gcg gcg tcg gac atc ctg 816

Gln Glu Leu Val Glu Arg Trp Asn Leu Ala Ala Ala Ser Asp Ile Leu
260 265 270

cgt atc agc gct ctg aaa gag atc ggc ggc atg tac ctg gac gtg gat 864

Arg Ile Ser Ala Leu Lys Glu Ile Gly Gly Met Tyr Leu Asp Val Asp
275 280 285

atg ctc cct ggc atc cag cct gat ctg ttt gaa tcg att gaa aag cca 912

Met Leu Pro Gly Ile Gln Pro Asp Leu Phe Glu Ser Ile Glu Lys Pro
290 295 300

tcg agc gtg acc gtc gac ttt tgg gag atg acc aag ctg gag gcg atc 960

Ser Ser Val Thr Val Asp Phe Trp Glu Met Thr Lys Leu Glu Ala Ile

305	310	315	320	
atg aag tac aaa gaa tac atc ccg gag tat acg agt gaa cac ttt gat				1008
Met Lys Tyr Lys Glu Tyr Ile Pro Glu Tyr Thr Ser Glu His Phe Asp				
325	330	335		
atg ctg gac gaa gaa gtg caa tcc tcg ttt gaa agc gtc ctg gcg agc				1056
Met Leu Asp Glu Glu Val Gln Ser Ser Phe Glu Ser Val Leu Ala Ser				
340	345	350		
aag agt gat aag agc gaa atc ttc tcg tcc ttg ggc gat atg gag gcg				1104
Lys Ser Asp Lys Ser Glu Ile Phe Ser Ser Leu Gly Asp Met Glu Ala				
355	360	365		
tcc cca ctg gag gtc aaa atc gcc ttc aac agc aag ggc att atc aat				1152
Ser Pro Leu Glu Val Lys Ile Ala Phe Asn Ser Lys Gly Ile Ile Asn				
370	375	380		
cag ggc ctg att tcg gtc aag gat agc tac tgc agc aac ctg atc gtc				1200
Gln Gly Leu Ile Ser Val Lys Asp Ser Tyr Cys Ser Asn Leu Ile Val				
385	390	395	400	
aag cag atc gag aac cgt tac aag atc ctg aac aac agt ctg aac ccc				1248
Lys Gln Ile Glu Asn Arg Tyr Lys Ile Leu Asn Asn Ser Leu Asn Pro				
405	410	415		
gcc atc agc gaa gat aat gac ttc aat acc acg acg aac acg ttt atc				1296
Ala Ile Ser Glu Asp Asn Asp Phe Asn Thr Thr Thr Asn Thr Phe Ile				
420	425	430		
gac tcc atc atg gcc gaa gcc aac gcg gac aac ggc cgc ttt atg atg				1344
Asp Ser Ile Met Ala Glu Ala Asn Ala Asp Asn Gly Arg Phe Met Met				
435	440	445		
gag ttg ggg aag tac ctg cgc gtg ggc ttc ttc ccg gac gtg aaa acc				1392
Glu Leu Gly Lys Tyr Leu Arg Val Gly Phe Phe Pro Asp Val Lys Thr				
450	455	460		
acg atc aat ctc tcc ggc cca gaa gcg tat gca gcc gca tac caa gat				1440
Thr Ile Asn Leu Ser Gly Pro Glu Ala Tyr Ala Ala Ala Tyr Gln Asp				
465	470	475	480	

ctg ctc atg ttc aaa gag ggc tcg atg aac atc cat ctg att gag gcg	1488
Leu Leu Met Phe Lys Glu Gly Ser Met Asn Ile His Leu Ile Glu Ala	
485 490 495	
gac ttg cgc aac ttc gaa atc tcg aaa acg aac atc agc caa tcg acg	1536
Asp Leu Arg Asn Phe Glu Ile Ser Lys Thr Asn Ile Ser Gln Ser Thr	
500 505 510	
gag cag gaa atg gcg agc ctg tgg tcc ttc gac gat gct cgc gcc aag	1584
Glu Gln Glu Met Ala Ser Leu Trp Ser Phe Asp Asp Ala Arg Ala Lys	
515 520 525	
gcc caa ttt gaa gag tac aaa cgg aac tac ttc gaa ggc tcg ctg ggt	1632
Ala Gln Phe Glu Glu Tyr Lys Arg Asn Tyr Phe Glu Gly Ser Leu Gly	
530 535 540	
gag gat gac aac ttg gac ttc tcg caa aac atc gtc gtg gac aaa gaa	1680
Glu Asp Asp Asn Leu Asp Phe Ser Gln Asn Ile Val Val Asp Lys Glu	
545 550 555 560	
tac ctg ttg gaa aag atc agc tcc ctg gcc cgg agc tcg gaa cgg ggc	1728
Tyr Leu Leu Glu Lys Ile Ser Ser Leu Ala Arg Ser Ser Glu Arg Gly	
565 570 575	
tac atc cac tac att gtt cag ctg caa ggg gat aag atc tcg tat gaa	1776
Tyr Ile His Tyr Ile Val Gln Leu Gln Gly Asp Lys Ile Ser Tyr Glu	
580 585 590	
gcg gcg tgc aat ctc ttc gcc aag acg ccg tac gac tcc gtg ctg ttc	1824
Ala Ala Cys Asn Leu Phe Ala Lys Thr Pro Tyr Asp Ser Val Leu Phe	
595 600 605	
cag aag aac atc gag gac agt gaa atc gcc tac tat tac aac ccc ggt	1872
Gln Lys Asn Ile Glu Asp Ser Glu Ile Ala Tyr Tyr Tyr Asn Pro Gly	
610 615 620	
gac ggc gaa atc caa gaa att gat aag tac aag atc ccg tcc att atc	1920
Asp Gly Glu Ile Gln Glu Ile Asp Lys Tyr Lys Ile Pro Ser Ile Ile	
625 630 635 640	
tcc gat cgt ccg aag atc aaa ctc acc ttc att ggc cac ggg aag gac	1968

Ser Asp Arg Pro Lys Ile Lys Leu Thr Phe Ile Gly His Gly Lys Asp

645	650	655	
gag ttc aac acc gat atc ttc gca ggt ttt gac gtg gat agt ctc tcg			2016
Glu Phe Asn Thr Asp Ile Phe Ala Gly Phe Asp Val Asp Ser Leu Ser			
660	665	670	
acc gag atc gag gcc gcg atc gac ctg gcg aaa gag gac att tcg cca			2064
Thr Glu Ile Glu Ala Ala Ile Asp Leu Ala Lys Glu Asp Ile Ser Pro			
675	680	685	
aag tcg atc gaa atc aac ctg ctg ggc tgc aat atg ttt tcg tat agt			2112
Lys Ser Ile Glu Ile Asn Leu Leu Gly Cys Asn Met Phe Ser Tyr Ser			
690	695	700	
atc aac gtt gaa gaa acc tat ccg ggc aaa ctc ctg ctg aag gtc aag			2160
Ile Asn Val Glu Glu Thr Tyr Pro Gly Lys Leu Leu Leu Lys Val Lys			
705	710	715	720
gac aag att agc gaa ctg atg ccg agc atc tcg cag gat agc atc atc			2208
Asp Lys Ile Ser Glu Leu Met Pro Ser Ile Ser Gln Asp Ser Ile Ile			
725	730	735	
gtg agt gct aac cag tat gag gtg cgt atc aac agc gag ggt cgc cgc			2256
Val Ser Ala Asn Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly Arg Arg			
740	745	750	
gag ctg ctc gat cac tcg ggc gag tgg atc aac aaa gaa gag agc atc			2304
Glu Leu Leu Asp His Ser Gly Glu Trp Ile Asn Lys Glu Glu Ser Ile			
755	760	765	
atc aaa gat atc agc agt aaa gaa tac att agt ttc aac ccg aaa gag			2352
Ile Lys Asp Ile Ser Ser Lys Glu Tyr Ile Ser Phe Asn Pro Lys Glu			
770	775	780	
aac aag atc acg gtg aaa tcg aag aat ttg ccg gag ttg agt acc ctg			2400
Asn Lys Ile Thr Val Lys Ser Lys Asn Leu Pro Glu Leu Ser Thr Leu			
785	790	795	800

ctg cag gag atc cgt aac aat tcc aac tcc agt gat atc gaa ctc gaa	2448
Leu Gln Glu Ile Arg Asn Asn Ser Asn Ser Ser Asp Ile Glu Leu Glu	
805 810 815	
gaa aag gtc atg ctg acc gag tgc gag atc aat gtc atc agc aac atc	2496
Glu Lys Val Met Leu Thr Glu Cys Glu Ile Asn Val Ile Ser Asn Ile	
820 825 830	
gac act cag atc gtg gaa gaa cgc att gaa gag gcc aag aac ctg acc	2544
Asp Thr Gln Ile Val Glu Glu Arg Ile Glu Glu Ala Lys Asn Leu Thr	
835 840 845	
tcg gac tcg atc aac tat atc aaa gac gag ttc aag ctg att gaa tcc	2592
Ser Asp Ser Ile Asn Tyr Ile Lys Asp Glu Phe Lys Leu Ile Glu Ser	
850 855 860	
atc tcg gat gcc ctg tgc gac ctg aag cag cag aac gag ttg gag gat	2640
Ile Ser Asp Ala Leu Cys Asp Leu Lys Gln Gln Asn Glu Leu Glu Asp	
865 870 875 880	
agc cac ttc atc agt ttc gag gat atc tcg gag act gac gag ggc ttc	2688
Ser His Phe Ile Ser Phe Glu Asp Ile Ser Glu Thr Asp Glu Gly Phe	
885 890 895	
agc atc cgt ttc atc aac aaa gaa acc ggt gaa tcc att ttc gtt gaa	2736
Ser Ile Arg Phe Ile Asn Lys Glu Thr Gly Glu Ser Ile Phe Val Glu	
900 905 910	
acg gag aaa acc att ttc agc gaa tac gcc aat cac atc acc gaa gag	2784
Thr Glu Lys Thr Ile Phe Ser Glu Tyr Ala Asn His Ile Thr Glu Glu	
915 920 925	
atc agc aag atc aag ggt act atc ttt gac acg gtg aat ggc aaa ctg	2832
Ile Ser Lys Ile Lys Gly Thr Ile Phe Asp Thr Val Asn Gly Lys Leu	
930 935 940	
gtg aag aaa gtc aac ctg gac acc acc cac gag gtc aac acc ctg aac	2880
Val Lys Lys Val Asn Leu Asp Thr Thr His Glu Val Asn Thr Leu Asn	
945 950 955 960	
gcc gcg ttc ttc atc caa agc ctg atc gaa tac aac agc tcc aaa gaa	2928

Ala Ala Phe Phe Ile Gln Ser Leu Ile Glu Tyr Asn Ser Ser Lys Glu	
965 970 975	
tcc ttg agc aac ttg agc gtg gcc atg aaa gtg cag gtg tat gcc cag	2976
Ser Leu Ser Asn Leu Ser Val Ala Met Lys Val Gln Val Tyr Ala Gln	
980 985 990	
ttg ttc tcg acc ggc ctg aac acc att acg gac gcg gca aag gtg gtt	3024
Leu Phe Ser Thr Gly Leu Asn Thr Ile Thr Asp Ala Ala Lys Val Val	
995 1000 1005	
gaa ctg gtc agc acc gcc ttg gac gaa acc atc gac ctg ctg ccg	3069
Glu Leu Val Ser Thr Ala Leu Asp Glu Thr Ile Asp Leu Leu Pro	
1010 1015 1020	
acc ctg agc gag ggc ctg ccc atc atc gcc acg atc att gac ggg	3114
Thr Leu Ser Glu Gly Leu Pro Ile Ile Ala Thr Ile Ile Asp Gly	
1025 1030 1035	
gtc agc ctg ggc gca gcc atc aaa gag ttg agc gag act tcc gac	3159
Val Ser Leu Gly Ala Ala Ile Lys Glu Leu Ser Glu Thr Ser Asp	
1040 1045 1050	
ccg ctg ctg cgc cag gaa atc gaa gct aag atc ggg atc atg gcc	3204
Pro Leu Leu Arg Gln Glu Ile Glu Ala Lys Ile Gly Ile Met Ala	
1055 1060 1065	
gtg aat ctg acc acc gcg acc acc gcg atc att acc tcc agc ctc	3249
Val Asn Leu Thr Thr Ala Thr Thr Ala Ile Ile Thr Ser Ser Leu	
1070 1075 1080	
ggc att gcg tcc ggc ttc tcc atc ctg ctg gtc ccc ttg gcg ggc	3294
Gly Ile Ala Ser Gly Phe Ser Ile Leu Leu Val Pro Leu Ala Gly	
1085 1090 1095	
atc agc gcc ggt atc cct agc ttg gtg aac aac gag ttg gtc ctg	3339
Ile Ser Ala Gly Ile Pro Ser Leu Val Asn Asn Glu Leu Val Leu	
1100 1105 1110	
cgt gat aaa gcg acc aag gtt gtg gat tac ttc aag cat gtc tcc	3384
Arg Asp Lys Ala Thr Lys Val Val Asp Tyr Phe Lys His Val Ser	

1115	1120	1125	
ctg gtc gaa acg gaa ggg gtg ttc acc ctg ctg gac gat aag att			3429
Leu Val Glu Thr Glu Gly Val Phe Thr Leu Leu Asp Asp Lys Ile			
1130	1135	1140	
atg atg ccc caa gac gat ttg gtc atc agc gaa att gac ttt aac			3474
Met Met Pro Gln Asp Asp Leu Val Ile Ser Glu Ile Asp Phe Asn			
1145	1150	1155	
aac aat tcg atc gtc ctc ggg aag tgt gaa atc tgg cgg atg gag			3519
Asn Asn Ser Ile Val Leu Gly Lys Cys Glu Ile Trp Arg Met Glu			
1160	1165	1170	
ggc ggt tcg ggc cac acc gtg acc gat gat atc gac cat ttc ttt			3564
Gly Gly Ser Gly His Thr Val Thr Asp Asp Ile Asp His Phe Phe			
1175	1180	1185	
agt gcg ccc tcc atc acg tac cgc gag ccg cac ctg agc atc tac			3609
Ser Ala Pro Ser Ile Thr Tyr Arg Glu Pro His Leu Ser Ile Tyr			
1190	1195	1200	
gac gtg ctc gag gtg cag aaa gaa gaa ctg gat ctg tcc aaa gat			3654
Asp Val Leu Glu Val Gln Lys Glu Glu Leu Asp Leu Ser Lys Asp			
1205	1210	1215	
ctg atg gtt ctg ccg aac gct ccc aat cgg gtg ttc gct tgg gaa			3699
Leu Met Val Leu Pro Asn Ala Pro Asn Arg Val Phe Ala Trp Glu			
1220	1225	1230	
acg ggc tgg acg cca ggc ctc cgc tcg ctg gag aac gac ggg act			3744
Thr Gly Trp Thr Pro Gly Leu Arg Ser Leu Glu Asn Asp Gly Thr			
1235	1240	1245	
aag ttg ctg gat cgc att cgc gac aat tac gaa ggc gaa ttc tac			3789
Lys Leu Leu Asp Arg Ile Arg Asp Asn Tyr Glu Gly Glu Phe Tyr			
1250	1255	1260	
tgg cgc tac ttt gcg ttt atc gcc gac gct ctg att acc acc ctg			3834

Trp Arg Tyr Phe Ala Phe Ile	Ala Asp Ala Leu Ile	Thr Thr Leu	
1265	1270	1275	
aag ccg cgg tac gag gac act	aac atc cgc atc aat	ctg gat agc	3879
Lys Pro Arg Tyr Glu Asp Thr	Asn Ile Arg Ile Asn	Leu Asp Ser	
1280	1285	1290	
aac act cgc agc ttc att gtc	ccg atc atc acc act	gaa tac att	3924
Asn Thr Arg Ser Phe Ile Val	Pro Ile Ile Thr Thr	Glu Tyr Ile	
1295	1300	1305	
cgc gaa aag ctg agc tac agc	ttt tac ggt agc ggt	ggt acc tat	3969
Arg Glu Lys Leu Ser Tyr Ser	Phe Tyr Gly Ser Gly	Gly Thr Tyr	
1310	1315	1320	
gcc ctg agc ctg tcc cag tac	aac atg ggc atc aac	att gaa ttg	4014
Ala Leu Ser Leu Ser Gln Tyr	Asn Met Gly Ile Asn	Ile Glu Leu	
1325	1330	1335	
tcg gag agc gat gtg tgg atc	att gac gtg gat aat	gtc gtc cgc	4059
Ser Glu Ser Asp Val Trp Ile	Ile Asp Val Asp Asn	Val Val Arg	
1340	1345	1350	
gat gtg acc att gag agc gac	aaa atc aag aaa ggg	gat ttg atc	4104
Asp Val Thr Ile Glu Ser Asp	Lys Ile Lys Lys Gly	Asp Leu Ile	
1355	1360	1365	
gaa ggc atc ctg tcg acc ttg	agc atc gaa gag aat	aag atc atc	4149
Glu Gly Ile Leu Ser Thr Leu	Ser Ile Glu Glu Asn	Lys Ile Ile	
1370	1375	1380	
ctg aat agc cac gaa atc aac	ttc tcc ggc gaa gtg	aac ggc agc	4194
Leu Asn Ser His Glu Ile Asn	Phe Ser Gly Glu Val	Asn Gly Ser	
1385	1390	1395	
aac ggc ttc gtc agc ctg acc	ttt tcc atc ctg gaa	ggt atc aac	4239
Asn Gly Phe Val Ser Leu Thr	Phe Ser Ile Leu Glu	Gly Ile Asn	
1400	1405	1410	
gcc atc att gag gtc gat ctg	ctc agc aaa agc tac	aag ctg ctc	4284
Ala Ile Ile Glu Val Asp Leu	Leu Ser Lys Ser Tyr	Lys Leu Leu	

1415	1420	1425	
atc tcg ggc gag ctg aaa atc	ctc atg ttg aat tcg	aat cac atc	4329
Ile Ser Gly Glu Leu Lys Ile	Leu Met Leu Asn Ser	Asn His Ile	
1430	1435	1440	
caa cag aaa atc gac tat atc	ggc ttc aac agc gag	ctg caa aag	4374
Gln Gln Lys Ile Asp Tyr Ile	Gly Phe Asn Ser Glu	Leu Gln Lys	
1445	1450	1455	
aac atc ccc tac tcg ttt gtg	gac agc gaa ggc aaa	gag aac ggc	4419
Asn Ile Pro Tyr Ser Phe Val	Asp Ser Glu Gly Lys	Glu Asn Gly	
1460	1465	1470	
ttc atc aac ggc tcc acc aaa	gag ggg ctg ttc gtg	agc gag ctg	4464
Phe Ile Asn Gly Ser Thr Lys	Glu Gly Leu Phe Val	Ser Glu Leu	
1475	1480	1485	
ccg gac gtg gtg ctc atc agc	aaa gtg tat atg gac	gac agc aag	4509
Pro Asp Val Val Leu Ile Ser	Lys Val Tyr Met Asp	Asp Ser Lys	
1490	1495	1500	
cct agt ttt ggc tac tac tcc	aat aac ctg aaa gat	gtt aaa gtg	4554
Pro Ser Phe Gly Tyr Tyr Ser	Asn Asn Leu Lys Asp	Val Lys Val	
1505	1510	1515	
att acc aag gat aac gtg aac	atc ctc acc ggt tac	tac ctg aag	4599
Ile Thr Lys Asp Asn Val Asn	Ile Leu Thr Gly Tyr	Tyr Leu Lys	
1520	1525	1530	
gac gac att aag atc agc ctg	tcc ctg acc ctg caa	gat gaa aag	4644
Asp Asp Ile Lys Ile Ser Leu	Ser Leu Thr Leu Gln	Asp Glu Lys	
1535	1540	1545	
acc atc aaa ttg aat agc gtc	cat ctc gac gag agt	ggg gtc gcc	4689
Thr Ile Lys Leu Asn Ser Val	His Leu Asp Glu Ser	Gly Val Ala	
1550	1555	1560	
gag atc ttg aag ttc atg aat	cgc aag ggt aat acc	aac acg agc	4734
Glu Ile Leu Lys Phe Met Asn	Arg Lys Gly Asn Thr	Asn Thr Ser	
1565	1570	1575	

gac tgc ttg atg tcc ttc ctg gaa agc atg aac atc aag agc att	4779
Asp Ser Leu Met Ser Phe Leu Glu Ser Met Asn Ile Lys Ser Ile	
1580 1585 1590	
ttc gtg aac ttc ctg caa agc aat atc aag ttc att ttg gac gcg	4824
Phe Val Asn Phe Leu Gln Ser Asn Ile Lys Phe Ile Leu Asp Ala	
1595 1600 1605	
aac ttt atc atc tcc ggg acc acg tcc atc ggc cag ttt gag ttc	4869
Asn Phe Ile Ile Ser Gly Thr Thr Ser Ile Gly Gln Phe Glu Phe	
1610 1615 1620	
atc tgt gac gag aac gac aac att cag ccg tat ttc atc aag ttc	4914
Ile Cys Asp Glu Asn Asp Asn Ile Gln Pro Tyr Phe Ile Lys Phe	
1625 1630 1635	
aat acc ttg gaa act aac tac acc ctg tac gtt ggc aac cgg caa	4959
Asn Thr Leu Glu Thr Asn Tyr Thr Leu Tyr Val Gly Asn Arg Gln	
1640 1645 1650	
aac atg att gtc gaa ccc aac tat gac ttg gat gac agt ggt gat	5004
Asn Met Ile Val Glu Pro Asn Tyr Asp Leu Asp Asp Ser Gly Asp	
1655 1660 1665	
atc agt agc acc gtg att aac ttt tcc cag aag tac ctg tat ggc	5049
Ile Ser Ser Thr Val Ile Asn Phe Ser Gln Lys Tyr Leu Tyr Gly	
1670 1675 1680	
atc gac tcc tgc gtg aac aaa gtg gtg atc tgc ccg aat atc tac	5094
Ile Asp Ser Cys Val Asn Lys Val Val Ile Ser Pro Asn Ile Tyr	
1685 1690 1695	
acg gac gaa atc aat atc act ccc gtc tat gaa acc aac aac acc	5139
Thr Asp Glu Ile Asn Ile Thr Pro Val Tyr Glu Thr Asn Asn Thr	
1700 1705 1710	
tac ccc gag gtg att gtc ttg gat gcg aac tac att aac gaa aag	5184
Tyr Pro Glu Val Ile Val Leu Asp Ala Asn Tyr Ile Asn Glu Lys	

1715	1720	1725	
att aac gtg aac att aac gac	ctg agc atc cgg tat	gtg tgg agt	5229
Ile Asn Val Asn Ile Asn Asp	Leu Ser Ile Arg Tyr	Val Trp Ser	
1730	1735	1740	
aat gac ggg aac gac ttc att	ctg atg agc acc tcc	gaa gaa aac	5274
Asn Asp Gly Asn Asp Phe Ile	Leu Met Ser Thr Ser	Glu Glu Asn	
1745	1750	1755	
aaa gtc tcg caa gtc aag atc	cgc ttc gtt aac gtt	ttc aaa gac	5319
Lys Val Ser Gln Val Lys Ile	Arg Phe Val Asn Val	Phe Lys Asp	
1760	1765	1770	
aag acc ttg gcc aac aaa ctc	agc ttc aat ttc tcg	gac aaa cag	5364
Lys Thr Leu Ala Asn Lys Leu	Ser Phe Asn Phe Ser	Asp Lys Gln	
1775	1780	1785	
gac gtg cct gtg tcg gag atc	att ctc agt ttc acc	ccg agc tac	5409
Asp Val Pro Val Ser Glu Ile	Ile Leu Ser Phe Thr	Pro Ser Tyr	
1790	1795	1800	
tac gag gac ggc ctg atc ggt	tac gac ctg ggc ctg	gtt agc ctc	5454
Tyr Glu Asp Gly Leu Ile Gly	Tyr Asp Leu Gly Leu	Val Ser Leu	
1805	1810	1815	
tac aac gaa aag ttc tat atc	aac aat ttc ggg atg	atg gtt tcg	5499
Tyr Asn Glu Lys Phe Tyr Ile	Asn Asn Phe Gly Met	Met Val Ser	
1820	1825	1830	
ggt ctg atc tat atc aat gac	agc ttg tac tac ttc	aaa cct ccg	5544
Gly Leu Ile Tyr Ile Asn Asp	Ser Leu Tyr Tyr Phe	Lys Pro Pro	
1835	1840	1845	
gtg aac aat ttg atc acc ggc	ttc gtg acc gtg ggc	gat gac aag	5589
Val Asn Asn Leu Ile Thr Gly	Phe Val Thr Val Gly	Asp Asp Lys	
1850	1855	1860	
tac tat ttc aac ccg att aac	ggt ggc gct gcg tcg	att ggc gaa	5634
Tyr Tyr Phe Asn Pro Ile Asn	Gly Gly Ala Ala Ser	Ile Gly Glu	
1865	1870	1875	

acc atc atc gac gac aag aac tac tat ttc aac cag agc ggt gtg	5679
Thr Ile Ile Asp Asp Lys Asn Tyr Tyr Phe Asn Gln Ser Gly Val	
1880 1885 1890	
ctg caa acc ggg gtc ttt agc acg gag gat ggg ttt aag tat ttc	5724
Leu Gln Thr Gly Val Phe Ser Thr Glu Asp Gly Phe Lys Tyr Phe	
1895 1900 1905	
gcg cct gcc aat acc ctg gac gag aat ttg gag ggc gaa gcc att	5769
Ala Pro Ala Asn Thr Leu Asp Glu Asn Leu Glu Gly Glu Ala Ile	
1910 1915 1920	
gac ttc acc ggc aaa ctg att atc gac gaa aac atc tac tat ttc	5814
Asp Phe Thr Gly Lys Leu Ile Ile Asp Glu Asn Ile Tyr Tyr Phe	
1925 1930 1935	
gat gac aac tac cgc ggt gcg gtc gaa tgg aaa gag ttg gac ggg	5859
Asp Asp Asn Tyr Arg Gly Ala Val Glu Trp Lys Glu Leu Asp Gly	
1940 1945 1950	
gag atg cac tat ttc tcg cca gag act ggt aag gcc ttc aag ggc	5904
Glu Met His Tyr Phe Ser Pro Glu Thr Gly Lys Ala Phe Lys Gly	
1955 1960 1965	
ctg aac cag atc ggc gac tac aag tac tat ttc aat agc gac ggc	5949
Leu Asn Gln Ile Gly Asp Tyr Lys Tyr Tyr Phe Asn Ser Asp Gly	
1970 1975 1980	
gtg atg cag aag ggg ttc gtc agc atc aac gac aac aag cat tac	5994
Val Met Gln Lys Gly Phe Val Ser Ile Asn Asp Asn Lys His Tyr	
1985 1990 1995	
ttc gac gat agc ggc gtt atg aag gtg ggc tat act gag atc gac	6039
Phe Asp Asp Ser Gly Val Met Lys Val Gly Tyr Thr Glu Ile Asp	
2000 2005 2010	
ggc aag cac ttc tat ttc gcc gag aac ggc gag atg caa atc ggc	6084
Gly Lys His Phe Tyr Phe Ala Glu Asn Gly Glu Met Gln Ile Gly	
2015 2020 2025	
gtg ttc aac acc gag gac ggc ttc aaa tac ttc gct cat cac aac	6129

Val Phe Asn Thr Glu Asp Gly Phe Lys Tyr Phe Ala His His Asn

2030	2035	2040	
gaa gat ctc ggt aat gaa gaa ggt gaa gag att tcc tat tcg ggc			6174
Glu Asp Leu Gly Asn Glu Glu Gly Glu Glu Ile Ser Tyr Ser Gly			
2045	2050	2055	
atc ctg aac ttc aac aat aag atc tac tac ttt gat gac tcg ttc			6219
Ile Leu Asn Phe Asn Asn Lys Ile Tyr Tyr Phe Asp Asp Ser Phe			
2060	2065	2070	
acc gcc gtg gtg ggt tgg aag gac ttg gag gac ggg agc aag tat			6264
Thr Ala Val Val Gly Trp Lys Asp Leu Glu Asp Gly Ser Lys Tyr			
2075	2080	2085	
tac ttc gac gaa gat acc gcc gag gca tac atc ggc ttg tcg ctc			6309
Tyr Phe Asp Glu Asp Thr Ala Glu Ala Tyr Ile Gly Leu Ser Leu			
2090	2095	2100	
atc aac gac ggt cag tac tac ttc aac gac gat ggc atc atg cag			6354
Ile Asn Asp Gly Gln Tyr Tyr Phe Asn Asp Asp Gly Ile Met Gln			
2105	2110	2115	
gtg ggc ttc gtg act atc aac gac aag gtg ttc tac ttc agt gac			6399
Val Gly Phe Val Thr Ile Asn Asp Lys Val Phe Tyr Phe Ser Asp			
2120	2125	2130	
agt ggc atc att gag tcg ggc gtg cag aat atc gac gat aac tat			6444
Ser Gly Ile Ile Glu Ser Gly Val Gln Asn Ile Asp Asp Asn Tyr			
2135	2140	2145	
ttc tat atc gat gac aat ggc att gtg cag atc ggc gtg ttt gat			6489
Phe Tyr Ile Asp Asp Asn Gly Ile Val Gln Ile Gly Val Phe Asp			
2150	2155	2160	
acc tcc gat ggg tat aag tat ttc gca cca gca aat acc gtc aat			6534
Thr Ser Asp Gly Tyr Lys Tyr Phe Ala Pro Ala Asn Thr Val Asn			
2165	2170	2175	

gac aac atc tac ggc cag gcc gtg gag tac agc ggg ctg gtt cgt	6579
Asp Asn Ile Tyr Gly Gln Ala Val Glu Tyr Ser Gly Leu Val Arg	
2180 2185 2190	
gtg ggc gag gac gtt tac tat ttc ggt gag act tat act atc gaa	6624
Val Gly Glu Asp Val Tyr Tyr Phe Gly Glu Thr Tyr Thr Ile Glu	
2195 2200 2205	
acc ggc tgg atc tat gat atg gaa aac gaa tcg gat aag tac tac	6669
Thr Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr	
2210 2215 2220	
ttt aac cca gaa acg aag aaa gcc tgc aag ggc atc aac ctc att	6714
Phe Asn Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile	
2225 2230 2235	
gat gac atc aag tat tac ttt gac gaa aag ggt atc atg cgc acc	6759
Asp Asp Ile Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr	
2240 2245 2250	
ggc ctg atc tcg ttt gag aac aac aac tac tat ttc aac gag aat	6804
Gly Leu Ile Ser Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn	
2255 2260 2265	
ggc gaa atg cag ttt ggg tac atc aat atc gag gat aag atg ttc	6849
Gly Glu Met Gln Phe Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe	
2270 2275 2280	
tac ttt ggg gag gac ggc gtc atg cag att ggt gtg ttt aac acc	6894
Tyr Phe Gly Glu Asp Gly Val Met Gln Ile Gly Val Phe Asn Thr	
2285 2290 2295	
ccg gat ggc ttc aag tac ttc gcc cat cag aac act ctg gac gaa	6939
Pro Asp Gly Phe Lys Tyr Phe Ala His Gln Asn Thr Leu Asp Glu	
2300 2305 2310	
aac ttc gag ggc gaa agc atc aat tac act ggc tgg ctg gac ctg	6984
Asn Phe Glu Gly Glu Ser Ile Asn Tyr Thr Gly Trp Leu Asp Leu	
2315 2320 2325	
gat gag aaa cgc tac tac ttc acc gac gag tac att gcc gcc acg	7029

Asp Glu Lys Arg Tyr Tyr Phe Thr Asp Glu Tyr Ile Ala Ala Thr
 2330 2335 2340
 ggc tcc gtg att atc gac ggc gaa gaa tac tat ttc gat ccc gac
 Gly Ser Val Ile Ile Asp Gly Glu Glu Tyr Tyr Phe Asp Pro Asp

7074

2345 2350 2355
 acc gcc cag ttg gtc att agc gaa
 Thr Ala Gln Leu Val Ile Ser Glu

7098

2360 2365

<210> 34

<211> 613

<212> PRT

<213> Pseudomonas aeruginosa

<220><221> MOD_RES

<222> (426)..(426)

<223> His or Tyr

<220><221> MOD_RES

<222> (553)..(553)

<223> May or may not be present

<400

> 34

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

Ser Ile Thr Ser Asp Gly Leu Thr Ile Arg Leu Glu Gly Gly Val Glu
 65 70 75 80
 Pro Asn Lys Pro Val Arg Tyr Ser Tyr Thr Arg Gln Ala Arg Gly Ser
 85 90 95
 Trp Ser Leu Asn Trp Leu Val Pro Ile Gly His Glu Lys Pro Ser Asn

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

Gln Gly Thr Gly Asn Asp Glu Ala Gly Ala Ala Ser Ala Asp Val Val
355 360 365

Ser Leu Thr Cys Pro Val Ala Ala Gly Glu Cys Ala Gly Pro Ala Asp
370 375 380

Ser Gly Asp Ala Leu Leu Glu Arg Asn Tyr Pro Thr Gly Ala Glu Phe
385 390 395 400

Leu Gly Asp Gly Gly Asp Ile Ser Phe Ser Thr Arg Gly Thr Gln Asn
405 410 415

Trp Thr Val Glu Arg Leu Leu Gln Ala Xaa Arg Gln Leu Glu Glu Arg
420 425 430

Gly Tyr Val Phe Val Gly Tyr His Gly Thr Phe Leu Glu Ala Ala Gln
435 440 445

Ser Ile Val Phe Gly Gly Val Arg Ala Arg Ser Gln Asp Leu Asp Ala
450 455 460

Ile Trp Arg Gly Phe Tyr Ile Ala Gly Asp Pro Ala Leu Ala Tyr Gly
465 470 475 480

Tyr Ala Gln Asp Gln Glu Pro Asp Ala Arg Gly Arg Ile Arg Asn Gly
485 490 495

Ala Leu Leu Arg Val Tyr Val Pro Arg Ser Ser Leu Pro Gly Phe Tyr
500 505 510

Arg Thr Gly Leu Thr Leu Ala Ala Pro Glu Ala Ala Gly Glu Val Glu
515 520 525

Arg Leu Ile Gly His Pro Leu Pro Leu Arg Leu Asp Ala Ile Thr Gly
530 535 540

Pro Glu Glu Glu Gly Gly Arg Leu Glu Thr Ile Leu Gly Trp Pro Leu
545 550 555 560

Ala Glu Arg Thr Val Val Ile Pro Ser Ala Ile Pro Thr Asp Pro Arg
565 570 575

Asn Val Gly Gly Asp Leu Asp Pro Ser Ser Ile Pro Asp Lys Glu Gln
580 585 590

Ala Ile Ser Ala Leu Pro Asp Tyr Ala Ser Gln Pro Gly Lys Pro Pro

595 Arg Glu Asp Leu Lys
610
<210> 35
<211> 3119
<212> DNA
<213> Bordetella pertussis
<220><221> CDS
<222> (1)..(807)
<220><221> CDS
<222> (850)..(1527)
<220><221> CDS
<222> (1587)..(1979)
<220><221> CDS
<222>
(1994)..(2353)
<220><221> CDS
<222> (2439)..(3119)
<400> 35
atg cgt tgc act cgg gca att cgc caa acc gca aga aca ggc tgg ctg 48
Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp Leu
1 5 10 15
acg tgg ctg gcg att ctt gcc gtc acg gcg ccc gtg act tcg ccg gca 96
Thr Trp Leu Ala Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala
20 25 30

tgg gcc gac gat cct ccc gcc acc gta tac cgc tat gac tcc cgc ccg 144
Trp Ala Asp Asp Pro Pro Ala Thr Val Tyr Arg Tyr Asp Ser Arg Pro
35 40 45
ccg gag gac gtt ttc cag aac gga ttc acg gcg tgg gga aac aac gac 192
Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp Gly Asn Asn Asp
50 55 60
aat gtg ctc gac cat ctg acc gga cgt tcc tgc cag gtc ggc agc agc 240

Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser

65	70	75	80	
aac agc gct ttc gtc tcc acc agc agc agc cgg cgc tat acc gag gtc				288
Asn Ser Ala Phe Val Ser Thr Ser Ser Ser Arg Arg Tyr Thr Glu Val				
	85	90	95	
tat ctc gaa cat cgc atg cag gaa gcg gtc gag gcc gaa cgc gcc ggc				336
Tyr Leu Glu His Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly				
	100	105	110	
agg ggc acc ggc cac ttc atc ggc tac atc tac gaa gtc cgc gcc gac				384
Arg Gly Thr Gly His Phe Ile Gly Tyr Ile Tyr Glu Val Arg Ala Asp				
	115	120	125	
aac aat ttc tac ggc gcc gcc agc tcg tac ttc gaa tac gtc gac act				432
Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu Tyr Val Asp Thr				
	130	135	140	
tat ggc gac aat gcc ggc cgt atc ctc gcc ggc gcg ctg gcc acc tac				480
Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr				
	145	150	155	160
cag agc gaa tat ctg gca cac cgg cgc att ccg ccc gaa aac atc cgc				528
Gln Ser Glu Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg				
	165	170	175	
agg gta acg cgg gtc tat cac aac ggc atc acc ggc gag acc acg acc				576
Arg Val Thr Arg Val Tyr His Asn Gly Ile Thr Gly Glu Thr Thr Thr				
	180	185	190	
acg gag tat tcc aac gct cgc tac gtc agc cag cag act cgc gcc aat				624
Thr Glu Tyr Ser Asn Ala Arg Tyr Val Ser Gln Gln Thr Arg Ala Asn				
	195	200	205	
ccc aac ccc tac aca tcg cga agg tcc gta gcg tcg atc gtc ggc aca				672
Pro Asn Pro Tyr Thr Ser Arg Arg Ser Val Ala Ser Ile Val Gly Thr				
	210	215	220	
ttg gtg cgc atg gcg ccg gtg ata ggc gct tgc atg gcg cgg cag gcc				720
Leu Val Arg Met Ala Pro Val Ile Gly Ala Cys Met Ala Arg Gln Ala				
	225	230	235	240

gaa agc tcc gag gcc atg gca gcc tgg tcc gaa cgc gcc ggc gag gcg	768
Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala	
245 250 255	
atg gtt ctc gtg tac tac gaa agc atc gcg tat tcg ttc tagacctggc	817
Met Val Leu Val Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe	
260 265	
ccagccccgc ccaactccgg taattgaaca gc atg ccg atc gac cgc aag acg	870
Met Pro Ile Asp Arg Lys Thr	
270 275	
ctc tgc cat ctc ctg tcc gtt ctg ccg ttg gcc ctc ctc gga tct cac	918
Leu Cys His Leu Leu Ser Val Leu Pro Leu Ala Leu Leu Gly Ser His	
280 285 290	
gtg gcg cgg gcc tcc acg cca ggc atc gtc att ccg ccg cag gaa cag	966
Val Ala Arg Ala Ser Thr Pro Gly Ile Val Ile Pro Pro Gln Glu Gln	
295 300 305	
att acc cag cat ggc agc ccc tat gga cgc tgc gcg aac aag acc cgt	1014
Ile Thr Gln His Gly Ser Pro Tyr Gly Arg Cys Ala Asn Lys Thr Arg	
310 315 320	
gcc ctg acc gtg gcg gaa ttg cgc ggc agc ggc gat ctg cag gag tac	1062
Ala Leu Thr Val Ala Glu Leu Arg Gly Ser Gly Asp Leu Gln Glu Tyr	
325 330 335 340	
ctg cgt cat gtg acg cgc ggc tgg tca ata ttt gcg ctc tac gat ggc	1110
Leu Arg His Val Thr Arg Gly Trp Ser Ile Phe Ala Leu Tyr Asp Gly	
345 350 355	
acc tat ctc ggc ggc gaa tat ggc ggc gtg atc aag gac gga aca ccc	1158
Thr Tyr Leu Gly Gly Glu Tyr Gly Gly Val Ile Lys Asp Gly Thr Pro	
360 365 370	
ggc ggc gca ttc gac ctg aaa acg acg ttc tgc atc atg acc acg cgc	1206
Gly Gly Ala Phe Asp Leu Lys Thr Thr Phe Cys Ile Met Thr Thr Arg	
375 380 385	

aat acg ggt caa ccc gca acg gat cac tac tac agc aac gtc acc gcc 1254

Asn Thr Gly Gln Pro Ala Thr Asp His Tyr Tyr Ser Asn Val Thr Ala
390 395 400

act cgc ctg ctc tcc agc acc aac agc agg cta tgc gcg gtc ttc gtc 1302

Thr Arg Leu Leu Ser Ser Thr Asn Ser Arg Leu Cys Ala Val Phe Val
405 410 415 420

aga agc ggg caa ccg gtc att ggc gcc tgc acc agc ccg tat gac ggc 1350

Arg Ser Gly Gln Pro Val Ile Gly Ala Cys Thr Ser Pro Tyr Asp Gly

425 430 435

aag tac tgg agc atg tac agc cgg ctg cgg aaa atg ctt tac ctg atc 1398

Lys Tyr Trp Ser Met Tyr Ser Arg Leu Arg Lys Met Leu Tyr Leu Ile
440 445 450

tac gtg gcc ggc atc tcc gta cgc gtc cat gtc agc aag gaa gaa cag 1446

Tyr Val Ala Gly Ile Ser Val Arg Val His Val Ser Lys Glu Glu Gln
455 460 465

tat tac gac tat gag gac gca acg ttc gag act tac gcc ctt acc ggc 1494

Tyr Tyr Asp Tyr Glu Asp Ala Thr Phe Glu Thr Tyr Ala Leu Thr Gly
470 475 480

atc tcc atc tgc aat cct gga tca tcc tta tgc tgagacgctt cccactcga 1547

Ile Ser Ile Cys Asn Pro Gly Ser Ser Leu Cys
485 490 495

accaccgccc cgggacaggg cggcgcccg cggtcgcgc gtg cgc gcc ctg gcg 1601

Val Arg Ala Leu Ala
500

tgg ttg ctg gca tcc ggc gcg atg acg cat ctt tcc ccc gcc ctg gcc 1649

Trp Leu Leu Ala Ser Gly Ala Met Thr His Leu Ser Pro Ala Leu Ala
505 510 515

gac gtt cct tat gtg ctg gtg aag acc aat atg gtg gtc acc agc gta 1697

Asp Val Pro Tyr Val Leu Val Lys Thr Asn Met Val Val Thr Ser Val

520	525	530	
gcc atg aag ccg tat gaa gtc acc ccg acg cgc atg ctg gtc tgc ggc			1745
Ala Met Lys Pro Tyr Glu Val Thr Pro Thr Arg Met Leu Val Cys Gly			
535	540	545	
atc gcc gcc aaa ctg ggc gcc gcg gcc agc agc ccg gac gcg cac gtg			1793
Ile Ala Ala Lys Leu Gly Ala Ala Ala Ser Ser Pro Asp Ala His Val			
550	555	560	
ccg ttc tgc ttc ggc aag gat ctc aag cgt ccc ggc agc agt ccc atg			1841
Pro Phe Cys Phe Gly Lys Asp Leu Lys Arg Pro Gly Ser Ser Pro Met			
565	570	575	580
gaa gtc atg ttg cgc gcc gtc ttc atg caa caa cgg ccg ctg cgc atg			1889
Glu Val Met Leu Arg Ala Val Phe Met Gln Gln Arg Pro Leu Arg Met			
585	590	595	
ttt ctg ggt ccc aag caa ctc act ttc gaa ggc aag ccc gcg ctc gaa			1937
Phe Leu Gly Pro Lys Gln Leu Thr Phe Glu Gly Lys Pro Ala Leu Glu			
600	605	610	
ctg atc cgg atg gtc gaa tgc agc ggc aag cag gat tgc ccc			1979
Leu Ile Arg Met Val Glu Cys Ser Gly Lys Gln Asp Cys Pro			
615	620	625	
tgaaggcgaa cccc atg cat acc atc gca tcc atc ctg ttg tcc gtg ctc			2029
Met His Thr Ile Ala Ser Ile Leu Leu Ser Val Leu			
630	635		
ggc ata tac agc ccg gct gac gtc gcc ggc ttg ccg acc cat ctg tac			2077
Gly Ile Tyr Ser Pro Ala Asp Val Ala Gly Leu Pro Thr His Leu Tyr			
640	645	650	
aag aac ttc act gtc cag gag ctg gcc ttg aaa ctg aag ggc aag aat			2125
Lys Asn Phe Thr Val Gln Glu Leu Ala Leu Lys Leu Lys Gly Lys Asn			
655	660	665	670
cag gag ttc tgc ctg acc gcc ttc atg tcg ggc aga agc ctg gtc cgg			2173
Gln Glu Phe Cys Leu Thr Ala Phe Met Ser Gly Arg Ser Leu Val Arg			
675	680	685	

gcg tgc ctg tcc gac gcg gga cac gag cac gac acg tgg ttc gac acc	2221
Ala Cys Leu Ser Asp Ala Gly His Glu His Asp Thr Trp Phe Asp Thr	
690 695 700	
atg ctt ggc ttt gcc ata tcc gcg tat gcg ctc aag agc cgg atc gcg	2269
Met Leu Gly Phe Ala Ile Ser Ala Tyr Ala Leu Lys Ser Arg Ile Ala	
705 710 715	
ctg acg gtg gaa gac tgc ccg tat ccg ggc act ccc ggc gat ctg ctc	2317
Leu Thr Val Glu Asp Ser Pro Tyr Pro Gly Thr Pro Gly Asp Leu Leu	
720 725 730	
gaa ctg cag atc tgc ccg ctc aac gga tat tgc gaa tgaacccttc	2363
Glu Leu Gln Ile Cys Pro Leu Asn Gly Tyr Cys Glu	
735 740 745	
cggagggtttc gacgtttccg cgcaatccgc ttgagacgat cttccgcct ggttcattc	2423
cgggaacacc gcaac atg ctg atc aac aac aag aag ctg ctt cat cac att	2474
Met Leu Ile Asn Asn Lys Lys Leu Leu His His Ile	
750 755	
ctg ccc atc ctg gtg ctc gcc ctg ctg ggc atg cgc acg gcc cag gcc	2522
Leu Pro Ile Leu Val Leu Ala Leu Leu Gly Met Arg Thr Ala Gln Ala	
760 765 770	
gtt gcg cca ggc atc gtc atc ccg ccg aag gca ctg ttc acc caa cag	2570
Val Ala Pro Gly Ile Val Ile Pro Pro Lys Ala Leu Phe Thr Gln Gln	
775 780 785 790	
ggc ggc gcc tat gga cgc tgc ccg aac gga acc cgc gcc ttg acc gtg	2618
Gly Gly Ala Tyr Gly Arg Cys Pro Asn Gly Thr Arg Ala Leu Thr Val	
795 800 805	
gcc gaa ctg cgc ggc aac gcc gaa ttg cag acg tat ttg cgc cag ata	2666
Ala Glu Leu Arg Gly Asn Ala Glu Leu Gln Thr Tyr Leu Arg Gln Ile	
810 815 820	
acg ccc ggc tgg tcc ata tac ggt ctc tat gac ggt acg tac ctg ggc	2714
Thr Pro Gly Trp Ser Ile Tyr Gly Leu Tyr Asp Gly Thr Tyr Leu Gly	
825 830 835	

cag gcg tac ggc ggc atc atc aag gac gcg ccg cca ggc gcg ggg ttc 2762

Gln Ala Tyr Gly Gly Ile Ile Lys Asp Ala Pro Pro Gly Ala Gly Phe

840 845 850

att tat cgc gaa act ttc tgc atc acg acc ata tac aag acc ggg caa 2810

Ile Tyr Arg Glu Thr Phe Cys Ile Thr Thr Ile Tyr Lys Thr Gly Gln

855 860 865 870

ccg gct gcg gat cac tac tac agc aag gtc acg gcc acg cgc ctg ctc 2858

Pro Ala Ala Asp His Tyr Tyr Ser Lys Val Thr Ala Thr Arg Leu Leu

875 880 885

gcc agc acc aac agc agg ctg tgc gcg gta ttc gtc agg gac ggg caa 2906

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

890 895 900

tcg gtc atc gga gcc tgc gcc agc ccg tat gaa ggc agg tac aga gac 2954

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

905 910 915

atg tac gac gcg ctg cgg cgc ctg ctg tac atg atc tat atg tcc ggc 3002

Met Tyr Asp Ala Leu Arg Arg Leu Leu Tyr Met Ile Tyr Met Ser Gly

920 925 930

ctt gcc gta cgc gtc cac gtc agc aag gag gaa cag tat tac gac tac 3050

Leu Ala Val Arg Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr

935 940 945 950

gag gac gcc aca ttc cag acc tat gcc ctc acc ggc att tcc ctc tgc 3098

Glu Asp Ala Thr Phe Gln Thr Tyr Ala Leu Thr Gly Ile Ser Leu Cys

955 960 965

aac ccg gca gcg tcg ata tgc 3119

Asn Pro Ala Ala Ser Ile Cys

970

<210> 36

<211> 535

<212> PRT

<213> Corynebacterium diphtheriae

<400> 36

Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn

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

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

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

195 200 205
Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly
210 215 220
Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu
225 230 235 240

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

485	490	495	
Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly Val Leu			
500	505	510	
Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys Leu Ser			
515	520	525	
Leu Phe Phe Glu Ile Lys Ser			
530	535		
<210> 37			
<211> 960			
<212> DNA			
<213> Artificial Sequence			
<220><223> Description of Artificial Sequence: Synthetic			
polynucleotide			
<220><221> CDS			
<222> (1)..(960)			
<400> 37			
ggg gcg gac gat gtg gtg gat tcc tcc aag tcg ttt gtc atg gaa aat	48		
Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn			
1 5 10 15			
ttc tcg tcg tac cat ggc act aag cca ggc tac gtg gat agc att caa	96		
Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser Ile Gln			
20 25 30			
aag ggc atc cag aag ccc aag agc ggt act cag ggg aac tat gac gac	144		
Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn Tyr Asp Asp			
35 40 45			
gac tgg aag gga ttt tac agc acc gac aat aag tac gat gct gcg ggc	192		
Asp Trp Lys Gly Phe Tyr Ser Thr Asp Asn Lys Tyr Asp Ala Ala Gly			
50 55 60			
tat agc gtg gac aac gaa aac cca ttg tcg ggc aag gcc ggt ggc gtg	240		
Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala Gly Gly Val			
65 70 75 80			
gtg aag gtg acc tat cct ggt ctg acg aaa gtt ctg gcg ttg aaa gtg	288		

Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu Lys Val

85	90	95	
gac aac gcc gag act atc aag aaa gaa ttg ggc ttg agt ttg acc gag			336
Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu Thr Glu			
100	105	110	
ccg ctg atg gaa cag gtg ggt acc gaa gaa ttc att aaa cgt ttt ggg			384
Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg Phe Gly			
115	120	125	
gac ggc gcg tcg cgc gtg gtc ctg tcg ttg ccg ttc gcc gaa ggg tcg			432
Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu Gly Ser			
130	135	140	
tcg tcg gtg gaa tat atc aac aac tgg gaa cag gcc aag gcg ctg tcg			480
Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys Ala Leu Ser			
145	150	155	160
gtg gaa ctg gaa att aac ttc gaa acg cgg ggc aaa cgg ggc cag gac			528
Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg Gly Gln Asp			
165	170	175	
gcc atg tac gaa tac atg gcg cag gcg tgc gcc ggg aac cgg gtg cgg			576
Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn Arg Val Arg			
180	185	190	
cgc agc gtg ggc agt tcc ttg agt tgc atc aat ctg gac tgg gac gtc			624
Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val			
195	200	205	
atc cgc gat aag acg aag acg aaa atc gag tcg ctc aaa gag cac ggc			672
Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly			
210	215	220	
ccg atc aaa aac aaa atg agc gag tcg ccg aat aaa acg gtg tcc gag			720
Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu			
225	230	235	240

gag aag gcg aag caa tac ctg gag gaa ttc cac cag acg gct ctg gag 768
 Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu
 245 250 255

cac ccg gag ctg agc gaa ctc aaa acc gtt acc ggt acg aac ccg gtg 816
 His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val
 260 265 270

ttt gcc ggg gca aac tat gca gct tgg gcc gtc aac gtg gcc caa gtg 864

Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val
 275 280 285

atc gac tcc gaa acg gcc gac aac ctg gaa aag act acc gcc gcg ttg 912
 Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu
 290 295 300

tcg atc ctc ccg ggc atc ggg agc gtc atg ggt att gcc gat ggt gcg 960
 Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala

305 310 315 320

<210> 38
 <211> 258
 <212> PRT
 <213> Vibrio cholerae
 <400> 38

Met Val Lys Ile Ile Phe Val Phe Phe Ile Phe Leu Ser Ser Phe Ser
 1 5 10 15

Tyr Ala Asn Asp Asp Lys Leu Tyr Arg Ala Asp Ser Arg Pro Pro Asp
 20 25 30

Glu Ile Lys Gln Ser Gly Gly Leu Met Pro Arg Gly Gln Ser Glu Tyr
 35 40 45

Phe Asp Arg Gly Thr Gln Met Asn Ile Asn Leu Tyr Asp His Ala Arg
 50 55 60

Gly Thr Gln Thr Gly Phe Val Arg His Asp Asp Gly Tyr Val Ser Thr
 65 70 75 80

Ser Ile Ser Leu Arg Ser Ala His Leu Val Gly Gln Thr Ile Leu Ser
 85 90 95

Gly His Ser Thr Tyr Tyr Ile Tyr Val Ile Ala Thr Ala Pro Asn Met
100 105 110

Phe Asn Val Asn Asp Val Leu Gly Ala Tyr Ser Pro His Pro Asp Glu
115 120 125

Gln Glu Val Ser Ala Leu Gly Gly Ile Pro Tyr Ser Gln Ile Tyr Gly
130 135 140

Trp Tyr Arg Val His Phe Gly Val Leu Asp Glu Gln Leu His Arg Asn
145 150 155 160

Arg Gly Tyr Arg Asp Arg Tyr Tyr Ser Asn Leu Asp Ile Ala Pro Ala
165 170 175

Ala Asp Gly Tyr Gly Leu Ala Gly Phe Pro Pro Glu His Arg Ala Trp
180 185 190

Arg Glu Glu Pro Trp Ile His His Ala Pro Pro Gly Cys Gly Asn Ala
195 200 205

Pro Arg Ser Ser Met Ser Asn Thr Cys Asp Glu Lys Thr Gln Ser Leu
210 215 220

Gly Val Lys Phe Leu Asp Glu Tyr Gln Ser Lys Val Lys Arg Gln Ile
225 230 235 240

Phe Ser Gly Tyr Gln Ser Asp Ile Asp Thr His Asn Arg Ile Lys Asp
245 250 255

Glu Leu

<210> 39

<211> 124

<212> PRT

<213> *Vibrio cholerae*

<400> 39

Met Ile Lys Leu Lys Phe Gly Val Phe Phe Thr Val Leu Leu Ser Ser
1 5 10 15

Ala Tyr Ala His Gly Thr Pro Gln Asn Ile Thr Asp Leu Cys Ala Glu
20 25 30

Tyr His Asn Thr Gln Ile Tyr Thr Leu Asn Asp Lys Ile Phe Ser Tyr

35

40

45

Thr Glu Ser Leu Ala Gly Lys Arg Glu Met Ala Ile Ile Thr Phe Lys

50

55

60

Asn Gly Ala Ile Phe Gln Val Glu Val Pro Gly Ser Gln His Ile Asp

65

70

75

80

Ser Gln Lys Lys Ala Ile Glu Arg Met Lys Asp Thr Leu Arg Ile Ala

85

90

95

Tyr Leu Thr Glu Ala Lys Val Glu Lys Leu Cys Val Trp Asn Asn Lys

100

105

110

Thr Pro His Ala Ile Ala Ala Ile Ser Met Ala Asn

115

120

<210> 40

<211> 1148

<212> DNA

<213> *Vibrio cholerae*

<400> 40

atggtaaaga taatatttgt gttttttatt ttcttatcat cattttcata tgcaaatgat 60

gataagttat atcgggcaga ttctagacct cctgatgaaa taaagcagtc aggtggcttt 120

atgccaagag gacagagtga gtactttgac cgagggtactc aaatgaatat caacctttat 180

gatcatgcaa gaggaactca gacgggatit gttaggcacg atgatggata tgtttccacc 240

tcaattagtt tgagaagtgc ccacttagtg ggtcaaacta tattgtctgg tcattctact 300

tattatata atgttatagc cactgcaccc aacatgttta acgttaatga tgtattaggg 360

gcatacagtc ctcacccaga tgaacaagaa gtttctgctt taggtgggat tccatactcc 420

caaatatatg gatggtatcg agttcatitt ggggtgcttg atgaacaatt acatcgtaat 480

aggggctaca gagatagata ttacagtaac ttagatattg ctccagcagc agatggttat 540

ggattggcag gtttcctcc ggagcataga gcttggaggg aagagccgtg gattcatcat 600

gcaccgccgg gttgtgggaa tgctccaaga tcatgatga gtaatacttg cgatgaaaaa 660

acccaaagtc taggtgtaaa attccttgac gaataccaat ctaaagttaa aagacaaata 720

ttttcaggct atcaatctga tattgatata cataatagaa ttaaggatga attatgatta 780

aattaaaatt tgggtgtttt ttacagttt tactatcttc agcatatgca catggaacac 840

ctcaaaatat tactgatttg tgtgcagaat accacaacac acaaatatat acgctaaatg 900
 ataagatatatt ttcgtataca gaatctctag ctggaaaaag agagatggct atcattactt 960
 ttaagaatgg tgcaattttt caagtagaag taccaggttag tcaacatata gattcacaaa 1020
 aaaaagcgat tgaaaggatg aaggataccc tgaggattgc atatcttact gaagctaaag 1080

tcgaaaagtt atgtgtatgg aataataaaa cgcctcatgc gattgccgca attagtatgg 1140
 caaatataa 1148

<210> 41

<211> 269

<212> PRT

<213> Bordetella pertussis

<400> 41

Met Arg Cys Thr Arg Ala Ile Arg Gln Thr Ala Arg Thr Gly Trp Leu

1 5 10 15

Thr Trp Leu Ala Ile Leu Ala Val Thr Ala Pro Val Thr Ser Pro Ala

20 25 30

Trp Ala Asp Asp Pro Pro Ala Thr Val Tyr Arg Tyr Asp Ser Arg Pro

35 40 45

Pro Glu Asp Val Phe Gln Asn Gly Phe Thr Ala Trp Gly Asn Asn Asp

50 55 60

Asn Val Leu Asp His Leu Thr Gly Arg Ser Cys Gln Val Gly Ser Ser

65 70 75 80

Asn Ser Ala Phe Val Ser Thr Ser Ser Arg Arg Tyr Thr Glu Val

85 90 95

Tyr Leu Glu His Arg Met Gln Glu Ala Val Glu Ala Glu Arg Ala Gly

100 105 110

Arg Gly Thr Gly His Phe Ile Gly Tyr Ile Tyr Glu Val Arg Ala Asp

115 120 125

Asn Asn Phe Tyr Gly Ala Ala Ser Ser Tyr Phe Glu Tyr Val Asp Thr

130 135 140

Tyr Gly Asp Asn Ala Gly Arg Ile Leu Ala Gly Ala Leu Ala Thr Tyr

145 150 155 160

Gln Ser Glu Tyr Leu Ala His Arg Arg Ile Pro Pro Glu Asn Ile Arg

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

225 230 235 240
Glu Ser Ser Glu Ala Met Ala Ala Trp Ser Glu Arg Ala Gly Glu Ala
245 250 255
Met Val Leu Val Tyr Tyr Glu Ser Ile Ala Tyr Ser Phe
260 265

<210> 42

<211> 226

<212> PRT

<213> Bordetella pertussis

<400> 42

Met Pro Ile Asp Arg Lys Thr Leu Cys His Leu Leu Ser Val Leu Pro
1 5 10 15

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

Ile Phe Ala Leu Tyr Asp Gly Thr Tyr Leu Gly Gly Glu Tyr Gly Gly
85 90 95

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

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

Glu Thr Tyr Ala Leu Thr Gly Ile Ser Ile Cys Asn Pro Gly Ser Ser
 210 215 220
 Leu Cys
 225
 <210> 43
 <211> 131
 <212> PRT
 <213> Bordetella pertussis
 <400> 43

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

65 70 75 80
 Gly Ser Ser Pro Met Glu Val Met Leu Arg Ala Val Phe Met Gln Gln
 85 90 95
 Arg Pro Leu Arg Met Phe Leu Gly Pro Lys Gln Leu Thr Phe Glu Gly

 100 105 110
 Lys Pro Ala Leu Glu Leu Ile Arg Met Val Glu Cys Ser Gly Lys Gln
 115 120 125
 Asp Cys Pro

130
 <210> 44

<211> 120

<212> PRT

<213> Bordetella pertussis

<400> 44

Met His Thr Ile Ala Ser Ile Leu Leu Ser Val Leu Gly Ile Tyr Ser
 1 5 10 15
 Pro Ala Asp Val Ala Gly Leu Pro Thr His Leu Tyr Lys Asn Phe Thr
 20 25 30

Val Gln Glu Leu Ala Leu Lys Leu Lys Gly Lys Asn Gln Glu Phe Cys
 35 40 45
 Leu Thr Ala Phe Met Ser Gly Arg Ser Leu Val Arg Ala Cys Leu Ser
 50 55 60

Asp Ala Gly His Glu His Asp Thr Trp Phe Asp Thr Met Leu Gly Phe
 65 70 75 80
 Ala Ile Ser Ala Tyr Ala Leu Lys Ser Arg Ile Ala Leu Thr Val Glu
 85 90 95

Asp Ser Pro Tyr Pro Gly Thr Pro Gly Asp Leu Leu Glu Leu Gln Ile
 100 105 110
 Cys Pro Leu Asn Gly Tyr Cys Glu
 115 120

<210> 45

<211> 227

<212> PRT

<213> Bordetella pertussis

<400> 45

Met Leu Ile Asn Asn Lys Lys Leu Leu His His Ile Leu Pro Ile Leu

1 5 10 15

Val Leu Ala Leu Leu Gly Met Arg Thr Ala Gln Ala Val Ala Pro Gly

20 25 30

Ile Val Ile Pro Pro Lys Ala Leu Phe Thr Gln Gln Gly Gly Ala Tyr

35 40 45

Gly Arg Cys Pro Asn Gly Thr Arg Ala Leu Thr Val Ala Glu Leu Arg

50 55 60

Gly Asn Ala Glu Leu Gln Thr Tyr Leu Arg Gln Ile Thr Pro Gly Trp

65 70 75 80

Ser Ile Tyr Gly Leu Tyr Asp Gly Thr Tyr Leu Gly Gln Ala Tyr Gly

85 90 95

Gly Ile Ile Lys Asp Ala Pro Pro Gly Ala Gly Phe Ile Tyr Arg Glu

100 105 110

Thr Phe Cys Ile Thr Thr Ile Tyr Lys Thr Gly Gln Pro Ala Ala Asp

115 120 125

His Tyr Tyr Ser Lys Val Thr Ala Thr Arg Leu Leu Ala Ser Thr Asn

130 135 140

Ser Arg Leu Cys Ala Val Phe Val Arg Asp Gly Gln Ser Val Ile Gly

145 150 155 160

Ala Cys Ala Ser Pro Tyr Glu Gly Arg Tyr Arg Asp Met Tyr Asp Ala

165 170 175

Leu Arg Arg Leu Leu Tyr Met Ile Tyr Met Ser Gly Leu Ala Val Arg

180 185 190

Val His Val Ser Lys Glu Glu Gln Tyr Tyr Asp Tyr Glu Asp Ala Thr

195 200 205

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

210 215 220

Ser Ile Cys

225

<210> 46

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

6xHis tag

<400> 46

His His His His His His

1 5