



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년02월10일

(11) 등록번호 10-1491867

(24) 등록일자 2015년02월03일

(51) 국제특허분류(Int. Cl.)

C12N 15/62 (2006.01) C12N 15/67 (2006.01)

(21) 출원번호 10-2009-7017960

(22) 출원일자(국제) 2008년01월30일

심사청구일자 2013년01월30일

(85) 번역문제출일자 2009년08월28일

(65) 공개번호 10-2009-0114422

(43) 공개일자 2009년11월03일

(86) 국제출원번호 PCT/US2008/052434

(87) 국제공개번호 WO 2008/094986

국제공개일자 2008년08월07일

(30) 우선권주장

60/887,476 2007년01월31일 미국(US)

60/887,486 2007년01월31일 미국(US)

(56) 선행기술조사문헌

WO2005089093 A2*

ANDREAS URBAN 등. J. BACTERIOL. Vol. 183, No.

2, 페이지 587-596 (2001)*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

피페넥스 임크.

미국 캘리포니아주 92121 샌디에고 로젤레 스트리트 10790

(72) 발명자

클랜, 러셀, 존

미국 92101 캘리포니아주 샌디에고 넘버519 아일랜드 애비뉴 1050

레탈락, 디안

미국 92064 캘리포니아주 포웨이 캐리지 로드 13420

(뒷면에 계속)

(74) 대리인

양영준, 김영

전체 청구항 수 : 총 27 항

심사관 : 최준호

(54) 발명의 명칭 증가된 발현을 위한 박테리아 리더 서열

(57) 요약

숙주 세포에서 관심 단백질 또는 폴리펩티드의 발현 및/또는 분비를 개선시키기 위한 조성물 및 방법이 제공된다. 박테리아 분비 신호 웨პ티드에 대한 코딩 서열을 포함하는 조성물이 제공된다. 숙주 세포에서의 관심 단백질 또는 폴리펩티드의 형질전환 및 발현을 위한 벡터 구축물 또는 발현 시스템에서 상기 코딩 서열을 사용할 수 있다. 본 발명의 조성물은 숙주 세포의 주변 세포질 내에 적절하게 프로세싱된 단백질의 촉매를 증가시키는데, 또는 숙주 세포로부터 적절하게 프로세싱된 단백질의 분비를 증가시키는데 유용하다. 특히, 분비 신호 웨პ티드는 단백질 디슬퍼드 이소머라제 C (dsbC), 돌연변이 포스페이트 결합 단백질 (pbp*), 단백질 디슬퍼드 이소머라제 A (dsbA), CupA2, CupB2, CupC2, NikA, FlgI, 테트라트리코펩티드 반복 패밀리 단백질 (ORF5550), 틀루엔 내성 단백질 (Ttg2C), 또는 메틸 수용성 화학주성 단백질 (ORF8124) 분비 신호를 포함한다.

(72) 발명자

쉬나이더, 제인, 씨.

미국 92104 캘리포니아주 샌디에고 서티세컨드 스트리트 2621

람세이어, 토마스, 엠.

미국 92064 캘리포니아주 포웨이 포마드 웨이 13033

허쉬버거, 찰스, 더글라스

미국 92064 캘리포니아주 포웨이 암소 스트리트 15042

리, 스테이시, 엘.

미국 92129 캘리포니아주 샌디에고 파세오 카디엘 13831

레스너, 솔

미국 92024 캘리포니아주 엔시니타스 나르도 로드 818

특허청구의 범위

청구항 1

- a) 서열 17의 뉴클레오티드 서열을 포함하는 핵산 분자; 및
- b) 서열 18의 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자

로 구성된 군으로부터 선택되는, FlgI 분비 폴리펩티드에 대한 분비 신호 서열을 코딩하는 단리된 핵산 분자.

청구항 2

FlgI 분비 폴리펩티드에 대한 분비 신호 코딩 서열을 포함하며,

상기 분비 신호 코딩 서열은

- a) 서열 17의 뉴클레오티드 서열을 포함하는 핵산 분자; 및
- b) 서열 18의 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자

로 구성된 군으로부터 선택되는 것인, 벡터.

청구항 3

FlgI 분비 폴리펩티드에 대한 분비 신호 코딩 서열을 포함하며,

상기 분비 신호 코딩 서열은

- a) 서열 17의 뉴클레오티드 서열을 포함하는 핵산 분자; 및
- b) 서열 18의 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자

로 구성된 군으로부터 선택되는 것인, 재조합 세포.

청구항 4

i) 숙주 세포; 및

ii) FlgI 분비 신호 폴리펩티드에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 코딩하는 핵산 분자를 포함하는 벡터를 포함하며,

상기 핵산 분자는

- a) 서열 17의 뉴클레오티드 서열을 포함하는 핵산 분자; 및
- b) 서열 18의 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자

로 구성된 군으로부터 선택되는 것인,

관심 단백질 또는 폴리펩티드의 발현을 위한 발현 조성물.

청구항 5

FlgI 분비 신호 폴리펩티드에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 코딩하는 벡터를 포함하는 숙주 세포를 제공하는 것을 포함하며,

상기 분비 신호 폴리펩티드는

- a) 서열 17의 뉴클레오티드 서열을 포함하는 핵산 분자; 및
- b) 서열 18의 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자

로 구성된 군으로부터 선택되는 핵산 분자에 의해 코딩되는 것인,

숙주 세포에서 재조합 단백질을 발현시키는 방법.

청구항 6

제1항에 있어서, 핵산 분자를 발현하도록 선택된 숙주 생물의 코돈 선호도를 반영하도록 조정된 핵산 분자.

청구항 7

제2항에 있어서, 상기 핵산 분자가, 핵산 분자를 발현하도록 선택된 숙주 생물의 코돈 선호도를 반영하도록 조정된 것인 벡터.

청구항 8

제4항에 있어서, 관심 단백질 또는 폴리펩티드가, 관심 단백질 또는 폴리펩티드가 발현되는 숙주 생물에 대해 천연인 발현 조성물.

청구항 9

제5항에 있어서, 관심 단백질 또는 폴리펩티드가, 관심 단백질 또는 폴리펩티드가 발현되는 숙주 생물에 대해 천연인 방법.

청구항 10

제4항에 있어서, 관심 단백질 또는 폴리펩티드가 슈도모나스(*Pseudomonas*) 종에 대해 천연인 발현 조성물.

청구항 11

제5항에 있어서, 관심 단백질 또는 폴리펩티드가 슈도모나스 종에 대해 천연인 방법.

청구항 12

제4항에 있어서, 관심 단백질 또는 폴리펩티드가 슈도모나스 플루오레센스(*Pseudomonas fluorescens*)에 대해 천연인 발현 조성물.

청구항 13

제5항에 있어서, 관심 단백질 또는 폴리펩티드가 슈도모나스 플루오레센스에 대해 천연인 방법.

청구항 14

제4항에 있어서, 관심 단백질 또는 폴리펩티드가, 관심 단백질 또는 폴리펩티드가 발현되는 숙주 생물에 대해 천연이 아닌 단백질 또는 폴리펩티드로부터 수득되는 것인 발현 조성물.

청구항 15

제5항에 있어서, 관심 단백질 또는 폴리펩티드가, 관심 단백질 또는 폴리펩티드가 발현되는 숙주 생물에 대해 천연이 아닌 단백질 또는 폴리펩티드로부터 수득되는 것인 방법.

청구항 16

제1항에 있어서, 프로모터를 추가로 포함하는 핵산 분자.

청구항 17

제2항에 있어서, 프로모터를 추가로 포함하는 벡터.

청구항 18

제1항에 있어서, 분비 신호 서열이 관심 단백질 또는 폴리펩티드를 코딩하는 서열에 작동가능하게 연결된 것인 핵산 분자.

청구항 19

제2항에 있어서, 분비 신호 코딩 서열이 관심 단백질 또는 폴리펩티드를 코딩하는 서열에 작동가능하게 연결된 것인 벡터.

청구항 20

제3항에 있어서, 분비 신호 코딩 서열이 관심 단백질 또는 폴리펩티드를 코딩하는 서열에 작동가능하게 연결된 것인 세포.

청구항 21

제20항에 있어서, 분비 신호 폴리펩티드에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 발현하는 세포.

청구항 22

제4항에 있어서, 세포가 분비 신호 폴리펩티드에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 발현하는 것인 밸현 조성물.

청구항 23

제5항에 있어서, 세포가 분비 신호 폴리펩티드에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 발현하는 것인 방법.

청구항 24

제4항에 있어서, 세포 내의 효소가 관심 단백질 또는 폴리펩티드로부터 분비 신호 폴리펩티드를 절단하는 것인 밸현 조성물.

청구항 25

제5항에 있어서, 세포 내의 효소가 관심 단백질 또는 폴리펩티드로부터 분비 신호 폴리펩티드를 절단하는 것인 방법.

청구항 26

제5항에 있어서, 관심 단백질 또는 폴리펩티드가 세포의 주변 세포질(periplasm) 구획에서 발현되는 것인 방법.

청구항 27

제5항에 있어서, 세포가 20 g/ℓ 이상의 세포 밀도로 성장하는 것인 방법.

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

명세서**기술 분야**

[0001]

본 발명은 단백질 생산 분야에 속하고, 특히 적절하게 프로세싱된 이종 단백질의 생산을 위한 표적화 폴리펩티드의 용도에 관한 것이다.

배경기술

[0002]

재조합으로 생산된 150가지를 초과하는 단백질 및 폴리펩티드가 생명공학 약물 및 백신으로서의 사용에 대해 미국 식품의약청 (FDA)에서 허가되었고, 또다른 370개가 임상 시험 중이다. 화학적 합성을 통해 생산되는 소형 분자 치료제와는 달리, 단백질 및 폴리펩티드는 살아 있는 세포에서 가장 효율적으로 생산된다. 그러나, 박테리아에서의 현재의 재조합 단백질 생산 방법으로는 부적절하게 폴딩(folding)되거나, 응집되거나 또는 불활성인 단백질이 종종 생산되고, 다수의 유형의 단백질은 공지된 방법을 사용하여 비효율적으로 달성되는 2차 변형을 필요로 한다.

[0003]

공지된 방법의 한 주요 문제점은 세포질 내에 응집된 단백질들로 이루어진 봉입체가 형성되는 것이고, 이는 과량의 단백질이 세포 내에 축적될 때 발생한다. 재조합 단백질 생산에서의 또다른 문제점은 발현된 단백질에 대해 적절한 2차 및 3차 형상을 수립하는 것이다. 한가지 장벽은 종종 적절한 단백질 폴딩의 기초가 되는 디슬퍼드 결합 형성을 박테리아 세포질이 활발하게 저지한다는 것이다 ([Derman et al. (1993) *Science* 262:1744-7]). 그 결과, 다수의 재조합 단백질, 특히 진핵생물 기원의 단백질이 박테리아에서 생산될 때 부적절하게 폴딩되고 불활성이다.

[0004]

재조합 시스템에서 적절하게 폴딩된 단백질의 생산을 증가시키기 위해 수많은 시도들이 개발되었다. 예를 들어, 연구원들은 발효 조건을 변화시키거나 ([Schein (1989) *Bio/Technology*, 7:1141-1149]), 프로모터 강도를 다르게 하거나, 봉입체의 형성을 방지하는 것을 도울 수 있는 과발현된 샤파론(chaperone) 단백질을 사용하였다 ([Hockney (1994) *Trends Biotechnol.* 12:456-463]).

[0005]

적절하게 폴딩된 단백질의 수확을 증가시키기 위한 대안적인 접근법은 단백질을 세포내 환경으로부터 분비시키는 것이다. 신호 서열이 있는 폴리펩티드의 가장 통상적인 형태의 분비는 Sec 시스템을 수반한다. Sec 시스템은 N-말단 신호 폴리펩티드가 있는 단백질의 세포질막을 가로지르는 수출을 담당한다 ([Agarraberes and Dice (2001) *Biochim Biophys Acta*, 1513:1-24]; [Muller et al. (2001) *Prog Nucleic Acid Res Mol. Biol.* 66:107-157] 참조).

[0006]

세포로부터 상등액 내로 단백질을 배출시키기 위한 전략들이 개발되었다. 예를 들어, 미국 특허 번호 5,348,867; 미국 특허 번호 6,329,172; PCT 공개 번호 WO 96/17943; PCT 공개 번호 WO 02/40696; 및 미국 출원

공보 2003/0013150. 증가된 발현을 위한 또다른 전략은 단백질을 주변 세포질(periplasm)로 표적화하는 것에 대해 지시된다. 일부 연구원들은 Sec 유형이 아닌 분비에 초점을 맞추었다 (예를 들어 PCT 공개 번호 WO 03/079007; 미국 공개 번호 2003/0180937; 미국 공개 번호 2003/0064435; 및 PCT 공개 번호 WO 00/59537 참조). 그러나, 대다수의 연구는 Sec-유형 분비 시스템으로의 외인성 단백질의 분비에 초점을 맞추었다.

[0007] 다수의 분비 신호가 재조합 폴리펩티드 또는 단백질을 발현시키는 것에서의 사용에 대해 기술되었다. 예를 들어, 미국 특허 번호 5,914,254; 미국 특허 번호 4,963,495; 유럽 특허 번호 0 177 343; 미국 특허 번호 5,082,783; PCT 공개 번호 WO 89/10971; 미국 특허 번호 6,156,552; 미국 특허 번호 6,495,357; 6,509,181; 6,524,827; 6,528,298; 6,558,939; 6,608,018; 6,617,143; 미국 특허 번호 5,595,898; 5,698,435; 및 6,204,023; 미국 특허 번호 6,258,560; PCT 공개 번호 WO 01/21662, WO 02/068660 및 미국 출원 공보 2003/0044906; 미국 특허 번호 5,641,671; 및 유럽 특허 번호 EP 0 121 352 참조.

[0008] 단백질을 세포질 밖으로 표적화시키기 위해 신호 서열에 의존하는 전략에서는 부적절하게 프로세싱된 단백질이 종종 생산된다. 이는 아미노-말단 분비 신호 예컨대 Sec 시스템을 통해 분비에 이르는 신호에 대해 특히 그러하다. 이러한 시스템을 통해 프로세싱된 단백질들은 종종 분비 신호의 일부를 유지하거나, 연결 요소 (종종 부적절하게 절단됨)를 필요로하거나, 또는 말단에서 절단된다.

[0009] 상기 기술된 기술로부터 명백하듯이, 단백질을 숙주 세포의 주변 세포질로 표적화하기 위해 다수의 전략들이 개발되었다. 그러나, 공지된 전략들로는 치료적 사용을 위해 경제될 수 있는, 지속적으로 높은 수율의 적절하게 프로세싱된 활성 재조합 단백질이 초래되지 않았다. 기존의 전략들에서의 한가지 주요 제한사항은 적당하지 않은 세포 시스템에서 불량한 분비 신호 서열이 있는 단백질이 발현되는 것이었다.

[0010] 그 결과, 적절하게 프로세싱된 형태의 트랜스제닉(transgenic) 단백질을 생산하기 위하여 재조합 폴리펩티드를 분비할 수 있고 적절하게 프로세싱할 수 있는 개선된 대규모 발현 시스템이 당업계에서 계속 요구된다.

발명의 개요

[0012] 본 발명은 세포 발현 시스템에서 높은 수준의 적절하게 프로세싱된 관심 단백질 또는 폴리펩티드를 생산하기 위한 개선된 조성물 및 방법을 제공한다. 특히, 본 발명은 박테리아 생물로부터 유래된 분비 신호에 대한 신규 아미노산 및 뉴클레오티드 서열을 제공한다. 한 실시양태에서, 본 발명의 분비 신호는 단백질 디솔피드 이소머라제 C (dsbC), 돌연변이 포스페이트 결합 단백질 (pbp*), 단백질 디솔피드 이소머라제 A (dsbA), CupA2, CupB2, CupC2, NikA, FlgI, 테트라트리코펩티드 반복 패밀리 단백질 (ORF5550), 틀루엔 내성 단백질 (Ttg2C), 또는 메틸 수용성 화학주성 단백질 (ORF8124) 분비 신호, 뿐만 아니라 이의 생물학적으로 활성인 변이체, 단편 및 유도체로부터 선택된 슈도모나스 플루오레센스(*Pseudomonas fluorescens*) 분비 폴리펩티드이거나 이에 대해 실질적으로 상동성인 서열을 갖는 단리된 폴리펩티드를 포함한다. 또다른 실시양태에서, 본 발명의 분비 신호는 바실루스 코아글란스(*Bacillus coagulans*) Bce 분비 신호 서열이거나 이에 대해 실질적으로 상동성인 서열을 갖는 단리된 폴리펩티드이다. 본 발명의 신호 서열을 코딩하는 뉴클레오티드 서열은 발현된 관심 단백질 또는 폴리펩티드를 그램(Gram)-음성 박테리아의 주변 세포질로 또는 세포외 환경 내로 표적화하는 것을 촉진하기 위한 벡터 및 발현 시스템에서 유용하다.

[0013] 분비 신호 서열을 포함하는 DNA 구축물은 재조합 단백질을 발현시키기 위한 숙주 세포에서 유용하다. 관심 단백질에 대한 뉴클레오티드 서열이 본원에 기술된 바와 같은 분비 신호에 작동가능하게 연결된다. 세포는 주변 세포질 구획에서 단백질을 발현할 수 있다. 특정 실시양태에서, 세포는 발현된 재조합 단백질을 외부 세포벽을 지나 세포외로 또한 분비할 수 있다. 숙주 세포는 효모 세포, 곤충 세포, 포유류 세포, 식물 세포 등이 포함되는 진핵생물 세포, 및 박테리아 세포 예컨대 슈도모나스 플루오레센스, 대장균 등이 포함되는 원핵생물 세포를 포함한다. 치료적 단백질, 호르몬, 성장 인자, 세포외 수용체 또는 리간드, 프로테아제, 키나제, 혈액 단백질, 케모카인, 사이토카인, 항체 등을 포함하는 임의의 관심 단백질이 본 발명의 분비 폴리펩티드 리더 서열을 사용하여 발현될 수 있다.

발명의 상세한 설명

I. 개관

[0024] 숙주 세포에서 높은 수준의 적절하게 프로세싱된 폴리펩티드를 생산하기 위한 조성물 및 방법이 제공된다. 특히, 작동가능하게 연결된 관심 폴리펩티드를 그램-음성 박테리아의 주변 세포질로 또는 세포외 환경 내로 표적화하는 것을 촉진하는 신규 분비 신호가 제공된다. 본 발명의 목적을 위해, "분비 신호", "분비 신호 폴리펩티드

드", "신호 웨티드", 또는 "리더 서열"은 작동가능하게 연결된 관심 단백질 또는 폴리웨티드를 그램-음성 박테리아의 주변 세포질로 또는 세포외 공간 내로 표적화하는데 유용한 웨티드 서열 (또는 웨티드 서열을 코딩하는 폴리뉴클레오티드)을 의미한다. 본 발명의 분비 신호 서열은 pbp*, dsbA, dsbC, Bce, CupA2, CupB2, CupC2, NikA, FlgI, ORF5550, Ttg2C, 및 ORF8124 분비 신호, 및 이의 단편 및 변이체로부터 선택된 분비 폴리웨티드를 포함한다. 분비 신호에 대한 아미노산 서열이 서열 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 및 24에 기재된다. 상응하는 뉴클레오티드 서열이 각각 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 및 23에서 제공된다. 본 발명은 이러한 서열들, 뿐만 아니라 이의 단편 및 변이체를 포함한다.

[0026] 본 발명의 방법은 부적절하게 폴딩되거나, 응집되거나 또는 불활성인 단백질이 종종 생산되는, 박테리아에서의 현재의 재조합 단백질 생산 방법의 개선점을 제공한다. 추가적으로, 다수의 유형의 단백질이 공지된 방법을 사용하여 비효율적으로 달성되는 2차 변형을 필요로 한다. 본원에서의 방법은 세포내 환경으로부터 단백질을 분비시킴으로써 적절하게 폴딩된 단백질의 수확을 증가시킨다. 그램-음성 박테리아에서, 세포질로부터 분비된 단백질은 원형질막 주위 공간에서, 외막에 부착되어, 또는 세포외 브로스에서 종착될 수 있다. 이 방법은 응집된 단백질로 이루어진 봉입체를 또한 방지한다. 주변 세포질 내로의 분비는 적합한 디슬퍼드 결합 형성을 용이하게 하는 주지된 효과를 또한 갖는다 ([Bardwell et al. (1994) *Phosphate Microorg.* 270-5]; [Manoil (2000) *Methods in Enzymol.* 326: 35-47]). 재조합 단백질의 분비의 또다른 이점에는 단백질의 더욱 효율적인 단리; 활성 형태의 단백질의 백분율에서의 증가에 이르는, 트랜스제너 단백질의 적절한 폴딩 및 디슬퍼드 결합 형성; 봉입체의 형성 감소 및 숙주 세포에 대한 독성 감소; 및 가용성 형태의 재조합 단백질의 백분율 증가가 포함된다. 배양 배지 내로의 관심 단백질의 배출에 대한 가능성은 단백질 생산용으로 뱃치(batch) 배양보다는 연속 배양을 또한 잠재적으로 촉진할 수 있다.

[0027] 그램-음성 박테리아에서 이의 이중막을 가로지르는 단백질의 활성 수출을 위한 수많은 시스템이 발달되었다. 이러한 분비 경로에는, 예를 들어, 형질막과 외막 양쪽 모두를 가로지르는 1-단계 전위를 위한, ABC (제I형) 경로, Path/Fla (제III형) 경로, 및 Path/Vir (제IV형) 경로; 형질막을 가로지르는 전위를 위한, Sec (제II형), Tat, MscL, 및 Holins 경로; 및 형질막과 외막을 가로지르는 2-단계 전위를 위한, Sec-플러스(plus)-펩브리알 어서 포린(fimbrial usher porin) (FUP), Sec-플러스-오토트랜스포터(autotransporter) (AT), Sec-플러스-2-파트너 분비 (TPS: two partner secretion), Sec-플러스-메인 터미널 브랜치(main terminal branch) (MTB), 및 Tat-플러스-MTB 경로가 포함된다. 모든 박테리아에 이러한 분비 경로 모두가 있지는 않다.

[0028] 3가지 단백질 시스템 (제I형, 제III형 및 제IV형)이 단일 에너지-커플링 단계에서 양쪽 막을 가로질러 단백질을 분비한다. 4가지 시스템 (Sec, Tat, MscL 및 Holins)는 내막만을 가로질러 분비하고, 또다른 4가지 시스템 (MTB, FUP, AT 및 TPS)은 외막만을 가로질러 분비한다.

[0029] 한 실시양태에서, 본 발명의 신호 서열은 Sec 분비 시스템을 이용한다. Sec 시스템은 세포질막을 가로지르는 N-말단 신호 폴리웨티드가 있는 단백질의 수출을 담당한다 ([Agarraberes and Dice (2001) *Biochim Biophys Acta.* 1513:1-24]; [Muller et al. (2001) *Prog Nucleic Acid Res Mol. Biol.* 66:107-157] 참조). Sec 패밀리의 단백질 복합체는 원핵생물 및 진핵생물에서 보편적으로 발견된다. 박테리아 Sec 시스템은 수송 단백질, 샤페론 단백질 (SecB) 또는 신호 인식 입자 (SRP) 및 신호 웨티다제 (SPase I 및 SPase II)로 구성된다. 대장균에서의 Sec 수송 복합체는 SecY, SecE 및 SecG의 3가지 통합형 내막 단백질, 및 세포질 ATPase인 SecA로 구성된다. SecA는 활성 전위 채널을 형성하도록 SecY/E/G 복합체를 동원한다. 샤페론 단백질 SecB는 신생 폴리웨티드 사슬에 결합하여, 이의 폴딩을 방지하고, 이를 SecA에 표적화시킨다. 이어서, 선형 폴리웨티드 사슬이 SecYEG 채널을 통해 운송되고, 신호 폴리웨티드의 절단 후, 단백질이 주변 세포질에서 폴딩된다. 3개의 보조 단백질 (SecD, SecF 및 YajC)은 분비에 필수적이지는 않지만 다수의 조건 하에, 특히 저온에서 10배까지 분비를 자극하는 복합체를 형성한다.

[0030] 주변 세포질 내로, 즉 제II형 분비 시스템을 통해 운송되는 단백질이 추가적인 단계에서 세포외 배지 내로 또한 수출될 수 있다. 메커니즘은 일반적으로 오토트랜스포터, 2-파트너 분비 시스템, 메인 터미널 브랜치 시스템 또는 펩브리알 어서 포린을 통한 것이다.

[0031] 그램-음성 박테리아에서의 12개의 공지된 분비 시스템 중에서, 8개가 발현된 단백질의 일부로서 발견된 표적화 신호 폴리웨티드를 이용하는 것으로 공지되어 있다. 이러한 신호 폴리웨티드가 분비 시스템의 단백질과 상호작용하여, 세포가 단백질을 이의 적합한 목적지로 적절하게 지시한다. 신호-폴리웨티드를 기초로 하는 이러한 8개의 분비 시스템 중 5개는 Sec 시스템을 수반하는 것이다. 이러한 5개는 Sec-의존적 세포질막 전위에 수반되는 것으로 지칭되고, 여기에서 작동성인 이들의 신호 폴리웨티드는 Sec 의존적 분비 신호로 지칭될 수 있다.

적합한 분비 신호를 개발하는 것에서의 쟁점들 중 하나는 신호가 적합하게 발현되고, 발현된 단백질로부터 절단되는 것을 확실하게 하는 것이다.

[0032] sec 경로를 위한 신호 폴리펩티드는 일반적으로 하기의 3개의 도메인으로 구성된다: (i) 양성 전하를 띠는 n-영역, (ii) 소수성 h-영역 및 (iii) 전하를 띠지 않지만 극성인 c-영역. 신호 웨პ티다제에 대한 절단 부위는 c-영역 내에 위치한다. 그러나, 신호 서열 보존 및 길이의 정도, 뿐만 아니라 절단 부위 위치는 여러 단백질들 간에 다를 수 있다.

[0033] Sec-의존적 단백질 수출에 대한 서명은 수출된 단백질 내에 짧은 (아미노산 약 30개), 주로 소수성인 아미노-말단 신호 서열이 존재하는 것이다. 신호 서열은 단백질 수출을 돋고, 수출된 단백질이 주변 세포질에 도달했을 때 주변 세포질 신호 웨პ티다제에 의해 절단되어 제거된다. 전형적인 N-말단 Sec 신호 폴리펩티드는 하나 이상의 아르기닌 또는 라이신 잔기가 있는 N-도메인에 이어지는 일련의 소수성 잔기의 도메인, 및 신호 웨პ티다제에 대한 절단 부위를 함유하는 C-도메인을 함유한다.

[0034] 단백질을 세포질 밖으로 표적화하기 위한 시도로 관심 단백질 및 분비 신호 양쪽 모두를 함유하는 융합 단백질로서 트랜스제닉 단백질 구축물이 조작된 박테리아 단백질 생산 시스템이 개발되었다.

[0035] 슈도모나스 플루오레센스가 다양한 단백질의 생산을 위한 개선된 플랫폼인 것으로 실연되었고, 여러 효율적인 분비 신호가 이러한 생물로부터 확인되었다 (거명에 의해 전체적으로 본원에 포함된 미국 출원 공보 번호 20060008877 참조). 슈도모나스 플루오레센스는 다른 박테리아 발현 시스템에서 전형적으로 나타나는 것보다 높은 수준으로 외인성 단백질을 정확하게 프로세싱된 형태로 생산하고, 이러한 단백질을 세포의 원형질막 주위 공간으로 더 높은 수준으로 운송하여, 완전히 프로세싱된 재조합 단백질의 회수를 증가시킨다. 따라서, 한 실시양태에서, 본 발명은 분비 신호에 연결된 표적 단백질을 발현시킴으로써 슈도모나스 플루오레센스 세포에서 외인성 단백질을 생산하는 방법을 제공한다.

[0036] 본 발명의 분비 신호 서열은 슈도모나스에서 유용하다. 슈도모나드(Pseudomonad) 시스템은 다른 박테리아 발현 시스템과 비교하여 폴리펩티드 및 효소의 상업적인 발현을 위한 장점을 제공한다. 특히, 슈도모나스 플루오레센스는 유리한 발현 시스템으로 확인되었다. 슈도모나스 플루오레센스는 토양, 물 및 식물 표면 환경에서 콜로니를 형성하는 일군의 일반적인 비-병원성 부패균(saprophyte)을 포함한다. 슈도모나스 플루오레센스로부터 유래된 상업적인 효소들이 환경 오염을 감소시키는데, 세제 첨가물로서, 그리고 입체선택적 가수분해를 위해 사용되었다. 슈도모나스 플루오레센스는 병원체를 제어하기 위해 농업적으로 또한 사용된다. 미국 특허 번호 4,695,462에는 슈도모나스 플루오레센스에서의 재조합 박테리아 단백질의 발현이 기술되어 있다. 1985년 내지 2004년 사이에, 다수의 회사들이 구충성, 살충성 및 선충구제성 독소의 생산을 위한 슈도모나스 플루오레센스의 농업적 사용, 뿐만 아니라 특이적 독소 서열 및 이들의 발현을 강화시키기 위한 유전자 조작에 투자하였다. 예를 들어, PCT 출원 번호 WO 03/068926 및 WO 03/068948; PCT 공개 번호 WO 03/089455; PCT 출원 번호 WO 04/005221; 및 미국 특허 공개 번호 20060008877 참조.

II. 조성물

A. 단리된 폴리펩티드

[0038] 본 발명의 한 실시양태에서, 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 그람-음성 박테리아의 주변 세포질로 또는 세포의 공간 내로 표적화하는데 유용한 신규 분비 신호인 단리된 폴리펩티드가 제공된다. 한 실시양태에서, 폴리펩티드는 pbp*, dsbA, dsbC, Bce, CupA2, CupB2, CupC2, NikA, FlgI, ORF5550, Ttg2C, 또는 ORF8124 분비 신호, 또는 이의 단편 또는 변이체이거나 이에 대해 실질적으로 상동성인 아미노산 서열을 갖는다. 또다른 실시양태에서, 이러한 단리된 폴리펩티드는 분비 신호와 관심 단백질 또는 폴리펩티드의 융합 단백질이다.

[0039] 또다른 실시양태에서, 폴리펩티드 서열은 서열 6, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 또는 24에 기재된 분비 신호 폴리펩티드이거나 이에 대해 실질적으로 상동성이거나, 또는 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 또는 23에 기재된 폴리뉴클레오티드 서열에 의해 코딩된다. 또다른 실시양태에서, 폴리펩티드 서열은 적어도 서열 2의 아미노산 2-24, 적어도 서열 4의 아미노산 2-22, 적어도 서열 6의 아미노산 2-21, 적어도 서열 8의 아미노산 2-33, 적어도 서열 10의 아미노산 2-25, 적어도 서열 12의 아미노산 2-24, 적어도 서열 14의 아미노산 2-23, 적어도 서열 16의 아미노산 2-21, 적어도 서열 18의 아미노산 2-21, 적어도 서열 20의 아미노산 2-21, 적어도 서열 22의 아미노산 2-33, 또는 적어도 서열 24의 아미노산 2-39를 포함한다. 또다른 실시양태에서, 폴리펩티드 서열은 서열 6, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 또는 24의 단편을 포함하고, 이는 아미노 말단

으로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개의 아미노산이 말단절단되지만 생물학적 활성, 즉, 분비 신호 활성을 유지한다.

[0041] 한 실시양태에서, 상동성 폴리펩티드의 아미노산 서열은 소정의 원래의 폴리펩티드의 변이체이고, 이러한 변이체의 서열은 원래의 폴리펩티드의 아미노산 잔기의 약 30% 이하 (약 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 또는 30%까지를 포함)를 또다른 아미노산 잔기(들)로 교체함으로써 수득가능하고, 단변이체는 원래의 폴리펩티드의 원하는 기능을 유지한다. 실질적인 상동성이 있는 변이체 아미노산은 소정의 폴리펩티드에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 또는 적어도 약 99% 상동성일 것이다. 변이체 아미노산은 서열 6, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 또는 24의 1개 이상의 아미노산의 아미노산 치환, 결실, 말단절단 및 삽입 (약 1개, 약 2개, 약 3개, 약 4개, 약 5개, 약 6개, 약 7개, 약 8개, 약 9개, 약 10개, 약 15개, 약 20개, 약 25개, 또는 그 이상까지의 아미노산 치환, 결실 또는 삽입 포함)을 포함하는 다양한 방식으로 수득할 수 있다.

[0042] "실질적으로 상동성인" 또는 "실질적으로 유사한"은 표준 파라메터를 사용하여 본원에 기술된 정렬 프로그램들 중 하나를 사용하여 기준 서열과 비교했을 때 적어도 약 60% 또는 65%의 서열 동일성, 약 70% 또는 75%의 서열 동일성, 약 80% 또는 85%의 서열 동일성, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98% 또는 약 99% 또는 그 이상의 서열 동일성을 갖는 아미노산 또는 뉴클레오티드 서열을 의미한다. 당업자는 이러한 값들이 코돈 축퇴성, 아미노산 유사성, 리딩 프레임 위치결정 등을 고려함으로써 2개의 뉴클레오티드 서열에 의해 코딩되는 단백질들의 상응하는 동일성을 결정하기 위해 적합하게 조정될 수 있다는 것을 인지할 것이다.

[0043] 예를 들어, 바람직하게는, 보존적 아미노산 치환이 하나 이상의 예상 아미노산 잔기, 바람직하게는 비-필수 아미노산 잔기에서 이루어질 수 있다. "비-필수" 아미노산 잔기는 생물학적 활성을 변경시키지 않으면서 분비 신호 폴리펩티드의 야생형 서열로부터 변경될 수 있는 잔기인 한편, "필수" 아미노산 잔기는 생물학적 활성에 필요하다. "보존적 아미노산 치환"은 아미노산 잔기가 유사한 측쇄가 있는 아미노산 잔기로 교체되는 것이다. 보존적 및 반-보존적 아미노산 잔기의 패밀리들이 표 1에 열거된다.

표 1

유사한 아미노산 치환 군	
보존성 군 (8)	반-보존성 군 (7)
Arg, Lys	Arg, lys, His
Asp, Glu	Asn, Asp, Glu, Gln
Asn, Gln	
Ile, Leu, Val	Ile, Leu, Val, Met, Phe
Ala, Gly	Ala, Gly, Pro, Ser, Thr
Ser, Thr	Ser, Thr, Tyr
Phe, Tyr	Phe, Typ, Tyr
Cys (비-시스테인), Ser	Cys (비-시스테인), Ser, Thr

[0045] 본 발명에 포함되는 변이체 단백질은 생물학적으로 활성이고, 즉 천연 단백질의 원하는 생물학적 활성을 계속 보유하며, 즉 분비 신호 활성을 유지한다. "활성을 유지한다"는 변이체가 천연 단백질의 분비 신호 활성의 적어도 약 30%, 적어도 약 50%, 적어도 약 70%, 적어도 약 80%, 약 90%, 약 95%, 약 100%, 약 110%, 약 125%, 약 150%, 적어도 약 200% 또는 그 이상을 갖는다는 것을 의미한다.

B. 단리된 폴리뉴클레오티드

[0046] 본 발명은 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 그램-음성 박테리아의 주변 세포질로 또는 세포 외 공간 내로 표적화하는데 유용한 신규 분비 신호를 코딩하는 서열을 갖는 단리된 핵산을 또한 포함한다. 한 실시양태에서, 단리된 폴리뉴클레오티드는 pbp*, dsbA, dsbC, Bce, CupA2, CupB2, CupC2, NikA, FlgI, ORF5550, Ttg2C, 또는 ORF8124 분비 신호 폴리펩티드에 대해 실질적으로 상동성이 폴리펩티드 서열을 코딩한다. 또다른 실시양태에서, 본 발명은 적어도 서열 2의 아미노산 2-24, 적어도 서열 4의 아미노산 2-22, 적어도 서열 6의 아미노산 2-21, 적어도 서열 8의 아미노산 2-33, 적어도 서열 10의 아미노산 2-25, 적어도 서열 12의 아미노산 2-24, 적어도 서열 14의 아미노산 2-23, 적어도 서열 16의 아미노산 2-21, 적어도 서열 18의 아미노산 2-

21, 적어도 서열 20의 아미노산 2-21, 적어도 서열 22의 아미노산 2-33, 또는 적어도 서열 24의 아미노산 2-39에 대해 실질적으로 상동성인 폴리펩티드 서열을 코딩하는 핵산을 제공하거나, 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 21, 또는 23에 대해 실질적으로 상동성인 핵산을 제공한다 (이의 생물학적으로 활성인 변이체 및 단편 포함). 또다른 실시양태에서, 핵산 서열은 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 21, 또는 23의 서열에 대해 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 약 75%, 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 또는 적어도 약 99% 동일하다. 또다른 실시양태에서, 핵산은 서열 6, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 또는 24의 아미노산 서열에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 또는 적어도 약 99% 동일한 폴리펩티드를 코딩한다.

[0048] 바람직한 본 발명의 분비 신호 폴리펩티드는 서열 1 또는 3의 뉴클레오티드 서열에 대해 실질적으로 상동성인 뉴클레오티드 서열에 의해 코딩된다. PCR, 혼성화 등과 같은 방법을 사용하여, 본 발명의 서열에 대한 실질적인 동일성이 있는, 상응하는 분비 신호 폴리펩티드 서열을 확인할 수 있다. 예를 들어, [Sambrook J., and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)] 및 [Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, NY)] 참조. 변이체 뉴클레오티드 서열은 부위-지정 돌연변이유발을 사용함으로써 예를 들어 생성되었지만 하기에 논의된 바와 같이 본 발명에 개시된 분비 신호 폴리펩티드를 여전히 코딩하는, 합성에 의해 유래된 뉴클레오티드 서열을 또한 포함한다. 본 발명에 포함되는 변이체 분비 신호 폴리펩티드는 생물학적으로 활성이고, 즉 천연 단백질의 원하는 생물학적 활성을 계속 보유하며, 즉 분비 신호 활성을 유지한다. "활성을 유지한다"는 변이체가 천연 분비 신호 폴리펩티드의 활성의 적어도 약 30%, 적어도 약 50%, 적어도 70%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 약 96%, 약 97%, 약 98%, 적어도 약 99% 또는 그 이상을 갖는다는 것을 의미한다. 분비 신호 폴리펩티드 활성을 측정하는 방법은 본원의 다른 곳에서 논의된다.

[0049] 당업자는 분비 신호 폴리펩티드의 생물학적 활성을 변경시키지 않으면서, 본 발명의 뉴클레오티드 서열 내로 돌연변이에 의해 변화가 도입됨으로써, 코딩되는 분비 신호 폴리펩티드의 아미노산 서열에서의 변화에 이를 수 있다는 것을 또한 이해할 것이다. 따라서, 코딩되는 단백질 내로 하나 이상의 아미노산 치환, 부가 또는 결실이 도입되도록 하나 이상의 뉴클레오티드 치환, 부가, 또는 결실을 본원에 개시된 상응하는 뉴클레오티드 서열 내로 도입함으로써 단리된 핵산 문자 변이체가 생성될 수 있다. 표준 기술, 예컨대 부위-지정 돌연변이유발 및 PCR-매개 돌연변이 유발에 의해 돌연변이가 도입될 수 있다. 이같은 변이체 뉴클레오티드 서열 또한 본 발명에 포함된다.

C. 핵산 및 아미노산 상동성

[0051] 당업계에 주지된 다양한 방법들 중 임의의 방법에 따라 핵산 및 아미노산 서열 상동성이 결정된다. 유용한 서열 정렬 및 상동성 결정 방법의 예로는 하기와 기술된 것들이 포함된다.

[0052] 미국 NCBI (National Center for Biotechnology Information)의 프로그램인 MegaBLAST (현재 <http://www.ncbi.nlm.nih.gov/BLAST/>에서 입수가능)를 사용하여 유사한 서열들에 대한 정렬 및 검색을 수행할 수 있다. 아미노산 서열에 대해 예를 들어 70%로 설정된, 또는 뉴클레오티드 서열에 대해 예를 들어 90%로 설정된 동일성 백분율에 대한 옵션과 함께 이러한 프로그램을 사용하여, 질의 서열에 대한 서열 동일성이 70%, 또는 90%, 또는 그 이상인 서열들이 확인될 것이다. 당업계에 공지된 또다른 소프트웨어가 유사한 서열, 예를 들어, 본 발명에 따른 분비 신호 서열을 함유하는 정보 문자열(string)에 대해 적어도 70% 또는 90% 동일한 서열을 정렬 및/또는 검색하는데 또한 이용가능하다. 예를 들어, 질의 서열에 대해 적어도 70% 또는 90% 동일한 서열을 확인하기 위한 비교용 서열 정렬을, 예를 들어, GCG 서열 분석 소프트웨어 패키지 (University of Wisconsin Biotechnology Center의 Genetics Computer Group (1710 University Avenue, Madison, Wis. 5370 5)으로부터 입수가능)에서 입수가능한 GAP, BESTFIT, BLAST, FASTA, 및 TFASTA 프로그램을 프로그램 내에 특정된 바와 같은 디폴트 파라메터 및 원하는 백분율로 설정된 서열 동일성 정도에 대한 파라메터와 함께 사용함으로써 수행할 수 있다. 또한, 예를 들어, CLUSTAL 프로그램 (Intelligenetics (Mountain View, Cal.)로부터의 PC/Gene 소프트웨어 패키지에서 입수가능)을 사용할 수 있다.

[0053] 이러한 서열 정렬 방법들 및 기타 서열 정렬 방법은 당업계에 주지되어 있고, 수동 정렬에 의해, 시각적인 검사에 의해, 또는 서열 정렬 알고리즘, 예컨대 상기 기술된 프로그램에 의해 구현되는 것들 중 임의의 것의 수동 또는 자동 적용에 의해 수행될 수 있다. 다양한 유용한 알고리즘에는, 예를 들어, [W. R. Pearson & D. J.

Lipman, Proc. Natl. Acad. Sci. USA 85:2444-48 (April 1988)]에 기술된 유사성 검색 방법; [T. F. Smith & M. S. Waterman, Adv. Appl. Math. 2:482-89 (1981)] 및 [T. F. Smith & M. S. Waterman, J. Molec. Biol. 147:195-97 (1981)]에 기술된 국소적인 상동성 방법; [S. B. Needleman & C. D. Wunsch, J. Molec. Biol. 48(3):443-53 (March 1970)]에 기술된 상동성 정렬 방법; 및 [W. R. Pearson, Genomics 11(3):635-50 (November 1991)]; [W. R. Pearson, Methods Molec. Biol. 24:307-31 & 25:365-89 (1994)]; [D. G. Higgins & P. M. Sharp, Comp. Appl'n's in Biosci. 5:151-53 (1989)] 및 [D. G. Higgins & P. M. Sharp, Gene 73(1):237-44 (15 Dec. 1988)]에 예를 들어 기술된 다양한 방법이 포함된다.

[0054] 달리 언급되지 않는 한, [Needleman and Wunsch (1970) 상기 문헌]의 알고리즘을 사용하는 GAP 버전 10이 하기의 파라메터를 사용하여 서열 동일성 또는 유사성을 결정하는데 사용될 것이다: GAP 가중치 50 및 길이 가중치 3, 및 nwsgapdna.cmp 채점 매트릭스를 사용하는 뉴클레오티드 서열에 대한 % 동일성 및 % 유사성; GAP 가중치 8 및 길이 가중치 2, 및 BLOSUM62 채점 프로그램을 사용하는 아미노산 서열에 대한 % 동일성 또는 % 유사성. 등가의 프로그램을 또한 사용할 수 있다. "등가의 프로그램"은 질의되는 임의의 2개의 서열에 대해, GAP 버전 10에 의해 생성된 상응하는 정렬에 비교했을 때, 뉴클레오티드 잔기 매치가 동일하고 서열 동일성 백분율이 동일한 정렬을 생성시키는 임의의 서열 비교 프로그램을 의미한다. 다양한 실시양태에서, 질의 서열 또는 대상 서열, 또는 양쪽 모두의 전체에 대해 서열 비교가 수행된다.

D. 혼성화 조건

[0056] 본 발명의 또다른 양상에서, ppb*, dsbA, dsbC, Bce, CupA2, CupB2, CupC2, NikA, FlgI, ORF5550, Ttg2C, 또는 ORF8124 분비 신호 폴리펩티드에 대해 실질적으로 유사한 서열을 갖는 폴리펩티드를 코딩하는 서열을 갖는 단리된 핵산에 혼성화하는 핵산이 제공된다. 특정 실시양태에서, 혼성화 핵산은 고도의 엄격성 조건 하에 결합할 것이다. 다양한 실시양태에서, 분비 신호 폴리펩티드를 코딩하는 뉴클레오티드 서열의 실질적으로 전체적인 길이에 걸쳐, 예를 들어, 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 또는 23 중 하나 이상의 실질적으로 전체 길이에 걸쳐 혼성화가 발생한다. 핵산 분자가 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 또는 23 중 하나 이상의 전체 길이의 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95%에 걸쳐 혼성화할 때 핵산 분자는 본원에 개시된 분비 신호-코딩 뉴클레오티드 서열의 "실질적으로 전체적인 길이"에 혼성화한다. 달리 특정되지 않는 한, "실질적으로 전체적인 길이"는 분비 신호-코딩 뉴클레오티드 서열의 전체 길이의 적어도 80%를 지칭하고, 이때 길이는 연속적인 뉴클레오티드들에서 측정된다 (예를 들어, 서열 3의 적어도 53개의 연속적인 뉴클레오티드, 서열 5의 적어도 51개의 연속적인 뉴클레오티드, 서열 7의 적어도 80개의 연속적인 뉴클레오티드 등에 혼성화한다).

[0057] 혼성화 방법에서, 분비 신호 폴리펩티드를 코딩하는 뉴클레오티드 서열 모두 또는 이의 일부분을 사용하여 cDNA 또는 게놈 라이브러리를 스크리닝할 수 있다. 이같은 cDNA 및 게놈 라이브러리의 구축 방법은 당업계에 일반적으로 공지되어 있고, [Sambrook and Russell, 2001]에 개시되어 있다. 소위 혼성화 프로브는 게놈 DNA 단편, cDNA 단편, RNA 단편, 또는 기타 올리고뉴클레오티드일 수 있고, 검출가능한 기 예컨대 ^{32}P , 또는 임의의 또다른 검출가능한 마커, 예컨대 기타 방사성동위원소, 형광 화합물, 효소, 또는 효소 보조인자로 표지될 수 있다. 본원에 개시된 공지된 분비 신호 폴리펩티드-코딩 뉴클레오티드 서열을 기초로 합성 올리고뉴클레오티드를 표지함으로써 혼성화용 프로브가 제조될 수 있다. 뉴클레오티드 서열 또는 코딩된 아미노산 서열 내의 보존된 뉴클레오티드 또는 아미노산 잔기를 기초로 디자인된 축퇴성 프라이머가 추가적으로 사용될 수 있다. 프로브는 본 발명의 분비 신호 폴리펩티드-코딩 뉴클레오티드 서열 또는 이의 단편 또는 변이체의 적어도 약 10개, 적어도 약 15개, 적어도 약 16개, 17개, 18개, 19개, 20개, 또는 그 이상의 연속적인 뉴클레오티드에 엄격한 조건 하에 혼성화하는 뉴클레오티드 서열의 영역을 전형적으로 포함한다. 혼성화용 프로브의 제조 방법은 당업계에 일반적으로 공지되어 있고, 거명에 의해 본원에 포함된 [Sambrook and Russell, 2001]에 개시되어 있다.

[0058] 혼성화 기술에서, 공지된 뉴클레오티드 서열 모두 또는 이의 일부분이 선택된 생물로부터의 클로닝된 DNA 단편들 또는 cDNA 단편들의 집단 (즉, 게놈 또는 cDNA 라이브러리) 내에 존재하는 또다른 상응하는 뉴클레오티드 서열에 선택적으로 혼성화하는 프로브로서 사용된다. 혼성화 프로브는 게놈 DNA 단편, cDNA 단편, RNA 단편, 또는 기타 올리고뉴클레오티드일 수 있고, 검출가능한 기 예컨대 ^{32}P , 또는 임의의 또다른 검출가능한 마커로 표지될 수 있다. 따라서, 예를 들어, 본 발명의 분비 신호 폴리펩티드-코딩 뉴클레오티드 서열을 기초로 합성 올리고뉴클레오티드를 표지함으로써 혼성화용 프로브가 제조될 수 있다. 혼성화용 프로브의 제조 방법 및 cDNA 및 게놈 라이브러리의 구축 방법은 당업계에 일반적으로 공지되어 있고, [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York)]에 개

시되어 있다.

[0059] 예를 들어, 본원에 개시된 전체 분비 신호 폴리펩티드-코딩 뉴클레오티드 서열, 또는 이의 하나 이상의 일부분을 분비 신호 폴리펩티드를 코딩하는 상응하는 뉴클레오티드 서열 및 전령 RNA에 특이적으로 혼성화할 수 있는 프로브로 사용할 수 있다. 다양한 조건 하에서의 특이적 혼성화를 달성하기 위해, 이같은 프로브는 독특한 서열을 포함하고, 길이가 뉴클레오티드 약 10개 이상이거나 또는 뉴클레오티드 약 15개 이상이다. 이같은 프로브를 사용하여, PCR에 의해 선택된 생물로부터 상응하는 분비 신호 폴리펩티드-코딩 뉴클레오티드 서열을 증폭시킬 수 있다. 이러한 기술을 원하는 생물로부터 추가적인 코딩 서열을 단리하는데 사용하거나, 또는 생물 내의 코딩 서열의 존재를 결정하기 위한 진단 분석법으로 사용할 수 있다. 혼성화 기술은 플레이팅된 DNA 라이브리 (플라크 또는 콜로니; 예를 들어, [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York)] 참조)의 혼성화 스크리닝을 포함한다.

[0060] 이같은 서열의 혼성화는 엄격한 조건 하에 수행될 수 있다. "엄격한 조건" 또는 "엄격한 혼성화 조건"은 프로브가 이의 표적 서열에 다른 서열보다 검출가능하게 더 큰 정도로 혼성화할 조건을 의미한다 (예를 들어, 배경에 비해 2배 이상). 엄격한 조건은 서열-의존적이고, 상이한 환경에서 상이할 것이다. 혼성화 및/또는 세정 조건의 엄격성을 제어함으로써, 프로브에 100% 상보적인 표적 서열이 확인될 수 있다 (상동성 프로빙). 별법적으로, 엄격성 조건을 조정하여, 더 낮은 정도의 유사성이 검출되도록 서열 내에 약간의 미스매칭을 허용할 수 있다 (이종성 프로빙). 일반적으로, 프로브는 길이가 뉴클레오티드 약 1000개 미만, 바람직하게는 뉴클레오티드 500개 미만이다.

[0061] 전형적으로, 엄격한 조건은 pH 7.0 내지 8.3에서 염 농도가 Na 이온 약 1.5 M 미만, 전형적으로는 Na 이온 농도 약 0.01 내지 1.0 M (또는 기타 염)이고, 온도가 적어도 약 60°C, 바람직하게는 약 68°C인 조건이다. 엄격한 조건은 포름아미드와 같은 불안정화제의 첨가에 의해 또한 달성될 수 있다. 예시적인 낮은 엄격성 조건은 37°C에서의 30 내지 35% 포름아미드, 1 M NaCl, 1% SDS (소듐 도데실 술페이트)의 완충제 용액으로의 혼성화, 및 50 내지 55°C에서의 1× 내지 2× SSC (20× SSC = 3.0 M NaCl/0.3 M 시트르산3나트륨)에서의 세정을 포함한다. 예시적인 중등도의 엄격성 조건은 37°C에서의 40 내지 45% 포름아미드, 1 M NaCl, 1% SDS에서의 혼성화, 및 55 내지 60°C에서의 0.5× 내지 1× SSC에서의 세정을 포함한다. 예시적인 높은 엄격성 조건은 37°C에서의 50% 포름아미드, 1 M NaCl, 1% SDS에서의 혼성화, 및 60 내지 68°C에서의 0.1× SSC에서의 세정을 포함한다. 임의적으로, 세정 완충제가 약 0.1% 내지 약 1% SDS를 포함할 수 있다. 혼성화 기간은 일반적으로 약 24시간 미만, 통상적으로 약 4시간 내지 약 12시간이다.

[0062] 특이성은 전형적으로 혼성화 후의 세정의 함수이고, 결정적인 인자는 최종 세정 용액의 이온 강도 및 온도이다. DNA-DNA 하이브리드에 대해, [Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284]의 식으로부터 T_m 의 근사값을 구할 수 있다: $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\% \text{GC}) - 0.61 (\% \text{form}) - 500/L$ [식중, M은 1가 양 이온의 물농도이고, %GC는 DNA 내의 구아노신 및 사이토신 뉴클레오티드의 백분율이며, % form는 혼성화 용액 내의 포름아미드의 백분율이고, L은 염기 쌍 내의 하이브리드의 길이이다. T_m 은 상보적인 표적 서열의 50%가 완벽하게 매칭된 프로브에 혼성화하는 온도 (규정된 이온 강도 및 pH 하의 온도)이다. T_m 은 각각 1%의 미스매칭에 대해 약 1°C만큼 감소된다; 따라서, 원하는 동일성의 서열에 혼성화하도록 T_m , 혼성화, 및/또는 세정 조건을 조정할 수 있다. 예를 들어, ≥90% 동일성의 서열을 찾는 경우, T_m 이 10°C 감소될 수 있다. 일반적으로, 엄격한 조건은 규정된 이온 강도 및 pH에서의 특정 서열 및 이의 상보물에 대한 열 용점 (T_m)보다 약 5°C 더 낮도록 선택된다. 그러나, 심하게 엄격한 조건은 열 용점 (T_m)보다 1, 2, 3, 또는 4°C 낮은 온도에서의 혼성화 및/또는 세정을 사용할 수 있고, 중등도로 엄격한 조건은 열 용점 (T_m)보다 6, 7, 8, 9, 또는 10°C 낮은 온도에서의 혼성화 및/또는 세정을 사용할 수 있으며, 낮은 엄격성 조건은 열 용점 (T_m)보다 11, 12, 13, 14, 15, 또는 20°C 낮은 온도에서의 혼성화 및/또는 세정을 사용할 수 있다. 식, 혼성화 및 세정 조성물, 및 원하는 T_m 을 사용하여, 당업자는 혼성화 및/또는 세정 용액의 엄격성에서의 변동이 고유하게 기술된다는 것을 이해할 것이다. 원하는 정도의 미스매칭으로 45°C (수용액) 또는 32°C (포름아미드 용액) 미만의 T_m 이 초래되는 경우, 더 높은 온도가 사용될 수 있도록 SSC 농도를 증가시키는 것이 바람직하다. 핵산의 혼성화에 대한 광범위한 안내가 [Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Acid Probes, Part I, Chapter 2 (Elsevier, New York)]; 및 [Ausubel et al., eds. (1995) Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, New York)]에서 발

견된다. [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York)] 참조.

[0063] **E. 코돈 용법**

숙주 생물의 코돈 용법을 기초로 본원에 개시된 핵산 서열이 조정될 수 있다. 코돈 용법 또는 코돈 선호도는 당업계에 주지되어 있다. 선택된 코딩 서열을 이의 유전자 코드를 박테리아 숙주 세포에 의해 사용되는 것에 매칭되도록 변경시킴으로써 변형시킬 수 있고, 이의 코돈 서열을 숙주가 사용하는 것에 더욱 양호하게 근접하도록 강화시킬 수 있다. 당업자에게 공지된 다양한 방법들 중 임의의 것, 예를 들어, 올리고뉴클레오티드-지정 돌연변이유발에 따라 유전자 코드 선택 및 코돈 빈도 강화를 수행할 수 있다. 이러한 프로세스를 보조하기 위한 유용한 온-라인 인터넷 리소스에는, 예를 들어: (1) www.kazusa.or.jp/codon에서 입수 가능한 Kazusa DNA Research Institute (2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818 Japan)의 코돈 용법 데이터베이스; 및 (2) www.ncbi.nlm.nih.gov/-Taxonomy/Utils/wprintgc.cgi?mode=c에서 NCBI 분류학(Taxonomy) 데이터베이스로부터 입수 가능한 유전자 코드 테이블이 포함된다. 예를 들어, 슈도모나스(Pseudomonas) 종은 NCBI 분류학 사이트의 유전자 코드 번역 표 11을 사용하는 것으로 보고되었고, Kazusa 사이트에서는 www.kazusa.or.jp/codon/cgibin에 제시된 표의 유전자 용법 빈도를 나타내는 것으로 보고되었다. 분비 신호 폴리펩티드, 본원의 다른 곳에 기술된 관심 폴리펩티드, 또는 양쪽 모두에 대한 코딩 서열이 코돈 용법에 대해 조정될 수 있는 것으로 인식된다.

[0065] **F. 발현 벡터**

본 발명의 또 다른 실시양태는 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 그램-음성 박테리아의 주변 세포질로 또는 세포의 공간 내로 표적화하는데 유용한 신규 분비 폴리펩티드를 코딩하는 핵산을 포함하는 발현 벡터를 포함한다. 한 실시양태에서, 벡터는 프로모터에 작동가능하게 연결된, 본원에 개시된 분비 신호 폴리펩티드와 실질적으로 유사한 폴리펩티드를 코딩하는 폴리뉴클레오티드 서열을 포함한다. 발현가능한 코딩 서열은 선택된 숙주 세포에서 기능할 수 있는 전사 프로모터, 뿐만 아니라 모든 또 다른 필요한 전사 및 번역 조절 요소에 작동적으로 부착될 것이다.

[0066] 용어 "작동가능하게 연결된"은 코딩 서열과 관련하여, 숙주 세포 내에서 숙주 세포의 작용에 의해, 조절 요소가 코딩 서열의 발현을 지시할 수 있는 배치로 전사 조절 요소 및 임의의 번역 조절 요소가 코딩 서열에 공유결합으로 부착되어 있는 임의의 형상을 지칭한다.

[0067] [0068] 벡터는 벡터의 유지를 확실하게 하고, 원하는 경우, 숙주 내에서의 증폭을 제공하기 위해 하나 이상의 선별가능 표현형 마커 및 복제 기원을 전형적으로 포함할 것이다. 본 발명에 따른 형질전환을 위한 적절한 숙주에는 슈도모나스 속에 속하는 다양한 종이 포함되고, 슈도모나스 플루오레센스의 숙주 세포 균주가 특히 바람직하다.

[0069] 한 실시양태에서, 벡터는 본원에 개시된 분비 신호에 작동가능하게 연결된, 관심 단백질 또는 폴리펩티드의 발현을 위한 코딩 서열을 추가로 포함한다. 표적 폴리펩티드 코딩 서열이 리더 서열 및 전사 및 번역 조절 요소에 작동가능하게 연결되어 기능성 유전자를 형성하고, 이러한 기능성 유전자로부터 숙주 세포가 단백질 또는 폴리펩티드를 발현할 수 있는 폴리뉴클레오티드로부터 재조합 단백질 및 폴리펩티드가 발현될 수 있다. 코딩 서열은 입수 가능하다면 표적 폴리펩티드에 대한 천연 코딩 서열일 수 있지만, 더욱 바람직하게는, 예를 들어, 숙주 종의 코돈 사용 경향을 반영하도록 유전자를 합성함으로써, 선택된 발현 숙주 세포에서의 사용에 대해 선택, 개선 또는 최적화된 코딩 서열일 것이다. 본 발명의 한 실시양태에서, 숙주 종은 슈도모나스 플루오레센스이고, 신호 서열 및/또는 단백질 또는 폴리펩티드 서열 양쪽 모두를 디자인할 때 슈도모나스 플루오레센스의 코돈 경향이 고려된다. 유전자(들)이 하나 이상의 벡터(들) 내에 구축되거나 또는 하나 이상의 벡터(들) 내로 삽입된 후, 벡터가 발현 숙주 세포 내로 형질전환될 수 있다.

[0070] 또 다른 조절 요소가 벡터 ("발현 구축물"로 또한 칭해짐) 내에 포함될 수 있다. 이같은 요소에는, 예를 들어, 전사 인핸서 서열, 번역 인핸서 서열, 기타 프로모터, 활성화제, 번역 개시 및 정지 신호, 전사 종결인자, 시스 트론성 조절인자, 폴리시스템트론성 조절인자, 태그 서열, 예컨대 뉴클레오티드 서열 "태그" 및 "태그" 폴리펩티드 코딩 서열이 포함되지만, 이에 한정되지 않고, 이들은 발현된 폴리펩티드의 확인, 분리, 정제, 및/또는 단리를 용이하게 한다.

[0071] 또 다른 실시양태에서, 발현 벡터는 분비 신호에 대한 코딩 서열 또는 관심 단백질 또는 폴리펩티드에 대한 코딩 서열에 인접한 태그 서열을 추가로 포함한다. 한 실시양태에서, 이러한 태그 서열은 단백질이 정제되도록 한다. 태그 서열은 친화성 태그, 예컨대 헥사-히스티딘 친화성 태그일 수 있다. 또 다른 실시양태에서, 친화성

태그는 글루타티온-S-트랜스페라제 분자일 수 있다. 또한 태그는 형광 분자, 예컨대 YFP 또는 GFP, 또는 이같은 형광 단백질의 유사체일 수 있다. 또한 태그는 항체 분자의 일부분, 또는 정제에 유용한 공지된 결합 파트너에 대한 공지된 항원 또는 리간드일 수 있다.

[0072] 본 발명에 따른 단백질-코딩 유전자는, 단백질 코딩 서열에 더하여, 여기에 작동가능하게 연결된 하기의 조절 요소들을 포함할 수 있다: 프로모터, 리보솜 결합 부위 (RBS), 전사 종결인자, 번역 개시 및 정지 신호. 유용한 RBS는 본 발명에 따른 발현 시스템에서 숙주 세포로 유용한 임의의 종으로부터, 바람직하게는 선택된 숙주 세포로부터 수득될 수 있다. 다수의 특이적이고 다양한 컨센서스 RBS가 공지되어 있고, 예를 들어, [D. Frishman et al., Starts of bacterial genes: estimating the reliability of computer predictions, *Gene* 234(2):257-65 (8 Jul. 1999)]; 및 [B. E. Suzek et al., A probabilistic method for identifying start codons in bacterial genomes, *Bioinformatics* 17(12):1123-30 (December 2001)]에 기술되고 참조되어 있다. 추가적으로, 천연 또는 합성 RBS, 예를 들어, EP 0207459 (합성 RBS); [O. Ikehata et al., Primary structure of nitrile hydratase deduced from the nucleotide sequence of a *Rhodococcus* species and its expression in *Escherichia coli*, *Eur. J. Biochem.* 181(3):563-70 (1989)] (AAGGAAG의 천연 RBS 서열)에 기술된 것들을 사용할 수 있다. 본 발명에서 유용한 방법, 벡터, 및 번역 및 전사 요소, 및 기타 요소의 추가적인 예가, 예를 들어, 미국 특허 번호 5,055,294 (Gilroy) 및 미국 특허 번호 5,128,130 (Gilroy 등); 미국 특허 번호 5,281,532 (Rammiller 등); 미국 특허 번호 4,695,455 및 4,861,595 (Barnes 등); 미국 특허 번호 4,755,465 (Gray et al 등); 및 미국 특허 번호 5,169,760 (Wilcox)에 기술되어 있다.

[0073] 벡터 또는 플라스미드 내로 인해서 서열을 삽입하는 것에 의해 본 발명의 단백질을 코딩하는 DNA의 전사가 증가된다. 전형적인 인해서는 프로모터 상에 작용하여 이의 전사를 증가시키는 작용을 하는, 일반적으로 bp 약 10개 내지 300개 크기의, DNA의 시스(cis)-작용 요소이다. 예로는 다양한 슈도모나스 인해서가 포함된다.

[0074] 일반적으로, 재조합 발현 벡터는 숙주 세포의 형질전환을 허용하는 복제 기원 및 선별가능 마커, 및 하류의 구조 서열의 전사를 지시하기 위한 고도로 발현된 유전자로부터 유래된 프로모터를 포함할 것이다. 이같은 프로모터는 3-포스포글리세레이트 키나제 (PGK), 산 포스파타제와 같은 효소, 또는 열 충격 단백질을 코딩하는 오페론으로부터 특히 유래될 수 있다. 번역 개시 및 종결 서열과 함께, 그리고 바람직하게는 번역된 폴리펩티드의 분비를 지시할 수 있는 분비 서열과 함께, 이종성 구조 서열이 적합한 단계에서 어셈블리된다. 임의적으로, 이종성 서열은 원하는 특성, 예를 들어, 발현된 재조합 생성물의 안정화 또는 간소화된 정제를 부여하는 N-말단 확인 폴리펩티드를 포함하는 융합 폴리펩티드를 코딩할 수 있다.

[0075] 숙주 세포에서 재조합 단백질을 발현시키기 위한 벡터들이 당업계에 공지되어 있고, 이들 중 임의의 것을 본 발명에 따른 유전자를 발현시키는데 사용할 수 있다. 이같은 벡터에는, 예를 들어, 플라스미드, 코스미드, 및 파지 발현 벡터가 포함된다. 유용한 플라스미드 벡터의 예로는 발현 플라스미드 pBBR1MCS, pDSK519, pKT240, pML122, pPS10, RK2, RK6, pRO1600, 및 RSF1010이 포함되지만, 이에 한정되지는 않는다. 이같은 유용한 벡터의 또 다른 예로는, 예를 들어, [N. Hayase, *Appl. Environ. Microbiol.* 60(9):3336-42 (September 1994)]; [A. Lushnikov et al., *Basic Life Sci.* 30:657-62 (1985)]; [S. Graupner & W. Wackemagel, *Biomolec. Eng.* 17(1):11-16. (October 2000)]; [H. P. Schweizer, *Curr. Opin. Biotech.* 12(5):439-45 (October 2001)]; [M. Bagdasarian & K. N. Timmis, *Curr. Topics Microbiol. Immunol.* 96:47-67 (1982)]; [T. Ishii et al., *FEMS Microbiol. Lett.* 116(3):307-13 (Mar. 1, 1994)]; [I. N. Olekhovich & Y. K. Fomichev, *Gene* 140(1):63-65 (Mar. 11, 1994)]; [M. Tsuda & T. Nakazawa, *Gene* 136(1-2):257-62 (Dec. 22, 1993)]; [C. Nieto et al., *Gene* 87(1):145-49 (Mar. 1, 1990)]; [J. D. Jones & N. Gutterson, *Gene* 61(3):299-306 (1987)]; [M. Bagdasarian et al., *Gene* 16(1-3):237-47 (December 1981)]; [H. P. Schweizer et al., *Genet. Eng. (NY)* 23:69-81 (2001)]; [P. Mukhopadhyay et al., *J. Bact.* 172(1):477-80 (January 1990)]; [D. O. Wood et al., *J. Bact.* 145(3):1448-51 (March 1981)]; 및 [R. Holtwick et al., *Microbiology* 147(Pt 2):337-44 (February 2001)]에 기술된 것들이 포함된다.

[0076] 본 발명의 분비 신호 구축물을 포함하는 숙주 세포에서 유용할 수 있는 발현 벡터의 추가적인 예로는 지시된 레플리콘(replicon)으로부터 유래된, 표 2에 열거된 것들이 포함된다.

표 2

유용한 발현 벡터의 예	
레플리콘	벡터(들)
PPS10	PCN39, PCN51
RSF1010	PKT261-3
	PMMB66EH
	PEB8
	PPLGN1
	PMYC1050
RK2/RPI	PRK415
	PJB653
PRO1600	PUCP
	PBSP

[0077]

예를 들어, [F. Heffron et al., Proc. Nat'l Acad. Sci. USA 72(9):3623-27 (September 1975)], 및 [K. Nagahashi & K. Sakaguchi, J. Bact. 133(3):1527-29 (March 1978)]에 발현 플라스미드 RSF1010이 기술되어 있다. 플라스미드 RSF1010 및 이의 유도체는 본 발명에서 특히 유용한 벡터이다. 당업계에 공지된, RSF1010의 예시적이고 유용한 유도체에는, 예를 들어, pKT212, pKT214, pKT231 및 관련된 플라스미드, 및 pMYC1050 및 관련된 플라스미드 (예를 들어, 미국 특히 번호 5,527,883 및 5,840,554 (Thompson 등) 참조), 예를 들어, pMYC1803이 포함된다. 플라스미드 pMYC1803은 조절된 테트라사이클린 저항성 마커 및 RSF1010 플라스미드로부터의 복제 및 동원 좌위를 보유하는 RSF1010계 플라스미드 pTJS260 (미국 특히 번호 5,169,760 (Wilcox 등) 참조)로부터 유래된다. 또 다른 예시적인 유용한 벡터에는 미국 특히 번호 4,680,264 (Puhler 등)에 기술된 것들이 포함된다.

[0079]

한 실시양태에서, 발현 플라스미드가 발현 벡터로 사용된다. 또 다른 실시양태에서, RSF1010 또는 이의 유도체가 발현 벡터로 사용된다. 또 다른 실시양태에서, pMYC1050 또는 이의 유도체, 또는 pMYC4803 또는 이의 유도체가 발현 벡터로 사용된다.

[0080]

플라스미드 내에 선별 마커 유전자가 포함됨으로써 플라스미드가 숙주 세포 내에서 유지될 수 있다. 이는 항생제 저항성 유전자(들)일 수 있거나 (이러한 경우, 상응하는 항생제(들)이 발효 배지에 첨가됨), 또는 당업계에 공지된 임의의 또 다른 유형의 선별 마커 유전자, 예를 들어, 원영양성(prototrophy)-복구 유전자일 수 있다 (이러한 경우, 상응하는 형질, 예를 들어, 생축매성 형질 예컨대 아미노산 생합성 또는 뉴클레오티드 생합성 형질, 또는 탄소 공급원 활용 형질에 대해 영양요구성(auxotrophic)인 숙주 세포에서 플라스미드가 사용됨).

[0081]

본 발명에 따라 사용되는 프로모터는 구성성 프로모터 또는 조절형 프로모터일 수 있다. 유용한 조절형 프로모터의 통상적인 예로는 lac 프로모터로부터 유래된 패밀리의 것들 (즉 lacZ 프로모터), 특히 미국 특히 번호 4,551,433 (DeBoer)에 기술된 tac 및 trc 프로모터, 뿐만 아니라 Ptac16, Ptac17, PtacII, PlacUV5, 및 T71ac 프로모터가 포함된다. 한 실시양태에서, 프로모터는 숙주 세포 생물로부터 유래되지 않는다. 특정 실시양태에서, 프로모터는 대장균 생물로부터 유래된다.

[0082]

본 발명에 따른 발현 시스템에서 유용한 lac-유형이 아닌 프로모터의 통상적인 예로는, 예를 들어, 표 3에 열거된 것들이 포함된다.

표 3

lac 유형이 아닌 프로모터의 예	
프로모터	유도인자
P _R	고온
P _L	고온
P _m	알킬- 또는 할로-벤조에이트
P _u	알킬- 또는 할로-톨루엔
P _{sal}	살리실레이트

[0083]

예를 들어, [J. Sanchez-Romero & V. De Lorenzo (1999) Genetic Engineering of Nonpathogenic Pseudomonas strains as Biocatalysts for Industrial and Environmental Processes, Manual of Industrial Microbiology and Biotechnology (A. Demain & J. Davies, eds.) pp. 460-74 (ASM Press, Washington, D.C.)]; [H. Schweizer (2001) Vectors to express foreign genes and techniques to monitor gene expression for Pseudomonads, Current Opinion in Biotechnology, 12:439-445]; 및 [R. Slater & R. Williams (2000) The Expression of Foreign DNA in Bacteria, Molecular Biology and Biotechnology (J. Walker & R. Rapley,

eds.) pp. 125-54 (The Royal Society of Chemistry, Cambridge, UK)] 참조. 선택된 박테리아 속주 세포에 천연인 프로모터의 뉴클레오티드 서열을 갖는 프로모터, 예를 들어, 슈도모나스 안트라닐레이트 또는 벤조에이트 오페론 프로모터 (Pant, Pben)가 표적 폴리펩티드를 코딩하는 트랜스진(transgene)의 발현을 제어하는데 또한 사용될 수 있다. 서열이 동일하거나 상이한지 여부와 상관 없이, 또는 동일한 또는 상이한 생물로부터 유래되는지 여부와 상관없이, 하나를 초과하는 프로모터들이 서로 공유결합으로 부착된 직렬 프로모터, 예를 들어, Pant-Pben 직렬 프로모터 (인터프로모터(interpromoter) 하이브리드) 또는 Plac-Plac 직렬 프로모터가 또한 사용될 수 있다.

[0085] 조절형 프로모터는 프로모터가 유전자의 일부분인 유전자의 전사를 제어하기 위해 프로모터 조절 단백질을 활용한다. 조절형 프로모터가 본원에서 사용되는 경우, 상응하는 프로모터 조절 단백질 또한 본 발명에 따른 발현 시스템의 일부일 것이다. 프로모터 조절 단백질의 예로는 활성화제 단백질, 예를 들어, 대장균 이화대사산물 활성화제 단백질, MaT 단백질; AraC 패밀리 전사 활성화제; 억제인자 단백질, 예를 들어, 대장균 LacI 단백질; 및 이중-기능 조절 단백질, 예를 들어, 대장균 NagC 단백질이 포함된다. 다수의 조절형-프로모터/프로모터-조절-단백질 쌍이 당업계에 공지되어 있다.

[0086] 프로모터 조절 단백질은 이펙터(effecter) 화합물, 즉 조절 단백질과 가역적으로 또는 비가역적으로 회합하여 조절 단백질이 프로모터의 제어 하에 있는 유전자의 하나 이상의 DNA 전사 조절 영역에서 방출되거나 이러한 영역에 결합될 수 있도록 함으로써 유전자 전사의 개시에서 전사효소의 작용을 허용 또는 차단하는 화합물과 상호 작용한다. 이펙터 화합물은 유도인자 또는 공동-억제인자(co-repressor)로 분류되고, 이러한 화합물에는 천연 이펙터 화합물 및 무상(gratuitous) 유도인자 화합물이 포함된다. 다수의 조절형-프로모터/프로모터-조절-단백질/이펙터-화합물 트리오가 당업계에 공지되어 있다. 세포 배양 또는 발효 전반에 걸쳐 이펙터 화합물이 사용될 수 있지만, 조절형 프로모터가 사용되는 바람직한 실시양태에서, 원하는 양 또는 밀도의 속주세포 생물량의 성장 후, 관심 단백질 또는 폴리펩티드를 코딩하는 원하는 유전자(들)의 발현이 직접적으로 또는 간접적으로 초래되도록 적합한 이펙터 화합물이 배양물에 침가된다.

[0087] 예를 들어, lac 패밀리 프로모터가 사용되는 경우, lacI 유전자가 또한 시스템 내에 존재할 수 있다. (정상적으로는) 구성적으로 발현되는 유전자인 lacI 유전자는 이러한 프로모터들의 lac 오퍼레이터에 결합하는 Lac 억제인자 단백질 (LacD 단백질)을 코딩한다. 따라서, lac 패밀리 프로모터가 사용되는 경우, lacI 유전자가 또한 발현 시스템에 포함되어 발현될 수 있다. lac 프로모터 패밀리 구성원, 예를 들어, tac 프로모터의 경우, 이펙터 화합물은 유도인자, 바람직하게는 무상 유도인자 예컨대 IPTG (이소프로필- β -D-1-티오갈락토파라노사이드; "이소프로필티오갈락토사이드"로 또한 칭해짐)이다.

[0088] 관심 단백질 또는 폴리펩티드의 발현을 위해, 임의의 식물 프로모터를 또한 사용할 수 있다. 프로모터는 식물 RNA 중합효소 II 프로모터일 수 있다. 식물 프로모터에 포함된 요소는 전사 개시 부위에서 약 25 내지 35 염기 쌍 상류 (5')에 전형적으로 위치하는 TATA 박스 또는 골드버그-호그네스(Goldberg-Hogness) 박스, 및 70 내지 100 염기쌍 상류에 위치하는 CCAAT 박스일 수 있다. 식물에서, CCAAT 박스는 포유류 프로모터의 기능적으로 유사한 서열과 상이한 컨센서스(concensus) 서열을 가질 수 있다 ([Messing et al. (1983) Genetic Engineering of Plants, Kosuge et al., eds., pp. 211-227]). 또한, 실제로 모든 프로모터는 전사 개시 부위에서 약 -100 bp 내지 -1,000 bp 이상 상류에서 연장된 추가적인 상류 활성화 서열 또는 인핸서를 포함한다 ([Benoist and Chambon (1981) Nature 290:304-310]; [Gruss et al. (1981) Proc. Nat. Acad. Sci. 78:943-947]; 및 [Khoury and Gruss (1983) Cell 27:313-314]).

G. 발현 시스템

[0089] 본 발명은 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 그램-음성 박테리아의 주변 세포질로 또는 세포 외 공간 내로 표적화하는데 유용한 개선된 발현 시스템을 추가로 제공한다. 한 실시양태에서, 상기 시스템은 속주 세포, 및 pbp*, dsbA, dsbC, Bce, CupA2, CupB2, CupC2, NikA, FlgI, ORF5550, Ttg2C, 및 ORF8124 분비 신호 서열로 구성된 군으로부터 선택된 분비 신호, 또는 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 21, 또는 23으로 본원에 개시된 분비 신호 서열에 대해 실질적으로 상동성인 서열, 또는 서열 6, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 또는 24를 코딩하는 뉴클레오티드 서열에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 코딩하는 뉴클레오티드 서열을 포함하는 상기 기술된 벡터를 포함한다. 일부 실시양태에서, 신호 서열과 관심 단백질 또는 폴리펩티드 사이에서 변형이 이루어지지 않는다. 그러나, 특정 실시양태에서, 추가적인 절단 신호가 혼입되어, 폴리펩티드의 아미노 말단의 적합한 프로세싱을 촉진한다.

[0090] 분비 시스템은 하기에 기술된 바와 같은 발효 배지를 또한 포함할 수 있다. 한 실시양태에서, 시스템은 무기염

배지를 포함한다. 또 다른 실시양태에서, 시스템은 화학 유도인자를 배지 내에 포함한다.

[0092]

ChampionTM pET 발현 시스템은 높은 수준의 단백질 생산을 제공한다. 강력한 T7lac 프로모터로부터 발현이 유도된다. 이러한 시스템은 관심 유전자의 높은 수준의 전사를 위해 박테리오파지 T7 RNA 중합효소의 높은 활성 및 특이성의 장점을 취한다. 프로모터 영역에 위치한 lac 오퍼레이터는 전통적인 T7계 벡터보다 엄격한 조절을 제공하여, 플라스미드 안정성 및 세포 생존력을 개선시킨다 ([Studier and Moffatt (1986) *J Molecular Biology* 189(1): 113-30]; [Rosenberg, et al. (1987) *Gene* 56(1): 125-35]). T7 발현 시스템은 관심 유전자의 높은 수준의 전사를 위해 T7 프로모터 및 T7 RNA 중합효소 (T7 RNAP)를 사용한다. T7 RNAP가 천연 대장균 RNAP보다 더 진행성(processive)이고 관심 유전자의 전사에 대해 전용이기 때문에, 높은 수준의 발현이 T7 발현 시스템 내에서 달성된다. 숙주 세포 내에 T7 RNAP의 공급원을 제공함으로써, 확인된 유전자의 발현이 유도된다. 이는 T7 RNAP 유전자의 염색체 카피를 함유하는 BL21 대장균 숙주를 사용함으로써 달성된다. T7 RNAP 유전자는 IPTG에 의해 유도될 수 있는 lacUV5 프로모터의 제어 하에 있다. 유도 시 T7 RNAP가 발현되고, 이에 의해 관심 유전자가 전사된다.

[0093]

pBAD 발현 시스템은 특정 탄소 공급원 예컨대 글루코스, 글리세롤 및 아라비노스의 존재를 통해 관심 단백질 또는 폴리펩티드의 엄격하게 제어되고 적정가능한 발현을 가능하게 한다 ([Guzman, et al. (1995) *J Bacteriology* 177(14): 4121-30]). pBAD 벡터는 발현 수준에 대한 정밀한 제어를 제공하도록 독특하게 디자인된다. pBAD 벡터로부터의 이종성 유전자 발현은 araBAD 프로모터에서 개시된다. 이러한 프로모터는 araC 유전자의 생성물에 의해 양성적으로 및 음성적으로 조절된다. AraC는 L-아라비노스와 복합체를 형성하는 전사 조절 인자이다. L-아라비노스가 없는 경우, AraC 이량체는 전사를 차단한다. 최대의 전사 활성화를 위해, 2개의 이벤트가 요구된다: (i) L-아라비노스가 AraC에 결합하여 전사가 시작되게 함, (ii) cAMP 활성화제 단백질 (CAP)-cAMP 복합체가 DNA에 결합하고, AraC가 프로모터 영역의 정확한 위치에 결합하는 것을 자극함.

[0094]

trc 발현 시스템은 trc 프로모터로부터의 대장균에서의 높은 수준의 조절형 발현을 허용한다. trc 발현 벡터는 대장균에서의 진핵생물 유전자의 발현에 대해 최적화되었다. trc 프로모터는 트립토판 (*trp*) 및 락토스 (*lac*) 프로모터로부터 유래된 강력한 하이브리드 프로모터이다. 이는 lacO 오퍼레이터 및 lacIQ 유전자 생성물에 의해 조절된다 ([Brosius, J. (1984) *Gene* 27(2): 161-72]).

[0095]

숙주 세포를 본원에 개시된 벡터(들)로 형질전환시키는 것은 당업계에 공지된 임의의 형질전환 방법을 사용하여 수행할 수 있고, 박테리아 숙주 세포는 무손상 세포로서 또는 원형질체 (즉 세포질 포함)로서 형질전환될 수 있다. 예시적인 형질전환 방법에는 천공 방법, 예를 들어, 전기천공, 원형질체 융합, 박테리아 접합, 및 2가 양이온 처리, 예를 들어, 염화칼슘 처리 또는 CaCl₂/Mg²⁺ 처리, 또는 당업계에 주지된 기타 방법이 포함된다. 예를 들어, [Morrison, J. *Bact.*, 132:349-351 (1977)]; [Clark-Curtiss & Curtiss, *Methods in Enzymology*, 101:347-362 (Wu et al., eds, 1983)], [Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989)]; [Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990)]; 및 [Current Protocols in Molecular Biology (Ausubel et al., eds., 1994)] 참조.

[0096]

H. 숙주 세포

[0097]

한 실시양태에서, 본 발명은 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 그람-음성 박테리아의 주변 세포질로 또는 세포의 공간 내로 표적화하는데 유용한 발현 시스템을 제공한다. 한 실시양태에서, 이러한 시스템은 분비 신호 웨티드를 이용한다. 또 다른 실시양태에서, 발현 시스템은 본원에 개시된 분비 신호를 포함하는 단백질의 발현을 위한 슈도모나스 플루오레센스 발현 시스템이다. 본 발명의 이러한 양상은 슈도모나스 플루오레센스가 슈도모나스 플루오레센스 및 비-슈도모나스 플루오레센스 시스템 양쪽 모두로부터의 분비 신호를 적절하게 프로세싱하고 표적화할 수 있다는 뜻밖의 발견을 기초로 한다.

[0098]

이러한 실시양태에서, 숙주 세포는 "그람-음성 프로테오박테리아(Proteobacteria) 아군 18"로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 18"은 하기에 예를 들어 속하는 것들이 포함되는, 슈도모나스 플루오레센스 종의 모든 아종, 변종, 군주 및 기타 하위-특정 유닛의 군으로 정의된다 (예시적인 군주(들)의 ATCC 또는 기타 기탁 번호가 괄호 안에 제시됨): 슈도모나스 플루오레센스 생물형(biotype) A (생물변이형(biovar) 1 또는 생물변이형 I로도 칭해짐) (ATCC 13525); 슈도모나스 플루오레센스 생물형 B (생물변이형 2 또는 생물변이형 II로도 칭해짐) (ATCC 17816); 슈도모나스 플루오레센스 생물형 C (생물변이형 3 또는 생물변이형 III으로도 칭해짐) (ATCC 17400); 슈도모나스 플루오레센스 생물형 F (생물변이형 4 또는 생물변이형 IV로도 칭해짐) (ATCC 12983); 슈도모나스 플루오레센스 생물형 G (생물변이형 5 또는 생물변이형 V로도 칭해짐) (ATCC 17518); 슈도모나스 플루오레센스 생물변이형 VI; 슈도모나스 플루오레센스 Pf0-1; 슈도모나스 플루오레센스 Pf-5 (ATCC

BAA-477); 슈도모나스 플루오레센스 SBW25; 및 슈도모나스 플루오레센스 아종 셀룰로사(cellulosa) (NCIMB 10462).

[0099] 숙주 세포는 "그람-음성 프로테오박테리아 아군 19"로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 19"는 슈도모나스 플루오레센스 생물형 A의 모든 균주의 균으로 정의된다. 이러한 생물형의 특히 바람직한 균주는 슈도모나스 플루오레센스 균주 MB101 (미국 특허 번호 5,169,760 (Wilcox) 참조), 및 이의 유도체이다. 이의 바람직한 유도체의 예는 MB101 염색체 *asd* (아스파테이트 탈수소효소 유전자) 좌위 내로 천연 대장균 *PlacI-lacI-lacZYA* 구축물 (즉, *PlacZ*가 결실됨)을 삽입함으로써 구축된 슈도모나스 플루오레센스 균주 MB214이다.

[0100] 본 발명에서 사용될 수 있는 추가적인 슈도모나스 플루오레센스 균주는 하기의 ATTC 명칭을 갖는 슈도모나스 플루오레센스 미글라(Migula) 및 슈도모나스 플루오레센스 로이토키토크(Loitokitok)를 포함한다: [NCIB 8286]; NRRL B-1244; NCIB 8865 균주 C01; NCIB 8866 균주 C02; 1291 [ATCC 17458; IFO 15837; NCIB 8917; LA; NRRL B-1864; 피롤리딘; PW2 [ICMP 3966; NCPPB 967; NRRL B-899]; 13475; NCTC 10038; NRRL B-1603 [6; IFO 15840]; 52-1C; CCEB 488-A [BU 140]; CCEB 553 [EM 15/47]; IAM 1008 [AHH-27]; IAM 1055 [AHH-23]; 1 [IFO 15842]; 12 [ATCC 25323; NIH 11; den Dooren de Jong 216]; 18 [IFO 15833; WRRL P-7]; 93 [TR-10]; 108 [52-22; IFO 15832]; 143 [IFO 15836; PL]; 149 [2-40-40; IFO 15838]; 182 [IFO 3081; PJ 73]; 184 [IFO 15830]; 185 [W2 L-1]; 186 [IFO 15829; PJ 79]; 187 [NCPPB 263]; 188 [NCPPB 316]; 189 [PJ227; 1208]; 191 [IFO 15834; PJ 236; 22/1]; 194 [Klinge R-60; PJ 253]; 196 [PJ 288]; 197 [PJ 290]; 198 [PJ 302]; 201 [PJ 368]; 202 [PJ 372]; 203 [PJ 376]; 204 [IFO 15835; PJ 682]; 205 [PJ 686]; 206 [PJ 692]; 207 [PJ 693]; 208 [PJ 722]; 212. [PJ 832]; 215 [PJ 849]; 216 [PJ 885]; 267 [B-9]; 271 [B-1612]; 401 [C71A; IFO 15831; PJ 187]; NRRL B-3178 [4; IFO 15841]; KY 8521; 3081; 30-21; [IFO 3081]; N; PYR; PW; D946-B83 [BU 2183; FERM-P 3328]; P-2563 [FERM-P 2894; IFO 13658]; IAM-1126 [43F]; M-1; A506 [A5-06]; A505 [A5-05-1]; A526 [A5-26]; B69; 72; NRRL B-4290; PMW6 [NCIB 11615]; SC 12936; A1 [IFO 15839]; F 1847 [CDC-EB]; F 1848 [CDC 93]; NCIB 10586; P17; F-12; AmMS 257; PRA25; 6133D02; 6519E01; Ni; SC15208; BNL-WVC; NCTC 2583 [NCIB 8194]; H13; 1013 [ATCC 11251; CCEB 295]; IFO 3903; 1062; 또는 Pf-5.

[0101] 한 실시양태에서, 숙주 세포는 상기 기술된 바와 같은 슈도모나스 플루오레센스 세포를 포함하는, 관심 단백질 또는 폴리펩티드를 생산할 수 있는 임의의 세포일 수 있다. 관심 단백질 또는 폴리펩티드를 생산하기 위한 가장 통상적으로 사용되는 시스템에는 특정 박테리아 세포, 특히 대장균이 포함되는데, 이들의 비교적 저렴한 성장 요건 및 대형 뱃치(batch) 배양에서 단백질을 생산하는 잠재적인 능력 때문이다. 생물학적으로 관련된 단백질 및 웹티드를 발현하기 위해, 특히 연구 목적으로 효모가 또한 사용된다. 사카로마이세스 세레비지아에 (*Saccharomyces cerevisiae*) 또는 피키아 파스토리스(*Pichia pastoris*)가 시스템에 포함된다. 이러한 시스템들은 잘 특성화되어 있고, 일반적으로 허용가능한 수준의 전체 단백질 발현을 제공하며, 비교적 빠르고 저비용이다. 곤충 세포 발현 시스템이 생물학적으로 활성인 형태의 재조합 단백질을 발현시키기 위한 별법으로서 또한 나타났다. 일부 경우에, 번역 후에 변형된 정확하게 풀딩된 단백질이 생산될 수 있다. 포유류 세포 발현 시스템, 예컨대 차이니즈 햄스터 난소 세포가 또한 관심 단백질 또는 폴리펩티드의 발현을 위해 사용되었다. 소규모인 경우, 이러한 발현 시스템들은 종종 효과적이다. 특정 생물체제가, 특히 동물 또는 인간 건강 용도에서, 단백질로부터 유도될 수 있다. 또 다른 실시양태에서, 숙주 세포는 담배 세포, 옥수수, 아라비돕시스 (*Arabidopsis*) 종으로부터의 세포, 감자 또는 쌀 세포가 포함되지만 이에 한정되지는 않는 식물 세포이다. 또 다른 실시양태에서, 트랜스제닉 생물이 포함되지만 이에 한정되지는 않는 다세포 생물이 본 방법에서 분석 또는 변형된다. 다세포 생물을 분석하고/하거나 변형시키기 위한 기술은 하기의 기술된 세포를 변형시키기 위해 기술된 기술을 일반적으로 기초로 한다.

[0102] 또 다른 실시양태에서, 숙주 세포는 에쉐리키아(*Escherichia*) 또는 슈도모나스 종이 포함되지만 이에 한정되지는 않는 박테리아 세포와 같은 원핵생물일 수 있다. 전형적인 박테리아 세포는, 예를 들어, Estrella Mountain Community College (Arizona, USA)의 M J Farabee 박사가 웹사이트 www.emc.maricopa.edu/faculty/farabee/BIOBK/BioBookDiversity에서 제공하는 온라인 생물학 저서의 "Biological Diversity: Bacteria and Archaeans" 챕터에 기술되어 있다. 특정 실시양태에서, 숙주 세포는 슈도모나드 세포일 수 있고, 전형적으로 슈도모나스 플루오레센스 세포일 수 있다. 또 다른 실시양태에서, 숙주 세포는 또한 대장균 세포일 수 있다. 또 다른 실시양태에서, 숙주 세포는 진핵생물 세포, 예를 들어, 스포둡테라(*Spodoptera*), 트리코플루시아(*Trichoplusia*), 드로소필라(*Drosophila*) 또는 에스티그멘(*Estigmene*) 종이 포함되지만 이에 한정되지는 않는 곤충 세포, 또는 뮤린(*murine*) 세포, 햄스터 세포, 원숭이, 영장류 또는 인간

세포가 포함되지만 이에 한정되지는 않는 포유류 세포일 수 있다.

[0103] 한 실시양태에서, 숙주 세포는 임의의 박테리아 분류군의 구성원일 수 있다. 세포는, 예를 들어, 유박테리아 (eubacteria)의 임의의 종의 구성원일 수 있다. 숙주는 하기 분류군 중 임의의 하나의 구성원일 수 있다: 애시도박테리아(Acidobacteria), 액티노박테리아(Actinobacteria), 애퀴피케(Aquificae), 박테로이데테스 (Bacteroidetes), 클로로비(Chlorobi), 클라미디에(Chlamydiae), 코로플렉시(Choroflexi), 크리시오제네테스 (Chrysiogenetes), 시아노박테리아(Cyanobacteria), 데페리박테레스(Deferrribacteres), 데이노코쿠스 (Deinococcus), 딕티요글로미(Dictyoglomi), 피브로박테레스(Fibrobacteres), 피르미쿠테스(Firmicutes), 푸소박테리아(Fusobacteria), 켐마티모나데테스(Gemmatimonadetes), 렌티스파레(Lentisphaerae), 니트로스파리에 (Nitrospirae), 플랑크토미세테스(Planctomycetes), 프로테오박테리아, 스피로캐테스(Spirochaetes), 테르모데술포박테리아(Thermodesulfobacteria), 테르모마이크로비아(Thermomicrobia), 테르모토가에(Thermotogae), 테르무스(Thermus) (테르말레스(Theamales)), 또는 베루코미크로비아(Verrucomicrobia). 유박테리아 숙주 세포의 한 실시양태에서, 세포는 시아노박테리아(Cyanobacteria)를 제외한 유박테리아의 임의의 종의 구성원일 수 있다.

[0104] 또한 박테리아 숙주는 프로테오박테리아의 임의의 종의 구성원일 수 있다. 프로테오박테리아 숙주 세포는 분류군 알파프로테오박테리아(Alphaproteobacteria), 베타프로테오박테리아(Betaproteobacteria), 감마프로테오박테리아(Gammaproteobacteria), 델타프로테오박테리아(Deltaproteobacteria), 또는 엡실론프로테오박테리아(Epsilonproteobacteria) 중 임의의 하나의 구성원일 수 있다. 또한, 숙주는 분류군 알파프로테오박테리아, 베타프로테오박테리아, 또는 감마프로테오박테리아 중 임의의 하나의 구성원, 및 감마프로테오박테리아의 임의의 종의 구성원일 수 있다.

[0105] 감마 프로테오박테리아 숙주의 한 실시양태에서, 숙주는 분류군 에어로모나달레스(Aeromonadales), 알테로모나달레스(Alteromonadales), 엔테로박테리알레스(Enterobacteriales), 슈도모나달레스(Pseudomonadales), 또는 잔토모나달레스(Xanthomonadales) 중 임의의 하나의 구성원; 또는 엔테로박테리알레스 또는 슈도모나달레스 중 임의의 종의 구성원일 수 있다. 한 실시양태에서, 숙주 세포가 엔테로박테리알레스 목일 수 있거나, 숙주 세포가 엔테로박테리아세아(Enterobacteriaceae) 과의 구성원일 것이거나, 또는 에르위니아(Erwinia) 속, 에쉐리키아 속, 또는 세라티아(Serratia) 속 중 임의의 하나의 구성원; 또는 에쉐리키아 속의 구성원일 수 있다. 숙주 세포가 슈도모나달레스 목의 세포인 경우, 숙주 세포는 슈도모나스 속을 포함하는 슈도모나다세아(Pseudomonadaceae) 과의 구성원일 수 있다. 감마 프로테오박테리아 숙주에는 대장균 종의 구성원 및 슈도모나스 플루오레센스 종의 구성원이 포함된다.

[0106] 또 다른 슈도모나스 생물이 또한 유용할 수 있다. 슈도모나드 및 밀접하게 관련된 종들에는 그램-음성 프로테오박테리아 아군 1이 포함되고, 이는 [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 의해 "그램-음성 호기성 간균 및 구균"으로 기술된 과 및/또는 속에 속하는 프로테오박테리아의 군을 포함한다. 표 4는 이러한 과 및 속의 생물들을 나타낸다.

표 4

"그램-음성 호기성 간균 및 구균" 파트 ([Bergey (1974)])에 열거된 과 및 속	
과 I. 슈도모나다세아에	글루코노박테르(Gluconobacter) 슈도모나스 잔토모나스(Xanthomonas) 주글로에아(Zoogloea)
과 II. 아조토박테리아세아(Azotobacteraceae)	아조모나스(Azomonas) 아조토박테르(Azotobacter) 베이제린키아(Beijerinckia) 데륵시아(Derxia)
과 III. 리조비아세아(Rhizobiaceae)	아그로박테리움(Agrobacterium) 리조비움(Rhizobium)
과 IV. 메틸로모나다세아(Methylomonadaceae)	메틸로코쿠스(Methylococcus) 메틸로모나스(Methylomonas)
과 V. 할로박테리아세아(Halobacteriaceae)	할로박테리움(Halobacterium) 할로코쿠스(Halococcus)

기타 속	아세토박테르(Aacetobacter) 알칼리게네스(Alcaligenes) 보르데렐라(Bordetella) 브루셀라(Brucella) 프란시셀라(Francisella) 테르무스(Thermus)
------	---

[0108]

"그람-음성 프로테오박테리아 아군 1"은 분류에 사용된 기준에 따라 이러한 표제로 분류되는 프로테오박테리아를 또한 포함한다. 상기 표제는 또한 종전에는 이러한 섹션으로 분류되었지만 더 이상은 아님 군, 예컨대 애시도보락스(Acidovorax) 속, 브레분디모나스(Brevundimonas) 속, 부르크홀데리아(Burkholderia) 속, 히드로게노파가(Hydrogenophaga) 속, 오세아니모나스(Oceanimonas) 속, 랄스토니아(Ralstonia) 속 및 스테노트로포모나스(Stenotrophomonas) 속; 잔토모나스(Xanthomonas) 속에 속하는 (그리고 종전에는 잔토모나스(Xanthomonas) 속의 종으로 칭해진) 생물을 재분류하여 생성된 스펑고모나스(Sphingomonas) 속 (및 이로부터 유래된 블라스토모나스(Blastomonas) 속); [Bergery (1974)]에 정의된 아세토박테르 속에 속하는 생물을 재분류하여 생성된 애시도모나스(Acidomonas) 속을 포함한다. 또한, 숙주는 슈도모나스 속, 슈도모나스 애날리아(*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 테스토스테로니(*Pseudomonas testosteroni*) (ATCC 11996)는 각각 코마모나스 애시도보란스(*Comamonas acidovorans*) 및 코마모나스 테스토스테로니(*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 아군 1"은 임의의 하기 과에 속하는 것으로 분류된 프로토박테리아를 또한 포함한다: 슈도모나다세아에, 아조토박테라세아에 (현재 슈도모나다세아에의 "아조토박테르 그룹"이라는 동의어로 종종 칭해짐), 리조비아세아에, 및 메틸로모나다세아에 (현재 "메틸로코카세아에(Methylococcaceae)"라는 동의어로 종종 칭해짐). 따라서, 본원에 달리 기술된 속들에 더하여, "그람-음성 프로테오박테리아 아군 1"에 속하는 추가적인 프로테오박테리아 속은: 1) 아조리조필루스(*Azorhizobius*) 속의 아조토박테르 그룹 박테리아; 2) 셀비브리오(*Cellvibrio*) 속, 올리겔라(*Oligella*) 속 및 테레디니박테르(*Teredinibacter*) 속의 슈도모나다세아에 과 박테리아; 3) 켈라토박테르(*Chelatobacter*) 속, 엔시페르(*Ensifer*) 속, 리베리박테르(*Liberibacter*) 속 ("칸디다투스 리베리박테르(*Candidatus Liberibacter*)"로 또한 칭해짐), 및 시노리조비움(*Sinorhizobium*) 속의 리조비아세아에 과 박테리아; 및 4) 메틸로박테르(*Methylobacter*) 속, 메틸로칼둠(*Methylocaldum*) 속, 메틸로마이크로비움(*Methylomicrobium*) 속, 메틸로사르시나(*Methylosarcina*) 속 및 메틸로스파에라(*Methylosphaera*) 속의 메틸로코카세아에 과 박테리아를 포함한다.

[0109]

또다른 실시양태에서, 숙주 세포는 "그람-음성 프로테오박테리아 아군 2"로부터 선택된다. "그람-음성 프로테오박테리아 아군 2"는 하기 속의 프로테오박테리아의 군으로 정의된다 (카탈로그에 열거되고, 공개적으로 입수 가능하며, 기탁된 군주의 전체 개수가 팔호 안에 지시되고, 달리 지시되지 않는 한 모두 ATCC에 기탁됨): 애시도모나스 (2); 아세토박테르 (93); 글루코노박테르 (37); 브레분디모나스 (23); 베이제린키아 (13); 데륵시아 (2); 브루셀라 (4); 아그로박테리움 (79); 켈라토박테르 (2); 엔시페르 (3); 리조비움 (144); 시노리조비움 (24); 블라스토모나스 (1); 스펑고모나스 (27); 알칼리게네스 (88); 보르데렐라 (43); 부르크홀데리아 (73); 랄스토니아 (33); 애시드보락스 (20); 히드로게노파가 (9); 주글로에아 (9); 메틸로박테르 (2); 메틸로칼둠 (1, NCIMB에 기탁됨); 메틸로코쿠스 (2); 메틸로마이크로비움 (2); 메틸로모나스 (9); 메틸로사르시나 (1); 메틸로스파에라; 아조모나스 (9); 아조리조필루스 (5); 아조토박테르 (64); 셀비브리오 (3); 올리겔라 (5); 슈도모나스 (1139); 프란시셀라 (4); 잔토모나스 (229); 스테노트로포모나스 (50); 및 오세아니모나스 (4).

[0110]

"그람-음성 프로테오박테리아 아군 2"의 예시적인 숙주 세포 종에는 하기의 박테리아가 포함되지만, 이에 한정되지는 않는다 (예시적인 군주(들)의 ATCC 또는 기타 기탁 번호가 팔호 안에 제시됨): 애시도모나스 메타놀리카(*Acidomonas methanolica*) (ATCC 43581); 아세토박테르 아세티(*Acetobacter aceti*) (ATCC 15973); 글루코노박테르 옥시단스(*Gluconobacter oxydans*) (ATCC 19357); 브레분디모나스 디미누타(*Brevundimonas diminuta*) (ATCC 11568); 베이제린키아 인디카(*Beijerinckia indica*) (ATCC 9039 및 ATCC 19361); 데륵시아 굽모사

(*Derxia gummosa*) (ATCC 15994); 브루셀라 멜리텐시스(*Bruceilla melitensis*) (ATCC 23456), 브루셀라 아보르투스(*Bruceilla abortus*) (ATCC 23448); 아그로박테리움 투메파시엔스(*Agrobacterium tumefaciens*) (ATCC 23308), 아그로박테리움 라디오박테르(*Agrobacterium radiobacter*) (ATCC 19358), 아그로박테리움 리조게네스(*Agrobacterium rhizogenes*) (ATCC 11325); 켈라토박테르 해인치이(*Chelatobacter heintzii*) (ATCC 29600); 엔시퍼 아드하에렌스(*Ensifer adhaerens*) (ATCC 33212); 리조비움 레구미노사룸(*Rhizobium leguminosarum*) (ATCC 10004); 시노리조비움 프레디(*Sinorhizobium fredii*) (ATCC 35423); 블라스토모나스 나타토리아(*Blastomonas natatoria*) (ATCC 35951); 스픽고모나스 파우시모빌리스(*Sphingomonas paucimobilis*) (ATCC 29837); 알칼리게네스 파에칼리스(*Alcaligenes faecalis*) (ATCC 8750); 보르데렐라 페르투시스(*Bordetella pertussis*) (ATCC 9797); 부르크홀데리아 세파시아(*Burkholderia cepacia*) (ATCC 25416); 랄스토니아 피케티이(*Ralstonia picketii*) (ATCC 27511); 애시도보락스 파실리스(*Acidovorax facilis*) (ATCC 11228); 히드로게노파가 플라바(*Hydrogenophaga flava*) (ATCC 33667); 주글로에 라미게라(*Zoogloea ramigera*) (ATCC 19544); 메틸로박테르 루테우스(*Methylobacter luteus*) (ATCC 49878); 메틸로칼둠 그라실레(*Methylocaldum gracile*) (NCIMB 11912); 메실로코쿠스 카프라투스(*Methylococcus capsulatus*) (ATCC 19069); 메틸로마이크로비움 아길레(*Methylomicrobium agile*) (ATCC 35068); 메틸로모나스 메타니카(*Methylomonas methanica*) (ATCC 35067); 메틸로사르시나 피브라타(*Methylosarcina fibrata*) (ATCC 700909); 메틸로스파에라 한소니이(*Methylosphaera hansonii*) (ACAM 549); 아조모나스 아길리스(*Azomonas agilis*) (ATCC 7494); 아조리조필루스 파스팔리(*Azorhizophilus paspali*) (ATCC 23833); 아조토박테르 크로오코쿰(*Azotobacter chroococcum*) (ATCC 9043); 셀비브리오 믹스투스(*Cellvibrio mixtus*) (UQM 2601); 올리겔라 우레트랄리스(*Oligella urethralis*) (ATCC 17960); 슈도모나스 아에루기노사(*Pseudomonas aeruginosa*) (ATCC 10145), 슈도모나스 플루오레센스 (ATCC 35858); 프란시셀라 툴라렌시스(*Francisella tularensis*) (ATCC 6223); 스테노트로포모나스 말토필리아(*Stenotrophomonas maltophilia*) (ATCC 13637); 잔토모나스 캄페스트리스(*Xanthomonas campestris*) (ATCC 33913); 및 오세아니모나스 도우도로피(*Oceanimonas doudoroffii*) (ATCC 27123).

[0111] 또 다른 실시양태에서, 숙주 세포는 "그람-음성 프로테오박테리아 아군 3"으로부터 선택된다. "그람-음성 프로테오박테리아 아군 3"은 하기 속의 프로테오박테리아의 군으로 정의된다: 브레분디모나스; 아그로박테리아; 리조비움; 시노리조비움; 블라스토모나스; 스픽고모나스; 알칼리게네스; 부르크홀데리아; 랄스토니아; 애시드보락스; 히드로게노파가; 메틸로박테르; 메틸로칼둠; 메틸로코쿠스; 메틸로마이크로비움; 메틸로모나스; 메틸로사르시나; 메틸로스파에라; 아조모나스; 아조리조필루스; 아조토박테르; 셀비브리오; 올리겔라; 슈도모나스; 테레디니박테르; 프란시셀라; 스테노트로포모나스; 잔토모나스; 및 오세아니모나스.

[0112] 또 다른 실시양태에서, 숙주 세포는 "그람-음성 프로테오박테리아 아군 4"로부터 선택된다. "그람-음성 프로테오박테리아 아군 4"는 하기 속의 프로테오박테리아의 군으로 정의된다: 브레분디모나스; 블라스토모나스; 스픽고모나스; 부르크홀데리아; 랄스토니아; 애시드보락스; 히드로게노파가; 메틸로박테르; 메틸로칼둠; 메틸로코쿠스; 메틸로마이크로비움; 메틸로모나스; 메틸로사르시나; 메틸로스파에라; 아조모나스; 아조리조필루스; 아조토박테르; 셀비브리오; 올리겔라; 슈도모나스; 테레디니박테르; 프란시셀라; 스테노트로포모나스; 잔토모나스; 및 오세아니모나스.

[0113] 한 실시양태에서, 숙주 세포는 "그람-음성 프로테오박테리아 아군 5"로부터 선택된다. "그람-음성 프로테오박테리아 아군 5"는 하기 속의 프로테오박테리아의 군으로 정의된다: 메틸로박테르; 메틸로칼둠; 메틸로코쿠스; 메틸로마이크로비움; 메틸로모나스; 메틸로사르시나; 메틸로스파에라; 아조모나스; 아조리조필루스; 아조토박테르; 셀비브리오; 올리겔라; 슈도모나스; 테레디니박테르; 프란시셀라; 스테노트로포모나스; 잔토모나스; 및 오세아니모나스.

[0114] 숙주 세포는 "그람-음성 프로테오박테리아 아군 6"로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 6"은 하기 속의 프로테오박테리아의 군으로 정의된다: 브레분디모나스; 블라스토모나스; 스픽고모나스; 부르크홀데리아; 랄스토니아; 애시드보락스; 히드로게노파가; 아조모나스; 아조리조필루스; 아조토박테르; 셀비브리오; 올리겔라; 슈도모나스; 테레디니박테르; 프란시셀라; 스테노트로포모나스; 잔토모나스; 및 오세아니모나스.

[0115] 숙주 세포는 "그람-음성 프로테오박테리아 아군 7"로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 7"은 하기 속의 프로테오박테리아의 군으로 정의된다: 아조모나스; 아조리조필루스; 아조토박테르; 셀비브리오; 올리겔라; 슈도모나스; 테레디니박테르; 스테노트로포모나스; 잔토모나스; 및 오세아니모나스.

[0116] 숙주 세포는 "그람-음성 프로테오박테리아 아군 8"로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 8"은 하기 속의 프로테오박테리아의 군으로 정의된다: 브레분디모나스; 블라스토모나스; 스픽고모나스; 부르크

홀데리아; 랄스토니아; 애시드보락스; 히드로게노파가; 슈도모나스; 스테노트로포모나스; 잔토모나스; 및 오세아니모나스.

[0117] 숙주 세포는 "그람-음성 프로테오박테리아 아군 9"로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 9"는 하기 속의 프로테오박테리아의 군으로 정의된다: 브레분디모나스; 부르크홀데리아; 랄스토니아; 애시드보락스; 히드로게노파가; 슈도모나스; 스테노트로포모나스; 및 오세아니모나스.

[0118] 숙주 세포는 "그람-음성 프로테오박테리아 아군 10"으로부터 선택될 수 있다. "그람-음성 프로테오박테리아 아군 10"은 하기 속의 프로테오박테리아의 군으로 정의된다: 부르크홀데리아; 랄스토니아; 슈도모나스; 스테노트로포모나스; 및 잔토모나스.

[0119] 숙주 세포는 "그람 음성 프로테오박테리아 아군 11"로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 11"은 하기 속의 프로테오박테리아의 군으로 정의된다: 슈도모나스; 스테노트로포모나스; 및 잔토모나스. 숙주 세포는 "그람 음성 프로테오박테리아 아군 12"로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 12"는 하기 속의 프로테오박테리아의 군으로 정의된다: 부르크홀데리아; 랄스토니아; 슈도모나스. 숙주 세포는 "그람 음성 프로테오박테리아 아군 13"으로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 13"은 하기 속의 프로테오박테리아의 군으로 정의된다: 부르크홀데리아; 랄스토니아; 슈도모나스; 및 잔토모나스. 숙주 세포는 "그람 음성 프로테오박테리아 아군 14"로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 14"는 하기 속의 프로테오박테리아의 군으로 정의된다: 슈도모나스 및 잔토모나스. 숙주 세포는 "그람 음성 프로테오박테리아 아군 15"로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 15"는 슈도모나스 속의 프로테오박테리아의 군으로 정의된다.

[0120] 숙주 세포는 "그람 음성 프로테오박테리아 아군 16"으로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 16"은 하기의 슈도모나스 종의 프로테오박테리아의 군으로 정의된다 (예시적인 군주(들)의 ATCC 또는 기타 기탁 번호가 괄호 안에 제시됨): 슈도모나스 아비에타니필라(*Pseudomonas abietaniphila*) (ATCC 700689); 슈도모나스 아에루지노사(*Pseudomonas aeruginosa*) (ATCC 10145); 슈도모나스 알칼리게네스(*Pseudomonas alcaligenes*) (ATCC 14909); 슈도모나스 안귈리셉티카(*Pseudomonas anguilliseptica*) (ATCC 33660); 슈도모나스 시트로넬로리스(*Pseudomonas citronellolis*) (ATCC 13674); 슈도모나스 플라베센스(*Pseudomonas fluorescens*) (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 아가리치(*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 beijerinckii*) (ATCC 19372); 슈도모나스 보레알리스(*Pseudomonas borealis*); 슈도모나스 보레오플리스(*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 클로로라파스(*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 cedreliae*); 슈도모나스 코루가타(*Pseudomonas corrugata*) (ATCC 29736); 슈도모나스 엑스트레모리엔탈리스(*Pseudomonas extremonatalis*); 슈도모나스 플루오레센스 (ATCC 35858); 슈도모나스 게사르디이(*Pseudomonas gessardii*); 슈도모나스 리바넨시스(*Pseudomonas libanensis*); 슈도모나스 만델리이(*Pseudomonas mandelii*) (ATCC 700871); 슈도모나스 마르기날리스(*Pseudomonas marginalis*) (ATCC 10844); 슈도모나스 미굴라에(*Pseudomonas migulae*); 슈도모나스 무시돌렌스(*Pseudomonas mucidolens*) (ATCC 4685); 슈도모나스 오리엔탈리스(*Pseudomonas orientalis*); 슈도모나스 로데시아에(*Pseudomonas rhodesiae*); 슈도모나스 신잔타(*Pseudomonas synxantha*) (ATCC 9890); 슈도모나스 틀라아시이(*Pseudomonas tolaasii*) (ATCC 33618); 슈도모나스 베로나이(*Pseudomonas*

veronii) (ATCC 700474); 슈도모나스 프레데릭스베르겐시스(*Pseudomonas frederiksbergensis*); 슈도모나스 게니콜라타(*Pseudomonas geniculata*) (ATCC 19374); 슈도모나스 진제리(*Pseudomonas gingeri*); 슈도모나스 그라미니스(*Pseudomonas graminis*); 슈도모나스 그리몬티(*Pseudomonas grimontii*); 슈도모나스 할로데니트리피칸스(*Pseudomonas halodenitrificans*); 슈도모나스 할로필라(*Pseudomonas halophila*); 슈도모나스 히비스시콜라(*Pseudomonas hibiscicola*) (ATCC 19867); 슈도모나스 후티엔시스(*Pseudomonas huttiensis*) (ATCC 14670); 슈도모나스 히드로게노보라(*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 데니트리피칸스(*Pseudomonas denitrificans*) (ATCC 19244); 슈도모나스 페르투치노게나(*Pseudomonas pertucinogena*) (ATCC 190); 슈도모나스 퍽토룸(*Pseudomonas pictorum*) (ATCC 23328); 슈도모나스 사이크로필라(*Pseudomonas psychrophila*); 슈도모나스 필바(*Pseudomonas filva*) (ATCC 31418); 슈도모나스 몬테일리이(*Pseudomonas monteili*) (ATCC 700476); 슈도모나스 모셀리이(*Pseudomonas mosseii*); 슈도모나스 오리지하비탄스(*Pseudomonas oryzihabitans*) (ATCC 43272); 슈도모나스 플레코글로시다(*Pseudomonas plecoglossicida*) (ATCC 700383); 슈도모나스 푸티다(*Pseudomonas putida*) (ATCC 12633); 슈도모나스 레액탄스(*Pseudomonas reactans*); 슈도모나스 스피노사(*Pseudomonas spinosa*) (ATCC 14606); 슈도모나스 발레아리카(*Pseudomonas balearica*); 슈도모나스 루테올라(*Pseudomonas luteola*) (ATCC 43273); 슈도모나스 스트트제리(*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 반코우베렌시스(*Pseudomonas vancouverensis*) (ATCC 700688); 슈도모나스 위스콘시엔시스(*Pseudomonas wisconsinensis*); 및 슈도모나스 시아메넨시스(*Pseudomonas xiamensis*).

[0121] 숙주 세포는 "그람 음성 프로테오박테리아 아군 17"로부터 선택될 수 있다. "그람 음성 프로테오박테리아 아군 17"은 하기의 슈도모나스 종에 예를 들어 속하는 것들을 포함하여 "형광성 슈도모나드"로 당업계에 공지된 프로테오박테리아의 군으로 정의된다: 슈도모나스 아조토포르만스; 슈도모나스 브렌네리; 슈도모나스 세드렐라; 슈도모나스 코루가타; 슈도모나스 엑스트레모리엔탈리스; 슈도모나스 플루오레센스; 슈도모나스 게사르디; 슈도모나스 리바넨시스; 슈도모나스 만델리이; 슈도모나스 마르지날리스; 슈도모나스 미쿨라에; 슈도모나스 무시돌렌스; 슈도모나스 오리엔탈리스; 슈도모나스 로데시아에; 슈도모나스 신잔타; 슈도모나스 톨라아시이; 및 슈도모나스 베로나이.

[0122] 또 다른 적절한 숙주에는 그람(+) 프로테오박테리아와 같이 참고문헌의 다른 부분에 분류된 것들이 포함된다. 한 실시양태에서, 숙주 세포는 대장균이다. 대장균의 계놈 서열이 대장균 MG1655에 대해 확립되어 있고 ([Blattner, et al. (1997) The complete genome sequence of *Escherichia coli* K-12, *Science* 277(5331): 1453-74]), 대장균 K12에 대한 DNA 마이크로어레이가 시판된다 (MWG Inc (High Point, N.C.)). 대장균을 루리아-베르타니(Luria-Bertani) (LB) (10 g/ℓ 트립تون, 5 g/ℓ NaCl, 5 g/ℓ 효모 추출물)와 같은 영양 배지 또는 M9 (6 g/ℓ Na₂HPO₄, 3 g/ℓ KH₂PO₄, 1 g/ℓ NH₄Cl, 0.5 g/ℓ NaCl, pH 7.4)와 같은 규정 최소 배지 + 1% 글루코스와 같은 적절한 탄소 공급원에서 배양할 수 있다. 통상적으로, 대장균 세포의 하룻밤 배양액을 회석하여, 진탕 플라스크 또는 발효기 내의 신선한 영양 또는 최소 배지에 접종하고, 37 ℃에서 성장시킨다.

[0123] 또한 숙주는 임의의 인간 또는 비-인간 포유동물이 포함되는 포유동물로부터 유래된 세포와 같이, 포유류 기원일 수 있다. 포유류에는 영장류, 원숭이, 돼지(porcine), 양, 소, 설치류, 유제류, 돼지(pig), 돼지(swine), 면양(sheep), 양(lamb), 염소, 소(cattle), 사슴, 노새, 말, 원숭이, 유인원(ape), 개, 고양이, 래트, 및 마우스가 포함될 수 있지만, 이에 한정되지는 않는다.

[0124] 숙주 세포는 또한 식물 기원일 수도 있다. 유전자 및 조절 서열의 확인을 위해 임의의 식물이 선택될 수 있다. 유전자 및 조절 서열의 단리를 위한 적절한 식물 표적의 예로는 알파파, 사과, 살구, 애기장대, 아티초크, 아루굴라, 아스파라거스, 아보카도, 바나나, 보리, 콩, 비트, 블랙베리, 블루베리, 브로콜리, 방울양배추, 양배추, 카놀라, 칸탈로프, 당근, 카사바, 페마자, 꽃양배추, 셀러리, 체리, 치커리, 고수, 감귤류, 클레멘타인,

토끼풀, 코코넛, 커피, 옥수수, 목화, 크랜베리, 오이, 미송, 가지, 꽃상추, 에스캐롤(escarole), 유칼리나무, 회향, 무화과, 마늘, 호리병박, 포도, 그레이프프루트, 감로멜론, 지카마(jicama), 키위, 양상추, 부추, 레몬, 라임, 테다소나무, 아마인, 망고, 멜론, 버섯, 승도 복숭아, 견과류, 귀리, 기름야자나무, 유채, 오크라, 올리브, 양파, 오렌지, 장식용 식물, 야자, 파파야, 파슬리, 파스닙, 완두콩, 복숭아, 땅콩, 배, 후추, 감, 솔, 파인애플, 질경이, 서양자두, 석류, 포플러, 감자, 호박, 모파, 라디아타 파인(radiata pine), 래디스치오(radiscchio), 무, 펑지씨, 나무딸기, 쌀, 호밀, 사탕수수(sorghum), 남부소나무, 대두, 시금치, 호박(squash), 딸기, 사탕무, 사탕수수(sugarcane), 해바라기, 고구마, 소합향, 탄제린, 찻잎, 담배, 토마토, 라이밀, 잔디, 순무, 포도나무, 수박, 밀, 암 및 주키니가 포함되지만, 이에 한정되지는 않는다. 일부 실시양태에서, 방법에 유용한 식물은 애기장대, 옥수수, 밀, 대두, 및 목화이다.

III. 방법

[0125] 본 발명의 방법은 *pbp**, *dsbA*, *dsbC*, *Bce*, *CupA2*, *CupB2*, *CupC2*, *NikA*, *F1gI*, *ORF5550*, *Ttg2C*, 또는 *ORF812* 분비 신호로부터 선택된 분비 신호 폴리펩티드를 포함하는 융합 단백질의 발현을 제공한다. 한 실시양태에서, 상기 방법은 본 발명의 분비 신호에 연결된 관심 단백질을 발현하는 숙주 세포를 포함한다. 상기 방법은 본원에 개시된 분비 신호 서열에 작동가능하게 연결된 관심 단백질 또는 폴리펩티드를 포함하는 재조합 단백질을 코딩하는 벡터를 포함하는 숙주 세포, 바람직하게는 슈도모나스 플루오레센스 숙주 세포를 제공하는 단계, 및 상기 세포를 단백질 또는 폴리펩티드의 발현을 초래하는 조건 하에 성장시키는 단계를 포함한다. 별법적으로, 확인된 분비 신호를 사용하여 단백질 또는 폴리펩티드를 발현시키는 방법은 진핵생물 또는 원핵생물 기원의 숙주 세포를 포함하는 임의의 소정의 숙주 시스템에서 사용될 수 있다. 벡터는 상기 기술된 특성들 중 임의의 특성을 가질 수 있다. 한 실시양태에서, 벡터는 서열 6, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 또는 24로서 본원에 개시된 분비 신호 폴리펩티드, 또는 이의 변이체 및 단편을 코딩하는 뉴클레오티드 서열을 포함한다. 또다른 실시양태에서, 벡터는 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 21, 또는 23을 포함하는 뉴클레오티드 서열을 포함한다.

[0126] 또다른 실시양태에서, 숙주 세포에는 주변 세포질이 있고, 분비 신호 폴리펩티드의 발현은 실질적으로 모든 관심 단백질 또는 폴리펩티드가 세포의 주변 세포질로 표적화되는 것을 초래한다. 주변 세포질에서 발현된 단백질의 일부가 실제로는 세포막을 지나 세포의 공간 내로 누출될 수 있지만, 대다수의 표적화된 단백질은 주변 세포질 내에서 유지될 것으로 인식된다.

[0127] 또한, 발현은 세포외 단백질의 생산에 이를 수 있다. 관심 단백질 또는 폴리펩티드를 주변 세포질로부터 또는 세포외 배지로부터 정제하는 단계가 방법에 추가로 포함될 수 있다. 분비 신호가 단백질에 연결되는 방식으로 분비 신호가 발현될 수 있고, 신호가 연결된 단백질이 세포로부터 정제될 수 있다. 따라서, 한 실시양태에서, 이러한 단리된 폴리펩티드는 분비 신호 및 관심 단백질 또는 폴리펩티드의 융합 단백질이다. 그러나, 단백질이 주변 세포질로 표적화될 때 분비 신호가 단백질로부터 절단될 수도 있다. 한 실시양태에서, 분비 신호의 절단을 증가시키도록 분비 신호와 단백질 또는 폴리펩티드 간의 연결이 변형된다.

[0128] 본 발명의 방법은 숙주 세포 내에서의 관심 단백질 또는 폴리펩티드의 증가된 생산에 또한 이를 수 있다. 별법적으로, 증가된 생산은 생산된 단백질 1 g 당 또는 숙주 단백질 1 g 당 적절하게 프로세싱된 단백질 또는 폴리펩티드의 증가된 수준일 수 있다. 또한, 증가된 생산은 재조합 단백질 1 g 당 또는 숙주 세포 단백질 1 g 당 생산된 회수 가능한 단백질 또는 폴리펩티드의 증가된 수준일 수도 있다. 또한, 증가된 생산은 전체 단백질의 증가된 수준, 적절하게 프로세싱된 단백질의 증가된 수준, 또는 활성 또는 가용성 단백질의 증가된 수준의 임의의 조합일 수도 있다. 이러한 실시양태에서, 용어 "증가된"은 관심 단백질 또는 폴리펩티드가 본 발명의 분비 신호 폴리펩티드 없이 세포에서 발현되는 경우의 생산되고/되거나, 적절하게 프로세싱되고/되거나, 가용성이고/이거나 회수 가능한 단백질 또는 폴리펩티드의 수준과 관련된다.

[0129] 관심 단백질 또는 폴리펩티드의 개선된 발현은 단백질의 용해도에서의 증가를 또한 지칭할 수 있다. 관심 단백질 또는 폴리펩티드는 숙주 세포의 세포질, 주변 세포질 또는 세포외 배지로부터 생산 및 회수될 수 있다. 단백질 또는 폴리펩티드는 불용성 또는 가용성일 수 있다. 단백질 또는 폴리펩티드는, 상기 논의된 바와 같이, 하나 이상의 표적화 서열 또는 정제를 보조하는 서열을 포함할 수 있다.

[0130] 본원에서 사용된 용어 "가용성"은 생리적인 조건 하에 완충제 내에서 10-30 분 동안 회전되었을 때 약 5,000× 내지 20,000× 중력에서의 원심분리에 의해 단백질이 침전되지 않는 것을 의미한다. 가용성 단백질은 봉입체 또는 기타 침전물의 일부가 아니다. 유사하게, "불용성"은 생리적인 조건 하에 완충제 내에서 10-30 분 동안 회전되었을 때 5,000× 내지 20,000× 중력에서의 원심분리에 의해 단백질 또는 폴리펩티드가 침전될 수 있는

것을 의미한다. 불용성 단백질 또는 폴리펩티드는 봉입체 또는 기타 침전물의 일부일 수 있다. 용어 "봉입체"는 단백질 또는 폴리펩티드의 응집물이 격리되어 있는, 세포 내에 함유된 임의의 세포내 바디(body)를 포함하도록 의도된다.

[0132] 본 발명의 방법은 숙주 세포의 주변 세포질로 국소화된 단백질을 생산할 수 있다. 한 실시양태에서, 이러한 방법은 적절하게 프로세싱된 관심 단백질 또는 폴리펩티드를 세포 내에서 생산할 수 있다. 또 다른 실시양태에서, 분비 신호 폴리펩티드의 발현은 활성인 관심 단백질 또는 폴리펩티드를 세포 내에서 생산할 수 있다. 또한 본 발명의 방법은 본 발명의 분비 신호 없이 단백질이 발현되는 경우와 비교하여 관심 단백질 또는 폴리펩티드의 증가된 수율에 이를 수 있다.

[0133] 한 실시양태에서, 이러한 방법으로 적어도 0.1 g/l 의 단백질이 주변 세포질 구획 내에서 생산된다. 또 다른 실시양태에서, 이러한 방법으로 세포 내에서 0.1 내지 10 g/l 의 주변 세포질 단백질, 또는 적어도 약 0.2 , 약 0.3 , 약 0.4 , 약 0.5 , 약 0.6 , 약 0.7 , 약 0.8 , 약 0.9 또는 적어도 약 1.0 g/l 의 주변 세포질 단백질이 생산된다. 한 실시양태에서, 생산된 전체 관심 단백질 또는 폴리펩티드는 적어도 1.0 g/l , 적어도 약 2 g/l , 적어도 약 3 g/l , 약 4 g/l , 약 5 g/l , 약 6 g/l , 약 7 g/l , 약 8 g/l , 약 10 g/l , 약 15 g/l , 약 20 g/l , 적어도 약 25 g/l , 또는 그 이상이다. 일부 실시양태에서, 생산된 주변 세포질 단백질의 양은 생산된 전체 관심 단백질 또는 폴리펩티드의 적어도 약 5% , 약 10% , 약 15% , 약 20% , 약 25% , 약 30% , 약 40% , 약 50% , 약 60% , 약 70% , 약 80% , 약 90% , 약 95% , 약 96% , 약 97% , 약 98% , 약 99% , 또는 그 이상이다.

[0134] 한 실시양태에서, 이러한 방법으로 적어도 0.1 g/l 의 정확하게 프로세싱된 단백질이 생산된다. 정확하게 프로세싱된 단백질은 천연 단백질의 아미노 말단을 갖는다. 일부 실시양태에서, 관심 단백질 또는 폴리펩티드의 적어도 50% 가 천연 아미노 말단을 포함한다. 또 다른 실시양태에서, 단백질의 적어도 60% , 적어도 70% , 적어도 80% , 적어도 90% , 또는 그 이상이 천연 단백질의 아미노 말단을 갖는다. 다양한 실시양태에서, 이러한 방법으로 세포 내에서 0.1 내지 10 g/l 의 정확하게 프로세싱된 단백질이 생산되고, 여기에는 적어도 약 0.2 , 약 0.3 , 약 0.4 , 약 0.5 , 약 0.6 , 약 0.7 , 약 0.8 , 약 0.9 또는 적어도 약 1.0 g/l 의 정확하게 프로세싱된 단백질이 포함된다. 또 다른 실시양태에서, 전체적인 정확하게 프로세싱된 관심 단백질 또는 폴리펩티드는 적어도 1.0 g/l , 적어도 약 2 g/l , 적어도 약 3 g/l , 약 4 g/l , 약 5 g/l , 약 6 g/l , 약 7 g/l , 약 8 g/l , 약 10 g/l , 약 15 g/l , 약 20 g/l , 약 25 g/l , 약 30 g/l , 약 35 g/l , 약 40 g/l , 약 45 g/l , 적어도 약 50 g/l , 또는 그 이상이다. 일부 실시양태에서, 생산된 정확하게 프로세싱된 단백질의 양은 정확하게 프로세싱된 형태의 전체 재조합 단백질의 적어도 약 5% , 약 10% , 약 15% , 약 20% , 약 25% , 약 30% , 약 40% , 약 50% , 약 60% , 약 70% , 약 80% , 약 90% , 약 95% , 약 96% , 약 97% , 약 98% , 약 99% , 또는 그 이상이다.

[0135] 본 발명의 방법은 관심 단백질 또는 폴리펩티드의 증가된 수율에 또한 이를 수 있다. 한 실시양태에서, 이러한 방법으로 관심 단백질 또는 폴리펩티드가 전체 세포 단백질 (tcp)의 적어도 약 5% , 적어도 약 10% , 약 15% , 약 20% , 약 25% , 약 30% , 약 40% , 약 45% , 약 50% , 약 55% , 약 60% , 약 65% , 약 70% , 약 75% , 또는 그 이상으로서 생산된다. "전체 세포 단백질 백분율"은 총 세포 단백질의 백분율로서의 숙주 세포 내의 단백질 또는 폴리펩티드의 양이다. 전체 세포 단백질 백분율의 결정은 당업계에 주지되어 있다.

[0136] 특정 실시양태에서, 무기염 배지에서 성장되었을 때 (즉, 약 10°C , 약 15°C , 약 20°C , 약 25°C , 약 30°C , 약 35°C , 약 40°C , 약 45°C , 및 약 50°C 를 포함하는 약 4°C 내지 약 55°C 의 온도 범위 내에서 성장되었을 때), 숙주 세포는 재조합 폴리펩티드, 폴리펩티드, 단백질, 또는 이의 단편의 발현 수준이 적어도 1% tcp이고 세포 밀도가 적어도 40 g/l 일 수 있다. 특히 바람직한 실시양태에서, 적어도 약 10 리터의 발효 규모로 무기염 배지에서 성장되었을 때 (즉, 약 4°C 내지 약 55°C 의 온도 범위 (경계 포함) 내에서 성장되었을 때), 발현 시스템은 단백질 또는 폴리펩티드 발현 수준이 적어도 5% tcp일 것이고, 세포 밀도가 적어도 40 g/l 일 것이다.

[0137] 실제로는, 주변 세포질로 표적화된 이종성 단백질이 종종 브로스에서 발견되는데 (유럽 특허 번호 EP 0 288 451 참조), 아마도 세포 외막에 대한 손상 또는 세포 외막의 유동성에서의 증가 때문일 것이다. 세포 외막을 투과 성이게 하는 다양한 메커니즘을 사용함으로 이러한 "수동" 분비의 비율이 증가될 수 있다: 콜리신 ([Mikscha et al. (1997) Arch. Microbiol. 167: 143-150]); 성장률 ([Shokri et al. (2002) App Miocrobiol Biotechnol 58:386-392]); TolIII 과발현 ([Wan and Baneyx (1998) Protein Expression Purif. 14: 13-22]); 박테리오신 방출 단백질 ([Hsiung et al. (1989) Bio/Technology 7: 267-71]), 콜리신 A 용해 단백질 ([Lloubes et al. (1993) Biochimie 75: 451-8])의 주변 세포질 단백질이 누출되는 돌연변이체 ([Furlong and Sundstrom (1989)

Developments in Indus. Microbio. 30: 141-8]); 융합 파트너 ([Jeong and Lee (2002) Appl. Environ. Microbio. 68: 4979-4985]); 삼투압 충격에 의한 회수 ([Taguchi et al. (1990) Biochimica Biophysica Acta 1049: 278-85]). 조작된 단백질의 주변 세포질로의 운송과 이어지는 브로스에서의 국소화를 사용하여, 대장균에서 적절하게 풀딩되고 활성인 단백질이 생산되었다 ([Wan and Baneyx (1998) Protein Expression Purif. 14: 13-22]; [Simmons et al. (2002) J. Immun. Meth. 263: 133-147]; [Lundell et al. (1990) J. Indust. Microbio. 5: 215-27]).

A. 활성 단백질의 생산

[0138] 일부 실시양태에서, 활성 형태로 단백질이 또한 생산될 수 있다. 용어 "활성"은 생물학적 활성의 존재를 의미하고, 이때 생물학적 활성은 상응하는 천연 단백질 또는 폴리펩티드의 생물학적 활성에 필적하거나 실질적으로 상응한다. 단백질의 정황에서, 이는 폴리뉴클레오티드 또는 폴리펩티드가 표준 파라메터를 사용하여 상응하는 천연 단백질 또는 폴리펩티드와 비교했을 때 적어도 약 20%, 약 50%, 바람직하게는 적어도 약 60-80%, 가장 바람직하게는 적어도 약 90-95% 활성을 갖는 생물학적 기능 또는 효과를 포함한다는 것을 전형적을 의미한다. 단백질 또는 폴리펩티드 활성의 결정은 특정 단백질 또는 폴리펩티드에 대한 상응하는 표적화된 표준 생물학적 비교 분석법을 사용하여 수행할 수 있다. 관심 단백질 또는 폴리펩티드가 생물학적 활성을 유지한다는 것의 한 지표는 폴리펩티드가 천연 폴리펩티드와 면역학적으로 교차 반응성이라는 것이다.

[0139] 본 발명은 활성 관심 단백질 또는 폴리펩티드의 회수를 또한 개선시킬 수 있다. 활성 단백질은 서열이 유래된 천연 단백질 또는 폴리펩티드의 활성의 적어도 약 20%, 적어도 약 30%, 적어도 약 40%, 약 50%, 약 60%, 적어도 약 70%, 약 80%, 약 90%, 또는 적어도 약 95%의 특이적 활성을 가질 수 있다. 또한, 임의적으로, 기질 특이성 (k_{cat}/K_m)이 천연 단백질 또는 폴리펩티드와 실질적으로 유사하다. 전형적으로, k_{cat}/K_m 은 적어도 약 30%, 약 40%, 약 50%, 약 60%, 약 70%, 약 80%, 적어도 약 90%, 적어도 약 95%, 또는 그 이상일 것이다. 단백질 및 폴리펩티드 활성 및 기질 특이성 (k_{cat}/K_m)의 치수를 분석 및 정량하는 방법은 당업자에게 주지되어 있다.

[0140] 관심 단백질 또는 폴리펩티드의 활성을 기준에 확립된 천연 단백질 또는 폴리펩티드 표준 활성과 또한 비교할 수 있다. 별법적으로, 관심 단백질 또는 폴리펩티드의 활성을 천연 단백질 또는 폴리펩티드와의 동시 비교 분석 또는 실질적으로 동시의 비교 분석으로 결정할 수 있다. 예를 들어, 시험관내 분석법을 사용하여 관심 단백질 또는 폴리펩티드와 표적 사이, 예를 들어 발현된 효소와 기질 사이, 발현된 호르몬과 호르몬 수용체 사이, 발현된 항체와 항원 사이 등의 임의의 검출가능한 상호작용을 결정할 수 있다. 이같은 검출에는 열량측정 변화, 증식 변화, 세포 사멸, 세포 리pellent(repelling), 방사성에서의 변화, 용해도에서의 변화, 젤 전기영동 및 /또는 젤 배제 방법에 의한 측정 시 분자량에서의 변화, 인산화 능력, 항체 특이성 분석법 예컨대 ELISA 분석법 등이 포함될 수 있다. 또한, 생체내 분석법에는 천연 단백질 또는 폴리펩티드의 생리학적 효과와 비교하여 슈도모나스에 의해 생산된 천연 단백질 또는 폴리펩티드의 생리학적 효과, 예를 들어, 중량 증가, 전해질 밸런스에서의 변화, 혈액 응고 시간에서의 변화, 혈병 용해에서의 변화 및 항원성 응답의 유도를 검출하는 분석법이 포함되지만, 이에 한정되지는 않는다. 일반적으로, 관심 단백질 또는 폴리펩티드의 활성 성질을 결정하기 위해, 이같은 활성이 분석가능한 한 천연 단백질 또는 폴리펩티드에 대한 비교 분석을 가능하게 하는 임의의 시험관내 또는 생체내 분석법을 사용할 수 있다. 별법적으로, 본 발명에서 생산된 단백질 또는 폴리펩티드를, 이러한 단백질 또는 폴리펩티드와 정상적으로 상호작용하는 분자, 예를 들어 천연 단백질이 정상적으로 상호작용하는 신호 경로의 성분 또는 기질과 단백질 또는 폴리펩티드 간의 상호작용을 자극 또는 억제하는 능력에 대해 분석할 수 있다. 이같은 분석법은 단백질 또는 폴리펩티드가 표적 분자와 상호작용하도록 하는 조건 하에 단백질을 기질 분자와 조합시키는 단계를 전형적으로 포함할 수 있고, 단백질과 표적 분자와의 상호작용의 생화학적 결과를 검출한다.

[0141] 단백질 또는 폴리펩티드 활성을 결정하기 위해 사용될 수 있는 분석법이, 예를 들어, [Ralph, P. J., et al (1984) J. Immunol. 132:1858] 또는 [Saiki et al (1981) J. Immunol. 127:1044], [Steward, W. E. II (1980) The Interferon Systems. Springer-Verlag, Vienna and New York], [Broxmeyer, H. E., et al (1982) Blood 60:595], [Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989], 및 [Methods in Enzymology: Guide to Molecular Cloning Techniques, Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987], [A K Patra et al., Protein Expr Purif, 18(2): p/182-92 (2000)], [Kodama et al., J. Biochem. 99: 1465-1472 (1986)]; [Stewart et al., Proc. Nat'l Acad. Sci. USA 90: 5209-5213 (1993)]; [Lombillo et al., J. Cell Biol.

128:107-115 (1995)]; [Vale et al., Cell 42:39-50 (1985)]에 기술되어 있다.

[0143] **B. 세포 성장 조건**

[0144] 본원에 기술된 수주 세포에 대한 세포 성장 조건은 관심 단백질의 발현을 용이하게 하는 조건 및/또는 발현된 관심 단백질의 발효를 용이하게 조건을 포함할 수 있다. 본원에서 사용된 용어 "발효"는 문자 그대로의 발효가 이용되는 실시양태와 기타 비-발효성 배양 방식이 이용되는 실시양태 양쪽 모두를 포함한다. 발효는 임의의 규모로 수행될 수 있다. 한 실시양태에서, 발효 배지는 영양 배지, 최소 배지, 및 무기염 배지로부터 선택될 수 있다; 영양 배지는 사용될 수 있지만, 전형적으로는 회피된다. 또다른 실시양태에서, 최소 배지 또는 무기염 배지가 선택된다. 또다른 실시양태에서는, 최소 배지가 선택된다. 또다른 실시양태에서, 무기염 배지가 선택된다. 무기염 배지가 특히 바람직하다.

[0145] 무기염 배지는 무기염, 및 탄소 공급원, 예를 들어 글루코스, 수크로스 또는 글리세롤로 구성된다. 무기염 배지의 예로는, 예를 들어, M9 배지, 슈도모나스 배지 (ATCC 179), 데이비스 앤드 맹지올리(Davis and Mingioli) 배지 ([BD Davis & ES Mingioli (1950) J. Bact. 60:17-28] 참조)가 포함된다. 무기염 배지를 제조하는데 사용되는 무기염에는, 예를 들어, 인산칼륨, 황산암모늄 또는 염화암모늄, 황산마그네슘 또는 염화마그네슘, 및 미량의 무기물 예컨대 염화칼슘, 철, 구리, 망간 및 아연의 황산염, 및 봉산염으로부터 선택되는 것들이 포함된다. 유기 질소 공급원, 예컨대 웨톤, 트립톤, 아미노산 또는 효모 추출물은 무기염 배지에 포함되지 않는다. 그 대신, 무기 질소 공급원이 사용되고, 이는, 예를 들어, 암모늄 염, 수성 암모니아 및 기체성 암모니아로부터 선택될 수 있다. 바람직한 무기염 배지는 글루코스를 탄소 공급원으로 함유할 것이다. 무기염 배지와 비교하여, 최소 배지 또한 무기염 및 탄소 공급원을 함유할 수 있지만, 예를 들어 낮은 수준의 아미노산, 비타민, 웨톤 또는 기타 성분이, 비록 매우 최소한도의 수준으로 첨가되지만, 첨가될 수 있다.

[0146] 한 실시양태에서, 하기의 표 5에 열거된 성분을 사용하여 배지를 제조할 수 있다. 성분들은 하기의 순서로 첨가될 수 있다: 먼저 $(\text{NH}_4)\text{HPO}_4$, KH_2PO_4 및 시트르산을 약 30 ℓ 의 중류수에 용해시킬 수 있다; 그후, 미량 원소의 용액을 첨가하고 나서, 소포제, 예컨대 Ucolub N 115를 첨가할 수 있다. 그후, 열 멸균 (예컨대 약 121 °C에서의 멸균) 후, 글루코스 MgSO_4 및 티아민-HCl의 무균성 용액을 첨가할 수 있다. pH를 약 6.8로 제어하는 것은 수성 암모니아를 사용하여 달성할 수 있다. 그후, 무균성 중류수를 첨가하여 초기 부피를 371 - 글리세롤 모액 (123 $\text{m}\ell$)으로 조정할 수 있다. 화학물질은 다양한 공급원, 예컨대 Merck로부터 시판된다. 이러한 배지는 슈도모나스 종 및 관련 박테리아의 성장을 위해 높은 세포 밀도 배양 (HCDC: high cell density cultivation)을 허용할 수 있다. HCDC는 뱃치 공정으로 시작될 수 있고, 2단계 페드-뱃치(fed-batch) 배양이 이어진다. 뱃치 파트에서의 무한한 성장 후, 생물량 농도가 수배 증가할 수 있는 3회의 배가 시간의 기간에 걸쳐, 감소된 특정 성장률로 성장이 제어될 수 있다. 이같은 배양 절차에 대한 상세사항은 [Riesenber, D.; Schulz, V.; Knorre, W. A.; Pohl, H. D.; Korz, D.; Sanders, E. A.; Ross, A.; Deckwer, W. D. (1991) "High cell density cultivation of *Escherichia coli* at controlled specific growth rate" J Biotechnol: 20(1) 17-27]에 기술되어 있다..

표 5

배지 조성	
성분	초기 농도
KH ₂ PO ₄	13.3 g l ⁻¹
(NH ₄) ₂ HPO ₄	4.0 g l ⁻¹
시트르산	1.7 g l ⁻¹
MgSO ₄ ·7H ₂ O	1.2g l ⁻¹
미량 금속 용액	10 ml l ⁻¹
티아민 HCl	4.5 mg l ⁻¹
글루코스·H ₂ O	27.3 g l ⁻¹
소포제 Ucolub N115	0.1 ml l ⁻¹
공급 용액	
MgSO ₄ ·7H ₂ O	19.7 g l ⁻¹
글루코스·H ₂ O	770 g l ⁻¹
NH ₃	23 g
미량 금속 용액	
6 g l ⁻¹ 시트르산철(III) 1.5 g l ⁻¹ MnCl ₂ ·4H ₂ O	
0.8 g l ⁻¹ ZnCH ₂ COO ₂ ·2H ₂ O 0.3 g l ⁻¹ H ₃ BO ₃	
0.25 g l ⁻¹ Na ₂ MoO ₄ ·2H ₂ O 0.25 g l ⁻¹ CoCl ₂ ·6H ₂ O	
0.15 g l ⁻¹ CuCl ₂ ·2H ₂ O 0.84 g l ⁻¹ 에틸렌	
디니트릴로-테트라아세트산 Na ₂ sah 2H ₂ O	
(Tritriplex III, Merck)	

[0147]

[0148]

본 발명에 따른 발현 시스템은 임의의 발효 형식으로 배양될 수 있다. 예를 들어, 뱃치, 페드-뱃치, 반-연속 및 연속 발효 방식이 본원에서 사용될 수 있다. 단백질이 세포외 배지 내로 배출되는 경우, 연속 발효가 바람직하다.

[0149]

본 발명에 따른 발현 시스템은 임의 규모 (즉, 부피)의 발효에서의 트랜스진 발현에 유용하다. 따라서, 예를 들어, 마이크로리터 규모, 센티리터 규모, 및 데시리터 규모의 발효 부피가 사용될 수 있고; 1 리터 규모 및 더 큰 발효 부피가 사용될 수 있다. 한 실시양태에서, 발효 부피는 1 리터 이상일 것이다. 또 다른 실시양태에서, 발효 부피는 5 리터 이상, 10 리터 이상, 15 리터 이상, 20 리터 이상, 25 리터 이상, 50 리터 이상, 75 리터 이상, 100 리터 이상, 200 리터 이상, 500 리터 이상, 1,000 리터 이상, 2,000 리터 이상, 5,000 리터 이상, 10,000 리터 이상 또는 50,000 리터 이상일 것이다.

[0150]

본 발명에서, 형질전환된 숙주 세포의 성장, 배양 및/또는 발효는 숙주 세포의 생존을 허용하는 온도 범위, 바람직하게는 약 4 °C 내지 약 55 °C (경계 포함) 범위 내의 온도에서 수행된다. 따라서, 예를 들어, 본 발명의 숙주 세포와 관련되어 본원에서 사용된 용어 "성장" (및 "성장하다", "성장하는"), "배양" (및 "배양하다"), 및 "발효" (및 "발효시키다", "발효시키는")은 약 4 °C 내지 약 55 °C (경계 포함)의 온도 범위에서의 "성장", "배양" 및 "발효"를 본질적으로 의미한다. 또한, "성장"은 활성 세포 분열 및/또는 확장의 생물학적 상태, 뿐만 아니라 분열하지 않고/않거나 확장되지 않는 세포가 대사적으로 유지되는 생물학적 상태 양쪽 모두를 지시하도록 사용되고, 후자의 생물학적 상태에 대해 사용될 때의 용어 "성장"은 용어 "유지"와 동의어이다.

[0151]

분비형 단백질을 발현시키는 것에서 슈도모나스 플루오레센스를 사용하는 것의 추가적인 장점에는 일부 또 다른 박테리아 발현 시스템과 비교하여 높은 세포 밀도로 성장되는 슈도모나스 플루오레센스의 능력이 포함된다. 이 때문에, 본 발명에 따른 슈도모나스 플루오레센스 발현 시스템은 약 20 g / l 이상의 세포 밀도를 제공할 수 있다. 게다가 본 발명에 따른 슈도모나스 플루오레센스 발현 시스템은 부피 당 생물량의 단위로 언급된다면 약 70 g / l 이상의 세포 밀도를 제공할 수 있고, 이때 생물량은 건조 세포 중량으로 측정된다.

[0152]

한 실시양태에서, 세포 밀도는 적어도 약 20 g/ℓ 일 것이다. 또다른 실시양태에서, 세포 밀도는 적어도 약 25 g/ℓ, 약 30 g/ℓ, 약 35 g/ℓ, 약 40 g/ℓ, 약 45 g/ℓ, 약 50 g/ℓ, 약 60 g/ℓ, 약 70 g/ℓ, 약 80 g/ℓ, 약 90 g/ℓ, 약 100 g/ℓ, 약 110 g/ℓ, 약 120 g/ℓ, 약 130 g/ℓ, 약 140 g/ℓ, 약 또는 적어도 약 150 g/ℓ 일 것이다.

[0153]

또다른 실시양태에서, 유도시의 세포 밀도는 약 20 g/ℓ 내지 약 150 g/ℓ 사이; 약 20 g/ℓ 내지 약 120 g/ℓ 사이; 약 20 g/ℓ 내지 약 80 g/ℓ 사이; 약 25 g/ℓ 내지 약 80 g/ℓ 사이; 약 30 g/ℓ 내지 약 80 g/ℓ 사이; 약 35 g/ℓ 내지 약 80 g/ℓ 사이; 약 40 g/ℓ 내지 약 80 g/ℓ 사이; 약 45 g/ℓ 내지 약 80 g/ℓ 사이; 약 50 g/ℓ 내지 약 80 g/ℓ 사이; 약 50 g/ℓ 내지 약 75 g/ℓ 사이; 약 50 g/ℓ 내지 약 70 g/ℓ 사이; 약 40 g/ℓ 내지 약 80 g/ℓ 사이일 것이다.

[0154]

C. 관심 단백질 또는 폴리펩티드의 단리

[0155]

관심 단백질의 수율, 용해도, 형상 및/또는 활성을 측정하기 위해, 숙주 세포 및/또는 세포외 배지로부터 단백질을 단리하는 것이 바람직할 수 있다. 적합하게 측정하도록 사용된 분석법의 요구사항에 따라, 단리는 미정제 단리, 반-정제 단리 또는 순수 단리일 수 있다. 단백질은 세포질에서 생산될 수 있거나, 주변 세포질로 표적화될 수 있거나, 배양 또는 발효 배지 내로 분비될 수 있다. 표적화된 단백질을 주변 세포질로부터 방출시키기 위해, 클로로포름 ([Ames et al. (1984) J. Bacteriol., 160: 1181-1183]), 구아니딘-HCl, 및 트리톤(Triton) X-100 ([Naglak and Wang (1990) Enzyme Microb. Technol., 12: 603-611])과 같은 화학물질을 수반하는 처리가 사용되었다. 그러나, 이러한 화학물질들은 불활성화되어 않고, 다수의 재조합 단백질 생성물 또는 이어지는 정제 절차에 대해 불리한 효과가 있을 수 있다. 외막을 투과성하게 만드는, 대장균 세포의 글리신 처리가 원형질막 주위공간 내용물을 방출시키는 것으로 또한 보고되었다 ([Ariga et al. (1989) J. Ferm. Bioeng., 68: 243-246]). 재조합 단백질의 가장 널리 사용되는 주변 세포질 방출 방법은 삼투압 충격 ([Nosal and Heppel (1966) J. Biol. Chem., 241: 3055-3062]; [Neu and Heppel (1965) J. Biol. Chem., 240: 3685-3692]), 암탉 혼자 (HEW)-라이소자임/에틸렌디아민 테트라아세트산 (EDTA) 처리 ([Neu and Heppel (1964) J. Biol. Chem., 239: 3893-3900]; [Witholt et al. (1976) Biochim. Biophys. Acta, 443: 534-544]; [Pierce et al. (1995) ICHEME Research. Event, 2: 995-997]), 및 조합형 HEW-라이소자임/삼투압 충격 처리 ([French et al. (1996) Enzyme and Microb. Tech., 19: 332-338])이다. 프렌치(French) 방법은 분획화 완충제에 세포를 재현탁시킨 후, 주변 세포질 분획을 회수하고, 이때 라이소자임 처리 직후 삼투압 충격이 이어지는 것을 수반한다.

[0156]

전형적으로, 이러한 절차는 삼투압적으로 안정화시키는 배지에서의 초기 파괴에 이어지는 비-안정화 배지에서의 선택적 방출을 포함한다. 이러한 배지의 조성 (pH, 보호제) 및 사용된 파괴 방법 (클로로포름, HEW-라이소자임, EDTA, 초음파처리)은 보고된 특정 절차들 간에 상이하다. EDTA 대신 2극성 이온 세제를 사용하는 HEW-라이소자임/EDTA 처리에 대한 변형이 [Stabel et al. (1994) Veterinary Microbiol., 38: 307-314]에 논의되어 있다. 대장균을 파괴하기 위한 세포내 용해 효소 시스템의 사용에 대한 일반적인 리뷰에 대해서, [Dabora and Cooney (1990), Advances in Biochemical Engineering/Biotechnology, Vol. 43, A. Fiechter, ed. (Springer-Verlag: Berlin), pp. 11-30] 참조.

[0157]

가용성 단백질 또는 굴절성 입자로서 세포질로부터 관심 단백질 또는 폴리펩티드를 회수하기 위한 통상적인 방법은 기계적인 파괴에 의한 박테리아 세포의 파쇄를 수반하였다. 기계적인 파열은 액체 혼탁액 내의 국소적인 공동화(cavitation)의 생성, 강성 비드와의 급속 진탕, 초음파 처리, 또는 세포 혼탁액의 분쇄를 전형적으로 수반한다 ([Bacterial Cell Surface Techniques, Hancock and Poxton (John Wiley & Sons Ltd, 1988), Chapter 3, p. 55]).

[0158]

생화학적으로, HEW-라이소자임은 세포벽의 펩티도글리칸 골격을 가수분해하는 작용을 한다. [Zinder and Arndt (1956) Proc. Natl. Acad. Sci. USA, 42: 586-590]에 의해 최초로 방법이 개발되었고, 이들은 대장균을 난 알부민 (HEW-라이소자임을 함유함)으로 처리하여, 스페로플라스트(spheroplast)로 추후에 공지된 둥글려진 세포 구체를 생산하였다. 이러한 구조물에서 일부 세포벽 성분이 유지되었지만, 세포질막이 노출된 넓은 표면적이 있었다. 미국 특허 번호 5,169,772에는 박테리아의 외피를 삼투압적으로 안정화된 배지, 예를 들어, 20% 수크로스 용액에서 EDTA, 라이소자임, 또는 유기 화합물을 예를 들어 사용하여 파괴시키는 단계, 박테리아를 이온 강도가 낮은 완충제에 노출시킴으로써 파괴된 박테리아의 주변 세포질로부터 비-헤파리나제 유사 단백질을 방출시키는 단계, 및 낮은 이온 강도로 세정된 박테리아를 완충된 염 용액에 노출시킴으로써 헤파리나제-유사 단백질을 방출시키는 단계를 포함하는, 박테리아로부터 헤파리나제를 정제하는 방법이 개시되어 있다.

[0159]

이러한 방법들의 다수의 상이한 변형들이 다양한 성공률로 광범위한 발현 시스템에 대해 사용되었다 ([Joseph-

Liazun et al. (1990) *Gene*, 86: 291-295]; [Carter et al. (1992) *Bio/Technology*, 10: 163-167]). 라이소자임을 생산하도록 재조합 세포 배양물을 유도하려는 시도가 보고되었다. EP 0 155 189에는 라이소자임을 생산하도록 재조합 세포 배양물을 유도하기 위한 수단이 개시되어 있고, 통상적으로 이는 세포벽 구조물을 파괴하거나 용해시키는 수단에 의해 이같은 숙주 세포를 사망시킬 것으로 예상될 것이다.

[0160] 미국 특허 번호 4,595,658에는 대장균의 주변 세포질로 운송된 단백질의 외부화를 용이하게 하는 방법이 개시되어 있다. 이러한 방법은 세포를 라이소자임으로 처리하거나, 기계적으로 분쇄하거나 삼투압 충격으로 처리할 필요 없이 주변 세포질 내에 위치하는 단백질이 선택적으로 단리되도록 한다. 미국 특허 번호 4,637,980에는 온도-민감성 용원(lysogen)을 생성물을 직접적으로 또는 간접적으로 코딩하는 DNA 분자로 형질전환시키는 단계, 형질전환체를 허용성 조건 하에 배양하여, 유전자 생성물을 세포내에서 발현시키는 단계, 및 파지에 의해 코딩되는 기능을 유도하도록 온도를 상승시킴으로써 생성물을 외부화시키는 단계에 의해 박테리아 생산물을 생산하는 것이 개시되어 있다. [Asami et al. (1997) *J. Ferment, and Bioeng.*, 83: 511-516]에는 T4 파지 감염에 의한 대장균 세포의 동기(synchronized) 파괴가 개시되어 있고, [Tanji et al. (1998) *J. Ferment, and Bioeng.*, 85: 74-78]에는 대장균 세포의 완만한 파괴를 위한 T4 파지에서 코딩되는 용해 유전자의 제어 발현이 개시되어 있다.

[0161] 세포 용해 시, 계놈 DNA가 세포질로부터 배지 내로 누출되고, 유체 점도에서의 상당한 증가가 초래되며, 이는 원심분리 필드에서 고체의 침강을 방해할 수 있다. 전단력, 예컨대 DNA 중합체를 파괴시키기 위한 기계적 봉괴 동안 가해지는 전단력의 부재 하에, 점성 유체를 통한 고체의 침강 속도 저하는 원심분리 동안 고체 및 액체의 불량한 분리를 초래한다. 기계적 전단력 이외에, DNA 중합체를 분해하는 뉴클레오라이티(nucleolytic) 효소가 존재한다. 대장균에서, 내인성 유전자 endA는 엔도뉴클레아제 (성숙형 단백질의 분자량은 약 24.5 kD임)를 코딩하고, 이는 일반적으로 주변 세포질로 분비되고, DNA를 엔도뉴클레오라이티(endonucleolytic) 방식으로 올리고데옥시리보뉴클레오티드로 절단한다. endA가 대장균에 의해 비교적 약하게 발현된다는 것이 제안되었다 ([Wackemagel et al. (1995) *Gene* 154: 55-59]).

[0162] 한 실시양태에서, 숙주 세포로부터 활성이고 가용성인 형태의 디슬퍼드 결합을 함유하는 확인된 폴리펩티드를 회수하기 위해, 추가적인 디슬퍼드 결합-촉진 조건 또는 작용제가 요구되지 않는다. 한 실시양태에서, 트랜스제닉 폴리펩티드, 폴리펩티드, 단백질, 또는 이의 단편은 이의 활성 상태에서 풀딩된 분자내 형상을 갖는다. 한 실시양태에서, 트랜스제닉 폴리펩티드, 폴리펩티드, 단백질, 또는 단편은 이의 활성 상태에서 하나 이상의 분자내 디슬퍼드 결합을 함유하고, 아마도 2개, 4개, 6개, 8개, 10개, 12개, 14개, 16개, 18개, 또는 20개 또는 그이상까지의 디슬퍼드 결합을 함유할 것이다.

[0163] 본 발명의 단백질은 황산암모늄 또는 에탄올 침전, 산 추출, 음이온 또는 양이온 교환 크로마토그래피, 포스포셀룰로스 크로마토그래피, 소수성 상호작용 크로마토그래피, 친화성 크로마토그래피, 니켈 크로마토그래피, 히드록실아파타이트 크로마토그래피, 역상 크로마토그래피, 렉틴 크로마토그래피, 분취용 전기영동, 세제 가용화, 컬럼 크로마토그래피와 같은 물질로의 선택 침전, 면역정제 방법 등이 포함되지만 이에 한정되지 않는 당업계에 주지된 표준 기술에 의해 상당한 순도로 단리 및 정제될 수 있다. 예를 들어, 분자 부착성이 확립된 단백질이 가역적으로 리간드와 융합될 수 있다. 적절한 리간드와 함께, 단백질이 정제 컬럼에 선택적으로 흡착된 후, 비교적 순수한 형태로 컬럼으로부터 유리될 수 있다. 이어서 융합된 단백질을 효소적 활성에 의해 제거한다. 또한, 면역친화성 컬럼 또는 Ni-NTA 컬럼을 사용하여 단백질을 정제할 수 있다. 일반적인 기술이, 예를 들어, [R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: N.Y. (1982)]; [Deutscher, *Guide to Protein Purification*, Academic Press (1990)]; 미국 특허 번호 4,511,503; [S. Roe, *Protein Purification Techniques: A Practical Approach (Practical Approach Series)*, Oxford Press (2001)]; [D. Bollag, et al., *Protein Methods*, Wiley-Lisa, Inc. (1996)]; [AK Patra et al., *Protein Expr Purifi*, 18(2): p/182-92 (2000)]; 및 [R. Mukhija, et al., *Gene* 165(2): p. 303-6 (1995)]에 추가로 기술되어 있다. 예를 들어, [Ausubel, et al (1987 & 정기 증보판)]; [Deutscher (1990) "Guide to Protein Purification", Methods in Enzymology vol. 182], 및 이러한 시리즈의 다른 권; [Coligan, et al. (1996 & 정기 증보판) *Current Protocols in Protein Science* Wiley/Greene, NY]; 및 단백질 정제 생성물의 사용에 대한 제조업자, 예를 들어, Pharmacia (Piscataway, N.J.), 또는 Bio-Rad (Richmond, Calif)의 문헌을 또한 참조한다. 재조합 기술과의 조합은 적합한 절편에, 예를 들어, 프로테아제로 제거가능한 서열을 통해 융합될 수 있는 FLAG 서열 또는 동가물에 융합되도록 한다. 예를 들어, [Hochuli (1989) *Chemische Industrie* 12:69-70]; [Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) *Genetic Engineering, Principle and Methods* 12:87-98, Plenum Press, NY]; 및 [Crowe, et al. (1992) *QIAexpress*:

The High Level Expression & Protein Purification System QUIAGEN, Inc., Chatsworth, Calif]를 또한 참조 한다.

[0164] 발현된 단백질의 검출은 당업계에 공지된 방법에 의해 달성되고, 예를 들어, 방사성면역분석법, 웨스턴 블로팅 기술 또는 면역침전이 포함된다.

[0165] 본 발명에서 발현된 특정 단백질은 불용성 응집물 ("봉입체")을 형성할 수 있다. 여러 프로토콜이 봉입체로부터 단백질을 정제하는데 적절하다. 예를 들어, 봉입체의 정제에는 숙주 세포의 파열에 의한, 예를 들어, 50 mM TRIS/HCl pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1 mM DTT, 0.1 mM ATP 및 1 mM PMSF의 완충액에서의 인큐베이션에 의한 봉입체의 추출, 분리 및/또는 정제가 전형적으로 수반된다. 전형적으로 세포 혼탁액을 프렌치 프레스 (French Press)에 2-3회 통과시키는 것을 사용하여 용균시킨다. 또한 세포 혼탁액을 Polytron (Brinkman Instruments)을 사용하여 균질화시키거나, 얼음 상에서 초음파처리할 수 있다. 박테리아의 별법적인 용해 방법들이 당업자에게 명백하다 (예를 들어, [Sambrook et al., 상기 문현]; [Ausubel et al., 상기 문현] 참조).

[0166] 필요하다면, 봉입체를 가용화시킬 수 있고, 용해된 세포 혼탁액을 전형적으로 원심분리하여 원치 않는 불용성 물질을 제거할 수 있다. 봉입체를 형성한 단백질은 상용성 완충제로의 회석 또는 투석에 의해 복원될 수 있다. 적절한 용매로는, 요소 (약 4 M 내지 8 M), 포름아미드 (부피/부피 기준으로 약 80 % 이상) 및 구아니딘 히드로클로라이드 (약 4 M 내지 약 8 M)가 포함되지만 이에 한정되지는 않는다. 구아니딘 히드로클로라이드 및 유사한 작용제는 변성제이지만, 이러한 변성은 비가역적이지 않고, 변성제의 제거 (예를 들어, 투석에 의한 제거) 또는 회석 시 복원이 일어나서, 면역학적으로 및/또는 생물학적으로 활성인 단백질이 재형성될 수 있다. 또다른 적절한 완충제들이 당업자에게 공지되어 있다.

[0167] 상등액 내에 존재하는 이종성으로 발현된 단백질을 당업자에게 주지된 표준 분리 기술에 의해 숙주 단백질로부터 분리할 수 있다. 예를 들어, 초기 염 분획화로 다수의 원치 않는 숙주 세포 단백질 (또는 세포 배양 배지로부터 유래된 단백질)을 관심 단백질 또는 폴리펩티드로부터 분리할 수 있다. 한 이같은 예는 황산암모늄일 수 있다. 황산암모늄은 단백질 혼합물을 내의 물의 양을 효과적으로 감소시킴으로써 단백질을 침전시킨다. 그러면, 단백질이 이의 용해도를 기초로 침전된다. 단백질이 더 소수성일수록, 더 낮은 황산암모늄 농도에서 더 쉽게 침전한다. 전형적인 프로토콜은 결과적인 황산암모늄 농도가 20-30 % 사이이도록 포화 황산암모늄을 단백질 용액에 첨가하는 것을 포함한다. 이러한 농도는 가장 소수성인 단백질을 침전시킬 것이다. 그후, 침전물을 폐기하고 (관심 단백질이 소수성이지 않다면), 관심 단백질을 침전시키는 것으로 알려진 농도로 상등액에 황산암모늄을 첨가한다. 그후, 침전물을 완충제에서 가용화시키고, 필요하다면 투석 또는 투석여과를 통해 과량의 염을 제거한다. 단백질의 용해도에 의존하는 기타 방법, 예컨대 저온 에탄올 침전이 당업자에게 주지되어 있고, 복합 단백질 혼합물을 분획화하는데 사용될 수 있다.

[0168] 관심 단백질 또는 폴리펩티드의 분자량을 사용하여, 구멍 크기가 상이한 막 (예를 들어, Amicon 또는 Millipore 막)을 통과시키는 초여과를 사용하여 더 큰 크기 또는 더 작은 크기의 단백질로부터 관심 단백질 또는 폴리펩티드를 단리할 수 있다. 첫번째 단계로서, 단백질 혼합물을 분자량 차단이 관심 단백질의 분자량보다 낮은 구멍 크기를 갖는 막을 통과시켜 초여과시킬 수 있다. 그후, 초여과 보전물을 분자량 차단이 관심 단백질의 분자량보다 큰 막에 대해 초여과시킬 수 있다. 관심 단백질 또는 폴리펩티드는 막을 통과하여 여과액이 될 것이다. 그후, 여과액을 하기 기술되는 바와 같이 크로마토그래프할 수 있다.

[0169] 분비된 관심 단백질 또는 폴리펩티드를 이의 크기, 표면 알짜 전하, 소수성, 및 리간드에 대한 친화력을 기초로 다른 단백질들로부터 또한 분리할 수 있다. 또한, 단백질에 대해 생성된 항체를 컬럼 매트릭스에 접합시켜서, 단백질을 면역정제할 수 있다. 모든 이러한 방법은 당업계에 주지되어 있다. 다수의 상이한 제조업자 (예를 들어, Pharmacia Biotech)로부터의 서비스를 사용하여 임의의 규모로 크로마토그래피 기술을 수행할 수 있다는 것이 당업자에게 명백할 것이다.

D. 재생 및 리풀딩

[0171] 본 발명의 일부 실시양태에서, 생산된 트랜스제닉 폴리펩티드, 폴리펩티드, 단백질, 또는 이의 단편 발현물의 50% 초과가 재생가능한 형태로 숙주 세포에서 생산될 수 있다. 또다른 실시양태에서, 발현된 단백질의 약 60 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %가 활성 형태로 수득되거나, 활성 형태로 재생될 수 있다.

[0172] 2차 및 3차 단백질 구조 형상이 생성되도록 불용성 단백질이 재생 또는 리풀딩될 수 있다. 재조합 생성물의 구성의 완성에서, 필요하다면 단백질 리풀딩 단계가 사용될 수 있다. 리풀딩 및 재생은 단백질의 해리/회합을 촉진하는 것으로 당업계에 공지된 작용제를 사용하여 달성할 수 있다. 예를 들어, 단백질을 디티오트레이틀과 함

께 인큐베이션한 후, 산화된 글루타티온 2나트륨 염과 함께 인큐베이션하고, 이어서 요소와 같은 리풀딩제를 함유하는 완충제와 함께 인큐베이션할 수 있다.

[0173] 예를 들어, 관심 단백질 또는 폴리펩티드를 포스페이트-완충 염수 (PBS) 또는 50 mM Na-아세테이트, pH 6 완충제 + 200 mM NaCl에 대해 투석함으로써 관심 단백질 또는 폴리펩티드가 또한 재생될 수 있다. 별법적으로, 단백질을 Ni-NTA 컬럼과 같은 컬럼 상에 고정시키고, 프로테아제 억제제를 함유하는 500 mM NaCl, 20% 글리세롤, 20 mM 트리스/HCl pH 7.4에서 선형 6M - 1M 요소 구배를 사용함으로써, 단백질이 리풀딩될 수 있다. 재생은 1.5 시간 이상의 기간에 걸쳐 수행될 수 있다. 재생 후, 250 mM 이미다졸의 첨가에 의해 단백질이 용출될 수 있다. PBS 또는 50 mM 아세트산나트륨 pH 6 완충제 + 200 mM NaCl에 대한 최종 투석 단계에 의해 이미다졸이 제거될 수 있다. 정제된 단백질을 4 °C에서 보관하거나 -80 °C에서 동결시킬 수 있다.

[0174] 기타 방법으로는 [M H Lee et al., Protein Expr. Purif., 25(1): p. 166-73 (2002)], [W. K. Cho et al., J. Biotechnology, 77(2-3): p. 169-78 (2000)], [Ausubel, et al (1987 & 정기 증보판)], [Deutscher (1990) "Guide to Protein Purification", Methods in Enzymology vol. 182], 및 이러한 시리즈의 다른 권, [Coligan, et al. (1996 & 정기 증보판) Current Protocols in Protein Science Wiley/Greene, NY], [S. Roe, Protein Purification Techniques: A Practical Approach (Practical Approach Series), Oxford Press (2001)]; [D. Bollag, et al., Protein Methods, Wiley-Lisa, Inc. (1996)]에 기술될 수 있는 것들이 예를 들어 포함된다.

E. 관심 단백질

[0175] 본 발명의 방법 및 조성물은 세포 발현 시스템에서 높은 수준의 적절하게 프로세싱된 관심 단백질 또는 폴리펩티드를 생산하는데 유용하다. 관심 단백질 또는 폴리펩티드 (본원에서 "표적 단백질" 또는 "표적 폴리펩티드"로 또한 지칭됨)는 임의의 종 및 임의의 크기일 수 있다. 그러나, 특정 실시양태에서, 관심 단백질 또는 폴리펩티드는 치료적으로 유용한 단백질 또는 폴리펩티드이다. 일부 실시양태에서, 단백질은 포유류 단백질, 예를 들어 인간 단백질일 수 있고, 예를 들어, 성장 인자, 사이토카인, 케모카인 또는 혈액 단백질일 수 있다. 관심 단백질 또는 폴리펩티드는 천연 단백질 또는 폴리펩티드와 유사한 방식으로 프로세싱될 수 있다. 특정 실시양태에서, 단백질 또는 폴리펩티드는 코딩 서열 내에 분비 신호를 포함하지 않는다. 특정 실시양태에서, 관심 단백질 또는 폴리펩티드는 크기가 100 kD 미만, 50 kD 미만, 또는 30 kD 미만이다. 특정 실시양태에서, 관심 단백질 또는 폴리펩티드는 아미노산 적어도 약 5개, 10개, 15개, 20개, 30개, 40개, 50개 또는 100개의 폴리펩티드이다.

[0176] 분자 유전학 및 유전 공학 기술에 필요한 광범위한 서열 정보는 널리 공개적으로 입수 가능하다. 포유류, 뿐만 아니라 인간의 완전한 뉴클레오티드 서열, 유전자, cDNA 서열, 아미노산 서열 및 게놈에 대한 액세스는 GenBank로부터 웹사이트 //www.ncbi.nlm.nih.gov/Entrez에서 수득할 수 있다. 추가적인 정보를 유전자 및 이의 생성물 및 생물의학 용도에 관한 정보가 통합되어 있는 전자 백과사전인 Weizmann Institute of Science Genome and Bioinformatics의 GeneCards (bioinformatics.weizmann.ac.il/cards)로부터 또한 수득할 수 있고, 뉴클레오티드 서열 정보를 EMBL Nucleotide Sequence Database (www.ebi.ac.uk/emb1/) 또는 DNA Databank of Japan (DDBJ, www.ddbj.nig.ac.jp/)으로부터 또한 수득할 수 있다; 아미노산 서열에 관한 정보에 대한 추가적인 사이트로는 Georgetown의 단백질 정보 리소스 웹사이트 (www-nbrf.Rgeorgetown.edu/pirl) 및 Swiss-Prot (au.expasy.org/sprot/sprot-top.html)이 포함된다.

[0177] 본 발명에서 발현될 수 있는 단백질의 예로는, 예를 들어, 레닌, 인간 성장 호르몬, 소 성장 호르몬을 포함하는 성장 호르몬; 성장 호르몬 방출 인자; 부갑상선 호르몬; 갑상선 자극 호르몬; 지단백질; α -1-항트립신; 인슐린 A-사슬; 인슐린 B-사슬; 프로인슐린; 트롬보포이에틴; 난포 자극 호르몬; 칼시토닌; 황체형성 호르몬; 글루카곤; 응고 인자 예컨대 VIIIC 인자, IX 인자, 조직 인자, 및 폰 빌레브란츠 인자; 항-응고 인자 예컨대 단백질 C; 심방 나트륨이뇨 인자; 폐 계면활성제; 폴리스미노겐 활성화제, 예컨대 유로키나제 또는 인간 소변 또는 조직-유형 폴라스미노겐 활성화제 (t-PA); 봄베신; 트롬빈; 조혈 성장 인자; 종양 피사 인자-알파 및 -베타; 엔케팔리나아제; 혈청 알부민 예컨대 인간 혈청 알부민; 물리관 억제 물질; 릴렉신 A-사슬; 릴렉신 B-사슬; 프로릴렉신; 마우스 생식샘자극호르몬-관련 폴리펩티드; 미생물 단백질 예컨대 베타-락타마제; DNase; 인히빈; 액티빈; 혈관 내피 성장 인자 (VEGF); 호르몬 또는 성장 인자에 대한 수용체; 인테그린; 단백질 A 또는 D; 류머티스 인자; 항신경 인자 예컨대 뇌-유래 항신경 인자 (BDNF), 뉴로트로핀-3, -4, -5, 또는 -6 (NT-3, NT-4, NT-5, 또는 NT-6), 또는 신경 성장 인자 예컨대 NGF- β ; 카디오토로핀 (심장 비대 인자) 예컨대 카디오토로핀-1 (CT-1); 혈소판-유래 성장 인자 (PDGF); 섬유모세포 성장 인자 예컨대 aFGF 및 bFGF; 유피 성장 인자 (EGF); 전환 성장 인자 (TGF) 예컨대 TGF-알파 및 TGF- β (TGF- β 1, TGF- β 2, TGF- β 3, TGF- β 4, 또는 TGF- β 5 포함);

인슐린 유사 성장 인자-I 및 -II (IGF-I 및 IGFII); des(1-3)-IGF-I (대뇌 IGF-I), 인슐린 유사 성장인자 결합 단백질; CD 단백질 예컨대 CD-3, CD-4, CD-8, 및 CD-19; 에리트로포이에틴; 골유도성 인자; 면역독소; 골 형태 형성 단백질 (BMP); 인터페론 예컨대 인터페론-알파, -베타 및 -감마; 콜로니 자극 인자 (CSF) 예컨대 M-CSF, GM-CSF, 및 G-CSF; 인터루킨 (IL), 예를 들어 IL-1 내지 IL-10; 항-HER-2 항체; 과산화물 디스뮤타제; T-세포 수용체; 표면 막 단백질; 붕괴(decay) 가속화 인자; 바이러스 항원, 예를 들어, AIDS 외피의 일부분; 운송 단백질; 귀소(homing) 수용체; 어드레신; 조절 단백질; 항체; 및 상기 열거된 폴리펩티드들 중 임의의 것의 단편과 같은 문자가 포함된다.

[0179] 특정 실시양태에서, 단백질 또는 폴리펩티드는 IL-1, IL-1a, IL-1b, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-12엘라스티(elasti), IL-13, IL-15, EL-16, IL-18, IL-18BPa, IL-23, IL-24, VIP, 에리트로포이에틴, GM-CSF, G-CSF, M-CSF, 혈소판 유래 성장인자 (PDGF), MSF, FLT-3 리간드, EGF, 섬유모세포 성장인자 (FGF; 예를 들어, α -FGF (FGF-1), β -FGF (FGF-2), FGF-3, FGF-4, FGF-5, FGF-6, 또는 FGF-7), 인슐린 유사 성장인자 (예를 들어, IGF-1, IGF-2); 종양 고사 인자 (예를 들어, TNF, 림프독소), 신경 성장인자 (예를 들어, NGF), 혈관 내피 성장 인자 (VEGF); 인터페론 (예를 들어, IFN- α , IFN- β , IFN- γ); 백혈병 억제 인자 (LIF); 섬모 향신경 인자 (CNTF); 옹코스타틴 M; 줄기 세포 인자 (SCF); 전환 성장 인자 (예를 들어, TGF- α , TGF- β 1, TGF- β 2, TGF- β 3); TNF 수퍼페밀리 (예를 들어, LIGHT/TNFSF14, STALL-1/TNFSF13B (BLY5, BAFF, THANK), TNF알파/TNFSF2 및 TWEAK/TNFSF12); 또는 캐모카인 (BCA-1/BLC-1, BRAK/Kec, CXCL16, CXCR3, ENA-78/LIX, 이오팩신(Eotaxin)-1, 이오팩신-2/MPIF-2, 엑소더스(Exodus)-2/SLC, 프락탈킨 (Fractalkine)/뉴로택틴(Neurotactin), GRO알파/MGSA, HCC-1, I-TAC, 림포택틴(Lymphotactin)/ATAC/SCM, MCP-1/MCAF, MCP-3, MCP-4, MDC/STCP-1/ABCD-1, MIP-1. 쿼드러츄어(quadrature), MIP-1. 쿼드러츄어., MIP-2. 쿼드러츄어./GRO. 쿼드러츄어., MIP-3. 쿼드러츄어./엑소더스/LARC, MIP-3/엑소더스-3/ELC, MIP-4/PARC/DC-CK1, PF-4, RANTES, SDF1, TARC, 또는 TECK)으로부터 선택될 수 있다.

[0180] 본 발명의 한 실시양태에서, 관심 단백질은 멀티-서브유닛 단백질 또는 폴리펩티드일 수 있다. 발현될 수 있는 멀티서브유닛 단백질에는 호모머성(homomeric) 및 혜테로머성(heteromeric) 단백질이 포함된다. 멀티서브유닛 단백질은 동일하거나 상이할 수 있는 2개 이상의 서브유닛을 포함할 수 있다. 예를 들어, 단백질은 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개 또는 그 이상의 서브유닛을 포함하는 호모머성 단백질일 수 있다. 또한 단백질은 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개 또는 그 이상의 서브유닛을 포함하는 혜테로머성 단백질일 수 있다. 예시적인 멀티서브유닛 단백질에는 수용체 (이온 채널 수용체 포함); 세포외 매트릭스 단백질 (콘드로이틴 포함); 콜라겐; 면역조정인자 (MHC 단백질, 전체 사슬 항체, 및 항체 단편 포함); 효소 (RNA 중합효소, 및 DNA 중합효소 포함); 및 막 단백질이 포함된다.

[0181] 또다른 실시양태에서, 관심 단백질은 혈액 단백질일 수 있다. 이러한 실시양태에서 발현되는 혈액 단백질에는 캐리어 단백질, 예컨대 알부민 (인간 및 소 알부민 포함), 트랜스페린, 재조합 트랜스페린 반-분자, 햅토글로빈, 피브리노겐 및 기타 응고 인자, 보체 성분, 면역글로불린, 효소 억제제, 안지오텐신 및 브라디키닌과 같은 물질의 전구체, 인슐린, 엔도텔린, 및 글로불린 (알파, 베타 및 감마-글로불린 포함), 및 포유동물의 혈액에서 주로 발견되는 기타 유형의 단백질, 폴리펩티드 및 이들의 단편이 포함되지만, 이에 한정되지는 않는다. 인간 혈청 알부민 ([Lawn, L. M., et al. (1981) Nucleic Acids Research, 9:6103-6114]) 및 인간 혈청 트랜스페린 ([Yang, F. et al. (1984) Proc. Natl. Acad. Sci. USA 81:2752- 2756])에 대한 아미노산 서열을 포함하여, 수많은 혈액 단백질에 대한 아미노산 서열이 보고되어 있다 ([S. S. Baldwin (1993) Comp. Biochem. Physiol. 106b:203-218] 참조).

[0182] 또다른 실시양태에서, 관심 단백질은 재조합 효소 또는 보조인자일 수 있다. 이러한 실시양태에서 발현되는 효소 및 보조인자에는 알돌라제(alcohol dehydrogenase), 아민 산화효소, 아미노산 산화효소, 아스파르타제(aspartate), B12 의 존성 효소, 카르복시펩티다제(carboxypeptidase), 카르복시에스테라제(carboxyesterase), 카르복실라제(carboxylyase), 케모트립신(chemotrypsin), CoA 요구 효소, 시아노히드린(cyanohydrin) 신태타제(synthetase), 시스탈티온(cystathione) 신태타제, 테카르복실라제(decarboxylase), 탈수소효소, 알코올 탈수소효소, 디히드라타제(dehydratase), 디아포라제(diaphorase), 디옥시게나제(dioxygenase), 에노에이트 환원효소, 에폭시드 히드라제(hydrase), 푸머라제(fumerase), 갈락토스 산화효소, 글루코스 이성화효소, 글루코스 산화효소, 글리코실트랜스페라제(glycosyltransferase), 메틸트랜스페라제(methyltransferase), 니트릴 히드라제, 뉴클레오시드 인산화효소, 산화환원효소, 옥시니틸라제(oxynitilase), 펩티다제(peptidase), 글리코실트랜스페라제, 과산화효소, 및 치료적으로 활성인 폴리펩티드에 용합된 효소; 유로키나제(urokinase), 렙틸라제(reptilase), 스트렙토키나제(streptokinase); 카탈라제(catalase), 과산화물 디스뮤타제; DNase, 아미노산

히드롤라제(hydrolase) (예를 들어, 아스파라기나제(asparaginase), 아미도히드롤라제(amidohydrolase)); 카르복시펩티다제; 프로테아제, 트립신, 웨신, 키모트립신, 파파인, 브로멜라인, 콜라제(collagenase); 뉴라미니다제(neuraminidase); 락타제(lactase), 말타제(maltase), 수크라제(sucrase), 및 아라비노푸라노시다제(arabinofuranosidases)가 포함되지만, 이에 한정되지는 않는다.

[0183] 또 다른 실시양태에서, 관심 단백질은 단일 사슬, Fab 단편 및/또는 전체 사슬 항체 또는 이의 단편 또는 일부분 일 수 있다. 단일 사슬 항체는 안정적으로 폴딩된 단일 폴리펩티드 사슬 상의 항체의 항원-결합 영역을 포함할 수 있다. Fab 단편은 특정 항체의 조각일 수 있다. Fab 단편은 항원 결합 부위를 함유할 수 있다. Fab 단편은 경쇄 및 중쇄 단편인 2개의 사슬을 함유할 수 있다. 이러한 단편들은 링커 또는 디술피드 결합을 통해 연결될 수 있다.

[0184] 관심 단백질 또는 폴리펩티드에 대한 코딩 서열은 입수 가능한 경우 표적 폴리펩티드에 대한 천연 코딩 서열일 수 있지만, 더욱 바람직하게는, 예를 들어, 슈도모나스 플루오레센스와 같은 슈도모나스 종 또는 기타 적절한 생물의 코돈 사용 경향을 반영하도록 유전자를 합성함으로써, 선택된 숙주 세포에서 사용하기 위해 선택, 개선, 또는 최적화된 코딩 서열일 것이다. 결과적인 유전자(들)은 하나 이상의 벡터 내에 구축되거나, 또는 벡터 내로 삽입될 것이고, 그 후 벡터가 발현 숙주 세포 내로 형질전환될 것이다. "발현가능한 형태"로 제공되는 것으로 언급되는 핵산 또는 폴리뉴클레오티드는 선택된 발현 숙주 세포에 의해 발현될 수 있는 하나 이상의 유전자를 함유하는 핵산 또는 폴리뉴클레오티드를 의미한다.

[0185] 특정 실시양태에서, 관심 단백질은 천연 단백질, 예컨대 천연 포유류 또는 인간 단백질이거나, 또는 이에 대해 실질적으로 상동성이다. 이러한 실시양태에서, 단백질은 콘카타머(concatamer) 형태로 발견되지 않고, 분비 신호, 및 정제 및/또는 인식을 위한 임의의 태그에만 연결된다.

[0186] 또 다른 실시양태에서, 관심 단백질은 약 20°C 내지 약 42°C의 온도에서 활성인 단백질이다. 한 실시양태에서, 단백질은 생리학적 온도에서 활성이 있고, 고온 또는 극한 온도, 예컨대 65°C를 초과하는 온도로 가열되는 경우 불활성화된다.

[0187] 한 실시양태에서, 관심 단백질은 약 20°C 내지 약 42°C의 온도에서 활성이 있고/이거나 고온 또는 극한 온도, 예컨대 65°C를 초과하는 온도로 가열되는 경우 불활성화되는 단백질이고; 천연 단백질, 예컨대 천연 포유류 또는 인간 단백질이거나 이에 대해 상동성이며, 콘카타머 형태로 핵산으로부터 발현되지 않고; 프로모터는 슈도모나스 플루오레센스에서의 천연 프로모터가 아니고 또 다른 생물, 예컨대 대장균으로부터 유래된다.

[0188] 또 다른 실시양태에서, 단백질은, 생산되었을 때, 추가적인 표적화 서열, 예를 들어, 단백질을 세포외 배지로 표적화하는 서열을 또한 포함한다. 한 실시양태에서, 추가적인 표적화 서열은 단백질의 카르복시-말단에 작동 가능하게 연결된다. 또 다른 실시양태에서, 단백질은 오토틴스포터, 2-파트너 분비 시스템, 메인 터미널 브랜치 시스템 또는 펌브리알 어셔 포린에 대한 분비 신호를 포함한다.

[0189] 하기의 실시예는 설명을 위해 제공되고, 제한을 위해 제공되는 것이 아니다.

실시예

실시예 1. *dsbC* 리더 서열의 확인

I. 재료 및 방법

A. *pDOW2258* 발현 플라스미드의 구축

표준 재조합 DNA 기술이 *DsbC* 리더 웨პ티드-Skp 융합 단백질의 발현에 사용되는 플라스미드 *pDOW2258*의 구축에 사용되었다 (도 1).

[0194] Herculase Master Mix (Stratagene, #600610-51), 프라이머 RC-322 (5'-AATTACTAGTAGGAGGTACATTGCGCTT-3', 서열 25) 및 RC-323 (5'-TATACTCGAGTTATTTAACCTGTTCAAGTA-3', 서열 26), 및 주형 플라스미드 *pDOW3001* (SOE PCR에 의해 생성된 클로닝된 *dsbC* 리더-skp 코딩 서열 융합물을 이미 함유함)를 사용하여 PCR 증폭 반응을 수행하여, 제조업자의 프로토콜을 사용하여 521 bp *dsbC-skp* 코딩 서열을 증폭시켰다. PCR 단편을 QIAQUICK® 젤 추출 키트 (Qiagen, #28704)를 사용하여 정제하고, *SpeI* 및 *XhoI* 제한 뉴클레아제 (New England Biolabs, R0 133 및 R0 146)로 소화시킨 후, 발현 플라스미드 *pDOW1169* (*SpeI* 및 *XhoI*로 이미 소화됨)에 T4 DNA 리가제 (New England Biolabs, M0202)를 사용하여 제조업자의 프로토콜에 따라 결찰시켰다. 결찰 반응물을 전기천공에 의해 슈도모나스 플루오레센스 DC454 (*Isc*::*IacI*^{Q1} *ΔpyrF*) 내로 형질전환시키고, SOC+Soy 배지에서 회수하고,

선별 배지 (M9 글루코스 한천) 상에 플레이팅하였다. 콜로니들을 플라스미드 DNA의 제한 소화에 의해 분석하였다 (Qiagen, 카탈로그# 27106). 삽입물을 함유하는 10개의 클론을 서열분석하여, 에러가 없는 *dsbC-skp* 코딩 서열의 존재를 확인하였다. 서열이 확인된 단리물로부터의 플라스미드를 pDOW2258로 칭하였다.

[0195] B. 진탕 플라스크에서의 성장 및 발현 분석

[0196] pDOW2258을 함유하는 슈도모나스 플루오레센스 균주 DC454 (*Isc::lacI^{Q1} ΔpyrF*) 단리물을 표준 Dow 1 ℓ-규모 진탕-플라스크 발현 프로토콜에 의해 분석하였다. 간략하게, 1% 글루코스 및 미량 원소가 보충된 M9 배지에서 성장된 종자 배양물을 사용하여, 탄소 공급원으로서의 5% 글리세롤이 있는 규정 무기염 배지 200 mL에 접종하였다. 초기의 24시간 성장 단계에 이어서, 0.3 mM 이소프로필-β-D-1-티오갈락토피라노시드 (IPTG)로 *Ptac* 프로모터를 통한 발현을 유도하였다.

[0197] 유도 시점 (I0), 유도 24시간 후 (I24), 및 유도 48시간 후 (I48)에 배양물을 샘플링하였다. 600 nm에서의 광학 밀도 (OD_{600})에 의해 세포 밀도를 측정하였다. 세포 밀도를 $OD_{600} = 20$ 으로 조정하고, 100 μl 의 분취량을 $14,000 \times \text{rpm}$ 에서 5분 동안 원심분리하고, 상등액을 제거하였다.

[0198] EASYLYSE™ (Epicentre Technologies)을 사용하여 진탕 플라스크로부터의 가용성 및 불용성 분획을 생성시켰다. 세포 펠렛을 용해 완충제에 재현탁시켜 1:4로 희석하고, 진탕하면서 실온에서 30분 동안 인큐베이션하였다. 용해물을 $14,000 \text{ rpm}$ 에서 20분 동안 원심분리하고 (4°C), 상등액을 제거하였다. 상등액을 가용성 분획으로서 보관하였다. 샘플을 β -메르캅토에탄올을 함유하는 $2 \times$ Laemmli 샘플 완충제 (BioRad 카탈로그# 161-0737)와 1:1 혼합하고, 5분 동안 끓인 후, 20 μl 를 Bio-Rad Criterion 12% Bis-Tris 젤 (BioRad 카탈로그# 45-0112 룩트# cx090505C2) 상에 로딩하고, 1× MES 완충제 (카탈로그# 161-0788 룩트# 210001188)에서 전기영동하였다. 젤을 SIMPLYBLUE™ SafeStain (Invitrogen 카탈로그# LC6060)으로 제조업자의 프로토콜에 따라 염색하고, Alpha Innotech Imaging 시스템을 사용하여 영상화하였다.

[0199] C. N-말단 서열 분석

[0200] SDS-PAGE에 의해 분리된 가용성 및 불용성 분획을 10 mM CAPS (2.21 g/ℓ), pH 11 (NaOH로 조정), 및 10% 메탄올을 이동 완충제로 사용하여 40V에서 1.5시간 동안 서열분석 등급의 PVDF 막 (Bio-Rad, 카탈로그# 162-0236)으로 이동시켰다. 염색 용액 (0.2% 쿠마시 브릴리언트 블루(Commassie Brilliant Blue) R-250, 40% 메탄올, 10% 아세트산)에서 10초 동안 블록을 염색한 후, 즉각적으로 3회 탈색시켰다 (각각 10초). 관심 단백질 랜드를 블록으로부터 절단하고, Applied Biosystems (Foster City, CA)로부터의 PROCISE® 단백질 서열분석기 (모델 494) 상에서 수행된 에드만(Edman) 분해를 사용하여 서열분석하였다.

[0201] II. 결과

[0202] 가용성 분획 및 불용성 분획 양쪽 모두에서 유도 후 24시 (I24) 및 48시 (I48) 양쪽 모두에서 Skp에 대한 예상 분자량 (~17kDa)에서의 단백질의 상당한 축적이 SDS-PAGE 분석으로 확인되었다 (도 2).

[0203] I24의 Skp 단백질에 대한 예상 크기의 유도된 가용성 랜드에서 *DsbC-Skp*의 프로세싱된 형태에 대한 예상 단백질 서열의 첫번째 5개의 아미노산 (ADKIA, 서열 27)이 생산되었음이 N-말단 서열분석으로 확인되었다. N-말단 분석은 I24의 불용성 분획 내에 축적된 2개의 랜드에서 프로세싱된 형태 및 프로세싱되지 않은 형태의 *DsbC-Skp* 양쪽 모두가 생산되었음을 또한 나타냈다. 더 높은 분자량의 랜드에서는 프로세싱되지 않은 형태의 *DsbC-Skp*에 대한 예상 단백질 서열의 첫번째 10개의 아미노산 (MRLTQIIIAAA, 서열 28)이 생산된 한편, 더 낮은 분자량의 랜드에서는 프로세싱된 형태의 *DsbC-Skp* 단백질에 대한 예상 단백질 서열의 첫번째 10개의 아미노산 (ADKIAIVNMG, 서열 29)이 생산되었다. 도 3A 및 3B 참조.

[0204] 실시예 2. pbp* 분비 신호의 확인

[0205] I. 재료 및 방법

[0206] A. 균주

[0207] *lacI^{Q1}* 프로모터 돌연변이 ([Calos et al. 1980])를 혼입시키는 프라이머를 사용하는 pCN51/*lacI* ([Schneider et al. 2005b])로부터의 대장균 *lacI* 유전자의 PCR 증폭, 및 대립유전자 교환에 의해 유전자를 MB101 Δ pyrF ([Schneider, Jenings et al. 2005b])의 *Isc* (레반 수크라제(levan sucrase)) 좌위 내로 재조합시키는 것에 의해 DC206 (Δ pyrF, *Isc::lacI^{Q1}*)을 구축하였다.

[0208] B. 슈도모나스 플루오레센스 신호 서열을 스크리닝하기 위한 트랜스포솜(transposome)의 구축

kanR 유전자 (카나마이신에 대한 저항성을 코딩함) 및 출발 코돈 및 N-말단 신호 서열이 없는 *phoA* 수용체 유전자를 트랜스포솜 벡터 pMOD-2<MCS> (Epicentre Technologies) 내의 벡터가 코딩하는 트랜스포사제 결합 부위 ("모자이크 말단")들 사이에 삽입함으로써 트랜스포솜 벡터를 조작하였다. 1.6 kB *kanR* 유전자를 *Xho*I으로의 제한 소화에 의해 pUC4-KIXX (Pharmacia)로부터 정제한 후, *Sal*I로 소화된 pMOD2<MCS> 내로 결찰시켜, pDOW1245를 형성시켰다. 신호 서열이 없는 *phoA* 유전자를 프라이머에 의해 *Bam*HI 및 *Xba*I 부위를 부가하면서 대장균 K12 (ATCC)로부터 PCR로 증폭시켰다. 제한 소화 후, 유전자를 *Bam*HI 및 *Xba*I으로 소화된 pDOW1245 내로 결찰시켜, pDOW1208을 제조하였다. pDOW1208의 *Psh*A1로의 제한 소화, 및 Ultrafree DA (Amicon)를 사용한 3.3 kb 모자이크-말단-플랭킹 단편의 젤 정제에 의해 선형 트랜스포솜이 제조되었다. MicroBioSpin 6 컬럼 (Biorad)을 통과시킨 후, 30 ng을 4 유닛의 트랜스포사제 (Epicentre)와 혼합하고, 분취량을 슈도모나스 플루오레센스 MB101 내로 전기천공시켰다.

[0210] C. 개선된 *pbp* 신호 서열의 확인

PhoA 활성에 대해 분석함으로써 주변 세포질에서의 프로인슐린-*phoA*의 촉적을 측정할 수 있고 개선된 촉적을 선택할 수 있도록, *pbp*-프로인슐린-*phoA* 발현 플라스미드를 디자인하여, *pbp*-프로인슐린 단백질을 성숙형 PhoA 효소에 융합시켰다. 분비 리더 및 프로인슐린에 대한 코딩 서열 및 프로인슐린 및 성숙형 형태의 PhoA (즉, 천연 분비 리더가 없는 형태)에 대한 코딩 서열을 중첩시키는 프라이머를 사용하여, SOE PCR ([Horton et al. 1990])에 의해 pINS-008 내의 *pbp* 신호 서열, 인간 프로인슐린, 및 *phoA* 간의 융합물을 구축하였다. 융합물을 *Spe*I 및 *Xho*I으로 제한 소화되고 새우 알칼리성 포스파타제로 처리된 pDOW1169 ([Schneider et al. 2005a]; [Schneider, Jenings et al. 2005b]) 내에서 *tac* 프로모터의 제어 하에 클로닝하였고, 결찰시키고, DC206 내로 전기천공시켜, pINS-008을 형성시켰다. 프로인슐린 유전자 주형을 슈도모나스 플루오레센스에서의 발현에 대해 코돈-최적화시켜 합성하였다 (DNA 2.0). 대장균 MG1655 게놈 DNA로부터 *phoA* 유전자를 증폭시켰다. *pbp*-프로인슐린-*phoA* 유전자의 발현을 유도하기 위한 IPTG와 함께 알칼리성 포스파타제 활성의 비색측정 지표인 BCIP를 함유하는 한천 플레이트 상에서 콜로니들을 스크리닝하였다. BCIP 가수분해를 나타낸 콜로니들 중에서, 1개가 다른 것들보다 훨씬 더 크게 성장하였다. 이러한 단리물은 분비 웨티드를 코딩하는 영역에 단일 C → T 돌연변이가 있는 것으로 발견되었고, 이는 아미노산 20에서 알라닌 → 발린으로의 변화를 야기하였다 (A20V, 서열 2; 표 6 참조).

[0212] 표준 진탕 플라스크 프로토콜에 의해 2개의 균주의 발현을 평가하였다. IPTG 유도인자의 첨가 직후 양쪽 모두의 성장이 안정화되었다. 돌연변이체 pINS-008-3 균주에서의 알칼리성 포스파타제 활성이 3-4배 더 높았고 (도 6), (가용성) 단백질의 촉적이 더 높았다 (도 7).

표 6

슈도모나스 플루오레센스에서 확인된 Sec 분비 신호

관리되는 단백질 기능	약자	예상 신호 서열 (SignalP- HMM)	서열 번호	DNA 서열	서열 번호
pbp (신호 서열 돌연변이체)	pbp*	MKLKRLMAAMTFV AAGVATVNAVA	2	atgaaactgaaacgttgatggcg gcaatgcctttgtcgctgcgtggcg ttggacccgtcaacgcgtggcc	1
포린 E1 전구체	PO	MKKSTLAVAVTGLA IAQQAGA	31	atgaaagaatccacccgtggcttg gctgtaaacgttgggcgcaatcgcc ccaccaacggccgt	30
외막 포린 F	OP	MKLKNTLGLAIGSLI AATSFVGvla	33	atgaaacgtgaaaacacccgtgg cttggccatgggttcctttatggcg ctacttttcggccgtctggca	32
원형질마 주위 공간 포스페이트 결합 단백질 (pbp)	PB	MKLKRLMAAMTFV AAGVATANAVA	35	atgaaactgaaacgttgatggcg gcaatgactttgtcgctgcgtggcg ttggacccgtcaacgcgtggcc	34
아주린	AZ	MFAKLVAVSLTLA SGQLLA	37	atgttgccaaacgttgctgttc cctgtcgactctggcgagggcc atgttgtgt	36
회귀 지질단백질 B 전구체	L	MIKRNLVMGLAVL LSA	39	atgtatcaaaccgcaatctgtcggtta tggccctggccgtcggttgcgcgt ct	38
라이신-아르기닌-오르니린-결합 단백질	LAO	MQNYKKFLAAAV SMAFSATAMA	41	atgcagaactataaaaaattccctc tggccggggccgtctcgatggcg ttcagccgcacggccatggca	40
철(III) 결합 단백질	IB	MIRDNRKTSLLRG LTLLTLLSLLSPAA HS	43	atgtatcgatcaacccgactcaa gacatcccttgcggccgtgac ccttacccctactcgacttgcacct gtctcgccgcggccatct	42

[0213]

[0214]

D. 계놈 서열확인

[0215]

DNA Easy 키트 (Invitrogen)에 의해 계놈 DNA를 정제하고, 10 μ g을 2 \times ABI PRISM BigDye Terminators v3.0 Cycle Kit (Applied Biosystems)를 사용하여 트랜스포손 특이적 프라이머로 서열확인하기 위한 주형으로 사용하였다. 반응물을 정제하고, ABI PRISM 3100 Genetic Analyzer (Applied Biosystems)에 제조업자의 지침에 따라 로딩하였다.

[0216]

E. 신호 서열 코딩 영역의 클로닝

[0217]

SPScan 소프트웨어에 의해 또는 [De et al. 1995]에서와 같이 신호 서열을 결정하였다. 이러한 실험의 결과가 공동 계류 중인 미국 특허 출원 번호 20060008877 (2004년 11월 22일 출원)에 개시되어 있다. 외막 포린 F (oprF) 포스페이트 결합 단백질 (pbp), 포린E1 (porE), 아주린, 지질단백질 B 및 철 결합 단백질 분비 리더를 중합효소 연쇄 반응 (PCR)에 의해 증폭시켰다. 생성된 PCR 생성물을 pCRII Blunt TOPO 벡터 내로 클로닝하고, 대장균 Top 10 세포 (Invitrogen) 내로 제조업자의 프로토콜에 따라 형질전환시켰다. 생성된 형질전환체를 M13 전방향 및 M13 역방향 프라이머로의 서열확인에 의해 정확한 삽입물에 대해 스크리닝하였다. 양성 클론을 하기와 같이 명명하였다: oprF (pDOW1112), pbp (pDOW1113), 포린E1 (pDOW1183), 아주린 (pDOW1180), 지질단백질 B (pDOW1182), 철 결합 단백질 (pDOW1181).

[0218]

F. 슈도모나스 플루오레센스에서의 분비를 위한 *gal2 scFv* 클론의 구축

[0219]

pDOW1112 또는 pDOW1113을 주형으로 사용하여 +2 위치에서 *Gal2* 코딩 서열에 융합되도록 OprF 및 pbp 신호 서열을 증폭시켰다. pGal2 ([Martineau et al. 1998])를 주형으로 사용하여 *gal2* 코딩 서열을 증폭시켰다. 837bp SOE-PCR 생성물을 PCR BLUNT II TOPO 벡터 내로 클로닝하고, 서열을 확인하였다. 표준 기술을 사용하여 ([Sambrook et al. 2001]), 신호 서열에 융합된 scFv 유전자를 TOPO 벡터로부터 *Xba*I 및 *Sai*I 제한 효소로 절제하고, pMYC1803의 *Spe*I 및 *Xho*I 부위 내로 클로닝하여, pDOW1122 (*oprF:gal2*) 및 pDOW1123 (*pbp:gal2*)를 생산하였다. 생성된 플라스미드를 30 μ g/ml 테트라사이클린 및 50 μ g/ml 카나마이신이 보충된 LB 한천 상에서 선별된 DC191 내로 형질전환시켰다.

[0220]

pDOW1183으로부터의 *porE* 신호 서열을 pDOW1123으로부터 증폭된 *gal2*에 SOE-PCR에 의해 융합시키고, PCR 생성물을 젤 추출에 의해 정제하였다. 생성된 PCR 생성물을 PCRII Blunt TOPO 내로 클로닝하고, 이어서 대장균 Top 10 세포 내로 제조업자의 사용설명서 (Invitrogen)에 따라 형질전환시켰다. 생성된 클론을 서열분석하고, 양성 클론 (pDOW1185)을 서브클로닝으로 선별하였다. pDOW1185를 *Spe*I 및 *Sai*I로 제한 소화시키고, *porE-gal2* 단편을 젤 정제하였다. 정제된 단편을 *Spe*I-*Sai*I로 소화된 pDOW1169에 T4 DNA 리가제 (NEB)를 사용하여 결찰시켰

다. 결찰 믹스를 전기수용성(electrocompetent) DC454 내로 형질전환시키고, M9 1% 글루코스 한천 플레이트 상에서 선별하였다. 형질전환체를 *SpeI* 및 *SaI*을 사용한 플라스미드 DNA의 제한 소화에 의해 스크리닝하였다. 양성 클론을 단리하고, pDOW1186으로 저장하였다.

[0221] 아주린, 철 결합 단백질 및 *lspB*의 신호 서열을 각각 클론 pDOW1180, pDOW1181 및 pDOW1182로부터 증폭시켰다. *galZ* 유전자를 각각의 분비 리더에 융합되도록 적합한 프라이머를 사용하여 pDOW1123으로부터 증폭시키고, 생성된 PCR 생성물들을 단리하고, 상기 기술된 바와 같이 SOE-PCR에 의해 융합시켰다. SOE-PCR 생성물을 pCR-BLUNT II TOPO 내로 클로닝하고, 생성된 클론을 서열분석하고, 각각의 융합물에 대한 양성 클론을 상기 기술된 바와 같이 pDOW1169 내로 서브클로닝하였다.

G. C-말단 히스티딘 태그가 있는 슈도모나스 플루오레센스 분비 베타의 구축

[0223] pbp 분비 리더, C-말단 His 태그가 있는 MCS, 및 *rrnT1T2* 전사 종결인자가 있는 삽입물을 함유하는 클론을 DNA 2.0 (pJ5:G03478)에 의해 합성하였다. 450bp 분비 카셋트를 *SpeI* 및 *NdeI*로의 제한 소화에 의해 단리하고, 젤 정제하였다. 단편을 동일한 효소로 소화된 pDOW1219 (pMYC1803 ([Shao et al. 2006])으로부터 유래됨)에 결찰시켰다. 결찰 생성물을 화학적으로 수용성인 대장균 JM109 내로 형질전환시켰다. 플라스미드 DNA를 제조하고, 베타 특이적 프라이머를 사용하여 PCR에 의해 삽입물에 대해 스크리닝하였다. 생성된 플라스미드의 서열을 확인하고, pDOW3718로 명명하였다. 그후, 전기수용성 슈도모나스 플루오레센스 DC454를 플라스미드로 형질전환시키고, 250 μ g/ml 우라실 및 30 μ g/ml 테트라사이클린이 보충된 LB 한천 상에서 선별하였다.

[0224] 인간 단백질을 코딩하는 오픈 리딩 프레임을 인간 ORFeome 선집 (Invitrogen)으로부터의 주형을 사용하여 증폭시켰다. PCR 생성물을 *NheI* 및 *XbaI*으로 제한 소화시킨 후, *NheI-XbaI*으로 소화된 pDOW3718에 결찰시켰다. 이어서, 결찰 생성물을 전기수용성 슈도모나스 플루오레센스 DC454 내로 형질전환시키고, 250 μ g/ml 우라실 및 30 μ g/ml 테트라사이클린이 보충된 LB 한천 상에서 형질전환체를 선별하였다. 양성 클론을 서열분석하여, 삽입물 서열을 확인하였다.

H. 대장균 분비 클론의 구축

[0226] 5' 프라이머 상의 *NcoI* 부위 및 3' 프라이머 상의 *XbaI* 부위가 있도록 프라이머들을 디자인하는 것을 제외하고는 상기와 같이 인간 ORF를 증폭시켰다. PCR 생성물을 *NcoI* 및 *XbaI* (NEB)로 제한 소화시킨 후, Qiaquick 추출 키트 (Qiagen)를 사용하여 정제하였다. 소화된 생성물을 *NcoI-XbaI*로 소화된 pET22b (Novagen)에 T4 DNA 리가제 (NEB)를 사용하여 결찰시키고, 결찰 생성물을 화학적으로 수용성인 대장균 Top 10 세포 내로 형질전환시켰다. 형질전환체를 LB 한천 엠피실린 플레이트 (Teknova)에서 선별하였다. 플라스미드 DNA를 제조하고 (Qiagen), 양성 클론을 서열분석하여, 삽입물 서열을 확인하였다. 이어서, 각각에 대한 확인된 클로닝된 플라스미드 1개를 발현 분석을 위해 BL21(DE3) (Invitrogen) 내로 형질전환시켰다.

I. DNA 서열분석

[0228] 서열분석 반응물 (Big dye 버전 3.1 (Applied Biosystems))을 G-50 (Sigma)을 사용하여 정제하고, ABI3100 서열분석기 상에 로딩하였다.

J. 고치리량 (HTP) 발현 분석

[0230] 슈도모나스 플루오레센스 균주를 표준 Dow HTP 발현 프로토콜을 사용하여 분석하였다. 간략하게, 1% 글루코스 및 미량 원소가 보충된 M9 배지에서 성장된 종자 배양물을 사용하여, 2.0 ml의 깊은 96웰 플레이트 내의 탄소 공급원으로서의 5% 글리세롤이 있는 규정 무기염 배지 0.5 ml에 접종하였다. 30°C에서의 초기 배양 단계에 이어서, 0.3 mM 이소프로필- β -D-1-티오갈락토피라노사이드 (IPTG)로 *Ptac* 프로모터를 통한 발현을 유도하였다. 600 nm에서의 광학 밀도 (OD₆₀₀)에 의해 세포 밀도를 측정하였다.

K. SDS-PAGE 분석을 위한 HTP 샘플의 제조

[0232] 배양 샘플로부터의 가용성 및 불용성 분획을 EASY LYSE™ (Epicentre Technologies 카탈로그# RP03750)을 사용하여 생성시켰다. 175 μ l의 EASY LYSE™ 완충제를 첨가하고 부드럽게 진동시키면서 실온에서 30분 동안 인큐베이션함으로써 25 μ l의 전체 브로스 샘플을 용해시켰다. 용해물을 14,000 rpm에서 20분 동안 원심분리하고 (4°C), 상동액을 제거하였다. 상동액을 가용성 분획으로서 보관하였다. 그후, 웰렛 (불용성 분획)을 동일한 부피의 용해 완충제에 재현탁시키고, 아래위로 피펫팅(pipetting)함으로써 재현탁시켰다. 선별된 클론에 대해, 무세포 브로스 샘플을 해동시키고, 희석 없이 분석하였다.

[0233] *L. 관심 단백질 또는 폴리펩티드의 발현, 및 분비 분석*

1% 글루코스 (Teknova)가 보충되고, 미량 원소 용액이 보충된 $1\times$ M9에서 성장된 종자 배양물을 사용하여, 2% 접종원으로 50 mL의 Dow 규정 무기염 배지에 접종하고, 진탕시키면서 30°C에서 인큐베이션하였다. 약 24시간의 경과 발효 시간 (EFT: elapsed fermentation time)에 0.3 mM IPTG (이소프로필 β -D-티오갈락토파라노시드)로 세포를 유도하였다. 유도 시점 (I0), 및 유도 후 16시 (I16), 24시 (I24), 또는 40시 (I40)에 분석을 위해 샘플을 취하였다. 600 nm에서의 광학 밀도 (OD_{600})에 의해 세포 밀도를 측정하였다. 세포 밀도를 $OD_{600} = 20$ 으로 조정하고, 1 mL의 분취량을 $14,000\times g$ 에서 5분 동안 원심분리하였다. 상등액 (무세포 브로스)을 새로운 미세원심분리 투브 내로 피펫팅한 후, 세포 웰렛 및 무세포 브로스 샘플을 추후의 프로세싱을 위해 -80°C에서 동결시켰다.

[0235] *M. SDS-PAGE 분석*

진탕 플라스크로부터의 가용성 및 불용성 분획을 EASY LYSE™ 완충제 (Epicentre Technologies)를 사용하여 생성시켰다. 동결된 웰렛을 1 mL의 용해 완충제에 재현탁시켰다. 50 μ L를 추가적인 150 μ L의 EASY LYSE™ 완충제에 첨가하고, 진탕시키면서 실온에서 30분 동안 인큐베이션하였다. 용해물을 14,000 rpm에서 20분 동안 원심분리하고 (4°C), 상등액을 제거하였다. 상등액을 가용성 분획으로서 보관하였다. 그후, 웰렛을 동일한 부피 (200 μ L)의 용해 완충제에 재현탁시키고, 아래위로 피펫팅함으로써 재현탁시켰다; 이것을 불용성 분획으로서 보관하였다. 무세포 브로스 샘플을 해동시키고, 전체적으로 사용하였다.

[0237] *N. 웨스턴 분석*

제조되어 SDS-PAGE에 의해 분리된 가용성 및 불용성 분획을 1시간 동안 100V에서, 제조업자의 프로토콜에 따라 제조된 $1\times$ 이동 완충제 (Invitrogen)를 사용하여 니트로셀룰로스 (BioRad)로 이동시켰다. 이동 후, 블록을 POLY-HRP 희석제 (Research Diagnostics, Inc.)로 차단하고, 항-His 태그 항체 (Sigma 또는 US Biologicals)의 1:5,000 희석물로 프로빙하였다. 블록을 $1\times$ PBS-Tween으로 세정하고, 이어서 Immunopure Metal Enhanced DAB Substrate Kit (Pierce)를 사용하여 발색시켰다.

[0239] *O. 20 ℥ 발효*

효모 추출물 및 글리세롤이 보충된 화학적으로 규정된 배지 600 mL를 함유하는 진탕 플라스크에 동결된 배양 모액을 접종함으로써 발효기 배양을 위한 접종원이 각각 생성되었다. 32°C에서 진탕시키면서 16-24시간 동안 인큐베이션한 후, 진탕 플라스크 배양물을 높은 생물량을 지지하도록 디자인된 배지를 함유하는 20 ℥ 발효기로 무균적으로 이동시켰다. 살포 공기 흐름 및 진탕 속도를 조절함으로써 액체 배양물 내에서 양성 수준으로 용해 산소를 유지시켰다. 수성 암모니아의 첨가를 통해 pH를 원하는 설정점에서 제어하였다. 페드-벳치 고밀도 발효 프로세스는 약 24시의 초기 성장 단계 및 재조합 유전자 발현이 시작되도록 IPTG가 첨가된 유전자 발현 단계로 분리되었다. 그후, 발효의 발현 단계가 24시간 동안 진행되도록 하였다.

[0241] *P. N-말단 아미노산 서열 분석*

상기 SDS-PAGE 분석에서 기술된 바와 같이 샘플을 러닝시키고, Criterion Sequi-Blot PVDF 막 (Biorad)으로 이동시켰다. 막을 GelCode Blue 염색 시약 (Pierce)으로 염색하고, 이어서 50% 메탄올 1% 아세트산으로 탈색시키고, 10% 메탄올에 이어서 탈이온수로 세정한 후, 건조시켰다. 관심 밴드를 막으로부터 잘라내어 추출하고, Procise 단백질 서열분석 시스템, 모델 494 (Applied Biosystems (Foster City, CA)) 상에서의 8회 사이클의 에드만 분해에 적용하였다. [P. Edman, Acta Chem. Scand. 4, 283 (1950)]; 리뷰 문헌 [R. A. Laursen et al., Methods Biochem. Anal. 26, 201-284 (1980)].

[0243] *II. 결과*[0244] *A. 트랜스포손 돌연변이유발에 의한 천연 분비 신호 서열의 확인*

이종성 단백질을 주변 세포질 또는 브로스로 분비시킬 슈도모나스 플루오레센스 신호 서열을 확인하기 위해, 분비 리포터 유전자를 트랜스포솜 내로 클로닝하였다. 사용된 분비 리포터 유전자는 출발 코돈 또는 N-말단 신호 서열이 없는 대장균 알칼리성 포스파타제 유전자 (*phoA*)였다. *PhoA*는 활성 형태로의 이량체화를 허용하는 분자 간 디설피드 결합의 형성으로 인해 주변 세포질에서 활성이다 (그러나, 세포질에서는 그렇지 않음) ([Derman et al. 1991]). "게놈 스캐닝"으로 지칭되는 유사한 방법을 사용하여 대장균에서의 분비형 단백질을 확인하였다 ([Bailey et al. 2002]). 대장균 ([Manoil et al. 1985]) 및 또 다른 박테리아 ([Gicquel et al. 1996])에서

주변 세포질 단백질, 막 단백질 및 수출형 단백질 내의 분비 신호를 분석하는데 *phoA* 유전자가 또한 사용되었다. 전기천공 및 지표물 배지 상에의 플레이팅 후, 8개의 청색 콜로니를 단리하였다. 게놈 내의 트랜스포솜의 삽입 부위를 서열분석하고, 슈도모나스 플루오레센스 MB101의 사유 게놈 데이터베이스를 검색하는데 사용하였다. 활성 PhoA를 발현할 수 있는 것으로 확인된 8개의 유전자 융합물이 표 6에 제시된다.

B. 신호 서열-*gal2* 융합물의 클로닝

[0246] **0247]** 상기 확인된 분비형 단백질인 외막 포린 F (OP), 포스페이트 결합 단백질 porE (PB), 철 결합 단백질 (IB), 아주린 (AZ), 지질단백질 B (L) 및 라이신-오르니틴-아르기닌 결합 단백질 (LOA)의 신호 서열을 SignalP 프로그램 ([J.D. Bendtsen 2004])을 사용하여 예측하였다. OP, PE 및 AZ에 대한 신호 서열은 기존에 또다른 시스템에서 확인되었다 ([Arvidsson, 1989 #25]; [De, 1995 #24]; [Yamano, 1993 #23]). 또다른 연구에서 확인된 추가적인 분비 리더인 *ppbA20V* ([Schneider et al. 2006])의 활성이 또한 병행 분석되었다. 이러한 연구에서, 신호 웹티드의 절단에 이어지는 *Gal2*의 N-말단의 4개의 아미노산이 AQVQ이도록 재료 및 방법에 기술된 바와 같은 중첩 확장에 의한 스플라이싱 PCR (SOE-PCR)을 사용하여 6개의 천연 슈도모나스 플루오레센스 신호 서열 및 슈도모나스 플루오레센스 포스페이트 결합 단백질 신호 서열의 돌연변이체의 코딩 영역 (표 6 참조)이 *gal2* scFv 유전자에 각각 융합되었다. LAO 신호 서열을 증폭시키려는 반복된 시도가 실패하였고, 이러한 신호 서열은 추가적인 분석에서 탈락되었다. 유전자 융합물을 슈도모나스 플루오레센스 발현 벡터 pDOW1169 내로 클로닝하고, DC454 숙주 균주 (Δ *pyrF*, *lsc*::*lacI*^{Q1}) 내로 형질전환시켰다. 이어서 생성된 균주를 *Gal2* scFv 발현 및 분비 리더의 적절한 프로세싱에 대해 평가하였다.

C. 분비형 *Gal2* scFv의 발현

[0248] **0249]** 진탕 플라스크 규모에서, 생산 배지 내로의 하위배양 후 성장하지 못한 L-*gal2* scFv를 제외하고는, *gal2* scFv에 대한 PB, OP, PO, AZ, IB, 및 L의 융합물은 예상 OD₆₀₀을 달성하였다 (데이터는 제시되지 않음). PB, OP, PO, AZ 및 IB 신호 서열이 *Gal2* scFv 융합물로부터 절단되었음이 웨스턴 블로트 분석으로 확인되었다. 그러나, 웨스턴 분석은 프로세싱되지 않은 PB-*Gal2* 및 OP-*Gal2*의 존재를 나타냈다. AZ 및 IB 융합물로부터 발현된 일부 가용성 *Gal2* scFv가 무세포 브로스에서 발견되었고, 이는 가용성 단백질이 발현되어 주변 세포질로부터 누출되었음을 가리킨다. 아미노 말단 서열 분석을 수행하여, 신호 서열의 절단을 확인하였다. 아주린 (pDOW1191) 융합물로부터 발현된 불용성 *Gal2* 단백질은 프로세싱된 분비 신호 및 프로세싱되지 않은 분비 신호가 있는 단백질의 혼합물을 나타냈다. 그러나, 신호 서열은 IB-*Gal2* 융합물로부터 완전하게 프로세싱된 것으로 관찰되었다.

[0250] 7개의 리더 각각에 융합된 *Gal2* scFv의 발현을 표준 발효 조건을 사용하여 20ℓ 발효 규모에서 평가하였다. 모든 균주들이 예상대로 성장하여, 18-24시에 유도 OD₆₀₀ (약 180 유닛)에 도달하였다. *lipB*-*Gal2* 균주는 다른 균주들보다 약간 더 느리게 성장하였다. 소규모 발효에서 *lipB*-*Gal2* 균주가 진탕 플라스크 배지의 접종 후 성장하지 않았기 때문에, 이는 전적으로 뜻밖이지는 않았다. *Gal2* scFv의 발현 및 프로세싱을 SDS-PAGE 및 웨스턴 블로트에 의해 평가하였다. SDS-PAGE 분석은 OP 또는 PB 분비 신호에 융합되었을 때 높은 수준의 *Gal2*가 발현되었음을 나타냈다. 그러나, OP-*Gal2* 융합 단백질의 일부 (약 50%)만이 신호 서열이 절단되면서 주변 세포질로 분비되는 것으로 나타났다. 소규모에서 관찰된 바와 같이, 가용성 단백질이 웨스턴 블로트에 의해 검출되었지만, *Gal2*는 불용성 분획에서 우세하게 발현되었다. 배양 상등액에서 소량의 단백질이 또한 검출되었고, 이는 주변 세포질로부터의 누출을 가리킨다 (도 7). *lipB* 및 아주린 리더가 예상대로 프로세싱되어 N-말단 아미노산 서열 AQVQL (서열 44)이 생성되었음이 N-말단 서열 분석으로 확인되었다. 유사하게, 웨스턴 분석에 의해 *PorE* 분비 리더가 프로세싱된 것으로 나타났고, N-말단 분석에 의해 확인되었다. 불용성 *PorE*-*Gal2* 발현 수준은 불용성 *lipB*-*Gal2* 및 아주린-*Gal2*의 수준보다 약간 더 낮았다. *lipB*-*Gal2*는 *PorE*-*Gal2*의 수준과 유사한 수준의 프로세싱된 *Gal2*의 발현을 나타냈다. 가장 많은 양의 단백질은 *ppb*-*Gal2* 및 *ppbA20V*-*Gal2*를 발현하는 균주로부터 관찰되었다. *ppbA20V*-*Gal2* 균주로부터 발현된 *Gal2*의 양은 *ppb*-*Gal2* 균주에 의해 생산된 것보다도 더 높은 것으로 나타났다 (도 6). 가용성의 프로세싱된 *Gal2*가 웨스턴 분석에 의해 검출되었고, 프로세싱되지 않은 불용성 단백질 및 프로세싱된 불용성 단백질의 혼합물도 그러하였다 (도 7). 불용성 단백질의 N-말단 서열 분석으로 프로세싱되지 않은 *Gal2* 및 정확하게 프로세싱된 *Gal2*의 혼합물이 확인되었다.

실시예 3. bce 리더 서열의 확인

I. 재료 및 방법

[0251] **0252]** **0253]** BceL는 바실루스 코아굴란스 CMC 104017로부터의 히드롤라제에 대한 유전자를 함유하는 DNA 삽입물의 일부분에 의해 코딩되는 것으로 확인된 분비 리더이다. 이러한 바실루스 코아굴란스 균주는 다양한 시판 배양 선집에서

NCIMB 8041, ATCC 10545 및 DSMZ 2311로 또한 공지되어 있고, NRS784로서 기원된다. NRS 784는 포자 형성 박테리아의 NR Smith 선집 ([Smith et al Aerobic spore forming bacteria US. Dep. Agr. Monogra. 16:1-148 (1952)])으로부터의 것이다. NCIMB에 의해 인용된 이러한 균주에 대한 또 다른 근원적인 참조문헌은 [Cambell, L.L. and Sniff E. E. (1959. J.Bacteriol. 78:267 An investigation of Folic acid requirements of *Bacillus coagulans*)]이다.

[0254] 서열 및 바이오인포매틱스 분석

히드롤라제 효소를 잠재적으로 코딩하는 코딩 서열을 국소화하도록 바실루스 코아글란스 CMC 104017로부터의 4,127 bp의 DNA 삽입물을 서열분석 및 분석하였다. CDS1으로 지정된 1개의 1,314 bp의 코딩 서열이 5' 말단에서 *lac* 프로모터 뒤에서 확인되었다. CDS1에 대한 DNA 및 예상 단백질 서열이 각각 서열 45 및 46에 기재된다. 예상 단백질 서열의 BLASTP 분석을 기초로 아마도 CDS1이 히드롤라제를 코딩하는 것으로 결정되었다. CDS1 서열은 슈도모나스 팔루스트리스(*Rhodopseudomonas palustris*) HaA2로부터의 베타-락타마제에 대한 상동성 (E-값: 2e-36)을 나타냈다. CDS1의 SignalP 3.0 은닉형 마르포크 모델 분석 ([Bendtsen JD, Nielson G, von Heijne G, Brunak S: Improved prediction of signal peptides: Signal 3.0. J. Mol. Biol 2004, 340:783])에서 그람-양성 박테리아의 생물 클래스에 대한 신호 서열의 존재가 예측되었고, 이때 서열 46의 잔기 33/34 사이에 신호펩티다제 절단 부위가 있었다.

[0255] 단백질 발현 플라스미드의 구축

표준 클로닝 방법이 발현 플라스미드의 구축에 사용되었다 ([Sambrook J, Russell D: Molecular Cloning a Laboratory Manual, third edn. Cold Spring Harbor: Cold Spring Harbor Press; 2001]). DNA 서열 융합을 SOE-PCR 방법 ([Horton, R. M., Z. Cai, S. N. Ho and L. R. Pease (1990). "Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction." BioTechniques 8(5): 528-30, 532, 534-5])를 사용하여 수행하였다. Phusion DNA 중합효소 (New England Biolabs 카탈로그# F531S)를 모든 PCR 반응에 사용하였다.

바실루스 코아글란스 CMC104017로부터의 에스테라제 단백질을 발현하고 슈도모나스 플루오레센스의 세포질 또는 주변 세포질 내로 국소화하도록 플라스미드를 디자인하였다. 최종 PCR 생성물을 *SpeI* 및 *XhoI* 제한 엔도뉴클레아제 (New England Biolabs 카탈로그# R0133 및 R0146)로 소화시킨 후, T4 DNA 리가제 (New England Biolabs 카탈로그# M0202S)를 사용하여 *SpeI* 및 *XhoI*으로 또한 소화된 발현 벡터 pDOW1169 내로 결찰시켜, 세포질 CMC 104641 CDS-1 발현 벡터 p484-001 및 천연 *Bce* 리더 CMC104641 CDS-1 발현 벡터 p484-002를 생산하였다. 그후 결찰 반응 혼합물을 슈도모나스 플루오레센스 균주 DC454 (Δ pyrF, *lsc*::*lacI*⁰¹) 내로 전기천공에 의해 형질전환시키고, SOC+Soy 배지 (Teknova 카탈로그# 2S2699)에서 회수하고, 선별 배지 (M9 글루코스 한천, Teknova 카탈로그# 2M1200) 상에 플레이팅하였다. 콜로니들을 미니프렙(miniprep) 플라스미드 DNA (Qiagen, 카탈로그# 27106)의 제한 소화에 의해 분석하였다. 각각의 형질전환으로부터의 10개의 콜로니들을 서열분석하여, 정확한 삽입물을 확인하였다.

[0256] 발현 분석

각각의 클론을 보유하는 슈도모나스 플루오레센스 균주 DC454를 200 μ l의 규정 무기염 배지 + 탄소 공급원으로서의 5% 글리세롤 ("Dow 배지")를 함유하는 진탕 플라스크에서 시험하였다. 초기 성장 단계에 이어서, *tac* 프로모터를 통한 발현을 0.3 mM 이소프로필- β -D-1-티오갈락토피라노시드 (IPTG)로 유도하였다. 유도 시점 (I0), 및 유도 24시간 후 (I24)에 배양물을 샘플링하였다. 600 nm에서의 광학 밀도 (OD_{600})에 의해 세포 밀도를 측정하였다. 진탕 플라스크 번호매김을 나타내는 표가 표 7에서 제시된다.

표 7

숙주 균주	플라스미드 번호 (리더-유전자)	플라스크 번호		
		EP484-001	EP484-002	EP484-003
DC454	P484-001 (세포질 484)	EP484-001	EP484-002	EP484-003
DC454	P484-002 (천연 리더 484)	EP484-004	EP484-005	EP484-006

[0261]

각각의 샘플링 시점에, 샘플의 세포 밀도를 $OD_{600} = 20$ 으로 조정하고, 1 μ l 분취량을 14000×g에서 5분 동안 원

심분리하였다. 상등액 (무세포 브로스)를 세로운 미세원심분리 투브 내로 피펫팅한 후, 세포 펠렛 및 무세포 브로스 샘플을 -20°C에서 동결시켰다.

[0263] 세포 용해 및 SDS-PAGE 분석

[0264] 진탕 플라스크 샘플로부터의 가용성 및 불용성 분획을 Easy Lyse (Epicentre Technologies)를 사용하여 생성시켰다. 동결된 펠렛을 용해 완충제에 재현탁시켜 1:4로 희석하고, 진탕하면서 실온에서 30분 동안 인큐베이션하였다. 용해물을 14,000 rpm에서 20분 동안 원심분리하고 (4°C), 상등액을 제거하였다. 상등액을 가용성 분획으로서 보관하였다. 그후, 펠렛 (불용성 분획)을 동일한 부피의 용해 완충제에 재현탁시키고, 아래위로 피펫팅 함으로써 재현탁시켰다. 무세포 브로스 샘플을 해동시키고, 전체적으로 사용하였다. 샘플을 β -메르캅토에탄올을 함유하는 2× Laemmli 샘플 완충제 (BioRad 카탈로그# 161-0737)와 1:1 혼합하고, 5분 동안 끓인 후, 20 μ l를 Bio-Rad Criterion 10% Criterion XT 젤 (BioRad 카탈로그# 45-0112) 상에 로딩하고, 권장된 1× MOPS 완충제 (카탈로그# 161-0788 룻트# 210001188)에서 전기영동하였다. 젤을 SIMPLYBLUE™ SafeStain (Invitrogen 카탈로그# LC6060)으로 제조업자의 프로토콜에 따라 염색하고, Alpha Innotech Imaging 시스템을 사용하여 영상화하였다. 관심 젤 빙드의 단백질 양을 동일한 젤에 로딩된 BSA 단백질 표준물에 대한 비교에 의해 추정하였다.

II. 결과

[0265] 총 6개의 진탕 플라스크 (균주 당 3개의 플라스크)를 사용하여 히드롤라제 발현을 평가하였다. 디자인된 주변 세포질 균주 및 세포질 균주의 성장은 슈도모나스 플루오레센스 균주에 대한 정상적인 성장과 일치하여, 유도 24시간 후 약 15의 OD₆₀₀에 도달하였다. SDS-PAGE 분석을 수행하여 유도 시점 및 유도 24시간 후의 히드롤라제 (CDS1 단백질) 발현을 평가하였다. 가용성, 불용성, 및 무세포 브로스 분획을 SDS-PAGE에 의해 분석하였다. 세포질 CDS-1 균주 (p484-001)에 대해, 세포질 히드롤라제에 대한 예상 크기 (44.1 kDa)의 단백질이 0.1 mg/ml의 추정 수율로 3개 모두의 단리물에서 I24 (IPTG 유도 24시간 후)에 거의 완전히 가용성 분획에 축적되었다. 도 8은 EP484-003으로서 평가된 세포질 균주에 대한 대표적인 결과를 나타낸다. 예상 크기의 무시할 수 있는 빙드가 불용성 분획에서 검출가능하였고, CDS1 단백질이 무세포 브로스에서 검출되지 않았다. 천연 Bce 리더-CDS1을 발현하는 주변 세포질 균주 (p484-002)에 대해, 천연 에스테라제에 대한 예상 크기의 단백질이 0.8 mg/ml의 추정 수율로 3개 모두의 단리물에서 I24에 거의 완전히 가용성 분획에 축적되었다. 도 8은 EP484-004로서 평가된 Bce 리더 용합물을 함유하는 주변 세포질 균주에 대한 대표적인 결과를 나타낸다. 사용된 젤 로딩이 예상된 47.6 kDa의 프로세싱되지 않은 크기와 44.1 kDa의 프로세싱된 크기를 분별하는 것을 어렵게 하였기 때문에, 발현된 천연 에스테라제가 완전히 프로세싱되었는지 여부는 불명확하였다. 세포질 발현 균주의 결과와 유사하게, 예상 크기의 무시할 수 있는 빙드가 불용성 분획에서 검출가능하였고, CDS1 단백질이 무세포 브로스에서 검출되지 않았다. 관심 Bce 리더의 번역된 서열이 서열 8에 기재된다.

실시예 4. 슈도모나스 플루오레센스 분비 리더의 확인 및 분석

[0266] MB214 게놈으로부터의 6,433개의 번역된 ORF를 신호 웹티드 예측 프로그램인 SignalP 2.0 ([Nielsen, H., et al. Protein Eng, 1997. 10(1): p. 1-6])으로 분석하였다. 1326개가 신호 웹티드를 함유하는 것으로 HMM 모델에 의해 예측되었다. 이러한 단백질들을 PsortB 2.0 ([Gardy, J.L., et al. Bioinformatics, 2005. 21(5): p. 617-23])으로 분석하였고, PsortB 최종 국소화가 세포질 또는 세포질막으로 확인된 모든 단백질을 제거하여, 891개가 남았다. 신호 웹티드를 함유할 SignalP HMM 확률이 0.79 미만인 82개의 단백질을 제거하여, 809개가 산출되었다. 0.79가 aprA (세포외 단백질로 공지된 RXF04304)를 제외하지 않은 최고값이었기 때문에 이러한 차단값을 선택하였다. SignalP Neural Network 알고리즘에 의해 예측된 바와 같이 신호 웹티드를 함유하는 이러한 809개의 번역된 ORF의 아미노 말단 서열 + 프로세싱된 단백질의 최초의 7개의 아미노산을 CLUSTALX 1.81 ([Thompson, J.D., et al. Nucleic Acids Res, 1997. 25(24): p. 4876-82])을 사용하여 정렬하였다.

[0267] Huber 등은 고도로 소수성인 신호 서열이 공동-번역에 의해 분비될 가능성이 높다는 것을 제안하였다 ([Huber, D., et al. J Bacteriol, 2005. 187(9): p. 2983-91]). 공동-번역에 의해 분비되는 단백질을 확인하기 위한 목적으로, Wertz-Scheraga (WS)의 아미노산 인덱스 ([Wertz, D. H. and H. A. Scheraga, Macromolecules, 1978. 11(1): p. 9-15]가 최상인 것으로 발견되었다. 이러한 연구를 위해, www.genome.jp/dbget-bin/www_bget?aaax1:WERD780101에서 월드와이드웹 상에서 AAindex로부터 이러한 인덱스를 수득하였다. Boyd에 의해 보고된 알고리즘 ([Boyd, D., C. Schierle, and J. Beckwith, Protein Sci, 1998. 7(1): p. 201-5])을 변형시켜, 소수성을 기초로 809개의 단백질의 등급을 매기는데 사용하였다. 알고리즘은 각각의 서열을 스캐닝하여, 윈도우 12 내에 WS 점수를 평균화하였다. 가장 소수성인 영역을 사용하여, 전체 단백질에 대한 WS 점수를

할당하였다. 이것으로 WS 접수가 [Huber et al.]에서 정의된 차단값인 0.69를 초과하는 142개의 신호 서열이 산출되었다. 이러한 더 작은 목록을 INCAPS (Indiana Centers for Applied Protein Sciences)에 의해 수행된 2D-LC 전체 프로테옴 실험으로부터의 데이터와 상호참조하였다. 이러한 실험들에서 다양한 성장 조건 하에 MB214 (슈도모나스 플루오레센스 MB101의 자손)에서 발현된 모든 단백질을 확인 및 정량하도록 시도하였다. 높은 최대 발현 수준으로 이러한 목록에 나타난 단백질은 고도로 발현될 것이다. 이러한 데이터에서, 1 또는 3의 우선순위 접수는 단백질의 확인에서의 높은 신뢰도를 가리킨다. 우선순위가 1 또는 3인 INCAPS 실험에서 확인된 142개의 목록으로부터의 단백질이 최대 발현 수준의 순서로 표 8에서 열거된다.

표 8

INCAPS 실험 동안 발견된 최대 발현 수준의 순서로 열거된 우선순위가 1 또는 3
(확인에서의 높은 신뢰도를 가리킴)인 142개의 목록으로부터의 7개의 특정 단백질

우선순위	단백질 번호	관리되는 기능	최대값
1	RXF05550.1	테트라트리코펩티드 반복 패밀리 단백질	377264.2
1	RXF08124.1	메틸-수용성 화학주성 단백질	134887.4
1	RXF07256.1	TolB 단백질	88429.16
3	RXF07256.1_a1	TolB 단백질	84020.51
3	RXF04046.2_a1	사이토크롬 c 산화효소, 모노헴(monoHEME) 서브유닛, 막-결합형 (ec 1.9.3.1)	79275.3
3	RXF03895.1_a1	asma	50164.08
3	RXF07256.1_pn	TolB 단백질	49215.09
1	RXF06792.1	보존된 가상 단백질	47485.35
3	RXF02291.1	톨루엔 내성 단백질 ttg2C	45703.08

[0270]

대장균 내의 공동-번역에 의해 분비되는 여러 단백질이 확인되었다. 이들 중 몇몇의 서열을 사용하여, 상동체에 대해 MB214 계놈을 검색하였다. 대장균 유전자는 DsbA, TorT, SfinC, FocC, CcmH, YraI, TolB, NikA, FlgI였다. BLASTP 알고리즘 ([Altschul, S.F., et al., J Mol Biol, 1990. 215(3): p. 403-10])을 사용하여, MB214 번역 ORF의 데이터베이스를 검색하였다. MB214 단백질을 대장균 대응물에 대해 나타낸 상동성 정도를 기초로 2개의 카테고리로 분류하였다. 높은 상동성의 단백질들은 $2e^{-84}$ 이상의 예상 접수로 매칭되었다. 낮은 상동성의 단백질들은 예상 접수가 $8e^{-17}$ 내지 $5e^{-32}$ 사이였다. 이러한 방법으로 11개의 특정한 잠재적인 상동체들이 산출되었고, 이들중 일부는 상기에서 수득된 7개의 표적과 중복되었다.

[0272]

18개의 특정 단백질의 조합된 목록을 SignalP를 사용하여 분석하였고, 1개의 가능한 신호 펩티다제 절단 부위가 있을 것으로 예측되는 9개의 최종 표적을 발현 연구용으로 선택하였다.

[0273]

분비 리더의 단리 및 서열 분석

[0274]

확인된 슈도모나스 플루오레센스 분비 리더를 DC454 (슈도모나스 플루오레센스 MB101의 자손) 계놈 DNA로부터 증폭시키고, DNA 서열 확인을 위해 pCRBLUNTII-TOPO (Invitrogen) 내로 클로닝하였다. 각각의 단리된 슈도모나스 플루오레센스 분비 리더의 DNA 및 추정 아미노산 서열이 표 9에서 제시된다.

표 9

슈도모나스 플루오레센스 분비 리더 서열

리더	DNA 서열 번호	아미노산 서열 번호
CupA2	9	10
CupB2	11	12
CupC2	13	14
TolB	49	50
NikA	15	16
FlgI	17	18
ORF5550	19	20
Ttg2C	21	22
ORF8124	23	24

[0275]

Ga12 scFv 및 대장균 티오레독신에 대한 분비 리더의 융합, 및 발현 분석

[0277]

중첩 확장에 의한 스플라이싱 PCR ([Horton R.M. et al. 1990 Biotechniques 8:528])을 사용하여 각각의 분비 리더 (표 9)를 인-프레임(in-frame)으로 Ga12 scFv 서열 ([Martineau, P. et al. 1998 J. Mol Bio. 280: 117]) 및/또는 대장균 티오레독신 (TrxA) 서열 (서열 46)에 융합시켰다. 생성된 절편을 정제하고, 이어서 제2 라운드의 PCR에 대한 주형으로 사용하여, NikA 분비 리더 코딩 서열을 *trxA* 서열에 융합시켰다. 그후, 융합물을 슈도모나스 플루오레센스 발현 벡터 pDOW1169 내로 *tac* 프로모터의 제어 하에 클로닝하였다. 각각의 구축물을 슈도모나스 플루오레센스 DC454 내로 형질전환시키고, 고처리량 포맷으로 발현을 평가하였다. 배양물을 0.5 mL의 배양 부피로 2 mL의 깊은 웰 플레이트에서 5% 글리세롤이 보충된 규정 무기염 배지에서 성장시켰다. 24시간의 성장 기간 후, 재조합 단백질을 0.3 mM IPTG로 유도하고, 24시간 동안 발현되도록 하였다. 초음파처리에 의해 배양물을 분획화하고, 단백질 발현 및 분비 리더 프로세싱을 SDS-CGE 및 웨스턴 블롯에 의해 평가하였다 (도 9). Bce 리더를 제외하고, 각각의 테스트된 리더는 Ga12 scFv 단백질 서열로부터 부분적으로 또는 완전하게 프로세싱된 것으로 발견되었다. 각각은 세포질 Ga12 scFv (없음)를 코딩하는 발현 군주와 비교하여 Ga12 scFv의 발현을 또한 매우 개선시켰고, 이는 세포하 위치선정을 지시하는 것에 더하여, 이러한 분비 리더들이 전체적인 발현을 또한 개선시킬 수 있다는 것을 가리킨다. 예상대로, Ga12 scFv의 다양한 수준의 발현 및 용해도가 또한 관찰되었다. CupA2, CupC2, NikA, FlgI 및 ORF 5550에 융합되었을 때 약간의 가용성 Ga12가 생산되었음을 웨스턴 분석으로 확인되었다 (도 10). Ga12에 융합된 TolB 리더의 발현이 다른 리더에서 관찰된 것보다 더 낮았지만, 웨스턴 분석은 발현된 모든 단백질이 가용성이었음을 나타냈다. N-말단 분석은 TolB, CupA2, CupC2, FlgI, NikA 및 ORF5550 리더가 예상대로 Ga12 scFv로부터 절단되었음을 나타냈다 (데이터는 제시되지 않음).

[0278]

Ga12 scFv로부터 프로세싱되지는 않았지만, Bce 리더는 TrxA로부터 프로세싱된 것으로 발견되었다. 세포질에서 급속하게 폴딩되기 때문에 티오레독신은 공동-변역 분비 리더의 확인을 위한 모델 단백질로서 기술되었다 ([Huber et al. 2005 J. Bacteriol. 187:2983]). Bce 리더를 사용한 가용성 TrxA의 성공적인 분비는 이러한 리더가 공동-변역 방식으로 작용하여 주변 세포질 분비를 용이하게 한다는 것을 가리킬 수 있다.

[0279]

본 명세서에서 언급된 모든 간행물 및 특허 출원은 본 발명이 속하는 분야의 당업자의 수준을 나타낸다. 모든 간행물 및 특허 출원은 각각의 개별적인 간행물 또는 특허 출원이 거명에 의해 포함되는 것으로 명확하게, 그리고 개별적으로 지시되는 것과 동일한 정도로 거명에 의해 본원에 포함된다.

[0280]

이해를 명확하게 하기 위한 목적으로 설명 및 예에 의해 상기의 발명이 일부 상세하게 기술되었지만, 첨부된 청구항의 범주 내에서 어느 정도의 변화 및 변형이 실행될 수 있다는 것이 명백할 것이다.

도면의 간단한 설명

[0014]

도 1은 *dsbC SS-skp* 융합 단백질에 대한 발현 구축물을 도해한다.

[0015]

도 2는 약 17 kDa의 Skp 단백질의 발현을 나타낸다 (화살표). 2 및 3으로 표지된 밴드는 Skp 단백질과 일치하였다. 밴드 1은 DNA 결합 단백질 (3691)과 Skp 양쪽 모두를 갖는 것으로 보인다.

[0016]

도 3은 2B-2 (도 3A) 및 2B-4 (도 3B)로 표지된 샘플에 대한 가용성 (S) 및 불용성 (I) 분획에서의 0시간 및

24시간 후의 슈도모나스 플루오레센스에서의 *dsbC SS-skp*의 발현 후의 분석이다. 도 3A에서, 밴드 5, 7 및 9는 불용성 분획 내의 프로세싱되지 않은 *dsbC-skp* 단백질이었다. 밴드 6, 8, 및 10은 불용성 분획 내의 프로세싱된 *dsbC-skp*였다. 밴드 1 및 3은 가용성 분획 내의 프로세싱된 *dsbC-skp*였다. 밴드 2 및 4는 미지의 단백질이었다. 도 3B에서, 밴드 15, 17, 및 19는 불용성 분획 내의 프로세싱되지 않은 *dsbC-skp* 단백질이었다. 밴드 16, 18, 및 20은 불용성 분획 내의 프로세싱된 *dsbC-skp*였다. 밴드 11 및 13은 가용성 분획 내의 프로세싱된 *dsbC-skp*였다. 밴드 12 및 14는 미지의 단백질이었다.

[0017] 도 4는 DC694 (*dsbA-PA83*) 발현 후의 단백질 축적의 웨스턴 분석을 나타낸다. 제0시 및 제24시의 가용성 (S), 불용성 (I), 및 무세포 브로스 (broth) (B)의 축적을 웨스턴 분석에 의해 평가하였다.

[0018] 도 5는 EP468-002.2 (*dsbA*) 발현 후의 단백질 축적의 웨스턴 분석을 나타낸다. 유도 후 제0시 및 제24시의 가용성 (S) 및 불용성 (I) 단백질의 축적을 웨스턴 분석에 의해 평가하였다.

[0019] 도 6은 pINS-008-5 (야생형 *pbp*) 분비 신호에 비교된 pINS-008-3 (*pbp**) 돌연변이체의 알칼리성 포스파타제 활성을 설명한다. 세포 배양물을 1 OD₆₀₀ 단위로 조정한 후, 4-메틸음밸리페론 (MUP)을 첨가하고, 제10분에 형광 생성물 형성을 측정함으로써 *PhoA* 활성을 측정하였다. 음성 대조군은 MUP는 함유하지만, 세포는 함유하지 않는다.

[0020] 도 7은 pINS-008-3 (*pbp**) 및 pINS0008-5 (야생형 *pbp*)의 발현 후의 단백질 축적의 웨스턴 분석을 나타낸다. I0시, I16시, 및 I40시의 가용성 (Sol), 불용성 (Insol), 및 세포외 분획 (Bro)에서의 프로인슐린-*phoA*의 축적을 웨스턴 분석에 의해 평가하였다. 배양물의 분취량을 20 OD₆₀₀ 단위로 조정하고, SDS-PAGE에 의해 분리하고, 필터로 이동시켜서, 인슐린에 대한 항체 (닭 폴리클로날 항체, Abcam 카탈로그# ab 14042)로 시각화하였다.

[0021] 도 8은 EP484-003 및 EP484-004 분획의 SDS-PAGE 분석을 나타낸다. SDS-PAGE 분석의 대표적인 결과가 제시된다. 분자량 마커 (L)가 중앙에 제시된다. BSA 표준 (BSA Stds.)이 제시된다. 화살표는 유도된 밴드를 지시한다. 각각의 레인 아래에 분획 유형이 제시된다: 가용성 (Sol), 불용성 (Ins), 또는 무세포 브로스 (CFB). 각각의 레인 위에 유도 시 (I0) 또는 유도 24시간 후 (I24)의 샘플 시점이 제시된다. 균주 번호가 각각의 샘플 군 아래에 제시된다. EP484-004의 I24 가용성 분획에서의 대형 단백질 밴드는 Bce 리더 서열에 의해 촉진된 강화된 유전자 발현에 상응한다.

[0022] 도 9는 Ga12 scFv 발현의 SDS-PAGE 및 웨스턴 분석을 나타낸다. 가용성 (S) 및 불용성 (I) 분획을 분석하였다. 각각의 레인 쌍 위에 Ga12에 융합된 분비 리더가 제시된다. 각각의 SDS-PAGE 젤 (상부) 또는 웨스턴 블롯 (하부)의 좌측에 분자량 마커가 기술된다. 화살표들은 Ga1의 이동을 가리킨다.

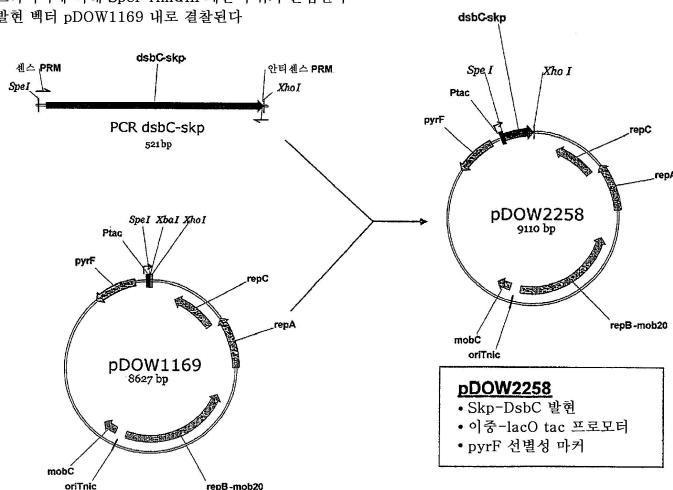
[0023] 도 10은 티오레독신 (TrxA) 발현의 SDS-PAGE 분석을 나타낸다. 가용성 분획들을 분석하였다. 각각의 레인 쌍 위에 TrxA에 융합된 분비 리더가 제시된다. SDS-PAGE 젤의 좌측에 분자량 마커가 기술된다. 화살표들은 프로세싱되지 않은 TrxA (위쪽 화살표) 및 프로세싱된 TrxA (아래쪽 화살표)의 이동을 가리킨다.

도면

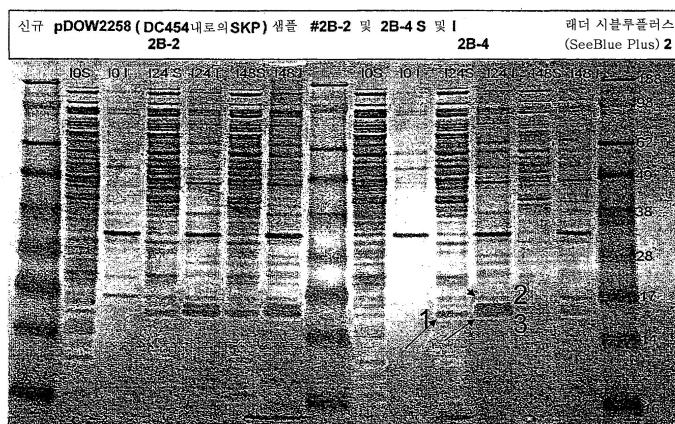
도면1

pDOW3001 주형으로부터의 PCR dsbC-skp

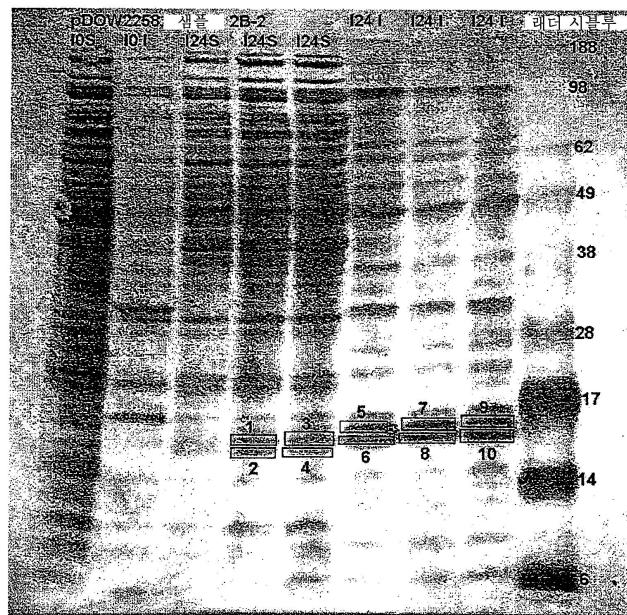
- 프라이머에 의해 SpeI-HindIII 제한 부위가 혼입된다
- 발현 벡터 pDOW1169 내로 결합된다



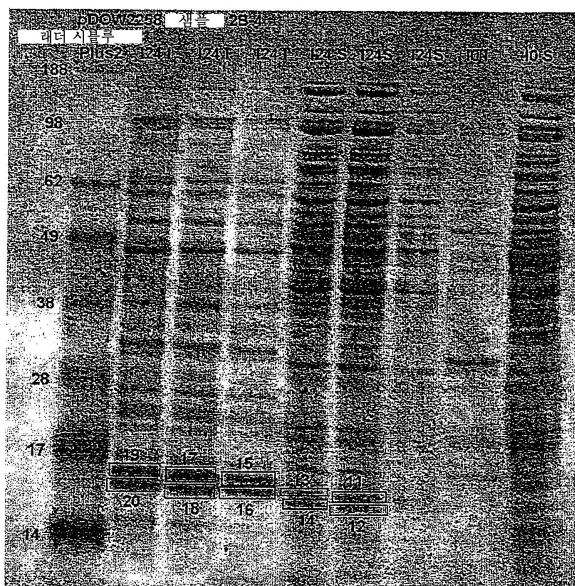
도면2



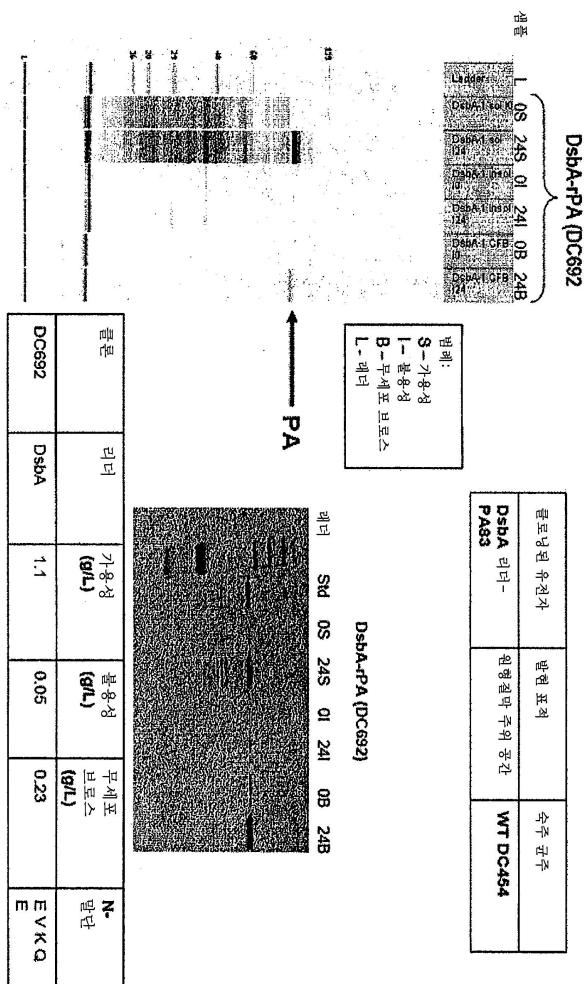
도면3



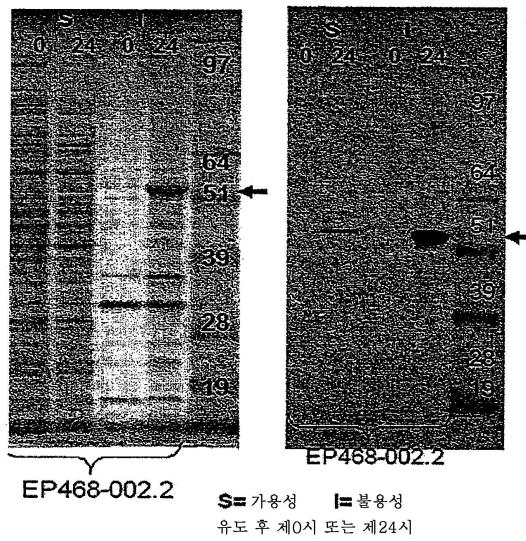
도면3B



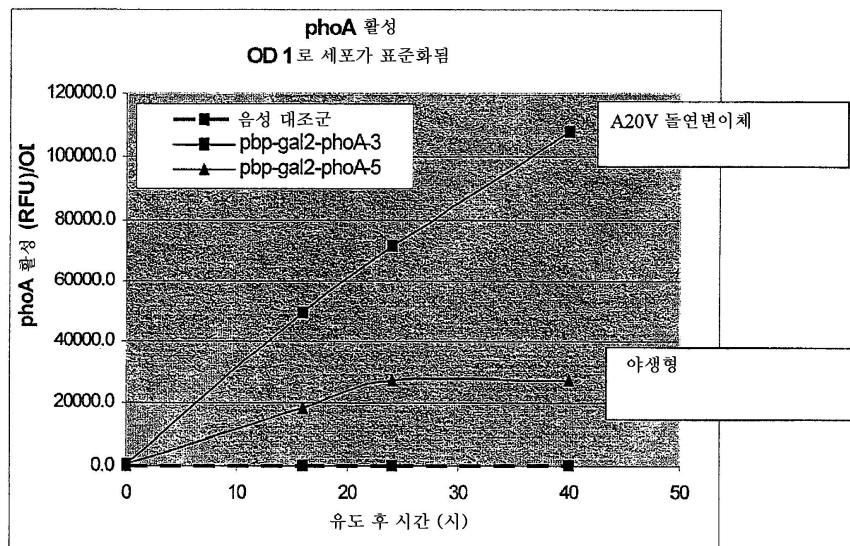
도면4



도면5

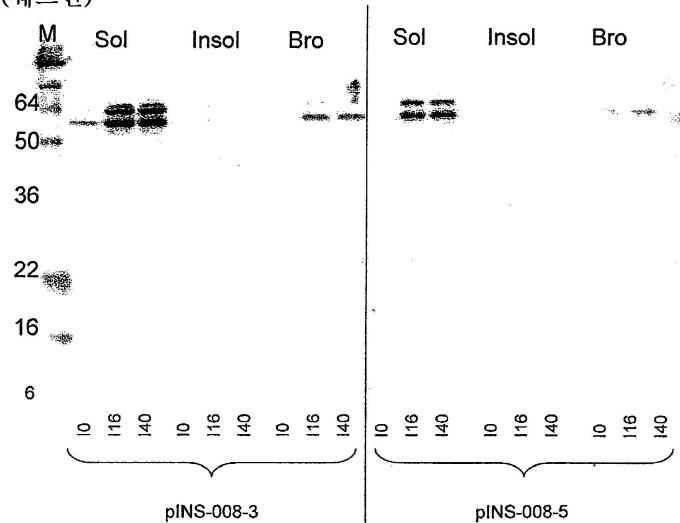


도면6

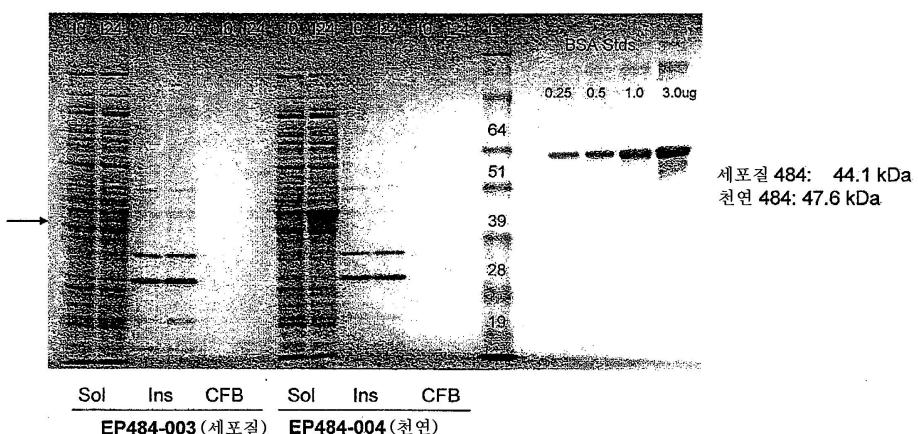


도면7

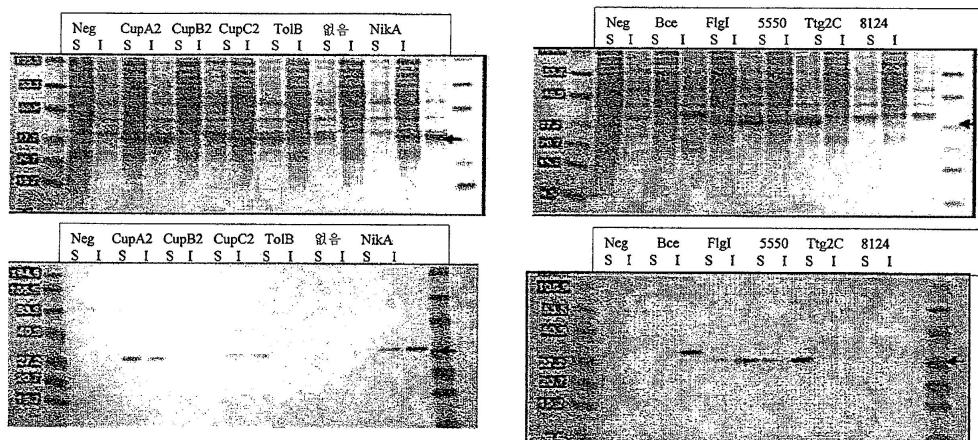
0.3mM IPTG로의 슈도모나스 플루오레센스에서의 **ppb-프로인슐린-phoA** 발현
(웨스턴)



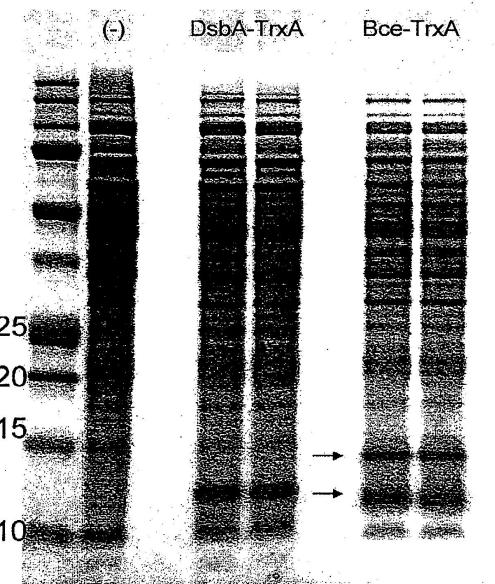
도면8



도면9



도면10



서열목록

SEQUENCE LISTING

<110> Coleman, Russell J.
Retallack, Diane M.
Schneider, Jane C.
Ramseier, Thomas M.

Hershberger, Charles D.
Lee, Stacey L.
Resnick, Sol M.

<120> BACTERIAL LEADER SEQUENCES FOR INCREASED
EXPRESSION

<130> 43292/339376

<160> 50

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> mutant phosphate binding protein leader sequence
(pbp*)

<221> CDS
<222> (1)...(72)

<400> 1
atg aaa ctg aaa cgt ttg atg gcg gca atg act ttt gtc gct gct ggc 48
Met Lys Leu Lys Arg Leu Met Ala Ala Met Thr Phe Val Ala Ala Gly
1 5 10 15
gtt gcg acc gtc aac gcg gtg gcc 72
Val Ala Thr Val Asn Ala Val Ala
20

<210> 2
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> mutant phosphate binding protein leader sequence
(pbp*)

<400> 2

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> 3
 <211> 66
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(66)

<400> 3
 atg cgt aat ctg atc ctc agc gcc gct ctc gtc act gcc agc ctc ttc 48
 Met Arg Asn Leu Ile Leu Ser Ala Ala Leu Val Thr Ala Ser Leu Phe
 1 5 10 15
 ggc atg acc gca caa gct 66
 Gly Met Thr Ala Gln Ala
 20

<210> 4
 <211> 22
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 4
 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> 5
 <211> 63
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(63)

<400> 5
 atg cgc ttg acc cag att att gcc gcc gca gcc att gcg ttg gtt tcc 48
 Met Arg Leu Thr Gln Ile Ile Ala Ala Ala Ile Ala Leu Val Ser

1	5	10	15
---	---	----	----

acc ttt gcg ctc gcc	63
Thr Phe Ala Leu Ala	
20	

<210> 6	
<211> 21	
<212> PRT	
<213> Pseudomonas fluorescens	

<400> 6	
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> 7	
<211> 99	
<212> DNA	
<213> Bacillus coagulans	

<220>	
<221> CDS	
<222> (1)...(99)	

<400> 7	
atg agc aca cga atc ccc cgc cga caa tgg ctg aaa ggc gcc tcg ggc	48
Met Ser Thr Arg Ile Pro Arg Arg Gln Trp Leu Lys Gly Ala Ser Gly	
1 5 10 15	
ctg ctg gcc gcc ggc agc ctg ggc cgg ttg gcc aac cgc gag ggc cgc	96
Leu Leu Ala Ala Ala Ser Leu Gly Arg Leu Ala Asn Arg Glu Ala Arg	
20 25 30	
gcc	99
Ala	

<210> 8	
<211> 33	
<212> PRT	
<213> Bacillus coagulans	

<400> 8	
Met Ser Thr Arg Ile Pro Arg Arg Gln Trp Leu Lys Gly Ala Ser Gly	

1	5	10	15
Leu	Leu	Ala	Ala
Ala	Ala	Ser	Leu
		Gly	Arg
		Leu	Ala
		Asn	Arg
		Glu	Ala
			Arg

20 25 30

Ala

<210> 9
 <211> 75
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(75)

<400> 9
 atg tcg tgc aca cgt gca ttc aaa cca ctg ctg ctg atc ggc ctg gcc 48
 Met Ser Cys Thr Arg Ala Phe Lys Pro Leu Leu Leu Ile Gly Leu Ala
 1 5 10 15

aca ctg atg tgt tcc cat gca ttc gct 75
 Thr Leu Met Cys Ser His Ala Phe Ala
 20 25

<210> 10
 <211> 25
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 10
 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> 11
 <211> 72
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(72)

<400> 11
 atg ctt ttt cgc aca tta ctg gcg agc ctt acc ttt gct gtc atc gcc 48

Met Leu Phe Arg Thr Leu Leu Ala Ser Leu Thr Phe Ala Val Ile Ala
 1 5 10 15

ggc tta ccg tcc acg gcc cac gcg 72
 Gly Leu Pro Ser Thr Ala His Ala
 20

<210> 12
 <211> 24
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 12
 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> 13
 <211> 69
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(69)

<400> 13
 atg ccg cct cgt tct atc gcc gca tgt ctg ggg ctg ctg ggc ttg ctc 48
 Met Pro Pro Arg Ser Ile Ala Ala Cys Leu Gly Leu Leu Gly Leu Leu
 1 5 10 15
 atg gct acc cag gcc gcc gcc 69
 Met Ala Thr Gln Ala Ala Ala
 20

<210> 14
 <211> 23
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 14
 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> 15
 <211> 63
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(63)

<400> 15
 atg cgc ctc gct gcc cta ccg cta ttg ctt gcc cct ctc ttt att gcg 48
 Met Arg Leu Ala Ala Leu Pro Leu Leu Ala Pro Leu Phe Ile Ala
 1 5 10 15
 ccg atg gcc gtt gcg 63
 Pro Met Ala Val Ala
 20

<210> 16
 <211> 21
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 16
 Met Arg Leu Ala Ala Leu Pro Leu Leu Ala Pro Leu Phe Ile Ala
 1 5 10 15
 Pro Met Ala Val Ala
 20

<210> 17
 <211> 63
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(63)

<400> 17
 atg aag ttc aaa cag ctg atg gcg atg gcg ctt ttg ttg gcc ttg agc 48
 Met Lys Phe Lys Gln Leu Met Ala Met Ala Leu Leu Leu Ala Leu Ser
 1 5 10 15
 gct gtg gcc cag gcc 63
 Ala Val Ala Gln Ala

<210> 18
 <211> 21
 <212> PRT
 <213> Pseudomonas fluorescens

<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> 63
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(63)

<400> 19
 atg aat aga tct tcc gcg ttg ctc ctc gct ttt gtc ttc ctc agc ggc 48
 Met Asn Arg Ser Ser Ala Leu Leu Leu Ala Phe Val Phe Leu Ser Gly
 1 5 10 15
 tgc cag gcc atg gcc 63
 Cys Gln Ala Met Ala
 20

<210> 20
 <211> 21
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 20
 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> 21
 <211> 99
 <212> DNA

<213> Pseudomonas fluorescens

<220>

<221> CDS

<222> (1)...(99)

<400> 21

atg	caa	aac	cgc	act	gtg	gaa	atc	ggt	gtc	ggc	ctt	ttc	ttg	ctg	gct	48
Met	Gln	Asn	Arg	Thr	Val	Glu	Ile	Gly	Val	Gly	Leu	Phe	Leu	Leu	Ala	
1															15	

ggc	atc	ctg	gct	tta	ctg	ttg	gcc	ctg	cga	gtc	agc	ggc	ctt	tcg	96
Gly	Ile	Leu	Ala	Leu	Leu	Leu	Ala	Leu	Leu	Leu	Ala	Leu	Arg	Val	Ser
20															30

gcc	99
Ala	

<210> 22

<211> 33

<212> PRT

<213> Pseudomonas fluorescens

<400> 22

Met	Gln	Asn	Arg	Thr	Val	Glu	Ile	Gly	Val	Gly	Leu	Phe	Leu	Leu	Ala
1															15
1															15

Gly	Ile	Leu	Ala	Leu	Leu	Leu	Ala	Leu	Leu	Leu	Ala	Leu	Arg	Val	Ser
20															30
20															30

Ala

<210> 23

<211> 117

<212> DNA

<213> Pseudomonas fluorescens

<220>

<221> CDS

<222> (1)...(117)

<400> 23

atg	tct	ctt	cgt	aat	atg	aat	atc	gcc	ccg	agg	gcc	ttc	ctc	ggc	ttc	48
Met	Ser	Leu	Arg	Asn	Met	Asn	Ile	Ala	Pro	Arg	Ala	Phe	Leu	Gly	Phe	
1															15	

gcg	ttt	att	ggc	gcc	ttg	atg	ttg	ttg	ctc	ggt	gtg	ttc	gcg	ctg	aac	96
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Ala Phe Ile Gly Ala Leu Met Leu Leu Leu Gly Val Phe Ala Leu Asn
 20 25 30

cag atg agc aaa att cgt gcg 117
 Gln Met Ser Lys Ile Arg Ala

35

<210> 24
 <211> 39
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 24
 Met Ser Leu Arg Asn Met Asn Ile Ala Pro Arg Ala Phe Leu Gly Phe
 1 5 10 15
 Ala Phe Ile Gly Ala Leu Met Leu Leu Leu Gly Val Phe Ala Leu Asn
 20 25 30
 Gln Met Ser Lys Ile Arg Ala
 35

<210> 25
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> oligonucleotide primer

<400> 25
 aattactagt aggaggtaca ttatgcgtt 30

<210> 26
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> oligonucleotide primer

<400> 26
 tataactcgag ttattnaacc tgtttcagta 30

<210> 27
 <211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> First 5 amino acids of the predicted protein
sequence for the processed form of dsbC-Skp

<400> 27

Ala Asp Lys Ile Ala

1 5

<210> 28

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> First 10 amino acids of the predicted protein
sequence for the unprocessed form of dsbC-Skp

<400> 28

Met Arg Leu Thr Gln Ile Ile Ala Ala Ala

1 5 10

<210> 29

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> First 10 amino acids of the predicted protein
sequence for the processed form of dsbC-Skp

<400> 29

Ala Asp Lys Ile Ala Ile Val Asn Met Gly

1 5 10

<210> 30

<211> 63

<212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(63)

<400> 30
 atg aag aag tcc acc ttg gct gtg gct gta acg ttg ggc gca atc gcc 48
 Met Lys Lys Ser Thr Leu Ala Val Ala Val Thr Leu Gly Ala Ile Ala
 1 5 10 15
 cag caa gca ggc gct 63
 Gln Gln Ala Gly Ala
 20

<210> 31
 <211> 21
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 31
 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> 32
 <211> 72
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(72)

<400> 32
 atg aaa ctg aaa aac acc ttg ggc ttg gcc att ggt tct ctt att gcc 48
 Met Lys Leu Lys Asn Thr Leu Gly Leu Ala Ile Gly Ser Leu Ile Ala
 1 5 10 15
 gct act tct ttc ggc gtt ctg gca 72
 Ala Thr Ser Phe Gly Val Leu Ala
 20

<210> 33

<211> 24
<212> PRT
<213> Pseudomonas fluorescens

<400> 33
Met Lys Leu Lys Asn Thr Leu Gly Leu Ala Ile Gly Ser Leu Ile Ala
1 5 10 15
Ala Thr Ser Phe Gly Val Leu Ala
20

<210> 34
<211> 72
<212> DNA
<213> Pseudomonas fluorescens

<220>
<221> CDS
<222> (1)...(72)

<400> 34
atg aaa ctg aaa cgt ttg atg gcg gca atg act ttt gtc gct gct ggc 48
Met Lys Leu Lys Arg Leu Met Ala Ala Met Thr Phe Val Ala Ala Gly
1 5 10 15
gtt gcg acc gcc aac gcg gtg gcc 72
Val Ala Thr Ala Asn Ala Val Ala
20

<210> 35
<211> 24
<212> PRT
<213> Pseudomonas fluorescens

<400> 35
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> 36
<211> 60
<212> DNA
<213> Pseudomonas fluorescens

<220>
<221> CDS

<222> (1)...(60)

<400> 36
 atg ttt gcc aaa ctc gtt gct tcc ctg ctg act ctg gcg agc ggc 48
 Met Phe Ala Lys Leu Val Ala Val Ser Leu Leu Thr Leu Ala Ser Gly
 1 5 10 15
 cag ttg ctt gct 60
 Gln Leu Leu Ala
 20

<210> 37
 <211> 20
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 37
 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> 38
 <211> 51
 <212> DNA
 <213> Pseudomonas fluorescens

<220>
 <221> CDS
 <222> (1)...(51)

<400> 38
 atg atc aaa cgc aat ctg ctg gtt atg ggc ctt gcc gtg ctg ttg agc 48
 Met Ile Lys Arg Asn Leu Leu Val Met Gly Leu Ala Val Leu Leu Ser
 1 5 10 15
 gct 51
 Ala

<210> 39
 <211> 17
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 39

Met	Ile	Lys	Arg	Asn	Leu	Leu	Val	Met	Gly	Leu	Ala	Val	Leu	Leu	Ser
1															
								10							15

Ala

<210> 40

<211> 69

<212> DNA

<213> Pseudomonas fluorescens

<220>

<221> CDS

<222> (1)...(69)

<400> 40

atg	cag	aac	tat	aaa	aaa	ttc	ctt	ctg	gcc	gct	gtc	tcg	atg	gct	48	
Met	Gln	Asn	Tyr	Lys	Lys	Phe	Leu	Leu	Ala	Ala	Ala	Ala	Val	Ser	Met	Ala
1															15	

ttc	agc	gcc	acg	gcc	atg	gca	69
Phe	Ser	Ala	Thr	Ala	Met	Ala	
							20

<210> 41

<211> 23

<212> PRT

<213> Pseudomonas fluorescens

<400> 41

Met	Gln	Asn	Tyr	Lys	Lys	Phe	Leu	Leu	Ala	Ala	Ala	Ala	Val	Ser	Met	Ala
1															15	
Phe	Ser	Ala	Thr	Ala	Met	Ala									20	

<210> 42

<211> 93

<212> DNA

<213> Pseudomonas fluorescens

<220>

<221> CDS

<222> (1)...(93)

<400> 42

atg	atc	cgt	gac	aac	cga	ctc	aag	aca	tcc	ctt	ctg	cgc	ggc	ctg	acc	48
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Met Ile Arg Asp Asn Arg Leu Lys Thr Ser Leu Leu Arg Gly Leu Thr
 1 5 10 15

ctc acc cta ctc agc ctg acc ctc tcg ccc gcg gcc cat tct 93
 Leu Thr Leu Leu Ser Leu Thr Leu Leu Ser Pro Ala Ala His Ser
 20 25 30

<210> 43
 <211> 31
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 43
 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 Ser
 20 25 30

<210> 44
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> N-terminal amino acid sequence of processed azurin
 and ibp

<400> 44
 Ala Gln Val Gln Leu
 1 5

<210> 45
 <211> 1317
 <212> DNA
 <213> Bacillus coagulans

<220>
 <221> CDS
 <222> (1)...(1317)

<400> 45
 atg agc aca cga atc ccc cgc cga caa tgg ctg aaa ggc gcc tcg ggc 48

Met Ser Thr Arg Ile Pro Arg Arg Gln Trp Leu Lys Gly Ala Ser Gly
 1 5 10 15
 ctg ctg gcc gcc gcg agc ctg ggc cgg ttg gcc aac cgc gag ggc ggc 96
 Leu Leu Ala Ala Ala Ser Leu Gly Arg Leu Ala Asn Arg Glu Ala Arg
 20 25 30
 gcc gcc gaa gcg agc gcc gcc gcg ctc gac act ggc tcg ctg ggc 144
 Ala Ala Glu Ala Ser Ala Ala Pro Leu Asp Thr Gly Ser Leu Gly
 35 40 45
 gcc tcg ccg cgc gcg acg ctc gac gcc tgc ctg caa aaa gcc gtc gac 192
 Ala Ser Pro Arg Ala Thr Leu Asp Ala Cys Leu Gln Lys Ala Val Asp
 50 55 60
 gac ggc acg ctc aag agc gtg gtg gcg atg gcc gcc acc gag cgc ggg 240
 Asp Gly Thr Leu Lys Ser Val Val Ala Met Ala Ala Thr Glu Arg Gly
 65 70 75 80
 ctc gcc tac cag ggc gcg cgc ggc ccg gcc aac gcg gcc ggc gag ccg 288
 Leu Ala Tyr Gln Gly Ala Arg Gly Pro Ala Asn Ala Ala Gly Glu Pro
 85 90 95
 atc ggc ccc gat acg gtg ttc tgg atg ctg tcg atg acc aag gcg atc 336
 Ile Gly Pro Asp Thr Val Phe Trp Met Leu Ser Met Thr Lys Ala Ile
 100 105 110
 acc gcc acc gcc tgc atg cag ctg atc gag cag ggc cgg ctc ggg ctc 384
 Thr Ala Thr Ala Cys Met Gln Leu Ile Glu Gln Gly Arg Leu Gly Leu
 115 120 125
 gac cag ccc gcc ggc gag atc ctg ccg caa ctg aag gcg ccg cag gtg 432
 Asp Gln Pro Ala Ala Glu Ile Leu Pro Gln Leu Lys Ala Pro Gln Val
 130 135 140
 ctg gag ggc ttc gac gcc ggc cag ccc agg ctg cgc ccg ggc cgc 480
 Leu Glu Gly Phe Asp Ala Ala Gly Gln Pro Arg Leu Arg Pro Ala Arg
 145 150 155 160
 cgc gcg atc acg gtg cgc cac ctc acg cat acc tcg ggc tat acc 528
 Arg Ala Ile Thr Val Arg His Leu Leu Thr His Thr Ser Gly Tyr Thr
 165 170 175
 tac agc atc tgg agc gag ggc ctg ggc cgc tac gaa cag gtc acg ggc 576
 Tyr Ser Ile Trp Ser Glu Ala Leu Gly Arg Tyr Glu Gln Val Thr Gly
 180 185 190
 atg ccc gac atc ggc tac tcg ctg aac ggc gcc ttc ggc gcc ccg ctc 624
 Met Pro Asp Ile Gly Tyr Ser Leu Asn Gly Ala Phe Ala Ala Pro Leu
 195 200 205
 gaa ttc gag ccc ggc gag cgc tgg caa tac ggc atc ggc atg gat tgg 672
 Glu Phe Glu Pro Gly Glu Arg Trp Gln Tyr Gly Ile Gly Met Asp Trp

210	215	220	
gtg ggc aag ctg gtg gag gcg gtg acc gac cag tcg ctg gaa gtg gcg Val Gly Lys Leu Val Glu Ala Val Thr Asp Gln Ser Leu Glu Val Ala			720
225	230	235	240
ttc cgc gag cgg atc ttc gcg ccg ctc ggc atg cac gat acg ggc ttc Phe Arg Glu Arg Ile Phe Ala Pro Leu Gly Met His Asp Thr Gly Phe			768
245	250	255	
ctg atc ggc agc gcg caa aag cgc cgc gtc gcc acg ctg cat cgg cgc Leu Ile Gly Ser Ala Gln Lys Arg Arg Val Ala Thr Leu His Arg Arg			816
260	265	270	
cag gcc gat ggc tcg ctg acg ccg gaa ccc ttc gag acc aac cag cgg Gln Ala Asp Gly Ser Leu Thr Pro Glu Pro Phe Glu Thr Asn Gln Arg			864
275	280	285	
ccc gag ttc ttc atg ggc ggc ggg ctg ttc agc acc ccg cgc gac Pro Glu Phe Phe Met Gly Gly Gly Leu Phe Ser Thr Pro Arg Asp			912
290	295	300	
tac ctc gcc ttc ctg cag atg ctg ctg aac ggc ggc gcc tgg cgc ggc Tyr Leu Ala Phe Leu Gln Met Leu Leu Asn Gly Gly Ala Trp Arg Gly			960
305	310	315	320
gag cgg ctg ctg cgg ccc gac acc gtc gcg agc atg ttc cgc aac cag Glu Arg Leu Leu Arg Pro Asp Thr Val Ala Ser Met Phe Arg Asn Gln			1008
325	330	335	
atc ggc gat ctt cag gtt cgc gaa atg aag acc gcc cag ccg gcc tgg Ile Gly Asp Leu Gln Val Arg Glu Met Lys Thr Ala Gln Pro Ala Trp			1056
340	345	350	
tcg aac agc ttc gac caa ttc ccc ggc gcg acg cac aag tgg ggg ctg Ser Asn Ser Phe Asp Gln Phe Pro Gly Ala Thr His Lys Trp Gly Leu			1104
355	360	365	
tcc ttc gat ctc aac agc gag ccg ggg ccg cac ggg cgc ggc ggc Ser Phe Asp Leu Asn Ser Glu Pro Gly Pro His Gly Arg Gly Ala Gly			1152
370	375	380	
tcg ggt agc tgg gcc ggc ctg ctg aac acc tac ttc tgg atc gat ccc Ser Gly Ser Trp Ala Gly Leu Leu Asn Thr Tyr Phe Trp Ile Asp Pro			1200
385	390	395	400
gcc aag cgc gtg acg ggg gcg ctg ttc acg cag atg ctg ccg ttc tac Ala Lys Arg Val Thr Gly Ala Leu Phe Thr Gln Met Leu Pro Phe Tyr			1248
405	410	415	
gac gcg cgc gtg gtc gat ctc tac ggg cgc ttc gag cgc ggg ctc tac Asp Ala Arg Val Val Asp Leu Tyr Gly Arg Phe Glu Arg Gly Leu Tyr			1296

420 425 430

gac ggg ctg ggc cgc gcc tga 1317
 Asp Gly Leu Gly Arg Ala *
 435

<210> 46
 <211> 438
 <212> PRT
 <213> Bacillus coagulans

<400> 46
 Met Ser Thr Arg Ile Pro Arg Arg Gln Trp Leu Lys Gly Ala Ser Gly
 1 5 10 15
 Leu Leu Ala Ala Ala Ser Leu Gly Arg Leu Ala Asn Arg Glu Ala Arg
 20 25 30
 Ala Ala Glu Ala Ser Ala Ala Pro Leu Asp Thr Gly Ser Leu Gly
 35 40 45
 Ala Ser Pro Arg Ala Thr Leu Asp Ala Cys Leu Gln Lys Ala Val Asp
 50 55 60

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

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

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

Leu Ile Gly Ser Ala Gln Lys Arg Arg Val Ala Thr Leu His Arg Arg
 260 265 270
 Gln Ala Asp Gly Ser Leu Thr Pro Glu Pro Phe Glu Thr Asn Gln Arg
 275 280 285
 Pro Glu Phe Phe Met Gly Gly Gly Leu Phe Ser Thr Pro Arg Asp
 290 295 300

Tyr Leu Ala Phe Leu Gln Met Leu Leu Asn Gly Gly Ala Trp Arg Gly
 305 310 315 320

Glu Arg Leu Leu Arg Pro Asp Thr Val Ala Ser Met Phe Arg Asn Gln
 325 330 335

Ile Gly Asp Leu Gln Val Arg Glu Met Lys Thr Ala Gln Pro Ala Trp
 340 345 350

Ser Asn Ser Phe Asp Gln Phe Pro Gly Ala Thr His Lys Trp Gly Leu
 355 360 365

Ser Phe Asp Leu Asn Ser Glu Pro Gly Pro His Gly Arg Gly Ala Gly
 370 375 380

Ser Gly Ser Trp Ala Gly Leu Leu Asn Thr Tyr Phe Trp Ile Asp Pro
 385 390 395 400

Ala Lys Arg Val Thr Gly Ala Leu Phe Thr Gln Met Leu Pro Phe Tyr
 405 410 415

Asp Ala Arg Val Val Asp Leu Tyr Gly Arg Phe Glu Arg Gly Leu Tyr
 420 425 430

Asp Gly Leu Gly Arg Ala
 435

<210> 47

<211> 324

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)...(324)

<400> 47

agc gat aaa att att cac ctg act gac gac agt ttt gac acg gat gta 48
 Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp Val
 1 5 10 15

ctc aaa gcg gac ggg gcg atc ctc gtc gat ttc tgg gca gag tgg tgc 96
 Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp Cys
 20 25 30

ggt ccg tgc aaa atg atc gcc ccg att ctg gat gaa atc gct gac gaa 144
 Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu

35 40 45

tat cag ggc aaa ctg acc gtt gca aaa ctg aac atc gat caa aac cct 192
 Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro
 50 55 60

ggc act gcg ccg aaa tat ggc atc cgt ggt atc ccg act ctg ctg ctg 240
 Gly Thr Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu Leu
 65 70 75 80

ttc aaa aac ggt gaa gtc gcg gca acc aaa gtc ggt gca ctg tct aaa 288

Phe Lys Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser Lys
85 90 95

ggt cag ttg aaa gag ttc ctc gac gct aac ctg gcg 324
Gly Gln Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
100 105

<210> 48
<211> 108
<212> PRT
<213> Escherichia coli

<400> 48
Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp Val
1 5 10 15
Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp Cys
20 25 30
Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu
35 40 45
Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro
50 55 60

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

<210> 49
<211> 63
<212> DNA
<213> Pseudomonas fluorescens

<220>
<221> CDS
<222> (1)...(63)

<400> 49
atg aga aac ctt ctt cga gga atg ctt gtc gtt att tgc tgt atg gca 48
Met Arg Asn Leu Leu Arg Gly Met Leu Val Val Ile Cys Cys Met Ala
1 5 10 15

ggg ata gct gct gct
Gly Ile Ala Ala Ala 63
20

<210> 50
<211> 21
<212> PRT
<213> Pseudomonas fluorescens

<400> 50
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